Fluid Approximations for Stochastic Telecommunication Models

Laura Aspirot

UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

Thesis submitted for the degree of
Doctor en Matemática

Title

Fluid Approximations for Stochastic Telecommunication Models

Author:
Laura Aspirot
laspirot@ccee.edu.uy

Advisor:
Gerardo Rubino
gerardo.rubino@inria.fr
Co-Advisor:
Ernesto Mordecki
mordecki@cmat.edu.uy

December 27, 2019

Abstract

Stochastic processes and differential equations are both widespread mathematical models in many disciplines. In particular they are both standard tools in telecommunications. The first motivation of this work is to analyze relationships between these stochastic and deterministic approaches, in the framework of telecommunications modeling. We consider both stochastic processes and differential equations as two suitable ways for modeling a system, and we study a link between both representations.

In the literature, the approximation of stochastic systems by deterministic ones, usually modeled by differential equations, appear with different flavors, under different assumptions and contexts and with many different names: fluid limits, hydrodynamic limits, mean field approximations. These ideas appear in many disciplines, for instance physics, biology, chemistry, queuing theory, game theory. In this work we will use fluid limits as the expression that best represents the topics addressed here.

The main objective of fluid limits is to analyze a complex stochastic system studying a simplified model, commonly deterministic, and represented by differential equations. This is a useful technique when the stochastic system is difficult to analyze, or even difficult to simulate. Once we obtain this approximation certain properties of the stochastic system can be derived from the behavior of the deterministic one.

Usually these limits are obtained as the limit of a sequence of stochastic processes in the fashion of the Law of Large Numbers, and represent the mean behavior of a stochastic process. Properties as fixed points and asymptotic behavior of the deterministic system obtained as limit are generally closely related with the stationary regime of the stochastic one. In addition, another natural question is about the asymptotic distribution of the system, in the sense of the relationship between the Law of Large Numbers and the Central Limit Theorem. Results in this context are also known as diffusion approximations.

In this work we will address some of the previous topics in three different

frameworks, all motivated by telecommunications modeling with stochastic processes. In all of them we consider a sequence of Markov processes and a limit process, obtained when features related with the size of the system tends to infinity.

We first explore fluid limits for modeling Peer to Peer BitTorrent-like systems. We study BitTorrent models proposed in the literature, both stochastic and deterministic, we compare those models, and we justify the passage, by a limit construction, from a stochastic process to a fluid approximation driven by a differential equation. We also analyze the asymptotic distribution.

As a second problem we deal with a Machine Repairman Model introducing phase-type distributions. This problem presents two main difficulties, that are two different time scales and discontinuous transition rates. We prove that the Markov process describing the system evolution converges to a deterministic process with piecewise smooth trajectories. We analyze the deterministic system by studying its fixed points, and we find different behaviors depending only on the expected values of the phase-type distributions involved. Concerning the asymptotic distribution we explore different scaling methods obtaining Gaussian and non-Gaussian distributions, depending on the scaling and on the parameters of the stochastic processes.

The last problem addressed concerns Cognitive Radio Networks, where we find a fluid approximation for the stochastic process that models the number of primary and secondary users. In this system we also have discontinuous transition rates that lead to a fluid limit with piecewise smooth trajectories. We analyze the fluid limit in transient and stationary regime. We also obtain the asymptotic distribution for the system in some cases, and we find different behaviors and obtain Gaussian and non-Gaussian asymptotic distributions.

Resumen

Los procesos estocásticos, y en particular los procesos de Markov y las cadenas de Markov, han sido modelos matemáticos masivamente utilizados para estudiar diversos fenómenos. Por su parte las ecuaciones diferenciales también son herramientas ampliamente utilizadas en el modelado matemático. En muchas de las aplicaciones de matemática que conocemos ambos tipos de modelos y de abordajes coexisten para analizar los mismos problemas.

La motivación de este trabajo surge de la existencia de estos dos tipos de acercamientos a diferentes problemas en telecomunicaciones, y la primera pregunta que planteamos es cómo se puede establecer una relación entre modelos estocásticos y determinísticos para un mismo objeto. Por otra parte, cuando existe esta relación, interesa saber qué tipo de características de uno de los modelos puede brindar información sobre el otro, así como medir, en algún sentido, qué tan exacta es esa aproximación. Para esto es necesario estudiar en qué marco podemos analizar la relación entre modelos estocásticos y determinísticos y cuáles son las herramientas y técnicas involucradas. En la literatura encontramos una amplia variedad de problemas y técnicas en este sentido, con diferentes nombres, y múltiples variantes, pero que comparten ciertas características esenciales. Así encontramos denominaciones como límites fluidos, aproximaciones tipo campo medio, límites hidrodinámicos. Estas denominaciones involucran ideas matemáticas usadas desde larga data en diferentes problemas, por ejemplo en física, biología, química, teoría de colas, teoría de juegos, que buscan simplificar modelos estocásticos complejos, planteando para ellos su aproximación determinística.

Un contexto general para analizar estas relaciones se conoce como límites fluidos. Este será el objeto de estudio en este trabajo, en particular su recorte a modelos de telecomunicaciones. La finalidad es entonces aproximar modelos, con diferentes tipos de complejidades a la hora de su análisis, mediante modelos más sencillos. La dificultad para tratar los modelos estocásticos puede estar dada por las dependencias internas en el sistema, por

la cantidad de individuos, y pueden ser difíciles de estudiar analíticamente o incluso mediante simulaciones, ya que estas pueden ser computacionalmente muy costosas. Sin embargo estos modelos muchas veces pueden simplificarse a modelos determinísticos gobernados por ejemplo por ecuaciones diferenciales. Mediante estas aproximaciones en gran parte de los casos el comportamiento del proceso estocástico original puede analizarse a partir de características del modelo determinístico.

En general estas aproximaciones de procesos estocásticos son asintóticas en algún parámetro del sistema, en muchos casos vinculado a su tamaño, y lo que se obtiene es un límite en media, en el sentido de la Ley de los Grandes Números. Entonces una de las preguntas que surge es la velocidad de convergencia. Por ese motivo, el otro tema que se aborda en esta tesis es la convergencia tipo Teorema Central del Límite, que también se denomina aproximación por difusiones. Así, un segundo objetivo es, una vez que un sistema estocástico se aproxima por uno determinístico, estudiar qué distribución tiene el error de la aproximación.

En lo que sigue estudiamos tres modelos de límites fluidos motivados en problemas que aparecen en telecomunicaciones. Estos tres modelos analizados permiten ver el funcionamiento de la técnica de límites fluidos en diferentes aplicaciones, y mostrar resultados del comportamiento asintótico de los sistemas a partir del análisis de sus límites determinísticos.

Este trabajo consta de tres partes, la primera dedicada al estudio de redes par a par, en particular al análisis de un modelo para el protocolo BitTorrent. Para ese modelo se estudian límites fluidos, se describe cómo se obtienen estos límites y se estudian aproximaciones Gaussianas. Parte de los resultados de este capítulo se encuentran en [AMR11].

La segunda parte de la tesis presenta un modelo de teoría de colas de fallas y reparaciones. Para ese modelo se introducen distribuciones tipo fase, y se obtiene un límite fluido y un límite en distribución. En este caso el sistema presenta diferentes escalas de tiempo, al mismo tiempo que da lugar a un límite determinístico que es un sistema dinámico diferenciable a tramos. A nivel de distribución asintótica también encontramos límites Gaussianos y no Gaussianos. Estos resultados fueron presentados en [AMR13].

El tercer problema abordado consiste en el estudio de límites fluidos y distribución asintótica en un modelo para redes cognitivas. Aquí tenemos un sistema dinámico diferenciable a tramos y para la distribución asintótica podemos obtener un resultado del tipo Teorema Central del Límite en algunos casos, mientras que en otros, con otro escalado, se obtiene una distribución asintótica no Gaussiana. Algunos resultados de este capítulo se presentaron en [RAB15].

Acknowledgements

This work was partially supported by the following projects:

STIC-AmSud 2018 “Algorithms for the Capacity Crunch problem in Optical Networks” (ACCON),

CSIC I+D 2016 “Límites fluidos, aproximación por difusión y grandes desvíos en sistemas de comunicación de gran dimensión”,

CSIC Grupos I+D: “Análisis de redes, tráfico y estadística de servicios” (ARTES).

Agradecimientos

A Gerardo y a Ernesto por toda la generosidad, la disponibilidad, la paciencia, el entusiasmo, la capacidad de acompañar.

A Claudina y a Pablo, que hacen que compartir el trabajo sea una alegría.

A los integrantes del tribunal por su disponibilidad, su tiempo, y sus devoluciones.

A todos las compañeras y compañeros del IMERL y de FING, y del DMC y de FCEA, que compartieron conmigo y me acompañaron en diferentes momentos de este trabajo.

A los integrantes del equipo Dionysos, que hicieron que trabajar en Rennes fuera un disfrute.

A las compañeras y compañeros de matemática, al Canario, y a quienes en el Juan XXIII acompañaron, ayudaron y alentaron con muchas ganas durante este trabajo.

A mis estudiantes del Juan XXIII, de FING y de FCEA.

A todas mis amigas y amigos que compartieron desde cerquita este tiempo y este trabajo.

A mi familia, que siempre está.

A la memoria del Flaco Peirano que me enseñó probabilidad por primera vez.

Contents

Introduction

There are several examples of complex stochastic systems for which analytic expressions cannot be derived, or that are even difficult to simulate, as a Markov chain with a huge number of states. The complexity of the system may be for example due to its size or due to its dependence structure. For large systems, numerical simulations with large state spaces would also be very costly. However, in many cases those systems can be studied by analyzing deterministic models, obtained as asymptotic approximations of the original ones. The deterministic systems in most cases are described by an ordinary differential equation (ODE), by a piecewise smooth dynamical system (PWSDS), by a partial differential equation (PDE) or by a difference equation (DE). These kind of limits are known as fluid limits, or, in some variants, as mean field approximations or hydrodynamic limits. The limits described before are deterministic systems, but there are also stochastic fluid limits, out of the scope of this work.

The study of fluid limits is a widely developed technique, very useful for the analysis of large Markov systems. As an example, let us consider the fluid limit for a $M/M/1$ queue [Rob03] with arrival rate λ and service rate μ. Let $X(t)$ be the number of units in the system at time t and let $\widehat{X}^N(t) = X(Nt)/N$ be the scaled number of units. Time is accelerated by a factor N, and the initial state is also scaled by the same factor, for example $X(0) = Nx_0$, and then $\widehat{X}^N(0) = X(0)/N = x_0$. Whenever the scaled initial condition converges to x_0 with N, then the process $\widehat{X}^N$ can be approximated, for large N, by the deterministic solution to $x' = \lambda - \mu$ if $x > 0$, $x' = 0$ if $x = 0$. For $\lambda < \mu$ the equation defines a piecewise smooth dynamical system, with a solution for the initial condition x_0 that is smooth on $[0, x_0/(\mu - \lambda))$ and $(x_0/(\mu - \lambda), \infty)$ (Figure 1). If the initial condition is 0, the solution remains at zero.

Other examples in [Rob03] are the $M/M/\infty$ queue and the $M/M/N/N$ queue. Let λN be the arrival rate and μ the service rate in both cases and let $\widetilde{X}^N$ be the number of units in the system. The scaling is different

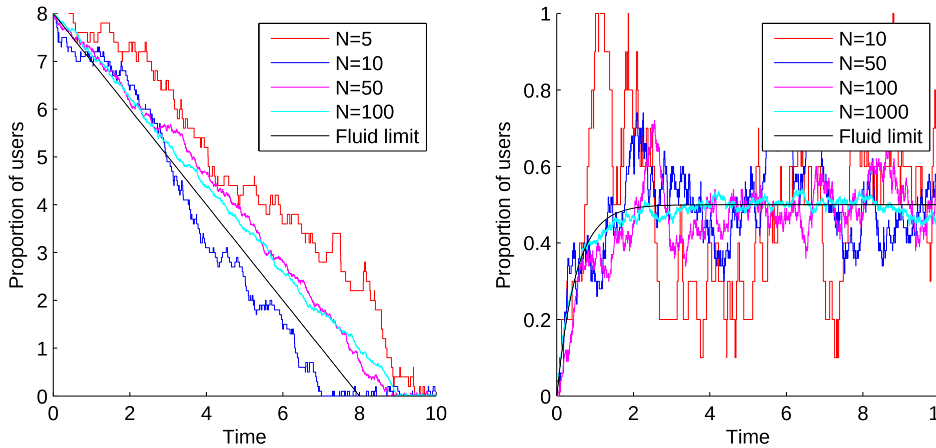

Figure 1: Left: $M/M/1$ queue. Scaled process $\widehat{X}^N(t)$ for different values of N and fluid limit ($\lambda = 1$, $\mu = 1$, $x_0 = 8$). Right: $M/M/\infty$ queue. Scaled process $X^N(t)$ for different values of N and fluid limit ($\lambda = 1$, $\mu = 1$, $x_0 = 0$).

from the $M/M/1$, as time is not scaled, only the arrival rate is accelerated, and the total service rate scales with the number of units in the system. The scaled number of units $X^N = \widetilde{X}^N/N$ converges, provided the initial condition converges to a constant, to the solution to $x' = \lambda - \mu x$ for the $M/M/\infty$ queue (Figure 1), and to the solution to $x' = \lambda - \mu x$, if $x < 1$, $x' = 0$ if $x = 1$, in the $M/M/N/N$ model. In the last case, we find again a piecewise smooth dynamical system, that converges exponentially fast to $\rho = \lambda/\mu$ if $\rho < 1$ and to 1 if $\rho \geq 1$.

Fluid limits have been used for particle systems, biology, epidemics, game theory, computer science models, as well as in the study of telecommunication networks. There are many examples of the latest, to mention some of the first works that addressed different topics in telecommunications modeling, we have for instance modeling internet protocols like TCP [TM06], wireless systems [BMP08, BJLM16], queuing models [MMR98, Gra00], game theory [BLB08]. We also find in this framework different approaches in the characterization of the models and limits obtained. There are studies of fluid limits and mean field approximations considering systems with discrete time and deterministic approximations in discrete time by means of difference equations [LBMM07], as well as systems in discrete time with approximations in continuous time, involving ODEs [BLB08] and systems in continuous time with approximations in continuous time, involving ODEs [Bor11a], PWSDS [Bor11a, GG12, Bor16] or PDEs [CLBR09, PF12].

Generally speaking, starting from a stochastic model the objective is to find a deterministic approximation for the original process. This introduces the problem of finding the suitable scale for the approximation. For example, as we mentioned previously, some classical results in queuing theory consider a sequence of stochastic processes indexed by an integer N, where some key state variable appears divided by N, and the time variable is multiplied or accelerated by the same factor, obtaining a deterministic limit when N goes to infinity (as in the case of the $M/M/1$ queue). In addition, in other queuing models (as in the $M/M/\infty$), or in other areas such as in biology or in the analysis of epidemic phenomena, a typical scaling consists in dividing by N, and in considering transition rates increasing with N (jumps are of order $1/N$ and transition rates of order N, which means that the product of jump and transitions remains or tends to a constant as N increases). We refer to [Rob03, HW81, MMR98] as references for this and other scaling regimes, under which there are several limit results.

Different mathematical tehniques are involved in order to obtain these results, as for example the random time change for stochastic processes introduced by Kurtz [EK86] or a martingale approach, via exponential or L^2 martingales [SW95, Rob03, DN08].

Complementary to fluid limits, and closely related, mean field approximations come from physics, used there to study systems with a large number of interacting particles. When the number of particles increases each particle behaves as if it were under the action of a global force, the mean field, due to the influence of the other particles in the system. One of the main applications of this approach is related with procedures for fast simulation of systems, where, instead of simulating the whole system, a single individual is simulated alone and the whole behavior of the rest of the system is replaced by its mean field limit. Many applications to telecommunications appeared in the literature. For instance [LBMM07, BLB08, CLBR09, BMP10] and the references therein cover a wide range of techniques and applications. Further, mean field methods have been applied to game theory and optimal control, all in the framework of telecommunication research (see for example [TLBEAA09, GGLB10] and references therein). In mean field approximations we can distinguish two steps: the first one is focused on the occupation measure limit, that is, the asymptotic proportion of individuals in each state, and the second one is focused on the decoupling assumption, which means that, asymptotically, the state of each individual is independent of the state of the others. A very frequent approach to the first step consists in proving the limit with the same techniques as in the fluid limits case. The proof of the asymptotic independence relies in limit theorems as in [Szn91].

As general references on the mathematical tools needed for fluid limits and mean field approximations, with emphasis in applications in telecommunications, whose techniques are taken in this work, we refer to [EK86, SW95, Rob03, DN08, LB10a, BHLM13]. [EK86] is a classical reference, that treats fluid limits and diffussion approximations, especially in Chapter 11 (Density dependent population process). [SW95] presents fluid limit closely related with large deviations techniques, but also addresses many problems stated here, as convergence in stationary regime and limits for Markov processes with discontinuous transition rates. [Rob03, DN08] present classical results, obtained mostly with martingale techniques. [LB10a, BHLM13] and references therein specially cover problems with discontinuous transition rates and convergence in stationary regime.

Outline

This thesis is divided in three chapters, motivated by three different problems in telecommunications, where in each case we study a stochastic model and its derived fluid limit and asymptotic distribution, finding different technical issues in each case and also considering in some cases different approaches from the literature. All problems and results are illustrated with simulations. The appendices summarizes some background material.

Peer to Peer

The first chapter studies fluid limits for modeling peer to peer networks, in particular BitTorrent-like systems.

Peer to peer (P2P) networks are decentralized networks that avoid the traditional client/server model, and where users are both, at the same time, servers and clients.

One of its most popular protocols is BitTorrent [Coh03] for file sharing applications. BitTorrent divides the file into small pieces called chunks. Each peer connects to others and downloads simultaneously different chunks. There are two types of peers: leechers and seeds. Leechers download parts of the file from other peers and upload parts of the file to other leechers. Seeds have all the file and only remain in the system to help leechers to get missing file parts. The problem addressed here consists on modeling the number of leechers and seeds in a BitTorrent network.

Based on well known models from the literature we analyze the passage from stochastic models to deterministic ones, both in transient and in stationary regime. We also study the asymptotic distribution of the difference between the stochastic process and its fluid limit in transient regime, and the limit of this distribution when time goes to infinity.

Machine Repairman Model

The second chapter analyzes a Machine Repairman Model. The Machine Repairman Model (MRM) is a basic Markovian queue representing a finite number N of machines that can fail independently and, then, be repaired by a repair facility. The latter, in the basic model, is composed of a single repairing server with a waiting room for failed machines managed in FIFO (first in, first out) order, in case the repairing server is busy when units fail. In Kendall's notation, this is the $M/M/1//N$ model, specifying that lifetimes and repair times are exponentially distributed.

This $M/M/1//N$ model is well known and widely studied in queuing theory and in many applications, as for example in telecommunications or in reliability, where most of the studies look at the queue in equilibrium.

The model is a precursor of the development of queuing network theory, motivated first in computer science. In particular, Scherr from IBM used it in 1972 for analyzing the S360 OS (see [Lav83]). Many extensions of the basic model have been studied, considering more than one repairing server, different queuing disciplines, and other probability distributions for the lifetime or repair time.

In this work we analyze a repairman problem with N working units that break randomly and independently according to a phase-type distribution. Broken units go to one repairman where the repair time also follows a phase-type distribution (that is, a $PH/PH/1//N$ model). We consider a scaled system, dividing the number of broken units and the number of working units in each phase by the total number of units N and accelerating the repairman with the same factor. The scaled process has a deterministic limit when N goes to infinity.

The first problem that the model presents is that there are two time scales: the repairman changes its phase at a rate of order N, whereas the total scaled number of working units changes at a rate of order 1.Another problem is that transition rates are discontinuous because of idle periods at the repairman. Thus, we obtain a piecewise smooth dynamical system as fluid limit, for which we study fixed points and stability. Finally we analyze the asymptotic distribution in the exponential case.

Cognitive Radio Networks

The third chapter studies fluid models in the context of cognitive networks. Cognitive Radio Networks (CR networks) have emerged in the last years as a solution of two problems: spectrum underutilization and spectrum scarcity.

With the rapid development of wireless communications, the demand on spectrum has been growing dramatically resulting in the spectrum scarcity problem: unlicensed bands are too crowded while licensed bands are vastly underutilized.

CR Networks has been proposed as a promising technology to solve that problem by an intelligent and efficient dynamic spectrum access. In this new paradigm we can identify two classes of users: primary (PU) and secondary (SU). PUs are the licensed users, they have allocated a certain portion of spectrum. SUs (also called cognitive users) are devices which are capable of detecting unused licensed bands and adapt their parameters for using them.

The main idea of CR networks is to dynamically reallocate unused licensed frequency bands to secondary users.

The focus of the third chapter is on the analysis and characterization of a dynamic spectrum sharing mechanism where primary users have strict priority over secondary ones.

We develop a stochastic model for primary and secondary users in cognitive radio networks and analyze it by fluid limits, with the objective of proposing some way of admission control for secondary users in order to guarantee certain quality of service requirements, in terms of loss probability, for those secondary users in the network.

Contributions

Peer to Peer

In the first chapter we explore models proposed for the analysis of BitTorrent P2P systems and we provide the arguments to justify the passage from the stochastic process, under adequate scaling, to a fluid approximation driven by an ODE.

We prove uniform almost sure convergence in compact time intervals. We also explore the link between the stationary regime of the stochastic models and the fixed points of the associated ODEs. Finally, we analyze the asymptotic distribution of the scaled process.

The contributions consist first in the mathematical justification that the deterministic fluid models presented in [QS04] and [RAR10] are fluid limits of stochastic models. En each case we define a stochastic model and construct a sequence of stochastic processes such that, under adequate scaling, converges to a deterministic model driven by an ODE.

The second contribution is the proof of the existence of a stationary regime for each process in the sequence and then of the fact that the sequence of processes in stationary regime converges to the ODE's fixed point.

Finally we describe the asymptotic distribution. We prove that the difference between the scaled process and the deterministic one can be approximated by a Gaussian process in transient regime and we also describe the limit of this process when time goes to infinity.

Some of the results presented in the first chapter are presented in [AMR11].

Machine Repairman Model

In the second chapter we analyze a repairman problem with N working units that break randomly and independently according to a phase-type distribution. Broken units go to one repairman where the repair time also follows a phase-type distribution (that is a $PH/PH/1//N$ model).

In our main result we prove that the scaled Markovian multidimensional process describing the system dynamics converges to an ODE solution when N tends to infinity. The convergence is in probability and takes place uniformly in compact time intervals (*u.c.p.* convergence). In this case the deterministic limit, the solution to the ODE, is only piecewise smooth.

We analyze the properties of this limit, and we prove the convergence in probability of the system in stationary regime to the deterministic fixed point. We also find that this fixed point only depends on the repair time by its mean. As a matter of fact, when in equilibrium, if the repair times are exponentially distributed, the distribution of the number of broken machines has the insensitivity property with respect to the lifetime distribution (only the latter's mean appears in the former). Although this behavior may be expected because of the scaling, it is not straightforward. In addition, the whole phase-type lifetime distribution takes part in the result. Different lifetimes with the same mean give different behaviors, both in the transient and in the stationary regimes.

The analysis of the asymptotic distribution of the difference between the stochastic process and its fluid limit presents technical problems because the limit is non-differentiable. We address these issues only in stationary regime for the exponential case.

Some results presented in the second chapter are included in [AMR13].

Cognitive Radio Networks

The main contribution of this chapter is the analysis and characterization of a possible model of spectrum sharing in CR networks. This spectrum sharing mechanism is designed in order to improve the average utilization of the spectrum while ensuring a small probability of interruption to secondary users.

We consider a Markov chain that represents the population of the different types of users in the system. We formulate the associated fluid model and we study its deterministic solutions.

We the Markov model for users and two simple admission control criteria, one deterministic and one probabilistic, that for a system with a large number of users, guarantees with high probability that secondary users which enters in the system will not have service interruptions.

We study the fluid limit and its fixed points, with and without admission control, together with its asymptotic distribution. We find non-differentiable trajectories for the limit process and, depending on the load of the system, Gaussian and non-Gaussian asymptotic distribution.

This chapter is part of a joint work with C. Rattaro and P. Belzarena, presented in [RAB15], some results are part of C. Rattaro's PhD thesis [Rat18] and of a submitted paper [RABM19].

Chapter 1

Peer to Peer

This chapter presents our work on modeling Peer to Peer (P2P) applications, in particular BitTorrent. We make a brief presentation of general P2P networks and then we describe the BitTorrent protocol. We focus on some aspects of BitTorrent modeling. In Section 1.1 we summarize and analyze different models and results from the literature that motivate and are related with this work. We state our stochastic model for BitTorrent networks in Section 1.2, then we show some numerical examples, simulated in order to illustrate different models, in Section 1.3, and present the results in Section 1.4 about fluid limits and in Section 1.5 about asymptotic distribution.

Peer to peer networks are decentralized networks that avoid the traditional client/server model, and where users are, at the same time, servers and clients. There is not a central entity that rules the network, and peers are organized in a distributed way. P2P networks are used as content delivery networks, and originally were thought for file sharing, but with some modifications P2P protocols have been adapted also for streaming. At the beginning P2P file sharing became very popular with the development of Napster by 2000, devoted to MP3 file sharing between users. After that many applications appeared, as Gnutella, Kazaa, eDonkey, BitTorrent [QS04].

There exist different P2P protocols, some based on unstructured networks, others with networks with some kind of structure, and hybrid ones. In unstructured P2P networks nodes are indistinguishable, that is the most pure version of P2P, whereas in structured P2P networks there are selected nodes that help to organize content delivery. For example in the BitTorrent case, there are some nodes that gather information from users and help to connect peers with each other.

There are also single-torrent and multi-torrent networks, that means that there is only one content distributed between users, or there are several different contents distributed for the same set of users.

In the last decades research about P2P networks was deeply developed. Concerning more theoretical models, in 2004, Yang and de Veciana [YdV04] and Qiu and Srikant [QS04] pointed out the relevance of P2P traffic, for file sharing applications. The major advantage of P2P networks, first revealed in practice, and then thoroughly studied with different models, is that its performance scales with the number of users. As for classical client/server applications performance degrades as the number of users increases, the goal of P2P networks is to maintain or even improve performance when this number increases. On the other hand, one of the main challenges of P2P networks is to guarantee the cooperation of users, and to deal with users that consume resources from the network but are not disposed to share their own.

The first works in P2P networks for file sharing were mostly focused in protocols, network design, and also in traffic measurement [IUKB+04, GCX+05, PGES05, GCX+07]. Then works on modeling and performance evaluation appeared, and these are the results that motivate our approach and are addressed later in this work. Streaming P2P networks also opened other research lines, related with real time applications and searching and caching content in P2P networks. (See for example [CRRB+08, SSY11] and references therein.)

BitTorrent is a peer to peer protocol, introduced by Bram Cohen [Coh03] for file sharing over a network. BitTorrent divides the target file into small files (chunks). Each peer connects to others and downloads from them different chunks. For this purpose there is a centralized controlling software (tracker). Peers ask the tracker for a file and the tracker returns a list of peers randomly chosen among those that have the file. Then each peer connects to them, see the chunks they have and ask for their missing chunks. Each peer can upload (unchock) only to four peers, the ones with best downloading bandwidth. Periodically each peer performs what is called optimistical unchocking, exploring other peers, and can upload chunks to a fifth peer. In addition, BitTorrent implements algorithms to search chunks, that prevent that some chunks become more difficult to obtain and that give preferences to peers that have almost finished downloading the file. These are the rarest-first policy and the endgame mode. There are two types of peers: leechers and seeds. Leechers download parts of the file from other peers and upload parts of the file to other leechers. Seeds have all the file and only remain in the system to help leechers to get missing file parts, they are altruist nodes.

In the literature authors point out several issues to be addressed when studying BitTorrent. One is peer evolution, that is the dynamics of leechers and seeds in the system. Other issue is the scalability of the network, that means the analysis of how performance of the network is affected by the number of users. The efficiency in file sharing is another topic of discussion, concerning how to match peers in order to improve the way of obtaining their missing chunks. In addition, we have the problem of incentives. P2P networks have, as one of their main drawbacks, what is called free riders, that are peers that download files from the network but do not upload their own files. Incentive policies try to discourage this type of users.

In this chapter we first describe some BitTorrent models in the literature in Section 1.1. We focus on the fluid models presented in [QS04] and [RAR10]. For these fluid models we aim to state a stochastic model and properly justify the passage from the stochastic model to the fluid one. We also want to find a fluid approximation for the BitTorrent model presented in [YdV04]. For this purpose in Section 1.2 we present stochastic models for BitTorrent, suitable to find a fluid limit, considering the different cases studied in [YdV04, QS04, RAR10]. In Section 1.3 we illustrate these models by simulations. In Section 1.4 we derive the fluid limit for the stochastic models proposed in Section 1.2, properly scaled by a parameter N that goes to infinity, and that can be interpreted as the scaling size of the system. We obtain as a result the fluid model from [QS04], and as a generalization the model in [RAR10]. We also stablish a fluid limit for the model presented in [YdV04]. In [QS04] and [RAR10] the initial model is deterministic, however both seem to assume an implicit stochastic model. We, instead, consider a sequences of stochastic models that properly scaled converge to the deterministic ones. On the other hand, in [YdV04] the system is described by a Markov chain, but there is not a deterministic approximation, so here we scale its original system and find its related fluid limit. Then we justify the study of stationary regime by means of the study of fixed points for the ODE obtained as fluid limit. In [QS04] the stationary regime is represented by the ODE's fixed point, and the ODE is also deeply analyzed in [QS08]; here we justify this approximation. Finally in Section 1.5 we describe the asymptotic distribution. We find a limit in the sense of the Central Limit Theorem for the difference between the stochastic process and the fluid limit. This Gaussian approximation also appears in [QS04] as an additional model, but not explicitly obtained as a limit. We formulate this result as a consequence of Section 1.4 and we also analyze the Gaussian limit distribution when time goes to infinity.

1.1 BitTorrent models in the literature

In this section we describe several BitTorrent models studied in the literature, that motivated our model.

1.1.1 A Markov chain model

We first describe a stochastic model by Yang and de Veciana in [YdV04]. The BitTorrent network is described first by a deterministic model, and then using a branching process for the transient regime and a Markov model for the stationary regime.

For the Markov model, the following parameters are considered:

$X(t)$: number of leechers at time t,

$Y(t)$: number of seeds at time t,

λ: arrival rate (Poisson) of peers,

μ: uploading/downloading rate for each peer,

$\eta \in [0,1]$: efficiency factor, a constant equal to the fraction of leechers uploading chunks,

γ: leaving rate for seeds.

They define the following transition rates, where $q((x,y),(x',y'))$ is the transition rate from state (x,y) to state (x',y'), for states in $\mathbb{N} \times \mathbb{N}^+$, where x is the number of leechers and y is the number of seeds:

$q((x,y),(x+1,y)) = \lambda$ (arrival of a new peer),

$q((x,y),(x-1,y+1)) = \mu(\eta x + y)$ (a leecher successfully finishes downloading the file),

$q((x,y),(x,y-1)) = \gamma y$ (a seed leaves the network).

The rationale behind transition $q((x,y),(x-1,y+1)) = \mu(\eta x + y)$ is that if all leechers and seeds have the whole file the total uploading/downloading rate is $\mu(x+y)$. The factor η is then introduced in order to take into account leechers that do not have the whole file, which are in general less efficient than seeds as servers.

Once stated the Markov chain model in [YdV04] the stationary distribution of this processes is computed numerically, and the analysis considers different metrics of performance, as throughput and delay. Results in transient and stationary regime are compared with traffic measurements. In addition the authors present incentives to users in order to improve the performance.

1.1.2 A deterministic fluid model

Now we describe the dterministic fluid model proposed by Qiu and Srikant in [QS04], based on the stochastic one of [YdV04]. In this model the evolution of leechers and seeds is described by an ordinary differential equation. The fluid model also considers two aspects that are not discussed in [YdV04]: the first one is that leechers may leave the system before finishing their download (this is called an abandon) and the second one is that a capacity restriction related to the time needed to finish a download is introduced. This time depends on the uploading capacity of peers (as in [YdV04]) but it also depends on their downloading capacity. This restriction makes the model a switched linear system.

The fluid model is built as follows:

$x(t)$: number of leechers at time t, seen as a real number,

$y(t)$: number of seeds at time t, seen as a real number,

λ: arrival rate of peers,

μ: uploading rate for each peer,

c: downloading rate for each peer,

θ: individual leaving rate for leechers,

γ: individual leaving rate for seeds,

$\eta \in [0, 1]$: efficiency factor, that takes into account the efficiency of the file sharing mechanism ([QS04] provides a detailed analysis of η),

The maximal total uploading rate is $\mu(\eta x + y)$, the maximal total downloading rate is cx, and the restriction may be in the upload or in the download. The effective downloading rate is thus $\min(cx, \mu(\eta x + y))$. The evolution of the number of leechers and seeds is described by the following ODE:

$$\begin{cases} x' = \lambda - \min(cx, \mu(\eta x + y)) - \theta x, \\ y' = \min(cx, \mu(\eta x + y)) - \gamma y. \end{cases} \tag{1.1}$$

There is a line $y = (c/\mu - \eta)x$ where the behavior of the system changes because of the term $\min(cx, \mu(\eta x+y))$, dividing the state space in two zones.

