



Bandwidth allocation in large stochastic networks

Mathieu Feuillet

► To cite this version:

Mathieu Feuillet. Bandwidth allocation in large stochastic networks. Networking and Internet Architecture [cs.NI]. Ecole Polytechnique X, 2012. English. NNT: . pastel-00717671

HAL Id: pastel-00717671

<https://pastel.hal.science/pastel-00717671>

Submitted on 13 Jul 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



École doctorale de l'École polytechnique
Inria Paris-Rocquencourt



Thèse présentée en vue de l'obtention du grade de
Docteur de l'École polytechnique
Spécialité Mathématiques Appliquées

Allocation de bande passante dans les grands réseaux stochastiques

Bandwidth Allocation in Large Stochastic Networks

Mathieu FEUILLET

Soutenue publiquement le 12 juillet 2012 devant un jury composé de :

Thomas BONALD	Telecom ParisTech	Co-directeur
Carl GRAHAM	École polytechnique	Examineur
Laurent MASSOULIÉ	Technicolor	Rapporteur
Ravi MAZUMDAR	University of Waterloo	Examineur
Philippe ROBERT	Inria	Directeur
Alain SIMONIAN	Orange Labs	Examineur

et s'appuyant sur les rapports de :

Sem BORST	Technische Universiteit Eindhoven/Bell Labs
Adam WIERMAN	Caltech

Remerciements

Rédiger ces remerciements est un plaisir à deux titres ; il s'agit du point final de la rédaction de cette thèse et de la seule page qui sera lue par tous.

Je ne peux que commencer par remercier Philippe Robert pour avoir accepté de m'accueillir dans son équipe et d'encadrer ma thèse. Ces trois années à son contact auront été particulièrement enrichissantes. Au delà des techniques de renormalisation et autres méthodes probabilistes, son encadrement s'est caractérisé par une exigence scientifique et une grande disponibilité bienvenues. Il a cependant su me laisser une grande liberté dans le choix des sujets que j'ai abordés. Indépendamment des sciences, Philippe m'a communiqué son goût pour la course à pied donnant lieu à de nombreuses réunions importantes de 11h30 à 13h ainsi que deux semi-marathons. Mon seul regret est de ne pas avoir réussi à lui faire redécouvrir les joies de la natation.

Je tiens également à remercier chaleureusement Thomas Bonald qui m'accompagne dans mes aventures scientifiques depuis mon stage chez Orange. J'ai eu le plaisir de travailler avec lui sur des sujets qui lui tiennent à cœur et j'ai été très heureux d'enseigner à Télécom ParisTech dans son cours ainsi que de coécrire *Performance des réseaux et des systèmes informatiques* avec lui. J'espère que nous aurons encore de nombreuses occasions de travailler ensemble.

Jim Roberts a également été très présent tout au long de cette thèse. Il a toujours été très disponible et a suivi mes travaux avec intérêt. Il m'a également beaucoup appris sur l'art, difficile, de la rédaction scientifique. Pour tout cela, je lui suis très reconnaissant.

Tous mes remerciements vont à Sem Borst, Laurent Massoulié et Adam Wierman qui ont accepté d'être rapporteurs de ma thèse. Je suis également très honoré que Carl Graham, Ravi Mazumdar et Alain Simonian aient accepté de faire partie de mon jury.

Pendant ces trois années, j'ai eu le plaisir, outre Philippe et Thomas, de collaborer avec Matthieu Jonckheere, Balakrishna Prabhu et Alexandre Proutière. Je remercie en particulier Alexandre de m'avoir accueilli à Cambridge en plusieurs occasions et Matthieu de m'avoir donné l'occasion de passer une semaine d'échange scientifique en Uruguay. Je tiens également à citer Sébastien Monnet, Pierre Sens et Véronique Simon de l'équipe REGAL avec qui Philippe et moi avons eu le plaisir d'échanger sur les tables de hachage distribuées.

Je suis très reconnaissant à tous les membres de l'équipe RAP, Yousra Chabchoub, Virginie Collette, Christine Fricker, Emanuele Leoncini, Nada Sbihi, Florian Simatos, Jim et Philippe, pour leur bonne humeur et leur gentillesse tout au long de ces trois années. Un grand merci à Virginie pour sa gestion efficace des tâches administratives qui nous permet de nous concentrer sur nos travaux de recherche. Grazie mille, Emanuele, per la tua gentilezza e il tuo buon umore italiano. Merci à tous les coureurs et nageurs pour tous ces bons moments sportifs ou non, Thierry Martinez, Benoit Montagu, Jonathan Protzenko, Alexandre Pilkiewicz, Philippe et Emanuele. Encore quelques semaines à Inria à fréquenter certains d'entre vous et j'allais me mettre à faire du Coq et des systèmes de types, il était temps de

partir. Merci à tous les membres de l'équipe Algorithms et notamment, Alexandre Benoit, Nicolas Broutin pour l'organisation du groupe de travail et Bruno Salvy pour avoir suggéré à Emanuele, Jonathan et moi-même de faire revivre le séminaire des doctorants.

Je termine par ma famille qui me soutient dans mes études depuis deux décennies en me faisant confiance même si ce n'était pas forcément facile de voir où cela allait me mener ces dernières années. J'ai une pensée toute particulière pour mon épouse, Jessica, dont l'affection, la présence et le soutien me sont précieux, indispensables.

Contents

Remerciements	v
Résumé	ix
Abstract	xi
Chapter I. Introduction	
1. Preamble	1
2. From Circuit Switching to Packet Switching	2
3. Bandwidth Sharing Networks Modeling	5
4. Wireless Networks Modeling	10
5. Distributed Storage with Failures	13
6. Mathematical Tools	14
7. Presentation of Subsequent Chapters	
References for the introduction	29
Chapter II. On the Flow-Level Stability of Data Networks without Congestion Control 35	
1. Introduction	35
2. Flow-Level Model	37
3. Stability Conditions for a Linear Network	39
4. Asymptotic Stability	47
Appendix: Proof of Proposition 3.5	60
References for Chapter II	66
Chapter III. On the Performance of CSMA in Multi-Channel Wireless Networks 69	
1. Introduction	
2. Related Work	70
3. Model	71
4. Packet-Level Dynamics	72
5. Flow-Level Dynamics	73
6. Time-Scale Separation	76
7. Infrastructure-Based Networks	80
8. Conclusion	
References for Chapter III	88
Chapter IV. On a Transient Stochastic Network with Failures . 91	
1. Introduction	91
2. The Stochastic Model	94
3. The Overloaded Network	95

4. The Critical Case	97
5. The Time Scales of the Stable Network	100
Appendix. Generalized Skorokhod Problems	
References for Chapter IV	107
Chapter V. On the Transient Behavior of Ehrenfest and Engset	
Processes	111
1. Introduction	111
2. The Stochastic Model	114
3. Positive Martingales	115
4. A Fluid Picture	120
5. Super-Critical Regime	121
6. Sub-Critical Regime	123
7. Critical Regime	126
References for Chapter V	128
Comprehensive list of references	137

Résumé

L'objectif de cette thèse est de traiter trois problèmes relatifs aux réseaux de grande taille. Les outils utilisés à cette fin sont issus des probabilités et plus spécifiquement de la théorie des files d'attente. En plus d'améliorer la compréhension des systèmes étudiés, les travaux réalisés dans cette thèse ont permis de prouver des résultats théoriques nouveaux ainsi que d'illustrer certains phénomènes probabilistes.

Dans le Chapitre II, un modèle de réseau à partage de bande passante est étudié. Contrairement à ce qui avait été étudié dans la littérature, les utilisateurs n'utilisent pas de contrôle de congestion. On suppose qu'ils envoient des données avec un débit maximum et protègent leur transmission à l'aide d'un mécanisme basé sur des codes correcteurs d'erreur. Le modèle obtenu est analysé pour deux topologies de réseaux spécifiques : les réseaux linéaires et les arbres montants. À l'aide de limites fluides, les conditions de stabilité de ces réseaux sont établies. Ces limites fluides donnent lieu à un phénomène intéressant de moyennage stochastique. Ensuite, une autre méthode de renormalisation est utilisée pour prouver que la région de stabilité de ces processus converge vers la région optimale lorsque que les débits maximaux des utilisateurs deviennent infiniment petits par rapport à la taille des liens du réseau.

Dans le Chapitre III, on se propose d'étudier CSMA/CA, un algorithme d'accès implémenté dans certains standards de réseaux sans fil. Chaque lien est constitué d'un émetteur et d'un récepteur et un graphe d'interférence modélise les collisions potentielles entre les liens. Les arrivées et les départs de ces derniers sont prises en compte. Une approximation est faite en supposant que la dynamique d'accès au canal est infiniment plus rapide que la dynamique des arrivées et départs de liens. Il est alors établi que le CSMA permet une utilisation optimale des ressources radio dans le cadre des réseaux ad-hoc. Cependant, il est également prouvé que ce même algorithme n'est pas efficace pour les réseaux avec une station de base; dans ce cas, un biais en faveur des transmissions vers la station de base est observé. À la fin du chapitre, l'hypothèse simplificatrice est discutée.

Les deux derniers chapitres de la thèse sont consacrés à l'étude d'un grand système distribué de stockage de données avec pertes. L'objectif est d'estimer la vitesse de perte des fichiers ou la durée de vie d'un fichier donné. Dans le Chapitre IV, c'est le premier point de vue qui est adopté. Le système est considéré de manière globale. Le système est constitué d'un grand nombre de fichiers qui peuvent avoir chacun deux copies au maximum. Chaque copie disparaît au bout d'un temps aléatoire. Un mécanisme centralisé de sauvegarde permet alors de restaurer les copies perdues. Un fichier dont les deux copies ont été détruites est définitivement perdu. Le système est étudié dans le cas limite où le nombre de fichiers tend vers l'infini. Afin de décrire correctement le système, trois échelles de temps différentes sont étudiées. Ralentir le temps permet de comprendre le mécanisme de sauvegarde; laisser le temps inchangé permet de définir la capacité du système; accélérer le temps permet d'évaluer la vitesse de perte des fichiers. Le principe de moyennage stochastique est également observé à l'échelle de temps la plus rapide.

Dans le chapitre V, le point de vue d'un fichier donné est adopté. Des liens sont établis avec les modèles classiques d'Ehrenfest, issu de la physique statistique, et d'Erlang, issu des télécommunications. Des méthodes basées sur les martingales sont utilisées pour calculer la transformée de Laplace des temps d'atteinte de ces deux processus. Ces transformées permettent alors d'estimer le comportement asymptotique de ces temps d'atteinte et notamment le temps de disparition d'un fichier.

Abstract

The purpose of this thesis is to tackle three problems inspired by large distributed systems. The tools used for that purpose come from probability and more specifically queueing theory. The studies led during this thesis allowed to understand the behavior of the observed systems and algorithms but also to prove some interesting theoretical results and to emphasize some probabilistic phenomena.

In Chapter II, a bandwidth sharing network model is analyzed. Contrary to what has been studied in the literature, users do not use a congestion control mechanism; they are assumed to send data at their maximum rate and to protect their transmission against loss thanks to some error code. The obtained model is analyzed for two different topologies: linear networks and upstream trees. Using fluid limits, the stability condition of these networks is obtained. Due to these fluid limits, an interesting stochastic averaging phenomenon arises. Another scaling method is used to prove that the stability condition of these networks tends to the optimal one when the maximal rate of users is infinitely smaller than the links capacity.

In Chapter III, CSMA/CA, a classical channel access algorithm for wireless networks, is studied. Each link consists in a sender/receiver pair and an interference graph models potential collisions between links. Link arrivals and departures are taken into account by the model. An approximation is made by assuming that the channel access dynamics is infinitely faster than link arrivals and departures dynamics. It is proved that CSMA results in an optimal use of radio resources for ad-hoc networks. However, it is also established that this algorithm is not efficient for infrastructure-based networks where a bias in favor of upstream traffic is observed. At the end of the chapter, the time-scale separation assumption is discussed.

The last two chapters of this thesis are dedicated to the study of a large distributed storage system with failures. The purpose is to estimate the decay rate of the system or the durability of a given file. In Chapter IV, the first point of view is studied. The system is then considered as a whole. It consists in a large set of files which can have two copies. Each copy is lost after a random time. A centralized back-up mechanism allows to restore lost copies. A file with no copy is definitively lost. The system is analyzed when the number of files grows to infinity. In order to describe correctly its behavior, three different time-scales are considered. The slow time-scale allows to describe the back-up mechanism; the normal time-scale to define the capacity of the system and the fast time-scale to evaluate the decay rate. The stochastic averaging principle is also observed on the fast time-scale.

In Chapter V, the point of view of a file is taken. Links are established with the classical Ehrenfest and Engset models, from statistical physics and telecommunication. Exponential martingales methods are used to compute the Laplace transform of hitting times. The asymptotic behavior of these hitting times, in particular the durability of a file, is then estimated.

CHAPTER I

Introduction

Contents

1. Preamble	1
2. From Circuit Switching to Packet Switching	2
3. Bandwidth Sharing Networks Modeling	5
4. Wireless Networks Modeling	10
5. Distributed Storage with Failures	13
6. Mathematical Tools	14
7. Presentation of Subsequent Chapters	23
References for the introduction	29

1. Preamble

The global telephone network was the first machine built on a very large scale. It can be considered as a single machine since it is tightly coupled and, after all, its aim is to connect two telephones anywhere in the world. However, the global telephone network is no longer the largest network in use today. Slowly, it has been integrated in a much larger network: the Internet, where the traffic is not limited to voice but also includes data and, notably, an exponentially increasing volume of video. One could think that the Internet is just a communication network like the telephone network with different communication technologies since the aim is still almost the same: connecting two devices anywhere in the world. However, the nature of those two networks is completely different. Typically, the telephone network is deployed and managed by a single entity, the network operator. It is then designed globally and most of the design patterns assume that there is a central authority. The Internet has been designed with a completely different paradigm since no central authority is assumed and the network is supposed to operate in a fully distributed way. On the long-time scale, the deployment of the infrastructure has been decided by each actor individually with essentially no global coordination. On a short time scale, traffic management is performed by all involved actors including the final users. The complexity of the Internet is then several orders of magnitude beyond the complexity of the telephone network.

Furthermore, if the initial aim of the Internet was to connect two devices anywhere in the network, it is clearly not the case anymore. New usages have emerged and with them new complex systems with a very large scale. One can think about peer-to-peer networks as the first example. Here the aim is not to connect two machines but to get a specific file whoever the person transmitting the file. The number of examples is growing quickly in the last years with data centers, content oriented networks and cloud computing for instance.

Operating this huge system requires to solve design issues of different natures. One of the main concerns of a network operator is *dimensioning*. It seems natural that the bigger the network, the better the quality of service. However, building

a network is very expensive and the goal of the network operator is to build the smallest possible network that ensures an acceptable quality of service. This last notion is not clearly defined and raises the question of defining meaningful and tractable *performance metrics*. Moreover, it appears that increasing the capacity of the network is not sufficient and it is necessary to design *implementable* and *scalable* traffic control mechanisms that ensure an *efficient use* of the available capacity, this last notion being loosely defined at this point.

The aim of network performance analysis is to establish mathematical tools that define precisely some questions in this domain. The analysis of those models can provide some insight into how to address these issues. It allows us to establish dimensioning rules for networks but also design rules for the algorithms which are in operation in those large-scale systems. In this thesis, we propose to tackle three problems which are representative of this method. The first problem consists in evaluating the impact on the Internet of replacing congestion control by an error recovery mechanism. The second problem consists in evaluating the efficiency of a channel allocation algorithm in wireless networks. Finally, the last problem is to evaluate file resilience in a large distributed storage system with failures.

The organization of the introduction is as follows. We begin with a short history of modeling in the telephone network for the purpose of illustrating network performance evaluation. The bandwidth sharing networks are then introduced as a good representation of packet-switching networks. A detailed description of some models of wireless networks is given, followed by some discussions on the modeling of recent large-scale systems. Finally, we introduce in detail the mathematical tools that are used in this thesis. We conclude this introduction by a short description of the content of the different chapters of this thesis.

2. From Circuit Switching to Packet Switching

2.1. Circuit Switching: Erlang and Engset. As we said, network performance analysis was born at the beginning of the 20th century with the development of the first telephone networks. Basically, a telephone network can be seen as a set of circuits (or lines) which users access and that they keep busy during all the communication. For several decades, a physical circuit was created by operators on the route between users and preserved until the end of the communication. Today, no physical circuit is created but the principle remains in both wired and cellular phone networks. An incoming call is then accepted if there is a free circuit on each link on the route between the two users. If not, the call is rejected and the caller can try later hoping for another communication to end in the meantime. If we put aside the quality of the transmission and the reliability of the multiple devices, the *quality of service* that is perceived by users is the frequency at which a call is rejected, in other words, the *blocking probability*.

It is clear that the more circuits are used in the network, the lower the blocking probability; however, building a network is expensive. For that reason, two Scandinavian engineers, Engset and Erlang proposed independently two mathematical models to evaluate the blocking probability of a phone switch connecting N users to the rest of the world through C lines, as represented on Figure 1. Each user makes on average θ calls per minute with an average duration of σ minutes. A rule of thumb would be to say that there are, on average, $N\theta\sigma$ simultaneous calls, so that it is sufficient to take the smallest number of lines C such that $C > N\theta\sigma$. However, with such a method, we do not take into account the *random* behavior of users and we are then unable to evaluate the blocking probability. For this reason, Engset and Erlang proposed two mathematical models based on the theory of Markov

processes for a telephone switch and two very similar formulas for calculating the blocking probability of a phone call.

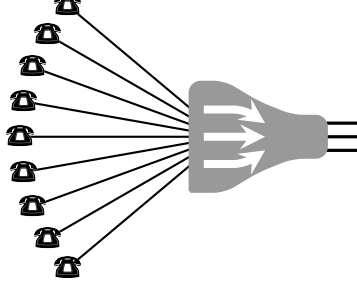


FIGURE 1. A telephone switch.

Engset proposed in [Eng98] what is maybe the most natural model. Indeed, he assumed that each user alternates between a communication of random duration with mean σ and an idle period of random duration with mean τ . At the end of each idle period, the user tries to make a call and if it is rejected, she starts a new idle period. In order to use the Markov theory, Engset assumed that the duration of communication and idle periods have exponential distributions. The behavior of such a system can then be described by a Markov process and Engset obtained the following formula for the blocking probability, with $\beta = \sigma/\tau$,

$$B = \frac{\binom{N-1}{C} \beta^C}{1 + \beta + \dots + \binom{N-1}{C} \beta^C}.$$

In fact, it has been proved later that the assumption of exponential distributions for communication durations and idle periods is not necessary and the previous formula remains true with general non-lattice distributions (see [Bon07], for instance).

Erlang had a slightly different approach and he proposed in [Er109] his famous model of a telephone switch. For that purpose, he made two assumptions. The first one is that users behave independently and that call arrivals can then be modeled by a Poisson process of intensity $\lambda = N/\tau$. The other one is that each call duration is an exponential random variable of mean σ . The system can then be mathematically described by a Markov process and Erlang obtained in [Er117] the following formula for the blocking probability, with the traffic intensity $\alpha = \lambda \times \sigma$:

$$(2.1) \quad B = E(\alpha, C) = \frac{\alpha^C / C!}{1 + \alpha + \dots + \alpha^C / C!}.$$

Inverting the previous formula, we are then able to derive the minimal number of circuits for a target blocking probability given the traffic intensity. It has been later established by Takács [Tak69] that the assumption of the exponential duration of calls is not necessary and the Erlang formula remains true in a more general framework. On the contrary, the assumption of independence between the users is crucial but verified in practice.

We define the load respectively as $\rho = N\beta/C$ and as $\rho = \alpha/C$ for the Engset and Erlang models. It appears that, at any given load, the Engset Formula always results in a lower blocking rate than the Erlang Formula: the Erlang Formula is the worst case and can always be used for dimensioning. Furthermore, at constant load, when the number of users N grows to infinity, the Engset Formula converges to the Erlang Formula. These last two statements are illustrated by Figure 2. This explains why the Erlang formula, simpler and more conservative, is commonly used in practice.

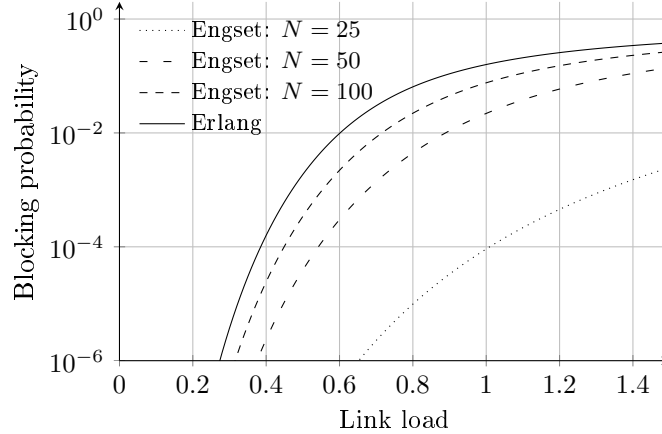


FIGURE 2. The Engset Formula and the Erlang Formula.

2.2. Extensions of Erlang Model. For a long period, the only existing networks were the telephone networks with the principle of circuit switching. The pioneering works of Engset and Erlang led to an entire field of research with many extensions of both models. We introduce here two extensions of the Erlang model and justify the need for them.

In 1965, Gimpelson proposed a multiclass Erlang model in [Gim65]. This model has been developed at a time where the Integrated Services Digital Network (ISDN) was emerging and the telephone network started to be used to transfer data flows over several circuits in parallel. In this model, users are divided into classes such that flows of each class arrive according to a Poisson process and have an exponential duration and require an arbitrary number of circuits. When a flow arrives and there are not enough free circuits, the flow is blocked and lost. As for the Erlang Model, we are able to derive an explicit expression for the blocking rate of each class. However, this formula is very difficult to compute since a naive approach has a complexity of order $O(C^K)$ where C is the number of circuits and K the number of classes. Fortunately, a recursive formula with a complexity $O(CK)$ has been derived by Kaufman in [Kau81] and by Roberts in [Rob81]. Due to its low complexity, this formula has been used extensively in the 80s for ISDN dimensioning.

Despite its more general scope, the previous model still applies to a single telephone switch and we reach the limits of this model when we want to understand the interactions between different classes of users on a more complex network topology. As soon as a network with more than a single link is considered, another problem arises: the routing. We do not address this issue and the routes are assumed to be known and static. This is typically the case in a telephone network where the routing tables are computed in a centralized way taking into account the topology of the network and traffic statistics. This assumption allows to build a simple class of models of telephone network with multiple links, called loss networks. This subject is very well introduced by Kelly in [Kel91].

In loss networks, users are divided into classes which are identified by the route in the network and the required number of circuits per user. A call is accepted in the network if there are enough available circuits in all links on its route. Otherwise, it is blocked and lost. As previously, calls arrive according to a Poisson process and have exponential durations. By studying the underlying Markov process, we are

able to derive a closed-form expression for the blocking probability of each class of calls.

The complexity to compute those probabilities is prohibitive in many practical situations. But contrary to the previous case, it is virtually impossible to find an efficient computation method. Despite having a closed-form expression for blocking rates, we are unable to use it directly. A usual approximation consists in considering the impact of each link individually and assuming that the blocking events of the different links are independent. Some studies have been done in order to justify this approximation. Whitt has proven in [Whi85] that this approximation is conservative, the obtained blocking rate is bigger than the actual one. However, this is not true in the multi-rate case. In [Kel91], Kelly gives an intuition of why this approximation is accurate when the load of each link is greater than 1 and the number of circuits very large. A proof has been provided by Theberge et al. in [TSM98] but only for some specific cases. Today, the general problem remains open.

2.3. Emergence of Distributed systems. Despite being the first large scale machine, the telephone network is still highly centralized. The deployment of the networks is planned and decided by the network operator who has a global vision of the situation. When operating, the routing schemes are typically chosen by finding the solution of a huge optimization problem in order to minimize the costs and use efficiently all the available resources.

The main drawback of such an architecture is that it is difficult to adapt to major changes such as a link failure. Furthermore, the bigger the network, the more intricate the different optimization problems to solve. It is then more and more difficult to increase the size of the network. For all these reasons, most of the large-scale systems in use today incorporate distributed mechanisms which allow to scale the system and to react rapidly to any major event.

In the following, we present three of these systems. Firstly, we introduce packet-switching networks which use distributed traffic control and routing mechanisms. However, most packet-switching networks are deployed and operated by network operators which perform some optimization in order to minimize the costs. The second example presented are wireless networks which are fully distributed since they are deployed and operated by users without any coordination, which means no centralized mechanism can be used in practice. Finally, we present a large-scale distributed mechanism which corresponds to different systems. However, due to its very large scale, any centralized mechanism seems unrealistic.

3. Bandwidth Sharing Networks Modeling

3.1. Queues and Queueing Networks. For a long period of time, telephone networks were the only real communication networks and, more generally, the only real information systems in service. After 1960, the emergence of computers and computer networks offered new systems to study and model. During that period, the most usual system consisted of a main frame able to compute tasks one after another and people were primarily interested in evaluating the *capacity* of the system and the *delays* encountered by a task submitted in the system. For studying such a system, the natural framework is the queueing theory which was already fairly developed at that time.

The simplest model for such a system is the $M/M/1$ queue, according to the classification introduced by Kendall in [Ken53]. In an $M/M/1$ queue, tasks arrive according to a Poisson process of parameter λ and each task needs a service time which is an exponential random variable of parameter μ . This is enough to define a Markov process describing the number of tasks waiting to be treated in the queue.

The first difference with the Erlang model is that the number of tasks is unbounded and it is natural to ask under what condition the number of waiting tasks will remain bounded. In fact, if $\lambda < \mu$, the Markov process is ergodic and admits for stationary distribution:

$$\pi(x) = (1 - \rho)\rho^x, \quad \forall x \geq 0,$$

where $\rho = \lambda/\mu$ is the load of the queue. The system is then said to be stable. On the contrary, if $\lambda > \mu$, the queue is transient and the number of clients tends to infinity almost surely when time tends to infinity. The system is then said to be unstable.

As in the Erlang model, we can question the validity of the Poisson and exponential assumptions. We then need to specify the service policy of the queue. The three most famous policies are: First In First Out (FIFO) where tasks are served in their arrival order; preemptive Last In First Out (LIFO) where the last customer in the queue interrupts the current service to be served; Processor Sharing (PS) where the service capacity is equally shared between all the tasks present in the queue. In the case of FIFO, as proved by Loynes in [Loy62], the notion of stability can easily be extended to a much more general queue, the $G/G/1$ FIFO queue, where inter-arrival times and service times are general stationary random variables with respective means λ^{-1} and μ^{-1} . Concerning the delay, the Laplace transform of the waiting time at equilibrium is known for the $GI/GI/1$ FIFO queue and notably reduces to the Pollaczek-Khinchin formula for the case of the $M/GI/1$ FIFO queue. The average delay is then given by

$$\delta = \frac{1}{\mu} \left(1 + \frac{\rho}{1 - \rho} \frac{\mathbb{E}(S^2)}{2\mathbb{E}(S)^2} \right),$$

where S is the service time of a task and $\mathbb{E}(S) = \mu^{-1}$. We can note that this model is sensitive to the service time distribution beyond the mean unlike the Erlang model which is insensitive. On the contrary, the $M/G/1$ PS and $M/G/1$ LIFO are insensitive and the mean delay is the same as in the $M/M/1$ case. For more details on those queues, one can refer to [Rob03, Chapter 2, Chapter 7].

A natural extension of queues are queueing networks. In this context, tasks join a queue when entering the network. When they leave a queue, they can leave the system or join another queue. Eventually, they reach their destinations and leave the network. A representative class of examples are Jackson networks. Incoming tasks arrive in queue i according to a Poisson process of parameter ν_i . After leaving queue i , a task joins queue j with probability p_{ij} or leaves the network with probability $1 - \sum_j p_{ij}$. In queue i , a task requires an exponential service time of parameter μ_i . An example of Jackson network is represented on Figure 3.

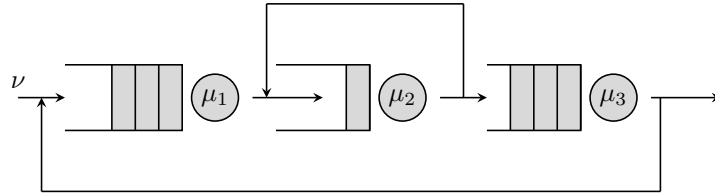


FIGURE 3. A Jackson network of $N = 3$ queues.

Jackson networks are a particularly interesting example since their behavior at equilibrium is well known. Consider the solutions of the traffic equations, i.e. the λ_i such that:

$$\lambda_i = \nu_i + \sum_j p_{ji} \lambda_j.$$

The quantity λ_i is the intensity of queue i arrivals. We know that the stability condition for a $M/M/1$ queue is $\lambda < \mu$ and, similarly, for a Jackson network, the stability condition is that $\lambda_i < \mu_i$ for any queue i . More surprisingly, Jackson proved in [Jac63] that at equilibrium, queue i is independent of other queues and behaves as an $M/M/1$ queue with arrival rate λ_i and departure rate μ_i and the stationary distribution of a Jackson network is

$$\pi(x_1, \dots, x_L) = \prod_{i=1}^L (1 - \rho_i) \rho_i^{x_i}, \quad \text{with } \rho_i = \frac{\lambda_i}{\mu_i}.$$

In particular, the delay experienced by a task in queue i is, on average, $1/(\mu_i - \lambda_i)$. The behavior of those networks can then be described very precisely and for that reason, Jackson networks (or a generalization) seem to be a good candidate to describe packet-switching networks. In the next Section, it is explained why it is not the case.

3.2. Packet Switching Networks and Congestion Control. In the early sixties, a new network design appeared. At that time, communication between computers was performed through the telephone network and, according to the circuit switching paradigm. There are several drawbacks with that principle, among which the fact that resources must be reserved along the route during the whole communication and the fact that routing is decided in a centralized way which is not robust to the destruction of a part of the network¹. This conducted Baran to propose a new network paradigm in [Bar62]: packet switching. Instead of allocating resources at the beginning of the communication, data is divided into packets which are sent into the network. When going through switches, routers and other networks nodes, packets are queued before being transmitted. The main advantage is that the network is organized with highly distributed mechanisms and the main drawback is that the throughput of a data transmission can vary over time according to the network load and there is no guarantee that a packet sent into the network reaches its destination. This is one of the issues that have been addressed by Cerf and Kahn in [CK74] by the creation of the Transport Control Protocol (TCP) which ensures end-to-end transmission error control. Basically, the receiver sends an acknowledgement for each received packet. If after some timeout, no acknowledgement has been received by the sender for a given packet, it is retransmitted.

If the original version of TCP seems to be sufficient to guarantee the integrity of a transmission, it does not guarantee an efficient use of available resources. In fact, a phenomenon called *congestion collapse* has been observed in 1986 in the Internet as described by Jacobson in [Jac88]. It appears that at certain times, most packets sent in the network were eventually dropped before reaching their destination resulting in a huge drop of the bandwidth and an inefficient use of the network resources. In order to avoid this phenomenon, Jacobson proposed the introduction of a congestion window in [Jac88]. The aim of that congestion window was to adapt the packet sending rate to the congestion in the network. TCP has been very well accepted and is used for the vast majority of the transmissions in the Internet.

The congestion control mechanism described above induces many correlations between packets in the network. Since data flows are cut into many packets, those packets arrive in bursts and the congestion control mechanisms induce correlations both in space and time: in space, since the congestion at a given bottleneck can change the sending rates of two flows coming from different places and in time since recovering from a congestion can be very long. These remarks are heuristic but it

1. This was a crucial issue during the Cold War and the Nuclear threat.

has been observed by Leland et al. in [LTWW93] and by Paxson and Floyd in [PF95] that packet traffic on an local area network is statistically very close to a self-similar process, excluding Poisson and Markov modeling of packets dynamics.

As a consequence, modeling packet-switching networks with queues is extremely intricate. For instance, Jackson networks, introduced above, or Kelly networks (see [Kel79]) seem perfectly adapted for modeling packet switching networks mainly due to their product-form stationary distribution. However, the Poisson arrivals assumption is crucial. Without it, those networks are very difficult to analyze.

3.3. Flow-Level Modeling. Looking at packet-switching networks at packet level may not be the right point of view for evaluating their performance. We saw in the previous section that mathematical models correctly describing packet traffic are very complex and despite extensive research in this area, there is no simple dimensioning rule like the Erlang formula for packet-switching networks, from the packet point of view. Fundamentally, we could think again of good performance metrics. Are packet delays the best metrics? Internet users are oblivious of packets and what they typically see is the average throughput obtained during a data transfer. If we want to evaluate the performance of a network correctly, we need to model networks at the flow level as proposed by Massoulié and Roberts in [MR00].

A flow in the Internet is a loosely defined object representing a stream of data with common characteristics (sender and receiver IP address, port number ...). Here, we use that term in order to represent the transfer of a particular document. The important property is that the flow is a continuous stream of packets using the same path in a network. It is then characterized by its starting time and its size.

We then represent network dynamics with a fluid model in order to get a more tractable mathematical model; flows become fluids of data instead of streams of packets. Thanks to this mechanism, we remove the granularity of flows. Here we assume that the sending rate of all flows is regulated by a congestion algorithm like TCP. In this model, it is assumed that TCP reacts instantly to any change in the network state and provides a fair bandwidth sharing between users. If the Poisson assumption is not reasonable at the packet level, it is more acceptable at the flow level (see [BBP⁺01] for instance).

We first illustrate this modeling principle with a single link of capacity C as illustrated on Figure 4. Flows are assumed to arrive according to a Poisson process of intensity λ . Each flow has an exponential size of mean σ and a maximum sending rate, which is typically the access rate of users, denoted by a . For simplicity, we assume that there exists an integer m such that $C = ma$. The congestion algorithm is assumed to achieve a perfectly fair sharing between flows so that the sending rate of a flow is a if $x \leq m$ and C/x otherwise, where x is the number of active flows.

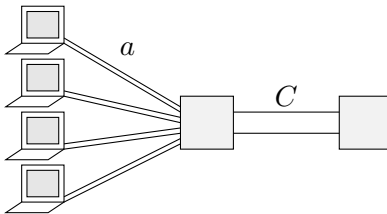


FIGURE 4. Users with peak rate a sharing a link of capacity C .

This model corresponds to an $M/M/m$ queue with arrival rate λ and departure rate $\mu = a/\sigma$. Denoting by $\alpha = \lambda/\mu$ the traffic intensity, we can prove that the

system is stable if $\alpha < m$, i.e. $\lambda\sigma < C$ which is again the natural condition. If this condition is satisfied, the stationary distribution is given by:

$$\begin{aligned}\pi(x) &= \pi(0) \frac{\alpha^x}{x!}, & \forall x \leq m, \\ \pi(x) &= \pi(m) \rho^{x-m}, & \forall x > m,\end{aligned}$$

with $\rho = \lambda\sigma/C < 1$ denoting the system load. With this notation, we are able to derive different metrics such as the congestion rate G which is the probability to have more than m active flows at a given time:

$$G = \frac{\rho E(\alpha, m)}{1 - \rho + \rho E(\alpha, m)},$$

where $E(\alpha, m)$ is the Erlang formula defined by equation (2.1). The mean throughput for an active flow is then given by:

$$\gamma = \frac{\rho(1 - \rho)m}{G + \rho(1 - \rho)m}.$$

The mean throughput is plotted for several values of m in Figure 5. We can remark that the smaller a is compared to C , the closer to the access rate the mean throughput is. This fact will be discussed in more detail in Chapter II and has been studied by Reed and Zwart in [RZ10].

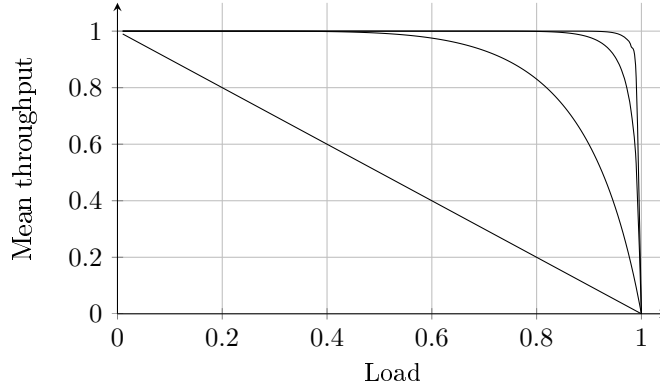


FIGURE 5. Mean throughput with respect to load ($m = 1, 10, 100, 500$, from bottom to top).

We now consider a network with L links and K classes of flows. Link l has a capacity C_l . A flow of class k has a fixed route R_k which is a sequence of links, a size which is an exponential random variable of mean σ_k , and a maximum rate a_k . Concerning the bandwidth sharing, the situation is not as straightforward as in the case of a single link. Flows of class k arrive according to a Poisson process of intensity λ_k . In [Kel99], Kelly gives heuristic arguments to justify that TCP approximately achieves *proportional fairness*. This means that for a given network state $(x_1, \dots, x_K)$, the throughputs of flows $(\varphi_1, \dots, \varphi_K)$ are the solution of the following optimization problem:

$$\begin{aligned}& \underset{(\varphi_1, \dots, \varphi_K)}{\text{maximize}} && \sum_{k=1}^K x_k \log(x_k \varphi_k) \\ & \text{subject to} && \sum_{k: l \in R_k} x_k \varphi_k \leq C_l, \quad \forall l \in \{1, \dots, L\}, \\ & && \varphi_k \leq a_k, \quad \forall k \in \{1, \dots, K\}.\end{aligned}$$

A more general class of fair sharing policies, the α -fair policies, has been proposed by Mo and Walrand in [MW00].

In this context, the natural stability condition would be

$$\sum_{k:l \in R_k} \lambda_k \sigma_k < C_l, \quad \forall l.$$

We can prove that if there exists l_0 such that $\sum_{k:l_0 \in R_k} \lambda_k \sigma_k > C_{l_0}$ then the system is unstable, as done by Bonald et al. in [BMPV06]. This means that this natural stability condition is in fact optimal. However, the proof that this is the actual stability condition is not straightforward. It has been proved for α -fair policies by Bonald and Massoulié in [BM01]. Massoulié extended this result in [Mas07]. A similar result has been obtained by de Veciana et al. in [dVKL01]. Performance of Proportional Fairness is evoked in Section 6.2.

The underlying assumption of all these works is that most of the traffic is regulated thanks to congestion control mechanisms. However, there is no clear reason why users should still enforce such a mechanism in the future. In [BFP09], a model of bandwidth sharing network where users do not use congestion control anymore is introduced. This model is studied in Chapter II.

So far, we have considered models of wired networks. In the next section, we explain why flow-level modeling is also useful in wireless networks.

4. Wireless Networks Modeling

4.1. Wireless Access Algorithms. In the previous section, packet-switching networks are presented from a macroscopic point of view. In particular, important design mechanisms occurring at link level are completely neglected. However, this is a domain where network performance analysis and queueing theory have proved to be very useful.

The problem is the following: there is a single communication channel which is shared between several users. Each of them generates messages to be transmitted on the channel to another user. If two users transmit messages at the same time, there is a collision and both transmissions are lost. Historically, this situation corresponds to a local area network where all the local stations are connected to a unique wired line linked to the Internet. Thanks to the development of some technologies and, in particular, switches, this design issue has disappeared since there is virtually no collision anymore. However, an increasing fraction of communications take place through wireless links. In that case, the communication channel is a frequency bandwidth and collisions should be considered.

There are some important constraints for the design of a wireless access algorithm. First, the number of users present in the network is unknown; it can vary over time and for some frequencies, there is no way to control it since anybody is allowed to use devices as they wish. As a consequence, it is impossible to define a centralized authority regulating access to the channel due to the fully distributed nature of the network; moreover the access control mechanism also has to be completely distributed. Secondly, in most cases, it is impossible to emit and listen on the channel at the same time since any communication of other users would be covered by one's own emitting. Thirdly, each user has only a partial vision of what happens in the network. In particular, without specific mechanism, she cannot even know if her own message has been received. Those different points make the design of an algorithm quite challenging.

4.2. A First Example: Aloha. The first algorithm which was designed for wireless communication is Aloha introduced by Abramson in [Amb70]. The motivation was to connect a central time-sharing computer based in Oahu, the main island of

Hawaii to users in other islands of the archipelago. In the initial design, two distinct frequencies were used. The first one was dedicated to the main computer for transmitting packets to users and the second frequency was shared by users for sending packets to the main computer. Since there was the possibility of a collision, the following mechanism was used. Each user transmits any incoming packet. If a collision is detected, the packet is retransmitted after a random time. The randomness allows to schedule the transmissions. Without it, the system would be blocked as soon as two users have packets to send. A first improvement of this algorithm has been proposed: it consists in dividing time in slots and allowing each user to start a transmission only at the beginning of a slot. The algorithm is then the following: at the beginning of each slot, if a user has a packet to send, she transmits it with probability p . This algorithm is called slotted-Aloha.

In that context, many analyses have been performed to evaluate the performance of this system. Assume that there are N users and each user i receives packets according to some stationary process of mean λ_i packets per time unit and the transmission probability of user i is p_i . When all users have a packet to send, the probability that user i successfully transmits her packet is

$$p_i \prod_{j \neq i} (1 - p_j).$$

One can think that the stability condition here is

$$\frac{\lambda_i}{p_i \prod_{j \neq i} (1 - p_j)} < 1.$$

This is true but only in a very specific case where the left member of the previous inequality does not depend on i as proved by Bonald et al. in [BBHP04]. In the general case, the stability region is known only for 2 users and it has been proved that, for $N \geq 3$, the stability condition depends not only on the intensity of the arrival rates but also on other statistical parameters as proved by Szpankowski in [Szp93]. Recently, Bordenave et al. proposed in [BMP08] a very accurate approximation of the stability region using mean field techniques.

Today, Aloha is still used in some specific situations such as a GSM cell where it regulates access to the control channel. However, the basic principles are still valid in the CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) which is used in the 802.11 standard for instance. The basic principle of CSMA is the following. When a user wants to send a packet, she listens to the channel. When the channel is free, she waits for a random period of time, the back-off period, during which she listens to the channel. If it is still free after the back-off period, she transmits its packet. The performance of this algorithm is analyzed in some specific cases in Chapter III.

4.3. Max-Weight Scheduling. In the previous section, we considered a very specific network where all users are connected to a central node. In practice, the topology of networks is more and more diverse and there is a need for a more general model of wireless networks. In particular, we need to take into account the topology induced by the positions of users. Indeed, if two users are close then they will interfere with each other. On the contrary, if they are distant, they can use the communication channel at the same time.

Here is the resulting model. The network is a set of wireless links, consisting in a transmitter-receiver pair. All these links use the same radio channel and, depending on the distance between users and radio conditions, they can use the channel at the same time without interfering with each other or not. We then define an interference graph (V, E) in the following way. Each vertex of this graph

is a link of the network and two links can transmit at the same time on the radio channel if and only if they are not connected in the interference graph. Each edge in the graph then represents a conflict. An example of an interference graph is given in Figure 6; links 1 and 4 are distant enough to transmit at the same time but link 3 is very close to all others links and cannot transmit simultaneously with any of them.

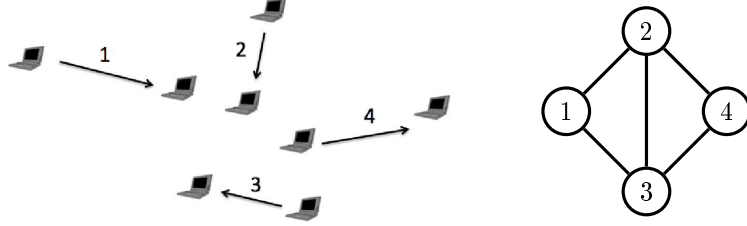


FIGURE 6. An ad-hoc wireless network with 4 links and its interference graph.

Packets are assumed to arrive according to Poisson processes with intensity λ_i for link i . The size of each packet is an exponential random variable of mean σ_i . The speed of transmission of each packet is assumed unitarian and we define the traffic intensity for link i as $\rho_i = \lambda_i \sigma_i$. As previously, we are interested in characterizing the stability region for a given algorithm. However, in order to evaluate the performance of these algorithms, it might be interesting to characterize the capacity region; the set of traffic intensities for which there exists an algorithm stabilizing the whole system. This capacity region has been characterized by Tassiulas and Ephremides in [TE92] as the convex hull of all probability distributions on feasible schedules where a feasible schedule is a set of active links without conflict. Tassiulas and Ephremides also provide an algorithm which stabilizes the system whenever possible: the Max-weight scheduling algorithm. Each time a packet is transmitted or arrives, a new schedule is chosen such that it maximizes the sum of the lengths of the queues of the active links. This algorithm is remarkable since it is the first throughput-optimal algorithm in this context. However, it suffers a main drawback: it is a centralized algorithm and finding a max-weight schedule is NP-hard for general interference graphs (see the book of Garey and Johnson [GJ90]).

Since then, several algorithms have been proposed to approximate Max-Weight Scheduling. In the context of switches, McKeown proposed in [MVW93] to use the Maximal Queue Scheduling algorithm. The principle is simple: each time a packet is transmitted or arrives, a new schedule is chosen in the following way. The link with the maximum number of packets in the queue is activated. All conflicting links are inactivated. We proceed recursively on the remaining links until all links are activated or inactivated. This algorithm performs quite well in the context of switches but, for wireless networks, it can result in an inefficient use of the available bandwidth as proved by Dimakis and Walrand in [DW06]. Furthermore this algorithm is not distributed. There exists a distributed algorithm approximately achieving Maximal Queue Scheduling, such as the one proposed by Alon et al. in [Alo86]. However, this algorithm is suboptimal as proved by Chaporkar et al. in [CKLS08].

More recently, some classes of optimal distributed algorithms have emerged. Those algorithms are based on CMA/CA scheme described in the previous section. In [JW08], Jiang and Walrand use a learning algorithm to adapt the back-off

parameter to the network conditions. This algorithm is distributed but requires a global starting time for learning the network conditions and is not able to adapt to arrival and departure of new nodes. Furthermore, the convergence rate of the algorithm is slow resulting in poor short-term performance. In [SS11], Shah and Shin propose a different approach by adapting the back-off rate of the link to its queue size; the more packets in the queue, the more aggressive the link to access the channel. The authors have proved that the algorithm is optimal assuming that the back-off rate depends on the queue size in a logarithmic way. Bouman et al. show in [BBLP11] that this assumption results in poor throughput performance for some topologies.

4.4. Flow-Level Modeling. All algorithms that are presented in the previous section are packet-based algorithms. Basically, the scheduling decisions are taken according to the arrival rate of packets or the number of packets waiting in the queue. Most of the proofs use strong assumptions on the traffic statistics of the arrival processes of packets. However the remarks made in Section 3.3 are still valid. Typically, packet arrivals are very bursty and state-dependent.

In [vdVBS09], Van de Ven et al. show that Max-Weight Scheduling can be inefficient if the packet arrivals are very bursty and if the flow dynamics are taken into account. This justifies that, again, algorithms need to be designed at flow level by taking into account the flow dynamics. In Chapter III, a new distributed algorithm is proposed and we prove that this algorithm is throughput-optimal under some mathematical assumptions.

5. Distributed Storage with Failures

New usages of the Internet have emerged and with them new complex systems with a very large scale. One can think of peer-to-peer networks as the first example. Here the aim is not to connect two machines but to get a specific file whoever the person transmitting the file. Peer-to-peer mechanism enjoys a huge success and represents a large fraction of the offered traffic in the Internet today, even if this fraction is decreasing over time due to the success of video streaming. The aim of peer-to-peer networks is to exchange files between users without requiring a large fixed infrastructure composed of servers. More recently, people proposed to build large and completely distributed systems such as distributed hash tables. The basic idea is to store data in nodes which are located anywhere in the network and form a logical ring independent of the Internet topology. For each file, an identifier is computed thanks to a hash function and the file is stored in the node which is the closest to its identifier on the logical ring. In order to ensure the resilience of data, the main server duplicates this file into its neighbors on the logical ring. All the difficulty comes from the fact that the topology is evolving rapidly due to node failures, arrivals and departures. Evaluating the durability of a file is then a very difficult task. A first approach for estimating this durability has been proposed by Picconi et al. in [PBS07a] and [PBS07b].

As mentioned above, video streaming demand has been increasing significantly for the last few years. Similarly the needs for computing capacities has drastically increased with the development of cloud computing services. For all these reasons, data centers have grown over the last years on a very large scale. For instance, Akamai Technologies runs more than 105,000 servers worldwide in March 2012 and Google is estimated to run more than 900,000 servers even if those figures have not been publicly confirmed by the company². As data centers scale up, the total storage capacity increases but it also makes failures more common. For those

2. Source: <http://www.datacenterknowledge.com/>

systems, failures arise on a daily basis and there is a real need for protection against data loss. Even if there already exist back-up solutions such as RAID [PGK88], it appears that those solutions are not adapted. For instance, rebuilding a 500GB RAID while still in service can take more than one day and there is typically a chance of a second failure of 0.1% in the meantime. Even if mechanisms have already been proposed such as [XMS⁺03] by Xin et al., there is today, to the best of our knowledge, no real analysis of very large storage systems with failures. There is a large literature on queues with failures but each time, a small number of queues are considered. We can notably evoke the work of Gallet et al. [GYJ⁺10] which considers a statistical model to analyze large failure traces.

Those two situations led us to study those large scale storage systems with failure. In Chapter V, we consider a single file and evaluate its durability with a very simple model: the Erlang model. We are then able to derive precise asymptotics when the number of copies in the system becomes very high. However, this model is very simple and does not take into account the fact that many files are stored by the same system. In Chapter IV, we introduce a model where two copies of each file are stored. It is assumed that there is a centralized back-up mechanism, we are then able to derive precise asymptotics of the system evolution with a scaling analysis. In an on-going work, we relax the assumption of the centralized back-up mechanism and we consider a fully distributed system. Asymptotics of the behavior of the system are derived thanks to a mean-field analysis.

6. Mathematical Tools

6.1. Markov Processes. In 1906, Andrei Markov introduced the chains named after him in order to produce a counter-example to Nekrasov who wanted to prove that the independence of random variables is necessary for the weak law of large numbers. The Markov chains appeared then in the middle of a pseudo-scientific and philosophical argument and for a purely mathematical purpose³. Today, Markov chains and processes are one of the very basic tools of mathematical modeling and a large amount of theory has been developed on the subject.

In the present work, most models are in fact (multi-dimensional) queues. In that particular context, in order to obtain Markov processes, two assumptions are necessary. The first one is that tasks (or clients, flows...) arrive according to Poisson processes. The second is that service times are exponential random variables.

On the one hand, Poisson arrivals is not an unrealistic assumption in many cases. For instance, in the case of telephone networks, each user behaves independently of the others and typically has a small activity. In the limit, when the number of users tends to infinity and the activity tends to 0, at constant load, the obtained arrival process is a Poisson process. This is why the Engset model tends to the Erlang model and the second one provides a very good approximation of the first one. In practice, this is true even for a relatively small number of users as illustrated by Figure 2. However, we mentioned that there are domains where Poisson arrivals are not realistic such as packet arrivals in packet-switching networks.

On the other hand, representing service times by an exponential random variable is something much more difficult to justify. In the example of telephone calls, Erlang and Engset already observed that this assumption is not realistic. In fact, the duration of a call is better represented by a power-law distribution which is very different from the exponential distribution. However, as explained above, the results obtained by Engset and Erlang are insensitive to the call duration distribution and are still valid for general service times. This is a very specific situation which arises in a significant number of cases: the stationary distribution of the

3. For more details, see the article of Seneta [Sen96]

process does not depend on detailed statistics but only on the mean. Many works have been carried on that subject. One can have a look at the work by Burman [Bur81] and Bonald [Bon07]. Of course, all Markov processes are not insensitive. In that case, one can consider looking at more general processes. However, it seems natural to look first at Markov processes to dissociate difficulties due to the original problem from those due to generalization.

6.2. Analysis at Equilibrium. In most cases, Markov processes and, more generally, stochastic processes are studied at equilibrium. Basically, a Markov process is stationary if its marginal distribution does not vary over time. The first question is then to determine in which cases this situation is possible. For instance, it is indicated in 3.1 that the $M/M/1$ queue is stable and admits a stationary distribution if and only if $\lambda < \mu$. In full generality, the stability of a Markov process can be obtained by proving that the general balance equations of the process admit a summable solution. However, despite specific classes of processes described below, this is not tractable in most cases. One of the most popular tools to prove the stability of a Markov process is Foster's criterion. Basically, it consists in proving that when the process is outside a compact set, there is a negative drift driving back the process in the compact set. Despite being quite powerful, this criterion requires to find a Lyapunov function which can be a difficult task. For more details on Foster's criterion, one can read the book of Brémaud [Bré01]. A proof of stability using Foster's criterion is detailed in Chapter III.

In fact, there is a large class of processes for which the global balance equations can be replaced by a much simpler system of equations; those are the reversible processes. Roughly speaking, a Markov process $(X(t))$ is said to be reversible if the process in reverse time $(X(T-t))$ has the same distribution as $(X(t))$ on $[0, T]$ for $T > 0$. In this case, we can prove that any stationary measure of the process must satisfy the local balance equations which state that the frequency of jumps from state i to state j is equal to the frequency of jumps from state j to state i . This is stronger than the global balance equations which simply state that the frequency of jumps into state i are equal to the frequency of jumps from state i . In particular, if the process is reversible, any stationary measure admits a closed form expression. Proving the stability is then reduced to proving the summability of this stationary measure. For more information on reversible networks, one can refer to the book of Kelly [Kel79]. Examples of reversible process are studied in Chapters III and V.

Even when the stability has been established, for most Markov processes, the stationary distribution does not admit a closed form expression and this limits any analysis that can be performed on the stationary regime. For instance, this is the case for bandwidth sharing networks introduced in Section 3.3. The fact that the stability condition of this model is the optimal one ensures that network resources are used in an efficient way. However, it does not necessarily imply a good quality of service, which typically is the mean throughput, depending on the stationary distribution. Bonald and Proutière then proposed in [BA03] to approach Proportional fairness and other usual bandwidth allocations by an allocation, *Balanced Fairness*, such that the underlying Markov process is reversible. It has been established that Balanced Fairness is a good approximation of Proportional Fairness. A survey on Balanced Fairness and its relation to Proportional fairness can be found in [BMPV06]. The approximation of Proportional Fairness by Balanced Fairness is discussed by Massoulié in [Mas07].

Reversible networks are not the only class of stochastic networks for which the stationary distribution is known. The Jackson networks introduced in Section 3.1 are generally not reversible. For instance, the network represented on Figure 3 is not reversible since there are customers migrating from queue 2 to queue 3

and not the contrary. However, the stationary distribution of Jackson networks is known and admits a simple product-form expression. There is in fact a large class of stochastic networks called quasi-reversible networks which also admit a product-form stationary distribution. An introduction on that subject can be found in the book of Serfozo [Ser99].

6.3. Martingale Methods. One of the first natural questions that arise when studying a Markov process is the evaluation of hitting times. Typically, starting from a point x , one wants to know with which probability the process reaches y and what is the probability distribution of T_y , the hitting time of y by the considered process. If the process is recurrent, the probability of hitting y starting from x is always 1 but we do not know anything about the distribution of T_y . Different methods have been developed to evaluate hitting times. Most of them rely on calculating the Laplace transform of hitting times. In the case of birth-death processes, Karlin and McGregor develop a method based on orthogonal polynomials in [KM57]. We present here a method to compute the Laplace transform based on a martingale method used by Kennedy in [Ken76]. An introduction to this method on different examples can be found in [Rob03].

For some processes, there exists a family of non-negative martingales which allow to get some useful information about hitting times. We illustrate this on a simple $M/M/1$ queue where clients arrive according to a Poisson process with parameter λ and each of them requires an exponential service time with parameter μ . If we consider the process $(L(t))$ describing the number of clients in the queue, it is a Markov process on $\mathbb{N}$ with transition rates

$$q(x, x+1) = \lambda, \quad q(x, x-1) = \mu \mathbb{1}_{\{x>0\}}.$$

The simplest example of a hitting time here is the time at which $L(t)$ hits 0 starting from $L(0) = N$. In this case, the discontinuity of the infinitesimal generator in 0 does not play a role and we can simply consider the biased random walk $(Z(t))$ with transition rates $q(x, x+1) = \lambda$ and $q(x, x-1) = \mu$. We denote by T_0 the hitting time of 0 by $(Z(t))$. Using a family of non-negative martingales, similar to the exponential martingale for the Brownian motion, we are able to compute the Laplace transform of T_0 . By analyzing it, we can obtain easily that T_0 is finite with probability $(\mu/\lambda)^N$ when $\lambda > \mu$. In order to obtain the asymptotic behavior of the hitting times when $\lambda < \mu$ and N tends to infinity, we just have to analyze the convergence of the Laplace transform. In particular, there is a central limit theorem for T_0 when N tends to infinity:

$$\frac{T_0 - N/(\mu - \lambda)}{\sqrt{N(\mu + \lambda)/(\mu - \lambda)^3}} \Rightarrow \mathcal{N}(0, 1).$$

We now consider the hitting time of N when $L(0) = N$. In that case, it is not possible to neglect the boundary effect in 0. However, it is possible to construct a martingale for $(L(t))$ by choosing a linear combination of two martingales of $(Z(t))$ which preserves the martingale property in 0. We are then able to derive an explicit expression for the Laplace transform of T_N . From this expression, we can derive the asymptotic distribution of $(\lambda/\mu)^N T_N$ when N tends to infinity and $\lambda < \mu$, this is an exponential random variable with parameter $(\mu - \lambda)^2/\mu$.

These methods for computing the Laplace transform and derive asymptotics of hitting times distributions are developed in Chapter V for the classical Ehrenfest model and the Engset model which has already been introduced in Section 2.

6.4. Scaling Methods. The method introduced in the previous section is quite powerful but requires to use an ad-hoc martingale family that can be difficult to find. Scaling techniques then provide a useful toolbox to get a first-order description

of such systems. During this PhD, different kinds of scalings in time and space have been considered.

The basic idea of the scaling in time and space is the following. We consider a sequence of Markov processes $(X^N(t))$ and we choose appropriate sequences (Ψ_N) and (Φ_N) in order to analyze the evolution of the sample path of

$$\left(\frac{X^N(\Psi_N(t))}{\Phi_N} \right)$$

when the parameter N tends to infinity. The sequence Φ_N defines the macroscopic scale of the process $(X^N(\Psi_N(t)))$, that is to say the area of the state space where the process lives. The time scale $t \mapsto \Psi_N(t)$ allows to focus on some specific part of the sample paths. As illustrated by Chapter IV, there could be several time scales of interest for the same sequence of processes $(X^N(t))$. We introduced in Sections 6.4.1 and 6.4.2 two representative examples of techniques used in this thesis.

6.4.1. Fluid Limits and the $M/M/1$ Queue. The goal here is to obtain a first order description of the process $(L(t))$. In particular, one would like to understand what is the evolution of the process when it starts from a very large initial state $L(0) = N \gg 1$. Since we increase the initial state, it is quite natural to renormalize the process with $\Phi_N = N$. Additionally, we must increase the time window in order to correctly evaluate the variations of the process. Since the transition rates are bounded, it is necessary to use the time scale $t \mapsto Nt$ to see variations of order N .

