



Speed of convergence of diffusion approximations

Eustache Besançon

► To cite this version:

Eustache Besançon. Speed of convergence of diffusion approximations. Probability [math.PR]. Institut Polytechnique de Paris, 2020. English. NNT : 2020IPPAT038 . tel-03128781

HAL Id: tel-03128781

<https://theses.hal.science/tel-03128781>

Submitted on 2 Feb 2021

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

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

Speed of Convergence of Diffusion Approximations

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à Télécom Paris

École doctorale n°574 École doctorale de mathématiques Hadamard
(EDMH)

Spécialité de doctorat: Mathématiques Appliquées

Thèse présentée et soutenue à Paris, le 8 décembre 2020, par

Eustache Besançon

Composition du Jury:

Laure Coutin	Président
Professeur, Université Paul Sabatier (– IM Toulouse)	
Kavita Ramanan	Rapporteur
Professor, Brown University (– Applied Mathematics)	
Ivan Nourdin	Rapporteur
Professeur, Université du Luxembourg (– Dpt Mathématiques)	
Philippe Robert	Examineur
Directeur de Recherche, INRIA	
François Roueff	Examineur
Professeur, Télécom Paris (– LTCI)	
Laurent Decreusefond	Directeur de thèse
Professeur, Télécom Paris (– LTCI)	
Pascal Moyal	Directeur de thèse
Professeur, Université de Lorraine (–Institut Elie Cartan)	

Speed of convergence of diffusion approximations

Eustache Besançon

December 8th, 2020

Contents

1	Introduction	7
1	Diffusion Approximations	7
2	Fields of Application	8
2.1	Queueing theory	8
2.2	Epidemiology	11
2.3	Genetics	14
3	Our contribution: a general approach to the rate of convergence in law	16
2	Roadmap	19
1	The goal	19
2	Overview of the Stein Method	20
3	Roadmap	23
3	Choice of a functional space	25
1	Fractional Sobolev spaces	25
2	Wiener structure on a fractional Sobolev space	27
4	Stein Method in Wiener fractional Sobolev Spaces	37
1	Malliavin calculus for Poisson processes	37
2	Stein Operator and linear interpolation	41
5	Application to a class of Markov Processes	49
1	A class of Markov processes	49
2	The scaling of the processes	50
3	Interpolation of these processes	53
3.1	Distance between the process and its interpolation	53
3.2	Distance between two affine interpolations	54
4	A central limit theorem	56
6	Application to Queueing Models	61
1	The $M/M/\infty$ queue	61
1.1	Short introduction to the $M/M/\infty$ model	61
1.2	Straight application of our model	63
2	The $M/M/1$ model	64
2.1	Short introduction to the $M/M/1$ model	64
2.2	Higher rate of convergence for a smaller class of functions	66

7	Other Applications	69
1	The ON/OFF model	69
2	The SIR model	71
3	The Moran model with selection	73
8	Appendix	75
1	Besov-Liouville Spaces	75
2	Moment bound for Poisson variables	79

Aknowledgement

All my thanks go to professors Laurent Decreusefond and Pascal Moyal, who were kind enough to accept to de supervise my doctoral thesis, and had the intelligence to invent a fascinating subjectas well as the patience to guide me in my research work. Without them this work would not have been possible. I thank Kavita Ramanan and Ivan Nourdin who have accepted to review my work and act as referees. I thank also professors Laure Coutin and Philippe Robert, who have accepted to be part of the jury and last but not least, professor François Roueff who greeted me in the laboratory, and guided me in the administrative arcanes of the school. I finally thank for their patience my wife Olga and my daughters Juliette and Enora whom I did not attend as much as expected while I was writing this thesis.

Résumé

Dans de nombreux champs d'applications (files d'attentes, épidémiologie, génétique, finance), les processus de Markov sont un outil privilégié de modélisation de processus aléatoires. Malheureusement, il est souvent nécessaire d'avoir recours à des espaces d'états très grands voire infinis, rendant l'analyse exacte des différentes caractéristiques (stabilité, loi stationnaire, temps d'atteinte de certains domaines, etc.) du processus délicate ou impossible. Depuis longtemps, grâce notamment à la théorie des martingales, on procède à des approximations par des diffusions browniennes. Celles-ci permettent souvent une analyse approchée du modèle d'origine.

Le principal défaut de cette approche est que l'on ne connaît pas l'erreur commise dans cette approximation. Il s'agit donc ici de développer une théorie du calcul d'erreur dans les approximations diffusion. Depuis quelques temps, le développement de la méthode de Stein-Malliavin a permis de préciser les vitesses de convergence dans les théorèmes classiques comme le théorème de Donsker (convergence fonctionnelle d'une marche aléatoire vers un mouvement brownien) ou la généralisation trajectorielle de l'approximation binomiale-Poisson.

Il s'agit dans ce travail de poursuivre le développement de cette théorie pour des processus de Markov comme ceux que l'on rencontre en théorie des files d'attente ou en épidémiologie et dans bien d'autres domaines appliqués. Or tous les travaux précédents utilisant la méthode de Stein en dimension infinie ont étudié la convergence de processus vers des mouvements Browniens standards ou vers des processus de Poisson alors que les limites fluides des approximations diffusions peuvent prendre des formes variées. Il s'agit donc d'inventer

une manière d'appliquer la méthode de Stein à un ensemble plus large de situations.

On se situe d'emblée dans un espace de dimension infinie : les espaces de Sobolev fractionnaires, qui sont suffisamment larges pour comprendre les fonctions continues Hölderiennes que sont les mouvements Browniens et les fonctions constantes par morceaux que sont les processus de Poisson, et qui présentent des caractéristiques fort commodes facilitant les calculs (espaces de Banach séparables, existence d'une injection canonique dans un espace de Hilbert permettant d'utiliser un produit scalaire notamment). Partant de la représentation des processus de Markov comme mesures de Poisson, on étend la méthode développée par Laurent Decreusefond et Laure Coutin pour estimer la vitesse de convergence dans les approximations diffusion. On munit ces espaces de Sobolev d'une mesure de Wiener abstraite et l'on construit un processus d'Ornstein Uhlenbeck en dimension infinie, ayant pour mesure invariante la mesure de Wiener. On est alors en mesure d'exprimer l'équation de Stein caractérisant la convergence vers le mouvement Brownien en utilisant les travaux de Shih. Cette équation reste valide pour toute fonctionnelle Lipschitzienne. Afin d'exploiter cette équation, il faut utiliser le calcul de Malliavin sur les processus de Poisson, et notamment la formule d'intégration par parties.

On étend la méthode de Stein-Malliavin à des vecteurs de processus dépendants plutôt qu'à un seul processus. Si la suite de processus converge vers une limite pouvant être exprimée comme limite d'une équation différentielle à paramètres déterministes, la limite est un processus Gaussien changé de temps. La méthode de Stein Malliavin étant développée surtout pour montrer la convergence vers le mouvement Brownien standard, on l'adapte à la convergence vers un processus changé de temps à travers des méthodes d'approximations linéaires, le processus changé de temps étant approximativement un mouvement Brownien standard sur un petit espace de temps. On utilise deux échelles de temps afin d'approcher au mieux le processus discret de départ et le mouvement continu à la limite. Un calcul d'optimisation permet enfin d'obtenir l'échelle de temps la plus appropriée pour évaluer la vitesse de convergence au plus juste. Celle s'avère être en définitive de $n^{-1/6} \ln n$.

Cette méthodologie est alors appliquée à quatre exemples d'approximations diffusion : en théorie des file d'attentes, les files $M/M/\infty$ et $M/M/1$. En épidémiologie le modèle SIR de catégorisation des populations entre susceptibles, infectés et remis. En génétique le modèle On/Off. Dans tous ces cas, la vitesse de convergence peut être bornée par la quantité $n^{-1/6} \ln n$. à un facteur multiplicatif constant près. Cette méthode n'est toutefois pas universelle, et le modèle de Moran, utilisé en génétique des populations, ne se prête pas à son utilisation, car il ne converge pas vers une équation différentielle à paramètres déterministes. D'autres méthodes restent à découvrir pour déterminer sa vitesse de convergence.

Chapter 1

Introduction

1 Diffusion Approximations

In the present manuscript we shall be interested in particular classes of stochastic models and their asymptotic behaviour. A stochastic process is just a model for a random quantity that evolves through time. Such a mathematical object is a collection of random variables $\{X_t : t \in \mathbb{T}\}$ indexed by a set $\mathbb{T}$ which we can interpret as a set of times. It will often be denoted simply by X in the sequel. In order to denote the value of the process X at time t , we shall use either $X(t)$ or X_t , depending on whether adding an index or parenthesis to the notation makes it easier to read, and these two notations should be considered as equivalent. For us $\mathbb{T}$ will always be a compact interval $[0, T]$ in $\mathbb{R}^+$ (continuous time setting). For all $t \in [0, T]$, we shall use the notation $\mathcal{F}_t$ to mean the history of the process up until time t , that is to say the information available to us if we watched the process up to time t . A stochastic process will be said to have the Markov property if its future evolution conditional on knowing all of $\mathcal{F}_t$ is the same as its evolution conditional on knowing just X_t . In other words, where it goes next may depend on its current value, but not on how it evolved before getting to that value.

A stochastic process X is said to be discrete if the random variables X_t are themselves discrete random variables, that is take their values in a finite or countable state space. There are said to be continuous if they can take any value in $\mathbb{R}$. In many fields of interest, for instance in queueing theory, in biology, information theory, finance or graph theory, one can find quite involved discrete stochastic processes which can be approximated by diffusions, i.e. continuous markovian processes obtained as solutions of a stochastic differential equation (i.e a differential equation whose unknown or whose variables are themselves random variables). Such processes are called "diffusion approximations". Most often, starting from a discrete process, a discrete parameter is introduced in order to change the scale of the sample size, of the population of interest or of time. When this parameter increases to infinity one obtains a limit which follows the same law as a continuous process.

In practical applications, this enables to replace the discrete process when observed over a sufficiently large population or when repeating observations at a high enough frequency, by a continuous approximation, supposed to be

easier to manipulate. The notion of approximation diffusion is built on that of weak convergence as the limit diffusion process is deemed to be a suitable approximation if the parameterized discrete process sequence converges in law towards the diffusion. Weak convergence towards stochastic processes was first correctly described by Prokhorov who uncovered the concept of tightness of measures ([98]) and devised the Prokhorov distance, and by Skorokhod, who proposed the Skorokhod topology ([110]) where tightness could be reduced to a form of the Arzela Ascoli theorem.

The most basic and easy to understand result in the diffusion approximation literature is Donsker's theorem which states that a sum of random variables, when viewed as a stochastic process, can be approximated by a Brownian motion. Donsker proved in [35] that if the $Z_n, n \in \mathbb{N}$ variables were independent and identically distributed, denoting $S_n = Z_1 + \dots + Z_n$ and for all $t \geq 0$,

$$X_n(t) = n^{1/2}(n^{-1}S_{\lfloor nt \rfloor} - \mathbf{E}[Z_n]t),$$

the sequence (X_n) of processes defined by $X_n(t)$ converges weakly towards a standard Brownian motion when n increases to infinity. A number of important extensions to this Donsker theorem have been made over the years in order to lift the i.i.d condition and to allow some kind of dependence or some form of non stationarity. For instance if Z_n is a stationary mixing process ([43]), or if Z_n is a real valued functional of a time homogeneous positive recurrent markov chain ([92]).

Most of the mathematical work made on approximation diffusions has been motivated by practical applications.

2 Fields of Application

2.1 Queueing theory

In 1909, Agner Krarup Erlang, a Danish engineer employed by the Copenhagen Telephone Exchange, published [42] where he modeled the number of telephone calls arriving at an exchange by a Poisson process. If there were more jobs at a node than there were servers, then jobs would queue and wait for service. In the 1950s queueing theory became an area of research interest to mathematicians. Queueing theory was systematized in 1953 by Kendall, who proposed describing queueing models using three factors written A/S/c ([67]) where A denotes the arrival rate of clients in the queue, S the service time distribution and c the number of service channels open at the node. The arrival process can be denoted (among others) by the symbol M for a Markovian (Poisson) process or G for a General, independent, distribution; the service time distribution can be denoted by the symbol M for exponential distribution, D for a deterministic service time, or G for a General, independent distribution. Erlang had solved the M/D/1 queue in 1917 and M/D/k queueing model in 1920 (see [72]). The M/G/1 queue was solved by Felix Pollaczek in 1930, ([95]) a solution later recast in probabilistic terms by Khinchin and now known as the Pollaczek–Khinchin formula ([69]). In 1953 David George Kendall solved the G/M/k queue ([67]). In 1957 Pollaczek studied the G/G/1 using an integral equation([96]). John Kingman gave a formula for the mean waiting time in a G/G/1 queue: Kingman's formula ([70]).

Leonard Kleinrock worked on the application of queueing theory to message switching in the early 1960s and packet switching in the early 1970s. His initial contribution to this field was his doctoral thesis at the Massachusetts Institute of Technology in 1961([73]) . His theoretical work published in the early 1970s underpinned the use of packet switching in the ARPANET, a forerunner to the Internet. The matrix geometric method, developed by Neuts and his students starting around 1975, and the more complicated matrix analytic method have allowed queues with phase-type distributed inter-arrival and service time distributions to be considered ([100]).

With time, more complex models have been studied. Networks of queues are systems in which a number of queues are connected, and a protocol called customer routing decides how customers go from one node to another once they are serviced. When a customer is serviced at one node it can join another node and queue for service, or leave the network. For networks of n nodes, the state of the system can be described by an n -dimensional vector $(x_1, x_2, \dots, x_n)$ where x_i denotes the number of customers queueing at each node.

The simplest non-trivial network of queues is called tandem queues: two queues with one server each, having independent service times where customers joining the first queue immediately enter the second queue on completing service. The theory of weak convergence for queueing networks has been largely based on the development of convenient representations for the corresponding queue length processes. By applying the continuous mapping principle, random time change theory and pairing techniques to the representation, diffusion limits for the network were successfully obtained. The first significant results in this area were obtained by Jackson ([61], [62]). He proved that an efficient product-form stationary distribution existed. A probability distribution π on $\mathbb{R}^d$ is said to be of product form if

$$\pi(\mathrm{d}x_1 \times \dots \times \mathrm{d}x_d) = \prod_{i=1}^d p_i(x_i) \mathrm{d}x_i$$

In the case where the total number of customers in the network remains constant the network is called a closed network and it has also been shown to have a product-form stationary distribution by Gordon and Newell ([49]). This result was extended to networks with very general service time, regimes and customer routing schemes ([6]). So called Kelly networks, where customers of different classes experience different priority levels at different service nodes, have also been investigated ([66]). Another type of network are G-networks first proposed by Gelenbe in 1993 ([48]). These networks do not assume exponential time distributions like the classic Jackson Network.

In those networks of queues, the complex behaviour of the queueing model is replaced by a more tractable Brownian network which is a diffusion process obtained by subjecting a Brownian motion to reflection at the boundaries of an appropriately defined region which spans at the limiting state space of the queueing network. In the case of a single server queue, this Brownian motion is unidimensional. If there are d servers, the reflected Brownian motion is a multidimensional object whose domain is an orthant. Skorokhod had ([111]) derived an equation for diffusion processes with a reflecting boundary and proved existence and uniqueness of the solution in the unidimensional case: specifically, if Y is a càdlàg function on $\mathbb{R}^+$ taking values in a domain $\mathbb{R}$ and such that

$Y(0) \geq 0$, Skorokhod built a function X such that X was constrained to remain within the closure of domain D , equal to Y inside this domain and being reflected at the border of that domain: He proved that in the case $D = \mathbb{R}^+$ there was a unique couple (X, R) in $\mathbb{D}$ which was a solution to what became known as the Sokhorod problem, i.e. such that for all $t \geq 0$,

1. $X(t) = Y(t) + R(t)$;
2. $X(t) \geq 0$, the map $t \rightarrow R(t)$ is non decreasing and $R(0) = 0$;
3. $\int_0^\infty X(s) \, dR(s) = 0$

In this case the solution (X, R) satisfies $X(t) = Y(t) \vee \sup_{0 \leq s \leq t} (Y(t) - Y(s))$ and $R(t) = 0 \vee \sup_{0 \leq s \leq t} (-Y(s))$ for all $t \geq 0$. Assuming that the domain D was convex, Tanaka extended this result to the multidimensional case ([114]).

Harrison and Reiman ([54]) managed to develop a complete analytical theory for a multidimensional diffusion process whose state space lied in a non negative orthant. Reiman ([101]) applied the multidimensional reflection map to establish heavy traffic limits with multidimensional reflected Brownian motion limit for single class open queueing networks. Peterson ([93]) provided limit theorems for feedforward multiclass networks. A more difficult problem is to prove such a limit theorem for multiclass networks with feedback because the conditions for stability of a multiclass network are not well understood (see for instance [79] or [108]). Despite the lack of a limit theorem in the multiclass case, the research focused on developing a theory for the diffusions. In particular a large body of research aimed at lifting the conditions put on the domain and proving existence and uniqueness of the Sokhorod problem solution with an approach based on stochastic differential equations with reflection. For instance Dupuis and Ishii ([39]) focused on the case when the domain was a convex polyhedron and showed that the solution mapping was Lipschitz continuous when a certain convex set, defined in terms of the normal directions to the faces of the polyhedron and the directions of the constraint mechanism, could be shown to exist. In a subsequent paper ([40]), they were able to show the strong existence and uniqueness of a solution to the Skorokhod problem for domains that might have corners and for which the allowed directions of reflection at a point on the boundary of the domain could possibly be oblique. Dupuis and Ramanan improved this result in [37] and [38]. Other contributions include [18], [115], or [14].

In the meantime, Stroock and Varadhan ([113]) studied multidimensional diffusions with reflecting boundaries by introducing the generator of the diffusion and formulating the problem in a submartingale setting. They adapted their well known approach amounting to describe a diffusion by its initial condition and the condition that the process M , defined for all $t \geq 0$ by

$$M_t = f(X_t) - f(X_0) + \int_0^t \mathcal{L}f(X_s) \, ds,$$

be a martingale, to the study of diffusions with boundary conditions, requiring that the latter expression be a submartingale. Using this line of reasoning, Varadhan and Williams proved in [116] that if the state space was an infinite two dimensional wedge with an angle of ξ , there existed a unique strong Markov

process behaving as a Brownian motion inside the wedge and reflecting instantaneously on the two borders with angles of θ_1 and θ_2 if and only if $(\theta_1 + \theta_2)/\xi < 2$. Williams later showed in [119] that under the assumption that the directions of reflection satisfied a skew symmetry condition, a d dimensional reflected brownian motion with constant drift contained in the interior of a polyhedron could be characterized in terms of a family of submartingales. Other examples of this stream of literature include [76], [77], or [24].

Kang and Ramanan managed in [63] to obtain a full characterization of reflected diffusions in those domains leading to explicit formulas for stationary reflected Brownian motions with state dependent drifts. They proved in [64] the equivalence between the submartingale and the stochastic differential equations approaches in piecewise smooth domains in a multidimensional setting.

Another stream of recent research has focused on modeling queueing systems as point measure valued processes, in order to provide an exhaustive information on the state of the system and the residual processing times of the customers present in the queue at a given time for the renormalized process representing the average behaviour of the queue. A Dirac measure is therefore put at all processing times. Such a framework is particularly adequate to describe particles or branching systems (see [23], [80], [17]), or queueing systems whose dynamics are too complex to be carried on with simple finite-dimensional processes: the processor sharing queue (see [52], [103]), queues with deadlines (see [36] for a queue under the earliest deadline first service discipline without reneging, [[27], [28]] for the same system with reneging and [51] for a processor sharing queue with reneging), or the Shortest Remaining Processing time queue ([2], [86]). In [32] Decreusefond and Moyal derived a functional fluid limit theorem and a functional central limit theorem for a queue with an infinity of servers $M/GI/\infty$. In contrast, in [99] Kaspi and Ramanan introduce for the study of the more complicated $M/GI/N$ queue, a different measure-valued representation keeping track of the ages of customers in service, this representation offering the advantage of yielding semimartingale representations that are more amenable to computation.

2.2 Epidemiology

In 1927 Kermack and Mc Kendrick [68] proposed a model for the dynamics of epidemics. The population was assigned to three mutually exclusive compartments Susceptible, Infectious, or Recovered. The model is dynamic in that the numbers in each compartment may fluctuate over time and are therefore a function of time. Each member of the population typically progresses from Susceptible to Infectious to Recovered. We are interested in the number of individuals in each of these categories, and this number is denoted S, I and R respectively. Between S and I , the transition rate is assumed to be $d(S/N)/dt = -\beta SI/N^2$, where N is the total population, β is the average number of contacts per person per time, multiplied by the probability of disease transmission in a contact between a susceptible and an infectious subject, and SI/N^2 is the fraction of those contacts between an infectious and susceptible individual. This is a consequence of the Law of mass action : if the individuals in a population mix homogeneously, the rate of interaction between two different subsets of the population is proportional to the product of the numbers in each of the subsets concerned. Between I and R , the transition rate is assumed to be proportional

to the number of infectious individuals which is γI . β is the infection parameter and γ the removal parameter giving the rate at which infectives become immune.