The authors state that the average number of leechers and seeds in stationary regime are the values of the ODE's fixed point (x^*, y^*), where they study local stability, and derive the average downloading time from an approximation of Little's law. They also show a good fitting with simulations of the BitTorrent protocol and with real traces, specially when the arrival rate λ is high. In a posterior paper global stability is proved by Qiu and Sang [QS08]. In addition they give a detailed analysis of the efficiency factor η and prove that in most systems this factor is close to 1. Their argument relies on the rarest first policy, that guarantees that chunks are distributed in a uniform fashion, and in the presence of a large number of chunks for each target file. Finally the variability is modeled by an Ornstein-Ulhenbeck process centered at the deterministic system. The incentive policy is studied by optimizing the selfish behavior of peers, obtaining a Nash equilibrium.

1.1.3 High and low tolerance leechers

Based on [QS04], Rivero and Rubino in [RAR10] consider a fluid model for a BitTorrent network with different classes of peers. There are two classes of leechers: high tolerance leechers and low tolerance ones.

The parameters are the following:

$x_a(t)$: number of high tolerance leechers at time t,

$x_b(t)$: number of low tolerance leechers at time t,

$y(t)$: number of seeds at time t,

λ_a: arrival rate of high tolerance leechers,

λ_b: arrival rate of low tolerance leechers,

μ: uploading rate for each peer,

c: downloading rate for each peer,

θ_a: leaving rate for high tolerance leechers,

θ_b: leaving rate for low tolerance leechers, with $\theta_b > \theta_a$,

γ: leaving rate for seeds,

the efficiency factor is $\eta = 1$,

The fluid model for this system is:

$$\begin{cases} x_a' &= \lambda_a - \theta_a x_a - u_a, \\ x_b' &= \lambda_b - \theta_b x_b - u_b, \\ y' &= u_a + u_b - \gamma y, \end{cases} \tag{1.2}$$

where

$$\begin{aligned} u_a &= \min\left(cx_a, \mu(x+y)\frac{x_a}{x}\right), \\ u_b &= \min\left(cx_b, \mu(x+y)\frac{x_b}{x}\right), \\ x &= x_a + x_b. \end{aligned}$$

From this equation there are also two different zones, divided by the plane $y = (c/\mu - 1)(x_a + x_b)$, in the x_a, x_b, y space, again due to the restriction in uploading or downloading capacities.

The authors propose a strategy to improve the performance by giving priority to peers that will probably stay more time in the system, specially in bad resource conditions. In order to define the policy, the space is divided by planes in three zones, according to the capacity. For each zone a server policy is defined, giving priority to high tolerance leechers when capacity becomes too low. A fluid model considering the different policies in each zone is stated. The analysis of the priority policy, compared with the non priority one, is obtained by computing the fixed points. The priorities policies for low tolerance and high tolerance leechers are adapted and exploited also for streaming P2P applications in [ERAR13, RARTM13].

1.1.4 More BitTorrent models

In this work we consider stochastic and fluid models that allow to analyze the number of peers in the system.

Another approach to study BitTorrent systems is to consider the number and type of chunks that each peer possesses as for instance in [MV05]. The authors study what they call a coupon replication system, that models a file sharing BitTorrent-like mechanism. Their model considers many users, each one aiming to complete a collection of coupons. At each time two users meet and obtain one missing coupon from the other by replication, if they do not have the same coupons. The model is motivated by the BitTorrent mechanism, where each chunk is a coupon. Results from [EK86] are used to prove the approximation by an asymptotic deterministic model for the number of coupons that each user holds, when the number of coupons goes to infinity. A closed form formula is obtained only in some particular cases.

There are also papers considering epidemiological models of BitTorrent-like systems as [KKS07, KKS09], where fluid limits and asymptotic distributions are obtained. The number of chunks is a crucial parameter for the model, as transition depends on this quantity. The authors analyze closed and open systems and also the presence of incentives. Incentives are based on splitting the file into more chunks, with the aim that pairs remain active in the system. Although diffusion approximation is mentioned, it is not considered further on the analysis.

Another way of managing the file sharing mechanisms is to consider a fluid model with a PDE that takes into account the residual downloading times. Such models are developed in [FKP11, FP12, PF12], where the authors combine the fluid limit approach via PDEs with queuing models. Stability of such systems is analyzed, using control theory tools.

A different approach appears in [BMNV13], that considers fluid limits combined with stochastic geometry taking into account the network topology. With these tools the authors prove scalability of P2P networks and obtain closed formulas for different performance metrics.

1.2 A stochastic model for BitTorrent

In this section we introduce a stochastic model that describes the number of leechers and seeds in a BitTorrent-like system. From this microscopic description we construct a sequence of processes that converges to a deterministic limit, governed by the ODE stated by Qiu and Srikant in [QS04]. We also formulate a stochastic model for the number of low tolerance leechers, high tolerance leechers and seeds, for the system studied by Rivero and Rubino in [RAR10] and the corresponding sequence of processes that converges to the ODE presented in their work. The case with a single type of leechers is a particular case of this one. Last, we present a sequence of processes as the scaled version of the model by Yang and de Veciana [YdV04] and obtain its fluid limit.

1.2.1 One class of leechers

We consider a sequence, indexed by N that is the scaling factor for the limit results, of two-dimensional continuous time Markov chains where the components are the number of leechers and the number of seeds. We describe the model's dynamics as follows.

$\widetilde{X}^N(t)$: number of leechers at time t,

$\widetilde{Y}^N(t)$: number of seeds at time t. We also assume that there is an additional fixed seed (the total number of seeds is thus $\widetilde{Y}^N(t)+1$), so that the system never dies,

$N\lambda$: arrival rate for peers (leechers),

μ: uploading rate for each peer,

c: downloading rate for each peer,

θ: abandonment rate for a leecher before completing its download,

γ: leaving rate for seeds,

$\eta \in [0,1]$: efficiency factor,

a leecher becomes a seed with rate

$$\min\left(c\widetilde{X}^N(t), \mu\left(\eta\widetilde{X}^N(t)+\widetilde{Y}^N(t)+1\right)\right).$$

The inclusion of a fixed seed simplifies the model in the sense that the system never dies, and we do not have to consider the extinction of seeds for the stochastic model. In addition, in the following sections we prove that this hypothesis does not affect the limit of the sequence, if there is not a fixed seed the asymptotic behavior is the same. The introduction of the fixed seed, or a number of fixed seeds that does not scale with N, only changes some technicalities of the proofs concerning the asymptotic behavior. This model corresponds to a stochastic description associated with the model in [QS04]. We prove in the following section that the ODE in [QS04] is the fluid limit of this sequence of processes as N goes to infinity.

1.2.2 Two classes of leechers

For two classes of leechers we consider a sequence of three-dimensional continuous time Markov chains where the components are the number of low tolerance leechers, the number of high tolerance leechers, and the number of seeds. The notation in this case is the following.

$\widetilde{X}_a^N(t)$: number of leechers of type a at time t,

$\widetilde{X}_b^N(t)$: number of leechers of type b at time t,

$\widetilde{Y}^N(t)$: number of seeds at time t. We also assume that there is an additional fixed seed (the total number of seeds is thus $\widetilde{Y}^N(t)+1$), so that the system never dies,

$N\lambda_a$, $N\lambda_b$: arrival rates for peers (leechers) of type a and b respectively,

μ: uploading rate for each peer,

c: downloading rate for each peer,

θ_a, θ_b: abandonment rates for a leecher before completing its download, for type a and b respectively, with $\theta_b > \theta_a$ (leechers of type a are high tolerance peers and leechers of type b are low tolerance ones),

γ: leaving rate for seeds,

$\eta \in [0,1]$: efficiency factor,

a leecher of type a becomes a seed with rate

$$\min\left(c\widetilde{X}_a^N(t), \mu\eta\widetilde{X}_a^N(t) + \mu\left(\widetilde{Y}^N(t)+1\right)\frac{\widetilde{X}_a^N(t)}{\widetilde{X}^N(t)}\right),$$

and a leecher of type b becomes a seed with rate

$$\min\left(c\widetilde{X}_b^N(t), \mu\eta\widetilde{X}_b^N(t) + \mu\left(\widetilde{Y}^N(t)+1\right)\frac{\widetilde{X}_b^N(t)}{\widetilde{X}^N(t)}\right).$$

The last transition rates model the system in the following way. The total uploading capacity is $\mu(\eta\widetilde{X}^N(t) + \widetilde{Y}^N(t) + 1)$ and this capacity is shared proportionally for high and low tolerance leechers, so for each class the capacity is multiplied by its proportion, respectively $\frac{\widetilde{X}_a^N(t)}{\widetilde{X}^N(t)}$ and $\frac{\widetilde{X}_b^N(t)}{\widetilde{X}^N(t)}$.

This model corresponds to the proposal in [RAR10] and the ODE in [RAR10] is the fluid limit of this sequence of processes as N goes to infinity. Let us note that the efficiency factor η may be taken to be equal to 1 as in [RAR10].

The model with a single class of leechers is a particular case of the previous one, considering $\lambda_a = \lambda_b = \lambda$, $\theta_a = \theta_b = \theta$, $\widetilde{X}_a^N + \widetilde{X}_b^N = \widetilde{X}^N$ and for one class of leechers we have a stochastic description associated with the fluid model proposed by Qiu and Srikant in [QS04].

1.2.3 Revisiting the Markov chain model

We want to compare the models proposed by Yang and de Veciana [YdV04] and by Qiu and Srikant [QS04], between them and with our model. For this purpose, we propose a natural scaling for the model of Yang and de

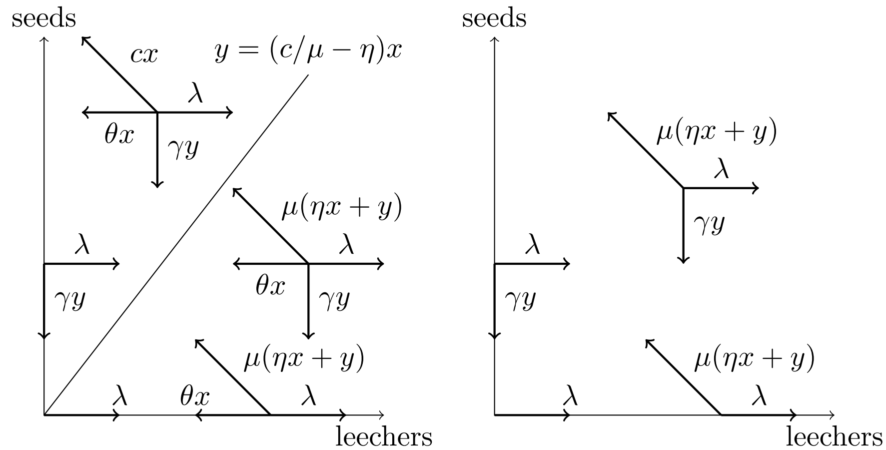

Figure 1.1: Transitions for P2P models. Left: transitions for the case of only one type of leechers (Subsection 1.2.1). Right: transitions for the model without abandons and without downloading capacity restriction (Subsection 1.2.3). In both pictures we do not consider the presence of the fixed seed, as it is negligible for large N.

Veciana, that consists in considering $(\widetilde{X}^N(t), \widetilde{Y}^N(t))$ the number of leechers and seeds at time t, with the following transitions:

a leecher arrives with rate $N\lambda$,

a leecher becomes seed with rate $\mu\left(\eta\widetilde{X}^N(t) + \widetilde{Y}^N(t) + 1\right)$,

a seed leaves the system with rate $\gamma\widetilde{Y}^N(t)$.

Later we analyze this model, that is very close to the one described in Subsection 1.2.1 with a single class of leechers, when c is large, but has discontinuous rates that change the characteristics of the limit.

1.3 Numerical examples

In this section we want to illustrate the models under consideration. Our objective of study is the scaled number of peers, we consider $\left(X^N(t), Y^N(t)\right) = \frac{1}{N}\left(\widetilde{X}^N(t), \widetilde{Y}^N(t)\right)$, the scaled number of leechers and seeds, when there is a single class of leechers.

In Figure 1.2 we work with the model described in Subsection 1.2.1 with a single type of leechers, and we show the evolution (a trajectory) of the scaled number of leechers and seeds $\big(X^N(t), Y^N(t)\big)$, for the parameter set in Table 1.1, with $\big(X^N(0), Y^N(0)\big) = (0, 0)$.

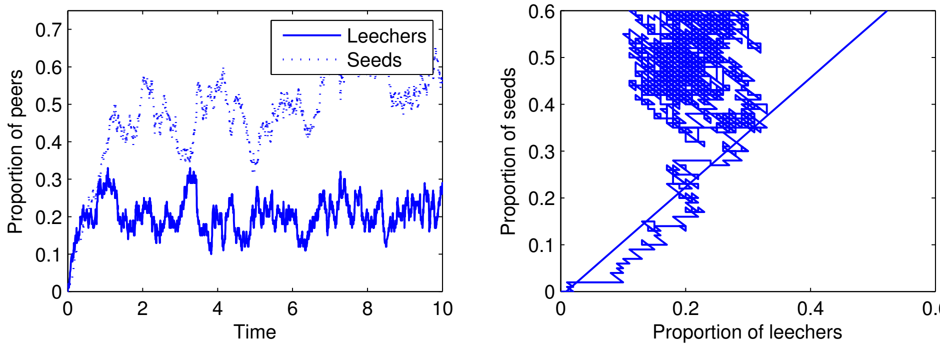

Figure 1.2: One class of leechers with downloading constraint. Evolution with time of the scaled number of leechers and seeds. (Model description in Subsection 1.2.1 and parameters in Table 1.1). The line divides the plane into two zones with different behaviors, considering constraints in uploading or in downloading (as in this example) capacity.

N	λ	μ	c	γ	θ	η
100	1	3	5	2	0.01	0.5

Table 1.1: One class of leechers with downloading constraint. Parameters for Figure 1.2.

Let us compare our model derived from the paper by Qiu and Srikant in the case of a single type of leechers with the BitTorrent Markov model of Yang and de Veciana and its scaled version described in Subsection 1.2.3. In [YdV04] the authors do not take into account the restriction in uploading or downloading capacity (their restriction only considers the uploading capacity) and the fact that peers may abandon the system before finishing their download. In our model (Subsection1.2.1) transition rates are continuous, whereas in [YdV04] (and its scaled version in Subsection 1.1.1) there are discontinuities when the number of leechers become zero. We will see later that this introduces difficulties when considering the fluid limit.

To get some intuition on both models we compare them, using similar parameters' values. We expect that without abandons and for very large values of the downloading capacity, with the other parameters equal, both models have similar behaviors, as in this case the restriction for model in Subsection 1.2.1 is always in the uploading capacity, as for the model in Subsection 1.2.3 that does not consider other capacity restrictions.

In Figure 1.3 and Figure 1.4, we show in each row, for large N, at the left: evolution with time of proportion of leechers and seeds, and at the right: relationship between proportion of leechers and seeds as time evolves. We notice that in the four cases we identify regions where such trajectories remains for large values of t.

In Figure 1.3, first row, we show a trajectory with one single type of leechers for the model in Subsection 1.2.1, for very large values of the downloading capacity (restriction is in uploading capacity). In the second row we consider the model of Subsection 1.2.3. The parameter set is given in Table 1.2. There is a high coincidence for the model with switching capacity constraints for large values of c and the model with only uploading capacity constraints when there are many leechers and few seeds.

In the first row of Figure 1.4, we show a trajectory with one single type of leechers for the model in Subsection 1.2.1, for very large values of the downloading capacity (restriction is in uploading capacity). In the second row we consider the model of Subsection 1.2.3. The parameter set is given in Table 1.3. However, in both rows the behavior of the system is quite different, as in the second row, for large values of time, the system remains over the y axis.

When there are few leechers and many seeds the behavior seems different, as in first and second rows in Figure 1.4. The number of leechers remains more time at zero for the case in the second row, and this situation is avoided in the case of the first row because the switching in transition rates forces the system to remain below the line that divides the plane into two zones. Because of this, trajectories in the first row cannot stay so long in the y axis, as it occurs in the example appearing in the second row.

	N	λ	μ	c	γ	θ	η
Model 1.2.1	100	1	3	100	10	0	1
Model 1.2.3	100	1	3	–	10	–	1

Table 1.2: Comparison between models with and without constraints with many leechers. Numerical values of the parameters for Figure 1.3.

	N	λ	μ	c	γ	θ	η
Model 1.2.1	100	1	3	100	3.15	0	1
Model 1.2.3	100	1	3	–	3.15	–	1

Table 1.3: Comparison between Markov chains with and without constraints with few leechers. Numerical values of the parameters for Figure 1.4.

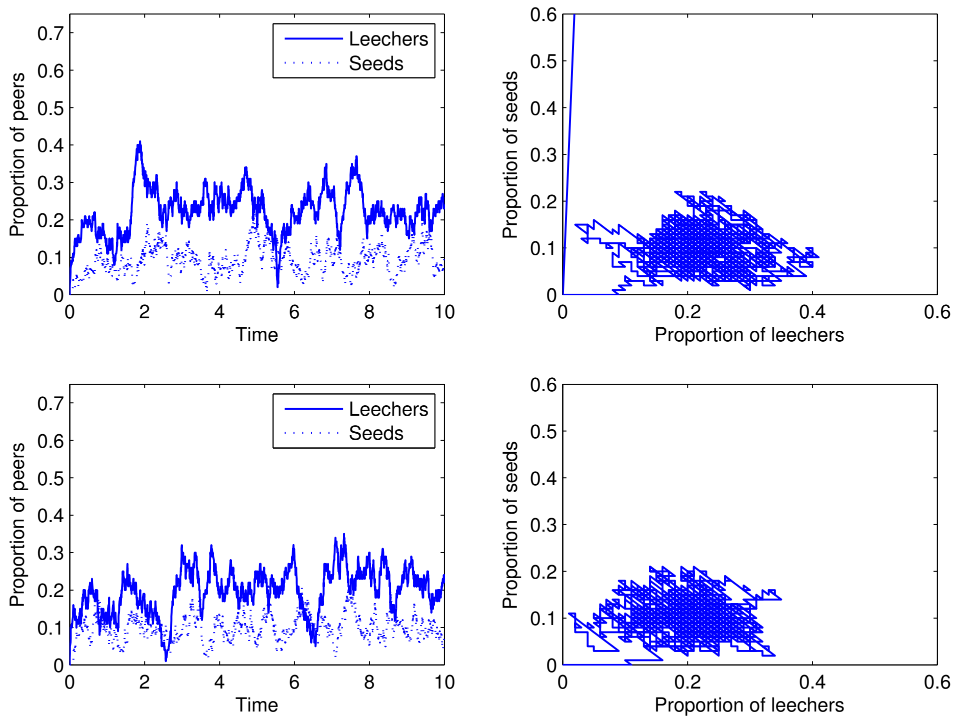

Figure 1.3: Comparison between Markov chains with and without constraints with many leechers and few seeds. Evolution of the scaled number of leechers and seeds. First row: model in Subsection 1.2.1; last row: model in Subsection 1.2.3. Parameters are in Table 1.2).

1.4 Fluid limits

In this section we present the results about fluid limits for the BitTorrent model.

We consider the model described in Subsection 1.2.1, in the the case of one class of leechers. For this model we find a fluid limit governed by the ODE stated in [QS04], where convergence is almost sure in $[0, T]$, using

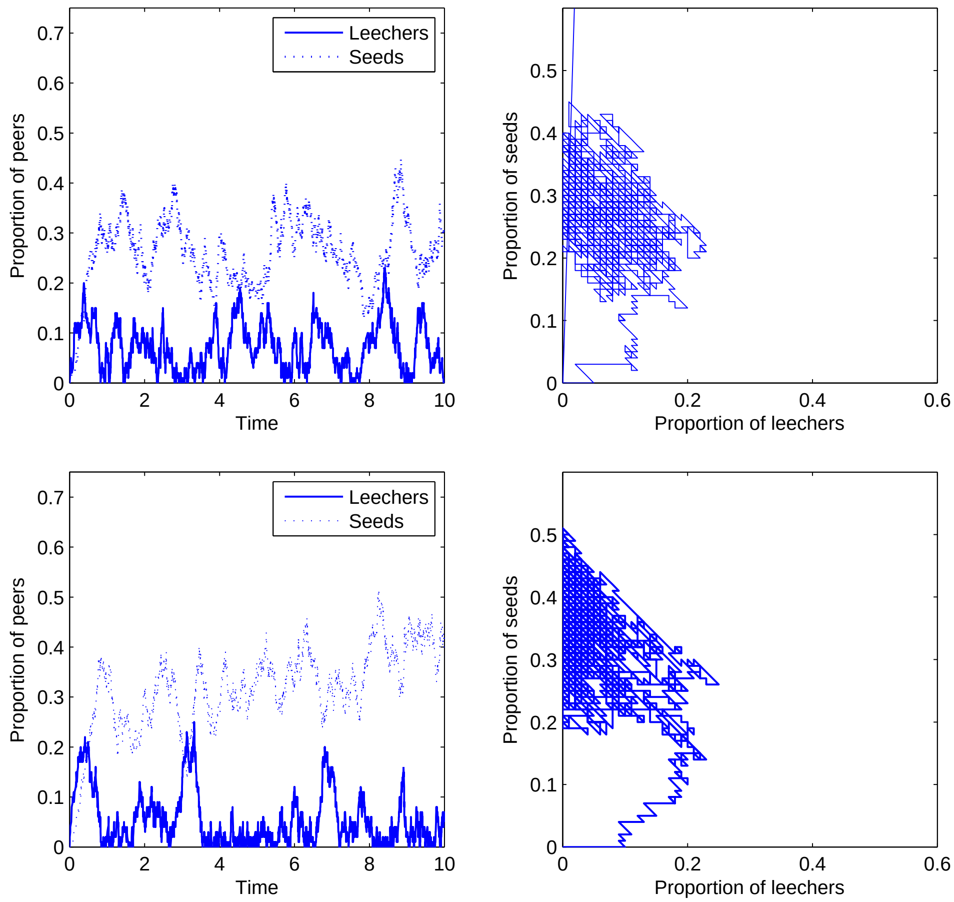

Figure 1.4: Comparison between Markov chains with and without constraints with few leechers and many seeds. Evolution of the scaled number of leechers and seeds. First row: model in Subsection 1.2.1; last row: model in Subsection 1.2.3. Parameters are in Table 1.3).

Kurtz's Theorem. Then we prove that there exists a stationary distribution for each N, by means of a Lyapunov function, and the convergence in probability for the processes under the stationary distribution to the ODE's fixed point. We recall here results about the behavior of the ODE, that is fully analyzed in [QS04, QS08]. These results can be obtained by different methods, and later in this work we will use other techniques in order to prove similar results in other contexts.

Results and proofs in the case of low and high tolerance leechers are straightforward from these ones. In the case with one class of leechers the notation is simpler and it is also more suitable to show numerical results.

We also describe the fluid limit for the sequence of processes presented in Subsection 1.1.1 following the model by Yang and de Veciana. In this case the proof is different from the results described above, due to discontinuities in the transition rates, that lead to non-differentiable fluid limits. This topic will be addressed further in this work, with different approaches. In this section we consider for this and all the other issues addressed the same approach that we presented in [AMR11].

1.4.1 Convergence for finite time

Theorem 1.1. *Let* $\left(\widetilde{X}^N(t), \widetilde{Y}^N(t)\right)$ *as in Subsection 1.2.1 Consider*

$$\left(X^N(t), Y^N(t)\right) = \frac{1}{N}\left(\widetilde{X}^N(t), \widetilde{Y}^N(t)\right)$$

and (x, y) *the solution to equation* (1.1) *with initial condition* $(x(0), y(0))$*. If*

$$\lim_{N\to\infty}(X^N(0), Y^N(0)) = (x(0), y(0))$$

then, for all $T > 0$*,*

$$\lim_{N\to\infty} \sup_{t\in[0,T]} \left|\left(X^N(t), Y^N(t)\right) - (x(t), y(t))\right| = 0 \text{ a.s.},$$

where a.s. means almost sure convergence.

Proof. The result follows directly from Kurtz's Theorem (Theorem 2.1, p. 456) in [EK86] (recalled in Appendix A), so we need to show that the hypotheses of Kurtz's Theorem are verified, and for this purpose we look at transitions rates.

The possible transitions in the N-th model $\left(\widetilde{X}^N, \widetilde{Y}^N\right)$, starting from state $\left(\widetilde{X}^N(t), \widetilde{Y}^N(t)\right)$ are the following:

a leecher arrives with rate $N\lambda$,

a leecher becomes seed with rate

$$N \min\left(cX^N(t), \mu\left(\eta X^N(t) + Y^N(t) + \frac{1}{N}\right)\right),$$

a leecher aborts before downloading with rate $N\theta X^N(t)$, a seed leaves the system with rate $N\gamma Y^N(t)$.

So, $(\widetilde{X}^N(t), \widetilde{Y}^N(t))$ is a jump Markov process with transition rates of the form

$$N\left[\beta_\ell\left(X^N(t), Y^N(t)\right) + O\left(\frac{1}{N}\right)\right]$$

for $\ell \in \mathbb{Z}^2$ (ℓ represents a possible transition). As in all the transitions β_ℓ is bounded and Lipschitz on compact subsets, the result follows directly from Kurtz's Theorem. □

The previous theorem is illustrated in Figure 1.5. We can see from the picture that, for large time values, the number of leechers is around the ODE's fixed point. We analyze this in Theorem 1.5.

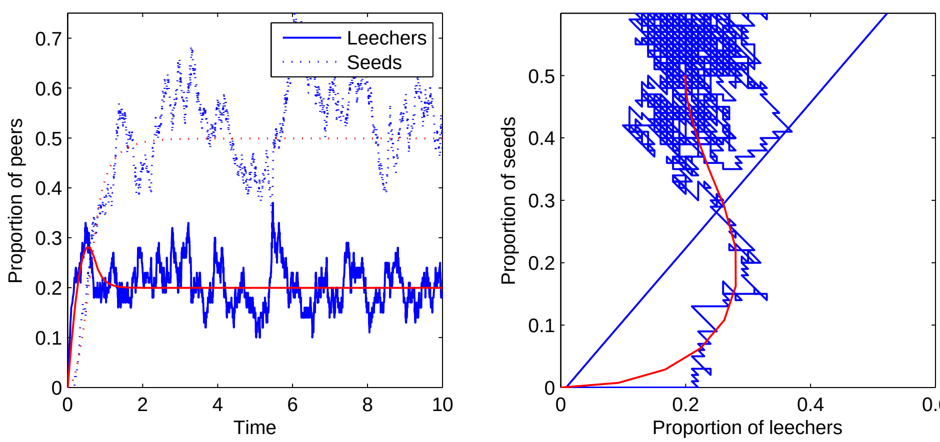

Figure 1.5: Markov chain for one class of leechers with downloading constraint and ODE. In the left we show the simulation of one trajectory of the scaled Markov chain (number of leechers and seeds) for large N and the trajectory of the ODE. In the right we show for the same simulation the evolution on the plane of the Markov chain and the ODE.(Model description in Subsection 1.2.1 and parameters in first row of Table 1.1.)

Remark 1.2. A result in the sense of Theorem 1.1 is also valid for the stochastic model associated with the system in [RAR10], as it verifies the same hypotheses and it is also valid for the priority scheme proposed in the same paper. The Markov chain that represents the priority scheme is of the same type as the previous ones, so the fluid approximation also holds.

Remark 1.3. Now we analyze the model presented in Subsection 1.2.3, as it is intrinsically different and it is not under the hypotheses of Kurtz' s Theorem. As before, let $(\widetilde{X}^N(t), \widetilde{Y}^N(t))$ be the number of leechers and seeds at time t and $(X^N(t), Y^N(t)) = \frac{1}{N}(\widetilde{X}^N(t), \widetilde{Y}^N(t))$. Consider the following transitions:

a leecher arrives with rate $N\lambda$,

a leecher becomes seed with rate $N\mu\left(X^N(t) + Y^N(t) + \frac{1}{N}\right)$,

a seed leaves the system with rate $N\gamma Y^N(t)$.

The convergence stated in Theorem 1.1 relies on the fact that transition rates from state $(\widetilde{X}^N(t), \widetilde{Y}^N(t))$ are of the form $N\left[\beta_\ell\left(X^N(t), Y^N(t)\right) + O\left(\frac{1}{N}\right)\right]$, with β a Lipschitz function. This assumption does not hold for the model in [YdV04], as transition rates are discontinuous in the boundary $x = 0$. This corresponds to a class of jump Markov processes studied in [SW95], called flat boundary processes, and in what follows we consider this framework. There are many different approaches for this kind of processes; later on this work, we will consider other ways of dealing with discontinuous transition rates following [Bor11a, GG12, JS14, Bor16].

From [SW95] (Chap. 8), in this case there is an analogous of Kurtz's Theorem, and the ODE that approximates the scaled process $(X^N(t), Y^N(t))$ is the following: if $x > 0$ or $\lambda - \mu(\eta x + y) \geq 0$,

$$\begin{cases} x' = \lambda - \mu(\eta x + y), \\ y' = \mu(\eta x + y) - \gamma y, \end{cases}$$

and if $x = 0$ and $y > \frac{\lambda}{\mu}$,

$$\begin{cases} x' = \pi_0\lambda + (1 - \pi_0)(\lambda - \mu(\eta x + y)), \\ y' = -\pi_0\gamma y + (1 - \pi_0)(\mu(\eta x + y) - \gamma y), \end{cases}$$

with

$$\pi_0 = \begin{cases} 1 - \dfrac{\lambda}{\mu y} & \text{if } y > \frac{\lambda}{\mu}, \\ 0 & \text{if } y \leq \frac{\lambda}{\mu}. \end{cases}$$

π_0 is obtained in [SW95] as the stationary distribution in the border, that is, when the system is under its stationary distribution considering two states, remaining at the border (probability π_0) and at the interior (probability $1 - \pi_0$).

Another interpretation is given by [Bor11a, Bor16], and will be addressed in more detail in this dissertation, but now we introduce it as a second explanation. We have two vector fields, given by $(\lambda-\mu(\eta x+y), \mu(\eta x+y)-\gamma y)$ that corresponds to the equation inside the first quadrant and $(\lambda, \gamma y)$ that corresponds to the equation in the border. When $(\lambda-\mu(\eta x+y), \mu(\eta x+y)-\gamma y)$ in the border points outside the first quadrant (which occurs when $y > \frac{\lambda}{\mu}$), as $(\lambda, \gamma y)$ points towards the first quadrant, the deterministic solution will stay at the border and should verify a equation given by a convex combination of the two vector fields.

Then we have the following equation describing the limit: if $x > 0$ or $y < \frac{\lambda}{\mu}$, then

$$\begin{cases} x' = \lambda - \mu(\eta x + y), \\ y' = \mu(\eta x + y) - \gamma y, \end{cases}$$

and if $x = 0$ and $y \geq \frac{\lambda}{\mu}$, then

$$\begin{cases} x' = 0, \\ y' = \lambda - \gamma y, \end{cases} .$$

The above equations show that it is possible to obtain fluid limits for this kind of models, despite discontinuities in transition rates. Convergence is illustrated in Figure 1.6. For the set the parameters chosen, trajectories stay some time in the axis, that is a typical behavior of PWSDS obtained as a solution of an ODE with discontinuous right hand side, and it is called in the literature sliding motion (see [Bor11a, GG12, JS14, Bor16] and references therein).

1.4.2 Convergence for stationary regime

Now we turn our attention to the ergodicity of $(\widetilde{X}^N(t), \widetilde{Y}^N(t))$. It seems not simple to find the stationary distribution explicitly. Classical sufficient conditions as reversibility are not verified, so we don't know if there is local balance. We prove ergodicity by using a Lyapunov function. The ergodicity result is also stated in [RAR10], as there is a Markov model that is compared with the fluid one by simulations, using queuing arguments that allow to reduce the analysis of the existence of a stationary regime to the study of a Jackson network.

Proposition 1.4. The process $(\widetilde{X}^N(t), \widetilde{Y}^N(t))$ is ergodic for each N.

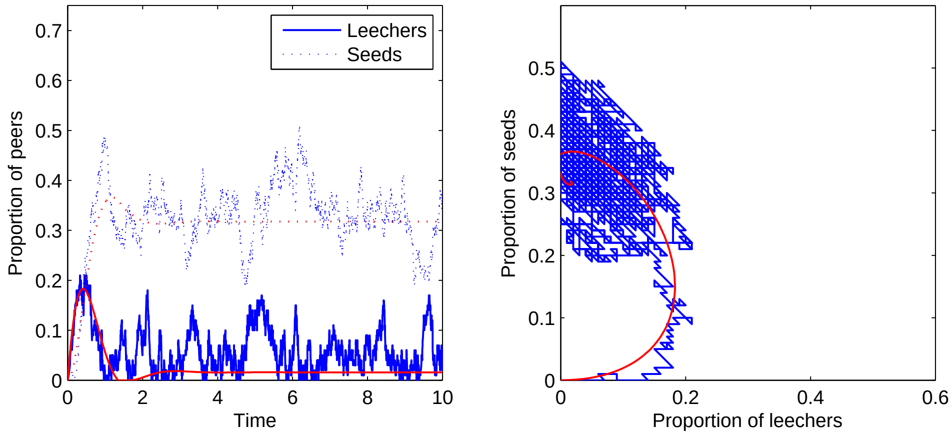

Figure 1.6: Markov chain for one class of leechers without constraints and ODE. In the left we show the simulation of one trajectory of the scaled Markov chain (number of leechers and seeds) for large N and the trajectory of the ODE. In the right we show for the same simulation the evolution on the plane of the Markov chain and the ODE.(Model description in Subsection 1.2.3 and parameters in second row of Table 1.3.)