If we denote by T_0 the first time where $(L(t))$ hits 0, it is clear from the transition rates that $(L(t))$ can be expressed as the difference between two Poisson processes of respective rates λ and μ :

$$L(t) = L(0) + \mathcal{N}_\lambda(t) - \mathcal{N}_\mu(t).$$

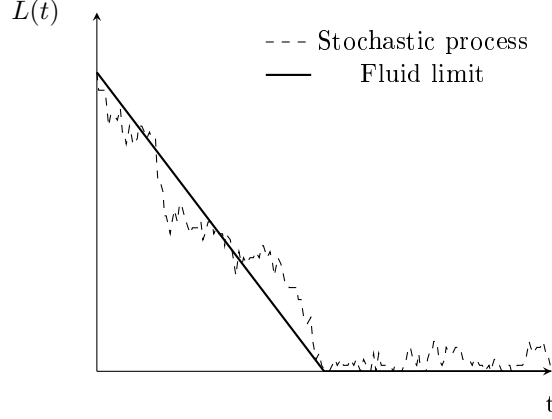
The strong law of large numbers gives that $\mathcal{N}_\lambda(Nt)/N$ tends to λt when n tends to infinity. Furthermore, we can prove that if $L(0) = N$ and $\lambda < \mu$, T_0/n converges in distribution to the constant $1/(\lambda - \mu)$. On the contrary, we can prove that, if $\lambda > \mu$, T_0/N is infinite with high probability for N sufficiently large. All this suggests that the right time scale is then $t \mapsto Nt$ and, indeed, we define the following sequence of processes:

$$\bar{L}_N(t) = \frac{L(Nt)}{N}, \quad \text{with } L(0) = N,$$

and we can prove that $\bar{L}_N(t)$ converges in distribution to the deterministic process $t \mapsto \max(0, 1 + (\lambda - \mu)t)$. Almost all ingredients of the proof have been given. The only missing part is for $\lambda < \mu$ and $t > T_0/N$, we have to prove that $\bar{L}_N(t)$ converges to 0. This comes from the fact that the time for an ergodic $M/M/1$ queue to reach the level εN is of order $(\mu/\lambda)^{\varepsilon N}$ and is typically much larger than Nt . For a rigorous presentation of this convergence, one can refer to the book of Robert [Rob03].

In fact, we can get more insight in the initial process by looking at the diffusion around the limiting process of $(L_N(t))$. Interestingly, the variations are of order $O(\sqrt{N})$ as long as the queue has not reached 0 and are of order $O(1)$ after. This fact is illustrated by Figure 7.

As explained in Section 6.2, proving the stability of a Markov process can be difficult and Foster's criterion can be very tricky to use. Therefore, we need a more qualitative approach, and this is what fluid limits have been designed for. In particular, one can remark that this scaling is very specific in the sense that the infinitesimal generator of the original process is left untouched. One of the first attempts to use fluid limits for establishing stability has been made by Malyshev in [Mal93]. In this article, he studied random walks on $\mathbb{N}^K$ where, in fact, transitions depend only on the number of empty queues. In particular, transition rates are

FIGURE 7. The dynamics of the fluid limit of an $M/M/1$ queue.

bounded. One can note that Jackson networks introduced above are an example of such a process. Malyshev established here relations between a random walk and some dynamical system which is a fluid limit of the considered process but the convergence is not established and no formal definition of a fluid limit is given.

This definition of fluid limits and their link with the stability of stochastic processes are introduced by Rybko and Stolyar in [RS92] where they proved the stability of FIFO queues under optimal stability conditions and studied some interesting example of a network with a priority mechanism. There are two links and two classes. Customers of classes 1 and 2 arrive according to independent Poisson processes of respective intensities λ_1 and λ_2 . Customers of class 1 are served by link 1 at rate μ_{11} and move to link 2 where they are served at rate μ_{12} . Customers of class 2 are served by link 2 at rate μ_{22} and then move to link 1 where they are served at rate μ_{21} . Link 1 gives priority to customers of class 2 and link 2 gives priority to customers of class 1. The principle of this network is summarized on Figure 8.

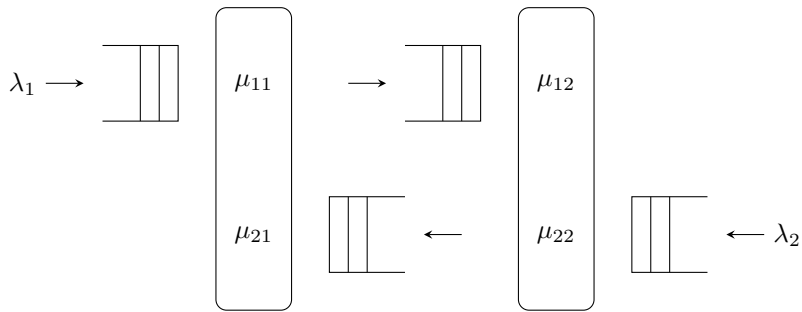


FIGURE 8. The Rybko-Stolyar network

For such a network, the optimal stability condition is

$$\frac{\lambda_1}{\mu_{11}} + \frac{\lambda_2}{\mu_{21}} < 1, \quad \frac{\lambda_1}{\mu_{11}} + \frac{\lambda_2}{\mu_{21}} < 1$$

However Rybko and Stolyar have proved that there is an additional cross condition:

$$\frac{\lambda_1}{\mu_{12}} + \frac{\lambda_2}{\mu_{21}} < 1$$

Notably, this is the first example of a random queueing network which is unstable although the optimal stability conditions are satisfied. The proof of instability is not strictly based on fluid limits but they used them to get an intuition of the behavior of the system; the proof is quite close to fluid limits methods. In fact, fluid limits describe very well the oscillations of the system: incoming customers of class 1 are queued while customers of class 2 are served and leave the system and vice-versa. This dynamics of the fluid limit is represented on Figure 9 in the case where $\mu_{11} = \mu_{22} = +\infty$. In that case, the system can be completely described by the number of customers in each class $L_1(t)$ and $L_2(t)$ and the active server. One can note that a very similar example has been proposed in a deterministic setting by Lu and Kumar in [LK91].

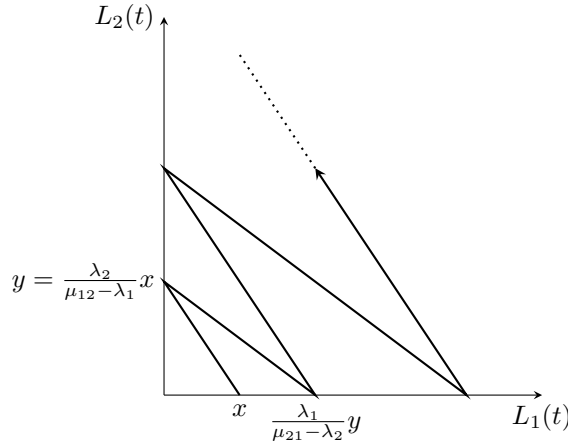


FIGURE 9. The dynamics of a fluid limit of a transient Rybko-Stolyar network

In [Dai95], Dai studied a large class of queueing networks. His aim was to propose a generic method for studying the stability of Markov processes using fluid limits. A criterion of transience has been proposed by Meyn in [Mey95]. A good introduction on fluid limits and a generalization of the criterion of stability can be found in [Rob03]. Many examples of fluid limits are studied in [Bra08]. Fluid limits are used to prove the stability of some networks in Chapters II and III.

6.4.2. Kelly's Scaling and the $M/M/\infty$ Queue. All stochastic processes considered in the previous section have bounded transition rates. For all of them, the fluid limit offers a good first-order description of their behavior and give a qualitative method to establish their stability. Here, we want to study a process with unbounded transition rates: an $M/M/\infty$ queue. Clients are assumed to arrive according to a Poisson process and since there is an infinity of servers, each client is served at rate μ . The process $(L(t))$, describing the number of clients in the queue, is a Markov process on $\mathbb{N}$ with transition rates

$$q(x, x+1) = \lambda, \quad q(x, x-1) = \mu x.$$

If we fix a maximum number C of clients in the system and if any incoming customer is blocked when the queue is full, this is the Erlang model introduced in Section 2.

If we compute the fluid limit of this queue, it results in a degenerated process whose initial value is 1 and which is null for any $t > 0$. This fact can be explained simply: if we consider a very high initial number of clients in the queue such that $L(t) = N$, the departure rate of the queue is μN which is much higher than the arrival rate λ . In the case of the $M/M/1$ queue, departure rates and arrival rates are

of the same order whatever the initial state. In order to compensate the departure rate, it is natural to scale the arrival rate by a factor N . This leads to Kelly's scaling which has been introduced in [Kel86] for loss networks. We define a sequence of processes $(L_N(t))$ such that $L_N(0) = N$, whose transition rates are

$$q(x, x+1) = \lambda N, \quad q(x, x-1) = \mu x.$$

We then renormalize the process by N and we can prove that the sequence of processes $(L_N(t)/N)$ converges in distribution to a deterministic process $(x(t))$ such that

$$\dot{x}(t) = \lambda - \mu x(t).$$

One can note that there is no need to scale time here since the number of transitions per time unit is already of order N . A major difference with the $M/M/1$ is that there is no discontinuity. However, if we consider the Erlang model such that $(L_N(t))$ has the following transition rates:

$$q(x, x+1) = \lambda N \mathbb{1}_{\{x < C_N\}}, \quad q(x, x-1) = \mu x,$$

for some $C_N \in \mathbb{N}$, there is a discontinuity in C_N . Assuming that $C_N \approx \eta N$ and $L_N(0) = 0$, one can prove that the sequence of processes $(L_N(t)/N)$ converges in distribution to $(x(t))$ such that

$$x(t) = \min(\rho(1 - e^{-\mu t}), \eta).$$

As in the case of the $M/M/1$, one can prove that, while the boundary has not been reached, there is a diffusion of order $O(\sqrt{N})$ around the limit. If $\eta < \rho$, then the boundary is reached after a finite time and the variations are then of order $O(1)$. Kelly's scaling has been introduced in the context of loss networks but arises naturally when the transition rates depend linearly on the state of the stochastic process. Examples coming from population dynamics and chemistry can be found in [EK86]. A scaling which can be seen as a variation of the Kelly's scaling is introduced in Chapter II.

6.4.3. Different Time Scales. As explained above, there can be several time scales of interest of the same process. For instance, consider again the $M/M/1$ queue with arrival rate λ and departure rate μ . We have seen that the fluid limit with time scale $t \mapsto Nt$ allows to describe the first-order dynamics of the system. From this fluid limit, we deduce that the process is stable if $\lambda < \mu$ and it increases linearly to infinity when $\lambda > \mu$.

In the case where $\lambda = \mu$, the fluid limit of the $M/M/1$ queue is the constant equal to 1 and it does not give much insights in the behavior of the system. Furthermore, since μ is the limit of stability, one could be very interested in describing the behavior of the system when approaching this limit. This is the Heavy-traffic approximation which has been introduced by Kingman in [Kin62]. The principle is the following. We consider a sequence of $M/M/1$ queues $(L_N(t))$ with arrival rates $\lambda_N = \mu - \alpha/N$ and departure rate μ and such that $L_N(0) = N$. On the time scale $t \mapsto Nt$, the sequence of processes $(L_N(Nt)/N)$ is almost their fluid limit and admits the same limiting process. However, on the time scale $N \mapsto N^2t$, the process $(L_N(Nt)/N)$ converges in distribution to a reflected Brownian motion. The Heavy-traffic approximation has been widely studied in much more general contexts. For instance, for multi-class networks, one could refer to the paper of William [Wil98], for the $M/G/1$ PS queue, one could refer to the paper by Gromoll [Gro04] and the recent paper by Lambert et al [LSZ11].

Similarly, in Chapter IV, we study a stochastic network with failures where several scaling methods with different time scales are studied.

6.4.4. Coexistence of Different Time Scales: Stochastic Averaging Principle.

For some scaling methods, it can happen that there is a coexistence of different time scales. There are two components evolving with different speeds. We present here a very simple example inspired by Hunt and Kurtz [HK94] where this phenomenon arises. We consider the Erlang model with two classes of clients and N circuits. The system is described by the stochastic process $(L_{1,N}(t), L_{2,N}(t))$ with transition rates:

$$q(x, x + e_i) = \lambda_i N \mathbb{1}_{\{x_1 + x_2 < N\}}, \quad q(x, x - e_i) = \mu_i x_i,$$

for $i = 1, 2$ where $x = (x_1, x_2)$, $e_1 = (1, 0)$ and $e_2 = (0, 1)$. We also assume that $L_{i,N}(0)/N$ converges to $x_i(0)$ for $i = 1, 2$ such that $\lambda_1 + \lambda_2 > \mu_1 x_1(0) + \mu_2 x_2(0)$ and $x_1(0) + x_2(0) = 1$. We can prove that the sequence of stochastic processes $(L_{1,N}(t), L_{2,N}(t))$ converges to a process $(x_1(t), x_2(t))$ and we want to understand its dynamics.

As long as $\lambda_1 + \lambda_2 > \mu_1 x_1(t) + \mu_2 x_2(t)$, the total arrival rate is bigger than the total departure rate and the system is saturated. In order to understand the global dynamics, we have to look at the evolution of the number of empty circuits since any incoming client is accepted if and only if there is an empty slot. We then define $Z_{N,t}(s) = N - (L_{1,N}(t+s) + L_{2,N}(t+s))$. In fact, the transition rates of $(Z_{N,t}(s))$ are of order N and this process evolves on the time scale $s \mapsto Ns$. We can easily prove that the sequence $(Z_{N,t}(s)/N)$ converges in distribution to the process $(Z_t(s))$, an $M/M/1$ queue with arrival rate $\mu_1 x_1(t) + \mu_2 x_2(t)$ and departure rate $\lambda_1 + \lambda_2$. In particular, at equilibrium, the probability that there is an empty slot is given by $p(t) = (\mu_1 x_1(0) + \mu_2 x_2(0))/(\lambda_1 + \lambda_2)$.

Since $(Z(t))$ evolves on the time scale $t \mapsto Nt$ and the processes $(x_1(t), x_2(t))$ evolve on the time scale $t \mapsto t$, they “see” $(Z_t(s))$ at equilibrium. In particular, their respective arrival rates are $\lambda_1 p(t)$ and $\lambda_2 p(t)$. We deduce that $(x_1(t))$ and $(x_2(t))$ satisfy the following ordinary differential equations:

$$\begin{aligned} \dot{x}_1(t) &= \frac{\lambda_1}{\lambda_1 + \lambda_2} (\mu_1 x_1(t) + \mu_2 x_2(t)) - \mu_1 x_1(t), \\ \dot{x}_2(t) &= \frac{\lambda_2}{\lambda_1 + \lambda_2} (\mu_1 x_1(t) + \mu_2 x_2(t)) - \mu_2 x_1(t), \end{aligned}$$

The interaction between the “slow” process and the “fast” process is then clear. The transition rates of $(Z_t(s))$ are determined by states $x_1(t)$ and $x_2(t)$. On the contrary, the differential equations of $(x_1(t))$ and $(x_2(t))$ are determined by the equilibrium of the processes $(Z_t(s))$.

In fact, the previous equations are valid as long as $\mu_1 x_1(t) + \mu_2 x_2(t) < \lambda_1 + \lambda_2$. In particular, if $\lambda_1/\mu_1 + \lambda_2/\mu_2 < 1$, there exists a time $T > 0$ such that $\mu_1 x_1(T) + \mu_2 x_2(T) = \lambda_1 + \lambda_2$. For $t \geq T$, the processes $(x_1(t), x_2(t))$ satisfy the equations:

$$\begin{aligned} \dot{x}_1(t) &= \lambda_1 - \mu_1 x_1(t), \\ \dot{x}_2(t) &= \lambda_2 - \mu_2 x_1(t). \end{aligned}$$

In this case, there is no stochastic averaging anymore. This is illustrated by Figure 10. One can note that all these equations could have been obtained very easily with the Skorokhod problem described below but this method gives much less intuition.

This phenomenon of stochastic averaging has been known for a long time in a deterministic framework (see Guckenheimer and Holmes [GH90]). In statistical mechanics, it has been introduced by Bogolyubov in [Bog61]. Later, it has been studied in the context of stochastic calculus by Khasminskii in [Kha68]. His results have been extended by Papanicolaou et al. [PSV77] and Freidlin and Wentzell [FW98]. More recently, the phenomenon of stochastic averaging has been observed in the context of loss networks by Hunt and Kurtz [HK94]. In Chapter II, we exhibit a phenomenon of stochastic averaging in fluid limits. The proof is based on

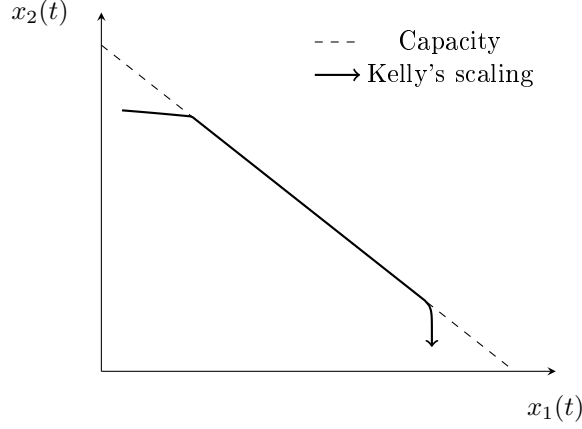


FIGURE 10. Stochastic averaging in the Kelly's scaling.

a martingale technique presented by Kurtz in [Kur92] but which is also present in the work of Papanicolaou et al. [PSV77]. Similarly, another stochastic averaging phenomenon is introduced in Chapter IV with a proof based on the same martingale technique.

6.4.5. Skorokhod Problem Representation. As observed in Section 6.4.1, for the $M/M/1$ queue, the difficulty of proving the convergence of the fluid limit is mainly due to the discontinuity at 0. Before that, the process can be simply expressed as the difference of two Poisson processes and the convergence follows easily. More generally, when the infinitesimal generator of the original process is sufficiently regular, the limit is the solution of an ordinary differential equation as it is the case for the $M/M/\infty$ queue. In the case of the fluid limit of an $M/M/1$ queue, the convergence can be proved directly but it is intricate and it uses some subtle properties of the $M/M/1$ process and this can hardly be used in the case of more complex processes with similar discontinuities. For a large class of problems, we can then use the Skorokhod problem.

We consider here the $M/M/1$ queue with arrival rate λ and departure rate μ . We can rewrite $(L(t))$ the process describing the number of customers in the following way:

$$\begin{aligned} L(t) &= L(0) + \mathcal{N}_\lambda([0, t]) - \int_0^t \mathbb{1}_{\{L(s-) > 0\}} \mathcal{N}_\mu(ds), \\ &= L(0) + \mathcal{N}_\lambda([0, t]) - \mathcal{N}_\mu([0, t]) + \int_0^t \mathbb{1}_{\{L(s-) = 0\}} \mathcal{N}_\mu(ds). \end{aligned}$$

Defining $Z(t) = L(0) + \mathcal{N}_\lambda([0, t]) - \mathcal{N}_\mu([0, t])$ and $R(t) = \int_0^t \mathbb{1}_{\{L(s-) = 0\}} \mathcal{N}_\mu(ds)$, we can see that $(L(t), R(t))$ is completely determined by $(Z(t))$ and the process $(R(t))$ is the pushing process which ensures that $(L(t))$ is non-negative. The couple $(L(t), R(t))$ is called the solution of the Skorokhod problem associated to $(Z(t))$ which is a biased random walk.

It is straightforward that, if $Z(0) = N$, the sequence of processes $(Z(Nt)/N)$ converges in distribution to $(1 + (\lambda - \mu)t)$. Using the continuity of the Skorokhod problem, this implies that the sequence $(X(Nt)/N, R(Nt)/N)$ converges to the solution of the Skorokhod problem associated to $(1 + (\lambda - \mu)t)$. This solution is simply $(t \mapsto \max(0, 1 + (\lambda - \mu)t), t \mapsto \max(0, -1 + (\mu - \lambda)t))$ and we obtain the fluid limit of the $M/M/1$ queue.

In fact, the Skorokhod problem can be generalized to the multi-dimensional case. For a short introduction, one can refer to the book of Robert [Rob03]. The previous demonstration has been generalized to Jackson networks by Chen and Mandelbaum in [CM91]. On the contrary, the demonstration presented at Section 6.4.1 would be hard to extend. There is a huge literature on Skorokhod problems which have been widely studied. For instance, it has been extended to the case where the constraints are a more general domain. For a review of those results, one can refer to the article of Ramanan [Ram06] and references herein.

As pointed out by Anderson and Orey in [AO76], the Skorokhod problem is not adapted to the case where the infinitesimal generator depends on the state of the stochastic process. The simplest example is the Erlang model. However, it is possible to adapt the Skorokhod method to a slightly more general context which applies to processes where the infinitesimal generator depends on the state of the stochastic process. These different techniques are used and detailed in Chapter IV.

7. Presentation of Subsequent Chapters

Each chapter of this thesis corresponds to a paper or a set of papers (one of them being under review):

Chapter II: Mathieu Feuillet. On the flow-level stability of data networks without congestion control: the case of linear networks and upstream trees. *Queueing Systems, Theory and Applications*, 70:105–143, 2012

Chapter III: T. Bonald and M. Feuillet. On the stability of flow-aware CSMA. *Perform. Eval.*, 67:1219–1229, November 2010

T. Bonald and M. Feuillet. On flow-aware CSMA in multi-channel wireless networks. In *CISS*, 2011

T. Bonald and M. Feuillet. Performance of CSMA in multi-channel wireless networks. *Queueing Systems: Theory & Applications*, 2012. To appear

Chapter IV: Mathieu Feuillet and Philippe Robert. A scaling analysis of a transient stochastic network (I). Preprint, 2012

Chapter V: Mathieu Feuillet and Philippe Robert. On the transient behavior of Ehrenfest and Engset processes. *Advances in Applied Probability*, 2012. To appear

7.1. Chapter II: a Bandwidth Sharing Network without Congestion Control. This chapter deals with a model of bandwidth sharing networks as introduced in Section 3.3. Contrary to the literature, users are not assumed to employ a congestion control algorithm but to transmit at maximum rate, their access rate, in the network. They recover from packet losses thanks to error codes. It is proved in [BFP09] that the bandwidth sharing is then characterized by the packet dropping policy in routers. Two policies have been studied in that paper: *Fair Dropping* and *Tail Dropping*. The first one consists in dropping packets in the biggest active flow while the second one consists in dropping packets at random. It has been proved in [BFP09] that Fair Dropping is stable under the optimal stability conditions. On the contrary, Tail Dropping results in an inefficient use of the bandwidth in many cases. Even worse, in the case of cyclic networks, all the available resources are wasted and the stability region reduces to the case where the traffic intensities are null. However, it has been conjectured that the stability region is not trivial for acyclic networks and that this stability region can be as close as wanted to the optimal stability region if the maximum rates of users are small enough compared to the capacity of the links in the network. The objective of Chapter II is to prove this conjecture for two representative examples of topologies: linear networks and upstream trees.

A linear network consists of L links and $L + 1$ classes such that flows of class 0 are going through all links (the long route) and flows of class k are going through link k only (the short routes). Flows of class k arrive according to a Poisson process of intensity λ_k and have an exponentially distributed size with parameter μ_k . The capacity of each link is assumed to be 1. For the sake of simplicity, the maximum rate of users in classes 0 and 1 is also equal to 1. The bandwidth allocation is then determined by the fact that output rates are proportional to input rates. The dynamics of the network is then represented by the process $(N_0(t), \dots, N_L(t))$ which is a Markov process with transition rates

$$\begin{cases} q(n, n + e_k) &= \lambda_k, \\ q(n, n - e_k) &= \mu_k \phi_k(n), \end{cases}$$

where e_k denotes the $L + 1$ -dimensional vector with every coordinate equal to 0 except the i th one equal to 1. The quantity $\phi_k(n)$ is the bandwidth allocation for class k when the state is n . For instance, when $L = 2$, the bandwidth allocation is given by:

$$\begin{aligned} \phi_0(n) &= \min \left(\frac{n_0}{n_0 + n_1}, \frac{n_0/(n_0 + n_1)}{n_0/(n_0 + n_1) + n_2 a_2} \right), \\ \phi_1(n) &= \frac{n_1}{n_0 + n_1}, \\ \phi_2(n) &= \min \left(n_2 a_2, \frac{n_2 a_2}{n_0/(n_0 + n_1) + n_2 a_2} \right), \end{aligned}$$

where a_2 is the access rate of flows of class 2.

In Chapter II, we first characterize the fluid limit of the Markov process $(N_0(t), \dots, N_L(t))$. If the initial state is inside the orthant, i.e. $N_k(0)/m$ tends to $Z_k(0) > 0$ for all k when m tends to infinity, then it is easy to see that the fluid limit is very classical. All the short classes except the first one use all the bandwidth of their link and the fluid limits $(Z(t))$ satisfy

$$\begin{aligned} \dot{Z}_0(t) &= \lambda_0, \\ \dot{Z}_1(t) &= \lambda_1 - \mu_1 \frac{Z_1(t)}{Z_0(t) + Z_1(t)}, \\ \dot{Z}_k(t) &= (\lambda_k - \mu_k) \mathbb{1}_{\{Z_k(t) > 0\}}, \quad \text{for } 2 \leq k \leq L \end{aligned}$$

as long as there is $2 \leq k \leq L$ such that $Z_k(t) > 0$. If $\lambda_k < \mu_k$, then all the short routes are null at some point, where there is a stochastic averaging phenomenon arising as described in Section 6.4.4. The “slow” process is the fluid limit of class 0 and the fast process is the stochastic limit of $(N_2(t), \dots, N_L(t))$ which reaches some local equilibrium depending on the ratio $Z_0(t)/(Z_0(t) + Z_1(t))$. On the contrary, the bandwidth allocation received by $(Z_0(t))$ is averaged with respect to the stationary distribution of the limit of $(N_2(t), \dots, N_L(t))$ and the fluid limit then satisfies the equation

$$\begin{aligned} \dot{Z}_0(t) &= \lambda_0 - \mu_0 \bar{\phi}_0 \left(\frac{Z_1(t)}{Z_0(t) + Z_1(t)} \right), \\ \dot{Z}_1(t) &= \lambda_1 - \mu_1 \frac{Z_1(t)}{Z_0(t) + Z_1(t)}, \\ \dot{Z}_k(t) &= 0, \quad \text{for } 2 \leq k \leq L \end{aligned}$$

where $\bar{\phi}_0$ is the averaged bandwidth allocation. Thanks to this allocation, we are able to derive non-trivial upper and lower bounds on the stability region of the process $(N_0(t), \dots, N_L(t))$.

The second part of the chapter is dedicated to evaluating the impact of the access rates on the stability region. For that purpose, we consider only the short routes $2 \leq k \leq L$ and we keep frozen the ratio $\alpha = n_0/(n_0 + n_1)$. We are then able to describe the dynamics of the short routes with a Markov process $(\tilde{N}_2^\alpha(t), \dots, \tilde{N}_L^\alpha(t))$. We then introduce a scaling on the access rates and the flow size. Basically, we consider the same process as previously but, for each class k , we replace the access rates by a_k/β , the arrival rate by $\lambda_k\beta$ and the departure rate $\mu_k\beta$. In that case, we can prove that the process $(\tilde{N}_2^\alpha(t)/\beta, \dots, \tilde{N}_L^\alpha(t)/\beta)$ converges to the solution of an ordinary differential equation. In particular, if $\lambda_k < \mu_k$ for $2 \leq k \leq L$ then this equation admits a fixed point which satisfies:

$$\gamma_k(\alpha) = \max \left(\rho_k, \frac{\rho_k}{1 - \rho_k} \min \left(\alpha, \min_{2 \leq j \leq k-1} (1 - \rho_j) \right) \right) \text{ for } 2 \leq k \leq L.$$

Since the stationary distributions of the processes $(\tilde{N}_2^\alpha(t), \dots, \tilde{N}_L^\alpha(t))$ can be controlled with some stochastic domination, we are able to prove that they converge in distribution to $(\delta_{\gamma_2}, \dots, \delta_{\gamma_L})$. This, in turn, implies that the stability region of the initial process $(N_0(t), \dots, N_L(t))$ converges to the optimal stability region when the access rates decrease to 0.

An upstream tree is a network where flows enter the network into a leaf then follow a path to the trunk where they leave the network. Those networks are very specific since the resulting stochastic processes describing the number of flows in each class are monotonic in the sense of the paper by Borst et al. [BJL08]. Using the same scaling on the access rates and a stochastic domination argument, we are able to prove recursively that the stability region of upstream trees tends to the optimal stability region when the access rates decrease to 0.

The two cases that are treated in this chapter tend to confirm that the conjecture is true. However, both proofs are ad-hoc and rely heavily on properties inherent to the topology and we have not been able to generalize them. For that reason, the conjecture is still open. From a technical point of view, we have been able to exhibit a non-trivial averaging phenomenon and an interesting application of scaling methods.

7.2. Chapter III: a Flow-Aware Access-Control Mechanism. This chapter deals with a model of wireless network as the one introduced in Section 4. Here, flow dynamics is taken into account and the objective is to design an access-control algorithm being optimal *at flow-level*.

The chapter deals with wireless ad-hoc networks in a multi-channel environment but, for the sake of simplicity, we present here the case of a single channel. In the model we introduced, the network consists of a random, dynamic set of wireless links in ad-hoc mode sharing a common radio channel. Each link consists of a transmitter-receiver pair; the transmitter is able to use at most one radio channel at a time. Links are grouped into a finite number of K classes sharing similar characteristics. Two links within the same class cannot be simultaneously active. The channel is associated with some conflict graph as the one introduced in Section 4. Links arrive in class k according to a Poisson process of parameter λ_k and have a flow with a geometric number of packets with mean $\sigma_k N$ to transmit where N is a positive integer. We call $X_k^N(t)$ the number of links in class k at time t .

In order to completely describe the network, we need to take into account the packet-level dynamics. Packets are assumed to have an exponential size with mean $1/N$ bits, which implies that the class- k mean flow size is constant and equal to σ_k bits. Links are assumed to use the standard CSMA algorithm where each transmitter waits for a period of random duration, the back-off time, before each transmission attempt. If the radio channel is sensed idle (in the sense that no

conflicting link is active), a packet is transmitted; otherwise, the transmitter waits for a new back-off time before the next attempt. The back-off times of class- k transmitters are random with mean $1/(N\nu_k)$. We denote by $Y^N(t)$ the activity of class k at time t ; $Y^N(t)$ is equal to 1 if there is a link of class k active at time t and 0 otherwise. The state of the network is then completely described by $(X^N(t), Y^N(t))$ which is a Markov process.

In order to evaluate the efficiency of standard CSMA in this ad-hoc context, we want to determine its stability region, i.e. the set of traffic intensities $(\rho_1, \dots, \rho_K)$ with $\rho_k = \lambda_k \sigma_k$ such that the process $(X^N(t), Y^N(t))$ is ergodic. However, the process $(X^N(t), Y^N(t))$ is very difficult to analyze and, apart from some very specific classes of interference graphs, we have not been able to derive its stability region. The idea is then to use a scaling to approximate $(X^N(t), Y^N(t))$ by a simpler system. For that purpose, we scale the parameter N to infinity and we prove that a stochastic averaging occurs (also called time-scale separation in this context). Indeed, the process $(Y^N(t))$ typically lives on the time scale $t \mapsto Nt$ and the process $(X^N(t))$ on the time scale $t \mapsto t$. The limiting process of $(Y^N(t))$ is then a reversible process whose stationary distribution admits a closed-form expression:

$$\pi_x(S) \propto \prod_{k \in S} \frac{x_k}{\nu_k},$$

where S is a feasible schedule; there is no conflict between two transmitting links. This stationary distribution allows to express explicitly the transition rates of the limiting process $(X(t))$ of the sequence $(X^N(t))$:

$$\begin{cases} q(x, x + e_k) &= \lambda_k, \\ q(x, x - e_k) &= \frac{1}{\sigma_k} \sum_{S: k \in S} \pi_x(S). \end{cases}$$

In fact, the stationary distribution π_x approximates the Max-Weight Scheduling algorithm where the weights are logarithmic in the sense that it chooses a schedule of maximal weight with high probability. Thanks to this and the Foster criterion, we are then able to prove that $(X(t))$ is stable if the traffic intensities lie in the optimal stability region. Here scaling methods allow to simplify a difficult system into a tractable one but, contrary to Chapter II, we are not able to obtain information on the original process even for large N . In particular, the stability of the process $(X^N(t), Y^N(t))$ remains an open problem.

In the second part of this chapter, we consider wireless networks with an infrastructure. The result of optimality obtained for ad-hoc networks is not true anymore and there is an asymmetry between uplink and downlink. Indeed, each user connected to the base station uses a back-off mechanism as the base station and this naturally favors the uplink while, in most cases, the downlink is more critical. The suboptimality of standard CSMA is proved on two examples using fluid limits. Again, a stochastic averaging phenomenon arises in the study of the fluid limits. In the second part, we then propose a small modification of CSMA to ensure the optimality of this algorithm: it consists in using a back-off mechanism for each flow. This also solves the problem of the asymmetry between the uplink and the downlink.

7.3. Chapter IV: a Large-Scale Stochastic Network with Failures.

We consider a large-scale storage system as presented in Section 5. In order to ensure persistence, files are duplicated on several servers. When the disk of a given server breaks down, its files are lost but can be retrieved on other servers if copies are available. In such an architecture, a fraction of the bandwidth of a server is devoted to the duplication mechanism of its files to other servers. On the one hand, there should be sufficiently many copies so that any file has a copy available on at least

one server at any time. On the other hand, in order to use the bandwidth in an optimal way, there should not be too many copies of a given file so that the network can accommodate a large number of distinct files. The purpose of this chapter is to investigate the evolution of the system when the number of files tends to infinity.

More precisely, we consider the following simple model. A file can have at most two copies, the total bandwidth allocated to file duplication is given by λN , for $\lambda > 0$ and N some positive integer. If at some moment there are $x \geq 1$ files with exactly one copy, a new copy of each of these files is created at rate $\lambda N/x$. It is assumed that initially F_N files are present in the system with two copies and each copy of a file disappears at rate μ . A file with 0 copies is lost. The system can then simply be described by the Markov process $(X_0^N(t), X_1^N(t))$ with transition rates

$$\begin{cases} q(x, x + e_0 - e_1) &= \mu x_1, \\ q(x, x + e_1) &= 2\mu(F_N - x_0 - x_1), \\ q(x, x - e_1) &= \lambda N \mathbb{1}_{\{x_1 > 0\}}. \end{cases}$$

In particular, one can note that the process admits $(F_N, 0)$ as an absorbing state. Furthermore, it is assumed that F_N is proportional to N such that F_N/N converges to $\beta > 0$ when N tends to infinity. For $\delta > 0$, there exists some finite random instant $T_N(\delta)$, such that a fraction $\lfloor \delta N \rfloor$ of the files are lost after time $T_N(\delta)$. This chapter investigates the order of magnitude in N of the variables $T_N(\delta)$ as N gets large and the role of the parameters λ , μ and β in these asymptotics.

For that purpose, we look at the system through different time scales in order to understand how it works. First, we slow down the time and look at the time scale $t \mapsto t/N$. It is easy to see that $(X_1^N(t/N))$ converges then to an $M/M/1$ queue with arrival rate $2\mu\beta$ and departure rate λ and $(X_0^N(t/N))$ converges to the constant equal to 0. It is clearly not the right time scale to see losses but it gives a first indication since the limiting process of $(X_1^N(t/N))$ is transient if $2\mu\beta > \lambda$ and ergodic if $2\mu\beta < \lambda$.

It is then natural to look at the normal time scale $t \mapsto t$. In this case, as suggested by the time scale $t \mapsto t/N$, there are three possible regimes. If $2\mu\beta > \lambda$ then the back-up mechanism is not sufficient to prevent losses and there is a significant fraction of files that are lost from the beginning. Indeed, the process $(X_0^N(t)/N, X_1^N(t)/N)$ converges to $(x_0(t), x_1(t))$ the solution of the ordinary differential equation:

$$\begin{cases} \dot{x}_0(t) &= -\mu x_1(t), \\ \dot{x}_1(t) &= -\lambda - \mu x_1(t) + 2\mu(\beta - x_0(t) - x_1(t)). \end{cases}$$

However, when t tends to infinity, the system $(x_0(t), x_1(t))$ tends to $(\beta - \lambda/(2\mu), 0)$, that is to say the system tends to the maximum sustainable load. If $2\mu\beta = \lambda$, the process $(X_0^N(t), X_1^N(t))$ is no more of order $O(N)$ but $O(\sqrt{N})$ and we can prove that $(X_0(t)/\sqrt{N}, X_1(t)/\sqrt{N})$ converges to a diffusion, solution of an unusual stochastic differential equation reflected at 0. Those two results can be proved using the Skorokhod problem presented in Section 6.4.5.

Finally, when $2\mu\beta < \lambda$, the number of files is sustainable and there are only finite losses. For all t , the random variable $X_1^N(t)$ converges in distribution to a geometrically distributed random variable with parameter $2\mu\beta/\lambda$. One can recall that at time scale $t \mapsto t/N$, $(X_1^N(t/N))$ converges to an $M/M/1$ queue and the limit of $X_1^N(t)$ can then be interpreted as the stationary distribution of an ergodic $M/M/1$ queue. The process $(X_0^N(t))$ converges in distribution to a Poisson process of parameter $\mu 2\mu\beta/(\lambda - 2\mu\beta)$. One can note that the intensity of this Poisson process is just the mean of the limiting geometric distribution of $X_1^N(t)$ multiplied

by the loss rate. This anticipates the stochastic averaging phenomenon happening at time scale $t \mapsto Nt$.

The time scale $t \mapsto t$ allows us to exhibit the capacity of the system but it does not say anything about the decay rate of the network which is given by the asymptotics of the random variables $T_N(\delta)$. That is why we accelerate the time and we consider the time scale $t \mapsto Nt$. The only interesting case here is when $2\mu\beta < \lambda$. Since the process $X_1^N(t/N)$ behaves like an $M/M/1$ queue when N tends to infinity, we can prove that $X_1^N(Nt)$ remains bounded on each finite interval and the number of files with 1 copy is negligible compared to F_N . Consequently, if we call $\Psi(t)$ the fraction of lost files “at time Nt ” when N tends to infinity, the fraction of files with two copies is $\beta - \Psi(t)$. As for the time scale $t \mapsto t$, the number of files with 1 copy behaves as an $M/M/1$ queue with arrival rate $2\mu(\beta - \Psi(t))$ and departure rate λ at equilibrium and the loss rate is then $\mu 2\mu(\beta - \Psi(t))/(\lambda - 2\mu(\beta - \Psi(t)))$ and we obtain the following equation for $(\Psi(t))$:

$$\Psi(t) = \mu \int_0^t \frac{2\mu(\beta - \Psi(s))}{\lambda - 2\mu(\beta - \Psi(s))} ds.$$

This is another example of the stochastic averaging phenomenon as described in Section 6.4.4, the fast process being $(X_1^N(t/N))$ and the slow process $(X_0^N(tN)/N)$. One can note that the factor between the two time scales is N^2 contrary to other examples developed in this thesis where it is typically N .

Finally, we are able to derive asymptotics on $T_N(\delta)$ with the previous equation and we obtain

$$\lim_{N \rightarrow +\infty} \frac{T_N(\delta)}{N} = -\frac{\rho}{2} \log(1 - \delta) - \delta\beta.$$

7.4. Chapter V: the Transient Behavior of the Engset Model. Contrary to Chapter IV where the storage system is modeled globally, this chapter focuses on the durability of a specific file. We assume that a file is stored in a set of N servers but can have at most C_N copies. As previously, servers can break down and each copy is lost at rate μ . There is a back-up mechanism downloading the file in the empty servers. We assume that each server can download a copy at rate ν . We denote by $(X_N(t))$ the process describing the number of copies of the file, it admits the following transition rates

$$\begin{cases} q(x, x+1) &= \nu(N-x)\mathbb{1}_{\{x < C_N\}}, \\ q(x, x-1) &= \mu x. \end{cases}$$

This corresponds to the classical Engset model which has been introduced in Section 2. The purpose of this chapter is to estimate the durability of the file and so we would like to evaluate T_0 , the hitting time of 0 when there are initially C_N copies of the file. In particular, we want to derive asymptotics on T_0 when N tends to infinity and C_N/N converges to $\eta > 0$.

This chapter relies on martingale techniques introduced in Section 6.3. For that purpose, we start by studying the classical Ehrenfest process $(E_N(t))$ whose transition rates are $q(x, x+1) = \nu(N-x)$ and $q(x, x-1) = \mu x$. Using classical techniques, we are then able to prove that the process

$$(M(t)) = \left(\left(1 - \beta\mu e^{(\mu+\nu)t}\right)^{E_N(t)} \left(1 + \beta\nu e^{(\mu+\nu)t}\right)^{N-E_N(t)} \right)$$

is a martingale for any $\beta \in \mathbb{R}$. With some calculation and the stopping time theorem, we are then able to derivate the Laplace transform of T_{C_N} the hitting

time of C_N when starting from 0 for the process $(E_N(t))$:

$$\mathbb{E}_0(e^{-\alpha T_{C_N}}) = \frac{\int_0^1 (1-u)^N u^{\alpha-1} du}{\int_0^1 (1-u)^{C_N} \left(1 + \frac{\nu}{\mu} u\right)^{N-C_N} u^{\alpha-1} du}.$$

Since conditioned on t being smaller than T_{C_N} , both processes $(E_N(t))$ and $(X_N(t))$ have the same distribution, this result is also valid for $(X_N(t))$.

On the contrary, the processes $(E_N(t))$ and $(X_N(t))$ do not have the same dynamics when we consider the hitting time T_0 because of the boundary effect at state C_N . For that reason, we have to build a family of non-negative martingales for the Engset process. The idea is the same as the one used for the $M/M/1$ queue by Kennedy in [Ken76] and described in Section 6.3. We construct a martingale as a linear combination of two martingales and use the fact that the infinitesimal generators of both processes are identical everywhere except in C_N . With this new martingale, the Laplace transform of T_0 can be obtained.

With those two Laplace transforms, we are then able to compute the asymptotics of these hitting times with respect to the parameters ν , μ and η . The basic ingredients of those estimates are the Laplace method and elementary (but careful) analysis.

We are then able to answer the initial question. Notably, if $\eta > \nu$, the random variable $N(1-\nu)^N T_0$ converges in distribution to an exponential random variable of parameter ν . If $\eta = \nu$, the random variable $N(1-\nu)^N T_0$ converges in distribution to an exponential random variable of parameter 2ν . On the contrary, if $\eta < \nu$, the scaling factor is strictly smaller than $N(1-\nu)^N$ and depends on η . So the optimum is reached as soon as $\eta = \nu + \varepsilon$. In practice, if the maximum number of copies is increased, the system is not able to increase the number of copies which stays around νN .

References for the introduction

- [Alo86] Noga Alon. A fast and simple randomized parallel algorithm for the maximal independent set problem. *Journal of Algorithms*, 7:567–583, 1986.
- [Amb70] Norman Abramson. The aloha system. In *FJCC AFIPS*, volume 37, pages 281–285, 1970.
- [AO76] Robert F. Anderson and Steven Orey. Small random perturbation of dynamical systems with reflecting boundary. *Nagoya Math. J.*, 60:189–216, 1976.
- [BA03] T. Bonald and Proutière A. Insensitive bandwidth sharing in data networks. *Queueing Systems: Theory and Applications*, 44(1):69–100, 2003.
- [Bar62] Paul Baran. On distributed communications networks. Technical Report P-2626, Rand Corporation, 1962.
- [BBHP04] T. Bonald, S.C. Borst, N. Hegde, and A. Proutière. Wireless data performance in multi-cell scenarios. In *Proceedings of ACM Sigmetrics*, pages 378–380, 2004.
- [BBLP11] N. Bouman, S. Borst, J. Van Leeuwen, and A. Proutière. Backlog-based random access in wireless networks: fluid limits and delay issues. In *ITC 23*, 2011.

- [BBP⁺01] S. Ben Fredj, T. Bonald, A. Proutière, G. Régnié, and J. W. Roberts. Statistical bandwidth sharing: a study of congestion at flow level. In *Proceedings of ACM SIGCOMM*, pages 111–122, 2001.
- [BF10] T. Bonald and M. Feuillet. On the stability of flow-aware CSMA. *Perform. Eval.*, 67:1219–1229, November 2010.
- [BF11] T. Bonald and M. Feuillet. On flow-aware CSMA in multi-channel wireless networks. In *CISS*, 2011.
- [BF12] T. Bonald and M. Feuillet. Performance of CSMA in multi-channel wireless networks. *Queueing Systems: Theory & Applications*, 2012. To appear.
- [BFP09] T. Bonald, M. Feuillet, and A. Proutière. Is the “Law of the Jungle” sustainable for the Internet? In *Proceedings of INFOCOM*, pages 28–36, april 2009.
- [BJL08] Sem Borst, Matthieu Jonckheere, and Lasse Leskelä. Stability of parallel queueing systems with coupled service rates. *Discrete Event Dynamic Systems*, 18(4):447–472, dec 2008.
- [BM01] T. Bonald and L. Massoulié. Impact of fairness on Internet performance. In *Proceedings of ACM SIGMETRICS/Performance*, pages 82–91, 2001.
- [BMP08] Charles Bordenave, David McDonald, and Alexandre Proutière. Asymptotic stability region of slotted-Aloha. Preprint, 2008.
- [BMPV06] T. Bonald, L. Massoulié, A. Proutière, and J. Virtamo. A queueing analysis of max-min fairness, proportional fairness, and balanced fairness. *Queueing Systems: Theory and Applications*, 53:65–84, 2006.
- [Bog61] N. N. Bogoliubov. The quasi-averages in problems of statistical mechanics (in russian). Technical report, Joint Institute for Nuclear Research, 1961. Preprint D-781.
- [Bon07] T. Bonald. Insensitive traffic models for communication networks. *Discrete Event Dynamic Systems*, 17(3):405–421, 2007.
- [Bra08] Maury Bramson. *Stability of Queueing Networks*, volume 1950 of *Lecture Notes in Mathematics*. Springer, 2008.
- [Bré01] Pierre Brémaud. *Markov chains, Gibbs fields, Monte Carlo simulation, and queues*. Texts in Applied Mathematics. Springer-Verlag, 2001.
- [Bur81] David Y. Burman. Insensitivity in queueing systems. *Advances in Applied Probability*, 13(4):846–859, 1981.
- [CK74] Vinton G. Cerf and Robert E. Kahn. A protocol for packet network intercommunication. *IEEE Transactions on Communications*, 22(5):637–648, 1974.
- [CKLS08] P. Chaporkar, K. Kar, Xiang Luo, and S. Sarkar. Throughput and fairness guarantees through maximal scheduling in wireless networks. *IEEE Transactions on Information Theory*, 54(2):572–594, feb 2008.
- [CM91] H. Chen and A. Mandelbaum. Discrete flow networks: bottleneck analysis and fluid approximations. *Mathematics of Operation Research*, 16(2):408–446, 1991.
- [Dai95] J. G. Dai. On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Annals of Applied Probability*, 5:49–77, 1995.

- [dVKL01] Gustavo de Veciana, Takis Konstantopoulos, and Tae-Jin Lee. Stability and performance analysis of networks supporting elastic services. *IEEE/ACM Trans. Netw.*, 9(1):2–14, 2001.
- [DW06] A. Dimakis and J. Walrand. Sufficient conditions for stability of longest-queue-first scheduling: second-order properties using fluid limits. *Advances in Applied Probability*, 38(2):505–521, 2006.
- [EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and convergence*. Wiley, 1986.
- [Eng98] T.O. Engset. On the calculation of switches in an automatic telephone system. *Teletronikk*, 94(2):99–142, 1998. Translation of a German version (1918).
- [Erl09] A.K. Erlang. The theory of probabilities and telephone conversations. *Nyt Tidsskrift for Matematik B*, 20:33, 1909.
- [Erl17] A.K. Erlang. Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges. *Elektroteknikerer*, 13:5, 1917.
- [Feu12] Mathieu Feuillet. On the flow-level stability of data networks without congestion control: the case of linear networks and upstream trees. *Queueing Systems, Theory and Applications*, 70:105–143, 2012.
- [FR12a] Mathieu Feuillet and Philippe Robert. On the transient behavior of Ehrenfest and Engset processes. *Advances in Applied Probability*, 2012. To Appear.
- [FR12b] Mathieu Feuillet and Philippe Robert. A scaling analysis of a transient stochastic network (I). Preprint, 2012.
- [FW98] M. Freidlin and A. Wentzell. *Random Perturbations of Dynamical Systems*. Springer, 1998. 2nd Edition.
- [GH90] John Guckenheimer and Philip Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, volume 42 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1990. Revised and corrected reprint of the 1983 original.
- [Gim65] L. Gimpelson. Analysis of mixtures of wide and narrow-band traffic. *IEEE Transactions on Communication Technology*, 13(3):258–266, 1965.
- [GJ90] Michael R. Garey and David S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., 1990.
- [Gro04] H. Christian Gromoll. Diffusion approximation for a processor sharing queue in heavy traffic. *Annals of Applied Probability*, 14(2):555–611, 2004.
- [GYJ⁺10] Matthieu Gallet, Nezih Yigitbasi, Bahman Javadi, Derrick Kondo, Alexandru Iosup, and Dick Epema. A model for space-correlated failures in large-scale distributed systems. In *EuroPar’10*, pages 88–100, 2010.
- [HK94] P. J. Hunt and T. G. Kurtz. Large loss networks. *Stochastic Processes and their Applications*, 53(2):363–378, 1994.
- [Jac63] James R. Jackson. Jobshop-like queueing systems. *Management Science*, 10(1):131–142, 1963.

- [Jac88] V. Jacobson. Congestion avoidance and control. In *SIGCOMM '88: Symposium proceedings on Communications architectures and protocols*, pages 314–329, New York, NY, USA, 1988. ACM.
- [JW08] Libin Jiang and Jean Walrand. A distributed CSMA algorithm for throughput and utility maximization in wireless networks. In *Allerton Conference on Communication, Control, and Computing*, 2008.
- [Kau81] Joseph S. Kaufman. Blocking in a shared resource environment. *IEEE Transactions on Communications*, 29(10):1474–1481, 1981.
- [Kel79] F.P. Kelly. *Reversibility and Stochastic Networks*. Wiley, Chichester, 1979.
- [Kel86] F.P. Kelly. Blocking probabilities in large circuit-switched networks. *Advances in Applied Probability*, 18:473–505, 1986.
- [Kel91] Frank P. Kelly. Loss networks. *Annals of Applied Probability*, 1(3):319–378, 1991.
- [Kel99] Frank Kelly. Mathematical modelling of the Internet. In *Proceedings of the Fourth International Congress on Industrial & Applied Mathematics*, 1999.
- [Ken53] David G. Kendall. Stochastic processes occurring in the theory of queues and their analysis by the method of the embedded markov chain. *Annals of Mathematical Statistics*, 24(3):338–354, 1953.
- [Ken76] D.P. Kennedy. Some martingales related to cumulative sum tests and single-server queues. *Stochastic Processes and their Applications*, 4:261–269, 1976.
- [Kha68] R. Khasminskii. On averaging principle for itô stochastic differential equations. *Kybernetika Chekhoslovakia*, 4(3):260–279, 1968. In Russian.
- [Kin62] J. F. C. Kingman. On queues in heavy traffic. *Journal of the Royal Statistical Society, Series B*, 24:383–392, 1962.
- [KM57] Samuel Karlin and James McGregor. The classification of birth and death processes. *Transactions of the AMS*, 86:366–400, 1957.
- [Kur92] T.G. Kurtz. Averaging for martingale problems and stochastic approximation. In *Applied Stochastic Analysis, US-French Workshop*, volume 177 of *Lecture notes in Control and Information sciences*, pages 186–209. Springer Verlag, 1992.
- [LK91] S.H. Lu and P.R. Kumar. Distributed scheduling based on due dates and buffer priorities. *IEEE Transactions on Automatic Control*, 36(12):1406–1416, 1991.
- [Loy62] R.M. Loynes. The stability of queues with non independent inter-arrival and service times. *Proceedings of Cambridge Philosophical Society*, 58:497–520, 1962.
- [LSZ11] Amaury Lambert, Florian Simatos, and Bert Zwart. Scaling limits via excursion theory: Interplay between crump-mode-jagers branching processes and processor-sharing queues. Submitted, 2011.
- [LTWW93] Will E. Leland, Murrad S. Taqqu, Walter Willinger, and Daniel V. Wilson. On the self-similar nature of ethernet traffic. In *Proceedings of Sigcomm*, pages 183–193, 1993.
- [Mal93] V. A. Malyshev. Networks and dynamical systems. *Advances in Applied Probability*, 25(1):140–175, 1993.

- [Mas07] Laurent Massoulié. Structural properties of proportional fairness: stability and insensitivity. *The Annals of Applied Probability*, 17(3):809–839, 2007.
- [Mey95] Sean Meyn. Transience of multiclass queueing networks via fluid limit models. *Annals of Applied Probability*, 5:946–957, 1995.
- [MR00] L. Massoulié and J. W. Roberts. Bandwidth sharing and admission control for elastic traffic. *Telecommunications Systems*, 15:185–201, 2000.
- [MVW93] N. McKeown, P. Varaiya, and J. Walrand. Scheduling cells in an input-queued switch. *Electronics Letters*, 29(25):2174–2175, 1993.
- [MW00] Jeonghoon Mo and Jean Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Trans. Netw.*, 8(5):556–567, 2000.
- [PBS07a] Fabio Picconi, Bruno Baynat, and Pierre Sens. An analytical estimation of durability in DHTs. In Tomasz Janowski and Hrushikesh Mohanty, editors, *Distributed Computing and Internet Technology*, volume 4882 of *Lecture Notes in Computer Science*, pages 184–196. Springer, 2007.
- [PBS07b] Fabio Picconi, Bruno Baynat, and Pierre Sens. Predicting durability in DHTs using Markov chains. In *Proceedings of the 2nd Intl. Conference on Digital Information Management (ICDIM)*, 2007.
- [PF95] Vern Paxson and Sally Floyd. Wide area traffic: The failure of poisson modeling. *IEEE/ACM Transactions on Networking*, 3(3):226–244, 1995.
- [PGK88] David A. Patterson, Garth Gibson, and Randy H. Katz. A case for redundant arrays of inexpensive disks (raid). In *SIGMOD*, pages 109–116, 1988.
- [PSV77] G. Papanicolaou, D. Stroock, and S.R.S Varadhan. Martingale approach to some limit theorems. *Statistical Mechanics, Dynamical Systems and the Duke Turbulence Conference*, 3, 1977.
- [Ram06] Kavita Ramanan. Reflected diffusions defined via the extended Skorokhod map. *Electronic Journal of Probability*, 11:934–992, 2006.
- [Rob81] James W. Roberts. A service system with heterogeneous user requirement. In Guy Pujolle, editor, *Performance of Data Communications Systems and Their Application*, pages 423–431, 1981.
- [Rob03] Philippe Robert. *Stochastic Networks and Queues*. Stochastic Modeling and Applied Probability Series. Springer-Verlag, New York, 2003. xvii+398 pp.
- [RS92] A. N. Rybko and A. L. Stolyar. Ergodicity of stochastic processes describing the operation of open queueing networks. *Problems of Information Transmission*, 28(3):3–26, 1992.
- [RZ10] J. E. Reed and B. Zwart. Limit theorems for bandwidth sharing networks with rate constraints. Preprint, 2010.
- [Sen96] E Seneta. Markov and the birth of chain dependence theory. *International Statistical Review*, 64(3):255–263, 1996.
- [Ser99] R. Serfozo. *Introduction to Stochastic Networks*. Springer-Verlag New York, 1999.
- [SS11] D. Shah and J. Shin. Randomized scheduling algorithm for queueing networks. *The Annals of Applied Probability*, page to appear, 2011.
- [Szp93] Wojciech Szpankowski. Stability conditions for some distributed systems: buffered random access systems. *Advances in Applied Probability*, 26:498–515, 1993.

- [Tak69] Lajos Takács. On Erlang's formula. *The Annals of Mathematical Statistics*, 40(1):71–78, 1969.
- [TE92] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, 37:1936–1948, 1992.
- [TSM98] François Theberge, Alain Simonian, and RaviR. Mazumdar. Upper bounds for blocking probabilities in large multi-rate loss networks. *Telecommunication Systems*, 9:23–39, 1998.
- [vdVBS09] P. van de Ven, S. Borst, and S. Shneer. Instability of MaxWeight Scheduling Algorithms. In *Proceedings of INFOCOM*, pages 1701–1709, april 2009.
- [Whi85] Ward Whitt. Blocking when service is required from several facilities simultaneously. *AT&T Technical Journal*, 64(8):1807–1856, 1985.
- [Wil98] R. J. Williams. Diffusion approximations for open multiclass queueing networks: sufficient conditions involving state space collapse. *Queueing Systems: Theory and Applications*, 30:27–88, 1998.
- [XMS⁺03] Qin Xin, Ethan L. Miller, Thomas Schwarz, Darrell D. E. Long, Scott A. Brandt, and Witold Litwin. Reliability mechanisms for very large storage systems. In *Proceedings of the 20th IEEE / 11th NASA Goddard Conference on Mass Storage Systems and Technologies*, pages 146–156, 2003.

CHAPTER II

On the Flow-Level Stability of Data Networks without Congestion Control

Contents

1. Introduction	35
2. Flow-Level Model	37
3. Stability Conditions for a Linear Network	39
4. Asymptotic Stability	47
Appendix: Proof of Proposition 3.5	60
References for Chapter II	66

1. Introduction

Most Internet traffic derives from the transfer of stored documents, corresponding to texts, music and movies. This traffic is elastic in the sense that the transmission rate varies depending on current network congestion. Each document transfer is a flow of packets whose rate is usually regulated by the Transmission Control Protocol (TCP). The rate is allowed to increase linearly until a packet loss is observed. The rate is then divided by two. These two mechanisms ensure that network bandwidth is shared approximately fairly between the varying number of flows in progress at any time. TCP also ensures that lost packets are retransmitted.

There is, however, no constraint on users to actually implement TCP algorithm and a more aggressive, greedy congestion control would be individually more profitable. Indeed, new, more aggressive versions of TCP are emerging to better exploit increasing link capacity [Flo03, JWL04, XHR04]. Some have suggested that the packet retransmission function of TCP could be replaced by forward error correction in the form of source coding [RS05]. It consists in the use of erasure codes in order to encode files before sending them into the network. As a consequence, if the file has been divided in K packets, the destination is able to rebuild the data with $(1+\varepsilon)K$ packets. A typical class of erasure codes which is adapted to source coding are the *Digital fountains* [Mit04]. The use of source coding enables the relaxation of control requirements and facilitates aggressive packet transmission policies.

If the congestion avoidance of TCP is no longer universally used, there is a danger that the Internet may again experience the congestion collapse observed in its early days [Nag84]. Packets dropped downstream and therefore retransmitted needlessly encumber upstream links, amplifying and spreading congestion. The consequence may be that the rate realized by concurrent flows is much less than the optimum (in terms of utility maximization) and may even tend to zero. The capacity of a network can be defined as the demand, flow arrival rate $\times$ mean flow size, that can be sustained. It depends on the way bandwidth is shared. It is important to understand what happens to capacity when the assumption that

users compliantly implement TCP is relaxed. This is our objective in the present chapter.

In this chapter, we assume users are greedy and do not implement congestion control. They send data at the greatest possible rate determined by an external constraint that we assimilate to an access rate. Bandwidth sharing is determined by this behavior. We consider a flow-level model, as described in [MR00], where the bandwidth allocation changes instantly as new flows arrive or existing flows terminate. Tail dropping is interpreted in this model such that, at each link, the output rate of flows is proportional to their input rate to the link in question. This is consistent with the assumption that packets are all equally likely to be dropped. The drop rate is such that the overall output rate is bounded by the link capacity. Flows are categorized into classes defined by their route and their access rate. We assume flows arrive according to a Poisson process and have an exponentially distributed size. The stochastic process describing the number of flows in each class is a Markov process. We say that the network is stable when this process is positive recurrent and unstable otherwise.

