STABILITY AND MOMENT BOUNDS UNDER UTILITY-MAXIMISING SERVICE ALLOCATIONS: FINITE AND INFINITE NETWORKS

SEVA SHNEER,* *Heriot-Watt University*
ALEXANDER STOLYAR,** *University of Illinois*

Abstract

We study networks of interacting queues governed by utility-maximising service-rate allocations in both discrete and continuous time. For *finite* networks we establish stability and some steady-state moment bounds under natural conditions and rather weak assumptions on utility functions. These results are obtained using direct applications of Lyapunov–Foster-type criteria, and apply to a wide class of systems, including those for which fluid-limit-based approaches are not applicable. We then establish stability and some steady-state moment bounds for two classes of *infinite* networks, with single-hop and multi-hop message routes. These results are proved by considering the infinite systems as limits of their truncated finite versions. The uniform moment bounds for the finite networks play a key role in these limit transitions.

Keywords: Stochastic stability; stationary moment bounds; utility-maximising service allocations; queueing networks; wireless networks; single-hop networks; multi-hop networks; infinite networks

2010 Mathematics Subject Classification: Primary 60K25
Secondary 68M12

1. Introduction

In this paper we consider networks of interacting queues. These models are primarily motivated by wireless systems, where the interference between simultaneous transmissions by different nodes imposes certain constraints. For example, 'neighbouring' nodes may not be allowed to transmit simultaneously, and/or a node's effective transmission rate may depend on the transmission powers of the node and its neighbours. However, the basic model studied in this paper takes a more abstract point of view; namely, it is an arbitrary network of queues such that individual instantaneous service rates may depend on the state of the entire system. The network state is represented as a set $\bar{X} = (X_i, i \in \mathcal{N})$ of the queue lengths X_i (of jobs, or messages, or customers) at the network nodes $i \in \mathcal{N}$. Each node receives exogenous arrivals of jobs (messages). We consider both *discrete-time* and *continuous-time* models, and both *finite* and *infinite* networks. Also, in addition to *single-hop* networks, where each job leaves the system after its service is completed, we consider one special class of *multi-hop* networks, where,

after a service completion at one node, a customer may either leave the network or be routed to another node, and these routing decisions are taken according to a certain random procedure.

In discrete-time models, time is divided into *slots* of the same (unit) size, and each job (or message) takes exactly one slot to complete service at a node. A service allocation algorithm (or rule) is any mapping (deterministic or random) of a network state $\bar{X}$ into the set of nodes that serve jobs (transmit messages) in a slot. (See, e.g., [15] for a recent model of an algorithm in discrete time employing a random procedure.) The instantaneous service rate μ_i of node i in a slot is the probability that it will serve a job. Thus, the deterministic mapping $\bar{\psi}(\bar{X}) = (\psi_i(\bar{X}), i \in \mathcal{N})$ of a network state $\bar{X}$ into a set of instantaneous service rates $\bar{\mu} = (\mu_i, i \in \mathcal{N})$ is determined by the service algorithm; the mapping $\bar{\psi}(\bar{X})$ is referred to as a service-rate allocation algorithm (or rule).

In continuous-time models, the instantaneous service rate μ_i of a node represents the intensity of the Poisson process modelling departures (service completions) of the node. In this case, the service-rate allocation algorithm $\bar{\psi}(\bar{X})$, mapping a network state $\bar{X}$ into a set of instantaneous service rates $\bar{\mu}$, is all that is needed to specify the service allocation algorithm (see, e.g., [11] for a recent model of an algorithm in continuous time).

In this paper we study service allocation algorithms (in both discrete and continuous time) such that the corresponding service-rate allocation $\bar{\psi}(\bar{X})$ maximises some utility function within some set $\mathcal{C}$. In some cases, the set $\mathcal{C}$ arises naturally as the set of all feasible instantaneous rates $\bar{\mu}$ *given the model structure*, but this is not necessarily true in all cases. For instance, in the networks considered in [15], as well as in Section 3 here, the set $\mathcal{C}$ is a subset of the set of all feasible rates. Our main goal is to obtain network *stability* conditions, in terms of the set $\mathcal{C}$. For example, our main stability results (Theorems 1 and 7) for *single-hop* networks show that the network is stable when the exogenous arrival rates $\bar{\lambda} = (\lambda_i, i \in \mathcal{N})$ are ('strictly') within the set $\mathcal{C}$. In addition to stability, we are able to obtain some steady-state moment bounds (Theorems 2 and 8). These moment bounds turn out to be key to establishing stability of infinite networks, because they allow a limit transition from finite to infinite networks (Theorems 4 and 9).

Service-rate allocations $\bar{\psi}(\bar{X})$, under many natural service allocation algorithms, are such that $\psi_i(\bar{X})$ is decreasing in each X_j for $j \neq i$, as for these algorithms a higher load in queue j usually leads to all other queues receiving less service. This property is in fact satisfied by the rates defined by algorithms introduced in [11] and [15] that we will study here as examples. We would like to emphasise, however, that for our general results we are *not* going to make this assumption. Our motivation for this stems, again, from wireless networks, where there are many competing factors at play, and in many situations ψ_i may not be decreasing in X_j for some $j \neq i$ (see, e.g., the model considered in [14], where the authors consider an algorithm designed to ensure avoidance of conflicts which gives advantage to a transmitter if its non-immediate neighbours are transmitting). This leads to a potentially wide range of possible assumptions on the dependence of service rates assigned to different queues on the state of the network.

We are interested in conditions guaranteeing stability. In finite networks stability, informally speaking, means the ability of all queues to complete service of all jobs, without the number of outstanding jobs building up infinitely. More formally, this means that the Markov chain $\bar{X}(\cdot)$ is positive recurrent. This also implies the existence and uniqueness of a stationary distribution.

In infinite networks, by stability we will understand the existence of a proper invariant distribution. In the cases when the system process is *monotone*, this implies that the process distribution converges to a proper steady state (namely, the lower invariant measure), starting from the 'empty' initial state, as time goes to infinity.

An important concept, explored extensively in the literature, is that of *maximum stability* (or *throughput optimality*). To illustrate this concept, consider a finite network and let $\mathcal{C}$ be the set of all feasible *long-term* rates that can be provided to the nodes, given model constraints. Such a set $\mathcal{C}$ is typically convex. Then, an algorithm is called maximally stable (or throughput-optimal) if it guarantees stability as long as $\bar{\lambda} < \bar{\nu}$ for some $\bar{\nu} \in \mathcal{C}$; in other words, essentially, as long as stability is feasible at all. For a large class of networks, the celebrated MaxWeight algorithm ([18]) and α-fair algorithm are known to be maximally stable. (See [7], [8], [9] for introduction of the fair-allocation concepts and [1] , [4] for stability proofs.) These algorithms, however, are *centralised* in that service-rate allocations are given by a solution to an optimisation problem that needs to be found by a certain central entity. There are also *decentralised* algorithms (where each node regulates its own behaviour according to its queue length) guaranteeing maximal stability (see [6], [12]), but they are known to suffer from large job delays. (This, in particular, prompted the introduction and analysis of algorithms which are not maximally stable, and instead ensure stability for $\bar{\lambda}$ within a 'smaller' set than the set of all feasible long-term rates, with this stability being not necessarily convex. See, e.g., [17].)

Some maximally stable algorithms are designed in such a way that the average service rates maximise a certain utility function. A notable example is presented by α-fair algorithms, where the rates ψ_i are such that

$$\overline{\psi} \in \arg\max_{\overline{\mu} \in \mathcal{C}} \sum_i X_i \frac{1}{1-\alpha} \left(\frac{\mu_i}{X_i} \right)^{1-\alpha} \quad \text{when} \quad \alpha > 0, \alpha \neq 1,$$

or

$$\overline{\psi} \in \arg\max_{\overline{\mu} \in \mathcal{C}} \sum_i X_i \log\left(\mu_i / X_i \right) \quad \text{when} \quad \alpha = 1,$$

where the set $\mathcal{C}$ is usually assumed to be convex. The known stability proofs are based on the fluid limit approach ([10], [2], [16]) and, in particular, implicitly use the fact that α-fair service-rate allocations are 0-homogeneous (or asymptotically 0-homogeneous), which allows a relatively simple characterisation of fluid limit dynamics.

In this paper we consider general utility-maximising algorithms, which, in particular, do not necessarily assign 0-homogeneous rates to queues. We also do *not* require that the maximisation set is necessarily convex. Our goal is threefold. First, we show that these very general algorithms for finite networks ensure stability when $\bar{\lambda}$ is within $\mathcal{C}$. Second, we also find some moment bounds for the stationary queue-length distributions. And finally, we demonstrate how our moment bounds may be used to extend the stability results and moment bounds to some infinite networks.

In the first part of our paper, we consider a class of general utility-maximising algorithms and prove that they are stable when $\bar{\lambda}$ is within $\mathcal{C}$. Namely, we study average service-rate allocations ψ_i such that

$$\overline{\psi} \in \arg\max_{\overline{\mu} \in \mathcal{C}} \sum_i g(X_i) h(\mu_i),$$

with some conditions on the functions g and h. Our conditions do not imply that the service-rate allocations are 0-homogeneous; hence the existing stability results, based on fluid limits, do not apply. Moreover, we do not even require that the function g be defined for non-integer values of the argument. Our results are valid for a large class of functions g such that $g(n+1)/g(n) \to 1$ as $n \to \infty$. This class includes the functions $g(n) = n^\alpha$ used in α-fair allocations, as well as functions of the form $g(n) = e^{\log^\beta n}$ with $\beta > 0$ and $g(n) = e^{n^\gamma}$ with $\gamma \in (0, 1)$, among others.

Our results are also valid for a very general class of functions h. Our stability proofs in both discrete- and continuous-time settings are based on the direct application of Lyapunov–Foster techniques.

In discrete time, for our general results (Theorem 1 and Theorem 2) we impose a strong additional assumption that the number of arrivals into each queue in a time slot is given by a Bernoulli random variable, while if we restrict our attention to a particular scenario of interest (see Theorem 3), we only need to assume a finite third moment of the per-slot number of arrivals. In continuous time, however (which is the standard setting for α-fair allocations), no additional assumptions are needed. We note again that we do not assume that the set $\mathcal{C}$ is convex.

Once stability is established, one is interested in characteristics of the stationary regime. For both discrete- and continuous-time settings, we demonstrate how essentially the same techniques used to prove stability may be employed to establish explicit bounds on the moments of queue states in stationarity.

These bounds are interesting in their own right, especially as very few results are known on the stationary regimes of networks governed by utility-maximising algorithms. We note [13] where an exponential bound has been established for the tail of the total stationary queue length of a system under an α-fair algorithm in a Markovian setting, and [3] where sufficient conditions for the existence of finite moments were established for general arrival streams. We note however that the results of both [3] and [13] imply finiteness of some moments of the stationary queue-length distributions but do not imply any bounds on them as the various constants are not explicit. In the second part of our paper, having explicit bounds is crucial for the analysis of some infinite networks.

In the second part of the paper, we apply the moment bounds established in this paper to obtain stability results for some infinite networks in discrete and continuous time that were considered in the recent papers [15] and [11], respectively. The models considered in the two papers are motivated by different wireless networks but share similar service-rate allocations. As our stability and moment analysis is based on service-rate allocations only, it allows us to handle both discrete and continuous cases, and the particular characteristics of the two models (which are very different), beyond the service-rate allocation, do not play any role in the proofs.

The simplest example of the two networks (results for more general settings are presented in the paper; we focus on a simple example in the introduction only) is given by nodes located on an infinite line $\mathbb{Z}$ and such that, given the state of the system $\bar{X}$, the service-rate allocation is given by

$$\psi_i = \frac{X_i}{X_{i-1} + X_i + X_{i+1}}.$$

The so-called *rate stability* (guaranteeing that the queue lengths do not grow linearly in time) is demonstrated in both discrete and continuous settings in [15] for arrival rates $\bar{\lambda}$ within some natural set $\mathcal{C}$. The authors of [11] considered a continuous-time model where arrival rates into all nodes are the same and equal to λ, say. They considered the system dynamics on intervals of the form $(-n, \ldots, n)$, viewed as a circle, with a growing n. These systems are stable for any n, provided $\lambda < 1/3$, and one can thus consider their stationary measures. Using the natural monotonicity of the corresponding process and the tightness of these measures, a stationary measure (in fact, the lower invariant measure) is constructed for the infinite network. To establish uniqueness of this stationary measure among those with finite second moments of the queue lengths, one needs a bound on the second moments of stationary measures of the

systems on the circle, independent of their size. This was not established in [11], but was left as a conjecture (Conjecture 1.12).