Proof. The proof is based on [Rob03] (Proposition 8.14, p. 225), considering Lyapunov functions. $f(x, y) = x + y$ is a Lyapunov function for $(\widetilde{X}^N(t), \widetilde{Y}^N(t))$. We must verify that there exist K and h such that the following conditions hold:

1. for $f(x, y) > K$, $Q(f)(x, y) \leq -h$, with

$$Q(f)(x,y) = \sum_{l \in \mathbb{Z}^2, l \neq 0} q\left((x,y),(x,y)+l\right)\left[f((x,y)+l) - f(x,y)\right]$$

   $(q((x, y), (x, y) + l)$ is the transition rate from (x, y) to $(x, y) + l)$;

2. the random variables

$$\sup\{f(\widetilde{X}^N(s), \widetilde{Y}^N(s)) : s \leq 1\},$$

$$\int_0^1 |Q(f)(\widetilde{X}^N(s), \widetilde{Y}^N(s))|\, ds$$

   are integrable;

3. $F = \{(x, y) : f(x, y) \leq K\}$ is finite.

These assumptions imply that the process is ergodic. Let us verify them:

1. $Q(f)(x,y) = \lambda N - \theta x - \gamma y \le -h$ for $x + y > K$; it suffices then to take $K \ge (\lambda N + h)/\min(\theta,\gamma)$.

2. The Poisson process $Z(s)$ with rate $N\lambda$ is an upper bound of the function $f(\widetilde{X}^N(s), \widetilde{Y}^N(s)) = \widetilde{X}^N(s) + \widetilde{Y}^N(s)$, and it is integrable on each bounded interval. Analogously $\lambda N + \max(\theta,\gamma) Z(s)$ is an upper bound of $|Q(f)(\widetilde{X}^N(s), \widetilde{Y}^N(s))|$; the integral of the former is thus bounded by $\lambda N + \max(\theta,\gamma)\int_0^1 Z(s)\,ds$ and it is then integrable.

3. Immediate.

From the previous assumptions $(\widetilde{X}^N(t), \widetilde{Y}^N(t))$ is ergodic for each N. □

The same result holds for the stochastic model considering two classes of leechers described above. In that case a Lyapunov function is $f(x_a, x_b, y) = x_a + x_b + y$. The assumptions are verified as above. It follows from noticing again that the only possible transition away from a region $\{(x_a, x_b, y) : x_a + x_b + y \le K\}$ is when a new peer arrives. As arrivals follow Poisson processes, the hypotheses about finite expectation hold.

In what follows we prove the convergence of the stationary regime to the ODE's fixed point. As the process $(X^N(t), Y^N(t))$ converges in bounded intervals and has a stationary distribution, one can expect that the stationary distribution converges to the ODE's fixed point. This result is used for the analysis in [QS04], and in different contexts in other works (see for example [MV05]), sometimes without a detailed proof. In [QS04] it is proven that the ODE has an unique fixed point

$$(x^*, y^*) = \left(\frac{\lambda}{\beta\left(1 + \frac{\theta}{\beta}\right)}, \frac{\lambda}{\gamma\left(1 + \frac{\theta}{\beta}\right)} \right),$$

with

$$\frac{1}{\beta} = \max\left\{ \frac{1}{c}, \frac{1}{\mu} - \frac{1}{\gamma} \right\}.$$

and that the system is locally stable. The work from Qiu and Sang [QS08] is devoted to the analysis of equation (1.1), and they prove that the unique fixed point is a global attractor. We show in Figure 1.7 the vector field associated with equation (1.1).

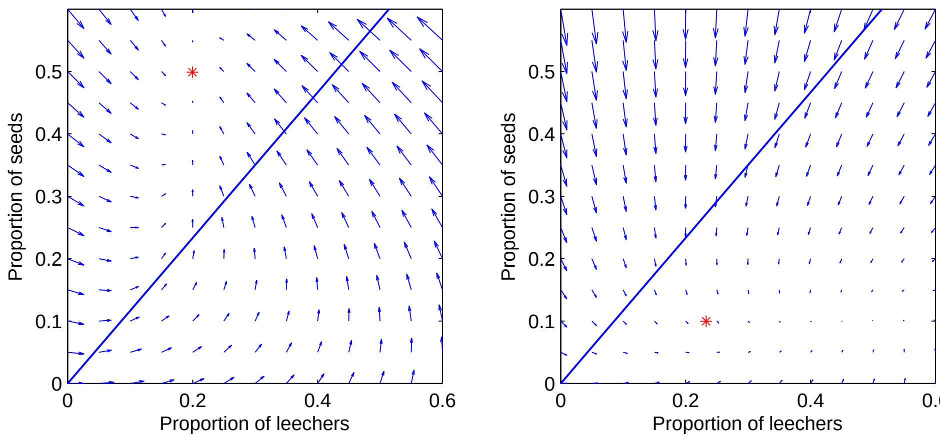

Figure 1.7: Vector fields and fixed points for equation (1.1) for different parameters sets. The parameters' values are in Table 1.4.)

λ	μ	c	γ	θ	η
1	3	5	2	0.01	0.5
1	3	10	2	0.01	0.5

Table 1.4: Parameters for vector fields for equation (1.1) in Figure 1.7.

Theorem 1.5. *Let $\left(X^N(\infty), Y^N(\infty)\right)$ be the scaled number of leechers and seeds in stationary regime. Let (x^*, y^*) be the fixed point of* (1.1)*. Then*

$$\lim_{N\to\infty} \left(X^N(\infty), Y^N(\infty)\right) = (x^*, y^*)$$

in probability.

Proof. Let $\mu^N(t)$ be the distribution of $\left(X^N(t), Y^N(t)\right)$ and let $\pi^N(\infty)$ be the stationary distribution of the process (we know from Proposition 1.4 that there exists a unique stationary distribution for each N). We will use for our proof Theorem 6.89, p. 165 in [SW95]. This theorem assures that under a set of hypotheses that will be verified, if (x^*, y^*) is a global attractor then

$$\lim_{N\to\infty} \int_{B_\varepsilon(q)} d\pi^N(\infty) = 1,$$

with $(x^*, y^*) = q$ and $B_\varepsilon(q) = \left\{y \in \mathbb{R}^2 : |y - q| < \varepsilon\right\}$, which implies that

$$\lim_{N\to\infty} \left(X^N(\infty), Y^N(\infty)\right) = (x^*, y^*)$$

in probability. To apply the result we must verify that:

1. the jumps of the Markov process take integer values in each direction,
2. the rates β_ℓ are uniformly Lipschitz continuous in a neighborhood of (x^*, y^*),
3. the process is positive recurrent,
4. if $\tau_\varepsilon(N) = \inf\left\{t : \left|\left(X^N(t), Y^N(t)\right) - (x^*, y^*)\right| < \varepsilon\right\}$, then for each K, ε and for all N, there exists a constant $C_{\varepsilon,M}$ such that

$$\sup_{|p-q|\le M} E_p\left[\tau_\varepsilon(N)\right] \le C_{\varepsilon,M} < \infty,$$

with $p = (x(0), y(0))$ and E_p the expected value starting from p.

The first three assumptions are immediately verified (the third one arises from Proposition 1.4). So, we focus on the fourth one. Also from [SW95] (Lemma 6.32 p. 143) the distribution of $\tau_\varepsilon(N)$ has geometric tails for large N, that is, there is a $T(\varepsilon) < \infty$ and a constant $C_0(\varepsilon)$ such that

$$P_p(\tau_\varepsilon(N) > kT) \le e^{-NC_0(\varepsilon)k}.$$

This implies the bound for $E_p\left[\tau_\varepsilon(N)\right]$ and thus completes the proof. □

The convergence for the stationary distribution to the ODE's fixed points is a widely discussed topic. The authors of [BLB08] prove this convergence in the case of the occupation measure of a system with N individuals and a finite state space. Our problem differs from that situation because of the compactness of the state space. However, our proof and the proof in [BLB08] rely in large deviations results. The proof in [BLB08] is based on [Ben99], where a very general result (considering the case with multiple invariant distributions and a much more complex asymptotic behavior for the ODE) is proven using large deviations arguments together with dynamical systems ones. In [BLB08] it is also discussed why the existence of a unique fixed point does not guarantee the convergence of a sequence of invariant distributions. It shows examples where there is only one fixed point but the support of accumulation points of invariant distributions lies on a subset that is a limit cycle for the ODE. In order to avoid the problem of proving asymptotic stability, [LB10b] presents a very general result of convergence for the stationary distribution when there is a unique fixed point in case of reversible processes, a strong assumption that is not valid in our model.

The convergence for the stationary distribution of the occupation measure is also discussed in [BMP08], in a more general framework (denumerable state spaces). The proof there is strongly related with the mean field decoupling assumption (the asymptotic independence and the convergence of the stationary distributions are proved together).

For the model in [RAR10] we do not have a proof of global stability of ODE's fixed point, so we cannot yet extend the previous theorem for that case. It is only observed in simulations in [RAR10] that the stationary regime converges to the fixed point of the associated ODEs, both in the priority and non priority schemes.

1.5 Asymptotic distribution

As the natural relation between the Law of Large Numbers and the Central Limit Theorem, once we have stated a fluid limit our aim is to find the velocity of that convergence. Thus we study the asymptotic distribution of the difference between the stochastic process and its fluid limit. We want to find a suitable scale for that difference in order to have a limit law. First, for a scaling as $\sqrt{N}$, we derive a Gaussian approximation for the transient regime in compact time intervals. Then we extend this approximation for the whole real line and we obtain the asymptotic distribution in stationary regime. This approximation describes in a precise way the system behavior for large values of N, simultaneously for all t, providing confidence intervals for the number of leechers and seeds. We use $\Rightarrow_N$ for convergence in distribution with $N \to +\infty$.

1.5.1 Gaussian approximation for finite time

Theorem 1.6. *Consider* $\left(X^N, Y^N\right)$ *as in Section 1.2 and let* (x, y) *be the solution to equation* (1.1) *with initial condition* $(x(0), y(0))$. *Let*

$$V^N = \sqrt{N}[(X^N, Y^N) - (x, y)].$$

If

$$\lim_{N\to\infty} \sqrt{N}\left[(X^N(0), Y^N(0)) - (x(0), y(0))\right] = \lim_{N\to\infty} V^N(0) = V(0)$$

in probability, with $V(0)$ *deterministic, then,*

$$\sqrt{N}\left[\left(X^N, Y^N\right) - (x, y)\right] = V^N \Rightarrow_N V,$$

where V is a Gaussian process with covariance matrix given by

$$Cov\left(V(t), V(r)\right) = \int_0^{t\wedge r} e^{M(t)(t-s)} G(x(s), y(s)) e^{M(r)^T(r-s)}\, ds, \qquad (1.3)$$

where

$$G(x(t), y(t)) = \begin{pmatrix} \lambda + \theta x(t) + r(x(t), y(t)) & -r(x(t), y(t)) \\ -r(x(t), y(t)) & \gamma y(t) + r(x(t), y(t)) \end{pmatrix},$$

$$M(t) = \begin{pmatrix} -(\mu\eta + \theta) & -\mu \\ \mu\eta & \mu - \gamma \end{pmatrix} \text{ if } cx(t) \geq \mu(\eta x(t) + y(t)),$$

$$M(t) = \begin{pmatrix} -(c + \theta) & 0 \\ c & -\gamma \end{pmatrix} \text{ if } cx(t) < \mu(\eta x(t) + y(t)),$$

$$r(x(t), y(t)) = \min\{cx(t), \mu(\eta x(t) + y(t))\}$$

and $M(t)^T$ denotes the transpose of $M(t)$.

Proof. The result follows as a consequence of Kurtz's Theorem (Theorem 2.3, p.458, in [EK86]. We use the explicit form of the covariance matrix presented there. The proof of that theorem relies on a representation of $V^N(t)$ and $V(t)$ by an integral involving the differential $dF(x(t), y(t))$, where F is the function defining the fluid limit, so the original theorem assumes that the transition rates evaluated at the fluid limit are C^1 functions. This assumption is not valid in our case, but there is only one t where the rate $\min\{cx(t), \mu(\eta x(t) + y(t))\}$ is not differentiable (that is when $cx(t) = \mu(\eta x(t) + y(t))$). As this happens at only one point, it does not affect the integral representation. The justification that there is only one t for which $cx(t) = \mu(\eta x(t) + y(t))$ follows from the fact that the fixed point is a global attractor, so the trajectories $(x(t), y(t))$ hit $\{(x, y) : cx = \mu(\eta x + y)\}$ only a finite number of times. □

In Figure 1.7 it can be seen that there is at most one hitting point in our case.

In Figures 1.8 and 1.9 we show, for different values of N, histograms and Q-Q plots of 100 independent samples of the scaled number of leechers $\sqrt{N}\left(X^N(t) - x(t)\right)$ and seeds $\sqrt{N}\left(Y^N(t) - y(t)\right)$ for a fixed t and in Figure 1.10 we show the 95% confidence interval for the scaled number of leechers.

[QS04] describes without a detailed proof the variability around the fluid limit (the solution to equation (1.1)). For a large arrival rate λ, the number of leechers and seeds are approximately $x(t) + \sqrt{\lambda}\widehat{x}(t)$ and $y(t) + \sqrt{\lambda}\widehat{y}(t)$, with $\widehat{x}(t)$ and $\widehat{y}(t)$ Gaussian processes (Ornstein-Uhlenbeck). In our framework

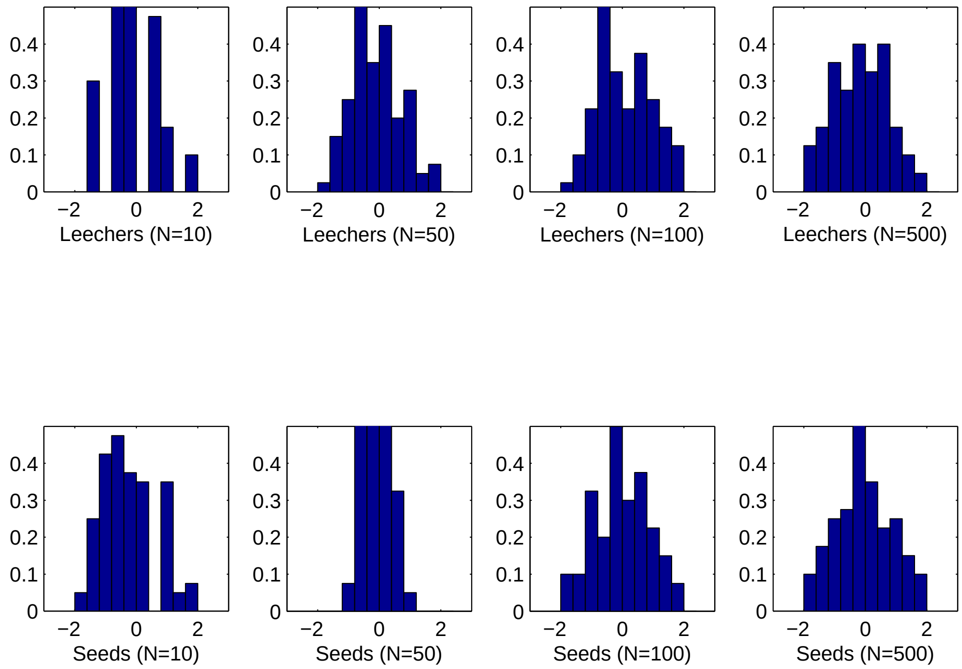

Figure 1.8: Histograms of 100 independent samples of $\sqrt{N}\left(X^N(1) - x(1)\right)$ and $\sqrt{N}\left(Y^N(1) - y(1)\right)$ for different values of N.

we have that the arrival rate is $N\lambda$ and the number of leechers and seeds are characterized by $\widetilde{X}^N(t) \approx Nx(t) + \sqrt{N}V_1(t)$ and $\widetilde{Y}^N(t) \approx Ny(t) + \sqrt{N}V_2(t)$, with $V = (V_1, V_2)$ the Gaussian process described in Theorem 1.6. We observe that the limit process $V(t)$ verifies a stochastic differential equation (see equation (2.18), p. 458 in [EK86]). The Gaussian process stated in [QS04] can be obtained from this stochastic differential equation by replacing $x(t)$ and $y(t)$ by its respective limits x^* and y^*, provided that the processes remains always in the same half-plane as the fixed point.

1.5.2 Gaussian approximation for stationary regime

We prove that the limit process $V(t)$ converges in distribution, when t goes to infinity $V(\infty)$ defined as a Gaussian variable with covariance matrix

$$\Sigma = \lim_{t\to+\infty} Cov\left(V(t), V(t)\right),$$

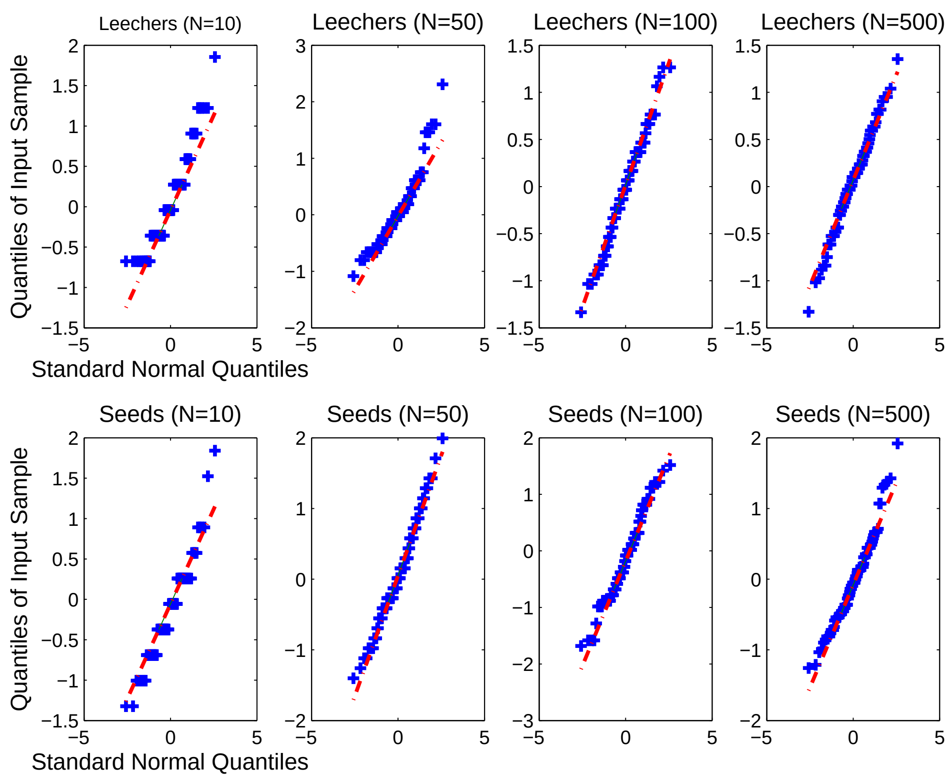

Figure 1.9: Normal Q-Q plots, obtained from 100 simulated independent samples of $\sqrt{N}\left(X^N(1) - x(1)\right)$ and $\sqrt{N}\left(Y^N(1) - y(1)\right)$ for different values of N.

defined by Equation (1.3). AS future work we want to analyze the distribution in stationary regime for

$$V^N = \sqrt{N}\left[\left(X^N, Y^N\right) - (x, y)\right]$$

, where the goal is to extend the convergence stated in Theorem 1.6, when time goes to infinity.

Proposition 1.7. Let $\Sigma(t) = Cov\left(V(t), V(t)\right)$ defined by Equation (1.3). The limit when $t \to \infty$ of matrix $\Sigma(t)$ is the covariance matrix Σ solution of the equation

$$M\Sigma + \Sigma M^T + G = 0,$$

where

$$M = \lim_{t\to\infty} M(t),$$

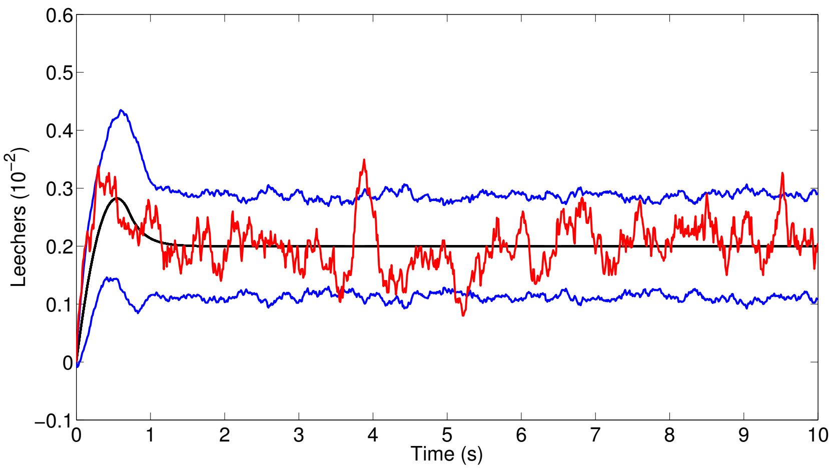

Figure 1.10: ODE trajectory, confidence interval for each t and one trajectory for the scaled number of leechers $X^N(t)$. The variance is computed for each t from 100 independent replications of the experiment (parameters in Table 1.1).

$$G = \lim_{t \to \infty} G(x(t), y(t)),$$

with $M(t)$ and $G(x(t), y(t))$ defined in Theorem 1.6.

We can write explicitly G and M in terms of the fixed point (x^*, y^*) of the ODE stated in Equation 1.1.

$$G = \begin{pmatrix} \lambda + \theta x^* + r(x^*, y^*) & -r(x^*, y^*) \\ -r(x^*, y^*) & \gamma y^* + r(x^*, y^*) \end{pmatrix},$$

$$M = \begin{pmatrix} -\mu\eta - \theta & -\mu \\ \mu\eta & \mu - \gamma \end{pmatrix} \text{ if } \frac{1}{c} \geq \frac{1}{\mu} - \frac{1}{\gamma},$$

$$M = \begin{pmatrix} -c - \theta & 0 \\ c & -\gamma \end{pmatrix} \text{ if } \frac{1}{c} < \frac{1}{\mu} - \frac{1}{\gamma},$$

$$r(x^*, y^*) = \min\{cx^*, \mu(\eta x^* + y^*)\}.$$

Proof. We consider $Cov\,(V(t), V(t))$ defined by equation (1.3)

$$\begin{aligned} \Sigma(t) &= \int_0^t e^{M(t)(t-s)} G(x(s), y(s)) e^{M(t)^T(t-s)}\, ds \\ &= e^{M(t)t} \left(\int_0^t e^{-M(t)s} G(x(s), y(s)) e^{-M(t)^T s}\, ds \right) e^{M(t)^T t}. \end{aligned}$$

We want to compute its derivative in each region of the plane divided by the line $y = (c/\mu - \eta)x$, that is assuming that $(x(t), y(t))$ is always in the same region and matrix $M(t)$ remains constant. This happens according with a time t_S that is the crossing time between regions for the solution $(x(t), y(t))$ or when $(x(t), y(t))$ has an initial condition that allows the solution to remain in the same region. As the fixed point is a global attractor, the previous assumption can be verified by considering an initial condition close enough to the fixed point. We consider $M(t) = M$, as dependence of time for $M(t)$ is only due to the region where the solution $(x(t), y(t))$ stays, and when it does not switch between regions the matrix $M(t)$ is a constant matrix. In addition we are interested only in the region that contains the fixed point. Then taking $M(t) = M$ and taking derivatives, we have

$$\begin{aligned}
\Sigma'(t) &= \left[e^{Mt}\left(\int_0^t e^{-Ms}G(x(s),y(s))e^{M^Ts}\,ds\right)e^{M^Tt}\right]' \\
&= Me^{Mt}\left(\int_0^t e^{-Ms}G(x(s),y(s))e^{-M^Ts}\,ds\right)e^{M^Tt} \\
&\quad +e^{Mt}e^{-Mt}G(x(t),y(t))e^{-M^Tt}e^{M^Tt} \\
&\quad +e^{Mt}\left(\int_0^t e^{-Ms}G(x(s),y(s))e^{M^Ts}\,ds\right)M^Te^{M^Tt} \\
&= Me^{Mt}\left(\int_0^t e^{-Ms}G(x(s),y(s))e^{M^Ts}\,ds\right)e^{M^Tt} \\
&\quad +G(x(t),y(t)) \\
&\quad +e^{Mt}\left(\int_0^t e^{-Ms}G(x(s),y(s))e^{M^Ts}\,ds\right)M^Te^{M^Tt} \\
&= M\Sigma(t) + G(x(t),y(t)) + \Sigma(t)M^T.
\end{aligned}$$

Matrix $\Sigma(t)$ verifies the matrix ODE

$$\Sigma'(t) = M\Sigma(t) + G(x(t), y(t)) + \Sigma(t)M^T,$$

whose fixed point is obtained by solving the Lyapunov equation

$$M\Sigma + G + \Sigma M^T = 0, \tag{1.4}$$

obtained from

$$\lim_{t\to+\infty} M\Sigma(t) + G(x(t), y(t)) + \Sigma(t)M^T = 0.$$

If the system $x' = Mx$ is globally asymptotically stable and $G = G^T$, then the Lyapunov Equation(1.4) has the unique solution [BEGFB94]

$$\Sigma = \int_0^{+\infty} e^{Mt} G e^{M^T t} dt. \tag{1.5}$$

□

Theorem 1.8. *For each t, $V(t)$ defined in Theorem1.6 converges in distribution, when $t \to +\infty$ to a bivariate centered Gaussian variable with covariance matrix Σ defined by Equations* (1.4) *and* (1.5)*.*

Proof. $V(t)$ is a centered process. We consider for each t the characteristic functions

$$\varphi(t) = Ee^{i\lambda^T V(t)} = e^{-\frac{1}{2}\lambda^T \Sigma(t)\lambda}.$$

We have that $\lim_{t\to+\infty} \Sigma(t) = \Sigma$ and then

$$\lim_{t\to+\infty} \varphi(t) = \lim_{t\to+\infty} Ee^{i\lambda^T V(t)} = \lim_{t\to+\infty} e^{-\frac{1}{2}\lambda^T \Sigma(t)\lambda} = e^{-\frac{1}{2}\lambda^T \Sigma\lambda}.$$

From this $V(t) \Rightarrow N(0, \Sigma)$ when $t \to +\infty$.

□

1.6 Conclusions Peer to Peer

Analyzing different models for BitTorrent-like systems in the literature we obtain that the deterministic fluid models presented in [QS04] and [RAR10] are fluid limits of stochastic models. Considering different type of leechers does not change the proof of the fluid limit, despite the ODE asymptotic behavior obtained in that case is more difficult to analyze. However, we prove there is a stationary regime for each process in the sequence and that this process converges to the ODE's fixed point. We find a Gaussian approximation in stationary regime and also obtain the limit of this Gaussian process when time goes to infinity. It is an issue for further work to prove that the process in stationary regime converges to this limit, following the lines sketched in [EK86], that derives results from [Nor77].

We also consider the fluid limit for the stochastic process presented in [YdV04], that has a piecewise smooth dynamical system as limit. The limit obtained shows sliding motion in the border of the state space, and the asymptotic distribution in that case cannot be obtained straightforward from classical results as Kurt'z theorem, so this is also a topic for future work.

Chapter 2

Machine Repairman Model

2.1 MRM models

The Machine Repairman Model (MRM) is a basic Markovian queue representing a finite number N of machines that can fail independently and, then, be repaired by a repair facility. The latter, in the basic model, is composed of a single repairing server with a waiting room for failed machines managed in FIFO (first in, first out) order, in case the repairing server is busy when units fail. In Kendall's notation, this is the $M/M/1//N$ model, specifying that lifetimes and repair times are exponentially distributed. This model is well known and widely studied in queuing theory and in many applications, as for example in telecommunications or in reliability. Almost all these studies look at the queue in equilibrium.

The model is a precursor of the development of queuing network theory, motivated first in computer science. In particular, Scherr from IBM used it in 1972 for analyzing the S360 operating system (see [Lav83]). Many extensions of the basic model have been studied, considering more than one repairing server, different queuing disciplines, and other probability distributions for the lifetime or for the repair time. We refer to [HA07] for further reference on the problem.

Looking for fluid limits is a suitable approach to repairman problems, as shown in [Kur82], where the MRM model with two repair facilities, studied by Iglehart and Lemoine in [IL73, IL74], is analyzed using these tools. In [IL73] there are N operating units subject to exponential failures with parameter λ. Failures are of type 1 (resp. 2) with probability p (resp. $1-p = q$). If failure is of type i ($i = 1, 2$) the unit goes to repair facility i that has s_N^i exponential servers, each one with exponential service rate μ_i. The goal is to study the stationary distribution when $s_i^N \sim N s_i$ as $N \to \infty$, $i = 1, 2$. The behavior of the system is characterized in terms of the parameter set that defines the model. In addition, the case with spares is presented in [IL74]. The paper's approach consists in approximating the number of units in each repair facility by binomial random variables, and then proving for them a law of large numbers and a Central Limit Theorem. Kurtz, in [Kur82] studies the same model with a fluid limit approach, proves convergence to a deterministic system, and goes a step beyond, considering the rate of this approximation through a central limit theorem-type result. The same discussion as in [IL73] in terms of the different parameter set follows from the study of the ODE's fixed point.

In this thesis we analyze a repairman problem with N working units that break randomly and independently according to a phase-type distribution. Broken units go to one repairman where the repair time also follows a phase-type distribution (that is, a $PH/PH/1//N$ model). We consider a scaled system, dividing the number of broken units and the number of working units in each phase by the total number of units N. The scaled process has a deterministic limit when N goes to infinity. The first problem that the model presents is that there are two time scales: the repairman changes its phase at a rate of order N, whereas the total scaled number of working units changes at a rate of order 1. Another problem is that transition rates are discontinuous because of idle periods at the repairman (this second issue is also present in simpler models as the $M/M/1$ and $M/M/N/N$).

In our main result we prove that the scaled Markovian multi-dimensional process describing the system dynamics converges to the solution of an ODE as $N \to \infty$. The convergence is in probability and takes place uniformly in compact time intervals (denoted *u.c.p.* convergence), and the deterministic limit, the solution to the ODE, is only picewise smooth. We analyze the properties of this limit, and we prove the convergence in probability of the system in stationary regime to the ODE's fixed point. We also find that this fixed point only depends on the repair time by its mean. As a matter of fact recall that, when in equilibrium, if the repair times are exponentially distributed, the distribution of the number of broken machines has the insensitivity property with respect to the lifetime distribution (only the latter's mean appears in the former). Although this behavior may be expected because of the scaling, it is not straightforward. In addition, the whole phase-type lifetime distribution takes part in the result. Different lifetimes with the same mean give different behaviors, both in the transient and in the stationary regime. We use phase-type distributions in both the units' lifetimes and in the repair facility, to be more general and to cover a larger number of situations.

As a general reference we refer again in this chapter to the monograph by Ethier and Kurtz [EK86] and references therein. The main approach there consists in a random change of time that allows to write the original Markov chain as a sum of independent unit Poisson processes evaluated at random times. Darling and Norris [DN08] present a survey about approximation of Markov chains by differential equations with an approach based on martingales. However, [DN08] does not deal with discontinuous transition rates. We refer to the books by Shwartz and Weiss [SW95], and Robert [Rob03] for extensive analysis of the $M/M/1$ and the $M/M/\infty$ queues, including deterministic limits, asymptotic distributions and large deviations results. In particular, in [SW95] discontinuous transition rates and different time scales are considered. The latter situation is also considered in [BKPR05] and [AEJV13]. We mostly follow the approach of [Bor11a] to deal with discontinuous transition rates, which considers hybrid limits for continuous time Markov chains with discontinuous transition rates, with examples in queuing and epidemic models. Discontinuous transition rates are also studied in [BT12, HB12]. Paper [BT12] analyzes queuing networks with batch services and batch arrivals, that lead to fluid limits represented by ODEs with discontinuous right-hand sides. Paper [HB12] models optical packet switches, where the queuing model lead to ODEs with discontinuous right-hand sides, and where they consider both exponential and phase-type distributions for packet lengths. Convergence to the fixed points is studied in several works (e.g. [SW95, BT12, BLB08]). However there are counterexamples where there is no convergence of invariant distributions to fixed points [BLB08]. There are general results with quite strong hypotheses as in [LB10b], where reversibility is assumed in order to prove convergence to the fixed point.

