

Mean-field limits for large-scale random-access networks

Citation for published version (APA):

Cecchi, F., Borst, S. C., van Leeuwen, J. S. H., & Whiting, P. A. (2016). Mean-field limits for large-scale random-access networks. *arXiv*, Article arXiv:1611.09723.

Document status and date:

Published: 29/11/2016

Document Version:

Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Mean-field limits for large-scale random-access networks

Fabio Cecchi ^{*1}, Sem C. Borst^{1,2}, Johan S.H. van Leeuwen¹ and Philip A. Whiting³

¹Eindhoven University of Technology, Eindhoven, 5600 MB, Netherlands

²Nokia Bell Labs, Murray Hill, NJ 07974, USA

³Macquarie University, North Ryde, NSW 2109, Australia

November 30, 2016

Abstract

We establish mean-field limits for large-scale random-access networks with buffer dynamics and arbitrary interference graphs. While saturated-buffer scenarios have been widely investigated and yield useful throughput estimates for persistent sessions, they fail to capture the fluctuations in buffer contents over time, and provide no insight in the delay performance of flows with intermittent packet arrivals. Motivated by that issue, we explore in the present paper random-access networks with buffer dynamics, where flows with empty buffers refrain from competition for the medium. The occurrence of empty buffers thus results in a complex dynamic interaction between activity states and buffer contents, which severely complicates the performance analysis. Hence we focus on a many-sources regime where the total number of nodes grows large, which not only offers mathematical tractability but is also highly relevant with the densification of wireless networks as the Internet of Things emerges. We exploit time scale separation properties to prove that the properly scaled buffer occupancy process converges to the solution of a deterministic initial-value problem, and establish the existence and uniqueness of the associated fixed point. This approach simplifies the performance analysis of networks with huge numbers of nodes to a low-dimensional fixed-point calculation. For the case of a complete interference graph, we demonstrate asymptotic stability, provide a simple closed-form expression for the fixed point, and prove interchange of the mean-field and steady-state limits. This yields asymptotically exact approximations for key performance metrics, in particular the stationary buffer content and packet delay distributions. The methodological framework that we develop easily extends to various model refinements as will be illustrated by several examples.

1 Introduction

1.1 Background and related work

Wireless networks are already large and complex today, and being at the heart of the so-called Internet of Things (IoT) [2], are expected to grow even denser in the future [21]. The rapid increase in wireless applications has led to an increasing demand for scarce wireless spectrum, and hence it is necessary to make the most efficient use of the limited available capacity. Obviously, when the number of nodes is large, in the hundreds or even thousands of nodes, a dedicated medium or channel cannot be assigned to each node, and nodes have to share the medium. Multiple simultaneous node transmissions on the same channel will however inevitably give rise to interference and loss of throughput. Medium access control (MAC) mechanisms are crucial to resolve this contention. However, in large networks, a centralized control mechanism is hard to implement and to maintain

*f.cecchi@tue.nl

since it would require constant status updates generating prohibitive communication overhead. For this reason the design of efficient distributed (local) MAC protocols has attracted a lot of attention.

A very popular distributed MAC mechanism is the CSMA (Carrier-Sense Multiple-Access) protocol, which is currently at the core of the IEEE 802.11 and 802.15.4 standards. Its popularity is mostly due to its simplicity and efficiency. The key feature of the CSMA protocol is that each node waits for a random back-off period before initiating a transmission. Interference is avoided since the back-off countdown is interrupted whenever potential interference is sensed, and only resumed once the medium is sensed idle again. This protocol, whilst extremely easy to understand on a local level, generates complex and interesting macroscopic network dynamics.

In the performance analysis of CSMA networks, a common assumption is the existence of an underlying graph that represents interference between the various nodes in the network. An edge between two nodes means that destructive interference is caused by simultaneous transmission. Both empirical and theoretical support for the notion of an interference graph is provided in [26, 46].

When the nodes always have packets to transmit, the network is said to be *saturated* and the macroscopic activity behavior is amenable to analysis under the assumption of an interference graph. In particular, the activity process has an elegant product-form stationary distribution [9, 29, 40]. The computation of the stationary distribution of the activity process reduces to the identification of all the subsets of nodes which may transmit simultaneously, namely the independent sets of the interference graph.

Real-life scenarios however involve *unsaturated* networks. Packets arrive at the various nodes according to exogenous processes, and buffers may drain from time to time as packets are transmitted. In particular, in IoT applications, sources are likely to generate packets only sporadically, with fairly tight delay constraints, and often have empty buffers. Since empty nodes temporarily refrain from the medium competition, the activity process is strictly intertwined with the buffer content process. In this situation, the product-form solution no longer holds [14, 40] and an exact stationary analysis does not seem tractable. Furthermore, not even stability conditions for the queueing dynamics in the nodes' buffers are known, let alone results for the stationary distribution of the buffer content process.

The analysis of unsaturated CSMA networks simplifies if certain symmetry conditions amongst the various nodes hold. An important instance is when there is a substantial number of nodes with similar traffic and placement in the network, so that the operation of one is equivalent to that of many others. More generally, nodes can be divided into classes with the symmetry conditions now applying to nodes of the same class. The asymptotic regime where the number of such nodes in each class grows to infinity, is often referred to as a *mean-field regime*. Mean-field theory originated in physics, where it is still largely used in the study of models involving a large number of interacting particles. The aggregated effect of all the other particles on any tagged particle is approximated by a single averaged effect (the mean-field), thus reducing a many-body problem to a tractable one-body problem. Mean-field analysis has since been exploited in various scientific areas, from epidemic models to queueing systems and from biology to game theory. In the context of random-access networks a mean-field regime, not only provides analytical tractability, but is also highly relevant in the context of the envisioned massive numbers of IoT devices.

A thorough survey of mean-field analysis of random-access protocols is presented in [19]. The work of Bianchi [6] is a landmark paper which assumed a propagation of chaos property to hold so as to derive tractable formulae for the key performance measures of the system. The papers surveyed in [19] mostly use mean-field theory to provide either evidence or objection for the propagation of chaos assumption of Bianchi. Among these papers, it is worth mentioning [16], where the authors investigated the existence of a global attractor for the mean-field system and provided sufficient conditions for its existence, deducing the validity of Bianchi's assumption. Further papers which deserve to be mentioned are [10, 36], where the authors exploited mean-field theory so as to obtain approximations for key performance measures of large systems. In particular [10] focuses on the characterization of the stability region, while [36] examines the throughput performance of the system. None of the above-mentioned papers considered scenarios with unsaturated buffers, with the exception of [10], which however dealt with systems evolving in discrete time and did not consider performance metrics like packet delays.

1.2 Key contributions and paper organization

In the present paper we examine the buffer dynamics in large-scale unsaturated random-access networks. Specifically, we analyze the buffer occupancy processes in a mean-field regime where the number of nodes grows large. Such a mean-field regime not only furnishes analytical tractability, but is also particularly relevant in understanding the collective behavior of massive numbers of interacting IoT sources.

The paper has two main interrelated threads. The first part in Section 4 deals with networks with general interference graphs. We prove that a suitably scaled version of the buffer occupancy processes converges in the mean-field limit to a tractable deterministic initial-value problem. We also establish necessary and sufficient conditions for the existence and uniqueness of a fixed point of the initial-value problem, and provide a characterization the fixed point as the solution of a low-dimensional equation.

The second part focuses on scenarios with a complete interference graph. We demonstrate in Section 5 global asymptotic stability of the initial-value problem, i.e., convergence to the unique fixed point from any initial state with finite mass, and then proceed to show positive recurrence of the pre-limit process and tightness. We combine these properties to prove interchange of the mean-field and steady-state limits. In Section 6 the interchange of limits is then leveraged to establish that the stationary buffer content distributions at the various nodes converge to geometric distributions while the stationary distributions of the scaled waiting time and sojourn time converge to exponential distributions. The parameters of these limiting distributions are directly expressed in terms of the fixed point of the initial-value problem. These results provide asymptotically exact approximations for the stationary waiting-time and sojourn time distributions, which in fact turn out to be remarkably accurate, even for a relatively moderate number of nodes.

Mean-field limits for various model extensions are briefly discussed in Section 7, illustrating the versatility of the methodology. In Section 8 we present some simulation experiments to illustrate the analytical results. Finally, in Section 9 we make a few concluding remarks and offer several suggestions for further research.

2 Model description

We consider a network of N nodes sharing a wireless medium according to a random medium access protocol. The various nodes may be grouped into a set of classes/clusters $\mathcal{C} = \{1, \dots, C\}$ such that nodes in the same class have the same statistical characteristics. Denote by $N_c^{(N)}$ the number of class- c nodes, with $\sum_{c \in \mathcal{C}} N_c^{(N)} = N$ and $p_{c,N} = N_c^{(N)}/N$.

Interference graph. Given a class-wise interference graph $G = (\mathcal{C}, \mathcal{E})$, two nodes interfere when they belong either to the same class or to two neighboring classes in G . A feasible class activity state can thus be represented by a vector $\omega \in \{0, 1\}^C$, with $\omega_c = 1$ if a class- c node is transmitting in state ω and $\omega_c = 0$ otherwise, and $\omega_c \omega_d = 0$ if $(c, d) \in \mathcal{E}$. Let Ω be the set of all feasible class activity states, which are in one-to-one correspondence with the independent sets of the graph G . For every class c , we define the following subsets of Ω :

$$\begin{aligned}\Omega_{-c} &\doteq \{\omega \in \Omega : \omega_c = 0, \omega_d = 0 \forall d \text{ s.t. } (c, d) \in \mathcal{E}\}, \\ \Omega_{+c} &\doteq \{\omega \in \Omega : \omega_c = 1\}.\end{aligned}$$

This means that $\omega \in \Omega_{-c}$ if and only if in the class activity state ω none of the nodes belonging to class c or to a class interfering with class c are active, while $\omega \in \Omega_{+c}$ if and only if a class- c node is active. We define the class-capacity region of the network as follows

$$\Gamma = \{\gamma \in \mathbb{R}_+^C : \exists \mathbf{a} = (a_\omega)_{\omega \in \Omega} \text{ s.t. } a_\omega \geq 0, \sum_{\omega \in \Omega} a_\omega \leq 1, \sum_{\omega \in \Omega} a_\omega \omega = \gamma\}.$$

As an illustration, Figure 1 shows a square interference graph G with $C = 4$ classes of nodes.

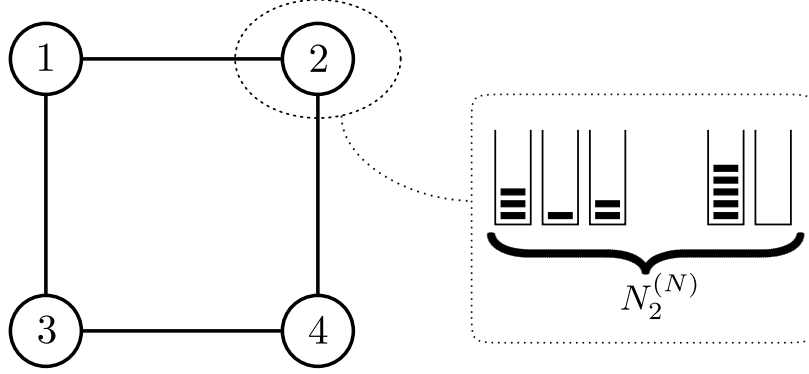


Figure 1: A square interference graph with $C = 4$.

The activity states are

$$\Omega = \{(0, 0, 0, 0), (1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1), (1, 0, 0, 1), (0, 1, 1, 0)\},$$

and the class-capacity region is given by

$$\Gamma = \{\boldsymbol{\gamma} \in \mathbb{R}_+^4 : \max\{\gamma_1, \gamma_4\} + \max\{\gamma_2, \gamma_3\} \leq 1\}.$$

Taking $c = 1$ the sets $\Omega_{-c}, \Omega_{+c} \subset \Omega$ are

$$\begin{aligned} \Omega_{-1} &= \{(0, 0, 0, 0), (0, 0, 0, 1)\} \\ \Omega_{+1} &= \{(1, 0, 0, 0), (1, 0, 0, 1)\}. \end{aligned}$$

Network dynamics. Packets arrive at the various class- c nodes as independent Poisson processes of rate $\lambda_c^{(N)} \doteq \lambda_c/N_c^{(N)}$. Once a class- c node obtains access to the medium, it transmits one packet, which takes an exponentially distributed time with parameter $\mu_c^{(N)} \doteq \mu_c$. In between two consecutive transmissions a class- c node must back-off for an exponentially distributed time period with parameter $\nu_c^{(N)} \doteq \nu_c/N_c^{(N)}$. Note that a node refrains from the back-off competition whenever its buffer is empty. Also, the back-off period of a class- c node is suspended (frozen) when the medium is occupied by an interfering node, i.e., the class activity state does not belong to Ω_{-c} .

Define the queue length process $\mathbf{Q}^{(N)}(t) = (Q_{c,k}^{(N)}(t))_{c \in \mathcal{C}, k \in \mathbb{N}}$, where $Q_{c,k}^{(N)}(t)$ represents the number of packets in the buffer of the k -th node belonging to class c at time t excluding any possible packet in transmission, and define by $\mathbf{Y}^{(N)}(t)$ the class activity process on the state space Ω . Observe that $(\mathbf{Q}^{(N)}(t), \mathbf{Y}^{(N)}(t))$ evolves as a Markov process.

A population process description. To ease the notation we write $N_c = N_c^{(N)}$ from now on. Due to the class symmetry, the nodes within a class are statistically exchangeable, and the system state may be described in terms of the numbers of nodes belonging to the same class and with the same number of packets in the buffer. In particular, define the population process $\tilde{\mathbf{X}}^{(N)}(t) = (\tilde{X}_{c,n}^{(N)}(t))_{c \in \mathcal{C}, n \in \mathbb{N}}$, where

$$\tilde{X}_{c,n}^{(N)}(t) = \frac{1}{N_c} \sum_{k=1}^{N_c} \mathbb{1}\{Q_{c,k}^{(N)}(t) = n\}.$$

The process $(\tilde{\mathbf{X}}^{(N)}(t), \mathbf{Y}^{(N)}(t))$ is itself Markovian with state space $E \times \Omega$, where

$$E \doteq \chi^{\mathcal{C}}, \quad \chi = \{\mathbf{x} = (x_n)_{n \in \mathbb{N}} : x_n \in [0, 1], \forall n \in \mathbb{N}\}. \quad (1)$$

Observe that the population process lies in $E^1 \subseteq E$ where

$$E^1 = \{\mathbf{x} = (x_{c,n})_{c \in \mathcal{C}, n \in \mathbb{N}} : \sum_{n=0}^{\infty} x_{c,n} = 1, \forall c \in \mathcal{C}\}. \quad (2)$$

The possible transitions for the process $(\tilde{\mathbf{X}}^{(N)}(t), \mathbf{Y}^{(N)}(t))$ may be described as follows:

- A packet arrives at a class- c node having n packets in its buffer: this happens at rate $\lambda_c^{(N)}$ times the number of class- c nodes in state n , i.e., $\lambda_c^{(N)} N_c \tilde{X}_{c,n}^{(N)} = \lambda_c \tilde{X}_{c,n}^{(N)}$, and generates the transition

$$\begin{cases} \tilde{X}_{c,n}^{(N)} \rightarrow \tilde{X}_{c,n}^{(N)} - \frac{1}{N_c} \\ \tilde{X}_{c,n+1}^{(N)} \rightarrow \tilde{X}_{c,n+1}^{(N)} + \frac{1}{N_c}, \end{cases}$$

i.e.,

$$(\tilde{\mathbf{X}}^{(N)}, \mathbf{Y}^{(N)}) \rightarrow (\tilde{\mathbf{X}}^{(N)} + \frac{1}{N_c}(\mathbf{e}_{c,n+1} - \mathbf{e}_{c,n}), \mathbf{Y}^{(N)}),$$

where $\mathbf{e}_{c,n} \in E^1$ has all 0 entries except a 1 in position (c, n) .

- A transmission is completed by a class- c node: this happens at rate $\mu_c^{(N)} = \mu_c$ and only if a class- c node is transmitting, i.e., $\mathbf{Y}^{(N)} \in \Omega_{+c}$. The transition generated is the following

$$(\tilde{\mathbf{X}}^{(N)}, \mathbf{Y}^{(N)}) \rightarrow (\tilde{\mathbf{X}}^{(N)}, \mathbf{Y}^{(N)} - \mathbf{e}_c).$$

- A back-off is completed by a class- c node having $n > 0$ packets in its buffer: this happens only if the class activity state allows class- c nodes to back-off, i.e., $\mathbf{Y}^{(N)} \in \Omega_{-c}$, and at rate $\nu_c^{(N)}$ times the number of class- c nodes in state n , i.e., $\nu_c^{(N)} N_c \tilde{X}_{c,n}^{(N)} = \nu_c \tilde{X}_{c,n}^{(N)}$, and generates the transition

$$\begin{cases} \mathbf{Y}^{(N)} \rightarrow \mathbf{Y}^{(N)} + \mathbf{e}_c \\ \tilde{X}_{c,n}^{(N)} \rightarrow \tilde{X}_{c,n}^{(N)} - \frac{1}{N_c} \\ \tilde{X}_{c,n-1}^{(N)} \rightarrow \tilde{X}_{c,n-1}^{(N)} + \frac{1}{N_c}, \end{cases}$$

i.e.,

$$(\tilde{\mathbf{X}}^{(N)}, \mathbf{Y}^{(N)}) \rightarrow (\tilde{\mathbf{X}}^{(N)} + \frac{1}{N_c}(\mathbf{e}_{c,n-1} - \mathbf{e}_{c,n}), \mathbf{Y}^{(N)} + \mathbf{e}_c).$$

Preliminary results for saturated scenario. As mentioned earlier, we focus on an unsaturated network where nodes with empty buffers refrain from competition for the medium. For later purposes, it is convenient to also consider a related saturated network where each class behaves as a single node *always* competing for the medium, and backing-off at rate $\eta_c \nu_c$, where $\boldsymbol{\eta} = (\eta_c)_{c \in \mathcal{C}} \in \mathbb{R}_+^{\mathcal{C}}$. The transmission times of the node associated with class c are exponentially distributed with parameter μ_c .

For compactness, denote $\sigma_c = \nu_c / \mu_c$, and for any $\boldsymbol{\eta} \in \mathbb{R}_+^{\mathcal{C}}$,

$$Z(\boldsymbol{\omega}; \boldsymbol{\eta}) \doteq \prod_{c \in \mathcal{C}} (\eta_c \sigma_c)^{\omega_c}, \quad \boldsymbol{\omega} \in \Omega,$$

and

$$Z(\boldsymbol{\eta}) \doteq \sum_{\boldsymbol{\omega} \in \Omega} Z(\boldsymbol{\omega}; \boldsymbol{\eta}).$$

The activity process in this fictitious saturated network has a product-form stationary distribution [9, 40, 42]

$$\pi(\boldsymbol{\omega}; \boldsymbol{\eta}) = \frac{Z(\boldsymbol{\omega}; \boldsymbol{\eta})}{Z(\boldsymbol{\eta})}, \quad \boldsymbol{\omega} \in \Omega. \quad (3)$$

Define now the throughput function $\boldsymbol{\theta} : \mathbb{R}_+^{\mathcal{C}} \rightarrow \Gamma$, where

$$\theta_c(\boldsymbol{\eta}) \doteq \sum_{\boldsymbol{\omega} \in \Omega} \omega_c \pi(\boldsymbol{\omega}; \boldsymbol{\eta}) = \sum_{\boldsymbol{\omega} \in \Omega_{+c}} \pi(\boldsymbol{\omega}; \boldsymbol{\eta}) = \pi(\Omega_{+c}; \boldsymbol{\eta}) \quad (4)$$

represents the fraction of time that the node associated with class c is active in the saturated network.

As proved in [27, 41], the throughput map θ is globally invertible, i.e., for any achievable throughput vector $\gamma \in \text{int}(\Gamma)$ there exists a unique vector $\eta(\gamma) \in \mathbb{R}_+^C$ such that

$$\theta(\eta(\gamma)) = \gamma. \quad (5)$$

Intuitively, the c -th coordinate of $\eta(\gamma)$ represents by what factor the back-off rate at node c needs to accelerate/decelerate in order for the target throughput vector γ to be achieved in stationarity.

3 Overview of the main results

In this section we provide an overview of the main results of the paper, along with an interpretation and high-level discussion of their ramifications, before presenting the proofs in the subsequent sections.

In the first part, we analyze networks with general class-based interference graphs. We consider the sequence of population processes $\tilde{\mathbf{X}}^{(N)}(t)$ as the number of nodes in the system increases, and rigorously establish its weak convergence to $\mathbf{x}(t)$, the solution of a tractable deterministic initial-value problem as stated in the next theorem.

Theorem 1. *Assume $\tilde{\mathbf{X}}^{(N)}(0) \xrightarrow{N \rightarrow \infty} \mathbf{x}^\infty \in E$ and $\lim_{N \rightarrow \infty} p_{c,N} = p_c > 0$ for every $c \in \mathcal{C}$. Then the sequence of processes*

$$(\tilde{\mathbf{X}}^{(N)}(Nt)) \in D_E[0, \infty)$$

has a continuous limit $\mathbf{x}(t) \in C_E[0, \infty)$ which is determined by the unique solution of the initial-value problem

$$\frac{d\mathbf{x}(t)}{dt} = H(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}^\infty, \quad (6)$$

where the function $H(\cdot)$ is defined by

$$H(\mathbf{x}) = \sum_{c \in \mathcal{C}} \frac{1}{p_c} \left(\lambda_c \sum_{n=0}^{\infty} x_{c,n} e_{c,n}^{c,n+1} + \pi_{\mathbf{x}_0}(\Omega_{-c}) \nu_c \sum_{n=1}^{\infty} x_{c,n} e_{c,n}^{c,n-1} \right)$$

and $\pi_{\mathbf{x}_0}(\Omega_{-c}) = \sum_{\omega \in \Omega_{-c}} \pi_{\mathbf{x}_0}(\omega)$.

The major difficulty in the analysis of unsaturated networks arises from the correlation between the population process $\tilde{\mathbf{X}}^{(N)}(t)$ and the activity process $\mathbf{Y}^{(N)}(t)$. A key observation in the proof of Theorem 1 is that these processes evolve on different time scales, i.e., the population process evolves N times slower than the activity process. Hence, in the mean-field regime, the rapidly changing activity process “converges to an instantaneous measure on the activity states” determined by the current fraction of queues which are empty. This is the reason why the initial-value problem (6) does not involve explicitly the activity process and the evolution of the population process at time t depends only on $\mathbf{x}(t)$. Specifically, fraction of class- c nodes with $n \geq 1$ packets in the buffer “increases” at rate λ_c and “decreases” at rate $\nu_c \pi_{\mathbf{x}_0}(\Omega_{-c})$, where the measure $\pi_{\mathbf{x}_0}(\Omega_{-c})$ represents the limiting “instantaneous measure” on the activity states that allows class- c nodes to back-off. This argument based on a “separation of time-scales” is highly flexible and applies to more general and complicated versions of random-access networks as will be illustrated in Section 7.

The initial-value problem (6) is certainly easier to analyze than the population process in a network with a finite number of nodes. Theorem 2 states that, under certain necessary and sufficient conditions, there exists a unique equilibrium point \mathbf{x}^* for the initial-value problem (6), and provides its characterization in terms of the load of the network.

Theorem 2. *If $\rho \in \text{int}(\Gamma)$ and $\xi = \eta(\rho) < \mathbf{e}$, where $\eta(\cdot)$ is defined in (5), then $\mathbf{x}^* = (x_{c,n}^*)_{c \in \mathcal{C}, n \in \mathbb{N}}$, with*

$$x_{c,n}^* = (1 - \xi_c) \xi_c^n, \quad (7)$$

is the unique equilibrium point with $H(\mathbf{x}^) = \mathbf{0}$ in E^1 of the initial-value problem (6).*

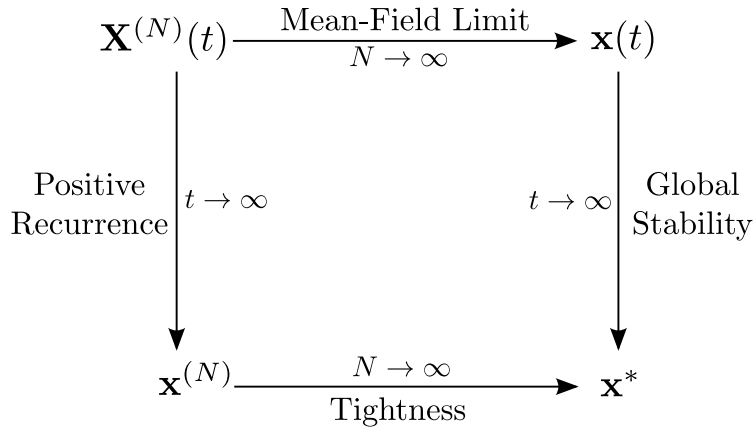


Figure 2: Diagram for the interchange of limits.