Our analysis is based on showing that the rates of [11] and [15] are in fact utility-maximising (or 2-fair in the α-fair terminology) in a certain natural set $\mathcal{C}$—a fact already mentioned in [15]. This allows us to use our results on stability and moment bounds for finite systems. In particular, our moment bounds immediately imply a uniform (not depending on the size of the network) bound for second moments. This, in turn, proves [11, Conjecture 1.12] in the case of identical arrival rates, along with all its implications, including the uniqueness of the stationary measure constructed there among stationary measures with finite second moments of the queue lengths.

Our analysis, however, allows us to demonstrate the existence of a stationary measure with a finite second moment in far more general settings, where arrival rates do not need to be the same at all nodes, but may be periodic (or dominated by periodic rates). Our analysis of systems with the specific service rates of [11] and [15] does rely on the existence and stability of fluid limits of the systems.

As our analysis is based on utility-maximisation and continuity properties of the processes (see Section 1.2 for the definition of the continuity property), it is not specific to the rates considered in [11] and [15] and may be applied to other infinite networks. We present further examples of models where the same strategy applies. The examples provided in this paper are however not exhaustive.

Finally, we also consider a multi-hop network from [15]. In a multi-hop network, jobs, after being served at one queue, may either leave the network or join another queue to be served there. The analysis of multi-hop networks is notoriously difficult. As in [15], we restrict our attention to symmetric routing. We use similar techniques to the ones we applied in the single-hop setting to first obtain moment bounds for finite networks and then apply these bounds to establish stability of an infinite network. Stability in this case is weaker than that obtained in the single-hop case, as the multi-hop network lacks monotonicity, which is at the core of the construction of the lower invariant measure in [11].

To summarise, our contributions are the following:

- We provide a proof of stability of utility-maximising algorithms in a general setting, covering cases in which fluid limit techniques cannot be applied. In particular, we do not assume that service-rate allocations are 0-homogeneous, and thus we use more general utility functions compared to the classical α-fair algorithms. This comes at the expense of additional assumptions on the arrival processes in discrete time. There are, however, no additional assumptions made in the case of a Markovian (driven by Poisson arrivals and departures) continuous-time system.

- Using a similar approach, we obtain steady-state moment bounds, provided stability conditions hold.

- The same ideas allow us to obtain further explicit steady-state moment bounds in some special cases of interest, which, in turn, allow us to establish stability and moment bounds of some infinite networks. When restricting our attention to some specific service rates, we do use the existence and stability of fluid limits.

- We use similar techniques to establish moment bounds for a certain finite multi-hop network and use these bounds to establish stability and moment bounds for its infinite version.

Single-hop networks   Multi-hop networks
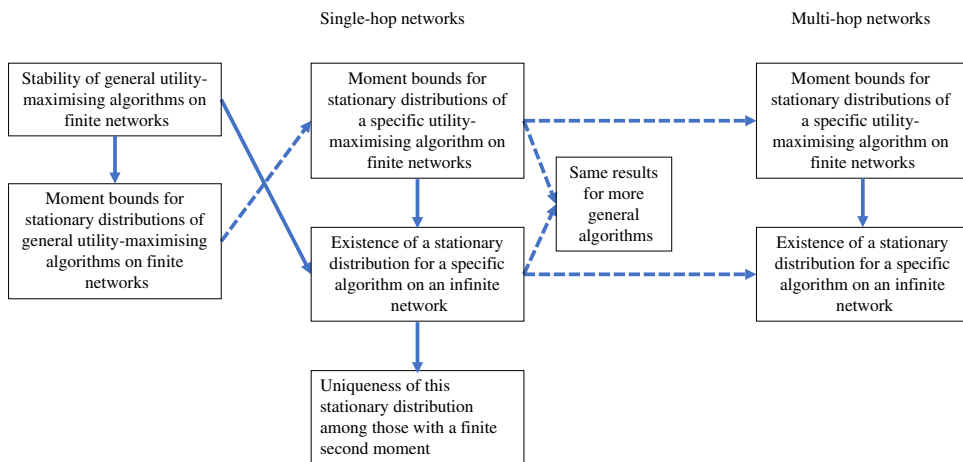

FIGURE 1: An illustration of connections between various parts of the paper. Solid lines represent strict logical connections, dashed lines connections in terms of ideas.

1.1. Structure of the paper

In order to facilitate a smooth presentation, we first present our main results in the discrete-time setting. Section 2 is devoted to finite networks. We describe the model and present our stability results and steady-state moment bounds. Sections 3 and 4 are devoted to infinite networks in the single- and multi-hop settings, respectively. In Section 3, we first analyse the model introduced in [11] and [15] and then demonstrate how our analysis may be extended to treat other related models. Section 4 is devoted to a symmetrical multi-hop extension of the model of [11] and [15] .

Section 5 is devoted to the continuous-time setting. We describe the model, explain why the analysis is a simplified version of our analysis in the discrete-time setting, and present continuous-time versions of our results.

The paper contains many parts, some of which are connected directly through statements, while others are only connected through ideas. The structure of the paper is thus nonlinear. We present an illustration of how the various parts are related in Figure 1.

1.2. Basic notation, conventions, and definitions

We will use the following notation throughout: $\mathbb{R}$ and $\mathbb{R}_+$ are the sets of real and real nonnegative numbers, respectively; $\mathbb{Z}^d$ is the d-dimensional lattice; $\mathbb{Z}_+$ is the set of nonnegative integers; $\bar{y}$ means the (finite- or infinite-dimensional) vector (y_i); for a finite-dimensional vector $\bar{y}$, $\|y\| = \sum_i |y_i|$; for a set of functions (f_i) and a vector $\bar{y}$, $\bar{f}(\bar{y})$ denotes the vector $(f_i(\bar{y}))$; vector inequalities are understood componentwise; we also use the convention that $0/0 = 0$.

The abbreviation *w.p.1* means *with probability* 1. Convergence in distribution of random elements is denoted by $\Rightarrow$. A discrete-time random process $(Y(k), k = 0, 1, 2, \dots)$ is often referred to as $Y(\cdot)$, and similarly for a continuous-time process $(Y(t), t \geq 0)$.

We will say that a sequence of random processes $Y^{(m)}(\cdot)$, $m = 1, 2, \dots$, and a random process $Y(\cdot)$ satisfy the *continuity* property if the following holds. Given any (random) initial state $Y(0)$ and any sequence of (random) initial states $Y^{(m)}(0)$, $m = 1, 2, \dots$, such that $Y^{(m)}(0) \Rightarrow Y(0)$, all processes can be coupled (constructed on a common probability space) so that $Y^{(m)}(k) \to Y(k)$ w.p.1 for any $k = 0, 1, \dots$ (or, for continuous time, $Y^{(m)}(t) \to Y(t)$ w.p.1

for any $t \geq 0$). This continuity property could be called generalised Feller-continuity, because in the special case when all $Y^{(m)}(\cdot)$ are copies of the same process $Y(\cdot)$, differing only in the initial state, the property defined above is Feller-continuity of $Y(\cdot)$; we call it continuity for short.

2. Finite single-hop networks: stability analysis and moment bounds

In this section we consider finite single-hop networks, where a job, after being served at any queue (node), leaves the system. Since the number of nodes is finite, the process describing system evolution is a countable (irreducible) Markov chain. The finite-network process *stability* is defined as positive recurrence of the Markov chain, which (because of irreducibility) is equivalent to the existence of a unique stationary distribution. We first introduce the model and make general assumptions, then state and prove stability results, and finally obtain moment bounds on stationary distributions.

2.1. Model and assumptions

Assume that there are N queues, each having its own arrival stream of jobs, and having an infinite buffer to store outstanding jobs. For models in discrete time, we will assume that all jobs require service that lasts 1 time unit, time is split into *slots* of length 1, and all arrivals and all service initiations happen at the beginning of a time slot, so that all services are completed by the end of a time slot. These assumptions are motivated mainly by wireless networks.

For convenience we assume that at the beginning of each time slot, first new services are started, and then new arrivals happen. We will denote time slots by $k = 0, 1, \ldots$. We can then write the evolution of the queue of node i as

$$X_i(k+1) = X_i(k) + \xi_i(k) - \eta_i(k), \tag{1}$$

where $\xi_i(k)$ denotes the number of new job arrivals into queue i at time slot k, and $\eta_i(k)$ denotes the number of service completions in queue i at time k.

We will assume that for each i, the sequence $\xi_i(k)$, $k = 0, 1, 2, \ldots$, consists of independent, identically distributed (i.i.d.) random variables such that $\mathbb{E}(\xi_i) = \lambda_i$, where, here and throughout, by ξ_i we denote a random variable with the distribution of any of $\xi_i(k)$. Note that arbitrary dependence between random variables with different values of i is allowed.

We will assume also that the random variables $\eta_i(k)$ take values 0 and 1 and are such that, on average, they maximise a global utility function in the following sense. Define

$$\psi_i(\bar{x}) = \mathbb{E}(\eta_i(k) | \bar{X}(k) = \bar{x}),$$

and assume that $\psi_i(\bar{x}) \in [0, 1]$ for all $\bar{x}$, $\psi_i(\bar{x}) = 0$ if $x_i = 0$, and

$$\bar{\psi}(\bar{x}) \in \arg \max_{\bar{\mu} \in \mathcal{C}} \sum_i g(x_i) h(\mu_i), \tag{2}$$

where the set $\mathcal{C}$ is compact and coordinate-convex (i.e. if a vector $\bar{\mu}$ belongs to $\mathcal{C}$ and $\bar{\mu}^* \leq \bar{\mu}$ coordinate-wise, then $\bar{\mu}^* \in \mathcal{C}$). We impose, in addition, the following conditions:

Condition (H): the function $h : [0, \infty) \to \mathbb{R}$ is strictly increasing, differentiable, and concave (both the cases $\lim_{y \downarrow 0} h(y) = h(0) > -\infty$ and $\lim_{y \downarrow 0} h(y) = -\infty$ are allowed); and

Condition (G): the function $g : \mathbb{Z}_+ \to [0, \infty)$ is strictly increasing and such that

$$\frac{g(y)}{\Delta(y)} \to \infty \tag{3}$$

as $y \to \infty$, where $\Delta(y) = g(y+1) - g(y)$. Note that (3) is equivalent to

$$\frac{g(y+1)}{g(y)} \to 1, \tag{4}$$

as $y \to \infty$.

Remark 1. Note that for what are usually referred to as α-fair algorithms, either $g(y) = y^\alpha$ and $h(y) = \frac{y^{1-\alpha}}{1-\alpha}$ with $\alpha > 0$, $\alpha \neq 1$, or $g(y) = y$ and $h(y) = \log y$, so all of the above conditions hold.

The class of functions satisfying the conditions above is however much wider. It includes, for instance, the functions $g(y) = e^{\log^\beta y}$ with $\beta > 0$ and $g(y) = e^{y^\gamma}$ with $0 < \gamma < 1$.

Throughout this section, we are going to assume the following:

$$\text{There exists } \bar{v} \in \mathcal{C} \text{ such that } \bar{\lambda} < \bar{v}. \tag{5}$$

We will also use the notation

$$G(z) = \sum_{y=0}^{z} g(y) \tag{6}$$

and

$$F(\bar{y}) = \sum_i h'(v_i) G(y_i). \tag{7}$$

2.2. Stability

In this section we prove that the utility-maximising algorithms described in the previous section are stable as long as the average arrivals $\bar{\lambda}$ are within the set $\mathcal{C}$. Our proof does not use fluid limits, which have been the standard tool for proving stability of algorithms of this type. The advantages and disadvantages of our approach are described in the introduction.