2.2 Stochastic model

We consider N identical units that work independently, as part of some system, that randomly fail and that get repaired. Broken units go to a repairman with one server, where the repair time is a random variable with a phase-type distribution. After being repaired units start working again. The units' lifetimes are independent identically distributed random variables also with phase-type distribution. We want to describe the number of working units in each phase before failure and the number of broken units in the system. We consider the system for large N, with the repair time scaled by N. This means that the repair time per unit decreases as N increases. We describe the limit behavior of the system when N goes to infinity. The assumption of phase-type distributions allows to represent a wide variety of systems, as phase-type distributions approximate well many positive distributions, allowing, at the same time, to exploit properties of exponential distributions and Markov structure. Concerning the repairing facility, we consider a single server with the service time also scaled according with the number of units, and we find a different behavior than for the model scaled both in the number of units and the number of servers.

A phase-type distribution with k phases is the distribution of the time to absorption in a finite Markov chain with $k+1$ states, where one state is absorbing and the remaining k states are transient. With an appropriate numbering of the states, the transient Markov chain has infinitesimal generator $\widehat{M} = \left(\begin{array}{c|c} M & m \\ \hline 0 & 0 \end{array}\right)$, where M is a $k \times k$ matrix, and $m = -M\mathbb{1}$, with $\mathbb{1}$ the column vector of ones in $\mathbb{R}^k$. The initial distribution for the transient Markov chain is a column vector $(r, 0) \in \mathbb{R}^{k+1}$, where r is the initial distribution among the transient states. We represent this phase-type distribution by (k, r, M). We refer to [AA10] for further background about phase-type distributions.

We describe the distributions and variables involved in the model. All vectors are column vectors.

Repair time. The repair time follows a phase-type distribution (m, p, NA), with m phases, matrix NA (where A is a fixed matrix and N is the scaling factor) and initial distribution p. We denote

$$Na = N(a_1, \ldots, a_m) = -NA\mathbb{1}.$$

Lifetime. The lifetime for each unit is phase-type (n, q, B), with n phases, matrix B and initial distribution q. We denote

$$b = (b_1, \ldots, b_n) = -B\mathbb{1}.$$

Working units. $\widetilde{X}_i^N(t)$ is the number of units working in phase i of their lifetimes at time t, for $i = 1, \ldots, n$, and

$$\widetilde{X}^N = (\widetilde{X}_1^N, \ldots, \widetilde{X}_n^N).$$

Repairman state. $\widetilde{Z}_i^N(t)$ is number of units being repaired in phase i of their repair times for $i = 1, \ldots, m$ ($\widetilde{Z}_i^N(t)$ is zero or one), and

$$\widetilde{Z}^N = (\widetilde{Z}_1^N, \ldots, \widetilde{Z}_m^N).$$

Waiting queue. $\widetilde{Y}^N(t)$ is the number of broken units waiting to be repaired.

Scaling. We consider the scaling:

$$X^N = \frac{1}{N}\widetilde{Y}^N,\ Y^N = \frac{1}{N}\widetilde{Y}^N,\ Z^N = \frac{1}{N}\widetilde{Z}^N.$$

Note that $\mathbb{1}^T\widetilde{Z}^N(t) + 1_{\{\mathbb{1}^T\widetilde{X}^N(t)=N\}} = 1$, where $1_{\mathcal{P}}$ is the indicator function of the predicate $\mathcal{P}$. That means that units are all working, or there is one unit being repaired at the server. In addition,

$$\mathbb{1}^T\widetilde{X}^N(t) + \widetilde{Y}^N(t) + \sum_{i=1}^{m}\widetilde{Z}_i^N(t) = N.$$

Let

$$\widetilde{U}^N = \left(\widetilde{X}_1^N, \ldots, \widetilde{X}_n^N, \widetilde{Y}^N, \widetilde{Z}_1^N, \ldots, \widetilde{Z}_m^N\right)$$

be a continuous time Markov chain.

We denote by $e^i \in \mathbb{R}^{n+m+1}$ the vector $e^i = (e_1^i, \ldots, e_{n+m+1}^i)$ with $e_i^i = 1$ and $e_j^i = 0$ for $i \neq j$, $i, j = 1, \ldots, n+m+1$. We call them direction vectors. We describe the possible transitions for this Markov chain from a vector $\widetilde{u}$ in the state space, with its corresponding transition rates. We identify transitions from state $\widetilde{u}$ to state $\widetilde{v}$ with the difference $\widetilde{v} - \widetilde{u}$, written as a difference of direction vectors. We consider the vector $\widetilde{u} = (\widetilde{x}, \widetilde{y}, \widetilde{z})$ with $\widetilde{x} = (\widetilde{x}_1, \ldots, \widetilde{x}_n)$, $\widetilde{z} = (\widetilde{z}_1, \ldots, \widetilde{z}_m)$, with $\widetilde{x}_i \in \{0, 1, \ldots, N\}$ for all $i = 1, \ldots n$, $\widetilde{y}_i \in \{0, 1, \ldots, N\}$ and $\widetilde{z}_i \in \{0, 1\}$ for all $i = 1, \ldots m$.

A working unit in phase i changes to phase j. For $i, j = 1, \ldots, n$, transition $e^j - e^i$ (that is from state $\widetilde{u} = (\widetilde{x}_1, \ldots, \widetilde{x}_i, \ldots, \widetilde{x}_j, \ldots, \widetilde{x}_n, \widetilde{y}, \widetilde{z})$ to state $\widetilde{v} = (\widetilde{x}_1, \ldots, \widetilde{x}_i - 1, \ldots, \widetilde{x}_j + 1, \ldots, \widetilde{x}_n, \widetilde{y}, \widetilde{z}))$ occurs with rate $b_{ij}\widetilde{x}_i$.

A working unit in phase i breaks and goes to the buffer. The unit goes to the buffer because there is one unit in service. For $i = 1, \ldots, n$, transition $e^{n+1} - e^i$ occurs at rate $b_i\widetilde{x}_i 1_{\{\mathbb{1}^T\widetilde{x}<N\}}$.

A working unit in phase i breaks and starts being repaired. The unit starts being repaired because the repairman is idle, at phase j. For $i = 1, \ldots, n$, $j = 1, \ldots, m$, transition $e^{n+1+j} - e^i$ occurs at rate $b_i p_j \widetilde{x}_i 1_{\{\mathbb{1}^T\widetilde{x}=N\}}$.

A unit that is being repaired in phase i changes to phase j. For $i, j = 1, \ldots, m$, transition $e^{n+1+j} - e^{n+1+i}$ occurs at rate $N a_{ij}\widetilde{z}_i$.

A unit that is being repaired in phase i ends its service and starts working at phase j with the buffer empty. If the buffer is empty, nobody starts being served and for $j = 1, \dots, n$, $i = 1, \dots, m$, the transition $e^j - e^{n+1+i}$ occurs at rate $N a_i q_j \widetilde{z}_i 1_{\{\widetilde{y}=0\}}$.

A unit that is being repaired in phase i ends its service and starts working at phase j with nonempty buffer. If the buffer is nonempty a unit in the buffer starts being served in phase k at the same time, then for $j = 1, \dots, n$ and $i, k = 1, \dots, m$, the transition $e^j + e^{n+1+k} - e^{n+1} - e^{n+1+i}$ occurs at rate $N a_i q_j p_k \widetilde{z}_i 1_{\{\widetilde{y}>0\}}$.

2.3 Fluid limit

In this section we compute the drift and analyze its limit, that defines a deterministic system, in this case a piecewise smooth dynamical system.

2.3.1 Drift computation and description of its limit

In order to understand an summarize the dynamics of the stochastic process we compute the drift for our model. We compute it and we analyze the behavior of the ODE that will define the limit. We obtain an ODE with discontinuous right-hand side.

Let us recall that for a Markov chain $V \in \mathbb{R}^d$, with transition rates $r_v(x)$ from x to $x + v$, the drift is defined by $\beta(x) = \sum_v v r_v(x)$, where the sum is in all possibles values of v. One possible representation of a Markov chain is in terms of the drift, where in a general way

$$V(t) = V(0) + \int_0^t \beta(V(s))ds + M(t),$$

with $M(t)$ a martingale. One approach to establish a fluid limit is to exploit this decomposition for the scaled process, to prove that there is a deterministic limit for the integral term and to prove that the martingale term converges to 0.

We write down the drift β of the scaled process $U^N = \widetilde{U}^N/N$, evaluated at $u = (x, y, z)$ with $x = (x_1, \dots, x_n)$, $z = (z_1, \dots, z_m)$, with $\widetilde{x}_i \in \{0, 1/N, \dots, 1\}$ for $i = 1, \dots n$, $\widetilde{y} \in \{0, 1/N, \dots, 1\}$ and $\widetilde{z}_i \in \{0, 1/N\}$ for $i = 1, \dots m$. Let $\beta = (\beta_1, \dots, \beta_{n+m+1})$. For $i = 1, \dots, n$ we have the following equations:

$$\beta_i(u) = \sum_{j=1}^{n} b_{ji} x_j + q_i \sum_{j=1}^{m} a_j N z_j.$$

Let us call β_X the first n coordinates of the drift. In matrix notation:

$$\beta_X(u) = B^T x + a^T Nzq.$$

For $i = n + 1$ the drift equation (the $(n+1)$th coordinate of the drift) is:

$$\beta_{n+1}(u) = \sum_{j=1}^{n} b_j x_j 1_{\{\mathbb{1}^T x<1\}} - \sum_{i=1}^{m} a_i N z_i 1_{\{y>0\}}$$

and in matrix notation (we also call this coordinate β_Y):

$$\beta_Y(u) = b^T x 1_{\{\mathbb{1}^T x<1\}} - a^T N z 1_{\{y>0\}}.$$

For $k = n + 1 + i$, with $i = 1, \dots, m$ we have:

$$\beta_k(u) = p_i 1_{\{\mathbb{1}^T x=1\}} \sum_{j=1}^{n} b_j x_j + \sum_{j=1}^{m} N z_j \left(a_{ji} + p_i 1_{\{y>0\}} a_j\right).$$

In matrix notation (we call these coordinates of the drift β_Z) we have:

$$\beta_Z(u) = b^T x 1_{\{\mathbb{1}^T x=1\}} p + A^T N z + a^T N z 1_{\{y>0\}} p.$$

We denote $\beta(u) = \beta(x, y, z)$. We have

$$\beta_X(x,y,z) = B^T x + a^T Nzq, \quad (2.1)$$
$$\beta_Y(x,y,z) = b^T x 1_{\{\mathbb{1}^T x<1\}} - a^T N z 1_{\{y>0\}}, \quad (2.2)$$
$$\beta_Z(x,y,z) = b^T x 1_{\{\mathbb{1}^T x=1\}} p + A^T N z + a^T N z 1_{\{y>0\}} p. \quad (2.3)$$

These equations suggest the ODE that should verify the deterministic limits (x, y) of (X^N, Y^N), if they exist. However, the drift depends on the values of Nz. The process $\widetilde{Z}^N$ varies at a rate of order N whereas the processes X^N and Y^N vary at a rate of order 1. So, we can assume that when N goes to infinity and for a fixed time the process $\widetilde{Z}^N$ has reached its stationary regime and then the limit of the last m coordinates of the drift is negligible. With this argument the candidate to the ODE defining the fluid limit is obtained by replacing in equations (2.1) and (2.2) Nz by $\widetilde{z}$, the solution to the n-dimensional equation

$$b^T x 1_{\{\mathbb{1}^T x=1\}} p + A^T \widetilde{z} + a^T \widetilde{z} 1_{\{y>0\}} p = 0.$$

Solving this last equation (multiplying by $\mathbb{1}^T$, by $\mathbb{1}^T \left(A^T\right)^{-1}$, and using the relationship $\mathbb{1}^T \widetilde{z} = 1_{\{\mathbb{1}^T x<1\}}$) we have $a^T \widetilde{z} = \mu 1_{\{\mathbb{1}^T x<1\}}$, with $1/\mu =$

$-\mathbb{1}^T \left(A^T\right)^{-1} p$, the mean time before absorption for the transient Markov chain defining the phase-type repair time distribution. We refer to [AA10] for properties of phase-type distributions. As we want to obtain an ODE for x, the candidate to ODE's vector field is $F(x) = B^T x + \mu 1_{\{\mathbb{1}^T x<1\}} q$. We observe that the equation $x' = B^T x + \mu q$ is valid when $\mathbb{1}^T x < 1$, or, in the border $\mathbb{1}^T x = 1$, when the vector field $B^T x + \mu q$ points towards the region $\mathbb{1}^T x < 1$, that is $\mathbb{1}^T (B^T x + \mu q) < 0$. Using that $B\mathbb{1} = -b$ the condition is $b^T x > \mu$. When $\mathbb{1}^T x = 1$ and $b^T x \leq \mu$ the equation presents what is called sliding motion. We follow the presentation of this topic in [Bor11a]. What happens is that the deterministic system has trajectories in the border surface $\mathbb{1}^T x = 1$. The vector field that drives the equation in the border is $G(x)$, where $\mathbb{1}^T G(x) = 0$ and $G(x)$ is a linear combination of $B^T x + \mu q$ and $B^T x$ (the vectors fields corresponding to the drift in the interior and in the border). Then $G(x) = (1 - \phi(x))(B^T x + \mu q) + \phi(x) B^T x$ with $\mathbb{1}^T G(x) = 0$ that leads to $\phi(x) = 1 - b^T x/\mu$ and then, computing $G(x)$,

$$x' = \begin{cases} B^T x + \mu q, & \text{if } b^T x > \mu \text{ or } \mathbb{1}^T x < 1, \\ B^T x + b^T x q, & \text{if } b^T x \leq \mu \text{ and } \mathbb{1}^T x = 1, \end{cases} \tag{2.4}$$

2.3.2 Convergence to fluid limit

In this section we state our main results. First we show that the scaled stochastic process $\left(X^N, Y^N\right)$ converges to a deterministic piecewise smooth dynamical system (x, y). Processes $\left(X^N, Y^N\right)$ and (x, y) are multidimensional (they live in $\mathbb{R}^{n+1}$), as the number of phases for working units is n. Convergence is in probability, uniformly in compact time intervals (*u.c.p.* convergence). From the calculus of the drift for the stochastic processes in Section 2.3 we have that the limit processes is driven by the vector field $B^T x + \mu q$ in the interior $\mathbb{1}^T x < 1$ and by the vector field $B^T x$ in the border $\mathbb{1}^T x = 1$. Very close to the border $\mathbb{1}^T x = 1$, when $b^T x \leq \mu$ the vector field $B^T x + \mu q$ points outside the region $\mathbb{1}^T x < 1$, but if we consider the vector field induced by transitions in the border, when $b^T x \leq \mu$ we have a vector field $B^T x$ that points towards the region $\mathbb{1}^T x < 1$. Because of this, the processes driven by those vector fields present a sliding motion, that means that, when $b^T x \leq \mu$, the trajectory remains in the border $\mathbb{1}^T x = 1$ driven by a convex combination of both fields. So, we must first define the piecewise smooth dynamical system (x, y), where $y = 1 - \mathbb{1}^T x$ and x is the solution (in the sense of Filippov) of the differential equation with discontinuous right-hand side (2.4).

We address the existence of solutions to Equation (2.4) and we prove

Lemma 2.1. For a differential equation $x' = F(x)$, with F discontinuous, solutions are defined in the set of absolutely continuous functions, instead of differentiable functions as in the classical case. The ODE is defined on the region $\{\mathbb{1}^T x \le 1\}$ by $B^T x + \mu q$ on $\{\mathbb{1}^T x < 1\}$ and $B^T x$ on $\{\mathbb{1}^T x = 1\}$. In order to consider the framework of differential equations with discontinuous right-hand sides, we extend the definition of the ODE. We define the equation by two continuous fields, $F_1(x) = B^T x + \mu q$ in the region $R_1 = \{\mathbb{1}^T x < 1\}$, $F_2(x) = B^T x$ in the region $R_2 = \{\mathbb{1}^T x > 1\}$, with a region $H = \{x \in \mathbb{R}^n : \ \mathbb{1}^T x = 1\}$ where the field is discontinuous. The field $B^T x$ always point towards R_1. The field $B^T x + \mu q$ points towards R_1 when $b^T x > \mu$ and point towards R_2 when $b^T x < \mu$. We find that the equation presents transversal crossing in H for $b^T x > \mu$ and a stable sliding motion in H for $b^T x < \mu$ (defined as in [Bor11a]). Transversal crossing occurs when both vector fields point towards R_1 and trajectories from R_2 cross H. If trajectories start in R_1 or in H, with $b^T x > \mu$ they go into R_1. Stable sliding motion occurs as F_1 points towards R_2 and F_2 points towards F_1 (in H, for $b^T x < \mu$). The ODE has trajectories in the border surface H. The vector field that drives the equation is $G(x)$, with $G = F_1$ in the interior and in the border, when $b^T x < \mu$, $G(x)$ verifies $\mathbb{1}^T G(x) = 0$ and it is a linear combination of $F_1(x) = B^T x + \mu q$ and $F_2(x) = B^T x$, leading to Equation (2.4). When $b^T x + \mu$, $B^T x + \mu q$ is tangential to H, whereas $B^T x$ points toward R_1, so the trajectories go into R_1. This is called first order exit condition of sliding motion.

Lemma 2.1. The differential Equation (2.4) has a unique solution for each initial condition x_0, with $\mathbb{1}^T x_0 \le 1$.

Proof. From [Bor11a], in order to prove the existence of solutions we need to verify that the field F defined as $F_1(x) = B^T x + \mu q$ in R_1 and $F_2(x) = B^T x$ in R_2 is continuous in each closure $\bar{R}_1$ and $\bar{R}_2$. In addition, if we consider the normal vector to H, $\mathbb{1}$, we can verify that (except in the region $\{b^T x = \mu\}$) we have $\mathbb{1}^T F_1(x) > 0$ and $\mathbb{1}^T F_2(x) < 0$. These conditions mean that there is a stable sliding motion, where the solution belongs to H, driven by G, the linear combination of F_1 and F_2. In addition, unique solutions are also defined for initial conditions in H. □

Once our limit candidate is defined, we state Theorem 2.2. We recall the definitions of the drift and the vector fields. Let $U^N = \left(X^N, Y^N, Z^N\right)$ and $u = (x, y, 0)$. We have that $(x(t), y(t)) = (x_0, y_0) + \int_0^t G(x(s), y(s)) \mathrm{d}s$.

Theorem 2.2. *Let* $\lim_{N\to\infty} X^N(0) = x_0$ *in probability, with* x_0 *deterministic. Then for all* $T > 0$

$$\lim_{N\to\infty} \sup_{[0,T]} \left|\left(X^N(t), Y^N(t)\right) - (x(t), y(t))\right| = 0,$$

in probability. The process (x, y) *is defined by* $y = 1 - \mathbb{1}^T x$ *and* x *the solution to Equation (2.4) with initial condition* x_0*.*

Proof. First, we observe that $\mathbb{1}^T X^N(t) + Y^N(t) + \mathbb{1}^T Z^N(t) = 1$, and that $\lim_{N\to+\infty} \sup_{[0,T]} Z^N(t) = 0$. Then, in order to prove that

$$\lim_{N\to+\infty} \sup_{[0,T]} \left|\left(X^N(t), Y^N(t)\right) - (x(t), y(t))\right| = 0$$

in probability, we only need to prove that

$$\lim_{N\to+\infty} \sup_{[0,T]} \left|X^N(t) - x(t)\right| = 0$$

in probability. In the proof of this theorem we follow the approach of [Bor11b].

$$\sup_{[0,T]} \left|X^N(t) - x(t)\right| \le \left|X^N(0) - x(0)\right| \tag{2.5}$$

$$+ \sup_{[0,T]} \left|X^N(t) - X^N(0) - \int_0^t \beta_X\left(U^N(s)\right) \mathrm{d}s\right| \tag{2.6}$$

$$+ \sup_{[0,T]} \left|\int_0^t \beta_X\left(U^N(s)\right) \mathrm{d}s - \int_0^t F\left(X^N(s)\right) \mathrm{d}s\right| \tag{2.7}$$

$$+ \sup_{[0,T]} \left|\int_0^t F\left(X^N(s)\right) \mathrm{d}s - \int_0^t G\left(X^N(s)\right) \mathrm{d}s\right| \tag{2.8}$$

$$+ \sup_{[0,T]} \left|\int_0^t G\left(X^N(s)\right) \mathrm{d}s - \int_0^t G\left(x(s)\right) \mathrm{d}s\right| \tag{2.9}$$

We want to prove that (2.6), (2.7) and (2.8) converge to 0 in probability. So, provided that the initial condition $X^N(0)$ converges to $x(0)$, we have that, with probability that tends to 1 with N,

$$\sup_{[0,T]} \left|X^N(t) - x(t)\right| \le \varepsilon + \sup_{[0,T]} \left|\int_0^t G\left(X^N(s)\right) \mathrm{d}s - \int_0^t G\left(x(s)\right) \mathrm{d}s\right|.$$

Using that G is piecewise linear and Gronwall inequality, we obtain the bound $\sup_{[0,T]} \left|X^N(t) - x(t)\right| \leq \varepsilon e^{KT}$, which leads to

$$\sup_{[0,T]} \left|X^N(t) - x(t)\right| \to 0$$

in probability. We study the convergence of (2.6), (2.7) and (2.8). To show the convergence of (2.6), we first notice that

$$(2.6) \leq \left|U^N(t) - U^N(0) - \int_0^t \beta\left(U^N(s)\right) \mathrm{d}s\right|.$$

Convergence follows from the representation of the process as the initial condition plus the integral of the drift plus a martingale term. The martingale term goes to 0 with N because of the scaling. Let us define for a Markov chain $V \in \mathbb{R}^d$, with transition rates $r_v(x)$ from x to $x+v$, $\alpha(x) = \sum_v |v|^2 r_v(x)$. Let us also call α the corresponding object for U^N. Convergence of (2.6) can be then proved using Proposition 8.7 in [DN08], that states that

$$E\left(\sup_{[0,T]} \left|U^N(t) - U^N(0) - \int_0^t \beta\left(U^N(s)\right) \mathrm{d}s\right|^2\right) \leq 4 \int_0^T \alpha\left(U^N(s)\right) \mathrm{d}s.$$

As in our scaling $\sup_{[0,T]} \alpha\left(U^N(t)\right) \sim \mathcal{O}(1/N)$, convergence holds.

To prove that (2.7) converges to 0 in probability we consider the last m coordinates of (2.6), corresponding to the phases at the repairman. As we have proved that (2.6) converges to 0 in probability, we conclude that

$$\int_0^t \left(b^T X^N(s) 1_{\{\mathbb{1}^T X^N(s)=1\}} p + A^T \widetilde{Z}^N(s) + a^T \widetilde{Z}^N(s) 1_{\{Y^N(s)>0\}} p\right) \mathrm{d}s$$

converges to 0 in probability. Multiplying by $\mu \mathbb{1}^T (A^T)^{-1}$,

$$\int_0^t \left(-b^T X^N(s) 1_{\{\mathbb{1}^T X^N(s)=1\}} + \mu 1_{\{\mathbb{1}^T X^N(s)<1\}} - a^T \widetilde{Z}^N(s) 1_{\{Y^N(s)>0\}}\right) \mathrm{d}s \tag{2.10}$$

goes to 0 in probability. In addition, $\mathbb{1}^T \beta_X + \beta_Y + \mathbb{1}^T \beta_Z = 0$. Then, as $\int_0^t \beta_Z\left(U^N(s)\right) \mathrm{d}s$ converges to 0 in probability,

$$\int_0^t \left(\mathbb{1}^T \beta_X\left(U^N(s)\right) + \beta_Y\left(U^N(s)\right)\right) \mathrm{d}s$$

also converges to 0 in probability and it is equal to

$$\int_0^t \left(-b^t X^N(s) 1_{\{\mathbb{1}^T X^N(s)=1\}} + a^T \widetilde{Z}^N(s) 1_{\{Y^N(s)=0\}}\right) \mathrm{d}s. \tag{2.11}$$

Then, considering the sum of Equations (2.10) and (2.11), we obtain

$$\lim_{N\to+\infty}(2.7) = \lim_{N\to+\infty} \sup_{[0,T]} \int_0^t \left(a^T \widetilde{Z}^N(s) q - \mu 1_{\{\mathbb{1}^T X^N(s)<1\}} q\right) \mathrm{d}s = 0.$$

The convergence of (2.8) can be proved by approximating the continuous process in the border by a discrete process with the same jumps. As our model verifies the hypotheses of [Bor11a], the same proof that in Lemma 3 of [Bor11b] holds.

□

Let us study the behavior of the system defined by Equation (2.4) by studying its fixed points. We observe that $1/\mu$ is the mean of the phase-type distribution (m, p, A). We define $1/\lambda = -\mathbb{1}^T \left(B^T\right)^{-1} q$, the expected value of the phase-type distribution (n, q, B), and

$$\rho = \frac{\mu}{\lambda}. \tag{2.12}$$

We identify three different behaviors, that we call (using the same definitions as in [Rob03] for the $M/M/N/N$ queue) sub-critical when $\rho < 1$, critical when $\rho = 1$ and super-critical when $\rho > 1$.

Sub-critical case, $\rho < 1$. The mean repair time per unit $1/\mu$ is greater than the mean lifetime, so we find an equilibrium with a positive number of broken units in the system. When we compute the fixed points in Equation (2.4) the fixed point is an interior point in $\mathbb{1}^T x \leq 1$, and it is a global attractor.

Super-critical case, $\rho > 1$. When $\rho > 1$, intuitively the repairman is more effective, and we have an equilibrium with all the units (in the deterministic approximation) working. The fixed point is also a global attractor, and it is in the border $\mathbb{1}^T x = 1$.

Critical case, $\rho = 1$. In this case the fixed points for the equation in the interior and in the border coincide, giving a fixed point in the border that is a global attractor.

We state these results in the following lemma.

Lemma 2.3. There are three different behaviors for Equation (2.4):

$\rho < 1$ (sub-critical). There is a unique fixed point x^* that is a global attractor and verifies $\mathbb{1}^T x^* = \rho < 1$:

$$x^* = -\mu(B^T)^{-1}q, \tag{2.13}$$

$\rho > 1$ (super-critical). There is a unique fixed point x^* that is a global attractor and verifies $\mathbb{1}^T x^* = 1$ and $b^T x^* < \mu$:

$$x^* = -\lambda(B^T)^{-1}q, \tag{2.14}$$

$\rho = 1$ (critical). There is unique fixed point x^* given by equations (2.13) or (2.14). It is a global attractor and verifies $\mathbb{1}^T x^* = 1$ and $b^T x^* = \mu = \lambda$.

Proof. First, we compute the fixed points of both fields: $B^T x + \mu q$ in $\mathbb{R}^n$, and $(B^T + qb^T)x$ in $\{\mathbb{1}^T x = 1\}$. We will exploit the linearity of the field in each region where it is continuous. Then we discuss in terms of ρ.

The fixed point for $B^T x + \mu q$ is $x_1^* = -\mu(B^T)^{-1}q$. We recall that B^T is regular due to the properties of phase type distributions. In addition, the eigenvalues of B^T are all negative (since B^T has the same eigenvalues than B). Matrix B has a negative diagonal, and the sum of all non diagonal entries per row (that are all positive) is less than or equal to the absolute value of the diagonal element. This is because the sum of each row of $\widehat{B}$ is 0 and the last column (that does not belong to B) has non-negative entries. Then, considering the field in $\mathbb{R}^n$, we have that $x_1^* = -\mu(B^T)^{-1}q$ is a global attractor. In addition, as $\mathbb{1}^T x_1^* = \rho$, we have that x_1^* is an interior point of R_1 iff $\rho < 1$, x_1^* is an interior point of R_2 iff $\rho > 1$, and $x_1^* \in H \cap \{b^T x = \mu\}$ iff $\rho = 1$. We observe that when $\rho \geq 1$, as x_1^* is the unique fixed point of $B^T x + \mu q$, and the vector field points outside R_1, the solution of $x' = B^T x + \mu q$ is pushed towards H.

Now we consider the fixed point of $(B^T + qb^T)x$ in $\{\mathbb{1}^T x = 1\}$. As $b = -B\mathbb{1}$, we have that $(B^T + qb^T)x = 0$ iff $(I - q\mathbb{1}^T)B^T x = 0$, where I is the identity matrix. As B^T is invertible, if v is an eigenvector of $q\mathbb{1}^T$ with eigenvalue 1, $x_1^* = (B^T)^{-1}v$ is a fixed point. Matrix $q\mathbb{1}^T$ has eigenvalues 0 and 1 and, as $q\mathbb{1}^T$ has range 1, the dimensions of the corresponding eigenspaces are respectively $n-1$ and 1. Then there is a one-dimensional space with eigenvalue 1, that intersects $\{\mathbb{1}^T x = 1\}$, giving the fixed point $x_2^* = -\lambda(B^T)^{-1}q$. Since $\mathbb{1}^T x_2^* = 1$, we have that $x_2^* \in H$. In addition $x_2^* \in H \cap \{b^T x < \mu\}$ iff $\rho > 1$, $x_2^* \in H \cap \{b^T x > \mu\}$ iff $\rho < 1$ and $x_2^* \in H \cap \{b^T x = \mu\}$ iff $\rho = 1$.

We also have that, restricted to H, x_2^* is a global attractor, so in the case of $\rho \leq 1$, the solution in H is pushed to the region $\{b^t x > \mu\}$, where the solution is again driven by the field $B^T x + \mu q$, so the solution that starts in H does not remain in H for $\rho < 1$. □

Theorem 2.4. *The system in stationary regime $X^N(\infty)$ converges in probability to x^*, where x^* is the unique fixed point of Equation (2.4).*

Proof. To prove the convergence in stationary regime, we use the results in [HB12]. In Theorem 5 of [HB12] it is proved convergence in stationary regime to the fixed point for piecewise smooth dynamical systems. This is an extension of a general result in [BLB08]. The hypotheses needed are that the fixed point is a unique global attractor and regularity assumptions for the trajectories. We use Lemma 2.3 to characterize the fixed point. The assumptions about the regularity of the trajectories are the following: the changes of the vector field (such as transversal crossings) are bounded; the number of points in which sliding motion starts or terminates is bounded; and sliding motion occurs on H with both vectors having a non-null component normal to H. These conditions are verified in our case, where solutions are piecewise linear and the discontinuity surface H is a hyperplane. Then applying Theorem 5 in [HB12] we obtain Theorem 2.4. □

We observe that the fixed point depends only on the repair process only through the mean repair time, but it does depend on matrix B. This means that different lifetime distributions with the same mean lead to different stationary behaviors.

2.4 Numerical examples

We consider the repairman problem with $N = 100$ units, for different phase type time distributions and for different values of ρ. The parameters are defined in Section 2.2 and ρ is defined in equation (2.12). As initial condition we fix the total number of working units and sample the number in each phase according to the phase type initial distribution. We show the scaled number of working units in each phase. We illustrate the convergence to the ODE's fixed point x^* and the sliding motion. The parameters for each example (figure) are given in Table 2.1.

Exponential life-time and hypoexponential repair time. In Figure 2.1 we consider exponential life-time and sums of independent exponentials for the repair time. We represent for each parameter set the evolution with time of the stochastic process X^N and we show the convergence to the fixed point x^*.

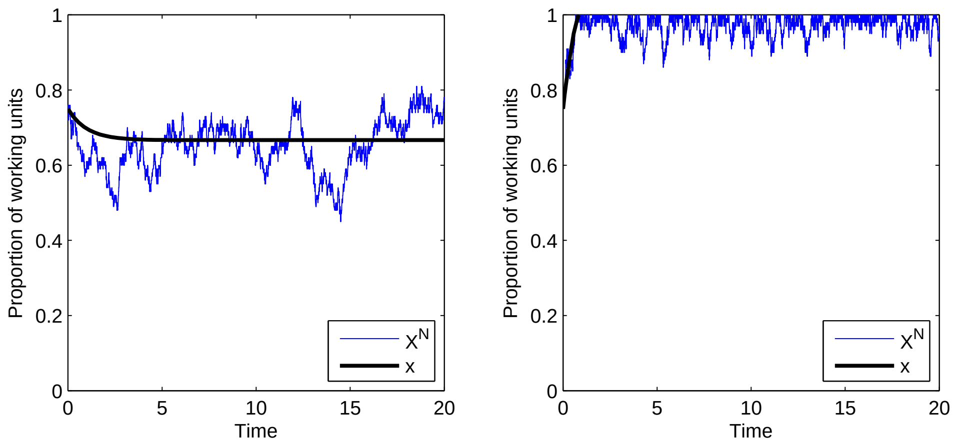

Figure 2.1: Exponential life-time and hypoexponential repair time. In the right we show the sliding motion.

Hypoexponential life and repair time, sliding motion. In Figure 2.2 we consider two phases both in the life and repair times (both distributions are hypoexponential). At the left we represent, for each parameter set, the evolution with time of the stochastic processes X_1^N and X_2^N and the ODE's solution (x_1, x_2). We also illustrate the convergence to the fixed point $x^* = (x_1^*, x_2^*)$. At the right we represent, for each parameter set, the trajectory of process X^N, the trajectory of the ODE's solution x and the fixed point. Depending on the initial condition, we find sliding motion (as shown in the bottom figures), but, as we have the same parameters, both trajectories converge to the fixed point.

Hypoexponential life-time with three phases and exponential repair time. In Figure 2.3 we consider three phases (hypoexponential) in the life-time and exponential repair time. At the left we represent the evolution with time of the stochastic processes X_1^N, X_2^N, X_3^N and we show the convergence to the fixed point $x^* = (x_1^*, x_2^*, x_3^*)$. At the right we represent the trajectory of process X^N and the fixed point. The example at the top has $\rho < 1$. The example at the bottom has $\rho > 1$, so we find sliding motion in the plane $\mathbb{1}^T x = 1$ and the fixed point is also in the same plane.