Kermack and McKendrick obtained two basic results, referred to as their Threshold Theorem. Solving for the system of differential equations they obtained first that

$$\frac{dI}{dt} = \beta I(S - \frac{\gamma}{\beta})$$

In order for the epidemic to grow, we must have $dI/dt \geq 0$ and therefore S must be greater than γ/β so that the initial number of susceptibles must exceed a threshold value. Second,

$$\frac{S_\infty}{N} = -R_0^{-1} W(-\frac{S_0 R_0 \exp(-R_0(1 - R_0/N))}{N})$$

where W is the Lambert function defined for $x \in [-1/e, \infty]$ by the equation

$$W(x)e^{W(x)} = x.$$

This shows that at the end of an epidemic, unless $S_0 = 0$, not all individuals of the population have been removed, so some must remain susceptible. This means that the end of an epidemic is caused by the decline in the number of infectious individuals rather than an absolute lack of susceptible subjects.

Epidemic modelling has three aims. The first is to understand better the mechanisms by which diseases spread; for this, a mathematical structure is important. For example, the simple insight provided by Kermack and McKendrick's model that the initial proportion of susceptibles must exceed the critical value β/γ for an epidemic to grow, could not have been reached without their mathematical equations. The second aim is to predict the future course of the epidemic. Again using Kermack and McKendrick's general epidemic model as an example, we can try to estimate the number of infected people at the end, in order to estimate the medical costs of the epidemic, or to assess the possible impact of any outbreak of the disease, or to schedule hospital capacity. The third aim is to understand how we may control the spread of the epidemic by education, vaccination or isolation. In order to make reasonable predictions and develop methods of control, we must be confident that the model captures the essential features of the course of an epidemic. Thus, it becomes important to validate models by checking whether they fit the observed data.

The model proposed by Kermack and McKendrick was deterministic and assumed that the sizes of the compartments were large enough that the mixing of members be homogeneous. However, at the beginning of a disease outbreak, there is a very small number of infective individuals and the transmission of infection is a stochastic event depending on the pattern of contacts between members of the population. A deterministic formulation is therefore not suitable. Fortunately, the parameters of the model can be interpreted in a probabilistic way. For instance, the probability of an infectious individual recovering in any time interval dt is γdt . If an individual is infectious for an average time period T , then $\gamma = 1/T$. This is also equivalent to the assumption that the length of time spent by an individual in the infectious state is a random variable with an exponential distribution.

In the later 1940s, Bartlett ([5]) formulated a stochastic epidemic model by analogy with the Kermack-McKendrick deterministic model and stochastic models for epidemic processes started to proliferate.

The "classical" SIR model may be modified by using more complex and realistic distributions for the transition rates between the compartments or by adding new compartments to represent for instance vaccinated individuals, individuals that have been exposed to the disease but are not yet contagious, or individuals who are infected but asymptomatic, with lower infectivity than symptomatic individuals. It can also be modified to deal with diseases where people who have recovered lose their immunity and become susceptible again. The development and analysis of compartmental models has grown rapidly since the early models. Many of these developments are due to Hethcote ([55],[56],[59],[57],[58]). Age is one of the most important characteristics in the modeling of populations and infectious diseases. Individuals with different ages may have different reproduction and survival capacities. Diseases may have different infection rates and mortality rates for different age groups. Individuals of different ages may also have different behaviors, and behavioral changes are crucial in control and prevention of many infectious diseases. Young individuals tend to be more active in interactions with or between populations, and in disease transmissions. These issues are addressed by increasing the number of compartments in the model and by introducing a contact matrix which relates each compartment with the others taking into account an intercompartment mixing parameter.

A stochastic branching process description of the beginning of a disease outbreak begins with the assumption that there is a network of contacts of individuals, which may be described by a graph with members of the population represented by vertices and with contacts between individuals represented by edges. The study of graphs originated with the abstract theory of Erdős and Rényi of the 1950's and 1960's ([102], [41]), and has become important more recently in many areas. An edge is a contact between vertices that can transmit infection. The number of edges of a graph at a vertex is called the degree of the vertex. The degree distribution of a graph is p_k , where p_k is the fraction of vertices having degree k . The degree distribution is fundamental in the description of the spread of disease. Of course, for the modelling of infectious diseases, networks are bi-directional, with disease transmission possible in either direction along an edge.

We can define the generating function for all $z \in [0, 1]$ by

$$F_0(z) = \sum_{k=0}^{\infty} p_k z^k.$$

Since $\sum_{k=0}^{\infty} p_k = 1$, this power series converges for $0 \leq z \leq 1$, and may be differentiated term by term. Thus

$$p_k = \frac{F_0^{(k)}(0)}{k!}, k = 0, 1, 2, \dots$$

It is easy to verify that the generating function has the properties

$$F_0(0) = p_0, F_0(1) = 1, F_0'(z) > 0, F_0''(z) > 0.$$

The mean number of secondary infections, often called R_0 , is $F'_0(1)$.

These types of models have been studied by Diekmann and Heesterbeek ([33]), or by Callaway ([15]) and Newman ([88], [87]). One possible approach to a realistic description of an epidemic would be to use a branching process model initially and then make a transition to a compartmental model when the epidemic has become established and there are enough infectives that mass action mixing in the population is a reasonable approximation. Another approach would be to continue to use a network model throughout the course of the epidemic ([84],[83]). It is possible to formulate this model dynamically, and the limiting case of this dynamic model as the population size becomes very large is the same as the compartmental model. The theoretical analysis of network models is a very active and rapidly developing field ([82],[81]). In more recent models the connectivity graph is heterogeneous (see for instance [31], where Decreusefond, Dhersin, Moyal and Tran provide a rigorous individual-based description of the epidemic on a random graph. They show that three degree distributions are sufficient to describe the epidemic dynamics and describe these distributions by equations in the space of measures on the set of nonnegative integers)

2.3 Genetics

Pioneering work in the modelization of population genetics dates back to the 1940's, before the discovery of DNA by Crick and Watson in 1950 which provided the molecular basis of evolution. The Wright-Fisher-model ([46],[120]) is used to describe the evolution of a population of individuals (or genes) of two different types (alleles), called A and a. These types are neutral, meaning their reproductive success does not depend on the type, and their reproduction is random. Consider N individuals, each carrying one copy of a specific genetic locus (a location of interest in the genome). Suppose that at each time unit each individual randomly chooses another individual (possibly itself) from the population and adopts its type ("parallel updating"). This is called resampling, and is a form of random reproduction. Suppose that all individuals update independently from each other and independently of how they updated at previous times. Each time unit represents one generation. We are interested in the evolution of the number of A's at time n which we denote X_n^N (N denotes the population size and n the generation) The assumption behind the Wright Fisher model is the gene pool approach. Every individual produces a large number of gametes of the same type as the individual itself. (A gamete is a cell that fuses with another cell during fertilisation in organisms that reproduce sexually.) The offspring generation is then formed by sampling N times without replacement from this gene pool. This is basically the same as sampling with replacement from the parent population, so that effectively the offspring individuals choose their parents from the parent population with replacement and inherit their type.

The sequence $(X_n)_{n \in \mathbb{N}}$ is the discrete-time Markov chain on the state space $\Omega = \{0, 1, \dots, N\}$ with transition kernel

$$p(i, j) = \binom{N}{j} \left(\frac{i}{N} \right)^j \left(\frac{N-i}{N} \right)^{N-j} \text{ for all } (i, j) \in \Omega^2$$

The states 0 and N are absorbing.

It is of interest to consider the following space-time rescaling of our process:

$$Y_t^n = \frac{1}{n} X_{\lceil nt \rceil}^n \text{ for all } t \in [0, T], n \in \mathbb{N}$$

Here, $\lceil \cdot \rceil$ denotes the upper integer part. $Y^n(t)$ represents the fraction of individuals of type A in the population at time t on time scale n . We expect that, in the limit as $n \rightarrow \infty$, if the initial condition scales properly, i.e.,

$$\lim_{n \rightarrow \infty} Y_0^n = Y_0$$

then the rescaled process converges in law to a limiting process, living on state space $[0, 1]$ and evolving in continuous time. This limiting process, which must also be a Markov process, turns out to be a diffusion given by the stochastic differential equation (SDE)

$$dY_t = \sqrt{Y_t(1 - Y_t)} dW_t$$

where $(W_t)_{t \geq 0}$ is a standard Brownian motion. This SDE has a unique strong solution for a given Y_0 . Indeed, the process Y^n can be approximated by a continuous-time analogue: instead of waiting a deterministic time $1/n$ before making a transition, wait an exponential time of rate n . This gives a Markov process with state space $\{0, 1/n, \dots, 1 - 1/n, 1\}$ and infinitesimal generator $(\mathfrak{L}_n f)$ given by

$$\mathfrak{L}_n f\left(\frac{i}{n}\right) = n \sum_{j=0}^n p_n(i, j) \left[f\left(\frac{j}{n}\right) - f\left(\frac{i}{n}\right) \right]$$

The limiting process Y is the diffusion with state space $[0, 1]$ and infinitesimal generator $\mathfrak{L}$ given by

$$\mathfrak{L}f(y) = \frac{1}{2} y(1 - y) f''(y).$$

Limiting ourselves to functions in $\mathcal{C}_0([0, 1])$ we can use the Taylor expansion of f around $i/2N$ up to second order to prove the convergence of generators. There is a continuous-time version of the Wright Fisher model, called the Moran model (see [85]), in which each individual chooses a random ancestor at rate 1 and adopts its type. In other words, the resampling is done sequentially rather than in parallel. The resulting process $X = (X_t)_{t \geq 0}$ is the birth-death process on the state space Ω with transition rates

$$i \rightarrow i + 1 \text{ at rate } b_i = (N - i)i/N$$

$$i \rightarrow i - 1 \text{ at rate } d_i = (N - i)i/N$$

Note that $b_i = d_i, i \in \Omega$, and $b_0 = d_0 = b_N = d_N = 0$. The limit of the scaled Moran Model is also a Wright Fisher diffusion.

A third model is of interest. The coalescent is a model of how gene variants sampled from a population may have originated from a common ancestor. In the simplest case, each variant is equally likely to have been passed from one generation to the next. The model looks backward in time, merging alleles into a single ancestral copy according to a random process in coalescence events. It was developed by John Kingman ([71]) and Donnelly and Kurtz ([34]) found

a way to extend the study of such genealogical trees to infinite populations in a rigorous manner.

We can modify the basic Wright Fisher model in the following manner. At each time unit each individual, immediately after it has chosen its ancestor and adopted its type (resampling), suffers a type mutation: type a spontaneously mutates into type A with probability u , and type A spontaneously mutates into type a with probability v . Here, $0 < u, v < 1$, and mutations occur independently for different individuals. $X = (X_n)_{n \in \mathbb{N}}$ is a Markov chain on the state space $\Omega = \{0, 1, \dots, N\}$ with transition kernel

$$p(i, j) = \binom{N}{j} (p_i)^j (1 - p_i)^{N-j} \text{ for all } (i, j) \in \Omega^2$$

with

$$p_i = \frac{i}{N}(1 - v) + \frac{N - i}{N}u$$

Indeed, either an A is drawn (probability i/N) and it does not mutate into an a (probability $1 - v$), or an a is drawn (probability $(N - i)/N$) and it does mutate into an A (probability u). A first consequence of the presence of mutation is that the states 0 and N are no longer absorbing: $p(i, j) > 0$ for all (i, j) . Griffiths ([50]) uses diffusion methods to study lines of descent in the Wright-Fisher process with mutation. Recent research has studied much more complex models where Darwinistic selection, mutations or population growth are allowed to play a role.

3 Our contribution: a general approach to the rate of convergence in law

In all the above examples, the focus has been on convergence and identification of the limit. However, although the literature on speed of convergence of moments is already wide (see for instance [104],[103]), work on the rate of convergence in law is much more scarce. The basic idea at the root of the study of convergence in law is that the sequence of random variables being studied becomes better and better modeled by a given probability distribution as n increases. For instance, the central limit theorem states that as n increases the average of n independent variables identically distributed becomes closer to a Gaussian distribution. The Berry Esseen theorem goes a step further and gives a bound on the maximal error of approximation between the normal distribution and the true distribution of the sample average: for independent variables the convergence rate is $n^{-1/2}$. If our variables are not single point variables but functions of time chosen randomly in a space of functions, i.e. stochastic processes, an equivalent of the central limit theorem is the Donsker theorem, which states that under certain circumstances, a random walk will converge to a Brownian motion. We have concentrated our work on finding equivalents to the Berry Esseen theorem for a class of stochastic processes, that is to say to assess the speed of convergence of a sequence of stochastic processes to a limit.

Several works from Dai and Braverman ([12],[11]) have focused on a somewhat related subject, which is the rate of convergence of the steady state diffusion for Erlang-C, Erlang-A, and $M/Ph/n + M$ queueing systems. They prove,

using Stein's method, that the distance between the stationary distribution of the normalized process and that of an Ornstein–Uhlenbeck process is bounded. Our work is not centered on the steady state process but focuses on the process itself, when properly renormalized. We therefore need by contrast to use the Stein method in an infinite dimensional setting as our state space of study is a set of functions.

We started our work with the study, published in *Queueing Systems: Theory and Applications* (see [7]), of the $M/M/1$ and $M/M/\infty$ queueing systems for which we have obtained relatively high speeds of convergence, of the order of $\log(n)/\sqrt{n}$. In both cases we made full use of an adequate representation of the underlying process. The speed of convergence was then derived with the help of the Stein method, which is thoroughly explained in the present document, and of linear interpolation of each process in the sequence and its supposed limit. The drawback of our method was its dependency on the exact representation of the sequence of processes. We have therefore managed to generalize our work to a wider class of Markovian processes, at the cost of a loss in accuracy of estimate. In the course of this work, we have been led to make explicit, for this class of processes, how to define scale parameters in space and time in order to obtain a diffusion approximation. As a matter of fact the available literature, which covers wide fields of application seems to propose at each occurrence ad hoc parameters in order to obtain convergence: in the $M/M/1$ queue the arrival intensity is rescaled as well as the service time, in the $M/M/\infty$ system only the arrival intensity is rescaled, in the On/Off model and in the SIR model it is the size of the population which is rescaled. In the Moran model, population is rescaled linearly but the intensity of mutations gets a quadratic rescaling (see for instance [65]). We have been therefore led to propose a unified method to determine scale parameters for the class of Markovian models we consider. Using the properties of the class of Markovian processes which we identified, and continuing to exploit the Stein method and linear interpolations, we assessed the speed of convergence to be of the order of $n^{-1/6}$. This loss in the accuracy of speed of convergence is fully attributable to the representation of the processes under study as belonging to the identified class of processes. This work is the subject of an article which is going to be submitted for publication soon.

Chapter 2

Roadmap

1 The goal

Our aim being to establish the speed of convergence in law, we were compelled to adopt a definition for the distance between probability laws. Such a notion of distance between probability measures appears naturally not only in the study of convergence in law but also in the calculation of errors in Bayesian statistics and in fields beyond probability and statistics, such as information theory, machine learning or cryptology. The variety of applications and the necessity to handle diverse mathematical objects has led mathematicians to propose a wide multiplicity of definitions for distances between measures, and various methods to bound them. For the study of convergence in law, we can mention for instance the Prokhorov distance valid in any Polish space (see [8]). For all measures ν and μ on this Polish space:

$$d(\mu, \nu) = \inf\{\varepsilon: \mu(A) \leq \nu(B_\varepsilon(A)) + \varepsilon, \nu(A) \leq \mu(B_\varepsilon(A)) + \varepsilon \text{ for every Borel set } A\},$$

where $B_\varepsilon(A)$ denotes the open ball of radius ε around A :

$$B_\varepsilon(A) = \{y : \exists x \in A, d(x, y) < \varepsilon\}$$

This distance is theoretically attractive because its convergence to zero is equivalent to convergence in law but its computation is not easy. One may therefore prefer to introduce the Wasserstein distance, also called Kantorovich Rubinstein distance (see [117],[118]):

$$W(\mu, \nu) = \sup \left\{ \int_E f d\mu - \int_E f d\nu, f \in \text{Lip}_1(E) \right\}$$

where $\text{Lip}_1(E)$ stands for the set of Lipschitz functions with a coefficient lower than 1. The convergence under the Wasserstein distance implies convergence in law and is easier to compute, all the more as the Kantorovich Rubinstein theorem gives a second equivalent definition for the Wasserstein distance as it states that

$$W(\mu, \nu) = \inf_{\pi} \int_{E \times E} d(x, y) d\pi(x, y)$$

where π is taken in the set of all laws on $E \times E$ with marginals μ and ν .

2 Overview of the Stein Method

The Stein method, invented by Charles Stein in 1972 (see [112]), enables to bound any distance d between two probabilistic measures ν and μ which can be written as

$$d(\mu, \nu) = \sup \left\{ \int_E f d\mu - \int_E f d\nu, f \in \mathcal{K} \right\},$$

where $\mathcal{K}$ denotes a sufficiently large class of functions. Consider for instance a random variable X with a probability distribution absolutely continuous with respect to the Lebesgue measure and a density denoted $f(x)$. If we assume that

$$\lim_{x \rightarrow -\infty} f(x) = \lim_{x \rightarrow +\infty} f(x) = 0,$$

for all absolutely continuous function g (and hence differentiable almost everywhere), growing at a reasonable pace at infinity (i.e. such that $\lim_{|x| \rightarrow \infty} f(x)g(x) = 0$) we see that

$$\mathbf{E} \left[g'(X) + g(X) \frac{f'(X)}{f(X)} \right] = \int_{-\infty}^{\infty} (g'(x)f(x) + g(x)f'(x)) dx = \left[f(x)g(x) \right]_{-\infty}^{\infty} = 0.$$

This class of absolutely continuous functions such that $\lim_{|x| \rightarrow \infty} f(x)g(x) = 0$ being sufficiently large, we can hope that the equation

$$\forall g, \mathbf{E} \left[g'(X) + g(X) \frac{f'(X)}{f(X)} \right] = 0 \quad (2.1)$$

characterises density f and it is indeed the case: let A be a Borel set and h_A the function defined by

$$h_A(y) = \mathbf{1}_A(y) - \mathbf{P}(X \in A).$$

Let $g_A(y) = 0$ if $y \notin \text{Supp}(X)$ and

$$\begin{aligned} g_A(y) &= \frac{1}{f(y)} \int_{-\infty}^y h_A(u) f(u) du \\ &= \frac{1}{f(y)} (\mathbf{P}(X < y, X \in A) - \mathbf{P}(X < y) \mathbf{P}(X \in A)) \end{aligned}$$

otherwise. It is then clear that $\lim_{|y| \rightarrow \infty} f(y)g_A(y) = 0$ and that g_A is absolutely continuous; furthermore, since

$$(g_A(y)f(y))' = h_A(y)f(y),$$

we get

$$g'_A(y) + g_A(y) \frac{f'(y)}{f(y)} = h_A(y) = \mathbf{1}_A(y) - \mathbf{P}(X \in A).$$

Taking the expected value

$$\mathbf{E} \left[g'_A(Y) + g_A(Y) \frac{f'(Y)}{f(Y)} \right] = \mathbf{P}(Y \in A) - \mathbf{P}(X \in A),$$

and if equation (2.1) is verified for all functions g_A associated to a Borel set A , it is obvious that Y and X have same law. The operator which maps g to

$\mathbf{E} \left[g'(x) + g(x) \frac{f'(x)}{f(x)} \right]$ is called the Stein operator.

The intuition in the Stein method is therefore as follows. Suppose that μ follows a known law with density f and that we aim at computing its distance with another law ν . Suppose that variables X and Z are distributed according to ν and μ respectively. We can write Stein's equation

$$g'(x) + g(x) \frac{f'(x)}{f(x)} = h(x) - \mathbf{E}_\mu [h(Z)] \text{ for all } x,$$

and try to find a solution g_h for a sufficiently large class of h functions. The idea is to bound

$$\mathbf{E}_\nu \left[g'_h(x) + g_h(x) \frac{f'(x)}{f(x)} \right],$$

which is equal by definition to $\mathbf{E}_\nu [h(X)] - \mathbf{E}_\mu [h(Z)]$. In the case where μ is a standard Gaussian distribution, the Stein equation is:

$$g'(x) - xg(x) = h(x) - \mathbf{E} [h(Z)].$$

This equation is by no way unique as it is possible to find other Stein operators vanishing over a sufficiently large class of functions (see [90]). In 1988 Barbour ([3]) proposed to use the generator of a Markovian semi group making the μ probability stationary. Indeed, if μ is a stationary probability for X with a generator denoted A , it is well known that there is an equivalence (see proposition 9.2 in [43]) between:

1. $X \stackrel{d}{=} \mu$, and
2. $\mathbf{E}_\mu [Af(X)] = 0$ for all f in $D(A)$.

For instance, staying with the Gaussian distribution, if B is a standard Brownian motion, it is a fact that the Ornstein Uhlenbeck process $(X_t)_{t \geq 0}$ defined by

$$dX_t = -X_t dt + \sqrt{2} dB_t$$

is stationary for the Gaussian distribution. Indeed, by the Ito formula,

$$X(t) = e^{-t}x + \sqrt{2} \int_0^t e^{-(t-s)} dB_s,$$

so that $X(t) \sim \mathcal{N}(e^{-t}x + \sqrt{1 - e^{-2t}})$ which implies the Mehler representation formula:

$$P_t F(x) = \int_{\mathbb{R}} F(e^{-t}x + \sqrt{1 - e^{-2t}}y) \mu(dy),$$

where μ is the Gaussian standard measure on $\mathbb{R}$. Then

$$\int_{\mathbb{R}} P_t F(x) \mu(dx) = \int_{\mathbb{R}} \int_{\mathbb{R}} F(e^{-t}x + \sqrt{1 - e^{-2t}}y) \mu(dy) \mu(dx) = \int_{\mathbb{R}} F(x) \mu(dx),$$

i.e. for all t , if X_0 has a Gaussian distribution,

$$\mathbf{E} [P_t F(X_0)] = \mathbf{E} [F(X_0)]$$

The generator A of the Ornstein Uhlenbeck being given by the equation $Af(x) = f''(x) - xf'(x)$, Barbour proposes to use the following expression as a Stein equation:

$$g_h''(x) - xg_h'(x) = h(x) - \mathbf{E}[h(Z)],$$

and we observe that we get a characterisation close to that obtained with the original Stein equation by differentiation.