To illustrate the realized bandwidth sharing without congestion control, consider the linear network depicted in Figure 2 with two links 1, 2 of capacity C_1 and C_2 , respectively, n_0 class-0 flows going through links 1 and 2, n_1 class-1 flows going through link 1 and n_2 class-2 flows going through link 2. The access rate for class- k flows is denoted by a_k . The aggregate input rates of classes 0 and 1 at the first link are respectively $n_0 a_0$ and $n_1 a_1$ and the total input rate at link 1 is $R_1 = n_0 a_0 + n_1 a_1$. If $R_1 > C_1$, then the first link is saturated and the aggregate output rates of classes 0 and 1 are respectively

$$\theta_0^1 = C_1 \frac{n_0 a_0}{n_0 a_0 + n_1 a_1} \quad \text{and} \quad \theta_1^1 = C_1 \frac{n_1 a_1}{n_0 a_0 + n_1 a_1}.$$

If $R_1 \leq C_1$ then the first link is not saturated and $\theta_0^1 = n_0 a_0$ and $\theta_1^1 = n_1 a_1$. The aggregate input rate of class 0 at link 2 is θ_0^1 and the aggregate input rate of class 2 is $n_2 a_2$. As above, we derive the respective aggregate output rates of class 0 and 2 after link 2:

$$\theta_0^2 = \theta_0^1 \min \left(1, \frac{C_2}{\theta_0^1 + n_2 a_2} \right) \quad \text{and} \quad \theta_2^2 = n_2 a_2 \min \left(1, \frac{C_2}{\theta_0^1 + n_2 a_2} \right).$$

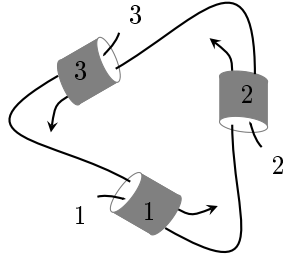


FIGURE 1. A cyclic network

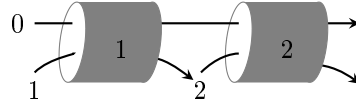


FIGURE 2. A linear network

The question of the capacity of networks without congestion control has already been addressed in [BFP09]. It was shown that congestion collapse does occur in cyclic networks like that depicted in Figure 1 where the system is unstable for any positive demand. In acyclic networks, however, capacity is not reduced to 0 and

this performance degradation is significantly mitigated by the access rates. It was conjectured in [BFP09] that capacity in acyclic networks increases as access rates tend to 0 and, in the limit, is determined only by the optimal per-link stability conditions. Here, we prove that the conjecture is indeed true for two particular acyclic topologies, linear networks and upstream trees. For linear networks, we derive bounds on stability conditions for arbitrary access rates.

Averaging Phenomenon. We evaluate stability conditions by applying fluid limit techniques. The end-to-end class in a linear network plays a particular role. In the fluid limit, all other flow classes are shown to empty in finite time for any initial conditions. When all these classes are empty, a phenomenon of *averaging* occurs and allows us to prove that the end-to-end class also empties in finite time. In that situation, the rate at which this class evolves depends on the stochastic evolution of the remaining local classes which in turn depends on the current fluid state of the end-to-end class. This dependence is decoupled by a time scale separation argument. The rate at which the fluid component (the end-to-end flow) evolves is infinitely slower than the rate at which the stochastic components evolve. It is therefore possible to apply a *quasi-stationary* model through which the *average* impact of the local traffic can be evaluated.

This is a non-trivial example of a local equilibrium of fluid limits. A simpler example of such a phenomenon can be found in [Rob03, 9.6, p271]. Averaging in fluid limits has also been considered in the context of wireless networks (see [BBP04] and subsequent papers [ST10] and [GLM⁺10]). The main difference here is that the local equilibrium depends on the state of the fluid limit. Note that this is somewhat related to the averaging phenomenon considered for loss networks [HK94].

To evaluate the impact of access rates on capacity, we introduce an ad-hoc scaling. Scaled processes are shown to converge to a simple deterministic process. Generally speaking, the convergence of processes does not imply the convergence of stationary distributions. However, in our case, we can prove this is true in two specific cases exhibiting an interesting example of limit inversion for stochastic processes. We use the same technique as in [DGR02] to prove this inversion. We use the convergence of stationary distributions combined with the averaging phenomenon to prove the asymptotic optimality of linear networks. For upstream trees, we use convergence and monotonicity.

The chapter is structured as follows. In Section 2, we give a complete description of the considered model. In Section 3, we study the stability conditions of linear networks with general access rates. In Section 4, we study the stability of linear networks and upstream trees when access rates decrease to 0 and prove their capacity is asymptotically optimal.

2. Flow-Level Model

2.1. Bandwidth Allocation. Consider a network of L links. Denote by C_l the capacity of link l . A number of flows compete to share the capacity of these links. The flows are categorized into a set of K classes. Each class- k flow has an access rate to the network denoted by a_k and follows a route of length d_k defined as an ordered set of distinct links $r_k = \{r_k(1), r_k(2), \dots, r_k(d_k)\}$. Let $x = (x_1, \dots, x_K)$ be the vector of input rates of all classes, $\psi_k(x)$ the aggregate throughput of class- k flows. We refer to the vector $\psi(x) = (\psi_1(x), \dots, \psi_K(x))$ as the bandwidth allocation.

Here, we consider allocations that result when users are greedy and transmit at their maximum input rate in the network without any congestion control. There can be losses on each link of the network and we need to describe the evolution of

the rate of each flow going through the network. Specifically, we assume that class- k flows initially transmit at full rate $\theta_k^0(x) = x_k = n_k a_k$ where n_k is the number of class- k active flows and we denote by $\theta_k^i(x)$ the class- k rate at the *output* of the i -th link on its route r_k , for $i = 1, 2, \dots, d_k$. The actual throughput $\psi_k(x)$ of class k corresponds to the aggregate throughput of class- k flows at the output of the last link on their route, $\psi_k(x) = \theta_k^{d_k}(x)$. Due to potential loss at each link, we have

$$(2.1) \quad \psi_k(x) = \theta_k^{d_k}(x) \leq \theta_k^{d_k-1}(x) \leq \dots \leq \theta_k^1(x) \leq \theta_k^0(x) = x_k = n_k a_k.$$

In particular, $\theta_k^{i-1}(x) - \theta_k^i(x)$ is the loss for class k occurring at the i -th link on its route.

The total rate at the output of each link cannot exceed the capacity of this link, so that

$$(2.2) \quad \forall l, \quad \sum_{k, i: r_k(i)=l} \theta_k^i(x) \leq C_l.$$

In order to characterize the allocation achieved in the absence of congestion control, it remains to determine how output rates depend on input rates at each link. This depends on the buffer management policy. In the considered fluid model, losses occur on saturated links only when the total input rate exceeds the capacity; the total output rate is then assumed to be equal to the capacity. In the following, we consider a single buffer management policy, Tail Dropping, which simply consists in dropping each incoming packet when the buffer is full. We deduce that, at the flow level, the output rate of classes from the link is proportional to the input rate of classes to the link.

Specifically, let $R_l(x)$ be the total input rate of link l :

$$R_l(x) = \sum_{k, i: r_k(i+1)=l} \theta_k^i(x).$$

If $R_l(x) \leq C_l$ then link l is not saturated and there are no losses; the output rate is equal to the input rate for each class. If $R_l(x) > C_l$, then the link is saturated and there are losses proportional to the input rate. The output rate of any class k then satisfies for all $i = 1, 2, 3, \dots, d_k$,

$$(2.3) \quad \theta_k^i(x) = \theta_k^{i-1}(x) \min\left(\frac{C_l}{R_l(x)}, 1\right) \quad \text{where } l = r_k(i).$$

When there is no ambiguity and when the access rates are fixed, the bandwidth allocation is determined by the number of flows in each class. We denote by $n = (n_1, \dots, n_K)$ the vector of the numbers of flows in progress and by the vector $\phi(n) = (\phi_1(n), \dots, \phi_K(n))$ the bandwidth allocation where, for each k , $\phi_k(n) = \psi_k(a \odot n)$ where $\odot$ denotes the componentwise product.

For example, we can consider a linear network under tail dropping (see Section 3) with two links of capacity 1 with access rates all equal to 1 and 1 flow in each class. We get the following bandwidth allocation:

$$\phi_0(1, 1, 1) = \frac{1}{3}, \quad \phi_1(1, 1, 1) = \frac{1}{2}, \quad \phi_2(1, 1, 1) = \frac{2}{3}.$$

It is not obvious that (2.3) defines a unique allocation for all networks. The definition is clearly non-ambiguous for *acyclic* networks, i.e., if links can be numbered in such a way that each route consists of an increasing sequence of link indexes. The allocation then directly follows from applying (2.3) to all $l = 1, \dots, L$, successively. For general networks, the allocation is still well-defined and unique as proved in [BFP09]. Here, we consider only acyclic networks.

2.2. Flow Dynamics. We assume that class- k flows are generated according to a Poisson process of intensity λ_k and have independent, exponentially distributed sizes with mean $1/\mu_k$. We define the traffic intensity of class- k flows as $\rho_k = \lambda_k/\mu_k$.

Under the above assumptions, let $N(t)$ denote the number of flows at time t , $(N(t), t \in \mathbb{R}_+)$ defines a Markov process with transition rates λ_k from state n to state $n + e_k$ and $\phi_k(n)\mu_k$ from state n to state $n - e_k$, where e_k denotes the vector with 1 in the k -th component and 0 elsewhere. We say that the network is *stable* when this Markov process is ergodic. There is a well-known necessary stability condition (see [BMPV06]) which is that the traffic intensity is less than the capacity for all links:

$$(2.4) \quad \forall l, \quad \sum_{k: l \in r_k} \rho_k < C_l.$$

An allocation will be called optimal if this condition is also sufficient. In the rest of this chapter, this condition will be further referred to as the *optimal stability condition*.

In the following, we study stability conditions for two specific classes of networks. In Section 3, we consider linear networks. In Section 4, we study the behavior of stability conditions when the access rates are small for linear networks and upstream trees.

3. Stability Conditions for a Linear Network

3.1. Linear Networks. We consider a linear network with L links as depicted in Figure 2. For the sake of simplicity, we assume that the capacity of all links is 1 but all the results are true in the general case and can be obtained by adapting the notation. The linear network has $K = L + 1$ different classes of flows. Class-0 flows go through all links. We label the links of the network from 1 to L . Link l is the l -th link on the route of class-0 flows. Class- l flows go through link l only. Equation (2.3) can be rewritten as follows for a given state n and for $1 \leq k \leq L$:

$$\begin{aligned} \theta_0^k(x) &= \min \left(\theta_0^{k-1}(x), \frac{\theta_0^{k-1}(x)}{\theta_0^{k-1}(x) + x_k} \right), \\ \theta_k^1(x) &= \min \left(x_k, \frac{x_k}{\theta_0^{k-1}(x) + x_k} \right). \end{aligned}$$

To simplify the presentation, we suppose that $a_0 = 1$ and $a_1 = 1$. In that case, a single flow of class 0 or 1 is enough to saturate the first link and we have

$$\theta_0^1(n \odot a) = \frac{n_0}{n_1 + n_0} \quad \text{and} \quad \theta_1^1(n \odot a) = \frac{n_1}{n_0 + n_1}.$$

All the results presented here remain true without this assumption. Only the proof of Proposition 4.12 has to be adapted.

In the rest of this section, we study the stability conditions of the stochastic process resulting from this linear network. The optimal stability conditions (2.4) are:

$$(3.1) \quad \rho_0 + \rho_k < 1 \text{ for } 1 \leq k \leq L.$$

3.2. General Fluid Limits. We show that classes $2, \dots, L$ are favored compared to classes 0 and 1 in the sense that the ergodicity conditions of classes $2, \dots, L$ do not depend on classes 0 and 1 but the contrary is not true.

In order to study the ergodicity of the system, we need to define fluid limits. In the following, we consider the norm such that for $x \in \mathbb{R}^K$, $\|x\| = \sum_{i=1}^K |x_i|$. Let $(m^i, i \in \mathbb{N})$ be a sequence of $\mathbb{N}^K$ such that $\lim_{i \rightarrow \infty} \|m^i\| = \infty$. Define for all $m \in \mathbb{N}^K$, the process $(N^m(t))$ describing the system evolution under tail dropping

when starting at $N^m(0) = m$. A fluid limit $\bar{Z}$ of the system is defined as an accumulation point of the laws of the processes in the set

$$\mathcal{C} = \left\{ (N^{m^i}(\|m^i\|t)/\|m^i\|, t \geq 0), i \in \mathbb{N} \right\}$$

which is regarded as a subset of the space $D_{\mathbb{R}_+^N}([0, \infty))$ with Skorohod topology (see [Bil99]). It can be proved, as in [Rob03, Prop 9.3 p 246], that the set $\mathcal{C}$ is relatively compact and that all the fluid limits $\bar{Z}$ are continuous.

In order to prove ergodicity of the process $(N(t))$, we just have to prove that there exists a time T such that for any initial condition of $(\bar{Z}(t))$, for all $t \geq T$, $\bar{Z}(t) = 0$ (see [Dai95]). On the contrary, in order to prove transience, we have to prove that there exists a time T such that after T , any fluid limit $(\bar{Z}(t))$ increases linearly (see [Mey95]). For that purpose, we have to characterize the fluid limits.

First, we consider fluid limits such that there exists $2 \leq k_0 \leq L$ with $\bar{Z}_{k_0}(0) > 0$. At the stochastic level, for each $k \geq 2$, class k can use an arbitrarily large part of link k leaving almost nothing to class 0; for any $\varepsilon > 0$, since $\theta_0^{k-1}(n \odot a) \leq 1$, if n_k is large enough, we have

$$\phi_k(n) \geq 1 - \varepsilon, \text{ and } \theta_0^k(n \odot a) \leq \varepsilon.$$

Thus, at fluid level, the throughput for $(\bar{Z}_k(t))$ is always 1 if $\bar{Z}_k(t) > 0$. On the contrary, the throughput for $(\bar{Z}_0(t))$ is 0 as long as there exists k_0 with $\bar{Z}_{k_0}(t) > 0$. We can prove, as in [Rob03, Prop 9.4, p247], the following proposition.

PROPOSITION 3.1. *If $(\bar{Z}(t))$ is a fluid limit of the system and there exists some time interval $[0, t_0]$ such that for any $t \in [0, t_0]$, there is some k_0 in $\{2, \dots, L\}$ such that $\bar{Z}_{k_0}(t) > 0$ then, almost surely, $(\bar{Z}(t))$ is differentiable on $[0, t_0]$ and satisfies for all $t \in [0, t_0]$,*

$$(3.2) \quad \dot{\bar{Z}}_0(t) = \lambda_0,$$

$$(3.3) \quad \dot{\bar{Z}}_1(t) = \lambda_1 - \mu_1 \frac{\bar{Z}_1(t)}{\bar{Z}_0(t) + \bar{Z}_1(t)},$$

$$(3.4) \quad \dot{\bar{Z}}_k(t) = (\lambda_k - \mu_k) \mathbb{1}_{\{\bar{Z}_k(t) > 0\}} \text{ for } 2 \leq k \leq L.$$

For $2 \leq k \leq L$, $\bar{Z}_k(t)$ thus decreases linearly to 0 if $\rho_k < 1$. Note also that $(\bar{Z}_k(t))$ for $2 \leq k \leq L$ does not depend on classes 0 and 1. This illustrates the strong asymmetry between classes 0 and 1 on one hand and classes $2, \dots, L$ on the other hand.

If there exists $2 \leq k_0 \leq L$ such that $\rho_{k_0} > 1$ then it is clear that the fluid limit will increase linearly to infinity and the process $(N(t))$ is transient. On the contrary, if $\rho_{k_0} < 1$ for all $2 \leq k \leq L$ then there exists a time T such that for $t \leq T$ and $2 \leq k \leq L$, $\bar{Z}_k(t) = 0$. In order to study ergodicity and transience of $(N(t))$ we then have to characterize the fluid limits such that $\bar{Z}_k(0) = 0$ for $2 \leq k \leq L$. This is the purpose of the next subsection.

3.3. Quasi-Stationary Fluid Limits. Before studying fluid limits, we need to define the average throughput of class 0 in the quasi-stationary case. This corresponds to the case where the number of flows in classes 0 and 1 is fixed and the number of flows in classes $2, \dots, L$ varies. In the following, we denote by $x_{2:L} = (x_2, \dots, x_L)$ the input rates of classes $2, \dots, L$. Fixing the number of flows in classes 0 and 1 is equivalent to considering a fixed rate α for class 0 after link 1.

Since link 1 is of capacity 1, $\alpha \in [0, 1]$. Define $\tilde{\theta}_0^1(\alpha, x_{2:L}) = \alpha$ and, for $2 \leq k \leq L$,

$$(3.5) \quad \tilde{\theta}_0^k(\alpha, x_{2:L}) = \min \left(\tilde{\theta}_0^{k-1}(\alpha, x_{2:L}), \frac{\tilde{\theta}_0^{k-1}(\alpha, x_{2:L})}{\tilde{\theta}_0^{k-1}(\alpha, x_{2:L}) + x_k} \right),$$

$$(3.6) \quad \tilde{\psi}_k(\alpha, x_{2:L}) = \min \left(x_k, \frac{x_k}{\tilde{\theta}_0^{k-1}(\alpha, x_{2:L}) + x_k} \right).$$

The quantity $\tilde{\theta}_0^k(\alpha, x_{2:L})$ is the output rate of class 0 after link k and the throughput of class 0 is $\tilde{\psi}_0(\alpha, x_{2:L}) = \tilde{\theta}_0^L(\alpha, x_{2:L})$. Moreover, $\tilde{\psi}_k(\alpha, x_{2:L})$ is the throughput of class k . For $0 \leq k \leq L$ and for $n_{2:L} \in \mathbb{N}^{K-2}$, we define $\tilde{\phi}_k(\alpha, n_{2:L}) = \tilde{\psi}_k(\alpha, n_{2:L} \odot a)$.

Denote by $(\tilde{N}_{2:L}^\alpha(t))$ the Markov process describing the evolution of classes $2, \dots, L$ in the quasi-stationary case. It is defined on $\mathbb{N}^{K-2}$ and its transition rates are, for $2 \leq k \leq L$,

$$\begin{aligned} n_k &\rightarrow n_k + 1 : \lambda_k, \\ n_k &\rightarrow n_k - 1 : \mu_k \tilde{\phi}_k(\alpha, n_{2:L}). \end{aligned}$$

PROPOSITION 3.2. *Consider a linear network with L links of capacity 1 in the quasi-stationary case, i.e., when the number of flows in classes 0 and 1 is fixed and the rate of class 0 after link 1 is α .*

The Markov process $(\tilde{N}_{2:L}^\alpha(t))$ describing the evolution of classes $2, \dots, L$ is ergodic if, for $2 \leq k \leq L$, $\rho_k < 1$. It is transient if there exists k_0 in $\{2, \dots, L\}$ such that $\rho_{k_0} > 1$.

PROOF. It is enough to see that for $2 \leq k \leq L$,

$$\min(n_k a_k, C_k) \geq \tilde{\phi}_k(\alpha, n_{2:L}) \geq \frac{n_k a_k}{C_{k-1} + n_k a_k}.$$

Since $C_k = 1$, if $\rho_k < 1$ for all k , there exists $\varepsilon > 0$ and η in $\mathbb{N}$ such that for all $n_{2:L}$ in $\mathbb{N}^{K-2} \setminus \{0, \dots, \eta\}^{K-2}$ and for $2 \leq k \leq L$ such that $n_k \geq \eta$,

$$\lambda_k - \mu_k \tilde{\phi}_k(\alpha, n_{2:L}) \leq -\varepsilon$$

and $(\tilde{N}_{2:L}^\alpha(t))$ is therefore ergodic.

Conversely, if there exists k_0 such that $\rho_{k_0} > 1$, then there exists $\varepsilon > 0$ and η such that for all $n_{2:L}$ with $n_{k_0} > \eta$,

$$\lambda_{k_0} - \mu_{k_0} \tilde{\phi}_{k_0}(\alpha, n_{2:L}) \geq \varepsilon$$

and $(\tilde{N}_{2:L}^\alpha(t))$ is therefore transient. ■

When $(\tilde{N}_{2:L}^\alpha(t))$ is ergodic, denote its unique stationary distribution by $\tilde{\pi}^\alpha$. The average throughput of class 0 in the quasi-stationary case is then defined as follows:

$$(3.7) \quad \forall \alpha \in [0, 1] \quad \bar{\phi}_0(\alpha) = \mathbb{E}_{\tilde{\pi}^\alpha}(\tilde{\phi}_0(\alpha, \cdot)) = \sum_{n_{2:L} \in \mathbb{N}^{K-2}} \tilde{\pi}^\alpha(n_{2:L}) \tilde{\phi}_0(\alpha, n_{2:L}).$$

The average $\bar{\phi}_0$ depends on the traffic intensities and access rates of classes $2, \dots, L$. In order to establish ergodicity and transience conditions for the linear network in Theorems 3.6 and 3.7 below, we need the continuity of the function $\bar{\phi}_0$ with respect to α .

Subsequent developments rely on the following notion of stochastic domination.

DEFINITION 3.3. *Let X and Y be two random variables in a partially ordered measurable space. We denote $X \leq_{\text{st}} Y$ and say that Y stochastically dominates X if, for any positive non-decreasing measurable function f , we have $\mathbb{E}[f(X)] \leq \mathbb{E}[f(Y)]$.*

On $\mathbb{R}^n$, we use the usual coordinate-wise partial orders. If x and y are in $\mathbb{R}^n$, $x \leq y$ if $x_k \leq y_k$ for all k . On the space $D(\mathbb{R}_+, \mathbb{R}^n)$, we say that $x \leq y$ if $x_k(t) \leq y_k(t)$ for all k and t .

For any two random variables X and Y with respective distributions π_X and π_Y and such that $X \leq_{\text{st}} Y$, we say that π_Y dominates π_X .

In particular, if $(X(t))$ and $(Y(t))$ are two Markov processes on $\mathbb{N}^n$ such that $(X(t)) \leq_{\text{st}} (Y(t))$ and $(X(t))$ is irreducible, the ergodicity of $(Y(t))$ implies the ergodicity of $(X(t))$ and the transience of $(X(t))$ implies the transience of $(Y(t))$. If $(Y(t))$ and $(X(t))$ are ergodic and $X(\infty)$ and $Y(\infty)$ are random variables with the stationary distributions of $(X(t))$ and $(Y(t))$ as distributions then $X(\infty) \leq_{\text{st}} Y(\infty)$. For more details on stochastic domination, see [Mas87].

LEMMA 3.4. *Consider a linear network with L links of capacity 1 in the quasi-stationary case and assume that $\rho_k < 1$ for $2 \leq k \leq L$, the function $\bar{\phi}_0$ defined by Equation (3.7) is continuous with respect to α on $[0, 1]$.*

PROOF. First, we prove that $\alpha \mapsto \tilde{\phi}_k(\alpha, n_{2:L})$ is 1-Lipschitz for all $n_{2:L}$ and k . By definition, $\alpha \mapsto \tilde{\theta}_0^1(\alpha, n_{2:L})$ is 1-Lipschitz. According to (3.5) and (3.6), $\alpha \mapsto \tilde{\theta}_0^2(\alpha, n_{2:L})$ and $\alpha \mapsto \tilde{\phi}_2(\alpha, n_{2:L})$ are defined as the minimum of two 1-Lipschitz functions and are then 1-Lipschitz functions. The result follows by recursion.

In particular, this is true for $\alpha \mapsto \tilde{\phi}_0(\alpha, n_{2:L})$. In order to prove the continuity of $\bar{\phi}_0$, we just have to prove that if a sequence (α_i) of $[0, 1]$ converges to α_∞ , then $\tilde{N}_{2:L}^{\alpha_i}(\infty)$ converges in distribution to $\tilde{N}_{2:L}^{\alpha_\infty}(\infty)$ and these random variables have $\tilde{\pi}^{\alpha_i}$ and $\tilde{\pi}^{\alpha_\infty}$ as respective distributions.

We first prove that $\{\tilde{\pi}^\alpha, \alpha \in [0, 1]\}$ is tight. Note that for $2 \leq k \leq L$ and α in $[0, 1]$,

$$(3.8) \quad \tilde{\phi}_k(\alpha, n_{2:L}) \geq \frac{n_k a_k}{1 + n_k a_k}.$$

Define the process $(\hat{N}_{2:L}(t))$ with the following transition rates for $2 \leq k \leq L$:

$$\begin{aligned} n_k &\mapsto n_k + 1 : \lambda_k, \\ n_k &\mapsto n_k - 1 : \mu_k \frac{n_k a_k}{1 + n_k a_k}, \end{aligned}$$

and such that $\hat{N}_{2:L}(0) = \tilde{N}_{2:L}^\alpha(0)$. The components of $(\hat{N}_{2:L}(t))$ are independent. For $2 \leq k \leq L$, we have $\rho_k < 1$ and there exist $\varepsilon > 0$ and η such that, for $n_k \geq \eta$, we have

$$\lambda_k - \mu_k \frac{n_k a_k}{1 + n_k a_k} < -\varepsilon,$$

which implies that $(\hat{N}_k(t))$ is ergodic; the ergodicity of $(\hat{N}_{2:L}(t))$ follows. Using Equation (3.8) and a standard coupling argument, see for instance [BJL08, Lemma 1], we can prove that, for α in $[0, 1]$ and $2 \leq k \leq L$, $(\hat{N}_k(t))$ stochastically dominates $(\tilde{N}_k^\alpha(t))$. It implies that $(\hat{N}_{2:L}(t))$ stochastically dominates $(\tilde{N}_{2:L}^\alpha(t))$ for any $\alpha \in [0, 1]$. This implies that the stationary distribution $\hat{\pi}$ of $(\hat{N}_{2:L}(t))$ dominates any distribution $\tilde{\pi}^\alpha$. In particular, we have

$$\forall \alpha \in [0, 1], \forall \kappa \in \mathbb{R}, \tilde{\pi}^\alpha(\mathbb{N}^{K-2} \setminus [0, \kappa]^{K-2}) \leq \hat{\pi}(\mathbb{N}^{K-2} \setminus [0, \kappa]^{K-2}).$$

Thus, the set $\{\tilde{\pi}^\alpha, \alpha \in [0, 1]\}$ is tight on $\mathbb{N}^{K-2}$ with the usual topology.

By tightness, we can suppose that the sequence $(\tilde{\pi}^{\alpha_i})_{i \in \mathbb{N}}$ is convergent. We call $\tilde{\pi}_\infty$ its limit. All we have to do now is to characterize this limit and prove its uniqueness.

Let f be a bounded function on $\mathbb{N}^{K-2}$. For any α in $[0, 1]$, let $\tilde{\Omega}^\alpha$ denote the infinitesimal generator of $(\tilde{N}_{2:L}^\alpha(t))$. For all $n_{2:L} \in \mathbb{N}^{K-2}$,

$$\begin{aligned} \tilde{\Omega}^\alpha(f)(n_{2:L}) &= \sum_{k=2}^L \lambda_k (f(n_{2:L} + e_k) - f(n_{2:L})) \\ &\quad - \sum_{k=2}^L \mu_k \tilde{\phi}_k(\alpha, n_{2:L}) (f(n_{2:L} - e_k) - f(n_{2:L})). \end{aligned}$$

Since $\alpha \mapsto \tilde{\phi}_k(\alpha, n_{2:L})$ is 1-Lipschitz for all $n_{2:L}$ and k and f is bounded, there exists η such that

$$(3.9) \quad \forall i \in \mathbb{N}, \forall n_{2:L} \in \mathbb{N}^{K-2}, |\tilde{\Omega}^{\alpha_\infty}(f)(n_{2:L}) - \tilde{\Omega}^{\alpha_i}(f)(n_{2:L})| \leq \eta |\alpha_\infty - \alpha_i|.$$

Because $(\tilde{\pi}^{\alpha_i})$ is tight, for any ε , there exists $\kappa > 0$ such that

$$(3.10) \quad \forall i \in \mathbb{N}, \tilde{\pi}^{\alpha_i}([0, \kappa]^{K-2}) \geq 1 - \varepsilon.$$

Using equations (3.9) and (3.10), we have

$$\lim_{i \rightarrow \infty} \int_{\mathbb{N}^{K-2}} \tilde{\Omega}^{\alpha_i}(f)(n_{2:L}) \tilde{\pi}^{\alpha_i}(dn_{2:L}) = \int_{\mathbb{N}^{K-2}} \tilde{\Omega}^{\alpha_\infty}(f)(n_{2:L}) \tilde{\pi}_\infty(dn_{2:L}) = 0.$$

We deduce that $\tilde{\pi}_\infty$ is an invariant distribution of $(\tilde{N}_{2:L}^{\alpha_\infty}(t))$ and by uniqueness, we have $\tilde{\pi}_\infty = \tilde{\pi}^{\alpha_\infty}$. Finally, we have

$$\lim_{i \rightarrow \infty} \tilde{\phi}_0(\alpha_i) = \tilde{\phi}_0(\alpha_\infty).$$

■

We are now able to characterize the fluid limits $(\bar{Z}(t))$ such that $Z_k(0) = 0$ for $2 \leq k \leq L$. The next proposition is proved in the appendix.

PROPOSITION 3.5. *Consider a linear network of L links of capacity 1 and assume that $a_0 = 1$ and $a_1 = 1$.*

If $(\bar{Z}(t))$ is a fluid limit of the system such that $\bar{Z}_k(0) = 0$ for $2 \leq k \leq L$, then almost surely

$$(3.11) \quad \dot{\bar{Z}}_0(t) = \lambda_0 - \mu_0 \bar{\phi}_0 \left(\frac{\bar{Z}_0(t)}{\bar{Z}_1(t) + \bar{Z}_0(t)} \right),$$

$$(3.12) \quad \dot{\bar{Z}}_1(t) = \lambda_1 - \mu_1 \frac{\bar{Z}_1(t)}{\bar{Z}_1(t) + \bar{Z}_0(t)},$$

$$(3.13) \quad \dot{\bar{Z}}_k(t) = 0 \quad \text{for } 2 \leq k \leq L$$

hold for all $t \in \mathbb{R}_+$ where $\bar{\phi}_0$ is the average throughput of class 0 in the quasi-stationary case defined by Equation (3.7).

This proposition shows that, at the fluid time scale, there is a separation of time scales between classes 0 and 1 and $2, \dots, L$. From the point of view of classes 0 and 1, classes $2, \dots, L$ are quasi-stationary and there is an averaging because they evolve infinitely faster. From the point of view of classes $2, \dots, L$, the ratio of classes 0 and 1 is constant. This is a further illustration of the asymmetry between classes 0, 1 on the one hand and classes $2, \dots, L$ on the other hand.

3.4. Stability of a Linear Network. Now that we have characterized the fluid limits of the system, we can study the stability conditions of the linear network. Theorem 3.6 gives a sufficient condition for stability and Theorem 3.7 gives a sufficient condition for transience of $(N(t))$.

THEOREM 3.6. *Consider a linear network of L links of capacity 1. If $\rho_k < 1$ for $2 \leq k \leq L$ and*

$$\rho_0 < \inf_{1-\rho_1 \leq x \leq 1} \bar{\phi}_0(x)$$

then the network is stable, i.e., $(N(t))$ is ergodic.

PROOF. In order to prove ergodicity, we just have to prove that there exists a time T such that for any initial condition of $(\bar{Z}(t))$, for all $t \geq T$, $\bar{Z}(t) = 0$ (see [Dai95]).

We consider a general fluid limit $(\bar{Z}(t))$. As proved in Proposition 3.1, as long as there exists k_0 such that $\bar{Z}_{k_0}(t) > 0$, $(\bar{Z}(t))$ satisfied Equations (3.2), (3.3) and (3.4). Since $\rho_k < 1$ for $2 \leq k \leq L$, there exists a finite time T_0 such that, for all $k \geq 0$ and $t \geq T_0$, $\bar{Z}_k(t) = 0$. According to the strong Markov property, the study is reduced to the case where the initial state of the fluid limit verifies $\bar{Z}_0(0) + \bar{Z}_1(0) = 1$ and $\bar{Z}_k(0) = 0$ for $2 \leq k \leq L$. According to Proposition 3.5, $(\bar{Z}(t))$ satisfies (3.11), (3.12) and (3.13). In particular, for all $t \geq 0$ and $2 \leq k \leq L$, we have $\bar{Z}_k(t) = 0$, we then just have to study the behavior of $(\bar{Z}_0(t), \bar{Z}_1(t))$.

We define the following function:

$$\forall \alpha \in [0, 1], f(\alpha) = \inf_{\alpha \leq x \leq 1} \bar{\phi}_0(x)$$

This function is continuous, non-decreasing and satisfies $f(\alpha) \leq \bar{\phi}_0(\alpha)$ for all $\alpha \in [0, 1]$. We define $(F(t))$ such that $F_0(0) = \bar{Z}_0(0)$, $F_1(0) = \bar{Z}_1(0)$ and for all $t \geq 0$,

$$(3.14) \quad \dot{F}_0(t) = \lambda_0 - \mu_0 f\left(\frac{F_0(t)}{F_0(t) + F_1(t)}\right),$$

$$(3.15) \quad \dot{F}_1(t) = \lambda_1 - \mu_1 \frac{F_1(t)}{F_0(t) + F_1(t)}.$$

Since we have $\bar{\phi}_0(\alpha) \geq f(\alpha)$, we can deduce for all $t \geq 0$, $\bar{Z}_0(t) \leq F_0(t)$ and $\bar{Z}_1(t) \leq F_1(t)$. This follows from the following.

- If $\bar{Z}_0(t) = F_0(t)$ and $\bar{Z}_1(t) = F_1(t)$ then $\dot{\bar{Z}}_0(t) \leq \dot{F}_0(t)$ and $\dot{\bar{Z}}_1(t) \leq \dot{F}_1(t)$;
- if $\bar{Z}_0(t) < F_0(t)$ and $\bar{Z}_1(t) = F_1(t)$ then $\dot{\bar{Z}}_1(t) \leq \dot{F}_1(t)$;
- if $\bar{Z}_0(t) = F_0(t)$ and $\bar{Z}_1(t) < F_1(t)$ then $\dot{\bar{Z}}_0(t) \leq \dot{F}_0(t)$.

To prove the stability of $(\bar{Z}(t))$, it is enough to prove that $(F(t))$ will return to 0 in finite time. Let $\alpha(t) = F_0(t)/(F_0(t) + F_1(t))$. We have

$$(3.16) \quad \dot{\alpha}(t) = \frac{1}{F_0(t) + F_1(t)} [(\lambda_0 - \mu_0 f(\alpha(t)))(1 - \alpha(t)) - (\lambda_1 - \mu_1(1 - \alpha(t)))\alpha(t)].$$

Let $\beta(t) = (\lambda_0 - \mu_0 f(\alpha(t)))(1 - \alpha(t)) - (\lambda_1 - \mu_1(1 - \alpha(t)))\alpha(t)$. So that

$$(3.17) \quad \dot{\alpha}(t) = \frac{\beta(t)}{F_0(t) + F_1(t)}.$$

We also define $\alpha_0 = \max\{\alpha \in [0, 1], f(\alpha) = \rho_0\}$. Since $\rho_0 < f(1 - \rho_1)$, we have $\alpha_0 < 1 - \rho_1$. Further, for t_0 such that $\alpha(t) = \alpha_0$ and t_1 such that $\alpha(t_1) = 1 - \rho_1$,

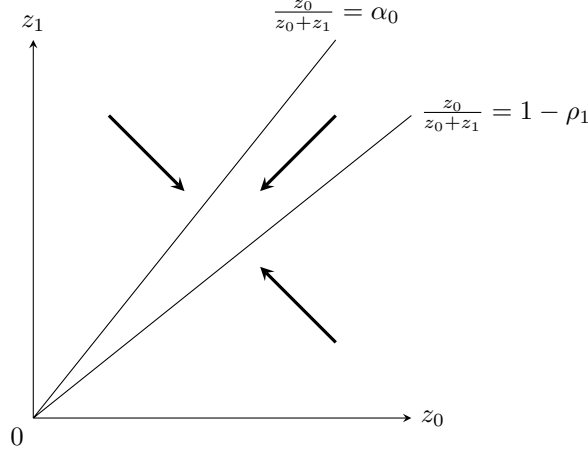
$$\beta(t_0) = \alpha_0(\mu_1(1 - \alpha_0) - \lambda_1) > 0,$$

$$\beta(t_1) = \rho_1(\lambda_0 - \mu_0 f(1 - \rho_1)) < 0.$$

Since f is continuous and non-decreasing, there exists $\kappa > 0$ and $\eta > 0$ such that if $\alpha(t) \leq \alpha_0 + \eta$, then $\beta(t) \geq \kappa$ and if $\alpha(t) \geq 1 - \rho_1 - \eta$, then $\beta(t) \leq -\kappa$.

As a consequence, we deduce from Equation (3.17) that

$$\begin{aligned} \dot{\alpha}(t) &\geq \frac{\kappa}{F_0(t) + F_1(t)} \text{ if } \alpha(t) \leq \alpha_0 + \eta, \\ \dot{\alpha}(t) &\leq -\frac{\kappa}{F_0(t) + F_1(t)} \text{ if } \alpha(t) \geq 1 - \rho_1 - \eta. \end{aligned}$$

FIGURE 3. Dynamics of $(F_0(t), F_1(t))$

Thus, if $\alpha(0) \in [\alpha_0 + \eta, 1 - \rho_1 - \eta]$, $\alpha(t) \in [\alpha_0 + \eta, 1 - \rho_1 - \eta]$ for all $t > 0$.

For all $t \geq 0$, we have

$$\begin{aligned} F_0(t) + F_1(t) &\leq F_0(0) + F_1(0) + (\lambda_0 + \lambda_1)t, \\ &\leq 1 + (\lambda_0 + \lambda_1)t, \end{aligned}$$

and we deduce that if $\alpha(0)$ is in $[0, \alpha_0 + \eta]$ and as long as $\alpha(t)$ stays in $[0, \alpha_0 + \eta]$, we have

$$\begin{aligned} \alpha(t) &\geq \int_0^t \frac{\kappa ds}{1 + (\lambda_0 + \lambda_1)s} \\ &\geq \frac{\kappa}{\lambda_0 + \lambda_1} \log(1 + (\lambda_0 + \lambda_1)t). \end{aligned}$$

Similarly, if $\alpha(0)$ is in $[1 - \rho_1 - \eta, 1]$ and $\alpha(t)$ stays in $[1 - \rho_1 - \eta, 1]$, we have

$$\begin{aligned} \alpha(t) &\leq 1 - \int_0^t \frac{\kappa ds}{1 + (\lambda_0 + \lambda_1)s} \\ &\leq 1 - \frac{\kappa}{\lambda_0 + \lambda_1} \log(1 + (\lambda_0 + \lambda_1)t). \end{aligned}$$

We then define

$$T_1 = \frac{1}{\lambda_0 + \lambda_1} \exp\left(\frac{\max(\alpha_0 + \eta, \rho_1 + \eta)(\lambda_0 + \lambda_1)}{\kappa}\right)$$

and if $\alpha(0) \in [0, \alpha_0 + \eta] \cup [1 - \rho_1 - \eta, 1]$ then $\alpha(T_1) \in [\alpha_0 + \eta, 1 - \rho_1 - \eta]$. Finally, for all $t \geq T_1$, $\alpha(t) \in [\alpha_0 + \eta, 1 - \rho_1 - \eta]$.

Now, we know that $\alpha(t)$ will reach $[\alpha_0 + \eta, 1 - \rho_1 - \eta]$ in finite time and stay in that interval; we just have to study the behavior of F when $\alpha(0)$ is in $[\alpha_0 + \eta, 1 - \rho_1 - \eta]$. Using equations (3.14) and (3.15), we have that, if $\alpha(0) \in [\alpha_0 + \eta, 1 - \rho_1 - \eta]$, then, for all $t \geq 0$,

$$\begin{aligned} \dot{F}_0(t) &\leq -\mu_0 (f(\alpha_0 + \eta) - \rho_0) < 0, \\ \dot{F}_1(t) &\leq -\mu_1 \eta \end{aligned}$$

and the two components decrease at least linearly to 0 in finite time. There exists T_2 such that for any fluid limit, for all $t \geq T_2$, $F(t) = 0$ and thus $\bar{Z}(t) = 0$. ■

The dynamics of $(F_0(t), F_1(t))$ described in this proof is represented in Figure 3. Theorem 3.6 gives a sufficient condition for stability. The next theorem gives a sufficient condition for transience. The arguments of the proof are very similar.

THEOREM 3.7. *Consider a linear network of L links of capacity 1. If there exists k_0 in $\{0, \dots, L\}$ such that $\rho_{k_0} > 1$ or, if for all k in $\{0, \dots, L\}$, $\rho_k < 1$ and*

$$\rho_0 > \sup_{0 \leq x \leq 1-\rho_1} \bar{\phi}_0(x),$$

then the network is unstable, i.e., $(N(t))$ is transient.

PROOF. In order to prove the transience of the process, we just have to prove that there exists a time T such that after T , any fluid limit $(\bar{Z}(t))$ increases linearly (see [Mey95]).

We consider a fluid limit $(\bar{Z}(t))$. If there is a k_0 such that $\bar{Z}_{k_0}(0) > 0$ then Equations (3.2), (3.3) and (3.4) are valid. It is obvious that, if there exists k_0 such that $\rho_{k_0} > 1$ then $(\bar{Z}_k(t))$ increases linearly and the process $(N(t))$ is transient.

We suppose now that $\rho_k < 1$ for all k in $\{0, \dots, L\}$. In that case, there is a time T_0 such that, for all k in $\{2, \dots, L\}$ and $t \geq T_0$, $\bar{Z}_k(t) = 0$. According to the strong Markov property, we just have to study the fluid limits such that $\bar{Z}_k(0) = 0$ for $2 \leq k \leq L$. In that case, $\bar{Z}_k(t) = 0$ for $2 \leq k \leq L$ and we just have to study the dynamics of $(\bar{Z}_0(t), \bar{Z}_1(t))$.

We define the following function

$$g(\alpha) = \sup_{0 \leq x \leq \alpha} \bar{\phi}_0(x).$$

This function is continuous, non-decreasing and, for any α in $[0, 1]$ satisfies $g(\alpha) \geq \bar{\phi}_0(\alpha)$. We also define the process $(G(t))$ such that $G(0) = \bar{Z}(0)$ and

$$(3.18) \quad \dot{G}_0(t) = \lambda_0 - \mu_0 g\left(\frac{G_0(t)}{G_0(t) + G_1(t)}\right),$$

$$(3.19) \quad \dot{G}_1(t) = \lambda_1 - \mu_1 \frac{G_0(t)}{G_0(t) + G_1(t)}.$$

As previously, we can prove that, for all $t \geq 0$, $G_0(t) \leq \bar{Z}_0(t)$ and $G_1(t) \leq \bar{Z}_1(t)$. So, we just have to prove that the process $(G(t))$ increases linearly to infinity in order to prove the transience of $(N(t))$.

We define $\alpha_0 = \min\{\alpha \in [0, 1], g(\alpha) = \rho_0\}$. Since $\rho_0 > g(1 - \rho_1)$, we have $\alpha_0 > 1 - \rho_1$. Define $\alpha(t) = G_0(t)/(G_0(t) + G_1(t))$.

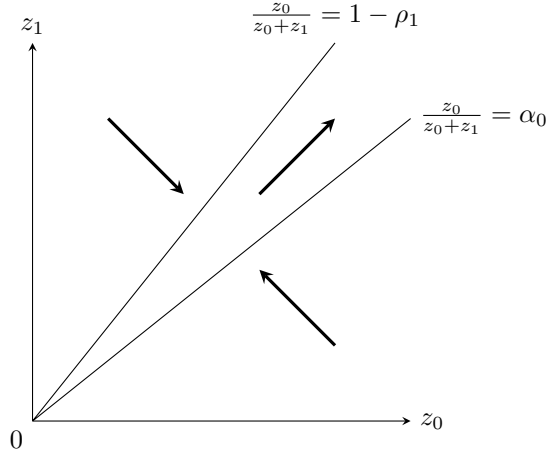
As in the proof of Theorem 3.6, we can show that there exists $\eta > 0$ such that, for any initial state $\alpha(0)$, after some finite time T_1 , $\alpha(t)$ reaches and stays in $[1 - \rho_1 + \eta, \alpha_0 - \eta]$. We then just have to study the dynamics of G when $\alpha(0)$ is in $[1 - \rho_1 + \eta, \alpha_0 - \eta]$.

Using equations (3.18) and (3.19), we have that, if $\alpha(t) \in [1 - \rho_1 + \eta, \alpha_0 - \eta]$, then

$$\begin{aligned} \dot{G}_0(t) &\geq \mu_0 (\rho_0 - g(\alpha_0 - \eta)) > 0, \\ \dot{G}_1(t) &\geq \mu_1 \eta. \end{aligned}$$

Thus, for any initial conditions, for all $t \geq T_1$, $G(t)$ increases linearly to infinity and the theorem is proved. ■

The dynamics of $(G_0(t), G_1(t))$ described in this proof is represented in Figure 4. In particular, we can remark that when the network is unstable, there is always a time after which both classes 0 and 1 of the fluid limit increase linearly to infinity.

FIGURE 4. Dynamics of $(G_0(t), G_1(t))$

REMARK 3.8. The stability region of the network is not reduced to $\{0\}$ since $\inf_{1-\rho_1 \leq x \leq 1} \bar{\phi}_0(x) > 0$ when $\rho_1 < 1$ and it is strictly included in the optimal stability region since $\sup_{0 \leq x \leq 1-\rho_1} \bar{\phi}_0(x) < \min_k(1 - \rho_k)$.

COROLLARY 3.9. Consider a linear network with two links of capacity 1.

This network is stable if $\rho_k < 1$ for all $0 \leq k \leq 2$ and $\rho_0 < \bar{\phi}_0(1 - \rho_1)$.

This network is unstable if there exists k_0 in $\{0, 1, 2\}$ such that $\rho_{k_0} > 1$ or if $\rho_k < 1$ for $0 \leq k \leq 2$ and $\rho_0 > \bar{\phi}_0(1 - \rho_1)$.

PROOF. All that we have to prove is that the function $\bar{\phi}_0$ is strictly increasing.

Consider the functions $\tilde{\phi}_0, \tilde{\phi}_2$, the processes $(\tilde{N}_2^\alpha(t))$ and their stationary distribution $\tilde{\pi}^\alpha$ as defined in Section 3.3. We assume that $\rho_2 < 1$.

The process $(\tilde{N}_2^\alpha(t))$ is one-dimensional and for $\alpha_1 \leq \alpha_2$, for all n_2 , $\tilde{\phi}_k(\alpha_1, n_2) \geq \tilde{\phi}_k(\alpha_2, n_2)$. This implies $(\tilde{N}_2^{\alpha_2}(t)) \geq_{st} (\tilde{N}_2^{\alpha_1}(t))$ and $\tilde{\pi}^{\alpha_2}$ dominates $\tilde{\pi}^{\alpha_1}$.

For any $\alpha \in [0, 1]$ and any $n_2 \in \mathbb{N}$,

$$\tilde{\phi}_0(\alpha, n_2) + \tilde{\phi}_2(\alpha, n_2) = \min(n_2 a_2 + \alpha, 1).$$

The function $n_2 \mapsto \tilde{\phi}_0(\alpha, n_2) + \tilde{\phi}_2(\alpha, n_2)$ is therefore non-decreasing.

Applying the definition of stochastic domination, we have

$$\mathbb{E}_{\tilde{\pi}^{\alpha_1}}(\tilde{\phi}_0(\alpha_1, \cdot) + \tilde{\phi}_2(\alpha_1, \cdot)) \leq \mathbb{E}_{\tilde{\pi}^{\alpha_2}}(\tilde{\phi}_0(\alpha_1, \cdot) + \tilde{\phi}_2(\alpha_1, \cdot)).$$

For any $n_2 \in \mathbb{N}$, we have $\tilde{\phi}_0(\alpha_2, n_2) + \tilde{\phi}_2(\alpha_2, n_2) \geq \tilde{\phi}_0(\alpha_1, n_2) + \tilde{\phi}_2(\alpha_1, n_2)$. Since $\alpha_1 < \alpha_2 \leq 1$, we have $\phi_0(\alpha_2, 0) > \phi_0(\alpha_1, 0)$ and we can conclude

$$\mathbb{E}_{\tilde{\pi}^{\alpha_1}}(\tilde{\phi}_0(\alpha_1, \cdot) + \tilde{\phi}_2(\alpha_1, \cdot)) < \mathbb{E}_{\tilde{\pi}^{\alpha_2}}(\tilde{\phi}_0(\alpha_2, \cdot) + \tilde{\phi}_2(\alpha_2, \cdot)).$$

Using the fact that $\mathbb{E}_{\tilde{\pi}^\alpha}(\tilde{\phi}_2(\alpha, \cdot)) = \rho_2$, we can conclude that if $\alpha_1 < \alpha_2$, then $\tilde{\phi}_0(\alpha_1) < \tilde{\phi}_0(\alpha_2)$. ■

4. Asymptotic Stability

From the results of the previous section, we know that ergodicity conditions in general depend on the access rates a_k . In this section, we qualify this dependence as the access rates become asymptotically small. In the next subsection, we introduce a scaling on the access rates in order to understand the qualitative behavior of networks when the access rates become small. In subsection 4.2, we use this scaling to determine the limit of $\bar{\phi}_0$ when the access rates tend to 0 and then deduce that

linear networks are asymptotically optimal. In subsection 4.3, we use the same scaling to determine the asymptotic optimality of upstream trees. Subsections 4.2 and 4.3 are independent.

4.1. Scaling over the Access Rates and Flow Sizes. We fix all network parameters except the arrival rates, the mean size of flows and the access rates. For each $\beta \in \mathbb{N}$, define the process $(N_\beta(t))$ describing the evolution of the number of flows in each class in a network where the access rate of each class k becomes a_k/β , the arrival rate becomes $\lambda_k\beta$ and the inverse of the mean size $\beta\mu_k$.

In this proof and in the proof in the appendix, we will need the definition of an increasing process of a martingale. If $(M(t))$ is a square-integrable martingale on $\mathbb{R}$ null at 0, $(\langle M \rangle_t)$ is the (essentially unique) increasing process such that $(M^2(t) - \langle M \rangle_t)$ is a martingale. If $(M(t))$ and $(N(t))$ are two square-integrable martingales on $\mathbb{R}$ null at 0, $(\langle M, N \rangle_t)$ is the (essentially unique) increasing process such that $(M(t)N(t) - \langle M, N \rangle_t)$ is a martingale. For more information on martingales and increasing processes, see [RW79].

THEOREM 4.1. *Consider an acyclic network with L links and K classes.*

If

$$\lim_{\beta \rightarrow \infty} \frac{1}{\beta} a \odot N_\beta(0) = \bar{X}(0),$$

then $(\beta^{-1}a \odot N_\beta(t))$ converges in probability uniformly on compact sets to $(\bar{X}(t))$ such that

$$(4.1) \quad \dot{\bar{X}}_k(t) = a_k \lambda_k t - a_k \mu_k \psi_k(\bar{X}(t)), \quad \text{for } 1 \leq k \leq K.$$

The convergence mentioned in this theorem is the uniform convergence in probability on compact sets for the space $D_{\mathbb{R}_+^N}([0, \infty))$ with Skorohod topology (see [Bil99]). The operator $\odot$ denotes the componentwise product.

PROOF. We define the process $(M_\beta(t))$ such that

$$(4.2) \quad M_{\beta,k}(t) = a_k \left(\frac{N_{\beta,k}(t)}{\beta} - \frac{N_{\beta,k}(0)}{\beta} - \lambda_k t + \mu_k \int_0^t \psi_k \left(\frac{1}{\beta} a \odot N_\beta(s) \right) ds \right),$$

for $1 \leq k \leq K$. The martingale characterization (see [RW87]) of the Markov jump process $(\beta^{-1}a \odot N_\beta(t))$ shows that $(M_\beta(t))$ is a martingale and its increasing processes are given by, for $1 \leq k, l \leq L$ and $k \neq l$,

$$(4.3) \quad \langle M_{\beta,k} \rangle_t = \frac{a_k}{\beta} \left(\lambda_k t + \mu_k \int_0^t \psi_k \left(\frac{1}{\beta} a \odot N_\beta(s) \right) ds \right),$$

$$(4.4) \quad \langle M_{\beta,k}, M_{\beta,j} \rangle_t = 0.$$

Since ψ_k is bounded by a constant C , we have $\langle M_{\beta,k} \rangle_t \leq a_k(\lambda_k + C\mu_k)t/\beta$, for $t \geq 0$.

As $(M_\beta(t))$ is a martingale, we can use Doob's inequality. For any $t > 0$ and $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P} \left(\sup_{0 \leq s \leq t} \|M_\beta\|_\infty(s) \geq \varepsilon \right) &\leq \sum_{k=1}^K \mathbb{P} \left(\sup_{0 \leq s \leq t} \|M_{\beta,k}\|_\infty(s) \geq \varepsilon \right) \\ &\leq \frac{K}{\varepsilon^2} \max_k \mathbb{E} (M_{\beta,k}^2(t)), \\ &\leq \frac{K}{\varepsilon^2} \max_k \langle M_{\beta,k} \rangle_t, \\ &\leq \frac{Kt}{\varepsilon^2 \beta} \max_k a_k (\lambda_k + C\mu_k). \end{aligned}$$

Thus, $(M_\beta(t))$ converges to 0 in probability uniformly on any compact when $\beta \rightarrow +\infty$.

We now consider the random variable $\sup_{0 \leq s \leq t} \|\beta^{-1}a \odot N_\beta(t) - \bar{X}(t)\|_\infty$ and show that it converges in distribution to 0. Define $Z_\beta(t) = \beta^{-1}a \odot N_\beta(t) - \bar{X}(t)$. Using (4.2),

$$\begin{aligned} Z_{\beta,k}(t) &= \frac{1}{\beta} a \odot N_{\beta,k}(0) - \bar{X}_k(0) + M_{\beta,k}(t) \\ &\quad - \mu_k \int_0^t \left(\psi_k \left(\frac{1}{\beta} a \odot N_\beta(s) \right) - \psi_k(\bar{X}(s)) \right) ds, \quad \text{for } 1 \leq k \leq K. \end{aligned}$$

The ψ_k are all κ -Lipschitz for some $\kappa > 0$ and, for $1 \leq k \leq K$,

$$\begin{aligned} \sup_{0 \leq s \leq t} |Z_{\beta,k}(s)| &\leq \left| \frac{a_k}{\beta} N_{\beta,k}(0) - \bar{X}_k(0) \right| + \sup_{0 \leq s \leq t} |M_{\beta,k}(s)| \\ (4.5) \quad &\quad + \kappa \mu_k \int_0^t \sup_{0 \leq u \leq s} |Z_\beta(u)| ds. \end{aligned}$$

Define the function

$$f_\beta(t) = \mathbb{E} \left(\sup_{0 \leq s \leq t} \|Z_\beta(s)\|_\infty \right).$$

Using Equations (4.5), (4.3) and (4.4), we deduce the inequality for all $s < t$,

$$f_\beta(s) \leq \left\| \frac{1}{\beta} a \odot N_\beta(0) - \bar{X}(0) \right\|_\infty + \frac{\hat{a}}{\beta} (\hat{\lambda} + \hat{\mu}) s + \hat{\mu} \int_0^s f_\beta(u) du$$

where $\hat{a} = \max_k a_k$, $\hat{\lambda} = \max_k \lambda_k$ and $\hat{\mu} = \max_k \mu_k$.

Applying Gronwall's lemma (see [Gro19]),

$$f_\beta(t) \leq \left(\left\| \frac{1}{\beta} a \odot N_\beta(0) - \bar{X}(0) \right\|_\infty + \frac{\hat{a}}{\beta} (\hat{\lambda} + \hat{\mu}) t \right) e^{\hat{\mu} t}.$$

Thus, $f_\beta \rightarrow 0$ as $\beta \rightarrow +\infty$ and $\sup_{0 \leq s \leq t} \left\| \frac{1}{\beta} a \odot N_\beta(s) - \bar{X}(s) \right\|_\infty$ converges in distribution to 0. In particular, for $\varepsilon > 0$,

$$\lim_{\beta \rightarrow +\infty} \mathbb{P} \left(\sup_{0 \leq s \leq t} \left\| \frac{1}{\beta} a \odot N_\beta(s) - \bar{X}(s) \right\|_\infty \geq \varepsilon \right) = 0.$$

Hence, $(\beta^{-1}a \odot N_\beta(t))$ converges to $(\bar{X}(t))$ in probability uniformly on compact sets. ■

The processes $(\beta^{-1}a \odot N_\beta(t))$ converge in distribution, as β tends to infinity, to a process $(\bar{X}(t))$ which is completely deterministic. This process has a fixed point if and only if there exists $x \in \mathbb{R}_+^K$ such that

$$(4.6) \quad \lambda_k - \mu_k \psi_k(x) = 0 \quad \text{for } 1 \leq k \leq K.$$

PROPOSITION 4.2. *Consider an acyclic network with L links and K classes.*

$(\bar{X}(t))$ admits a unique fixed point if the optimal stability conditions (2.4) are satisfied:

$$\sum_{k:l \in r_k} \rho_k < C_l \quad \text{for } 1 \leq l \leq L.$$

Conversely, if there exists a link l_0 such that

$$\sum_{k:l_0 \in r_k} \rho_k > C_{l_0}$$

then $(\bar{X}(t))$ does not admit a fixed point.

PROOF. By definition of acyclic networks, we suppose that the links are numbered in a such a way that the links on a route of any class are an increasing sequence. By definition, any fixed point $x = (x_1, \dots, x_K)$ satisfies

$$\psi_k(x) = \rho_k, \quad \text{for } 1 \leq k \leq K.$$

We first suppose that the optimal stability conditions (2.4) are satisfied. First, note that if $x = (x_1, \dots, x_K)$ satisfies the optimal stability conditions, then for all k , $\psi_k(x) = x_k$. The traffic intensities $(\rho_1, \dots, \rho_K)$ are then a fixed point and this is the only fixed point which satisfies the optimal stability conditions.

On the contrary, if x is such that there exists l_0 with

$$\sum_{k:l_0 \in r_k} x_k \geq C_{l_0},$$

then, there exists l_1 which is saturated i.e.

$$\sum_{k:l_1 \in r_k} \theta_k^{l_1}(x) = C_{l_1}.$$

Since the network is acyclic, without loss of generality, we can assume that l_1 is such that, for all $l > l_1$, link l is not saturated i.e.

$$\sum_{k:l \in r_k} \theta_k^l(x) < C_l.$$

In that case, for all class k such that $l \in r_k$, we have $\psi_k(x) = \theta_k^{l_1}(x)$ and

$$\begin{aligned} \sum_{k:l_1 \in r_k} \psi_k(x) &= C_{l_1} \\ &> \sum_{k:l_1 \in r_k} \rho_k. \end{aligned}$$

The vector x cannot be a fixed point and the uniqueness of the fixed point is proved.

We now suppose that there exists a link l_0 such that

$$\sum_{k:l_0 \in r_k} \rho_k > C_{l_0}.$$

In that case, it is enough to remark that, for all $x = (x_1, \dots, x_K)$,

$$\begin{aligned} \sum_{k:l_0 \in r_k} \psi_k(x) &\leq C_{l_0} \\ &< \sum_{k:l_0 \in r_k} \rho_k. \end{aligned}$$

There is no fixed point. ■

This proposition brings the intuition that the stability conditions for any acyclic network can be arbitrarily close to the optimal ones if the access rates are small enough. In the remainder of this section, we prove that this intuition is true in the case of linear networks and upstream trees.