2.5 Asymptotic distribution

The next issue is, given that the processes has a deterministic limit, the asymptotic of the difference behavior between the process and its limit, where we distinguish different behaviors.

2.5.1 Sub-critical regime

Proposition 2.5. If for $T > 0$ the trajectories of the solution x to the Equation (2.4) remains in the interior of the state space and $\sqrt{N}(X^N(0) - x(0)) \Rightarrow V(0)$, with $V(0)$ deterministic, then the system has a Gaussian limit, that is $\sqrt{N}(X^N(t) - x(t)) \Rightarrow V(t)$, with V a Gaussian processes in $\mathbb{R}^k$,

Proof. The proposition derives from a general result from Ethier y Kurtz (Theore 2.3, p. 458 [EK86]), averaging the variables that characterize the repairing processes. However the variables involving the phase at the repairman appear when computing the covariance matrix for limit process, deterring to obtain an analytic expression. □

2.5.2 Critical regime for exponential case

Now we consider the case when fluid limit trajectories remain part of the time at the border, that may be due to the presence at the border of the fixed point, or due to the initial conditions if we start at the border, or trajectories that hit the border, remain some time there and then converge to the fixed point. In this case the asymptotic distribution is not Gaussian.

We study this phenomenon by simulation, considering the total time of working units $\mathbb{1}^T X^N(t)$ when the trajectories verify $\mathbb{1}^T x(t) = 1$. Theoretically we only can show a geometric distribution, with a different scaling, in the case of exponential life and repair times. We expect from simulations that this holds with phase type distributions but we don't have a proof of this fact.

Proposition 2.6. If life and repair times are exponentially distributed with parameters λ and μ, respectively, and $\rho > 1$, we have that $N - X^N(t)$ converges to a geometric distribution with parameter $1/\rho$.

Proof. In the exponential case we observe that the system has the same drift as a $M/M/N/N$ queue with service time λ and interarrival times exponentially distributed with mean $1/\mu$. We consider that life time is service time in the $M/M/N/N$ queue, whereas service time at the repairman becomes interarrival time for the $M/M/N/N$ model. In this analysis, when N units are working at the MRM model, we have that the N servers are busy in the $M/M/N/N$ queue. To conclude, we use results from [Rob03] (Prop. 6.19, p. 169) for this queue. □

2.6 Conclusions Machine Repairman Model

In this chapter we study the $PH/PH/1//N$ model, that gives more generality in repair and life-times, at the same time that presents the difficulty to deal with different time scales in the fluid limit. We obtain a fluid limit considering these different time scales following the approaches from [BKPR05, AEJV13]. The averaging phenomena due to the different time scales also appears when analyzing the fixed point for the fluid limits, where we find that it only depends on the repair time by its mean, showing an insensitivity property with respect to the life-time distribution.

We treat also the problem of discontinuous transition rates, where we obtain a piecewise smooth dynamical system. It is a switched linear system and fixed points are obtained and analyzed using arguments for linear ODEs. In the case where asymptotically all units are working we find sliding motion in the border border of the state space. In this case, different from the two other chapters, the region where sliding motion lies has higher dimension and depends on the number of phases in the original model. This fact hinders to analyze the asymptotic distribution around the limit, that it is only considered in this work for the exponential case. As in the exponential case, we expect a non- Gaussian behavior.

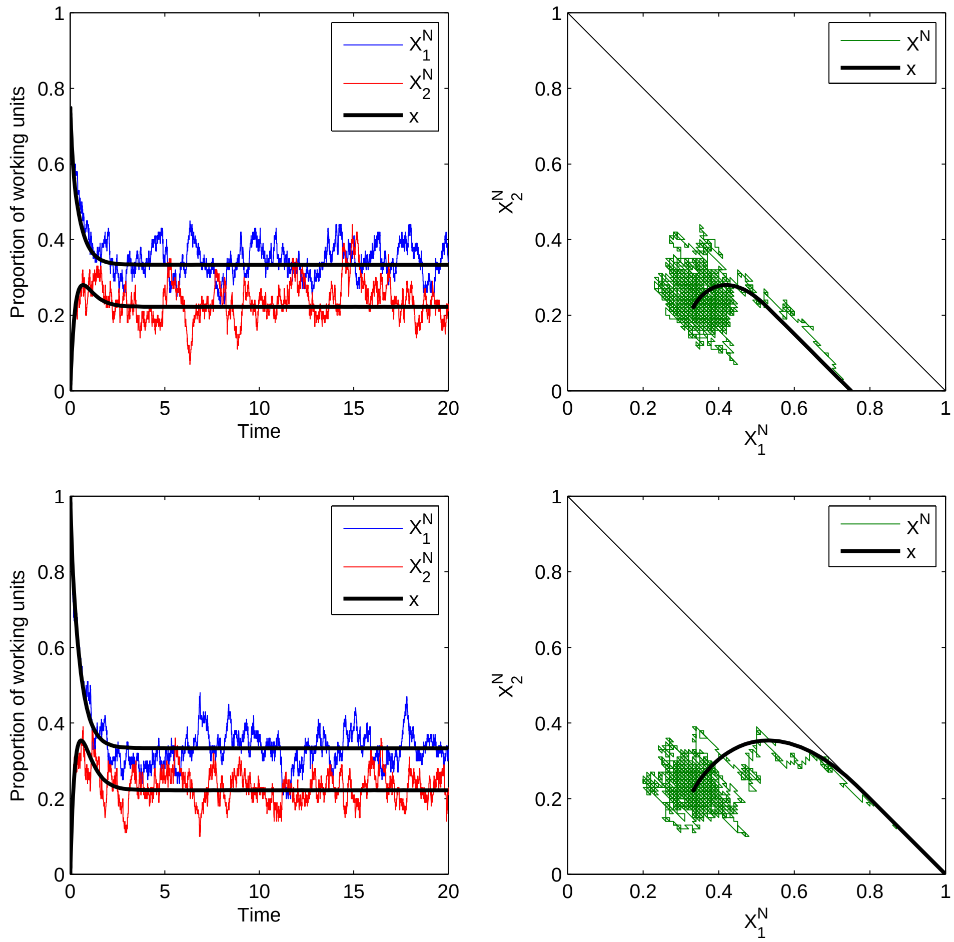

Figure 2.2: Hypoexponential life and repair time (parameters in Table 2.1). The bottom figure shows the sliding motion.

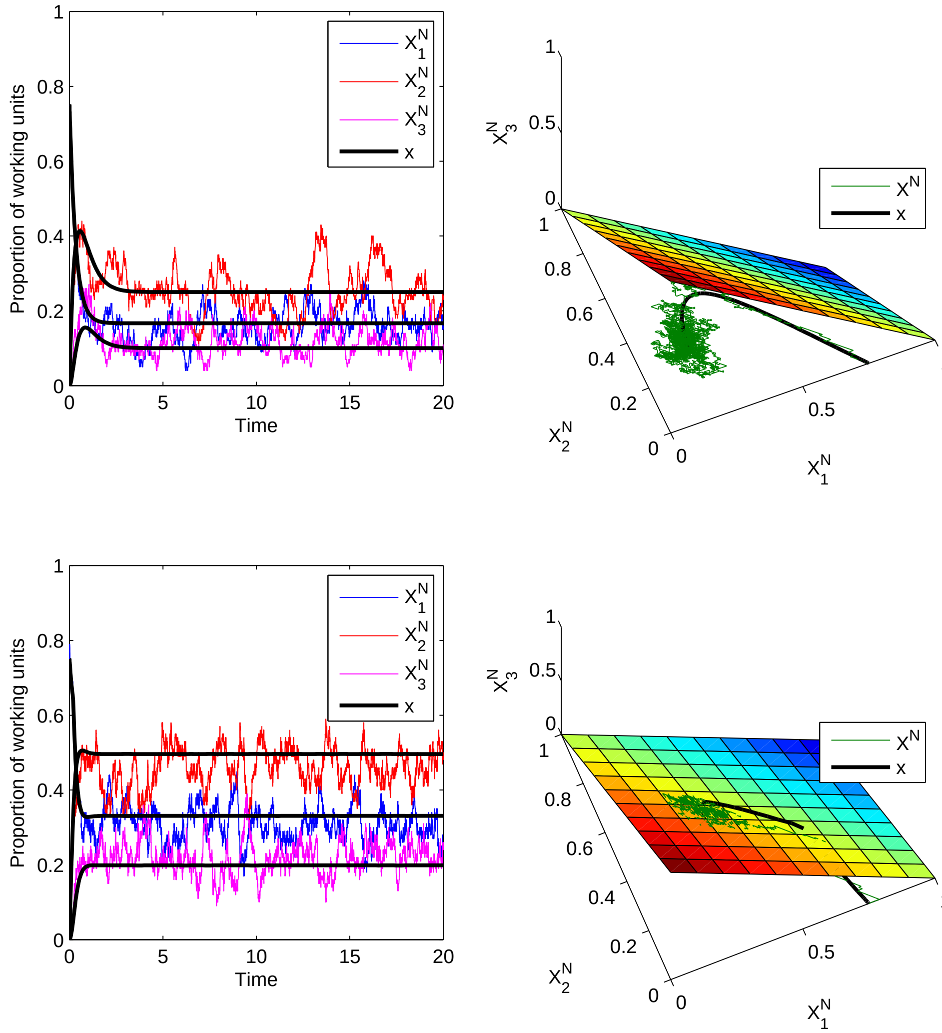

Figure 2.3: Hypoexponential life-time and exponential repair times (parameters in Table 2.1). The bottom figures shows the sliding motion and the fixed point in the plane $\mathbb{1}^T x = 1$ at the right.

Table 2.1: Parameters for Figures 2.1, 2.2 and 2.3.

	Fig.2.1(left)	Fig.2.1(right)
A	$\begin{pmatrix} -1 & 1 \\ 0 & -2 \end{pmatrix}$	$\begin{pmatrix} -2 & 2 \\ 0 & -3 \end{pmatrix}$
B	-1	-1
p	$(1,0)$	$(1,0)$
q	1	1
$X^N(0)$	$0.75N$	$0.75N$
ρ	0.6667	1.2

	Fig.2.2(top)	Fig.2.2(bottom)
A	$\begin{pmatrix} -1 & 1 \\ 0 & -2 \end{pmatrix}$	$\begin{pmatrix} -1 & 1 \\ 0 & -2 \end{pmatrix}$
B	$\begin{pmatrix} -2 & 2 \\ 0 & -3 \end{pmatrix}$	$\begin{pmatrix} -2 & 2 \\ 0 & -3 \end{pmatrix}$
p	$(1,0)$	$(1,0)$
q	$(1,0)$	$(1,0)$
$X^N(0)$	$0.75N$	N
ρ	0.5556	0.5556

	Fig.2.3(top)	Fig.2.3(bottom)
A	-0.5	-1.5
B	$\begin{pmatrix} -3 & 3 & 0 \\ 0 & -2 & 2 \\ 0 & 0 & -5 \end{pmatrix}$	$\begin{pmatrix} -3 & 3 & 0 \\ 0 & -2 & 2 \\ 0 & 0 & -5 \end{pmatrix}$
p	1	1
q	$(1,0,0)$	$(1,0,0)$
$X^N(0)$	$0.75N$	$0.75N$
ρ	0.5167	1.55

Chapter 3

Cognitive Radio

In this chapter we analyze in the fluid limit framework a model describing Cognitive Radio (abbreviated CR in the sequel) networks, where we present its basic principles as a solution of spectrum underutilization and spectrum scarcity. We propose and analyze a dynamic spectrum sharing mechanism, where primary users have strict priority over secondary ones, with the aim of providing to secondary users a satisfactory grade of service with a small interruption probability. We describe the model by means of its fluid limit, dealing with a stochastic system with discontinuous transition rates. We study the non-differentiable deterministic approximation that we obtain, inferring properties of the stochastic model by the analysis of fixed points. Finally, working with the asymptotic distribution associated with the fluid limit our main findings consist in a Gaussian limit theorem in the sub-critical case (when the system is not loaded), and a non-Gaussian limit theorem (under a different scaling scheme) in the critical one.

Nowadays, with the rapid development of wireless communications, the demand on spectrum has been growing dramatically resulting in the spectrum scarcity problem: unlicensed bands are too crowded while licensed bands are vastly underutilized, as it is described by several works devoted to measuremnts [BJR$^+$07, XFC13, JJ16].

Cognitive Radio networks, introduced by Mitola [MM99, Mit00], have been proposed as a promising technology to solve this problem by an intelligent and efficient dynamic spectrum access, with different attempts for protocols and implementation [ARR16, SV16, HHK$^+$17, HKKS17, MH17].

In this new paradigm we can identify two classes of users: primary (PU) and secondary (SU). PUs are the licensed users, they have allocated a certain portion of spectrum. SUs (also called cognitive users) are devices which are capable of detecting unused licensed bands and adapt their parameters for using them. The main idea is to dynamically re-allocate unused licensed frequency bands to secondary users. One challenge then is to distribute the spectrum holes efficiently and fairly. Another goal is to guarantee quality of service (QoS) to the SUs.

In what follows we focus on the analysis and characterization of a dynamic spectrum sharing mechanism where PUs have strict priority over secondary ones, and the key variable is the number of users in the system. We present some tools and criteria that can be used in order to improve the mean spectrum utilization with the commitment of providing to SUs a satisfactory grade of service and a small interruption probability.

We are interested in SUs whose service cannot be interrupted with high probability (like a phone call or other interactive services). For these services it is preferable to be rejected and to avoid the situation where the connection is established and then interrupted. These decisions (enter or not) represent a mechanism that can be adopted by the SUs as a sort of admission control policy. We analyze two features of these types of systems: the mean spectrum utilization and the probability that the SUs services can be interrupted. Associated with this last issue we analyze a possible policy in order to reduce this probability. For this purpose we analyze the whole system with fluid limits techniques.

In more detail we consider a scenario with C subchannels to be distributed between SUs and PUs, and where PUs have strict priority. That is to say, if a PU arrives when all the subchannels are in use, one of the SUs will be deallocated immediately.

As an example we consider a cellular network that employs frequency division duplexing where the operator has C frequency bands (subchannels) to be assigned to its users (PUs). Another example is given by the digital TV spectrum bands. In both scenarios, if there are free subchannels, the SUs could use them with the constraint that their communications can be interrupted at any time.

However, there are some differences between both examples: in a cellular network PUs can use any of the C subchannels, but in the case of digital TV each TV channel has its own frequency band. Then we assume in our model that when a PU arrives while a SU is using its subchannel and there are free subchannels, this SU can be moved instantly to another unused subchannel, without any consequence to its service. If there is not a free subchannel (the C subchannels are busy), the SUs communication will be interrupted with consequences to its quality of service. As our model takes into account only the number of subchannels that are being used by PUs and SUs, if there are free subchannels, the case of a PU that arrives to its own and a SU must be moved instantly to another free one will be modeled as if the PU arrives to a free subchannel.

We model the cognitive radio network as a two dimensional continuous time Markov chain. A fluid limit approach is used to analyze the stochastic system which is approximated by piecewise smooth dynamical system. For this purpose we use some results that generalize, for systems with discontinuous transition rates, the most classical theorems on fluid limits. One of the main is that the position of the fixed points is decisive in defining an effective operating point for the system. In addition we show that in many cases, an admission control mechanism for SUs is required in order to ensure a low probability of service interruption.

We also describe the asymptotic distribution related with the fluid limit. This distribution depends strongly on the fixed points of the deterministic approximation, both with and without admission control, and we find, depending on the parameters of the model, Gausssian and non-Gaussian asymptotic distributions. The asymptotic distribution allows to analyze the interruption probability for SUs, giving some kind of confidence bounds valid when the number of users is large. In the case of a non-Gaussian distribution the stationary regime and its limit are described explicitly in a simplified case, in the one-dimensional instead of two-dimensional situation.

There are related works on Cognitive Radio that use fluid model approaches, being representative examples [SFDS12, AA14, KB15, RLBB17, RBB18]. In [AA14], although the authors study an admission control mechanism over SUs, they do no obtain an analytical expression of QoS metrics, they only evaluate them through simulations. In [KB15] the authors study preemptive and non-preemptive priority queuing. The affected SU (which is deallocated when a PU needs a subchannel and they are all in use) has to wait in the system until a subchannel is available again (in the paper it is assumed the existence of a buffer), in our case we consider a total interruption of the communication, with no buffer. In [SFDS12] the authors study a spectrum sharing allocation for PUs and SUs where preemption can occur in three different schemes by studying a two-dimensional $M/M/C/C$. They obtain the mean number of PUs and SUs, the blocking probabilities, and the dropping probability for SUs, but they do not investigate the behavior of the system when admission control mechanisms are applied. Looking to other features of CR networks, [RLBB17] approximates by means of fluid limits the medium access probability for PUs and also for SUs, and [RBB18] characterizes the CR network using fluid limits but the focus is on the economical problem in order to optimize the profit.

This is a joint work with C. Rattaro and P. Belzarena, and many issues exposed here are also presented by C. Rattaro in [Rat18].

3.1 Model description

In this section we introduce our stochastic model for the number of primary and secondary users in the system as a two-dimensional CTMC. This is a comprehensive and although simple model for the system under study, however, it does not have an analytical solution. Therefore, we introduce in Section 3.2 a scaled version of the CTMC, in order to find a fluid limit that allows us to study the system analytically.

We model the arrival processes for both type of users as independent Poisson processes, and the service times are also independent and exponentially distributed random variables. We also model the possibility of admission control decisions when a SU arrives to the system (SUs shall decide, depending on the state of the system, whether to enter or not). We associate one user with one channel.

In this context a general model and its variations due to the admission control policy assumptions are stated in the following definitions.

Definition 3.1 (General model)**.** Consider a CTMC (X_1, X_2) defined as follows.

$X_1(t)$, $X_2(t)$: number of PUs and SUs at time t respectively,

C: total number of identical subchannels, therefore, the state space is the subset:

$$E = \{(x_1, x_2) \in \mathbb{N}^2 : 0 \leq x_1 \leq C, 0 \leq x_2 \leq C, x_1 + x_2 \leq C\}, \quad (3.1)$$

λ_1, λ_2: arrival rates for PUs and SUs respectively (independent Poisson arrivals),

μ_1, μ_2: service rates for PUs and SUs respectively (independent exponentially distributed service times),

$a(x_1, x_2)\colon E \to \mathbb{R}$, admission decision function for SUs, and represents the probability that a SU that arrives starts being served when there are x_1 PUs and x_2 SUs.

Thus the stochastic process $(X_1(t), X_2(t))$ has transition rates $q((x_1, x_2), (x_1', x_2'))$, from state (x_1, x_2) to state (x_1', x_2'), defined by:

$q((x_1, x_2), (x_1 + 1, x_2)) = \lambda_1$, if $x_1 + x_2 < C$,

$q((x_1, x_2), (x_1 - 1, x_2)) = \mu_1 x_1$,

$q((x_1, x_2), (x_1, x_2 + 1)) = a(x_1, x_2)\lambda_2$, if $x_1 + x_2 < C$,

$q((x_1, x_2), (x_1, x_2 - 1)) = \mu_2 x_2$,

$q((x_1, x_2), (x_1 + 1, x_2 - 1)) = \lambda_1$, if $x_1 + x_2 = C$ and $x_2 \neq 0$.

Definition 3.2 (Free admission control policy)**.** We call free admission control model when in the previous definition we consider no admission control policy, then $a(x_1, x_2) = 1$ if $x_1 + x_2 < C$ and $a(x_1, x_2) = 0$ if $x_1 + x_2 = C$.

Definition 3.3 (Deterministic admission control policy)**.** In the deterministic case $a(x_1, x_2) \in \{0, 1\}$; if $a(x_1, x_2) = 1$ and a SU arrives, it will start being served, and when $a(x_1, x_2) = 0$, it will not. In this work we assume $x_1 + x_2 = \delta$, with $0 < \delta < C$, as an admission control boundary. That is to say, $a(x_1, x_2) = 1$ if $x_1 + x_2 < \delta$ and $a(x_1, x_2) = 0$ if $x_1 + x_2 \geq \delta$.

Definition 3.4 (Probabilistic admission control policy)**.** In the probabilistic admission control SUs can access the system with a probability related with the number of users in the system. Let us assume that $a(x_1, x_2)$ is a continuous function that vanishes close to the border $\gamma = \{(x_1, x_2) : x_1 + x_2 = C\}$. In this work we consider $a(x_1, x_2) = 1 - (x_1 + x_2)/C$.

In figure 3.1 we show transitions both in the case where there is not admission control policy and in the case where there is a deterministic admission control.

Concerning the above formulation we may make some remarks. First notice that PUs behave as a $M/M/C/C$ queue, independent of the behavior of the SUs and of the admission policy. Secondly, in this work we consider two approaches for the admission control mechanisms: a deterministic and a probabilistic one. Thirdly, it is important to highlight that in the case when $\mu_1 = \mu_2$, for some type of admission policies $a(x_1, x_2)$, the process is a one-dimensional CTMC and the stationary distribution can be computed explicitly. However, when $\mu_1 \neq \mu_2$ (which represents the natural situation in cognitive radio networks) it is not possible to obtain a closed form expression of its stationary distribution (see for example [Zha12] and the references therein). Finally, although in the general case the stationary distribution can be computed numerically, our approach consists in formulating the corresponding fluid limit in order to characterize the system behavior and study

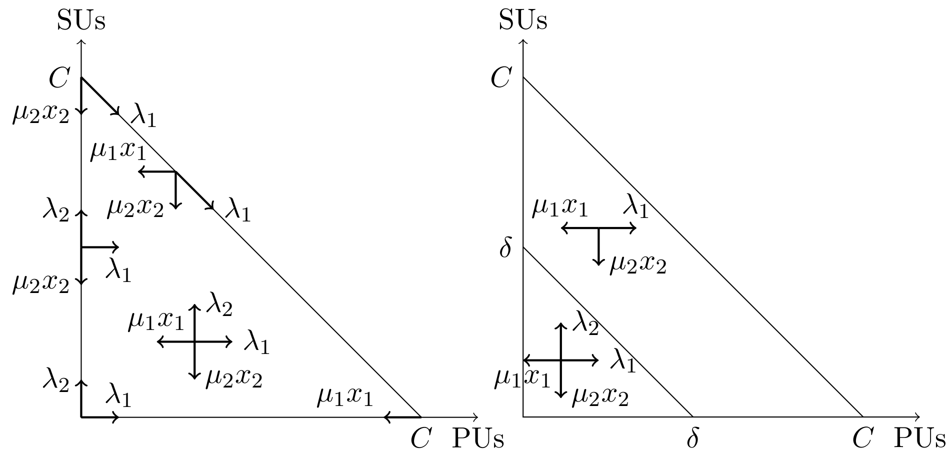

Figure 3.1: Transitions for CR model. Left: transitions without admission control. Right: transitions with admission control border $x_1 + x_2 = \delta < C$.

the influence of the admission control decisions in a more feasible and efficient way. We infer properties of the stochastic system from the study of fixed points of the deterministic fluid limit and the asymptotic distribution around it, and we define practical QoS criteria for sharing spectrum policies.

3.2 Fluid model

Let N be the scaling factor, and we define, as for the original model presented in Definition 3.1, the following sequence of CTMCs $(\widetilde{X}_1^N, \widetilde{X}_2^N)$ and whose state space is:

$$\widetilde{E}^N = \{(Ni, Nj) : (i, j) \in E\}$$

where E is defined in Equation (3.1).

Definition 3.5. Consider the sequence of CTMCs $(\widetilde{X}_1^N, \widetilde{X}_2^N)$ defined as follows.

$\widetilde{X}_1^N(t)$, $\widetilde{X}_2^N(t)$: number of PUs and SUs at time t, respectively,

CN: total number of subchannels,

$\lambda_1 N$, $\lambda_2 N$: arrival rates for PUs and SUs, respectively,

μ_1, μ_2: service rates for PUs and SUs, respectively,

$\widetilde{a}^N(x_1, x_2)$: admission decision for SUs in each state.

The admission decision in each state should verify:

$$\lim_{N \to +\infty} \widetilde{a}^N(Nx_1, Nx_2) = a(x_1, x_2).$$

This scaling concerns both admission control schemes presented in Section 3.1. In the case of a deterministic admission control $\widetilde{a}^N(x_1, x_2) = 1$ if $x_1 + x_2 < N\delta$ and $\widetilde{a}^N(x_1, x_2) = 0$ if $x_1 + x_2 \geq N\delta$. For the probabilistic admission control let $\widetilde{a}^N(x_1, x_2) = 1 - (x_1 + x_2)/CN$.

The sequence of scaled stochastic process $(\widetilde{X}_1^N(t), \widetilde{X}_2^N(t))$ has transition rates $\widetilde{q}^N((x_1, x_2), (x_1', x_2'))$, from state (x_1, x_2) to state (x_1', x_2'), defined by:

$\widetilde{q}^N((x_1, x_2), (x_1 + 1, x_2)) = \lambda_1 N$, if $x_1 + x_2 < CN$,

$\widetilde{q}^N((x_1, x_2), (x_1 - 1, x_2)) = \mu_1 x_1$,

$\widetilde{q}^N((x_1, x_2), (x_1, x_2 + 1)) = \widetilde{a}^N(x_, x_2)\lambda_2 N$, if $x_1 + x_2 < CN$,

$\widetilde{q}^N((x_1, x_2), (x_1, x_2 - 1)) = \mu_2 x_2$,

$\widetilde{q}^N((x_1, x_2), (x_1 + 1, x_2 - 1)) = \lambda_1 N$, if $x_1 + x_2 = CN$ and $x_2 \neq 0$.

In Table 3.1 we summarize the scaled parameters and their relationship with the original ones.

Table 3.1: Original and scaled parameters for the two-dimensional process (X_1, X_2) and $(\widetilde{X}_1^N, \widetilde{X}_2^N)$.

(X_1, X_2)	λ_1	λ_2	μ_1	μ_2	C	a
$(\widetilde{X}_1^N, \widetilde{X}_2^N)$	$\lambda_1 N$	$\lambda_2 N$	μ_1	μ_2	CN	$\widetilde{a}^N$

In this scaling scheme we go from our original system in Section 3.1 to a system where arrival rates and capacity are multiplied by N, that can be interpreted as a large network, with many users (both PUs and SUs) and large capacity. On the other hand, service rates in each channel are not scaled, as for instance they depend on the service type, and they do not increases individually, despite the total service time increases with the number of users as $\mu_1 \widetilde{X}_1^N(t)$ for PUs and $\mu_2 \widetilde{X}_2^N(t)$ for SUs. Finally the admission control in the large system depends only on the proportion of resources occupied.

We consider now the scaled process (X_1^N, X_2^N), defined by

$$(X_1^N, X_2^N) = (\widetilde{X}_1^N, \widetilde{X}_2^N)/N \tag{3.2}$$

and whose state space is:

$$E^N = \{(i/N, j/N) : (i,j) \in E\}$$

where E is defined in Equation (3.1).

This scaled process will converge to the deterministic fluid limit. In order to state the fluid limit result we verify that our process satisfies

$$\lim_{N\to\infty} \sup_{t\in[0,T]} \left| \frac{1}{N}\widetilde{Q}^N\left(\widetilde{X}_1^N(t), \widetilde{X}_2^N(t)\right) - Q\left(X_1^N(t), X_2^N(t)\right)\right| = 0$$

in probability (equation (A.3) in Appendix A).

We compute the drift in the following proposition.

Proposition 3.6. The drift for the process stated in Definition 3.1 is

$$Q(x_1,x_2) = \begin{pmatrix} \lambda_1 - \mu_1 x_1 \\ \lambda_2 a(x_1,x_2) - \mu_2 x_2 \end{pmatrix} \text{ if } x_1 + x_2 > C;$$

$$Q(x_1,x_2) = \begin{pmatrix} \lambda_1 - \mu_1 x_1 \\ -\lambda_1 - \mu_2 x_2 \end{pmatrix} \text{ if } x_1 + x_2 = C,\ x_2 > 0;$$

$$Q(x_1,x_2) = \begin{pmatrix} -\mu_1 x_1 \\ 0 \end{pmatrix} \text{ if } (x_1,x_2) = (C,0).$$

The drift $\widetilde{Q}^N(x_1,x_2)$ for the corresponding scaled process presented in Definition 3.5 verifies

$$\frac{1}{N}\widetilde{Q}^N\left(\widetilde{X}_1^N, \widetilde{X}_2^N\right) = Q\left(X_1^N, X_2^N\right).$$

Proof. We compute the drift of (X_1, X_2) defined as:

$$Q(x_1,x_2) = \sum_{(x_1',x_2')\in E} q((x_1,x_2),(x_1',x_2'))\left[(x_1',x_2') - (x_1,x_2)\right].$$

$$Q(x_1,x_2) = \lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a(x_1,x_2)\lambda_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \mu_1 x_1 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \mu_2 x_2 \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$
$$\text{if } x_1 + x_2 < NC,$$
$$Q(x_1,x_2) = \lambda_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \mu_1 x_1 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \mu_2 x_2 \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$
$$\text{if } x_1 + x_2 = C \text{ and } x_2 > 0,$$
$$Q(x_1,0) = \mu_1 x_1 \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$
$$\text{if } x_1 = C \text{ and } x_2 = 0.$$

We compute the drift of $(\widetilde{X}_1^N, \widetilde{X}_2^N)$ defined as:

$$\widetilde{Q}^N(x_1,x_2) = \sum_{(x'_1,x'_2)\in\widetilde{E}^N} \widetilde{q}^N((x_1,x_2),(x'_1,x'_2)) \left[(x'_1,x'_2) - (x_1,x_2)\right].$$

$$\widetilde{Q}^N(x_1,x_2) = N\lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \widetilde{a}^N(x_1,x_2)N\lambda_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \mu_1 x_1 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \mu_2 x_2 \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$
$$\text{if } x_1 + x_2 < NC,$$
$$\widetilde{Q}^N(x_1,x_2) = N\lambda_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \mu_1 x_1 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \mu_2 x_2 \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$
$$\text{if } x_1 + x_2 = NC \text{ and } x_2 > 0,$$
$$\widetilde{Q}^N(x_1,0) = \mu_1 x_1 \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$
$$\text{if } x_1 = NC \text{ and } x_2 = 0.$$

Then we replace (x_1,x_2) by (Nx_1,Nx_2) in the previous equations and divide by N. We have that $\widetilde{a}^N(Nx_1,Nx_2) = a(x_1,x_2)$ for the three different admission control schemes presented in Section 3.1.

In deterministic admission control $a(x_1,x_2) = 1$ if $x_1+x_2 < \delta$ and $a(x_1,x_2) = 0$ if $x_1+x_2 \geq \delta$ and $\widetilde{a}^N(x_1,x_2) = 1$ if $x_1+x_2 < N\delta$ and $\widetilde{a}^N(x_1,x_2) = 0$ if $x_1+x_2 \geq N\delta$. Then $\widetilde{a}^N(Nx_1,Nx_2) = 1$ if $x_1+x_2 < \delta$ and $\widetilde{a}^N(Nx_1,Nx_2) = 0$ if $x_1+x_2 \geq \delta$. The same is true for the case of free admission control but considering $a(x_1,x_2) = 1$ if $x_1+x_2 < C$.

In the probabilistic admission control we have that, for all N,

$$\widetilde{a}^N(Nx_1,Nx_2) = a(x_1,x_2) = 1 - (x_1+x_2)/C.$$

According to that we can conclude that

$$\frac{1}{N}\widetilde{Q}^N\left(\widetilde{X}_1^N,\widetilde{X}_2^N\right)=Q\left(X_1^N,X_2^N\right).$$

□

Classical results on convergence of Markov processes assume some regularity properties of the fluid ODE, i.e. the vector field $Q(x_1,x_2)$ defining the ODE must be a Lipschitz continuous function in the domain of interest. It is a sufficient condition for existence and uniqueness of solutions given initial conditions. In our system this regularity condition does not always hold. In this context, using results obtained by Bortolussi in [Bor11a, Bor16], it is possible to determine a PWSDS that is the fluid limit.

In all cases (free, deterministic and probabilistic admission control) we have a discontinuous drift that leads to a differential equations with discontinuous right-hand side and to the presence of sliding motion, (see Appendix B) as we show in the following subsections.

3.2.1 Free admission control policy

In this subsection we assume that $a(x_1,x_2)=1$ for all $(x_1,x_2)\in R_1$ such that $R_1=\{(x_1,x_2):x_1+x_2-C<0\}$, 0 otherwise. The goal is to study the behavior of the system without any intervention: if a SU arrives and there is at least one idle subchannel, the SU will be served. Here we have discontinuous transition rates, as in the border of the state space $\gamma=\{(x_1,x_2):x_1+x_2-C=0\}$ we have $a(x_1,x_2)=0$. As we explained, these discontinuities lead to a deterministic limit whose trajectories are continuous but not differentiable. Moreover, limit trajectories will stay in the border of the state space then we will show the existence of sliding motion.