In a nutshell, the Stein method enables to bound the distance between measures ν and μ by setting, if ν is the stationary law of X and A its generator,

$$d(\mu, \nu) = \sup \left\{ \int_E h d\mu - \int_E h d\nu, h \in \mathcal{C} \right\} = \sup_{h \in \mathcal{C}} (\mathbf{E}_\nu [Ag_h(X)]),$$

where g_h is the solution to the Stein equation

$$Ag_h(x) = h(x) - \mathbf{E}_\mu[h(Z)] \text{ for all } x.$$

The general outline of the Stein method to put a bound on the distance between a probability $\mathbf{P}$ that we try to approximate and a known probability measure denoted $\mathbf{Q}$ is therefore the following:

1. Choice of a space and of a metric which can be expressed as the superior bound of a difference between two expectations ;
 2. Choice of a Stein operator which vanishes for all functions of a variable distributed according to $\mathbf{Q}$ for a sufficiently large class of functions;
 3. Computation of a solution to the corresponding Stein equation ;
 4. Computation of a higher (upper) bound to the expected value of the generator of the solution (i.e. $\mathbf{E}_\nu [Ag_h(X)]$) in order to get a bound for the distance.
- This last step is the trickiest and many methods have been proposed based on coupling techniques or on concentration inequalities. These methods depend on the law of $\mathbf{P}$ but also on the class of functions on which a bound is sought, and ultimately on the definition chosen for the distance.

Since 2005 the Stein method has enjoyed a new interest and a significant development thanks to the introduction of the Malliavin calculus suggested by Nourdin and Peccati (see [91],[89]). Indeed, Malliavin calculus supplies integration by parts tools enabling simultaneously to make explicit the generator to use in the Stein equation, and to find more systematically a higher bound to the solution of such equation. It has become possible to extend this method to the case of abstract Wiener spaces ([74]) and to infinite-dimensional problems ([4], [109], [19])). The purpose of our work is to study how and to what extent the Stein method can be put to use to quantify, with the support of the tools offered by the Malliavin calculus, the speed of convergence of diffusion approximation processes towards their limit.

Note:

The Stein method can be extended to discrete probability laws. In 1974 Chen (see [16]) showed it was perfectly adapted to the approximation of a Poisson law. In fact, if X is a random variable having $\mathbb{N}$ or $\mathbb{Z}$ as support set, and denoting $p(i) = P(Y = i)$, we get the telescopic sum

$$\mathbf{E} \left[g(Y+1) - \frac{p(Y-1)}{p(Y)} g(Y) \right] = \sum_{i=0}^{\infty} p(i) \left(g(i+1) - \frac{p(i-1)}{p(i)} g(i) \right) = 0$$

for all function g such that $\lim_{n \rightarrow \infty} g(n)p(n-1) = 0$ if and only if Y is distributed according to the same law as X . This equation characterises the law of X since if $A \subset \mathbb{N}$ by defining g inductively by $g(i+1) - \frac{p(i-1)}{p(i)}g(i) = \mathbf{1}_A(i) - \mathbf{P}(X \in A)$ we can show as above that $\mathbf{E} \left[g(Y+1) - \frac{p(Y-1)}{p(Y)}g(Y) \right] = \mathbf{P}(Y \in A) - \mathbf{P}(X \in A)$. Therefore we can use as Stein equation

$$g(i+1) - \frac{p(i-1)}{p(i)}g(i) = h(i) - \mathbf{E}[h(X)] \text{ for all } i \geq 0$$

for instance, for a Poisson law of intensity λ , for all $i \geq 0$,

$$g(i+1) - \frac{i}{\lambda}g(i) = h(i) - \mathbf{E}[h(X)],$$

but the approach relying on the generator works as well.

3 Roadmap

Our goal is to use the Stein method to bound the rate of convergence of an approximation diffusion towards a diffusion depending on Poisson processes. This raises various difficulties: choice of a common functional space where to compare continuous diffusions and discrete processes, definition of a Stein operator for such infinite dimensional objects as stochastic processes with the aim to characterize convergence to Brownian motion, definition of a gradient and of an integration by parts for Poisson processes. Once all these building blocks will be gathered, we shall be able to calculate rates of convergence towards Brownian motions and we will need to use linear interpolation to make calculations in infinite dimensions tractable. So the roadmap of chapters 2 to 4 is as follows:

1. choice of a common functional space where we can define a distance between step function discrete processes and Hölder continuous Brownian motions
2. review of abstract Wiener space and identification of a generator for abstract Wiener processes
3. presentation of Malliavin calculus for Poisson processes and determination of an integration by parts formula
4. adaptation, through linear interpolation, of our theoretical apparatus, to reduce the infinite dimensional calculations to finite dimension

We shall then give some applications in the subsequent chapters, and illustrate the limitations of our model in the last chapter.

Chapter 3

Choice of a functional space

In order to compare the distribution of a Poisson process, having piecewise differentiable paths, with that of a Brownian motion, having continuous but nowhere differentiable paths, one needs to find a functional space which includes all possible trajectories for both processes. We are therefore going to select a space smaller than $\mathbb{D}$ but large enough to contain the linear interpolations of the step function trajectories of Poisson processes as well as the η -Hölderian trajectories (with $\eta < 1/2$) of the Brownian motion. Mathematical literature provides functional spaces which meet these constraints, and this section is devoted to a short introduction to those. We position ourselves from the start on a compact time interval, for instance $[0, T]$ where $T \in \mathbb{R}$.

1 Fractional Sobolev spaces

The Fractional Sobolev space (also called Slobodeckij space) $W_{\eta,p}$ for $\eta \in (0, 1]$ and $p \geq 1$ is defined as the closure of the set of $\mathcal{C}^1$ functions with respect to the norm

$$\|f\|_{\eta,p}^p = \int_0^T |f(t)|^p dt + \iint_{[0,T]^2} \frac{|f(t) - f(s)|^p}{|t - s|^{1+p\eta}} dt ds. \quad (3.1)$$

For $\eta = 1$, $W_{1,p}$ is the completion of $\mathcal{C}^1$ for the norm:

$$\|f\|_{1,p}^p = \int_0^T |f(t)|^p dt + \int_0^T |f'(t)|^p dt.$$

It is a Banach space which verifies the Sobolev inclusions ([1, 45]):

$$W_{\eta,p} \subset W_{\alpha,q} \text{ for } 1 \geq \eta \geq \alpha \text{ and } \eta - 1/p \geq \alpha - 1/q.$$

and

$$W_{\eta,p} \subset \text{Hol}(\eta - 1/p) \text{ for } \eta - 1/p > 0,$$

where $\text{Hol}(\eta - 1/p)$ denotes the space of Hölder continuous functions of parameter $\eta - 1/p$. Therefore, since $W_{1,p}$ is separable (see [13]), so is $W_{\eta,p}$. Let us compute the norm of the indefinite integral of a step functions in $W_{\eta,p}$.

Lemma 3.1. *Let $0 \leq s_1 < s_2 \leq T$ and h_{s_1, s_2} the function defined for all $t \geq 0$ by*

$$h_{s_1, s_2}(t) = \int_0^t \mathbf{1}_{[s_1, s_2]}(r) \, dr.$$

There exists $c > 0$ such that for all s_1, s_2 ,

$$\|h_{s_1, s_2}\|_{W_{\eta, p}} \leq c |s_2 - s_1|^{1-\eta}.$$

Proof. Observe that for all $s, t \in [0, T]$,

$$|h_{s_1, s_2}(t) - h_{s_1, s_2}(s)| \leq |t - s| \wedge (s_2 - s_1).$$

The result follows from the definition of the norm $W_{\eta, p}$. $\square$

This framework is necessary to prove the following result.

Corollary 3.2. *There exists $c > 0$ such that for any $m \in \mathbb{N}$, for any $i \in \{0, \dots, m-1\}$, we have*

$$\|\sqrt{m} \int_0^{\cdot} \mathbf{1}_{[i/m, (i+1)/m)}(s) \, ds\|_{W_{\eta, p}} \leq \frac{c}{m^{1/2-\eta}},$$

In the Appendix, we describe spaces of functions called Besov Liouville Spaces and denoted $\mathcal{I}_{\beta, p}^+$ where $0 < \alpha < 1$ and $p \geq 1$. It is a fact, proven in [44] (Theorem 27) that for $1 \geq \alpha > \beta > \gamma > \delta > 0$, the following embeddings are continuous and even compact:

$$W_{\alpha, p} \subset \mathcal{I}_{\beta, p}^+ \subset W_{\gamma, p} \subset \mathcal{I}_{\delta, p}^+.$$

We can therefore use indistinctively Besov Liouville and Sobolev-Slobodeckij spaces. It is a simple matter to show that for $\alpha < 1/2$ and $p = 2$, $\mathcal{I}_{\alpha, p}^+$ (and therefore $W_{\alpha, p}$) contains both Brownian motion and step functions representative of random Poisson process trajectories.

In what follows, we shall denote, as it is customary in the study of Besov Liouville spaces,

$$(I_{0+}^\alpha f)(x) = \frac{1}{\Gamma(\alpha)} \int_0^x f(t)(x-t)^{\alpha-1} dt, \quad x \geq 0,$$

$$(I_{T-}^\alpha f)(x) = \frac{1}{\Gamma(\alpha)} \int_x^T f(t)(t-x)^{\alpha-1} dt, \quad x \leq T,$$

where $\alpha \geq 0$, and Γ is the Euler function.

It is useful to keep in mind the following chart where arrows represent continuous embeddings. For all $\epsilon < 1/4$ and $\eta \in (1-\epsilon, 1)$,

$$\begin{array}{ccccccc} (\mathcal{I}_{1/2-2\epsilon, p}^+)^* & \xrightarrow{i_{\eta, p}^*} & (\mathcal{I}_{1,2}^+)^* & & & & \\ & & \downarrow \simeq & & & & \\ L^2 & \xrightarrow{I_{0+}^1} & \mathcal{I}_{1,2}^+ & \longrightarrow & \mathcal{I}_{\eta, \frac{1}{\eta-1/2+\epsilon}}^+ & \longrightarrow & \text{Hol}_0(1/2-\epsilon) \longrightarrow \mathcal{I}_{1/2-2\epsilon, p}^+ \end{array}$$

The pivot space, i.e. the Hilbert space being identified to its dual, is, in that context, the space $\mathcal{I}_{1,2}^+$, and not L^2 as is usually the case. This implies that $\mathfrak{i}_{\eta,p}^*$ is the adjoint of $\mathfrak{i}_{\eta,p}$ in this identification, where $\mathfrak{i}_{\eta,p}$ is the embedding of $\mathcal{I}_{1,2}^+$ into $\mathcal{I}_{\eta,p}^+$.

2 Wiener structure on a fractional Sobolev space

In order to use the Stein method, it is necessary to bestow a Gaussian structure on the space. Any diffusion being a semi martingale, our purpose to study convergence towards a diffusion approximation would imply at least the ability to demonstrate the convergence of the process towards a continuous martingale, i.e. a Brownian motion changed in time in the case of one-dimensional processes, even if it will be made clear later on that this can also be extended to multi-dimensional problems (see remark 4 after lemma 5.2). It is therefore in order to find a Stein operator, a generator, for a time changed Brownian motion. Remark first that if the change in time denoted γ , which is a positive, increasing and deterministic function is also Hölderian of exponent α , the time changed Brownian motion exhibits Hölderian trajectories of any exponent strictly lower than $\alpha/2$, so that its trajectories belong to $\mathcal{I}_{\beta,\infty} \subset \mathcal{I}_{\beta,2}$ for all $\beta < \alpha/2$. In order to build a suitable Stein operator we therefore need:

- to bestow a Gaussian structure on the space
- to define a differential calculus in order to describe a generator
- to identify the said generator

Abstract Gaussian Structure

The simplest way to construct the Wiener measure on $\mathcal{I}_{\beta,2}$, is to start with the Ito-Nisio theorem (see [60]). Let $(X_n, n \geq 1)$ be a sequence of random standard and independent Gaussian variables, defined over the same probabilistic space $(\Omega, \mathcal{A}, \mathbb{P})$. Let $(e_n, n \geq 1)$ be an orthonormal, complete basis of $\mathcal{L}^2([0,1]^n)$. Then the process defined by

$$\sum_{n \geq 1} X_n I_{0+}^1(e_n)(t), \text{ for all } t \in [0,1],$$

converges almost surely towards a process which is a standard Brownian motion B . Ito and Nisio have shown that such convergence is uniform with respect to t , which ensures the continuity of B . In order to show the convergence in $L^2(\Omega; \mathcal{I}_{\beta,2})$, it is enough to show that

$$\sum_{n \geq 1} \|I_{0+}^1 e_n\|_{\mathcal{I}_{\beta,2}}^2 = \sum_{n \geq 1} \|I_{0+}^{1-\beta} e_n\|_{\mathcal{L}^2}^2 = \|I_{0+}^{1-\beta}\|_{HS}^2 < \infty. \quad (3.2)$$

According to theorem 8.4, $I^{1-\beta}$ is a Hilbert-Schmidt operator from $\mathcal{L}^2$ into itself if and only if $1 - \beta > 1/2$, i.e. $\beta < 1/2$. Hence, for $\beta < 1/2$, the distribution of B defines the Wiener measure on $\mathcal{I}_{\beta,2}$ which we shall denote μ . Notice that (3.2) implies that the canonical embedding from $\mathcal{I}_{1-\beta,2}$ into $\mathcal{L}^2$ is also Hilbert-Schmidt and that its Hilbert-Schmidt norm is $\|I_{0+}^{1-\beta}\|_{HS}$. We equip

therefore $E = \mathcal{I}_{\alpha,p}^+$ (resp. $E = W_{\alpha,p}$) with a probability measure μ such that for all functions η in E^* ,

$$\mathbf{E} \left[e^{i \langle \omega, \eta \rangle_{E, E^*}} \right] = e^{-\frac{1}{2} \|\mathbf{i}^*(\eta)\|_H^2},$$

where H is the Hilbert space $\mathcal{I}_{\alpha,2}^+$ (resp. $W_{\alpha,2}$) and $\mathbf{i}^*$ the embedding of E^* in $H^* = H$. A result from Leonard Gross ([53]) asserts that μ exists and is unique as long as $\mathbf{i}$ is a Hilbert Schmidt operator (see [53], [106], or [107], theorem 7.1): there exists a random variable X in $\mathcal{I}_{\alpha,p}^+$ such that

$$\langle \mathbf{i}^*(\eta), \omega \rangle_H = \langle \eta, X \rangle_{\mathcal{I}_{\alpha,p}^{+*}, \mathcal{I}_{\alpha,p}^+}$$

for all $\eta \in \mathcal{I}_{\alpha,p}^{+*}$.

Remark 1. We can show that $\mathbf{i}^*$ is injective and that its image is dense in H^* . Moreover we can show that every linear function f from E^* into $\mathbb{R}$ is also square integrable and that in addition, $\|f\|_{L^2(E^*, \mathbb{R})} = \|\mathbf{i}^*(f)\|_H$. This injective isometry is continuous, from $\mathbf{i}^*(E^*)$ into L^2 and can be extended in a unique way into a linear mapping from H to L^2 denoted δ and called Wiener integral. For all $\eta \in E^*$, $\delta(\mathbf{i}^*(\eta)) = \langle \mathbf{i}(\eta), \omega \rangle_{E^*, E}$. We can show that when their definition domain coincide, the Wiener integral is equal to the Ito integral

$$\int_0^T \dot{\eta}(s) dB_s.$$

Gross-Sobolev derivative

Let us introduce now the notion of Gross-Sobolev derivative. Denote $\mathbf{P}_{\eta,p}$ the measure of B in $\mathcal{I}_{\eta,p}^+$. The $\mathcal{I}_{\eta,p}^+$ are Banach spaces, for which we can define a Fréchet derivative. A function $F : \mathcal{I}_{\eta,p}^+ \rightarrow \mathbf{R}$, is differentiable when

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (F(x + \varepsilon h) - F(x)) \quad (3.3)$$

exists for all $h \in \mathcal{I}_{\eta,p}^+$ and this defines a member of $(\mathcal{I}_{\eta,p}^+)^*$,

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (F(x + \varepsilon h) - F(x)) = \langle DF(x), h \rangle_{(\mathcal{I}_{\eta,p}^+)^*, \mathcal{I}_{\eta,p}^+}.$$

In particular, as in finite dimension, the Fréchet differentiability implies continuity. But the functions we consider are random variables, defined up to a negligible set, so that no continuity assumption can be satisfied. Moreover as expression (3.3) suggests, if $F = G$ almost surely, we need to ensure that $F(\cdot + h) = G(\cdot + h)$ almost surely for all $h \in W_{\eta,p}$, which implies that for $\bar{A} = \{F = G\}$,

$$\mathbf{P}(A) = 0 \Rightarrow \mathbf{E}[1_A] = 0 \Rightarrow \mathbf{E}[1_A \circ \tau_h] = 0 \Rightarrow \mathbf{P}(\tau_h^{-1}(A)) = 0;$$

i.e. the image measure of $\mathbf{P}_{\eta,p}$ by the mapping $\tau_h : x \mapsto x + h$ is absolutely continuous with respect to $\mathbf{P}_{\eta,p}$. In order for this property to be verified, the Cameron-Martin theorem imposes to restrain the perturbation domain of h to $\mathcal{I}_{1,2}^+$.

Theorem 3.3 (Cameron-Martin). *for all $h \in \mathcal{I}_{1,2}^+$, and all bounded functional $F : W_{\eta,p} \rightarrow \mathbf{R}$*

$$\mathbf{E}[F(B+h)] = \mathbf{E}\left[F(B) \exp\left(\int_0^T \dot{h}(s) dB(s) - \frac{1}{2}\|h\|_{\mathcal{I}_{1,2}^+}^2\right)\right], \quad (3.4)$$

where $\dot{h}$ is the derivative of $h \in \mathcal{I}_{1,2}^+$, so that $\dot{h}$ belongs to $L^2([0, T])$ the stochastic integral being understood as an Ito integral. In other words equation (3.4) means that the distribution of $B+h$ is absolutely continuous with respect to $\mathbf{P}_{\eta,p}$ and that the Radon-Nykodim derivative is given by the exponential factor in the right-hand part of equation (3.4).

Because of this theorem, $\mathcal{I}_{1,2}^+$ plays a crucial role in Malliavin calculus. We have the following structure:

$$\begin{array}{ccc} W_{\eta,p}^* & \xrightarrow{i_{\eta,p}^*} & (\mathcal{I}_{1,2}^+)^* \\ & \downarrow \simeq & \\ L^2 & \xrightarrow{I_{0+}^1} & \mathcal{I}_{1,2}^+ \xrightarrow{i_{\eta,p}} W_{\eta,p} \end{array}$$

$\mathcal{I}_{1,2}$, is called the Cameron-Martin space and is equipped with scalar product

$$\langle f, g \rangle_{\mathcal{I}_{1,2}} = \int_0^T \dot{f}(s) \dot{g}(s) ds.$$

and with the corresponding seminorm

$$\|f\|_{\mathcal{I}_{1,2}} = \|\dot{f}\|_{L^2}.$$

We would like to extend the notion of Fréchet derivative to larger sets than $\mathcal{I}_{1,2}^+$. With this purpose in mind we introduce the following definition.

Definition 3.4. *Let X be a Banach space. A function $F : W_{\eta,p} \rightarrow X$ is said to be cylindrical if it can take the following form:*

$$F(\omega) = \sum_{j=1}^k f_j(\delta h_1, \dots, \delta h_k) x_j, \quad (3.5)$$

Where for all $j \in \{1, \dots, k\}$, f_j belongs to the Schwartz space over $\mathbf{R}^k$, $(h_1, \dots, h_k)$ are members of $\mathcal{I}_{1,2}$, $(x_1, \dots, x_k)$ belongs to X and δh_k is the Wiener integral of $\dot{h}_k$:

$$\delta h_k = \int_0^T \dot{h}_k(s) dB(s).$$

The set of such functions is denoted $\mathbb{X}$.

We can now introduce the Gross Sobolev derivative

Definition 3.5. *For $h \in \mathcal{I}_{1,2}$,*

$$\langle \nabla F, h \rangle_{\mathcal{I}_{1,2}} = \sum_{j=1}^k \sum_{l=1}^k \partial_l f(\delta h_1, \dots, \delta h_k) \langle h_l, h \rangle_{\mathcal{I}_{1,2}} x_j,$$

Which amounts to write

$$\nabla F = \sum_{j,l=1}^k \partial_j f(\delta h_1, \dots, \delta h_k) h_l \otimes x_j.$$

Space $\mathbb{D}_{1,2}(X)$ is the closure of cylindrical functions with respect to the norm of $L^2(W_{\eta,p}; \mathcal{I}_{1,2} \otimes X)$. A member of $\mathbb{D}_{1,2}(X)$ is said to be Gross-Sobolev differentiable and ∇F belongs to $I_{1,2} \otimes X$ with probability 1.

