

Performance of Random Medium Access Control

An asymptotic approach

Charles Bordenave
CNRS & Math. Dept. of
Toulouse Univ., France
charles.bordenave@math.univ-
toulouse.fr

David McDonald
Math. Dept. of
Ottawa Univ., Canada
dmdsg@uottawa.ca

Alexandre Proutiere
Microsoft Research
Cambridge, UK
alexandre.proutiere@microsoft.com

ABSTRACT

Random Medium-Access-Control (MAC) algorithms have played an increasingly important role in the development of wired and wireless Local Area Networks (LANs) and yet the performance of even the simplest of these algorithms, such as slotted-Aloha, are still not clearly understood. In this paper we provide a general and accurate method to analyze networks where interfering users share a resource using random MAC algorithms. We show that this method is asymptotically exact when the number of users grows large, and explain why it also provides extremely accurate performance estimates even for small systems. We apply this analysis to solve two open problems: (a) We address the stability region of non-adaptive Aloha-like systems. Specifically, we consider a fixed number of buffered users receiving packets from independent exogenous processes and accessing the resource using Aloha-like algorithms. We provide an explicit expression to approximate the stability region of this system, and prove its accuracy. (b) We outline how to apply the analysis to predict the performance of adaptive MAC algorithms, such as the exponential back-off algorithm, in a system where saturated users interact through interference. In general, our analysis may be used to quantify how far from optimality the simple MAC algorithms used in LANs today are, and to determine if more complicated (e.g. queue-based) algorithms proposed in the literature could provide significant improvement in performance.

Categories and Subject Descriptors

C.2.1 [Network Architecture and Design]: Wireless communication

General Terms

Algorithms, Design, Performance

Keywords

Aloha/CSMA, exponential back-off, stability

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SIGMETRICS'08, June 2–6, 2008, Annapolis, Maryland, USA.
Copyright 2008 ACM 978-1-60558-005-0/08/06 ...\$5.00.

1. INTRODUCTION

Distributed Medium Access Control (MAC) algorithms, starting from the first version of Abramson's Aloha to the most recent algorithms used in IEEE802.11, have enabled a rapid growth of wired and wireless Local Area Networks (LANs). They aim at efficiently and fairly sharing a resource among users even though each user must decide independently when and how to attempt to use the resource. MAC algorithms have generated a lot of research interest, especially recently in attempts to use multi-hop wireless networks (Mesh and Ad-Hoc networks) to provide low-cost high-speed access to the Internet.

In this paper, we restrict our attention to wireless LANs, where the resource is a single radio channel and where users interact because of interference, i.e., two interfering users cannot transmit simultaneously (in such case we say that they experience a collision). Users willing to use the channel must first probe (or listen to) the channel, and when it is not busy, they can decide to transmit. In other words, we consider CSMA (Carrier Sense Multiple Access) algorithms. However as explained later, the analysis also applies to classical slotted-Aloha.

We may roughly categorize distributed MAC algorithms into two types: (1) Traffic-oblivious algorithms, where users do not use any information on the contents of their own or their neighbors' buffers to decide whether to transmit; (2) Traffic-aware algorithms, where users exchange information on their buffers before deciding to attempt to use the channel. In today's LANs, only simple traffic-oblivious algorithms are used. In spite of their simplicity, these algorithms are extremely difficult to analyze and their performance remains largely open (we outline recent efforts to analyze these algorithms in Sections 3 and 5). As a consequence, we do not know whether they can provide any performance guarantees. This has motivated the recent development of traffic-aware algorithms. For example, some researchers have proposed distributed implementations (see e.g. [16] and references therein) of Tassiulas-Ephremides optimal centralized algorithm [18]. In general, the performance of traffic-aware algorithms is easier to analyze than that of traffic-oblivious algorithms, and they should perform better. However they require signaling procedures that are sometimes complicated and that in turn affect the overall performance.

In this paper, we aim at analyzing the performance of traffic-oblivious algorithms used in today's LANs. This analysis is crucial to understand whether the performance gap between these algorithms and optimal schemes is large enough to motivate practical implementations of more elaborated

algorithms. We consider a time-slotted system with a fixed number of users sharing a radio channel according to some traffic-oblivious distributed MAC algorithms, namely Aloha-like or exponential back-off algorithms. These users interact through interference, and can be either unsaturated, which means that the packets they need to transmit are generated according to some exogenous random processes with fixed rates, or saturated, which means that they always have packets to transmit in their buffers. Unsaturated users model streaming applications, whereas saturated users are more appropriate to model elastic data traffic. For streaming applications, we are interested in packet delays, and then primarily in buffer stability. For data applications, the QoS metric is the user throughput, i.e., the number of packets per second the user can successfully transmit.

Contributions. We provide a generic analysis for predicting the performance of systems with unsaturated or saturated users. We show that this analysis is asymptotically exact when the number of users grows large, and we explain why it also provides extremely accurate performance estimates even for small systems.

(a) *Stability of non-adaptive (Aloha-like) MAC algorithms in unsaturated systems.* Characterizing the stability region of these kind of systems is notoriously difficult (see a detailed description of the state-of-the-art in Section 3). The difficulty can be illustrated by the fact that the stability condition is *sensitive*, i.e., it depends not only on the average packet arrival rates, but also on the detailed statistical characteristics of the arrival processes. The stability region has been characterized for small systems only, with two (or sometimes three) users. The main contribution of this paper is to provide a general approximate expression of the stability region for systems with an arbitrary number of users. This approximate stability region is shown to be exact when the number grows large and very accurate when this number is small; it also proves to be *insensitive*, i.e., it depends on the mean packet arrival rates only. We give approximate expressions in various scenarios depending on the type of interference considered, and on the duration of packet transmissions and collisions:

- In the case of full interference, i.e., when all users interfere with each other, the approximate stability region is particularly simple, even when the packet transmission and the collision durations are greater than one slot.
- In the case of partial interference, i.e., when a user does not necessarily interfere with all other users, the approximate stability region is more complicated and is expressed as a function of the stationary performance of certain loss networks.

(b) *Throughput of adaptive MAC algorithms in saturated systems.* We outline how to apply the analysis to predict the throughput achieved by saturated users running adaptive algorithms, such as the exponential back-off algorithm, in systems with partial interference. This generalizes and justifies theoretically similar analysis used for systems with full interference, see [3, 5].

The paper is organized as follows. In Section 2, we present the network and traffic models. In Sections 3 and 4, we analyze the stability of systems with unsaturated users. Section 5 is devoted to the analysis of systems with saturated users. We conclude in Section 6.

2. NETWORK AND TRAFFIC MODELS

For convenience we consider a system where time is measured in slots so any time interval is measured in multiples of slots.

2.1 Network model

We consider a wireless network with a fixed number N of links or users sharing the radio resources (a single channel) using some random MAC protocol. Users are classified by location: users of the same class interfere with and are interfered with by the same set of users.

For simplicity, we assume that if user i interferes with user j , then j interferes with user i but our results could be generalized to *asymmetric* interference. Two interfering users can not simultaneously transmit successfully. If they transmit at the same time, they both experience a collision. Denote by \mathcal{C} the set of user classes, and $i \in c$ means that user i is of class c . Also denote by \mathcal{N}_c the set of classes such that users of these classes interfere with users of class c . If $d \in \mathcal{N}_c$, we say that d is a neighbor of class c . Finally we define by \mathcal{I}_i the set of users interfering with user i .

An example of such network is the uplink of a cellular system with overlapping cells. In Figure 1, we present a network with two overlapping cells: it has three different geographic regions, each corresponding to a user class. Class-2 users transmit packets to the access point (AP) 1, so they are neighbors of class-1 users. They also neighbors of class-3 users due to the fact that AP 2 has to acknowledge packets of class-3 users, and these acknowledgment (ack) transmissions interfere with the transmissions by AP 1 of acks of packets of class-2 users. The acks make the interference symmetric.

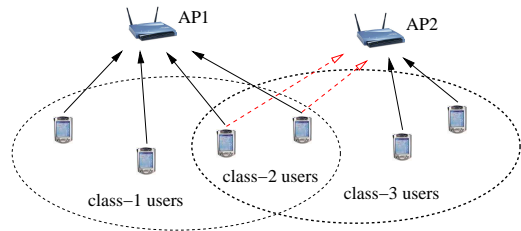


Figure 1: A network of 2 overlapping cells.

Finally, in this paper, we say that the system has *full interference* when each user interferes with all other users. Otherwise, we say that the system has *partial interference*.

2.2 Traffic models and performance metrics

We propose two traffic models, depending on the type of applications handled by the network.

2.2.1 Unsaturated users

Each user is assumed to have an infinite buffer. For user i , packets of fixed sizes are stored in a FIFO manner and arrive in this buffer according to an exogenous Markov modulated process that can be represented by a stationary ergodic Markov chain $A_i(t)$ with stationary probability $\pi_i(a)$ of being in state a . The Markov chains $A_i(t)$ are independent across users and take values in a finite space \mathcal{A} . When $A_i(t) = a$, a new packet arrives in the buffer of user i with probability $\lambda_{i,a}$. The average arrival rate of packets per slot

at user i is then $\lambda_i = \sum_a \lambda_{i,a} \pi_i(a)$. We use these chains to represent various types of packet inter-arrival processes. The simplest example is that of Bernoulli arrivals, i.e., when the inter-arrivals are geometrically distributed with mean $1/\lambda_i$: this can be represented by the Markov chain $A_i(t)$ with one state. We could also represent inter-arrivals that are sums or any combinations of geometric random variables. In the following we denote by $\alpha_i = \lambda_i / \sum_j \lambda_j$ the proportion of traffic generated by user i .

Unsaturated users constitute a convenient model for streaming applications such as voice or video, whose rate is intrinsic. For such applications, the main QoS requirement can be expressed in terms of packet delays, which remain bounded if the system is stable. Here, we consider that the system can be described by a Markov chain, so that stability means the ergodicity of this chain.