Theorem 1. *Consider the discrete-time model in Section 2.1 and assume that ξ_i is a Bernoulli random variable with $\mathbb{E}(\xi_i) = \lambda_i$. Assume that the vector $\bar{\lambda}$ is such that Condition (5) holds. Then the Markov chain $\{\bar{X}(k), k = 0, 1, \ldots\}$ is positive recurrent.*

Proof of Theorem 1. We will use the standard Lyapunov–Foster criterion [5]. Fix $\varepsilon > 0$ such that $\lambda_i < v_i - \varepsilon$ for all i. Note that, by (2) and the concavity of the function h,

$$0 \leq \sum_i g(x_i)(h(\psi_i(\bar{x})) - h(v_i)) \leq \sum_i g(x_i) h'(v_i)(\psi_i(\bar{x}) - v_i). \tag{8}$$

We are going to consider

$$\mathbb{E}\left(F(\bar{X}(1)) - F(\bar{X}(0))|\bar{X}(0) = \bar{x}\right) = \mathbb{E}\left(F(\bar{x} + \bar{\xi}(0) - \bar{\eta}(0))|\bar{X}(0) = \bar{x}\right) - F(\bar{x}).$$

In what follows we are going to assume that $\bar{X}(0) = \bar{x}$ is fixed and will drop the dependence on this event. We will also write ξ_i and η_i instead of $\xi_i(0)$ and $\eta_i(0)$, for simplicity. We can write

$$\mathbb{E}(F(\bar{x} + \bar{\xi} - \bar{\eta})) - F(\bar{x}) = \mathbb{E}\left(\sum_i G(x_i + \xi_i - \eta_i) h'(v_i)\right) - \sum_i G(x_i) h'(v_i)$$

$$= \sum_i h'(v_i)\Big(\lambda_i \psi_i(\bar{x}) G(x_i) + (1 - \lambda_i)(1 - \psi_i(\bar{x})) G(x_i)$$

$$+ \lambda_i (1 - \psi_i(\bar{x})) G(x_i + 1) + (1 - \lambda_i) \psi_i(\bar{x}) G(x_i - 1) - G(x_i)\Big)$$

$$= \sum_i h'(v_i)\Big(\lambda_i(1 - \psi_i(\bar{x}))(G(x_i + 1) - G(x_i)) + (1 - \lambda_i)\psi_i(\bar{x})(G(x_i - 1) - G(x_i))\Big)$$

$$= -\sum_i h'(v_i)\lambda_i\psi_i(\bar{x})(G(x_i + 1) + G(x_i - 1) - 2G(x_i))$$

$$\quad + \sum_i h'(v_i)\Big(\lambda_i(G(x_i + 1) - G(x_i)) + \psi_i(\bar{x})(G(x_i - 1) - G(x_i))\Big)$$

$$= -\sum_i h'(v_i)\lambda_i\psi_i(\bar{x})(g(x_i + 1) - g(x_i)) + \sum_i h'(v_i)\Big(\lambda_i g(x_i + 1) - \psi_i(\bar{x})g(x_i)\Big)$$

$$\leq \sum_i h'(v_i)\Big(\lambda_i g(x_i + 1) - \psi_i(\bar{x})g(x_i)\Big) \tag{9}$$

$$= \sum_i h'(v_i)(\lambda_i - \psi_i(\bar{x}))g(x_i) + \sum_i h'(v_i)\lambda_i(g(x_i + 1) - g(x_i))$$

$$= \sum_i h'(v_i)(\lambda_i - v_i)g(x_i) + \sum_i h'(v_i)(v_i - \psi_i(\bar{x}))g(x_i) + \sum_i h'(v_i)\lambda_i\Delta(x_i)$$

$$\leq -\varepsilon \sum_i h'(v_i)g(x_i) + \sum_i h'(v_i)\Delta(x_i) = \sum_i h'(v_i)(-\varepsilon g(x_i) + \Delta(x_i)),$$

where in the last inequality we used (8).

This, together with (3), implies that if $F(\bar{x})$ is large, then its expected drift may be made arbitrarily small, and certainly negative and bounded away from zero (and then the classical Lyapunov–Foster stability criterion in [5] applies).

Indeed, if $F(\bar{x})$ is larger than a certain constant, say C, then there exists i such that $h'(v_i)G(x_i) > C/N$, which, since G is strictly increasing, implies that $x_i > C_1$ for a certain constant C_1.

The condition (3) implies that there exists a constant C_2 such that

$$\frac{g(y)}{\Delta(y)} \geq C_2$$

for all y. If $\varepsilon C_2 > 1$, then the drift of F is always negative. Assume now $\varepsilon C_2 < 1$. For any C_3 there exists Y such that

$$\frac{g(y)}{\Delta(y)} \geq C_3$$

for all $y > Y$. We can always choose C_3 and C so that $C_3 > 1/\varepsilon$ and $C_1 > Y$, and then the drift of F may be bounded from above by

$$\sum_i h'(v_i)g(x_i)\left(-\varepsilon + \frac{1}{C_2}\mathbb{I}(x_i \leq Y) + \frac{1}{C_3}\mathbb{I}(x_i > Y)\right)$$

$$\leq (N - 1)h_u g(Y)\left(-\varepsilon + \frac{1}{C_2}\right) + h_l g(C_1)\left(-\varepsilon + \frac{1}{C_3}\right),$$

where $h_u = \max_i h'(v_i)$ and $h_l = \min_i h'(v_i)$. It is clear that we can always choose C_1 such that the above is negative. This completes the proof.

2.3. Moment bounds

Once stability is established, one can employ arguments similar to those used in the proof of Theorem 1 to obtain bounds on some moments of the stationary distributions of queue states.

The stationary regime exists under the conditions of Theorem 1; in this section we will write $\bar{X}$ to represent a random vector with distribution equal to that of $\bar{X}(k)$ in the stationary regime.

Theorem 2. *Assume that all conditions of Theorem 1 hold and fix $\varepsilon > 0$ such that $\lambda_i < \nu_i - \varepsilon$ for each i. Then*

$$\sum_i h'(\nu_i)\mathbb{E}g(X_i) \leq \frac{1}{\varepsilon}\sum_i h'(\nu_i)\mathbb{E}\Delta(X_i).$$

Proof of Theorem 2. Consider the process with an arbitrary fixed initial state $\bar{X}(0)$. Then, because of the assumptions on the input flows, $\mathbb{E}(F(\bar{X}(k))) < \infty$ for any $k \geq 0$. By Theorem 1 the process is stable, and therefore $\bar{X}(k)$ converges in distribution to $\bar{X}$. Following the lines of (9), we have the following drift estimate:

$$\mathbb{E}\big(F(\bar{X}(k+1)) - F(\bar{X}(k))|\bar{X}(k) = \bar{x}\big)$$
$$\leq \sum_i h'(\nu_i)\big(\lambda_i g(x_i + 1) - \psi_i(\bar{x})g(x_i)\big) \tag{10}$$
$$\leq -\varepsilon \sum_i h'(\nu_i)g(x_i) + \sum_i h'(\nu_i)\Delta(x_i) \leq c,$$

where c is a fixed finite constant, and the last inequality follows from the properties of the function g. Indeed, thanks to (3), there exists a finite Y such that

$$\frac{g(y)}{\Delta(y)} \geq 1/\varepsilon$$

for all $y \geq Y$. Define

$$C = \min_{y \leq Y}\frac{g(y)}{\Delta(y)}.$$

Then

$$\mathbb{E}\big(F(\bar{X}(k+1)) - F(\bar{X}(k))|\bar{X}(k) = \bar{x}\big) \leq \left(-\varepsilon + \frac{1}{C}\right)\sum_i h'(\nu_i)g(x_i)\mathbb{I}(x_i < Y).$$

Now c may be taken to be

$$c = \begin{cases} 0, & C \geq 1/\varepsilon, \\ N\left(-\varepsilon + \frac{1}{C}\right)g(Y)\max_i h'(\nu_i), & C < 1/\varepsilon. \end{cases}$$

Then we obtain

$$\mathbb{E}[F(\bar{X}(k+1)) - F(\bar{X}(k))] \leq \mathbb{E}\left[-\varepsilon\sum_i h'(\nu_i)g(X_i(k)) + \sum_i h'(\nu_i)\Delta(X_i(k))\right].$$

Recall that $\bar{X}(k)$ converges in distribution to $\bar{X}$. Using Skorohod representation and the last inequality in (10), we can apply Fatou's lemma to take $\limsup_{k\to\infty}$ of the right-hand side of the above inequality. Thus, we obtain

$$\limsup_{k\to\infty}\mathbb{E}[F(\bar{X}(k+1)) - F(\bar{X}(k))] \leq \mathbb{E}\left[-\varepsilon\sum_i h'(\nu_i)g(X_i) + \sum_i h'(\nu_i)\Delta(X_i)\right].$$

The left-hand side of the above must be greater than or equal to 0 (because otherwise we would have $\mathbb{E}F(\bar{X}(k)) \to -\infty$). This completes the proof.

In certain cases (such as, e.g., when $g(x) = x^\alpha$ with α an integer) one can significantly weaken the assumptions on the random variables ξ_i. We provide the following theorem in order to illustrate this, and also to introduce the specific choice of functions g and h which we will use in Section 3 to prove stability of an infinite network. We note however that we do rely on fluid limits in the proof of the following result.

Theorem 3. *For a discrete-time model as defined in Section 2.1, assume that $g(y) = y^2$ and $h(y) = -y^{-1}$. Assume also that ξ_i is a nonnegative integer-valued random variable with $\mathbb{E}(\xi_i^3) < \infty$ and $\mathbb{E}(\xi_i) = \lambda_i$, where the vector $\bar{\lambda}$ is such that Condition (5) holds. Then the Markov chain $\{\bar{X}(k), k = 0, 1, \ldots\}$ is stable.*

Moreover, fix $\varepsilon > 0$ such that $\lambda_i < \nu_i - \varepsilon$ for all i. Then

$$\varepsilon \sum_i \frac{\mathbb{E}(X_i^2)}{\nu_i^2} \leq A \sum_i \frac{\mathbb{E}(X_i)}{\nu_i^2} + B,$$

where $\bar{X}$ denotes a random element with the stationary distribution of $\bar{X}(\cdot)$,

$$A = 3 \sum_i \frac{\mathbb{E}(\xi_i^2) + \lambda_i(1 - 2\lambda_i)}{\nu_i^2},$$

and

$$B = \sum_i \frac{\mathbb{E}(\xi_i^3) - \lambda_i + 3\lambda_i^2 - 3(1 - 2\lambda_i)(\lambda_i^2 - \lambda_i/2 + \mathbb{E}(\xi_i^2)/2)}{\nu_i^2}.$$

The specific functions $g(\cdot)$ and $h(\cdot)$ are such that the fluid limits of the process are well-defined. We use this fact in applying previous results on stability and existence of moments, as will be seen shortly.

Proof of Theorem 3. Under the assumptions of the theorem (in fact the existence of only the first moments of the ξ_i is sufficient for stability), positive recurrence of the Markov chain $\bar{X}(\cdot)$ holds by [15, Lemma 12], where the stability of the corresponding fluid limits is established. (Note that if one assumed convexity of the set $\mathcal{C}$, stability would follow from earlier results (see, e.g., [1], [4]); however, as pointed out in [15], the convexity of the set $\mathcal{C}$ is not in fact necessary for stability results). We will consider the stationary version of the process. The finiteness of the third moment of ξ_i (along with stability of fluid limits) guarantees that $\mathbb{E}(X_i^2) < \infty$ (see [3]).

By stationarity, $\mathbb{E}(X_i(k+1)) = \mathbb{E}(X_i(k))$, and hence

$$\mathbb{E}(\psi_i) = \lambda_i, \tag{11}$$

where for simplicity we write ψ_i instead of $\psi_i(\bar{X})$.