We can iterate the construction to higher level gradients, thus defining $\nabla^{(k)} F$ for all $k \geq 1$, if F is regular enough. For all $p \geq 1$, the set of cylindrical functions is dense in L^p , which motivates the following definition

Definition 3.6. Let $\nabla^{(2)} F$ be the member of $L^2(W_{\eta,p}; \mathcal{I}_{1,2}^+ \otimes \mathcal{I}_{1,2}^+)$ defined by

$$\nabla^{(2)} F = \sum_{j,l=1}^k \partial_{jl}^{(2)} f(\delta h_1, \dots, \delta h_k) h_j \otimes h_l.$$

Consider the norm

$$\|F\|_{2,2}^2 = \|F\|_{L^2}^2 + \mathbf{E} \left[\|\nabla F\|_{\mathcal{I}_{1,2}^+}^2 \right] + \mathbf{E} \left[\|\nabla^{(2)} F\|_{\mathcal{I}_{1,2}^+ \otimes \mathcal{I}_{1,2}^+}^2 \right],$$

where

$$\|\nabla F\|_{\mathcal{I}_{1,2}^+}^2 = \int_0^T \left(\sum_{j=1}^k \partial_j f(\delta h_1, \dots, \delta h_k) \dot{h}_j(s) \right)^2 ds,$$

and

$$\|\nabla^{(2)} F\|_{(\mathcal{I}_{1,2}^+)^{\otimes 2}}^2 = \int_0^T \int_0^T \left(\sum_{j,l=1}^k \partial_{jl}^{(2)} f(\delta h_1, \dots, \delta h_k) \dot{h}_j(s) \dot{h}_l(r) \right)^2 ds dr.$$

The set $\mathbb{D}_{2,2}$ is the completion of the set of cylindrical functions with respect to norm $\|\cdot\|_{2,2}$.

From the definition of $\mathcal{I}_{1,2}^+$, it is clear that for $h \in \mathcal{I}_{1,2}^+$, $\dot{h}$ belongs to L^2 , and since

$$\delta h(\omega + \xi) = \langle h, \omega + \xi \rangle_{E^*, E} = \langle h, \omega \rangle_{E^*, E} + \langle h, \xi \rangle_{E^*, E} = \delta h(\omega) + \langle \mathbf{i}^*(h), \xi \rangle_{\mathcal{I}_{1,2}^+},$$

we get:

$$\begin{aligned} \int_0^t \dot{h}(s) d(B + k)(s) &= \int_0^t \dot{h}(s) dB(s) + \int_0^t \dot{h}(s) \dot{k}(s) ds \\ &= \int_0^t \dot{h}(s) dB(s) + \langle \dot{h}, \dot{k} \rangle_{L^2} = \int_0^t \dot{h}(s) dB(s) + \langle h, k \rangle_{\mathcal{I}_{1,2}^+}. \end{aligned}$$

so that

$$\langle \nabla F, k \rangle = \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (F(h + \varepsilon k) - F(h)).$$

Hence, if F is cylindrical,

$$\left. \frac{d}{d\varepsilon} F(B + \varepsilon k) \right|_{\varepsilon=0} = \sum_{j=1}^k \partial_j f(\delta h_1, \dots, \delta h_k) \langle h_j, k \rangle_{\mathcal{I}_{1,2}^+}.$$

Remark 2. If h belongs to $\mathcal{I}_{2,2}^\pm = (I_{0+}^1 \circ I_{T-}^1)(L^2) \subset \mathcal{I}_{1,2}^+$ then

$$\nabla f(\delta h) = f'(\delta h) h$$

belongs to $L^2(W_{\eta,p}; \mathcal{I}_{2,2}^\pm)$. This amounts to say that for such functional, its gradient is more regular, meaning that it belongs to a smaller space than other members of $\mathbb{D}_{2,2}$. Since we have identified $\mathcal{I}_{1,2}^+$ with its dual, space $\mathcal{I}_{2,2}^\pm$ is dual to L^2 : for $h \in \mathcal{I}_{2,2}^\pm$, there exists $\ddot{h} \in L^2$ such that $h = I_{0+}^1(I_{T-}^1(\ddot{h}))$. Then for $k \in \mathcal{I}_{1,2}^+$,

$$\begin{aligned} |\langle h, k \rangle_{\mathcal{I}_{1,2}^+}| &= \left| \int_0^T I_{T-}^1(\ddot{h})(s) \dot{k}(s) ds \right| = \left| \int_0^T \ddot{h}(s) I_{0+}^1(\dot{k})(s) ds \right| \\ &= \left| \int_0^T \ddot{h}(s) k(s) ds \right| \leq \|\ddot{h}\|_{L^2} \|k\|_{L^2}. \end{aligned}$$

Since $\mathcal{I}_{1,2}^+$ is dense within L^2 , we can extend the duality relationship to $h \in \mathcal{I}_{2,2}^\pm$ and $k \in L^2$.

Generator of Brownian motion

We are now in a position to describe the generator of Brownian motion which we are looking for. Consider the Ornstein-Uhlenbeck process on $\mathbb{R}$ solution to the following stochastic differential equation

$$dX(t) = -X_t dt + \sqrt{2} dB(t), \quad X(0) = x.$$

We can deduce from the Ito formula that

$$f(X_t) - f(X_0) = - \int_0^t X_s f'(X_s) ds + \sqrt{2} \int_0^t f'(X_s) dB_s + \int_0^t f''(X_s) ds.$$

If we set

$$P_t f(x) = \mathbf{E}_x[f(X)] \text{ and } Lf(x) = -xf'(x) + f''(x),$$

then, since the expectation of a stochastic integral is zero,

$$P_t f(x) = f(x) + \int_0^t \mathbf{E}_x[Lf] ds = f(x) + \int_0^t P_s[Lf] ds. \quad (3.6)$$

So that by differentiation

$$\frac{d}{dt} P_t f(x) = P_t Lf(x).$$

In particular, for $t = 0$, L is the generator of the P_t semi group and we have the classical equality

$$P_t Lf(x) = L P_t f(x). \quad (3.7)$$

Moreover the distribution of X knowing that $X(0) = x$ is Gaussian with an expectation of $e^{-t}x$ and a variance of $(1 - e^{-2t})$ hence if $\mu \sim \mathcal{N}(0, 1)$,

$$P_t f(x) = \int_{\mathbf{R}} f(e^{-t}x + \sqrt{1 - e^{-2t}}y) d\mu(y).$$

We can deduce easily that if $X(0)$ follows a standard normal distribution, $X(t)$ is Gaussian with zero expectation and unit variance: $\mathcal{N}(0, 1)$ is the invariant law of X . We can check easily that

$$P_t f(x) \xrightarrow{t \rightarrow \infty} \int_{\mathbf{R}} f \, d\mu.$$

By combining this last equality with (3.6) with (3.7) we get

$$\int_{\mathbf{R}} f \, d\mu - f(x) = \int_0^\infty LP_t f(x) \, dt.$$

Which gives us effectively the following Stein equation :

$$\int_{\mathbf{R}} f \, d\mu - \mathbf{E}[f(X)] = \mathbf{E} \left[\int_0^\infty LP_t f(X) \, dt \right].$$

We now transfer this construction to infinite dimension in an abstract Wiener space. Following Barbour ([4]) we define an Ornstein Uhlenbeck process W on an abstract space by taking a collection (X_k) of independent Ornstein Uhlenbeck processes identically distributed on $[0, \infty[$ as in the unidimensional case by setting

$$W(t, u, \omega) = \sum_{k \geq 0} X_k(u, \omega) S_k(t),$$

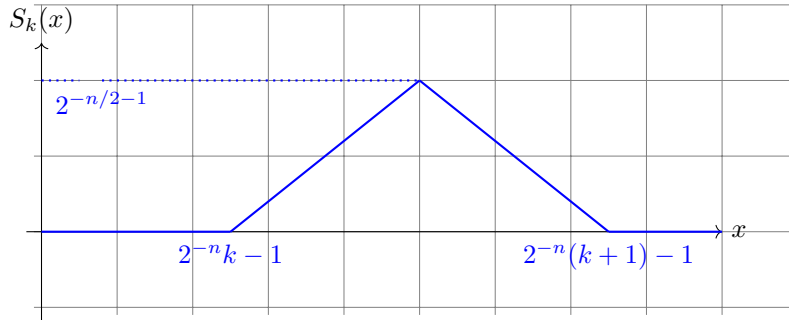
where S_k is the k -th Schauder function defined for all $t \in \mathbb{R}^+$ by $S_0(t) = t$, and for $2^n \leq k < 2^{n+1}$,

$$S_k(t) = \int_0^t H_k(s) \, ds,$$

where H_k is the Haar function

$$H_k(s) = 2^{n/2} (\mathbf{1}_{[2^{-n}k-1; 2^{-n}(k+1/2)-1]}(s) - \mathbf{1}_{[2^{-n}(k+1/2)-1; 2^{-n}(k+1)-1]}(s)).$$

Graph of S_k



The series used to define W is uniformly convergent since (see [78]), the S_k

functions are positive and bounded by $2^{-n/2-1}$ where n is the integer such that $2^n \leq k < 2^{n+1}$. There support set are disjoint and if we set

$$b_n(\omega) = \max_{2^n \leq k < 2^{n+1}} |X_k(u, \omega)|, \quad (3.8)$$

we can easily see by the Cauchy criterion that the series is uniformly converging and even absolutely as long as

$$\sum_n b_n 2^{-n/2} < \infty, \quad (3.9)$$

which is almost surely the case for the Ornstein Uhlenbeck processes which are Gaussian, so that

$$\mathbf{P}(|X_k(u, \omega)| = O(k^\varepsilon)) = 1 \text{ for all } \varepsilon > 0 \quad (3.10)$$

(A classical argument makes use of Tchebichev inequality to show that, because X is Gaussian with even order moments $\mathbf{E}[X_k^{2N}]$ equal to $1.3.5...N$, for all N

$$\mathbf{P}(|X_k(u, \omega)| \geq k^\varepsilon) \leq \frac{\mathbf{E}[X_k^{2N}]}{k^{2\varepsilon N}} = \frac{1.3.5...N}{k^{2\varepsilon N}},$$

and by summation over k , this converges if we choose a sufficiently large N (i.e $N > 1/(2\varepsilon)$), so that thanks to the Borel Cantelli lemma, with probability 1, there is only a finite number of events $\{|X_k(u, \omega)| \geq k^\varepsilon\}$. This proves (3.10) and according to the definition of b_n given in (3.8), $b_n(\omega) < C2^{n\varepsilon}$ so that for $\varepsilon < 1/2$ (3.9) is verified. (Note: such convergence can also be proved through the Ito Nisio theorem [60])

Remark at last that functions $(H_k)_{k \in \mathbb{N}}$ form a complete orthonormal basis and therefore,

$$\begin{aligned} \text{Cov}[W(t, u), W(s, u)] &= \mathbf{E} \left[\sum_k X_k(u) S_k(t) \sum_l X_l(u) S_l(s) \right] \\ &= \sum_{k, l} \mathbf{E}[X_k(u) X_l(u)] S_k(t) S_l(s) \\ &= \sum_k S_k(t) S_k(s) \\ &= \sum_k \int_0^t H_k(v) dv \int_0^s H_l(v) dv \\ &= \sum_k \langle \mathbf{1}_{[0, t]}, H_k \rangle \langle \mathbf{1}_{[0, s]}, H_k \rangle \\ &= \langle \mathbf{1}_{[0, t]}, \mathbf{1}_{[0, s]} \rangle \\ &= s \wedge t. \end{aligned}$$

Thus, W is almost surely a continuous, Gaussian function of (t, u) and for all u , $W(., u)$ is distributed according to the Wiener measure. Moreover, if we set $\mathcal{F}_u = \sigma(W(., v) v \leq u)$, W has the characteristic Ornstein Uhlenbeck process property:

$$W(., u + v) = e^{-v} W(., u) + (\sqrt{1 - e^{-2v}}) Z(.),$$

where Z is a standard Brownian motion independent from $\mathcal{F}_u$. If $u = 0$ the initial measure is the Wiener measure, $W(., v)$ is also distributed according to a Wiener process which shows that this measure is invariant. Let f be a continuous function from E to $\mathbf{R}$. The operator P_u is defined such that for all $\omega \in E$

$$(P_u f)(\omega) = \mathbb{E}[f(W(., u)) | W(., 0) = \omega] = \mathbf{E} \left[f(\omega e^{-u} + \sqrt{1 - e^{-2u}} Z) \right].$$

Here again, thanks to the dominated convergence theorem,

$$P_u F(x) \xrightarrow{u \rightarrow \infty} \int_E F \, d\mu_Z.$$

The generator of this Ornstein Uhlenbeck process in infinite dimension has been computed by Piech in L^2 (see [94]). Barbour proves in [4]

Theorem 3.7. *For $F \in \mathcal{C}(E)$, twice Fréchet differentiable and such that*

$$\mathbf{E} \left[\|\nabla^{(2)} F(Z)\|_{trH} \right] < \infty, \text{ for all } x \in E,$$

the generator of W denoted L is

$$LF(x) = -\langle x, \nabla F(x) \rangle_{E, E^*} + \sum_{i=1}^{\infty} \langle \nabla^{(2)} F(x), h_i \otimes h_i \rangle_H,$$

where $(h_n, n \geq 1)$ is a complete orthonormal basis of H so that the Stein equation is

$$-\langle x, \nabla F(x) \rangle_{E, E^*} + \sum_{i=1}^{\infty} \langle \nabla^{(2)} F(x), h_i \otimes h_i \rangle_H = h(x) - \mathbf{E}[h(Z)],$$

and its solution is

$$\begin{aligned} F_h(x) &= - \int_0^{\infty} (P_t h(x) - \mathbf{E}[h(Z)]) \, dt \\ &= - \int_0^{\infty} \int_B (h(xe^{-t} + \sqrt{1 - e^{-2t}} y) - \mathbf{E}[h(Z)]) \mu_Z(dy) \, dt, \end{aligned}$$

so that we get the identity

$$\begin{aligned} -\frac{d}{dt} P_t h(x) &= L P_t h(x) \\ &= -\langle x, \nabla(P_t h)(x) \rangle_{E, E^*} + \sum_{i=1}^{\infty} \langle \nabla^{(2)} P_t h(x), h_i \otimes h_i \rangle_H. \end{aligned} \tag{3.11}$$

Shih [109] (Theorems 4.8 and 4.10) shows that this Theorem remains true for the larger class of Lipschitz functions $F \in Lip_1(E)$. In particular, for F Lipschitz, $\nabla P_t F$ takes its values in $E^* \subset H$ and $\nabla^{(2)} P_t F(x)$ is trace class. In addition we get the two following semigroup properties ([109], equations (4.5) and (4.7)):

$$\langle \nabla P_t F(x), z \rangle = \frac{e^{-t}}{\sqrt{1 - e^{-2t}}} \int_B \langle y, z \rangle \nabla F(e^{-t} x + \sqrt{1 - e^{-2t}} y) \, d\mu_B(y) \tag{3.12}$$

and

$$\begin{aligned} & \langle \nabla^{(2)} P_t F(x), z_1 \otimes z_2 \rangle \\ &= \frac{e^{-t}}{\sqrt{1-e^{-t}}} \int_B \langle y, z_1 \rangle \langle \nabla P_{t/2} F(e^{-t/2}x + \sqrt{1-e^{-t}}y), z_2 \rangle d\mu_B(y) \end{aligned} \quad (3.13)$$

In order to bound the distance between the diffusion approximation and its limit, we can make use of the identity:

$$\begin{aligned} \mathbf{E}_\nu[h(x)] - \mathbf{E}_{m_B}h(Z) &= -\mathbf{E}_\nu \left[\int_0^\infty \langle x, \nabla F_h(x) \rangle_{E, E^*} dt \right] \\ &+ \sum_{i=1}^\infty \mathbf{E}_\nu \left[\int_0^\infty \langle \nabla^{(2)} F_h(x), h_i \otimes h_i \rangle_H dt \right], \end{aligned} \quad (3.14)$$

where ν is the law of the diffusion approximation whose limit is being sought and m_B is the law of a standard Brownian motion.

Chapter 4

Stein Method in Wiener fractional Sobolev Spaces

1 Malliavin calculus for Poisson processes

Finding a solution to the Stein equation presumes therefore the computation of (3.14) whereas our initial variable follows a Poisson process. We introduce now the elementary requirements to have an integration by parts for point Poisson processes which will enable us to compute the first term on the left hand side of (3.14). Malliavin calculus for Poisson processes is for instance made fully explicit in [29, 97]. Let $\mathbb{X}$ be a metric space, separable and complete equipped with a measure σ -finite ν and $E = \mathbf{R}^+ \times \mathbb{X}$. Let $\mathfrak{N}_E$ be the space of locally finite configurations in E , i.e. the collection of at most countable subsets of E without limit points. Such a set ϕ can be described as a set or as a sum of atomic measures:

$$\phi \simeq \sum_{x \in \phi} \varepsilon_x,$$

where ε_x is the Dirac measure on x , so that for all $\psi : E \rightarrow \mathbf{R}$,

$$\int_E \psi \, d\phi = \sum_{x \in \phi} \psi(x).$$

For a σ -finite measure ν on E , a Poisson point process with control measure ν is a random variable taking values in $\mathfrak{N}_E$, denoted for instance $\mathcal{N}_\nu$, such that for all $\psi : E \rightarrow \mathbf{R}$, with a compact support,

$$\mathbf{E} \left[\exp \left(- \sum_{x \in \mathcal{N}_\nu} \psi(x) \right) \right] = \exp \left(- \int_E 1 - e^{-\psi(x)} \, d\nu(x) \right).$$

Then for all functions $u = (u(s, z), s \in [0, T], z \in \mathbf{R}^+)$ in $L^2(\nu)$,

$$\int_0^T \int_{\mathbf{R}^+} u(s, z)^2 \, d\nu(s, z) < \infty, \text{ i.e. } u \in L^2(\nu),$$

the process

$$\begin{aligned} t \mapsto (\delta_\nu u)(t) &= \sum_{T_n \leq t} u(T_n, Z_n) - \int_0^t \int_{\mathbf{R}^+} u(s, z) \, d\nu(s, z) \\ &= \int_E u \, d\mathcal{N}_\nu - \int_0^t \int_{\mathbf{R}^+} u(s, z) \, d\nu(s, z) \end{aligned}$$

is a square integrable martingale.

A Poisson measure is simple, meaning its control measure does not have atoms, i.e. for all $x \in E, \nu(\{x\}) = 0$. The set of simple measures $\mathcal{N}_s$ is a measurable subset of $\mathfrak{N}_E$ and the implied sigma algebra $\mathcal{N}_s$ is generated by the family of sets taking the form

$$N_{C,0} = \{\eta \in \mathfrak{N}_E, \eta(C) = 0\} \text{ where } C \text{ runs through } \mathcal{B}.$$

This family is stable by intersection.

Let f be a function defined on $\mathfrak{N}_E$. We are interested in the so-called discrete gradient [29, 97]:

$$D_{s,z}f(N_\nu) = f(N_\nu + \varepsilon_{s,z}) - f(N_\nu), \quad s \in [0, T], \quad z \in E,$$

where $N_\nu + \varepsilon_{s,z}$ represents the configuration N_ν to which an atom is added at time s with mark z . Since ν is diffuse, there is a zero probability that an atom at time s already exists in N_ν . Similarly we denote $N_\nu - \varepsilon_{s,z}$ the configuration N_ν from which an $\varepsilon_{s,z}$ atom is removed provided it is present in N_ν , otherwise N_ν remains unchanged.

The Campbell-Mecke formula asserts that if X is a Poisson point process whose control measure is non atomic,

$$\mathbf{E} \left[\sum_{x \in \mathcal{N}_\nu} F(x, \mathcal{N}_\nu) \right] = \int_E \mathbf{E} \left[F(x, \mathcal{N}_\nu + \varepsilon_x) \right] \, d\nu(x),$$

for all positive functions F in $E \times \mathfrak{N}_E$.

Proof: Since X is simple and since the sets $N_{C,0}, C \in \mathcal{B}$, generate the sigma algebra $\mathcal{N}_s$, it is enough to consider the case $F = \mathbf{1}_{B \times N_{C,0}}$, where $B, C \in \mathcal{B}$ are

Borel sets and $\nu(B) < \infty$. Then, because of the independence of measures

$$\begin{aligned}
\int_E \mathbf{E} \left[F(x, \mathcal{N}_\nu + \varepsilon_x) \right] d\nu(x) &= \int_B \mathbf{E} \left[\mathbf{1}_{\{(\mathcal{N}_\nu + \varepsilon_x)(C)=0\}} \right] d\nu(x) \\
&= \int_{B \cap C} \mathbf{P}[\mathcal{N}_\nu(C) + 1 = 0] d\nu(x) \\
&+ \int_{B \setminus C} \mathbf{P}[\mathcal{N}_\nu(C) = 0] d\nu(x) \\
&= \mathbf{P}[\mathcal{N}_\nu(C) = 0] \nu(B \setminus C) \\
&= \mathbf{E} \left[\mathbf{1}_{\{\mathcal{N}_\nu(C)=0\}} \right] \mathbf{E}[\mathcal{N}_\nu(B \setminus C)] \\
&= \mathbf{E} \left[\mathbf{1}_{\{\mathcal{N}_\nu(C)=0\}} \mathcal{N}_\nu(B) \right] \\
&= \mathbf{E} \left[\sum_{x \in \mathcal{N}_\nu} \mathbf{1}_{B \times N_{C,0}}(x, \mathcal{N}_\nu) \right] \\
&= \mathbf{E} \left[\sum_{x \in \mathcal{N}_\nu} F(x, \mathcal{N}_\nu) \right].
\end{aligned}$$

The multivariate Campbell-Mecke formula asserts that for all natural number $k \geq 1$, for all positive $F : E^k \times \mathfrak{N}_E$,

$$\begin{aligned}
\mathbf{E} \left[\sum_{x_1, \dots, x_k \in \mathcal{N}_\nu^\neq} F(x_1, \dots, x_k, \mathcal{N}_\nu) \right] \\
= \int_{E^k} \mathbf{E} \left[F(x_1, \dots, x_k, \mathcal{N}_\nu + \sum_{j=1}^k \varepsilon_{x_j}) \right] \otimes_{j=1}^k d\nu(x_j),
\end{aligned}$$

where the left hand side sum is extended to all k -uples of distinct points in configuration $\mathcal{N}_\nu$.