Proposition 3.7. Let $Q_1(x_1,x_2)$ and $Q_2(x_1,x_2)$ be vector fields, both smooth in R_1 and γ respectively such that

$$Q_1(x_1,x_2)=\begin{pmatrix}\lambda_1-\mu_1x_1\\ \lambda_2-\mu_2x_2\end{pmatrix},\qquad Q_2(x_1,x_2)=\begin{pmatrix}\lambda_1-\mu_1x_1\\ -\lambda_1-\mu_2x_2\end{pmatrix},$$

and let $n(x_1,x_2)$ be the normal vector to the boundary γ ($n^T(x_1,x_2)=(1,1)$ for all $(x_1,x_2)\in\gamma$) then we have a PWSDS driven by the following equations.

If $x_1+x_2<C$ or $\lambda_1+\lambda_2\le\mu_1x_1+\mu_2x_2$:

$$\begin{cases} x_1' &= \lambda_1-\mu_1x_1,\\ x_2' &= \lambda_2-\mu_2x_2.\end{cases},$$

and if $x_1 + x_2 = C$ and $\lambda_1 + \lambda_2 > \mu_1 x_1 + \mu_2 x_2$:

$$\begin{cases} x_1' &= \lambda_1 - \mu_1 x_1, \\ x_2' &= -\lambda_1 + \mu_1 x_1. \end{cases}.$$

Proof. Consider x on the border γ and $n(x)$ the normal vector to the border. We have that $n^T(x_1,x_2)Q_1(x_1,x_2) = 0 \Leftrightarrow \lambda_1 + \lambda_2 - \mu_1 x_1 - \mu_2 x_2 = 0$ and $n^T(x_1,x_2)Q_2(x_1,x_2) = 0 \Leftrightarrow -\mu_1 x_1 - \mu_2 x_2 = 0$. Then, for studying $n^T(x_1,x_2)Q_i(x_1,x_2)$ we have several cases depending on the position of the line $\lambda_1 + \lambda_2 - \mu_1 x_1 - \mu_2 x_2 = 0$. It is clear that it depends on the values of λ_1, λ_2, μ_1 and μ_2. In particular, we have that $n^T(x_1,x_2)Q_1(x_1,x_2) > 0$ if $\lambda_1 + \lambda_2 - \mu_1 x_1 - \mu_2 x_2 > 0$, Q_1 and n are tangent in the points over the line $\lambda_1 + \lambda_2 - \mu_1 x_1 - \mu_2 x_2 = 0$ and $n^T(x_1,x_2)Q_1(x_1,x_2) < 0$ if $\lambda_1 + \lambda_2 - \mu_1 x_1 - \mu_2 x_2 < 0$. On the other hand, $n^T(x_1,x_2)Q_2(x_1,x_2) < 0$ in γ independently of the parameters λ_i and μ_i. Because of this we are in the presence of sliding motion when $-\mu_1 x_1 + \mu_2 x_2 < \lambda_1 + \lambda_2$ and $x_1 + x_2 = C$. In that case we define the differential inclusion by a convex combination $g(x_1,x_2) = (1 - \alpha(x_1,x_2))Q_1 + \alpha(x_1,x_2)Q_2$ verifying $n^T(x_1,x_2)g(x_1,x_2) = 0$ (the solution cannot escape from the border). Computing $\alpha(x_1,x_2)$ we obtain

$$\alpha(x_1,x_2) = \frac{\mu_1 x_1 + \mu_2 x_2}{\lambda_1 + \lambda_2}$$

then substituting in $g(x_1,x_2)$ the result is proved. □

Let

$$(x_1(t), x_2(t)) \tag{3.3}$$

be the PWSDS that is the solution to the previous equations with initial condition $(x_1(0), x_2(0))$. Next we present the fluid limits results.

Theorem 3.8. *Consider the process $(\widetilde{X}_1^N, \widetilde{X}_2^N)$ with transition rates defined in Table 3.1, define:*

$$(X_1^N(t), X_2^N(t)) = (\widetilde{X}_1^N(t), \widetilde{X}_2^N(t))/N$$

and let (x_1,x_2) be the PWSDS defined in Equation (3.3) with initial condition $(x_1(0), x_2(0))$. If

$$\lim_{N\to\infty}(X_1^N(0), X_2^N(0)) = (x_1(0), x_2(0))$$

then, for all $T > 0$,

$$\lim_{N\to\infty} \sup_{t\in[0,T]} \left|(X_1^N(t), X_2^N(t)) - (x_1(t), x_2(t))\right| = 0$$

in probability.

Proof. The proof follows straightforward from Theorem IV.2 in [Bor11a]. The hypotheses of this Theorem are verified in our case; i.e. the scaling scheme of the process, as we have defined in Section 3.2 and sumarized in Table 3.1, and the existence of a unique PWSDS with regular trajectories, presented Subsection 3.2.1 in order to define Equation (3.3). Theorem IV.2 in [Bor11a] considers two different regions R_1 and R_2 with a border between them, where the process and the PWSDS may change many times from one region to the other and there may be several different pieces of sliding motion at the border. In our case we have only one region and the border, but the proof in [Bor11a] is suitable for this case. More specifically, the proof there consists in splitting the whole PWSDS trajectory in pieces in each region (where classical results, for example from [EK86], hold) and in sliding trajectories at the border. For that case the proof in [Bor11a] consists in replacing the discontinuous drift by the sliding vector field and prove convergence by an uniformization procedure. The proof of Lemma 3 in [Bor11b] that considers two different regions is valid for the case with a region and its border.

In our case, as we in fact have a switched linear system we only have one piece of sliding motion, and as our initial condition can only be in R_1 we have three cases. In the first case the deterministic trajectory stays all the time in R_1 and classical results hold. In the second case the trajectory starts in R_1 and then presents sliding motion and stays at the border for $t \to \infty$. In the last case the trajectory PSWDS starts on R_1 or at the border, then presents sliding motion and exits at the border, and then remains in R_1. In that case we need to check the exit conditions for the sliding motion, that are guaranteed because both vector fields are not tangential to the border at the same time (let us recall that Q_2 always points towards R_1). □

An alternative approach can be done following [SW95]. In Chapter 8 the authors define what they call flat boundary process, that is a Markov process where transitions at the border are different from transitions in the interior of the state space, and transition rates are discontinuous at the border. Under some regularity conditions the authors of [SW95] obtained very similar results than [Bor11a] they prove existence and uniqueness for the solution of the PWSDS and they prove convergence to such systems. In their presentation the coefficients of the convex combination defining the sliding vector field have a probabilistic interpretation. They consider a process with three coordinates, where the first two correspond to PUs and SUs as ours, and the third, valued 0 or 1 indicates if the system is at the border or not. With the same scaling as ours the third coordinate tends to

0 when scaled. In addition, as transition rates in the interior are smooth, while process stays in the interior the scaled process (dividing by N) changes a little. On the other hand, the number of jumps from the interior to the border per time unit increases with N, so the third coordinate that indicates where the process is reaches its invariant distribution before the number of PUs and SUs. In that interpretation the coefficient α is the probability or the proportion of time that the process spends at the border and $(1-\alpha)$ the the proportion of time that the process spends in the interior.

For further work concerning fluid limits, including systems with discontinuous rates we also refer to [GG12, JS14, Bor16].

In the context of fluid limits it is usual to infer from the fixed point analysis of the deterministic system the behavior of the stochastic one in the stationary regime. If there is a unique fixed point that is a global attractor, the stochastic invariant distributions converges in probability to this fixed point [BLB08, LB10b, TT17]. In what follows we will exploit this general result.

Proposition 3.9. Considering $a(x_1, x_2) = 1$ for all $(x_1, x_2) : x_1 + x_2 < C$, and 0 otherwise, letting R_1 and γ be the above defined zone and border and setting (x_1^*, x_2^*) as the PWSDS fixed point, then:

a. If $\frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\mu_2} < C$, then the fixed point $(x_1^*, x_2^*) = \left(\frac{\lambda_1}{\mu_1}, \frac{\lambda_2}{\mu_2}\right) \in R_1$ and the mean system utilization is $\frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\mu_2}$ (sub-critical case).

b. If $\frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\mu_2} \geq C$ and $\frac{\lambda_1}{\mu_1} < C$, then the fixed point $(x_1^*, x_2^*) = \left(\frac{\lambda_1}{\mu_1}, C - \frac{\lambda_1}{\mu_1}\right) \in \gamma$ and the mean system utilization is C (critical case).

c. If $\frac{\lambda_1}{\mu_1} \geq C$, then the fixed point $(x_1^*, x_2^*) = (C, 0) \in \gamma$ and the mean system utilization is C (critical case).

Proof. Let Q_1 and Q_2 be the velocity vectors, both continuous in R_1 and γ respectively:

$$Q_1(x_1, x_2) = \begin{pmatrix} \lambda_1 - \mu_1 x_1 \\ \lambda_2 - \mu_2 x_2 \end{pmatrix}, \; Q_2(x_1, x_2) = \begin{pmatrix} \lambda_1 - \mu_1 x_1 \\ -\lambda_1 - \mu_2 x_2 \end{pmatrix},$$

and let $n(x_1, x_2)$ be a normal vector to the line $\gamma = \{(x_1, x_2) : x_1 + x_2 = C\}$, so that $n^T(x_1, x_2)$ is collinear with $(1, 1)$ for all $(x_1, x_2) \in \gamma$.

As it is explained in B.1, in Appendix B when summarizing the different possible behaviors of both vector fields we have that all possible cases, for different values of the parameters, can be categorized into two groups represented by Case 1 and Case 2 from Figure 3.2. In that figure

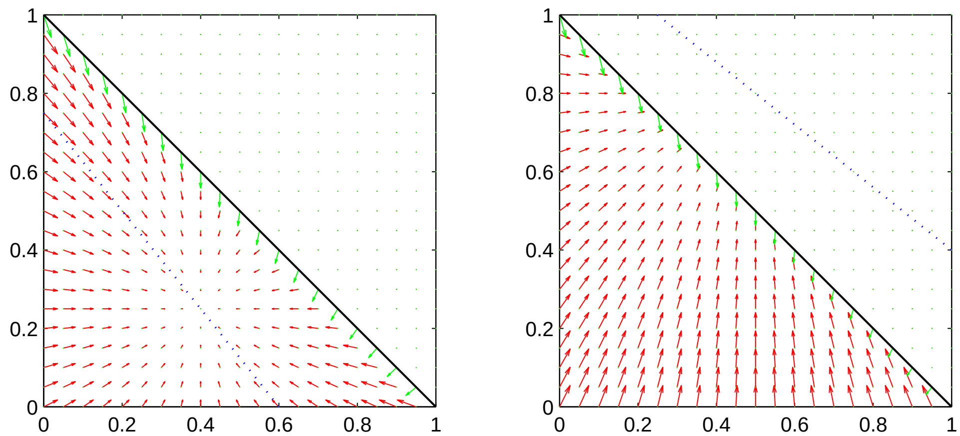

Figure 3.2: Vector field for Case 1 (left): $C = 1$, $\lambda_1 = 2$, $\lambda_2 = 1$, $\mu_1 = 5$, $\mu_2 = 4$ and Case 2 (right): $C = 1$, $\lambda_1 = 2$, $\lambda_2 = 4$, $\mu_1 = 4$, $\mu_2 = 5$. The solid line represents γ and the dotted line is $\lambda_1 + \lambda_2 - \mu_1 x_1 - \mu_2 x_2 = 0$. The region R_1 is defined by $R_1 = \{(x_1, x_2) : x_1 + x_2 < 1\}$.

the solid line represents γ, the dotted line is $\lambda_1 + \lambda_2 - \mu_1 x_1 - \mu_2 x_2 = 0$ and the vectors are Q_1 and Q_2. In Case 1 the PWSDS fixed point is in R_1 (Proposition 3.9.a); we call it sub-critical case. It is easy to note that $(x_1^*, x_2^*) = (\lambda_1/\mu_1, \lambda_2/\mu_2)$. On the other hand, in Case 2, the fixed point is on γ and its value is $(x_1^*, x_2^*) = (\lambda_1/\mu_1, C - \lambda_1/\mu_1)$ representing a critical case, (Proposition 3.9.b). When $\lambda_1/\mu_1 + \lambda_2/\mu_2 \geq C$ we can identify a sliding motion behavior near the fixed point, more precisely we can affirm that the equation solution will live on γ most of the time. Both examples of Figure 3.2 consider $\lambda_1/\mu_1 < C$. Finally, when the system is saturated by PUs ($\lambda_1/\mu_1 \geq C$), as a corollary from Proposition 3.9.b the fixed point is $(x_1^*, x_2^*) = (C, 0)$ (Proposition 3.9.c). □

In Figures 3.3 and 3.4 we show the deterministic approximation and a trajectory of the stochastic processes for the Cases 1 and 2 of Figure 3.2. In each one, in the left graphic we show the simulation of one trajectory of the scaled Markov process and the PWSDS. In the right we show for the same simulation the evolution on the plane of the Markov chain and the corresponding PWSDS. In Figure 3.3 the fixed point is in R_1 and in Figure 3.4 it is on γ. It is important to note that in both cases, for large time values, the scaled number of users in the stochastic process is around the PWSDS fixed point (x_1^*, x_2^*).

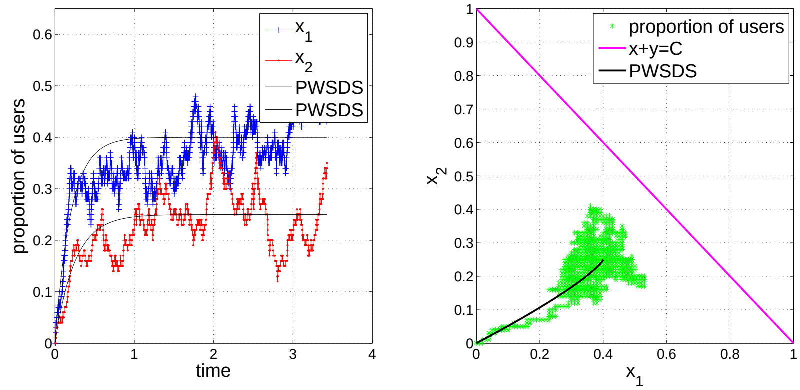

Figure 3.3: Case 1 with parameters: $N = 100$, $C = 1$, $\lambda_1 = 2$, $\lambda_2 = 1$, $\mu_1 = 5$ and $\mu_2 = 4$. The PWSDS fixed point is $(x_1^*, x_2^*) = (\lambda_1/\mu_1, \lambda_2/\mu_2) = (2/5, 1/4)$.

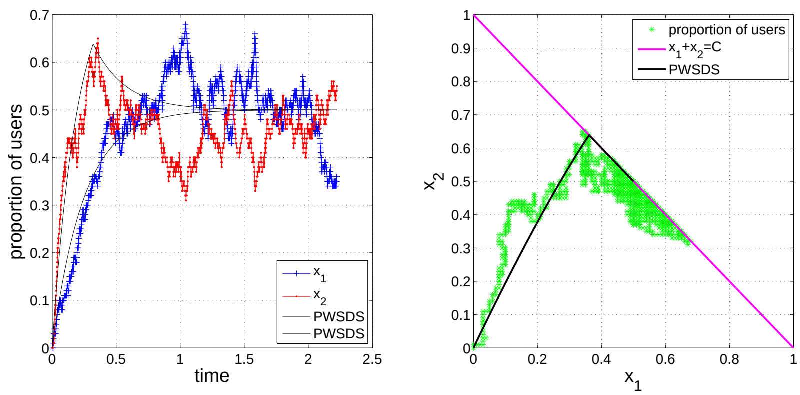

Figure 3.4: Case 2 with parameters: $N = 100$, $C = 1$, $\lambda_1 = 2$, $\lambda_2 = 4$, $\mu_1 = 4$ and $\mu_2 = 5$. The PWSDS fixed point is $(x_1^*, x_2^*) = (\lambda_1/\mu_1, C - \lambda_1/\mu_1) = (1/2, 1/2)$.

Recalling the description of the system, if $x_1 + x_2 = C$ and a PU arrives, a SU will be immediately deallocated giving the subchannel to the new PU. In this case, the QoS perceived by the SU will be affected because of the interruption of its communication. We are interested in SUs whose service cannot be interrupted with high probability (like a phone call or other interactive services). With this in mind, we can relate the interruption probability with the probability that the process lives on γ. We can conclude that the PWSDS fixed point has to be far enough from γ to avoid a strong impact on secondary communications. However, it has to be as close as possible to γ to permit more spectrum utilization and a high access probability for SUs. According to Proposition 3.9 and observing Figures 3.3 and 3.4 we can identify two cases: when the PWSDS fixed point is in R_1 and when it is on γ. For the last one, the system in stationary regime works near γ, so the probability of service interruption is too large. Then, the question in this case is how can we move the fixed-point? The analysis in the subsections 3.2.2 and 3.2.3 will be concentrated on answering that question. In particular we move the fixed point using admission control decisions. On the other hand, in the next two subsections we concentrate in cases like Case 1 (when the PWSDS fixed point is in R_1). More specifically, we concentrate our efforts on answering the question: is the fixed point far enough from γ to assure a small interruption probability?

We study the interruption probability by means of the probability that the system is full $P(X_1 + X_2 = C)$. Both probabilities are highly related by

$$P(\text{deallocate a SU}) + P(\text{block a PU}) = P(X_1 + X_2 = C|\text{PU arrives}),$$

by PASTA property (Poisson Arrivals See Time Averages) the last probability is $P(X_1 + X_2 = C)$ and $P(\text{block a PU})$ is the blocking probability for a $M/M/C/C$ queue.

In Section 3.3 we analyze the asymptotic distribution and provide some practical criteria for admission control.

3.2.2 Deterministic admission control policy

In this subsection we analyze the system when a deterministic admission control is applied. In this situation we consider $\gamma' = \{(x_1, x_2) : x_1 + x_2 = \delta\}$ as the admission control border, then the question is: what is a reasonable value of δ to guarantee certain level of QoS to SUs?

In this case the fluid limit is obtained in the same way as in Subsection 3.2.1, but the sliding motion occurs in γ'.

Proposition 3.10. Let $Q_1(x_1, x_2)$ and $Q_2(x_1, x_2)$ be vector fields, both smooth in $R_1 = \{(x_1, x_2) : x_1 + x_2 - \delta < 0\}$ and $R_2 = \{(x_1, x_2) : x_1 + x_2 - \delta > 0\}$ respectively such that

$$Q_1(x_1, x_2) = \begin{pmatrix} \lambda_1 - \mu_1 x_1 \\ \lambda_2 - \mu_2 x_2 \end{pmatrix}, \qquad Q_2(x_1, x_2) = \begin{pmatrix} \lambda_1 - \mu_1 x_1 \\ -\mu_2 x_2 \end{pmatrix},$$

and let $n(x_1,x_2)$ be the normal vector to the boundary γ' ($n^T(x_1,x_2) = (1,1)$ for all $(x_1,x_2) \in \gamma'$) then we have a PWSDS driven by the following equations.

If $x_1 + x_2 < \delta$:

$$\begin{cases} x_1' &= \lambda_1 - \mu_1 x_1, \\ x_2' &= \lambda_2 - \mu_2 x_2. \end{cases},$$

else, if $x_1 + x_2 - \delta = 0$ (γ'):

$$\begin{cases} x_1' &= \lambda_1 - \mu_1 x_1, \\ x_2' &= -\lambda_1 + \mu_1 x_1. \end{cases},$$

and if $x_1 + x_2 > \delta$:

$$\begin{cases} x_1' &= \lambda_1 - \mu_1 x_1, \\ x_2' &= -\mu_2 x_2. \end{cases}.$$

Notice that we consider the case $\frac{\lambda_1}{\mu_1} < C$ because its practical importance.

The behavior is very similar as in the free admission control case, and we can think in this case as a sort of translation of the free case where the border that presents sliding motion is shifted from γ to γ'. Then, the proof is totally analogous to Proposition 3.7.

A convergence theorem analogous to Theorem 3.8 holds, in this case using directly Theorem IV.2 in [Bor11a], as for the deterministic policy there are two regions $R_1' = \{(x_1,x_2) : x_1 + x_2 < \delta\}$, $R_2' = \{(x_1,x_2) : x_1 + x_2 > \delta\}$, and the border $\gamma' = \{(x_1,x_2) : x_1 + x_2 = \delta\}$, where the drift is discontinuous.

In Figure 3.5 we show the scaled Markov chain an its fluid limit.

In addition we can study fixed points, by an analysis similar to Proposition 3.9. However, the fixed point of the PWSDS in the most interesting cases lies in γ'. This fact hinders the development of a design criterion like in Subsection 3.3.1, as we do not have the same results for the asymptotic distribution.

3.2.3 Probabilistic admission control policy

Let us consider another class of admission control: a probabilistic admission mechanism where secondary users can access the system with a probability related to the number of users in the system. Let $a(x_1,x_2)$ be the probability that a secondary user that arrives enters the system when there are x_1 primary users and x_2 secondary users. Then, when computing the entry rates for the whole system, the arrival rates of secondary users appear multiplied by this probability. Let us assume that $a(x_1,x_2)$ is a Lipschitz

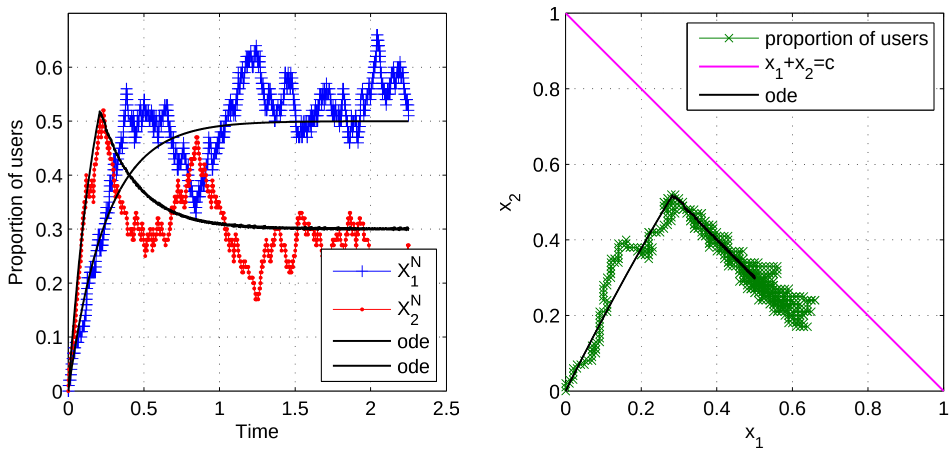

Figure 3.5: CR critical regime with deterministic admission control. Parameters: $N = 100$, $\lambda_1 = 2$, $\lambda_2 = 4$, $\mu_1 = 4$ and $\mu_2 = 5$, $\delta = 4/5$. The PWSDS fixed point is $(x_1^*, x_2^*) = (\lambda_1/\mu_1, \delta - \lambda_1/\mu_1) = (1/2, 3/10)$.

function that vanishes close to the border $\{(x_1, x_2) : x_1 + x_2 = C\}$ (as an example we could choose $a(x_1, x_2) = 1 - (x_1 + x_2)/C$), that is:

$$\lim_{x_1+x_2 \to C} a(x_1, x_2) = 0.$$

In this case we also have discontinuous transition rates in the border of the state space. Following the same lines as for the free admission case in Subsection 3.2.1 Proposition 3.7 we have:

Proposition 3.11. Let $Q_1(x_1, x_2)$ and $Q_2(x_1, x_2)$ be vector fields, both smooth in $R_1 = \{(x_1, x_2) : x_1 + x_2 - C < 0\}$ and γ respectively such that

$$Q_1(x_1, x_2) = \begin{pmatrix} \lambda_1 - \mu_1 x_1 \\ \lambda_2 a(x_1, x_2) - \mu_2 x_2 \end{pmatrix}, \qquad Q_2(x_1, x_2) = \begin{pmatrix} \lambda_1 - \mu_1 x_1 \\ -\lambda_1 - \mu_2 x_2 \end{pmatrix},$$

and let $n(x_1, x_2)$ be the normal vector to the boundary γ ($n^T(x_1, x_2) = (1, 1)$ for all $(x_1, x_2) \in \gamma$) then we have a PWSDS driven by the following equations.

If $x_1 + x_2 < C$ or $\lambda_1 \leq \mu_1 x_1 + \mu_2 x_2$:

$$\begin{cases} x_1' &= \lambda_1 - \mu_1 x_1, \\ x_2' &= \lambda_2 a(x_1, x_2) - \mu_2 x_2. \end{cases}$$

If $x_1 + x_2 = C$ and $\lambda_1 > \mu_1 x_1 + \mu_2 x_2$:

$$\begin{cases} x_1' &= \lambda_1 - \mu_1 x_1, \\ x_2' &= -\lambda_1 + \mu_1 x_1. \end{cases}$$

This system has a unique solution with initial condition $(x_1(0), x_2(0))$. Let:

$$(x_1(t), x_2(t)) \tag{3.4}$$

be the piecewise smooth dynamical system that is the solution to the previous equations with initial condition $(x_1(0), x_2(0))$. Theorem 3.8 holds in this context.

This probabilistic admission control is different from the deterministic one. The solution presents sliding motion when $\lambda_1/\mu_1 \geq C$ or when $\lambda_1/\mu_2 \geq C$. In the first case the system in stationary regime is always saturated by PUs, despite of the admission control. In the second case sliding motion depends on the initial condition but it does not influence the stationary regime.

Let us consider the fixed point (x_1^*, x_2^*) for the processes defined by (3.4). For $\theta_1 = \lambda_1/\mu_1 < C$ and $\theta_2 = \lambda_2/\mu_2$, we have $x_1^* = \theta_1$ and x_2^* verifies the equation $x_2^* = \theta_2 a(\theta_1, x_2^*)$. Therefore, x_2^* is unique if we assume that the equation $\theta_2 a(\theta_1, x) - x = 0$ has an unique solution. In the example, with $a(x_1, x_2) = 1 - x_1 + x_2/C$, x_2^* can be obtained explicitly.

Proposition 3.12. Let $a(x_1, x_2)$ continuous with $\lim_{x_1+x_2 \to C} a(x_1, x_2) = 0$ and (x_1^*, x_2^*) as the PWSDS fixed point, then:

a. If $\theta_1 < C$, then $x_1^* = \frac{\lambda_1}{\mu_1}$ and $x_2^* = \theta_2 a(\theta_1, x_2^*)$ where $(x_1^*, x_2^*) \in R_1$ and the mean system utilization will be $\theta_1 + \theta_2 a(\theta_1, x_2^*)$.

b. If $\theta_1 \geq C$, then $(x_1^*, x_2^*) = (C, 0)$ and the mean system utilization will be C.

Proof. Results follows from similar arguments than in Proposition 3.9. In addition we could assume hypotheses about the probabilistic admission control function $a(x_1, x_2)$ that ensures asymptotic stability of the solutions (for example negative real part of eigenvalues for the linearized system). This condition follows in the example $a(x_1, x_2) = 1 - x_1 + x_2/C$, where we have a linear ODE. (See Figure 3.6 in order to analyze the vector field behavior in two different cases.) □

In particular in Figure 3.7 we show the simulation of one trajectory of the scaled Markov process and the trajectory of the ODE.

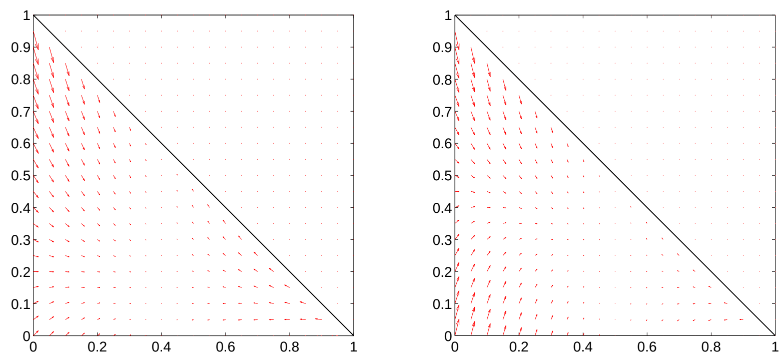

Figure 3.6: Vector field for Case 1 (left): $N = 100$, $C = 1$, $\lambda_1 = 2$, $\lambda_2 = 1$, $\mu_1 = 5$, $\mu_2 = 4$ and Case 2 (right): $N = 100$, $C = 1$, $\lambda_1 = 2$, $\lambda_2 = 4$, $\mu_1 = 4$, $\mu_2 = 5$. The continuous line represents γ.

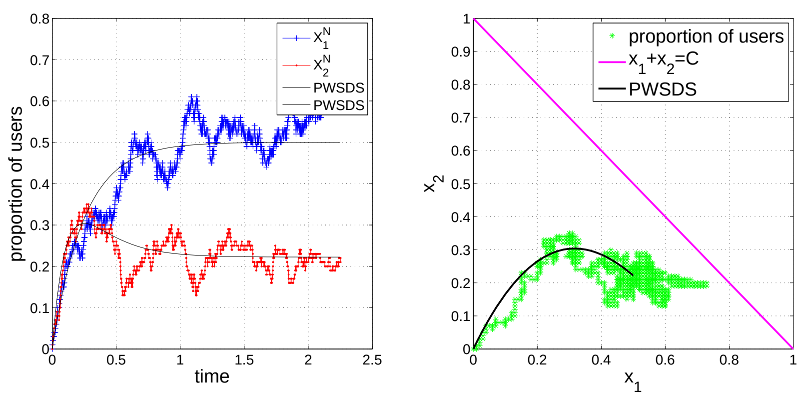

Figure 3.7: Probabilistic control ($a(x_1, x_2) = 1 - x_1 - x_2$), parameters: $N = 100$, $C = 1$, $\lambda_1 = 2$, $\lambda_2 = 4$, $\mu_1 = 4$ and $\mu_2 = 5$. The PWSDS fixed point, with $\theta_1 = \frac{\lambda_1}{\mu_1}$ and $\theta_2 = \frac{\lambda_2}{\mu_2}$, is $(x_1^*, x_2^*) = \left(\theta_1, \theta_2\left(\frac{C-\theta_1}{C+\theta_2}\right)\right) = (\frac{1}{2}, \frac{2}{9})$.

3.3 Asymptotic distribution

3.3.1 Gaussian asymptotic distribution in sub-critical cases

Another issue studied in the context of fluid limits is the velocity of this convergence, by looking at the fluctuations of the process around the limit. We refer to [EK86, Rob03, HW81, SW95] for the analysis of different scaling regimes and limit theorems in that sense, concerning different kinds of limit distributions.

If we consider cases when $\lambda_1/\mu_1 + \lambda_2/\mu_2 < C$ and the PWSDS trajectory remains all the time in R_1, it is possible to apply known results (see Theorem 2.3 of Chapter 11 in [EK86]) in order to obtain the asymptotic distribution. Let (x_1, x_2) be the trajectory of the PWSDS with initial condition $(x_1(0), x_2(0))$. If:

$$\lim_{N\to+\infty} \sqrt{N}[(X_1^N(0), X_2^N(0)) - (x_1(0), x_2(0))] = \chi(0)$$

with $\chi(0)$ deterministic, then:

$$\sqrt{N}[(X_1^N, X_2^N) - (x_1, x_2)] \Rightarrow \chi,$$

where $(\chi(t))$ is a two-dimensional Gaussian process and $\Rightarrow$ means convergence in distribution. $(\chi(t))$ has a covariance matrix determined explicitly by:

$$Cov(\chi(t), \chi(r)) = \int_0^{t\wedge r} e^{A(t-s)} G(x_1(s), x_2(s)) e^{A(r-s)} ds$$

where

$$A = \begin{pmatrix} -\mu_1 & 0 \\ 0 & -\mu_2 \end{pmatrix}$$

and

$$G(x_1(s), x_2(s)) = \begin{pmatrix} \lambda_1 + \mu_1 x_1(s) & 0 \\ 0 & \lambda_2 + \mu_2 x_2(s) \end{pmatrix}.$$

In sub-critical cases, we can conclude that $\lim_{N\to+\infty} P(X_1^N(t) + X_2^N(t) = C) = 0$ for all t. However, considering the defined Gaussian process and a finite large N we can present a practical criterion to analyze if the PWSDS fixed point is far enough from γ. In particular, we can obtain confidence bounds and also infer an adequate number of subchannels in order to avoid a high interruption probability for SUs.

Practical QoS design criterion in sub-critical cases

As we have shown in the simulated examples of Figures 3.3 and 3.4, the fluid limit is an excellent approximation when N is large. Then, in practice, a possible criterion to determine whether the PWSDS fixed point is far enough from γ would be to consider a certain confidence region of $(X_1^N(t), X_2^N(t))$ assuming a large value of t. If the resulting confidence ellipse is entirely inside R_1, certain probability of non-interruption is guaranteed. Otherwise, we should try to move the fixed point. In particular we consider $\lim_{t \to +\infty} G(x_1(s), x_2(s)) = G(x_1^*, x_2^*)$ and define this limit matrix as $G(\infty) = G(x_1^*, x_2^*)$. The, following the development in [QS04] we can obtain the covariance matrix $\Sigma(\infty)$ by solving:

$$A\Sigma(\infty) + \Sigma(\infty)A^T = -G(\infty). \tag{3.5}$$

Considering a fixed relation between both classes $\frac{\lambda_1/\mu_1}{\lambda_2/\mu_2} = constant$, using the deterministic confidence ellipse we can infer which is the ideal scaling parameter. In other words, we can obtain an idea of the optimal number of resources (subchannels) necessary to guarantee a small interruption probability for SUs. In Figure 3.8 we show the theoretical confidence ellipses considering different values of N for two different parameter sets (Case A and Case B). In particular, we have considered $C = 1$, then N represents the number of channels of the system $(\widetilde{X}_1^N(t), \widetilde{X}_2^N(t))$. For Case A, we have that the ellipse is tangent to γ when $N = 180$. Therefore, we can conclude that an admission control does not make sense in the system $(\widetilde{X}_1^N(t), \widetilde{X}_2^N(t))$ when $N \geq 180$. In Table 3.2 we confirm that the interruption probability of SUs can be analyzed studying the probability that the system is full of users (blocking probability of PUs can be approximated by 0). For Case B, we have an analogous conclusion when $N \geq 120$.