2.2.2 Saturated users

A user is said to be saturated when it always has packets in its buffer. This scenario is common in case of *elastic* data applications when the wireless link is a bottleneck. In this case, the congestion control algorithm tends to saturate the buffer. For such applications, we will be interested in the long term throughput of the users. By definition the throughput of a user is the long term average packet transmission rate achieved by this user.

2.3 Medium Access Control algorithms

We consider CSMA (Carrier-Sense-Multiple-Access) protocols, which means that before transmitting, a user senses the channel, and if it is busy it does not start transmitting. In other words, a user cannot interrupt a packet transmission of another user. Nevertheless, collisions are possible when two users have sensed the channel idle, and decided to transmit simultaneously. After the beginning of the successful transmission of a packet, the channel can not be used for L slots (note that we assumed packets with fixed sizes, so that L represents the time to transmit the packet and its MAC ack plus the additional time required for the users to find out and declare that the channel is idle again, in IEEE802.11, this duration is DIFS). Similarly, we denote by L_c the duration of a collision. L_c might be smaller than L (e.g. as in IEEE802.11 with RTS/CTS). Note that our models could be easily generalized to the case where L and L_c are random variables. When needed, to keep the system in a Markovian setting, we will keep track of the remaining time of packet transmissions or collisions. A particular case of the above model is slotted-Aloha [14], where $L = L_c = 1$.

In Section 3, we consider non-adaptive random MAC algorithms, where, when the channel is sensed idle, a user attempts to use the channel with a fixed probability. In Section 5, we discuss the performance of adaptive algorithms where the transmission probability depends of the number of successive collisions previously experienced by the user. Today's wireless LANs use the exponential back-off algorithm: after a successful transmission the transmission probability is set to p_0 , and after experiencing a collision, the transmission probability is divided by 2 (unless it has already reached the minimum probability $p_0 2^{-m}$, where usually $m = 7$).

3. STABILITY REGION OF NON-ADAPTIVE MAC ALGORITHMS

Consider a fixed number N of unsaturated users. Each user generates packets according to a random process that can be represented by a Markov chain as explained in Section 2.2. When user i has a packet in its buffer, it attempts to use the channel with fixed probability p_i . When its buffer is empty, it does not attempt to use the channel.

3.1 Full interference - Unit packet duration

We first investigate of the stability region in the case all users interfere with each other and when the packet transmission and collision durations are equal to one slot. This problem has received a lot of attention in the literature in the three last decades. An exact characterization of the stability region has been provided in [13, 15, 19] under general traffic assumptions but only for $N = 2$ users. For two users, the stability region Λ^2 is defined by: $\lambda \in \Lambda^2$ if and only if:

$$\text{either } \lambda_1 < p_1(1 - p_2), \lambda_2 < p_2(1 - \lambda_1/(1 - p_2)),$$

$$\text{or } \lambda_2 < p_2(1 - p_1), \lambda_1 < p_1(1 - \lambda_2/(1 - p_1)).$$

The first (resp. second) condition is obtained assuming that buffer 2 (resp. buffer 1) is saturated. When the number of users is greater than three, the stability region depends not only on the mean arrival rates λ_i , but also on the other detailed statistical properties of the arrival processes. For example, when $N = 3$, this is due to the fact that the stability condition for a particular buffer depends on the probability that the two other buffers are empty separately or simultaneously. These probabilities actually depend on the detailed characteristics of the arrival processes, see e.g. [4, 7]. For $N = 3$ and Bernoulli arrivals, the stability region can be characterized [17]. When the arrivals are not Bernoulli, the system stability region is unknown. When the number of users N exceeds 3, it becomes impossible to derive explicit stability conditions. As shown in [7, 17], the stability region Λ^N is then recursively described as a function of the various stability regions of systems with $N - 1$ users, Λ^{N-1} , and of the probabilities that in these systems, some buffers are simultaneously empty. These probabilities are unknown in general, and so is the stability region. The only previous explicit stability result for arbitrary N is given in [1]; unfortunately, to obtain this stability condition, the author has to assume that the arrival processes of the different users are correlated, which is unrealistic in practice. Some other authors have proposed bounds on the stability region, see e.g. [4, 12]. The basic idea behind all the proposed bounds is that the stability region is decreased when one assumes that some buffers are saturated and this can be quite conservative. Finally, note that when the system is *homogeneous* in the sense that $\lambda_i / [p_i \prod_{j \neq i} (1 - p_j)]$ does not depend on i , then one can show as in [4] that the stability condition is: $\lambda_i < p_i \prod_{j \neq i} (1 - p_j)$ for all i (in this very specific case, all buffers saturate simultaneously at the stability limit).

3.1.1 Approximate stability region $\hat{\Lambda}^N$

We now provide an approximate expression of the stability region for a system with an arbitrary number of users. We prove that this approximation is exact when the number of users grows large. The approximate expression is valid

for any arrival processes, which indicates that the stability region becomes insensitive when N grows. Finally, we show through theoretical arguments and numerical results that the approximate stability region is actually very tight even for a small number of users, say three.

Roughly speaking, the approximate stability region is obtained assuming that the evolutions of the queues of the various users are independent. Let $\partial_j[0, 1]^N$ be the set of $\rho \in \mathbb{R}_+^N$ such that $\forall i, \rho_i \leq 1$, and $\rho_j = 1$. The approximate stability region is the region lying below one of N boundaries $\partial_j \hat{\Lambda}^N$ defined by:

$$\partial_j \hat{\Lambda}^N = \left\{ \lambda : \exists \rho \in \partial_j[0, 1]^N, \forall i, \lambda_i = \rho_i p_i \prod_{k \neq i} (1 - \rho_k p_k) \right\}.$$

More precisely, $\hat{\Lambda}^N$ is the set of positive vectors λ such that there exist j and $\sigma \in \partial_j \hat{\Lambda}^N$ with $\lambda_i < \sigma_i$ for all i . Note that $\hat{\Lambda}^2 = \Lambda^2$, so the proposed approximation is exact when $N = 2$.

Let $(\lambda^N, p^N) = ((\lambda_1^N, p_1^N), \dots, (\lambda_N^N, p_N^N))$ be a sequence of vectors of arrival intensities and transmission probabilities. For this sequence, we assume that for any N , $\inf_{i,j} \lambda_i^N / \lambda_j^N > 0$ and $\inf_{i,j} p_i^N / p_j^N > 0$. These assumptions implies that when investigating stability, the arrival and transmission rates have to scale as $1/N$. Without loss of generality, we assume that $(N\lambda_i^N, Np_i^N)$ can take a finite number of values, indexed by a set \mathcal{V} , and that the proportion of users such that $(\lambda_i^N, p_i^N) = (\lambda_v/N, p_v/N)$ tends to β_v . We finally assume that for all N , $\sum_i p_i^N \leq 1$. This assumption is made so as to keep the approximation expression of the stability region simple. Note that as Kleinrock already noticed [11], the assumption is needed to guarantee a certain efficiency of the system.

The following result compares the stability region Λ^N with our proposed approximation $\hat{\Lambda}^N$ as N gets large. Define $1^N := (1/N, \dots, 1/N)$.

THEOREM 1. *For $\epsilon > 0$ small enough, there exists N_ϵ such that: for $N > N_\epsilon$,*
 (a) *if $\lambda^N + \epsilon \cdot 1^N \in \hat{\Lambda}^N$, then the system is stable;*
 (b) *if $\lambda^N - \epsilon \cdot 1^N \notin \hat{\Lambda}^N$, then the system is unstable.*

When $\hat{\Lambda}^N$ satisfies the property described in Theorem 1, we say that it is an *asymptotically exact approximation* of the stability region Λ^N . A consequence of the above result is that when N grows large, and whatever the arrival processes considered, the set traffic intensities $(\lambda_1, \dots, \lambda_N)$ such that there exist transmission probabilities (p_1, \dots, p_N) stabilizing the system is the set \mathcal{M} with boundary $\partial \mathcal{M}$:

$$\partial \mathcal{M} = \left\{ \lambda : \exists p_1, \dots, p_n \in (0, 1) : \forall i, \lambda_i = p_i \prod_{j \neq i} (1 - p_j) \right\}.$$

This result has been conjectured by Tsybakov and Mikhailov in [19].

3.1.2 How to compute $\hat{\Lambda}^N$

Assume that the distribution of traffic demand $\alpha = (\alpha_i, i = 1, \dots, N)$, is fixed and let us find the limit total arrival rate $\hat{s}^* = \sum_i \lambda_i$ such that $\lambda = \hat{s}^* \alpha \in \hat{\Lambda}^N$. It can be easily shown that at this limit, the user i^* such that $\rho_{i^*} = 1$ is: $i^* = \arg \max_i \alpha_i (1 - p_i) / p_i$. We deduce the limit arrival

rate:

$$\hat{s}^* = \frac{p_{i^*}}{\alpha_{i^*}} \prod_{i \neq i^*} \left(1 - \frac{\alpha_i p_{i^*}}{\alpha_i p_{i^*} + \alpha_{i^*} (1 - p_{i^*})} \right).$$

3.1.3 Accuracy of $\hat{\Lambda}^N$

How far is the approximate region $\hat{\Lambda}^N$ from the actual stability region? Theorem 1 says that the gap tends to 0 when the number N of users grows large. But even for small N , $\hat{\Lambda}^N$ is quite an accurate approximation as illustrated in the numerical examples provided later. We can explain this accuracy by remarking that the boundaries of the regions Λ^N and $\hat{\Lambda}^N$ coincide in many scenarios. Remember that $\hat{\Lambda}^N$ can be interpreted as the stability region one would get if the evolutions of the different buffers were independent. As a consequence, it provides the exact stability condition for scenarios where, in the stability limit, the buffers become independent.