Note that

$$\mathbb{E}(X_i^l \eta_i) = \mathbb{E}(\mathbb{E}(X_i^l \eta_i | \bar{X})) = \mathbb{E}(X_i^l \mathbb{E}(\eta_i | \bar{X})) = \mathbb{E}(X_i^l \psi_i) \tag{12}$$

for any l. Note also that $\eta_i^l = \eta$ almost surely for any $l > 0$. By stationarity, we also have $\mathbb{E}(X_i^2(k+1)) = \mathbb{E}(X_i^2(k))$, which is equivalent to

$$0 = \mathbb{E}(\xi_i^2) + \mathbb{E}(\psi_i) - 2\mathbb{E}(\xi_i)\mathbb{E}(\psi_i) + 2\mathbb{E}(\xi_i)\mathbb{E}(X_i) - 2\mathbb{E}(X_i\psi_i)$$
$$= \mathbb{E}(\xi_i^2) + \lambda_i - 2\lambda_i^2 + 2\lambda_i\mathbb{E}(X_i) - 2\mathbb{E}(X_i\psi_i),$$

where we used (11), and hence

$$\mathbb{E}(X_i \psi_i) = \lambda_i \mathbb{E}(X_i) - \lambda_i^2 + \lambda_i/2 - \mathbb{E}(\xi_i^2)/2. \tag{13}$$

Assume now that $\mathbb{E}(X_i^3) < \infty$ (we will demonstrate how to drop this additional assumption at the end of the proof). Then the equality of the third moments in stationarity implies

$$
\begin{aligned}
0 &= \mathbb{E}(\xi_i^3) - \mathbb{E}(\psi_i) + 3\mathbb{E}(\xi_i)\mathbb{E}(\psi_i) - 3\mathbb{E}(\xi_i^2)\mathbb{E}(\psi_i) + 3\mathbb{E}(X_i\psi_i) - 3\mathbb{E}(X_i^2\psi_i) + 3\mathbb{E}(X_i)\mathbb{E}(\xi_i^2) \\
&\quad + 3\mathbb{E}(X_i^2)\mathbb{E}(\xi_i) - 6\mathbb{E}(X_i\psi_i)\mathbb{E}(\xi_i) \\
&= \mathbb{E}(\xi_i^3) - \lambda_i + 3\lambda_i^2 - 3\lambda_i\mathbb{E}(\xi_i^2) - 3\mathbb{E}(X_i^2\psi_i) + 3\mathbb{E}(X_i)\mathbb{E}(\xi_i^2) + 3\lambda_i\mathbb{E}(X_i^2) \\
&\quad + 3(1 - 2\lambda_i)(\lambda_i\mathbb{E}(X_i) - \lambda_i^2 + \lambda_i/2 - \mathbb{E}(\xi_i^2)/2) \\
&= 3(\lambda_i\mathbb{E}(X_i^2) - \mathbb{E}(X_i^2\psi_i)) + A_i\mathbb{E}(X_i) + B_i,
\end{aligned} \tag{14}
$$

where we used (11) and (13), and where

$$A_i = 3\mathbb{E}(\xi_i^2) + 3\lambda_i(1 - 2\lambda_i)$$

and

$$B_i = \mathbb{E}(\xi_i^3) - \lambda_i + 3\lambda_i^2 - 3(1 - 2\lambda_i)(\lambda_i^2 - \lambda_i/2 + \mathbb{E}(\xi_i^2)/2).$$

By (8),

$$0 \le \sum_i \frac{x_i^2}{v_i^2}(\psi_i(\bar{x}) - v_i)$$

for any $\bar{x}$, and hence

$$
\sum_i \frac{\lambda_i \mathbb{E}(X_i^2) - \mathbb{E}(X_i^2\psi_i)}{v_i^2} = \sum_i \frac{(\lambda_i - v_i)\mathbb{E}(X_i^2)}{v_i^2} + \sum_i \frac{\mathbb{E}((v_i - \psi_i)X_i^2)}{v_i^2}
$$

$$
\le -\varepsilon \sum_i \frac{\mathbb{E}(X_i^2)}{v_i^2}.
$$

The statement of the theorem now follows from dividing (14) by v_i^2 and summing over all i.

We now show that the assumption $\mathbb{E}(X_i^3) < \infty$ can be dropped. Let $M < \infty$ and consider the system with arrivals given by $\xi_i^{(M)} = \min\{\xi_i, M\}$ instead of ξ_i. Of course, $\mathbb{E}\xi_i^{(M)} \le \mathbb{E}\xi_i$. Therefore the system is stable for each M; let us denote by $\bar{X}^{(M)}$ a random element which has its stationary distribution. For each M, $\mathbb{E}((X_i^{(M)})^3) < \infty$ (because $\mathbb{E}((\xi_i^{(M)})^4) < \infty$; see [3]), and the derivations above imply that

$$\sum_i \frac{\mathbb{E}((X_i^{(M)})^2)}{v_i^2} \le A^{(M)} \sum_i \frac{\mathbb{E}(X_i^{(M)})}{v_i^2} + B^{(M)},$$

with the obvious expressions for $A^{(M)}$ and $B^{(M)}$. Since $\mathbb{E}((\xi_i^{(M)})^l) \to \mathbb{E}(\xi_i^l)$ as $M \to \infty$ for $l = 1, 2, 3$, we have $A^{(M)} \to A$ and $B^{(M)} \to B$ as $M \to \infty$.

It is easy to check that the sequences $\bar{X}^{(M)}(\cdot)$ and $\bar{X}(\cdot)$ satisfy the continuity property. Indeed, if $\bar{X}^{(M)}(0) \Rightarrow \bar{X}(0)$, we can use Skorohod representation to construct all (random) initial states $\bar{X}^{(M)}(0)$ and $\bar{X}(0)$ on a common probability space so that $\bar{X}^{(M)}(0) \to \bar{X}(0)$ w.p.1. Now, we augment that probability space to the following natural probability space, on which all processes

$\bar{X}^{(M)}(\cdot)$ and $\bar{X}(\cdot)$ are constructed. The service process is driven by the (independent) sequence of i.i.d. random mappings from a queue length vector $\bar{X}$ to a service vector (η_i). The arrival process is driven by an (independent) sequence of i.i.d. arrival vectors (ξ_i). All processes $\bar{X}^{(M)}(\cdot)$ and $\bar{X}(\cdot)$ are then constructed in exactly the same natural way, except that in the Mth processes the number of job arrivals is 'clipped' at M; i.e. it is $\xi_i^{(M)} = \min\{\xi_i, M\}$. We directly observe that, w.p.1, for any time k and all sufficiently large M, $\bar{X}^{(M)}(k) = \bar{X}(k)$.

Note that all $X_i^{(M)}$ have uniformly (in M) bounded second—and then also first—moments. Therefore we can choose a subsequence of values of M along which $\bar{X}^{(M)} \Rightarrow \tilde{X}$. We now construct the stationary versions of the processes $\bar{X}^{(M)}(\cdot)$ (along the subsequence) and a process $\bar{X}(\cdot)$ with $\bar{X}(0)$ distributed as $\tilde{X}$, on a common probability space as described above. Since the sequences $\bar{X}^{(M)}(\cdot)$ and $\bar{X}(\cdot)$ satisfy the continuity property, we see that, w.p.1, $\bar{X}^{(M)}(k) \to \bar{X}(k)$ for any k, and the $\bar{X}(\cdot)$ thus constructed is in fact stationary. This, in turn, implies that $\bar{X}^{(M)} \Rightarrow \bar{X}$.

It remains to rewrite the last display as

$$\sum_i \frac{1}{v_i^2} \mathbb{E}[X_i^{(M)} - A^{(M)}/2]^2 \le B^{(M)} + \sum_i \frac{(A^{(M)})^2}{4v_i^2}$$

and apply Fatou's lemma to obtain

$$\sum_i \frac{1}{v_i^2} \mathbb{E}[X_i - A/2]^2 \le \liminf_{M \to \infty} \sum_i \frac{1}{v_i^2} \mathbb{E}[X_i^{(M)} - A^M/2]^2 \le B + \sum_i \frac{A^2}{4v_i^2}.$$

This completes the proof.

3. Stability analysis of infinite single-hop networks

In this section we provide an application of our moment bounds to establishing the stability of infinite networks considered in [11] and [15]. The *stability of an infinite network* we define as the existence of a proper stationary distribution (with all queues finite with probability 1).

In our analysis of finite systems in the previous sections, only the average service rates (at a given time, given the system state) were of importance, and any dependencies between the departures from different queues were not relevant. When we move to analysis of infinite systems, we still will not require that the departures in each time slot be independent (given the system state), but we do have to specify the departure (service) mechanism to make sure that the processes we consider satisfy the continuity and/or monotonicity properties. In particular, continuity will be the key property which we need in order to make limit transitions from finite systems to infinite ones.

For the motivation of the specific service mechanisms that we consider (in particular related to wireless networks), we refer the reader to [11] and [15].

3.1. Model

The queues (or nodes) are assumed to be located on a d-dimensional lattice, with the service rates given by

$$\psi_i(\bar{x}) = \frac{x_i}{\sum_{j \in \mathbb{Z}^d} a_{j-i} x_j}, \tag{15}$$

where $a_0 = 1$, $a_i = a_{-i} \ge 0$ for all $i \in \mathbb{Z}^d$, and $L = \sup\{|i| : a_i > 0\} < \infty$. For each i, the nodes j within the finite set $\mathcal{N}_i = \{j | a_{j-i} > 0\}$ are called *neighbours* of i. Note that $i \in \mathcal{N}_i$.

Recall that the arrivals are driven by the set of independent random variables $\xi_i(k)$, which represent the number of arrivals into node i at time k. The sets $\{\xi_i(k)\}$ are i.i.d. across k, and for

each fixed i, the $\xi_i(k)$ are i.i.d. across k. As before, we denote by ξ_i the generic $\xi_i(k)$, and we assume

$$\mathbb{E}\xi_i^3 < \infty.$$

We consider the following two service algorithms for the discrete-time case. Our results apply to both. (Again, see [15] for the motivation of the algorithms.) Recall that the $X_i(k)$ are the queue lengths at time k.

Discrete-time service algorithm 1 (D1). The algorithm is driven by the set of i.i.d. (across node indices i and times k) random variables $\nu_i(k)$, distributed uniformly in $[0, 1]$. The *access priority* of node i at time k is $\tau_i(k) = [-\log \nu_i]/X_i(k)$; it is exponentially distributed with mean $1/X_i(k)$. (The smaller the $\tau_i(k)$, the 'higher' the priority.) Node i transmits in slot k if $X_i(k) > 0$ and

$$\tau_i(k) < \tau_j(k)/a_{j-i} \text{ for all } j \in \mathcal{N}_i \setminus i.$$

Note that the probability of node i transmitting, conditioned on $\bar{X}(k)$, is exactly $\psi_i(\bar{X}(k))$, as required by (15). At the same time, the transmissions of the nodes at time k, even conditioned on $\bar{X}(k)$, are *not* independent (except in the degenerate case $\mathcal{N}_i = i$). In fact, in the case when all a_i are either 1 or 0, neighbouring nodes can *never* transmit simultaneously.

Discrete-time service algorithm 2 (D2). This algorithm is much simpler. It is also driven by the set of i.i.d. (across node indices i and times k) random variables $\nu_i(k)$, distributed uniformly in $[0, 1]$. Node i transmits in slot k if $X_i(k) > 0$ and

$$\nu_i(k) < \psi_i(\bar{X}(k)).$$

In other words, conditioned on $\bar{X}(k)$, the probabilities of nodes transmitting are exactly $\psi_i(\bar{X}(k))$ (as required by (15)), and the transmissions are independent.

3.2. Continuity and monotonicity

For the infinite system process, we will use the continuity property (defined in Section 1.2) and the monotonicity property (defined below).

For a continuity property to be well-defined, a topology on the process state space needs to be specified. A state of the process is a set $\bar{X} = \{X_i\}$ of the queue lengths, i.e. a function of i. On this state space (which is uncountable for an infinite system), we consider the natural topology of componentwise convergence.

We also consider the natural componentwise order relation $\bar{X} \leq \bar{X}^*$ on the state space. With respect to this partial order, it is easy to see that the process for the system defined above has the following *monotonicity* property: two versions of the process such that $\bar{X}^*(0) \leq \bar{X}(0)$ can be coupled (constructed on a common probability space) so that $\bar{X}^*(k) \leq \bar{X}(k)$ at all times $k \geq 0$. We note that this monotonicity only holds for single-hop systems; it does *not* hold for the multi-hop system which we will consider later in Section 4.

3.3. Auxiliary system on a finite torus

For a vector $\bar{N} = (N_1, \ldots, N_d)$, denote by $\mathcal{T} = \{i : i_k \in \{-N_k, \ldots, N_k - 1\}\} \subset \mathbb{Z}^d$ the finite subset of points, 'wrapped around' to form a torus. Consider the auxiliary version of our system, defined on the finite torus $\mathcal{T}$, with the node neighborhood structure being that of the torus. Denote by $n = \prod_k (2N_k)$ the number of points in $\mathcal{T}$. We will only consider vectors $\bar{N}$ such that $N_k > l$ for each k.