Table 3.2: Full system and blocking probabilities. Values of full system probability and blocking PU probability for Case A of Figure 3.8. $P(\widetilde{X}_1^N(t) + \widetilde{X}_2^N(t) > C)$ is approximated by a Gaussian distribution.

N	100	120	150
$P(\widetilde{X}_1^N(t) + \widetilde{X}_2^N(t) > C)$	0.031	0.0205	0.0111
P(block a PU)	0.5×10^{-40}	0.8×10^{-48}	0

3.3.2 Non-Gaussian invariant distribution

In Subsection 3.2.2 we said that for the deterministic admission control policy we can study fixed points by an analysis similar to Proposition 3.9. In the sub-critical regime, that is when the fixed point lies in the interior, we have the same asymptotic Gaussian distribution that for the free admission control case, so the proposal of Subsection 3.3.1 can be applied also for a

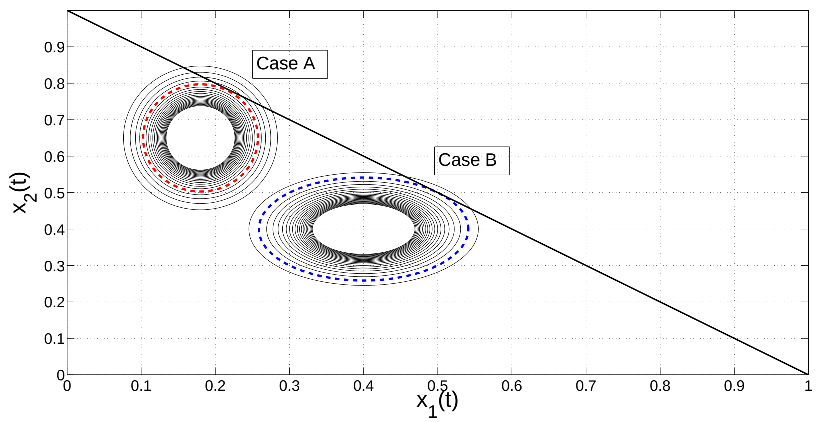

Figure 3.8: Theoretical 95% confidence ellipses of $(X_1^N(t), X_2^N(t))$, considering different N values ($N = 100, 120, 140, 160, \ldots, 500$). Parameters of Case A: $C = 1$, $\lambda_1 = 9$, $\lambda_2 = 13$, $\mu_1 = 50$ and $\mu_2 = 20$. Parameters of Case B: $C = 1$, $\lambda_1 = 10$, $\lambda_2 = 12$, $\mu_1 = 25$ and $\mu_2 = 30$.

deterministic admission control in the sub-critical case. However, the fixed point of the PWSDS in the most interesting cases lies in the border γ', where a design criterion like in Subsection 3.3.1 is not suitable, as we do not have the same results for the asymptotic distribution.

As in the deterministic policy scheme the hypotheses of Theorem 2.3 of Chapter 11 in [EK86] are not verified, due to the non-smoothness of the PWSDS at the border γ', we consider a different asymptotic regime.

We have convergence of the stationary regime to the PWSDS fixed point but there is not a general framework that allows to state a Gaussian asymptotic distribution. Therefore, we proceed to compute the asymptotic stationary distribution of the total number of users in the particular case when service rates are the same for primary and secondary users.

The limitation to a particular case comes from our proof method that depends on the explicit stationary distribution, that can only be explicitly obtained for equal service rates. However, we will give some insight on the asymptotic distribution in this case, for different service rates, based on simulations. Further study of the general case is object of future work.

Using the limit of the invariant distribution for the particular case, we can build some design criteria in order to find an estimation of the optimal δ value for certain maximum level of interruption probability. In˘citeRattaro2018 and [RABM19] it is shown that the criteria can be extended to the case

$\mu_1 \neq \mu_2$ with good results for practical purposes.

We consider the original system (when $\mu_1 = \mu_2$) and compute its invariant distribution, both for the original system and the scaled one. Then for the scaled system that depends on N we find the limit of the invariant distribution when N goes to infinity. For practical purposes as in this case, the invariant distribution can be computed explicitly. There is no need to compute the asymptotic distribution, but theoretically we want to show that the asymptotic distribution in the case where the equilibrium lies in the admission control border is not Gaussian; with a different scaling, we find a geometric asymptotic distribution. The proof is only in the particular case when service rates are the same for primary and secondary users, that is when the problem is one-dimensional. The scaling, different from $\sqrt{N}$, and the geometric distribution are not frequent in related works.

Consider the Markov chain defined in Section 3.1 with $\mu_1 = \mu_2$ where the access control is defined considering γ'. In this case $X = X_1 + X_2$ is a one dimensional Markov chain with state space $E = \{0, 1, \ldots, C\}$, and non-zero transition rates from i to j, $q(i,j)$, given by:

$$q(i,i+1) = \begin{cases} \lambda_1 + \lambda_2 & \text{for } 0 \leq i < \delta \\ \lambda_1 & \text{for } \delta \leq i < C \end{cases}, \qquad q(i,i-1) = i\mu \text{ for } 0 < i \leq C,$$

where $\delta \in E$ denotes the border of the admission control, that is when we have δ or more users we prohibit the access of new secondary users. Let us observe that in the general case described in Section 3.1 the total number of users $X = X_1 + X_2$ is Markovian only if $\mu_1 = \mu_2$, so the reduction to a one dimensional Markov chain follows only in this case. In order to simplify notation in what follows let us call $\nu_2 = \lambda_1 + \lambda_2$, and $\nu_1 = \lambda_1$. We then have the following transition rates, where $\nu_1 < \nu_2$:

$$q(i,i+1) = \begin{cases} \nu_2 & \text{for } 0 \leq i < \delta \\ \nu_1 & \text{for } \delta \leq i < C \end{cases}, \qquad q(i,i-1) = i\mu \text{ for } 0 < i \leq C,$$

Table 3.3: Scaling for the one dimensional processes (X) and $(\widetilde{X}^N)$.

(X)	ν_1	ν_2	C	δ
$(\widetilde{X}^N)$	$\nu_1 N$	$\nu_2 N$	CN	δN

We consider the scaled process $\widetilde{X}^N = \widetilde{X}_1^N + \widetilde{X}_2^N$ with the scaling in Table 3.3. Theorem 3.8 holds in this case and we also have that $X^N =$

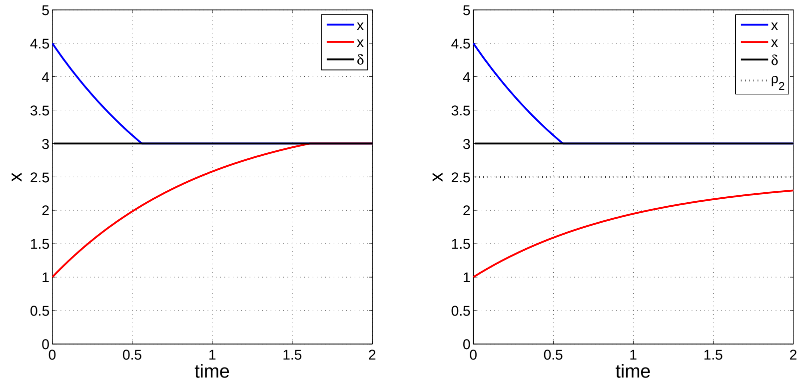

Figure 3.9: PWSDS solution in the one dimensional case for two different initial conditions. Parameters: $C = 5$, $\delta = 3$, $\rho_1 = 1$, $\rho_2 = 3.5$ (left) and $C = 5$, $\delta = 3$, $\rho_1 = 1$, $\rho_2 = 2.5$ (right).

$\widetilde{X}^N/N$ converges in probability, uniformly over compact time intervals to x, given by the following equations, for $0 \leq \delta \leq C$.

If the initial condition is $x(0) < \delta$:

$$x(t) = \begin{cases} \rho_2 + (x(0) - \rho_2)e^{-\mu t} & \text{if } t < \tau_2 \\ \delta & \text{if } t \geq \tau_2. \end{cases}$$

If the initial condition is $x(0) > \delta$:

$$x(t) = \begin{cases} \rho_1 + (x(0) - \rho_1)e^{-\mu t} & \text{if } t < \tau_1 \\ \delta & \text{if } t \geq \tau_1, \end{cases}$$

where $\rho_1 = \dfrac{\nu_1}{\mu}$, $\rho_2 = \dfrac{\nu_2}{\mu}$, $\tau_1 = \dfrac{1}{\mu} \log\left(\dfrac{x(0) - \rho_1}{\delta - \rho_1}\right)$ and $\tau_2 = \dfrac{1}{\mu} \log\left(\dfrac{x(0) - \rho_2}{\delta - \rho_2}\right)$. (See Figure 3.9 for different behaviors of these solutions.)

Now we come back to the original system. As the problem now is one-dimensional we will explicitly compute the stationary distribution of X, and then for the scaled system $\widetilde{X}^N$ for fixed N and obtain its limit when $N \to \infty$. X is a modification of a $M/M/C/C$ queue, where the arrivals changes their rate depending on the number of clients in the queue. We will analyze our system as in [Rob03] for the $M/M/C/C$ queue. Computing the stationary distribution π for this Markov chain:

$$\begin{aligned} \pi(i)\nu_2 &= \pi(i+1)(i+1)\mu \text{ for } 0 \leq i < \delta, \\ \pi(i)\nu_1 &= \pi(i+1)(i+1)\mu \text{ for } \delta \leq i < C. \end{aligned}$$

Then,

$$\pi(i) = \frac{\rho_2^i}{i!}\pi(0) \quad \text{for} \quad 0 \le i \le \delta,$$
$$\pi(i) = \left(\frac{\rho_2}{\rho_1}\right)^{\delta} \frac{\rho_1^i}{i!}\pi(0) \quad \text{for} \quad \delta < i \le C,$$

where:

$$\pi(0)^{-1} = \sum_{i=0}^{\delta} \frac{\rho_2^i}{i!} + \left(\frac{\rho_2}{\rho_1}\right)^{\delta} \sum_{i=\delta+1}^{C} \frac{\rho_1^i}{i!}.$$

The stationary distribution $\widetilde{\pi}^N$ for the process $\widetilde{X}^N$ is:

$$\widetilde{\pi}^N(i) = \frac{(N\rho_2)^i}{i!}\widetilde{\pi}^N(0) \quad \text{for} \quad 0 \le i \le N\delta,$$
$$\widetilde{\pi}^N(i) = \left(\frac{\rho_2}{\rho_1}\right)^{N\delta} \frac{(N\rho_1)^i}{i!}\widetilde{\pi}^N(0) \quad \text{for} \quad N\delta < i \le C,$$

with:

$$\widetilde{\pi}^N(0)^{-1} = \sum_{i=0}^{N\delta} \frac{(N\rho_2)^i}{i!} + \left(\frac{\rho_2}{\rho_1}\right)^{N\delta} \sum_{i=N\delta+1}^{NC} \frac{(N\rho_1)^i}{i!}.$$

Theorem 3.13. *Consider the original processes with equal service rates $\mu_1 = \mu_2$ and X^N defined as before. The, the stationary distribution of $\widetilde{X}^N - N\delta$ converges to the distribution of an integer variable Z given by*

$$P(Z = j) = \begin{cases} \rho\left(\frac{\rho_2}{\delta}\right)^j & \text{if } j < 0, \\ \\ \rho\left(\frac{\rho_1}{\delta}\right)^j & \text{if } j \ge 0, \end{cases}$$

where $\rho_1 = \dfrac{\nu_1}{\mu}$, $\rho_2 = \dfrac{\nu_2}{\mu}$, $\rho = \left(\dfrac{\rho_1}{\delta - \rho_1} + \dfrac{\rho_2}{\rho_2 - \delta}\right)^{-1}$.

Proof. Let us compute, for $0 \le k < N\delta$ the stationary distribution $\widetilde{\pi}^N(N\delta -$

k). Computing its inverse:

$$\begin{aligned}
\widetilde{\pi}^N(N\delta-k)^{-1} &= \frac{(N\delta-k)!}{(N\rho_2)^{N\delta-k}}\widetilde{\pi}^N(0)^{-1}\\
&= \frac{(N\delta-k)!}{(N\rho_2)^{N\delta-k}}\left(\sum_{i=0}^{N\delta}\frac{(N\rho_2)^i}{i!}+\left(\frac{\rho_2}{\rho_1}\right)^{N\delta}\sum_{i=N\delta+1}^{NC}\frac{(N\rho_1)^i}{i!}\right)\\
&= \frac{(N\delta-k)!}{(N\rho_2)^{N\delta-k}}\left(\sum_{j=0}^{N\delta}\frac{(N\rho_2)^{N\delta-j}}{(N\delta-j)!}+\left(\frac{\rho_2}{\rho_1}\right)^{N\delta}\sum_{i=N\delta+1}^{NC}\frac{(N\rho_1)^i}{i!}\right)\\
&= \sum_{j=0}^{N\delta}\frac{(N\rho_2)^k(N\delta-k)!}{(N\rho_2)^j(N\delta-j)!}+\sum_{j=1}^{N(C-\delta)}\frac{(N\rho_1)^j(N\rho_2)^k(N\delta-k)!}{(N\delta+j)!}\\
&= \rho_2^k\sum_{j=0}^{N\delta}\rho_2^{-j}\frac{N^k(N\delta-k)!}{N^j(N\delta-j)!}+\sum_{j=1}^{N(C-\delta)}\frac{(N\rho_1)^j(N\rho_2)^k(N\delta-k)!}{(N\delta+j)!}.
\end{aligned}$$

Using Stirling's formula we obtain:

$$\lim_{N\to\infty}\frac{N^k(N\delta-k)!}{N^j(N\delta-j)!}=\frac{\delta^j}{\delta^k}$$

and then for the first term, using dominated convergence we have:

$$\lim_{N\to\infty}\rho_2^k\sum_{j=0}^{N\delta}\rho_2^{-j}\frac{N^k(N\delta-k)!}{N^j(N\delta-j)!}=\left(\frac{\rho_2}{\delta}\right)^k\frac{1}{1-\dfrac{\delta}{\rho_2}}$$

For the second term we have using Stirling's formula that:

$$\lim_{N\to\infty}\frac{(N\rho_1)^j(N\rho_2)^k(N\delta-k)!}{(N\delta+j)!}=\left(\frac{\rho_1}{\delta}\right)^j\left(\frac{\rho_2}{\delta}\right)^k,$$

then, with dominated convergence:

$$\lim_{N\to\infty}\sum_{j=1}^{N(C-\delta)}\frac{(N\rho_1)^j(N\rho_2)^k(N\delta-k)!}{(N\delta+j)!}=\frac{\rho_1}{\delta-\rho_1}\left(\frac{\rho_2}{\delta}\right)^k$$

and, for $0\le k<N\delta$, we obtain:

$$\lim_{N\to\infty}\widetilde{\pi}^N(N\delta-k)=\rho\left(\frac{\delta}{\rho_2}\right)^k.$$

If $\widetilde{Z}^N = \widetilde{X}^N - N\delta$ we have that the stationary distribution $\widetilde{\mu}^N$ of $\widetilde{Z}^N$ verifies, for $j < 0$, that:

$$\lim_{N\to\infty} \widetilde{\mu}^N(j) = \lim_{N\to\infty} \widetilde{\pi}^N(N\delta - (-j)) = \rho\left(\frac{\rho_2}{\delta}\right)^j .$$

In the same way we compute, for $0 \leq k \leq NC - N\delta$, $\widetilde{\pi}^N(N\delta + k)$. □

In Figure 3.10 we show the limit distribution and the stationary distribution for different values of N. We see that the limit distribution represents a good estimation of the stationary one for all N values.

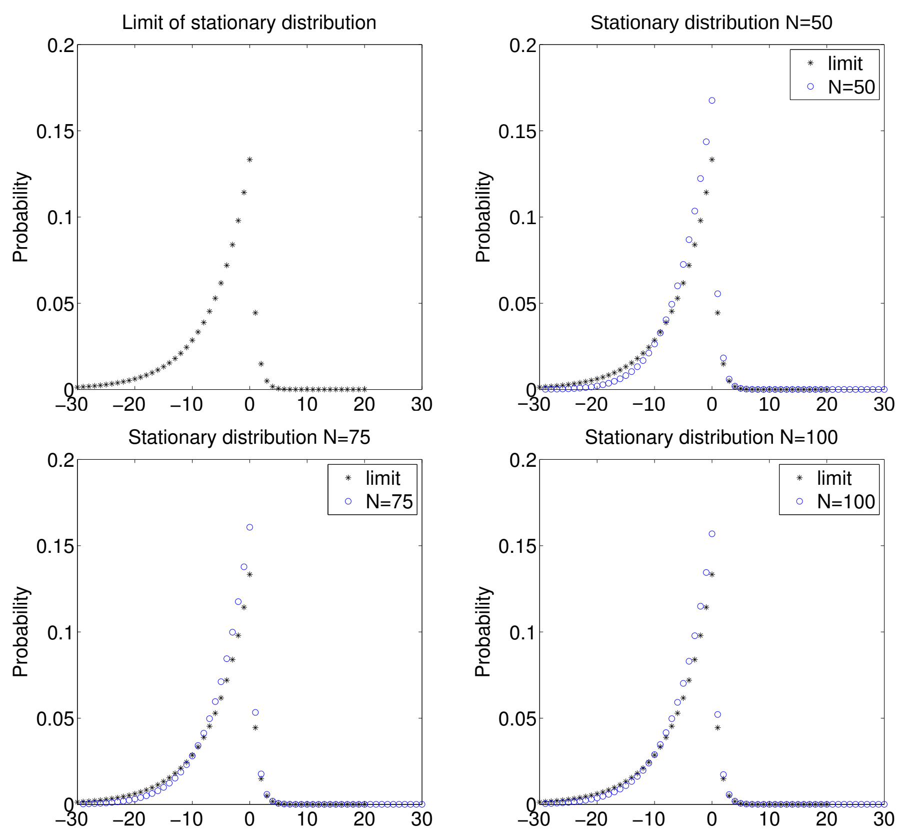

Figure 3.10: Limit ($N \to \infty$) of stationary distribution and stationary distributions for different N values. Parameters: $C = 5$, $\delta = 3$, $\rho_1 = 1$, $\rho_2 = 3.5$.

Simulations for the general case with different service rates

The asymptotic distribution obtained in Theorem 3.13 has the drawback that assumes equal services rates for PUs and SUs. This hypothesis is needed because of the technique that we used to prove Theorem 3.13. Our proof, based on [Rob03], relies on the explicit computation of the stationary distribution for the original system and for the scaled one, and in the calculus of the limit when the scaling factor N goes to infinity. This approach is not suitable for different service rates as there is not a closed formula for the stationary distribution. Then, a different approach would be necessary to obtain a non-Gaussian asymptotic distribution in the general case.

Despite of this restriction, in [Rat18] the author provides a practical QoS criterion for δ selection, applying then our results about geometric distribution to the case of different services rates in order to design an admission control policy. In addition, we notice that for different service rates, when the fixed point is in the admission control border, the asymptotic distribution is non-Gaussian.

Here we analyze with simulations the case with different service rates. We consider a deterministic admission control where the fixed point (x_1^*, x_2^*) of the PWSDS is at the admission control border γ'. We observe that in this case we can not obtain a Gaussian distribution as it is obtained when the fixed point is in the interior the region $\{(x_1, x_2) : x_1 + x_2 < \delta\}$. In this last case it actually does not matter if there is a deterministic admission control at $x_1 + x_2 = \delta$ or we consider the free admission control policy, in both cases, if the fixed point is in the interior the behavior is similar. On the other hand, when the fixed point lies in γ', we observe that fluctuations around the fixed point are asymmetric and much smaller than in the case of the interior fixed point, so that scaled by $\sqrt{N}$ as in the Gaussian case we would not obtain such limit.

We simulate two cases, with different parameters and the same deterministic admission control, one with the fixed point in the interior an the other at γ'. In both cases the starting point is the fixed point. For both case we show the trajectories in Figure 3.11. The case at the right is more asymmetrical and shows less dispersion around the fixed point than the left one. In Figure 3.12 we show the kernel density estimations of the normalized fluctuations of $X_1^N + X_2^N$ around the mean system utilization $x_1^* + x_2^*$ and in Figure 3.13 the corresponding QQ-plots. In both pictures it is more visible the asymmetry but specially that the distribution at the right is concentrated at zero, and cannot be approximated by a Gaussian.

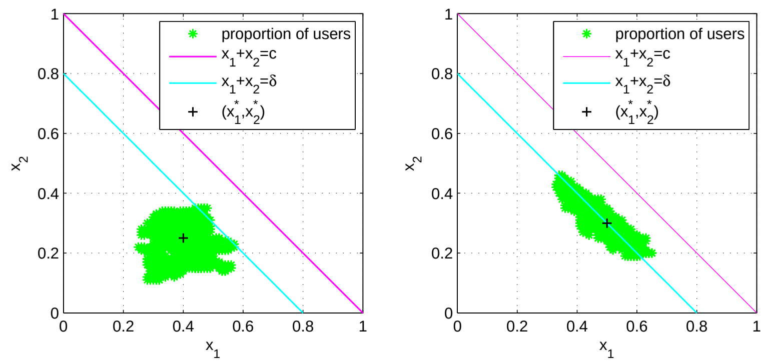

Figure 3.11: Left: $N = 100$, $C = 1$, $\delta = 0.8$ $\lambda_1 = 2$, $\lambda_2 = 1$, $\mu_1 = 5$ and $\mu_2 = 4$. The PWSDS fixed point is $(2/5, 1/4)$. Right: $N = 100$, $C = 1$, $\delta = 0.8$ $\lambda_1 = 2$, $\lambda_2 = 4$, $\mu_1 = 4$ and $\mu_2 = 5$. The PWSDS fixed point is $(1/2, 3/10)$.

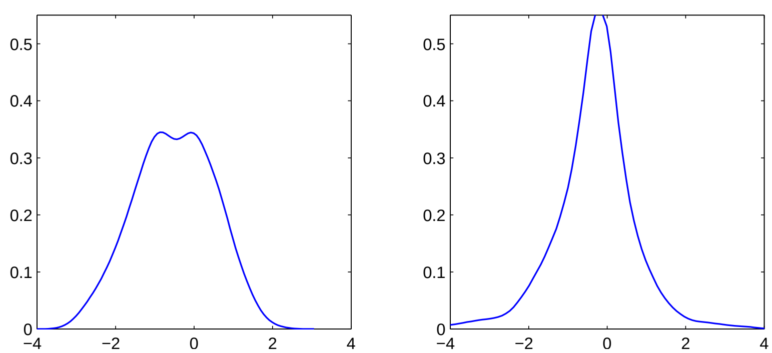

Figure 3.12: Left: $N = 100$, $C = 1$, $\delta = 0.8$ $\lambda_1 = 2$, $\lambda_2 = 1$, $\mu_1 = 5$ and $\mu_2 = 4$, kernel density estimation for the normalized variable $X_1^N + X_2^N - 0.65$. Right: $N = 100$, $C = 1$, $\delta = 0.8$ $\lambda_1 = 2$, $\lambda_2 = 4$, $\mu_1 = 4$ and $\mu_2 = 5$, kernel density estimation for the normalized variable $X_1^N + X_2^N - 0.8$.

Asymptotic distribution for probabilistic admission control

The previous discussion can also be extended for probabilistic admission control.

Remark 3.14. For case a. of Proposition 3.12, for initial conditions where there is not sliding motion, it is possible to apply the same results about Gaussian limits. Let (x_1, x_2) be the trajectory of the solution with initial

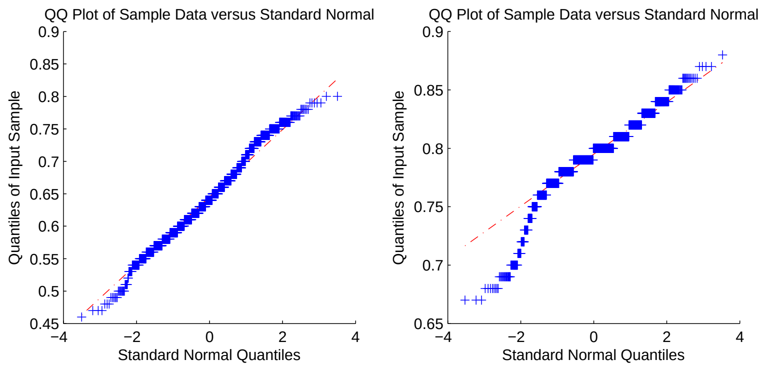

Figure 3.13: Left: $N = 100$, $C = 1$, $\delta = 0.8$ $\lambda_1 = 2$, $\lambda_2 = 1$, $\mu_1 = 5$ and $\mu_2 = 4$. Right: $N = 100$, $C = 1$, $\delta = 0.8$ $\lambda_1 = 2$, $\lambda_2 = 4$, $\mu_1 = 4$ and $\mu_2 = 5$.

condition $(x_1(0), x_2(0))$. If:

$$\lim_{N\to+\infty} \sqrt{N}[(X_1^N(0), X_2^N(0)) - (x_1(0), x_2(0))] = \chi(0)$$

with $\chi(0)$ deterministic, then $\sqrt{N}[(X_1^N, X_2^N) - (x_1, x_2)] \Rightarrow \chi$, where $(\chi(t))$ is a Gaussian process whose matrix can be determined explicitly by:

$$Cov(\chi(t), \chi(r)) = \int_0^{t\wedge r} e^{A(t-s)} G(x_1(s), x_2(s)) e^{A(r-s)} ds$$

where

$$A = \begin{pmatrix} -\mu_1 & 0 \\ 0 & -\mu_2 \end{pmatrix},$$

$$G(x_1(s), x_2(s)) = \begin{pmatrix} \lambda_1 + \mu_1 x_1(s) & 0 \\ 0 & \lambda_2 a(x_1(s), x_2(s)) + \mu_2 x_2(s) \end{pmatrix}.$$

In the same way as in Section 3.2.1, considering a large finite value of N we can calculate the covariance matrix $\Sigma(\infty)$ solving Equation (3.5). For instance in Figure 3.14 we illustrate different confidence ellipses (for different N values) for the case of Figure 3.7.

Remark 3.15. For case b. of Proposition 3.12, where there is sliding motion, it is possible to apply a similar approach as in Theorem 3.13.

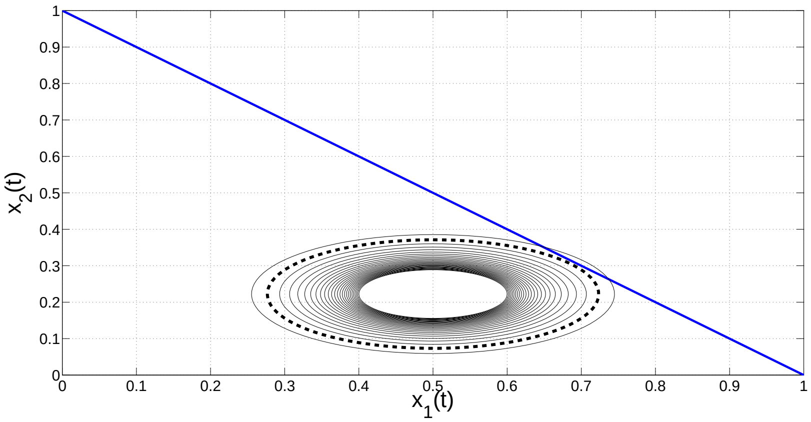

Figure 3.14: Theoretical 95% confidence ellipses of $(X_1^N(t), X_2^N(t))$, considering different N values ($N = 50, 60, 70, \ldots, 300$). Parameters: $C = 1$, $\lambda_1 = 2$, $\lambda_2 = 4$, $\mu_1 = 4$ and $\mu_2 = 5$.

3.4 Conclusions Cognitive Radio

We analyze a model for Cognitive Radio, with a dynamic spectrum sharing mechanism, where primary users have strict priority over secondary ones. We study the system without admission control policies and then propose two types of admission control policies (deterministic and probabilistic).

For the system without admission control policies we find the fluid limit using results form [EK86], and in particular its generalizations like in [Bor11a].

We find that, depending on the parameters of the system, trajectories presents sliding motion in the border of the state space. In addition we study fixed points, that can be in the border of the state space or in the interior. As we have convergence to fixed points, we consider different admission control policies, especially for the case where the fixed point is in the border of the state space. When the fixed point is in the interior we also derived the asymptotic Gaussian distribution, as a result from [EK86]. In this part we only prove convergence on compact time intervals. Despite of this, for a practical design we consider the limit when times goes to infinity for the limit process in order to analyze the behavior around the fixed point. It is object of further study the proof that in stationary regime the asymptotic distribution is the limit when time goes to infinity of the Gaussian process. This issue was remarked also in Chapter 1.

The same approach holds for the system with deterministic admission and probabilistic admission control policies. In the first case we find, that when the system is most charged the fixed point lies in the border of the admission control region and the limit presents sliding motion in the border of the admission control region. In this case there is not a Gaussian asymptotic distribution. In the unidimensional case (with the same service rates for PUs and SUs) we obtain a explicit geometric distribution under a different scaling scheme, and we conjecture through simulations that in the general case the distribution is non-Gaussian. On the other hand, for the probabilistic admission control we obtain a Gaussian distribution in compact time intervals, and the same as for the free admission control scheme, it it is needed to prove in future work that the limit in stationary regime is Gaussian, despite we also uses the Gaussian distribution for practical design.

Conclusions and future work

The objective of the present thesis is to apply the stochastic methodology to the analysis of the asymptotic behavior of some telecommunications systems.

In the review of the literature we found a large amount of research in those topics, devoted to many disciplines. Our selection focused on fluid limits, in particular for telecommunication models, leaving aside other similar or complementary approaches, as for example the mean field literature most based in particle dynamics, in the sense of [Szn91].

We studied three different models from telecommunications, where the three have some different characteristics, or different questions that we wanted to answer, that hindered to express all models in an unique framework that at the same time could be also simple to interpret. The paper from Bortolussi [Bor16] seems to address this issue in the context of stochastic Concurrent Constraint Programming. on the other hand, in this thesis we considered separated models and sometimes we also followed different approaches for similar problems appearing in those three models.

Our first model considered a BitTorrent-like P2P network. Our aim was to understand carefully the passage from a stochastic model that allow to obtain the deterministic description for leechers and seeds in a BitTorrent network presented in the classical work of Qiu and Srikant in [QS04]. In their work some issues about the relationship between a stochastic model and the associated deterministic one are sketched, but are not fully treated. In the same way we wanted to study convergence issues from stochastic to deterministic models in the case of the work presented by Rivero and Rubino [RAR10], that considers two types of leechers and implement different policies in order to improve the performance.

For both cases we defined here a stochastic system and under a proper scaling we found its fluid limit, using Kurt'z Theorem, obtaining almost sure convergence in compact time intervals. We also compared the stochastic model that converges to the deterministic model stated in [QS04] with the stochastic model previously stated by de Veciana and Yang in [YdV04].

The convergence to a fluid limit in the latter involves discontinuous transition rates, that in the first chapter were treated considering flat boundary processes presented in [SW95]. However, in the first chapter, this is mostly a remark but in the two other models in the thesis we needed to take explicitly into account discontinuous transition rates, following the approach of Bortolussi in [Bor11a] and [Bor16].

When looking at the stationary regime this model is different from the two other treated in the thesis, as for the model without scaling or for the model with a fixed parameter N we have a Markov chain with an infinite state space, so we needed to analyze previously the existence of a stationary regime. We addressed this problem using a Lyapunov function, as we could not compute an explicit stationary distribution. Once we proved the existence of a stationary distribution for each value of the scaled processes, we studied the convergence in stationary regime. For this problem there are different approaches in the literature, and for this first model we used results presented in [SW95], whereas for the same problem in the other two models studied here we considered results in [BLB08] and [LB10b].

Then we focused on the asymptotic distribution when we scale the process as in the Central Limit Theorem. From [EK86] there is convergence to a Gaussian process with an explicit covariance matrix. For our model we also obtained a Gaussian random variable as the limit in distribution when time goes to infinity.

The second model studied here is a Machine Repairman Model, where we introduced phase-type distributions both in the failure and the life-times. For this model we found a fluid limit, and in the case of exponential distributions we recovered classical results obtained in [IL73, IL74]. The main difficulties in this model were the different scales, and we treated this issue following [BKPR05], and discontinuous transition rates in the border of the state space, where we used results from [Bor11a, Bor16].