DEFINITION 1 (*k*-HOMOGENEOUS DIRECTIONS). *A direction (a vector with unit L_1 -norm) $\alpha \in \mathbb{R}_+^N$ is *k*-homogeneous for the system considered if there exists a permutation σ of $\{1, \dots, N\}$ such that, for all $i = 1, \dots, k$, $\alpha_{\sigma(i)}(1 - p_{\sigma(i)}) / p_{\sigma(i)}$ does not depend on i .*

In the following, without loss of generality, when a direction is *k*-homogeneous, the corresponding permutation is given by $\sigma(i) = i$ for all i .

LEMMA 1. *Assume that $\lambda = s \times \alpha$, where α is a *k*-homogeneous direction for the system considered. Define $s^* = \sup\{s \geq 0 : s\alpha \in \Lambda^N\}$ and similarly \hat{s}^* . Then if $1_{k+1 \leq N} \alpha_{k+1} (1 - p_{k+1}) / p_{k+1} \leq \alpha_1 (1 - p_1) / p_1$ and $\alpha_l = 0$ for $l > k + 1$, then $s^* = \hat{s}^*$.*

Proof. For $k+1 = N$, the lemma is equivalent to Proposition 6 in [4] (a less general version of this result was given in [17], Corollary 4). For $k \leq N$, the proof can be made as in [4], remarking that along direction α , when we increase the overall arrival rate s , then due to *k*-homogeneity, buffers $1, \dots, k$ are going to saturate first. So at the stability limit, the buffer evolutions are independent. \square

The above lemma states that the boundaries of $\hat{\Lambda}^N$ and Λ^N coincide on a set of curves corresponding to particular directions. In case $N = 3$, Figure 2 gives a schematic illustration of these curves.

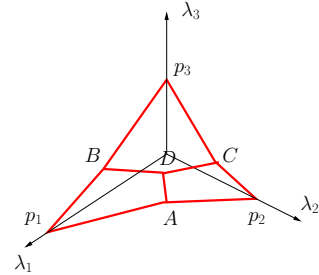


Figure 2: Curves where $\partial \hat{\Lambda}^3$ and $\partial \Lambda^3$ coincide. $A = (p_1(1 - p_2), p_2(1 - p_1), 0)$; $B = (p_1(1 - p_3), 0, p_3(1 - p_1))$; $C = (0, p_2(1 - p_3), p_3(1 - p_2))$; $D = (p_1(1 - p_2(1 - p_3)), p_2(1 - p_1(1 - p_3)), p_3(1 - p_1)(1 - p_2))$.

3.1.4 Numerical examples

We now illustrate the accuracy of $\hat{\Lambda}^N$ using numerical experiments.

Example 1: First, we consider the case of $N = 3$ sources, each transmitting with probability $1/3$. We vary the relative values of the arrival rates at the various queues: $\lambda_1 = \lambda$, $\lambda_2 = \lambda \times \frac{(1+1/x)}{2}$ and $\lambda_3 = \lambda/x$. We vary x from 1 to 50. It can be shown that the approximate stability condition is

$$\sum_{i=1}^3 \lambda_i < \hat{s}^* = \frac{4x(x+1)}{(2x+1)(5x+1)}.$$

In Figure 3 (left), we compare this limit to the actual stability limit found by simulation with Bernoulli arrivals (Simulation 1) and hyper-geometric arrivals (Simulation 2). In the latter case, the inter-arrivals for each user i are i.i.d., and an inter-arrival is a geometric random variable with parameter $a\lambda_i$ with probability $1/2$, and $(1-a)\lambda_i$ with probability $1/2$. This increases the variance of inter-arrivals (when a is small the variance scales as $1/a$). In the numerical experiment, we chose $a = 1/5$. Remark that the stability region is roughly insensitive to the distribution of inter-arrivals. This insensitivity has been also observed in the other examples presented in this section. The simulation results have been obtained running the system for about 10^7 packet arrivals. Note finally that the arrival rates are chosen so that the system is not k -homogeneous.

Example 2: We make a similar numerical experiment when $(p_1, p_2, p_3) = (0.6, 0.3, 0.1)$. The arrival rates at the three queues are as in Example 1. We vary x from 0.1 and 10. For $x < x_0 = 47/7$, at the boundary of $\hat{\Lambda}^3$, queue 3 is saturated ($\rho_3 = 1$); whereas for $x \geq x_0$, queue 2 is saturated ($\rho_2 = 1$). The approximate stability condition is:

$$\sum_{i=1}^3 \lambda_i < \hat{s}^* = \begin{cases} \frac{7.2x(x+1)}{(7x+3)(2x+3)}, & \text{if } x < 47/7, \\ \frac{44.1(x+1)^2}{(13x+7)(7x+13)}, & \text{if } x \geq 47/7. \end{cases}$$

Figure 3 (center) illustrates the accuracy of $\hat{\Lambda}^3$.

Example 3: Finally, we illustrate the accuracy of $\hat{\Lambda}^N$ when the number of users N grows. Each user is assumed to transmit with probability $1/N$ and the traffic distribution is such that $\alpha_1 > \alpha_i$ for all $i \geq 2$. Hence, again the system is not k -homogeneous. One can easily show that in this direction, the approximate stability condition is:

$$\sum_{i=1}^N \lambda_i < \hat{s}^* = \frac{1}{N\alpha_1} \prod_{i=2}^N \left(1 - \frac{\alpha_i}{\alpha_i + (N-1)\alpha_1} \right).$$

In Figure 3 (right), we compare the boundary of $\hat{\Lambda}^N$ with that of Λ^N when the distribution α_i is linearly decreasing with i . Again as expected, \hat{s}^* provides an excellent approximation of the saturation level in the actual system.

3.2 Full interference - Arbitrary packet duration

When the packet transmission and collision durations L and L_c are not equal to 1 slot, the challenges of deriving an exact stability region are similar to those arising when they are equal to 1 slot. As previously, we can provide an approximation $\hat{\Lambda}_L^N$ of the actual stability region Λ_L^N .

For $\rho = (\rho_1, \dots, \rho_N) \in \mathbb{R}_+^N$, define $\gamma_i(\rho)$:

$$\gamma_i(\rho) = \frac{P_i}{L \sum_j P_j + L_c C + E},$$

where

$$\begin{cases} P_i = \rho_i p_i \prod_{j \neq i} (1 - \rho_j p_j), \\ E = \prod_k (1 - \rho_k p_k), \\ C = 1 - E - \sum_j P_j. \end{cases}$$

The approximate stability region $\hat{\Lambda}_L^N$ is the set of points lying below one of the boundaries $\partial_j \hat{\Lambda}_L^N$ defined by:

$$\partial_j \hat{\Lambda}_L^N = \left\{ \lambda : \exists \rho \in \partial_j [0, 1]^N, \forall i, \lambda_i = \gamma_i(\rho) \right\}.$$

Under the assumptions of Theorem 1 (see the paragraph above Theorem 1), we can show that:

THEOREM 2. $\hat{\Lambda}_L^N$ is an asymptotically exact approximation of the stability region Λ_L^N .

As in the case of unit packet duration, we can show using theoretical arguments and numerical experiments that the approximation is extremely accurate. For example, the notion of k -homogenous directions can be easily extended, and $\hat{\Lambda}_L^N$ is exact in those directions. Due to space limitations, we do not discuss this further here.

3.3 Partial interference - Unit packet duration

We now consider the case with partial interference, and denote by \mathcal{I}_i the set of users interfering with user i (including i). The packet transmission duration is 1 slot. Again we can provide an accurate approximation of the stability region. This approximate region is denoted by $\hat{\Lambda}^N((\mathcal{I}_i)_{i=1}^N)$ and defined by the set of points lying below one of the boundaries $\partial_j \hat{\Lambda}^N((\mathcal{I}_i)_{i=1}^N)$:

$$\begin{aligned} \partial_j \hat{\Lambda}^N((\mathcal{I}_i)_{i=1}^N) = \\ \left\{ \lambda : \exists \rho \in \partial_j [0, 1]^N, \forall i, \lambda_i = \rho_i p_i \prod_{k \in \mathcal{I}_i \setminus \{i\}} (1 - \rho_k p_k) \right\}. \end{aligned}$$

In the case of partial interference, to state the analog of Theorem 1, we group users according to their arrival and transmission rates, and to their interference properties: a user i is in group $v \in \mathcal{V}$ if $(N\lambda_i^N, Np_i^N) = (\lambda_v, p_v)$ and if its class is identified by v , i.e., $i \in c(v)$. The analog of Theorem 1 is:

THEOREM 3. $\hat{\Lambda}^N((\mathcal{I}_i)_{i=1}^N)$ is an asymptotically exact approximation of the stability region $\Lambda^N((\mathcal{I}_i)_{i=1}^N)$.

Once again the approximation turns out to be very accurate.

3.4 Partial interference - Arbitrary packet duration

The case of partial interference and packet transmission and collision durations greater than 1 slot is more complicated than all previous cases. This is due to the fact that the users of various classes do not see the same channel busy periods. For example, the users of class 2 in the network of Figure 1 typically see longer periods where the channel is busy than the users of classes 1 and 3: users of the latter

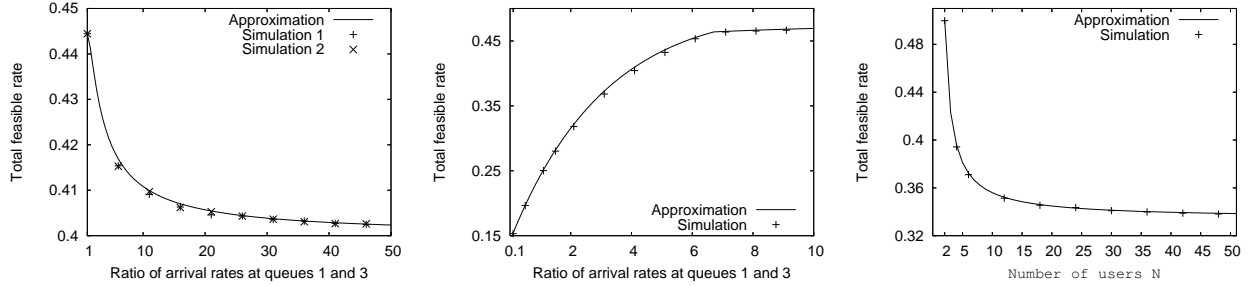


Figure 3: Maximum total rate compatible with stability - Left: Example 1, $N = 3$, $(p_1, p_2, p_3) = (0.6, 0.3, 0.1)$ - Center: Example 2, $N = 3$, $(p_1, p_2, p_3) = (0.6, 0.3, 0.1)$ - Right: Example 3, N varies, linearly decreasing traffic distribution β_i .

classes do not hear each other so a user of class 1 may start to transmit in the middle of a transmission made by a user of class 3. In such case, the channel is declared busy by users of class 2 for a period longer than L .