Specifically, the condition $\boldsymbol{\rho} \in \text{int}(\Gamma)$ is necessary and ensures the load $\boldsymbol{\rho}$ of the system is sustainable. On the other hand, the condition $\boldsymbol{\xi} = \boldsymbol{\eta}(\boldsymbol{\rho}) < \mathbf{e}$ ensures that the desired throughput vector $\boldsymbol{\rho}$ can be achieved by a feasible back-off vector. Although Theorem 2 only concerns the scaled population process, combined with Theorem 1 it yields that the stationary distribution of the associated class activity process is given by $\pi_{\mathbf{x}_0^*}(\cdot)$ in the limit. In other words, in the limit the stationary distribution of the class activity process is the same as in a scenario where the aggregate back-off rate of class c is a *constant fraction* ξ_c of the nominal back-off rate ν_c . This is consistent with the fact that ξ_c is the stationary fraction of class- c nodes that have non-empty buffers and compete for the medium, and $\boldsymbol{\xi}$ will therefore in the sequel be referred to as the vector of *activity factors*. The equation $\boldsymbol{\xi} = \boldsymbol{\eta}(\boldsymbol{\rho})$ reflects that the activity factors must be such that each class c is active a fraction ρ_c of the time.

The second part of the paper focuses on deriving asymptotically exact results for the key performance measures of the system. For a rigorous treatment, we focus on the scenario where all the classes mutually interfere, i.e., the interference graph $G = (\mathcal{C}, \mathcal{E})$ is complete, in which case

$$\Gamma = \{\boldsymbol{\gamma} \in \mathbb{R}^{\mathcal{C}} : \sum_{c \in \mathcal{C}} \gamma_c \leq 1\}.$$

In this framework, the nodes of every class are allowed to back-off only when the activity process is in the idle state. We will show that the vector of activity factors $\boldsymbol{\xi}$ in Theorem 2 simplifies to

$$\xi_c = \frac{\lambda_c}{\nu_c (1 - \sum_{d \in \mathcal{C}} \frac{\lambda_d}{\mu_d})},$$

so that the condition $\boldsymbol{\xi} = \boldsymbol{\eta}(\boldsymbol{\rho}) < \mathbf{e}$ can be expressed in explicit form as

$$\max_{c \in \mathcal{C}} \frac{\lambda_c}{\nu_c} < 1 - \sum_{d \in \mathcal{C}} \frac{\lambda_d}{\mu_d}, \quad (8)$$

which forces the right-hand side to be positive, i.e., $\boldsymbol{\rho} \in \text{int}(\Gamma)$. When condition (8) applies, Theorem 3 holds and asymptotically characterizes the stationary distribution of the population process as the equilibrium point of the initial-value problem (6).

Theorem 3. *The sequence of stationary distributions $(\mathbf{x}^{(N)})_{N \in \mathbb{N}} \subseteq E^1$ converges to $\mathbf{x}^* \in E^1$, i.e.,*

$$\mathbf{X}^{(N)}(t) \xrightarrow{t \rightarrow \infty} \mathbf{x}^{(N)} \xrightarrow{N \rightarrow \infty} \mathbf{x}^*. \quad (9)$$

In order to prove Theorem 3, we show that the interchange of limits displayed in Figure 2 holds. Specifically, our methodology reduces the derivation of the stationary distribution of an intractable N -dimensional Markov process to the computation of the equilibrium point of a low-dimensional fixed-point equation.

Theorem 3 is exploited so as to obtain approximations for the performance measures of the system. In particular, denote by $Q_c^{(N)}$ the stationary queue length at a class- c node, and by $W_c^{(N)}$ and $S_c^{(N)}$ the stationary waiting time and sojourn time of a packet at a class- c node, respectively. Assume condition (8) holds.

Theorem 4. *For every class c ,*

$$Q_c^{(N)} \Rightarrow \bar{Q}_c, \quad \frac{\lambda_c}{N_c} S_c^{(N)} \Rightarrow \bar{S}_c, \quad \frac{\lambda_c}{N_c} W_c^{(N)} \Rightarrow \bar{W}_c,$$

where

$$\bar{Q}_c \sim \text{Geo}(\xi_c), \quad \bar{W}_c \sim \bar{S}_c \sim \text{Exp}\left(\frac{1 - \xi_c}{\xi_c}\right), \quad c \in \mathcal{C}.$$

Theorem 5. *For every class c ,*

$$\mathbb{E}[Q_c^{(N)}] \rightarrow \mathbb{E}[\bar{Q}_c], \quad \frac{\lambda_c}{N_c} \mathbb{E}[W_c^{(N)}] \rightarrow \mathbb{E}[\bar{W}_c], \quad \frac{\lambda_c}{N_c} \mathbb{E}[S_c^{(N)}] \rightarrow \mathbb{E}[\bar{S}_c].$$

These theorems show that, given any node in the system, its stationary queue length converges in distribution and expectation to a geometric random variable, while the scaled waiting time and sojourn time of a packet in its buffer converge to exponentially distributed random variables. These limits can be used as approximations for the performance of finite-node systems, which are provably exact as the number of nodes in the network grows large, and remain very accurate when the size of the network is moderate.

4 Mean-field limit in a general framework

In this section we present the proofs of Theorems 1 and 2 as stated in Section 3.

4.1 Derivation of the mean-field limit

Denote

$$\mathbf{X}^{(N)}(t) \doteq \tilde{\mathbf{X}}^{(N)}(Nt)$$

as the fluid version of the population process. Theorem 1 describes the weak limit of the sequence of processes $\{\mathbf{X}^{(N)}(t)\}_{N \in \mathbb{N}}$, which is known as the mean-field limit of the population process. In order to prove Theorem 1, we need to go through various steps which we briefly outline here.

1. In Section 4.1.1, a Poisson representation for the prelimit process

$$(\mathbf{X}^{(N)}(t), \mathbf{Y}^{(N)}(t))$$

is provided. This representation allows us to easily derive the dynamics of the process on a fluid time scale.

2. In Section 4.1.2 a detour is needed in order to ensure the existence of the limiting process. Specifically, the class activity process $\mathbf{Y}^{(N)}$ is replaced by a measure-valued activity process $\boldsymbol{\alpha}^{(N)}$, so that the evolution of the model on fluid time scale can be equivalently described via the process

$$(\mathbf{X}^{(N)}(t), \boldsymbol{\alpha}^{(N)}(t)). \tag{10}$$

Finally, the sequence of processes (10) is shown to weakly converge as $N \rightarrow \infty$ to a limiting process

$$(\mathbf{x}(t), \boldsymbol{\alpha}(t)). \tag{11}$$

3. In Section 4.1.3 we characterize the limiting process (11).
4. As a last step, in Section 4.1.4, we describe how the limiting process $\alpha(t)$ at time t is uniquely determined by the value $\mathbf{x}(t)$. Hence a self-contained deterministic initial-value problem governing the behavior of $\mathbf{x}(t)$ is obtained.

In preparation for the analysis, we briefly recall few useful properties of the topological space where the population process evolves. The sample paths of $\mathbf{X}^{(N)}(t)$ lie in $D_E[0, \infty)$, i.e., the set of the cadlag functions from $[0, \infty)$ in $E = \chi^{\mathcal{C}}$, where the space $E \subseteq \prod_{c \in \mathcal{C}} \mathbb{R}^\infty$ is defined in (1). Define on E the metric $\rho(\cdot, \cdot)$ such that

$$\rho((\mathbf{x}_c^{(1)})_{c \in \mathcal{C}}, (\mathbf{x}_c^{(2)})_{c \in \mathcal{C}}) = \sum_{c \in \mathcal{C}} \rho_1(\mathbf{x}_c^{(1)}, \mathbf{x}_c^{(2)}),$$

where $\mathbf{x}_c = (x_{c,n})_{n \in \mathbb{N}}$ and ρ_1 is a metric on \mathbb{R}^∞ , i.e.,

$$\rho_1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = \sum_{n \in \mathbb{N}} 2^{-n} \frac{|x_n^{(1)} - x_n^{(2)}|}{1 + |x_n^{(1)} - x_n^{(2)}|}.$$

Note that $(\mathbb{R}^\infty, \rho_1)$ is separable and complete, and χ is compact in $(\mathbb{R}^\infty, \rho_1)$ as each coordinate lies in $[0, 1]$, see [7, page 219]. From [8, Section M6, pages 240-241], the following lemma is therefore obtained.

Lemma 1. *The subset $E = \chi^{\mathcal{C}}$ is complete, separable, and compact under the product topology induced by the metric ρ in the space $\prod_{c \in \mathcal{C}} \mathbb{R}^\infty$.*

In the subsequent analysis we will work with $D_E[0, \infty)$ using the metric d , an exponentially weighted version of the Skorohod metric, see [20, page 117]. Under this metric $D_E[0, \infty)$ is complete and separable as (E, ρ) is itself complete and separable, see [20, Theorem 5.6, page 121].

4.1.1 Unit Poisson process representation

In this subsection we describe the fluid-time marginals of the prelimit Markov population process $(\tilde{\mathbf{X}}^{(N)}, \mathbf{Y}^{(N)})$ in terms of infinite sequences of unit Poisson processes. For each $c \in \mathcal{C}$, and then for each $n \in \mathbb{N}_0$, the random quantity $\mathcal{N}_{A,c,n}(t)$ determines the class- c arrivals to queues with n packets in the buffer during the time interval $[0, t)$. We similarly define $\mathcal{N}_{B,c,n}(t)$ for each $c \in \mathcal{C}$, $n \in \mathbb{N}$, to determine the back-offs and finally $\mathcal{N}_{T,c}(t)$ determines the process of transmissions for each $c \in \mathcal{C}$. These processes are supposed mutually independent and defined on a common probability space $(\Omega_F, \mathcal{F}^F, \mathbb{P})$. We also define the centered versions of these processes to be $\tilde{\mathcal{N}}_{A,c,n}(s) = \mathcal{N}_{A,c,n}(s) - s$, $c \in \mathcal{C}$, $n \in \mathbb{N}_0$, with corresponding definitions for the back-offs and the transmissions. These are martingales with respect to their natural filtrations.

To obtain the Poisson process representation, we first fix the initial conditions, which are deterministic. Realizations of the above unit Poisson processes are then drawn and the (unscaled) population random variables $\tilde{\mathbf{X}}, \mathbf{Y}$ are obtained as solutions to the following equations, where fluid time is t . That is,

$$\tilde{\mathbf{X}}^{(N)}(t) = \tilde{\mathbf{X}}^{(N)}(0) + \sum_{c \in \mathcal{C}} \sum_{n=0}^{\infty} \frac{e_{c,n+1}}{N_c^{(N)}} \tilde{A}_{c,n}(t) + \sum_{c \in \mathcal{C}} \sum_{n=1}^{\infty} \frac{e_{c,n-1}}{N_c^{(N)}} \tilde{D}_{c,n}(t)$$

$$\mathbf{Y}^{(N)}(t) = \mathbf{Y}^{(N)}(0) - \sum_{c \in \mathcal{C}} e_c \tilde{V}_c(t) + \sum_{c \in \mathcal{C}} \sum_{n=1}^{\infty} e_c \tilde{D}_{c,n}(t),$$

where

$$\begin{aligned}\tilde{A}_{c,n}(t) &= \mathcal{N}_{A,c,n} \left(\lambda_c^{(N)} \int_0^t N_c^{(N)} \tilde{X}_{c,n}^{(N)}(s) ds \right), \\ \tilde{D}_{c,n}(t) &= \int_0^t \mathbb{1}\{\mathbf{Y}^{(N)}(s^-) \in \Omega_{-c}\} d\mathcal{N}_{B,c,n} \left(\nu_c^{(N)} \int_0^s N_c^{(N)} \tilde{X}_{c,n}^{(N)}(u) du \right), \\ \tilde{V}_c(t) &= \int_0^t \mathbb{1}\{\mathbf{Y}^{(N)}(s^-) \in \Omega_{+c}\} d\mathcal{N}_{T,c} \left(\mu_c^{(N)} \int_0^s du \right),\end{aligned}$$

and $e_{c,n}^{d,m} = e_{d,m} - e_{c,n}$ with $e_{c,n} \in E$ denoting the vector having zero entries everywhere but a 1 in its (c, n) coordinate.

As in the discussion in [32, Section 2.2], the solutions to these equations can be obtained by fixing the sample paths of the Poisson processes $\mathcal{N}_{A,c,n}$, $\mathcal{N}_{B,c,n}$, $\mathcal{N}_{T,c}$ above. It can then be seen inductively that the sample paths of the process $(\tilde{\mathbf{X}}^{(N)}(t), \mathbf{Y}^{(N)}(t))$ lie in $D_{E^1 \times \Omega}$. Note that the total rate of the above Poisson processes is bounded above by $\sum_{c \in \mathcal{C}} (\lambda_c + \nu_c + \mu_c) < \infty$. Hence there can only be finitely many jumps in any finite interval $[0, t]$ w.p. 1 and the sample paths are by construction piecewise constant cadlag functions.

To confirm that the distribution corresponds to that of the multi-dimensional process defined in Section 2, let \mathcal{G}_t denote the history up to time t , that is

$$\mathcal{G}_t \doteq \sigma \left((\tilde{\mathbf{X}}(s), \mathbf{Y}(s)) : 0 \leq s \leq t \right).$$

It is straightforward to show that the infinitesimal transition rates obtained by conditioning on \mathcal{G}_t and the current state, yield those of the multi-dimensional process detailed in Section 2, see also [32]. But these conditional rates are the defining property of the process and so the representation follows the same statistical law. Equivalently, as the rates are bounded, the representation could also be verified by considering the uniformized version of the process and then showing that the underlying jump chain matches that of Section 2.

We now rewrite these equations in a more compact form, working componentwise. The N superscript has been omitted for conciseness.

$$X_{c,n}(t) = X_{c,n}(0) + A_{c,n-1}(t) - A_{c,n}(t) + D_{c,n+1}(t) - D_{c,n}(t), \quad (12)$$

$$Y_c(Nt) = Y_c(0) + N_c \sum_{n=1}^{\infty} D_{c,n}(t) - V_c(t), \quad (13)$$

where

$$A_{c,n}(t) = \frac{1}{N_c} \mathcal{N}_{A,c,n} \left(N \lambda_c \int_0^t X_{c,n}(s) ds \right), \quad (14)$$

$$D_{c,n}(t) = \frac{1}{N_c} \int_0^t \mathbb{1}\{\mathbf{Y}(Ns^-) \in \Omega_{-c}\} d\mathcal{N}_{B,c,n} \left(N \nu_c \int_0^s X_{c,n}(u) du \right), \quad (15)$$

$$V_c(t) = \int_0^t \mathbb{1}\{\mathbf{Y}(Ns^-) \in \Omega_{+c}\} d\mathcal{N}_{T,c} \left(N \mu_c \int_0^s du \right), \quad (16)$$

for each $n \in \mathbb{N}_0$ and $c \in \mathcal{C}$.

Equation (12) expresses the change in the fraction of class- c nodes with n packets, i.e., arrivals to a queue with $n - 1$ packets or departures from a queue with $n + 1$ packets increment this fraction whereas arrivals or departures to component n result in a decrement. Equation (13) describes the dynamics of the class activity process and Y_c is incremented whenever there is a back-off from any non-empty class- c queue, and returns to inactivity when the corresponding packet has been transmitted. Equation (14) defines the sequence of arrivals by fluid time t , the (stochastic) intensity is proportional to the fraction of queues for each component, with the convention $A_{c,-1}(t) \equiv 0$ for all t . Equation (15) expresses the back-offs as an integral of a previsible process over a Poisson process,

again with intensity varying according to the fraction of the corresponding component. Note that $D_{c,0}(t) = 0$ as no back-offs can occur from empty queues. Finally the transmission process (16) is again an integral of a previsible process over a scaled version of the original Poisson process.

The expressions (12)-(16) are preferably written in the following martingale form:

$$\begin{aligned} \mathbf{X}(t) = & \mathbf{X}(0) + \sum_{c \in \mathcal{C}} \sum_{n=0}^{\infty} e_{c,n}^{c,n+1} \frac{\lambda_c}{p_{c,N}} \int_0^t X_{c,n}(s) ds \\ & + \sum_{c \in \mathcal{C}} \sum_{n=1}^{\infty} e_{c,n}^{c,n-1} \frac{\nu_c}{p_{c,N}} \int_0^t \mathbb{1}\{\mathbf{Y}(Ns^-) \in \Omega_{-c}\} X_{c,n}(s) ds + \Xi(t), \end{aligned} \quad (17)$$

$$\begin{aligned} Y_c(Nt) = & Y_c(0) + N\nu_c \int_0^t \mathbb{1}\{\mathbf{Y}(Ns^-) \in \Omega_{-c}\} (1 - X_{c,0}(s)) ds \\ & - N\mu_c \int_0^t \mathbb{1}\{\mathbf{Y}(Ns^-) \in \Omega_{+c}\} ds + \psi_c(t), \end{aligned} \quad (18)$$

where

$$\Xi_{c,n}(t) = \frac{1}{N_c} [M_{A,c,n-1}(t) - M_{A,c,n}(t) + M_{B,c,n+1}(t) - M_{B,c,n}(t)], \quad (19)$$

$$\psi_c(t) = \sum_{n=1}^{\infty} M_{B,c,n}(t) - M_{T,c}(t), \quad (20)$$

$$M_{A,c,n}(t) = \tilde{\mathcal{N}}_{A,c,n} \left(N\lambda_c \int_0^t X_{c,n}(s) ds \right), \quad (21)$$

$$M_{B,c,n}(t) = \int_0^t \mathbb{1}\{\mathbf{Y}(Ns^-) \in \Omega_{-c}\} d\tilde{\mathcal{N}}_{B,c,n} \left(N\nu_c \int_0^s X_{c,n}(u) du \right), \quad (22)$$

$$M_{T,c}(t) = \int_0^t \mathbb{1}\{\mathbf{Y}(Ns^-) \in \Omega_{+c}\} d\tilde{\mathcal{N}}_{T,c} \left(N\mu_c \int_0^s du \right). \quad (23)$$

The following technical lemma guarantees that the processes $M_{A,n}$, $M_{B,n}$, M_T , and Ξ are martingales, a property that we will use throughout the following analysis. The proof of the lemma is given in Appendix B.1. For ease of exposition we will consider the single-class case only, but the extension of the argument to the multi-class case is straightforward.

Lemma 2. *Given the common probability space $(\Omega_F, \mathcal{F}^F, \mathbb{P})$, there is a filtration \mathbf{F} defined on it, such that $M_{A,n}, n \in \mathbb{N}_0$ and $M_{B,n}, n \in \mathbb{N}$, are sequences of locally square integrable martingales adapted to \mathbf{F} together with M_T which is also a locally square integrable \mathbf{F} -martingale. In addition Ξ is a \mathcal{L}^2 - \mathbf{F} martingale.*

Observe now equations (17) and (18). Note that on a fluid time scale there are order N transitions in the activity process, which thus evolves much faster than the fluid population process. Hence, in the mean-field regime, the activity process reaches its stationary distribution before the state of the population process changes. Intuitively, the stationary distribution of the activity process is influenced by the current state of the population process, in particular only a fraction $(1 - x_{c,0}) \in [0, 1]$ of class- c nodes has packets and so competes for the medium access. Hence, using the notation introduced in (3), given that the population process is in state $\mathbf{x} \in E$, the probability for the activity process to be in state $\omega \in \Omega$ in stationarity is

$$\pi_{\mathbf{x}_0}(\omega) \doteq \pi(\omega; \mathbf{e} - \mathbf{x}_0).$$

4.1.2 Representing \mathbf{Y} as a random measure

Since the weak limit of the sequence $\mathbf{Y}^{(N)}$ does not exist in D , we introduce the cumulative time process $\alpha^{(N)} \in C_{\mathbb{R}_+^c, \uparrow}[0, \infty)$, which has components $\alpha_\omega^{(N)}(t)$ for $\omega \in \Omega$, as the cumulative fluid time

spent in state ω by the activity process $\mathbf{Y}^{(N)}$ in the interval $[0, t]$. That is, for each $\omega \in \Omega, t \geq 0$,

$$\alpha_\omega^{(N)}(t) \doteq \frac{1}{N} \int_0^{Nt} \mathbb{1}\{\mathbf{Y}^{(N)}(s) = \omega\} ds. \quad (24)$$

Observe that $\sum_\omega \alpha_\omega^{(N)}(t) = t$, $\alpha_\omega^{(N)}(0) = 0$, and $\alpha_\omega^{(N)}(t)$ is a continuous, increasing and unit Lipschitz function for every $\omega \in \Omega$. Indeed, $\boldsymbol{\alpha}^{(N)}$ is unit Lipschitz under the sup norm metric.

Thus $\mathbf{Y}^{(N)}$ induces a sequence of random measures $\gamma^{(N)}$, as determined by $\boldsymbol{\alpha}^{(N)} \in C_{\mathbb{R}_+^c, \uparrow}[0, \infty)$, on the space of increasing, nonnegative functions with initial value 0. This holds since the mapping from $D_\Omega[0, \infty) \rightarrow C_{\mathbb{R}_+^c, \uparrow}[0, \infty)$ as defined by (24) is continuous. For any realization, $\boldsymbol{\alpha}^{(N)}$ the measure on $\mathcal{B}([0, \infty) \times \Omega)$ is obtained by integrating the derivatives of $\boldsymbol{\alpha}^{(N)}$. In fact, $\alpha_\omega^{(N)}(t)$ has left and right derivatives everywhere taking values in $\{0, 1\}$. Note that $\alpha_\omega^{(N)}(t)$ increases at rate 1 if $\mathbf{Y}^{(N)}(t) = \omega$, and 0 otherwise. Taking left derivatives along a sample path, it follows that

$$\dot{\alpha}_\omega^{(N)}(t) = \begin{cases} \mathbb{1}\{\mathbf{Y}^{(N)}(Nt) = \omega\}, & \text{if } \mathbf{Y}^{(N)}(Nt) \text{ is continuous at } Nt, \\ \mathbb{1}\{\mathbf{Y}^{(N)}(Nt^-) = \omega\}, & \text{otherwise.} \end{cases}$$

We may therefore rewrite (17) by substituting the left derivatives of $\boldsymbol{\alpha}^{(N)}$ in the intensity function (compensator),

$$\begin{aligned} \mathbf{X}^{(N)}(t) &= \mathbf{X}^{(N)}(0) + \sum_{c \in \mathcal{C}} \sum_{n=0}^{\infty} e_{c,n}^{c,n+1} \frac{\lambda_c}{p_{c,N}} \int_0^t X_{c,n}^{(N)}(s) ds \\ &\quad + \sum_{c \in \mathcal{C}} \sum_{n=1}^{\infty} e_{c,n}^{c,n-1} \frac{\nu_c}{p_{c,N}} \int_0^t \sum_{\omega \in \Omega_{-c}} \dot{\alpha}_\omega(s) X_{c,n}^{(N)}(s) ds + \boldsymbol{\Xi}^{(N)}(t). \end{aligned} \quad (25)$$

We now proceed to show that a weak limit for the sequence of joint processes

$$(\mathbf{X}^{(N)}(t), \boldsymbol{\alpha}^{(N)}(t))$$

exists. This follows from Proposition 1, whose proof is given in Appendix B.2.

Proposition 1. *The sequence of joint measures $(\mathbf{X}^{(N)}(t), \boldsymbol{\alpha}^{(N)}(t))$ is relatively compact.*

Observe that, as a consequence of [20, Theorem 2.4, page 107] (see also [7, Exercise 6, page 41]), in order to prove Proposition 1 it suffices to show the tightness for the marginals, i.e., the tightness of $\boldsymbol{\alpha}^{(N)}$ and of $\mathbf{X}^{(N)}$. The proof of the tightness of $\boldsymbol{\alpha}^{(N)}$ is quite simple, but establishing the tightness of $\mathbf{X}^{(N)}$ requires substantial effort. In particular, the argument in Appendix B.2 relies on Lemma 1 and on the sequence of processes $\mathbf{X}^{(N)}$ being *asymptotically Lipschitz in probability*, an easy-to-verify property which we detail separately in Appendix A.

4.1.3 Mean-field limit characterization

We just established the existence of $(\mathbf{x}(t), \boldsymbol{\alpha}(t))$, i.e., a weak limit for the sequence $(\mathbf{X}^{(N)}(t), \boldsymbol{\alpha}^{(N)}(t))$. In this section, as a further step towards Theorem 1 we prove an intermediate mean-field limit, which will be further developed in the next section. Specifically we describe the steps needed to obtain the following result.

Proposition 2. *Given $(\mathbf{x}, \boldsymbol{\alpha})$ a weak limit for the sequence $(\mathbf{X}^{(N)}, \boldsymbol{\alpha}^{(N)})$ then \mathbf{x} and $\boldsymbol{\alpha}$ satisfy the differential equation*

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{x}(0) + \sum_{c \in \mathcal{C}} \frac{\lambda_c}{p_c} \sum_{n=0}^{\infty} e_{c,n}^{c,n+1} \int_0^t x_{c,n}(s) ds \\ &\quad + \sum_{c \in \mathcal{C}} \frac{\nu_c}{p_c} \sum_{n=1}^{\infty} e_{c,n}^{c,n-1} \int_0^t x_{c,n}(s) \sum_{\omega \in \Omega_{-c}} \dot{\alpha}_\omega(s) ds. \end{aligned}$$

Observe that Equation (25) determines $\mathbf{X}^{(N)}(t)$ as in the following sum, of an initial term plus sample paths lying in D :

$$\mathbf{X}^{(N)}(t) = \mathbf{X}^{(N)}(0) + \mathbf{I}_A^{(N)}(t) + \mathbf{I}_B^{(N)}(t) + \mathbf{\Xi}^{(N)}(t), \quad (26)$$

where

$$\begin{aligned} \mathbf{I}_A^{(N)}(t) &\doteq \sum_{c \in \mathcal{C}} \frac{\lambda_c}{p_{c,N}} \sum_{n=0}^{\infty} e_{c,n}^{c,n+1} \int_0^t X_{c,n}^{(N)}(s) ds \\ \mathbf{I}_B^{(N)}(t) &\doteq \sum_{c \in \mathcal{C}} \frac{\nu_c}{p_{c,N}} \sum_{n=1}^{\infty} e_{c,n}^{c,n-1} \int_0^t \sum_{\omega \in \Omega_{-c}} \dot{\alpha}_\omega(s) X_{c,n}^{(N)}(s) ds. \end{aligned}$$

Furthermore if $\mathbf{x}, \boldsymbol{\alpha}$ are weak limits of $\mathbf{X}^{(N)}, \boldsymbol{\alpha}^{(N)}$ then we make corresponding definitions for $\mathbf{I}_A, \mathbf{I}_B$. Further define $\mathbf{I}^{(N)}(t) \doteq \mathbf{X}^{(N)}(0) + \mathbf{I}_A^{(N)}(t) + \mathbf{I}_B^{(N)}(t)$ and let \mathbf{I} be the corresponding limit.