For this problem we also analyzed the deterministic limit obtained and studied its fixed point, proving that there is a global attractor. We found different regimes (sub-critical, critical and super-critical) and when scaling as in the Central Limit Theorem we obtained a Gaussian process in the limit in the case of sub-critical regime. For the other cases we only found the asymptotic distribution in the exponential case in stationary regime, using results from [Rob03], obtaining under a different scaling (N, instead on $\sqrt{N}$) a geometric asymptotic distribution.

The third model was motivated by Cognitive Radio Networks. We proposed a Markov chain model and studied its fluid limit and asymptotic distribution with the objective of defining a sort of admission control policy that improves the performance of the network. The objective was to avoid high interruption probabilities for secondary users, once they are active in the network. In this model we also have discontinuous transition rates, treated as in [Bor11a].

We considered both deterministic and probabilistic admission controls, based on Gaussian asymptotic distributions obtained for the sub-critical case as in [EK86]. The rationale behind this is that we approximated the interruption probability using the Gaussian distribution. Another possible approach that we did not consider here is to study these probabilities as large deviations. In fact we proved that in the stationary regime the scaled process converges in probability to a fixed point, the fixed point of the deterministic system obtained as limit. This means that the interruption probability is negligible, and the Gaussian approximation is a conservative approach.

For the critical case we analyzed the problem when it is in dimension one, for a deterministic admission control policy, and we obtained, based on [Rob03], a mixture of geometric distributions for the limit in the stationary regime.

In the three models there are issues that may be explored further, as the convergence in the sense of the Central Limit Theorem in stationary regime, where the difficulty comes from the fact that interchanging limits in the scaling parameter N and time requires some kind of uniformity of convergence in the whole real line or in the scaling parameter. For this purpose a possibility is to follow [EK86], that addressed for this issues results from [Nor77].

In addition, as we mentioned before, we have approximated probabilities using the Gaussian process obtained as limit, it seems that an approach based on large deviations could complementary to these estimations.

Other issue to be addressed is the case of discontinuous transition rates, when system switches and there are not Gaussian asymptotic distributions, as we found in the second chapter and in the third chapter. In both cases our approach is based in the explicit calculus of the stationary distribution. This was possible in the unidimensional case. It is necessary to explore different scaling schemes to look for other limiting distributions, when the problem is not unidimensional, both in transient and stationary regime, when there is no explicit distribution for the scaled process.

Appendix A

Density dependent population processes

In this appendix we summarize some of the principal results presented in [EK86] that are used in this work. In particular results are from Chapter 11 (Density dependent population process), devoted for families of Markov jump processes depending on a parameter, where the parameter has different interpretations, as for example: total population size, area, or volume. Theorem A.5 is often known in the literature as Kurtz's Theorem. There are many different flavours of this theorem, as for example theorems presented in [DN08] or in [Rob03]. We also summarize here some of their results.

Definition A.1 (Markov jump process in $\mathbb{Z}^d$)**.** Consider a finite set in $\mathbb{Z}^d$ $\{e_1, \ldots, e_k\}$. A Markov jump process in $\mathbb{Z}^d$ is a process U with jump directions e_i and transition rates, from state u to state $u+e_i$, $q(u, u+e_i) = \beta_i(u)$ for each $i = 1 \ldots, k$.

Definition A.2 (Generator and drift)**.** The generator of a jump Markov process is defined by

$$Q(f)(u) = \sum_{i=1}^{k} q(u, u+e_i) \left(f(u+e_i) - f(u) \right),$$

where $f : \mathbb{Z}^d \to \mathbb{R}$ is a function with compact support. The drift is defined as

$$F(u) = Q(I)(u) = \sum_{i=1}^{k} q(u, u+e_i) e_i = \sum_{i=1}^{k} e_i \beta_i(u).$$

Remark A.3. Markov jump processes can be written in terms of independent Poisson processes, using the random time change presented in [EK86]. If U is a jump Markov process,

$$U(t) = U(0) + \sum_{i=1}^{k} e_i M_i \left(\int_0^t \beta_i \left(U(s) \right) ds \right)$$

where $M_1, \ldots, M_k$ are independent standard Poisson processes.

Definition A.4 (Density dependent Markov jump process)**.** Consider a finite set of vectors $\{e_1, \ldots, e_k\} \subset \mathbb{Z}^d$, non-negative functions $\beta_1, \ldots, \beta_k$ defined in $\mathbb{R}^d$. A sequence of density dependent jump Markov processes is a sequence of processes $U^N : \mathbb{R}^d \to \mathbb{R}$ with jump directions e_i/N for each $i = 1 \ldots, k$ and transition rates $q^N(u, u + e_i/N) = N\beta_i(u)$ for each $i = 1 \ldots, k$.

Theorem A.5 (Theorem 2.1, p. 456, [EK86])**.** *Consider a sequence of process as in Definition A.4. Assume that, for each compact $K \subset \mathbb{Z}^d$,*

$$\sum_{i=1}^{k} |e_i| \beta_i(z) < +\infty$$

and there exists a constant M_k such that

$$|F(u) - F(v)| \leq M_K |u - v|.$$

Suppose also that $U^N(0) \to z_0$ and z verifies

$$z(t) = z_0 + \int_0^t F(z(s)) ds \tag{A.1}$$

for every $t \geq 0$. Then, for every $T \geq 0$,

$$\lim_{N \to +\infty} \sup_{t \in [0,T]} |Z^N(t) - z(t)| = 0 \ a.s.$$

The proof of Kurtz's Theorem relies on a characterization on the process (X^N, Y^N) as a sum of independent Poisson processes (one for each direction of possible transitions) evaluated in a random time change. Under this characterization the theorem follows from Gronwall's inequality and the Law of Large Numbers for the Poisson process.

In a addition it is possible to state a Central Limit Theorem. Let Z^N be a solution of the following equation

$$Z^N(t) = Z^N(0)+\sum_l \frac{l}{N} W_l \left(N \int_0^t \beta_l \left(Z^N(s)\right) ds \right) + \int_0^t F\left(Z^N(s)\right) ds \quad \text{(A.2)}$$

where the W_l are independent standard Brownian motions. In [EK86] it is proved that there the solution Z^N exists.

Theorem A.6 (Theorem 2.3, p. 458, [EK86]). *Suppose for each compact subset $K \subset E$,*

$$\sum_l |l|^2 \sup_{z \in K} \beta_l(z) < +\infty$$

and that the β_l and dF are continuous. Suppose Z^N satisfies (A.2)*, z satisfies A.1, $V^N = \sqrt{N}(Z^N - z)$, and $\lim_{N \to +\infty} V^N(0) = V(0)$ constant. Then $V^N \Rightarrow_N V$ where V is a Gaussian process with covariance matrix*

$$Cov\left(V(t), V(r)\right) = \int_0^{t \wedge r} \Phi(t,s) G(z(s)) \Phi(r,s)^T \, ds,$$

with $\Phi(t,s)$ the solution to the matrix equation

$$\frac{\partial}{\partial t}\Phi(t,s) = dF(z(t))\Phi(t,s), \;\; \Phi(s,s) = I,$$

$$G(z) = \sum_\ell \beta_\ell(z) \ell \ell^T.$$

As a very simplified description of the presentation in [DN08, Rob03], the proof of this approximation result is generally based on a martingale decomposition of the Markov process, which shows that the average behavior of the stochastic process is captured by the drift part while the stochastic fluctuation of second order (corresponding to the martingale) vanishes with the scaling and limit procedure. More specifically, consider a Markov process $\widetilde{Z}^N(t)$ parametric in N and its martingale decomposition:

$$\widetilde{Z}^N(t) = \widetilde{Z}^N(0) + \int_0^t \widetilde{Q}^N(\widetilde{Z}^N(s)) ds + M^N(t),$$

where $\widetilde{Q}^N(l)$ is the so-called drift of the process at state l, which is calculated as $\sum_{m \in S}(l-m)q(l,m)$, being $q(l,m)$ the transition rate from state l to m, S

the state space, and where $M^N(t)$ is a martingale. Consider now the scaled process $Z^N(t) = \widetilde{Z}^N(t)/N$, then:

$$Z^N(t) = Z^N(0) + \frac{1}{N}\int_0^t \widetilde{Q}^N(\widetilde{Z}^N(s))ds + \frac{M^N(t)}{N}.$$

If there exists a Lipschitz function Q such that:

$$\lim_{N\to\infty} \sup_{t\in[0,T]} \left| \frac{\widetilde{Q}^N(\widetilde{Z}^N(t))}{N} - Q(Z^N(t)) \right| = 0 \tag{A.3}$$

in probability, where $|\cdot|$ is the Euclidean norm, and $M^N(t)/N$ converges to zero in probability, then $Z^N(t)$ converges in probability over compact time intervals to a deterministic process $z(t)$, described by the ODE:

$$z'(t) = Q(z(t)).$$

The drift Q may be interpreted as the expected rate of change of the stochastic process.

This introduces the problem of finding the suitable scale for the approximation. A typical scaling procedure consists in dividing the process by N and considering transition rates multiples of N; jumps are of order $1/N$ and transition rates are of order N, which means that the product remains or tends to a constant as N increases. We refer to [Rob03, HW81] as references for this and other scaling regimes, under which there are different limit results.

Appendix B

Hybrid limits

In this appendix we present some results about fluid limits for CTMCs with discontinuous transition rates, where the limit becomes a PWSDS. In this work the follow the approach of [Bor11a] and [Bor11b]. More general results are presented by the author in [Bor16]. For other references for fluid limits for CTMCs with discontinuous transition rate we address for example [SW95].

The main idea is to deal with an ODE with discontinuous right hand side. That is, instead of considering an ODE $x' = F(x)$ with F a locally Lipschitz function, that guarantees the existence and uniqueness of a solution given an initial condition, we consider differential equations where F is a discontinuous function. This systems are known in the literature as PWSDS or switching systems.

Bortolussi summarizes basic methods and results of the theory of differential equations with discontinuous right-hand sides from [FA88].

We give an informal explanation of results in [Bor11a] restricting our attention to a system with two regions (R_1 and R_2). In this context we have f_1 and f_2 the velocity vectors, both continuous in R_1 and R_2 respectively and we define γ as the boundary between R_1 and R_2. In R_1 and R_2 we can apply the classical results on convergence of Markov processes. The question is what happen on γ.

B.1 Filippov solutions

In the case where the drift is discontinuous the fluid limit can be obtained in the framework of differential equations with discontinuous right-hand side [FA88]. In this context the differential equation is replaced by what is called a differential inclusion, this means that Equation (A.4) is replaced by

$$x'(t) \in \overline{Q}(x(t)) \tag{B.1}$$

where $\overline{Q}$ is a set-valued mapping known as Filippov extension of Q defined as the convex hull of the accumulation points of the drift. We define a Filippov solution as an absolutely continuous function $x(t)$ such that $x(0) = x_0$ and $x'(t) \in \overline{Q}(x(t))$ almost everywhere. We refer to [Bor11a, GG12] and references therein for a more detailed exposition.

Concretely, considering $x'(t) = Q(x)$, $Q : E \to \mathbb{R}^n$, $E \subseteq \mathbb{R}^n, \bigcup R_i \supseteq E$ (R_i $i = 1 \dots s$ is a finite set of disjoint regions), where Q is smooth on R_i and can be discontinuous only on the boundaries of R_i. We restrict our attention to a system with two regions (R_1 and R_2).

Filippov proved results about existence and uniqueness of solutions. If Q_1 and Q_2 are C^1, $Q_1 - Q_2$ is C^1 in γ (or γ' as appropriate), h is C^2 in γ (or γ') and at least one of $n^T(x)Q_1(x) > 0$ or $n^T(x)Q_2(x) < 0$ holds, with x on the border γ (or γ') and $n(x)$ the normal vector to the border, then there exists a unique Filippov solution from each initial condition. Considering x on the border γ' (or γ), there are different behaviors of a solution starting in x depending on the value of $n^T(x)Q_1(x)$ and $n^T(x)Q_2(x)$:

Transversal crossing: if $n^T(x)Q_1(x)$ and $n^T(x)Q_2(x)$ have the same sign, e.g. if $n^T(x)Q_1(x) > 0$ and $n^T(x)Q_2(x) > 0$, a solution starting in R_1 crosses the border and stays in R_2;

Sliding motion: if $(n^T(x)Q_1(x))(n^T(x)Q_2(x)) < 0$ there is sliding motion, there are two cases: unstable sliding motion when $n^T(x)Q_1(x) < 0$ and $n^T(x)Q_2(x) > 0$ (in this case there is no uniqueness for solutions) and stable sliding motion when $n^T(x)Q_1(x) > 0$ and $n^T(x)Q_2(x) < 0$. In this last case, that is ours, the system cannot escape from the border, then the solution follows a vector field obtained as convex combination of Q_1 and Q_2, obtaining a new vector field

$$g(x) = (1 - \alpha(x))Q_1 + \alpha(x)Q_2 \tag{B.2}$$

with $\alpha(x) \in [0, 1]$ that verifies $n^T(x)g(x) = 0$;

Tangential crossing: if $n^T(x)Q_1(x) = 0$ (or $n^T(x)Q_2(x) = 0$), then the trajectory continues in the region pointed by the non-zero vector field.

B.2 Piecewise smooth fluid limit

Now we present the model considered in [Bor11a]

Consider a model (X, D, τ, X_0) where:

1. $X = (X_1 \ldots, X_n)$ is a set of variables.
2. Each X_i takes values in a finite or countable domain $Di \subset \mathbb{R}$. Usually, but not necessarily, D_i is a subset of the integers. Hence, $D = \prod_{i=1}^{n} D_i$ is the state space of the model.
3. $X_0 \in D$ is the initial state of the model.
4. $\tau = \{\tau_1, \ldots, \tau_m\}$ is the set of transitions, of the form $\tau_i = (\varphi(X), v, R(X))$ where:
   (a) $\varphi(X)$ is a conjunction of inequalities of the form $h(X) \geq 0$ or $h(X) > 0$, where h's are suitably smooth functions on D, usually linear.
   (b) $v \in \mathbb{R}^n$, is the update vector, i.e. a vector giving the net change on each variable caused by the transition. We require that $X + v \in D$ whenever $\varphi(X)$ is true.
   (c) $r : D \to \mathbb{R}^+ \cup \{0\}$ is the rate function, which specifies the rate of the transition as a function of the current state of the system. We require each r to be Lipschitz continuous and bounded on D.

Consider a sequence of models $(X^N, D^N, \tau^N, X_0^N)$ and the normalized variable X^N/N with the corresponding scaling assumptions:

initial conditions scale properly: $x_0^N = X_0^N/N$;

domains scale properly: $\hat{D}^N = \{X/N : x \in D^N\}$;

for each transition $(\varphi_i^N(X), v_i^N, r_i^N(X))$ of the nonnormalized model, there exists a predicate $\varphi_i(\hat{X})$ on normalized variables, a vector v_i, and a bounded and Lipschitz function $f_i(\hat{X}) : E \to \mathbb{R}^n$ on normalized variables, all independent of N, such that $\varphi^N(X) = \varphi_i(X/N)$, $v_i^N = v_i$ and $r_i^N(X) = N f_i(X/N)$. The corresponding transition in the normalized model is $(\varphi_i(\hat{X}), v_i/N, N f_i(\hat{X}))$.

The drift is:

$F^N(\hat{X}) = F(\hat{X}) = \sum_{i=1}^{m} v_i f_i(\hat{X}) I_{\varphi_i}(\hat{X})$ where $I_{\varphi_i}(\hat{X})$ is the indicator function of the predicate φ_i

Theorem B.1. *Let the sequence X^N of CTMC models satisfy the scaling assumptions and consider the PWS system $x' = F(x)$, assumed to have only two continuity regions separated by a smooth manifold. Let $\mathcal{H}$ be the discontinuity surface of the PWSDS, an assume $x_0 \in \mathcal{H}$ is such that $n^T(x_0)f_1(x_0) < 0$ and $n^T(x_0)f_2(x_0) > 0$. Let $x(t)$, $t \leq T_S$ be the unique solution of the PWSDS starting from x_0, defined as the solution of the ODE $x' = G(x)$, where G is the sliding vector field defined as in* (B.2) *and T_S is the time at which sliding motion terminates with first order exit conditions.Fix $T \leq T_S$, $T < \infty$. If $X^N(0) \to x_0$ in probability, then* $\lim_{N\to+\infty} \sup_{t\leq T} |\hat{X}^N(t) - x(t)| = 0$ *in probability.*

List of Figures

List of Tables

Bibliography

[AA10] S. Asmussen and H. Albrecher. *Ruin probabilities*. Advanced Series on Statistical Science & Applied Probability, 14. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, second edition, 2010.

[AA14] S. A. AlQahtani and H. Ahmed. An admission control scheme for secondary users in cognitive radio networks. In *Proceedings of the 11th IEEE International Conference on Networking, Sensing and Control*, pages 245–250, April 2014.

[AEJV13] U. Ayesta, M. Erausquin, M. Jonckheere, and I. M. Verloop. Scheduling in a Random Environment: Stability and Asymptotic Optimality. *IEEE/ACM Trans. Netw.*, 21(1):258–271, 2013.

[AMR11] L. Aspirot, E. Mordecki, and G. Rubino. Fluid Limits Applied to Peer to Peer Network Analysis. In *Quantitative Evaluation of Systems (QEST), 2011 Eighth International Conference on*, pages 13 –20, sept. 2011.

[AMR13] L. Aspirot, E. Mordecki, and G. Rubino. Fluid Limit for the Machine Repairman Model with Phase-Type Distributions. In *Quantitative Evaluation of Systems - 10th International Conference, QEST 2013, Buenos Aires, Argentina, August 27-30, 2013. Proceedings*, pages 139–154, 2013.

[ARR16] F. Akhtar, M. H. Rehmani, and M. Reisslein. White space: Definitional perspectives and their role in exploiting spectrum opportunities. *Telecommunications Policy*, 40(4):319 – 331, 2016.

[BEGFB94] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory.* Society for Industrial and Applied Mathematics, 1994.

[Ben99] M. Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 1–68. Springer, Berlin, 1999.

[BHLM13] L. Bortolussi, J. Hillston, D. Latella, and M. Massink. Continuous approximation of collective system behaviour: A tutorial. *Performance Evaluation*, 70:317–349, 2013.

[BJLM16] P. Bermolen, M. Jonckheere, F. Larroca, and P. Moyal. Estimating the transmission probability in wireless networks with configuration models. *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, 1(2):1–23, apr 2016.

[BJR+07] V. Blaschke, H. Jaekel, T. Renk, C. Kloeck, and F. K. Jondral. Occupation Measurements Supporting Dynamic Spectrum Allocation for Cognitive Radio Design. In *Cognitive Radio Oriented Wireless Networks and Communications, 2007. CrownCom 2007. 2nd International Conference on*, pages 50–57, Aug 2007.

[BKPR05] K. Ball, T. G. Kurtz, L. Popovic, and G. Rempala. Asymptotic analysis of multiscale approximations to reaction networks. *The Annals of Applied Probability*, 16(4):1925–1961, 2005.

[BLB08] M. Benaïm and J-Y. Le Boudec. A class of mean field interaction models for computer and communication systems. *Perform. Eval.*, 65(11-12):823–838, 2008.

[BMNV13] F. Baccelli, F. Mathieu, I. Norros, and R. Varloot. Can P2P networks be super-scalable? In *2013 Proceedings IEEE INFOCOM*, pages 1753–1761, April 2013.

[BMP08] C. Bordenave, D. McDonald, and A. Proutiere. Performance of random medium access control, an asymptotic approach. In *SIGMETRICS '08: Proceedings of the 2008 ACM SIGMETRICS international conference on Measurement and modeling* *of computer systems*, pages 1–12, New York, NY, USA, 2008. ACM.

[BMP10] C. Bordenave, D. McDonald, and A. Proutiere. A particle system in interaction with a rapidly varying environment: Mean field limits and applications. *Networks and Heterogeneous Media*, 5(1):31–62, 2010.

[Bor11a] L. Bortolussi. Hybrid Limits of Continuous Time Markov Chains. In *Proceedings of Eighth International Conference on Quantitative Evaluation of Systems (QEST), 2011*, pages 3 –12, sept. 2011.

[Bor11b] L. Bortolussi. Supplementary material of Hybrid Limits of Continuous Time Markov Chains. `www.dmi.units.it/~bortolu/files/qest2011supp.pdf`, 2011. [Online; accesed 15-January-2018].

[Bor16] L. Bortolussi. Hybrid behaviour of Markov population models. *Information and Computation*, 247(Supplement C):37 – 86, 2016.

[BT12] L. Bortolussi and M. Tribastone. Fluid limits of queueing networks with batches. In *Proceedings of the third joint WOSP/SIPEW international conference on Performance Engineering*, ICPE '12, pages 45–56, New York, NY, USA, 2012. ACM.

[CLBR09] A. Chaintreau, J-Y. Le Boudec, and N. Ristanovic. The age of gossip: Spatial mean-field regime. In *ACM Sigmetrics*, ACM Sigmetrics, 2009.

[Coh03] B. Cohen. Incentives build robustness in BitTorrent. In *Workshop on Economics of Peer-to-Peer systems*, volume 6, pages 68–72, 2003.

[CRRB+08] H. Cancela, F. Robledo, P. Rodríguez-Bocca, G. Rubino, and A. Sabiguero. A robust P2P streaming architecture and its application to a high quality live-video service. *Electronic Notes in Discrete Mathematics*, 30:219–224, 2008.

[DN08] R. W. R. Darling and J. R. Norris. Differential equation approximations for Markov chains. *Probab. Surv.*, 5:37–79, 2008.

[EK86] S. N. Ethier and T. G. Kurtz. *Markov processes: Characterization and convergence*. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., New York, 1986.

[ERAR13] E. Báez Esquivel, M. E. Rivero-Angeles, and G. Rubino. Priority scheme for window-based video-on-demand transmission on BitTorrent-like Peer-to-Peer networks. In *IEEE International Conference on Communications (ICC)*, pages 3000–3005, 2013.

[FA88] A. F. Filippov and F. M. Arscott. *Differential equations with discontinuous righthand sides*. Mathematics and its applications. Soviet series. Kluwer Academic Publ, Dordrecht, 1988.

[FKP11] A. Ferragut, F. Kozynski, and F. Paganini. Dynamics of content propagation in BitTorrent-like P2P file exchange systems. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 3116–3121, 2011.

[FP12] A. Ferragut and F. Paganini. Content Dynamics in P2P Networks from Queueing and Fluid Perspectives. In *Proceedings of the 24th International Teletraffic Congress*, ITC '12, pages 11:1–11:8. International Teletraffic Congress, 2012.

[GCX+05] L. Guo, S. Chen, Z. Xiao, E. Tan, X. Ding, and X. Zhang. Measurements, analysis, and modeling of BitTorrent-like systems. In *IMC '05: Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement*, pages 4–4, Berkeley, CA, USA, 2005. USENIX Association.

[GCX+07] L. Guo, S. Chen, Z. Xiao, E. Tan, X. Ding, and X. Zhang. A performance study of BitTorrent-like peer-to-peer systems. *Selected Areas in Communications, IEEE Journal on*, 25(1):155–169, 2007.

[GG12] N. Gast and B. Gaujal. Markov chains with discontinuous drifts have differential inclusion limits. *Perform. Eval.*, 69:623–642, 2012.

[GGLB10] N. Gast, B. Gaujal, and J-Y. Le Boudec. Mean field for Markov Decision Processes: from Discrete to Continuous Optimization. Technical report, EPFL, 2010.

[Gra00] C. Graham. Chaoticity on path space for a queueing network with selection of the shortest queue among several. *J. Appl. Probab.*, 37(1):198–211, 2000.

[HA07] L. Haque and M. J. Armstrong. A survey of the machine interference problem. *European Journal of Operational Research*, 179(2):469–482, 2007.

[HB12] B. Van Houdt and L. Bortolussi. Fluid limit of an asynchronous optical packet switch with shared per link full range wavelength conversion. In *SIGMETRICS'12*, pages 113–124, 2012.

[HHK+17] Z. Haider, R. Hussain, I. L. Khan, A. Shakeel, B. Ijaz, and S. A. Malik. Evaluation of capabilities of open source Cognitive radio network simulators. In *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1814–1817, June 2017.

[HKKS17] M. R. Hassan, G. Karmakar, J. Kamruzzaman, and B. Srinivasan. Exclusive Use Spectrum Access Trading Models in Cognitive Radio Networks: A Survey. *IEEE Communications Surveys Tutorials*, PP(99):1–1, 2017.

[HW81] S. Halfin and W. Whitt. Heavy-traffic limits for queues with many exponential servers. *Operations Research*, 29:567–588, 06 1981.

[IL73] D. L. Iglehart and A. J. Lemoine. Approximations for the Repairman Problem with Two Repair Facilities, I: No Spares. *Advances in Applied Probability*, 5(3):pp. 595–613, 1973.

[IL74] D. L. Iglehart and A. J. Lemoine. Approximations for the repairman problem with two repair facilities, II: Spares. *Advances in Appl. Probability*, 6:147–158, 1974.

[IUKB+04] M. Izal, G. Urvoy-Keller, E. W. Biersack, P. Felber, A. Al Hamra, and L. Garcés-Erice. Dissecting BitTorrent: Five Months in a Torrent's Lifetime. In *PAM*, pages 1–11, 2004.

[JJ16] J. Jacob and B. R. Jose. Spectrum occupancy measurement and analysis in Kochi-India from a cognitive radio perspective. In *2016 Sixth International Symposium on Embedded*

Computing and System Design (ISED), pages 328–333, Dec 2016.

[JS14] M. Jonckheere and S. Shneer. Stability of Multi-Dimensional Birth-and-Death Processes with State-Dependent 0-Homogeneous Jumps. *Advances in Applied Probability*, 46(1):59 – 75, 2014.

[KB15] N. Khedun and V. Bassoo. Analysis of priority queueing with multichannel in cognitive radio network. In *IEEE EUROCON 2015 - International Conference on Computer as a Tool (EUROCON)*, pages 1–6, Sept 2015.

[KKS07] G. Kesidis, T. Konstantopoulos, and P. Sousi. Modeling file-sharing with Bittorrent-like incentives. In *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2007*, volume 4, pages IV–1333 –IV–1336, apr. 2007.

[KKS09] G. Kesidis, T. Konstantopoulos, and P. Sousi. A Stochastic Epidemiological Model and a Deterministic Limit for BitTorrent-Like Peer-to-Peer File-Sharing Networks. In E. Altman and A. Chaintreau, editors, *Network Control and Optimization*, volume 5425 of *Lecture Notes in Computer Science*, pages 26–36. Springer Berlin, Heidelberg, 2009.

[Kur82] T. G. Kurtz. Representation and approximation of counting processes. In *Advances in filtering and optimal stochastic control (Cocoyoc, 1982)*, volume 42 of *Lecture Notes in Control and Inform. Sci.*, pages 177–191. Springer, Berlin, 1982.

[Lav83] S.S. Lavenberg. *Computer performance modeling handbook*. Notes and reports in computer science and applied mathematics. Academic Press, 1983.

[LB10a] J-Y. Le Boudec. Mean Field Methods for Computer and Communication Systems (A Tutorial). Performance 2010 Tutorials, November 2010.

[LB10b] J-Y. Le Boudec. The Stationary Behaviour of Fluid Limits of Reversible Processes is Concentrated on Stationary Points. Technical report, EPFL, 2010.

[LBMM07] J-Y. Le Boudec, D. McDonald, and J. Mundinger. A Generic Mean Field Convergence Result for Systems of Interacting Objects. In *QEST '07, Proceedings of the Fourth International Conference on Quantitative Evaluation of Systems*, pages 3–18, Washington, DC, USA, 2007. IEEE Computer Society.

[MH17] H. Mu and T. Hu. *Cognitive Radio and the New Spectrum Paradigm for 5G*, pages 265–286. Springer Singapore, Singapore, 2017.

[Mit00] J. Mitola. *Cognitive radio*. PhD thesis, KTH, Teleinformatics, 2000. NR 20140805.

[MM99] J. Mitola and G. Q. Maguire. Cognitive radio: making software radios more personal. *IEEE Personal Communications*, 6(4):13–18, Aug 1999.

[MMR98] A. Mandelbaum, W. A. Massey, and M. I. Reiman. Strong Approximations for Markovian Service Networks. *Queueing Syst. Theory Appl.*, 30(1/2):149–201, June 1998.

[MV05] L. Massoulié and M. Vojnović. Coupon replication systems. *SIGMETRICS Perform. Eval. Rev.*, 33(1):2–13, 2005.

[Nor77] M. F. Norman. Ergodicity of Diffusion and Temporal Uniformity of Diffusion Approximation. *Journal of Applied Probability*, 14(2):399–404, 1977.

[PF12] F. Paganini and A. Ferragut. PDE models for population and residual work applied to peer-to-peer networks. In *Information Sciences and Systems (CISS), 2012 46th Annual Conference on*, pages 1–6, 2012.

[PGES05] J. A. Pouwelse, P. Garbacki, D. H. J. Epema, and H. J. Sips. The Bittorrent P2P File-Sharing System: Measurements and Analysis. In *IPTPS*, pages 205–216, 2005.

[QS04] D. Qiu and R. Srikant. Modeling and performance analysis of BitTorrent-like peer-to-peer networks. In *SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 367–378, New York, NY, USA, 2004. ACM.

[QS08] D. Qiu and W. Sang. Global stability of peer-to-peer file sharing systems. *Comput. Commun.*, 31(2):212–219, 2008.

[RAB15] C. Rattaro, L. Aspirot, and P. Belzarena. Analysis and characterization of dynamic spectrum sharing in Cognitive Radio Networks. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2015 International*, pages 166–171, Aug 2015.

[RABM19] C. Rattaro, L. Aspirot, P. Belzarena, and E. Mordecki. QoS provision in a dynamic channel allocation based on admission control decisions. *Accepted in ACM Transactions on Modeling and Performance Evaluation of Computing Systems (TOMPECS)*, 2019.

[RAR10] M. E. Rivero-Angeles and G. Rubino. Priority-Based Scheme for File Distribution in Peer-to-Peer Networks. In *IEEE International Conference on Communications (ICC), 2010*, pages 1–6, 2010.

[RARTM13] M. E. Rivero-Angeles, G. Rubino, I. O. Olguín Torres, and L. A. Martinez. Window-based streaming Video-on-Demand transmission on BitTorrent-like Peer-to-Peer networks. In *IIIE 10th Consumer Communications and Networking Conference (CCNC)*, pages 1–6, 2013.

[Rat18] C. Rattaro. *Stochastic models for cognitive radio networks.* Phd thesis, Universidad de la República (Uruguay), Facultad de Ingeniería, IIE, 2018.

[RBB18] C. Rattaro, P. Bermolen, and P. Belzarena. Multi-resource allocation : Analysis of a paid spectrum sharing approach based on fluid models. *IEEE Transactions on Cognitive Communications and Networking*, pages 1–14, 2018.

[RLBB17] C. Rattaro, F. Larroca, P. Bermolen, and P. Belzarena. Estimating the medium access probability in large cognitive radio networks. *Ad Hoc Networks*, 63:1–13, aug 2017.

[Rob03] P. Robert. *Stochastic Networks and Queues.* Stochastic Modelling and Applied Probability Series. Springer-Verlag, New York, 2003.

[SFDS12] P. J. Smith, A. Firag, P. A. Dmochowski, and M. Shafi. Analysis of the M/M/N/N Queue with Two Types of Arrival Process: Applications to Future Mobile Radio Systems. *J. Applied Mathematics*, 2012:123808:1–123808:14, 2012.

[SSY11] S. Shakkottai, R. Srikant, and L. Ying. The Asymptotic Behavior of Minimum Buffer Size Requirements in Large P2P Streaming Networks. *IEEE Journal on Selected Areas in Communications*, 29(5):928–937, 2011.

[SV16] S. Sasipriya and R. Vigneshram. An overview of cognitive radio in 5G wireless communications. In *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pages 1–5, Dec 2016.

[SW95] A. Shwartz and A. Weiss. *Large deviations for performance analysis*. Stochastic Modeling Series. Chapman & Hall, London, 1995.

[Szn91] A-S. Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[TLBEAA09] H. Tembine, J-Y. Le Boudec, R. El-Azouzi, and E. Altman. Mean Field Asymptotic of Markov Decision Evolutionary Games and Teams. In *Gamenets 2009*, 2009. Invited Paper.

[TM06] P. Tinnakornsrisuphap and A. M. Makowski. On the behavior of ECN/RED gateways under a large number of TCP flows: Limit theorems. *Queueing Syst. Theory Appl.*, 52(4):287–304, 2006.

[TT17] M. Tschaikowski and M. Tribastone. Spatial fluid limits for stochastic mobile networks. *Performance Evaluation*, 109:52 – 76, 2017.

[XFC13] J. Xue, Z. Feng, and K. Chen. Beijing Spectrum Survey for Cognitive Radio Applications. In *2013 IEEE 78th Vehicular Technology Conference (VTC Fall)*, pages 1–5, Sept 2013.

[YdV04] X. Yang and G. de Veciana. Service Capacity of Peer to Peer Networks. In *IEEE INFOCOM*, pages 1–11, 2004.

[Zha12] Z. Zhao. A preemptive connection pool manager for web-based application collaboration. In *8th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, pages 326–334, Oct 2012.