4.2. Asymptotic Stability of Linear Networks. We return here to the linear networks discussed in Section 3. Theorems 3.6 and 3.7 show that stability conditions depend on the function $\bar{\phi}_0$ and consequently on stationary distributions $(\bar{\pi}^\alpha)$. In order to characterize the ergodicity conditions, we have to study the evolution of these distributions when the access rates decrease to 0.

Consider the process $(\tilde{N}_{2:L}^\alpha(t))$ introduced in Section 3.3 for $\alpha \in [0, 1]$ under the scaling defined in Section 4.1. We call $(\tilde{N}_{2:L,\beta}^\alpha(t))$ the scaled process whose transition rates are given by, for $2 \leq k \leq L$,

$$\begin{aligned} n_k &\rightarrow n_k + 1 : \beta \lambda_k, \\ n_k &\rightarrow n_k - 1 : \beta \mu_k \tilde{\psi}_k(n \odot \beta^{-1} a) \end{aligned}$$

where the functions $\tilde{\psi}_k$ are defined by Equation (3.6). We suppose that the ergodicity conditions of this process are fulfilled, i.e., $\rho_k < 1$ for $2 \leq k \leq L$, and denote by $\bar{\pi}_\beta^\alpha$ the stationary distribution of $(\beta^{-1} a_{2:L} \odot \tilde{N}_{2:L,\beta}^\alpha(t))$.

As in Section 4.1, when β tends to infinity, the processes $(\beta^{-1} a_{2:L} \odot \tilde{N}_{2:L,\beta}^\alpha(t))$ converge in distribution to a deterministic limit that we call $(\bar{X}^\alpha(t))$. This process has a unique fixed point that we call $\gamma(\alpha)$ which is the solution of the following equations:

$$\lambda_k - \mu_k \tilde{\psi}_k(\alpha, x_{2:L}) = 0 \text{ for } 2 \leq k \leq L.$$

Using the definition of $\tilde{\psi}_k$, $\gamma(\alpha)$ can be built recursively yielding:

$$(4.7) \quad \gamma_k(\alpha) = \max \left(\rho_k, \frac{\rho_k}{1 - \rho_k} \min \left(\alpha, \min_{2 \leq j \leq k-1} (1 - \rho_j) \right) \right) \text{ for } 2 \leq k \leq L.$$

The next proposition states the convergence of the stationary distributions $\bar{\pi}_\beta^\alpha$ to the Dirac measure at $\gamma(\alpha)$ for all $\alpha \in [0, 1]$ and the convergence of $\bar{\phi}_0$.

PROPOSITION 4.3. *Consider a linear network with L links in the quasi stationary regime. The set $\{\bar{\pi}_\beta^\alpha, \alpha \in [0, 1], \beta \in \mathbb{N}\}$ is tight. For any $\alpha \in [0, 1]$, the stationary distribution $(\bar{\pi}_\beta^\alpha)$ converges to $\delta_{\gamma(\alpha)}$ when $\beta \rightarrow +\infty$.*

If $\bar{\phi}_{0,\beta}$ is the function such that

$$(4.8) \quad \bar{\phi}_{0,\beta}(\alpha) = \int_{\mathbb{R}^{L-1}} \tilde{\psi}_0(\alpha, x_{2:L}) \bar{\pi}_\beta^\alpha(dx_{2:L}), \quad \text{for } \alpha \in [0, 1],$$

then $\bar{\phi}_{0,\beta}$ converges pointwise on $[0, 1]$ to

$$(4.9) \quad \bar{\phi}_{0,\infty} : \alpha \mapsto \min \left(\alpha, \min_{2 \leq k \leq L} (1 - \rho_k) \right), \quad \text{when } \beta \rightarrow +\infty.$$

Moreover, for all $\alpha \in [0, 1]$,

$$\lim_{\beta \rightarrow +\infty} \inf_{\alpha \leq u \leq 1} \bar{\phi}_{0,\beta}(u) = \bar{\phi}_{0,\infty}(\alpha).$$

The tightness and the convergence mentioned in this proposition is the convergence of probability distributions on $\mathbb{R}_+^{K-2}$ with the usual topology.

PROOF. First, we have to prove that the set $\{\tilde{\pi}_\beta^\alpha, \alpha \in [0, 1], \beta \in \mathbb{N}\}$ is tight.

Note that for $\alpha \in [0, 1]$ and for $k \geq 2$,

$$\tilde{\psi}_k(\alpha, \beta x_{2:L}) \geq \frac{n_k a_k}{\beta + n_k a_k}.$$

Define the Markov processes $(\hat{N}_{2:L,\beta}(t))$ with transition rates:

$$\begin{aligned} n_k &\rightarrow n_k + 1 : \lambda_k \beta, \\ n_k &\rightarrow n_k - 1 : \mu_k \beta \frac{n_k a_k}{\beta + n_k a_k}. \end{aligned}$$

We assume that $\hat{N}_{2:L,\beta}(0) = \tilde{N}_{2:L,\beta}^\alpha(0)$. We then have $(\beta^{-1} a_{2:L} \odot \hat{N}_{2:L,\beta}(t)) \geq_{\text{st}} (\beta^{-1} a_{2:L} \odot \tilde{N}_{2:L,\beta}^\alpha(t))$. In order to prove that the set $\{\tilde{\pi}_\beta^\alpha, \alpha \in [0, 1], \beta \in \mathbb{N}\}$ is relatively compact, it is enough to prove that $\{\hat{\pi}_\beta, \beta \in \mathbb{N}\}$ is tight where $\hat{\pi}_\beta$ is the invariant measure of $(\beta^{-1} a_{2:L} \odot \hat{N}_{2:L,\beta}(t))$. These are birth-death processes whose components are independent. Their invariant measures thus have a product form and it is sufficient to study the one-dimensional case:

$$\hat{\pi}_\beta(n_2 \beta a_2) = (1 - \rho_2)^{\beta a_2^{-1} + 1} \rho_2^{n_2} \frac{\Gamma\left(\frac{\beta}{a_2} + 1 + n_2\right)}{\Gamma\left(\frac{\beta}{a_2} + 1\right) n_2!}$$

where Γ is the usual gamma function

$$\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt, \quad \text{for } z > 0.$$

$\hat{\pi}_\beta$ is a negative binomial distribution and we have

$$\hat{\pi}_\beta([A, +\infty[) = \frac{\int_0^1 t^{\beta a_2^{-1}} (1-t)^{\lfloor \beta A/a_2 \rfloor} dt}{\int_0^1 t^{\beta a_2^{-1}} (1-t)^{\lfloor \beta A/a_2 \rfloor} dt}.$$

Thus, if $A > \rho_2/(1 - \rho_2)$,

$$(4.10) \quad \lim_{\beta \rightarrow +\infty} \hat{\pi}_\beta([A, +\infty[) = 0.$$

Fix $\varepsilon > 0$ and $A > \rho_2/(1 - \rho_2)$. According to (4.10), there exists $\beta_0 \in \mathbb{N}$ such that for all $\beta \geq \beta_0$, $\hat{\pi}_\beta([0, A]) \geq 1 - \varepsilon$. Additionally, there exists a compact $\Gamma_1 \subset \mathbb{R}$ such that for $\beta < \beta_0$, $\hat{\pi}_\beta(\Gamma_1) \geq 1 - \varepsilon$. By choosing $\Gamma = [0, A] \cup \Gamma_1$, we have $\hat{\pi}_\beta(\Gamma) \geq 1 - \varepsilon$ for all $\beta \in \mathbb{N}$ and the set $\{\hat{\pi}_\beta, \beta \in \mathbb{N}\}$ is tight on $\mathbb{R}_+$ with the usual topology. Finally, the set $\{\tilde{\pi}_\beta^\alpha, \alpha \in [0, 1], \beta \in \mathbb{N}\}$ is tight and, thus, relatively compact.

We now prove the convergence of the invariant measures. The infinitesimal generator $\tilde{\Omega}_\beta^\alpha$ of $(\beta^{-1} a_{2:L} \odot \tilde{N}_{2:L,\beta}^\alpha(t))$ is defined by

$$\begin{aligned} \tilde{\Omega}_\beta^\alpha(f)(x_{2:L}) &= \sum_{k=2}^L \beta \lambda_k \left(f\left(x_{2:L} + \frac{a_k}{\beta} e_k\right) - f(x_{2:L}) \right) \\ &\quad + \sum_{k=2}^L \beta \mu_k \tilde{\psi}_k(\alpha, x_{2:L}) \left(f\left(x_{2:L} - \frac{a_k}{\beta} e_k\right) - f(x_{2:L}) \right). \end{aligned}$$

The “infinitesimal generator” $\bar{\Omega}^\alpha$ of $(\bar{X}^\alpha(t))$ is defined by

$$\bar{\Omega}^\alpha(f)(x_{2:L}) = \sum_{k=2}^L \lambda_k \frac{\partial f}{\partial x_k}(x_{2:L}) - \sum_{k=2}^L \mu_k \tilde{\psi}_k(\alpha, x_{2:L}) \frac{\partial f}{\partial x_k}(x_{2:L}).$$

Let f be a bounded smooth function with a bounded derivative. Since $\{\tilde{\pi}_\beta, \beta \in [0, 1], \alpha \in [0, 1]\}$ is tight, for any $\varepsilon > 0$, there exists a compact set $\Gamma \subset \mathbb{R}^{K-2}$ such

that for all β in $\mathbb{N}$ and α in $[0, 1]$,

$$(4.11) \quad \int_{\mathbb{R}_+^{K-2} \setminus \Gamma} f(x_{2:L}) \tilde{\pi}_\beta^\alpha(dx_{2:L}) \leq \varepsilon.$$

Since f and its derivative are bounded and the $\tilde{\psi}_k$ are γ -Lipschitz for some γ , there exists η such that for x in Γ ,

$$(4.12) \quad |\tilde{\Omega}_\beta^\alpha(f)(x_{2:L}) - \bar{\Omega}^\alpha(f)(x_{2:L})| \leq \frac{\eta}{\beta}.$$

The equilibrium equations give, for all α ,

$$(4.13) \quad \int_{\mathbb{R}_+^{K-2}} \tilde{\Omega}_\beta^\alpha(f)(x_{2:L}) \tilde{\pi}_\beta^\alpha(dx_{2:L}) = 0.$$

Since $\{\tilde{\pi}_\beta^\alpha, \beta \in \mathbb{N}, \alpha \in [0, 1]\}$ is relatively compact, we can extract a convergent subsequence $(\tilde{\pi}_{\beta_i}^{\alpha_i})$ such that $\alpha_i \rightarrow \alpha$ and $\beta_i \rightarrow +\infty$. Thanks to relations (4.11), (4.12), (4.13) and the continuity of f and its derivative, we can write that

$$\lim_{i \rightarrow +\infty} \int_{\mathbb{R}_+^{K-2}} \tilde{\Omega}_{\beta_i}^{\alpha_i}(f)(x_{2:L}) \tilde{\pi}_{\beta_i}^{\alpha_i}(dx_{2:L}) = \int_{\mathbb{R}_+^{K-2}} \bar{\Omega}^\alpha(f)(x_{2:L}) \bar{\pi}^\alpha(dx_{2:L}).$$

Finally, $(\tilde{\pi}_{\beta_i}^{\alpha_i})$ converges to $\bar{\pi}^\alpha$ when $\beta \rightarrow +\infty$.

The convergence pointwise of $\bar{\phi}_{0,\beta}$ follows by choosing $\alpha_i = \alpha$.

Finally, for the last part of the proposition, we consider a sequence (α_i, β_i) such that $\alpha_i \rightarrow \alpha_\infty$, $\beta_i \rightarrow +\infty$ when $i \rightarrow \infty$ and, for all i , $\bar{\phi}_{0,\beta_i}(u_i) = \inf_{\alpha \leq u \leq 1} \bar{\phi}_{0,\beta_i}(u)$. We have that

$$\lim_{i \rightarrow \infty} \bar{\phi}_{0,\beta_i}(\alpha_i) = \bar{\phi}_{0,\infty}(\alpha_\infty).$$

By construction of α_i , we also have

$$\lim_{i \rightarrow \infty} \bar{\phi}_{0,\beta_i}(\alpha_i) \leq \inf_{\alpha \leq u \leq 1} \bar{\phi}_{0,\infty}(u).$$

Since $\bar{\phi}_{0,0}$ is a non-decreasing function, it implies that

$$\lim_{i \rightarrow \infty} \bar{\phi}_{0,\beta_i}(\alpha_i) = \bar{\phi}_{0,\infty}(\alpha).$$

■

We finally deduce the following theorem which is a consequence of Theorem 3.6 and Proposition 4.3.

THEOREM 4.4. *In a linear network with L links, for any traffic intensities satisfying the optimal stability conditions (3.1), if the access rates are small enough, the resulting stochastic process is ergodic.*

PROOF. We consider the scaled version of the process introduced in Subsection 4.1 ($N_\beta(t)$). According to Theorem 3.6, a sufficient condition of stability is

$$\begin{aligned} \rho_k &< 1, \quad \text{for } 0 \leq k \leq L, \\ \rho_0 &< \inf_{1-\rho_1 \leq \alpha \leq 1} \bar{\phi}_{0,\beta}(\alpha). \end{aligned}$$

Since the traffic intensities $(\rho_0, \dots, \rho_L)$ satisfy the optimal stability conditions (3.1), there exists $\varepsilon > 0$ such that

$$\begin{aligned} \rho_k &< 1, \quad \text{for } 0 \leq k \leq L, \\ \rho_0 &< \min_{1 \leq k \leq L} (1 - \rho_k) - \varepsilon. \end{aligned}$$

According to Proposition 4.3, there exists β_0 such that for all $\beta \geq \beta_0$, we have

$$\inf_{1-\rho_1 \leq \alpha \leq 1} \bar{\phi}_{0,\beta}(\alpha) \geq \min_{1 \leq k \leq L} (1 - \rho_k) - \varepsilon/2.$$

This implies that, for all $\beta \geq \beta_0$, the process $(N_\beta(t))$ is ergodic.

Finally, to conclude, it is enough to remark that the process with transition rates

$$\begin{aligned} n_k \rightarrow n_k + 1 : & \quad \lambda_k, \\ n_k \rightarrow n_k - 1 : & \quad \mu_k \psi_k(n \odot \beta^{-1} a) \end{aligned}$$

admits the same stationary measures as $(N_\beta(t))$ and is then ergodic if and only if $(N_\beta(t))$ is ergodic. ■

This result means that the Tail Dropping policy is asymptotically optimal in linear networks. Figures 5 and 6 represent the stability region for classes 0 and 1, obtained by simulation, for two particular networks and illustrate this result.

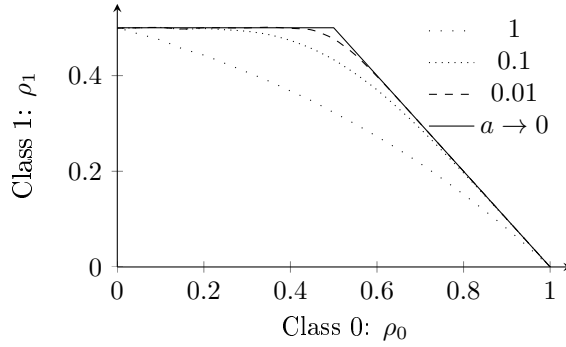


FIGURE 5. Stability region for classes 0 and 1 when $L = 2$ and $\rho_2 = 0.5$

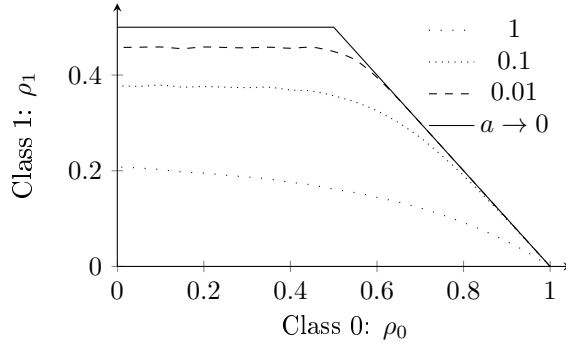


FIGURE 6. Stability region for classes 0 and 1 when $L = 4$ and $\rho_2, \rho_3, \rho_4 = 0.5$

4.3. Asymptotic Stability of Upstream Trees. Upstream trees are a specific class of networks because they are monotonic and this will allow us to use stochastic domination to establish the asymptotic stability of upstream trees. Thanks to the monotonicity, the worst case, in the sense of stochastic domination, for a given class is when there is an infinite number of flows in all other classes. When the optimal stability conditions are satisfied, we can prove that there exists a class k_0 which can get enough bandwidth to be stable even if there is an infinite number of flows in all other classes. Thanks to the same scaling that we used in subsections 4.1 and 4.2, we are able to quantify the bandwidth used by the class k_0 when the access rates are decreasing to 0 and to prove that there exists a class k_1 which is stable when

the access rates are small enough even if there is an infinite number of flows in all classes except k_0 and k_1 . We conclude by using a recursion.

In upstream trees, when two different classes go through the same link, they share the same path until they exit the network. The last link on the path of all classes is the same and it is called the root. Formally, upstream trees are defined as follows:

- (i): a common root link, say 1, so that $r_k(d_k) = 1$ for all classes k ;
- (ii): for any two classes j and k , there exists m such that they only have their last m links in common; i.e, for i in $\{0, \dots, m-1\}$, $r_j(d_j - i) = r_k(d_k - i)$ and for i_k in $\{0, \dots, d_k - m\}$ and i_j in $\{0, \dots, d_j - m\}$, $r_j(i_j) \neq r_k(i_k)$.

Two examples of upstream trees are given in Figure 7. In [BFP09], two very specific examples were shown not to be ergodic under optimal stability conditions except when the access rates tend to 0. We generalize this result.

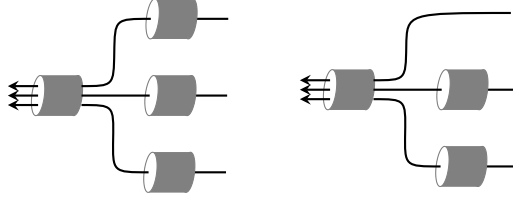


FIGURE 7. Examples of upstream trees

For any link l , let S_l denote the set of children of l , the links just *before* l on the path of some class:

$$S_l = \{j \in \{1, \dots, L\}, \exists k \in \{1, \dots, K\}, \exists i, r_k(i) = j, r_k(i+1) = l\}.$$

In particular, S_1 is the set of links which are the links just before the root link on the path of all classes with a route of length longer than 1.

We suppose that there are no two classes with exactly the same path. We also suppose without loss of generality that all the links of the network can be saturated. For any link l , this implies that there is a class k_l directly entering in the network through that link, i.e., $r_{k_l}(1) = l$ or $\sum_{j \in S_l} C_j > C_l$. If a link l_0 cannot be saturated then the bandwidth allocation is the same with or without l_0 and we can remove it.

Upstream trees are monotonic in the sense that if $n \in \mathbb{N}^K$ and $m \in \mathbb{N}^K$ and $m_k = n_k$ for a given k and $m_j \geq n_j$ for $j \neq k$ then $\phi_k(m) \leq \phi_k(n)$. Thanks to this monotonicity property, the processes we study fit into the framework of [BJL08]. We suppose that the traffic intensities satisfy the optimal stability conditions (2.4).

Because of the monotonicity of this network, we know that the worst case for class k_0 is when the numbers of flows in all other classes increase to infinity. We define the following allocation:

$$(4.14) \quad \forall x_{k_0} \in \mathbb{R}_+, \check{\psi}_{k_0}^1(x_{k_0}) = \lim_{\eta \rightarrow \infty} \inf_{\substack{y: y_{k_0} = x_{k_0}, \\ k \neq k_0, y_k > \eta}} \psi_{k_0}(y)$$

and $\check{\phi}_{k_0}^1$ is such that $\check{\phi}_{k_0}^1(n_{k_0}) = \check{\psi}_{k_0}^1(a_{k_0} n_{k_0})$ for all $n_{k_0} \in \mathbb{N}$. We then define the Markov process $(\check{N}_{k_0}^1(t))$ such that $\check{N}_{k_0}^1(0) = N_{k_0}(0)$ with the transition rates

$$\begin{aligned} n_{k_0} &\rightarrow n_{k_0} + 1 : \lambda_{k_0}, \\ n_{k_0} &\rightarrow n_{k_0} - 1 : \mu_{k_0} \check{\phi}_{k_0}^1(n_{k_0}). \end{aligned}$$

Because of the monotonicity of the considered allocations, we have that the process $(\check{N}_{k_0}^1(t))$ stochastically dominates $(N_{k_0}(t))$ (see [BJL08]).

In the sense of the stochastic domination, this process represents the worst case for class k_0 . The next lemma shows that there is a class which is stable even in the worst case.

LEMMA 4.5. *Consider an upstream tree and suppose the traffic intensities of this network satisfy the optimal stability conditions (2.4). There exists a class k_0 such that the Markov process $(\check{N}_{k_0}^1(t))$ is ergodic.*

PROOF. Because of the monotonicity of the network all that we have to prove is that there exists a class k_0 such that

$$(4.15) \quad \lim_{\eta \rightarrow +\infty} \inf_{n_1 \geq \eta, \dots, n_K \geq \eta} \phi_{k_0}(n) > \rho_{k_0}.$$

If k_0 satisfies (4.15) then there exists $\varepsilon > 0$ and η_0 such that, for all $n_{k_0} \geq \eta_0$,

$$\lambda_{k_0} - \mu_{k_0} \check{\phi}_{k_0}^1(n_{k_0}) \leq -\varepsilon.$$

This implies the ergodicity of $(\check{N}_{k_0}^1(t))$. We proceed by recursion on the depth of the tree.

If the depth is 1 then there is a single class and a single link and condition (4.15) is obviously true.

We suppose now that the depth of the tree is $d \geq 2$. If there is a class k_0 such that $r_{k_0}(1) = 1$ then the root link is the entry point of k_0 which is the only one entering the network through 1. The worst case for class k_0 is when all links in S_1 are saturated; we have for all $n \in \mathbb{N}^K$,

$$\phi_{k_0}(n) \geq \min \left(a_{k_0} n_{k_0}, C_1 \frac{a_{k_0} n_{k_0}}{\sum_{l \in S_1} C_l + a_{k_0} n_{k_0}} \right).$$

Because the optimal conditions are satisfied, we have $\rho_{k_0} < C_1$ and there exist $\delta > 0$ and η_0 such that for all $n \in \mathbb{N}^K$ with $n_{k_0} \geq \eta_0$, we have

$$\phi_{k_0}(n) \geq \rho_{k_0} + \delta$$

which is enough to conclude.

We now suppose that there is no class directly connected to the root, i.e. there is no class k with a path of length 1. This implies that S_1 is not empty. Since the root link can be saturated, we have $\sum_{l \in S_1} C_l > C_1$. Due to optimal stability conditions, we have

$$\sum_{k=1}^K \rho_k < C_1$$

and we can deduce that there is a link l_0 in S_1 such that

$$\sum_{k: r_k(d_k-1)=l_0} \rho_k < \frac{C_{l_0}}{\sum_{j \in S_1} C_j} C_1.$$

Since the link l_0 can be saturated, there is a constant η_0 such that, if for any class k going through l_0 (i.e. satisfying $r_k(d_k-1) = l_0$), we have $n_k > \eta_0$, we have

$$\sum_{k: r_k(d_k-1)=l_0} \theta_k^{d_k-1}(n \odot a) = C_{l_0}.$$

Using the monotonicity of upstream trees, we can deduce that the worst case for classes going through l_0 is when all the other links in S_1 are saturated. We deduce immediately that when $n_k > \eta_0$ for all k going through l_0 , we have

$$\phi_k(n) \geq \frac{C_1}{\sum_{j \in S_1} C_j} \theta_k^{d_k-1}(n \odot a).$$

It remains to consider the case where all the links in S_1 are saturated corresponding to the case of equality in the previous equation. With these conditions, the subtree whose root link is l_0 is equivalent to an upstream tree where the capacity of link l_0 is replaced by $C_1 C_{l_0} / (\sum_{j \in S_1} C_j)$.

If there is a class k_0 such that $d_{k_0} = l_0$ then, as previously, we have

$$\lim_{\eta \rightarrow \infty} \inf_{n: n_k \geq \eta} \phi_{k_0}(n) \geq \frac{C_1 C_{l_0}}{\sum_{j \in S_1} C_j} > \rho_{k_0}$$

and k_0 satisfies (4.15).

If there is no class k_0 such that $d_{k_0} = l_0$ then S_{l_0} is not empty and, as previously, we can find l_1 such that

$$\sum_{k: r_k(d_k-2)=l_1} \rho_k < C_{l_1} \frac{C_{l_0}}{\sum_{j \in S_{l_0}} C_j} \frac{C_1}{\sum_{j \in S_1} C_j}.$$

Using a recursion, we finally find k_0 satisfying (4.15). In addition, the links on the path of k_0 satisfy

$$(4.16) \quad \forall i \in \{1, \dots, d_{k_0} - 1\}, \quad \sum_{k: r_{k_0}(i) \in r_k} \rho_k < C_{r_{k_0}(i)} \prod_{j=i+1}^{d_{k_0}} \frac{C_{r_{k_0}(j)}}{\sum_{l \in S_{r_{k_0}(j)}} C_l}.$$

■

We now perform the scaling of Section 4.1 on the process $(\check{N}_{k_0}^1(t))$ and call the scaled process $(\check{N}_{k_0, \beta}^1(t))$. We deduce from the above lemma that the scaled process is ergodic for all $\beta \in \mathbb{N}$. For each β , we call $\check{\pi}_{k_0, \beta}^1$ the stationary distribution of $(\beta^{-1} a_{k_0} \check{N}_{k_0, \beta}^1(t))$. As in Section 4.1, the process $(\beta^{-1} a_{k_0} \check{N}_{k_0, \beta}^1(t))$ converges in probability uniformly on compact sets to a deterministic process $(\check{X}_{k_0}^1(t))$. The next lemma shows that there is also convergence for the stationary distribution. The proof of this lemma is similar to that of Proposition 4.3 and is omitted. The tightness and convergence in distribution mentioned in this lemma are on the space $\mathbb{R}_+$ with its usual topology.

LEMMA 4.6. *Consider an upstream tree and suppose that the optimal stability conditions (2.4) are satisfied by the traffic intensities. Consider k_0 such that $(\check{N}_{k_0}^1(t))$ is ergodic. The set $\{\check{\pi}_{k_0, \beta}^1, \beta \in \mathbb{N}\}$ is tight and $\check{\pi}_{k_0, \beta}^1$ converges to $\delta_{\alpha_{k_0}^1}$ where $\alpha_{k_0}^1$ is the unique solution of the following equation*

$$(4.17) \quad \lambda_{k_0} - \mu_{k_0} \check{\psi}_{k_0}^1(\alpha_{k_0}^1) = 0.$$

The function $\check{\psi}_{k_0}^1$ is defined by Equation (4.14).

Now, we proceed by recursion. First, we give a class k_1 which is also stable when the access rates of classes k_0 and k_1 are small enough whatever the state of other classes. For any classes k_0 and k_1 , define the following allocation

$$(4.18) \quad \forall (x_{k_0}, x_{k_1}) \in \mathbb{R}_+^2, \quad \check{\psi}_{k_0}^2(x_{k_0}, x_{k_1}) = \lim_{\eta \rightarrow \infty} \inf_{\substack{y: y_{k_0} = x_{k_0} \\ y_{k_1} = x_{k_1} \\ k \neq k_i, y_k \geq \eta}} \psi_{k_0}(x)$$

and $\check{\psi}_{k_1}^2$ is defined similarly. Now define $\check{\phi}_{k_0}^2$ and $\check{\phi}_{k_1}^2$ such that $\check{\phi}_{k_i}^2(n_{k_0}, n_{k_1}) = \check{\psi}_{k_i}^2(a_{k_0} n_{k_0}, a_{k_1} n_{k_1})$ for all (n_{k_0}, n_{k_1}) in $\mathbb{N}^2$. Let $(\check{N}_{k_0}^2(t), \check{N}_{k_1}^2(t))$ be the Markov process with the transition rates for $i = 0, 1$:

$$\begin{aligned} n_{k_i} &\rightarrow n_{k_i} + 1 : \lambda_{k_i}, \\ n_{k_i} &\rightarrow n_{k_i} - 1 : \mu_{k_i} \check{\phi}_{k_i}^2(n_{k_0}, n_{k_1}), \end{aligned}$$

and such that $(\check{N}_{k_0}^2(0), \check{N}_{k_1}^2(0)) = (N_{k_0}(0), N_{k_1}(0))$. Using the monotonicity of the considered allocations, we know that the process $(\check{N}_{k_0}^2(t), \check{N}_{k_1}^2(t))$ stochastically dominates $(N_{k_0}(t), N_{k_1}(t))$. Again, let $(\check{N}_{k_0,\beta}^2(t), \check{N}_{k_1,\beta}^2(t))$ denote the scaled version of $(\check{N}_{k_0}^2(t), \check{N}_{k_1}^2(t))$.

LEMMA 4.7. *Consider an upstream tree and assume that the optimal stability conditions (2.4) are satisfied by the traffic intensities of this tree. Consider k_0 as defined in Lemmas 4.5 and 4.6.*

There exist k_1 and $\beta_1 \in \mathbb{N}$ such that, for all $\beta \geq \beta_1$, the Markov process $(\check{N}_{k_0,\beta}^2(t), \check{N}_{k_1,\beta}^2(t))$ is ergodic. When $\beta \geq \beta_1$, the Markov process

$$(\beta^{-1}a_{k_0}\check{N}_{k_0,\beta}^2(t), \beta^{-1}a_{k_1}\check{N}_{k_1,\beta}^2(t))$$

has a unique stationary distribution that we denote by $\check{\pi}_{k_0,k_1,\beta}^2$. The set

$$\{\check{\pi}_{k_0,k_1,\beta}^2, \beta \in \mathbb{N}\}$$

is tight and the stationary distribution $\check{\pi}_{k_0,k_1,\beta}^2$ converges to $\delta_{(\alpha_{k_0}^2, \alpha_{k_1}^2)}$ where $(\alpha_{k_0}^2, \alpha_{k_1}^2)$ is the unique solution of the equations

$$(4.19) \quad \begin{cases} \lambda_{k_0} - \mu_{k_0}\check{\psi}_{k_0}^2(\alpha_{k_0}^2, \alpha_{k_1}^2) = 0, \\ \lambda_{k_1} - \mu_{k_1}\check{\psi}_{k_1}^2(\alpha_{k_0}^2, \alpha_{k_1}^2) = 0. \end{cases}$$

The functions $\check{\psi}_{k_0}^2$ and $\check{\psi}_{k_1}^2$ are defined by Equation (4.18).

The tightness and convergence in distribution mentioned in this lemma are on the space $\mathbb{R}_+^2$ with its usual topology.

PROOF. First, we prove that there exists k_1 such that

$$(4.20) \quad \lim_{x_{k_1} \rightarrow \infty} \check{\psi}_{k_1}^2(\alpha_{k_0}^1, x_{k_1}) > \rho_{k_1}$$

where $\alpha_{k_0}^1$ is the unique solution of (4.17).

For all i in $\{1, \dots, d_{k_0}\}$, We define

$$\check{\tau}_{k_0}^i(x_{k_0}) = \lim_{\eta \rightarrow \infty} \inf_{\substack{y: y_{k_0} = x_{k_0} \\ k \neq k_0, y_k \geq \eta}} \theta_{k_0}^i(y).$$

Using Lemma 4.6, we can prove

$$\sigma_i = \check{\tau}_{k_0}^i(\alpha_{k_0}^1) = \lim_{\beta \rightarrow +\infty} \mathbb{E}_{\check{\pi}_{k_0,\beta}^1} \left(\check{\tau}_{k_0}^i \left(\beta^{-1}a_{k_0}\check{N}_{k_0,\beta}^1 \right) \right)$$

The quantity σ_i is the bandwidth used asymptotically by class k_0 in link $r_{k_0}(i)$ when all the other classes are saturated and when the access rate a_{k_0} decreases to 0. We want to prove that when a_{k_0} is small enough, class k_0 leaves enough bandwidth for other classes. First, we have the following result which states that the optimal stability conditions are met by all classes except k_0 when we remove the bandwidth asymptotically used by class k_0 in the saturated case:

$$\forall i \in \{1, \dots, d_{k_0}\}, \quad \sum_{k: r_{k_0}(i) \in r_k} \rho_k < C_{r_{k_0}(i)} - \sigma_i.$$

Indeed, for link 1, note that $\sigma_{d_{k_0}} = \rho_{k_0}$ because $\alpha_{k_0}^1$ is the solution of Equation (4.17) and the previous inequality is obviously true for link 1. If the route of k_0 is of length 1 the lemma is proved.

Suppose that the route of k_0 is at least of length 2. By construction of k_0 , there is no class entering the network through $r_{k_0}(i)$ for i in $\{2, \dots, d_{k_0}\}$. Since any link can be saturated, this implies

$$\forall i \in \{2, \dots, d_{k_0}\}, C_{r_{k_0}(i)} < \sum_{l \in S_{r_{k_0}(i)}} C_l.$$

Moreover, we can prove that

$$\forall i \in \{1, \dots, d_{k_0} - 1\}, \sigma_i = \prod_{j=i+1}^{d_{k_0}} \frac{\sum_{l \in S_{r_{k_0}(j)}} C_l}{C_{r_{k_0}(j)}} \rho_{k_0}.$$

Using these last two equations and Equation (4.16), we deduce

$$\forall i \in \{1, \dots, d_{k_0} - 1\}, \sum_{k: r_{k_0}(i) \in r_k} \rho_k < C_{r_{k_0}(i)} - \sigma_i.$$

Considering, for all $k \neq k_0$, the allocation

$$(x_0, \dots, x_{k_0-1}, x_{k_0+1}, \dots, x_K) \mapsto \psi_k(x_0, \dots, \alpha_{k_0}^1, \dots, x_K)$$

and using the same procedure used to find k_0 , we can find k_1 such that

$$\lim_{\eta \rightarrow \infty} \inf_{\substack{k \neq k_0, x_k \geq \eta \\ x_{k_0} = \alpha_{k_0}^1}} \psi_{k_1}(x) > \rho_{k_1}$$

which is equivalent to Equation (4.20).

We are now able to prove that there exists $\beta_1 \in \mathbb{N}$ such that for all $\beta \geq \beta_1$, the Markov process $(\check{N}_{k_0, \beta}^2(t), \check{N}_{k_1, \beta}^2(t))$ is ergodic. By construction of $(\check{N}_{k_0, \beta}^1(t))$ and $(\check{N}_{k_0, \beta}^2(t), \check{N}_{k_1, \beta}^2(t))$, according to Theorem 2 of [BJL08], we just have to prove that there exists $\beta_1 \in \mathbb{N}$ such that, for all $\beta \geq \beta_1$,

$$(4.21) \quad \rho_{k_1} < \mathbb{E}_{\check{\pi}_{k_0, \beta}^1} \left(\liminf_{x_{k_1} \rightarrow \infty} \check{\psi}_{k_1}^2 \left(\beta^{-1} a_{k_0} \check{N}_{k_0, \beta}^1(0), x_{k_1} \right) \right).$$

We deduce from (4.20) that there exists $b > \alpha_{k_0}^1$, $\varsigma > 0$ and η_0 such that

$$(4.22) \quad \forall x_{k_0} \leq b, \forall x_{k_1} \geq \eta_0, \check{\psi}_{k_1}^2(x_{k_0}, x_{k_1}) \geq \rho_{k_1} + \varsigma.$$

We know that $\check{\pi}_{k_0, \beta}^1$ converges to $\delta_{\alpha_{k_0}^1}$ when $\beta \rightarrow +\infty$. Thus, there exists $\beta_1 \in \mathbb{N}$, such that, for all $\beta \geq \beta_1$,

$$(4.23) \quad \check{\pi}_{k_0, \beta}^1([b, \infty]) \leq \frac{\varsigma}{2(\rho_{k_1} + \varsigma)}.$$

Finally, we have that, for all $\beta \geq \beta_1$, Equation (4.21) is satisfied and the process $(\check{N}_{k_0, \beta}^2(t), \check{N}_{k_1, \beta}^2(t))$ is ergodic.

We now prove tightness. By construction, we have $(\check{N}_{k_0, \beta}^1(t)) \geq_{\text{st}} (\check{N}_{k_0, \beta}^2(t))$ and we deduce immediately that

$$(4.24) \quad \forall c \in \mathbb{R}, \check{\pi}_{k_0, k_1, \beta}^2([c, \infty[\times \mathbb{R}) \leq \check{\pi}_{k_0, \beta}^1([c, \infty]).$$

Moreover, according to Equations (4.22) and (4.23), for all $\beta \geq \beta_1$, there exists a process $(Y_\beta(t))$ such that $(Y_\beta(t)) \geq_{\text{st}} (\check{N}_{k_1, \beta}^2(t))$ with the following transition rates

$$\begin{aligned} n_{k_1} &\rightarrow n_{k_1} + 1 : \beta \lambda_{k_1}, \\ n_{k_1} &\rightarrow n_{k_1} + 1 : \beta \left(\lambda_{k_1} + \frac{\varsigma}{3} \mu_{k_1} \right) \mathbb{1}_{\{Y_\beta(t) \geq \eta_0\}}. \end{aligned}$$

$(Y_\beta(t))$ has the transition rates of a $M/M/1$ queue with $\lfloor \eta_0 \beta \rfloor$ permanent clients. Let $\pi_{Y,\beta}$ denote the stationary distribution of $(\beta^{-1}Y_\beta(t))$ and we obtain immediately that

$$\begin{aligned} \forall \beta \geq \beta_1, \forall \eta \geq \eta_0, \pi_{Y,\beta}([\eta, \infty]) &\leq \left(\frac{\lambda_{k_1}}{\lambda_{k_1} + \frac{\xi}{3}\mu_{k_1}} \right)^{\lfloor (\eta - \eta_0)\beta \rfloor} \\ &\leq \left(\frac{\lambda_{k_1}}{\lambda_{k_1} + \frac{\xi}{3}\mu_{k_1}} \right)^{\lfloor (\eta - \eta_0)\beta_1 \rfloor} \end{aligned}$$

We deduce from the previous equation that

$$(4.25) \quad \forall \beta \geq \beta_1, \forall \eta \geq \eta_0, \check{\pi}_{k_0, k_1, \beta}^2(\mathbb{R} \times [c, \infty]) \leq \left(\frac{\lambda_{k_1}}{\lambda_{k_1} + \frac{\xi}{3}\mu_{k_1}} \right)^{\lfloor (\eta - \eta_0)\beta_1 \rfloor}$$

The tightness of $\{\check{\pi}_{k_0, k_1, \beta}^2, \beta \geq \beta_1\}$ follows from Equations (4.24) and (4.25) and the tightness of $\{\check{\pi}_{k_0, \beta}^1, \beta \in \mathbb{N}\}$.

We can remark that existence and uniqueness of the solution of Equation (4.19) are due to the monotonicity of $(\check{\psi}_{k_0}^2, \check{\psi}_{k_1}^2)$ and from Equations (4.15) and (4.20).

The proof of convergence is similar to that used in the proof of Proposition 4.3. $\blacksquare$

Using a recursion we can easily generalize the previous lemma. The principle is still the same: little by little, we decrease the access rates and the classes become stable one by one. In particular, there exists $\beta_{K-1} \in \mathbb{N}$ such that for all $\beta \geq \beta_{K-1}$, the process $(N_\beta(t))$ is ergodic where $(N_\beta(t))$ is the scaled version of $(N(t))$. We can therefore state the following.

THEOREM 4.8. *In an upstream tree, for any traffic intensities satisfying the optimal stability conditions (2.4), there exist positive access rates small enough such that the resulting stochastic process is ergodic.*

Appendix: Proof of Proposition 3.5

We first give a sketch of the proof. The technical details are proved in the lemmas which follow.

We consider the process $(\bar{N}(m, t)) = ((\bar{N}_0(m, t), \bar{N}_1(m, t), \dots, \bar{N}_k(m, t)))$ such that, for all $0 \leq k \leq L$,

$$\bar{N}_k(m, 0) = \frac{m_k}{\|m\|} \quad \text{and} \quad \bar{N}_k(m, t) = \frac{1}{\|m\|} N_k(\|m\|t), \quad \text{for } t \geq 0.$$

We consider a sequence $(m^i, i \in \mathbb{N})$ such that

$$(4.26) \quad \begin{aligned} \lim_{i \rightarrow +\infty} \|m^i\| &= +\infty, & \lim_{i \rightarrow +\infty} \frac{m_0^i}{\|m^i\|} &= \alpha, \\ \lim_{i \rightarrow +\infty} \frac{m_1^i}{\|m^i\|} &= 1 - \alpha, & \lim_{i \rightarrow +\infty} \frac{m_k^i}{\|m^i\|} &= 0, \quad \text{for } 2 \leq k \leq L. \end{aligned}$$

We want to prove that the sequence of processes $(\bar{N}(m^i, t))$ converges and to characterize the limit. For that purpose, we write them as the sum of a martingale term and a continuous term and we define the process $(M(m, t))$ such that for $t \geq 0$ and $0 \leq k \leq L$, we have

$$(4.27) \quad \begin{aligned} \bar{M}_k(m, t) &= \bar{N}_k(m, t) - \frac{m_k}{\|m\|} - \lambda_k t \\ &\quad + \mu_k \int_0^t \phi_k(\bar{N}_0(m, s), \bar{N}_1(m, s), N_2(\|m\|s), \dots, N_L(\|m\|s)) ds. \end{aligned}$$

Because the access rates a_0 and a_1 are equal to 1, we have

$$\forall n \in \mathbb{N}^K, \gamma \in \mathbb{R}, \phi_k(\gamma n_0, \gamma n_1, n_2, \dots, n_L) = \phi_k(n_0, n_1, n_2, \dots, n_L)$$

If a_0 and a_1 are not equal to 1, the previous expression is true outside a compact set and it is possible to prove all the results in this section without this assumption because, at a fluid level, the first link is always saturated.

In Lemma 4.9, we prove that the set

$$\mathcal{C} = \{(\bar{N}(m^i, t)), i \in \mathbb{N}\}$$

is relatively compact and its limiting points are continuous. We can now extract a converging subsequence and we suppose that (m^i) is such that $(\bar{N}(m^i, t))$ converges in distribution to a limit that we denote by $(\bar{Z}(t)) = (\bar{Z}_0(t), \bar{Z}_1(t), \dots, \bar{Z}_L(t))$ and which is continuous. In Lemma 4.10, we characterize $(\bar{Z}_2(t), \dots, \bar{Z}_L(t))$ by proving that, for $2 \leq k \leq L$ and $t \geq 0$, $\bar{Z}_k(t) = 0$.

Proposition 4.12 is the key result in order to understand the time scale separation between classes 0 and 1 and classes $2, \dots, L$. We deduce from this lemma that

$$\left(\int_0^t \phi_k(\bar{N}_0(m^i, s), \bar{N}_1(m^i, s), N_2(\|m^i\|s), \dots, N_L(\|m^i\|s)) ds, k = 0, 1 \right)$$

converges in distribution to

$$\left(\int_0^t \bar{\phi}_0 \left(\frac{\bar{Z}_0(s)}{\bar{Z}_0(s) + \bar{Z}_1(s)} \right) ds, \int_0^t \frac{\bar{Z}_1(s)}{\bar{Z}_0(s) + \bar{Z}_1(s)} ds \right)$$

where $\bar{\phi}_0$ characterizes the average throughput of class 0 in the quasi-stationary case as defined by Equation (3.7).

We can deduce from that and the fact that $(\bar{M}(m^i, t))$ converges to 0 in distribution that $(\bar{N}_0(m^i, t), \bar{N}_1(m^i, t))$ converges in distribution to

$$\left(\bar{Z}_0(0) + \lambda_0 t - \mu_0 \int_0^t \bar{\phi}_0 \left(\frac{\bar{Z}_0(s)}{\bar{Z}_0(s) + \bar{Z}_1(s)} \right) ds, \right. \\ \left. \bar{Z}_1(0) + \lambda_1 t - \mu_1 \int_0^t \frac{\bar{Z}_1(s)}{\bar{Z}_0(s) + \bar{Z}_1(s)} ds \right)$$

and we conclude that, almost surely,

$$\begin{aligned} \bar{Z}_0(t) &= \bar{Z}_0(0) + \lambda_0 t - \mu_0 \int_0^t \bar{\phi}_0 \left(\frac{\bar{Z}_0(s)}{\bar{Z}_1(s) + \bar{Z}_0(s)} \right) ds, \\ \bar{Z}_1(t) &= \bar{Z}_1(0) + \lambda_1 t - \mu_0 \int_0^t \frac{\bar{Z}_1(s)}{\bar{Z}_1(s) + \bar{Z}_0(s)} ds, \\ \bar{Z}_k(t) &= 0, \quad \text{for } 2 \leq k \leq L \end{aligned}$$

holds for all t in $\mathbb{R}_+$.

LEMMA 4.9.

$$\mathcal{C} = \{(\bar{N}(m^i, t)), i \in \mathbb{N}\}$$

is relatively compact and its limiting points are continuous processes.

PROOF. The method used here is also standard. We define w_h as the modulus of continuity for any function h defined on $[0, T]$:

$$w_h(\delta) = \sup_{s, t \leq T; |t-s| < \delta} |h(s) - h(t)|.$$

As in the proof of Theorem 4.1, we can prove that $(\bar{M}(m, t))$ is a martingale and its increasing processes are given by, for $0 \leq k, l \leq K$ with $k \neq l$,

$$\langle \bar{M}_k(m, t) \rangle = \frac{\lambda_k t}{\|m\|} + \frac{\mu_k t}{\|m\|} \int_0^t \phi_k(\bar{N}_0(m, s), \bar{N}_1(m, s), N_2(\|m\|s), \dots, N_L(\|m\|s)) ds,$$

$$\langle \bar{M}_k(m, t), \bar{M}_l(m, t) \rangle = 0.$$

Since the functions ϕ_k are bounded by 1, we deduce that

$$\langle \bar{M}_k(m, t) \rangle \leq \frac{(\lambda_k + \mu_k)t}{\|m\|}.$$

By using Doob's inequality, we have that, for $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P} \left(\sup_{0 \leq s \leq t} \|\bar{M}(m, s)\| \geq \varepsilon \right) &\leq \sum_{k=0}^L \mathbb{P} \left(\sup_{0 \leq s \leq t} |\bar{M}_k(m, s)| \geq \frac{\varepsilon}{L} \right), \\ &\leq \frac{L^2 t \sum_{k=0}^L \lambda_k + \mu_k}{\varepsilon^2 \|m\|}. \end{aligned}$$

We deduce that if (m^i) is a sequence satisfying Equation (4.26) then $(\bar{M}(m^i, t))$ converges in probability to 0 uniformly on compact sets when i tends to infinity; for any $T \leq 0$ and any $\varepsilon > 0$,

$$\lim_{i \rightarrow \infty} \mathbb{P} \left(\sup_{0 \leq t \leq T} \|\bar{M}(m^i, t)\| \geq \varepsilon \right) = 0.$$

Using Equation (4.27) and the previous equation, we can prove that for any $\varepsilon > 0$ and $\eta > 0$, there exist $\delta > 0$ and A such that for all $i \geq A$,

$$\mathbb{P}(w_{\bar{N}_k(m^i, \cdot)}(\delta) > \eta) \leq \varepsilon.$$

The conditions of [Bil99, 7.2 p81] are then fulfilled and the set $\mathcal{C}$ is relatively compact and its limiting points are continuous processes. $\blacksquare$

We are now able to characterize $(\bar{Z}_2, \dots, \bar{Z}_L(t))$.

LEMMA 4.10. *We consider a sequence (m^i) satisfying Equation (4.26) and such that $(N(m^i, t))$ converges in distribution to $(\bar{Z}(t))$. Then the process $(\bar{Z}(t))$ satisfies $\bar{Z}_k(t) = 0$ for $2 \leq k \leq L$ and $t \geq 0$.*

PROOF. Define the process $(\hat{N}_{2:L}(t))$ with the following transition rates for $2 \leq k \leq L$:

$$\begin{aligned} n_k &\mapsto n_k + 1 : \lambda_k, \\ n_k &\mapsto n_k - 1 : \mu_k \frac{n_k a_k}{1 + n_k a_k}. \end{aligned}$$

and such that $\hat{N}_{2:L}(0) \geq \tilde{N}_{2:L}^\alpha(0)$. Clearly, $(\hat{N}_{2:L}(t))$ is ergodic and stochastically dominates $(N_{2:L}(t))$. Moreover, the components evolve independently. If we call $(\hat{Z}_{2:L}(t))$ a fluid limit of $(\hat{N}_{2:L}(t))$, we can prove, as in [Rob03, Prop 5.16 p125], that it satisfies

$$\hat{Z}_k(t) = (\hat{Z}_k(0) + (\lambda_k - \mu_k)t)_+, \quad \text{for } t \geq 0 \text{ and } 2 \leq k \leq L.$$

Moreover, by stochastic domination, we have $\bar{Z}_k(t) \leq \hat{Z}_k(t)$ for all $t \geq 0$ and $2 \leq k \leq L$. Since (m^i) satisfies Equation (4.26), we have that $\bar{Z}_k(0) = 0$, for $2 \leq k \leq L$ and we conclude that $\bar{Z}_k(t) = 0$ for all $t \geq 0$ and $2 \leq k \leq L$. $\blacksquare$

We define $\bar{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$ and we consider $\bar{\mathbb{N}}^{K-2}$. We endow $\bar{\mathbb{N}}^{K-2}$ with the metric induced from the L_1 -norm on $\mathbb{R}^{K-2}$ by the mapping $(x_2, \dots, x_L) \mapsto (1/(x_2 + 1), \dots, 1/(x_L + 1))$. We can note that, in particular, $\bar{\mathbb{N}}^{K-2}$ is compact. In the same way as in [HK94], we then define a family of random measures on $[0, \infty) \times \bar{\mathbb{N}}^{K-2}$, for any $m \in \mathbb{N}^2$, any $\Gamma \subset \bar{\mathbb{N}}^{K-2}$ and $t \geq 0$,

$$\nu_m((0, t) \times \Gamma) = \int_0^t \mathbb{1}_{\{(N_2(\|m\|s), \dots, N_L(\|m\|s)) \in \Gamma\}} ds.$$

We then consider $\mathcal{L}_0(\bar{\mathbb{N}}^{K-2})$, the set of measures γ defined on $[0, t] \times \bar{\mathbb{N}}^{K-2}$ and such that $\gamma([0, t] \times \bar{\mathbb{N}}^{K-2}) = t$ for all $t \geq 0$. Under the topology induced by weak convergence on every compact and since $\bar{\mathbb{N}}^{K-2}$ is compact, $\mathcal{L}_0(\bar{\mathbb{N}}^{K-2})$ is compact. The next lemma shows that a random measure in $\mathcal{L}_0(\bar{\mathbb{N}}^{K-2})$ can be expressed as the sum of probability measures of $\bar{\mathbb{N}}^{K-2}$ indexed by s in $\mathbb{R}_+$.

LEMMA 4.11. *We consider a sequence $(m^i, i \in \mathbb{N})$ satisfying Equation (4.26). The set*

$$\{((\bar{N}(m^i, t)), \nu_{m^i}), i \in \mathbb{N}\}$$

is relatively compact. If $((\bar{Z}(t)), \nu)$ is a limit process then there exists a process ϑ such that for all t , $\vartheta(t, \cdot)$ is a random probability measure on $\bar{\mathbb{N}}^{K-2}$ and

$$\forall t \geq 0, \forall \Gamma \subset \bar{\mathbb{N}}^{K-2}, \nu([0, t] \times \Gamma) = \int_0^t \vartheta(s, \Gamma) ds.$$

PROOF. One can find a related result in a slightly different context in [Kur92] and [HK94].

In order to prove the relative compactness of $\{((\bar{N}(m^i, t)), \nu_{m^i}), i \in \mathbb{N}\}$, we just need to prove that $\{(\bar{N}(m^i, t)), i \in \mathbb{N}\}$ and $\{\nu_{m^i}, i \in \mathbb{N}\}$ are relatively compact. We have already proved the relative compactness of the first one in lemma 4.9. $\mathcal{L}_0(\bar{\mathbb{N}}^{K-2})$ is compact then the second one is relatively compact.

We consider a convergent sequence $((\bar{N}(m^i, t)), \nu_{m^i})$ and its limiting process $((\bar{Z}(t)), \nu)$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space on which they are defined. We call $\{\mathcal{F}_t\}$ the natural filtration of $(\bar{Z}(t), \nu)$.

We then define γ such that

$$\forall B \in \mathcal{F}, \forall C \in \mathcal{B}([0, \infty)) \otimes \mathcal{B}(\bar{\mathbb{N}}^{K-2}) \quad \gamma(B \times C) = \mathbb{E}(\mathbb{1}_{\{B\}} \nu(C)).$$

According to [EK86, appendix 8], γ can be extended to a measure on $\mathcal{F} \otimes \mathcal{B}([0, \infty)) \otimes \mathcal{B}(\bar{\mathbb{N}}^{K-2})$ and there exists ϑ such that for all t , $\vartheta(t, \cdot)$ is a random probability measure on $\bar{\mathbb{N}}^{K-2}$ and for any $B \in \mathcal{B}(\bar{\mathbb{N}}^{K-2})$, $(\vartheta(t, B), t \geq 0)$ is $\{\mathcal{F}_t\}$ -adapted and for any $C \in \mathcal{F} \otimes \mathcal{B}([0, \infty))$,

$$\gamma(C \times B) = \mathbb{E} \left(\int_0^{+\infty} \mathbb{1}_{\{C\}}(s) \vartheta(s, B) ds \right).$$

We define

$$M_B(t) = \nu([0, t] \times B) - \int_0^t \vartheta(s, B) ds.$$

$(M_B(t))$ is $\{\mathcal{F}_t\}$ -adapted and continuous. We consider $D \in \mathcal{F}_t$. We define $\mathbb{1}_{\{C\}}(\omega, s) = \mathbb{1}_{\{D\}}(\omega) \mathbb{1}_{\{[t, +\infty)\}}(s)$ and we have

$$\begin{aligned} \mathbb{E}(\mathbb{1}_{\{D\}} \nu([t, +\infty) \times B)) &= \gamma(C \times B), \\ &= \mathbb{E} \left(\mathbb{1}_{\{D\}} \int_t^{+\infty} \vartheta(s, B) ds \right). \end{aligned}$$

It follows that

$$\mathbb{E}(\nu([t, +\infty) \times B) | \mathcal{F}_t) = \mathbb{E} \left(\int_t^{+\infty} \vartheta(s, B) ds | \mathcal{F}_t \right).$$

Then, $(M_B(t))$ is a continuous $\{\mathcal{F}_t\}$ -martingale. It has finite sample paths and then is almost surely identically null. Almost surely, the following equation holds for all t ,

$$\forall B \subset \bar{\mathbb{N}}^{K-2}, \nu([0, t] \times B) = \int_0^t \vartheta(s, B) \, ds.$$

■

The previous lemma gives a convenient way to express the integral of linear combinations of indicator functions against a limit measure of (ν_m) . We will use this technical lemma in the next one to characterize the limit of (ν_m) . This is the most important part of the theorem and we prove the time scale separation between classes 0 and 1 and $2, \dots, L$ here.

PROPOSITION 4.12. *We consider a sequence $(m^i, i \in \mathbb{N})$ satisfying Equation (4.26) and such that $((\bar{N}(m^i, t)), \nu_{m^i})$ is a converging sequence and h a continuous function on $[0, 1]^2 \times \bar{\mathbb{N}}^{K-2}$. Then, we have that*

$$\left(\int_0^t h(\bar{N}_0(m^i, s), \bar{N}_1(m^i, s), N_2(\|m^i\|s), \dots, N_L(\|m^i\|s)) \, ds \right)$$

converges in distribution to

$$\left(\int_0^t \sum_{y \in \bar{\mathbb{N}}^{K-2}} h(\bar{Z}_0(s), \bar{Z}_1(s), y) \tilde{\pi}^{\alpha(s)}(y) \, ds \right)$$

with

$$\forall t \in \mathbb{R}_+, \alpha(t) = \frac{\bar{Z}_0(t)}{\bar{Z}_0(t) + \bar{Z}_1(t)}$$

and where, $\tilde{\pi}^{\alpha(t)}$ is the stationary distribution of the process $(\tilde{N}_{2:L}^{\alpha(t)}(s))$ defined in Section 3.3.

In particular, the function

$$(x_0, x_1, n_2, \dots, n_L) \mapsto \psi_0(x_0, x_1, n_2 a_2, \dots, n_L a_L)$$

is continuous on $[0, 1]^2 \times \bar{\mathbb{N}}^{K-2}$ and

$$\left(\int_0^t \phi_0(\bar{N}_0(m^i, s), \bar{N}_1(m^i, s), N_2(\|m^i\|s), \dots, N_L(\|m^i\|s)) \, ds \right)$$

converges in distribution to

$$\left(\int_0^t \bar{\phi}_0(\alpha(s)) \, ds \right)$$

PROOF. The functions which are continuous on $[0, 1]^2 \times \bar{\mathbb{N}}^{K-2}$ for the topology induced by the natural topology on $[0, 1]^2$ and the topology induced by the mapping $(x_2, \dots, x_L) \mapsto (1/(x_2 + 1), \dots, 1/(x_L + 1))$ are the bounded functions h which are continuous on $[0, 1]^2 \times \bar{\mathbb{N}}^{K-2}$ for the natural topology and such that for each $x \in [0, 1]^2$, $y \mapsto h(x, y)$ admits a unique limit $h(x, \infty)$ in all directions such that $\|y\| \rightarrow \infty$ and the function $x \mapsto h(x, \infty)$ has to be continuous on $[0, 1]^2$. We can remark that, for all $(x_0, x_1) \in [0, 1]^2$ $\psi_0(x_0, x_1, n_{2:L} \odot a_{2:L}) \rightarrow 0$ when $\|n_{2:L}\| \rightarrow +\infty$ and the function

$$(x_0, x_1, n_2, \dots, n_L) \mapsto \psi_0(x_0, x_1, n_2 a_2, \dots, n_L a_L)$$

is then continuous on $[0, 1]^2 \times \bar{\mathbb{N}}^{K-2}$.