Along the lines of [15, Lemma 11], we can show that for any such $\mathcal{T}$, the rates (15) are in fact utility-maximising in a certain set. Indeed, let

$$\mathcal{C} = \{\bar{\mu} : \quad \text{there exists} \quad \bar{p} \quad \text{such that} \quad \bar{\mu} \leq \bar{\psi}(\bar{p})\}.$$

We can prove the following optimality result.

Lemma 1. *The rates* (15) *are utility-maximising (they satisfy the relation* (2)*) for the functions* $g(y) = y^2$ *and* $h(y) = -y^{-1}$ *and the set* $\mathcal{C}$.

Remark 2. Using the standard terminology of α-fairness, Lemma 1 states that the rates (15) are 2-fair in the set $\mathcal{C}$.

Proof of Lemma 1. Indeed, by the definition of the set $\mathcal{C}$, for any $\bar{\mu} \in \mathcal{C}$,

$$\sum_i x_i \left(\frac{\mu_i}{x_i} \right)^{-1} \geq \sum_i x_i \left(\frac{\psi_i(\bar{p})}{x_i} \right)^{-1}$$

for the corresponding vector $\bar{p}$. Hence it is sufficient to show that

$$\sum_i x_i \left(\frac{\psi_i(\bar{x})}{x_i} \right)^{-1} \leq \sum_i x_i \left(\frac{\psi_i(\bar{p})}{x_i} \right)^{-1}$$

for all vectors $\bar{p}$. Note that the left-hand side of the above is equal to

$$\sum_i x_i \sum_{j \in \mathcal{T}} a_{j-i} x_j = \sum_i x_i^2 + \sum_i \sum_{j \in \mathcal{T}, j \neq i} a_{j-i} x_i x_j.$$

Consider now

$$\sum_i x_i \left(\frac{p_i}{(\sum_{j \in \mathcal{T}} a_{j-i} p_j) x_i} \right)^{-1} = \sum_i x_i^2 \left(1 + \sum_{j \in \mathcal{T}, j \neq i} \frac{a_{j-i} p_j}{p_i} \right)$$

$$= \sum_i x_i^2 + \frac{1}{2} \sum_i \sum_{j \in \mathcal{T}, j \neq i} \left(x_i^2 \frac{a_{j-i} p_j}{p_i} + x_j^2 \frac{a_{i-j} p_i}{p_j} \right).$$

For any i and j,

$$x_i^2 \frac{a_{j-i} p_j}{p_i} + x_j^2 \frac{a_{i-j} p_i}{p_j} = a_{j-i} \left(x_i^2 \frac{p_j}{p_i} + x_j^2 \frac{p_i}{p_j} \right) \geq 2 a_{j-i} x_i x_j,$$

where we used the symmetry of the sequence a_i. Equality in the above is possible if and only if $x_i^2 \frac{p_j}{p_i} = x_j^2 \frac{p_i}{p_j}$, which is equivalent to $\frac{p_i}{x_i} = \frac{p_j}{x_j}$. Therefore we obtain

$$\sum_i x_i \left(\frac{p_i}{(\sum_{j \in \mathcal{T}} a_{j-i} p_j) x_i} \right)^{-1} \geq \sum_i x_i^2 + \sum_i \sum_{j \in \mathcal{T}, j \neq i} a_{j-i} x_i x_j,$$

and equality is possible if and only if $\frac{p_i}{x_i} = \frac{p_j}{x_j}$ for all i and j. This implies that $\frac{p_i}{x_i}$ has to be a constant for each i. This completes the proof.

Lemma 2. *If* $\lambda_i = \lambda$ *for each i, then the existence of* $\bar{v} \in \mathcal{C}$ *such that* $\bar{\lambda} < \bar{v}$ *is equivalent to the inequality*

$$\lambda < \frac{1}{\sum_{j \in \mathcal{T}} a_j} = \frac{1}{\sum_{j \in \mathbb{Z}^d} a_j}$$

for any vector $\bar{N}$ *such that* $N_k > L$ *for all k.*

Proof of Lemma 2. Indeed, if

$$\lambda < \frac{1}{\sum_{j \in \mathcal{T}} a_j},$$

we can take $\bar{p} = (1, \ldots, 1)$ and $\bar{\nu} = \bar{\psi}(\bar{p})$—such a vector clearly belongs to $\mathcal{C}$, and it is also clear that $\bar{\lambda} < \bar{\nu}$. In the opposite direction, assume that $\bar{\lambda} < \bar{\nu}$ for some $\bar{\nu} \in \mathcal{C}$, and fix the corresponding vector $\bar{p}$. Then

$$\frac{1}{\lambda} > \frac{\sum_{j \in \mathcal{T}} a_{j-i} p_j}{p_i}$$

for each $i \in \mathcal{T}$. If we add up these inequalities over all $i \in \mathcal{T}$, we obtain

$$\frac{n}{\lambda} > \sum_{i \in \mathcal{T}} \frac{\sum_{j \in \mathcal{T}} a_{j-i} p_j}{p_i} = n + \sum_{i} \sum_{j \in \mathcal{T}, j \neq i} \frac{a_{j-i} p_j}{p_i}$$

$$= n + \frac{1}{2} \sum_{i} \sum_{j \in \mathcal{T}, j \neq i} \left(\frac{a_{j-i} p_j}{p_i} + \frac{a_{i-j} p_i}{p_j} \right)$$

$$= n + \frac{1}{2} \sum_{i} \sum_{j \in \mathcal{T}, j \neq i} a_{j-i} \left(\frac{p_j}{p_i} + \frac{p_i}{p_j} \right)$$

$$\geq n + n \sum_{j \in \mathbb{Z}^d, j \neq 0} a_j,$$

which concludes the proof (recall that $a_0 = 1$).

3.4. Stability analysis

In this section we demonstrate how our results on moment bounds allow us to obtain a stability result for an infinite network, along with a second moment bound for a stationary distribution.

We will say that the arrival rates λ_i are *periodic* if the following properties hold: (a) the values of λ_i are given for i within the rectangle $\mathcal{I} = [0, \ldots, C_1 - 1] \times \ldots \times [0, \ldots, C_d - 1]$, where $C_1, \ldots, C_d$ are fixed positive integers; (b) for any $i \in \mathbb{Z}^d$ and any $k = 1, \ldots, d$, $\lambda_{i + C_k e_k} = \lambda_i$, where e_k is the kth unit coordinate vector (with kth entry equal to 1 and all other entries equal to zero). We define periodicity of any other function of i similarly. We will say that random variables ξ_i are *i.i.d. up to periodicity* if they are all independent, and $\xi_{i + C_k e_k}$ and ξ_i have identical distribution for any i and k.

Theorem 4. *Consider periodic rates $\bar{\lambda}$. Assume that ξ_i are i.i.d. up to periodicity, and $\mathbb{E} \xi_i^3 < \infty$ for all i. Assume that there exists a periodic $\bar{\nu}$ from the set*

$$\mathcal{C} = \{\bar{\mu} : \quad \text{there exists} \quad \bar{p} \quad \text{such that} \quad \bar{\mu} \leq \bar{\psi}(\bar{p})\}$$

such that $\bar{\lambda} < \bar{\nu}$. Consider an infinite network with arrival rates $\bar{\lambda}' \leq \bar{\lambda}$, and assume in addition that the per-slot (random) number of arrivals ξ_i' is dominated by ξ_i w.p.1: $\xi_i' \leq \xi_i$. Then this infinite network is stable, and there exists a stationary regime with finite second moments $\mathbb{E} X_i^2$ of the queue lengths.

Proof of Theorem 4. Because of the monotonicity of the process, it suffices to prove the theorem for the periodic arrival rates $\bar{\lambda}$ and the arrival process $\bar{\xi}$.

Consider the auxiliary finite version of our system on the sets $\mathcal{R}_n = \{i : i_k \in \{-nC_k, \ldots, nC_k - 1\}\}$ 'wrapped around' to form a torus (the node neighbourhood structure being that of the torus). Then the conditions of the theorem, along with Lemma 1, imply stability and therefore existence of the (unique) stationary measure for the process on $\mathcal{R}_n$. Lemma 1 and Theorem 3, along with the periodicity, imply that

$$\sum_{i \in \mathcal{I}} \frac{\mathbb{E}\left((X_i^{(n)})^2\right)}{v_i^2} \leq A_1 \sum_{i \in \mathcal{I}} \frac{\mathbb{E}(X_i^{(n)})}{v_i^2} + A_2 \tag{16}$$

for some constants A_1 and A_2, where the upper index n is used to indicate the finite system on the torus $\mathcal{R}_n$. Note that A_1 and A_2 do not depend on n. This implies a second moment bound

$$\mathbb{E}\left((X_i^{(n)})^2\right) \leq C < \infty \tag{17}$$

which is uniform in n and $i \in \mathcal{I}$. Let us view each process $\bar{X}^{(n)}(\cdot)$ as a process on the entire infinite lattice $\mathbb{Z}^d$; say, by letting $X_i(\cdot) \equiv 0$ for $i \notin \mathcal{R}_n$. (We note that the node neighbourhood structure remains that of the torus, and so the process is still equivalent to that on the torus.) Correspondingly, we will view the (stationary) distributions of $\bar{X}^{(n)}$ as distributions on the entire infinite lattice $\mathbb{Z}^d$; we see from (17) that these distributions are tight (as distributions on $\mathbb{Z}^d$). Then there exists a subsequence of (stationary) processes $\bar{X}^{(n)}(\cdot)$ along which $\bar{X}^{(n)}(0) \Rightarrow \bar{X}^*$, where $\bar{X}^*$ is some proper random element (with all components being finite w.p.1), and then $\bar{X}^{(n)}(k) \Rightarrow \bar{X}^*$ for each k.

It is easy to observe that the sequence of processes $\bar{X}^{(n)}(\cdot)$ and the process $\bar{X}(\cdot)$ (which is the true infinite system process) satisfy the continuity property (from Section 1.2). This means the subsequence of (stationary) processes $\bar{X}^{(n)}(\cdot)$ and the process $\bar{X}(\cdot)$ with $\bar{X}(0)$ distributed as $\bar{X}^*$ can be coupled in a way such that $\bar{X}^{(n)}(k) \to \bar{X}(k)$ w.p.1 for each $k \geq 0$. This means that $\bar{X}(k)$ is equal in distribution to $\bar{X}^*$ for each k; i.e. we have constructed a stationary version of $\bar{X}(\cdot)$. Since $X_i^{(n)} \Rightarrow X_i^*$, Fatou's lemma and (17) imply that $\mathbb{E}X_i^2 \leq C < \infty$. This completes the proof.

Consider now a special case—a symmetric infinite system, which means that the random variables ξ_i are i.i.d. From Theorem 4 we obtain the following.

Corollary 1. *Consider the symmetric system and assume*

$$\lambda < \frac{1}{\sum_{j \in \mathbb{Z}^d} a_j}. \tag{18}$$

Assume additionally that $\mathbb{E}\xi_i^3 < \infty$. Then the system is stable, and its lower invariant measure (i.e. the stationary distribution dominated by any other) is such that $\mathbb{E}X_i^2 < \infty$.

Indeed, by Lemmas 1 and 2 and Theorem 4, the condition (18) ensures stability. By Theorem 4, $\mathbb{E}X_i^2 < \infty$ holds for *some* stationary distribution, and therefore it holds for the lower invariant measure as well.

Remark 3. The simplest case where the λ_i are not all the same is when we consider a network on the line such that the rates are periodic with period 2. Denote the different values by λ_1 and λ_2. Theorem 4 implies stability as long as there exist p_1 and p_2 such that

$$\lambda_1 < \frac{p_1}{p_1 + 2p_2}$$

and

$$\lambda_2 < \frac{p_2}{p_2 + 2p_1}.$$

One can see that by taking, for instance, $p_1 = 1$ and $p_2 = \delta > 0$, the values of λ_1 for which stability holds may be taken arbitrarily close to 1 (of course at the expense of very low values of λ_2).

Remark 4. The proof of Theorem 4 uses only the utility-maximising properties of the rates and the process continuity properties. It is clear that similar arguments may be used to demonstrate stability of infinite networks in many other cases (see the next section for some examples).