In order to prove Proposition 2, we need to go through the following steps:

- (a) Show that the weak limits of $\mathbf{X}^{(N)}(t)$ and $\mathbf{I}^{(N)}(t)$ coincide. This is established through Proposition 3 and the following remark.
- (b) Derive the weak limit of $\mathbf{I}^{(N)}(t)$ as in Proposition 4 and Lemma 3.

Focus on step (a), we apply the Continuous Mapping Theorem, see [7, Theorem 5.1, page 30] to the weak limits for the terms on the right hand side of (26) as follows. First recall [8, Theorem 3.1, page 27] which states that if $(\mathbf{I}^{(N)}, \mathbf{X}^{(N)})$ are random elements of $D_E[0, \infty) \times D_E[0, \infty)$ (defined on some common probability space) and taking values in a metric space with metric d then if $\mathbf{I}^{(N)} \Rightarrow \mathbf{I}$ and it also holds that $d(\mathbf{I}^{(N)}, \mathbf{X}^{(N)}) \Rightarrow 0$, then it follows that $\mathbf{X}^{(N)} \Rightarrow \mathbf{I}$. To apply the above result to (26), we will take d to be the Skorohod metric on $D_E[0, \infty)$ as defined in [20, Section 3.5]. In order to show that $\mathbf{X}^{(N)}$ and $\mathbf{I}^{(N)}$ have the same weak limit, it will be enough to establish the following proposition which is proved in Appendix B.3.

Proposition 3. *Given any $\eta > 0$ and $T > 0$,*

$$\lim_{N \rightarrow \infty} \mathbb{P} \left\{ \sup_{t \in [0, T]} \rho(\mathbf{\Xi}^{(N)}(t), \mathbf{0}) > \eta \right\} = 0.$$

To see that Proposition 3 implies that $d(\mathbf{X}^{(N)}, \mathbf{I}^{(N)}) \Rightarrow 0$, first note that we may choose $\lambda(s)$ to be $\iota(s)$ the identity map in [20, Equation (5.2), page 117]. In this case the term on the right hand side of [20, Equation (5.3)] becomes

$$q(\mathbf{X}^{(N)}(t \wedge u), \mathbf{I}^{(N)}(t \wedge u)) \leq \sup_{t \leq u} \rho(\mathbf{\Xi}^{(N)}(t), \mathbf{0}),$$

where $q(\cdot, \cdot) \doteq \rho(\cdot, \cdot) \wedge 1$. Finally, given $\zeta > 0$, choose T_ζ so that $e^{-T_\zeta} < \zeta/2$ and set $u := T_\zeta$ in the above upper bound which then shows the required result.

We now continue with step (b). Observe that $\mathbf{I}^{(N)}(t)$ is a sequence of continuous paths. We will establish convergence of this latter sequence in $C_E[0, \infty)$ under the uniform metric using the Continuous Mapping Theorem. This establishes convergence in D (the Skorohod topology relativized to C coincides with the uniform topology).

Given a subsequence $(\mathbf{X}^{(N_m)}, \boldsymbol{\alpha}^{(N_m)}) \Rightarrow (\mathbf{x}, \boldsymbol{\alpha})$ (without loss of generality we will take this subsequence to be the whole sequence), we obtain the corresponding weak limit of $\mathbf{I}^{(N)}$ provided that the given mapping is continuous.

Proposition 4. *Consider a sequence $(\mathbf{X}^{(N)}, \boldsymbol{\alpha}^{(N)}) \in D_E[0, \infty) \times L[0, \infty)$ such that $(\mathbf{X}^{(N)}, \boldsymbol{\alpha}^{(N)}) \rightarrow (\mathbf{x}, \boldsymbol{\alpha})$ in the product topology. Then for all $t > 0$,*

$$\int_0^t X_{c,n}^{(N)}(s) ds \rightarrow \int_0^t x_{c,n}(s) ds, \quad (27)$$

and

$$\sum_{\omega \in \Omega_{-c}} \int_0^t X_{c,n}^{(N)}(s) \dot{\alpha}_\omega^{(N)}(s) ds \rightarrow \sum_{\omega \in \Omega_{-c}} \int_0^t x_{c,n}(s) \dot{\alpha}_\omega(s) ds. \quad (28)$$

The proof of Proposition 4 is given in Appendix B.4. Note that $X_{c,n}^{(N)}(s) \in D_{\mathbb{R}}[0, \infty)$ is Lebesgue integrable over any finite interval and so the first integral is well-defined. Since the $\alpha_\omega^{(N)}$ are unit Lipschitz and increasing, the integrals in (28) also exist. Proposition 4 only establishes pointwise convergence to the continuous functions on the right hand side of (27) and of (28). However since these functions are non-negative, non-decreasing as well as unit Lipschitz continuous, then given any fixed $T > 0$ it follows that convergence is uniform in the interval $[0, T]$, for each component (c, n) as a consequence of the following lemma which is proved in Appendix B.5.

Lemma 3. *Suppose that $u^{(N)}, u \in L_{\uparrow}[0, T]$ with a common Lipschitz constant, which we take to be unity. Suppose that $\forall t \in [0, T], u^{(N)}(t) \rightarrow u(t)$ i.e. the sequence of functions converges pointwise. Then convergence is uniform, i.e. given $\epsilon > 0, \exists N_\epsilon$ such that $\sup_{t \in [0, T]} |u^{(N)}(t) - u(t)| < \epsilon, \forall N \geq N_\epsilon$.*

However uniform convergence of the components implies the same for the sequence so that

$$\begin{aligned} \sup_{t \in [0, T]} \rho(I_A^{(N)}(t), I_A(t)) &\rightarrow 0, \\ \sup_{t \in [0, T]} \rho(I_B^{(N)}(t), I_B(t)) &\rightarrow 0. \end{aligned}$$

But since T is arbitrary this implies convergence in $C_E[0, \infty)$ and hence in $D_E[0, \infty)$. This yields

$$\mathbf{I}^{(N)}(t) \Rightarrow \mathbf{I}(t).$$

4.1.4 Weak limit characterization of α

Proposition 2 determines any weak limit \mathbf{x} as the solution to a differential equation for which the corresponding limit occupancy measure α is given. We now proceed to characterize this latter limit in terms of the stationary measure of the activity process.

Recall the properties satisfied by the processes $\alpha_\omega^{(N)}$. In particular, for every $t \geq 0$,

$$\sum_{\omega \in \Omega} \alpha_\omega^{(N)}(t) = t, \quad \sum_{\omega \in \Omega} \dot{\alpha}_\omega^{(N)}(t) = 1. \quad (29)$$

We begin by substituting α into (18) to obtain that

$$\begin{aligned} Y_c^{(N)}(Nt) - Y_c^{(N)}(0) + N\mu_c \sum_{\omega \in \Omega_{+c}} \int_0^t \dot{\alpha}_\omega^{(N)}(s) ds \\ - N\nu_c \sum_{\omega \in \Omega_{-c}} \int_0^t \dot{\alpha}_\omega^{(N)}(s) (1 - X_{c,0}^{(N)}(s)) ds \end{aligned}$$

is an \mathbf{F} -martingale. Dividing by N and taking weak limits for (\mathbf{x}, α) , we obtain that

$$\int_0^t \nu_c (1 - X_{c,0}^{(N)}(s)) \sum_{\omega \in \Omega_{-c}} \dot{\alpha}_\omega^{(N)}(s) ds - \int_0^t \mu_c \sum_{\omega \in \Omega_{+c}} \dot{\alpha}_\omega^{(N)}(s) ds \Rightarrow 0, \quad (30)$$

for every $t \geq 0$.

We apply to (30) the Continuous Mapping Theorem as explained in the discussion before Proposition 3, and use the limits established in Proposition 4, to obtain

$$0 = \int_0^t \nu_c (1 - x_{c,0}(s)) \sum_{\omega \in \Omega_{-c}} \dot{\alpha}_\omega(s) ds - \int_0^t \mu_c \sum_{\omega \in \Omega_{+c}} \dot{\alpha}_\omega(s) ds, \quad (31)$$

for every $t \geq 0$.

Corollary 1. *The function $\alpha_\omega(t) \in C[0, \infty)$ is differentiable almost everywhere and*

$$\dot{\alpha}_\omega(t) = \frac{\prod_{c \in \mathcal{C}} \left(\frac{\nu_c(1-x_{c,0})}{\mu_c} \right)^{\omega_c}}{\sum_{\omega' \in \Omega} \prod_{c \in \mathcal{C}} \left(\frac{\nu_c(1-x_{c,0})}{\mu_c} \right)^{\omega'_c}}. \quad (32)$$

Proof. Thanks to the fundamental theorem of calculus applied to (31), it holds that

$$0 = \nu_c(1-x_{c,0}(t)) \sum_{\omega \in \Omega_{-c}} \dot{\alpha}_\omega(t) - \mu_c \sum_{\omega \in \Omega_{+c}} \dot{\alpha}_\omega(t), \quad \text{a.e.} \quad (33)$$

Hence, at every time t , we have to look for $(z_\omega^t)_{\omega \in \Omega}$, solutions of

$$\nu_c(1-x_{c,0}(t)) \sum_{\omega \in \Omega_{-c}} z_\omega^t = \mu_c \sum_{\omega \in \Omega_{+c}} z_\omega^t, \quad \forall c \in \mathcal{C}, \quad (34)$$

where in view of (29)

$$\sum_{\omega \in \Omega} z_\omega^t = 1, \quad z_\omega^t \geq 0.$$

Observe that, z_ω^t corresponds to the stationary fraction of time spent by the activity process in state $\omega \in \Omega$ in the saturated network where node c transmits at rate μ_c and back-offs at rate $\nu_c(1-x_{c,0}(t))$, i.e.,

$$\dot{\alpha}_\omega(t) = z_\omega^t = \pi(\omega; \mathbf{e} - \mathbf{x}_0(t)) = \pi_{\mathbf{x}_0(t)}(\omega) = \frac{Z(\omega; \mathbf{e} - \mathbf{x}_0(t))}{Z(\mathbf{e} - \mathbf{x}_0(t))},$$

which is equivalent to (32). In particular, existence and uniqueness of the stationary distribution $\pi_{\mathbf{x}_0(t)}(\omega)$ is ensured whenever $\mathbf{x}_0(t) \leq \mathbf{e}$.

Hence, $\dot{\alpha}_\omega(t) = z_\omega^t$ is well-defined and, since an increasing unit Lipschitz function starting from 0 is determined by its derivatives, the equality holds for all t , determining α_ω . Observe that α_ω exists everywhere since it coincides with the integral of a continuous function. \square

Having characterized the weak limit of α , we obtain Theorem 1 by substituting (32) in Proposition 2.

4.2 Analysis of the mean-field limit

In the previous section we proved Theorem 1, establishing the convergence of $\mathbf{X}^{(N)}(t)$ to $\mathbf{x}(t) \in C_E[0, \infty)$, the solution of the initial-value problem (6). In this section we present the proof of Theorem 2. Specifically, we show that a solution to (6) exists and is unique. Moreover, we provide a general condition yielding the existence of a unique equilibrium point \mathbf{x}^* .

Transient behavior. Observe that by definition $\mathbf{X}^{(N)}(t) \in E^1$ for every $N \in \mathbb{N}$ and $t \geq 0$. Such property is carried over as $N \rightarrow \infty$, indeed if $\mathbf{x}(0) \in E^1$, then $\mathbf{x}(t) \in E^1$ for every $t \geq 0$. Hence, the set E^1 is positive invariant for the initial-value problem (6). It is easy to show that the function $H(\cdot)$ defined in (6) is Lipschitz continuous in the metric space (E^1, ρ) , so that uniqueness and existence of the solution of (6) follows.

Limiting behavior. In order to prove Theorem 2, we observe that given that an equilibrium point \mathbf{x}^* exists, it must satisfy the relation $H(\mathbf{x}^*) = \mathbf{0}$ component-wise, i.e., for every $c \in \mathcal{C}$

$$\lambda_c(x_{c,n-1}^* - x_{c,n}^*) + \nu_c \pi_{\mathbf{x}^*}(\Omega_{-c})(x_{c,n+1}^* - x_{c,n}^*) = 0,$$

for all $n > 0$, and

$$-\lambda_c x_{c,0}^* + \nu_c \pi_{\mathbf{x}^*}(\Omega_{-c}) x_{c,1}^* = 0.$$

Thus

$$\lambda_c x_{c,n}^* = \nu_c \pi_{\mathbf{x}^*}(\Omega_{-c}) x_{c,n+1}^*, \quad \forall n \in \mathbb{N},$$

yielding

$$x_{c,n}^* = \xi_c^n x_{c,0}^*, \quad \xi_c = \frac{\lambda_c}{\nu_c \pi_{\mathbf{x}_0^*}(\Omega_{-c})} \quad (35)$$

for every $n \in \mathbb{N}$. In other words, any equilibrium point \mathbf{x}^* must satisfy the above geometric relation. In addition, such an equilibrium point \mathbf{x}^* lies in E^1 if and only if $\xi_c = 1 - x_{c,0}^* < 1$ for every $c \in \mathcal{C}$, i.e., $\boldsymbol{\xi} = \mathbf{e} - \mathbf{x}_0^* < \mathbf{e}$. Hence, thanks to (35), the uniqueness of $\boldsymbol{\xi}$ yields the same for \mathbf{x}^* .

According to the product-form solution in (3) and the definition of throughput (4), it holds that

$$\theta_c(\boldsymbol{\xi}) = \pi_{\mathbf{x}_0^*}(\Omega_{+c}) = \sum_{\omega \in \Omega_{+c}} \pi(\omega; \boldsymbol{\xi}).$$

In order for \mathbf{x}^* to be an equilibrium point, $\boldsymbol{\xi}$ must solve

$$\theta_c(\boldsymbol{\xi}) = \rho_c, \quad \forall c \in \mathcal{C},$$

i.e., the fraction of time that class c is active in equilibrium must be equal to its load. Noting that $\boldsymbol{\rho} \in \text{int}(\Gamma)$, the global invertibility of the throughput map (5) implies that $\boldsymbol{\xi}$ has to satisfy

$$\boldsymbol{\xi} = \boldsymbol{\eta}(\boldsymbol{\rho}).$$

Observe that the existence and uniqueness of $\boldsymbol{\xi}$ follows from the global invertibility property of the map $\boldsymbol{\theta}(\cdot)$. We conclude that \mathbf{x}^* as in (7) is the unique equilibrium point of the initial-value problem (6), and that $\mathbf{x}^* \in E^1$ if and only if $\boldsymbol{\eta}(\boldsymbol{\rho}) < \mathbf{e}$. (As a side-remark, we mention that the initial-value problem would not have any equilibrium point, if the assumption $\boldsymbol{\rho} \in \text{int}(\Gamma)$ were not satisfied. This makes sense since the latter condition is necessary for the queue lengths to be stable.)

As an example, consider the square network displayed in Figure 1 and assume that

$$\max\{\rho_1, \rho_4\} + \max\{\rho_2, \rho_3\} < 1,$$

i.e., $\boldsymbol{\rho} \in \text{int}(\Gamma)$. In order to identify the unique equilibrium point of the initial-value problem (6), the following system of equations has to be solved:

$$\xi_c = \frac{\lambda_c}{\nu_c \pi_{\mathbf{x}_0^*}(\Omega_{-c})}, \quad c = 1, 2, 3, 4, \quad (36)$$

where

$$\begin{aligned} \pi_{\mathbf{x}_0^*}(\Omega_{-1}) &= \frac{1}{Z(\boldsymbol{\xi})} (1 + \sigma_4 \xi_4), & \pi_{\mathbf{x}_0^*}(\Omega_{-2}) &= \frac{1}{Z(\boldsymbol{\xi})} (1 + \sigma_3 \xi_3), \\ \pi_{\mathbf{x}_0^*}(\Omega_{-3}) &= \frac{1}{Z(\boldsymbol{\xi})} (1 + \sigma_2 \xi_2), & \pi_{\mathbf{x}_0^*}(\Omega_{-4}) &= \frac{1}{Z(\boldsymbol{\xi})} (1 + \sigma_1 \xi_1), \end{aligned}$$

and

$$Z(\boldsymbol{\xi}) = 1 + \sum_{c=1}^4 \sigma_c \xi_c + \sigma_1 \sigma_4 \xi_1 \xi_4 + \sigma_2 \sigma_3 \xi_2 \xi_3.$$

The equilibrium point exists if and only if $\boldsymbol{\xi} < \mathbf{e}$. Methods for solving these kinds of equations are discussed more extensively in [15, 41].

In this section, we proved the convergence of the population process towards the solution of a deterministic initial-value problem. Necessary and sufficient conditions for the existence and uniqueness of an equilibrium point have been established. In the following section we aim to exploit the equilibrium point so as to obtain an asymptotically exact approximation for the stationary performance measures of the system. The analysis will focus on the case of a complete interference graph where all classes mutually interfere.

5 The complete interference graph case

The second part of the paper focuses on the analysis of important stationary performance measures such as the queue length, the waiting-time, and the sojourn-time distributions. In particular, we will build upon the equilibrium point \mathbf{x}^* described in Theorem 2 to obtain approximations which are asymptotically exact in the mean-field regime for the case of a complete interference graph and accurate even in moderately large networks. Specifically, we will prove Theorem 3, i.e., we will show that for any initial state $\mathbf{x}^\infty \in E^1$ with finite mass the solution $\mathbf{x}(t)$ of the initial value problem satisfies $\mathbf{x}(t) \rightarrow \mathbf{x}^*$ as $t \rightarrow \infty$, which implies

$$\lim_{N \rightarrow \infty} \lim_{t \rightarrow \infty} \mathbb{P}\{\rho(\mathbf{X}^{(N)}(t), \mathbf{x}^*) > \epsilon\} = 0, \quad \forall \epsilon > 0. \quad (37)$$

Due to the class symmetry, relation (37) implies that

$$\lim_{N \rightarrow \infty} \lim_{t \rightarrow \infty} \mathbb{P}\{Q_{c,k}^{(N)}(t) = n\} = x_{c,n}^*, \quad \forall c \in \mathcal{C}, k = 1, \dots, N_c.$$

In order to establish (37), we need several steps.

- In Section 5.1 we will show that

$$\lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} \mathbb{P}\{\rho(\mathbf{X}^{(N)}(t), \mathbf{x}^*) > \epsilon\} = 0, \quad \forall \epsilon > 0.$$

- Equation (37) is well-posed only if the population process in the finite system possesses a stationary distribution, hence conditions for the positive recurrence of the prelimit processes are established in Section 5.2.
- In Section 5.3 the interchange of limits is then proved so that (37) follows.

Throughout Sections 5 and 6 we assume the interference graph G to be complete and condition (8) to hold. Simulation results suggest that (37) remains valid even in the case of a general interference graph. In the course of the analysis we will point out the difficulties that would need to be resolved in order to obtain a rigorous proof in that case.

For a complete interference graph, it holds that

$$\theta_c(\boldsymbol{\xi}) = \pi(\mathbf{e}_c; \boldsymbol{\xi}) = \frac{\xi_c \sigma_c}{1 + \sum_{d \in \mathcal{C}} \xi_d \sigma_d}.$$

Thus the equilibrium point \mathbf{x}^* has the following closed-form expression, and (8) is indeed necessary and sufficient so as to guarantee $\boldsymbol{\xi} < \mathbf{e}$.

Corollary 2. *The unique equilibrium point in E^1 of the initial-value problem (6) is given by $\mathbf{x}^* = (x_{c,n}^*)_{c \in \mathcal{C}, n \in \mathbb{N}}$, where*

$$x_{c,n}^* = (1 - \xi_c) \xi_c^n, \quad \xi_c = \frac{\lambda_c}{\nu_c (1 - \sum_{d \in \mathcal{C}} \frac{\lambda_d}{\mu_d})}. \quad (38)$$

5.1 Global stability of \mathbf{x}^*

We begin by observing that, when the interference graph is complete, Theorem 1 specializes as follows.

Corollary 3. *Assume $G = (\mathcal{C}, \mathcal{E})$ to be complete, $\mathbf{X}^{(N)}(0) \xrightarrow{N \rightarrow \infty} \mathbf{x}^\infty \in E^1$ and $\lim_{N \rightarrow \infty} p_{c,N} = p_c > 0$ for every $c \in \mathcal{C}$. Then the sequence of Markov processes $(\mathbf{X}^{(N)}(Nt))_{N \geq 1} \in D_{E^1}[0, \infty)$ has a continuous limit $\mathbf{x}(t) \in C_{E^1}[0, \infty)$ which is determined by the unique solution of the initial-value problem*

$$\frac{d\mathbf{x}(t)}{dt} = H(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}^\infty, \quad (39)$$

where the function $H(\cdot)$ is defined by

$$H(\mathbf{x}) = \sum_{c \in \mathcal{C}} \frac{1}{p_c} \left(\lambda_c \sum_{n=0}^{\infty} x_{c,n} e_{c,n}^{c,n+1} + \pi_{\mathbf{x}_0}^b \nu_c \sum_{n=1}^{\infty} x_{c,n} e_{c,n}^{c,n-1} \right),$$

and $\pi_{\mathbf{x}_0}^b = \frac{1}{1 + \sum_{c \in \mathcal{C}} \frac{\nu_c}{\mu_c} (1 - x_{c,0})}$.

In this section we will show that for any initial state $\mathbf{x}^\infty \in E^1$ with finite mass the solution $\mathbf{x}(t)$ of the initial value problem satisfies $\mathbf{x}(t) \rightarrow \mathbf{x}^*$ as $t \rightarrow \infty$, which implies

$$\lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} \mathbb{P}\{\rho(\mathbf{X}^{(N)}(t), \mathbf{x}^*) > \epsilon\} = 0, \quad \forall \epsilon > 0, \quad (40)$$

i.e., the fixed point $\mathbf{x}^* \in E^1$ as described in Corollary 2 is a stable equilibrium point of the initial-value problem described in Corollary 3.

Denote by $B(\mathbf{x}^*) \subseteq E^1$ the *basin of attraction* of \mathbf{x}^* , where $\mathbf{x} \in B(\mathbf{x}^*)$ if and only if for any open neighborhood $N(\mathbf{x}^*)$ of \mathbf{x}^* , there exists a finite time T such that the solution of the initial-value problem (39) starting from \mathbf{x} hits $N(\mathbf{x}^*)$ and remains there for every $t > T$. Of course $\mathbf{x}^* \in B(\mathbf{x}^*)$. Moreover we say that \mathbf{x}^* is *locally stable* if and only if $\{\mathbf{x}^*\} \neq B(\mathbf{x}^*)$. In this section we explicitly characterize $B(\mathbf{x}^*)$, and we show that it contains a nonempty open subset of E^1 .

Here we briefly outline the steps necessary in order to prove (40):

- In Section 5.1.1 we introduce the concept of mass of a population vector as the average number of packets in the buffer of the various nodes. Observe that the mass of the equilibrium point \mathbf{x}^* is finite and the mass of a solution $\mathbf{x}(t)$ of (39) is a continuous Lipschitz function.
- Then, in Section 5.1.2, we show that stochastic dominance is an invariant property for the initial-value problem (39), see Proposition 5. This property allows us to couple different solutions of (39) and will play a key role in the proof of (40).
- We leverage the previous steps in Section 5.1.3 so as to show that, if $\mathbf{x}(0)$ has finite mass, $\mathbf{x}(t) \rightarrow \mathbf{x}^*$. This is proved in Proposition 6 and implies that $\mathbf{x}(0) \in B(\mathbf{x}^*)$. We will say that \mathbf{x}^* is a *globally stable* equilibrium point.

5.1.1 The concept of mass

Definition 1. Given $\mathbf{x} \in E^1$, we define the class- c mass of \mathbf{x} as

$$m_c(\mathbf{x}) = \sum_{n=0}^{\infty} n x_{c,n},$$

and the total mass as $m(\mathbf{x}) = \sum_{c \in \mathcal{C}} m_c(\mathbf{x})$. Denote $\mathbf{m}(\mathbf{x}) = (m_c(\mathbf{x}))_{c \in \mathcal{C}}$.