We consider h a continuous function on $[0, 1]^2 \times \bar{\mathbb{N}}^{K-2}$, since the space $[0, 1]^2 \times \bar{\mathbb{N}}^{K-2}$ is compact, Lemma 4.11 implies directly that

$$\left(\int_0^t h(\bar{N}_0(m^i, s), \bar{N}_1(m^i, s), N_2(\|m^i\|s), \dots, N_L(\|m^i\|s)) \, ds \right)$$

converges in distribution to

$$\left(\int_0^t \sum_{y \in \bar{\mathbb{N}}^{K-2}} h(\bar{Z}_0(s), \bar{Z}_1(s), y) \vartheta(s, y) \, ds \right).$$

We are now able to fully characterize the random measures $(\vartheta(\cdot, t))$. For any continuous bounded function f on $\bar{\mathbb{N}}^{K-2}$ and any $m \in \mathbb{N}^{K-2}$, we define,

$$\begin{aligned} \bar{M}_f(m, t) &= \frac{1}{\|m\|} \left(f(N_2(\|m\|t), \dots, N_L(\|m\|t)) - f(0) \right) \\ &\quad - \sum_{k=2}^L \lambda_k \int_0^t \left(f((N_2(\|m\|s), \dots, N_L(\|m\|s)) + e_k) \right. \\ &\quad \left. - f(N_2(\|m\|s), \dots, N_L(\|m\|s)) \right) \, ds \\ &\quad - \sum_{k=2}^L \mu_k \int_0^t \left((f((N_2(\|m\|s), \dots, N_L(\|m\|s)) - e_k) \right. \\ &\quad \left. - f(N_2(\|m\|s), \dots, N_L(\|m\|s))) \right. \\ &\quad \left. \phi_k(\bar{N}_0(m, s), \bar{N}_1(m, s), N_2(\|m\|s), \dots, N_L(\|m\|s)) \right) \, ds. \end{aligned}$$

As $(\bar{M}(m, t))$ defined by equation (4.27) is a martingale, $(\bar{M}_f(m, t))$ is a martingale. We consider a convergent sequence $((\bar{N}(m^i, t), \nu_{m^i}))$. We have that the sequence $(\bar{M}_f(m^i, t))$ converges in distribution to 0. This in turn implies that $\|m^i\|^{-1}(f(N_2(\|m^i\|t), \dots, N_L(\|m^i\|t)) - f(0))$ also converges to 0 because f is bounded. As a consequence, the following term

$$\begin{aligned} &\sum_{k=2}^L \lambda_k \int_0^t \left(f((N_2(\|m^i\|s), \dots, N_L(\|m^i\|s)) + e_k) \right. \\ &\quad \left. - f(N_2(\|m^i\|s), \dots, N_L(\|m^i\|s)) \right) \, ds \\ &- \sum_{k=2}^L \mu_k \int_0^t \left((f((N_2(\|m^i\|s), \dots, N_L(\|m^i\|s)) - e_k) \right. \\ &\quad \left. - f(N_2(\|m^i\|s), \dots, N_L(\|m^i\|s))) \right. \\ &\quad \left. \phi_k(\bar{N}_0(m^i, s), \bar{N}_1(m^i, s), N_2(\|m^i\|s), \dots, N_L(\|m^i\|s)) \right) \, ds \end{aligned}$$

also converges in distribution to 0. But, by the continuous mapping theorem and Lemma 4.11, it converges in distribution to

$$\begin{aligned} &\int_0^t \sum_{k=2}^L \left(\lambda_k \sum_{y \in \bar{\mathbb{N}}^{K-2}} f(y + e_k) - f(y) \right. \\ &\quad \left. + \mu_k \sum_{y \in \bar{\mathbb{N}}^{K-2}} (f(y - e_k) - f(y)) \varphi_k(\bar{Z}_0(s), \bar{Z}_1(s), y) \right) \vartheta(s, y) \, ds. \end{aligned}$$

Consequently, this is null almost surely for all t and we have then, for Lebesgue-almost every t ,

$$\sum_{k=2}^L \left(\lambda_k \sum_{y \in \mathbb{N}^{K-2}} f(y + e_k) - f(y) + \mu_k \sum_{y \in \mathbb{N}^{K-2}} (f(y - e_k) - f(y)) \varphi_k(\bar{Z}_0(t), \bar{Z}_1(t), y) \right) \vartheta(t, y) = 0.$$

We deduce immediately that

$$\int_{\mathbb{N}^{K-2}} \tilde{\Omega}^{\alpha(t)}(f)(y) \vartheta(t, dy) = 0$$

where $\alpha(t) = \bar{Z}_0(t)/(\bar{Z}_0(t) + \bar{Z}_1(t))$ and $\tilde{\Omega}^{\alpha(t)}$ is the infinitesimal generator of $(\tilde{N}_{2:L}^{\alpha(t)}(s))$, defined in Section 3.3. This proves exactly that $\vartheta(t, \cdot)$ is invariant for $(\tilde{N}_{2:L}^{\alpha(t)}(s))$. By uniqueness of the invariant distribution of such a process, we have that

$$\vartheta(t, \cdot) = \tilde{\pi}^{\alpha(t)}.$$

■

References for Chapter II

- [BBP04] T. Bonald, S.C. Borst, and A. Proutiere. How mobility impacts the flow-level performance of wireless data systems. In *Proceedings of INFOCOM*, volume 3, pages 1872–1881, 2004.
- [BFP09] T. Bonald, M. Feuillet, and A. Proutiere. Is the “Law of the Jungle” sustainable for the Internet? In *Proceedings of INFOCOM*, pages 28–36, april 2009.
- [Bil99] Patrick Billingsley. *Convergence of Probability Measures (second edition)*. Wiley Series in Probability and Statistics. Wiley-Interscience, 1999.
- [BJL08] Sem Borst, Matthieu Jonckheere, and Lasse Leskelä. Stability of parallel queueing systems with coupled service rates. *Discrete Event Dynamic Systems*, 18(4):447–472, dec 2008.
- [BMPV06] T. Bonald, L. Massoulié, A. Proutiere, and J. Virtamo. A queueing analysis of max-min fairness, proportional fairness, and balanced fairness. *Queueing Systems: Theory and Applications*, 53:65–84, 2006.
- [Dai95] J. G. Dai. On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Annals of Applied Probability*, 5:49–77, 1995.
- [DGR02] Vincent Dumas, Fabrice Guillemin, and Philippe Robert. A Markovian analysis of Additive-Increase Multiplicative-Decrease (AIMD) algorithms. *Advances in Applied Probability*, 34(1):85–111, 2002.
- [EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and convergence*. Wiley, 1986.
- [Flo03] S. Floyd. Highspeed TCP for large congestion windows. RFC 896, 2003.
- [GLM⁺10] Ayalvadi Ganesh, Sarah Lilienthal, D. Manjunath, Alexandre Proutiere, and Florian Simatos. Load balancing via random local search in closed and open systems. In *SIGMETRICS*, pages 287–298, 2010.

- [Gro19] Thomas H. Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *The Annals of Mathematics, Second Series*, 20(4):292–296, July 1919.
- [HK94] P. J. Hunt and T. G. Kurtz. Large loss networks. *Stochastic Processes and their Applications*, 53(2):363–378, 1994.
- [JWL04] Cheng Jin, D.X. Wei, and S.H. Low. FAST TCP: motivation, architecture, algorithms, performance. In *INFOCOM 2004*, volume 4, pages 2490 – 2501, jul 2004.
- [Kur92] T.G. Kurtz. Averaging for martingale problems and stochastic approximation. In *Applied Stochastic Analysis, US-French Workshop*, volume 177 of *Lecture notes in Control and Information sciences*, pages 186–209. Springer Verlag, 1992.
- [Mas87] W. A. Massey. Stochastic orderings for Markov processes on partially ordered spaces. *Mathematics of Operations Research*, 12(2):350–367, May 1987.
- [Mey95] Sean Meyn. Transience of multiclass queueing networks via fluid limit models. *Annals of Applied Probability*, 5:946–957, 1995.
- [Mit04] M. Mitzenmacher. Digital fountains: a survey and look forward. *Information Theory Workshop, 2004. IEEE*, pages 271–276, 2004.
- [MR00] L. Massoulié and J. W. Roberts. Bandwidth sharing and admission control for elastic traffic. *Telecommunications Systems*, 15:185–201, 2000.
- [Nag84] J. Nagle. Congestion Control in IP/TCP Internetworks. RFC 896, January 1984.
- [Rob03] Philippe Robert. *Stochastic Networks and Queues*. Stochastic Modeling and Applied Probability Series. Springer-Verlag, New York, 2003. xvii+398 pp.
- [RS05] Barath Raghavan and Alex C. Snoeren. Decongestion control. In *Proceedings of the 5th ACM Workshop on Hot Topics in Networks*, november 2005.
- [RW79] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes & martingales vol. 1: Foundations*. Cambridge University Press, 2000 (1979).
- [RW87] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes & martingales vol. 2: Itô Calculus*. Cambridge University Press, 2000 (1987).
- [ST10] Florian Simatos and Danielle Tibi. Spatial homogenization in a stochastic network with mobility. *The Annals of Applied Probability*, 20(1):312–355, 2010.
- [XHR04] Lisong Xu, K. Harfoush, and Injong Rhee. Binary increase congestion control (BIC) for fast long-distance networks. In *INFOCOM 2004*, volume 4, pages 2514 – 2524, jul 2004.

CHAPTER III

On the Performance of CSMA in Multi-Channel Wireless Networks

Contents

1. Introduction	69
2. Related Work	70
3. Model	71
4. Packet-Level Dynamics	72
5. Flow-Level Dynamics	73
6. Time-Scale Separation	76
7. Infrastructure-Based Networks	80
8. Conclusion	87
References for Chapter III	88

1. Introduction

The CSMA (Carrier Sense Multiple Access) algorithm is a key component of IEEE 802.11 networks. While it proves successful in sharing a single radio channel between a limited number of stations, its efficiency is questionable in more involved environments with multiple radio channels and a large number of stations having different interference constraints. In this chapter, we analyze the ability of CSMA to fully utilize the radio resources in such environments, in both ad-hoc and infrastructure modes, accounting for the random nature of traffic. Specifically, each station attempts to access a randomly chosen radio channel after some random backoff time and transmits a packet over this channel if it is sensed idle. We study the random variations of the number of active wireless links induced by this random access algorithm and the random activity of users. In particular, we analyze the ergodicity of the associated Markov process, which characterizes the ability of CSMA to stabilize the network.

It turns out that, while CSMA is always efficient in ad-hoc mode, in the sense that the network is stable whenever possible, it is generally inefficient in infrastructure mode, when all data flows originate from or are destined to some finite set of access points. This is due to the fact that, in ad-hoc mode, each new flow adds a new link to the network, which competes fairly with other links for accessing the radio channels. In infrastructure mode, each access point attempts to access the radio channels at some constant rate, independently of the number of uplink flows (from the stations) and downlink flows (to the stations) at this access point. This inherent bias against access points results in some loss of capacity in the sense that the network may be unstable at load strictly less than 1. We prove that a slight modification of CSMA, which consists in running one instance of CSMA per flow at each access point, corrects this bias and makes the algorithm optimal. We refer to this algorithm, introduced in [BF10], as *flow-aware CSMA*.

The rest of the chapter is organized as follows. We present some related work in the next section. The network model in ad-hoc mode is described in section 3. Sections 4 and 5 are devoted to the packet- and flow-level dynamics, respectively, assuming time-scale separation. The main result of the chapter, given in Theorem 5.1, shows in particular the optimality of CSMA in ad-hoc mode. The validity of the time-scale separation assumption is discussed in section 6. The infrastructure mode is considered in section 7, where we prove the suboptimality of standard CSMA and the optimality of flow-aware CSMA. Section 8 concludes the chapter.

2. Related Work

The present work is related to the problem of throughput-optimal scheduling in wireless networks. A centralized solution is known since the seminal work of Tassiulas and Ephremides, who proved in [TE92] the optimality of the *maximum weight* policy. A number of distributed implementations of this policy have then been proposed, all relying on some message passing protocol between stations, see e.g. [MSZ06]. It is only recently that fully distributed solutions without message passing have been found. Jiang and Walrand introduced in [JW08] a distributed CSMA algorithm where the rate at which each link attempts to access the radio channel is adapted to the packet arrival rate and departure rate at this link, so as to meet the demand. This approach is based on a time-scale separation assumption whereby the attempt rates evolve at a much slower timescale than the packet arrivals and departures. This assumption turns out to be critical and the algorithm must actually be carefully designed in order to guarantee convergence and throughput-optimality [PYLC10, JSSW10]. Another approach that does not rely on the time-scale separation assumption was proposed independently by Rajagopalan, Shah and Shin [RS08, RSS09]. In their algorithm, the attempt rate of each station is some slowly varying function of the queue size, which essentially enforces the time-scale separation. The proof of optimality then relies on the fact that this adaptive version of CSMA asymptotically achieves some form of maximum weight scheduling. Similar ideas are used by Ni, Tan and Srikant in [NTS10].

All these papers focus on the packet-level dynamics, assuming packets are generated by some fixed number of flows over some fixed set of links. The flow-level dynamics are ignored, whereas they are known to be critical, see for instance [BTI⁺03, BBP⁺01, BK00, MR00] in the context of wireline networks. A flow-level model using adaptive CSMA is considered by Shah and Shin [SS11]. This model is very different from ours in that it applies to optical networks and, in particular, does not involve any packet-level dynamics: a flow that is allocated a circuit in the network uses this circuit until the completion of its service requirement. As in [BF10], we consider both the packet-level and flow-level dynamics, under the usual assumption that the former are much faster than the latter. Specifically, we extend the results of [BF10] to multi-channel wireless networks in both ad-hoc and infrastructure modes, accounting for the dynamic nature of the network topology.

Surprisingly, little attention has been paid so far to multi-channel networks. A notable exception is the adaptive, multi-channel version of CSMA introduced in [PYLC10], which is shown to maximize the network utility when combined with some appropriate virtual queue mechanism. We here prove the optimality of CSMA in the sense of flow-level stability for a very general model where the interference constraints may depend on the considered channel and each transmitter may only use a subset of the channels. Specifically, we show that it is sufficient for each transmitter to probe one of its channels at random, without any further information on the network state.

Another salient feature is the observation of the key difference between the ad-hoc and infrastructure modes. In the former, the number of transmitters grows with the congestion, which increases the channel attempt rate and in turn stabilizes the network. This is not the case in the latter since the channel access opportunities of each access point must be shared by all downlink flows at this access point. This inherent bias of CSMA against downlink traffic is well known, see e.g. [HRBSD03, KKF05], and can be easily corrected by letting the attempt rate of each access point depend on the number of downlink flows, a scheme we refer to as *flow-aware* CSMA [BF10]. It is worth noting that this algorithm does not depend on the queue *length* (number of packets) but on the queue *width* (number of flows). As such, the nature of the algorithm and the significance of the underlying time-scale separation assumption (packet-level vs. flow-level dynamics) are very different from those considered so far, based on packet-level models [JW08, PYLC10, JSSW10, BBLP11].

3. Model

3.1. A Multi-Channel Wireless Network. The network consists of a random, dynamic set of wireless links in ad-hoc mode (there is no access point at this stage). These links must share some finite number J of non-interfering radio channels. Each link consists of a transmitter-receiver pair; the transmitter is able to use at most one radio channel at a time. We group links into a finite number of K classes, as illustrated by Figure 1. All links within the same class have the same radio conditions, the same interference constraints and the same CSMA parameters. We denote by x_k the number of class- k links and by x the corresponding vector, which we refer to as the *network state*. Two links within the same class cannot be simultaneously active on the same channel. An active class- k link on channel j transmits data at the physical rate φ_k bit/s, independently of j . We say that class k is active on channel j if there is an active class- k link on channel j .

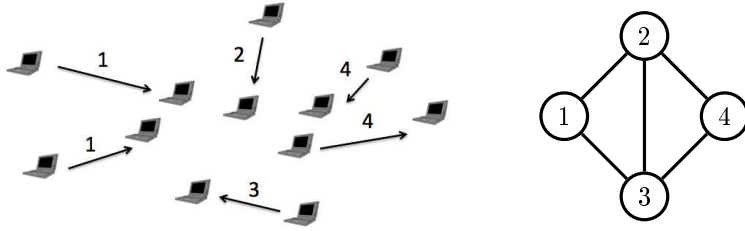


FIGURE 1. An ad-hoc wireless network with 4 classes of links and its interference graph.

Each channel j is associated with some *conflict graph* $G_j = (V_j, E_j)$, where $V_j \subset \{1, \dots, K\}$ is the set of classes that are able to transmit on channel j and E_j is the set of edges, each representing a conflict. Specifically, two classes $k, l \in V_j$ can be simultaneously active on channel j if and only if they do not conflict with each other, that is if $(k, l) \notin E_j$. The J conflict graphs are typically the same but could differ due to different radio propagation environments on the J channels, or to different transmission capabilities of the K classes.

3.2. Feasible Schedules. We refer to a schedule as any vector $y \in \{0, 1\}^{K \times J}$, where $y_{kj} = 1$ if class k is active on channel j . We denote by y_k the number of

active class- k links:

$$y_k = \sum_{j=1}^J y_{kj}.$$

The schedule is *feasible* if for all $j = 1, \dots, J$, the active classes on channel j belong to V_j and do not conflict with each other, that is $y_{kj}y_{lj} = 0$ for all $(k, l) \in E_j$. Moreover, we must have:

$$(3.1) \quad \forall k = 1, \dots, K, \quad y_k \leq x_k.$$

We denote by $\mathcal{Y}(x)$ the set of feasible schedules. Note that if $x_k \geq J$ for all $k = 1, \dots, K$, the constraint (3.1) is no longer limiting (since the number of active class- k links is limited by the number of radio channels J) and the set of feasible schedules becomes independent of the network state. We denote by $\mathcal{Y}$ the corresponding set, which is the union of $\mathcal{Y}(x)$ over all network states x .

3.3. Capacity Region. Assume that each feasible schedule y is selected with probability $\pi(y)$, with $\sum_{y \in \mathcal{Y}} \pi(y) = 1$. The mean throughput of class k is then given by:

$$(3.2) \quad \phi_k = \varphi_k \sum_{y \in \mathcal{Y}} y_k \pi(y).$$

Let ϕ be the corresponding throughput vector. We refer to the *capacity region* as the set of vectors ϕ generated by all probability measures $\pi(y)$, $y \in \mathcal{Y}$. Note that the capacity region depends both on the physical rates and on the interference constraints of all wireless links.

4. Packet-Level Dynamics

We first analyze the packet-level dynamics induced by CSMA for a static network state x . The flow-level dynamics that make x vary are introduced in section 5.

4.1. Random Access. We consider the standard CSMA algorithm where each transmitter waits for a period of random duration referred to as the *backoff time* before each transmission attempt. At each attempt, the transmitter chooses a radio channel at random and probes it. If the radio channel is sensed idle (in the sense that no conflicting link is active), a packet is transmitted; otherwise, the transmitter waits for a new backoff time before the next attempt.

Packets have random sizes of unit mean and are transmitted at the physical rate φ_k on class- k links; the backoff times of class- k transmitters are random with mean $1/\nu_k$, where $\nu_k > 0$ is the corresponding attempt rate. We denote by $\alpha_k = \nu_k/\varphi_k$ the ratio of the mean packet transmission time to the mean backoff time of class- k links. Channel j is chosen with probability β_{kj} , with $\sum_{j=1}^J \beta_{kj} = 1$ and $\beta_{kj} > 0$ if and only if $k \in V_j$, so that all accessible channels are attempted with positive probability.

4.2. Stationary Distribution. Let $(Y(t))$ be the schedule selected by the above random access algorithm at time t . We look for the stationary distribution of $(Y(t))$, which we denote by $\pi(x, y)$ to highlight the fact that it depends on the network state x . We have:

PROPOSITION 4.1. *If both the packet sizes and the backoff times have exponential distributions, then $(Y(t))$ is a reversible Markov process, with stationary measure:*

$$(4.1) \quad w(x, y) = \prod_{k: x_k > 0} \frac{x_k!}{(x_k - y_k)!} \alpha_k^{y_k} \prod_{j=1}^J \beta_{kj}^{y_{kj}}, \quad y \in \mathcal{Y}(x).$$

PROOF. Let e_{kj} be the unit vector on component k, j on $\{0, 1\}^{K \times J}$. The Markov process $(Y(t))$ jumps from state y to state $y + e_{kj}$ with rate $(x_k - y_k)\nu_k\beta_{kj}$ (since all idle links attempt to access the channel) and from state $y + e_{kj}$ to state y with rate φ_k (since all class- k links have the same physical rate φ_k , independently of the used channel), for any state y such that $y + e_{kj} \in \mathcal{Y}(x)$. The proof then follows from the local balance equations:

$$w(x, y)(x_k - y_k)\nu_k\beta_{kj} = w(x, y + e_{kj})\varphi_k.$$

■

The stationary distribution $\pi(x, y)$ follows from the normalization of the stationary measure $w(x, y)$ over all $y \in \mathcal{Y}(x)$. We deduce the mean throughput of class k in state x :

$$(4.2) \quad \phi_k(x) = \varphi_k \sum_{y \in \mathcal{Y}} y_k \pi(x, y).$$

It turns out that, by the insensitivity property of the underlying loss network [Bon07], these expressions are in fact valid for any *phase-type* distributions of packet sizes and backoff times; such distributions are known to form a dense subset within the set of all distributions with real, non-negative support [Ser99], so that the results hold for virtually any distributions of packet sizes and backoff times. We refer the reader to [vdVBvLP10] for further details on this insensitivity property.

5. Flow-Level Dynamics

We now introduce the flow-level dynamics under the assumption of infinitely fast packet-level dynamics; the validity of this time-scale separation assumption is discussed in section 6.

5.1. Traffic Characteristics. We assume that flows using class- k links are generated according to a Poisson process of intensity λ_k . Each such flow has an exponential size with mean σ_k bits and leaves the network once the corresponding data transfer is completed. There is a one-to-one correspondence between flows and links so that both terms are used interchangeably in the following. We denote by $\rho_k = \lambda_k \sigma_k$ the traffic intensity of class k (in bit/s) and by ρ the corresponding vector.

Under the time-scale separation assumption, the flow-level dynamics are much slower than the packet-level dynamics so that, at the time scale of a flow, everything happens as if the stationary distribution (4.1) of the packet-level dynamics were reached instantaneously. In particular, the mean throughput of class k is given by (4.2) in state x .

5.2. Stability Region. Let $X_k(t)$ be the number of class- k flows at time t . The corresponding vector $(X(t))$ describes the evolution of the network state. This is a Markov process with transition rates λ_k from state x to state $x + e_k$ and $\phi_k(x)/\sigma_k$ from state x to state $x - e_k$ (provided $x_k > 0$), where e_k denotes the unit vector on component k .

We say that the network is stable if this Markov process is ergodic. Clearly, a necessary condition for stability is that the vector of traffic intensities ρ lies in the capacity region. The following key result shows that this condition is in fact sufficient, up to the critical case where ρ lies on the boundary of the capacity region. In this sense, CSMA is optimal in the considered ad-hoc mode.

THEOREM 5.1. *The network is stable for all vectors of traffic intensities ρ in the interior of the capacity region.*

The proof is based on the fact that the random access algorithm selects schedules in proportion to their weights (4.1). For large x , this is equivalent to selecting schedules in proportion to the following uniform weight, which is independent of the channel probing distribution:

$$(5.1) \quad u(x, y) = \prod_{k: x_k > 0} (x_k \alpha_k)^{y_k}, \quad y \in \mathcal{Y}(x).$$

Defining:

$$u(x) = \max_{y \in \mathcal{Y}(x)} u(x, y),$$

the following result shows that those schedules of maximum weight are actually selected with probability close to 1:

LEMMA 5.2. *For any $\epsilon > 0$, we have:*

$$\sum_{y \in \mathcal{Y}(x)} \pi(x, y) \log(u(x, y)) \geq (1 - \epsilon) \log(u(x))$$

for all states x but some finite number.

PROOF. For any class k , let:

$$\beta_k = \min_{j: k \in V_j} \beta_{kj}.$$

Note that $\beta_k > 0$. We have for all $y \in \mathcal{Y}(x)$:

$$w(x, y) \geq \prod_{k: x_k > 0} \frac{x_k(x_k - 1) \dots (x_k - y_k + 1)}{x_k^{y_k}} \beta_k^J u(x, y).$$

If $x_k \leq 2J$, we have:

$$\frac{x_k(x_k - 1) \dots (x_k - y_k + 1)}{x_k^{y_k}} \geq \frac{1}{x_k^{y_k}} \geq \frac{1}{(2J)^J}.$$

Otherwise, we have using the fact that $y_k \leq J$ for all $k = 1, \dots, K$:

$$\frac{x_k(x_k - 1) \dots (x_k - y_k + 1)}{x_k^{y_k}} \geq \left(\frac{x_k - y_k + 1}{x_k} \right)^{y_k} \geq \frac{1}{2^J}.$$

Combining these results, we obtain the existence of some constant $m > 0$ such that:

$$\forall y \in \mathcal{Y}(x), \quad w(x, y) \geq m u(x, y).$$

Now let:

$$\mathcal{Z}(x) = \left\{ y \in \mathcal{Y}(x) : \log(u(x, y)) \geq (1 - \frac{\epsilon}{2}) \log(u(x)) \right\}.$$

We have:

$$\sum_{y \in \mathcal{Z}(x)} \pi(x, y) \log(u(x, y)) \geq (1 - \frac{\epsilon}{2}) \log(u(x)) \sum_{y \in \mathcal{Z}(x)} \pi(x, y).$$

Using the fact that $w(x, y) \leq u(x, y)$ for all $y \in \mathcal{Y}(x)$, we get:

$$\begin{aligned} \sum_{y \in \mathcal{Y}(x) \setminus \mathcal{Z}(x)} \pi(x, y) &= \frac{\sum_{y \in \mathcal{Y}(x) \setminus \mathcal{Z}(x)} w(x, y)}{\sum_{y \in \mathcal{Y}(x)} w(x, y)}, \\ &\leq \frac{1}{m} \frac{\sum_{y \in \mathcal{Y}(x) \setminus \mathcal{Z}(x)} u(x, y)}{\sum_{y \in \mathcal{Y}(x)} u(x, y)}, \\ &\leq \frac{1}{m} \frac{M u(x)^{1 - \frac{\epsilon}{2}}}{\max_{y \in \mathcal{Y}(x)} u(x, y)}, \\ &= \frac{1}{m} \frac{M}{u(x)^{\frac{\epsilon}{2}}}, \end{aligned}$$

where M denotes the total number of schedules (that is, the cardinality of $\mathcal{Y}$). Since $u(x)$ tends to $+\infty$ when $|x| \equiv \sum_k x_k$ tends to $+\infty$, this quantity is less than $\epsilon/2$ for all states x but some finite number. In those states, we have:

$$\sum_{y \in \mathcal{Z}(x)} \pi(x, y) \geq 1 - \frac{\epsilon}{2}.$$

We deduce that in all states x but some finite number:

$$\begin{aligned} \sum_{y \in \mathcal{Y}(x)} \pi(x, y) \log(u(x, y)) &\geq (1 - \frac{\epsilon}{2})^2 \log(u(x)), \\ &\geq (1 - \epsilon) \log(u(x)). \end{aligned}$$

■

The result then follows from the stable behavior of maximum weight scheduling, except that the latter is defined over the set of *all* feasible schedules. Defining the corresponding weight by:

$$v(x) = \max_{y \in \mathcal{Y}} u(x, y),$$

the following result shows that it is essentially the same as $u(x)$:

LEMMA 5.3. *We have:*

$$\sup_{x \in \mathcal{X}} \frac{v(x)}{u(x)} < \infty.$$

PROOF. Let:

$$v(x, y) = \prod_{k: x_k \geq J} (x_k \alpha_k)^{y_k}.$$

There are some positive constants m, M such that:

$$\forall x \in \mathbb{N}^K, \forall y \in \mathcal{Y}, \quad m \leq \frac{u(x, y)}{v(x, y)} \leq M.$$

The proof then follows from the fact that:

$$v(x) = \max_{y \in \mathcal{Y}} u(x, y) \leq M \max_{y \in \mathcal{Y}} v(x, y) = M \max_{y \in \mathcal{Y}(x)} v(x, y) \leq \frac{M}{m} \max_{y \in \mathcal{Y}(x)} u(x, y) = \frac{M}{m} u(x).$$

■

The proof of Theorem 1, based on Lemmas 1 and 2, then follows from Foster's criterion.

PROOF. If the vector of traffic intensities lies in the interior of the capacity region, there exist some $\epsilon > 0$ and some probability measure π on $\mathcal{Y}$ such that:

$$(5.2) \quad \forall k = 1, \dots, K, \quad \rho_k = \varphi_k(1 - 2\epsilon) \sum_{y \in \mathcal{Y}} \pi(y) y_k.$$

Note that we can choose $\pi(y) > 0$ for all $y \in \mathcal{Y}$.

Define the Lyapunov function:

$$F(x) = \sum_{k: x_k > 0} \frac{x_k \sigma_k}{\varphi_k} \log(x_k \alpha_k).$$

The corresponding drift is given by:

$$\begin{aligned}\Delta F(x) &= \sum_k \lambda_k (F(x + e_k) - F(x)) + \sum_{k:x_k > 0} \frac{\phi_k(x)}{\sigma_k} (F(x - e_k) - F(x)), \\ &= \sum_{k:x_k=0} \frac{\rho_k}{\varphi_k} \log(\alpha_k) + \sum_{k:x_k > 0} \frac{\rho_k}{\varphi_k} ((x_k + 1) \log((x_k + 1)\alpha_k) - x_k \log(x_k \alpha_k)) \\ &\quad + \sum_{k:x_k > 0} \frac{\phi_k(x)}{\varphi_k} ((x_k - 1) \log((x_k - 1)\alpha_k) - x_k \log(x_k \alpha_k)).\end{aligned}$$

In particular, we have $\Delta F(x) = G(x) + H(x)$ with:

$$\begin{aligned}G(x) &= \sum_{k:x_k > 0} \frac{\rho_k - \phi_k(x)}{\varphi_k} \log(x_k \alpha_k), \\ H(x) &= \sum_{k:x_k > 0} \frac{\rho_k}{\varphi_k} (x_k + 1) \log(1 + \frac{1}{x_k}) + \sum_{k:x_k > 0} \frac{\phi_k(x)}{\varphi_k} (x_k - 1) \log(1 - \frac{1}{x_k}) + \sum_{k:x_k=0} \frac{\rho_k}{\varphi_k} \log(\alpha_k),\end{aligned}$$

where we use the convention $0 \log(0) \equiv 0$. Since $\phi_k(x) \leq J\varphi_k$, the function $H(x)$ is bounded. Regarding $G(x)$, it follows from (5.1) and (5.2) that:

$$\begin{aligned}G(x) &= \sum_{y \in \mathcal{Y}} ((1 - 2\varepsilon)\pi(y) - \pi(x, y)) \sum_{k:x_k > 0} y_k \log(x_k \alpha_k), \\ &= \sum_{y \in \mathcal{Y}} ((1 - 2\varepsilon)\pi(y) - \pi(x, y)) \log(u(x, y)).\end{aligned}$$

By Lemma 1, we have for all states x but some finite number:

$$\begin{aligned}G(x) &\leq -\epsilon \sum_{y \in \mathcal{Y}} \pi(y) \log(u(x, y)) + (1 - \epsilon) \left(\sum_{y \in \mathcal{Y}} \pi(y) \log(u(x, y)) - \log(u(x)) \right), \\ &\leq -\epsilon \sum_{y \in \mathcal{Y}} \pi(y) \log(u(x, y)) + (1 - \epsilon) \log \left(\frac{v(x)}{u(x)} \right).\end{aligned}$$

Since $\pi(y) > 0$ for all $y \in \mathcal{Y}$, the first term tends to $-\infty$ when $|x| \equiv \sum_k x_k$ tends to $+\infty$. By Lemma 2, the second term is bounded. We deduce the existence of some $\delta > 0$ such that $\Delta F(x) \leq -\delta$ for all states x but some finite number. The proof then follows from Foster's criterion. $\blacksquare$

6. Time-Scale Separation

Theorem 5.1 is based on the time-scale separation assumption: in the packet-level model of section 4, packets “see” a fixed number of flows, while in the flow-level model of section 5, flows “see” the equilibrium state of packet-level dynamics. In this section, we remove this assumption. Specifically, we prove that when the size of the flows grows, the model without time-scale separation converges to the model with time-scale separation, over any finite-time horizon. While not proving it, this suggests that CSMA is optimal for sufficiently large flow sizes. We actually conjecture that CSMA is optimal for any flow size, which we prove at the end of the section for a specific class of networks.

6.1. Scaling. As in section 5, class- k flows are assumed to arrive according to a Poisson process of intensity λ_k . The number of packets per class- k flow has a geometric distribution with mean $N\sigma_k$, where N is some positive integer, that we refer to as the scaling parameter. In particular, each class- k flow terminates with probability $1/(\sigma_k N)$ after each packet transmission. Packets are assumed to have an exponential size with mean $1/N$ bits, so as to keep the class- k mean flow size

constant and equal to σ_k bits. In particular, the corresponding traffic intensity $\rho_k = \lambda_k \sigma_k$ is independent of N .

The random access algorithm is that described in section 4.1. The only difference is that the attempt rates must be scaled so as to keep the ratio of mean packet transmission time to mean backoff time constant. Thus each class- k link now attempts to access the channels at rate $N\nu_k$.

6.2. Asymptotic Time-Scale Separation. The state of the network is now described by the tuple $(X^N(t), Y^N(t))$, where $X^N(t)$ gives the number of flows of each class at time t and $Y^N(t)$ the schedule that is selected at time t . This is a Markov process with transition rates λ_k from state (x, y) to state $(x + e_k, y)$ (class- k flow arrival), $N(x_k - y_k)\nu_k\beta_{kj}$ from state (x, y) to state $(x, y + e_{kj})$ (access to channel j by a class- k flow), $Ny_{kj}\varphi_k(1 - 1/(\sigma_k N))$ from state (x, y) to state $(x, y - e_{kj})$ (packet transmission of a class- k flow over channel j , without flow completion), $y_{kj}\varphi_k/\sigma_k$ from state (x, y) to state $(x - e_k, y - e_{kj})$ (packet transmission of a class- k flow over channel j , with flow completion).

When N grows, the packet-level dynamics, represented by $(Y^N(t))$, are accelerated with respect to the flow-level dynamics, represented by $(X^N(t))$. The following result, proved in the appendix, shows that there is indeed time-scale separation between the packet level and the flow level in the limit. We assume that $X^N(0) = X(0)$ for all $N \geq 1$.

THEOREM 6.1. *When $N \rightarrow \infty$, the stochastic process $(X^N(t))$ converges in distribution to the Markov process $(X(t))$, which describes the network state under the time-scale separation assumption.*

PROOF. In the following, we consider $((X^N(t)))_{N \geq 1}$ as a sequence of stochastic processes in the space $\mathcal{D}_{\mathbb{N}^K}([0, \infty])$ of càd-làg functions with values in $\mathbb{N}^K$ with the Skorohod topology.

First, we have to prove the tightness of the sequence $((X^N(t)))$. It is enough to remark that, for all $N \geq 1$, $(X_k^N(t))$ is stochastically dominated by a Poisson process of intensity λ_k and stochastically dominates an $M/M/1$ queue with arrival rate λ_k and service rate φ_k/σ_k . Thus, the conditions of the Arzelà-Ascoli theorem are fulfilled and the sequence $((X^N(t)))$ is tight (see [Bil99, Th 12.3]).

We now consider a bounded function f on $\mathbb{N}^K$. Denote by Ω^N the infinitesimal generator of the Markov process $(X^N(t), Y^N(t))$. For all $x \in \mathbb{N}^K$ and $y \in \mathcal{Y}$, we have

$$\Omega^N(f)(x, y) = \sum_{k=1}^K \lambda_k (f(x + e_k) - f(x)) - \sum_{k=1}^K \varphi_k / \sigma_k \sum_{j=1}^J y_{kj} (f(x - e_k) - f(x)).$$

Note that $\Omega^N(f)(x, y)$ does not depend on N . We then define $\Omega^\infty(f)(x, y) = \Omega^N(f)(x, y)$. According to the Martingale characterization of Markov jump processes (see [RW87]), the process:

$$(M_f^N(t)) = \left(f(X^N(t)) - f(X^N(0)) - \int_0^t \Omega^N(f)(X^N(s), Y^N(s)) ds \right)$$

is a locale martingale and, since the process $(X^N(t))$ is not exploding on $[0, t]$ (it is stochastically dominated by a Poisson process), it is a martingale.

For each $N \geq 1$, define the random measure:

$$\Gamma^N([0, t] \times B) = \int_0^t \mathbb{1}_{\{Y(s) \in B\}} ds, \quad \text{for all } B \subset \mathcal{Y}.$$

Γ^N is a random variable with value in the set $\mathcal{L}(\mathcal{Y})$ of the random measures on $[0, \infty[\times \mathcal{Y}$ such that if $\mu \in \mathcal{L}(\mathcal{Y})$ then $\mu([0, t] \times \mathcal{Y}) = t$ for all $t \geq 0$. Since $\mathcal{Y}$ is finite, the set $\mathcal{L}(\mathcal{Y})$ is compact and then the sequence $(\Gamma^N)_{N \geq 1}$ is relatively compact.

Assume that the sequence $((X^N(t)), \Gamma^N)_{N \geq 1}$ tends to some limit $((X(t)), \Gamma)$. Since:

$$\int_0^t \Omega^N(f)(X^N(s), Y^N(s)) \, ds = \int_0^t \sum_{y \in \mathcal{Y}} \Omega^N(f)(X^N(s), y) \Gamma^N(ds \times dy)$$

and f is bounded, this random variable tends in distribution to:

$$\int_0^t \sum_{y \in \mathcal{Y}} \Omega^\infty(f)(X(s), y) \Gamma(ds \times dy).$$

It remains to characterize Γ . According to Lemma 1.3 of [Kur92], there exists a set of random probability measures $\vartheta(t, \cdot)$ on $\mathcal{Y}$ such that:

$$\Gamma([0, t] \times B) = \int_0^t \vartheta(s, B) \, ds, \quad \text{for } B \subset \mathcal{Y}.$$

For any function g on $\mathcal{Y}$, we define the martingale:

$$(\bar{M}_g^N(t)) = \left(\frac{1}{N} \left(g(Y^N(t)) - g(Y^N(0)) - \int_0^t \Omega^N(g)(X^N(s), Y^N(s)) \, ds \right) \right).$$

For $x \in \mathbb{N}^K$ and $y \in \mathcal{Y}$, we have:

$$\begin{aligned} \Omega^N(g)(x, y) &= \sum_{k=1}^K \sum_{j=1}^J N(x_k - y_k) \nu_k \beta_{kj} (g(y + e_{kj}) - g(y)) \\ &\quad + \left(N y_{kj} \varphi_k \left(1 - \frac{1}{\sigma_k N} \right) + \frac{\varphi_k}{\sigma_k} \right) (g(y - e_{kj}) - g(y)). \end{aligned}$$

The increasing process of this martingale is:

$$\begin{aligned} \langle \bar{M}_g^N(t) \rangle &= \frac{1}{N^2} \int_0^t \Omega^N(g)(X^N(s), Y^N(s)) \, ds, \\ &\leq \frac{2t}{N} \max_{y \in \mathcal{Y}} |g(y)| \left(\max_k \varphi_k + \max_k \nu_k \max_{k,j} \beta_{kj} \right). \end{aligned}$$

It tends to 0 on all compact sets so that the martingale tends in distribution to 0. Since $\mathcal{Y}$ is finite, g is bounded and $(g(Y^N(t)) - g(Y^N(0)))/N$ also tends to 0. Finally, we get that:

$$\frac{1}{N} \int_0^t \Omega^N(g)(X^N(s), Y^N(s)) \, ds$$

converges in distribution to 0. This implies:

$$\begin{aligned} \int_0^t \sum_{y \in \mathcal{Y}} \left(\sum_{k=1}^K \sum_{j=1}^J (X_k(s) - y_k) \nu_k \beta_{kj} (g(y + e_{kj}) - g(y)) \right. \\ \left. + y_{kj} \varphi_k (g(y - e_{kj}) - g(y)) \right) \vartheta(s, y) \, ds = 0 \end{aligned}$$

and for almost every s in $[0, t]$, we have:

$$\begin{aligned} \sum_{y \in \mathcal{Y}} \left(\sum_{k=1}^K \sum_{j=1}^J (X_k(s) - y_k) \nu_k \beta_{kj} (g(y + e_{kj}) - g(y)) \right. \\ \left. + y_{kj} \varphi_k (g(y - e_{kj}) - g(y)) \right) \vartheta(s, y) = 0 \end{aligned}$$

The probability distribution $\vartheta(s, \cdot)$ is then the stationary distribution given by (4.1).

It follows that:

$$\int_0^t \Omega^N(f)(X^N(s), Y^N(s)) \, ds$$

converges in distribution to:

$$\int_0^t \Omega(f)(X(s)) \, ds$$

where Ω is the infinitesimal generator of the Markov process described in section 5. For $x \in \mathbb{N}^K$, we have

$$\Omega(f)(x) = \sum_{k=1}^K \lambda_k (f(x + e_k) - f(x)) + \phi_k(x) / \sigma_k (f(x - e_k) - f(x)),$$

where $\phi_k(x)$ is the mean throughput of class k in state x , given by (4.2).

By dominated convergence, $(M_f^N(t))$ tends in distribution to:

$$(M_f(t)) = \left(f(X(t)) - f(X(0)) - \int_0^t \Omega(f)(X(s)) \, ds \right),$$

and $(M_f(t))$ is a martingale. Using the characterization of the Markov jump processes, we get that the process $(X(t))$ is a Markov process with infinitesimal generator Ω .

This concludes the proof. ■

6.3. Stability of some Class of Networks. Theorem 6.1 justifies the time scale separation for sufficiently large flows at finite time horizon and thus does not implies anything about stability of CSMA without time-scale separation which typically occurs at infinite time horizon. Theorems 5.1 and 6.1 suggest that CSMA is optimal for sufficiently large flow sizes. We conjecture that CSMA is actually optimal for any flow size, in the sense that the Markov process $(X^N(t), Y^N(t))$ is ergodic for any scaling parameter $N \geq 1$ provided the vector of traffic intensities ρ lies in the interior of the capacity region. To support this conjecture, consider the following class of networks. We assume that all links have access to the J channels. The interference graph is the same on all channels and given by some L -partite graph, i.e. there exists a partition $\{C_1, \dots, C_L\}$ of $\{1, \dots, K\}$ such that two classes in C_l do not interfere with each other but a class in C_l does interfere with all classes in $\{1, \dots, K\} \setminus C_l$. Examples of L -partite graphs are given in figure 2. The following result shows that CSMA is optimal independently of the scaling parameter N :

PROPOSITION 6.2. *Any network with a L -partite interference graph is stable for all vectors of traffic intensities ρ in the interior of the capacity region.*

PROOF. For this proof, we need fluid limits. For any $x \in \mathbb{R}_+^K$, we define $|x| \equiv \sum_{k=1}^K x_k$. For all $n \geq 1$, let $(X^{N,n}(t), Y^{N,n}(t))$ be the evolution of the Markov process when starting from some initial state such that $|X^{N,n}(0)| = n$. A fluid limit is a limiting point $(\bar{X}^N(t))$ of the laws of the processes $\{(X^{N,n}(nt)/n)\}_{n \geq 1}$ in the set of probability measures on the space $\mathcal{D}$ of *càd-làg* functions on $\mathbb{R}_+$ with values in $\mathbb{R}_+^K$ with the Skorohod topology [Bil99]. It is not difficult to show that the set of processes $\{(X^{N,n}(nt)/n), n \geq 1\}$ is tight in the set of probability distributions on the space $\mathcal{D}$ endowed with the metric associated with the uniform norm on compact sets. Therefore, there exists at least one fluid limit and any fluid limit is continuous. Since the process $(Y^{N,n}(nt))$ has its values in a finite space for all $n \geq 1$, it can be proved as in [Dai95, Rob03] that, if there exists a deterministic time $T > 0$ such that $\bar{X}^N(t) = 0$ for all $t \geq T$, then the Markov process $(X^N(t), Y^N(t))$ is ergodic.

The proof is then very similar to that given in [FPR10] for random capture algorithms. Consider a fluid limit $(\bar{X}^N(t))$. We say that a class k is non-empty at time t if $\bar{X}_k^N(t) > 0$. As long as there is a non-empty class, the J channels are used. Moreover, if some class in C_l takes channel j , all other non-empty classes in C_l use this channel. Let $\alpha_{jl}(t)$ be the fraction of time channel j is used by classes in C_l at time t . For any non-empty class $k \in C_l$, we have:

$$(6.1) \quad \frac{d\bar{X}_k^N}{dt}(t) = \lambda_k - \frac{\varphi_k}{\sigma_k} \sum_{j=1}^J \alpha_{jl}(t).$$

Now define:

$$W^N(t) = \sum_{l=1}^L \max_{k \in C_l} \left(\bar{X}_k^N(t) \frac{\sigma_k}{\varphi_k} \right).$$

Note that $W^N(t) = 0$ if and only if $\bar{X}^N(t) = 0$. Moreover,

$$W^N(0) \leq M \equiv \sum_{l=1}^L \max_{k \in C_l} \left(\frac{\sigma_k}{\varphi_k} \right).$$

Using (6.1) and the fact that:

$$\sum_{j=1}^J \sum_{l=1}^L \alpha_{jl}(t) = J,$$

at any time t such that $W^N(t) > 0$, we deduce:

$$(6.2) \quad W^N(t) \leq \max \left(0, M + \left(\sum_{l=1}^L \max_{k \in C_l} \frac{\rho_k}{\varphi_k} \right) t - Jt \right).$$

In L -partite networks, the capacity region is given by the throughput vectors ϕ for which there exist a set $\{\pi_1, \dots, \pi_J\}$ of probability distributions over $\{C_1, \dots, C_L\}$ such that:

$$\forall k \in C_l, \quad \phi_k \leq \varphi_k \sum_{j=1}^J \pi_j(C_l).$$

In particular,

$$\sum_{l=1}^L \max_{k \in C_l} \frac{\phi_k}{\varphi_k} \leq \sum_{l=1}^L \sum_{j=1}^J \pi_j(C_l) = J.$$

Since ρ lies inside the capacity region, it follows from (6.2) that $W^N(t) = 0$ for all $t \geq T$, with:

$$T = \frac{1}{J - \sum_{l=1}^L \max_{k \in C_l} \rho_k},$$

which implies the ergodicity of the Markov process $(X^N(t), Y^N(t))$. ■

7. Infrastructure-Based Networks

We have so far considered a network in ad-hoc mode, without infrastructure. We now consider A access points to which users must connect. In particular, each class now corresponds either to uplink traffic (from the users to an access point) or to downlink traffic (from an access point to the users). We study the flow-level dynamics of CSMA under the time-scale separation assumption. Specifically, we prove the suboptimality of standard CSMA in this context and introduce a slight modification of CSMA, we refer to as *flow-aware* CSMA, which makes the algorithm optimal.

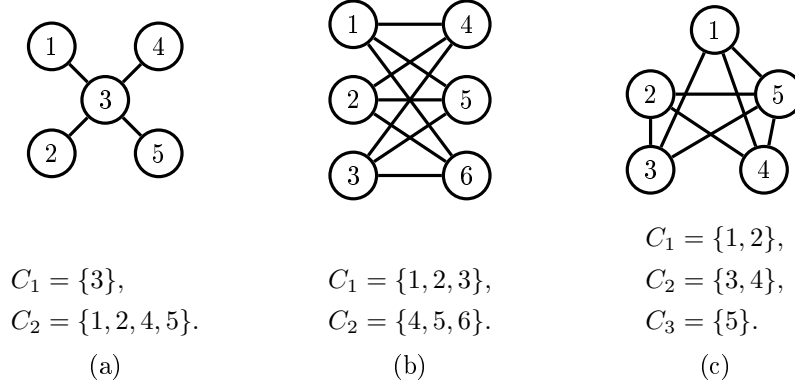


FIGURE 2. Examples of 2-partite (a)-(b) and 3-partite (c) graphs.

7.1. Uplink vs. Downlink. For all $i = 1, \dots, A$, we denote by U_i and D_i the sets of uplink and downlink classes, respectively, associated with access point i . In the example of figure 3, for instance, there are $A = 2$ access points and $K = 6$ classes, with $U_1 = \{2\}$, $D_1 = \{1, 3\}$, $U_2 = \{5\}$ and $D_2 = \{4, 6\}$. An access point cannot transmit and receive simultaneously on the same channel. In particular, those classes sharing the same access point, either in uplink or downlink, conflict with each other. Formally, for all access points $i = 1, \dots, A$ and all classes $k, l \in U_i \cup D_i$, we have $(k, l) \in E_j$ for each channel j such that $k, l \in V_j$. We assume that an access point cannot transmit data on more than one channel at a time but is able to receive data on the J channels simultaneously (or on a subset of these channels, i.e. those channels j such that $U_i \cap V_j \neq \emptyset$ for access point i).

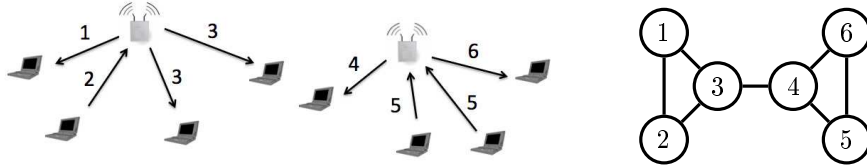


FIGURE 3. A network of 2 access points with 6 classes of links and its interference graph.

The feasible schedules are those defined in section 3.2, with the additional constraint that each access point cannot transmit data on more than one channel at a time, that is:

$$(7.1) \quad \forall i = 1, \dots, A, \quad \sum_{k \in D_i} y_k \leq 1.$$

We denote by $\mathcal{Y}(x)$ the set of feasible schedules and by $\mathcal{Y}$ the union of $\mathcal{Y}(x)$ over all network states x . The corresponding capacity region is defined in section 3.3.

7.2. Standard CSMA. We first consider the standard CSMA algorithm: each transmitter waits for a period of random duration before attempting transmission on some randomly chosen channel. The key difference with the ad-hoc wireless network considered so far is that each access point runs a single instance of the CSMA algorithm for all its downlink traffic. In particular, for each access point i , the attempt rates ν_k are the same for all classes $k \in D_i$. At each attempt, the access point i selects a class- k flow with some probability proportional to x_k and

probes channel j with probability β_{kj} . If the probed channel is sensed idle, a packet of this flow is transmitted.

It is worth noting that the attempt rate of each access point is independent of its congestion level, in terms of the number of ongoing downlink flows at this access point. This breaks the natural stabilizing effect of CSMA we have proven in Theorem 5.1 in the context of ad-hoc networks, where those classes with a higher number of flows get preferential access to the radio channels. In the following, we illustrate the suboptimality of standard CSMA on two examples with downlink traffic only. Note that, in the presence of uplink traffic only, the model is in fact equivalent to the ad-hoc network considered so far.

For this purpose, we give the distribution of feasible schedules achieved by the algorithm under the time-scale separation assumption. Denoting by $(Y(t))$ the schedule at time t , we have the analogue of Proposition 4.1:

PROPOSITION 7.1. *If both the packet sizes and the backoff times have exponential distributions, then $Y(t)$ is a reversible Markov process, with stationary measure:*

$$(7.2) \quad w(x, y) = \prod_{i=1}^A \prod_{k \in U_i: x_k > 0} \frac{x_k!}{(x_k - y_k)!} \alpha_k^{y_k} \prod_{j=1}^J \beta_{kj}^{y_{kj}} \\ \times \prod_{k \in D_i: x_k > 0} \left(\frac{x_k \alpha_k}{\sum_{l \in D_i} x_l} \right)^{y_k} \prod_{j=1}^J \beta_{kj}^{y_{kj}}, \quad y \in \mathcal{Y}(x).$$

PROOF. As for Proposition 4.1, the proof follows from the local balance equations. For all $i, \dots, A$, we have for uplink classes:

$$\forall k \in U_i, \quad w(x, y)(x_k - y_k)\nu_k\beta_{kj} = w(x, y + e_{kj})\varphi_k,$$

while for downlink classes:

$$\forall k \in D_i, \quad w(x, y) \frac{x_k}{\sum_{l \in D_i} x_l} \nu_k \beta_{kj} = w(x, y + e_{kj})\varphi_k.$$

■

The stationary distribution of the schedules $\pi(x, y)$ follows from normalization. Again, it is insensitive to the packet size and backoff time distributions beyond the means. The throughput of class k is given by (4.2).

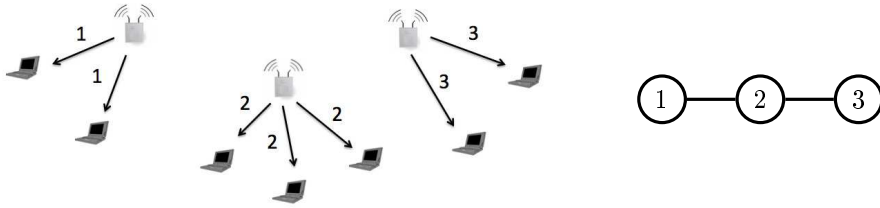


FIGURE 4. Network of 3 access points with a single downlink class per access point and its interference graph.

7.2.1. Example 1. The simplest example showing the suboptimality of CSMA is shown in Figure 4. It consists of $A = 3$ access points, a single class per access point and a single channel. Taking unit physical rates, the optimal stability region is $\rho_1 + \rho_2 < 1$ and $\rho_2 + \rho_3 < 1$ where 1 and 3 are the edge classes and 2 is the center class. Assume for simplicity all links have the same mean packet sizes and

mean backoff times, so that $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$ for some $\alpha > 0$. In view of (4.2) and (7.2), the throughput of the links in state x are given by:

$$\phi_1(x) = \begin{cases} \frac{\alpha}{1+\alpha} & \text{if } x_2 = 0, \\ \frac{\alpha}{1+2\alpha} & \text{if } x_2 > 0, x_3 = 0, \\ \frac{\alpha + \alpha^2}{1+3\alpha + \alpha^2} & \text{if } x_2 > 0, x_3 > 0, \end{cases}$$

and

$$\phi_2(x) = \begin{cases} \frac{\alpha}{1+\alpha} & \text{if } x_1 = 0, x_3 = 0, \\ \frac{\alpha}{1+2\alpha} & \text{if } x_1 > 0, x_3 = 0, \text{ or } x_1 = 0, x_3 > 0, \\ \frac{\alpha}{1+3\alpha + \alpha^2} & \text{if } x_1 > 0, x_3 > 0. \end{cases}$$

The throughput of link 3 follows by symmetry. As for a single link, the backoff times must be chosen sufficiently small to limit the overhead of the algorithm. In the limit $\alpha \rightarrow \infty$, we get:

$$(7.3) \quad \phi(x) = \begin{cases} (1, 0, 1) & \text{if } x_1 > 0, x_3 > 0, \\ (1/2, 1/2, 0) & \text{if } x_1 > 0, x_2 > 0, x_3 = 0, \\ (1, 0, 0) & \text{if } x_1 > 0, x_2 = 0, x_3 = 0, \\ (0, 1, 0) & \text{if } x_1 = 0, x_2 > 0, x_3 = 0, \end{cases}$$

the other cases following by symmetry. Note that link 2 is not served when both links 1 and 3 are active. This is due to the fact that link 2 is in conflict with both links 1 and 3 and thus cannot access the channel for an infinitely small backoff time. This results in a suboptimal stability region:

PROPOSITION 7.2. *The stability region is given by:*

$$\rho_1 < \frac{1+\rho_3}{2}, \quad \rho_3 < \frac{1+\rho_1}{2}, \quad \rho_2 < \pi_0 + \frac{\pi_{1,3}}{2},$$

or

$$\rho_1 < \frac{1+\rho_3}{2}, \quad \frac{1+\rho_1}{2} \leq \rho_3 < \frac{1+\rho_1}{2} + \frac{1-\rho_1}{2}\pi_{2,1}, \quad \rho_2 < \frac{1-\rho_1}{2},$$

or

$$\rho_3 < \frac{1+\rho_1}{2}, \quad \frac{1+\rho_3}{2} \leq \rho_1 < \frac{1+\rho_3}{2} + \frac{1-\rho_3}{2}\pi_{2,3}, \quad \rho_2 < \frac{1-\rho_3}{2},$$

where $\pi_0, \pi_{1,3}, \pi_{2,1}$ and $\pi_{2,3}$ are the respective probabilities that:

- both links 1 and 3 are idle when link 2 is always active;
- one of the links 1 or 3 is idle when link 2 is always active;
- link 2 is idle given that link 1 is idle, when link 3 is always active;
- link 2 is idle given that link 3 is idle, when link 1 is always active.

More precisely, the Markov process $X(t)$ is positive recurrent if the vector of traffic intensities ρ lies in this region and transient if it lies outside its closure.

PROOF. This example is similar to the one studied in [Rob03, p274]. We consider the fluid limits of the Markov process $(X(t))$. Specifically, we define $(X^{(n)}(t))$ as the Markov process $X(t)$ whose initial state is $X^{(n)}(0) = (\lfloor \beta_1 n \rfloor, \lfloor \beta_2 n \rfloor, \lfloor \beta_3 n \rfloor)$ for some non-negative real numbers $\beta_1, \beta_2, \beta_3$ such that $\beta_1 + \beta_2 + \beta_3 = 1$. We then define:

$$\bar{X}^{(n)}(t) = \frac{1}{n} X^{(n)}(nt).$$

The fluid limits of the Markov process $(X(t))$, if they exist, are the limiting points of this set of processes when $n \rightarrow +\infty$. It is easy to check that the Markov process $(X(t))$ belongs to the class $(\mathcal{C})$ defined in [Rob03, p241] and that the associated

Proposition 9.3 applies. In particular, the set $\{(\bar{X}^{(n)}(t)), n \in \mathbb{N}\}$ is tight and the fluid limits are continuous. The Markov process $(X(t))$ is then positive recurrent if there exists some finite time after which all fluid limits are null, cf. [Rob03, Theorem 9.7, p259]; it is transient if, after some finite time, for any initial state $\beta_1, \beta_2, \beta_3$, there are some components of the fluid limits grow at least linearly to infinity [Mey95].

We first calculate the fluid limit until the first time where one component reaches 0, if any, for all possible initial states. The three components of the process $(X^{(n)}(t))$ behave as three coupled $M/M/1$ queues, with arrival rates $\lambda_1, \lambda_2, \lambda_3$ and state-dependent service rates. We denote by $\mu_k = 1/\sigma_k$ the maximum service rate of queue k , so that $\rho_k = \lambda_k/\mu_k$. The Markov process is positive recurrent if all queues empty in finite time in the limit and transient if, starting from any initial state, at least one queue grows linearly to infinity after some finite time.

We start with the case $\beta_1 > 0, \beta_2 > 0, \beta_3 > 0$. The three queues are then mutually independent, with respective service rates $\mu_1, 0, \mu_3$. The scaling property of the $M/M/1$ queue shows that the process $(\bar{X}^{(n)}(t))$ weakly converges to the function:

$$(\beta_1 + (\lambda_1 - \mu_1)t, \beta_2 + \lambda_2 t, \beta_3 + (\lambda_3 - \mu_3)t),$$

until one of the components reaches 0, if any.

We now consider the case $\beta_1 = 0, \beta_2 > 0, \beta_3 > 0$. In view of (7.3), queue 1 has service rate μ_1 and is empty with probability $1 - \rho_1$. Queues 2 and 3 have service rates 0, μ_3 with probability ρ_1 and $\mu_2/2, \mu_3/2$ with probability $1 - \rho_1$. Proposition 9.14 of [Rob03] applies and the process $(\bar{X}^{(n)}(t))$ weakly converges to the function:

$$(0, \beta_2 + (\lambda_2 - \mu_2 \frac{1 - \rho_1}{2})t, \beta_3 + (\lambda_3 - \mu_3 \frac{1 + \rho_1}{2})t),$$

until one of the components reaches 0, if any.

Next, we consider the case $\beta_1 = \beta_2 = 0, \beta_3 > 0$. In view of (7.3), queue 1 has service rate μ_1 . Queue 2 has service rate $\mu_2/2$ if queue 1 is empty and 0 otherwise. This queue is stable if $\rho_2 < (1 - \rho_1)/2$, which we assume. Queue 2 then remains empty in the limit, and the service rate of queue 3 is equal to μ_3 with probability $\rho_1 + (1 - \rho_1)\pi_{2,1}$ and to $\mu_3/2$ otherwise. We deduce that the process $(\bar{X}^{(n)}(t))$ weakly converges to the function:

$$(0, 0, \beta_3 + (\lambda_3 - \mu_3(\frac{\rho_1 + 1}{2} - \frac{1 - \rho_1}{2}\pi_{2,1}))t),$$

whenever component 3 is positive.

Finally, we consider the case $\beta_1 = \beta_3 = 0, \beta_2 > 0$. In view of (7.3), the service rates of queues 1 and 3 are equal to μ_1 and μ_3 when both are non-empty and to $\mu_1/2$ and $\mu_3/2$ otherwise. This system is stable if $\rho_1 < (1 + \rho_3)/2$ and $\rho_3 < (1 + \rho_1)/2$, which we assume. Queues 1 and 3 then remain empty in the limit. The service rate of queue 3 is equal to μ_2 with probability π_0 and to $\mu_2/2$ with probability $\pi_{1,3}$. The process $\bar{X}^{(n)}(t)$ weakly converges to the function:

$$(0, \beta_2 + (\lambda_2 - \mu_2(\pi_0 - \frac{\pi_{1,3}}{2}))t, 0),$$

whenever component 2 is positive.