3.5. Other networks

So far in Section 3 we have only looked at a particular infinite network, motivated by [11] and [15]. In this subsection we demonstrate that the methods developed above are rather general and may be applied to a much wider class of networks. We present some examples below but would like to stress that we think this list is not exhaustive.

For simplicity, we restrict ourselves here to the case when queues are located on a one-dimensional lattice (i.e. at the integers). One can easily see that the results described below also hold for the more general location and neighbourhood structures described in the rest of Section 3.

Consider rates given by

$$\psi_i^*(\bar{x}) = \log\left(1 + \frac{x_i}{x_{i-1} + x_i + x_{i+1}}\right) = \log\left(1 + \psi_i(\bar{x})\right), \tag{19}$$

which provide a better approximation to the wireless channel capacity than the rates considered in Section 3 so far. (The ratio $x_i/(x_{i-1} + x_i + x_{i+1})$ represents the signal-to-interference ratio (SIR).) Consider the auxiliary finite version of our system on the finite set $\{-n, \ldots, n-1\}$ (with the $2n$ nodes 'wrapped around' to form a torus, the node neighborhood structure being that of the torus). Define

$$\mathcal{C}^* = \left\{\bar{\mu} : \mu_i \le \log\left(1 + \frac{p_i}{\sum_{j \in \mathcal{N}_i} p_j}\right) \quad \text{for some} \quad \bar{p} \in \mathbb{R}_+^{2n}\right\}. \tag{20}$$

The following result is immediate.

Lemma 3. *We have that*

$$\bar{\psi}^* \in \arg\max_{\bar{\mu} \in \mathcal{C}^*} \sum_i g(x_i)\tilde{h}(\mu_i),$$

where $g(x) = x^2$ *and*

$$\tilde{h}(y) = -\frac{1}{e^y - 1}.$$

Indeed, the statement of the lemma reads

$$\sum_i g(x_i)\tilde{h}(\mu_i) \le \sum_i g(x_i)\tilde{h}(\psi_i^*)$$

for all $\bar{\mu} \in \mathcal{C}^*$. This is equivalent to

$$\sum_i g(x_i)\tilde{h}\left(\log\left(1 + \frac{p_i}{p_{i-1} + p_i + p_{i+1}}\right)\right) \le \sum_i g(x_i)\tilde{h}(\tilde{\psi}_i^*)$$

for all $\bar{p} \in \mathbb{R}_+^{2n}$. Taking into account the definitions of g, $\tilde{h}$, and $\tilde{\psi}_i$, the above is equivalent to

$$- \sum_i x_i^2 \left(\frac{p_i}{p_{i-1} + p_i + p_{i+1}} \right)^{-1} \leq - \sum_i x_i^2 \psi_i^{-1},$$

which follows immediately from Lemma 1.

As the functions g and $\tilde{h}$ in the statement of Lemma 3 satisfy all the conditions imposed in Section 2, we obtain *stability for any finite network,* as long as $\bar{\lambda}$ is within $\mathcal{C}^*$. Moreover, one can also readily see that fluid limits for such a system are well-defined and stable. From here we obtain that the conclusion of Theorem 3 holds, with the same *finite-system second-moment bound*, except with $1/v_i^2$ replaced by $\tilde{h}'_i(v_i)$. The strategy of Section 3.4 can thus be followed to obtain *stability of the infinite version of the network with rates $\bar{\psi}^*$.*

More generally, assume that

$$\sum g(x_i) h(\psi_i(\bar{x})) \geq \sum g(x_i) h(\mu_i)$$

for all $\bar{\mu} \in \mathcal{C}$. We can rewrite this as

$$\sum g(x_i) \tilde{h}(f(\psi_i(\bar{x}))) \geq \sum g(x_i) \tilde{h}(f(\mu_i))$$

for all $\bar{\mu} \in \mathcal{C}$, where $\tilde{h} = h \circ f^{-1}$ and f^{-1} is the inverse of f (which we assume to be increasing). The above may be rewritten again as

$$\sum g(x_i) \tilde{h}(f(\psi_i(\bar{x}))) \geq \sum g(x_i) \tilde{h}(v_i)$$

for all $\bar{v} \in \mathcal{C}^*$, where

$$\mathcal{C}^* = f(\mathcal{C}) = \{\bar{v} : (f^{-1}(v_1), \ldots, f^{-1}(v_N)) \in \mathcal{C}\}.$$

Therefore, as long as $\tilde{h}$ is concave (which in general does not necessarily hold), all conditions of Section 2 are satisfied, and we obtain stability of a finite network with rates $f(\psi_i)$, provided $\bar{\lambda}$ is within $\mathcal{C}^*$. If, further, fluid limits of finite networks are well-defined and stable, additional steady-state moment bounds can be obtained, which in turn leads to stability of infinite networks.

It is important to note that, *if one is only interested in the stability of finite networks* with the rates considered so far in Section 3 (including the rates considered in this subsection), this may be established *without* the use of utility maximisation. In fact, for finite networks, one can demonstrate stability for a still wider class of rates, including rates of the form

$$\log \left(1 + \frac{x_i}{x_{i-1} + x_i + x_{i+1} + B} \right).$$

(The proof is a slight generalisation of that of Lemma 9 in [15].) Here $B > 0$ represents the 'background noise', and $x_i/(x_{i-1} + x_i + x_{i+1} + B)$ is the signal-to-interference-plus-noise ratio (SINR). These rates provide an even more realistic approximation of the wireless channel capacity. We are not aware of any utility-maximisation properties satisfied by these rates, yet (a slight generalisation of) [15, Lemma 9] provides stability of a finite network with these transmission rates, as long as the vector of arrival intensities belongs to the set $\mathcal{C}^*$ defined in

(20). This is thanks to the fact that fluid limits for the system with these rates are the 'same' (satisfy the same properties) as for the system with rates (19).

We emphasise again that it is the utility-maximisation property enjoyed by the rates (19) that allows us to obtain not only stability, but also steady-state moment bounds, for finite networks. The finite network moment bounds, in turn, allow us to apply the methods developed in this paper to construct a stationary distribution for an infinite network, and moreover to obtain moment bounds for this stationary distribution.

4. Stability analysis of an infinite multi-hop network

In this section we demonstrate how techniques similar to the ones we used in the single-hop case may be used to demonstrate the existence of invariant measures for infinite multi-hop networks, where, upon a service completion at a given queue, a job may either leave the system or enter the queue of a neighbouring node.

Multi-hop networks have an additional layer of difficulty as the movement of jobs between different queues complicates the dependence structure of the queue states further. Multi-hop networks are notoriously difficult to analyse, and we consider significantly stronger assumptions on the structure of the network, with strictly i.i.d. arrival processes and with symmetric routing (see [15]).

Specifically, the model is as follows. Just as in Section 3, the nodes (queues) are located on the d-dimensional lattice. The exogenous arrival processes are strictly i.i.d.; that is, the random variables ξ_i representing the numbers of new arrivals are i.i.d., with $\mathbb{E}(\xi_i) = \lambda q$ for a fixed $q \in (0, 1)$, and with $\mathbb{E}(\xi_i^2) < \infty$. The service is governed by either the algorithm (D1) or the algorithm (D2) specified in Section 3. Upon a service completion at any node i, a job either leaves the system, with probability q, or joins the queue of a neighbour of node i (i.e. a node connected to node i by one lattice edge) chosen independently at random (i.e., each neighbour is chosen with probability $(1 - q)/2^d$). All the routing decisions are taken independently of everything else. The service rates are given by

$$\psi_i(\bar{x}) = \frac{x_i}{\sum_{j \in \mathcal{N}_i} x_j},$$

where $\mathcal{N}_i$ is the neighbourhood of node i on the $\mathbb{Z}^d$ lattice, which by convention includes node i itself. (In other words, the service rates are a special case of those considered in Section 3, with the neighbourhood $\mathcal{N}_i$ of node i including specifically the neighbours in terms of the lattice, and with $a_{j-i} = 1$ for all $j \in \mathcal{N}_i$.)

Consider the auxiliary finite version of our system on the set $\mathcal{T}_n = \{i : i_k \in \{-n, \ldots, n-1\}\}$, 'wrapped around' to form a torus. (The node neighbourhood structure is that of the torus.) Then $\mathcal{T}_n$ is a finite 2^d-regular graph, and [15, Theorem 5] implies that if $\lambda < 1/(2^d + 1)$, then the system is stable. Therefore, there exists a stationary distribution of the number of messages in each queue. By symmetry, the (stationary) numbers of messages in any two queues are identically distributed; we will consider queue 0 for simplicity. Denote the stationary number of messages in queue 0 in the system on the torus $\mathcal{T}_n$ by $X^{(n)}$.

We want to emphasise that the described multi-hop process (for both the infinite system and a finite torus) is *not monotone* (unlike in the single-hop model of Section 3), and this is in fact one of the key challenges of the multi-hop system analysis. Versions of this process, however, do have continuity properties, which we will exploit, just as in the single-hop case.

Theorem 5. *Consider the multi-hop model on the torus $\mathcal{T}_n$ described above, and denote by ξ a random variable with the distribution of $\xi_i(k)$ for any i and k. Assume that $\mathbb{E}(\xi^2) < \infty$ and $\mathbb{E}(\xi) = \lambda < \frac{1}{2^d+1}$. Then*

$$\mathbb{E}(X^{(n)}) \leq \frac{\mathbb{E}(\xi^2) + 2^d(1-q)\lambda + \lambda - 2\lambda^2 q^2}{2q\left(\frac{1}{2^d+1} - \lambda\right)}. \tag{21}$$

Proof of Theorem 5. Consider the stationary version $\bar{X}^{(n)}(\cdot)$ of the Markov chain. For ease of notation, in this proof we will drop the superscript (n), and write $\bar{X}(\cdot)$ instead of $\bar{X}^{(n)}(\cdot)$. Similarly, we will write X instead of $X^{(n)}$ for a random vector distributed as $X_i^{(n)}(k)$ (for any i and k, by stationarity and symmetry). We can describe the evolution of $X_i(k)$ as

$$X_i(k+1) = X_i(k) + \xi_i(k) + \sum_{j \in \mathcal{N}_i, j \neq i} I_{ji}(k)\eta_j(k) - \eta_i(k), \tag{22}$$

where the random variable $I_{ji}(k)$ is the indicator function of the event that a message potentially leaving node j in time slot k will choose node i as its destination. For ease of notation, as we only consider a single time slot in what follows, we are going to simply write ξ_i, η_l, and I_{ji}.

As in the previous sections, note that the random variables η_l can only take values 0 and 1, and

$$\mathbb{E}(\eta_l|\bar{X}) = \mathbb{P}(\eta_l = 1|\bar{X}) = \psi_l(\bar{X}) \quad \text{almost surely.}$$

Note also that

$$\mathbb{E}(I_{ji}) = (1-q)\frac{1}{2^d}$$

for all j and i. By the stationarity of the process $\bar{X}(\cdot)$, $X_i(k)$ and $X_i(k+1)$ have the same distributions.

Moreover, this distribution has a finite mean:

$$\mathbb{E}(X_i(k+1)) = \mathbb{E}(X_i(k)) = \mathbb{E}X < \infty. \tag{23}$$

Indeed, $\mathbb{E}\xi^2 < \infty$. We also know from [15] that the system *fluid limits are stable*. (For our system, the definitions of fluid limits—called fluid sample paths in [15]—and of fluid limit stability for a finite system are given in Section 4 of [15]. The proof of fluid limit stability is in Section 6 of [15].) By results of [3], the combination of the finite second moment of the arrival process $\mathbb{E}\xi^2 < \infty$, the finite second (in fact, any positive) moment of the number of departures from any node at any time, and the fluid limit stability implies that $\mathbb{E}X < \infty$.

From (22) and (23),

$$0 = \lambda q + \sum_{j \in \mathcal{N}_i, j \neq i} (1-q)\frac{1}{2^d}\mathbb{E}(\eta_j) - \mathbb{E}(\eta_i), \tag{24}$$

where we used the fact that I_{ji} and η_j are independent. Note now that

$$\mathbb{E}(\eta_l) = \mathbb{E}(\mathbb{E}(\eta|\bar{X})) = \mathbb{E}(\psi_l(\bar{X})) = \mathbb{E}\left(\frac{X_l}{\sum_{j \in \mathcal{N}_l} X_j}\right)$$

for any l, and, because of the symmetry of the model, this quantity does not depend on l. Hence, continuing (24),

$$0 = \lambda q + 2^d(1-q)\frac{1}{2^d}\mathbb{E}(\eta_i) - \mathbb{E}(\eta_i),$$

which implies

$$\mathbb{E}(\eta_i) = \lambda \tag{25}$$

for any i. Let $A_i = \sum_{j \in \mathcal{N}_i, j \neq i} I_{ji} \eta_j$.