Given a solution $\mathbf{x}(t) \in C_{E^1}[0, \infty)$ of (39) such that $m_c(\mathbf{x}(0)) < \infty$, we observe that $m_c(\mathbf{x}(t))$ is differentiable in t and the following differential equation holds

$$\begin{aligned} \dot{m}_c(\mathbf{x}(t)) &= \frac{1}{p_c} \sum_{n=1}^{\infty} n \left(\lambda_c (x_{c,n-1}(t) - x_{c,n}(t)) - \nu_c \pi_{\mathbf{x}_0(t)}^b (x_{c,n}(t) - x_{c,n+1}(t)) \right) \\ &= \frac{1}{p_c} \left(\lambda_c \sum_{n=0}^{\infty} x_{c,n}(t) - \nu_c \pi_{\mathbf{x}_0(t)}^b \sum_{n=1}^{\infty} x_{c,n}(t) \right) \\ &= \frac{1}{p_c} (\lambda_c - \nu_c (1 - x_{c,0}(t)) \pi_{\mathbf{x}_0(t)}^b). \end{aligned} \quad (41)$$

Hence, the evolution of $\mathbf{m}(\mathbf{x}(t))$ is governed only by the value of $\mathbf{x}_0(t)$, i.e., by the fraction of nodes with empty buffers. Moreover, observe that $\dot{\mathbf{m}}(\mathbf{x}(t)) = \mathbf{0}$ if and only if

$$\pi_{\mathbf{x}_0(t)}^b = \frac{\lambda_c}{\nu_c (1 - x_{c,0}(t))}, \quad \forall c \in \mathcal{C}.$$

In particular, if $\mathbf{x}_0 = \mathbf{x}_0^*$, i.e., $x_{c,0} = 1 - \xi_c$, then, by definition of ξ_c (see (35)), \mathbf{x} is an equilibrium point for $\mathbf{m}(\mathbf{x}(t))$.

Furthermore, observe that $\dot{m}_c(\mathbf{x}(t))$ is increasing in $x_{c,0}$ (at least as long as $\mathbf{x} \in E^1$). In fact

$$\frac{\partial}{\partial x_{c,0}} \dot{m}_c(\mathbf{x}(t)) = \frac{\nu_c}{p_c} \pi_{\mathbf{x}_0}^b (1 + (1 - x_{c,0}) \frac{\nu_c}{\mu_c} \pi_{\mathbf{x}_0}^b) > 0.$$

Therefore, given two solutions $\mathbf{x}^L(t)$ and $\mathbf{x}^U(t)$ such that $\mathbf{x}_0^L(t) \geq \mathbf{x}_0^U(t)$ (where the inequality is componentwise), we have

$$\dot{m}_c(\mathbf{x}^L(t)) \geq \dot{m}_c(\mathbf{x}^U(t)), \quad \forall c \in \mathcal{C},$$

and consequently, if $\mathbf{x}_0^L(t) \geq \mathbf{x}_0^* \geq \mathbf{x}_0^U(t)$,

$$\dot{m}_c(\mathbf{x}^L(t)) \geq 0 \geq \dot{m}_c(\mathbf{x}^U(t)), \quad \forall c \in \mathcal{C}.$$

5.1.2 Stochastic dominance

The set $E^1 \subseteq E$ may be partially ordered via the stochastic dominance relation.

Definition 2. Given $\mathbf{x}^L, \mathbf{x}^U \in E^1$, then \mathbf{x}^L is stochastically dominated by \mathbf{x}^U if

$$\sum_{n=0}^m x_{c,n}^L \geq \sum_{n=0}^m x_{c,n}^U, \quad \text{for every } m \in \mathbb{N}_0, c \in \mathcal{C}.$$

We denote this relation by $\mathbf{x}^L \leq_s \mathbf{x}^U$.

This relation allows us to couple different solutions of the initial-value problem governed by (39) as stated in the following proposition, whose proof is provided in Appendix B.6.

Proposition 5. The stochastic dominance is a positive invariant relation for the initial-value problem governed by (39). In particular, given solutions $\mathbf{x}^L(t), \mathbf{x}^U(t) \in C_{E^1}[0, \infty)$ of (39), and $t_0 \geq 0$ such that $\mathbf{x}^L(t_0) \leq_s \mathbf{x}^U(t_0)$, we have $\mathbf{x}^L(t) \leq_s \mathbf{x}^U(t)$ for every $t \geq t_0$.

Hence, by exploiting that if $\mathbf{x}(0) = \mathbf{x}^*$ then $\mathbf{x}(t) = \mathbf{x}^*$ for every $t \geq 0$, we identify the following invariant sets.

Corollary 4. The sets

$$\underline{D} = \{\mathbf{x} \in E^1 : \mathbf{x} \leq_s \mathbf{x}^*\}, \quad \overline{D} = \{\mathbf{x} \in E^1 : \mathbf{x} \geq_s \mathbf{x}^*\}$$

are positive invariant for the initial-value problem (39).

Moreover, stochastic dominance provides valuable information about how the mass of a certain solution evolves.

Corollary 5. Consider solutions $\mathbf{x}^L(t), \mathbf{x}^U(t) \in C_{E^1}[0, \infty)$ of (39) such that $\mathbf{x}^L(t) \leq_s \mathbf{x}^U(t)$. Then

$$m_c(\mathbf{x}^L(t)) \leq m_c(\mathbf{x}^U(t)), \quad \dot{m}_c^L(\mathbf{x}(t)) \geq \dot{m}_c^U(\mathbf{x}(t)), \quad (42)$$

for every $c \in \mathcal{C}$.

The first relation is a direct consequence of the mass definition, the second follows from the formula for the mass derivative (41) and from $\mathbf{x}_{c,0}^L(t) \geq \mathbf{x}_{c,0}^U(t)$ for $c \in \mathcal{C}$, which is a direct implication of the stochastic dominance property. In particular, observe that if $\mathbf{x}(0) \in \underline{D}$ (resp. $\mathbf{x}(0) \in \overline{D}$) then $m_c(\mathbf{x}(t)) \leq m_c(\mathbf{x}^*)$ (resp. $m_c(\mathbf{x}(t)) \geq m_c(\mathbf{x}^*)$) and $\dot{m}_c(\mathbf{x}(t)) \geq 0$ (resp. $\dot{m}_c(\mathbf{x}(t)) \leq 0$) for every $c \in \mathcal{C}$ and $t \geq 0$.

5.1.3 Characterization of $B(\mathbf{x}^*)$

Consider a solution $\mathbf{x}(t)$ of (39) with $\mathbf{x}(0) \in E^1$. Observe that if $\mathbf{x}(t) \rightarrow \mathbf{x}^*$, then $\mathbf{x}(0) \in B(\mathbf{x}^*)$. The following proposition, whose proof is provided in Appendix B.7, characterizes $B(\mathbf{x}^*)$.

Proposition 6. *The basin of attraction $B(\mathbf{x}^*)$ of \mathbf{x}^* coincides with the set $E_{<\infty}^1 \subseteq E^1$, where*

$$E_{<\infty}^1 = \{\mathbf{x} \in E^1 : m(\mathbf{x}) < \infty\},$$

i.e., $\mathbf{x}(0) \in E_{<\infty}^1$ implies that $\mathbf{x}(t) \rightarrow \mathbf{x}^$.*

Observe that the stated condition is necessary, indeed $m(\mathbf{x}^*)$ is finite and a solution of (39) cannot hit $E_{<\infty}^1$ in finite time starting from an infinite-mass state. However, apart from this necessary exclusion, all the other solutions are attracted by the equilibrium point $\mathbf{x}^* \in E^1$, which is therefore *globally stable* in $E_{<\infty}^1$.

The same argument, without the assumption for the interference graph to be complete, would not lead to the proof of Proposition 6. Indeed the argument relies on Proposition 5 which is not true in general, i.e., given solutions $\mathbf{x}^L(t), \mathbf{x}^U(t)$ of the initial-value problem governed by (6) for general class-interference graphs $\mathbf{x}^L(t_0) \leq_s \mathbf{x}^U(t_0)$ does not imply that $\mathbf{x}^L(t) \leq_s \mathbf{x}^U(t)$ for all $t \geq t_0$. However, in the inductive argument used in the proof of Proposition 6, we rely on the invariance of the stochastic dominance *only* in the basic step, i.e., Lemma 5 in Appendix B.7, where we show that $\mathbf{x}_0(t) \rightarrow \mathbf{x}_0^*$. Hence the following proposition holds.

Proposition 7. *Given a general class-interference graph, and assume the existence of an equilibrium point for the initial-value problem (6). Then, if a solution $\mathbf{x}(t)$ of (6) is such that*

$$\mathbf{x}(0) \in E_{<\infty}^1, \quad \mathbf{x}_0(t) \rightarrow \mathbf{x}_0^*,$$

it holds that $\mathbf{x}(0) \in B(\mathbf{x}^)$.*

Propagation of chaos and consequences of Proposition 6. In dealing with many-particle interaction problems, a common assumption is the propagation of chaos, i.e., the assumption that the particles become pairwise independent when their number grows and their evolution is influenced only by an aggregated averaged effect of the entire population state. Mean-field theory provides a powerful machinery by which the propagation of chaos assumption can be validated. In our setting, the particles are the nodes and their state is given by the number of packets in their buffer. Consider a chaotic initial condition, i.e., the initial number of packets in the buffers of the various nodes are mutually independent, then the multi-exchangeability property [25, Section 5] and the convergence towards the initial-value problem presented in Theorem 1 suffice to guarantee that the independence propagates over any finite horizon, see [3, 39]. However, as described in [3], in order for the independence to hold in the limit (as the time grows to infinity), it is necessary that the solution of the initial-value problem (39) converges to an equilibrium point. We conclude that, in a complete interference graph setting, the queues of the various nodes in the system behave independently and undergo a common averaged effect depending on the equilibrium configuration of the network.

In the case of slotted CSMA networks, Bianchi in [6] assumed chaos propagates provided the network is large enough, and so obtained an analytical expression for the network throughput. Many papers discuss this assumption, including [16], which validated the propagation of chaos under certain assumptions via a mean-field approach. Here we just provided another supporting argument for the case of a complete interference graph.

5.2 Positive recurrence

As mentioned earlier, the final objective of this section is to obtain an approximation for the stationary distribution of the queue length processes at the various nodes in large systems. In particular, we will show that $\mathbf{x}^* \in E^1$ provides an approximation for the stationary distribution $\mathbf{x}^{(N)}$ of the system with N nodes when N is large. First, however, it must be ensured that the population process has a stationary distribution $\mathbf{x}^{(N)}$ under condition (8) for each system with a finite number

of N nodes. Hence, we will establish the positive recurrence of the queue length process $\mathbf{Q}^{(N)}$, from which the positive recurrence of $\mathbf{X}^{(N)}$ follows.

The complete interference graph assumption plays an important role in this analysis. Indeed, since at every moment in time at most one node can be active, the CSMA model can be described as a single-server polling system with random routing. Specifically, the switchover periods of the server correspond to the idle-state of the activity process, i.e., the intervals in which each node is allowed to back-off. When a node completes its back-off period, the server visits that node and serves a packet, when present. Specifically, consider the polling system with the following features:

- Single server and N nodes. The nodes are partitioned in $N_1^{(N)} + \dots + N_C^{(N)}$ nodes with class-dependent features;
- Packets arrive at the k -th node belonging to class c as a Poisson process of rate $\lambda_c^{(N)}$.
- The service policy is 1-limited. In particular, a single packet is served if present, otherwise the server leaves the node immediately;
- When a class- c node is visited and its buffer is nonempty, a packet is served for an exponentially distributed time with parameter $\mu_c^{(N)}$;
- The switchover time of the server between two consecutive visits is exponentially distributed with parameter $\nu := \sum_{c \in \mathcal{C}} N_c^{(N)} \nu_c^{(N)} = \sum_{c \in \mathcal{C}} \nu_c$;
- At the end of a switchover period, the probability for the server to pick the k -th node belonging to class c is equal to $p_{c,k} = \nu_c^{(N)} / \nu$.

Observe that the above-described polling system behaves exactly as the CSMA model with complete interference and that the latter property is crucial in establishing this connection.

Polling systems have been extensively studied in the literature. In particular, the condition for the positive recurrence of the above-described model has already been investigated. In [17, 18, 22], the authors provide positive recurrence conditions for general classes of polling systems via the fluid limit analysis of the queue length processes [17, 18] and via a direct approach [22]. The following result is a special case of Theorem 2.3 in [18], and in Appendix B.8 we adapted our notation to the setting in [18], so as to further illuminate the connection.

Proposition 8 (Theorem 2.3 [18]). *Define*

$$\varrho \doteq \sum_{d \in \mathcal{C}} \frac{\lambda_d}{\mu_d} + \max_{c \in \mathcal{C}} \frac{\lambda_c}{\nu_c}.$$

- (i) *If $\varrho < 1$, then the queue length process $\mathbf{Q}^{(N)}$ is positive recurrent for every $N \in \mathbb{N}$.*
- (ii) *If $\varrho > 1$, then the queue length process $\mathbf{Q}^{(N)}$ is transient for every $N \in \mathbb{N}$.*

The proof in [18, Theorem 2.3] consists of analyzing the fluid limit of a polling system, and we can conclude the positive recurrence of the CSMA model due to the above-described equivalence.

Remark: Observe that condition $\varrho < 1$ is not dependent on N , and is equivalent to condition (8), i.e.,

$$\varrho < 1 \iff \xi_c < 1, \text{ for every } c \in \mathcal{C}.$$

In Appendix B.9 we provide an alternative and constructive proof of Proposition 8 for the single-class case. The proof consists in exhibiting explicitly a quadratic Lyapunov function.

When the interference graph is not complete, the connection with polling systems does not apply, and an explicit condition for positive recurrence is an open problem.

5.3 Tightness and interchange of limits

In Section 5.2 we showed that the queue length process $\mathbf{Q}^{(N)}(t)$ has a stationary distribution if condition (8) is satisfied. Denote it by the probability measure $\pi^{(N)}$, i.e.,

$$\lim_{t \rightarrow \infty} \mathbb{P}\{\mathbf{Q}^{(N)}(t) = \mathbf{q}\} = \pi^{(N)}(\mathbf{q}), \quad \mathbf{q} \in \mathbb{N}_0^N.$$

For finite N the process $\mathbf{Q}^{(N)}(t)$ completely characterizes $\mathbf{X}^{(N)}(t)$, hence the population process $\mathbf{X}^{(N)}(t)$ also has a stationary distribution. In particular, define $\mathbf{x}^{(N)} \in E^1$ as

$$\mathbf{x}^{(N)} = \sum_{\mathbf{q} \in \mathbb{N}_0^N} \pi^{(N)}(\mathbf{q}) \mathbf{e}(\mathbf{q}),$$

where $\mathbf{e}(\mathbf{q}) \in E^1$ is such that

$$(\mathbf{e}(\mathbf{q}))_{c,n} = \frac{1}{N_c} \sum_{k=1}^{N_c} \mathbb{1}_{\{q_{c,k}=n\}}.$$

Then we know that

$$\lim_{t \rightarrow \infty} \mathbb{P}(\rho(\mathbf{X}^{(N)}(t), \mathbf{x}^{(N)}) > \epsilon) = 0, \quad \forall \epsilon > 0.$$

At this point we are close to the target we set at the start of this section, that is proving equation (37). Indeed, in Section 5.1 we characterized $\mathbf{x}^* \in E^1$, and thanks to Proposition 8, we proved the existence of the sequence of limiting distributions $\{\mathbf{x}^{(N)}\}_{N \in \mathbb{N}} \in E^1$.

Two further steps are needed in order to show (37). First, we need to actually show that the sequence $(\mathbf{x}^{(N)})_{N \in \mathbb{N}} \subseteq E^1$ has at least one limiting point. Then we need to show that each possible limiting point necessarily coincides with \mathbf{x}^* . Once we settle these issues, we can finally state that the limits can indeed be interchanged and that equation (37) holds.

The proof of the following proposition is given in Appendix B.10.

Proposition 9. *The sequence $(\mathbf{x}^{(N)})_{N \in \mathbb{N}} \in E^1$ is tight.*

At this point, via Prohorov's theorem [7, Theorem 6.2, page 37] we deduce that the sequence $(\mathbf{x}^{(N)})_{N \in \mathbb{N}} \subseteq E^1$ possesses converging subsequences. In order to complete the proof of Theorem 3, it suffices to show that these subsequences all converge to $\mathbf{x}^* \in E^1$.

Given any of the converging subsequences $(\mathbf{x}^{(N_m)})_{m > 0}$, denote by \mathbf{d} its limit. The elements of the sequence $(\mathbf{x}^{(N_m)})_{m > 0}$ may be interpreted as a sequence of initial values for the mean-field limit described in Theorem 1. Consider the sequence of solutions of (39) along this sequence of initial-values. Observe that, as $N_m \rightarrow \infty$, such a solution has to converge to the solution $\mathbf{x}(t)$ of the differential equation (39) with initial condition \mathbf{d} . However, \mathbf{d} is a stationary distribution for the system with $N \rightarrow \infty$ nodes, and therefore \mathbf{d} must be a stationary point of the initial value problem (39). Due to Theorem 2, the only possibility is $\mathbf{d} = \mathbf{x}^*$. We conclude that the sequence $(\mathbf{x}^{(N)})_{N \in \mathbb{N}}$ converges to \mathbf{x}^* .

The above argument is well-known in the literature, and has been detailed in [4]. It has been applied in many contexts, see for instance [23, 31, 38]. Observe that, apart from the tightness of the sequence $(\mathbf{x}^{(N)})_{N \in \mathbb{N}}$, we relied only on the fact that there exists a unique equilibrium point for (6) (stability of \mathbf{x}^* is not required). In [28, Section 7.2] the authors provide an example where it is shown that if multiple limiting points exist, then the interchange of limits may lead to drastically different results.

6 Performance measures

In this section, we leverage the asymptotic results of Section 5 for the population process to obtain various performance measures of interest. In particular, we derive approximations for the stationary distribution and expectation of the queue length of a node, the sojourn time and waiting time of a packet in the system. Denote by $Q_c^{(N)}$ the stationary queue length at a class- c node, and by $W_c^{(N)}$ and $S_c^{(N)}$ the stationary waiting time and sojourn time of a packet at a class- c node. For ease of notation we denote $N_c = N_c^{(N)}$.

6.1 Convergence in distribution

Consider a tagged class- c node. In order to obtain an approximation for the stationary performance measures, we rely on the symmetry of the class- c nodes and we leverage the results established in the previous section. In particular, for N sufficiently large, we use the approximation

$$\lim_{t \rightarrow \infty} \mathbb{P}\{Q_{c,1}^{(N)}(t) = k\} \approx \lim_{N \rightarrow \infty} \lim_{t \rightarrow \infty} \mathbb{P}\{Q_{c,1}^{(N)}(t) = k\}, \quad (43)$$

and similar approximations for the waiting time and the sojourn time, and derive exact expressions for the limit on the right-hand side as stated in Theorem 4.

In order to prove Theorem 4, we begin with the queue length distribution. Theorem 3 implies that

$$\lim_{N \rightarrow \infty} \lim_{t \rightarrow \infty} \mathbb{P}\{Q_{c,1}^{(N)}(t) = k\} = \lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} \mathbb{P}\{Q_{c,1}^{(N)}(t) = k\},$$

and in turn, via the mean-field limit in Theorem 1, we argue that

$$\lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} \mathbb{P}\{Q_{c,1}^{(N)}(t) = k\} = \lim_{t \rightarrow \infty} x_{c,k}(t),$$

while Proposition 6 implies

$$\lim_{t \rightarrow \infty} x_{c,k}(t) = x_{c,k}^* = (1 - \xi_c)\xi_c^k.$$

Hence,

$$Q_c^{(N)} \Rightarrow \bar{Q}_c, \quad \bar{Q}_c \sim \text{Geo}(\xi_c), \quad c \in \mathcal{C}.$$

We now continue with the waiting-time distribution. Denote by $\phi_X(\cdot)$ and $F_X(\cdot)$ the Laplace-Stieltjes transform and the probability generating function of a random variable X . Define the random variable $\bar{W}_c^{(N)} = \frac{\lambda_c}{N_c} W_c^{(N)}$. Observe that the Laplace-Stieltjes transform $\phi_{W_c^{(N)}}$ and the probability generating function $F_{Q_c^{(N)}}(z)$ satisfy the distributional form of Little's law, i.e.

$$\phi_{\bar{W}_c^{(N)}}(z) = \phi_{\frac{\lambda_c}{N_c} W_c^{(N)}}(z) = F_{Q_c^{(N)}}(1 - z), \quad z \in [0, 1], \quad (44)$$

see [5]. Moreover, we already proved that

$$F_{Q_c^{(N)}}(1 - z) \rightarrow \frac{1 - \xi_c}{1 - \xi_c(1 - z)}, \quad z \in [0, 1].$$

Hence, it only remains to be shown that the sequence of random variables $\{\bar{W}_c^{(N)}\}_N$ converges. We state the following proposition, which is a consequence of Little's law.

Proposition 10. *The sequence of random variables $\{\bar{W}_c^{(N)}\}_N$ is tight.*

Then, it follows from Prohorov's theorem, that the sequence $\{\bar{W}_c^{(N)}\}_N$ is relatively compact. Let \bar{W}_c be a limit random variable for an arbitrary convergent subsequence. It follows from (44) that

$$\phi_{\bar{W}_c}(z) = \frac{1 - \xi_c}{1 - \xi_c(1 - z)}, \quad z \in [0, 1]. \quad (45)$$

However it is known that the Laplace-Stieltjes transform of a non-negative random variable is analytic in the positive open half plane (excluding the imaginary axis) [45, Chapter 8]. Moreover, from (45), it coincides with the Laplace-Stieltjes transform of an exponential random variable on a line segment. It is known that two analytic functions agreeing on a line segment must be the same on the entire domain, that is the non-negative half plane, see [1], so

$$\phi_{\bar{W}_c}(s) = \frac{1 - \xi_c}{1 - \xi_c(1 - s)}, \quad \text{Re}(s) \geq 0.$$

Therefore, we have shown that

$$\frac{\lambda_c}{N_c} W_c^{(N)} \Rightarrow \bar{W}_c, \quad \bar{W}_c \sim \text{Exp}\left(\frac{1 - \xi_c}{\xi_c}\right), \quad c \in \mathcal{C}.$$

We now conclude with the sojourn time distribution. Define the random variable $\bar{S}_c^{(N)} = \frac{\lambda_c}{N_c} S_c^{(N)}$, and observe that

$$\bar{S}_c^{(N)} \sim \bar{W}_c^{(N)} + \frac{\lambda_c}{N_c} U_c^{(N)}, \quad c \in \mathcal{C},$$

where $U_c^{(N)} \sim \text{Exp}(\mu_c^{(N)}) \sim \text{Exp}(\mu_c)$ denotes the transmission time of a class- c packet. Therefore

$$\frac{\lambda_c}{N_c} S_c^{(N)} \Rightarrow \bar{S}_c, \quad \bar{S}_c \sim \text{Exp}\left(\frac{1 - \xi_c}{\xi_c}\right), \quad c \in \mathcal{C}.$$

Combining (43) and Theorem 4 we obtain an approximation for the stationary distribution of the queue length at a tagged class- c node given by

$$\lim_{t \rightarrow \infty} \mathbb{P}\{Q_{c,1}^{(N)}(t) = k\} \approx (1 - \xi_c) \xi_c^k,$$

i.e., a geometric distribution with parameter ξ_c , for N sufficiently large, and similar approximations are obtained for the waiting and sojourn time distributions.

6.2 Convergence in expectation

In the previous subsection, we proved Theorem 4 obtaining the limiting distributions of the stationary queue length, waiting-time, and sojourn time distributions of the various classes. Such convergence results do not allow any immediate conclusions for the corresponding expectations as stated in Theorem 5. Here we show the convergence of the stationary expected waiting time from which the convergence of the stationary expected sojourn time and queue length will follow easily.

Thanks to Theorem 4, we know that the rescaled stationary waiting time of a packet at a class- c node converges in distribution to $\bar{W}_c \sim \text{Exp}\left(\frac{1 - \xi_c}{\xi_c}\right)$. It is well-known [8] that

$$\mathbb{E}[\bar{W}_c] \leq \liminf_{N \rightarrow \infty} \frac{\lambda_c}{N_c} \mathbb{E}[W_c^{(N)}], \quad (46)$$

and we want to show that

$$\mathbb{E}[\bar{W}_c] = \lim_{N \rightarrow \infty} \frac{\lambda_c}{N_c} \mathbb{E}[W_c^{(N)}].$$

In order to prove that, we will exploit again the equivalence with a 1-limited polling system with random routing described in Section 5.2.

Specifically, the following formula is the pseudo-conservation law proved in [12, Eqn. (5.34)] specialized to a CSMA setting, i.e., $p_{c,k} = \nu_c^{(N)}/\nu$ and $\sigma_{c,k} = \nu$, with $\nu = \sum_{c=1}^C \nu_c$:

$$\begin{aligned} & \sum_{c \in \mathcal{C}} \sum_{k=1}^{N_c} \frac{\lambda_c}{N_c \mu_c} \left(1 - \frac{\lambda_c}{\nu_c} \frac{1}{1 - \rho}\right) \mathbb{E}[W_c^{(N)}] \\ &= \frac{\rho}{1 - \rho} \sum_{c \in \mathcal{C}} \sum_{k=1}^{N_c} \frac{\lambda_c}{N_c \mu_c^2} + \frac{1}{1 - \rho} \sum_{c \in \mathcal{C}} \sum_{k=1}^{N_c} \frac{\lambda_c}{\mu_c \nu_c} - \sum_{c \in \mathcal{C}} \sum_{k=1}^{N_c} \frac{\lambda_c}{\nu N_c \mu_c} + \frac{\rho}{\nu}, \end{aligned} \quad (47)$$

where $\rho = \sum_{c=1}^N \rho_c$. The above pseudo-conservation law is valid as long as the stationary expectations of the waiting times are finite, which is granted by [17, Theorem 5.5] as long as the stability condition (8) is satisfied.