To conclude the proof, we consider the evolution of the fluid limit in the following five cases (the others follow by symmetry):

- (1) Assume $\rho_1 < (1 + \rho_3)/2$ and $\rho_3 < (1 + \rho_1)/2$. Note that this implies $\rho_1 < 1$ and $\rho_3 < 1$. Queues 1 and 3 empty in finite time, independently of queue 2. Queue 2 then empties in finite time if $\rho_2 < \pi_0 + \pi_{1,3}/2$; it grows linearly to infinity if $\rho_2 > \pi_0 + \pi_{1,3}/2$.

- (2) Assume $\rho_1 < (1 + \rho_3)/2$ and $\rho_3 > (1 + \rho_1)/2$. If $\rho_1 \geq 1$ then $\rho_3 > 1$ and queue 3 grows linearly to infinity. We now assume $\rho_1 < 1$. If $\rho_2 > (1 - \rho_1)/2$ then queue 2 grows linearly to infinity. If $\rho_2 = (1 - \rho_1)/2$ then starting from a state where $\beta_1 = 0$, $\beta_2 > 0$ and $\beta_3 > 0$, queue 1 stays empty, queue 2 is constant and queue 3 grows linearly to infinity. We assume that $\rho_1 < 1$ and $\rho_2 < (1 - \rho_1)/2$. Starting from the initial state $\beta_1 = \beta_2 = 0$, $\beta_3 > 0$, queue 3 grows linearly to infinity if $\rho_3 > (1 + \rho_1)/2 + \pi_{2,1}(1 - \rho_1)/2$. We assume that $\rho_3 < (1 + \rho_1)/2 + \pi_{2,1}(1 - \rho_1)/2$. Starting from the initial state $\beta_1 = \beta_2 = 0$, $\beta_3 > 0$, queue 3 then empties in finite time. It remains to prove that, starting from any initial state, queues 1 and 2 empty in finite time. We first note that, since $\rho_1 < 1$ and $\rho_3 < 1$, queue 1 or queue 3 empties in finite time. Moreover, if both queues 1 and 3 are empty but not queue 2, then queue 3 grows linearly. Thus we can assume that queue 1 empties before queue 3. We know that queue 2 empties in finite time in this case.
- (3) Assume $\rho_1 < (1 + \rho_3)/2$ and $\rho_3 = (1 + \rho_1)/2$. Note that $\rho_1 < 1$ and $\rho_3 < 1$ in this case. Moreover, we have $\pi_0 = 0$ and $\pi_{1,3} = 1 - \rho_1$, so that the inequality $\rho_2 < \pi_0 + \pi_{1,3}/2$ is equivalent to $\rho_2 < (1 - \rho_1)/2$. If the latter is satisfied, then if queue 1 is non-empty then queue 2 empties in finite time independently of queue 3. We just have to consider the case where $\beta_1 = \beta_2 = 0$ and $\beta_3 > 0$. Because $\rho_3 = (1 + \rho_1)/2 < (1 + \rho_1)/2 + \pi_{2,1}(1 - \rho_1)/2$, queue 3 empties in finite time. If $\rho_2 > (1 - \rho_1)/2$, we choose an initial state such that queue 1 empties before 3. When queue 1 is empty, queue 3 is constant and queue 2 grows linearly to infinity.
- (4) Assume $\rho_1 \geq (1 + \rho_3)/2$ and $\rho_3 > (1 + \rho_1)/2$. Then $\rho_1 > 1$ and $\rho_3 > 1$ so that queues 1 and 3 grow linearly to infinity.
- (5) Assume $\rho_1 = (1 + \rho_3)/2$ and $\rho_3 = (1 + \rho_1)/2$. Then $\rho_1 = \rho_3 = 1$ and $\pi_0 = \pi_{1,3} = 0$. If $\rho_2 = 0$, the vector ρ lies on the boundary of the stability region. If $\rho_2 > 0$, queue 2 grows linearly to infinity.

■

Note that, when one of the links is always active, the two other links form a coupled system of two queues as considered by Fayolle and Iasnogorodski [F179]. In particular, the stability region can be calculated exactly. In the symmetric case $\rho_1 = \rho_3$, the stability condition reduces to $\rho_1 < 1$, $\rho_2 < \pi_0 + \pi_{1,3}/2$. Figure 5 shows the corresponding stability region for equal mean flow sizes.

7.2.2. Example 2. Consider the multi-channel network of Figure 6 with $A = 5$ access points, a single class per access point and $J = 2$ channels, further referred to as the *bow tie network*. The conflict graph is the same for both channels. We refer to class 3 as the center class and to the other classes as the edge classes. We assume that the mean packet sizes and the mean backoff times are the same for all classes, so that $\alpha_k = \alpha$ for all $k = 1, \dots, 5$, for some $\alpha > 0$. We also assume that all classes except class 3 have the same traffic intensities. The optimal stability condition is then given by:

$$(7.4) \quad \rho_3 < 1 \quad \text{and} \quad 2\rho_1 + \rho_3 < 2.$$

We consider the limiting case where $\alpha \rightarrow \infty$ and we assume that the two channels are chosen uniformly at random. We then deduce from (4.1)-(4.2) the

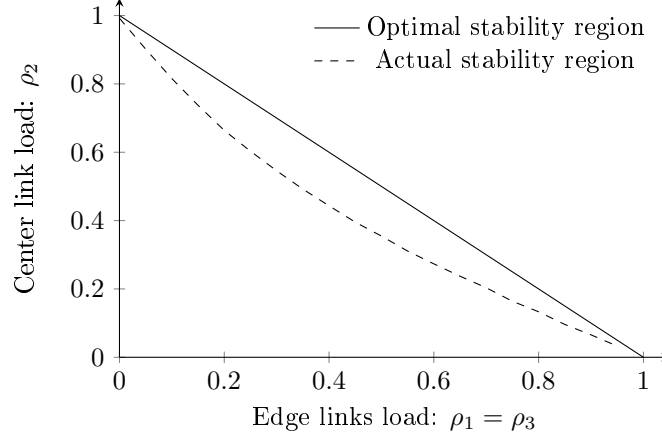


FIGURE 5. Stability condition for the network of Fig. 1 under standard CSMA ($\rho_1 = \rho_3$).

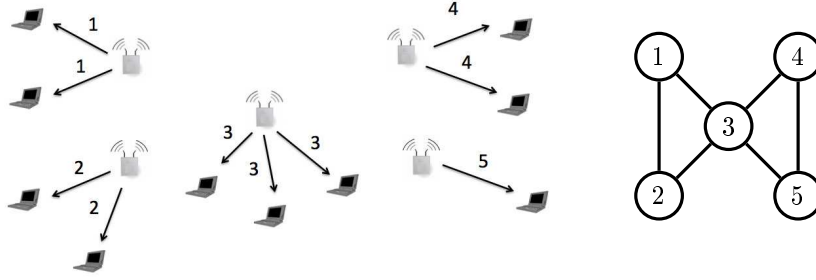


FIGURE 6. Network of 5 access points with a single downlink class per access point and its interference graph.

following throughput vector:

$$(7.5) \quad \phi(x) = \begin{cases} (1, 1, 0, 1, 1) & \text{if } x_1, x_2, x_4, x_5 > 0, \\ (\frac{3}{4}, \frac{3}{4}, \frac{1}{2}, 1, 0) & \text{if } x_1, x_2, x_3, x_4 > 0, x_5 = 0, \\ (\frac{2}{3}, \frac{2}{3}, \frac{2}{3}, 0, 0) & \text{if } x_1, x_2, x_3 > 0, x_4 = x_5 = 0, \\ (0, 1, 1, 1, 0) & \text{if } x_2, x_3, x_4 > 0, x_1 = x_5 = 0, \\ (1, 1, 0, 0, 0) & \text{if } x_1, x_2 > 0, x_3 = x_4 = x_5 = 0, \\ (1, 0, 0, 0, 0) & \text{if } x_1 > 0, x_2 = 0, x_3 = x_4 = x_5 = 0. \end{cases}$$

The other cases follow by symmetry. The center class is in conflict with all other classes for accessing the channels and is either not served when the 4 other classes are active or served at a low rate when 3 other classes are active. This also results in a suboptimal stability region:

PROPOSITION 7.3. *The bow tie network is unstable whenever:*

$$(7.6) \quad \rho_3 > \frac{1}{3}\rho_1^4 - \frac{2}{3}\rho_1^3 - \frac{2}{3}\rho_1^2 + 1.$$

PROOF. Define the throughput vector $\tilde{\phi}$ such that $\tilde{\phi}_3(x) = \phi_3(x)$ and $\tilde{\phi}_k(x) = \mathbb{1}_{\{x_k > 0\}}$ for all $k \neq 3$. The two following inequalities can be verified:

$$(7.7) \quad \tilde{\phi}_k(x) \geq \phi_k(x), \quad \text{for all } k, x,$$

$$(7.8) \quad \tilde{\phi}_k(x) \geq \tilde{\phi}_k(y), \quad \text{for all } x \leq y \text{ and } k \text{ such that } x_k > 0.$$

Now consider the coupling of the stochastic processes $(X(t))$ and $(\tilde{X}(t))$ describing the evolution of the queues for the throughput vectors ϕ and $\tilde{\phi}$, respectively, starting from the same initial state $X(0) = \tilde{X}(0)$. The inequality (7.7) and the monotonicity property (7.8) imply that $(\tilde{X}(t)) \leq (X(t))$ a.s. at any time $t \geq 0$. In particular, the transience or the null recurrence of $(\tilde{X}(t))$ implies that of $(X(t))$.

For the modified system, queues 1,2,4 and 5 are independent $M/M/1$ queues with load ρ_1 . If $\rho_1 \geq 1$, the Markov process $(\tilde{X}(t))$ is null recurrent or transient. Note that (7.6) then reduces to $\rho_3 \geq 0$.

Assume now that $\rho_1 < 1$. To prove the transience of $(\tilde{X}(t))$, we use fluid limits. Since $\rho_1 < 1$ and queues 1,2,4 and 5 are independent $M/M/1$ queues with load ρ_1 in the modified system, there exists some finite time after which, for any initial conditions, the corresponding components of the fluid limit are null. We then consider the fluid limits with the initial condition $\bar{X}_3(0) = 1$ and $\bar{X}_k(0) = 0$ for all $k \neq 3$. In this case, Proposition 9.14 of [Rob03, p.241] applies and the fluid limit satisfies:

$$\bar{X}_3(t) = 1 + \left(\lambda_3 - \frac{\bar{\phi}_3}{\sigma_3} \right) t,$$

as long as this function is positive, where $\bar{\phi}_3$ is the throughput of link 3 averaged over the states of other links. Since each other link is active with probability ρ_1 , it follows from (7.5) that:

$$\bar{\phi}_3 = \frac{1}{3}\rho_1^4 - \frac{2}{3}\rho_1^3 - \frac{2}{3}\rho_1^2 + 1.$$

In particular, $(\bar{X}_3(t))$ increases linearly to infinity whenever inequality (7.6) is satisfied and, according to [Mey95], the Markov process $(\tilde{X}(t))$ is transient. ■

In the homogeneous case $\rho_1 = \rho_3$ for instance, Proposition 7.3 implies that the network is unstable whenever $\rho_1 > 0.63$. In view of (7.4), the optimal stability condition is $\rho_1 < 2/3$, which shows that the standard CSMA algorithm is not optimal. This suboptimality is illustrated by Fig. 7, the actual stability condition being obtained by the simulation of the underlying Markov process. In the homogeneous case for instance, the loss of efficiency is around 15%.

7.3. Flow-aware CSMA. The flow-aware CSMA algorithm consists for each access point to run one standard CSMA algorithm per flow. This compensates for the inherent bias of standard CSMA against downlink flows and stabilizes the network whenever possible. Indeed, the stationary measure of the schedules is now given by (4.1). The only difference with the ad-hoc wireless network considered in section 5 is the additional constraint (7.1) on the set of feasible schedules. This does not change the proof of Theorem 5.1, showing the optimality of flow-aware CSMA.

8. Conclusion

We have proved that, under the time-scale separation assumption, the distributed scheduling achieved by standard CSMA exploits the radio resources in an optimal way in ad-hoc wireless networks. This is due to the fact that each new flow adds a link to the network, which competes fairly with the other links for accessing the radio channels. This is not the case in the presence of access points, due to the inherent bias of CSMA against downlink flows. A slight modification of CSMA we

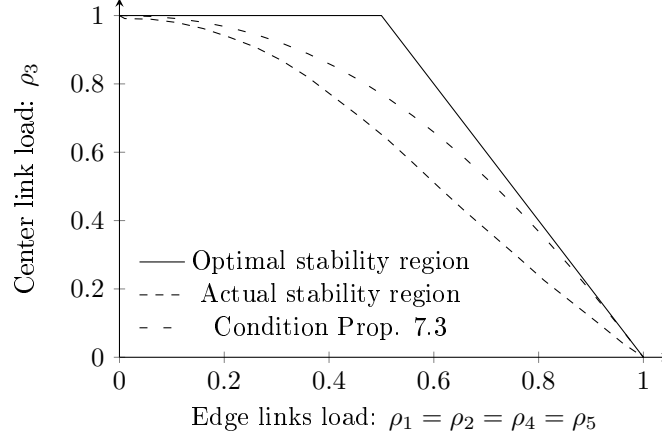


FIGURE 7. Stability region of the bow-tie network with two channels under standard CSMA.

refer to as flow-aware CSMA is then sufficient to correct this bias and to make the algorithm optimal.

Proving stability in the absence of the time-scale separation assumption is a major issue that we plan to address in future work. From a more practical perspective, a number of simplifying assumptions also need to be relaxed. First, we have neglected the impact of packet collisions; these could be included in the model, as done in [JW09] for rate-based adaptive CSMA for instance. One may also account for the adaptive backoff of the IEEE 802.11 protocol, which is essential in practice to limit the number of collisions, and for the presence of TCP acknowledgements. Finally, one may think of multi-hop networks where the flows of some source-destination pairs must go through one or several relay nodes. Although we believe that flow-aware CSMA is still optimal in this more general setting, we have not yet been able to prove this result.

References for Chapter III

- [BBLP11] N. Bouman, S. Borst, J. Van Leeuwen, and A. Proutiere. Backlog-based random access in wireless networks: fluid limits and delay issues. In *ITC 23*, 2011.
- [BBP⁺01] S. Ben Fredj, T. Bonald, A. Proutiere, G. Régnier, and J. W. Roberts. Statistical bandwidth sharing: a study of congestion at flow level. In *Proceedings of ACM SIGCOMM*, pages 111–122, 2001.
- [BF10] T. Bonald and M. Feuillet. On the stability of flow-aware CSMA. *Perform. Eval.*, 67:1219–1229, November 2010.
- [Bil99] Patrick Billingsley. *Convergence of Probability Measures (second edition)*. Wiley Series in Probability and Statistics. Wiley-Interscience, 1999.
- [BK00] A. W. Berger and Y. Kogan. Dimensioning bandwidth for elastic traffic in high-speed data networks. *IEEE/ACM Trans. Netw.*, 8(5):643–654, 2000.
- [Bon07] T. Bonald. Insensitive traffic models for communication networks. *Discrete Event Dynamic Systems*, 17(3):405–421, 2007.

- [BTI⁺03] C. Barakat, P. Thiran, G. Iannaccone, C. Diot, and P. Owezarski. Modeling internet backbone traffic at the flow level. *IEEE Transactions on Signal processing*, 51:2003, 2003.
- [Dai95] J. G. Dai. On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Annals of Applied Probability*, 5:49–77, 1995.
- [FI79] Guy Fayolle and Roudolf Iasnogorodski. Two coupled processors: The reduction to a riemann-hilbert problem. *Probability Theory and Related Fields*, 47(3):325–351, 1979.
- [FPR10] M. Feuillet, A. Proutiere, and P. Robert. Random capture algorithms: Fluid limits and stability. In *Information Theory and Applications Workshop*, February 2010.
- [HRBSD03] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda. Performance anomaly of 802.11b. In *INFOCOM*, volume 2, pages 836–843, 2003.
- [JSSW10] L. Jiang, D. Shah, J. Shin, and J. Walrand. Distributed random access algorithm: Scheduling and congestion control. *IEEE Transactions on Information Theory*, 56(12):6182 – 6207, 2010.
- [JW08] L. Jiang and J. Walrand. A distributed CSMA algorithm for throughput and utility maximization in wireless networks. In *the 46th Annual Allerton Conference on Communication, Control, and Computing*, 2008.
- [JW09] L. Jiang and J. Walrand. Approaching throughput-optimality in a distributed CSMA algorithm: collisions and stability. In *MobiHoc S3’09*, pages 5–8. ACM, 2009.
- [KKF05] S. W. Kim, B.-S. Kim, and Y. Fang. Downlink and uplink resource allocation in IEEE 802.11 wireless LANs. *IEEE Trans. Veh. Technol.*, 54:320–327, Jan. 2005.
- [Kur92] T.G. Kurtz. Averaging for martingale problems and stochastic approximation. In *Applied Stochastic Analysis, US-French Workshop*, volume 177 of *Lecture notes in Control and Information sciences*, pages 186–209. Springer Verlag, 1992.
- [Mey95] Sean Meyn. Transience of multiclass queueing networks via fluid limit models. *Annals of Applied Probability*, 5:946–957, 1995.
- [MR00] L. Massoulié and J. W. Roberts. Bandwidth sharing and admission control for elastic traffic. *Telecommunications Systems*, 15:185–201, 2000.
- [MSZ06] E. Modiano, D. Shah, and G. Zussman. Maximizing throughput in wireless networks via gossiping. *SIGMETRICS Perform. Eval. Rev.*, 34(1):27–38, 2006.
- [NTS10] J. Ni, B. Tan, and R. Srikant. Q-CSMA: Queue-length based CSMA/CA algorithms for achieving maximum throughput and low delay in wireless networks. In *IEEE INFOCOM*, 2010.
- [PYLC10] A. Proutiere, Y. Yi, T. Lan, and M. Chiang. Resource allocation over network dynamics without timescale separation. In *IEEE INFOCOM*, 2010.
- [Rob03] Philippe Robert. *Stochastic Networks and Queues*. Stochastic Modeling and Applied Probability Series. Springer-Verlag, New York, 2003. xvii+398 pp.

- [RS08] S. Rajagopalan and D. Shah. Distributed algorithm and reversible network. In *CISS*, pages 498–502, 2008.
- [RSS09] Shreevatsa Rajagopalan, Devavrat Shah, and Jinwoo Shin. Network Adiabatic Theorem: An Efficient Randomized Protocol for Contention Resolution. In *Proceedings of SIGMETRICS*, pages 133–144, New York, NY, USA, 2009. ACM.
- [RW87] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes & martingales vol. 2: Itô Calculus*. Cambridge University Press, 2000 (1987).
- [Ser99] R. Serfozo. *Introduction to Stochastic Networks*. Springer-Verlag New York, 1999.
- [SS11] D. Shah and J. Shin. Randomized scheduling algorithm for queueing networks. *The Annals of Applied Probability*, page to appear, 2011.
- [TE92] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, 37:1936–1948, 1992.
- [vdVBvLP10] P. M. van de Ven, S. C. Borst, J. S. H. van Leeuwaarden, and A. Proutiere. Insensitivity and stability of random-access networks. *Perform. Eval.*, 67:1230–1242, November 2010.

CHAPTER IV

On a Transient Stochastic Network with Failures

Contents

1. Introduction	91
2. The Stochastic Model	94
3. The Overloaded Network	95
4. The Critical Case	97
5. The Time Scales of the Stable Network	100
Appendix. Generalized Skorokhod Problems	106
References for Chapter IV	107

1. Introduction

Storage systems. One considers a large scale storage system, it is a set of file servers in a communication network. In order to ensure persistence, files are duplicated on several servers. When the disk of a given server breaks down, its files are lost but can be retrieved on the other servers if copies are available. For these architectures a fraction of the bandwidth of a server is devoted to the duplication mechanism of its files to other servers. On the one hand, there should be sufficiently many copies so that any file has a copy available on at least one server at any time. On the other hand, in order to use the bandwidth in an optimal way, there should not be too many copies of a given file so that the network can accommodate a large number of distinct files. These systems are known as distributed hash tables (DHTs), they play an important role in the development of some large scale distributed systems, see Rhea et al. [RGK⁺05] and Rowstron and Druschel [RD01] for a more detailed presentation.

Failures of disks occur naturally randomly, these events are quite rare but, given the large number of nodes of these distributed systems, this is not a negligible phenomenon at the level of the network. If, for a short period of time, several of the servers break down, it may happen that files will be lost for good just because all the available copies were on these servers and because the recovery procedure was not completed before the last copy disappeared. To design such a system, it is therefore desirable to find a convenient duplication policy and to dimension the system so that all files will have at least a copy as long as possible. The natural critical parameters of the network are the failure rates of servers, the bandwidth allocated to duplication, the number of files and the number of servers. The ratio of the two last quantities being a measure of the storage capacity of the system. It is important to understand the impact of each of these parameters on the efficiency of the storage system.

Stochastic Models. This network can be seen as a classical set of queues with breakdowns. Numerous stochastic models of such systems have been investigated in the literature, see Chapter 6 of King [Kin90] for example and the references

therein. Related models concern queues with retrial and queues with servers of walking types, see Artalejo and Gómez-Corral [AGC08] and Falin and Templeton [FT97]. For most of the systems analyzed, there are, in general, one or two nodes which are subject to breakdowns. A queueing analysis is generally done in this context: convergence in distribution of the associated Markov model and analysis of the distribution of the availability of the system, of the delays and of queue sizes, ... For DHTs, there are very few stochastic models which investigate their performances; the existing ones describe the evolution of the number of copies of a single file. See Chun et al. [CDH⁺06], Picconi et al. [PBS07] and Ramabhadran and Pasquale [RP06]. See also Chapter V. In most of these studies the interaction between different files, due to the bandwidth sharing limitations, has not been really considered, except through simulations. The purpose of this chapter is to investigate the impact of this interaction. The second important aspect is that a large system, i.e. with a large number of files, will be considered instead of a small number of elements. This assumption is quite natural for current distributed systems.

More precisely, the following simple model is considered: A file can have at most two copies, the total bandwidth allocated to file duplication is given by λN , for $\lambda > 0$ and $N \in \mathbb{N}$. If at some moment there are $x \geq 1$ files with exactly one copy, a new copy of each of these files is created at rate $\lambda N/x$. It is assumed that initially F_N files are present in the system with two copies and each copy of a file disappears at rate μ . Recall that a file with 0 copies is lost. It will be assumed that the total number of files F_N is proportional to N , i.e. that F_N/N converges to some $\beta > 0$. Clearly enough, this system is transient and the empty state, all files are lost, is an absorbing state. The aim of this chapter is of describing the decay of the network, i.e. how the set of lost files is increasing. For $\delta > 0$, there exists some finite random instant $T_N(\delta)$, such that a fraction $\lfloor \delta N \rfloor$ of the files are lost after time $T_N(\delta)$. The chapter investigates the order of magnitude in N of the variables $T_N(\delta)$ as N gets large and the role of the parameters λ , μ and β in these asymptotics.

In practice, if there are N servers and that each of them has an available bandwidth λ to duplicate files, then the maximal capacity for duplication is λN . The model described above has therefore an optimal use of the duplication mechanism since the maximal duplication capacity is always available. For this reason this model provides upper bounds on the optimal performances of such a system. In particular, for any duplication mechanisms, after a duration of time with the same distribution as $T_N(\delta)$, at least $\lfloor \delta N \rfloor$ files will be lost for good. A more realistic model, when the total duplication bandwidth is not anymore centralized, is investigated in Feuillet and Robert [FR12] via mean-field limit asymptotics. It turns out that the corresponding mean-field limit can in fact be expressed in terms of the simple model analyzed in this chapter. The more general case when there are at most $d \geq 2$ copies of a given file will be investigated in another article to come.

Time Scales of Transient Markov Processes. If, for $i \in \{0, 1, 2\}$, $X_i^N(t)$ denotes the number of files with i copies in the network, then, under Poisson assumptions for failures and for duplication processes, $(X_0^N(t), X_1^N(t))$ is clearly a finite Markov process with $(F_N, 0)$ as an absorbing state. Contrary to previous works mentioned above, there is clearly no question of equilibrium here since the system *dies* at $(F_N, 0)$. A possible approach to investigate the decay of such a system could be of considering the associated quasi-stationary distributions of the Markov process. See Darroch and Seneta [DS65] and Ferrari et al. [FKMP95] for example. It would give a description of the system conditionally on the event that only a fraction of the files has been lost. These quantities are generally expressed in

terms of the spectral characteristics of the jump matrix. For this reason, explicit description of these distributions is quite rare outside one-dimensional birth and death processes. In this chapter, different time scales will be used to investigate the qualitative behavior of these transient processes. Times scales can be thought as “lenses”. Two of them will focus on the stable part of the sample path of the process (if any) and then describe the quasi-stationary behavior of the process.. Finally, a third time scale will focus on the decaying part of the sample paths, i.e. when the proportion of lost files is significant and the system tends to extinction.

Stochastic Averaging Principles. It is shown that in some cases, a stochastic averaging principle (SAP) occurs for this transient process: roughly speaking its dynamics can be decomposed into two components, one evolving on a fast time scale and the other one on a slower time scale. The system is fully coupled in the sense that the jump rates of the slow process depends on the equilibrium of the fast process, and the jump rates of the fast process depends of the state of the slow process. See Khasminskii [Kha80] and Freidlin and Wentzell [FW98]. This phenomenon is known to occur for the classical example of loss networks. In this case the vector of the number of free places of the congested links is the fast component, see Kelly [Kel91] and Hunt and Kurtz [HK94]. Outside this class of networks, there are, up to now, few examples of stochastic networks for which a fully coupled SAP occurs. See Chapter II and Perry and Whitt [PW11] for recent examples of SAP.

This SAP phenomenon is already well known in the framework of deterministic dynamical systems, see Guckenheimer and Holmes [GH90]. In a stochastic context, an additional difficulty, sometimes underestimated, is of controlling the regularity properties of the family of invariant distributions indexed by the states of the slow process, instead of the family of fixed points in the deterministic case. This can be done through a kind of uniform control of some ergodic averages, see Freidlin and Wentzell [FW98] or by using a martingale representation of the associated Markov processes, see Kurtz [Kur92]. In any case, there are several delicate technical issues to address: a convenient tightness result for a set of random measures and the rate of convergence of ergodic averages. In this chapter, a martingale formulation is also used but with a technical background significantly reduced. By taking a convenient state space for random measures, technical results related to extensions of random measures with specific measurability properties are not necessary. Furthermore, the tightness of the family of invariant distributions of fast processes is obtained as a consequence of a simple monotonicity property. If the monotonicity property is quite specific, it seems that the method to avoid extension results can be used in a quite general framework. This will be the subject of further investigations.

Outline of the Chapter. Section 2 introduces the Markov process investigated and its corresponding martingale representation. Section 3 studies a fluid picture of the network, i.e. the limit of the sequence of processes $(X_0^N(t)/N, X_1^N(t)/N)$, it is shown in Theorem 3.1 that its limit, the solution of an ODE, is not trivial when $\lambda < 2\mu\beta$ and is $(0,0)$ when $\lambda > 2\mu\beta$. The storage system is therefore properly designed when $\lambda > 2\mu\beta$, otherwise it is inefficient since it is losing a significant number of files right from the beginning. Section 4 is devoted to the critical case $\lambda = 2\mu\beta$, Theorem 4.1 shows that the sequence of processes $(X_0^N(t)/\sqrt{N}, X_1^N(t)/\sqrt{N})$ is converging in distribution and that its limit can be expressed in terms of a non-Markovian one-dimensional process, which is solution of an unusual stochastic differential equation with reflection at 0. In Section 5, the stable case $\lambda > 2\mu\beta$ is investigated. It is shown that the capacity of the system remains intact at the normal time scale: For $t \geq 0$, Theorem 5.3 proves that the variable $(X_0^N(t))$

converges in distribution to a Poisson process. Only a finite number of files is lost as N goes to infinity. More interesting, Theorem 5.5 shows that on the time scale $t \rightarrow Nt$ the transience of the Markov process shows up: at “time” Nt a fraction $\psi(t)N$ of the files is lost where $\psi(t)$ is the solution of some fixed point equation. This is the case where a stochastic averaging principle holds: around time Nt there is a local equilibrium for which $(\beta - \Psi(t))N$ files are still available. As a consequence, $t \rightarrow Nt$ is the convenient time scale to observe the degradation of the storage system. The proof of the convergence results uses a more or less straightforward extension of the classical Skorokhod problem formulation, see Skorokhod [Sko61]. The necessary material is gathered in the appendix to keep the chapter self-contained.

2. The Stochastic Model

Recall that F_N is the total number of distinct files initially present in the network and $X_1^N(t)$, resp. $X_0^N(t)$ is the number of files with one copy at time t , the number of lost files at this instant. The number $X_2^N(t)$ of files with two copies at time t is defined by $X_2^N(t) = F_N - X_0^N(t) - X_1^N(t)$. In general it will be assumed that all files have the maximum number of copies initially. The copy of a file is lost with rate μ and, conditionally on $X_1^N(t) = x$, a file with only one copy gets an additional copy with rate $\lambda N/x$. All events are supposed to occur after an exponentially distributed amount of time. Under these assumptions $(X(t)) \stackrel{\text{def.}}{=} (X_0^N(t), X_1^N(t))$ is a Markov process on the state space

$$\mathcal{S} = \{x = (x_0, x_1) \in \mathbb{N}^2 : x_0 + x_1 \leq F_N\},$$

as mentioned above, with these assumptions, the state $(F_N, 0)$ is an absorbing point of the process $(X^N(t))$.

For $x \in \mathbb{N}^2$, the Q -matrix $Q^N = (q^N(\cdot, \cdot))$ of the process $(X(t))$ is defined by

$$(2.1) \quad \begin{cases} q^N(x, x + e_1) = 2\mu(F_N - x_0 - x_1), \\ q^N(x, x - e_1) = \lambda N \mathbb{1}_{\{x_1 > 0\}}, \\ q^N(x, x - e_1 + e_0) = \mu x_1. \end{cases}$$

It is assumed that

$$(2.2) \quad \lim_{N \rightarrow +\infty} F_N/N = \beta,$$

and one denotes $\rho = \lambda/\mu$.

The stochastic differential equations associated with this transient Markov process can be written as

$$(2.3) \quad X_0^N(t) = X_0^N(0) + \sum_{i=1}^{+\infty} \int_0^t \mathbb{1}_{\{i \leq X_1^N(u-)\}} \mathcal{N}_{\mu,i}(du),$$

$$(2.4) \quad \begin{aligned} X_1^N(t) = X_1^N(0) &- \int_0^t \mathbb{1}_{\{X_1^N(u-) > 0\}} \mathcal{N}_{\lambda N}(du) - \sum_{i=1}^{+\infty} \int_0^t \mathbb{1}_{\{i \leq X_1^N(u-)\}} \mathcal{N}_{\mu,i}(du) \\ &+ \sum_{i=1}^{+\infty} \int_0^t \mathbb{1}_{\{i \leq F_N - X_0^N(u-) - X_1^N(u-)\}} \mathcal{N}_{2\mu,i}(du), \end{aligned}$$

where $(\mathcal{N}_{\mu,i})$ and $(\mathcal{N}_{2\mu,i})$ are two i.i.d. independent sequence of Poisson processes with respective parameters μ and 2μ , $\mathcal{N}_{\lambda N}$ is an independent Poisson process with parameter λN . For the i th file having only one copy, the integrand of the right hand side of Relation (2.3) corresponds to its definitive loss and the first term of the right hand side of Relation (2.4) is associated with its duplication. The last term of Relation (2.4) represents the loss of a copy of files with two copies.

Relation (2.3) can be rewritten as

$$(2.5) \quad X_0^N(t) = X_0^N(0) + \mu \int_0^t X_1^N(u) \, du + M_0^N(t),$$

where $(M_0^N(t))$ is the martingale defined by

$$M_0^N(t) = \sum_{i=1}^{+\infty} \int_0^t \mathbb{1}_{\{i \leq X_1^N(u-)\}} [\mathcal{N}_{\mu,i}(du) - \mu \, du],$$

its increasing process is given by

$$\langle M_0^N(t) \rangle = \mu \int_0^t X_1^N(u) \, du,$$

in particular, since $X_1^N(u) \leq F_N$, there exists some constant C_0 such that

$$\mathbb{E}(M_0^N(t)^2) = E(\langle M_0^N(t) \rangle) \leq C_0 N t$$

holds for all $t \geq 0$ and $N \geq 1$.

Similarly, if f is in $C_c(\mathbb{N})$, the set of functions with finite support on $\mathbb{N}$, Relation (2.4) gives the representation

$$(2.6) \quad f(X_1^N(t)) = f(X_1^N(0)) + \mu \int_0^t [f(X_1^N(u) - 1) - f(X_1^N(u))] X_1^N(u) \, du \\ + N \int_0^t \Omega \left[\frac{F_N}{N} - \frac{X_1^N(u) + X_0^N(u)}{N} \right] (f)(X_1^N(u)) \, du + M_1^N(t),$$

where, for $y \geq 0$, $\Omega[y]$ is the functional operator defined by

$$(2.7) \quad \Omega[y](f)(x) = 2\mu y(f(x+1) - f(x)) + \lambda \mathbb{1}_{\{x>0\}}(f(x-1) - f(x)), \quad x \in \mathbb{N},$$

and $(M_1^N(t))$ is a martingale such that, for some constant C_1 ,

$$\mathbb{E}(M_1^N(t)^2) \leq C_1 N \|f\|_{\infty, F_N} t,$$

holds for all $t \geq 0$ and $N \geq 1$, where $\|f\|_{\infty, F_N} = \max\{|f(x)| : 0 \leq x \leq F_N\}$.

3. The Overloaded Network

In this section, it is proved that a significant fraction of files is lost quickly if the network is not correctly dimensioned, i.e. when the ratio $\rho = \lambda/\mu$ is less than 2β . In this case, for a large N , the fraction of files with two copies at time t , $(F_N - X_0^N(t) - X_1^N(t))/N$ is close to $\rho/2$ if t is large enough. As a consequence $(\beta - \rho/2)N$ files are lost and the network stabilizes with a subset of files with two copies whose cardinality is of the order of $\rho/2$. This is the critical case which is analyzed in Section 4 where it is proved that the number of lost files is of the order of $\sqrt{N}$. When $\rho > 2\beta$, no file is lost at the fluid level. This case is investigated precisely in Section 5.

THEOREM 3.1 (Fluid Equations). *If $(X_0^N(0), X_1^N(0))$ is some fixed element of $\mathcal{S}$ and $\lim_{N \rightarrow +\infty} F_N/N = \beta$ then the sequence of processes $(X_0^N(t)/N, X_1^N(t)/N)$ converges in distribution to*

$$\begin{cases} [(\beta - \rho/2)(1 - 2e^{-\mu t} + e^{-2\mu t}), (2\beta - \rho)(e^{-\mu t} - e^{-2\mu t})] & \text{if } \rho \leq 2\beta, \\ (0, 0) & \text{if } \rho > 2\beta. \end{cases}$$

PROOF. Let $(X_0^N(0), X_1^N(0)) = (y_0, y_1) \in \mathcal{S}$. Equations (2.5) and (2.6), with the function $f \equiv \text{Id}$ on $[0, F_N]$, can be written as

$$(3.1) \quad \frac{X_0^N(t)}{N} = \frac{y_0}{N} + \mu \int_0^t \frac{X_1^N(u)}{N} du + \frac{M_0^N(t)}{N},$$

$$(3.2) \quad \begin{aligned} \frac{X_1^N(t)}{N} = & \frac{y_1}{N} + 2\mu \int_0^t \left(\frac{F_N}{N} - \frac{X_1^N(u) + X_0^N(u)}{N} \right) du - \lambda t \\ & - \mu \int_0^t \frac{X_1^N(u)}{N} du + \frac{M_1^N(t)}{N} + \lambda \int_0^t \mathbb{1}_{\{X_1^N(u)=0\}} du. \end{aligned}$$

Doob's Inequality and the bounds on the second moments of the associated martingales show that, for $i = 0, 1$ and $t \geq 0$,

$$\mathbb{P} \left(\sup_{0 \leq s \leq t} \frac{M_i^N(s)}{N} \geq \varepsilon \right) \leq \frac{1}{\varepsilon^2} \mathbb{E}(M_i(t)^2) \leq \frac{1}{N} \frac{C_i t}{\varepsilon^2}.$$

Therefore, the two sequences of processes $(M_0^N(t)/N)$ and $(M_1^N(t)/N)$ converge in distribution to 0 uniformly on compact sets.

For $T > 0$, $\delta > 0$ and for $i = 0, 1$, define $w_{X_i^N}^T(\delta)$ as the modulus of continuity of the process $(X_i^N(t))$ on the interval $[0, T]$,

$$(3.3) \quad w_{X_i^N}^T(\delta) = \sup_{0 \leq s \leq t \leq T, |t-s| \leq \delta} |X_i^N(t) - X_i^N(s)|.$$

By using the fact that, for some constant C , $X_i^N(t) \leq F_N \leq CN$ for all $N \in \mathbb{N}$ and $t \geq 0$, the above equations and the convergence of the martingales to 0 give that, for any $\varepsilon > 0$ and $\eta > 0$, there exists $\delta > 0$ such that the relation $\mathbb{P}(w_{X_i^N}^T(\delta) \geq \eta) \leq \varepsilon$ holds for all N .

This implies that the sequence of stochastic processes $(X_0^N(t)/N, X_1^N(t)/N)$ is tight. See Billingsley [Bil99] for example. One denotes by $(x_0(t), x_1(t))$ a limiting value for some subsequence (N_k) . From Equation (3.1), one gets the relation

$$x_0(t) = \mu \int_0^t x_1(u) du.$$

Define

$$(3.4) \quad Z^N(t) = \mu \int_0^t \left(\frac{2F_N}{N} - \frac{3X_1^N(u)}{N} - \frac{2X_0^N(u)}{N} \right) du - \lambda t + \frac{M_1^N(t)}{N}.$$

Equation (3.2) can be also interpreted as the fact that

$$(3.5) \quad (X_Z^N(t), R_Z^N(t)) \stackrel{\text{def.}}{=} \left(X_1^N(t)/N, \lambda \int_0^t \mathbb{1}_{\{X_1^N(u)=0\}} du \right)$$

is the unique solution of the Skorokhod problem associated with the process $(Z^N(t))$. See Appendix for a definition.

The sequence $(Z^{N_k}(t))$ is converging in distribution and by the continuous mapping theorem

$$(3.6) \quad \begin{aligned} \lim_{k \rightarrow +\infty} (Z^{N_k}(t)) = y(t) & \stackrel{\text{def.}}{=} \mu \int_0^t (2\beta - 2x_0(u) - 3x_1(u)) du - \lambda t \\ & = (2\mu\beta - \lambda)t - \mu \int_0^t \left(3x_1(u) + 2\mu \int_0^u x_1(v) dv \right) du \end{aligned}$$

The solutions of Skorokhod problems being continuous with respect to the process $(Z^N(t))$, see Appendix D of Robert [Rob03] for example, one gets that

$(X_Z^N(t), R_Z^N(t))$ converges in distribution to the solution $(x_y(t), r_y(t))$ of the Skorokhod problem associated with $(y(t))$. Since $x_y(t) = x_1(t)$ and $y(t) = F(x_1(t))$ with

$$F(x)(t) = (2\mu\beta - \lambda)t - \mu \int_0^t \left(3x(u) + 2\mu \int_0^u x(v) \, dv \right) du,$$

the process $(x_1(t))$ is a solution of the generalized Skorokhod problem (GSP) associated with the functional F . See Appendix. Proposition 5.8 shows that such a solution exists and is unique. This implies that there is a unique, deterministic limiting value for the sequence $(X_0^N(t)/N, X_1^N(t)/N)$. It is easy to check that the explicit expressions for $(x_0(t))$ and $(x_1(t))$ given in the statement of the theorem are indeed the solutions of the GSP. The convergence in distribution is therefore established. $\blacksquare$

4. The Critical Case

To complete the picture of the overloaded network $\rho \leq 2\beta$, one considers the critical case $\rho = 2\beta$. As it will be seen, the convergence result is expressed in terms of a reflected stochastic differential equation. The appendix presents the corresponding definition and a result of existence and uniqueness.

THEOREM 4.1. *If $\lambda/\mu = 2\beta$ and, for some $\gamma \in \mathbb{R}$,*

$$\lim_{N \rightarrow +\infty} \frac{1}{\sqrt{N}} \left(F_N - N \frac{\rho}{2} \right) = \gamma \text{ and } \lim_{N \rightarrow +\infty} \frac{X_1^N(0)}{\sqrt{N}} = y,$$

and $X_0^N(0) = 0$, then for the convergence in distribution

$$\lim_{N \rightarrow +\infty} \left(\frac{X_0^N(t)}{\sqrt{N}}, \frac{X_1^N(t)}{\sqrt{N}} \right) = \left(\mu \int_0^t Y(u) \, du, Y(t) \right),$$

where $(Y(t))$ is the solution starting at y of the stochastic differential equation

$$(4.1) \quad dY(t) = \sqrt{2\lambda} \, dB(t) + \mu \left(2\gamma - 3Y(t) - 2\mu \int_0^t Y(u) \, du \right) dt$$

reflected at 0, i.e. with the constraint that $Y(t) \geq 0$, for all $t \geq 0$. The process $(B(t))$ is a standard Brownian motion on $\mathbb{R}$.

The solution of SDE (4.1) is non-Markovian due to the integral term in the drift.

PROOF. Equations (2.5) and (2.6), with the function $f \equiv \text{Id}$ on $[0, F_N]$, can be written as

$$(4.2) \quad \bar{X}_0^N(t) \stackrel{\text{def.}}{=} \frac{X_0^N(t)}{\sqrt{N}} = \mu \int_0^t \frac{X_1^N(u)}{\sqrt{N}} \, du + \frac{M_0^N(t)}{\sqrt{N}},$$

$$(4.3) \quad \begin{aligned} \bar{X}_1^N(t) \stackrel{\text{def.}}{=} \frac{X_1^N(t)}{\sqrt{N}} &= \frac{X_1^N(0)}{\sqrt{N}} + 2\mu \int_0^t \left(\gamma_N - \frac{X_1^N(u)}{\sqrt{N}} - \frac{X_0^N(u)}{\sqrt{N}} \right) du \\ &\quad - \mu \int_0^t \frac{X_1^N(u)}{\sqrt{N}} \, du + \frac{M_1^N(t)}{\sqrt{N}} + \lambda \sqrt{N} \int_0^t \mathbb{1}_{\{X_1^N(u)=0\}} \, du, \end{aligned}$$

with $\gamma_N = (F_N - N\rho/2)/\sqrt{N}$. With the same notations as in Section 2, the martingales $(M_0^N(t))$ and $(M_1^N(t))$ are

$$\begin{aligned} M_0^N(t) &= \sum_{i=1}^{+\infty} \int_0^t \mathbb{1}_{\{i \leq X_1^N(u-)\}} [\mathcal{N}_{\mu,i}(du) - \mu \, du] \\ M_1^N(t) &= \sum_{i=1}^{+\infty} \int_0^t \mathbb{1}_{\{i \leq F_N - X_1^N(u-) - X_0^N(u-)\}} [\mathcal{N}_{2\mu}(du) - 2\mu \, du] \\ &\quad - M_0^N(t) - \int_0^t \mathbb{1}_{\{X_1^N(u) > 0\}} [\mathcal{N}_{\lambda N}(du) - \lambda N \, du]. \end{aligned}$$

Their increasing processes are given by

$$\begin{aligned} \left\langle \frac{1}{\sqrt{N}} M_0^N \right\rangle(t) &= \mu \int_0^t \frac{X_1^N(u)}{N} \, du, \\ \left\langle \frac{1}{\sqrt{N}} M_1^N \right\rangle(t) &= 2\mu \int_0^t \left(\frac{F_N}{N} - \frac{X_1^N(u)}{N} - \frac{X_0^N(0)}{N} \right) \, du + \left\langle \frac{1}{\sqrt{N}} M_0^N \right\rangle(t) \\ &\quad + \lambda \int_0^t \mathbb{1}_{\{X_1^N(u) > 0\}} \, du. \end{aligned}$$

The last term of the right hand side of the above equation is $(R^N(t))$ defined by Equation (3.5) in the proof of the previous theorem. It is the second component of the solution to the Skorokhod problem associated with the process $(Z^N(t))$ of Relation (3.4). It has been seen that the sequence of processes $(Z^N(t))$ is converging to $(y(t))$ defined in Equation (3.6). In this case $(y(t))$ is identically 0, the solution of the corresponding Skorokhod problem associated with $(y(t))$ is therefore $(0, 0)$. The continuity properties of the solutions of the Skorokhod problem imply that the process $(R^N(t))$ converges to 0. Consequently, by Theorem 3.1 one gets the convergence in distribution

$$\lim_{N \rightarrow +\infty} \left(\int_0^t \mathbb{1}_{\{X_1^N(u)=0\}} \, du \right) = 0$$

and therefore

$$\lim_{N \rightarrow +\infty} \left(\left\langle \frac{1}{\sqrt{N}} M_0^N \right\rangle(t) \right) = 0 \text{ and } \lim_{N \rightarrow +\infty} \left(\left\langle \frac{1}{\sqrt{N}} M_1^N \right\rangle(t) \right) = (2\lambda t).$$

One deduces that $(\overline{M}_1^N(t)) \stackrel{\text{def.}}{=} (M_1^N(t)/\sqrt{N})$ converges to $(\sqrt{2\lambda}B(t))$ where $(B(t))$ is a standard Brownian motion and that $(\overline{M}_0^N(t)) \stackrel{\text{def.}}{=} (M_0^N(t)/\sqrt{N})$ converges to 0. See Ethier and Kurtz [EK86] for example.

One now proves that the processes

$$\left(\overline{X}_0^N(t) \right) \stackrel{\text{def.}}{=} \left(\frac{X_0^N(t)}{\sqrt{N}} \right) \text{ and } \left(\overline{X}_1^N(t) \right) \stackrel{\text{def.}}{=} \left(\frac{X_1^N(t)}{\sqrt{N}} \right)$$

are tight. If $(h(t))$ is a function on $\mathbb{R}_+$, one denotes,

$$\|h\|_{\infty,t} = \sup_{0 \leq s \leq t} |h(s)|$$

and $w_h^t(\cdot)$ is the modulus of continuity of h defined by Equation (3.3). Equation (4.2) gives, for $0 \leq t \leq T$,

$$\|\overline{X}_0^N\|_{\infty,t} \leq \|\overline{M}_0^N\|_{\infty,T} + \mu \int_0^t \|\overline{X}_1^N\|_{\infty,u} \, du.$$

Equation (4.3) shows that $(X_1^N(t)/\sqrt{N})$ is the first coordinate of the solution of the Skorokhod problem associated with $(Z_1^N(t))$ defined by

$$(4.4) \quad Z_1^N(t) \stackrel{\text{def.}}{=} y_N + \mu \int_0^t \left(2\gamma_N - 3\frac{X_1^N(u)}{\sqrt{N}} + 2\frac{X_0^N(u)}{\sqrt{N}} \right) du + \frac{M_1^N(t)}{\sqrt{N}},$$

with $y_N = X_1^N(0)/\sqrt{N}$. By using the explicit representation of the solution of a Skorokhod problem in dimension 1, one has

$$\|\bar{X}_1^N\|_{\infty,t} \leq 2\|Z_1^N\|_{\infty,t}, \text{ for } 0 \leq t \leq T,$$

see Appendix D of Robert [**Rob03**] for example, then

$$\begin{aligned} \|\bar{X}_1^N\|_{\infty,t} &\leq 2y_N + 4\mu\gamma_N T + 2\|\bar{M}_1^N\|_{\infty,T} + 4\mu \int_0^t \left(\|\bar{X}_0^N\|_{\infty,u} + \|\bar{X}_1^N\|_{\infty,u} \right) du \\ &\leq U^N(T) + (4 + \mu T)\mu \int_0^t \|\bar{X}_1^N\|_{\infty,u} du, \end{aligned}$$

with $U^N(T) \stackrel{\text{def.}}{=} 2y_N + 4\mu\gamma_N T + 2\|\bar{M}_1^N\|_{\infty,T} + 4\mu T\|\bar{M}_0^N\|_{\infty,T}$.

Gronwall's Inequality gives that the relation $\|\bar{X}_1^N\|_{\infty,t} \leq U^N(T) \exp((4+\mu T)\mu t)$ holds for $0 \leq t \leq T$, and, consequently,

$$\|\bar{X}_0^N\|_{\infty,t} \leq \mu T U^N(T) e^{(4+\mu T)\mu t} + \|\bar{M}_0^N\|_{\infty,T}.$$

The convergence of martingales shows that the two sequences of random variables $(U^N(T))$ and $\|\bar{M}_0^N\|_{\infty,T}$ converge in distribution. Consequently, for $\varepsilon > 0$, there exists some $K > 0$ such that for $i = 0, 1$ and all $N \geq 0$,

$$\mathbb{P} \left(\|\bar{X}_i^N\|_{\infty,t} > K \right) \leq \varepsilon.$$

If $\eta > 0$, there exists N_0 and δ sufficiently small so that, for all $N \geq N_0$,

$$2\mu\delta T(\gamma_N + 2K) < \eta/2 \text{ and } \mathbb{P} \left(w_{\bar{M}_1^N}^T \geq \eta \right) \leq \varepsilon.$$

The last relation coming from the fact that the sequence $(\bar{M}_1^N(t))$ is converging in distribution to a continuous process. One gets finally

$$\begin{aligned} \mathbb{P} \left(w_{\bar{Z}_1^N}^T(\delta) \geq \eta \right) &\leq \mathbb{P} \left(2\mu\delta T \left[\gamma_N + \|\bar{X}_0^N\|_{\infty,T} + \|\bar{X}_1^N\|_{\infty,T} \right] + w_{\bar{M}_1^N}^T(\delta) \geq \eta \right) \\ &\leq \mathbb{P} \left(\|\bar{X}_0^N\|_{\infty,T} \geq K \right) + \mathbb{P} \left(\|\bar{X}_1^N\|_{\infty,T} \geq K \right) + \mathbb{P} \left(w_{\bar{M}_1^N}^T(\delta) \geq \eta/2 \right) \leq 3\varepsilon. \end{aligned}$$

The sequence $(Z_1^N(t))$ is therefore tight, by continuity of the solution of the Skorokhod problem the same property holds for $(X_1^N(t)/\sqrt{N})$ and consequently for $(X_0^N(t)/\sqrt{N})$.

If $(Y_0(t), Y_1(t))$ is a limit of a subsequence $[(X_0^{N_k}(t)/\sqrt{N_k}, X_1^{N_k}(t)/\sqrt{N_k})]$. By Equation (4.2) and (4.3), one gets that

$$\left(\frac{X_1^{N_k}(t)}{\sqrt{N_k}}, \lambda\sqrt{N} \int_0^t \mathbb{1}_{\{X_1^N(u)=0\}} du \right)$$

converges in distribution to the solution of the Skorokhod problem associated with the process

$$\left(y + \sqrt{2\lambda}B(t) + \mu \int_0^t \left(2\gamma - 3Y_1(u) - 2\mu \int_0^u Y_1(v) dv \right) du \right).$$

One concludes that $(Y_1(t))$ is the solution of the generalized Skorokhod problem for the functional F defined by

$$F(h)(t) = y + \sqrt{2\lambda}B(t) + \mu \int_0^t \left(2\gamma - 3h(u) - 2\mu \int_0^u h(v) \, dv \right) du.$$

Proposition 5.8 in the appendix shows that there is a unique solution $(Y_1(t))$ and consequently a unique limit $(Y_0(t), Y_1(t))$. The theorem is proved. $\blacksquare$

5. The Time Scales of the Stable Network

The asymptotic properties of the network are investigated under the condition $\rho = \lambda/\mu > 2\beta$. In Section 3 it has been shown that, in this case, the system is stable at the fluid level, i.e. that the fraction of lost files is 0. Of course this does not change the fact that the system is still transient with the absorbing state $(F_N, 0)$. To have a precise idea on how the system reaches this state, there are three interesting time scales to consider:

- (1) Slow time scale: $t \rightarrow t/N$,
- (2) Normal time scale: $t \rightarrow t$,
- (3) Linear time scale: $t \rightarrow Nt$,

they are investigated successively in this section. The following elementary lemma will be used throughout the section.

LEMMA 5.1. *If $\rho = \lambda/\mu > 2\beta$, for any $\beta_0 > \beta$ such that $\lambda/\mu > 2\beta_0$, $\varepsilon > 0$, $\eta > 0$ and $T > 0$, there exists $N_0 \in \mathbb{N}$ such that*

- (1) *Coupling: there exists a probability space where the relation*

$$X_1^N(t) \leq L_{\beta_0}(Nt), \forall t \geq 0,$$

holds for all $N \geq N_0$ and $t \geq 0$, with $(L_{\beta_0}(t))$ the process of the number of customers of an ergodic $M/M/1$ queue with arrival rate $2\mu\beta_0$ and service rate λ and with initial condition $L_{\beta_0}(0) = X_1^N(0)$.

- (2) *The relation*

$$\mathbb{P} \left[\sup_{\substack{0 \leq s, t \leq T, \\ |t-s| \leq \delta}} \frac{1}{N} \int_{sN}^{tN} L_{\beta_0}(u) \, du > \eta \right] \leq \varepsilon$$

holds.

PROOF. There exists some $\beta_0 \geq \beta$ and $N_0 \geq 1$ such that $\lambda > 2\beta_0\mu$ and that $F_N \leq N\beta_0$ for $N \geq N_0$. It is enough to take the $M/M/1$ queue with arrival rate $2\mu\beta_0$ and service rate λ .

Denote by $\mathcal{A}$ the event on the left hand side of the last relation to prove. If, for $x \in \mathbb{N}$, τ_x denotes the hitting time of x by the process $(L_{\beta_0}(t))$, for $\delta < 1/2$, one has

$$\mathbb{P}(\mathcal{A}) \leq \mathbb{P}(\tau_{\lfloor \eta N \rfloor} \leq NT).$$

By ergodicity of this process and Proposition 5.11 of Robert [Rob03] for example, there exists some $0 < \alpha < 1$ such that the sequence $(\alpha^N \tau_{\lfloor \eta N \rfloor})$ converges in distribution. The last term of the above relation is thus arbitrarily small as N gets large. $\blacksquare$

The slow time scale. A description of the asymptotic behavior for the slow time scale is presented informally. From Relation (2.1), one can see that the Q -matrix of the process on the slow time scale $(X_0^N(t/N), X_1^N(t/N))$ has the following asymptotic expansion

$$\lim_{N \rightarrow +\infty} \begin{cases} q^N(x, x + e_1) & = 2\mu\beta, \\ q^N(x, x - e_1) & = \lambda \mathbb{1}_{\{x_1 > 0\}}, \\ q^N(x, x - e_1 + e_0) & = 0. \end{cases}$$

With elementary arguments which are skipped one can easily get the following proposition. It states that, on the slow time scale, with probability 1 no file is lost at all in the limit.

PROPOSITION 5.2. *The sequence of processes $(X_0^N(t/N), X_1^N(t/N))$ converges in distribution to the process $(0, L_\beta(t))$, where $(L_\beta(t))$ is the process of the number of jobs of an $M/M/1$ queue with arrival rate $2\mu\beta$ and service rate λ .*

The normal time scale. It is shown that, on the normal time scale, the stability does not only hold on the fluid level: almost surely there is a finite number of losses in any finite time interval, more precisely losses occur as a Poisson process, see Theorem 5.3. The capacity λN of the network is thus able to maintain an almost complete set of files. The following proposition shows in particular that the number of definitive losses at time $t > 0$ is finite with a Poisson distribution.

THEOREM 5.3. *If $\rho = \lambda/\mu > 2\beta$,*

- *the sequence of processes $(X_0^N(t))$ converges in distribution to a Poisson point process on $\mathbb{R}_+$ with rate $2\mu\beta/(\rho - 2\beta)$.*
- *For $t > 0$, as N goes to infinity, the random variable $X_1^N(t)$ converges in distribution to a geometric distribution with parameter $2\beta/\rho$.*

The second convergence is for the marginal distribution of $(X_1^N(s))$ at time t . One cannot expect a convergence in distribution of the sequence of processes $(X_1^N(t))$. Indeed, since the sequence of processes $(X_1^N(t/N))$ is converging in distribution to the law of the $M/M/1$ process $(L_\beta(t))$, for $0 \leq s < t$, the distribution of $(X_1^N(s), X_1^N(t))$ and of $(L_\beta(Ns), L_\beta(Nt))$ are close. Between time Ns and Nt , the $M/M/1$ “forgets” its location at time Ns (just because it hits 0 with probability 1) so that when N goes to infinity the couple $(X_1^N(s), X_1^N(t))$ converges in distribution to the distribution of two independent geometric distributions. The sample paths of a possible limit of $(X_1^N(s))$ would not have regularity properties.

PROOF. Define

$$\eta_N(t) \stackrel{\text{def.}}{=} \int_0^t X_1^N(u) \, du.$$

For $0 \leq s \leq t$, the above lemma gives that,

$$\eta_N(t) - \eta_N(s) = \int_s^t X_1^N(u) \, du \leq \int_s^t L_{\beta_0}(Nu) \, du = \frac{1}{N} \int_{Ns}^{Nt} L_{\beta_0}(u) \, du.$$

The criteria of the modulus of continuity and Lemma 5.1 give that the sequence of processes $(\eta^N(t))$ is tight. The above inequality and the ergodic theorem applied to the ergodic Markov process $(L_{\beta_0}(t))$ show also that, almost surely,

$$(5.1) \quad \limsup_{N \rightarrow +\infty} \int_0^t X_1^N(u) \, du \leq \frac{2\beta_0}{\rho - 2\beta_0} t.$$

For $T > 0$ fixed and $K > 0$,

$$\begin{aligned} \mathbb{P}(X_0^N(T) \geq K) &\leq \mathbb{P}\left(\mu \int_0^T X_1^N(u) \geq K/2\right) + \mathbb{P}(M_0^N(T) \geq K/2) \\ &\leq \mathbb{P}\left(\mu \int_0^T X_1^N(u) \geq K/2\right) + \frac{4}{K^2} \mathbb{E}\left(\mu \int_0^T X_1^N(u) \, du\right). \end{aligned}$$

One can thus choose K so that $\mathbb{P}(X_0^N(T) \geq K) \leq \varepsilon$ holds for $N \geq N_0$ for some $N_0 \in \mathbb{N}$. As in the proof of Lemma 5.1, for $\delta > 0$, there exists some $N_1 \in \mathbb{N}$ such that if $N \geq N_1$ then

$$\mathbb{P}\left(\sup_{0 \leq s \leq T} L_{\beta_0}(s) \geq \delta N\right) \leq \varepsilon.$$

In the same way as in the proof of the above lemma, one can construct an $M/M/1$ process $(Z^N(t))$ whose arrival and service rates are respectively

$$2\mu \left(\frac{F_N}{N} - \frac{K}{N} - \delta \right) \text{ and } \lambda,$$

and such that, on the event,

$$\mathcal{A}_T \stackrel{\text{def.}}{=} \left\{ X_0^N(T) \leq K, \sup_{0 \leq t \leq NT} L_{\beta_0}(t) \leq \delta N \right\},$$

the relation $X_1^N(t) \geq Z^N(Nt)$ holds for all $t \leq T$. Hence, almost surely,

$$(5.2) \quad \liminf_{N \rightarrow +\infty} \eta^N(t) \geq \liminf_{N \rightarrow +\infty} \frac{1}{N} \int_0^{Nt} Z^N(u) \, du = \frac{2(\beta - \delta)}{\rho - 2(\beta - \delta)} t.$$

holds on $\mathcal{A}_T$. By letting δ go to 0 and β_0 to β in Equations (5.1) and (5.2) respectively, one gets that the variable $\eta^N(t)$ converges almost surely to αt with $\alpha = 2\beta\mu/(\rho - 2\beta)$. Consequently, the tightness of the sequence of processes $(\eta^N(t))$ implies that it is converging in distribution to (αt) .