3.4.1 An intermediate loss network

To circumvent this difficulty, we introduce the following loss network (see e.g. [9] for a survey on loss networks). The construction of this network is as follows. For each user class $c \in \mathcal{C}$, create a link c of unit capacity. The network has a set $\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2$ of routes.

- Routes in \mathcal{R}_1 go through a single link, and there is one of such routes per link $c \in \mathcal{C}$. These routes represent successful transmissions, i.e., when a client is active on route $r = \{c\} \in \mathcal{R}_1$, it means that in the original system, a class- c user is successfully transmitting a packet. When a client starts being active on a route of \mathcal{R}_1 , it remains active for L slots.
- Routes in \mathcal{R}_2 are any set of neighboring classes, and represent collisions. These routes can be of any length. For example, when there is a client on route $r = \{c\} \in \mathcal{R}_2$, it means that in the original system, at least 2 users of class c are transmitting simultaneously and that there is no user of one of the neighboring classes of c currently transmitting. When there is a client on route $r = \{c, c'\} \in \mathcal{R}_2$, it means that there is at least one user of classes c and c' transmitting, and that no user of one of the neighboring classes of c and c' (except c, c') is currently transmitting. When a client starts being active on a route of \mathcal{R}_2 , it remains active for L_c slots.

For the system of Figure 1, the corresponding loss network is presented in Figure 4. It has 3 unit-capacity links 1, 2, and 3; $\mathcal{R}_1 = \{\{1\}, \{2\}, \{3\}\}$, $\mathcal{R}_2 = \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 2, 3\}\}$.

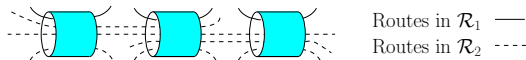


Figure 4: The loss network corresponding to the network of Figure 1.

The state of the loss network n belongs to $\{0, 1\}^{|\mathcal{R}|}$. We write $n_c = 0$ when there is no client on any route r containing link c , and $n_c = 1$ if there is at least one client on a route

going through link c . We also write $n_r = 0$ or 1 depending on whether or not there is a client on route r . We introduce the *clear-to-send* functions $S_c(n)$ as follows: $S_c(n) = 1$ if for all $d \in \mathcal{N}_c$, $n_d = 0$, $S_c(n) = 0$ otherwise. In the original system, when $S_c(n) = 1$, a class- c user observes an idle channel and none of its neighbors is currently transmitting. In this case, this user can attempt to use the channel.

The evolution of n is discrete in time and is parametrized by a vector $\rho = (\rho_1, \dots, \rho_N) \in \mathbb{R}_+^N$, where ρ_i will be naturally interpreted as the proportion of time the buffer of user i is not empty. First introduce P_c (resp. C_c) the probability that only one (resp. more than one) class- c user starts transmitting, and E_c the probability that none of these users starts transmitting:

$$\begin{cases} P_c = \sum_{i \in \mathcal{C}} \rho_i p_i \prod_{j \neq i, j \in \mathcal{C}} (1 - \rho_j p_j), \\ E_c = \prod_{i \in \mathcal{C}} (1 - \rho_i p_i), \\ C_c = 1 - E_c - P_c. \end{cases}$$

The service time of a client on a route is deterministic so we have to include the remaining service time to make the loss network Markovian. Denote by d_r the residual service time of the client on route r , and assume that $d_r = 0$ means that there is no client on this route. We are now ready to write the transition kernel of the loss network. Consider a transition from state (n, d) to (n', d') . First, the state n' has to be feasible: this means that if for any couple of routes r, s such that $n'_r = 1 = n'_s$, then user classes $c \in r$ and $c' \in s$ can not be neighboring classes. Formally, for all $c \in r$ and all $c' \in s$, then $c' \notin \mathcal{N}_c$. We gather these conditions together through the function $1_f(n')$ equal to 1 if all the conditions are met and 0 otherwise. The transition probability $K((n, d), (n', d'))$ can then be written:

$$K((n, d), (n', d')) = 1_f(n') \times K_1 \times K_2 \times K_e \times K_s.$$

K_s represents the services of clients in the network:

$$K_s = \prod_{r: n_r=1, n'_r=0} 1_{\{d_r=1\}} \prod_{r: n_r=1=n'_r} 1_{\{d'_r=d_r-1>0\}}.$$

K_1 represents the probability that clients arrive on routes in \mathcal{R}_1 . Denote by $R_1(n, n')$ the set of routes in \mathcal{R}_1 such that if $r \in R_1(n, n')$, $n_r = 0$, $n'_r = 1$, and $d'_r = L$. Then:

$$K_1 = \prod_{c: \{c\} \in R_1(n, n')} S_c(n) P_c.$$

Similarly, K_2 is the probability that routes in \mathcal{R}_2 become active. Define by $R_2(n, n')$ the set of routes of \mathcal{R}_2 becoming

active in state n' (note that here we must have $d'_r = L_c$). Then:

$$K_2 = \prod_{c:\{c\} \in \mathcal{R}_2(n, n')} S_c(n) C_c \times \prod_{\substack{r \in \mathcal{R}_2(n, n') \\ |r| \geq 2}} \prod_{c \in r} S_c(n) (P_c + C_c).$$

Finally, K_e is the probability that routes initially empty in state n remain empty in state n' . Define by $R_e(n, n')$ the set of classes c such that there exists no route r in $\mathcal{R}_1(n, n') \cup \mathcal{R}_2(n, n')$ such that $c \in r$. Then:

$$K_e = \prod_{c \in R_e(n, n')} (S_c(n) E_c + 1 - S_c(n)).$$

Remark that the expression of the transition kernel is complicated by the fact that the loss network evolves in discrete time. Note also that the arrival rates in the loss network depend on the network state, so that in general, the network loses its reversibility [9]. Hence, its stationary distribution π_ρ can not be explicitly written. Anyway, the state space of the loss network is rather restricted, so that it is easy to compute π_ρ numerically. Note that there are cases where the stationary distribution can be accurately approximated. For example when $L_c \ll L$, then we can neglect routes of \mathcal{R}_2 . In such a scenario, writing the balance equations of the loss network with geometrically distributed service times (instead of deterministic), we observe that the network is reversible. This implies that its stationary distribution can be explicitly written and is insensitive to the service time distribution provided that the mean service time L is not changed (see e.g. [10]). This distribution is then valid for deterministic service times, and it is such that, for all n with $n_r = 0$ for $r \in \mathcal{R}_2$, $\pi_\rho(n) = \pi_\rho(0) \prod_{c: n_c=1} L \times P_c \prod_{r \in \mathcal{R}_2} 1_{\{n_r=0\}}$. With this approximation, we recover the stationary distribution identified in [8] to understand the performance of Aloha-like algorithms in networks with saturated users, and with negligible collision probability.

From the stationary distribution π_ρ , we can deduce the service rate of user $i \in c$, $\gamma_i(\rho)$:

$$\gamma_i(\rho) = \sum_{n: S_c(n)=1} \pi_\rho(n) \frac{\rho_i p_i}{1 - \rho_i p_i} E_c \prod_{d \in \mathcal{N}_c, S_d(n)=1} E_d.$$

The above expression is obtained using the cycle formula (see 1.3.2 [2]).

3.4.2 Approximate stability region

We are now able to define $\hat{\Lambda}_L^N((\mathcal{I}_i)_{i=1}^N)$, the proposed approximation of the stability region. It is the set of points lying below one of the N boundaries $\partial_j \hat{\Lambda}_L^N((\mathcal{I}_i)_{i=1}^N)$ defined by:

$$\partial_j \hat{\Lambda}_L^N((\mathcal{I}_i)_{i=1}^N) = \left\{ \lambda : \exists \rho \in \partial_j [0, 1]^N, \forall i, \lambda_i = \gamma_i(\rho) \right\}.$$

The analog of Theorem 1 is:

THEOREM 4. $\hat{\Lambda}_L^N((\mathcal{I}_i)_{i=1}^N)$ is an asymptotically exact approximation of the stability region $\Lambda_L^N((\mathcal{I}_i)_{i=1}^N)$.

3.4.3 Stability region of the network of Figure 1

In Figure 5, we compare $\hat{\Lambda}_L^N((\mathcal{I}_i)_{i=1}^N)$ and $\Lambda_L^N((\mathcal{I}_i)_{i=1}^N)$ in the case of the network of Figure 1. There are 2 users per class, and they all transmit with probability 0.1. We assume that the arrival rates λ_i for users of classes 1 and 3 are identical, i.e., $\lambda_1 = \lambda_2 = \lambda_5 = \lambda_6$. The arrival rates for

users of class 2 are also identical, $\lambda_3 = \lambda_4$. We investigate the cases where $L = L_c = 2, 10, \text{ and } 100$. Again the approximation is quite accurate. Note also that they are some points in the curves representing the boundary of $\hat{\Lambda}^N$ that actually correspond exactly to stability limits of the system (analog of the notion of k -homogenous systems discussed in Section 3.1). This is the case for example for the elbow of the curve for $L = 2$ (at that point all queues get saturated simultaneously).

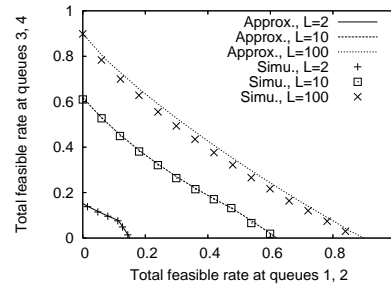


Figure 5: Stability region of the network of Figure 1 - 2 users per class, $p_i = 0.1$ for all i .