The proof is achieved by setting

$$\sum_{x_1, \dots, x_k \in \mathcal{N}_\nu^\neq} F(x_1, \dots, x_k, \mathcal{N}_\nu) = \sum_{x \in \mathcal{N}_\nu} g(x, \mathcal{N}_\nu),$$

with

$$g(x, \mathcal{N}_\nu) = \sum_{\substack{(x_2, \dots, x_k) \in \mathcal{N}_\nu^\neq \\ x_2, \dots, x_k \neq x}} F(x, x_2, \dots, x_k, \mathcal{N}_\nu),$$

and by induction through the Fubini Theorem.

$F : E \rightarrow \mathbf{R}$ is said to belong to $\text{dom } D$ if

$$\mathbf{E} \left[\int_E \left(F(\mathcal{N}_\nu + \varepsilon_x) - F(\mathcal{N}_\nu) \right)^2 d\nu(x) \right] < \infty,$$

and we set, for all $x \in E$,

$$D_x F(\mathcal{N}_\nu) = F(\mathcal{N}_\nu + \varepsilon_x) - F(\mathcal{N}_\nu - \varepsilon_x),$$

where $\mathcal{N}_\nu + \varepsilon_x$ must be identified to $\mathcal{N}_\nu$ each time $x \in \mathcal{N}_\nu$, and $\mathcal{N}_\nu - \varepsilon_x$ must be identified to $\mathcal{N}_\nu$ each time $x \notin \mathcal{N}_\nu$. Let

$$\text{dom } \delta_\nu = \left\{ u : \mathfrak{N}_E \times E \rightarrow \mathbf{R}, \mathbf{E} \left[\int_E |u(\mathcal{N}_\nu, x)|^2 d\nu(x) \right] < \infty \right\}.$$

We set, for $u \in \text{dom } \delta_\nu$,

$$\delta_\nu u = \int_E u(\mathcal{N}_\nu - \varepsilon_x, x) d\mathcal{N}_\nu(x) - \int_E u(\mathcal{N}_\nu, x) d\nu(x).$$

Then, for $F \in \text{dom } D$ and $u \in \text{dom } \delta_\nu$, the integration by parts formula asserts that

$$\mathbf{E} [F(\mathcal{N}_\nu) \delta_\nu u] = \mathbf{E} \left[\int_E D_x F(\mathcal{N}_\nu) u(\mathcal{N}_\nu, x) d\nu(x) \right], \quad (4.1)$$

Indeed by definition of gradient D

$$\begin{aligned} & \mathbf{E} \left[\int_E D_{s,z} f(N_\nu) u(s, z) d\nu(s, z) \right] \\ &= \mathbf{E} \left[\int_E f(N_\nu + \varepsilon_{s,z}) u(s, z) d\nu(s, z) \right] - \mathbf{E} \left[\int_E f(N_\nu) u(s, z) d\nu(s, z) \right]. \end{aligned} \quad (4.2)$$

But the Campbell-Mecke formula for Poisson process applied to function $F(s, z, \mathcal{N}) = f(\mathcal{N}_\nu).u(s, z)$ enables to write

$$\begin{aligned} \mathbf{E} \left[\int_E f(N_\nu + \varepsilon_{s,z}) u(s, z) d\nu(s, z) \right] &= \mathbf{E} \left[f(\mathcal{N}_\nu) \sum_n u(T_n, Z_n) \right] \\ &= \mathbf{E} \left[f(\mathcal{N}_\nu) \int_E u d\mathcal{N}_\nu \right]. \end{aligned}$$

By making use of this equality in the right-hand side of (4.2)

$$\begin{aligned} & \mathbf{E} \left[\int_E D_{s,z} f(N_\nu) u(s, z) d\nu(s, z) \right] \\ &= \mathbf{E} \left[f(\mathcal{N}_\nu) \int_E u d\mathcal{N}_\nu \right] - \mathbf{E} \left[\int_E f(N_\nu) u(s, z) d\nu(s, z) \right]. \end{aligned}$$

Recall that since ν is diffuse, for all $(s, z) \in E$, $\mathbf{P}(N_\nu(\{(s, z)\}) \geq 1) = 0$ and

$$u(N_\nu, (s, z)) = u(N_\nu - \varepsilon_{s,z}, (s, z)) \quad \mathbf{P} \otimes \nu\text{-a.s.}$$

Which enables to obtain (4.1) (see [97] for a proof by another method). Moreover,

$$\mathbf{E} [\delta_\nu u] = 0 \text{ et } \mathbf{E} [(\delta_\nu u)^2] = \int_E |u(x)|^2 d\nu(x).$$

In addition, if u is deterministic,

$$\delta_\nu u = \int_E u(x) (d\mathcal{N}_\nu(x) - d\nu(x)) \text{ and } D_x \delta_\nu u = u(x) \quad (4.3)$$

The integration by parts formula (4.1) enables to compute the first term on the left of (3.14) :

$$\mathbf{E} [\langle \delta_\nu u, \nabla F_h(\mathcal{N}_\nu) \rangle] = \mathbf{E} \left[\int_E \nabla^{(2)} F_h(\mathcal{N}_\nu) u(\mathcal{N}_\nu, x) d\nu(x) \right],$$

and to prove the convergence towards a standard Brownian motion.

2 Stein Operator and linear interpolation

We have all the necessary components and need now to reduce to a finite dimensional setting to make the calculations. Hence the idea to cut time in small periods of time, to approximate the process of interest by linear interpolation over each small period, and to compare such interpolated process during each period with a process taking the form of the increment of a Brownian motion, possibly multiplied by a coefficient, which follows the same law as the increment of a time changed Brownian motion.

For fixed $T > 0$ and $n \in \mathbf{N}^*$, we consider the discretization of $[0, T]$ of constant mesh T/n . For a function $F \in \mathbb{D}$, the space of functions being right continuous with left limits, denote by $\pi_n F$ its affine interpolation on the latter grid, that is, for all $t \in [0, T]$, for $1 \leq k \leq n$ such that $t \in \left[\frac{(k-1)T}{n}, \frac{kT}{n}\right]$,

$$\pi_n F(t) = \frac{n}{T} \left(F\left(\frac{kT}{n}\right) - F\left(\frac{(k-1)T}{n}\right) \right) \left(t - \frac{(k-1)T}{n} \right) + F\left(\frac{(k-1)T}{n}\right).$$

As will be made clear later on, considering a mesh of $T/\varphi(n)$ instead of T/n will enable to group increments of $\pi_{\varphi(n)} F$ into packets of size $\varphi(n)/n$ and to make use of a local central limit theorem inside each packet. An immediate computation shows that for all $t \leq T$ and for k as above, we have that

$$\begin{aligned} \pi_{\varphi(n)} F(t) &= \frac{\varphi(n)}{T} \left(\left(F\left(\frac{kT}{\varphi(n)}\right) - F\left(\frac{(k-1)T}{\varphi(n)}\right) \right) \left(t - \frac{(k-1)T}{\varphi(n)} \right) \right. \\ &\quad \left. + \sum_{i=1}^{k-1} \left(F\left(\frac{iT}{\varphi(n)}\right) - F\left(\frac{(i-1)T}{\varphi(n)}\right) \right) \frac{T}{\varphi(n)} \right) + F(0) \end{aligned}$$

So that omitting the term $F(0)$,

$$\pi_{\varphi(n)} F(t) = \sum_{i=0}^{\varphi(n)-1} \frac{\varphi(n)}{T} \left(F\left(\frac{(i+1)T}{\varphi(n)}\right) - F\left(\frac{iT}{\varphi(n)}\right) \right) I_i^n(t). \quad (4.4)$$

where

$$I_i^n(t) = I_{0+}^1 \left(\mathbf{1}_{\left[\frac{iT}{\varphi(n)}, \frac{(i+1)T}{\varphi(n)}\right)} \right) : t \mapsto \int_0^t \mathbf{1}_{\left[\frac{(i-1)T}{\varphi(n)}, \frac{iT}{\varphi(n)}\right)}(s) ds, \quad i = 1, \dots, n.$$

Notice that a stochastic process and its linear interpolation exhibit similar dynamics since they coincide on each point of the subdivision whose step size decreases to 0. Denoting the initial process L_n , our plan consists in computing

1. the distance $d(L_n, \pi_n L_n)$ between the distribution of L_n and that of its linear interpolation $\pi_n L_n$,
2. the distance between the distribution of the linear interpolation of the initial stochastic process $\pi_n L_n$ and its coarse linear interpolation $\pi_{\varphi(n)} L_n$ where $\varphi(n) < n$
3. the distance $d(\pi_{\varphi(n)} L_n, \pi_{\varphi(n)} B)$ between $\pi_{\varphi(n)} L_n$ and $\pi_{\varphi(n)} B$ where B is a standard Brownian motion, and then

4. the distance $d(\pi_{\varphi(n)}B, B)$ between $\pi_{\varphi(n)}B$ and B .

Taking into account the preceding discussion, these distances are computed in the topology of fractional Sobolev space $W_{\eta,p}$. At the end, by triangular inequality we shall get the distance between the distribution of L_n and the Gaussian measure. For all $n \in \mathbf{N}$,

$$\begin{aligned} & d(L_n, B) \\ & \leq d(L_n, \pi_n L_n) + d(\pi_n L_n, \pi_{\varphi(n)} L_n) + d(\pi_{\varphi(n)} L_n, \pi_{\varphi(n)} B) + d(\pi_{\varphi(n)} B, B), \end{aligned} \quad (4.5)$$

In the course of this study we may hope to get an intuition on the various factors which explain such distance: that between L_n and $\pi_{\varphi(n)} L_n$ and that between $\pi_{\varphi(n)} B$ and B are due to the difference in regularity of the trajectories, whereas the distance between the laws of $\pi_{\varphi(n)} L_n$ and $\pi_{\varphi(n)} B$ is owed to their distinct stochastic behaviour.

Leaving aside for the time being distances 1., 2. and 4., we focus here on bounding $d(\pi_{\varphi(n)} L_n, \pi_{\varphi(n)} B)$ as this is where the interest in the Stein method lies. Let us first fix some notations.

Let φ be a function of $\mathbf{N}$ into $\mathbf{N}$ such that

$$\lim_{n \rightarrow \infty} \varphi(n) = \infty \text{ and } \varphi(n) \leq n \text{ for any } n \in \mathbf{N}. \quad (4.6)$$

for all $n \geq 0$, let $\mathcal{D}_n^\varphi = \{i/\varphi(n), i = 0, \dots, \varphi(n) - 1\}$ the subdivision of $[0, T]$ of step size $T/\varphi(n)$, and $\mathcal{D}^n = \{i/n, i = 0, \dots, n - 1\}$ be the natural subdivision. For all $m \in \mathbf{N}$ and $s \in [0, T]$, we define the function

$$h_i^m : s \mapsto \sqrt{\frac{m}{T}} I^1 \left(\mathbf{1}_{\left[\frac{i}{m}T, \frac{i+1}{m}T\right)} \right) (s) = \sqrt{\frac{m}{T}} \left(\left(s - \frac{i}{m}T\right)^+ \wedge \frac{T}{m} \right).$$

Simple calculations show that for all n , and all $a \in \mathcal{D}^n$, $b \in \mathcal{D}^n$, $\|h_a^n\|_{I_{1,2}} = \|h_b^{\varphi(n)}\|_{I_{1,2}} = 1$ and

$$\langle h_a^n, h_b^{\varphi(n)} \rangle_{I_{1,2}} \begin{cases} \leq \sqrt{\frac{\varphi(n)}{n}} & \text{if } a \in \left[b, b + \frac{1}{\varphi(n)}\right); \\ = 0 & \text{otherwise.} \end{cases} \quad (4.7)$$

Remark 3. In the remainder of this document, as we are studying speed of convergence, we shall often use the notation $Y_n \leq cf(n)$ as a shortcut for the sentence "there exists $c > 0$ such that for n sufficiently large, $Y_n \leq cf(n)$." c should not be understood as denoting the same quantity each time it is used and the statement should not be deemed true for all $n \in \mathbf{N}$, but only if it is sufficiently large.

As a direct consequence of corollary 3.2,

$$\|h_j^{\varphi(n)}\|_{W_{\eta,p}} \leq c \varphi(n)^{\eta - \frac{1}{2}}.$$

Thus, keeping in mind the above remark,

$$\|\pi_{\varphi(n)}(h_i^n)\|_{W_{\eta,p}} \leq c \sqrt{\frac{\varphi(n)}{n}} \varphi(n)^{\eta - \frac{1}{2}}. \quad (4.8)$$

Set

$$\mathcal{V}_n = \text{vect}\{h_a^n, a \in \mathcal{D}^n\},$$

equipped with the trace topology of $W_{\eta,p}$ for a certain $(\eta, p) \in (0, 1] \times [1, \infty)$ such that $0 < \eta - 1/p < 1/2$. If B is a standard Brownian motion on $[0, T]$,

$$\pi_n B = \text{Proj}_{\mathcal{V}_n}(B) \stackrel{\text{law}}{=} \sum_{a \in \mathcal{D}^n} Y_a h_a^n,$$

where $(Y_a, a \in \mathcal{D}^n)$ is a sequence of Gaussian independent random variables. Let $\mu_{\mathcal{V}_n}$ be the distribution of $\pi_n B$ on $\mathcal{V}_n$. It is characterized by the identity

$$\left(\int_{\mathcal{V}_n} L^n F(x) \, d\mu(x) = 0, \forall F \in \mathcal{F} \right) \iff \mu = \mu_{\mathcal{V}_n},$$

where $\mathcal{F}$ is the set of twice Fréchet differentiable functions from $\mathcal{V}_n$ into $\mathbf{R}$ and for all $x \in I_{1,2}$, as a direct application of equation (3.11),

$$\begin{aligned} L^n F(x) = & - \sum_{a \in \mathcal{D}^n} \langle h_a^n, x \rangle_{I_{1,2}} \langle h_a^n, \nabla F(x) \rangle_{I_{1,2}} \\ & + \sum_{a \in \mathcal{D}^n} \langle h_a^n \otimes h_a^n, \nabla^{(2)} F(x) \rangle_{I_{1,2} \otimes I_{1,2}}. \end{aligned} \quad (4.9)$$

For all $t > 0$, let P_t^n be the Ornstein-Uhlenbeck operator defined for all $x \in I_{1,2}$ by

$$P_t^n F(x) = \int_{\mathcal{V}_n} F(e^{-t}x + \beta_t y) \, d\mu_{\mathcal{V}_n}(y),$$

where $\beta_t = \sqrt{1 - e^{-2t}}$. Using the Stein-Dirichlet representation formula (see [19, 30]), we know that

$$\int_{\mathcal{V}_n} F \, d\mu_{\mathcal{V}_n} - F(x) = \int_0^T L^n P_t^n F(x) \, dt, \quad (4.10)$$

for all continuous $F : W_{\eta,p} \rightarrow \mathbf{R}$ such that $P_t^n F$ belongs to $\mathcal{F}$.

We have the following higher bound:

Theorem 4.1. *There exists $c_{\eta,p} > 0$ such that for all $x \in W_{\eta,p}$, $m < n$, $t > 0$ and $a \in \mathcal{D}^n$,*

$$\begin{aligned} & \left| \left\langle h_a^n \otimes h_a^n, \nabla^{(2)} P_t^n (F \circ \pi_{\varphi(n)})(x + r h_a^{\varphi(n)}) - \nabla^{(2)} P_t^n (F \circ \pi_{\varphi(n)})(x) \right\rangle_{I_{1,2}} \right| \\ & \leq c_{\eta,p} \frac{e^{-5t/2}}{\beta_{t/2}^2} \frac{\varphi(n)^{1+\eta}}{n^{3/2}} |r|. \end{aligned}$$

Remember that in [109] (Theorems 4.8 and 4.10) Shih proved that Theorem 3.7 remained true for Lipschitz functions $F \in \text{Lip}_1(E)$. Therefore for $F \in \text{Lip}_1(W_{\eta,p})$, for any $t > 0$, $P_t F$ is twice Fréchet differentiable, and for any $x \in W_{\eta,p}$, for any $h, k \in I_{1,2}$ and n , according to equation (3.13),

$$\begin{aligned} & \langle \nabla^{(2)} P_t F(x), h \otimes k \rangle \\ & = \frac{e^{-t}}{\sqrt{1 - e^{-t}}} \int_B \langle y, h \rangle \langle \nabla P_{t/2} F(e^{-t/2}x + \sqrt{1 - e^{-t}}y), k \rangle \, d\mu_B(y) \end{aligned}$$

Now using the representation for $\langle \nabla P_{t/2} F(\cdot), k \rangle$ given by equation (3.12) we end up with

$$\begin{aligned} & \langle h \otimes k, \nabla^{(2)} P_t F \circ \pi_{\varphi(n)}(x) \rangle_{I_{1,2}^{\otimes(2)}} = \\ & \frac{e^{-\frac{3t}{2}}}{\beta_{\frac{t}{2}}^2} \mathbf{E} \left[\nabla P_{\frac{t}{2}} F \circ \pi_{\varphi(n)} \left(e^{-t} x + e^{-\frac{t}{2}} \beta_{\frac{t}{2}} B + \beta_{\frac{t}{2}} \tilde{B} \right) \langle h, B \rangle_{I_{1,2}} \langle k, \tilde{B} \rangle_{I_{1,2}} \right], \quad (4.11) \end{aligned}$$

where $\tilde{B}$ is an independent copy of B .

We are now in position to prove Theorem 4.1:

Proof of Theorem 4.1. For further use we set

$$w_t(x, \zeta, \xi) = e^{-t} x + e^{-t/2} \beta_{t/2} \zeta + \beta_{t/2} \xi.$$

First remark that h_i^n belongs to $\mathcal{V}_n$ hence according to the very definition of the divergence as an extension of the scalar product on $I_{1,2}$, we have

$$\delta_B h_i^n(B) = \delta_B h_i^n(\pi_n B).$$

Furthermore,

$$\begin{aligned} & F \circ \pi_{\varphi(n)} \left(e^{-t} x + e^{-t/2} \beta_{t/2} B + \beta_{t/2} \tilde{B} \right) \\ &= F \left(w_t(\pi_{\varphi(n)}(x), \pi_{\varphi(n)}(B), \pi_{\varphi(n)}(\tilde{B})) \right). \end{aligned}$$

This means that we have

$$\begin{aligned} & \left(\frac{e^{-3t/2}}{\beta_{t/2}^2} \right)^{-1} \left\langle \nabla^{(2)} P_t F \circ \pi_{\varphi(n)}(x), h_i^n \otimes h_i^n \right\rangle_{I_{1,2}^{\otimes(2)}} \\ &= \mathbf{E} \left[F \left(w_t(\pi_{\varphi(n)}(x), \pi_{\varphi(n)}(B), \pi_{\varphi(n)}(\tilde{B})) \right) \right. \\ & \quad \left. \times \mathbf{E} [\delta_B h_i^n(\pi_n B) \mid \pi_{\varphi(n)}(B)] \mathbf{E} [\delta_B h_i^n(\pi_n(\tilde{B})) \mid \pi_{\varphi(n)}(\tilde{B})] \right]. \quad (4.12) \end{aligned}$$

According to lemma 4.7 in [21]

$$\text{var} \left(\mathbf{E} [\delta_B h_i^n(\pi_n B) \mid \pi_{\varphi(n)}(B)] \right) \leq \frac{\varphi(n)}{n}. \quad (4.13)$$

Now remark that since F is Lipschitz continuous from $W_{\eta,p}$ into $\mathbf{R}$, we have

$$\begin{aligned} & \left| F(w_t(\pi_{\varphi(n)}(x + r h_i^n), \pi_{\varphi(n)}(B), \pi_{\varphi(n)}(\tilde{B}))) - F(w_t(\pi_{\varphi(n)}(x), \pi_{\varphi(n)}(B), \pi_{\varphi(n)}(\tilde{B}))) \right| \\ & \leq e^{-t} \|F\|_{\text{Lip}_1} \|\pi_{\varphi(n)} h_i^n\|_{W_{\eta,p}} |r| \\ & \leq e^{-t} \|F\|_{\text{Lip}_1} |r| \sqrt{\frac{\varphi(n)}{n}} \varphi(n)^{\eta-1/2}. \quad (4.14) \end{aligned}$$

Consequently, plug (4.14) into (4.12) and use Cauchy-Schwarz inequality in conjunction with (4.13) to get

$$\begin{aligned} & \left| \left\langle \nabla^{(2)} P_t f_{\varphi}(x + r h_i^n) - \nabla^{(2)} P_t f_{\varphi}(x), h_i^n \otimes h_i^n \right\rangle_{I_{1,2}^{\otimes(2)}} \right| \\ & \leq c \frac{e^{-5t/2}}{\beta_{t/2}^2} |r| \frac{\varphi(n)^{1+\eta}}{n^{3/2}}. \end{aligned}$$

The proof is thus complete. $\square$

In order to control the variations of $L^n P_t^n F$ for small values of t , we need the following result:

Lemma 4.2. *There exists $c > 0$ such that for any $F \in \text{Lip}_1(W_{\eta,p})$, $t > 0$, $n \in \mathbf{N}$ and any $x \in W_{\eta,p}$,*

$$|P_t^n(F \circ \pi_{\varphi(n)})(x) - (F \circ \pi_{\varphi(n)})(x)| \leq c\beta_t(1 + \|x\|_{W_{\eta,p}}).$$

Proof. According to the Mehler formula,

$$\begin{aligned} |P_t^n(F \circ \pi_{\varphi(n)})(x) - (F \circ \pi_{\varphi(n)})(x)| &= |\mathbf{E}[F(e^{-t}\pi_{\varphi(n)}(x) + \beta_t\pi_{\varphi(n)}B) - F(\pi_{\varphi(n)}(x))]| \\ &\leq (1 - e^{-t})\|\pi_{\varphi(n)}(x)\|_{W_{\eta,p}} + \beta_t \mathbf{E}[\|\pi_{\varphi(n)}B\|_{W_{\eta,p}}]. \end{aligned}$$

Since $\pi_{\varphi(n)}B$ converges to B in $W_{\eta,p}$ in L^p hence in L^1 , the sequence of norms is bounded, hence the result. $\square$

Assumption 1 (Assumption Orth_p). *A family $(u_a, a \in \mathcal{D}^n)$ of elements of $L^p([0, T] \times \mathbb{X}, \text{ds} \otimes \text{d}\nu)$ satisfies Assumption Orth_p whenever for any $a \neq b \in \mathcal{D}^n$,*

$$u_a(s, z)u_b(s, z) = 0, \text{ds} \otimes \text{d}\nu(z) \text{ a.s.}$$

and

$$\int u_a(s, z)u_a(s, z) \text{ds} \otimes \text{d}\nu(z) = 1.$$

It is well known (see [97]) that under hypothesis Orth_p , the random variables $(\delta_\nu u_a, a \in \mathcal{D}^n)$ are independent hence we have the following result which is proved in [20].