Assume first that $\mathbb{E}X^2 < \infty$. (We will show later in the proof how to get rid of this additional assumption.) Stationarity of the process $\bar{X}(\cdot)$ implies that $\mathbb{E}(X_i^2(k+1)) = \mathbb{E}(X_i^2(k))$, and hence, from (22),

$$
\begin{aligned}
0 &= \mathbb{E}(\xi_i^2) + \mathbb{E}(A_i^2) + \mathbb{E}(\eta_i^2) - 2\mathbb{E}(A_i \eta_i) + 2\mathbb{E}(\xi_i(A_i - \eta_i)) + 2\mathbb{E}(X_i \xi_i) + 2\mathbb{E}(X_i(A_i - \eta_i)) \\
&\leq \mathbb{E}(\xi_i^2) + \mathbb{E}(A_i^2) + \lambda + 2\lambda q \left(\mathbb{E}(A_i) - \mathbb{E}(\eta_i)\right) + 2\lambda q \mathbb{E}(X_i) + 2\mathbb{E}(X_i(A_i - \eta_i)) \\
&= \mathbb{E}(\xi_i^2) + \mathbb{E}(A_i^2) + \lambda - 2\lambda^2 q^2 + 2\lambda q \mathbb{E}(X_i) + 2\mathbb{E}(X_i(A_i - \eta_i)).
\end{aligned} \tag{26}
$$

In the derivations above we used the independence of ξ_i from all other random variables, the fact that $\mathbb{E}(\eta_i^2) = \mathbb{E}(\eta_i) = \mathbb{P}(\eta_i = 1)$, Equation (25), and finally, in the last equality, a simple calculation of $\mathbb{E}(A_i)$ already performed earlier in this proof (see (24)).

We consider some of the terms above separately. First,

$$
\mathbb{E}(A_i^2) = \mathbb{E}\left(\left(\sum_{j \in \mathcal{N}_i, j \neq i} I_{ji} \eta_j\right)^2\right) \leq 2^d \mathbb{E}\left(\sum I_{ji}^2 \eta_j^2\right) = 2^d \mathbb{E}\left(\sum I_{ji} \eta_j\right) = 2^d (1-q)\lambda, \tag{27}
$$

where we used the convexity of the function x^2, the independence of the I's and η's, and the fact that all the random variables concerned only take values 0 and 1 and therefore are equal to their squares.

Let us note now that

$$
\mathbb{E}(X_i \eta_j) = \mathbb{E}(\mathbb{E}(X_i \eta_j | \bar{X})) = \mathbb{E}(X_i \mathbb{E}(\eta_j | \bar{X})) = \mathbb{E}(X_i \psi_j) = \mathbb{E}\left(X_i \frac{X_j}{\sum_{l \in \mathcal{N}_j} X_l}\right)
$$

for any i and j. It is clear from the symmetry of the model that for any $j \in \mathcal{N}_i$, the pairs (X_i, η_j) and (X_j, η_i) have identical distributions, which implies, in particular, that

$$
\mathbb{E}(X_i \eta_j) = \mathbb{E}(X_j \eta_i).
$$

Consider now

$$
\begin{aligned}
\mathbb{E}(X_i(A_i - \eta_i)) &= \sum_{j \in \mathcal{N}_i, j \neq i} \mathbb{E}(X_i I_{ji} \eta_j) - \mathbb{E}(X_i \eta_i) = (1-q)\frac{1}{2^d} \sum_{j \in \mathcal{N}_i, j \neq i} \mathbb{E}(X_i \eta_j) - \mathbb{E}(X_i \eta_i) \\
&= (1-q)\frac{1}{2^d} \sum_{j \in \mathcal{N}_i, j \neq i} \mathbb{E}(X_j \eta_i) - \mathbb{E}(X_i \eta_i) = (1-q)\frac{1}{2^d} \sum_{j \in \mathcal{N}_i, j \neq i} \mathbb{E}(X_j \psi_i) - \mathbb{E}(X_i \psi_i) \\
&= \mathbb{E}\left(\psi_i \left((1-q)\frac{1}{2^d} \sum_{j \in \mathcal{N}_i, j \neq i} X_j - X_i\right)\right) \\
&= \mathbb{E}\left(\frac{X_i}{\sum_{j \in \mathcal{N}_i} X_j} \left((1-q)\frac{1}{2^d} \sum_{j \in \mathcal{N}_i, j \neq i} X_j + (1-q)\frac{1}{2^d} X_i - \left(1 + (1-q)\frac{1}{2^d}\right) X_i\right)\right) \\
&= (1-q)\frac{1}{2^d} \mathbb{E}(X_i) - \left(\frac{2^d + 1}{2^d} - q\frac{1}{2^d}\right) \mathbb{E}(X_i \psi_i). \tag{28}
\end{aligned}
$$

The convexity of the function $1/x$ implies that

$$\frac{1}{\sum_i X_i} \sum_i X_i \psi_i = \sum_i \frac{X_i}{\sum_i X_i} \frac{1}{(\sum_{j\in\mathcal{N}_i} X_j)/X_i} \geq \frac{\sum_i X_i}{\sum_i \sum_{j\in\mathcal{N}_i} X_j} = \frac{1}{2^d+1}$$

for any $\bar{X}$. The summation above is taken over all nodes i in our graph. We can rewrite the above as

$$\sum_i X_i \psi_i \geq \frac{1}{2^d+1} \sum_i X_i$$

and take the expectations on both sides of the inequality. Noting that, because of the symmetry of the model, $\mathbb{E}(X_i\psi_i)$ does not depend on i, we conclude that

$$\mathbb{E}(X_i\psi_i) \geq \frac{1}{2^d+1}\mathbb{E}(X_i).$$

Plugging this into (28), we obtain

$$\mathbb{E}(X_i(A_i - \eta_i)) \leq -q\frac{1}{2^d+1}\mathbb{E}(X_i). \tag{29}$$

We can now plug (27) and (29) into (26) to obtain

$$0 \leq \mathbb{E}(\xi_i^2) + 2^d(1-q)\lambda + \lambda - 2\lambda^2 q^2 + 2q\mathbb{E}(X_i)\left(\lambda - \frac{1}{2^d+1}\right),$$

which implies the statement of Theorem 5.

We now show how to remove the additional assumption $\mathbb{E}X^2 < \infty$. Consider the system with truncated arrival quantities $\xi^{(M)} = \max\{\xi, M\}$. The corresponding process is stable for any $M > 0$; we denote by $\bar{X}^{(M)}(\cdot)$ its stationary version, and by $X^{(M)}$ a generic $X_i^{(M)}(k)$. For each M, $\mathbb{E}[\xi^{(M)}]^3 < \infty$ (in fact, of course, $\mathbb{E}[\xi^{(M)}]^m < \infty$ for any $m \geq 0$). As already described above in this proof, it is proved in [15] that the system fluid limits are stable. Using again the results of [3], the combination of the finite third moment of the arrival process $\mathbb{E}[\xi^{(M)}]^3 < \infty$, the finite third (in fact, any positive) moment of the number of departures from any node at any time, and the fluid limit stability implies that $\mathbb{E}[X^{(M)}]^2 < \infty$ for any M. (In fact, $\mathbb{E}[X^{(M)}]^m < \infty$ for any M and any $m \geq 0$).

Therefore, for $\mathbb{E}X^{(M)}$, we have the upper bound (21) with λ replaced by $\mathbb{E}\xi^{(M)}$ and $\mathbb{E}\xi^2$ replaced by $\mathbb{E}[\xi^{(M)}]^2$. Choose a sequence $M \uparrow \infty$. Then $\mathbb{E}X^{(M)}$ is uniformly bounded above along this sequence, and therefore the sequence of distributions of $X^{(M)}$ is tight. We further observe that the sequence of processes $\bar{X}^{(M)}(\cdot)$ and the process $\bar{X}(\cdot)$ satisfy the continuity property (defined in Section 1.2). Proceeding analogously to the argument we used at the end of the proof of Theorem 4, we obtain that $X^{(M)} \Rightarrow X$. By Fatou's lemma, $\mathbb{E}X \leq \liminf_{M\to\infty} \mathbb{E}X^{(M)}$, and the lim inf is bounded above by the right-hand side of (21). This completes the proof.

The fact that the bound in (21) does not depend on n allows us to prove a result on an infinite-lattice multi-hop model.

Theorem 6. *Consider the multi-hop model of this section defined for the entire lattice $\mathbb{Z}^d$ and assume that $\lambda < 1/(2^d + 1)$. Then the process is stable. Moreover, there is a translation-invariant stationary distribution, for which*

$$\mathbb{E}X \leq \frac{\mathbb{E}(\xi^2) + 2^d(1-q)\lambda + \lambda - 2\lambda^2 q^2}{2q\left(\frac{1}{2^d+1} - \lambda\right)}, \tag{30}$$

where X has the distribution of $X_i(k)$ (for any i and k) in steady-state.

Remark 6. Since the multi-hop process is not monotone, constructions similar to that of [11] cannot be applied. We provide a different construction, based on continuity alone. Note that Theorem 6 does *not* claim any form of uniqueness for the stationary distribution. The uniqueness properties (among stationary distributions with finite second moments of the queue lengths) derived in [11] and in this paper for single-hop models relied in essential way on the process monotonicity.

Proof of Theorem 6. We already know that for each torus $\mathcal{T}_n$ there exists a (unique) stationary distribution of the corresponding process $\bar{X}^{(n)}(\cdot)$. (It is translation-invariant, of course, by symmetry.) We can view this distribution as a distribution on the entire lattice $\mathbb{Z}^d$. Moreover, the uniform-in-n bound (21) on the expected queue length implies that these distributions (viewed as distributions on the entire lattice) are tight. It is easy to see that the sequence of processes $\bar{X}^{(n)}(\cdot)$ and the process $\bar{X}(\cdot)$ (i.e. the 'true' infinite-lattice process) satisfy the continuity property of Section 1.2. Proceeding analogously to the argument we used in the last two paragraphs of the proof of Theorem 4, we can construct a proper stationary process $\bar{X}(\cdot)$ for the infinite system. The constructed stationary distribution of $\bar{X}(\cdot)$ is a limit of those of $\bar{X}^{(n)}(\cdot)$, and therefore translation-invariant. Finally, (30) follows from (21) and Fatou's lemma. This completes the proof.

5. Continuous-time setting

In this section we describe how our results translate to the continuous-time Markovian setting. The proofs are along the same lines as those provided in Sections 2–4 for the discrete-time setting, with minor changes. In the case of finite networks, the proofs of stability and moment bounds in continuous time (results corresponding to those in Section 2) turn out to be simplified versions of the proofs in discrete time, as, thanks to Markovian assumptions, the probability that two or more events happen in a small time interval is negligible. We thus restrict ourselves to short descriptions of the proofs, pointing out where they start to repeat the discrete-time proofs.

As far as construction of stationary distributions for infinite networks is concerned, the proofs in continuous time are based on the same continuity and monotonicity properties as their discrete-time analogues and thus repeat them almost verbatim. For this reason, for infinite networks, we present only statements, not proofs, of the results corresponding to those in Sections 3 and 4.

5.1. Finite networks: model

Assume, as before, that there are N interacting queues. Arrivals into queue i occur according to a Poisson process with a constant rate λ_i, independent of all the other processes. The instantaneous departure rate from queue i at time t, conditioned on the state $\bar{X}(t)$, is $\psi_i(\bar{X}(t))$; more precisely, the number of departures up to time t is $\Pi_i(\int_0^t \psi_i(\bar{X}(\tau))d\tau)$, where the $\Pi_i(\cdot)$ are independent unit-rate Poisson processes.

We assume that all the conditions imposed in Section 2.1 on the functions $\psi_i(\cdot)$, $i = 1, \ldots, N$, and on the arrival intensities $\bar{\lambda}$ hold. More precisely, we assume that the functions $\psi_i(\cdot)$, $i = 1, \ldots, N$, satisfy the condition (2) with functions h and g satisfying Conditions *(H)* and *(G)*, respectively; and we assume that the arrival intensities $\bar{\lambda}$ satisfy Condition (5). We will also use the functions G and F defined in (6) and (7), respectively.