Note that $t \mapsto X_0^N(t)$ can also be seen as a point process with jumps of size 1. By Equation (2.5), one has

$$\left(X_0^N(t) - \mu \int_0^t X_1^N(u) \, du \right)$$

is a martingale with respect to the natural filtration of the associated Poisson processes. The random measure

$$\Lambda^N([0, t]) = \mu \int_0^t X_1^N(u) \, du$$

is a *compensator* of the point process $t \mapsto X_0^N(t)$. See Kasahara and Watanabe [KW86]. It has therefore been shown that the sequence of compensators is converging to the deterministic measure $\alpha \, dx$. Theorem 5.1 of [KW86], see also Brown [Bro78], gives the convergence in distribution of $(X_0^N(t))$ to a Poisson process with rate α .

In a similar way as before, through the convergence of the Q -matrix, the asymptotic distribution of $X_1^N(t)$ can be easily obtained by conditioning on the event $\{X_0^N(t) \leq K\}$ for K large and by using arbitrarily close $M/M/1$ processes at equilibrium as upper and lower stochastic bounds for $X_1^N(t)$. Details are skipped. ■

The linear time scale $t \rightarrow Nt$. On the linear time scale, it will be shown that a fraction $\Psi(t)$ of the files is lost at time t . In some way the linear time scale gives a picture of the decay of the network.

For $N \geq 1$, the random measure μ_N on $\mathbb{N} \times \mathbb{R}_+$ is defined as, for a measurable function $g : \mathbb{N} \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$,

$$\langle \mu_N, g \rangle = \int_{\mathbb{R}_+} g(X_1^N(Nt), t) dt.$$

Note that if $g(x, t) = h(x) \mathbb{1}_{\{[0, T]\}}(t)$ for $T > 0$, then

$$\langle \mu_N, g \rangle = \sum_{x \in \mathbb{N}} h(x) \frac{1}{N} \int_0^{NT} \mathbb{1}_{\{X_1^N(t)=x\}} dt.$$

Consequently (μ_N) is a relatively compact sequence of random Radon measures on $\mathbb{N} \times \mathbb{R}_+$. See Dawson [Daw93] for example. Note that the measure identically null can be a possible limit of this sequence.

From now on, one fixes (N_k) such that (μ_{N_k}) is a converging subsequence whose limit is ν . By taking a convenient probability space, one can assume that the convergence of (μ_{N_k}) holds almost surely for the weak convergence of Radon measures.

Since, for $N \geq 1$, μ_N is absolutely continuous with respect to the product of the counting measure on $\mathbb{N}$ and Lebesgue measure on $\mathbb{R}_+$, the same property holds for the limiting measure ν . Let $(x, t) \rightarrow \pi_t(x)$ denote its (random) density. It should be remarked that, one can choose a version of $\pi_t(x)$ such that the map $(\omega, x, t) \rightarrow \pi_t(x)(\omega)$ on the product of the probability space and $\mathbb{N} \times \mathbb{R}_+$ is measurable by taking $\pi_t(x)$ as a limit of measurable maps,

$$\pi_t(x) = \limsup_{s \rightarrow 0} \frac{1}{s} \nu(\{x\} \times [t, t+s]).$$

See Chapter 8 of Rudin [Rud87] for example.

PROPOSITION 5.4. *For the convergence in distribution of continuous processes*

$$\lim_{k \rightarrow +\infty} \left(\frac{\mu}{N_k} \int_0^{N_k t} X_1^{N_k}(u) du \right) = (\Psi(t)) \stackrel{\text{def.}}{=} \left(\mu \int_0^t \langle \pi_u, I \rangle du \right),$$

where $I(x) = x$ for $x \in \mathbb{N}$. Moreover, almost surely, for all $t \geq 0$,

$$\int_0^t \pi_u(\mathbb{N}) du = t.$$

It must be noted that the last relation is crucial, it shows that the masses of the measures μ_{N_k} , for $k \geq 1$, do not vanish at infinity. This property is sometimes absent from the proofs of stochastic averaging principles, it is nevertheless mandatory to identify π_u as an invariant distribution of a Markov process.

PROOF. The criteria of the modulus of continuity is used to prove the tightness of

$$(\Psi_N(t)) \stackrel{\text{def.}}{=} \left(\frac{\mu}{N} \int_0^{Nt} X_1^N(u) du \right).$$

By Lemma 5.1

$$\Psi_N(t) - \Psi_N(s) = \frac{\mu}{N} \int_{Ns}^{Nt} X_1^N(u) du \leq \frac{\mu}{N^2} \int_{N^2 s}^{N^2 t} L_{\beta_0}(u) du.$$

As in the proof of Theorem 5.3, one concludes that the sequence of processes $(\Psi_N(t))$ is tight.

For $K > 0$ and $t \geq 0$, the almost sure convergence of the measures (μ_{N_k}) gives the convergence

$$\lim_{k \rightarrow +\infty} \frac{1}{N_k} \int_0^{N_k t} X_1^{N_k}(u) \mathbb{1}_{[0, K]}(X_1^{N_k}(u)) \, du = \int_0^t \langle \pi_u, I \mathbb{1}_{[0, K]} \rangle \, du,$$

where $I(x) = x$. By using again Lemma 5.1, one gets that

$$\frac{1}{N_k} \int_0^{N_k t} X_1^{N_k}(u) \mathbb{1}_{\{X_1^{N_k}(u) \geq K\}} \, du \leq \frac{1}{N_k^2} \int_0^{N_k^2 t} L_{\beta_0}(u) \mathbb{1}_{\{L_{\beta_0}(u) \geq K\}} \, du,$$

and the ergodic theorem applied to $(L_{\beta_0}(t))$ shows that the last quantity is converging in distribution to

$$\left(t \mathbb{E} \left(L_{\beta_0}(\infty) \mathbb{1}_{\{L_{\beta_0}(\infty) \geq K\}} \right) \right)$$

where $L_{\beta_0}(\infty)$ is the limit in distribution of $(L_{\beta_0}(t))$, a geometrically distributed random variable. For $\varepsilon > 0$, K is chosen sufficiently large so that the last quantity is less than $\varepsilon/2$, consequently if k is large enough, one has

$$\frac{1}{N_k} \int_0^{N_k t} X_1^{N_k}(u) \mathbb{1}_{\{X_1^{N_k}(u) \geq K\}} \, du \leq \varepsilon.$$

One deduces that $(\Psi(t))$ is the only possible limiting process for $(\Psi_{N_k}(t))$. This proves the first half of the proposition.

For $K \geq 1$, the convergence of (μ_{N_k}) gives the relation

$$\lim_{k \rightarrow +\infty} \frac{1}{N_k} \int_0^{N_k t} \mathbb{1}_{\{X_1^{N_k}(u) \leq K\}} \, du = \nu([0, K] \times [0, t]) = \int_0^t \pi_u([0, K]) \, du.$$

By using again the stochastic domination by an ergodic $M/M/1$ queue,

$$\frac{1}{N_k} \int_0^{N_k t} \mathbb{1}_{\{L_{\beta_0}(u) \leq K\}} \, du \leq \frac{1}{N_k} \int_0^{N_k t} \mathbb{1}_{\{X_1^{N_k}(u) \leq K\}} \, du.$$

by letting k go to infinity one gets that, almost surely,

$$t \mathbb{P}(L_{\beta_0}(\infty) \leq K) \leq \int_0^t \pi_u([0, K]) \, du \leq \int_0^t \pi_u(\mathbb{N}) \, du,$$

now if K go to infinity, one obtains the relation

$$\int_0^t \pi_s(\mathbb{N}) \, ds = t$$

holds for all $t \in \mathbb{N}$ and consequently for all $t \geq 0$. The proposition is proved. $\blacksquare$

THEOREM 5.5 (Rate of Decay of the Network). *If $\rho = \lambda/\mu > 2\beta$, then, as N goes to infinity, the process $(X_0^N(Nt)/N)$ converges to $(\Psi(t))$ where $\Psi(t)$ is the unique solution $y \in [0, \beta]$ of the equation*

$$(5.3) \quad (1 - y/\beta)^{\rho/2} e^{y+\mu t} = 1.$$

For $t \geq 0$, the process $(X_1^N(Nt + u), u > 0)$ converges in distribution to the stationary process of the number of jobs of an $M/M/1$ queue with service rate λ and arrival rate $2\mu(\beta - \Psi(t))$.

It is easily seen that the asymptotic expansion $\Psi(t) - \beta \sim \beta \exp(-2(\beta + \mu t)/\rho)$ holds as t goes to infinity. The last part of the theorem states that, “around” time Nt , the process X_1^N has a local equilibrium.

PROOF. Equation (2.6) gives that, for $f \in C_c(\mathbb{N})$

$$(5.4) \quad f(X_1^N(Nt)) - f(X_1^N(0)) - M_{f,1}^N(Nt) = N^2 \int_0^t \Omega[Y_N(u)](f)(X_1^N(Nu)) \, du \\ + \mu N \int_0^t \Delta^-(f)(X_1^N(Nu)) X_1^N(Nu) \, du$$

with, from Equation (2.5),

$$Y_N(u) = \frac{F_N}{N} - \frac{X_1^N(Nu)}{N} - \frac{M_0^N(Nu)}{N} - \frac{\mu}{N} \int_0^{Nu} X_1^N(v) \, dv,$$

and $\Delta^-(f)(x) = (f(x-1) - f(x)) \mathbb{1}_{\{x \geq 1\}}$. The bound on the increasing process of the martingale $(M_{f,1}^N(t))$ at the end of Section 2, Doob's Inequality and Lemma 5.1 show that the sequence of processes

$$\left(\frac{1}{N^2} \left[f(X_1^N(Nt)) - f(X_1^N(0)) - M_{f,1}^N(Nt) \right. \right. \\ \left. \left. - \mu N \int_0^t (\Delta^-(f)(X_1^N(Nu)) X_1^N(Nu)) \, du \right] \right)$$

converges to 0 for the topology of the uniform norm on compact sets.

By using Lemma 5.1, one gets that

$$\frac{X_1^N(Nu)}{N} \leq \frac{L_{\beta_0}(N^2u)}{N},$$

hence the sequence of processes $(X_1^N(Nu)/N)$ converges in distribution to 0.

The bound on the increasing process and Proposition 5.4 show that the sequence of processes $(Y_{N_k}(t))$ converges in distribution to $(\beta - \Psi(t))$. One deduces from Equation (5.4) that the sequence of processes

$$\left(\int_0^t \Omega[\beta - \Psi(u)](f)(X_1^{N_k}(N_k u)) \, du \right) = \left(\int_{\mathbb{N} \times [0,t]} \Omega[\beta - \Psi(u)](f)(x) \mu_{N_k}(dx, du) \right)$$

converges to 0.

The convergence of the (μ_{N_k}) and Proposition 5.4 give therefore that, almost surely, the relations

$$\int_0^t \langle \pi_u, \Omega[\beta - \Psi(u)](f) \rangle \, du = 0 \text{ and } \int_0^t \pi_u(\mathbb{N}) \, du = t,$$

hold for all $t \geq 0$ and all functions $f \in C_c(\mathbb{N})$. Note that one has used the fact that $C_c(\mathbb{N})$ has a countable dense subset for the uniform norm.

If Δ is the subset of all real numbers $u \geq 0$ such that one of the relations

$$\begin{cases} \pi_u(\mathbb{N}) \neq 1, \\ \langle \pi_u, \Omega[\beta - \Psi(u)](f) \rangle \neq 0, \text{ for some } f \in C_c(\mathbb{N}), \end{cases}$$

holds, then the Lebesgue measure of Δ is 0. Hence if $u \notin \Delta$, then $\pi_u(\mathbb{N}) = 1$ and $\langle \pi_u, \Omega[\beta - \Psi(u)](f) \rangle = 0$ for all $f \in C_c(\mathbb{N})$. Since $\Omega[\beta - \Psi(u)]$ is the infinitesimal generator of an $M/M/1$ queue with arrival rate $2\mu(\beta - \Psi(u))$ and service rate λ , one gets that π_u is a geometric distribution on $\mathbb{N}$ with parameter $2\mu(\beta - \Psi(u))/\lambda$.

From Proposition 5.4 one gets that, for $t \geq 0$,

$$\Psi(t) = \mu \int_{[0,t] \setminus \Delta} \langle \pi_u, I \rangle \, du = \mu \int_0^t \frac{2\mu(\beta - \Psi(u))}{\lambda - 2\mu(\beta - \Psi(u))} \, du.$$

Straightforward calculus gives the relation

$$(\beta - \psi(u))^{\rho/2} e^{\psi(u)} = \beta^{\rho/2} e^{-\mu u}.$$

It is easily checked that since $2\beta < \rho$, there is a unique $\Psi(u) < \beta$ satisfying the above equation. The theorem is proved. $\blacksquare$

The theorem gives directly the following corollary on the asymptotic behavior of $T_N(\delta)$, the first time when a fraction δ of the files has been lost.

COROLLARY 5.6. *If $\rho = \lambda/\mu > 2\beta$ then, for $N \geq 1$ and $\delta \in (0, 1)$,*

$$T_N(\delta) = \inf\{t \geq 0 : X_0^N(t) \geq \delta F_N\},$$

then, for the convergence in distribution,

$$\lim_{N \rightarrow +\infty} \frac{T_N(\delta)}{N} = \frac{1}{\mu} \left(-\frac{\rho}{2} \log(1 - \delta) - \delta\beta \right).$$

Appendix. Generalized Skorokhod Problems

For the sake of self-containedness, this section presents quickly the more or less classical material necessary to state and prove the convergence results used in this chapter. The general theme concerns the rigorous definition of a solution of a stochastic differential equation constrained to stay in some domain and also the proof of the existence and uniqueness of such a solution. See Skorokhod [Sko61], Anderson and Orey [AO76], Chaleyat-Maurel and El Karoui [EKCM78] and, in a multi-dimensional context, Harrison and Reiman [HR81] and Taylor and Williams [TW93] and, in a more general context, Ramanan [Ram06]. See Appendix D of Robert [Rob03] for a brief account.

We first recall the classical definition of Skorokhod problem. If $(Z(t))$ is some function of the set $\mathcal{D}(\mathbb{R}_+, \mathbb{R})$ of càdlàg functions defined on $\mathbb{R}_+$, the couple of functions $[(X(t)), (R(t))]$ is said to be a solution of the Skorokhod problem associated with $(Z(t))$ whenever

- (1) $X(t) = Z(t) + R(t)$, for all $t \geq 0$,
- (2) $X(t) \geq 0$, for all $t \geq 0$,
- (3) $t \rightarrow R(t)$ is non-decreasing, $R(0) = 0$ and

$$\int_{\mathbb{R}_+} X(t) dR(t) = 0.$$

The generalization used in this chapter corresponds to the case when $(Z(t))$ is itself a functional of $(X(t))$.

DEFINITION 5.7 (Generalized Skorokhod Problem).

If $G : \mathcal{D}(\mathbb{R}_+, \mathbb{R}) \rightarrow \mathcal{D}(\mathbb{R}_+, \mathbb{R})$ is a Borelian function, $((X(t)), (R(t)))$ is a solution of the generalized Skorokhod Problem (GSP) associated with G if $((X(t)), (R(t)))$ is the solution of the Skorokhod Problem associated with $G(X)$, in particular, for all $t \geq 0$,

$$X(t) = G(X)(t) + R(t) \text{ and } \int_{\mathbb{R}_+} X(t) dR(t) = 0.$$

The classical Skorokhod problem described above corresponds to the case when the functional G is constant and equal to $(Z(t))$. If one takes

$$G(x)(t) = \int_0^t \sigma(x(u)) dB(u) + \int_0^t \delta(x(u)) du,$$

where $(B(t))$ is a standard Brownian motion and σ and δ are Lipschitz functions on $\mathbb{R}$, the first coordinate $(X(t))$ of a possible solution to the corresponding GSP can be described as the solution of the SDE

$$dX(t) = \sigma(X(t)) dB(t) + \delta(X(t)) dt$$

reflected at 0.

PROPOSITION 5.8. *If $G : \mathcal{D}(\mathbb{R}_+, \mathbb{R}) \rightarrow \mathcal{D}(\mathbb{R}_+, \mathbb{R})$ is such that, for any $T > 0$, there exists a constant C_T such that, for all $(x(t)) \in \mathcal{D}(\mathbb{R}_+, \mathbb{R})$ and $0 \leq t \leq T$,*

$$(5.5) \quad \sup_{0 \leq s \leq t} |G(x)(s) - G(y)(s)| \leq C_T \int_0^t |x(u) - y(u)| du,$$

then there exists a unique solution to the generalized Skorokhod problem associated with the functional G and the matrix P .

PROOF. Define the sequence $(X_N(t))$ by induction $(X^0(t), R^0(t)) = 0$ and, for $N \geq 1$, (X^{N+1}, R^{N+1}) is the solution of the Skorokhod problem (SP) associated with $G(X^N)$, in particular,

$$X^{N+1}(t) = G(X^N)(t) + R^{N+1}(t) \text{ and } \int_{\mathbb{R}_+} X^{N+1}(u) dR^{N+1}(u) = 0.$$

The existence of such a solution is guaranteed as well as the Lipschitz property of the solutions of a classical Skorokhod problem, see Proposition D.4 of Robert [Rob03], this gives the existence of some constant K_T such that, for all $N \geq 1$ and $0 \leq t \leq T$,

$$\|X^{N+1} - X^N\|_{\infty, t} \leq K_T \|G(X^N) - G(X^{N-1})\|_{\infty, t},$$

where $\|h\|_{\infty, T} = \sup\{|h(s)| : 0 \leq s \leq T\}$. From Relation (5.5), this implies that

$$\|X^{N+1} - X^N\|_{\infty, t} \leq \alpha \int_0^t \|X^N - X^{N-1}\|_{\infty, u} du,$$

with $\alpha = K_T C_T$. The iteration of the last relation yields the inequality

$$\|X^{N+1} - X^N\|_{\infty, t} \leq \frac{\alpha^N t^{N-1}}{(N-1)!} \int_0^t \|X^1\|_{\infty, u} du, \quad 0 \leq t \leq T.$$

One concludes that the sequence $(X^N(t))$ is converging uniformly on compact sets and consequently the same is true for the sequence $(R^N(t))$. Let $(X(t))$ and $(R(t))$ be the limit of these sequences. By continuity of the SP, the couple $((X(t)), (R(t)))$ is the solution of the SP associated with $G(X)$, and hence a solution of the GSP associated with G .

Uniqueness. Let $(Y(t))$ be another solution of the GSP associated with G . In the same way as before, one gets by induction, for $0 \leq t \leq T$,

$$\|X - Y\|_{\infty, t} \leq \frac{(\alpha t)^N}{N!} \int_0^t \|X - Y\|_{\infty, u} du,$$

and by letting N go to infinity, one concludes that $X = Y$. The proposition is proved. ■

References for Chapter IV

- [AGC08] Jesús R. Artalejo and Antonio Gómez-Corral. *Retrial queueing systems*. Springer-Verlag, Berlin, 2008. A computational approach.
- [AO76] Robert F. Anderson and Steven Orey. Small random perturbation of dynamical systems with reflecting boundary. *Nagoya Math. J.*, 60:189–216, 1976.
- [Bil99] Patrick Billingsley. *Convergence of Probability Measures (second edition)*. Wiley Series in Probability and Statistics. Wiley-Interscience, 1999.
- [Bro78] Tim Brown. A martingale approach to the Poisson convergence of simple point processes. *The Annals of Probability*, 6(4):615–628, 1978.

- [CDH⁺06] Byung-Gon Chun, Frank Dabek, Andreas Haeberlen, Emil Sit, Hakim Weatherspoon, M. Frans Kaashoek, John Kubiatowicz, and Robert Morris. Efficient replica maintenance for distributed storage systems. In *NSDI*, pages 4–4, 2006.
- [Daw93] Donald A. Dawson. Measure-valued Markov processes. In *École d'Été de Probabilités de Saint-Flour XXI—1991*, volume 1541 of *Lecture Notes in Math.*, pages 1–260. Springer, Berlin, 1993.
- [DS65] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *Journal of Applied Probability*, 2(1):88–100, 1965.
- [EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and convergence*. Wiley, 1986.
- [EKCM78] N. El Karoui and M. Chaleyat-Maurel. *Temps locaux*, volume 52-53, chapter Un problème de réflexion et ses applications au temps local et aux équations différentielles stochastiques sur $\mathbb{R}$, pages 117–144. Société Mathématique de France, 1978. Exposés du Séminaire J. Azéma-M. Yor, Paris, 1976–1977.
- [FKMP95] P. A. Ferrari, H. Kesten, S. Martinez, and P. Picco. Existence of quasi-stationary distributions. A renewal dynamical approach. *The Annals of Probability*, 23(2):501–521, 1995.
- [FR12] Mathieu Feuillet and Philippe Robert. A scaling analysis of a transient stochastic network (II): A mean-field picture. Preprint, 2012.
- [FT97] Guennadi I. Falin and James G. C. Templeton. *Retrials Queues*. Chapman & Hall, 1997.
- [FW98] M. Freidlin and A. Wentzell. *Random Perturbations of Dynamical Systems*. Springer, 1998. 2nd Edition.
- [GH90] John Guckenheimer and Philip Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, volume 42 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1990. Revised and corrected reprint of the 1983 original.
- [HK94] P. J. Hunt and T. G. Kurtz. Large loss networks. *Stochastic Processes and their Applications*, 53(2):363–378, 1994.
- [HR81] J.M. Harrison and M.I. Reiman. Reflected Brownian motion on an orthant. *Annals of Probability*, 9(2):302–308, 1981.
- [Kel91] Frank P. Kelly. Loss networks. *Annals of Applied Probability*, 1(3):319–378, 1991.
- [Kha80] R. Z. Khasminski. *Stochastic stability of differential equations*. Sijthoff & Noordhoff, Alphen aan den Rijn, 1980. Translated from the Russian by D. Louvish.
- [Kin90] Peter J.B. King. *Computer and communication systems performance modelling*. Computer-Communication Networks. Prentice Hall, 1990.
- [Kur92] T.G. Kurtz. Averaging for martingale problems and stochastic approximation. In *Applied Stochastic Analysis, US-French Workshop*, volume 177 of *Lecture notes in Control and Information sciences*, pages 186–209. Springer Verlag, 1992.
- [KW86] Yuji Kasahara and Shinzo Watanabe. Limit theorems for point processes and their functionals. *Journal of the Mathematical Society of Japan*, 38(3):543–574, 1986.

- [PBS07] Fabio Picconi, Bruno Baynat, and Pierre Sens. An analytical estimation of durability in DHTs. In Tomasz Janowski and Hrushikesh Mohanty, editors, *Distributed Computing and Internet Technology*, volume 4882 of *Lecture Notes in Computer Science*, pages 184–196. Springer, 2007.
- [PW11] Ohad Perry and Ward Whitt. An ODE for an overloaded x model involving a stochastic averaging principle. *Stochastic Systems*, 1(1):59–108, 2011.
- [Ram06] Kavita Ramanan. Reflected diffusions defined via the extended Skorokhod map. *Electronic Journal of Probability*, 11:934–992, 2006.
- [RD01] Antony Rowstron and Peter Druschel. Storage management and caching in PAST, a large-scale, persistent peer-to-peer storage utility. In *Proceedings of SOSR*, pages 188–201. ACM, 2001.
- [RGK⁺05] Sean Rhea, Brighten Godfrey, Brad Karp, John Kubiatowicz, Sylvia Ratnasamy, Scott Shenker, Ion Stoica, and Harlan Yu. OpenDHT: a public DHT service and its uses. In *Proceedings of SIGCOMM*, pages 73–84, 2005.
- [Rob03] Philippe Robert. *Stochastic Networks and Queues*. Stochastic Modeling and Applied Probability Series. Springer-Verlag, New York, 2003. xvii+398 pp.
- [RP06] S. Ramabhadran and J. Pasquale. Analysis of Long-Running Replicated Systems. In *INFOCOM 2006*, pages 1–9, 2006.
- [Rud87] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.
- [Sko61] A.V. Skorokhod. Stochastic equations for diffusions in a bounded region. *Theory of Probability and Its Applications*, 6:264–274, 1961.
- [TW93] L.M. Taylor and R.J. Williams. Existence and uniqueness of semi-martingale reflecting Brownian motions in an orthant. *Probability Theory and Related Fields*, 96:283–317, 1993.

CHAPTER V

On the Transient Behavior of Ehrenfest and Engset Processes

Contents

1. Introduction	111
2. The Stochastic Model	114
3. Positive Martingales	115
4. A Fluid Picture	120
5. Super-Critical Regime	121
6. Sub-Critical Regime	123
7. Critical Regime	126
References for Chapter V	128

1. Introduction

The Ehrenfest Process. In this chapter we consider the following continuous-time version of the classical Ehrenfest urn model. This process has been introduced to study the heat exchange between bodies. One assumes that each particle of a set of N particles is located in one of two boxes (bodies), 0 and 1 say. A particle in box 0 [resp. 1] goes into box 1 [resp. 0] at rate ν [resp. μ]. One denotes by $E_N(t)$ the number of particles in box 1 at time $t \geq 0$. This birth-and-death process can also be represented as $E_N(t) = Y_1(t) + \dots + Y_N(t)$, where $(Y_i(t), 1 \leq i \leq N)$ are N i.i.d. Markov jump processes with values in $\{0, 1\}$.

Originally, the model is a discrete-time process $(Z_N(k))$ and each unit of time a particle is taken at random (i.e. all particles equally likely) to be moved from one box to the other; $Z_N(k)$ is the number of particles in box 1 at time k . This corresponds to the symmetrical case $\mu = \nu$ and, clearly, $E_N(t)$ can be represented as $Z_N(\mathcal{N}_{N\mu}((0, t]))$ if $\mathcal{N}_{N\mu}$ is a Poisson process with rate $N\mu$. The process $(E_N(t))$ follows the same path as $(Z_N(k))$ but on a time scale with a factor $N\mu$. The Markov chain $(Z_N(k))$ is also a random walk on the graph of the *hypercube* $\{0, 1\}^N$, where edges connect elements that differ at only one coordinate. The equilibrium properties of $(E_N(t))$ and $(Z_N(k))$ are fairly well known, in particular a quite precise estimate of the duration of time to reach equilibrium is available. See Diaconis et al. [DGM90].

Results on transient quantities of this process, like the distribution of the first time when box 0 is empty, are rarer. There are generic results on birth-and-death process which describe some of these distributions in terms of spectral characteristics of the associated infinitesimal generator: the spectral measure and a family of orthogonal polynomials. See Karlin and McGregor [KM57, KM59]. In practice, the corresponding orthogonal polynomials and, sometimes, their spectral measure do not have a simple representation; this complicates significantly detailed investigations of these hitting times. In the symmetrical case $\mu = \nu$, Palacios [Pal93]

gives closed-form expressions for the averages of hitting times and Bingham [Bin91] and Flajolet and Huillet [FH08] obtains a representation of their distributions. In the general case Crescenzo [DC98] and Flegg et al. [FGP08] provide expressions of the densities. It turns out that the expressions obtained in these papers involve, in general, sums of combinatorial terms for which asymptotic results (when N goes to infinity) may be difficult to obtain. Note that this is nevertheless done in the symmetrical case in Flajolet and Huillet [FH08].

The Engset Process. This is one of the oldest stochastic models of communication networks, see Engset [Eng98]. For this model there are N sources of communication which are active (state 1) or inactive (state 0). An active source becomes inactive at rate μ . The total number of simultaneous active sources cannot exceed the quantity C_N , the number of circuits of the network. An inactive source can therefore become active only if there are already strictly less than C_N active sources; In this case it occurs at rate ν . If $X_N(t)$ is the number of active communications at time t , when $C_N = N$ the process is just the Ehrenfest process. Otherwise $(X_N(t))$ can be described as a reflected version of $(E_N(t))$. At equilibrium the expression for the probability that X_N is equal to C_N is known under the name of *Engset Formula*. For transient characteristics, the important quantity is the time it takes to have the full capacity C_N of the network used. For this reason, the distribution of the hitting time of C_N by $(X_N(t))$ is of special interest. To the best of our knowledge, results concerning this hitting time are quite rare, in particular for possible asymptotics when N goes to infinity.

A Storage System. Another, more recent, motivation for considering $(X_N(t))$ is the stochastic analysis of a storage system where files are duplicated on C_N servers. Each server breaks down independently at rate μ , in which case it is repaired but all its files are lost. As a simplified model, $X_N(t)$ is defined as the number of copies of a specified file, if $X_N(t) = x$ then a copy of the file is lost if one of the servers breaks down, i.e. at rate $x\mu$. If $X_N(t) = 0$, there is no copy of the file in the system, it is lost so that 0 is an absorbing point. If $0 < x < C_N$ then a new copy of the file may be added but at rate $\nu(N - x)$; $N\nu$ is the maximal capacity of duplication of the system. It is easily seen that as long as $(X_N(t))$ does not hit 0, $(X_N(t))$ is precisely the Engset process. In this context it is of special interest to study the distribution of the first time when the file is lost, i.e. the hitting time of 0. Another point of view is studied in Chapter IV.

A Collection of Exponential Martingales. This chapter relies heavily on the use of martingales to derive explicit, simple, expressions of the Laplace transforms of the hitting times of a state of the system. One obtains expressions of these transforms as ratios of simple integrals for which various asymptotic results, when N goes to infinity, can be derived quite easily with standard technical tools. In particular one does not need to cope with the asymptotic behavior of sums of combinatorial expressions. Quite surprisingly, up to now, martingales did not play a major role in the previous studies of the Ehrenfest process. One can mention Simatos and Tibi [ST10] where a martingale approach is used to estimate certain exit times for multi-dimensional Ehrenfest processes. It is one of the results of this chapter to show that a simple and important family of martingales allows a quite detailed investigation of this process, and also of its variants like the Engset process.

The key ingredient of this chapter is a set of non-negative martingales which will be called exponential martingales. If $(M(t))$ is a martingale on some probability space, the associated *exponential martingale* is the solution $(Z(t))$ of the stochastic

differential equation (SDE)

$$dZ(t) = Z(t-) dM(t), \quad t \geq 0,$$

where $Y(t-)$ is the left limit of Y at t and $dY(t)$ is the limit on the right of t of $s \mapsto Y(s) - Y(t-)$. It is called the *Doléans exponential* of $(M(t))$. See Chapter IV of Rogers and Williams [RW87] for example. Although there is an exponential martingale for each martingale, a small subset of these martingales plays an important role. For the standard Brownian motion $(B(t))$ this is the martingale

$$\left(\exp \left(\beta B(t) - \beta^2 t/2 \right) \right),$$

for a fixed $\beta \in \mathbb{R}$. It is very helpful to derive the explicit expressions for Laplace transforms of hitting times associated with Brownian motion. See Revuz and Yor [RY99]. For jump processes, this is less clear. It does not seem that a “classification” of exponential martingales exists in general, even for birth-and-death processes. See Chapter V of Feinsilver [Fei78] and Schoutens [Sch00] for related questions. Some examples of important processes are reviewed.

For $\xi \in \mathbb{R}_+$, $\mathcal{N}_\xi$ denotes a Poisson process with rate ξ and $(\mathcal{N}_{\xi,i})$ an i.i.d. sequence of such Poisson processes. All Poisson processes are assumed to be independent.

- **Random Walks.** The classical exponential martingale associated with the random walk $(S(t)) = (\mathcal{N}_\lambda([0, t]) - \mathcal{N}_\mu([0, t]))$ is given by, for $\beta \in \mathbb{R}$,

$$(1.1) \quad \left(\exp \left[-\beta S(t) - t \left(\lambda (1 - e^{-\beta}) + \mu (1 - e^\beta) \right) \right] \right).$$

It is the exponential martingale associated with the martingale

$$\left(\beta(S(t) - (\lambda - \mu)t) \right).$$

The corresponding reflected process is the $M/M/1$ queue with input rate λ and service rate μ .

- **The $M/M/\infty$ Process.** This is a classical Markov process on $\mathbb{N}$ whose Q -matrix $Q = (q(x, y))$ is, for $x \in \mathbb{N}$, $q(x, x+1) = \lambda$ and $q^N(x, x-1) = \mu x$. It can be also seen as a kind of discrete Ornstein-Uhlenbeck process, defined as the solution $(L(t))$ of the following SDE:

$$dL(t) = \mathcal{N}_\lambda(dt) - \sum_{i=1}^{L(t-)} \mathcal{N}_{\mu,i}(dt).$$

The following martingale was introduced in Fricker et al. [FRT99], for $\beta \in \mathbb{R}$:

$$(1.2) \quad \left((1 + \beta e^{\mu t})^{L(t)} \exp(-\beta e^{\mu t} \lambda / \mu) \right).$$

It is the exponential martingale associated with the martingale

$$\left(\int_0^t (1 + \beta e^{\mu s}) [\mathcal{N}_\lambda(ds) - \lambda ds] - \sum_{i=1}^{+\infty} \int_0^t \frac{1}{1 + \beta e^{\mu s}} \mathbb{1}_{\{i < L(s-)\}} [\mathcal{N}_{\mu,i}(ds) - \mu ds] \right).$$

- **The Ehrenfest Process.** Such a process $(E_N(t))$ with N particles can be seen as the solution of the SDE

$$dL(t) = \sum_{i=1}^{N-L(t-)} \mathcal{N}_{\nu,i}(dt) - \sum_{i=1}^{L(t-)} \mathcal{N}_{\mu,i}(dt).$$

It will be seen that the corresponding exponential martingale is given, for $\beta \in \mathbb{R}$, by

$$(1.3) \quad \left(\left(1 - \beta \mu e^{(\mu+\nu)t} \right)^{E_N(t)} \left(1 + \beta \nu e^{(\mu+\nu)t} \right)^{N-E_N(t)} \right).$$

It is the exponential martingale associated with the martingale defined by, up to the multiplicative factor $\beta(\mu+\nu)$,

$$\left(\sum_{i=1}^N \int_0^t e^{(\mu+\nu)s} \left(\mathbb{1}_{\{Y_i(s-)=1\}} [\mathcal{N}_{\mu,i}(ds) - \mu ds] - \mathbb{1}_{\{Y_i(s-)=0\}} [\mathcal{N}_{\nu,i}(ds) - \nu ds] \right) \right),$$

where $(Y_i(t))$ are such that $E_N(t) = Y_1(t) + \dots + Y_N(t)$. Recall that the Engset process is a reflected version of this process.

From these exponential martingales, explicit expressions of Laplace transforms of the distribution of hitting times associated with these processes can be derived. It may be not be as straightforward as in the case of Brownian motion since the space variable t is not separated from the space variable, but a convenient integration with respect to the free parameter β solves the problem. See Chapters 5 and 6 of Robert [Rob03] for the $M/M/1$ and $M/M/\infty$ processes and Section 3 for the Ehrenfest process. See also Schoutens and Teugels [ST98].

From the point of view of potential theory, these martingales are associated with the set of extreme harmonic functions. This statement can be made precise in terms of space-time Martin boundary. See Lamperti and Snell [LS63] and the discussion in Section 3.

Organization of the Chapter. In Section 2, the two stochastic processes are defined precisely. In Section 3, the exponential martingale for the Ehrenfest process is introduced, and, based on it, several interesting martingales for the Ehrenfest process and the Engset process are constructed. As a corollary, closed-form expressions of the Laplace transform of the hitting time of a given state are obtained as the ratio of simple integrals. This holds in particular for the blocking time for the Engset process. The last three sections are devoted to the analysis of the asymptotic behavior of the distribution of the hitting time of C_N and 0 when N goes to infinity in such a way that $C_N \sim \eta N$ for some $0 < \eta \leq 1$. Each section considers one of the three possible regimes: *sub-critical* when $\nu < \eta$, the process “lives” in the neighborhood of νN , *super-critical* when the difference $C_N - X_N(t)$ converges to a finite process, and, finally, *critical* when $C_N - X_N(t)$ is of the order of $\sqrt{N}$. For each regime, by taking advantage of the simple expressions of the corresponding Laplace transforms obtained, various convergence-in-distribution results are derived.

2. The Stochastic Model

The Ehrenfest process. Let $(Y(t))$ be the simple Markov process on $\{0, 1\}$ whose Q -matrix Q_Y is given by

$$Q_Y = \begin{pmatrix} -\nu & \nu \\ \mu & -\mu \end{pmatrix}.$$

For $N \in \mathbb{N}$, if $(Y_i(t))$, $1 \leq i \leq N$, are N independent copies of $(Y(t))$, the Ehrenfest process $(E_N(t))$ is also a birth-and-death process but on the state space $\{0, 1, \dots, N\}$, defined as

$$(2.1) \quad E_N(t) = Y_1(t) + Y_2(t) + \dots + Y_N(t).$$

The Q -matrix of $(E_N(t))$ will be denoted as Q_{E_N} , for $x \in \{0, \dots, N\}$,

$$(2.2) \quad q_{E_N}(x, x-1) = \mu x \text{ and } q_{E_N}(x, x+1) = \nu(N-x).$$

The Engset process. For $1 \leq C_N \leq N$, the Engset process $(X_N(t))$ is a birth-and-death process on $\{1, \dots, C_N\}$ which can be seen as a reflected version of $(E_N(t))$ at the boundary C_N , i.e. its Q -matrix $Q_{X_N} = (q_N(x, y))$ is given by, for $0 \leq x \leq C_N$,

$$(2.3) \quad q_{X_N}(x, x-1) = \mu x \text{ and } q_{X_N}(x, x+1) = \nu(N-x) \text{ if } x < C_N.$$

In particular the process $(X_N(t))$ has the same distribution as the process $(E_N(t))$ constrained to the state space $\{0, \dots, C_N\}$. In particular, when $C_N = N$, the two processes $(X_N(t))$ and $(E_N(t))$ starting from the same initial state have the same distribution.

As ergodic birth-and-death processes, the Markov processes $(X_N(t))$ and $(E_N(t))$ are reversible and their stationary distribution at x is, up to a normalization constant, given by

$$\binom{N}{x} \left(\frac{\nu}{\mu}\right)^x,$$

if x is an element of their respective state space.

Normalization of the time scale. By considering the time scale $t \rightarrow t/(\nu + \mu)$ in the analysis of the processes $(E_N(t))$ and $(X_N(t))$, it can be assumed without any loss of generality that $\nu + \mu = 1$. This will be the case in this chapter.

A limiting regime. In the following it will be assumed that the constant C_N is asymptotically of the order of N , i.e. that

$$(2.4) \quad \eta := \lim_{N \rightarrow +\infty} C_N/N,$$

holds for some $\eta \in (0, 1]$.

For t large the probability that the variable $Y(t)$ defined above is at 1 is given by its equilibrium distribution at 1, that is ν . The law of large numbers gives that $E_N(t)$ is of the order of $N\nu$. Roughly speaking, if $N\nu < C_N$ for N large, i.e. $\nu < \eta$, then the boundary at C_N should not play a significant role for first-order quantities related to $(X_N(t))$ and therefore the processes $(X_N(t))$ and $(E_N(t))$ should have the same behavior in the limit. On the contrary if $\nu \geq \eta$, due to the reflecting boundary at C_N for $(X_N(t))$, the Ehrenfest and the Engset processes should behave differently. This phenomenon will be stated more precisely in the last three sections.

3. Positive Martingales

Several families of positive martingales for the Ehrenfest and the Engset processes are introduced in this section. More specifically, when $(Z(t))$ is either $(X_N(t))$ or $(E_N(t))$, one identifies a set of functions $f : \mathbb{N} \times \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that the process $(f(Z(t), t))$ is a martingale, i.e. that, for $t \geq 0$, the relation

$$\mathbb{E}(f(Z(t), t) \mid \mathcal{F}_t) = f(Z(s), s), \text{ for } s \leq t,$$

holds almost surely, where $(\mathcal{F}_t)$ is the natural filtration associated with $(Z(t))$.

If $Q_Z = (q_Z(\cdot, \cdot))$ is the Q -matrix of $(Z(t))$, this probabilistic property is equivalent to the fact that the function f is *space-time harmonic* with respect to Q_Z , i.e. that the relation

$$(3.1) \quad \frac{\partial}{\partial t} f(x, t) + Q_Z(f(\cdot, t))(x) = 0$$

holds for $x \in \mathbb{N}$ and $t \geq 0$, where for $h : \mathbb{N} \mapsto \mathbb{R}_+$,

$$Q_Z(h)(x) = \sum_{y \in \mathbb{N}} q(x, y)h(y).$$

A space-time harmonic function of the Markov process $(Z(t))$ is just an harmonic function of the transient Markov process $((Z(t), t))$. See Appendix B of Robert [Rob03] for example. When $(Z(t))$ is $(E_N(t))$, one will prove that there is a family f_β , $\beta \in \mathbb{R}$ of such functions. As it will be seen, these martingales

can be interpreted as exponential martingales. They will give in particular an explicit expression for the Laplace transform of the hitting times associated with both processes $(E_N(t))$ and $(X_N(t))$.

For a given birth-and-death process, there is already a complete description of all such positive martingales. This is the (space-time) Martin boundary of the birth-and-death process, see Lamperti and Snell [LS63]. This description is expressed in terms of the orthogonal polynomials associated with the birth-and-death process which may be defined by an induction relation or by the measure with respect to which they are orthogonal, see Karlin and Mc Gregor [KM59] and Chapter 2 of Szegő [Sze75]. As long as moments of some transient characteristics are investigated, these martingales can be used but, in general, they do not seem to be really helpful to analyze the distributions of hitting times.

This situation is quite classical, for Brownian motion for example, for which there is a family of martingales indexed by $N \in \mathbb{N}$: if H_N is the Hermite polynomial of degree N , then $(M_N(t)) = (t^{N/2}H_N(B(t)/\sqrt{t}))$ is a martingale. Another family of martingales is provided by the exponential martingale $(\exp(\beta B(t) - \beta^2 t/2))$ indexed by $\beta \in \mathbb{R}$. This exponential martingale can be expressed as a weighted sum of the martingales $(M_N(t))$, but to get explicit expressions of the distributions of hitting times, it is the really useful martingale. See Propositions 3.4 and 3.8 of Chapter 3 of Revuz and Yor [RY99] for example. In the case of birth-and-death processes, a general result concerning the construction of such exponential martingales from the martingales associated with the orthogonal polynomials does not seem to exist.

3.1. Exponential Martingales for the Ehrenfest Process. Due to the simple structure of the Ehrenfest process, these martingales are really elementary. Nevertheless they play a fundamental role; most of the asymptotic results obtained in this chapter are based on these martingales. A more general version in a multi-dimensional context has been introduced by Simatos and Tibi [ST10].

PROPOSITION 3.1 (Exponential martingales). *For $\beta \in \mathbb{R}$, the process*

$$(3.2) \quad (M_N^\beta(t)) = \left((1 - \beta\mu e^t)^{E_N(t)} (1 + \beta\nu e^t)^{N - E_N(t)} \right)$$

is a martingale.

PROOF. Define, for $N \geq 1$, $t \geq 0$ and $0 \leq x \leq N$,

$$(3.3) \quad h_N^\beta(x, t) = (1 - \beta\mu e^t)^x (1 + \beta\nu e^t)^{N-x}.$$

Clearly, the relations

$$\frac{\partial}{\partial t} h_1^\beta(0, t) = \beta\nu e^t = -Q_Y(h_1^\beta(\cdot, t))(0) \quad \text{and} \quad \frac{\partial}{\partial t} h_1^\beta(1, t) = -\beta\mu e^t = -Q_Y(h_1^\beta(\cdot, t))(1),$$

hold, where Q_Y is the Q -matrix of $(Y(t))$ introduced in Section 2. Consequently, the function h_1^β is space-time harmonic for the matrix Q_Y , equivalently $(f(Y(t), t))$ is a martingale.

If $((Y_i(t)), 1 \leq i \leq N)$ are N independent copies of $(Y(t))$, then, by using the independence of the processes $(Y_i(t))$, $i = 1, \dots, N$ and Equation (2.1), one gets that the process

$$\left(\prod_{i=1}^N f(Y_i(t), t) \right) \stackrel{\text{dist.}}{=} (M_N^\beta(t)) = (h_N^\beta(E_N(t), t))$$

is also a martingale with respect to the filtration $(\mathcal{F}_t) = (\sigma(Y_i(s), s \leq t, 1 \leq i \leq N))$. In particular the function h_N^β is space-time harmonic with respect to Q_{E_N} . ■

Martingales Associated with Orthogonal Polynomials. As remarked by Karlin and McGregor [KM57, KM65], the polynomials associated with the Ehrenfest process are the $N + 1$ *Krawtchouk polynomials* $(\mathcal{K}_n^N; 0 \leq n \leq N)$ defined by

$$(3.4) \quad \mathcal{K}_n^N(x) = \binom{N}{n}^{-1} \sum_{\ell=0}^n (-1)^\ell \binom{x}{\ell} \binom{N-x}{n-\ell} \left(\frac{\mu}{\nu}\right)^\ell, \quad 0 \leq n, x \leq N.$$

These polynomials are orthogonal with respect to the binomial distribution

$$\left(\binom{N}{k} \nu^k \mu^{N-k}, 0 \leq k \leq N \right).$$

The classical identity, see Karlin and McGregor [KM65] for example,

$$(3.5) \quad \sum_{n=0}^N \binom{N}{n} \mathcal{K}_n^N(x) u^n = (1+u)^{N-x} \left(1 - \frac{\mu}{\nu} u\right)^x, \quad u \in \mathbb{R},$$

and the above proposition give that

$$\left(\sum_{n=0}^N \binom{N}{n} \mathcal{K}_n^N(E_N(t)) \beta^n e^{nt} \right) = \left(M_N^{\beta/\nu}(t) \right)$$

is a martingale. As a consequence one gets that, for any $0 \leq n \leq N$, the process $(\mathcal{K}_n^N(E_N(t)) \exp(nt))$ is a martingale. The martingale $(M_N^\beta(t))$ can thus be seen as an encoding of these $(N+1)$ martingales in the same way as the exponential martingale of the Brownian motion with the Hermite polynomials, or the martingale (1.2) with the Poisson-Charlier polynomials. See Schoutens [Sch00] and Schoutens and Teugel [ST98] and see Robert [Rob03].

Note that the space variable $E_N(t)$ and the time variable t are not separated in Expression (3.2) of the exponential martingale. Provided that it can be used, Doob's optional stopping theorem applied to some hitting time of some specified state x does not give useful information on the distribution of this variable. But given that there is a free parameter $\beta \in \mathbb{R}$ in Expression (3.2) and that the martingale property is clearly preserved by integration with respect to β , one may try to find a measure on $\mathbb{R}_+$ that will "separate" the space and time variables. The following proposition uses such a method.

PROPOSITION 3.2. *For any $\alpha > 0$ and $t \geq 0$, if*

$$(3.6) \quad \begin{cases} I_\alpha^N(t) = e^{-\alpha t} \int_0^1 (1-u)^{E_N(t)} \left(1 + \frac{\nu}{\mu} u\right)^{N-E_N(t)} u^{\alpha-1} du, \\ J_\alpha^N(t) = e^{-\alpha t} \int_0^1 (1-u)^{N-E_N(t)} \left(1 + \frac{\mu}{\nu} u\right)^{E_N(t)} u^{\alpha-1} du, \end{cases}$$

and

$$T_x^{E_N} = \inf\{t > 0 : E_N(t) = x\}, \quad 0 \leq x \leq N,$$

then $(I_\alpha(t \wedge T_0^{E_N}))$ and $(J_\alpha(t \wedge T_N^{E_N}))$ are martingales.

PROOF. Since $(N - E_N(t))$ is also an Ehrenfest process but with the two parameters μ and ν exchanged, one needs only to prove that the process $(I_\alpha(t \wedge T_0^{E_N}))$ is a martingale.

Define

$$f_N(x, t) := \int_0^{e^{-t}/\mu} h_N^\beta(x, t) \beta^{\alpha-1} d\beta,$$

where h_N^β is defined by Equation (3.3); then

$$\frac{\partial f_N}{\partial t}(x, t) = \int_0^{e^{-t}/\mu} \frac{\partial h_N^\beta}{\partial t}(x, t) \beta^{\alpha-1} d\beta - \frac{e^{-\alpha t}}{\mu^\alpha} h_N^{e^{-t}/\mu}(x, t).$$

Note that the last term of the above expression is 0 if $x \neq 0$. Consequently, for $x \neq 0$ and $t \geq 0$,

$$\frac{\partial f_N}{\partial t}(x, t) + Q_{E_N}(f_N)(x, t) = \int_0^{e^{-t}/\mu} \left[\frac{\partial h_N^\beta}{\partial t}(x, t) + Q_{E_N}(h_N^\beta)(x, t) \right] \beta^{\alpha-1} d\beta = 0,$$

because h_N^β is space-time harmonic with respect to Q_{E_N} as in the proof of Proposition 3.1. In other words, the function f_N is space-time harmonic for the Q -matrix of the stopped process $(E_N(t \wedge T_0^{E_N}))$, hence

$$(f_N(E_N(t \wedge T_0^{E_N}), t \wedge T_0^{E_N})) = (I_\alpha^N(t \wedge T_0^{E_N}))$$

is a martingale. ■

It is now easy to get a representation of the Laplace transform of the hitting times for the Ehrenfest process.

PROPOSITION 3.3 (Laplace Transform of Hitting Times). *For $0 \leq x \leq y \leq N$ and if $T_x^{E_N} = \inf\{t > 0 : E_N(t) = x\}$, the relations*

$$(3.7) \quad \mathbb{E}_y(e^{-\alpha T_x^{E_N}}) = \frac{B_x(\alpha)}{B_y(\alpha)}, \quad \text{and} \quad \mathbb{E}_x(e^{-\alpha T_y^{E_N}}) = \frac{D_x(\alpha)}{D_y(\alpha)}$$

hold, with

$$(3.8) \quad \begin{cases} B_x(\alpha) = \int_0^1 (1-u)^x \left(1 + \frac{\nu}{\mu}u\right)^{N-x} u^{\alpha-1} du, \\ D_x(\alpha) = \int_0^1 (1-u)^{N-x} \left(1 + \frac{\mu}{\nu}u\right)^x u^{\alpha-1} du. \end{cases}$$

There is in fact only one result here since $(N - E_N(t))$ is, as already remarked, an Ehrenfest process with the parameters μ and ν exchanged. The second relation of (3.7) is therefore a consequence of the first one.

PROOF. The martingale $I_\alpha(t \wedge T_0)$ is bounded and so uniformly integrable. Therefore, Doob's optional stopping Theorem gives the relation

$$\mathbb{E}_y(I_\alpha(0)) = \mathbb{E}_y(I_\alpha(T_x)),$$

the first relation of (3.7) follows. ■

By expanding one of the terms of the integrand of (3.8), one gets

$$\begin{aligned} B_x(\alpha) &= \sum_{k=0}^{N-x} \binom{N-x}{k} \left(\frac{\nu}{\mu}\right)^k \int_0^1 (1-u)^x u^{\alpha+k-1} du \\ &= \sum_{k=0}^{N-x} \binom{N-x}{k} \left(\frac{\nu}{\mu}\right)^k \frac{\Gamma(x+1)\Gamma(\alpha+k)}{\Gamma(\alpha+x+k+1)}, \end{aligned}$$

by Euler's integral for Beta functions and Gamma functions. See Whittaker and Watson [WW96] page 254 for example. Laplace transforms of hitting times can therefore also be expressed as ratio of such sums, as is quite common for hitting times of birth-and-death processes. See Equation (4.4) of Karlin and McGregor [KM65] for example. Flajolet and Huillet [FH08] use this kind of representations in the symmetrical case. As will be seen, from the compact representation (3.8) with integrals, one will get asymptotic results for the distribution of these variables with standard techniques.

3.2. Martingales for the Engset Process. It has been seen that the Engset process $(X_N(t))$ is a reflected version of the process $(E_N(t))$ at the boundary C_N . The two families of martingales of Proposition 3.2 cannot be used directly if the sample path of $(E_N(t))$ may exceed C_N , when the hitting times of 0 is analyzed for example. The idea is to construct a linear combination of the martingales $(I_\alpha(t))$ and $(J_\alpha(t))$ such that the space-time harmonicity of the corresponding function which is valid when the space variable in $\{1, \dots, C_N - 1\}$ holds also at the boundary C_N . This method has been used in Kennedy [Ken76] in the case of reflected random walks.

PROPOSITION 3.4. *For $\alpha > 0$, define*

$$(3.9) \quad \begin{cases} b_N(\alpha) = \nu \int_0^1 (1-u)^{C_N} \left(1 + \frac{\nu}{\mu}u\right)^{N-C_N-1} u^\alpha du, \\ d_N(\alpha) = \mu \int_0^1 (1-u)^{N-C_N-1} \left(1 + \frac{\mu}{\nu}u\right)^{C_N} u^\alpha du, \end{cases}$$

and

$$(K_\alpha^N(t)) = (d_N(\alpha)I_\alpha^N(t) + b_N(\alpha)J_\alpha^N(t)),$$

where $(I_\alpha^N(t))$ and $(J_\alpha^N(t))$ are defined by (3.6) with $E_N(t)$ replaced by $X_N(t)$. Then, if $T_0^{X_N}$ is the hitting time of 0 by $(X_N(t))$, the process $(K_\alpha^N(t \wedge T_0^{X_N}))$ is a martingale.

PROOF. Define the function g_α such that, for all $t \geq 0$ and $0 \leq x \leq C_N$,

$$\begin{aligned} g_\alpha(x, t) = d_N(\alpha) e^{-\alpha t} \int_0^1 (1-u)^x \left(1 + \frac{\nu}{\mu}u\right)^{N-x} u^{\alpha-1} du \\ + b_N(\alpha) e^{-\alpha t} \int_0^1 (1-u)^{N-x} \left(1 + \frac{\mu}{\nu}u\right)^x u^{\alpha-1} du. \end{aligned}$$

The function g_α is space-time harmonic for the matrix Q_{X_N} on $\{1, \dots, C_N - 1\}$, that is

$$\left[\frac{\partial g_\alpha}{\partial t} + Q_{X_N}(g_\alpha) \right] (x, t) = 0, \quad 0 < x < C_N,$$

since the two matrices Q_{X_N} and Q_{E_N} are identical as long as the starting point is in $\{1, \dots, C_N - 1\}$, and $(I_\alpha^N(t \wedge T_0^{E_N}))$ and $(J_\alpha^N(t \wedge T_0^{E_N}))$ are martingales by Proposition 3.2.

The space-time harmonicity of g_α for the matrix Q_{E_N} at $C_N < N$ gives

$$\left[\frac{\partial g_\alpha}{\partial t} + Q_{X_N}(g_\alpha) \right] (C_N, t) = -\nu(N - C_N) [g_\alpha(C_N + 1, t) - g_\alpha(C_N, t)].$$

For $0 \leq y \leq 1$, one has

$$(1-y)^{C_N+1} \left(1 + \frac{\nu}{\mu}y\right)^{N-C_N-1} - (1-y)^{C_N} \left(1 + \frac{\nu}{\mu}y\right)^{N-C_N} = -\frac{y}{\mu} (1-y)^{C_N} \left(1 + \frac{\nu}{\mu}y\right)^{N-C_N-1},$$

and, similarly

$$\left[\frac{\partial g_\alpha}{\partial t} + Q_{X_N}(g_\alpha) \right] (C_N, t) = \frac{(N - C_N)}{\mu} (d_N(\alpha)b_N(\alpha) - b_N(\alpha)d_N(\alpha)) = 0.$$

The function g_α is therefore space-time harmonic for the Q -matrix of the stopped process $(X_N(t \wedge T_0^{X_N}))$, hence the process $(K_\alpha^N(t \wedge T_0^{X_N}))$ is a martingale. ■

PROPOSITION 3.5 (Laplace Transform of Hitting Times for Engset Process). *For $0 \leq x \leq y \leq C_N$, if $T_x^{X_N} = \inf\{s \geq 0 : X_N(s) = x\}$, then, for $\alpha \geq 0$,*

$$(3.10) \quad \mathbb{E}_x \left(e^{-\alpha T_y^{X_N}} \right) = \frac{D_x(\alpha)}{D_y(\alpha)} \text{ and } \mathbb{E}_y \left(e^{-\alpha T_x^{X_N}} \right) = \frac{d_N(\alpha)B_y(\alpha) + b_N(\alpha)D_y(\alpha)}{d_N(\alpha)B_x(\alpha) + b_N(\alpha)D_x(\alpha)},$$

with the notations of Propositions 3.3 and 3.4.

PROOF. The first identity comes from the fact that the two processes $(X_N(t))$ and $(E_N(t))$ starting from the same initial state are identical in distribution as long as they do not reach C_N . In particular, if $X_N(0) = E_N(0) = x$, the variables $T_y^{E_N}$ and $T_y^{X_N}$ have the same distribution. The second identity is a direct consequence of the martingale property of $(K_\alpha^N(t \wedge T_0^{X_N}))$ proved in the above proposition. ■

4. A Fluid Picture

This section gives a quick description of the first-order properties of the Ehrenfest and Engset processes as N goes to infinity. Its purpose is mainly to introduce the three natural possible asymptotic regimes that will be investigated in detail in the last sections. The proofs of the asymptotic results are quite standard and therefore will be skipped.

From now on, it is assumed that Relation (2.4) holds, that is

$$\lim_{N \rightarrow +\infty} C_N/N = \eta > 0.$$

The Engset process $(X_N(t))$ can also be seen as the unique solution of the following stochastic differential equation:

$$dX_N(t) = \mathbb{1}_{\{X_N(t-) < C_N\}} \sum_{i=1}^{N-X_N(t-)} \mathcal{N}_{\nu,i}(dt) - \sum_{i=1}^{X_N(t-)} \mathcal{N}_{\mu,i}(dt),$$

starting from $X_N(0)$, where, for $\xi > 0$, $(\mathcal{N}_{\xi,k})$ are independent Poisson processes with rate ξ .

The initial state is assumed to satisfy

$$\lim_{N \rightarrow +\infty} X_N(0)/N = x_0 \in [0, \eta];$$

then, by complementing Poisson processes in order to get martingales, the above equation can be rewritten as

$$dX_N(t) = dM_N(t) + [\nu(N - X_N(t))\mathbb{1}_{\{X_N(t) < C_N\}} - \mu X_N(t)] dt,$$

where $(M_N(t))$ is a martingale of the order of $\sqrt{N}$. In the same way as for the Erlang process, see Chapter 6 of Robert [Rob03] for example, one can prove the following convergence in distribution of processes:

$$\lim_{N \rightarrow +\infty} (X_N(t)/N) = (\min(\eta, \nu + (x_0 - \nu)e^{-t})).$$

This first-order description of the Engset process shows that there are three different asymptotic regimes.

– *Super-Critical Regime:* $\nu > \eta$.

Under this condition the renormalized process is at the boundary C_N at time

$$(4.1) \quad t^* := \log((\nu - x_0)/(\nu - \eta)).$$

A more detailed picture can be obtained by looking at the process

$$(Z_N(t)) = (C_N - X_N(t)/N)$$

of empty spaces with a “slow” time scale. As N goes to infinity, it is easily seen that the Q -matrix of this birth-and-death process converges to the Q -matrix of an ergodic $M/M/1$ process with input rate η and service rate ν . In particular, this gives the asymptotic expression of the Engset formula: for $t \in \mathbb{R}_+$,

$$\lim_{N \rightarrow +\infty} \mathbb{P}(X_N(t) = C_N) = 1 - \eta/\nu.$$

– *Sub-Critical Regime: $\nu < \eta$.*

In this case, one has

$$\lim_{N \rightarrow +\infty} (X_N(t)/N) = (\nu + (x_0 - \nu)e^{-t}) = \lim_{N \rightarrow +\infty} (E_N(t)/N).$$

As expected, the boundary at C_N does not play a role: at first order, the Engset process and the Ehrenfest process are identical.