Theorem 4.3. *Assume that $(u_a, a \in \mathcal{D}^n)$ satisfies Orth_p for some $p \geq 2$. Then for any $\eta < 1/2$,*

$$\sup_{n \geq 1} \mathbf{E} \left[\left\| \sum \delta u_a h_a^n \right\|_{W_{\eta,p}}^p \right] < \infty.$$

We need the following result:

Lemma 4.4. *Let $(u_a, a \in \mathcal{D}^n)$ be a family of functions satisfying Assumption Orth_p . Then, denoting*

$$\pi_n X = \sum_{a \in \mathcal{D}^n} \delta_\nu u_a h_a^n,$$

there exists $c > 0$ such that for all n ,

$$\begin{aligned} \mathbf{E}[L^n P_t^n F_n(\pi_n X)] &\leq c \sum_{a \in \mathcal{D}^n} \int_0^1 \left| \left\langle h_a^n \otimes h_a^n, \nabla^{(2)} P_t^n(F \circ \pi_{\varphi(n)})(X + rh_a^{\varphi(n)}) \right. \right. \\ &\quad \left. \left. - \nabla^{(2)} P_t^n(F \circ \pi_{\varphi(n)})(X) \right\rangle_{I_{1,2}} \right| dr \quad (4.15) \end{aligned}$$

Proof. Fix $n \in \mathbf{N}$. In what follows we denote $F_n = F \circ \pi_{\varphi(n)}$. According to (4.9), we have for all t that

$$\begin{aligned} \mathbf{E} [L^n P_t^n F_n(\pi_n X)] &= - \sum_{a \in \mathcal{D}^n} \mathbf{E} [\delta_\nu u_a \langle h_a^n, \nabla P_t^n F_n(\pi_n X) \rangle_{I_{1,2}}] \\ &\quad + \sum_{a \in \mathcal{D}^n} \mathbf{E} [\langle h_a^n \otimes h_a^n, \nabla^{(2)} P_t^n F_n(\pi_n X) \rangle_{I_{1,2} \otimes I_{1,2}}]. \end{aligned} \quad (4.16)$$

The first term on the right-hand side of (4.16) can be calculated with the integration by parts formula for Poisson measures (4.1) which entails

$$\begin{aligned} &\mathbf{E} [\delta_\nu u_a \langle h_a^n, \nabla P_t^n F_n(\pi_n X) \rangle_{I_{1,2}}] \\ &= \mathbf{E} \left[\int_E u_a(s, x) \langle h_a^n, D_{s,x} \nabla P_t^n F_n(\pi_n X) \rangle_{I_{1,2}} \, ds \, d\nu(x) \right] \end{aligned} \quad (4.17)$$

Remembering that if u is deterministic, $D_{s,x} \delta_\nu u = u(s, x)$ (see equation (4.3)) so that

$$D_{s,x} \pi_n X = \sum_{a \in \mathcal{D}^n} u_a(s, x) h_a^n$$

and the right-hand side of equation (4.17) is equal to

$$\begin{aligned} &\mathbf{E} \left[\sum_{a \in \mathcal{D}^n} \int_E u_a(s, x) \right. \\ &\quad \left. \langle h_a^n, \nabla P_t^n F_n(\pi_n X + u_a(s, x) h_a^n) - \nabla P_t^n F_n(\pi_n X) \rangle_{I_{1,2}} \, ds \, d\nu(x) \right] \\ &= \int_0^1 \mathbf{E} \left[\sum_{a \in \mathcal{D}^n} \int_E u_a(s, x)^2 \langle h_a^n \otimes h_a^n, \right. \\ &\quad \left. \nabla^{(2)} P_t^n F_n(\pi_n X + r h_a^n) \rangle_{I_{1,2}^{\otimes 2}} \, ds \, d\nu(x) \right] \, dr, \end{aligned} \quad (4.18)$$

according to the fundamental theorem of calculus. Now the second term on the right-hand side of (4.16), because of Assumption Orth_p is equal to

$$\sum_{a \in \mathcal{D}^n} \int_0^1 \mathbf{E} \left[\int_E u_a(s, x)^2 \langle h_a^n \otimes h_a^n, \nabla^{(2)} P_t^n F_n(\pi_n X) \rangle_{I_{1,2}^{\otimes 2}} \, ds \, d\nu(x) \right] \, dr$$

Hence, gathering the two terms of (4.16) we have for all t ,

$$\begin{aligned}
& \mathbf{E} [L^n P_t^n F_n(\pi_n X)] \\
&= \sum_{a \in \mathcal{D}^n} \int_0^1 \mathbf{E} \left[\int_E u_a(s, x)^2 \left\langle h_a^n \otimes h_a^n, \nabla^{(2)} P_t^n F_n(\pi_n X) \right. \right. \\
&\quad \left. \left. - \nabla^{(2)} P_t^n F_n(\pi_n X + r h_a^n) \right\rangle_{I_{1,2}^{\otimes 2}} ds d\nu(x) \right] dr, \\
&= \sum_{a \in \mathcal{D}^n} \int_0^1 \left| \left\langle h_a^n \otimes h_a^n, \nabla^{(2)} P_t^n (F \circ \pi_{\varphi(n)})(X + r h_a^{\varphi(n)}) \right. \right. \\
&\quad \left. \left. - \nabla^{(2)} P_t^n (F \circ \pi_{\varphi(n)})(X) \right\rangle_{I_{1,2}} \right| dr
\end{aligned}$$

which yields the desired result $\square$

We can then state:

Theorem 4.5. *Let $(u_a, a \in \mathcal{D}^n)$ be a family of functions satisfying Assumption Orth_p . Then, there exists $c > 0$ such that for all n ,*

$$dist_{W_{\eta,p}} \left(\pi_{\varphi(n)}(\pi_n X), \pi_{\varphi(n)}(\pi_n B) \right) \leq c \frac{\varphi(n)^{1+\eta}}{n^{1/2}} \log \left(\frac{\varphi(n)^{1+\eta}}{n^{1/2}} \right).$$

Proof. By the very definition of the Kantorovitch-Rubinstein distance, we have

$$\begin{aligned}
& dist_{W_{\eta,p}} \left(\pi_{\varphi(n)}(\pi_n X), \pi_{\varphi(n)}(\pi_n B) \right) \\
&= \sup_{F \in \text{Lip}_1(W_{\eta,p})} (\mathbf{E} [F_n(\pi_n X)] - \mathbf{E} [F_n(\pi_n B)]),
\end{aligned}$$

which, applying 4.10 to F_n is equal to

$$\sup_{F \in \text{Lip}_1(W_{\eta,p})} \mathbf{E} \left[\int_0^T L^n P_t^n F_n(\pi_n X) dt \right];$$

Or in other words, using Dynkin's lemma

$$\begin{aligned}
& \sup_{F \in \text{Lip}_1(W_{\eta,p})} \mathbf{E} [P_\tau^n F_n(\pi_n X) - F_n(\pi_n X)] + \mathbf{E} \left[\int_\tau^T L^n P_t^n F_n(\pi_n X) dt \right] \\
&= \sup_{F \in \text{Lip}_1(W_{\eta,p})} [|A_1(F_n)| + |A_2(F_n)|],
\end{aligned}$$

for all $\tau > 0$. First, it follows from lemma (4.2) that

$$|A_1(F_n)| \leq c\beta_\tau \left(1 + \mathbf{E} \left[\left\| \sum_{a \in \mathcal{D}^n} \delta_\nu u_a h_a^n \right\|_{W_{\eta,p}} \right] \right) \leq c\beta_\tau, \quad (4.19)$$

in view of Theorem 4.3.

And combining Theorem 4.1 together with Lemma 4.4 entails that

$$\mathbf{E} [L^n P_t^n F_n(\pi_n X)] \leq c_{\eta,p} \frac{e^{-5t/2}}{\beta_{t/2}^2} \frac{\varphi(n)^{1+\eta}}{n^{1/2}}.$$

and in turn

$$|A_2(F_m)| \leq c \frac{\varphi(n)^{1+\eta}}{n^{1/2}} |\log(1 - e^{-\tau})|. \quad (4.20)$$

Combine (4.19) and (4.20) to obtain that there exists a universal constant c such that for any $\tau > 0$,

$$\begin{aligned} & \text{dist}_{W_{\eta,p}} \left(\pi_{\varphi(n)}(\pi_n X), \pi_{\varphi(n)}(\pi_n B) \right) \\ & \leq c \left(\sqrt{1 - e^{-2\tau}} - \frac{\varphi(n)^{1+\eta}}{n^{1/2}} \log(1 - e^{-\tau}) \right) \\ & = c \left(\sqrt{1 - e^{-\tau}} \sqrt{1 + e^{-\tau}} - \frac{\varphi(n)^{1+\eta}}{n^{1/2}} \log(1 - e^{-\tau}) \right) \\ & \leq c \left(\sqrt{1 - e^{-\tau}} \sqrt{2} - \frac{\varphi(n)^{1+\eta}}{n^{1/2}} \log(1 - e^{-\tau}) \right) \\ & \leq (c \vee \sqrt{2}) \left(\sqrt{1 - e^{-\tau}} - \frac{\varphi(n)^{1+\eta}}{n^{1/2}} \log(1 - e^{-\tau}) \right). \end{aligned}$$

The optimal value of τ is such that $\sqrt{1 - e^{-\tau}} = \frac{\varphi(n)^{1+\eta}}{n^{1/2}}$, hence there exists a universal constant c such that

$$\text{dist}_{W_{\eta,p}} \left(\pi_{\varphi(n)}(\pi_n X), \pi_{\varphi(n)}(\pi_n B) \right) \leq c \frac{\varphi(n)^{1+\eta}}{n^{1/2}} \log \left(\frac{\varphi(n)^{1+\eta}}{n^{1/2}} \right).$$

The proof is thus complete. $\square$

Chapter 5

Application to a class of Markov Processes

1 A class of Markov processes

We are now equipped with a Stein method to prove the convergence of interpolated Poisson process to interpolated Brownian processes. How can this be used to assess the convergence speed of Markov processes ? The present section is devoted to describing a class of Markov processes denoted X for which we are able to answer two questions:

1. How to rescale process X so that the rescaled process $\bar{X}$ converges to a limit Λ ?
2. How to rescale the difference $\bar{X} - \Lambda$ to make sure that $U_n = n^\kappa \|\bar{X} - \Lambda\|$ converges towards a non degenerate limit at a speed that can be bounded ?

We are going to show that for a wide class of Markov models we shall have

$$\|U_n - \Upsilon\| = o(n^{-\epsilon})$$

where $\Upsilon = \Theta(B \circ \gamma)$, Θ is a linear, on one mapping of $\mathbb{D}$, B a standard brownian motion and γ a time change and where ϵ can be calculated as a function of the scaling parameters of X identified in (1) and (2).

The d -dimensional Markov processes investigated in this paper can be represented as a linear combination of stochastic integrals with respect to a number of independent Poisson measures. Fix a time horizon $T > 0$ throughout. For any $m \in \mathbb{N}^*$, any family $(\zeta_1, \dots, \zeta_m)_{1 \leq k \leq m}$ of elements of $\mathbb{R}^d$ and any array $(\rho_k)_{1 \leq k \leq m}$ of mappings from $\mathbb{R}^+ \times \mathbb{R}^d$ to $\mathbf{R}$, consider the $\mathbf{R}^d$ -valued process X defined as the solution of the SDE

$$X(t) = X(0) + \sum_{k=1}^m \left(\int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \rho_k(s, X(s^-))\}} d\mathcal{N}_k(s, z) \right) \cdot \zeta_k, \quad t \leq T,$$

where $X(0) \in \mathbf{R}^d$ is fixed, and $(\mathcal{N}^k)_{1 \leq k \leq m}$ denote m independent Poisson measures of unit intensity $ds \otimes dz$.

2 The scaling of the processes

Fix $n \in \mathbf{N}^*$, and an array $\alpha := (\alpha_1, \dots, \alpha_m) \in (\mathbf{R}^+)^m$. We scale the process X by replacing for all k , the measure $\mathcal{N}_k$ by a Poisson measure $\mathcal{N}_k^{n^{\alpha_k}}$ of intensity $(n^{\alpha_k} ds) \otimes dz$, and normalizing in space by n so as to define the process $\bar{X}_n$ as the solution of the SDE

$$\bar{X}_n(t) = \frac{X(0)}{n} + \sum_{k=1}^m \left(\frac{1}{n} \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \rho_k(s, \frac{\bar{X}_n(s-)}{n})\}} d\mathcal{N}_k^{n^{\alpha_k}}(s, z) \right) \cdot \zeta_k. \quad (5.1)$$

The key assumption on our scaling (i.e. the choice of the α_k 's) is the following,

Assumption 2 (LLN Assumption). *For any $k = 1, \dots, m$, there exists a mapping $\bar{\rho}_k : [0, T] \times \mathbf{R}^d \rightarrow \mathbf{R}$ such that for all $n \in \mathbf{N}^*$, any $x \in \mathbf{R}^d$ and $t \leq T$,*

$$n^{\alpha_k-1} \rho_k(t, \frac{x}{n}) = \bar{\rho}_k(t, x).$$

We furthermore assume that the mapping $\bar{\rho}_k$ is uniformly Lipschitz continuous in its second variable, i.s. for some $a > 0$, for all $x, y \in \mathbf{R}^d$ and all $t \leq T$,

$$|\bar{\rho}_k(t, x) - \bar{\rho}_k(t, y)| \leq a \|x - y\|. \quad (5.2)$$

Some examples will be provided further on.

Observe that crucially, under the LLN Assumption the $\bar{\rho}_k$'s do not depend on n . We denote for any k , by $\tilde{\mathcal{N}}_k^{n^{\alpha_k}}$, the compensated Poisson measures of $\mathcal{N}_k^{n^{\alpha_k}}$, that is, we let $d\tilde{\mathcal{N}}_k^{n^{\alpha_k}}(s, z) = d\mathcal{N}_k^{n^{\alpha_k}}(s, z) - n^{\alpha_k}(ds \otimes dz)$. Then, for all n , the $\mathbf{R}^d$ -valued processes $(M_{n,k,\bar{X}_n})_{k=1,\dots,m}$, defined for all $t \leq T$ by

$$M_{n,k,\bar{X}_n}(t) = \left(\int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n^{1-\alpha_k} \bar{\rho}_k(s, \bar{X}_n(s-))\}} d\tilde{\mathcal{N}}_k^{n^{\alpha_k}}(s, z) \right) \cdot \zeta_k \quad (5.3)$$

are martingales with respect to the natural filtration of the Poisson measures. It follows from (5.1) that for all n and t , for all t ,

$$\bar{X}_n(t) = \bar{X}_n(0) + \sum_{k=1}^m \left(\int_0^t \bar{\rho}_k(s, \bar{X}_n(s-)) ds \right) \cdot \zeta_k + \sum_{k=1}^m n^{-1} M_{n,k,\bar{X}_n}(t). \quad (5.4)$$

Now, denote by Λ the solution in $\mathbb{C}([0, T])$ of the integral equation

$$\Lambda(t) = \Lambda(0) + \sum_{k=1}^m \left(\int_0^t \bar{\rho}_k(s, \Lambda(s)) ds \right) \cdot \zeta_k, \quad t \geq 0. \quad (5.5)$$

(The existence and uniqueness of such Λ readily follow from the Cauchy-Lipschitz Theorem in view of (5.2).)

Lemma 5.1. *Let Λ be the unique solution of (5.5) on $[0, T]$. Suppose that the LLN Assumption holds, and that*

$$\lim_{n \rightarrow \infty} \mathbf{E} [\|\bar{X}_n(0) - \Lambda(0)\|] = 0. \quad (5.6)$$

Then, $(\bar{X}_n)$ converges in L^1 and uniformly on $[0, T]$ to Λ , that is,

$$\lim_{n \rightarrow \infty} \mathbf{E} [\|\bar{X}_n - \Lambda\|_T] = 0.$$

where $\|\cdot\|_T$ denotes the uniform convergence norm on $[0, T]$.

Proof. Fix $t \leq T$. It follows from (5.4), the triangular inequality and the Lipschitz continuity assumption (5.2) that

$$\begin{aligned} & \sup_{s \leq t} \|\bar{X}_n(s^-) - \Lambda(s)\| \\ & \leq \|\bar{X}_n(0) - \Lambda(0)\| + m \max_{k \leq m} \|\zeta_k\| a \int_0^t \sup_{s \leq u} \|\bar{X}_n(s^-) - \Lambda(s)\| \, du \\ & \quad + \frac{1}{n} \sum_{k=1}^m \|\zeta_k\| \left(\int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n^{1-\alpha_k} \bar{\rho}_k(u, \bar{X}_n(u^-))\}} \, d\tilde{\mathcal{N}}_k^{n^{\alpha_k}}(u, z) \right) \end{aligned}$$

Thus, from the Burkholder Davis Gundy inequality, there exists C such that

$$\begin{aligned} & \mathbf{E} \left[\sup_{s \leq t} \|\bar{X}_n(s^-) - \Lambda(s)\| \right] \\ & \leq \mathbf{E} [\|\bar{X}_n(0) - \Lambda(0)\|] + m \max_{k \leq m} \|\zeta_k\| a \int_0^t \mathbf{E} \left[\sup_{s \leq u} \|\bar{X}_n(s^-) - \Lambda(s)\| \right] \, du \\ & \quad + \frac{C}{n} \sum_{k=1}^m \|\zeta_k\| \mathbf{E} \left[\sqrt{\alpha_k \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n^{1-\alpha_k} \bar{\rho}_k(u, \bar{X}_n(u^-))\}} \, du \, dz} \right], \end{aligned}$$

so in view of the Gromwall lemma, for some positive constants K and C' we have that

$$\begin{aligned} & \mathbf{E} \left[\sup_{s \leq t} \|\bar{X}_n(s^-) - \Lambda(s)\| \right] \\ & \leq \left(\mathbf{E} [\|\bar{X}_n(0) - \Lambda(0)\|] + \frac{C'}{\sqrt{n}} \sum_{k=1}^m \mathbf{E} \left[\sqrt{\int_0^t \bar{\rho}_k(u, \bar{X}_n(u^-)) \, du} \right] \right) e^{Kt}, \end{aligned} \tag{5.7}$$

whence the result. $\square$

This lemma provides a systematic method to rescale a Markov process so that it gets a fluid limit: one simply has to find the sequence of α_k 's (scaling of intensities) that make the $\bar{\rho}_k$ independent of n (see the examples in chapters 6 and 7 below).