4. PROOFS OF THEOREMS 1-4

We now prove the results of the previous section. Informally speaking, we prove that when the number of users grows large, the queues attached to the various sources become independent. Assuming independence, we can derive the system stability, and then show that the stability condition obtained approximates that of the original system. The proofs are based on interpreting the system as a system of interacting particles for which we analyze the asymptotic behavior. All the results related to the particle system are postponed to the appendix, and proved in [6]. In the present section, we explain how to use these results. We will prove Theorem 1, and briefly explain how to obtain Theorems 2-4.

4.1 Proof of Theorem 1

4.1.1 The limiting regime

Consider a system with N users, and for user i , denote by λ_i^N and p_i^N the arrival and transmission rates, respectively. The arrivals of packets for user i are driven by a \mathcal{A} -valued Markov chain A_i^N whose transition kernel is K_i^N , and is initially in steady state π_i . By assumption, for each i , there is a $v \in \mathcal{V}$ such that $(\lambda_i^N, p_i^N) = (\lambda_v/N, p_v/N)$. In which case we say that user i is in group v . Without loss of generality, we assume that v also characterizes the arrival process of users, i.e., for a user i in group v , $K_i^N = K_v$, $\pi_i = \pi_v$, $\lambda_{i,a}^N = \lambda_{v,a}/N$, and $\lambda_v = \sum_{a \in \mathcal{A}} \pi_v(a) \lambda_{v,a}$ (refer to Section 2.2 for the notation).

The system can be represented as a system of interacting particles as described in Appendix. Each user i corresponds to a particle whose state $X_i^N(t)$ at time slot t represents the group v it belongs to, and the length $B_i^N(t)$ of its buffer: $X_i^N(t) = (v, B_i^N(t))$. The individual environment of particle i is $A_i^N(t)$, and the global environment process Z^N is not needed here. Denote by $\nu^N(t)$ the empirical measure of the system at time slot t : $\nu^N(t) = \frac{1}{N} \sum_{i=1}^N \delta_{X_i^N(t)}$.

Assume that at time slot t , the empirical measure is ν . If user/particle i in group v is in state $x = (v, k)$, and if its individual environment $A_i^N(t)$ is a , the probabilities of transition for the next slot are given as follows. The state becomes $(v, k + 1)$ with probability:

$$F_b^N(x, \nu, a)/N = \lambda_{v,a}/N + o(1/N),$$

and $(v, k - 1)$ with probability:

$$F_d^N(x, \nu, a)/N = 1_{k>0} \frac{p_v}{N(1 - \frac{p_v}{N})} \prod_{v'} (1 - \frac{p_{v'}}{N})^{\beta_{v'}^N \nu_{v'}^+} + o(1/N),$$

where β_v^N is the proportion of users in group v and ν_v^+ is the proportion of users in group v with non-empty buffers. Denote by β_v the proportion of users in group v at the limit when N grows large. When $N \rightarrow \infty$, the functions F_b^N, F_d^N converge to F_b, F_d where:

$$F_b(x, \nu, a) = \lambda_{v,a}, \quad F_d(x, \nu, a) = 1_{k>0} p_v \exp(-\sum_{v'} \beta_{v'} \nu_{v'}^+ p_{v'}).$$

At time 0, we apply a random and uniformly distributed permutation to the users so that their initial states become i.i.d.. This operation does not change the stability of the system. Since the system evolves slowly (at speed $1/N$), we accelerate time and define $q_i^N(t) = X_i^N(\lfloor Nt \rfloor)$. The limiting averaged functions defining the particle transitions (see (9) in Appendix) are then:

$$\bar{F}_b(x, \nu) = \lambda_v, \quad \bar{F}_d(x, \nu) = 1_{k>0} p_v \exp(-\sum_{v'} \beta_{v'} \nu_{v'}^+ p_{v'}).$$

We can now apply Theorem 6 in Appendix, and conclude that when N grows large, the evolutions of the users become independent. Furthermore, at time t , if $Q_{(v,k)}(t)$ denotes the limiting probability that a user in group v has k packets in its buffer, then thanks to Theorem 7, we have:

$$\begin{aligned} \frac{\partial}{\partial t} Q_{(v,k)}(t) &= \lambda_v (1_{k>0} Q_{(v,k-1)}(t) - Q_{(v,k)}(t)) \quad (1) \\ &+ p_v \exp(-\gamma(t)) (Q_{(v,k+1)}(t) - 1_{k>0} Q_{(v,k)}(t)). \end{aligned}$$

with

$$\gamma(t) = \sum_v \beta_v p_v \nu_v^+(t) = \sum_v \beta_v p_v (1 - Q_{(v,0)}(t)). \quad (2)$$

For a given v , equations (1) are the Kolmogorov equations corresponding to an M/M/1 queue with time-varying capacity equal to $p_v \exp(-\gamma(t))$ at time t . One can also write the evolution of the workload $V_v(t) = \sum_k k Q_{(v,k)}(t)$ of a queue of group v :

$$\frac{\partial V_v}{\partial t}(t) = \lambda_v - p_v e^{-\gamma(t)} (1 - Q_{(v,0)}(t)). \quad (3)$$

Finally, multiplying by β_v and summing over v , we can characterize the evolution of the total workload $V(t) = \sum_v \beta_v V_v(t)$ as:

$$\frac{\partial}{\partial t} V(t) = \sum_v \beta_v \lambda_v - \gamma(t) \exp(-\gamma(t)). \quad (4)$$

4.1.2 Stability at the limiting regime

We now investigate the stability of the dynamical system (1)-(2). Assume that $\sum_v \beta_v \lambda_v < e^{-1}$, in the following we denote by $\underline{\gamma}(\lambda)$ and $\bar{\gamma}(\lambda)$ the unique solutions in $(0, 1)$ and

in $(1, \infty)$, respectively, of:

$$\gamma e^{-\gamma} = \lambda := \sum_v \beta_v \lambda_v. \quad (5)$$

Define the function ξ from $[0, \infty)$ to $[0, e^{-1}]$ by $\xi(x) = x e^{-x}$. Let $p = \sum_v \beta_v p_v$. The stability of the dynamical system is given by:

THEOREM 5. (a) Assume that:

$$p < \bar{\gamma}(\lambda) \text{ and } \forall v \in \mathcal{V}, \lambda_v < p_v \exp(-\underline{\gamma}(\lambda)), \quad (6)$$

then the dynamical system (1)-(2) is globally stable, and $p > \underline{\gamma}(\lambda)$.

(b) If for some $v \in \mathcal{V}, \lambda_v > p_v \exp(-\underline{\gamma}(\lambda))$ or if $p < \underline{\gamma}(\lambda)$ then the dynamical system (1)-(2) is unstable.

(c) If $p > \bar{\gamma}(\lambda)$, then the system is not globally stable.

The above theorem states that the stability region of (1)-(2) is $\Gamma(1, V)$, where for $b \in [0 : 1]$, $\Gamma(b, V)$ is the following subset of \mathbb{R}_+^V :

$$\{\lambda \in \mathbb{R}_+^V : \exists \rho \in [0 : 1]^V : \forall v, \lambda_v = p_v \rho_v b e^{-\sum_u \beta_u \rho_u p_u}\}.$$

Actually, one can easily prove that $\Gamma(b, V)$ is the stability region of a generalized system obtained from (1)-(2) by adding a slot availability probability b to the service rate of group v users; i.e. this service rate becomes $b p_v \exp(-\gamma(t))$. We now provide an alternative representation of $\Gamma(b, V)$. Define $\Lambda(b, V)$ as the subset of \mathbb{R}_+^V whose upper Pareto-boundary is the union of the following surfaces $\partial_v \Lambda(b, V)$:

$$\{\lambda \in \mathbb{R}_+^V : \exists \rho \in \partial_v [0 : 1]^V : \forall u, \lambda_u = \rho_u p_u b e^{-\sum_w \beta_w \rho_w p_w}\}.$$

We prove that when $\langle \beta, p \rangle := \sum_v \beta_v p_v < 1$, then $\Lambda(b, V) = \Gamma(b, V)$. Let component v of the function f be $f_v(g) = b g_v \exp(-\langle \beta, g \rangle)$, and let $\mathcal{P} = [0, p_1] \times \dots \times [0, p_V]$. The derivative df of f is $b \exp(-\langle \beta, g \rangle) (I - g \beta^T)$ where $g^T = (g_1, \dots, g_V)$ and $\beta^T = (\beta_1, \dots, \beta_V)$. $(I - g \beta^T)$ is a rank one matrix with one nonzero eigenvalue, $\langle \beta, g \rangle$ associated with the eigenvector g . Since $\sum_u \beta_u g_u \leq \sum_v \beta_v p_v < 1$, the inverse of df is the positive matrix $b^{-1} \exp(\langle \beta, g \rangle) (I + (1 - \langle \beta, g \rangle)^{-1} g \beta^T)$ by inspection. f is clearly one-to-one from \mathcal{P} to the star-like domain $f(\mathcal{P}) = \Gamma(b, V)$. Since df is nonsingular it follows that the image of points g in the interior of \mathcal{P} are mapped to the interior of $f(\mathcal{P})$ (since $f(g+h) - f(g)$ includes a ball around g by first order approximation) and that points g on the boundary of \mathcal{P} are mapped to the boundary of $f(\mathcal{P})$ (again by first order approximation). It is also clear that the upper, respectively lower, boundary of \mathcal{P} is mapped to the upper boundary (the union of the $\partial_v \Lambda(b, V)$), respectively lower boundary of $f(\mathcal{P})$. Moreover, since the inverse of df is a positive matrix, it follows that if $\alpha < \lambda \in \Gamma(b, V)$ then $\alpha \in \Gamma(b, V)$. It follows that the boundaries of $f(\mathcal{P})$ are Pareto boundaries so $\Lambda(b, V) = \Gamma(b, V)$.