– *Critical Regime: $\nu = \eta$.*

The fluid limit picture gives that the system saturates “at infinity” which, as we shall see, is a too rough description of its evolution.

The next sections are devoted to the asymptotic analysis of the distributions of hitting times. For simplicity, it is assumed that the initial state is on the boundary, either 0 or C_N . Similar results could be obtained without any additional difficulty when the initial state is in the neighborhood of some $\lfloor zN \rfloor$ for $0 \leq z \leq \eta$.

5. Super-Critical Regime

As we have seen, under the condition $\nu > \eta$ and at time t^* defined by Equation (4.1), the system is saturated in the fluid limit. It implies in particular that the hitting time of the boundary C_N ,

$$T_{C_N}^{X_N} = \inf\{s \geq 0; X_N(s) = C_N\},$$

converges in distribution to t^* . The following proposition gives a more precise asymptotic result. See Theorem 3 of Flajolet and Huillet [FH08] for a related result in the symmetrical case.

PROPOSITION 5.1. *If $C_N = \eta N + O(1)$, $\eta < \nu$ and $X_N(0) = 0$, then the sequence of random variables*

$$\left(\sqrt{N} \left[T_{C_N}^{X_N} - \log(\nu/(\nu - \eta)) \right] \right)$$

converges in distribution to a centered normal random variable with variance

$$\sqrt{\eta(1 - \eta)/(\nu - \eta)}.$$

PROOF. Proposition 3.3 gives the equation

$$(5.1) \quad \mathbb{E}_0 \left(e^{-\alpha \sqrt{N} T_{C_N}^{X_N}} \right) = \int_0^{\sqrt{N}} \left(1 - \frac{u}{\sqrt{N}} \right)^N u^{\sqrt{N}\alpha - 1} du \\ \Bigg/ \int_0^{\sqrt{N}} \left(1 - \frac{u}{\sqrt{N}} \right)^{N - C_N} \left(1 + \frac{\mu}{\nu} \frac{u}{\sqrt{N}} \right)^{C_N} u^{\sqrt{N}\alpha - 1} du,$$

for $\alpha \geq 0$. The integrand of the numerator of the right-hand side of the above equation can be expressed as $\exp(f_N(u))$, with

$$f_N(u) = N \log \left(1 - u/\sqrt{N} \right) + \left(\sqrt{N}\alpha - 1 \right) \log u.$$

The function has a unique maximum at

$$y_N = \frac{\alpha \sqrt{N} - 1}{\sqrt{N} + \alpha - 1/\sqrt{N}} = \alpha - \frac{1 + \alpha^2}{\sqrt{N}} + o\left(1/\sqrt{N}\right),$$

and

$$\begin{cases} f_N(y_N) = (\alpha \log(\alpha) - \alpha)\sqrt{N} - \frac{\alpha^2}{2} - \log(\alpha) + o(1), \\ f_N''(y_N) = -\frac{(1 + \alpha^2)}{\alpha^2} - \frac{\sqrt{N}}{\alpha} + o(1). \end{cases}$$

Laplace's method, see Section 3 of Chapter VIII of Flajolet and Sedgewick [FS09] for example, gives therefore the relation

$$(5.2) \quad \int_0^{\sqrt{N}} e^{f_N(u)} du \sim \frac{\sqrt{2\pi}}{\sqrt{-f_N''(y_N)}} e^{f_N(y_N)} \\ \sim \frac{\sqrt{2\pi\alpha}}{N^{1/4}} \exp\left((\alpha \log(\alpha) - \alpha)\sqrt{N} - \frac{\alpha^2}{2} - \log(\alpha)\right).$$

Similarly, the integrand of the denominator of the right-hand side of Equation (5.1) is $\exp(g_N(u))$, with

$$g_N(u) = (N - C_N) \log\left(1 - \frac{u}{\sqrt{N}}\right) + C_N \log\left(1 + \frac{\mu}{\nu} \frac{u}{\sqrt{N}}\right) + (\sqrt{N}\alpha - 1) \log u.$$

This concave function on the interval $(0, \sqrt{N})$ has a unique maximum at $z_N = z_0 - \delta/\sqrt{N} + o(1/\sqrt{N})$, with

$$z_0 = \alpha \frac{\nu}{\nu - \eta} \text{ and } \delta = \nu \frac{\alpha^2 (\nu^2 + \eta - 2\eta\nu) + (\nu - \eta)^2}{(\nu - \eta)^3},$$

and, with some calculations, one finds

$$g_N(z_N) = \left(-\alpha + \alpha \log(\alpha) + \alpha \log\left(\frac{\nu}{\nu - \eta}\right)\right) \sqrt{N} \\ + \frac{(2\nu\eta - \nu^2 - \eta)}{(\nu - \eta)^2} \frac{\alpha^2}{2} - \log(\alpha) - \log\left(\frac{\nu}{\nu - \eta}\right) + o(1)$$

and

$$g_N''(z_N) = -\left(1 - \eta + \eta \frac{(1 - \nu)^2}{\nu^2}\right) - \frac{\alpha\sqrt{N} - 1}{z_0^2} + o(1).$$

Again by Laplace's method, this gives

$$(5.3) \quad \exp\left(-\alpha\sqrt{N} \log \frac{\nu}{\nu - \eta}\right) \int_0^{\sqrt{N}} e^{g_N(u)} du \\ \sim \frac{\sqrt{2\pi\alpha}}{N^{1/4}} \exp\left((\alpha \log(\alpha) - \alpha)\sqrt{N} + \frac{(2\nu\eta - \nu^2 - \eta)}{(\nu - \eta)^2} \frac{\alpha^2}{2} - \log(\alpha)\right).$$

Equation (5.1) together with (5.2) and (5.3) give finally

$$\lim_{N \rightarrow +\infty} \mathbb{E} \left(\exp \left(-\alpha\sqrt{N} \left[T_{C_N}^{X_N} - \log(\nu/(\nu - \eta)) \right] \right) \right) = \exp \left(\frac{\eta(1 - \eta)}{(\nu - \eta)^2} \frac{\alpha^2}{2} \right).$$

■

An Informal Proof. The limit theorem obtained in Proposition 5.1 is a consequence of some detailed, annoying, but simple, calculations used to apply Laplace method. One can get quite quickly an idea of the possible limit with the help of the exponential martingale $(M_N^\beta(t))$ of Proposition 3.1 through a non-rigorous derivation. As will be seen, it gives the correct answer but its justification seems to be difficult. The main problem comes from the fact that, in this martingale, the term e^t stopped at some random time may not be integrable. For example, it is easily seen that the first jump of the martingale $(M_1^\beta(t))$ is not a regular stopping time for this martingale, i.e. the optional stopping theorem is not valid here.

Denote $Z_N = \sqrt{N} \left(\exp(T_{C_N}^{X_N}) - \exp(t^*) \right)$, where t^* is, as before, $\log(\nu/(\nu-\eta))$. By using the martingale (3.2) of Proposition 3.1 by assuming that the stopping time $T_{C_N}^{X_N}$ is regular for it, one gets

$$\mathbb{E} \left[\left(1 - \mu e^{T_{C_N}^{X_N}} \beta / \sqrt{N} \right)^{C_N} \left(1 + \nu e^{T_{C_N}^{X_N}} \beta / \sqrt{N} \right)^{N-C_N} \right] = \left(1 + \nu \beta / \sqrt{N} \right)^N.$$

This can be written as $\mathbb{E}(\exp(U_N)) = 1$, with

$$\begin{aligned} U_N &= C_N \log \left(1 - \mu e^{T_{C_N}^{X_N}} \beta / \sqrt{N} \right) \\ &\quad + (N - C_N) \log \left(1 + \nu e^{T_{C_N}^{X_N}} \beta / \sqrt{N} \right) - N \log \left(1 + \nu \beta / \sqrt{N} \right), \end{aligned}$$

hence,

$$\begin{aligned} U_N &= -\beta(\eta - \nu)Z_N - \left((\eta\mu^2 + (1-\eta)\nu^2) e^{2t^*} - \nu^2 \right) \beta^2/2 + o(1/N) \\ &= -\beta(\eta - \nu)Z_N - \frac{\nu^2\eta(1-\eta)}{(\nu-\eta)^2} \beta^2/2 + o(1/N), \end{aligned}$$

provided that the limit can be taken under the integral. One gets finally

$$\lim_{N \rightarrow +\infty} \mathbb{E}(e^{-\beta Z_N}) = \exp \left(\frac{\beta^2 \nu^2 \eta (1-\eta)}{2 (\nu-\eta)^4} \right).$$

Expressed as a limit theorem for $T_{C_N}^{X_N}$, this is precisely the above proposition.

6. Sub-Critical Regime

It is assumed in this section that $\nu < \eta$, so that the Ehrenfest process “lives” in the interior of the state space; the hitting time of the boundary C_N should be therefore quite large. The following propositions give asymptotic results concerning this phenomenon.

The first result concerns the time it takes from the Ehrenfest process to have all particles in one box when, initially, they are all in the other box. This is of course a very natural quantity for this process. In the discrete-time case, representations of the average of this quantity have been obtained in a symmetrical setting. See Bingham [Bin91] and Palacios [Pal93] and the references therein.

PROPOSITION 6.1. *If $\nu < 1$, $X_N(0) = 0$ and $C_N = N$, then the sequence of random variables $(N\nu^N T_{C_N}^{X_N})$ converges in distribution to an exponentially distributed random variable with parameter $1 - \nu$.*

PROOF. One uses Equation (3.7) of Proposition 3.3 to get that, for $N \geq 1$,

$$(6.1) \quad \mathbb{E} \left(e^{-\alpha_N T_N^{X_N}} \right) = \int_0^1 (1-u)^N u^{\alpha_N-1} du \Big/ \int_0^1 \left(1 + \frac{\mu}{\nu} u \right)^N u^{\alpha_N-1} du,$$

with $\alpha_N = \alpha N \nu^N$, for some $\alpha > 0$.

The numerator of this expression can be written, after an integration by parts, as

$$\begin{aligned} \int_0^1 (1-u)^N u^{\alpha_N-1} du &= \int_0^1 N(1-u)^{N-1} \frac{u^{\alpha_N}}{\alpha_N} du \\ &= N^{-\alpha_N} \int_0^N \left(1 - \frac{u}{N} \right)^{N-1} \frac{u^{\alpha_N}}{\alpha_N} du \sim \frac{1}{\alpha_N}, \end{aligned}$$

by dominated convergence.

By subtracting $1/\alpha_N$ from the denominator of the right-hand side of Equation (6.1), one gets

$$\begin{aligned}\Delta_N &:= \int_0^1 \left(\left(1 + \frac{\mu}{\nu} u\right)^N - 1 \right) u^{\alpha_N-1} du = \frac{N\mu}{\nu} \int_0^1 \left(1 + \frac{\mu}{\nu} u\right)^{N-1} \frac{1 - u^{\alpha_N}}{\alpha_N} du \\ &= \frac{1}{\nu^N} \int_0^N \mu \left(1 - \mu \frac{u}{N}\right)^{N-1} \frac{1 - (1 - u/N)^{\alpha_N}}{\alpha_N} du,\end{aligned}$$

hence,

$$\begin{aligned}\alpha_N \Delta_N &= \frac{\alpha_N}{N\nu^N} \int_0^N \mu \left(1 - \mu \frac{u}{N}\right)^{N-1} \frac{1 - (1 - u/N)^{\alpha_N}}{\alpha_N/N} du \\ &\sim \alpha \int_0^{+\infty} \mu e^{-\mu u} u du = \frac{\alpha}{\mu}.\end{aligned}$$

These two asymptotic results plugged into Equation (6.1) give the desired convergence in distribution. ■

Theorem 2 of Bingham [Bin91] provides a similar result in the symmetrical case $\mu = \nu$ and in discrete time. In the present case, there is an additional factor N in the scaling of $T_{C_N}^{X_N}$ which is due to the fact that the continuous-time dynamics are N times faster than the discrete-time case.

PROPOSITION 6.2. *If $C_N = \eta N + O(1)$, $\nu < \eta < 1$ and $X_N(0) = 0$, then if*

$$H = (1 - \eta) \log \left(\frac{1 - \eta}{1 - \nu} \right) + \eta \log \left(\frac{\eta}{\nu} \right),$$

the sequence of random variables

$$(6.2) \quad \left(\frac{\sqrt{\eta(1 - \eta)}}{(\eta - \nu)\sqrt{2\pi}} \sqrt{N} e^{-NH} T_{C_N}^{X_N} \right),$$

converges in distribution to an exponentially distributed random variable with parameter 1.

One remarks that the exponential decay factor H of the above proposition is in fact a relative entropy of Bernoulli random variables with respective parameters η and ν . Despite the fact that similar “entropy” expressions appeared on several occasions in the study of these processes, we have not been able to find a simple explanation for the occurrences of these constants.

PROOF. For $\alpha > 0$, denote by α_N the product of α and the coefficient of $T_{C_N}^{X_N}$ in (6.2). Equation (3.7) of Proposition 3.3 is again used

$$(6.3) \quad \mathbb{E} \left(e^{-\alpha_N T_N^{X_N}} \right) = \int_0^1 (1-u)^N u^{\alpha_N-1} du \Big/ \int_0^1 (1-u)^{N-C_N} \left(1 + \frac{\mu}{\nu} u\right)^{C_N} u^{\alpha_N-1} du.$$

The asymptotic behavior of the numerator of this Laplace transform has already been obtained in the proof of the above proposition.

To study the denominator, we proceed as before. For $u \in [0, 1]$, write

$$f_N(u) = (N - C_N) \log(1 - u) + C_N \log \left(1 + \mu u / \nu \right).$$

This function has a unique maximum at

$$y_0 := \frac{C_N/N - \nu}{1 - \nu} = \frac{(\eta - \nu)}{(1 - \nu)} + O(1/N),$$

given by

$$f_N(y_0) = \left[(1-\eta) \log \left(\frac{1-\eta}{1-\nu} \right) + \eta \log \left(\frac{\eta}{\nu} \right) \right] N + o(1),$$

and

$$f_N''(y_0) = -\frac{(1-\nu)^2}{\eta(1-\eta)} N + o(1).$$

The denominator of (6.3) is

$$\begin{aligned} \int_0^1 \left[(1-u)^{N-C_N} \left(1 + \frac{\mu}{\nu} u \right)^{C_N} - 1 \right] u^{\alpha_N-1} du &= \int_0^1 \left[e^{f_N(u)} - 1 \right] u^{\alpha_N-1} du \\ &= \int_0^1 f_N'(u) e^{f_N(u)} \frac{(y_0^{\alpha_N} - u^{\alpha_N})}{\alpha_N} du + \frac{1 - y_0^{\alpha_N}}{\alpha_N}, \end{aligned}$$

by integration by parts. The integral I_N in the right-hand side can be written as

$$I_N = \int_{-y_0\sqrt{N}}^{(1-y_0)\sqrt{N}} \frac{1}{\sqrt{N}} f_N' \left(y_0 + \frac{u}{\sqrt{N}} \right) e^{f_N(y_0 + u/\sqrt{N})} \frac{(y_0^{\alpha_N} - (y_0 + u/\sqrt{N})^{\alpha_N})}{\alpha_N} du,$$

hence,

$$\begin{aligned} I_N &= y_0^{\alpha_N-1} \frac{-f_N''(y_0)}{N^{3/2}} e^{f_N(y_0)} \int_{-\infty}^{+\infty} u^2 \exp \left(\frac{f_N''(y_0)}{N} \frac{u^2}{2} \right) du + o(1/N) \\ &= \frac{1}{y_0} \sqrt{\frac{2\pi}{-f_N''(0)}} e^{f_N(y_0)} + o(1/N). \end{aligned}$$

Combining,

$$\lim_{N \rightarrow +\infty} \mathbb{E} \left(\exp \left(-\alpha_N T_N^{X_N} \right) \right) = 1/(1+\alpha).$$

■

PROPOSITION 6.3 (Hitting time of the empty state). *Under the condition $\nu < \eta$ and if $C_N = \eta N + O(1)$ for $\eta > 0$ and $X_N(0) = C_N$, then the sequence of variables*

$$\left(N(1-\nu)^N T_0^{X_N} \right)$$

converges in distribution to an exponential random variable with parameter ν .

Note that this result can, informally, be justified by the result of Proposition 6.1. Without the boundary C_N , one could obtain the result by exchanging μ and ν and using Proposition 6.1. This result shows in particular that the boundary does not change the limiting behavior of $T_0^{X_N}$ in the sub-critical regime.

PROOF. Denote $\alpha_N = N(1-\nu)^N$. Proposition 3.5 gives

$$(6.4) \quad \mathbb{E}_{C_N} \left(e^{-\alpha_N T_0^{X_N}} \right) = \frac{d_N(\alpha_N) B_{C_N}(\alpha_N) + b_N(\alpha_N) D_{C_N}(\alpha_N)}{d_N(\alpha_N) B_0(\alpha_N) + b_N(\alpha_N) D_0(\alpha_N)}.$$

One starts with the asymptotic behavior of $(d_N(\alpha_N))$:

$$\begin{aligned} d_N(\alpha_N) &= \mu \int_0^1 (1-u)^{N-C_N-1} \left(1 + \frac{\mu}{\nu} u \right)^{C_N} u^{\alpha_N} du \\ &= \frac{\mu}{\sqrt{N}} \int_0^{\sqrt{N}} \left(1 - \frac{u}{\sqrt{N}} \right)^{N-C_N-1} \left(1 + \frac{\mu}{\nu} \frac{u}{\sqrt{N}} \right)^{C_N} du + o(1/\sqrt{N}) \\ &= \frac{1}{\sqrt{N}} \exp \left(\frac{\eta-\nu}{\nu} \sqrt{N} \right) \int_0^{+\infty} \exp \left(- \left(\frac{\eta}{\nu^2} + \frac{(1-\eta)}{(1-\nu)^2} \right) \frac{u^2}{2} \right) du + o(1/\sqrt{N}). \end{aligned}$$

The other coefficient $b_N(\alpha_N)$ is such that

$$\begin{aligned} b_N(\alpha_N) &= \nu \int_0^1 (1-u)^{C_N} \left(1 + \frac{\nu}{\mu}u\right)^{N-C_N} u^{\alpha_N} du \\ &= \frac{\nu}{N} \int_0^N \left(1 - \frac{u}{N}\right)^{C_N} \left(1 + \frac{\nu}{\mu N}u\right)^{N-C_N-1} du + o(1/N) \\ &= \frac{\nu(1-\nu)}{\nu-\eta} \frac{1}{N} + o(1/N). \end{aligned}$$

The proof of Proposition 6.1 provides the relations

$$D_0(\alpha_N) = \int_0^1 (1-u)^N u^{\alpha_N-1} du \sim \frac{1}{\alpha_N}$$

and

$$B_0(\alpha_N) = \int_0^1 \left(1 + \frac{\nu}{\mu}u\right)^N u^{\alpha_N-1} du \sim \frac{1}{\alpha_N} + \frac{1}{\nu N(1-\nu)^N}.$$

The two remaining terms to estimate are

$$\begin{aligned} B_{C_N}(\alpha_N) &= \int_0^1 (1-u)^{C_N} \left(1 + \frac{\nu}{\mu}u\right)^{N-C_N} u^{\alpha_N-1} du, \\ D_{C_N}(\alpha_N) &= \int_0^1 (1-u)^{N-C_N} \left(1 + \frac{\mu}{\nu}u\right)^{C_N} u^{\alpha_N-1} du. \end{aligned}$$

As in the proof of Proposition 6.2, one can show that $B_{C_N}(\alpha_N)$ and $D_{C_N}(\alpha_N)$ can be written as $1/\alpha_N + o(B_0(\alpha_N) - 1/\alpha_N)$. More informally, the term $(1-u)^N$ under the integral for these two expressions reduces their asymptotic behavior by an exponential factor.

These various estimates give finally that

$$\lim_{N \rightarrow +\infty} \mathbb{E} \left(e^{-\alpha_N T_0^{X_N}} \right) = \lim_{N \rightarrow +\infty} \frac{B_{C_N}(\alpha_N)}{B_0(\alpha_N)} = \frac{1}{1 + \alpha/\nu}.$$

■

7. Critical Regime

In this section, it is assumed that $C_N \sim \nu N$, if $X_N(0) = 0$, the fluid limit of the process is given by $\nu(1 - \exp(-t))$, the fluid boundary ν is reached at time $t = +\infty$. In fact, with a second-order description, the process $X_N(t)$ can be written as $X_N(t) \sim \nu(1 - \exp(-t))N + Y(t)\sqrt{N}$ for some ergodic diffusion process $(Y(t))$, so the hitting time $T_{C_N}^{X_N}$ of the boundary is such that

$$\exp \left(-T_{C_N}^{X_N} \right) \sim \frac{Y(T_{C_N}^{X_N})}{\nu\sqrt{N}},$$

which gives a rough estimate $T_{C_N}^{X_N} \sim \log(\sqrt{N})$. The following proposition shows that this approximation is fact quite precise. See also Theorem 4 of Flajolet and Huillet [FH08].

PROPOSITION 7.1. *If $C_N = \nu N + \delta\sqrt{N} + o(\sqrt{N})$ with $\nu < 1$, $\delta \in \mathbb{R}$, and $X_N(0) = 0$, then the sequence of random variables*

$$\left(T_{C_N}^{X_N} - \log(N)/2 \right)$$

converges in distribution to a random variable Z on $\mathbb{R}$ whose Laplace transform at $\alpha > 0$ is given by:

$$(7.1) \quad \mathbb{E} \left(e^{-\alpha Z} \right) = \Gamma(\alpha) \left/ \int_0^{+\infty} \exp \left(u \frac{\delta}{\nu} - \frac{u^2(1-\nu)}{2\nu} \right) u^{\alpha-1} du \right.$$

If $\delta = 0$, then the variable $Z - \log(\nu/(1-\nu))/2$ has the following density on $\mathbb{R}$:

$$x \mapsto \sqrt{2/\pi} \exp(-x - e^{-2x}/2).$$

Note that the Laplace transform of the limit in distribution is the ratio of the Mellin transforms of the functions

$$u \mapsto \exp(-u) \quad \text{and} \quad u \mapsto \exp\left(-u \frac{\delta}{\nu} - \frac{u^2(1-\nu)}{2\nu}\right).$$

See Section B.7 of Flajolet and Sedgewick [FS09] on Mellin transforms.

PROOF. Proposition 3.3 gives the equation

$$\mathbb{E}_0 \left(e^{-\alpha T_{C_N}^{x_N}} \right) = \int_0^1 (1-u)^N u^{\alpha-1} du \Big/ \int_0^1 (1-u)^{N-C_N} \left(1 + \frac{\mu}{\nu} u\right)^{C_N} u^{\alpha-1} du,$$

for $\alpha > 0$. The asymptotic behavior of the numerator is easy since

$$\int_0^1 (1-u)^N u^{\alpha-1} du = \frac{1}{N^\alpha} \int_0^N \left(1 - \frac{u}{N}\right)^N u^{\alpha-1} du \sim \frac{\Gamma(\alpha)}{N^\alpha}.$$

The denominator can be expressed as

$$\frac{1}{N^{\alpha/2}} \int_0^{\sqrt{N}} e^{f_N(u)} u^{\alpha-1} du,$$

with

$$\begin{aligned} f_N(u) &= (N - C_N) \log \left(1 - \frac{u}{\sqrt{N}}\right) + C_N \log \left(1 + \frac{\mu}{\nu} \frac{u}{\sqrt{N}}\right) \\ &= -\frac{(1-\nu)}{\nu} \frac{u^2}{2} + \frac{\delta}{\nu} u + o(1/\sqrt{N}). \end{aligned}$$

By dominated convergence, one gets therefore that, for $\alpha > 0$,

$$\lim_{N \rightarrow +\infty} \mathbb{E}_0 \left(e^{-\alpha [T_{C_N}^{x_N} - \log N/2]} \right) = \Gamma(\alpha) \Big/ \int_0^{+\infty} \exp \left(-\frac{(1-\nu)}{\nu} \frac{u^2}{2} + \frac{\delta}{\nu} u \right) u^{\alpha-1} du,$$

and hence the first part of the proposition.

For $\delta = 0$, a change of variable gives

$$\int_0^{+\infty} \exp \left(-\frac{u^2(1-\nu)}{2\nu} \right) u^{\alpha-1} du = \frac{1}{2} \left(\frac{2\nu}{1-\nu} \right)^{\alpha/2} \Gamma(\alpha/2).$$

The Laplace transform of Z can therefore be expressed as

$$\mathbb{E} (e^{-\alpha Z}) = 2 \left(\frac{1-\nu}{2\nu} \right)^{\alpha/2} \frac{\Gamma(\alpha)}{\Gamma(\alpha/2)} = \left(\frac{1-\nu}{\nu} \right)^{\alpha/2} \frac{2^{\alpha/2}}{\sqrt{\pi}} \Gamma((\alpha+1)/2),$$

by using Legendre's duplication Formula for Gamma functions (see e.g. Whittaker and Watson [WW96] page 240). Since

$$\begin{aligned} \frac{2^{\alpha/2}}{\sqrt{\pi}} \Gamma \left(\frac{\alpha+1}{2} \right) &= \frac{1}{\sqrt{\pi}} \int_0^{+\infty} \exp(\alpha \log(2u)/2 - \log(u)/2 - u) du \\ &= \sqrt{\frac{2}{\pi}} \int_{-\infty}^{+\infty} e^{-\alpha u} \exp(-u - e^{-2u}/2) du, \end{aligned}$$

with a change of variables, one gets the desired result on the distribution of Z . ■

We conclude with the hitting time of empty state. We remark that, at the correct time scale, the time is half of the corresponding variable in the sub-critical case. See Proposition 6.3. A simple, naive, explanation is as follows. For sub-critical regime the process lives in a region centered at νN and whose width is of the order of $\sqrt{N}$ and therefore makes many excursions in this region before reaching C_N . In

the critical case the process lives near, but only on one side of νN . In particular it cannot go above C_N and, consequently, does not waste time on such excursions.

PROPOSITION 7.2 (Hitting Time of Empty State). *If $C_N = \nu N + o(\sqrt{N})$ with $\nu > 0$ and $X_N(0) = C_N$, then the sequence of random variables*

$$\left(N(1 - \nu)^N T_0^{X_N} \right)$$

converges in distribution to an exponential random variable with parameter 2ν .

PROOF. Denote $\alpha_N = N(1 - \nu)^N$. Recall that

$$\mathbb{E}_{C_N} \left(e^{-\alpha_N T_0^{X_N}} \right) = \frac{d_N(\alpha) B_{C_N}(\alpha) + b_N(\alpha) D_{C_N}(\alpha)}{d_N(\alpha) B_0(\alpha) + b_N(\alpha) D_0(\alpha)}.$$

One starts with the asymptotic behavior of $(d_N(\alpha_N))$. By definition,

$$\begin{aligned} d_N(\alpha_N) &= \mu \int_0^1 (1 - u)^{N - C_N - 1} \left(1 + \frac{\mu}{\nu} u \right)^{C_N} u^{\alpha_N} du \\ &= \mu \frac{1}{\sqrt{N}} \int_0^{\sqrt{N}} \left(1 - \frac{u}{\sqrt{N}} \right)^{N - C_N - 1} \left(1 + \frac{\mu}{\nu} \frac{u}{\sqrt{N}} \right)^{C_N} du + o\left(1/\sqrt{N}\right). \end{aligned}$$

Since $\eta = \nu$ and $\mu = 1 - \nu$, $\nu(1 - \eta) = \mu\eta$, so

$$\begin{aligned} d_N(\alpha_N) &= (1 - \nu) \frac{1}{\sqrt{N}} \int_0^{+\infty} \exp\left(-\frac{(1 - \nu) u^2}{\nu} \frac{1}{2}\right) du + o\left(1/\sqrt{N}\right) \\ &= \frac{1}{\sqrt{N}} \sqrt{\frac{\pi}{2}} \sqrt{\nu(1 - \nu)} + o\left(1/\sqrt{N}\right). \end{aligned}$$

Note that, up to a term -1 in an exponent which does not play a role in the limiting behavior, the quantity $b_N(\alpha)$ is almost $d_N(\alpha)$ with ν replaced by $(1 - \nu)$. Consequently $b_N(\alpha)$ has the same asymptotic expansion as $d_N(\alpha)$.

The asymptotic behaviors of the quantities $B_0(\alpha_N)$, $D_0(\alpha_N)$, $B_{C_N}(\alpha_N)$ and $D_{C_N}(\alpha_N)$ are the same as those obtained in the proof of Proposition 6.3. By gathering these various estimates one finds

$$\lim_{N \rightarrow +\infty} \mathbb{E} \left(\exp \left(-\alpha_N T_0^{X_N} \right) \right) = 2/(2 + \alpha/\nu).$$

■

References for Chapter V

- [Bin91] N. H. Bingham, *Fluctuation theory for the Ehrenfest urn*, Advances in Applied Probability **23** (1991), no. 3, 598–611.
- [DC98] Antonio Di Crescenzo, *First-passage-time densities and avoiding probabilities for birth-and-death processes with symmetric sample paths*, J. Appl. Probab. **35** (1998), no. 2, 383–394.
- [DGM90] Persi Diaconis, R. L. Graham, and J. A. Morrison, *Asymptotic analysis of a random walk on a hypercube with many dimensions*, Random Structures Algorithms **1** (1990), no. 1, 51–72.
- [Eng98] T.O. Engset, *On the calculation of switches in an automatic telephone system*, Teletronikk **94** (1998), no. 2, 99–142, Translation of a German version (1918).
- [Fei78] Philip J. Feinsilver, *Special functions, probability semigroups, and Hamiltonian flows*, Lecture Notes in Mathematics, vol. 696, Springer, Berlin, 1978.

- [FGP08] Mark Flegg, Dmitri Gramotnev, and P Pollett, *Ehrenfest model for condensation and evaporation processes in degrading aggregates with multiple bonds*, Physical Review E **78** (2008), 031117–1–031117–9.
- [FH08] Philippe Flajolet and Thierry Huillet, *Analytic combinatorics of the Mabinogion urn*, Proceedings of Fifth Colloquium on Mathematics and Computer Science (Blaubeuren) (U. Rösler, ed.), September 2008, Discrete Mathematics and Theoretical Computer Science (DMTCS).
- [FJP11] Mathieu Feuillet, Matthieu Jonckheere, and Balakrishna J. Prabhu, *Bandwidth sharing networks with priority scaling*, submitted (2011).
- [FRT99] Christine Fricker, Philippe Robert, and Danielle Tibi, *On the rates of convergence of Erlang’s model.*, Journal of Applied Probability **36** (1999), no. 4, 1167–1184.
- [FS09] Philippe Flajolet and Robert Sedgewick, *Analytic combinatorics*, Cambridge University Press, 2009.
- [Ken76] D.P. Kennedy, *Some martingales related to cumulative sum tests and single-server queues*, Stochastic Processes and their Applications **4** (1976), 261–269.
- [KM57] Samuel Karlin and James McGregor, *The classification of birth and death processes*, Transactions of the AMS **86** (1957), 366–400.
- [KM59] ———, *Coincidence properties of birth and death processes*, Pacific Journal of Mathematics **9** (1959), 1109–1140.
- [KM65] ———, *Ehrenfest urn models*, Journal of Applied Probability **2** (1965), 352–376.
- [LS63] J. Lamperti and J. L. Snell, *Martin boundaries for certain Markov chains*, Journal of the Mathematical Society of Japan **15** (1963), 113–128.
- [Pal93] José Luis Palacios, *Fluctuation theory for the Ehrenfest urn via electric networks*, Advances in Applied Probability **25** (1993), no. 2, 472–476.
- [Rob03] Philippe Robert, *Stochastic networks and queues*, Stochastic Modeling and Applied Probability Series, Springer-Verlag, New York, 2003, xvii+398 pp.
- [RW87] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes & martingales vol. 2: Itô calculus*, Cambridge University Press, 2000 (1987).
- [RY99] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999.
- [Sch00] Wim Schoutens, *Stochastic processes and orthogonal polynomials*, Lecture Notes in Statistics, vol. 146, Springer-Verlag, New York, 2000.
- [ST98] Wim Schoutens and Jozef L. Teugels, *Lévy processes, polynomials and martingales*, Comm. Statist. Stochastic Models **14** (1998), no. 1-2, 335–349, Special issue in honor of Marcel F. Neuts.
- [ST10] Florian Simatos and Danielle Tibi, *Spatial homogenization in a stochastic network with mobility*, The Annals of Applied Probability **20** (2010), no. 1, 312–355.
- [Sze75] Gábor Szegő, *Orthogonal polynomials*, fourth ed., American Mathematical Society, Providence, R.I., 1975, American Mathematical Society, Colloquium Publications, Vol. XXIII.

- [WW96] E. T. Whittaker and G. N. Watson, *A course of modern analysis*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1996, Reprint of the fourth (1927) edition.

Comprehensive list of references

- [AGC08] Jesús R. Artalejo and Antonio Gómez-Corral. *Retrial queueing systems*. Springer-Verlag, Berlin, 2008. A computational approach.
- [Alo86] Noga Alon. A fast and simple randomized parallel algorithm for the maximal independent set problem. *Journal of Algorithms*, 7:567–583, 1986.
- [Amb70] Norman Abramson. The aloha system. In *FJCC AFIPS*, volume 37, pages 281–285, 1970.
- [AO76] Robert F. Anderson and Steven Orey. Small random perturbation of dynamical systems with reflecting boundary. *Nagoya Math. J.*, 60:189–216, 1976.
- [BA03] T. Bonald and Proutière A. Insensitive bandwidth sharing in data networks. *Queueing Systems: Theory and Applications*, 44(1):69–100, 2003.
- [Bar62] Paul Baran. On distributed communications networks. Technical Report P-2626, Rand Corporation, 1962.
- [BBHP04] T. Bonald, S.C. Borst, N. Hegde, and A. Proutière. Wireless data performance in multi-cell scenarios. In *Proceedings of ACM Sigmetrics*, pages 378–380, 2004.
- [BBLP11] N. Bouman, S. Borst, J. Van Leeuwen, and A. Proutière. Backlog-based random access in wireless networks: fluid limits and delay issues. In *ITC 23*, 2011.
- [BBP⁺01] S. Ben Fredj, T. Bonald, A. Proutière, G. Régnié, and J. W. Roberts. Statistical bandwidth sharing: a study of congestion at flow level. In *Proceedings of ACM SIGCOMM*, pages 111–122, 2001.
- [BBP04] T. Bonald, S.C. Borst, and A. Proutière. How mobility impacts the flow-level performance of wireless data systems. In *Proceedings of INFOCOM*, volume 3, pages 1872–1881, 2004.
- [BF10] T. Bonald and M. Feuillet. On the stability of flow-aware CSMA. *Perform. Eval.*, 67:1219–1229, November 2010.
- [BF11] T. Bonald and M. Feuillet. On flow-aware CSMA in multi-channel wireless networks. In *CISS*, 2011.
- [BF12] T. Bonald and M. Feuillet. Performance of CSMA in multi-channel wireless networks. *Queueing Systems: Theory & Applications*, 2012. To appear.
- [BFP09] T. Bonald, M. Feuillet, and A. Proutière. Is the “Law of the Jungle” sustainable for the Internet? In *Proceedings of INFOCOM*, pages 28–36, april 2009.
- [Bil99] Patrick Billingsley. *Convergence of Probability Measures (second edition)*. Wiley Series in Probability and Statistics. Wiley-Interscience, 1999.
- [Bin91] N. H. Bingham. Fluctuation theory for the Ehrenfest urn. *Advances in Applied Probability*, 23(3):598–611, 1991.
- [BJL08] Sem Borst, Matthieu Jonckheere, and Lasse Leskelä. Stability of parallel queueing systems with coupled service rates. *Discrete Event Dynamic Systems*, 18(4):447–472, dec 2008.
- [BK00] A. W. Berger and Y. Kogan. Dimensioning bandwidth for elastic traffic in high-speed data networks. *IEEE/ACM Trans. Netw.*, 8(5):643–654, 2000.
- [BM01] T. Bonald and L. Massoulié. Impact of fairness on Internet performance. In *Proceedings of ACM SIGMETRICS/Performance*, pages 82–91, 2001.
- [BMP08] Charles Bordenave, David McDonald, and Alexandre Proutière. Asymptotic stability region of slotted-Aloha. Preprint, 2008.
- [BMPV06] T. Bonald, L. Massoulié, A. Proutière, and J. Virtamo. A queueing analysis of max-min fairness, proportional fairness, and balanced fairness. *Queueing Systems: Theory and Applications*, 53:65–84, 2006.

- [Bog61] N. N. Bogoliubov. The quasi-averages in problems of statistical mechanics (in russian). Technical report, Joint Institute for Nuclear Research, 1961. Preprint D-781.
- [Bon07] T. Bonald. Insensitive traffic models for communication networks. *Discrete Event Dynamic Systems*, 17(3):405–421, 2007.
- [Bra08] Maury Bramson. *Stability of Queueing Networks*, volume 1950 of *Lecture Notes in Mathematics*. Springer, 2008.
- [Bré01] Pierre Brémaud. *Markov chains, Gibbs fields, Monte Carlo simulation, and queues*. Texts in Applied Mathematics. Springer-Verlag, 2001.
- [Bro78] Tim Brown. A martingale approach to the Poisson convergence of simple point processes. *The Annals of Probability*, 6(4):615–628, 1978.
- [BTI⁺03] C. Barakat, P. Thiran, G. Iannaccone, C. Diot, and P. Owezarski. Modeling internet backbone traffic at the flow level. *IEEE Transactions on Signal processing*, 51:2003, 2003.
- [Bur81] David Y. Burman. Insensitivity in queueing systems. *Advances in Applied Probability*, 13(4):846–859, 1981.
- [CDH⁺06] Byung-Gon Chun, Frank Dabek, Andreas Haeberlen, Emil Sit, Hakim Weatherspoon, M. Frans Kaashoek, John Kubiatowicz, and Robert Morris. Efficient replica maintenance for distributed storage systems. In *NSDI*, pages 4–4, 2006.
- [CK74] Vinton G. Cerf and Robert E. Kahn. A protocol for packet network intercommunication. *IEEE Transactions on Communications*, 22(5):637–648, 1974.
- [CKLS08] P. Chaporkar, K. Kar, Xiang Luo, and S. Sarkar. Throughput and fairness guarantees through maximal scheduling in wireless networks. *IEEE Transactions on Information Theory*, 54(2):572–594, feb 2008.
- [CM91] H. Chen and A. Mandelbaum. Discrete flow networks: bottleneck analysis and fluid approximations. *Mathematics of Operation Research*, 16(2):408–446, 1991.
- [Dai95] J. G. Dai. On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Annals of Applied Probability*, 5:49–77, 1995.
- [Daw93] Donald A. Dawson. Measure-valued Markov processes. In *École d'Été de Probabilités de Saint-Flour XXI—1991*, volume 1541 of *Lecture Notes in Math.*, pages 1–260. Springer, Berlin, 1993.
- [DC98] Antonio Di Crescenzo. First-passage-time densities and avoiding probabilities for birth-and-death processes with symmetric sample paths. *J. Appl. Probab.*, 35(2):383–394, 1998.
- [DGM90] Persi Diaconis, R. L. Graham, and J. A. Morrison. Asymptotic analysis of a random walk on a hypercube with many dimensions. *Random Structures Algorithms*, 1(1):51–72, 1990.
- [DGR02] Vincent Dumas, Fabrice Guillemin, and Philippe Robert. A Markovian analysis of Additive-Increase Multiplicative-Decrease (AIMD) algorithms. *Advances in Applied Probability*, 34(1):85–111, 2002.
- [DS65] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *Journal of Applied Probability*, 2(1):88–100, 1965.
- [dVKL01] Gustavo de Veciana, Takis Konstantopoulos, and Tae-Jin Lee. Stability and performance analysis of networks supporting elastic services. *IEEE/ACM Trans. Netw.*, 9(1):2–14, 2001.
- [DW06] A. Dimakis and J. Walrand. Sufficient conditions for stability of longest-queue-first scheduling: second-order properties using fluid limits. *Advances in Applied Probability*, 38(2):505–521, 2006.
- [EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and convergence*. Wiley, 1986.
- [EKCM78] N. El Karoui and M. Chaleyat-Maurel. *Temps locaux*, volume 52-53, chapter Un problème de réflexion et ses applications au temps local et aux équations différentielles stochastiques sur $\mathbb{R}$, pages 117–144. Société Mathématique de France, 1978. Exposés du Séminaire J. Azéma-M. Yor, Paris, 1976–1977.
- [Eng98] T.O. Engset. On the calculation of switches in an automatic telephone system. *Teletronikk*, 94(2):99–142, 1998. Translation of a German version (1918).
- [Erl09] A.K. Erlang. The theory of probabilities and telephone conversations. *Nyt Tidsskrift for Matematik B*, 20:33, 1909.

- [Erl17] A.K. Erlang. Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges. *Elektroteknikerer*, 13:5, 1917.
- [Fei78] Philip J. Feinsilver. *Special functions, probability semigroups, and Hamiltonian flows*, volume 696 of *Lecture Notes in Mathematics*. Springer, Berlin, 1978.
- [Feu12] Mathieu Feuillet. On the flow-level stability of data networks without congestion control: the case of linear networks and upstream trees. *Queueing Systems, Theory and Applications*, 70:105–143, 2012.
- [FGP08] Mark Flegg, Dmitri Gramotnev, and P. Pollett. Ehrenfest model for condensation and evaporation processes in degrading aggregates with multiple bonds. *Physical Review E*, 78:031117–1–031117–9, 2008.
- [FH08] Philippe Flajolet and Thierry Huillet. Analytic combinatorics of the Mabinogion urn. In U. Rösler, editor, *Proceedings of Fifth Colloquium on Mathematics and Computer Science*, Blaubeuren, September 2008. Discrete Mathematics and Theoretical Computer Science (DMTCS).
- [FI79] Guy Fayolle and Roudolf Iasnogorodski. Two coupled processors: The reduction to a riemann-hilbert problem. *Probability Theory and Related Fields*, 47(3):325–351, 1979.
- [FKMP95] P. A. Ferrari, H. Kesten, S. Martinez, and P. Picco. Existence of quasi-stationary distributions. A renewal dynamical approach. *The Annals of Probability*, 23(2):501–521, 1995.
- [Flo03] S. Floyd. Highspeed TCP for large congestion windows. RFC 896, 2003.
- [FPR10] M. Feuillet, A. Proutière, and P. Robert. Random capture algorithms: Fluid limits and stability. In *Information Theory and Applications Workshop*, February 2010.
- [FR12a] Mathieu Feuillet and Philippe Robert. On the transient behavior of Ehrenfest and Engset processes. *Advances in Applied Probability*, 2012. To Appear.
- [FR12b] Mathieu Feuillet and Philippe Robert. A scaling analysis of a transient stochastic network (I). Preprint, 2012.
- [FR12c] Mathieu Feuillet and Philippe Robert. A scaling analysis of a transient stochastic network (II): A mean-field picture. Preprint, 2012.
- [FRT99] Christine Fricker, Philippe Robert, and Danielle Tibi. On the rates of convergence of Erlang’s model. *Journal of Applied Probability*, 36(4):1167–1184, 1999.
- [FS09] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.
- [FT97] Guennadi I. Falin and James G. C. Templeton. *Retrials Queues*. Chapman & Hall, 1997.
- [FW98] M. Freidlin and A. Wentzell. *Random Perturbations of Dynamical Systems*. Springer, 1998. 2nd Edition.
- [GH90] John Guckenheimer and Philip Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, volume 42 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1990. Revised and corrected reprint of the 1983 original.
- [Gim65] L. Gimpelson. Analysis of mixtures of wide and narrow-band traffic. *IEEE Transactions on Communication Technology*, 13(3):258–266, 1965.
- [GJ90] Michael R. Garey and David S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., 1990.
- [GLM⁺10] Ayalvadi Ganesh, Sarah Lilienthal, D. Manjunath, Alexandre Proutière, and Florian Simatos. Load balancing via random local search in closed and open systems. In *SIGMETRICS*, pages 287–298, 2010.
- [Gro19] Thomas H. Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *The Annals of Mathematics, Second Series*, 20(4):292–296, July 1919.
- [Gro04] H. Christian Gromoll. Diffusion approximation for a processor sharing queue in heavy traffic. *Annals of Applied Probability*, 14(2):555–611, 2004.
- [GYJ⁺10] Matthieu Gallet, Nezih Yigitbasi, Bahman Javadi, Derrick Kondo, Alexandru Iosup, and Dick Epema. A model for space-correlated failures in large-scale distributed systems. In *EuroPar’10*, pages 88–100, 2010.
- [HK94] P. J. Hunt and T. G. Kurtz. Large loss networks. *Stochastic Processes and their Applications*, 53(2):363–378, 1994.

- [HR81] J.M. Harrison and M.I. Reiman. Reflected Brownian motion on an orthant. *Annals of Probability*, 9(2):302–308, 1981.
- [HRBSD03] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda. Performance anomaly of 802.11b. In *INFOCOM*, volume 2, pages 836–843, 2003.
- [Jac63] James R. Jackson. Jobshop-like queueing systems. *Management Science*, 10(1):131–142, 1963.
- [Jac88] V. Jacobson. Congestion avoidance and control. In *SIGCOMM '88: Symposium proceedings on Communications architectures and protocols*, pages 314–329, New York, NY, USA, 1988. ACM.
- [JSSW10] L. Jiang, D. Shah, J. Shin, and J. Walrand. Distributed random access algorithm: Scheduling and congestion control. *IEEE Transactions on Information Theory*, 56(12):6182 – 6207, 2010.
- [JW08a] L. Jiang and J. Walrand. A distributed CSMA algorithm for throughput and utility maximization in wireless networks. In *the 46th Annual Allerton Conference on Communication, Control, and Computing*, 2008.
- [JW08b] Libin Jiang and Jean Walrand. A distributed CSMA algorithm for throughput and utility maximization in wireless networks. In *Allerton Conference on Communication, Control, and Computing*, 2008.
- [JW09] L. Jiang and J. Walrand. Approaching throughput-optimality in a distributed CSMA algorithm: collisions and stability. In *MobiHoc S3'09*, pages 5–8. ACM, 2009.
- [JWL04] Cheng Jin, D.X. Wei, and S.H. Low. FAST TCP: motivation, architecture, algorithms, performance. In *INFOCOM 2004*, volume 4, pages 2490 – 2501, jul 2004.
- [Kau81] Joseph S. Kaufman. Blocking in a shared resource environment. *IEEE Transactions on Communications*, 29(10):1474–1481, 1981.
- [Kel79] F.P. Kelly. *Reversibility and Stochastic Networks*. Wiley, Chichester, 1979.
- [Kel86] F.P. Kelly. Blocking probabilities in large circuit-switched networks. *Advances in Applied Probability*, 18:473–505, 1986.
- [Kel91] Frank P. Kelly. Loss networks. *Annals of Applied Probability*, 1(3):319–378, 1991.
- [Kel99] Frank Kelly. Mathematical modelling of the Internet. In *Proceedings of the Fourth International Congress on Industrial & Applied Mathematics*, 1999.
- [Ken53] David G. Kendall. Stochastic processes occurring in the theory of queues and their analysis by the method of the embedded markov chain. *Annals of Mathematical Statistics*, 24(3):338–354, 1953.
- [Ken76] D.P. Kennedy. Some martingales related to cumulative sum tests and single-server queues. *Stochastic Processes and their Applications*, 4:261–269, 1976.
- [Kha68] R. Khasminskii. On averaging principle for itô stochastic differential equations. *Kybernetika Chekhoslovakia*, 4(3):260–279, 1968. In Russian.
- [Kha80] R. Z. Khasminski. *Stochastic stability of differential equations*. Sijthoff & Noordhoff, Alphen aan den Rijn, 1980. Translated from the Russian by D. Louvish.
- [Kin62] J. F. C. Kingman. On queues in heavy traffic. *Journal of the Royal Statistical Society, Series B*, 24:383–392, 1962.
- [Kin90] Peter J.B. King. *Computer and communication systems performance modelling*. Computer-Communication Networks. Prentice Hall, 1990.
- [KKF05] S. W. Kim, B.-S. Kim, and Y. Fang. Downlink and uplink resource allocation in IEEE 802.11 wireless LANs. *IEEE Trans. Veh. Technol.*, 54:320–327, Jan. 2005.
- [KM57] Samuel Karlin and James McGregor. The classification of birth and death processes. *Transactions of the AMS*, 86:366–400, 1957.
- [KM59] Samuel Karlin and James McGregor. Coincidence properties of birth and death processes. *Pacific Journal of Mathematics*, 9:1109–1140, 1959.
- [KM65] Samuel Karlin and James McGregor. Ehrenfest urn models. *Journal of Applied Probability*, 2:352–376, 1965.
- [Kur92] T.G. Kurtz. Averaging for martingale problems and stochastic approximation. In *Applied Stochastic Analysis, US-French Workshop*, volume 177 of *Lecture notes in Control and Information sciences*, pages 186–209. Springer Verlag, 1992.
- [KW86] Yuji Kasahara and Shinzo Watanabe. Limit theorems for point processes and their functionals. *Journal of the Mathematical Society of Japan*, 38(3):543–574, 1986.

- [LK91] S.H. Lu and P.R. Kumar. Distributed scheduling based on due dates and buffer priorities. *IEEE Transactions on Automatic Control*, 36(12):1406–1416, 1991.
- [Loy62] R.M. Loynes. The stability of queues with non independent inter-arrival and service times. *Proceedings of Cambridge Philosophical Society*, 58:497–520, 1962.
- [LS63] J. Lamperti and J. L. Snell. Martin boundaries for certain Markov chains. *Journal of the Mathematical Society of Japan*, 15:113–128, 1963.
- [LSZ11] Amaury Lambert, Florian Simatos, and Bert Zwart. Scaling limits via excursion theory: Interplay between crump-mode-jagers branching processes and processor-sharing queues. Submitted, 2011.
- [LTWW93] Will E. Leland, Murrad S. Taqqu, Walter Willinger, and Daniel V. Wilson. On the self-similar nature of ethernet traffic. In *Proceedings of Sigcomm*, pages 183–193, 1993.
- [Mal93] V. A. Malyshev. Networks and dynamical systems. *Advances in Applied Probability*, 25(1):140–175, 1993.
- [Mas87] W. A. Massey. Stochastic orderings for Markov processes on partially ordered spaces. *Mathematics of Operations Research*, 12(2):350–367, May 1987.
- [Mas07] Laurent Massoulié. Structural properties of proportional fairness: stability and insensitivity. *The Annals of Applied Probability*, 17(3):809–839, 2007.
- [Mey95] Sean Meyn. Transience of multiclass queueing networks via fluid limit models. *Annals of Applied Probability*, 5:946–957, 1995.
- [Mit04] M. Mitzenmacher. Digital fountains: a survey and look forward. *Information Theory Workshop, 2004. IEEE*, pages 271–276, 2004.
- [MR00] L. Massoulié and J. W. Roberts. Bandwidth sharing and admission control for elastic traffic. *Telecommunications Systems*, 15:185–201, 2000.
- [MSZ06] E. Modiano, D. Shah, and G. Zussman. Maximizing throughput in wireless networks via gossiping. *SIGMETRICS Perform. Eval. Rev.*, 34(1):27–38, 2006.
- [MVW93] N. McKeown, P. Varaiya, and J. Walrand. Scheduling cells in an input-queued switch. *Electronics Letters*, 29(25):2174–2175, 1993.
- [MW00] Jeonghoon Mo and Jean Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Trans. Netw.*, 8(5):556–567, 2000.
- [Nag84] J. Nagle. Congestion Control in IP/TCP Internetworks. RFC 896, January 1984.
- [NTS10] J. Ni, B. Tan, and R. Srikant. Q-CSMA: Queue-length based CSMA/CA algorithms for achieving maximum throughput and low delay in wireless networks. In *IEEE INFOCOM*, 2010.
- [Pal93] José Luis Palacios. Fluctuation theory for the Ehrenfest urn via electric networks. *Advances in Applied Probability*, 25(2):472–476, 1993.
- [PBS07a] Fabio Picconi, Bruno Baynat, and Pierre Sens. An analytical estimation of durability in DHTs. In Tomasz Janowski and Hrushikesh Mohanty, editors, *Distributed Computing and Internet Technology*, volume 4882 of *Lecture Notes in Computer Science*, pages 184–196. Springer, 2007.
- [PBS07b] Fabio Picconi, Bruno Baynat, and Pierre Sens. Predicting durability in DHTs using Markov chains. In *Proceedings of the 2nd Intl. Conference on Digital Information Management (ICDIM)*, 2007.
- [PF95] Vern Paxson and Sally Floyd. Wide area traffic: The failure of poisson modeling. *IEEE/ACM Transactions on Networking*, 3(3):226–244, 1995.
- [PGK88] David A. Patterson, Garth Gibson, and Randy H. Katz. A case for redundant arrays of inexpensive disks (raid). In *SIGMOD*, pages 109–116, 1988.
- [PSV77] G. Papanicolaou, D. Stroock, and S.R.S Varadhan. Martingale approach to some limit theorems. *Statistical Mechanics, Dynamical Systems and the Duke Turbulence Conference*, 3, 1977.
- [PW11] Ohad Perry and Ward Whitt. An ODE for an overloaded x model involving a stochastic averaging principle. *Stochastic Systems*, 1(1):59–108, 2011.
- [PYLC10] A. Proutière, Y. Yi, T. Lan, and M. Chiang. Resource allocation over network dynamics without timescale separation. In *IEEE INFOCOM*, 2010.
- [Ram06] Kavita Ramanan. Reflected diffusions defined via the extended Skorokhod map. *Electronic Journal of Probability*, 11:934–992, 2006.

- [RD01] Antony Rowstron and Peter Druschel. Storage management and caching in PAST, a large-scale, persistent peer-to-peer storage utility. In *Proceedings of SOSR*, pages 188–201. ACM, 2001.
- [RGK⁺05] Sean Rhea, Brighten Godfrey, Brad Karp, John Kubiatowicz, Sylvia Ratnasamy, Scott Shenker, Ion Stoica, and Harlan Yu. OpenDHT: a public DHT service and its uses. In *Proceedings of SIGCOMM*, pages 73–84, 2005.
- [Rob81] James W. Roberts. A service system with heterogeneous user requirement. In Guy Pujolle, editor, *Performance of Data Communications Systems and Their Application*, pages 423–431, 1981.
- [Rob03] Philippe Robert. *Stochastic Networks and Queues*. Stochastic Modeling and Applied Probability Series. Springer-Verlag, New York, 2003. xvii+398 pp.
- [RP06] S. Ramabhadran and J. Pasquale. Analysis of Long-Running Replicated Systems. In *INFOCOM 2006*, pages 1–9, 2006.
- [RS92] A. N. Rybko and A. L. Stolyar. Ergodicity of stochastic processes describing the operation of open queueing networks. *Problems of Information Transmission*, 28(3):3–26, 1992.
- [RS05] Barath Raghavan and Alex C. Snoeren. Decongestion control. In *Proceedings of the 5th ACM Workshop on Hot Topics in Networks*, november 2005.
- [RS08] S. Rajagopalan and D. Shah. Distributed algorithm and reversible network. In *CISS*, pages 498–502, 2008.
- [RSS09] Shreevatsa Rajagopalan, Devavrat Shah, and Jinwoo Shin. Network Adiabatic Theorem: An Efficient Randomized Protocol for Contention Resolution. In *Proceedings of SIGMETRICS*, pages 133–144, New York, NY, USA, 2009. ACM.
- [Rud87] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.
- [RW79] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes & martingales vol. 1: Foundations*. Cambridge University Press, 2000 (1979).
- [RW87] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes & martingales vol. 2: Itô Calculus*. Cambridge University Press, 2000 (1987).
- [RY99] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.
- [RZ10] J. E. Reed and B. Zwart. Limit theorems for bandwidth sharing networks with rate constraints. Preprint, 2010.
- [Sch00] Wim Schoutens. *Stochastic processes and orthogonal polynomials*, volume 146 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 2000.
- [Sen96] E Seneta. Markov and the birth of chain dependence theory. *International Statistical Review*, 64(3):255–263, 1996.
- [Ser99] R. Serfozo. *Introduction to Stochastic Networks*. Springer-Verlag New York, 1999.
- [Sko61] A.V. Skorokhod. Stochastic equations for diffusions in a bounded region. *Theory of Probability and Its Applications*, 6:264–274, 1961.
- [SS11] D. Shah and J. Shin. Randomized scheduling algorithm for queueing networks. *The Annals of Applied Probability*, page to appear, 2011.
- [ST98] Wim Schoutens and Jozef L. Teugels. Lévy processes, polynomials and martingales. *Comm. Statist. Stochastic Models*, 14(1-2):335–349, 1998. Special issue in honor of Marcel F. Neuts.
- [ST10] Florian Simatos and Danielle Tibi. Spatial homogenization in a stochastic network with mobility. *The Annals of Applied Probability*, 20(1):312–355, 2010.
- [Sze75] Gábor Szegő. *Orthogonal polynomials*. American Mathematical Society, Providence, R.I., fourth edition, 1975. American Mathematical Society, Colloquium Publications, Vol. XXIII.
- [Szp93] Wojciech Szpankowski. Stability conditions for some distributed systems: buffered random access systems. *Advances in Applied Probability*, 26:498–515, 1993.
- [Tak69] Lajos Takács. On Erlang’s formula. *The Annals of Mathematical Statistics*, 40(1):71–78, 1969.
- [TE92] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, 37:1936–1948, 1992.

- [TSM98] François Theberge, Alain Simonian, and RaviR. Mazumdar. Upper bounds for blocking probabilities in large multi-rate loss networks. *Telecommunication Systems*, 9:23–39, 1998.
- [TW93] L.M. Taylor and R.J. Williams. Existence and uniqueness of semimartingale reflecting Brownian motions in an orthant. *Probability Theory and Related Fields*, 96:283–317, 1993.
- [vdVBS09] P. van de Ven, S. Borst, and S. Shneer. Instability of MaxWeight Scheduling Algorithms. In *Proceedings of INFOCOM*, pages 1701–1709, april 2009.
- [vdVBvLP10] P. M. van de Ven, S. C. Borst, J. S. H. van Leeuwen, and A. Proutière. Insensitivity and stability of random-access networks. *Perform. Eval.*, 67:1230–1242, November 2010.
- [Whi85] Ward Whitt. Blocking when service is required from several facilities simultaneously. *AT&T Technical Journal*, 64(8):1807–1856, 1985.
- [Wil98] R. J. Williams. Diffusion approximations for open multiclass queueing networks: sufficient conditions involving state space collapse. *Queueing Systems: Theory and Applications*, 30:27–88, 1998.
- [WW96] E. T. Whittaker and G. N. Watson. *A course of modern analysis*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1996. Reprint of the fourth (1927) edition.
- [XHR04] Lisong Xu, K. Harfoush, and Injong Rhee. Binary increase congestion control (BIC) for fast long-distance networks. In *INFOCOM 2004*, volume 4, pages 2514 – 2524, jul 2004.
- [XMS⁺03] Qin Xin, Ethan L. Miller, Thomas Schwarz, Darrell D. E. Long, Scott A. Brandt, and Witold Litwin. Reliability mechanisms for very large storage systems. In *Proceedings of the 20th IEEE / 11th NASA Goddard Conference on Mass Storage Systems and Technologies*, pages 146–156, 2003.