We will show that there exist a sequence of constants $K_{c,N} > 0$ for $c \in \mathcal{C}, N > 0$ converging to $K_c > 0$, such that

$$\sum_{c \in \mathcal{C}} K_c \mathbb{E}[\bar{W}_c] = \lim_{N \rightarrow \infty} \sup_{c \in \mathcal{C}} K_{c,N} \frac{\lambda_c}{N_c} \mathbb{E}[W_c^{(N)}]. \quad (48)$$

Since $K_c, K_{c,N} > 0$ for every $c \in \mathcal{C}$, $N > 0$, (46) and (48) imply

$$\lim_{N \rightarrow \infty} \frac{\lambda_c}{N_c} \mathbb{E}[W_c^{(N)}] = \mathbb{E}[\bar{W}_c], \quad c \in \mathcal{C}.$$

It remains to show that (48) holds. As the pseudo-conservation law (47) suggests, we consider

$$K_{c,N} = \frac{p_{c,N}}{\mu_c} \left(1 - \frac{\lambda_c}{\nu_c} \frac{1}{1-\rho}\right),$$

and consequently $K_c = \frac{p_c}{\mu_c} \left(1 - \frac{\lambda_c}{\nu_c} \frac{1}{1-\rho}\right)$. In fact,

$$\begin{aligned} & \limsup_{N \rightarrow \infty} \sum_{c \in \mathcal{C}} \frac{p_{c,N}}{\mu_c} \left(1 - \frac{\lambda_c}{\nu_c} \frac{1}{1-\rho}\right) \frac{\lambda_c}{N_c} \mathbb{E}[W_c^{(N)}] \\ &= \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{c \in \mathcal{C}} \frac{\lambda_c}{\mu_c} \left(1 - \frac{\lambda_c}{\nu_c} \frac{1}{1-\rho}\right) \mathbb{E}[W_c^{(N)}] \\ &= \limsup_{N \rightarrow \infty} \frac{1}{N} \left(\frac{\rho}{1-\rho} \sum_{c \in \mathcal{C}} \frac{\lambda_c}{\mu_c^2} + \frac{1}{1-\rho} \sum_{c \in \mathcal{C}} \frac{N_c \lambda_c}{\mu_c \nu_c} - \sum_{c \in \mathcal{C}} \frac{\lambda_c}{\nu \mu_c} + \frac{\rho}{\nu} \right) \\ &= \frac{1}{1-\rho} \sum_{c \in \mathcal{C}} \frac{p_c \lambda_c}{\mu_c \nu_c}, \end{aligned}$$

where the second equality is due to (47). On the other hand,

$$\begin{aligned} \sum_{c \in \mathcal{C}} \frac{p_c}{\mu_c} \left(1 - \frac{\lambda_c}{\nu_c} \frac{1}{1-\rho}\right) \mathbb{E}[\bar{W}_c] &= \sum_{c \in \mathcal{C}} \frac{p_c}{\mu_c} \left(1 - \frac{\lambda_c}{\nu_c} \frac{1}{1-\rho}\right) \frac{\xi_c}{1-\xi_c} \\ &= \sum_{c \in \mathcal{C}} \frac{p_c}{\mu_c} \frac{\nu_c(1-\rho) - \lambda_c}{\nu_c} \frac{1}{1-\rho} \frac{\lambda_c}{\nu_c(1-\rho) - \lambda_c} \\ &= \frac{1}{1-\rho} \sum_{c \in \mathcal{C}} \frac{p_c \lambda_c}{\mu_c \nu_c}, \end{aligned}$$

hence (48) holds, and we just showed

$$\frac{\lambda_c}{N_c} \mathbb{E}[W_c^{(N)}] \rightarrow \mathbb{E}[\bar{W}_c].$$

The result for the sojourn time follows immediately since

$$W_c^{(N)} = S_c^{(N)} - U_c^{(N)}, \quad U_c^{(N)} \sim \text{Exp}(\mu_c).$$

The argument we just presented implies that the sequence of random variables $\{\frac{\lambda_c}{N_c} W_c^{(N)}\}_N$ is uniformly integrable. Via the conventional Little's law we can then deduce that, for every $c \in \mathcal{C}$, the sequence $\{Q_c^{(N)}\}_N$ is itself uniformly integrable. Hence, it holds that

$$\mathbb{E}[Q_c^{(N)}] \rightarrow \mathbb{E}[\bar{Q}_c].$$

7 Model extensions

The theoretical machinery for the derivation of mean-field limits developed in Section 4 is highly flexible and may be applied to a wide range of models. In order to illustrate that, we outline in this section some possible extensions of the model described in Section 2, and state the associated mean-field limits. The proof of these limit results resembles that for the original model, and the arguments are not repeated. In particular, the time-scale separation between the activity process and the population process plays a key role once again. However, the expression for the stationary distribution of the activity process with a fixed population is model-dependent and this is reflected in the corresponding limiting initial-value problems which differ only in the expression for $\pi_{\mathbf{x}}(\Omega_{-c})$. We recall that $\pi_{\mathbf{x}}(\Omega_{-c})$ may be obtained from the product-form solution for the stationary distribution of the activity process of saturated CSMA models.

7.1 Models with multiple transmission rates

In this model, the transmission of the packets may occur via different modes, which are associated with different distributions of the transmission times. Denote the set of modes which a class- c node may use by $M_c = \{m_1^c, \dots, m_{n_c}^c\}$. In particular, mode m_l^c is picked with probability \tilde{p}_l^c and the transmission time with that mode is phase-type distributed with mean $(\mu_c^l)^{-1} > 0$.

Observe that a weighted sum of phase-type distributions is itself phase-type. This extension influences our original analysis only in Section 4.1.4, where we characterize the limit of $\alpha_\omega^{(N)}$, yielding a more complicated version of the system of equations (34). However, it is shown in [40] that such a system can be uniquely solved and its solution only depends on the mean of the phase-type distributions involved, due to an insensitivity property for the stationary distribution of the activity process in saturated networks.

Hence, the mean-field limit of this generalized model turns out to be very similar to Theorem 1, and we only need to characterize the counterpart of $\pi_{\mathbf{x}}(\Omega_{-c})$. Denote by U_c the expected value of the transmission time of a class- c node, i.e.,

$$U_c = \sum_{l=1}^{n_c} \tilde{p}_l^c \frac{1}{\mu_c^l}.$$

Theorem 6. *Assume $\tilde{\mathbf{X}}^{(N)}(0) \xrightarrow{N \rightarrow \infty} \mathbf{x}^\infty \in E$ and $\lim_{N \rightarrow \infty} p_{c,N} = p_c > 0$ for every $c \in \mathcal{C}$. Then the sequence of processes*

$$(\tilde{\mathbf{X}}^{(N)}(Nt)) \in D_E[0, \infty)$$

has a continuous limit $\mathbf{x}(t) \in C_E[0, \infty)$ which is determined by the unique solution of the initial-value problem

$$\frac{d\mathbf{x}(t)}{dt} = H^{mtr}(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}^\infty,$$

where the function $H^{mtr}(\cdot)$ is defined by

$$H^{mtr}(\mathbf{x}) = \sum_{c \in \mathcal{C}} \frac{1}{p_c} \left(\lambda_c \sum_{n=0}^{\infty} x_{c,n} \mathbf{e}_{c,n}^{c,n+1} + \pi_{\mathbf{x}_0}^{mtr}(\Omega_{-c}) \nu_c \sum_{n=1}^{\infty} x_{c,n} \mathbf{e}_{c,n}^{c,n-1} \right)$$

and $\pi_{\mathbf{x}_0}^{mtr}(\Omega_{-c}) = \sum_{\omega \in \Omega_{-c}} \pi_{\mathbf{x}_0}^{mtr}(\omega)$ with

$$\pi_{\mathbf{x}_0}^{mtr}(\omega) = \frac{1}{Z^{mtr}} \prod_{c \in \mathcal{C}} \left(U_c \nu_c (1 - x_{c,0}) \right)^{\omega_c}$$

and Z^{mtr} a normalizing constant.

Note that in the original case we have only one rate per class, yielding $U_c = 1/\mu_c$ for every $c \in \mathcal{C}$, in which case Theorem 6 reduces to Theorem 1.

Observe that all the results in Sections 4.2 and 5 may be extended to the model with multiple transmission rates, by replacing μ_c with $1/U_c$. Moreover, since phase-type distributions approximate any nonnegative distribution, we deduce that the mean-field limit is insensitive with respect to the transmission time distribution.

7.2 Models with queue-based back-off rates

In the basic model, nodes with nonempty buffers back-off at a rate that depends only on their class. However, in [24, 35], the authors showed that CSMA networks with suitable queue-based back-off rates are maximally stable without any explicit knowledge of the arrival rates in arbitrary interference graphs.

Consider the model described in Section 2, where the back-off rate of a node is a function of the number of packets in its buffer, i.e., the back-off rate of a class- c node having n packets in the buffer

is $\nu_c^{(N)}(n) = \nu_c(n)/N_c$, see [11]. Note that when $\nu_c(n) = \nu_c$ for every n we recover the basic model. The system dynamics are very similar to those of the original model. Specifically, we only have to replace the back-off transitions, as follows:

A back-off is completed by a class- c node having $n > 0$ packets in its buffer: this happens only if the class activity state allows class- c nodes to back off, i.e., $\mathbf{Y}^{(N)} \in \Omega_{-c}$, and at rate $\nu_c^{(N)}(n)N_c\tilde{X}_{c,n}^{(N)} = \nu_c(n)\tilde{X}_{c,n}^{(N)}$, i.e., $\nu_c^{(N)}$ times the number of class- c nodes in state n , and generates the transition

$$\begin{cases} \mathbf{Y}^{(N)} \rightarrow \mathbf{Y}^{(N)} + \mathbf{e}_c \\ \tilde{X}_{c,n}^{(N)} \rightarrow \tilde{X}_{c,n}^{(N)} - \frac{1}{N_c} \\ \tilde{X}_{c,n-1}^{(N)} \rightarrow \tilde{X}_{c,n-1}^{(N)} + \frac{1}{N_c}, \end{cases}$$

i.e.,

$$(\tilde{\mathbf{X}}^{(N)}, \mathbf{Y}^{(N)}) \rightarrow (\tilde{\mathbf{X}}^{(N)} + \frac{1}{N_c}\mathbf{e}_{c,n-1}, \mathbf{Y}^{(N)} + \mathbf{e}_c).$$

In accordance with the basic framework, the following theorem provides a description for the mean-field limit of the CSMA model with queue-based back-off rates.

Theorem 7. *Assume $\tilde{\mathbf{X}}^{(N)}(0) \xrightarrow{N \rightarrow \infty} \mathbf{x}^\infty \in E^1$ and $\lim_{N \rightarrow \infty} p_{c,N} = p_c > 0$ for every $c \in \mathcal{C}$. Then the sequence of processes*

$$(\tilde{\mathbf{X}}^{(N)}(Nt)) \in D_E[0, \infty)$$

has a continuous limit $\mathbf{x}(t) \in C_E[0, \infty)$ which is determined by the unique solution of the initial-value problem

$$\frac{d\mathbf{x}(t)}{dt} = H^{qb}(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}^\infty, \quad (49)$$

where the function $H(\cdot)$ is defined by

$$H^{qb}(\mathbf{x}) = \sum_{c \in \mathcal{C}} \frac{1}{p_c} \left(\lambda_c \sum_{n=0}^{\infty} x_{c,n} \mathbf{e}_{c,n}^{c,n+1} + \pi_{\mathbf{x}}^{qb}(\Omega_{-c}) \sum_{n=1}^{\infty} \nu_c(n) x_{c,n} \mathbf{e}_{c,n}^{c,n-1} \right)$$

and $\pi_{\mathbf{x}}^{qb}(\Omega_{-c}) = \sum_{\omega \in \Omega_{-c}} \pi_{\mathbf{x}}^{qb}(\omega)$ with

$$\pi_{\mathbf{x}}^{qb}(\omega) = \frac{1}{Z^{qb}} \prod_{c \in \mathcal{C}} \left(\frac{\sum_{n=1}^{\infty} \nu_c(n) x_{c,n}}{\mu_c} \right)^{\omega_c}$$

and Z^{qb} a normalizing constant.

Note that if $\nu_c(n) = \nu_c$ for every $n \in \mathbb{N}$, then

$$\sum_{n=1}^{\infty} \nu_c(n) x_{c,n} = \nu_c \sum_{n=1}^{\infty} x_{c,n} = \nu_c(1 - x_{c,0})$$

and the original mean-field limit in Theorem 1 may be interpreted as a special case of Theorem 7.

In order to understand the formula for $\pi_{\mathbf{x}}^{qb}(\omega)$, consider a fictitious saturated CSMA model, with C nodes and the original class-interference graph and a time-invariant population $\mathbf{x} \in E$. The time required to transmit a packet originating from node c is exponentially distributed with parameter μ_c , while the length of a back-off period of node c is captured by the population-dependent random variable $V_c(\mathbf{x})$, where

$$V_c(\mathbf{x}) \sim \min_{n \in \mathbb{N}} V_{c,n}(\mathbf{x}), \quad V_{c,n}(\mathbf{x}) \sim \text{Exp}(x_{c,n} \nu_c(n)),$$

and therefore $V_c(\mathbf{x}) \sim \text{Exp}(\sum_{n=1}^{\infty} \nu_c(n) x_{c,n})$. In [40] the authors proved the insensitivity of the stationary distribution of the activity process for saturated CSMA networks with respect to the back-off distribution. Hence, we know that the expression for the stationary probability of the activity process to be in state ω corresponds to $\pi_{\mathbf{x}}^{qb}(\omega)$. Intuitively, the cumulative back-off time needed by class- c nodes is given by the minimum of the back-off time required by each node in the class, whose distribution depends on the number of packets in their buffer.

Theorem 8. *If $\rho \in \text{int}(\Gamma)$ and*

$$\sum_{m=0}^{\infty} \tilde{\nu}_c \prod_{n=1}^m \frac{1}{\nu_c(n)} = L_c < \infty, \quad \forall c \in \mathcal{C},$$

where $\tilde{\nu}_c = \nu_c \eta(\rho)_c$, then there exists a unique equilibrium point $\mathbf{x}^* = (x_{c,n}^*)_{c \in \mathcal{C}, n \in \mathbb{N}} \in E$ for the initial-value problem (49). In particular

$$x_{c,0}^* = \frac{1}{L_c}, \quad x_{c,n}^* = \frac{\tilde{\nu}_c}{\nu_c(n)} x_{c,n-1}^*. \quad (50)$$

We now briefly sketch the main argument needed to prove Theorem 8. The relation $H^{qb}(\mathbf{x}^*) = 0$ must be satisfied component-wise, i.e.,

$$x_{c,n}^* = \xi_c(n) x_{c,n-1}^*, \quad \xi_c(n) \doteq \frac{\lambda_c}{\nu_c(n) \pi_{\mathbf{x}^*}^{qb}(\Omega_{-c})}$$

for every $c \in \mathcal{C}$ and $n > 0$. Hence we only need to show that necessarily $\xi_c(n) = \tilde{\nu}_c / \nu_c(n)$.

Given $\mathbf{x} \in S$, define the vector of effective class-back-off rates $\tilde{\nu}(\mathbf{x}) \in \mathbb{R}_+^{\mathcal{C}}$ as follows

$$\tilde{\nu}_c(\mathbf{x}) \doteq \sum_{k=1}^{\infty} \nu_c(k) x_{c,k}, \quad c \in \mathcal{C}.$$

Observe that

$$\pi_{\mathbf{x}}^{qb}(\Omega_{+c}) = \frac{\tilde{\nu}_c(\mathbf{x})}{\mu_c} \pi_{\mathbf{x}}^{qb}(\Omega_{-c}),$$

and $\mathbf{x}^* \in E$ is an equilibrium point only if $\tilde{\nu}(\mathbf{x}^*) = \tilde{\nu}$ where $\tilde{\nu}_c = \eta(\rho)_c$ with $\sigma_c = 1/\mu_c$ for every $c \in \mathcal{C}$. Hence, the uniqueness of $\tilde{\nu}$ is ensured by the global invertibility of the map $\theta(\cdot)$ over Γ . Therefore, if \mathbf{x}^* is such that $H^{qb}(\mathbf{x}^*) = 0$, then

$$\pi_{\mathbf{x}^*}^{qb}(\Omega_{-c}) = \frac{\mu_c}{\tilde{\nu}_c} \pi_{\mathbf{x}^*}^{qb}(\Omega_{+c}) = \frac{\lambda_c}{\tilde{\nu}_c},$$

and consequently

$$\xi_c(n) = \frac{\tilde{\nu}_c}{\nu_c(n)}.$$

As a simple example, consider the case with a single class and $\nu(n) = n\nu$. Observe that $\tilde{\nu}$ must solve

$$\frac{\lambda}{\mu} = \frac{\tilde{\nu}/\mu}{1 + \tilde{\nu}/\mu}$$

and therefore

$$\tilde{\nu} = \frac{\lambda}{1 - \frac{\lambda}{\mu}}.$$

We now need to verify that $L < \infty$,

$$L = \sum_{m=0}^{\infty} \tilde{\nu} \prod_{n=1}^m \frac{1}{\nu(n)} = \frac{\tilde{\nu}}{\nu} \sum_{m=0}^{\infty} \frac{1}{m!} = \frac{\lambda e}{\nu(1 - \frac{\lambda}{\mu})}.$$

Hence, we conclude that

$$x_0^* = \frac{\nu(1 - \frac{\lambda}{\mu})}{\lambda e}, \quad x_n^* = \frac{\lambda}{\nu n(1 - \frac{\lambda}{\mu})} x_{n-1}^*.$$

7.3 Models with finite capacity

A practical constraint in the implementation of CSMA is that the number of packets stored in the buffer of a node must not exceed a certain capacity. In this section, we assume that the buffer capacity of every node in class c is limited to be no more than a certain value $K_c \in \mathbb{N}$. In particular, this feature yields immediately

$$X_{c,n}^{(N)} = 0, \quad \forall n > K_c, c \in \mathcal{C}.$$

Theorem 9. *Assume $\tilde{\mathbf{X}}^{(N)}(0) \xrightarrow{N \rightarrow \infty} \mathbf{x}^\infty \in E^1$ and $\lim_{N \rightarrow \infty} p_{c,N} = p_c > 0$ for every $c \in \mathcal{C}$. Then the sequence of Markov processes*

$$(\tilde{\mathbf{X}}^{(N)}(Nt)) \in D_E[0, \infty)$$

has a continuous limit $\mathbf{x}(t) \in C_E[0, \infty)$ which is determined by the unique solution of the initial-value problem

$$\frac{d\mathbf{x}(t)}{dt} = H^{fb}(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}^\infty, \quad (51)$$

where the function $H^{fb}(\cdot)$ is defined by

$$H^{fb}(\mathbf{x}) = \sum_{c \in \mathcal{C}} \frac{1}{p_c} \left(\lambda_c \sum_{n=0}^{K_c-1} x_{c,n} e_{c,n}^{c,n+1} + \pi_{\mathbf{x}}(\Omega_{-c}) \nu_c \sum_{n=1}^{K_c} x_{c,n} e_{c,n}^{c,n-1} \right).$$

Observe that when $K_c = \infty$ for every $c \in \mathcal{C}$ we recover Theorem 1 as a special case of Theorem 9.

Focus now on the analysis of the equilibrium points for (51). For any $\mathbf{x} \in E^1$, denote

$$\xi_c^{\mathbf{x}} = \frac{\lambda_c}{\nu_c \pi_{\mathbf{x}_0}^{fb}(\Omega_{-c})}.$$

Consider an equilibrium point $\mathbf{x}^* \in E^1$ for (51), i.e., $H^{fb}(\mathbf{x}^*) = 0$, then

$$x_{c,n}^* = (\xi_c^{\mathbf{x}^*})^n x_{c,0}^*, \quad \forall n = 0, \dots, K_c, c \in \mathcal{C},$$

and $x_{c,n}^*$ for every $n > K_c$. Hence, $\mathbf{x}^* \in E^1$ is an equilibrium point, if and only if

$$x_{c,0}^* = \frac{1}{\sum_{n=0}^{K_c} (\xi_c^{\mathbf{x}^*})^n}, \quad x_{c,n}^* = (\xi_c^{\mathbf{x}^*})^n x_{c,0}^* \mathbb{1}_{\{n \leq K_c\}}, \quad \forall n = 0, \dots, K_c, c \in \mathcal{C}.$$

We now identify sufficient and necessary conditions on $\boldsymbol{\xi} \in \mathbb{R}_+^{\mathcal{C}}$ such that the population vector $\mathbf{x}^\boldsymbol{\xi} \in E^1$ is an equilibrium point for (51), where

$$x_{c,0}^\boldsymbol{\xi} = \frac{1}{\sum_{n=0}^{K_c} \xi_c^n}, \quad x_{c,n}^\boldsymbol{\xi} = \xi_c^n x_{c,0}^\boldsymbol{\xi} \mathbb{1}_{\{n \leq K_c\}}.$$

Assume the population vector $\mathbf{x}^\boldsymbol{\xi} \in E^1$ to remain constant over time, i.e., $\mathbf{x}^\boldsymbol{\xi}$ is an equilibrium point for (51). Then, the fraction of time allocated to the activity of class- c nodes is given by

$$\tilde{\theta}_c(\boldsymbol{\xi}) = \frac{1}{Z^{fb}} \sum_{\omega \in \Omega_{+c}} \prod_{c' \in \mathcal{C}} \left(\frac{\nu_{c'}(1 - x_{c',0}^\boldsymbol{\xi})}{\mu_{c'}} \right)^{\omega_{c'}} = \theta_c(\mathbf{e} - \mathbf{x}^\boldsymbol{\xi}).$$

On the other hand, the effective stationary load of class- c nodes is given by

$$\tilde{\rho}_c(\boldsymbol{\xi}) = \frac{\lambda_c(1 - x_{c,K_c}^\boldsymbol{\xi})}{\mu_c},$$

i.e., a packet arriving at a class- c node with a full buffer is lost and does not enter the system so that the effective arrival rate at class- c nodes is $\lambda_c(1 - x_{c,K_c}^\xi)$. We conclude that $\mathbf{x}^\xi \in E^1$ is an equilibrium point for (51) if and only if

$$\tilde{\theta}(\xi) = \tilde{\rho}(\xi). \quad (52)$$

Consider the single-class case, i.e., $C = 1$, and drop the c subscript. Observe that

$$\tilde{\theta}(\xi) = \frac{\frac{\nu}{\mu} \left(1 - \frac{1}{\sum_{n=0}^K \xi^n}\right)}{1 + \frac{\nu}{\mu} \left(1 - \frac{1}{\sum_{n=0}^K \xi^n}\right)},$$

so that $\tilde{\theta}(0) = 0$, $\lim_{\xi \rightarrow \infty} \tilde{\theta}(\xi) = (\nu/\mu)/(1 + \nu/\mu) > 0$, and $\frac{\partial}{\partial \xi} \tilde{\theta}(\xi) > 0$. On the other hand,

$$\tilde{\rho}(\xi) = \frac{\lambda}{\mu} \left(1 - \frac{\xi^K}{\sum_{n=0}^K \xi^n}\right),$$

and $\tilde{\rho}(0) = \lambda/\mu > 0$, $\lim_{\xi \rightarrow \infty} \tilde{\rho}(\xi) = 0$, and $\frac{\partial}{\partial \xi} \tilde{\rho}(\xi) < 0$. Hence, the uniqueness of ξ satisfying (52) is ensured for every λ, μ, ν, K , and in turn yields the uniqueness of the equilibrium point for the initial-value problem (51).

8 Numerical experiments

In this section we perform numerical experiments in order to complement the analytical results obtained in the previous sections. In particular, we provide evidence that the results proved in Sections 5-7 for complete interference graphs in fact extend to more general graphs.

Throughout this section we focus on the square network displayed in Figure 1 and considered at the end of Section 4 with the following parameters:

$$\lambda = (0.4, 0.2, 0.3, 0.4), \quad \nu = (4, 3, 3, 5), \quad \mu = (1, 1, 1, 1),$$

and an equal number of nodes in each class, i.e., $p_c = 1/4$ for $c = 1, 2, 3, 4$. First of all, we verify that $\xi = \boldsymbol{\eta}(\boldsymbol{\rho}) < \mathbf{e}$ to ensure that a unique equilibrium point \mathbf{x}^* of (6) exists, by virtue of Theorem 2. In view of (36), we need to solve the system of equations

$$\xi_c \nu_c (1 + \xi_{5-c} \nu_{5-c}) = \lambda_c Z(\xi) \quad c = 1, 2, 3, 4,$$

where

$$Z(\xi) = 1 + \sum_{c=1}^4 \xi_c \nu_c + \sum_{c=1}^2 \xi_c \xi_{5-c} \nu_c \nu_{5-c}.$$

Observe that the above system of equations cannot be solved explicitly in general. Hence, we use the numerical algorithm described in [15], which is an adaptation of the algorithm presented in [41], and obtain

$$\xi \approx (0.4302, 0.2635, 0.6537, 0.3442).$$

Theorem 2 thus yields that

$$\mathbf{x}^* = (x_{c,n}^*), \quad x_{c,n}^* = (1 - \xi_c) \xi_c^n,$$

is the unique equilibrium point.