Proof of Theorem 5. The proof is based on the probabilistic interpretation of the dynamical system (1)-(2) as a collection of M/M/1 queues with time-varying capacities.

For two probability measures σ, σ' on \mathbb{N} , we write $\sigma \leq_{st} \sigma'$ if for all $k \in \mathbb{N}$, $\sum_{l=0}^k \sigma_l \geq \sum_{l=0}^k \sigma'_l$. For a collection $\alpha = (\alpha_v, v \in \mathcal{V})$ of probability measures on \mathbb{N} , we also define $\gamma^\alpha = \sum_v \beta_v p_v (1 - \alpha_{v,0})$. For two sets of measures α, α' , if for all v $\alpha_v \leq_{st} \alpha'_v$, then $\gamma^\alpha \leq \gamma^{\alpha'}$. Let us now denote by $Q^\alpha(t)$ the set of probability measures solution of (1)-(2) with for all v , $Q_v^\alpha(0) = \alpha_v$. $Q^0(t)$ is obtained when we start with an empty system, i.e., $Q_{(v,k)}^0(0) = 1_{k=0}$ for all v .

LEMMA 2. *If for all v , $\alpha_v \leq_{st} \alpha'_v$, then*

$$\forall t \geq 0, \quad Q^\alpha(t) \leq_{st} Q^{\alpha'}(t).$$

Furthermore: $\forall t, h \geq 0$, $Q^0(t) \leq_{st} Q^0(t+h)$.

Proof. The proof of the first statement can be made using standard coupling arguments. It suffices to observe that the arrivals are exogenous so we can make the arrival process identical in both copies of the coupled chains. To prove the second statement, observe that for all v , $Q_v^{Q^0(h)}(0) \geq_{st} Q_v^0(0)$. Hence by monotonicity, $Q_v^0(t+h) \stackrel{\leq}{\leq} Q_v^{Q^0(h)}(t) \geq_{st} Q_v^0(t)$. \square

Part (a): Stability starting from an empty system. We have $\gamma^0(0) = 0$. From Lemma 2, $Q_v^0(t)$ is stochastically increasing in time, and $\gamma^0(t)$ is a non-decreasing function. This also implies that $V^0(t)$ increases, and then, by (4):

$$\forall t, \lambda \geq \gamma^0(t) \exp(-\gamma^0(t)).$$

Remark also that $\gamma^0(t)$ converges to some G when $t \rightarrow \infty$. From the above equation, we deduce that $G \leq \underline{\gamma}(\lambda)$ since by (4), $V^0(t)$ decreases if $\gamma^0(t) > \underline{\gamma}(\lambda)$. Next $\lambda_v < p_v \exp(-\underline{\gamma}(\lambda)) \leq p_v \exp(-G)$ so the workload V_v^0 is stable as $\gamma^0(t) \rightarrow G$ and the distribution of queue v is that of an M/M/1 queue with service rate $p_v \exp(-G)$. Hence, $G = \sum_v \beta_v p_v \lambda_v / (p_v \exp(-G))$ so $\xi(G) = \lambda$ and finally $G = \underline{\gamma}(\lambda)$.

Part (a): Arbitrary initial condition. Denote by α the initial condition, and note that for all t , $\gamma(t) \leq \bar{\gamma}(\lambda)$. We first state a further property of a system starting at 0: $Q_{(v,0)}^0(t)$ converges exponentially fast to $Q_{(v,0)}^0(\infty)$. To see that, it suffices to consider (1) and then compare with the same equation where $\gamma^0(t)$ replaced by $\underline{\gamma}(\lambda)$. This gives a linear system with solution $Q^{0,e}(t)$. The linearized system has a slower service rate so $Q_{(v,0)}^{0,e}(t)$ decays faster than $Q_{(v,0)}^0(t)$ from 1 to $Q_{(v,0)}^0(\infty)$. Consequently the associated $\gamma^0(t)$ increases faster than $\gamma^{0,e}(t)$ to $\underline{\gamma}(\lambda)$. However a linear system decays exponentially fast to steady state so $\gamma^{0,e}(t)$ converges exponentially fast to $\underline{\gamma}(\lambda)$ and so does $\gamma^0(t)$.

By monotonicity, $Q_{(v,0)}^\alpha(t) \leq Q_{(v,0)}^0(t)$ for all v . This implies: $\gamma^\alpha(t) \geq \gamma^0(t)$. We also know that $\gamma^0(t)$ converges exponentially fast to $\underline{\gamma}(\lambda)$ so $\xi(\gamma^\alpha(t)) - \xi(\underline{\gamma}(\lambda)) \geq -C \exp(-ct)$ for some C and c (recall that $\gamma_v^\alpha(t) \leq p \leq \bar{\gamma}(\lambda)$). Also by monotonicity $V^\alpha(t) \geq V^0(t)$ so by (4):

$$\begin{aligned} V^0(\infty) - V^\alpha(0) &\leq \int_0^\infty \frac{\partial}{\partial t} V^\alpha(t) dt \\ &= \int_0^\infty [\xi(\underline{\gamma}(\lambda)) - \xi(\gamma^\alpha(t))] dt \\ &= \int_0^\infty [\xi(\gamma^\alpha(t)) \\ &\quad - \xi(\underline{\gamma}(\lambda))]^- dt - \int_0^\infty [\xi(\gamma^\alpha(t)) - \xi(\underline{\gamma}(\lambda))]^+ dt \\ &\leq \int_0^\infty C \exp(-ct) dt - \int_0^\infty [\xi(\gamma^\alpha(t)) - \xi(\underline{\gamma}(\lambda))]^+ dt. \end{aligned}$$

Using our lower bound on $\xi(\gamma^\alpha(t))$, we conclude that:

$$\int_0^\infty [\xi(\gamma^\alpha(t)) - \xi(\underline{\gamma}(\lambda))]^+ dt < \infty.$$

However the derivative of $\gamma^\alpha(t)$ is bounded (see (1)), so we conclude $\xi(\gamma^\alpha(t)) \rightarrow \xi(\underline{\gamma}(\lambda))$ so $\gamma^\alpha(t) \rightarrow \underline{\gamma}(\lambda)$. Now as before, if $\lambda_v < p_v \exp(-\underline{\gamma}(\lambda))$, queue v is stable. Summing over v gives $\bar{\gamma}(\lambda) < p$.

Part (b): If we can show that Q_v^0 is unstable then Q_v^α will also be unstable. As previously, $\gamma^0(t) \rightarrow G \leq \underline{\gamma}(\lambda)$. Suppose some queues in the set S^c are unstable while the rest in S are stable. Consequently, $v \in S$ iff $\lambda_v < p_v e^{-G}$, and

$$\begin{aligned} G &= \sum_{v \in S^c} \beta_v p_v + \sum_{v \in S} \beta_v p_v \left(1 - \left(1 - \frac{\lambda_v}{p_v \exp(-G)}\right)\right) \\ &= p' + (\lambda - \lambda') e^G \end{aligned}$$

where $p' = \sum_{v \in S^c} \beta_v p_v$ and $\lambda' = \sum_{v \in S^c} \beta_v \lambda_v$. If there is a v such that $\lambda_v > p_v \exp(-\underline{\gamma}(\lambda))$, then $G < \underline{\gamma}(\lambda)$ and the workload V^0 diverges. If $p < \underline{\gamma}(\lambda)$ then $\gamma^\alpha(t) \leq p < \underline{\gamma}(\lambda)$ so it follows from (4) that the workload tends to infinity and $G < \underline{\gamma}(\lambda)$.

Part (c): We just show here that the dynamical system has two fixed points. We have already shown that, if for all v , $\lambda_v < p_v \exp(-\underline{\gamma}(\lambda))$, and if the system start at 0, then it converges to a fixed point where $\gamma(t) = \underline{\gamma}(\lambda)$ and where a queue of group v has the same distribution as a stationary M/M/1 queue of arrival rate λ_v and capacity $p_v \exp(-\underline{\gamma}(\lambda))$. Now the second fixed point is obtained as follows. Assume $\lambda_v < p_v \exp(-\bar{\gamma}(\lambda))$ for all v . Suppose also that the initial condition for queues of group v is the stationary distribution of an M/M/1 queue with arrival rate λ_v and capacity $p_v \exp(-\bar{\gamma}(\lambda))$. Then the derivatives in (1) are all 0 and we have identified a second fixed point. \square

4.1.3 Stability of the finite system of queues

To conclude the proof of Theorem 1, we need to relate the stability region of the dynamical system to $\hat{\Lambda}^N$. We discuss the case of stability only. The case of instability is proved similarly and left to the reader.

The arrival (resp. transmission) rate of a user i in group $v \in \mathcal{V}$ is λ_v/N (resp. p_v/N). Then, in this setting, $\lambda^N + \epsilon \cdot 1^N \in \hat{\Lambda}^N$ iff $\lambda + \epsilon \cdot 1 = (\lambda_1, \dots, \lambda_V) + \epsilon \cdot 1 \in \hat{\Lambda}^N(1, V)$, where for $b \in [0 : 1]$, $\hat{\Lambda}^N(b, V)$ is the subset of \mathbb{R}_+^V whose Pareto-boundary is the union (over v) of the following surfaces:

$$\{\lambda : \exists \rho \in \partial_v [0 : 1]^V : \forall v', \lambda_{v'} = \frac{b p_{v'} \rho_{v'}}{1 - \rho_{v'} \frac{p_{v'}}{N}} \prod_u (1 - \rho_u \frac{p_u}{N})^{\beta_u^N N}\}.$$

One can easily see that for N large enough, $\hat{\Lambda}^N(b, V)$ is very close to $\Lambda(b, V)$ (their Hausdorff distance is of order $1/N$). From this we deduce that there exists N_ϵ , such that for all $N > N_\epsilon$, $\lambda + \epsilon \cdot 1 \in \hat{\Lambda}^N(1, V)$. Define $\hat{\Gamma}^N(b, V)$ as:

$$\{\lambda \in \mathbb{R}_+^V : \exists \rho \in [0 : 1]^V : \forall v, \lambda_v = \frac{b p_v \rho_v}{1 - \rho_v \frac{p_v}{N}} \prod_u (1 - \rho_u \frac{p_u}{N})^{\beta_u^N N}\}.$$

We can prove (as done after Theorem 5) that when $\sum_v \beta_v p_v < 1$, $\hat{\Lambda}^N(b, V) = \hat{\Gamma}^N(b, V)$.