5.2. Finite networks: stability analysis

Theorem 7. *For the continuous-time model defined in Section* 5.1 *such that Condition* (5) *holds, the Markov process* $\{\bar{X}(t)\}_{t\geq 0}$ *is stable.*

Proof of Theorem 7. The proof is a simplified version of that of Theorem 1. By the Lyapunov–Foster criterion, in order to show positive recurrence, it is sufficient to show that for the function F defined in (7),

$$\sum_i (\lambda_i(F(x_i+1) - F(x_i) + \psi_i(\bar{x})(F(x_i-1) - F(x_i))) < -\delta$$

for some $\delta > 0$, for values of $\bar{x}$ outside of a compact set. This may be found e.g. in [19].

Note that the expression on the left-hand side of the above may be written as

$$\sum_i h'(v_i)(\lambda_i(G(x_i+1) - G(x_i)) + \psi_i(\bar{x})(G(x_i-1) - G(x_i)),$$

which is equal to (9), and then the rest of the proof of Theorem 1 applies.

5.3. Finite networks: moment bounds

As in discrete time, once stability is established, one can use similar arguments to establish steady-state moment bounds.

Since the stationary regime exists under the conditions of Theorem 7, in this section we will write $\bar{X}$ to represent a random vector with distribution equal to that of $\bar{X}(t)$ in the stationary regime.

Theorem 8. *Assume that all conditions of Theorem* 7 *hold. Fix $\varepsilon > 0$ such that $\lambda_i < v_i - \varepsilon$ for all i. Then*

$$\sum_i h'(v_i)\mathbb{E}(g(X_i)) \leq \frac{1}{\varepsilon} \sum_i h'(v_i)\mathbb{E}(\Delta(X_i)).$$

Proof of Theorem 8. Note that the condition (4) implies that $g(x) = o(e^{ax})$ as $x \to \infty$, for any $a > 0$. This, in turn, implies that $G(x) = o(e^{ax})$ as $x \to \infty$, for any $a > 0$. Note also that, as arrival flows are given by Poisson processes, for any t and $\bar{x}$ there exists $a > 0$ such that

$$\mathbb{E}(e^{a\bar{X}(t)}|\bar{X}(0) = \bar{x}) < \infty.$$

Then $\mathbb{E}(F(\bar{X}(t))|\bar{X}(0) = \bar{x}) < \infty$ for any t and any $\bar{x}$.

Fix $\varepsilon > 0$ such that $\lambda_i < v_i - \varepsilon$ for all i. Standard Poisson-process arguments imply that

$$\frac{d(\mathbb{E}(F(\bar{X}(t))|\bar{X}(0) = \bar{x}))}{dt}$$
$$= \sum_i h'(v_i)(\lambda_i\mathbb{E}(G(X_i+1) - G(X_i)) + \mathbb{E}(\psi_i(\bar{x})(G(x_i-1) - G(x_i)))),$$

and the right-hand side of the above is equal to (10). The rest of the proof is the same as that of Theorem 2, with $\mathbb{E}\big(F(\bar{X}(k+1)) - F(\bar{X}(k))|\bar{X}(k) = \bar{x}\big)$ replaced by $\dfrac{d(\mathbb{E}(F(\bar{X}(t))|\bar{X}(0) = \bar{x}))}{dt}$.

5.4. Stability analysis of infinite single-hop networks

We now turn our attention to infinite networks. In the single-hop setting, as in the discrete case, the queues (or nodes) are assumed to be located on a d-dimensional lattice, with the service rates given by

$$\psi_i(\bar{x}) = \frac{x_i}{\sum_{j \in \mathbb{Z}^d} a_{j-i} x_j}, \tag{31}$$

where $a_0 = 1$, $a_i = a_{-i} \geq 0$ for all $i \in \mathbb{Z}^d$, and $L = \sup\{|i| : a_i > 0\} < \infty$. For each i, the nodes j within the finite set $\mathcal{N}_i = \{j | a_{j-i} > 0\}$ are called *neighbours* of i. Note that $i \in \mathcal{N}_i$.

In the continuous-time case, unlike the discrete-time case, these assumptions describe the system completely and no additional structural constructions are needed.

We say that the arrival rates λ_i are *periodic* if the following properties hold: (a) the values of λ_i are given for i within the rectangle $\mathcal{I} = [0, \ldots, C_1 - 1] \times \ldots \times [0, \ldots, C_d - 1]$, where $C_1, \ldots, C_d$ are fixed positive integers; (b) for any $i \in \mathbb{Z}^d$ and any $k = 1, \ldots, d$, $\lambda_{i+C_k e_k} = \lambda_i$, where e_k is the kth unit coordinate vector (with kth entry equal to 1 and all other entries equal to zero).

The following result is a continuous-time analogue of Theorem 4.

Theorem 9. *Consider periodic rates* $\bar{\lambda}$. *Assume that there exists a periodic* $\bar{\nu}$ *from the set*

$$\mathcal{C} = \{\bar{\mu} : \quad \text{there exists} \quad \bar{p} \quad \text{such that} \quad \bar{\mu} \leq \bar{\psi}(\bar{p})\}$$

such that $\bar{\lambda} < \bar{\nu}$. *Consider an infinite network with arrival rates* $\bar{\lambda}' \leq \bar{\lambda}$. *Then this infinite network is stable, and there exists a stationary regime with finite second moments* $\mathbb{E}X_i^2$ *of the queue lengths.*

As in the discrete-time case, we immediately obtain the following corollary.

Corollary 2. *Consider the symmetric system where* $\lambda_i = \lambda$ *for all i, and assume*

$$\lambda < \frac{1}{\sum_{j \in \mathbb{Z}^d} a_j}. \tag{32}$$

Then the system is stable and its lower invariant measure (i.e. the stationary distribution dominated by any other) is such that $\mathbb{E}X_i^2 < \infty$.

In the continuous-time setting, the corollary above proves Conjecture 1.12 in [11], along with all the implications stated in [11, Section 1.1]; in particular it proves the uniqueness of the stationary regime *with finite second moments of the queue lengths*, for the infinite symmetric network.

We note also that our results allow us to consider networks which are not necessarily symmetric, and Remark 3 applies to the continuous-time setting too, significantly expanding the range of arrival intensities for which stability of infinite networks holds.

5.5. Stability analysis of infinite multi-hop networks

We now consider infinite multi-hop networks. Just as in the single-hop setting, the nodes (queues) are located on the d-dimensional lattice.

The arrival processes are such that $\lambda_i = \lambda q$ for a fixed $q \in (0, 1)$, for all i.

Upon a service completion at any node i, a job either leaves the system, with probability q, or joins the queue of a neighbour of node i (i.e. a node connected to node i by one lattice edge)

chosen independently at random (i.e., each neighbour is chosen with probability $(1 - q)/2^d$)). All the routing decisions are taken independently of everything else. The service rates are given by

$$\psi_i(\bar{x}) = \frac{x_i}{\sum_{j \in \mathcal{N}_i} x_j},$$

where $\mathcal{N}_i$ is the neighbourhood of node i on the $\mathbb{Z}^d$ lattice, which by convention includes node i itself.

Consider the auxiliary finite version of our system on the set $\mathcal{T}_n = \{i : i_k \in \{-n, \ldots, n-1\}\}$, 'wrapped around' to form a torus. (The node neighbourhood structure is that of the torus.) Then $\mathcal{T}_n$ is a finite 2^d-regular graph, and [15, Theorem 5] implies that if $\lambda < 1/(2^d + 1)$, then the system is stable. Therefore, there exists a stationary distribution of the number of messages in each queue. By symmetry, the (stationary) numbers of messages in any two queues are identically distributed; we will consider queue 0 for simplicity. Denote the stationary number of messages in queue 0 in the system on the torus $\mathcal{T}_n$ by $X^{(n)}$.

Theorem 10. *Consider the multi-hop model on the torus $\mathcal{T}_n$ described above. Assume that $\lambda < \frac{1}{2^d+1}$. Then*

$$\mathbb{E}(X^{(n)}) \leq \frac{\lambda}{q\left(\frac{1}{2^d+1} - \lambda\right)}. \tag{33}$$

The fact that the bound in (33) does not depend on n allows us to prove a result on an infinite-lattice multi-hop model.

Theorem 11. *Consider the multi-hop model of this section defined for the entire lattice $\mathbb{Z}^d$ and assume that $\lambda < 1/(2^d + 1)$. Then the process is stable. Moreover, there is a translation-invariant stationary distribution, for which*

$$\mathbb{E}X \leq \frac{\lambda}{q\left(\frac{1}{2^d+1} - \lambda\right)}, \tag{34}$$

where X has the distribution of $X_i(k)$ (for any i and k) in steady-state.

References

[1] BONALD, T. AND MASSOULIE, L. (2001). Impact of fairness on Internet performance. *ACM SIGMETRICS Performance Evaluation Rev.* **29**, 82–91.

[2] DAI, J. G. (1995). On the positive Harris recurrence for open multiclass queueing networks: a unified approach via fluid limit models. *Ann. Appl. Prob.* **5**, 49–77.

[3] DAI, J. G. AND MEYN, S. P. (1995). Stability and convergence of moments for multiclass queueing networks via fluid limit models. *IEEE Trans. Automatic Control* **40**, 1889–1904.

[4] DE VECIANA, G., KONSTANTOPOULOS, T. AND LEE, T. (2001). Stability and performance analysis of networks supporting elastic services. *IEEE/ACM Trans. Networking* **9**, 2–14.

[5] FOSTER, F. G. (1953). On the stochastic matrices associated with certain queueing processes. *Ann. Math. Statist.* **24**, 355–360.

[6] JIANG, L. AND WALRAND, J. (2010). A distributed CSMA algorithm for throughput and utility maximization in wireless networks. *IEEE/ACM Trans. Networking* **18**, 960–972.

[7] KELLY, F. P., MAULLOO, A. K. AND TAN, D. K. H. (1998). Rate control for communication networks: Shadow prices, proportional fairness and stability. *J. Operat. Res. Soc.* **49**, 237–252.

[8] MO, J. AND WALRAND, J. (2000). Fair end-to-end window-based congestion control. *IEEE/ACM Trans. Networking* **8**, 556–567.

[9] ROBERTS, J. AND MASSOULIE, L. (2000). Bandwidth sharing and admission control for elastic traffic. *Telecommun. Systems* **15**, 185–201.

[10] RYBKO, A. N. AND STOLYAR, A. L. (1992). Ergodicity of stochastic processes describing the operation of open queueing networks. *Probl. Inf. Transm.* **28**, 199–220.

[11] SANKARARAMAN, A., BACCELLI, F. AND FOSS, S. (2018). Interference queueing networks on grids. Preprint. Available at http://arxiv.org/abs/1710.09797.

[12] SHAH, D. AND SHIN, J. (2012). Randomized scheduling algorithm for queueing networks. *Ann. Appl. Prob.* **22**, 128–171.

[13] SHAH, D., TSITSIKLIS, J. AND ZHONG, Y. (2014). Qualitative properties of α-fair policies in bandwidth-sharing networks. *Ann. Appl. Prob.* **24**, 76–113.

[14] SHNEER, S. AND STOLYAR, A. L. (2018). Stability conditions for a discrete-time decentralised medium access algorithm. *Ann. Appl. Prob.* **28**, 3600–3628.

[15] SHNEER, S. AND STOLYAR, A. L. (2018). Stability conditions for a decentralised medium access algorithm: single- and multi-hop networks. Submitted.

[16] STOLYAR, A. L. (1995). On the stability of multiclass queueing networks: a relaxed sufficient condition via limiting fluid processes. *Markov Process. Relat. Fields* **1**, 491–512.

[17] STOLYAR, A. L. (2008). Dynamic distributed scheduling in random access networks. *J. Appl. Prob.* **45**, 297–313.

[18] TASSIULAS, L. AND EPHREMIDES, A. (1992). Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Trans. Automatic Control* **37**, 1936–1948.

[19] TWEEDIE, R. (1981). Criteria for ergodicity, exponential ergodicity and strong ergodicity of Markov processes. *J. Appl. Prob.* **18**, 122–130.