Convergence of population process. In Theorem 1 we proved the convergence of the properly scaled population process $\{\mathbf{X}^{(N)}(Nt)\}$ to the solution $\mathbf{x}(t)$ of (6) for arbitrary interference graphs. In Proposition 6 we showed global stability of \mathbf{x}^* , the unique fixed point of (6), for complete interference graphs. In Figure 3 we display a numerical solution of (6) and observe that $\mathbf{x}(t) \rightarrow \mathbf{x}^*$. This supports the hypothesis that the solution $\mathbf{x}(t)$ of (6) converges to \mathbf{x}^* even when the interference

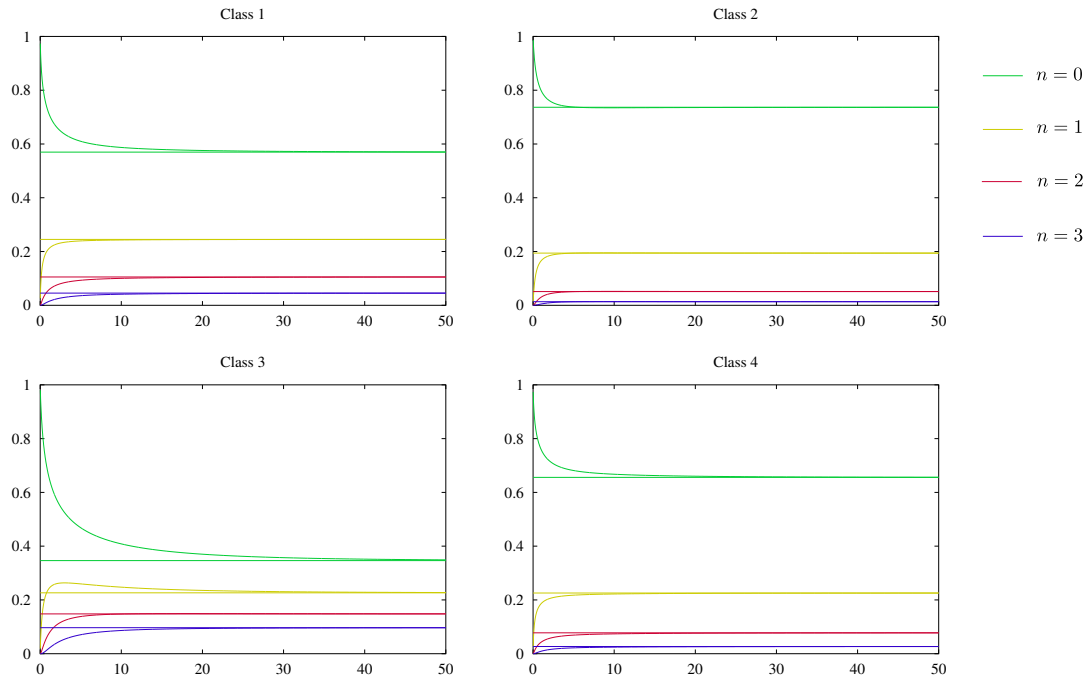


Figure 3: Plot of the solution $x_{c,n}(t)$ of (6) for $n = 0, 1, 2, 3$.

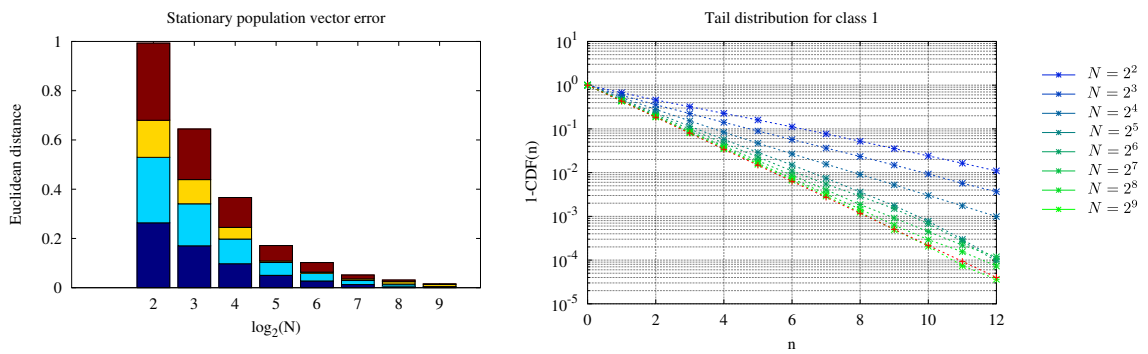


Figure 4: Difference between $\tilde{\mathbf{x}}^{(N)}$ and \mathbf{x}^* . (a) Euclidean distance. (b) Tail distributions for class 1.

graph is not complete.

Convergence of the stationary distribution of the population and the queue length process. In Theorem 3 we proved that $\lim_{N \rightarrow \infty} \mathbf{x}^{(N)}(\infty) = \mathbf{x}^*$, where $\mathbf{x}^{(N)}(\infty)$ is a random variable with the stationary distribution of $\mathbf{X}^{(N)}(t)$ as $t \rightarrow \infty$, for complete interference graphs. Also, in Theorem 5, we showed that

$$\lim_{N \rightarrow \infty} \mathbb{P}\{Q_c^{(N)}(\infty) = n\} = x_{c,n}^*,$$

where $Q_c^{(N)}(\infty)$ is a random variable with the stationary distribution of the queue-length $Q_c^{(N)}$ at a class- c node as $t \rightarrow \infty$. In order to examine these two properties, we conducted 10000 simulation runs and collected 10000 observations $(\tilde{\mathbf{x}}_c^{(N),h})_{h=1,\dots,10000}$ of the population process for large t and values $N = 2^l$ for $l = 2, \dots, 9$. We use the average observed distribution of the population process for large t in a network with N nodes, i.e.,

$$\tilde{\mathbf{x}}^{(N)} = \frac{1}{10000} \sum_{h=1}^{10000} \tilde{\mathbf{x}}^{(N),h}$$

as an empirical estimate for $\mathbf{x}^{(N)}$, and

$$CDF_c^{(N)}(n) = \sum_{m=n}^{\infty} \tilde{x}_{c,m}^{(N)}$$

as a proxy for

$$\lim_{t \rightarrow \infty} \mathbb{P} \sum_{k=m}^{\infty} \{Q_c^{(N)}(t) = k\}.$$

In Figure 4 we display in the bar plot the euclidean distance $d_c(\mathbf{x}_c^*, \tilde{\mathbf{x}}_c^{(N)})$ between the asymptotic approximation \mathbf{x}_c^* and the empirical estimate $\tilde{\mathbf{x}}_c^{(N)}$ for values $N = 2^l$ with $l = 2, \dots, 9$, where

$$d_c(\mathbf{x}_c^*, \tilde{\mathbf{x}}_c^{(N)}) = \sqrt{\sum_{n=0}^{\infty} (x_{c,n}^* - \tilde{x}_{c,n}^{(N)})^2}.$$

The different colors in the plot correspond to the various classes. Observe that the distance is small even for moderate values of N . In the plot on the right we display the empirical estimate $CDF_1^{(N)}(n)$ along with the asymptotic value $(\xi_1^*)^n$ for $n = 0, \dots, 12$ and $N = 2^l$ with $l = 2, \dots, 9$. Observe that as N increases, the accuracy of the asymptotic approximation $(\xi_1^*)^n$ for the tail probability $\mathbb{P}\{Q_1^{(N)}(\infty) \geq n\}$ (in red) improves, and the difference becomes almost indiscernible. In order to save space, we do not display the tail distribution convergence for the classes $c = 2, 3, 4$ which show similar behavior. These observations suggest that the convergence of the stationary distribution of the population process and the queue length holds in non-complete interference graphs as well, and that the asymptotic values provide accurate approximations even for moderate values of N .

Convergence of mean waiting time. In Theorem 5 we proved that a properly scaled version of the mean waiting time $\lambda_c \mathbb{E}[W_c^{(N)}]/N_c$ converges to $\xi_c/(1 - \xi_c)$ as $N \rightarrow \infty$ for complete interference graphs. In order to obtain an empirical estimate for the mean waiting time from the simulation runs, we invoke Little's law

$$\lambda_c \mathbb{E}[\tilde{W}_c^{(N)}]/N_c = \tilde{Y}_c^{(N)} = \frac{1}{10000} \sum_{h=1}^{10000} \sum_{n=0}^{\infty} n \tilde{x}_c^{(N),h}.$$

In Figure 5 we display the empirical estimates $\mathbb{E}[\tilde{W}_c^{(N)}]/N_c$, $c = 1, 2, 3, 4$, along with the asymptotic approximations $\xi_c^*/\lambda_c(1 - \xi_c^*)$ for values $N = 2^l$ with $l = 2, \dots, 9$. The asymptotic approximations turn out to be extremely accurate even for $N = 2^6 = 64$, i.e., $N_c = 16$. The results in Figure 5 support the conjecture that the mean waiting-time approximations in Theorem 5 extend to non-complete interference graphs, and are accurate even for moderate values of N .

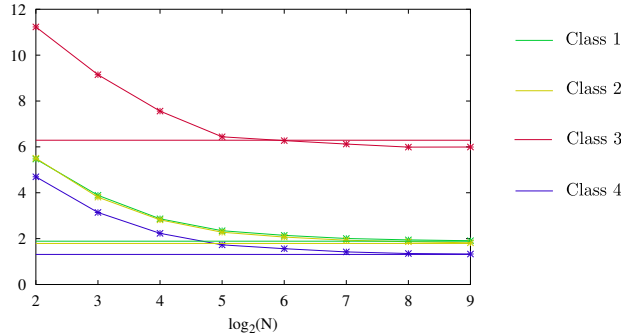


Figure 5: Comparison of $\mathbb{E}[\tilde{W}_c^{(N)}]/N_c$ with $\xi_c/\lambda_c(1 - \xi_c)$.

9 Conclusion

We analyzed distributed random-access networks in a many-sources regime with buffer-dynamics. The difficulty in the analysis is the complex interaction between the medium activity state and the buffer content process. However, in the mean-field regime, these processes are shown to decouple and a tractable initial-value problem describing the buffer content process is obtained. As illustrated through several examples, the methodological framework that we developed is fairly generic, and expected to be applicable to a broader range of problems where a similar decoupling occurs.

In the second part of the paper we focus on the analysis of the key performance measures for the case of a complete interference graph. The stationary buffer content and the packet waiting and sojourn time are shown to converge in distribution and expectation to random variables whose parameters depend on the unique equilibrium point of the initial-value problem and are explicitly provided. These results are obtained via an interchange of limits approach which ensures that the resulting approximations are asymptotically exact. Numerical experiments suggest that these results extend to general interference graphs, but a rigorous proof remains as a challenging problem for further research.

It would be of interest to develop a mean-field analysis for random-access networks in different regimes. In this work we assumed both arrival and back-off rates to scale as the number of nodes increases. Such scaling allows analytical tractability thanks to the decoupling between the medium activity state and the buffer content process. However, the back-off rate scaling entails a degradation in the delay performance which scales with the number of nodes as well. The study of large networks where only the arrival rate scales would be important since we expect this degradation would not happen. However, different analytical tools are required and we will investigate this problem in future work.

A Asymptotically Lipschitz in probability

In the course of obtaining the mean-field limit, we will use the property of a sequence in $D_E[0, \infty)$ of being asymptotically Lipschitz in probability and its relation with the relative compactness of a sequence of processes taking values in a metric space E with metric ρ and sample paths in D . Various versions of this property have been described in the literature, e.g., [13, 37], and we will use the following definition.

Definition 3. A sequence of random paths $\mathbf{X}^{(N)} \in D_E[0, \infty)$ is asymptotically Lipschitz in probability for a given $T > 0$ if there exists $K \geq 0$ such that for every $t_1, t_2 \in [0, T]$ and for every $\epsilon > 0$ the following is satisfied:

$$\mathbb{P}\{\mathbf{X}^{(N)} : \rho(\mathbf{X}^{(N)}(t_1), \mathbf{X}^{(N)}(t_2)) \leq K|t_2 - t_1| + \epsilon\} \xrightarrow{N \rightarrow \infty} 1.$$

We now reproduce [20, Theorem 3.7.2], and also refer to [8, Theorem 16.4], and the subsequent discussion.

Theorem 10 ([20] Theorem 3.7.2). *Let (E, ρ) be complete and separable, and let $\{\mathbf{X}^{(N)}\}$ be a family of processes with sample paths in $D_E[0, \infty)$. Then $\{\mathbf{X}^{(N)}\}$ is relatively compact if and only if the following two conditions hold:*

(a) *For every $\eta \in \mathbb{R}_{>0}$ and $b \in \mathbb{Q}_{\geq 0}$, there exists a compact set $\Gamma_{\eta, b} \subset E$ such that*

$$\inf_N \mathbb{P}\{\mathbf{X}^{(N)}(b) \in \Gamma_{\eta, b}\} \geq 1 - \eta.$$

(b) *For every $\eta \in \mathbb{R}_{>0}$ and $T \in \mathbb{R}_{>0}$, there exists $\delta \in \mathbb{R}_{>0}$ such that*

$$\sup_N \mathbb{P}\{w'(\mathbf{X}^{(N)}, \delta, T) \geq \eta\} \leq \eta,$$

where $w'(\mathbf{X}^{(N)}, \delta, T)$ represents the modulus of continuity, i.e.,

$$w'(\mathbf{X}^{(N)}, \delta, T) = \inf_{\{t_i\}_{i=0}^n} \max_i \sup_{s, t \in [t_{i-1}, t_i]} \rho(\mathbf{X}^{(N)}(s), \mathbf{X}^{(N)}(t))$$

and the sequence $\{t_i\}_{i=0}^n$ is nondecreasing and such that $t_{n-1} < T$.

Remark: In order to show relative compactness in the space $D_E[0, \infty)$ we therefore need to show tightness in $D_E[0, T]$ for each T . However, in many cases the space E is itself compact so that only condition (b) of Theorem 10 has to be shown. ([20, Theorem 6.3] characterizes the compact closure in $D_E[0, \infty)$.)

This leaves us with condition (b) and we therefore proceed to show the next proposition.

Proposition 11. *If $\mathbf{X}^{(N)} \in D_E[0, \infty)$ is asymptotically Lipschitz in probability, then $\mathbf{X}^{(N)} \in D_E[0, \infty)$ satisfies condition (b) in Theorem 10.*

Proof. Recall the definition of the modulus of continuity for paths in D . see [7, pages 109-110], and also [20, page 122]. Given $\delta > 0, T > 0$,

$$w'_{\mathbf{x}}(\delta) \doteq \inf_{\Pi} \max_{0 < i \leq n_{\Pi}} w_{\mathbf{x}}[t_{i-1}^{\Pi}, t_i^{\Pi}],$$

where Π is a partition of $[0, T]$ such that $0 = t_0^{\Pi} < t_1^{\Pi} < \dots < t_{n_{\Pi}}^{\Pi} = T$ with $|t_i^{\Pi} - t_{i-1}^{\Pi}| > \delta$ for every $i = 1, \dots, n_{\Pi}$, and

$$w_{\mathbf{x}}[s, t] \doteq \sup_{u, v \in [s, t]} (\rho(\mathbf{x}(u), \mathbf{x}(v))).$$

Consider condition (b) in Theorem 10 and assume η and T given. Note that it suffices to verify this condition for $N \geq N_0$, as [20, Lemma 6.2 (a)] shows that w' is right continuous and therefore we may choose $\delta < \delta_{N_0}$ to ensure (b) holds for the first $n \leq N_0$ cases.

Fix $\epsilon < \eta/2$. Now choose $N_0(\epsilon, \eta)$ such that

$$\mathbb{P}\{\mathbf{X}^{(N)} : \rho(\mathbf{X}^{(N)}(s_1), \mathbf{X}^{(N)}(s_2)) \leq K|s_2 - s_1| + \epsilon, \forall s_1, s_2 \in [0, T]\} > 1 - \eta$$

for every $N > N_0(\epsilon, \eta)$. Observe that $N_0(\epsilon, \eta)$ exists due to Definition 3. Now choose a uniform and sufficiently fine partition so that the spacing is δ_1 where $\delta_1 K < \eta/2$ and denote this partition as Π^{δ_1} . Observe that for $N > N_0(\epsilon, \eta)$,

$$\mathbb{P}\{\mathbf{X}^{(N)} : w_{\mathbf{X}^{(N)}}[t_{i-1}^{\Pi^{\delta_1}}, t_i^{\Pi^{\delta_1}}] \leq \underbrace{K\delta_1 + \epsilon}_{< \eta} > 1 - \eta$$

for every $i = 1, \dots, n^{\Pi^{\delta_1}}$. We now choose $\delta = \delta_1/2$ and observe that by definition of $w'(\mathbf{X}^{(N)}, \delta, T)$,

$$\mathbb{P}\{\mathbf{X}^{(N)} : w'(\mathbf{X}^{(N)}, \delta, T) \leq \eta\} > 1 - \eta.$$

Indeed, $w'(\mathbf{X}^{(N)}, \delta, T)$ is given by the infimum over the suitable partitions and therefore is certainly lower than the outcome obtained by considering the suitable partition Π^{δ_1} , hence

$$\mathbb{P}\{\mathbf{X}^{(N)} : w'(\mathbf{X}^{(N)}, \delta, T) \geq \eta\} < \eta.$$

For the remaining $n < N_0$ make δ smaller as described above to obtain the required result as w' is non-decreasing in δ . \square

B Extended proofs

B.1 Proof of Lemma 2

As previously stated, we consider the single-class case only. The following arguments are along the same lines as given in [30]. First observe that the unit Poisson processes for the arrivals, the back-offs, and the transmissions are all mutually independent and by assumption are defined on a common probability space which satisfies the usual conditions. Moreover, the arrival rates, back-off rates and the transmission rate are all bounded. In particular, the natural filtrations

$$\mathcal{H}_n^A(t) \doteq \sigma(\mathcal{N}_{A,n}(u), 0 \leq u \leq t), \forall n \in \mathbb{N}_0, \quad (53)$$

are defined with the σ -algebra generated by $\mathcal{N}_{A,n}(u)$ for all $0 \leq u \leq t$. We make corresponding definitions for the back-offs $\mathcal{H}_n^B(t)$, $n \in \mathbb{N}$, and for the transmission process denoted by $\mathcal{H}^T(t)$.

Given a multi-parameter time $\mathbf{s} = (s_0, s_1, s_2, \dots) \in [0, \infty)^\infty$, define $\mathcal{H}(\mathbf{s})$ to be the σ -algebra

$$\mathcal{H}(\mathbf{s}) \doteq \left[\bigvee_{n \in \mathbb{N}_0} \mathcal{H}_n^A(s_{2n+1}) \right] \vee \left[\bigvee_{n \in \mathbb{N}} \mathcal{H}_n^B(s_{2n}) \right] \vee \mathcal{H}^T(s_0) \vee \sigma(\mathcal{N}_{\text{Null}}^C),$$

where $\mathcal{N}_{\text{Null}}^C$ is the collection of all null sets in \mathcal{F}^C . With this construction we thus obtain the multi-parameter filtration

$$\mathcal{H} \doteq \{ \mathcal{H}(\mathbf{s}) : \mathbf{s} \in \mathbb{R}_+^\infty \},$$

see [20, page 85], where the metric space for the index set is \mathbb{R}_+^∞ with the usual metric. (We may suppose that this filtration is right continuous as otherwise we may work with its right continuous augmentation, see [34, Lemma 67.4] for example.) Also the process

$$\tilde{\mathcal{N}}_T, \tilde{\mathcal{N}}_{A,0}, \tilde{\mathcal{N}}_{A,1}, \tilde{\mathcal{N}}_{B,1}, \dots$$

is a right continuous multi-parameter martingale adapted to the filtration \mathcal{H} .

Since \mathbf{X}, Y are the unique solutions to (12)-(16), it follows that we can apply [20, Theorem 2.2, page 313], to show that τ is a multi-parameter \mathcal{H} -stopping time, where

$$\begin{aligned} \tau_n^A(t) &\doteq \int_0^t N \lambda X_n(s) ds, & n \in \mathbb{N}_0 \\ \tau_n^B(t) &\doteq \int_0^t N \nu X_n(s) ds, & n \in \mathbb{N} \\ \tau^T(t) &\doteq \int_0^t N \mu ds, \end{aligned}$$

see [20, Chapter 2, Section 8] for the definition.

As in [20, Proposition 8.6, page 87], define $\mathcal{F}(t) \doteq \mathcal{H}(\tau(t))$ where τ is the above \mathcal{H} stopping time. Since $\tau(t)$ is a right continuous function of t and $\mathcal{H}(\mathbf{s})$ is a right continuous filtration, it follows that if η is a $\mathcal{F}(t)$ stopping time, then so is $\tau(\eta)$.

It follows from the Optional Sampling Theorem for martingales on a metric lattice [20, Theorem 8.7, page 87] that \mathbf{M}_A is a right continuous \mathbf{F} multi-parameter martingale. This martingale is locally square integrable, because the rates are uniformly bounded over time. In particular if $T > 0$ the

version of these martingales stopped at T is a \mathcal{L}^2 \mathbf{F} -martingale. The same arguments show that the vector back-off process,

$$\left(\tilde{\mathcal{N}}_{\mathbf{B},n} \left(N \nu \int_0^t X_n(s) ds \right) \right)_{n \in \mathbb{N}}$$

is also a locally square integrable \mathbf{F} -martingale as of course is the process $\tilde{\mathcal{N}}_T \left(N \mu \int_0^t ds \right)$. However $Y \in m\mathcal{F}(t)$ which implies that \mathbf{M}_B is also a locally square integrable \mathbf{F} -martingale as the integrand is previsible and similarly for \mathbf{M}_T . Finally Ξ is adapted, being a countable sum of \mathbf{F} martingales, and it is straightforward to show that it is bounded in \mathcal{L}^2 and thus a \mathcal{L}^2 martingale.

B.2 Proof of Proposition 1

Observe that $\alpha^{(N)}(t)$ lies in the space of unit, increasing Lipschitz functions $L[0, T]$, where

$$L[0, T] = \left\{ l : [0, T] \rightarrow \mathbb{R}_+^{|\Omega|}, \|l(t_1) - l(t_2)\| \leq |t_1 - t_2| \right\},$$

so that $|l(t)| \leq t$. It follows from the Arzela-Ascoli theorem [7, page 221], that $L[0, T]$ has compact closure restricted to $[0, T]$, since the space $L[0, T]$ is compact within the topology of $C[0, T]$. It follows from [43, Lemma 2.1], (see also [44]) that we may extend this result to deduce that $L[0, \infty)$ is compact within $C_{\mathbb{R}_+^c, \uparrow}[0, \infty)$. Therefore a sample path of $\alpha^{(N)}$ lies in a compact set within $C_{\mathbb{R}_+^c, \uparrow}[0, \infty)$. This implies that the sequence of measures on $\alpha^{(N)}$ is tight.

We turn to the sequence of population processes $\mathbf{X}^{(N)}$. First observe that the space E is compact as a consequence of Tychonoff's Theorem as it is a finite product of compact spaces, see Lemma 1 and the preceding discussion. In Appendix A the concept of asymptotically Lipschitz in probability is introduced and its relation with relatively compactness is stated in Proposition 11. Lemma 4, which is stated and proved below implies that the sequence of population processes is asymptotically Lipschitz in probability.

Lemma 4. *The sequence of processes $(\mathbf{X}^{(N)})_N$ is asymptotically Lipschitz in probability.*

However since the space (E, ρ) is complete and separable, it follows from the converse half of Prohorov's Theorem that the sequence $\mathbf{X}^{(N)}$ is tight, [7, Theorem 6.2, page 37]. It follows, as already stated, that the sequence of measures $(\mathbf{X}^{(N)}, \alpha^{(N)})$ is therefore tight. Hence, the sequence $(\mathbf{X}^{(N)}, \alpha^{(N)})$ is relatively compact, again as a consequence of Prohorov's Theorem, [7, Theorem 6.2, page 37].

Proof. (of Lemma 4) Recall that $E = \chi^C$ and $C < \infty$, see (1). Focus on (25), and observe that its (c, n) component can be written as

$$\begin{aligned} X_{c,n}^{(N)}(t) = & X_{c,n}^{(N)}(0) - \frac{1}{N_c} \tilde{\mathcal{N}}_{\mathbf{A},c,n} \left(N \lambda_c \int_0^t X_{c,n}^{(N)}(s) ds \right) \\ & + \frac{1}{N_c} \tilde{\mathcal{N}}_{\mathbf{A},c,n-1} \left(N \lambda_c \int_0^t X_{c,n-1}^{(N)}(s) ds \right) \\ & - \frac{1}{N_c} \tilde{\mathcal{N}}_{\mathbf{B},c,n} \left(N \nu_c \int_0^t X_{c,n}^{(N)}(s) \mathbb{1}\{Y_c^{(N)}(Ns) \in \Omega_{-c}\} ds \right) \\ & + \frac{1}{N_c} \tilde{\mathcal{N}}_{\mathbf{B},c,n+1} \left(N \nu_c \int_0^t X_{c,n+1}^{(N)}(s) \mathbb{1}\{Y_c^{(N)}(Ns) \in \Omega_{-c}\} ds \right) \\ & + \frac{1}{p_{c,N}} \left(\lambda_c \int_0^t X_{c,n-1}^{(N)}(s) ds + \nu_c \int_0^t X_{c,n+1}^{(N)}(s) \mathbb{1}\{Y_c^{(N)}(Ns) \in \Omega_{-c}\} ds \right) \\ & - \lambda_c \int_0^t X_{c,n}^{(N)}(s) ds + \nu_c \int_0^t X_{c,n}^{(N)}(s) \mathbb{1}\{Y_c^{(N)}(Ns) \in \Omega_{-c}\} ds, \end{aligned} \quad (54)$$

where the centered Poisson processes are martingales bounded in \mathcal{L}^2 . We recall that Doob's submartingale inequality, see [33, Theorem 1.7], yields that

$$\mathbb{P}\left\{\sup_{t \in [0, T]} |\tilde{\mathcal{N}}_{\cdot, c, n}(t)| \geq N\eta\right\} \leq \frac{\mathbb{E}[\tilde{\mathcal{N}}_{\cdot, c, n}^2(T)]}{N^2\eta^2}, \quad (55)$$

and it follows that $\tilde{\mathcal{N}}_{\cdot, c, n}$ are non-negative submartingales. Also,

$$\mathbb{E}[\mathcal{N}_{A, c, n}(N\lambda_c \int_0^t X_{c, n}^{(N)}(s) ds) | \mathcal{F}_t] = N\lambda_c \int_0^t X_{c, n}^{(N)}(s) ds,$$

and

$$\begin{aligned} & \mathbb{E}[\mathcal{N}_{A, c, n}^2(N\lambda_c \int_0^t X_{c, n}^{(N)}(s) ds) | \mathcal{F}_t] \\ &= (N\lambda_c \int_0^t X_{c, n}^{(N)}(s) ds)^2 + N\lambda_c \int_0^t X_{c, n}^{(N)}(s) ds, \end{aligned}$$

yielding that

$$\mathbb{E}[\tilde{\mathcal{N}}_{A, c, n}^2(N\lambda_c \int_0^t X_{c, n}^{(N)}(s) ds) | \mathcal{F}_t] = N\lambda_c \int_0^t X_{c, n}^{(N)}(s) ds \leq N\lambda_c t,$$

and a similar argument applies for $\mathcal{N}_{B, c, n}$.