We now consider systems built from our original systems but such that each slot is available for transmission with probability b , i.i.d. over slots. We show the following result by induction on V , and deduce Theorem 1 applying it for $b = 1$.

“If there exists an $\epsilon > 0$ small enough, such that for N sufficiently large, $\lambda + \epsilon \cdot 1 \in \hat{\Lambda}^N(b, V)$, then the system with N queues is stable. Furthermore in such a case, the stationary

distributions π_{st}^N of such systems constitute a tight family of probability measures.”

Let us first prove the result when $V = 1$. In such case, all the queues are similar, and the system is then homogeneous. We have $\hat{\Lambda}^N(b, 1) = \Lambda^N(b, 1)$ and the system is stable iff $\lambda_v < p_v b(1 - p_v/N)^{N-1}$. Now assume that $\lambda_v < p_v b(1 - p_v/N)^{N-1} - \epsilon$. Then the system with N queues is stochastically bounded by the one we would obtain assuming that all the other queues are saturated. In the latter system, the stationary distribution is that of a Markovian queue of load $\lambda_v/(p_v b(1 - p_v/N)^{N-1}) < 1 - \alpha\epsilon$ for some $\alpha > 0$. Tightness follows.

Now let us assume that the result is true when $|\mathcal{V}| \leq V$, and let us prove it when $|\mathcal{V}| = V + 1$. Assume that for N large enough, $\lambda + \epsilon \cdot 1 \in \hat{\Lambda}^N(b, V + 1)$. Denote by λ^v the V -dimensional vector built from λ where the v -th component has been removed. Since $\hat{\Lambda}^N(b, V + 1) = \hat{\Gamma}^N(b, V + 1)$, there exists v such that: for N large enough,

$$\lambda^v + \epsilon \cdot 1 \in \hat{\Lambda}^N\left(b\left(1 - \frac{p_v}{N}\right)^{\beta_v^N}, V\right). \quad (7)$$

Consider the stochastically dominant system where all queues of type different than v see saturated queues of type v . For the latter sub-system, in view of (7), we can apply the induction result. We conclude that for N large enough, the dominant system without queues of type v is stable, and that the family of the corresponding stationary distributions $\pi_{st}^{N,v}$ is tight.

From Theorem 5 applied to the dominant systems without queues of type v , we know that the corresponding limiting system is globally stable. We can then apply Theorem 8 to these systems to characterize the average proportion of slots left idle by the queues of type different than v : when $N \rightarrow \infty$, this proportion tends to $\exp(-\sigma)$ where σ is the lower solution of $\sigma e^{-\sigma} \cdot b e^{-\beta_v p_v} = \sum_{u \neq v} \beta_u \lambda_u$. Now consider a queue of type v in the dominant system. Denote by $(S_t^N, t \geq 0)$ its service process. We can make this process stationary ergodic just assuming that initially the system without queues of type v is in stationary regime. The service rate of a queue of type v converges to $b p_v \exp(-\sigma) \exp(-\beta_v p_v)$ when $N \rightarrow \infty$. Hence when N is large enough, we have:

$$E[S_t^N] \geq b p_v \exp(-\sigma) \exp(-\beta_v p_v) - \epsilon/4.$$

Now since $\lambda + \epsilon \cdot 1 \in \hat{\Lambda}^N(b, V + 1)$, we have for N large enough:

$$\lambda_v < b p_v \exp(-\sigma) \exp(-\beta_v p_v) - \epsilon/2 \leq E[S_t^N] - \epsilon/4.$$

We deduce that in the dominant system, the queue of type v are stable for N large enough, and that their stationary distributions are tight. We conclude the proof noting that the original systems are stochastically dominated by systems that are stable for N large enough, and such that their stationary distributions are tight.

4.2 Sketch of proofs of Theorems 2-4

Due to space limitations, the complete proofs of Theorems 2-4 will be given in a separate manuscript. They are similar to that of Theorem 1, and consist in interpreting the system of users as a system of particles. When the packet transmission and collision durations are equal to 1, we use the same particle system as in Theorem 1. In the case of full interaction with arbitrary packet duration, the system

of particles in the same as that used in the proof of Theorem 1, but the state of particles is recorded only at certain times, i.e., at the beginning of idle slots, of packet transmissions or collisions. Now in the case of partial interference and arbitrary packet durations, we introduce a global environment process that is exactly the loss network discussed in Section 3.4.1. The state of this loss network represents the set of active classes (a class is active if there is one user of this class currently transmitting). It evolves very rapidly (at rate ~ 1) compared to the users (who evolve at rate $\sim 1/N$). The system then corresponds exactly to one of the particles systems introduced in Appendix.

5. THROUGHPUT PERFORMANCE OF MULTI-ACCESS ALGORITHMS

In this section we outline the performance analysis of a system of saturated users running the exponential back-off algorithm, and with partial interference. This generalizes the analysis of systems with full interference done in [5]. Here we only sketch the main idea behind the analysis. Refer to [6] for more details.

The wireless system is represented as a system of particles as described in Appendix. Each user corresponds to a particle whose state $X_i^N(t)$ records at time t , its class $c_i \in \mathcal{C}$ and its transmission probability $p_i^N(t)$. There is no individual environment attached to the users, but particles interact with a global environment process Z^N which records the set of active classes, i.e., classes such that there is at least one user currently transmitting. We can represent this process as the loss network as defined in Section 3.4.1. The only difference is that here the transition kernel of the loss network evolves, and at time t , it is written as in Section 3.4.1, but replacing the terms E_c, P_c, C_c as follows:

$$\begin{cases} P_c = \sum_{i \in c} p_i^N(t) \prod_{j \neq i, j \in c} (1 - p_j^N(t)), \\ E_c = \prod_{i \in c} (1 - p_i^N(t)), \\ C_c = 1 - E_c - P_c. \end{cases}$$

Simple manipulations of the above equations show that the transition kernel of the loss network depends on the empirical measure of the particles. Similarly we can write the transition probabilities of the particles. For example, the probability that a particle i corresponding to a user of class c starts a successful transmission when in state $(c, p_i^N(t))$ is:

$$\frac{p_i^N(t) S_c(n)}{1 - p_i^N(t)} \prod_{d \in \mathcal{N}_c} \left(S_d(n) \prod_{j \in d} (1 - p_j^N(t)) + (1 - S_d(n)) \right).$$

The transitions depend on the current state of the particle, on the empirical measure of the other particles and on the state of the environment process. The framework is then that of the systems considered in the Appendix. Hence we can state that when N tends to ∞ , the users become independent, and one may write the differential equations corresponding to the evolution of the distribution of a particle. We can also solve the resulting dynamical system, whose fixed point can lead to an expression of the stationary throughput of users. Unfortunately, for this complicated system, simple performance formulas can not be derived. However, the network performance can be very easily computed numerically, especially if the network has a small number of classes.

In Figure 6, we present the resulting throughput in the case of the network of Figure 1. We vary the duration of packet transmission L ($L_c = L$). Observe that when $L = 100$ (roughly corresponding to a transmission of a 1000-bytes packet at a rate 54 Mbit/s in IEEE802.11g systems), users of class 2 achieve 5 times less throughput than users of classes 1 and 3. This illustrates and quantifies the known fairness issues of the exponential back-off algorithm in multi-cell WLANs.

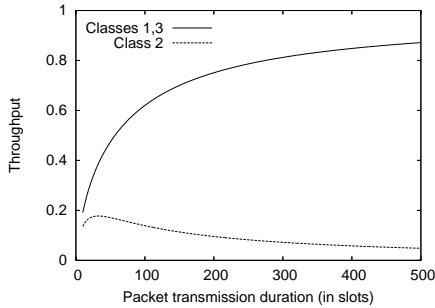


Figure 6: Throughputs achieved using exponential back-off algorithm in the network of Figure 1 - two users per class.

6. CONCLUSION

The main contribution of this work was to provide simple approximate expressions for the stability region of non-adaptive MAC protocols. These expressions are proved to be exact when the number of users grows large, and are very accurate in the case of small systems. The problem of characterizing the stability region of MAC protocols has been open for a long time, actually since the introduction of the first version of Aloha systems in 1970, and we believe that we bring considerable light into this issue.

The method used to derive stability regions of non-adaptive algorithms can be applied to predict the performance of adaptive MAC protocols, such as the exponential back-off algorithms. Unfortunately, the analysis does not lead to simple formulas, although it considerably simplifies the problem and allows us to rapidly compute the stationary network performance.

7. REFERENCES

- [1] V. Anantharam. The stability region of the finite-user slotted Aloha protocol. *IEEE Trans. on Information Theory*, 37:535–540, 1991.
- [2] F. Baccelli and P. Bremaud. *Elements of Queueing Theory, 2dn edition*. Springer-verlag, 2003.
- [3] G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3):535–547, 2000.
- [4] T. Bonald, S. Borst, N. Hegde, and A. Proutiere. Wireless data performance in multi-cell scenarios. In *Proceedings of ACM Sigmetrics*, 2004.
- [5] C. Bordenave, D. McDonald, and A. Proutiere. Random multi-access algorithms: A mean field analysis. In *Proceedings of Allerton conference*, 2005.