We now turn to the so-called diffusion scaling of the process X . We study the sequence of processes $\{U_n\}$ defined for all n by

$$U_n = U_n^\alpha = \sqrt{n}(\bar{X}_n - \Lambda).$$

To obtain a formal functional central limit Theorem for the sequence $\{U_n\}$, we make the following additional assumptions,

Assumption 3 (FCLT). (i) For all n , all $k = 1, \dots, m$ and all $t \in [0, T]$,

$$\bar{\rho}_k(t, \bar{X}_n(t)) - \bar{\rho}_k(t, \Lambda(t)) = \frac{1}{\sqrt{n}} L_k(t, U_n(t)) + E_{n,k}(t), \tag{5.8}$$

where L is linear with respect to its second variable and

$$\sqrt{n} \lim_{n \rightarrow \infty} \mathbf{E} [\|E_{n,k}\|_T] = 0; \tag{5.9}$$

(ii) The initial conditions are such that

$$\lim_{n \rightarrow \infty} \sqrt{n} \mathbf{E} [\|\bar{X}_n(0) - \Lambda(0)\|] = 0. \quad (5.10)$$

Recall (5.3), and let us introduce for all n , the $\mathbf{R}^d$ -valued martingales $(M_{n,k,\bar{X}_n})_{k=1,\dots,m}$, $\bar{X}_{n,\bar{X}_n}$, $(M_{n,k,\Lambda})_{k=1,\dots,m}$ and $\bar{M}_{n,\Lambda}$, defined respectively for all $t \leq T$ by

$$\bar{M}_{n,\bar{X}_n}(t) = \frac{1}{\sqrt{n}} \sum_{k=1}^m M_{n,k,\bar{X}_n}(t); \quad (5.11)$$

$$M_{n,k,\Lambda}(t) = \left(\int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n^{1-\alpha_k} \bar{\rho}_k(s, \Lambda(s))\}} d\tilde{\mathcal{N}}_k^{n^{\alpha_k}}(s, z) \right) \cdot \zeta_k; \quad (5.12)$$

$$\bar{M}_{n,\Lambda}(t) = \frac{1}{\sqrt{n}} \sum_{k=1}^m M_{n,k,\Lambda}(t). \quad (5.13)$$

From (5.4) and (5.5), for all n and all $t \leq T$ we have that

$$\begin{aligned} U_n(t) &= \sqrt{n} (\bar{X}_n(0) - \Lambda(0)) \\ &+ \sum_{k=1}^m \left(\sqrt{n} \int_0^t (\bar{\rho}_k(s, \bar{X}_n) - \bar{\rho}_k(s, \Lambda(s))) ds \right) \cdot \zeta_k + \bar{M}_{n,\bar{X}_n}(t). \end{aligned} \quad (5.14)$$

Under the FCLT Assumption, for any $f \in \mathbb{D}([0, T], \mathbf{R}^d)$, there exists a unique solution in $f \in \mathbb{D}([0, T], \mathbf{R}^d)$ to the equation

$$g(t) = \sum_{k=1}^m \left(\int_0^t L_k(s, g(s)) ds \right) \cdot \zeta_k + f(t), \quad t \leq T. \quad (5.15)$$

So we can define the bijective mapping

$$\Theta : \begin{cases} \mathbb{D}([0, T], \mathbf{R}^d) & \longrightarrow \mathbb{D}([0, T], \mathbf{R}^d), \\ f & \longmapsto g \text{ solution to (5.15)}. \end{cases} \quad (5.16)$$

We have the following result:

Lemma 5.2. *Suppose that the LLN and FCLT Assumptions hold. Suppose also that for some process $W \in \mathbb{D}([0, T], \mathbf{R}^d)$,*

$$\mathbf{E} [\|\bar{M}_{n,\Lambda} - W\|_T] \leq n^{-3/4}, \quad (5.17)$$

where $\bar{M}_{n,\Lambda}$ is defined by (5.13). Then,

$$\mathbf{E} [\|U_n - \Theta(W)\|_T] \leq n^{-3/4},$$

for Θ defined in (5.16).

Proof. For all n , let $W_n = \Theta^{-1}(U_n)$. Then it follows from (5.14) and (5.8) that for all n and all $t \leq T$,

$$\begin{aligned} U_n(t) &= \sum_{k=1}^m \left(\int_0^t L_k(s, U_n(s)) ds \right) \cdot \zeta_k + \sqrt{n} \left(\bar{X}_n(0) - \Lambda(0) + \sum_{k=1}^m E_{n,k}(t) \cdot \zeta_k \right) \\ &+ \frac{1}{\sqrt{n}} \sum_{k=1}^m M_{n,k,\bar{X}_n}(t). \end{aligned}$$

and by definition of Θ

$$W_n(t) = \sqrt{n} \left(\bar{X}_n(0) - \Lambda(0) + \sum_{k=1}^m E_{n,k}(t) \cdot \zeta_k \right) + \frac{1}{\sqrt{n}} \sum_{k=1}^m M_{n,k,\bar{X}_n}(t). \quad (5.18)$$

But from the Burkholder Davis Gundy inequality, for any k there exists a constant C_k such that

$$\begin{aligned} & \mathbf{E} \left[\left\| M_{n,k,\bar{X}_n} - M_{n,k,\Lambda} \right\|_T \right] \\ & \leq C_k \|\zeta_k\| \mathbf{E} \left[\left(\alpha_k \int_0^T \int_{\mathbb{R}} \left(\mathbf{1}_{\{z \leq n^{1-\alpha_k} \bar{\rho}_k(t, \bar{X}_n(t^-))\}} - \mathbf{1}_{\{z \leq n^{1-\alpha_k} \bar{\rho}_k(t, \Lambda(t))\}} \right)^2 dt dz \right)^{1/2} \right] \\ & = C_k \sqrt{n^{\alpha_k}} \|\zeta_k\| \mathbf{E} \left[\left(\int_0^T |\bar{\rho}_k(t, \bar{X}_n(t^-)) - \bar{\rho}_k(t, \Lambda(t))| dt \right)^{1/2} \right]. \end{aligned}$$

Hence, in view of (5.2) there exists a constant D_k such that this last term is less than

$$\begin{aligned} & D_k \sqrt{n} \mathbf{E} \left[\left(\int_0^T \|\bar{X}_n(t^-) - \Lambda(t)\| dt \right)^{1/2} \right] \\ & \leq D \sqrt{n} \left(\int_0^t \mathbf{E} [\|\bar{X}_n(t^-) - \Lambda(t)\|] dt \right)^{1/2}, \end{aligned}$$

in view of Jensen's inequality and Fubini's Theorem. In view of (5.7), the above term is thus at most of order $n^{-3/4}$, which, using assumptions (5.9) and (5.10), entails that

$$\mathbf{E} [\|W_n - \bar{M}_{n,\Lambda}\|_T] \leq n^{-3/4}. \quad (5.19)$$

To conclude, observe that, as it is linear, the mapping Θ is Lipschitz continuous. Thus there exists a constant a such that $\|U_n - \Theta(W)\|_T \leq a\|W_n - W\|_T$ for all n . This, together with (5.19) and Assumption (5.17), yields the result. $\square$

In other words, this lemma enables to bound the speed of convergence of U_n towards $\Theta(\lim_{n \rightarrow \infty} \bar{M}_{n,\Lambda})$ by that of W_n towards $\lim_{n \rightarrow \infty} \bar{M}_{n,\Lambda}$.

Remark 4. Thanks to lemma 5.2 we can focus on the study of the convergence of the family of martingales $\bar{M}_{n,\Lambda}$ defined by (5.13) where the indicator function is indeed deterministic and the family of Poisson measures $\tilde{N}^{n^{\alpha_k}}$ being considered are independent of each other. We will therefore be able to use the Dubins Schwarz theorem for each k and to assess the speed of convergence by retaining the slowest rate of convergence among the members of the family of processes indexed by k .

3 Interpolation of these processes

3.1 Distance between the process and its interpolation

We prove the following theorem:

Theorem 5.3. *Let N be a Poisson process of intensity λ and consider for some $\kappa > 0$,*

$$X(t) = \kappa(N(t) - \lambda t).$$

Then, for all $q \in \mathbb{R}^+$

$$\mathbf{E} [\|X - \pi_q X\|_T] \leq 2\kappa \left| \frac{\log(qe^{-\lambda T/q})}{\log\left(\frac{q \log(qe^{-\lambda T/q})}{\lambda T e}\right)} \right|.$$

where $\|\cdot\|_T$ denotes the norm of $L^\infty([0, T])$.

Proof. For any $t \in [0, T]$ there exists $i \leq q-1$ such that $t \in [\frac{iT}{q}, \frac{(i+1)T}{q})$ and

$$\begin{aligned} & \|X(t) - \pi_q X(t)\| \\ &= \left\| X(t) - X\left(\frac{iT}{q}\right) - \frac{q}{T} \left(t - \frac{iT}{q}\right) \left(X\left(\frac{(i+1)T}{q}\right) - X\left(\frac{iT}{q}\right)\right) \right\| \\ &\leq 2 \sup_{t \in [\frac{iT}{q}, \frac{(i+1)T}{q}]} \left\| X(t) - X\left(\frac{iT}{q}\right) \right\|, \end{aligned} \quad (5.20)$$

so that

$$\mathbf{E} [\|X - \pi_q X\|_T] \leq 2 \mathbf{E} \left[\max_{i \in [0, q-1]} \left\| X\left(\frac{(i+1)T}{q}\right) - X\left(\frac{iT}{q}\right) \right\| \right]. \quad (5.21)$$

The increments of X are independent and distributed as κ times a Poisson random variable of parameter $\lambda T/q$. The result follows from Proposition 8.6 in section 2 of the Appendix. $\square$

3.2 Distance between two affine interpolations

Let (λ_n) be a sequence of numbers in $\mathbb{R}^+$ and let us consider a sequence S_n of processes such that

$$S_n(t) = \sqrt{\lambda_n} (N^n(t) - \lambda_n t).$$

Where N^n is a Poisson process of intensity λ_n . According to the definition of the interpolated process,

$$\begin{aligned} \pi_n S_n(t) &= \frac{1}{\sqrt{T}} \sum_{i=0}^{n-1} \sqrt{n} \left(S_n\left(\frac{(i+1)T}{n}\right) - S_n\left(\frac{iT}{n}\right) \right) h_i^n(t) \\ &= \frac{1}{\sqrt{T}} \sum_{i=0}^{n-1} \delta_n(\mathbf{1}_{[iT/n, (i+1)T/n]}) h_i^n(t) \end{aligned} \quad (5.22)$$

Note that the random variables $\delta_n(\mathbf{1}_{[iT/n, (i+1)T/n]})$ are independent and have the distribution of a centered Poisson random variable of parameter T . When divided by $\sqrt{T}$, it has a unit variance.

In what follows, for the sake of simplicity, we assume $T = 1$ without loss of generality. Let $(\varphi(n))_{n \geq 1}$ be a sequence of numbers which tends to infinity and such that $n/\varphi(n)$ is an integer greater than 1. Fix $n \geq 0$, and let

$$\mathcal{D}_n^\varphi = \left\{ i/\varphi(n), i = 0, \dots, \varphi(n) - 1 \right\}$$

be the subdivision of $[0, 1]$ of mesh $1/\varphi(n)$. We have the following theorem

Theorem 5.4. For any integer n ,

$$\mathbf{E} \left[\|\pi_n S_n - \pi_{\varphi(n)} S_n\|_{W_{n,p}}^p \right]^{1/p} \leq c \frac{\varphi(n)^\eta}{\sqrt{\varphi(n)}}.$$

Before proving this theorem we need the following result

Lemma 5.5. For all $p \geq 2$, there exists a constant c_p such that for any sequence of independent, centered, identically distributed, $\mathbb{R}$ -valued random variables $(X_i; i \in \mathbb{N})$ with $X \in L^p$ and any sequence $(\alpha_i; i \in \mathbb{N})$

$$\mathbf{E} \left[\left| \sum_{i=1}^n \alpha_i X_i \right|^p \right] \leq c_p |\{i \leq n, \alpha_i \neq 0\}|^{p/2} \left(\sum_{i=1}^n |\alpha_i|^p \right) \mathbf{E} [|X|^p]$$

where $|A|$ is the cardinality of the set A .

Proof. The Marcinkiewicz–Zygmund inequality yields

$$\mathbf{E} \left[\left| \sum_{i=1}^n \alpha_i X_i \right|^p \right] \leq c_p \mathbf{E} \left[\left| \sum_{i=1}^n \alpha_i^2 X_i^2 \right|^{p/2} \right].$$

Using Jensen's inequality $\psi(\sum p_i x_i / \sum p_i) \leq \sum p_i \psi(x_i) / \sum p_i$, applied to the convex function $\psi : x \rightarrow x^{p/2}$, weights $p_i = \mathbf{1}_{|\alpha_i| > 0}$ and variables $x_i = \alpha_i^2 X_i^2$, we obtain

$$\mathbf{E} \left[\left| \sum_{i=1}^n \alpha_i X_i \right|^p \right] \leq c_p |\{i \leq n, \alpha_i \neq 0\}|^{p/2-1} \mathbf{E} \left[\sum_{i=1}^n |\alpha_i|^p |X_i|^p \right].$$

We then use the fact that

$$\mathbf{E} \left[\sum_{i=1}^n |\alpha_i|^p |X_i|^p \right] \leq |\{i \leq n, \alpha_i \neq 0\}| \left(\sum_{i=1}^n |\alpha_i|^p \right) \mathbf{E} [|X|^p],$$

which completes the proof. $\square$

Proof of Theorem 5.4. Actually, it is already proved in [22] that for all $m \geq 1$,

$$\mathbf{E} [\|S_m(s) - S_m(t)\|] \leq c \|X\|_{L^p} (\sqrt{t-s} \wedge m^{-1/2}), \quad (5.23)$$

where $X = \delta_n(\mathbf{1}_{[iT/n, (i+1)T/n]})$ is a random centered Poisson variable. Assume that s and t belong to the same sub-interval: There exists $l \in \{1, \dots, n\}$ such that

$$\frac{l-1}{n} \leq s < t \leq \frac{l}{n}.$$

Then we have

$$\pi_n S_m(t) - \pi_n S_m(s) = \sqrt{n} \left(\sum_{k=1}^m X_k \langle h_k^m, h_l^n \rangle_H \right) (t-s)$$

. Using lemma 5.5, there exists a constant c such that

$$\frac{\|\pi_n S_m(t) - \pi_n S_m(s)\|_{L^p}}{\sqrt{n}|t-s|} \leq c \|X\|_{L^p} |\{k, \langle h_k^m, h_l^n \rangle_H \neq 0\}|^{1/2} \sup_k |\langle h_k^m, h_l^n \rangle_H|.$$

Note that $|\langle h_k^m, h_l^n \rangle_H| \leq \sqrt{\frac{n}{m}}$ and there are at most $m/n + 2$ terms such that $\langle h_k^m, h_l^n \rangle_H$ is non zero. Thus,

$$\frac{\|\pi_n S_m(t) - \pi_n S_m(s)\|_{L^p}}{\sqrt{n}|t-s|} \leq c\|X\|_{L^p} \left(\frac{m}{n} + 2\right)^{1/2} \sqrt{\frac{n}{m}} \leq c\|X\|_{L^p},$$

as m/n tends to infinity. Since $|t-s| \leq 1/n$,

$$\|S_m(s) - S_m(t)\| \leq c\|X\|_{L^p} \sqrt{|t-s|}. \quad (5.24)$$

For $0 \leq s < t \leq 1$ let $s_+^n := \min\{l, s \leq \frac{l}{n}\}$ and $t_-^n := \sup\{l, t \geq \frac{l}{n}\}$. We have, denoting $f^{s,t} = f(t) - f(s)$,

$$\pi_n(S_m)^{s,t} - S_m^{s,t} = (\pi_n(S_m)^{s,s_+^n} - S_m^{s,s_+^n}) + (\pi_n(S_m)^{s_+^n,t_-^n} - S_m^{s_+^n,t_-^n}) + (\pi_n(S_m)^{t_-^n,t} - S_m^{t_-^n,t})$$

Note that for all $f \in W$, $\pi_n(f)$ is the linear interpolation of f along the subdivision $\mathcal{D}^n$; hence, for $s, t \in \mathcal{D}^n$, $\pi_n(S_m)^{s,t} = S_m^{s,t}$. Thus the median term vanishes and we obtain

$$\begin{aligned} \mathbf{E} [\|\pi_n(S_m)^{s,t} - S_m^{s,t}\|^p] &\leq c \left(\mathbf{E} [\|\pi_n(S_m)^{s,s_+^n}\|^p] \right. \\ &\quad \left. + \mathbf{E} [\|S_m^{s,s_+^n}\|^p] + \mathbf{E} [\|\pi_n(S_m)^{t_-^n,t}\|^p] + \mathbf{E} [\|S_m^{t_-^n,t}\|^p] \right) \end{aligned} \quad (5.25)$$

From (5.24), we deduce that

$$\mathbf{E} [\|\pi_n(S_m)^{s,s_+^n}\|^p]^{1/p} \leq c\|X\|_{L^p} \sqrt{|s_+^n - s|} \leq c\|X\|_{L^p} n^{-1/2}, \quad (5.26)$$

and the same holds for the t and t_-^n , but also for $\mathbf{E} [\|S_m^{s,s_+^n}\|^p]$ and $\mathbf{E} [\|S_m^{t_-^n,t}\|^p]$. We infer from (5.23), (5.24) and (5.26) that

$$\mathbf{E} [\|\pi_n(S_m)^{s,t} - S_m^{s,t}\|^p]^{1/p} \leq c\|X\|_{L^p} \left(\sqrt{t-s} \wedge n^{-1/2} \right). \quad (5.27)$$

Remembering equation 3.1, a straightforward computation shows that

$$\int \int_{[0,1]^2} \frac{[|t-s| \wedge n^{-1}]^{p/2}}{|t-s|^{1+\eta p}} ds dt \leq cn^{-p(1/2-\eta)}. \quad (5.28)$$

The result follows from (5.27) and (5.28). $\square$

4 A central limit theorem

We are going to show in what follows that

$$\lim_{n \rightarrow \infty} \bar{M}_{n,\Lambda} \stackrel{d}{=} \sum_{k=0}^m \zeta_k \int_0^t \bar{\rho}_k(s, \Lambda(s)) dB_k(s).$$

and assess its speed of convergence. In order to do so, notice that

$$\bar{M}_{n,\Lambda} = \frac{1}{\sqrt{n}} \sum_{k=0}^m \zeta_k \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n^{\beta-\alpha_k} \bar{\rho}_k(s, \Lambda(s))\}} d\tilde{\mathcal{N}}_k^{n^{\alpha_k}}(s, z)$$

We know that

$$R_n^k = \frac{1}{\sqrt{n}} \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n^{\beta-\alpha_k} \bar{\rho}_k(s, \Lambda(s))\}} d\tilde{\mathcal{N}}_k^{n^{\alpha_k}}(s, z)$$

converges in distribution, to a Gaussian martingale of square bracket

$$\gamma_k(t) = \int_0^t \bar{\rho}_k(s, \Lambda(s)) ds,$$

which according to the Dubins-Schwarz theorem can be represented as $B \circ \gamma_k$ where B is a one dimensional standard Brownian motion. Our goal now is to precise the rate of this convergence in $W_{\eta,p}$. In order to use the Stein method presented in the previous section, we need to reduce the problem to a convergence to Brownian motion. As a consequence, we must first analyze the rate of convergence of $R_n^k \circ \gamma_k^{-1}$ to B .

Lemma 5.6. *The process $R_n^k \circ \gamma_k^{-1}$ has the distribution of $n^{-1/2}$ times a Poisson process of intensity n .*

Proof. Let $\mathcal{F}_t^k$ the σ -field generated by $\tilde{\mathcal{N}}^{\alpha_k}$:

$$\mathcal{F}_t^k = \sigma\{\tilde{\mathcal{N}}^{\alpha_k}([0, s] \times A), s \leq t, A \in \mathcal{B}(\mathbf{R}^+)\}$$

and $\mathcal{G}_t^k = \mathcal{F}_{\gamma_k^{-1}(t)}^k$. Consider $T_n^k = n^{1/2} R_n^k$. According to the Watanabe characterization of Poisson point processes, it is sufficient to prove that

$$t \mapsto (T_n^k \circ \gamma_k^{-1})(t) - nt$$

is a $\mathcal{G}_t^k$ -martingale. Let $0 \leq s < t$ and $A \in \mathcal{G}_s^k$,

$$\begin{aligned} \mathbf{E} [\mathbf{1}_A (T_n^k \circ \gamma_k^{-1}(t) - T_n^k \circ \gamma_k^{-1}(s))] &= \mathbf{E} \left[\mathbf{1}_A \int_{\gamma_k^{-1}(s)}^{\gamma_k^{-1}(t)} n \bar{\rho}_k(v, \Lambda(v)) dv \right] \\ &= n \mathbf{E} \left[\mathbf{1}_A \left(\gamma(\gamma_k^{-1}(t)) - \gamma(\gamma_k^{-1}(s)) \right) \right] \\ &= n \mathbf{E} [\mathbf{1}_A (t - s)]. \end{aligned}$$

The proof is thus complete. $\square$

In the sequel we denote $S_n^k = R_n^k \circ \gamma_k^{-1} \stackrel{d}{=} n^{-1/2}(\mathcal{N}^n(t) - nt)$ where $\mathcal{N}^n$ is a Poisson process on the half-line of intensity n . We show it converges to a Brownian motion and provide a bound on the convergence rate.