From (54) we obtain that

$$\begin{aligned} & |X_{c, n}^{(N)}(t) - X_{c, n}^{(N)}(u)| \\ & \leq |Z_{c, n}^{(N)} + Z_{c, n-1}^{(N)}| + |V_{c, n}^{(N)} + V_{c, n+1}^{(N)}| + \frac{2}{p_{c, N}}(\lambda_c + \nu_c)|t - u| \end{aligned} \quad (56)$$

for every $t, u \in [0, T]$, where $Z_{c, n}^{(N)}$ and $V_{c, n}^{(N)}$ are the maximum values of the martingales in magnitude over $[0, T]$. Observe that (55) yields

$$\mathbb{P}\{Z_{c, n}^{(N)}(T) > \eta\} \leq \frac{\lambda_c T}{N\eta^2}, \quad \mathbb{P}\{V_{c, n}^{(N)}(T) > \eta\} \leq \frac{\nu_c T}{N\eta^2}. \quad (57)$$

Denote $T_{c, n}^{(N)} = |Z_{c, n}^{(N)} + Z_{c, n-1}^{(N)}| + |V_{c, n}^{(N)} + V_{c, n+1}^{(N)}|$.

Now, fix $\epsilon > 0$ and choose $n(\epsilon)$ so that

$$\sum_{n=n(\epsilon)+1}^{\infty} 2^{-n} < \frac{\epsilon}{2C}.$$

From (56), we obtain

$$\rho(\mathbf{X}^{(N)}(t), \mathbf{X}^{(N)}(u)) \leq \sum_{c \in \mathcal{C}} \left(\sum_{n=0}^{n(\epsilon)} 2^{-n} |T_{c, n}^{(N)}| + \frac{4}{p_{c, N}}(\lambda_c + \nu_c)|t - u| + \frac{\epsilon}{2C} \right)$$

for every $t, u \in [0, T]$. Observe that, thanks to (57), it holds that

$$\mathbb{P}\{|T_{c, n}^{(N)}| > \epsilon/2n(\epsilon)\} \xrightarrow{N \rightarrow \infty} 0.$$

Hence,

$$\lim_{N \rightarrow \infty} \mathbb{P}\{\mathbf{X}^{(N)} : \rho(\mathbf{X}^{(N)}(t), \mathbf{X}^{(N)}(u)) \leq \epsilon + 4 \sum_{c \in \mathcal{C}} \frac{1}{p_c}(\lambda_c + \nu_c)|t - u|\} \rightarrow 1$$

for every $t, u \in [0, T]$. □

B.3 Proof of Proposition 3

As shown in Lemma 2, the process Ξ is a multi-parameter martingale adapted to \mathbf{F} . Since C is finite, there exists K_η such that

$$\sum_{c \in \mathcal{C}} \sum_{n > K_\eta} \frac{|\Xi_{c,n}^{(N)}|}{1 + |\Xi_{c,n}^{(N)}|} < \eta/2.$$

For each $c \in \mathcal{C}, n \leq K_\eta$, $\Xi_{c,n}^{(N)}(t)$ is a \mathcal{L}^2 \mathbf{F} martingale so that $(\Xi^{(N)}(t))^2$ is a \mathcal{L}^1 \mathbf{F} submartingale. Applying Doob's submartingale inequality with $p = 1$ (see [33]), we obtain

$$\mathbb{P} \left\{ \sup_{t \in [0, T]} |\Xi_{c,n}^{(N)}(t)|^2 > \varepsilon \right\} \leq \frac{\mathbb{E} \left\{ |\Xi_{c,n}^{(N)}(T)|^2 \right\}}{\varepsilon}, \quad c \in \mathcal{C}, n \leq K_\eta. \quad (58)$$

However $\mathbb{E} \left\{ |\Xi_{c,n}^{(N)}(T)|^2 \right\} = O(1/N)$ from which the result follows.

B.4 Proof of Proposition 4

Consider

$$\mathbf{X}^{(N)} \rightarrow \mathbf{x}, \quad \alpha_0^{(N)} \rightarrow \alpha_0,$$

in $D_E[0, \infty)$ and $C[0, \infty)$ respectively. Note that $x_{c,n}^{(N)} \in D_{[0,1]}[0, \infty)$ has only countably many jumps in a finite interval $[0, t]$, see [20, Lemma 5.1, page 116], and is bounded by definition. Hence, $\int_0^t x_{c,n}^{(N)}(s) ds$ exists for all $t \geq 0$. Moreover, at all points of continuity of $x_{c,n}(s)$ we have that $x_{c,n}^{(N)}(s) \rightarrow x_{c,n}(s)$. Therefore, for every $t \geq 0$

$$\int_0^t x_{c,n}^{(N)}(s) ds \rightarrow \int_0^t x_{c,n}(s) ds$$

by dominated convergence, since $x_{c,n}^{(N)}$ is bounded to be in $[0, 1]$.

Now we show that for every $t \geq 0$

$$\int_0^t x_{c,n}^{(N)}(s) \dot{\alpha}_\omega^{(N)}(s) ds \rightarrow \int_0^t x_{c,n}(s) \dot{\alpha}_\omega(s) ds.$$

First of all, these integrals exist since $x_{c,n}^{(N)}(s)$ may be approximated by step functions $h^{(m)}(s)$ on $[0, t]$ converging to $x_{c,n}^{(N)}(s)$ almost everywhere. Then

$$\int_0^t h^{(m)}(s) \dot{\alpha}_\omega^{(N)}(s) ds$$

exists, so that the existence of

$$\int_0^t x_{c,n}^{(N)}(s) \dot{\alpha}_\omega^{(N)}(s) ds$$

follows by dominated convergence. We can conclude similarly that

$$\int_0^t x_{c,n}(s) \dot{\alpha}_\omega(s) ds$$

exists.

Consider the difference

$$\begin{aligned} & \int_0^t x_{c,n}^{(N)}(s) \dot{\alpha}_\omega^{(N)}(s) ds - \int_0^t x_{c,n}(s) \dot{\alpha}_\omega(s) ds \\ &= \int_0^t (x_{c,n}^{(N)}(s) - x_{c,n}(s)) \dot{\alpha}_\omega^{(N)}(s) ds + \int_0^t x_{c,n}(s) \dot{\alpha}_\omega^{(N)}(s) ds - \int_0^t x_{c,n}(s) \dot{\alpha}_\omega(s) ds. \end{aligned}$$

Then the first term

$$\begin{aligned} \left| \int_0^t (x_{c,n}^{(N)}(s) - x_{c,n}(s)) \dot{\alpha}_\omega^{(N)}(s) ds \right| &\leq \int_0^t |x_{c,n}^{(N)}(s) - x_{c,n}(s)| \dot{\alpha}_\omega^{(N)}(s) ds \\ &\leq \int_0^t |x_{c,n}^{(N)}(s) - x_{c,n}(s)| ds \rightarrow 0 \end{aligned}$$

by dominated convergence since $x_{c,n}^{(N)}(s) \rightarrow x_{c,n}(s)$ at all but countably many points. As for the second term

$$\left| \int_0^t x_{c,n}(s) \dot{\alpha}_\omega^{(N)}(s) ds - \int_0^t x_{c,n}(s) \dot{\alpha}_\omega(s) ds \right|,$$

observe that for any step function $h(s)$,

$$\left| \int_0^t h(s) \dot{\alpha}_\omega^{(N)}(s) ds - \int_0^t h(s) \dot{\alpha}_\omega(s) ds \right| \rightarrow 0,$$

since $|\dot{\alpha}_\omega^{(N)}(s) - \dot{\alpha}_\omega(s)| \rightarrow 0$ uniformly in $[0, t]$. Now choose a step function $h_\epsilon(s)$ such that $|x_{c,n}(s) - h_\epsilon(s)| < \epsilon$ for every $s \in [0, t]$ ($h_\epsilon(s)$ exists as shown by [7, page 110]). Hence,

$$\begin{aligned} &\left| \int_0^t x_{c,n}(s) \dot{\alpha}_\omega^{(N)}(s) ds - \int_0^t x_{c,n}(s) \dot{\alpha}_\omega(s) ds \right| \\ &\leq \left| \int_0^t h_\epsilon(s) \dot{\alpha}_\omega^{(N)}(s) ds - \int_0^t h_\epsilon(s) \dot{\alpha}_\omega(s) ds \right| + 2 \int_0^t |x_{c,n}(s) - h_\epsilon(s)| ds, \end{aligned}$$

and taking limits as $N \rightarrow \infty$, we obtain

$$\lim_{N \rightarrow \infty} \left| \int_0^t x_{c,n}(s) \dot{\alpha}_\omega^{(N)}(s) ds - \int_0^t x_{c,n}(s) \dot{\alpha}_\omega(s) ds \right| \leq 2t\epsilon.$$

However, t is fixed and ϵ is arbitrary, hence

$$\lim_{N \rightarrow \infty} \left| \int_0^t x_{c,n}(s) \dot{\alpha}_\omega^{(N)}(s) ds - \int_0^t x_{c,n}(s) \dot{\alpha}_\omega(s) ds \right| \rightarrow 0.$$

It follows that for every $t \geq 0$

$$\int_0^t x_{c,n}^{(N)}(s) \dot{\alpha}_\omega^{(N)}(s) ds \rightarrow \int_0^t x_{c,n}(s) \dot{\alpha}_\omega(s) ds.$$

B.5 Proof of Lemma 3

Given $\epsilon > 0$, we may find a partition of $[0, T]$, $t_0 = 0 < t_1 < \dots < t_K = T$ such that,

$$0 \leq u(t_k) - u(t_{k-1}) \leq \epsilon/2, \quad k = 1, \dots, K,$$

since u is Lipschitz continuous by assumption.

Next given any $t \in (0, T)$ the following inequality holds, for some $k < K$,

$$|u^{(N)}(t) - u(t)| \leq |u^{(N)}(t_k) - u(t_{k+1})| \vee |u^{(N)}(t_{k+1}) - u(t_k)|,$$

since if on the one hand, $u^{(N)}(t) \leq u(t)$, then the first term on the right is a bound and on the other hand if the inequality is reversed then the second term is a bound. Using the triangle inequality on each respective term, we obtain

$$\begin{aligned} &|u^{(N)}(t) - u(t)| \\ &\leq \left(|u^{(N)}(t_k) - u(t_k)| \vee |u^{(N)}(t_{k+1}) - u(t_{k+1})| \right) + u(t_{k+1}) - u(t_k). \end{aligned}$$

Since the partition is fixed, we can find N_ϵ so that the right hand side is smaller than ϵ using pointwise convergence and the properties of the partition. Clearly, this N_ϵ can be taken to cover the case $t = 0, T$ as well and the result is proved.

B.6 Proof of Proposition 5

For notational convenience, the proof is only presented for the single-class case, but the arguments easily extend to the multi-class case.

Consider solutions $\mathbf{x}^L(t)$ and $\mathbf{x}^U(t)$ of the initial-value problem (39). These solutions are continuous and differentiable in t , therefore it is sufficient to show that, when the solutions are on the boundary of the stochastic dominance region, the solutions are forced to remain inside such region. Specifically, consider a time $t \geq t_0$ such that the stochastic dominance relation $\mathbf{x}^L(t) \leq_s \mathbf{x}^U(t)$ holds, i.e., the solutions are inside the stochastic dominance region, and there exists $\bar{k} \in \mathbb{N}_0$ such that

$$\sum_{n=0}^{\bar{k}} x_n^L(t) = \sum_{n=0}^{\bar{k}} x_n^U(t), \quad (59)$$

i.e., the solutions are on the boundary of the region, then we need to show that

$$\frac{\partial}{\partial t} \left(\sum_{n=0}^{\bar{k}} x_n^L(t) - \sum_{n=0}^{\bar{k}} x_n^U(t) \right) \geq 0,$$

i.e., the solutions are forced to remain inside the stochastic dominance region.

From (39) we obtain

$$\begin{aligned} & \frac{\partial}{\partial t} \left(\sum_{n=0}^{\bar{k}} x_n^L(t) - \sum_{n=0}^{\bar{k}} x_n^U(t) \right) \\ &= -\lambda x_{\bar{k}}^L(t) + \nu \pi_{\mathbf{x}_0^L(t)}(0) x_{\bar{k}+1}^L(t) + \lambda x_{\bar{k}}^U(t) - \nu \pi_{\mathbf{x}_0^U(t)}(0) x_{\bar{k}+1}^U(t) \\ &\geq -\lambda x_{\bar{k}}^L(t) + \nu \pi_{\mathbf{x}_0^U(t)}(0) x_{\bar{k}+1}^L(t) + \lambda x_{\bar{k}}^U(t) - \nu \pi_{\mathbf{x}_0^U(t)}(0) x_{\bar{k}+1}^U(t) \\ &= \lambda (x_{\bar{k}}^U(t) - x_{\bar{k}}^L(t)) + \nu \pi_{\mathbf{x}_0^U(t)}(0) (x_{\bar{k}+1}^L(t) - x_{\bar{k}+1}^U(t)), \end{aligned} \quad (60)$$

and in fact $\pi_{\mathbf{x}_0^L(t)}(0) \geq \pi_{\mathbf{x}_0^U(t)}(0)$ since $x_0^L(t) \geq x_0^U(t)$.

Observe that the stochastic dominance property and (59) yield

$$\begin{cases} \sum_{n=0}^{\bar{k}-1} x_n^L(t) \geq \sum_{n=0}^{\bar{k}-1} x_n^U(t), \\ \sum_{n=0}^{\bar{k}} x_n^L(t) = \sum_{n=0}^{\bar{k}} x_n^U(t), \\ \sum_{n=0}^{\bar{k}+1} x_n^L(t) \geq \sum_{n=0}^{\bar{k}+1} x_n^U(t), \end{cases}$$

and therefore

$$x_{\bar{k}+1}^L(t) - x_{\bar{k}+1}^U(t) \geq 0, \quad x_{\bar{k}}^L(t) - x_{\bar{k}}^U(t) \leq 0.$$

Hence, the nonnegativity of (60) follows.

B.7 Proof of Proposition 6

For notational convenience, the proof is only presented for the single-class case, but the arguments easily extend to the multi-class case.

Define $\mathbf{x}^0(t)$ as the solution of the initial-value problem (39) with an *empty* initial state, i.e., $x_0^0(0) = 1$ and $x_n^0(0) = 0$ for every $n \in \mathbb{N}$. Proposition 5 implies that $\mathbf{x}^0(t) \leq_s \mathbf{x}(t)$ for any solution $\mathbf{x}(t)$ of (39).

Assume now that $\mathbf{x}(0) \in E_{<\infty}^1$, we will prove that the solution $\mathbf{x}(t)$ of the initial-value problem (39) is then such that $x_n(t) \rightarrow x_n^*$ for every $n \in \mathbb{N}_0$. The proof relies on induction on n , and the base case is proved in the following lemma.

Lemma 5. *Consider a solution $\mathbf{x}(t) = (x_n(t))_{n \geq 0}$ of (39) with $\mathbf{x}(0) \in E_{<\infty}^1$. It holds that $x_0(t) \rightarrow x_0^*$.*

Proof. Define $\Delta(t) \doteq m(\mathbf{x}(t)) - m(\mathbf{x}^0(t))$, and observe that $\Delta(0) < \infty$ because $\mathbf{x}(0) \in E_{<\infty}^1$. Thanks to the fundamental theorem of calculus,

$$\Delta(T) = \Delta(0) - \int_0^T (\dot{m}(\mathbf{x}^0(t)) - \dot{m}(\mathbf{x}(t))) dt = \Delta(0) - \Delta(0, T),$$

where

$$\Delta(t_1, t_2) = \int_{t_1}^{t_2} (\dot{m}(\mathbf{x}^0(t)) - \dot{m}(\mathbf{x}(t))) dt.$$

Observe that, according to (42), $\Delta(t) \geq 0$ and $\dot{m}(\mathbf{x}^0(t)) - \dot{m}(\mathbf{x}(t)) \geq 0$ for every $t \geq 0$, therefore $\lim_{T \rightarrow \infty} \Delta(0, T) \leq \Delta(0)$. By means of a contradiction argument, we deduce that

$$\dot{m}(\mathbf{x}(t)) \rightarrow \dot{m}(\mathbf{x}^0(t)).$$

In fact, assume there exist $\epsilon > 0$ and two intertwined sequences $\{t_n^\epsilon\}_{n \in \mathbb{N}}$, $\{\bar{t}_n^\epsilon\}_{n \in \mathbb{N}}$, where $\bar{t}_n^\epsilon \in [t_n^\epsilon, t_{n+1}^\epsilon]$, such that for every $n \in \mathbb{N}$

1. $\dot{m}(\mathbf{x}^0(t_n^\epsilon)) - \dot{m}(\mathbf{x}(t_n^\epsilon)) > \epsilon$,
2. $\dot{m}(\mathbf{x}^0(\bar{t}_n^\epsilon)) - \dot{m}(\mathbf{x}(\bar{t}_n^\epsilon)) = \epsilon/2$,
3. $\dot{m}(\mathbf{x}^0(t)) - \dot{m}(\mathbf{x}(t)) \geq \epsilon/2$, for every $t \in [t_n^\epsilon, \bar{t}_n^\epsilon]$.

Since $\dot{m}(\mathbf{x}^0(t)) - \dot{m}(\mathbf{x}(t)) \geq 0$ for every $t \geq 0$, we have that

$$\lim_{T \rightarrow \infty} \Delta(0, T) \geq \sum_{n=0}^{\infty} \Delta(t_n^\epsilon, \bar{t}_n^\epsilon).$$

Note that each term of the sum is strictly positive. In fact, since the function $\dot{m}(\cdot)$ taking values in E^1 is Lipschitz continuous,

$$\Delta(t_n^\epsilon, \bar{t}_n^\epsilon) \geq \frac{\epsilon}{2} (\bar{t}_n^\epsilon - t_n^\epsilon) > 0.$$

Hence,

$$\lim_{T \rightarrow \infty} \Delta(0, T) = \infty > \Delta(0),$$

which is a contradiction. By letting $\epsilon \downarrow 0$, we conclude that $\dot{m}(\mathbf{x}^0(t)) \rightarrow \dot{m}(\mathbf{x}(t))$.

Since $\dot{m}(\mathbf{x}^0(t)) \rightarrow \dot{m}(\mathbf{x}(t))$, no matter the choice of $\mathbf{x}(0) \in E_{<\infty}^1$, it is legitimate to pick $\mathbf{x}(0) = \mathbf{x}^* \in E_{<\infty}^1$. In this way,

$$\dot{m}(\mathbf{x}^0(t)) \rightarrow \dot{m}(\mathbf{x}^*) = 0,$$

and due to the transitive property of the convergence relation,

$$\dot{m}(\mathbf{x}(t)) \rightarrow 0$$

for every solution $\mathbf{x}(t)$ such that $\mathbf{x}(0) \in E_{<\infty}^1$.

Now, due to (41), we may write $\dot{m}(\mathbf{x}(t))$ explicitly, so that

$$\lambda - \nu(1 - x_0(t))\pi_{\mathbf{x}_0(t)}(0) \rightarrow 0,$$

or equivalently,

$$x_0(t) \rightarrow 1 - \frac{\lambda}{\nu\pi_{\mathbf{x}_0(t)}(0)}. \tag{61}$$

Plugging $\pi_{\mathbf{x}_0(t)}(0) = \frac{1}{1 + \frac{\nu}{\mu}(1 - x_0(t))}$ in (61), we then conclude that

$$x_0(t) \rightarrow 1 - \frac{\lambda}{\nu(1 - \frac{\lambda}{\mu})} = 1 - \xi = x_0^*.$$

□

The inductive step is proved in the following lemma.

Lemma 6. *Consider a solution $\mathbf{x}(t) = (x_n(t))_{n \geq 0}$ of (39) with $\mathbf{x}(0) \in E_{< \infty}^1$ and such that $\mathbf{x}_m(t) \rightarrow \mathbf{x}_m^*$ for every $m \leq n - 1$. It holds that $\mathbf{x}_n(t) \rightarrow \mathbf{x}_n^*$.*

Proof. According to (6),

$$\dot{x}_0(t) = -R_1(t), \quad \dot{x}_{m-1}(t) = R_{m-1}(t) - R_m(t), \quad \text{for } m \geq 2,$$

where

$$R_m(t) = \lambda x_{m-1}(t) - \nu \pi_{\mathbf{x}_0(t)}(0) x_m(t).$$

The fact that $\mathbf{x}_m(t) \rightarrow \mathbf{x}_m^*$ for every $m \leq n - 1$ implies that there exists T_ϵ such that for every $t \geq T_\epsilon$, we have that $|\pi_{\mathbf{x}_0(t)}(0) - \pi_{\mathbf{x}_0^*}(0)| \leq \epsilon$ and $|x_m(t) - x_m^*| \leq \epsilon$ for every $m = 0, \dots, n - 1$. Consider now $t \geq T_\epsilon$, then there exist functions $(\tilde{g}_m(\epsilon))_{m=2, \dots, n}$ such that

$$|R_{m-1}(t)| \leq \tilde{g}_m(\epsilon),$$

and $\lim_{\epsilon \rightarrow 0} \tilde{g}_m(\epsilon) = 0$. Define $\tilde{g}_1(\epsilon) = 0$.

Recall that the fundamental theorem of calculus yields

$$x_{m-1}(t+h) - x_{m-1}(t) = \int_t^{t+h} \dot{x}_{m-1}(s) ds.$$

Now consider $t \geq T_\epsilon$, and note that

$$-2\epsilon - \tilde{g}_m(\epsilon) \leq \int_t^{t+h} R_m(s) ds \leq 2\epsilon + \tilde{g}_m(\epsilon),$$

for every $m = 1, \dots, n$. Therefore, for $m = n$, we obtain that

$$\frac{-2\epsilon - \tilde{g}_n(\epsilon) + \lambda(x_{n-1}^* - \epsilon)h}{\nu(\pi_{\mathbf{x}_0^*}(0) + \epsilon)} \leq \int_t^{t+h} x_n(s) ds \leq \frac{2\epsilon + \tilde{g}_n(\epsilon) + \lambda(x_{n-1}^* + \epsilon)h}{\nu(\pi_{\mathbf{x}_0^*}(0) - \epsilon)},$$

which may be rewritten as

$$h \frac{\lambda x_{n-1}^*}{\nu \pi_{\mathbf{x}_0^*}(0)} - L_1(\epsilon, h) \leq \int_t^{t+h} x_n(s) ds \leq h \frac{\lambda x_{n-1}^*}{\nu \pi_{\mathbf{x}_0^*}(0)} + U_1(\epsilon, h),$$

where $L_1(\epsilon, h), U_1(\epsilon, h) \xrightarrow{\epsilon \rightarrow 0} 0$. At this point, observe that $\frac{\lambda x_{n-1}^*}{\nu \pi_{\mathbf{x}_0^*}(0)} = x_n^*$ due to the definition of \mathbf{x}^* in Corollary 2. Since these inequalities hold for every $h > 0$, thanks to the Lipschitz continuity of $x_n(t)$, we conclude that

$$x_n(t) \rightarrow x_n^*.$$

□

In order to complete the proof of Proposition 6, we have to show that given any solution $\mathbf{x}(t)$ of (6) with initial-value $\mathbf{x}(0) \in E^1 \setminus E_{< \infty}^1$ it holds that $\mathbf{x}(t) \not\rightarrow \mathbf{x}^*$. Note that a necessary condition for $\mathbf{x}(t) \rightarrow \mathbf{x}^*$ is the convergence of the masses, i.e., $m(\mathbf{x}(t)) \rightarrow m(\mathbf{x}^*)$. Note that $m(\mathbf{x}(0)) = \infty$ and $m(\mathbf{x}^*) < \infty$. Hence, since the function $m(\mathbf{x}(\cdot))$ is Lipschitz continuous, the convergence of the masses cannot happen in finite time, and in turn $\mathbf{x}(0) \notin B(\mathbf{x}^*)$.

B.8 Proof of Proposition 8

Here we relate the notation in our framework with the notation in [18], so as to make the proof of Proposition 8 as transparent as possible. For ease of notation we omit the superscript N .