- [6] C. Bordenave, D. McDonald, and A. Proutiere. A particle system in interaction with a rapidly varying environment: Mean field limits and applications. *to appear in NHM*, available at <http://arxiv.org/abs/math/0701363>, 2007.
- [7] S. Borst, M. Jonckheere, and L. Leskela. Stability of parallel queueing systems with coupled rates. *Journal of Discrete Event Dynamic Systems*, 2007.
- [8] M. Durvy and P. Thiran. Packing approach to compare slotted and non-slotted medium access control. In *Proc. of IEEE Infocom*, 2006.
- [9] F. Kelly. Loss networks. *Annals of Applied Probability*, 1:319–378, 1991.
- [10] F. P. Kelly. *Reversibility and Stochastic Networks*. Wiley, Chichester, 1979.
- [11] L. Kleinrock. *Queueing Systems, volume 2*. Wiley, 1976.
- [12] W. Luo and A. Ephremides. Stability of n interacting queues in random-access systems. *IEEE trans. on Information Theory*, 45(5):1579–1587, 1999.
- [13] P. Nain. Analysis of a two-node Aloha network with infinite capacity buffers. In *Proc. of the Int. Seminar on Computer Networking and Performance Evaluation*, 1985.
- [14] L. Roberts. Aloha packet system with and without slot and capture. *ASS Note 8, ARPA, Stanford Res. Inst.*, 1972.
- [15] T. Saadawi and A. Ephremides. Analysis, stability and optimization of slotted Aloha with a finite number of buffered users. *IEEE Transactions on Automatic Control*, 26:680–689, 1981.
- [16] S. Sanghavi, L. Bui, and R. Srikant. Distributed link scheduling with constant overhead. In *Proceedings of ACM Sigmetrics*, 2007.
- [17] W. Szpankowski. Stability conditions for some multiqueue distributed systems: Buffered random access systems. *Advances in Applied Probability*, 26:498–515, 1994.
- [18] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Trans. on Automatic Control*, 37:1936–1949, 1992.
- [19] B. Tsybakov and W. Mikhailov. Ergodicity of slotted Aloha system. *Probl. Peredachii Infor.*, 15:73–87, 1979.

APPENDIX

A. A SYSTEM OF PARTICLES IN VARYING ENVIRONMENTS

In this section, we present the asymptotic analysis of a particle systems interacting with a random environment. The results derived here are used in the core of the paper. Their proofs can be found in [6].

A.1 The particle system

We consider N particles evolving in a state space $\mathcal{V} \times \mathcal{X}$ at discrete time slots $t \in \mathbb{N}$. \mathcal{V} is a finite set, and \mathcal{X} is at most countable. For simplicity we assume the particles are exchangeable. At time t , the state of particle i is $X_i^N(t) = (v_i^N, Y_i^N(t)) \in \mathcal{V} \times \mathcal{X}$. The first component v_i^N of

$X_i^N(t)$ is fixed, and is used to represent a type of *individual environment* as explained below. The state of the system at time t is described by the empirical measure $\nu^N(t) \in \mathcal{P}(\mathcal{X})$: $\nu^N(t) = \frac{1}{N} \sum_{i=1}^N \delta_{X_i^N(t)}$.

Each particle i is attached an individual environment whose state $A_i^N(t)$ at time t belongs to a finite space \mathcal{A} . $(A_i^N(t), t \geq 0)$ is a Markov chain independent of everything else and with kernel K_v^N where $v \in \mathcal{V}$ is defined by $X_i^N(t) = (v, Y_i^N(t))$. We assume that for all v , K_v^N converges to K_v when $N \rightarrow \infty$ and that $A_i^N(t)$ is in stationary regime for all t . Let π_v be the stationary distribution of a chain with kernel K_v .

The particles interact with a global environment whose state is denoted by $Z^N(t)$ at time t . The environment process takes its values in an at most countable state space \mathcal{Z} . We assume that the joint evolution of the particles, their individual environment and the global environment is Markovian, i.e., $(X_i^N(t), A_i^N(t), i = 1, \dots, N, Z^N(t), t \geq 0)$ is a Markov chain. Next we describe the marginals of the transitions of this Markov chain. In the system we consider, these marginals are sufficient to completely characterize the asymptotic system behaviour when the number of particles grows large. Nevertheless, when N is fixed, the transitions of the particles and of the global environment process may be correlated.

Global environment evolution. The process Z^N evolves as follows:

$$P(Z^N(t+1) = z | \mathcal{F}_t) = K_{\nu^N(t)}^N(Z^N(t), z),$$

where K_μ^N is a transition kernel on \mathcal{Z} depending on a probability measure μ on $\mathcal{P}(\mathcal{X})$, and where $\mathcal{F}_t = \sigma(\nu^N(0), Z^N(0), \dots, \nu^N(t), Z^N(t))$. In words, Z^N is a Markov chain whose transition kernel evolves with the empirical measure of the state of the particle system.

Evolution of the particles. We represent the possible transitions for a particle by a finite set \mathcal{S} of mappings from \mathcal{X} to \mathcal{X} . A s -transition for a particle in state $x = (v, y)$ leads this particle to the state $s(x) = (v, s(y))$. In each time slot the state of a particle has a transition with probability $1/N$ independently of everything else. If a transition occurs for a particle whose individual environment is in state $a \in \mathcal{A}$, this transition is a s -transition with probability $F_s^N(x, \nu, a, z)$, where x, ν , and z are respectively the state of the particle, the empirical measure, and the state of the environment process before the transition. Hence, in this state, a s -transition occurs with probability:

$$\frac{1}{N} F_s^N(x, \nu, a, z). \quad (8)$$

with $\sum_{s \in \mathcal{S}} F_s^N(x, \nu, a, z) = 1$ for all (x, ν, a, z) .

Note that, due to (8), the processes $A_i^N(t)$, Z^N evolve quickly while the empirical measure $\nu^N(t)$ evolves slowly. Remark that the s -transitions of the various particles may be correlated. Finally the process Z^N may depend on the transitions of the particles. The particle system is thus in interaction with its environment.

We make the assumptions that the transitions F_s^N converges to limiting transitions F_s when N tends to ∞ . We assume as well that the transition kernels K_α^N converges to K_α . These technical assumptions can be found in [6].

In what follows, we characterize the evolution of the system when the number of particles grows. According to (8), as $N \rightarrow \infty$, the chains $X_i^N(t)$ slow down hence to derive a limiting behavior we define: $q_i^N(t) = X_i^N([Nt])$. When

$N \rightarrow \infty$, the environment processes evolve rapidly, and the particles see an average of the environments. We define the average transition rates for a particle in state $x = (v, y)$ by

$$\bar{F}_s(x, \alpha) = \sum_{z \in \mathcal{Z}} \sum_{a \in \mathcal{A}} F_s(x, \alpha, a, z) \pi_v(a) \pi_\alpha(z). \quad (9)$$

A.2 Transient regimes

THEOREM 6. *Suppose that the initial values $q_i^N(0)$, $i = 1, \dots, N$, are i.i.d. and such that their empirical measure μ_0^N converges in distribution to a deterministic limit $Q_0 \in \mathcal{P}(\mathcal{X})$. There exists a probability measure Q on $D(\mathbb{R}^+, \mathcal{X})^1$ such that for all finite set \mathcal{I} of I particles:*

$$\lim_{N \rightarrow \infty} \mathcal{L}(q_i^N(\cdot), i \in \mathcal{I}) = Q^{\otimes I}, \quad \text{weakly.}$$

The above theorem states that the trajectories of the particles becomes independent when the number of particles grows large. The independence allows us to derive an explicit expression for the system state evolution. Define $Q^n(t) = Q(t)(\{x_n\})$ where $\mathcal{X} = \{x_n, n \in \mathbb{N}\}$. $Q^n(t)$ is the limiting (when $N \rightarrow \infty$) probability that a particle is in state x_n at time t .

THEOREM 7. *For all time $t > 0$, for all $n \in \mathbb{N}$,*

$$\frac{dQ^n}{dt} = \sum_{s \in \mathcal{S}} \sum_{m: s(x_m) = x_n} Q^m(t) \bar{F}_s(x_m, Q(t)) - \sum_{s \in \mathcal{S}} Q^n(t) \bar{F}_s(x_n, Q(t)). \quad (10)$$

The differential equations (10) have a natural simple interpretation, e.g. $\sum_{s \in \mathcal{S}} \sum_{m: s(x_m) = x_n} Q^m(t) \bar{F}_s(x_m, Q(t))$, is the total mean incoming flow of particles to state x_n .

A.3 Stationary regime

We now characterize the stationary behavior of the system in the mean field limit. To do so, we make two additional assumptions:

- (i) For all N , the Markov chain $((X_i^N(k), Z^N(k))_{k \in \mathbb{N}}$ is positive recurrent. The set of the stationary distributions π_{st}^N of a particle is tight.
- (ii) The dynamical system (10) is globally stable: there exists a measure $Q_{st} = (Q_{st}^n) \in \mathcal{P}(\mathcal{X})$ satisfying for all n :

$$\sum_{s \in \mathcal{S}} \sum_{m: s(x_m) = x_n} Q_{st}^m \bar{F}_s(x_m, Q_{st}) = Q_{st}^n \sum_{s \in \mathcal{S}} \bar{F}_s(x_n, Q_{st}), \quad (11)$$

and such that for all $Q \in \mathcal{P}(D(\mathbb{R}^+, \mathcal{X}))$ satisfying (10), for all n , $\lim_{t \rightarrow +\infty} Q^n(t) = Q_{st}^n$. Then the asymptotic independence of the particles also holds in the stationary regime, and Q_{st} is the limiting distribution of a particle:

THEOREM 8. *For all subsets $I \subset \mathbb{N}$ of finite cardinal $|I|$,*

$$\lim_{N \rightarrow \infty} \mathcal{L}_{st} \left((q_i^N(\cdot))_{i \in I} \right) = Q_{st}^{\otimes |I|} \quad \text{weakly in } \mathcal{P}(D(\mathbb{R}^+, \mathcal{X})^{|I|}).$$

¹ $D(\mathbb{R}^+, \mathcal{X})$ denotes the space of right-continuous with left limits functions with values in \mathcal{X} .