The N queues in [18] are identified in our framework by (c, k) for $c \in \mathcal{C}$ and $k = 1, \dots, N_c$. The load of queue (c, k) is $\rho_{c,k} = \frac{\lambda_{c,k}}{\mu_{c,k}}$, where $\lambda_{c,k} = \lambda_c/N_c$ and $\mu_{c,k} = \mu_c$. In our framework when the server visits a queue with m packets in the buffer $S(m)$ packets are served, where

$$S(m) = \begin{cases} 0, & \text{if } m = 0, \\ 1, & \text{if } m > 0. \end{cases}$$

Hence $\bar{S} = \lim_{m \rightarrow \infty} S(m) = 1$. Denote by $r_{(c,k)(d,h)}$ the probability that the server visits queue (d, h) after having visited queue (c, k) . In our case, the queue visited by a server is not dependent on the past visits, so that

$$r_{(c,k)(d,h)} = r_{(d,h)} = \frac{\nu_d}{N_d \nu},$$

where $\nu = \sum_{c \in \mathcal{C}} \nu_c$. Also, the switchover times $\{\delta_{(c,k)(d,h)}(n)\}_{n \geq 0}$, i.e., the time the server needs for switching from queue (c, k) to (d, h) at the n -th occurrence of such a switch, are independent of the queue just left. More explicitly, the random variables $\delta_{(c,k)(d,h)}(n)$ are i.i.d. copies of the random variable

$$\delta_{(c,k)(d,h)} \sim \text{Exp}(\nu).$$

In particular, $\delta_{(c,k)(d,h)}$ is distributed as the minimum of N exponential random variables each corresponding to the back-off time of a different node whose parameter depends on the class in which it belongs. The routing probabilities $r_{(c,k)(d,h)}$ define a Markov chain on the state space $\{(c, k) : c \in \mathcal{C}, k = 1, \dots, N_c\}$. Let $p_{(c,k)}$ be its unique invariant distribution. In our framework, due to the independence of $r_{(c,k)(d,h)}$ from (c, k) , it holds that

$$p_{(c,k)} = \frac{\nu_c}{N_c \nu}.$$

Lastly, define

$$\delta^* \doteq \sum_{(c,k),(d,h)} p_{(c,k)} r_{(c,k)(d,h)} \mathbb{E}[\delta_{(c,k)(d,h)}],$$

which in our framework yields

$$\begin{aligned} \delta^* &= \sum_{(c,k)} \sum_{(d,h)} \frac{\nu_c}{N_c \nu} \frac{\nu_d}{N_d \nu} \frac{1}{\nu} = \frac{1}{\nu^3} \sum_{(c,k)} \sum_{(d,h)} \frac{\nu_c}{N_c} \frac{\nu_d}{N_d} \\ &= \frac{1}{\nu^3} \sum_{(c,k)} \frac{\nu_c}{N_c} \sum_{(d,h)} \frac{\nu_d}{N_d} = \frac{1}{\nu^3} \sum_{(c,k)} \frac{\nu_c}{N_c} \nu \\ &= \frac{1}{\nu^2} \sum_{(c,k)} \frac{\nu_c}{N_c} = \frac{1}{\nu}. \end{aligned}$$

At this point, in [18, Theorem 2.3], the condition for positive recurrence of the polling system is proved to be the following

$$\sum_{(c,k)} \rho_{c,k} + \max_{(c,k)} \frac{\lambda_{(c,k)}}{p_{(c,k)} \bar{S}} \delta^* < 1. \quad (62)$$

Observe that in our framework

$$\frac{\lambda_{(c,k)}}{p_{(c,k)} \bar{S}} = \frac{\lambda_c}{N_c} \frac{N_c \nu}{\nu_c} = \nu \frac{\lambda_c}{\nu_c}.$$

Hence, (62) can be rewritten as

$$\max_{(c,k)} \nu \frac{\lambda_c}{\nu_c} \frac{1}{\nu} < 1 - \sum_{c \in \mathcal{C}} \frac{\lambda_c}{\mu_c},$$

which yields Proposition 8.

B.9 Alternative proof for Proposition 8 - Single-class case

Fix the number of nodes N . For every node k , define a Poisson process of rate ν/N , which represents the back-off *opportunities*. By opportunity, we mean that the corresponding event (back-off in this case) happens if the condition of the system allows for it. Denote by $b_n^{(k,N)}$ a realization of the back-off opportunity process at node k , i.e., at time $b_n^{(k,N)}$, a back-off is accomplished if $Q_k^{(N)}(b_n^{(k,N)}) \geq 1$ and $Y(b_n^{(k,N)}) = 0$. In parallel, define a Poisson process of rate μ , which identifies the transmission *opportunities*. In particular, denote by $s_n^{(N)}$ a realization of the transmission opportunities process. At time $s_n^{(N)}$, a transmission is completed if $Y^{(N)}(s_n^{(N)}) = 1$. The arrivals at node k are governed by a Poisson process of rate λ/N , and in this case we do not need to consider opportunities. Indeed at every realization of the process a new packet arrives, no matter the state of the system. These processes are sufficient for describing the stationary behavior in the single-class case with N nodes, a special case of the model described in Section 2.

Consider now the following random sets of times

$$\mathcal{B}^{(N)} = \bigcup_{k=1}^N \bigcup_{n \in \mathbb{N}} \{b_n^{(k,N)}\}, \quad \mathcal{S}^{(N)} = \bigcup_{n \in \mathbb{N}} \{s_n^{(N)}\},$$

and define the subset $\mathcal{T}^{(N)} \subseteq \mathcal{B}^{(N)}$ such that $\bar{t} \in \mathcal{T}^{(N)}$ if and only if there exists $s_n^{(N)} \in \mathcal{S}^{(N)}$ such that

$$\bar{t} = \inf_{t \in \mathcal{B}^{(N)}} \{t > s_n^{(N)}\}.$$

Observe that different elements of $\mathcal{S}^{(N)}$ may identify the same element in $\mathcal{T}^{(N)}$. Basically, the back-off opportunities in $\mathcal{T}^{(N)}$ are the *only* opportunities taken by nodes which are not blocked by the activity process. Observe that the absence of packets in their buffer might still prevent them from initiating a transmission.

Drop the superscript (N) , and denote the sequence of elements of $\mathcal{T}^{(N)} = \mathcal{T}$ by $\{t_n\}_{n \in \mathbb{N}}$. We want to analyze the embedded queue length Markov chain $(\mathbf{Q}^n)_{n \in \mathbb{N}}$ on the state space \mathbb{N}^N where

$$\mathbf{Q}^n = \mathbf{Q}(t_n),$$

and $Q_k(t)$ is the length of the k -th queue at time t .

Define the following random variables

- $T_n = t_{n+1} - t_n$ represents the length of the n -th time slot. Observe that $(T_n)_{n \in \mathbb{N}}$ is a sequence of i.i.d. random variables whose distribution is given by the sum of two exponentially distributed random variables of rate μ and ν . Observe that $\mathbb{E}[T_n] = \bar{T} = \frac{1}{\mu} + \frac{1}{\nu}$.
- $\mathbf{A}^n = (A_k^n)_{k=1, \dots, N}$ counts the number of packets arriving at the various queues during the n -th time slot. Observe that $A_k^n |_{T_n=s} \sim \text{Poisson}(\frac{\lambda}{N}s)$ for every $k = 1, \dots, N$.
- $\mathbf{D}^n = (D_k^n)_{k=1, \dots, N}$ counts the number of packets leaving the various queues during the n -th time slot. Note that $D_k^n \in \{0, 1\}$ and $\sum_k D_k^n \leq 1$.

Proposition 12. *If $\frac{\lambda}{\mu} + \frac{\lambda}{\nu} < 1$, then the queue length Markov process $(\mathbf{Q}^{(N)}(t))_{t \geq 0}$ is positive recurrent for every $N > 0$.*

Define $L(\mathbf{Q}^n) = \sum_{k=1}^N (Q_k^n)^2$. We will show that $L(\cdot)$ is a Lyapunov function for the embedded queue length Markov chain $(\mathbf{Q}^n)_{n \in \mathbb{N}}$. In order to prove that, it suffices to show that $\mathbb{E}[L(\mathbf{Q}^{n+1}) | \mathbf{Q}^n] - L(\mathbf{Q}^n) < 0$ unless $\mathbf{Q}^n \in \mathcal{K}$, where $\mathcal{K} \in \mathbb{R}_+^N$ is a compact set.

Since

$$Q_k^{n+1} = Q_k^n + A_k^n - D_k^n,$$

it holds that

$$\begin{aligned}
& \mathbb{E}[L(\mathbf{Q}^{n+1})|\mathbf{Q}^n] - L(\mathbf{Q}^n) \\
&= L(\mathbf{Q}^n) + 2 \sum_{k=1}^N Q_k^n \mathbb{E}[(A_k^n - D_k^n)|\mathbf{Q}^n] + \sum_{k=1}^N \mathbb{E}[(A_k^n - D_k^n)^2|\mathbf{Q}^n] - L(\mathbf{Q}^n) \\
&= 2 \underbrace{\sum_{k=1}^N Q_k^n \mathbb{E}[(A_k^n - D_k^n)|\mathbf{Q}^n]}_{S_1} + \underbrace{\sum_{k=1}^N \mathbb{E}[(A_k^n - D_k^n)^2|\mathbf{Q}^n]}_{S_2}.
\end{aligned}$$

Consider S_1 , and observe its terms individually. As for the first term, it holds that

$$2 \sum_{k=1}^N Q_k^n \mathbb{E}[A_k^n|\mathbf{Q}^n] = 2 \sum_{k=1}^N Q_k^n \mathbb{E}[A_k^n] = 2 \sum_{k=1}^N Q_k^n \frac{\lambda}{N} \bar{T} = 2\lambda \left(\frac{1}{\mu} + \frac{1}{\nu} \right) \bar{Q}^n,$$

where $\bar{Q}^n = \frac{1}{N} \sum_{k=1}^N Q_k^n$. Focus now on $\mathbb{E}[D_k^n|\mathbf{Q}^n]$ for a generic node $k \in \{1, \dots, N\}$. Denote by C_j^n the random variable indicating the node associated with the j -th back-off *opportunity* during the n -th time slot, i.e., in $[t_n, t_{n+1})$. Denote by J^n the random variable counting the number of back-off opportunities during the n -th time slot. Observe that $J^n \geq 1$ due to the time-slot definition. Finally, define the random variables

$$\bar{C}_j^n = \begin{cases} 1, & \text{if the } j\text{-th back-off opportunity is successful,} \\ 0, & \text{otherwise.} \end{cases}$$

It holds that

$$\begin{aligned}
\mathbb{E}[D_i^n|\mathbf{Q}^n] &= \sum_{j=1}^{J^n} \mathbb{P}[C_j^n = i, \bar{C}_j^n = 1|\mathbf{Q}^n] \\
&\geq \mathbb{P}[C_1^n = i, \bar{C}_1^n = 1|\mathbf{Q}^n] \\
&= \mathbb{P}[\bar{C}_1^n = 1|C_1^n = i, \mathbf{Q}^n] \mathbb{P}[C_1^n = i] \\
&\geq \mathbb{1}_{\{Q_i^n > 0\}} \frac{1}{N}.
\end{aligned}$$

Thus

$$S_1 \leq 2\lambda \left(\frac{1}{\mu} + \frac{1}{\nu} \right) \bar{Q}^n - 2 \sum_{k=1}^N Q_k^n \frac{1}{N} = 2\bar{Q}^n \left(\frac{\lambda}{\mu} + \frac{\lambda}{\nu} - 1 \right). \quad (63)$$

Consider now S_2 , and observe that its value is bounded by a constant. Indeed,

$$\begin{aligned}
S_2 &\leq \sum_{i=1}^N \mathbb{E}[(A_i^n)^2] + 1 = \sum_{i=1}^N \mathbb{E}[\mathbb{E}[(A_i^n)^2|T_n]] + 1 \\
&= \sum_{i=1}^N \mathbb{E}[\mathbb{E}[(\frac{\lambda}{N} T_n)^2 + \frac{\lambda}{N} T_n]] + 1 \\
&= \frac{\lambda^2}{N} \mathbb{E}[T_n^2] + \lambda \mathbb{E}[T_n] + 1 \\
&= \frac{\lambda^2}{N} \frac{2(\mu^2 + \mu\nu + \nu^2)}{\mu^2\nu^2} + \lambda \left(\frac{1}{\mu} + \frac{1}{\nu} \right) + 1 = c_{N,\lambda,\mu,\nu} < \infty.
\end{aligned} \quad (64)$$

Combining (63) and (64) we obtain

$$\mathbb{E}[L(\mathbf{Q}^{n+1})|\mathbf{Q}^n] - L(\mathbf{Q}^n) \leq 2\bar{Q}^n \left(\frac{\lambda}{\mu} + \frac{\lambda}{\nu} - 1 \right) + c_{N,\lambda,\mu,\nu}. \quad (65)$$

Since $\lambda/\mu + \lambda/\nu < 1$, there exists a compact set of the form of $\mathcal{K} = \{\mathbf{Q} : \frac{1}{N} \sum_{i=1}^N Q_i < c\}$ such that if $\mathbf{Q} \notin \mathcal{K}$ then

$$\mathbb{E}[L(\mathbf{Q}^{n+1})|\mathbf{Q}^n] - L(\mathbf{Q}^n) < 0.$$

B.10 Proof of Proposition 9

Recall that, given $\mathbf{x}^{(N)}$, the stationary distribution of the population process, we would like the following property to hold for every class $c \in \mathcal{C}$: for every $\epsilon > 0$, there exist $n_\epsilon, N_\epsilon > 0$ such that

$$\sum_{n=0}^{n_\epsilon} x_{c,n}^{(N)} > 1 - \epsilon, \quad \forall N > N_\epsilon.$$

Observe that this property would be implied by: for every class $c \in \mathcal{C}$ there exists a constant $K_c < \infty$ such that $m_c(\mathbf{x}^{(N)}) < K_c$ for every $N > N_\epsilon$, where

$$m_c(\mathbf{x}^{(N)}) = \sum_{n=0}^{\infty} n x_{c,n}^{(N)}.$$

We would like to relate $m_c(\mathbf{x}^{(N)})$ with $\mathbb{E}[\hat{W}_c^{(N)}]$, i.e., the expected waiting time at a class- c node in stationarity in a system with N nodes. For the sake of notation we write N_c in place of $N_c^{(N)}$. Observe that

$$\begin{aligned} \mathbb{E}[\hat{W}_c^{(N)}] &\geq \left(\frac{1}{\mu_c} + \frac{N_c}{\nu_c}\right) \mathbb{E}[Q_c^{(N)}] \\ &= \left(\frac{1}{\mu_c} + \frac{N_c}{\nu_c}\right) \sum_{n=0}^{\infty} n \mathbb{P}(Q_c^{(N)}(\infty) = n) \\ &= \left(\frac{1}{\mu_c} + \frac{N_c}{\nu_c}\right) \sum_{n=0}^{\infty} n x_c^{(N)}(n), \end{aligned}$$

where $Q_c^{(N)}$ is the random variable denoting the number of packets waiting in the buffer of a class- c node in stationarity. Therefore

$$m_c(\mathbf{x}^{(N)}) \leq \frac{\mu_c \nu_c}{\nu_c + N_c \mu_c} \mathbb{E}[\hat{W}_c^{(N)}] \leq \frac{\nu_c}{N_c} \mathbb{E}[\hat{W}_c^{(N)}].$$

Thus it suffices to show that $\mathbb{E}[\hat{W}_c^{(N)}]$ grows at most linearly in N as $N \rightarrow \infty$.

The following formula is the pseudo-conservation law proved in [12, Eqn. (5.34)] specialized to a CSMA setting, i.e., $p_{c,k} = \frac{\nu_c^{(N)}}{\nu}$ and $\sigma_{c,k} = \nu$, with $\nu = \sum_{c=1}^N \nu_c$. Hence, equation [12, Eqn. (5.34)] becomes equation (47) which can be further reduced to

$$\begin{aligned} &\sum_{c \in \mathcal{C}} \frac{\lambda_c}{\mu_c} \left(1 - \frac{\lambda_c}{\nu_c} \frac{1}{1 - \rho}\right) \mathbb{E}[\hat{W}_c^{(N)}] \\ &= \frac{\rho}{1 - \rho} \sum_{c \in \mathcal{C}} \frac{\lambda_c}{\mu_c^2} + \frac{1}{1 - \rho} \sum_{c \in \mathcal{C}} \frac{N_c \lambda_c}{\mu_c \nu_c} - \sum_{c \in \mathcal{C}} \frac{\lambda_c}{\nu \mu_c} + \frac{\rho}{\nu}, \end{aligned} \quad (66)$$

where $\rho = \sum_{c=1}^N \rho_c$. The above pseudo-conservation law is valid as long as the stationary expectations of the waiting times are finite, which is granted by [17, Theorem 5.5] as long as the stability condition (8) is satisfied.

From (66) we can deduce that

$$\begin{aligned} \frac{\nu_d}{N_d} \mathbb{E}[\hat{W}_d^{(N)}] &\leq \frac{\nu_d}{N_d} \frac{\frac{\rho}{1 - \rho} \sum_{c \in \mathcal{C}} \frac{\lambda_c}{\mu_c^2} + \frac{1}{1 - \rho} \sum_{c \in \mathcal{C}} \frac{N_c \lambda_c}{\mu_c \nu_c} - \sum_{c \in \mathcal{C}} \frac{\lambda_c}{\nu \mu_c} + \frac{\rho}{\nu}}{\frac{\lambda_d}{\mu_d} \left(1 - \frac{\lambda_d}{\nu_d} \frac{1}{1 - \rho}\right)} \\ &= \frac{\frac{\nu_d}{1 - \rho} \sum_{c \in \mathcal{C}} \frac{p_c \lambda_c}{p_d \mu_c \nu_c}}{\frac{\lambda_d}{\mu_d} \left(1 - \frac{\lambda_d}{\nu_d} \frac{1}{1 - \rho}\right)} + K_N, \end{aligned}$$

where $\lim_{N \rightarrow \infty} K_N = 0$.

References

- [1] L. V. Ahlfors (1953). Complex Analysis: an Introduction to the Theory of Analytic Functions of One Complex Variable.
- [2] L. Atzori, A. Iera, G. Morabito (2010). The Internet of Things: a survey. *Computer Networks* **54** (15), 2787–2805.
- [3] M. Benaïm, J.Y. Le Boudec (2008). A class of mean field interaction models for computer and communication systems. *Performance Evaluation* **65** (11), 823–838.
- [4] M. Benaïm, J.Y. Le Boudec (2011). On mean field convergence and stationary regime. *arXiv:1111.5710*.
- [5] D. Bertsimas, D. Nazakato (1995). The distributional Little’s law and its applications. *Operations Research* **43** (2), 298–310.
- [6] G. Bianchi (2000). Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications* **18** (3), 535–547.
- [7] P. Billingsley (1968). Convergence of Probability Measures. 1st Edition.
- [8] P. Billingsley (1999). Convergence of Probability Measures. 2nd Edition.
- [9] R.R. Boorstyn, A. Kershenbaum, B. Maglaris, V. Sahin (1987). Throughput analysis in multihop CSMA packet radio networks. *IEEE Transactions on Communications* **35**, 267–274.
- [10] C. Bordenave, D. McDonald, A. Proutiere (2008). Performance of random medium access control, an asymptotic approach. *ACM SIGMETRICS Performance Evaluation Review* **36** (1), 1–12.
- [11] N. Bouman, S.C. Borst, J.S.H. van Leeuwaarden, A. Proutiere (2011). Backlog-based random access in wireless networks: fluid limits and delay issues. In: *Proc. International Teletraffic Congress*, 39–46.
- [12] O.J. Boxma, J.A. Weststrate (1989). Waiting times in polling systems with Markovian server routing. In: *Proc. Messung, Modellierung und Bewertung von Rechensystemen und Netzen*. 89–105.
- [13] M. Bramson (1998). State space collapse with application to heavy traffic limits for multiclass queueing networks. *Queueing Systems* **30** (1-2), 89–140.
- [14] F. Cecchi, S.C. Borst, J.S.H. van Leeuwaarden (2014). Throughput of CSMA networks with buffer dynamics. *Performance Evaluation* **79**, 216–234.
- [15] F. Cecchi, S.C. Borst, J.S.H. van Leeuwaarden, P.A. Whiting (2016). CSMA networks in a many-sources regime: A mean-field approach. In: *Proc. IEEE Infocom*.
- [16] J. Cho, J.Y. Le Boudec, Y. Jiang (2012). On the asymptotic validity of the decoupling assumption for analyzing 802.11 MAC protocol. *IEEE Transactions on Information Theory* **58** (11), 6879–6893.
- [17] J.G. Dai, S.P. Meyn (1995). Stability and convergence of moments for multiclass queueing networks via fluid limit models. *IEEE Transactions on Automatic Control* **40** (11), 1889–1904.
- [18] D. Down (1998). On the stability of polling models with multiple servers. *Journal of Applied Probability* **34** (4), 925–935.
- [19] K.R. Duffy (2010). Mean field Markov models of wireless local area networks. *Markov Processes and Related Fields* **16** (2), 295–328.

- [20] S.N. Ethier, T.G. Kurtz (1986). Markov Processes: Characterization and Convergence.
- [21] D. Evans (2011). The internet of things. *How the next evolution of the internet is changing everything*, Whitepaper, Cisco (IBSG).
- [22] C. Fricker, M.R. Jaibi (1998). Stability of multi-server polling models. *INRIA Research report*, No. 3347.
- [23] D. Gamarnik, A. Zeevi (2006). Validity of heavy traffic steady-state approximations in generalized Jackson networks. *Annals of Applied Probability* **16** (1), 56–90.
- [24] J. Ghaderi, R. Srikant (2010). On the design of efficient CSMA algorithms for wireless networks. In: *Proc. IEEE Decision and Control*, 954–959.
- [25] C. Graham, P. Robert (2009). Interacting multi-class transmissions in large stochastic networks. *Annals of Applied Probability* **19** (6), 2334–2361.
- [26] M. Halldorsson, T. Tonoyan (2015). How well can graphs represent wireless interference? In: *Proc. ACM Symposium on Theory of Computing*.
- [27] L. Jiang, J. Walrand (2008). A distributed CSMA algorithm for throughput and utility maximization in wireless networks. In: *Proc. Allerton Conference*, 1511–1519.
- [28] W. Kang, K. Ramanan (2012). Asymptotic approximations for stationary distributions of many-server queues with abandonment. *Annals of Applied Probability* **22** (2), 477–521.
- [29] S.C. Liew, C.H. Kai, J. Leung, B. Wong (2010). Back-of-the-envelope computation of throughput distributions in CSMA wireless networks. *IEEE Transactions on Mobile Computing* **9** (9), 1319–1331.
- [30] A. Mandelbaum, G. Pats (1998). State-dependent stochastic networks. Part I: Approximations and applications with continuous diffusion limits. *Annals of Applied Probability* **8** (2), 569–646.
- [31] A. Mukhopadhyay, A. Karthik, R. Mazumdar, F. Guillemin (2015). Mean field and propagation of chaos in multi-class heterogeneous loss models. *Performance Evaluation* **91**, 117–131.
- [32] G. Pang, R. Talreja, W. Whitt (2007). Martingale proofs of many-server heavy-traffic limits for Markovian queues. *Probability Surveys* **4**, 193–267.
- [33] D. Revuz, M. Yor (2013). Continuous Martingales and Brownian Motion.
- [34] L.C.G Rogers, D. Williams (2000). Diffusions, Markov Processes and Martingales: Volume 1, Foundations.
- [35] D. Shah, J. Shin, P. Tetali (2011). Medium access using queues. In: *Proc. FOCS Conference*
- [36] G. Sharma, A. Ganesh, P.B. Key (2009). Performance analysis of contention based medium access control protocols. *IEEE Transactions on Information Theory* **55** (4), 1665–1682.
- [37] A. Stolyar (2005). On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation. *Operations Research* **53** (1), 12–25.
- [38] A. Stolyar (2015). Pull-based load distribution in large-scale heterogeneous service systems. *Queueing Systems* **80** (4), 341–361.
- [39] A.S. Sznitman (1991). Topics in propagation of chaos. *Ecole d’ete de probabilites de Saint-Flour XIX - 1989*, 165–251.
- [40] P.M. van de Ven, S.C. Borst, J.S.H. van Leeuwen, A. Proutiere (2010). Insensitivity and stability of random-access networks. *Performance Evaluation* **67** (11), 1230–1242.

- [41] P.M. van de Ven, A.J.E.M. Janssen, J.S.H. van Leeuwaarden, S.C. Borst (2011). Achieving target throughputs in random-access networks. *Performance Evaluation* **68** (11), 1103–1117.
- [42] X. Wang, K. Kar (2005). Throughput modeling and fairness issues in CSMA/CA based ad-hoc networks. In: *Proc. IEEE Infocom*.
- [43] W. Whitt (1969). Weak convergence of probability measures on the function space $C[0, \infty)$. *Technical Report No. 120, Stanford University*.
- [44] W. Whitt (1970). Weak convergence of probability measures on the function space $C[0, \infty)$. *The Annals of Mathematical Statistics*, **41** (3), 939-944.
- [45] D.V. Widder (1942). The Laplace transform.
- [46] X. Zhou, Z. Zhang, G. Wang, X. Yu, B.Y. Zhao, H. Zheng (2013). Practical conflict graphs for dynamic spectrum distribution. In: *Proc. ACM SIGMETRICS*, **41** (1), 5–16.