Theorem 5.7. *Using the notations defined above, for $\eta < 1/2$, we have*

$$d(S_n^k, B) \leq cn^{\eta/3-1/6} \log(n).$$

Proof. We make use of an interpolation of process S_n^k with step size $\varphi(n)$ and use triangular inequality (4.5) to show that for all k ,

$$d(\pi_n S_n^k, B) \leq d(S_n^k, \pi_n S_n^k) + d(\pi_n S_n^k, \pi_{\varphi(n)} S_n^k) + d(\pi_{\varphi(n)} S_n^k, \pi_{\varphi(n)} B) + d(\pi_{\varphi(n)} B, B),$$

Bound on $d(\pi_{\varphi(n)}S_n^k, \pi_{\varphi(n)}B)$: Since $\pi_n S_n$ belongs to $I_{1,2}$, its affine interpolation on $\mathcal{D}_n^\varphi$ is nothing but the orthogonal projection on $\text{span}\{h_j^{\varphi(n)}, j \in \mathcal{D}_n^\varphi\}$. This means that

$$\begin{aligned}\pi_{\varphi(n)}S_n &= \pi_{\varphi(n)}(\Pi_n S_n) \\ &= \sum_{j \in \mathcal{D}_n^\varphi} \sum_{i=0}^{n-1} \delta_n(\mathbf{1}_{[\frac{i}{n}, \frac{i+1}{n}]}) \langle h_i^n, h_j^{\varphi(n)} \rangle_{I_{1,2}} h_j^{\varphi(n)}.\end{aligned}$$

In view of (4.7), we have

$$\begin{aligned}\pi_{\varphi(n)}S_n &= \sqrt{\frac{\varphi(n)}{n}} \sum_{j \in \mathcal{D}_n^\varphi} \left(\sum_{i: i_\varphi=j} \delta_n(\mathbf{1}_{[\frac{i}{n}, \frac{i+1}{n}]}) \right) h_j^{\varphi(n)} \\ &= \sqrt{\frac{\varphi(n)}{n}} \sum_{j \in \mathcal{D}_n^\varphi} \delta_n(\mathbf{1}_{[\frac{j}{\varphi(n)}, \frac{j+1}{\varphi(n)]}) h_j^{\varphi(n)},\end{aligned}$$

since $\cup_{i, i_\varphi=j} [\frac{i}{n}, \frac{i+1}{n}] = [\frac{j}{\varphi(n)}, \frac{j+1}{\varphi(n)}]$. Therefore, for any $\varphi(n) < n$, we have

$$\pi_{\varphi(n)}S_n = \sqrt{\frac{\varphi(n)}{n}} \sum_{j \in \mathcal{D}_n^\varphi} \delta_n(\mathbf{1}_{[\frac{j}{\varphi(n)}, \frac{j+1}{\varphi(n)]}) h_j^{\varphi(n)}. \quad (5.29)$$

Clearly the random variables $\sqrt{\frac{\varphi(n)}{n}} \delta_n(\mathbf{1}_{[\frac{j}{\varphi(n)}, \frac{j+1}{\varphi(n)]})$ are orthogonal because of the $\mathbf{1}_{[i/n, (i+1)/n]}$ term and of unit variance. We can apply Theorem 4.5 and thus

$$\text{dist}_{\text{KR}}(\pi_{\varphi(n)}S_n, \pi_{\varphi(n)}B) \leq c \frac{\varphi(n)^{1+\eta}}{n^{1/2}} \log \left(\frac{\varphi(n)^{1+\eta}}{n^{1/2}} \right). \quad (5.30)$$

Bound on $d(\pi_{\varphi(n)}B, B)$: Since $\pi_n B$ and B are defined on the same probability space, we can make use of the result in Proposition 13.20 of [47], which shows the existence of a constant c such that, for all $\eta < 1/2$,

$$\text{dist}_{W_{n,p}}(\pi_{\varphi(n)}B, B) \leq c \frac{\varphi(n)^\eta}{\sqrt{\varphi(n)}}, \quad n \in \mathbf{N}. \quad (5.31)$$

Bound on $d(\pi_n S_n^k, \pi_{\varphi(n)} S_n^k)$: according to theorem 5.4,

$$d(\pi_n S_n^k, \pi_{\varphi(n)} S_n^k) \leq c \frac{\varphi(n)^\eta}{\sqrt{\varphi(n)}}. \quad (5.32)$$

Bound on $d(S_n^k, \pi_n S_n^k)$: According to Theorem 5.3, for our particular values of $\lambda = n, \kappa = n^{-1/2}$, and $q = n$ we get

$$d(S_n^k, \pi_n S_n^k) \leq \frac{c}{\sqrt{n}} \frac{\log(n) - T}{\log(\log(n) - T) - T - 1} \quad (5.33)$$

Combining equations (5.30), (5.31), (5.33) and (5.32), we get that

$$d(S_n^k, B) \leq c \left(\frac{\varphi(n)^{1+\eta}}{n^{1/2}} \log \left(\frac{\varphi(n)^{1+\eta}}{n^{1/2}} \right) + \frac{\varphi(n)^\eta}{\sqrt{\varphi(n)}} + \frac{\log(n)}{\sqrt{n} \log(\log n)} \right) \quad (5.34)$$

The later expression is minimized for $\varphi(n) \sim n^{1/3}$: after factorizing $\varphi(n)^\eta$, it is a matter of optimizing

$$x \rightarrow \frac{x}{n^{1/2}} \log \left(\frac{x}{n^{1/2}} \right) + \frac{1}{\sqrt{x}}$$

where we can neglect the terms in $\log$.

As n must be an integer multiple of $\varphi(n)$ we can choose $\varphi(n) = n/[n^{2/3}]$. This leads to

$$d(S_n^k, B) \leq c \left(\frac{n^{\eta/3}}{n^{1/6}} \log \left(\frac{n^{\eta/3}}{n^{1/6}} \right) + \frac{n^{\eta/3}}{n^{1/6}} + \frac{\log(n)}{\sqrt{n} \log(\log n)} \right) \sim cn^{\eta/3-1/6} \log(n)$$

□

If we want to compare the distributions of R_n and B , we must evaluate the effect of the time change $\gamma(t) = \int_0^t r(s) ds$ where r is upper and lower bounded. We have the following corollary:

Corollary 5.8. *For an initial process belonging to the class of Markov processes we consider, rescaled according to the LLN and FCLT Assumptions, and for all test functions $f : \mathbb{D} \rightarrow \mathbf{R}$ which are Lipschitz on $\mathbb{D}$ and whose restriction to $W_{\eta,p}$ belongs to $\text{Lip}_1(W_{\eta,p})$ there exists $c > 0$ such that*

$$|\mathbf{E}[f(R_n)] - \mathbf{E}[f(B \circ \gamma)]| \leq c n^{-\frac{1}{6} + \frac{\eta}{3}} \log(n)$$

Proof of Corollary 5.8. We have to estimate

$$\begin{aligned} & \iint_{[0,T]^2} \frac{|\psi \circ \gamma(t) - \psi \circ \gamma(s)|^p}{|t-s|^{1+\eta p}} ds dt \\ &= \iint_{[0,T]^2} \frac{|\psi \circ \gamma(t) - \psi \circ \gamma(s)|^p}{|\gamma(t) - \gamma(s)|^{1+\eta p}} \frac{|\gamma(t) - \gamma(s)|^{1+\eta p}}{|t-s|^{1+\eta p}} ds dt \\ &\leq r_{\max}^{1+\eta p} \iint_{[0,T]^2} \frac{|\psi \circ \gamma(t) - \psi \circ \gamma(s)|^p}{|\gamma(t) - \gamma(s)|^{1+\eta p}} ds dt \\ &= r_{\max}^{1+\eta p} \iint_{[0,\gamma(T)]^2} \frac{|\psi(v) - \psi(u)|^p}{|u-v|^{1+\eta p}} \frac{1}{r(\gamma^{-1}(v))} \frac{1}{r(\gamma^{-1}(u))} du dv \\ &\leq \frac{r_{\max}^{1+\eta p}}{r_{\min}^2} \|\psi\|_{W_{\eta,p}([0,\gamma(T)])}, \end{aligned}$$

thanks to the change of variables formula. □

This implies that there exists $c > 0$ such that

$$|\mathbf{E}[f(\Pi_n R_n)] - \mathbf{E}[f(B \circ \gamma)]| \leq c n^{-\frac{1}{6} + \frac{\eta}{3}} \log n.$$

If we want to compare the distributions of R_n and $B \circ \gamma$, we must use test functions $f : \mathbb{D} \rightarrow \mathbf{R}$ which are Lipschitz on $\mathbb{D}$ and whose restriction to $W_{\eta,p}$ belongs to $\text{Lip}_1(W_{\eta,p})$. In view of Theorem 5.3, the slower rate of convergence is induced by the comparison between $\Pi_n R_n$ and $B \circ \gamma$, thus for such functionals, there exists $c > 0$ such that

$$|\mathbf{E}[f(R_n)] - \mathbf{E}[f(B \circ \gamma)]| \leq c n^{-\frac{1}{6} + \frac{\eta}{3}} \log n \text{ if } \beta \neq 1.$$

Keeping in mind Lemma 5.2 we can state

Corollary 5.9. *For all test functions $f : \mathbb{D} \rightarrow \mathbf{R}$ which are Lipschitz on $\mathbb{D}$ and whose restriction to $W_{\eta,p}$ belongs to $\text{Lip}_1(W_{\eta,p})$ there exists $c > 0$ such that*

$$\mathbf{E} [\|f(U_n) - f(\Theta(W))\|_T] \leq cn^{-\frac{1}{6} + \frac{\eta}{3}} \log(n).$$

Chapter 6

Application to Queueing Models

We show in this chapter how our model can be applied to bound the speed of convergence of the simplest queueing models.

1 The $M/M/\infty$ queue

1.1 Short introduction to the $M/M/\infty$ model

We consider an $M/M/\infty$ queue: a potentially unlimited number of servers attend customers that enter the system following a Poisson process of intensity λ , requesting service times that are exponentially distributed with intensity μ (where $\lambda, \mu > 0$).

If $L^\#(t)$ denotes the number of customers in the system at time t , $L^\#$ is an ergodic Markov process which can be represented as

$$L^\#(t) = L^\#(0) + \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \lambda\}} d\mathcal{N}_{s,z}^1 - \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \mu L^\#(s^-)\}} d\mathcal{N}_{s,z}^{-1}, \quad t \geq 0,$$

where $\mathcal{N}_{s,z}^1$ and $\mathcal{N}_{s,z}^{-1}$ denote two independent marked Poisson measures of $\mathbb{R} \times [0, n]$ with intensities of $ds \otimes dz$, and x_0 is the initial number of customers at time 0.

This process is rescaled by accelerating the time scale of the arrival process while increasing by the same factor n , the number of customers in the initial state and dividing both by n . The new process can be represented as

$$L_n^\#(t) = \frac{L^\#(0)}{n} + \frac{1}{n} \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n\lambda\}} d\mathcal{N}_{s,z}^1 - \frac{1}{n} \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \mu n L_n^\#(s^-)\}} d\mathcal{N}_{s,z}^{-1}, \quad t \geq 0,$$

Based on the latest representation

$$\begin{aligned} L_n^\# - L_n^\#(0) &= \frac{1}{n} \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n\lambda\}} d\tilde{\mathcal{N}}_{s,z}^1 - \frac{1}{n} \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \mu n L_n^\#(s^-)\}} d\tilde{\mathcal{N}}_{s,z}^{-1} \\ &\quad + \lambda t - \mu \int_0^t L_n^\#(s^-) ds \end{aligned}$$

where $\tilde{\mathcal{N}}_{s,z}^1$ and $\tilde{\mathcal{N}}_{s,z}^{-1}$ denote the compensated marked Poisson measures.

It is a well known fact (see for instance [9] or [104]) that the sequence of processes $\{\bar{L}_N^\#\}$ converges to the deterministic process

$$t \mapsto \bar{L}^\#(t) = \rho + (x_0 - \rho)e^{-\mu t}, \quad (6.1)$$

where $\rho = \lambda/\mu$. Moreover, if we define for all N ,

$$t \mapsto U_n(t) = \sqrt{n}(\bar{L}_n^\#(t) - \bar{L}^\#(t)), \quad (6.2)$$

and $U_n(0)$ tends to 0 in a suitable sense, the sequence $\{U_n\}$ converges in distribution to the process X defined by

$$t \mapsto X(t) = X(0)e^{-\mu t} + \int_0^t e^{-\mu(t-s)} \sqrt{h(s)} \, dB(s), \quad (6.3)$$

where $h(t) = 2\lambda + \mu(x - \rho)e^{-\mu t}$ for all $t \geq 0$ and B is the standard Brownian motion. As shown in [104] (eq. (6.28) in Chapter 6), X is an Ornstein-Uhlenbeck process and

$$X(t) = X(0)e^{-\mu t} + \sqrt{2\lambda} \int_0^t e^{-\mu(t-s)} \, dB(s), \quad t \geq 0.$$

A planar representation of the $M/M/\infty$ queue can be given: assume throughout that the initial number of customers $L^\#(0)$ in the queue is zero. A point (x, z) represents a customer arriving at time x and requiring a service of duration z . Let $\mathcal{N}_{\lambda,\mu}$ be a Poisson process on $\mathbb{R}^+ \times \mathbb{R}^+$ of intensity $\lambda \otimes \mu$. The number of busy servers at time t then corresponds to the number of points located in a trapeze bounded by the axes of equation $x = 0$ and $x = t$, and above the line $z = t - x$.

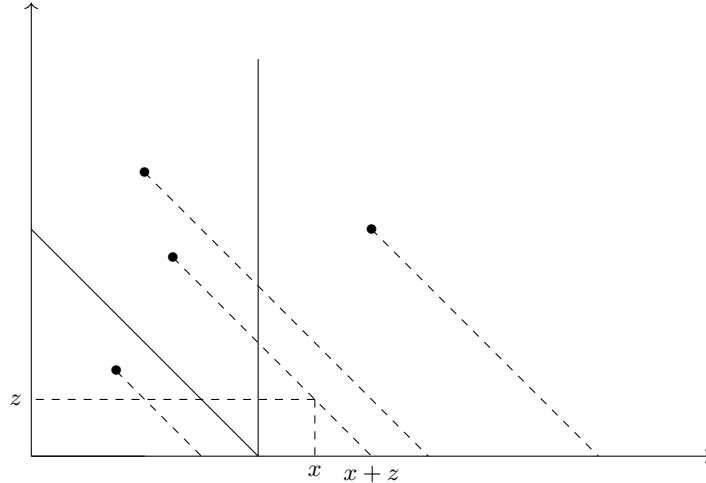


Figure 6.1: Representation of the $M/M/\infty$ queue

Thus for all $t \geq 0$,

$$L^\#(t) = \int \int_{C_t} d\mathcal{N}_{\lambda,\mu}(x, z),$$

where

$$C_t = \{(x, z), 0 \leq x \leq t, z \geq t - x\}.$$

Hence after scaling, for any positive integer n ,

$$\bar{L}_n^\#(t) = \frac{1}{n} \mathcal{N}_{\lambda n, \mu}(C_t).$$

As readily follows from (6.1, the fluid limit $\bar{L}^\#$ can be written for all $t \geq 0$ as

$$\bar{L}^\#(t) = \int \mathbf{1}_{C_t}(x, z) \lambda x e^{-\mu z} \mu \, dx \, dz,$$

so that for all n and t ,

$$U_n(t) = \frac{1}{\sqrt{n}} \int \int \mathbf{1}_{C_t}(x, z) \, d\mathcal{N}_{\lambda n, \mu}(x, z) - \mathbf{1}_{C_t}(x, z) \lambda x e^{-\mu z} \mu \, dx \, dz, \quad (6.4)$$

where U_n is defined by (6.2).

1.2 Straight application of our model

Denote Λ the solution of the following differential equation

$$\Lambda(t) = \lambda t - \mu \int_0^t \Lambda(s) \, ds + \Lambda(0)$$

with initial condition $\Lambda(0) = \lim_{n \rightarrow \infty} L_n^\#(0)$. It can be shown that

$$\Lambda(t) = \frac{\lambda}{\mu} - \left(\Lambda(0) - \frac{\lambda}{\mu} \right) \exp(-\mu t)$$

Defining U_n as $U_n = \sqrt{n}(L_n^\# - \Lambda)$ and U as

$$U(t) = \sqrt{\lambda} B_1(t) - \mu \int_0^t \Lambda(s) \, dB_2(s),$$

B_1 and B_2 being two independent Brownian motions.

We have

Theorem 6.1. *If condition (ii) of the FCLT Assumption is satisfied, for all $n \in \mathbf{N}^*$ and all Lipschitz function F on $\mathbb{D}$ whose restriction to $W_{\eta,p}$ belongs to $\text{Lip}_1(W_{\eta,p})$, we have, defining Θ as in (5.16), there exists $C > 0$ such that*

$$|\mathbf{E}[F(\Theta(U))] - \mathbf{E}[F(U_n)]| \leq \frac{C}{n^{1/6}} \log(n).$$

Proof. Using $\alpha_1 = 1$ and $\alpha_{-1} = 0$, it is clear that all conditions of Lemma 5.1 are met: first

$$\overline{\rho}_1(s, f) = \lambda \text{ and } \overline{\rho}_{-1}(s, f) = \mu f$$

do not depend on n and $\overline{\rho_1}$ and $\overline{\rho_{-1}}$ are obviously Lipschitz continuous with respect to their second variable. In addition, condition (i) of the FCLT Assumption is obviously met with $L_1 = 0, L_2 = \mu, E_{n,1} = E_{n,2} = 0$ so that if condition (ii) is also met, Denoting

$$\tilde{M}_n = \frac{1}{\sqrt{n}} \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n\lambda\}} d\tilde{\mathcal{N}}_{s,z}^1 - \frac{1}{\sqrt{n}} \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \mu n \Psi(s)\}} d\tilde{\mathcal{N}}_{s,z}^{-1},$$

we know from Lemma 5.2 that we have

$$\mathbf{E} \left[\|U_n - \Theta(\tilde{M}_n)\|_T \right] \leq cn^{-3/4},$$

and we can deduce from Corollary 5.9 that

$$\mathbf{E} [\|U_n - \Theta(U)\|_T] \leq cn^{-1/6} \log(n).$$

□

Remark 5. *Using the planar representation of the $M/M\infty$ queue, it is possible, using similar calculations, to compute a more precise speed for the convergence, and to replace the $1/6$ factor in the equation above by a $1/2$ as we showed in [7].*

2 The $M/M/1$ model

2.1 Short introduction to the $M/M/1$ model

The $M/M/1$ queue consists in a single server with infinite queue, where the service times are independently and identically distributed, according to an exponential distribution of parameter μ . The customers arrive at the time epochs of a Poisson process of intensity $\lambda > 0$. Let $L^\dagger(t)$ denote the number of customers in the system (including the one in service, if any) at time $t \geq 0$. The process $(L^\dagger(t), t \geq 0)$ counting the number of customers in the system is clearly a birth and death process, that is ergodic if and only if $\lambda/\mu < 1$. This model can be represented as the Skorokhod reflection (see [103]) of the following process:

$$L^\dagger(t) = R(X) \text{ where } X(t) = x_0 + \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \lambda\}} d\mathcal{N}_{s,z}^1 - \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \mu\}} d\mathcal{N}_{s,z}^{-1}.$$

These processes are rescaled by accelerating the time scale by an arbitrarily large factor n , while increasing the number of clients in the initial state by the same multiplicative factor and dividing the number of clients by n : For all $t \geq 0$,

$$\overline{X}_n(t) = \frac{x_0}{n} + \frac{1}{n} \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n\lambda\}} d\mathcal{N}_{s,z}^1 - \frac{1}{n} \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n\mu\}} d\mathcal{N}_{s,z}^{-1}$$

and $\overline{L}_n^\dagger(t) = R(\overline{X}_n)$.

Denote Λ the solution of the following differential equation

$$\Lambda(t) = x_0 + \lambda t - \mu \int_0^t \mathbf{1}_{\{\Lambda(s^-) > 0\}} ds$$

with initial condition $\Lambda(0) = \lim_{n \rightarrow \infty} L_n^\dagger(0)$. It can be shown that

$$\Lambda(t) = x_0 + (\lambda - \mu)t$$

Defining U_n as $U_n = \sqrt{n}(X_n - \Lambda)$ and U as

$$U(t) = \sqrt{\lambda}B_1(t) - \mu \int_0^t \Lambda(s) \, dB_2(s),$$

B_1 and B_2 being two independent Brownian motions.

We have

Theorem 6.2. *If condition (ii) of the FCLT Assumption is satisfied, for all $n \in \mathbf{N}^*$ and all Lipschitz function F on $\mathbb{D}$ whose restriction to $W_{\eta,p}$ belongs to $\text{Lip}_1(W_{\eta,p})$, we have*

$$|\mathbf{E}[F(U)] - \mathbf{E}[F(U_n)]| \leq \frac{C}{n^{1/6}} \log(n).$$

As a corollary,

$$|\mathbf{E}[F(R(U))] - \mathbf{E}[F(R(U_n))]| \leq \frac{C}{n^{1/6}} \log(n).$$

Proof. Using $\alpha_1 = \alpha_{-1} = 1$, it is clear that all conditions of lemma 5.1 are met:

$$\overline{\rho}_1(s, f) = \lambda \text{ and } \overline{\rho}_{-1}(s, f) = \mu$$

do not depend on n . ρ_1 and ρ_{-1} are obviously Lipschitz continuous with respect to their second variable. As $L_1 = L_{-1} = 0$, Θ defined as in (5.16) is the identity. Therefore using lemma 5.1, for some $C > 0$ X_n converges to Λ and $\overline{L}_n^\dagger$ converges in L^2 to

$$R(\Lambda)(t) = (x_0 + (\lambda - \mu)t)^+$$

Denoting

$$M_n^\dagger = \frac{1}{\sqrt{n}} \int_0^t \left(\int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n\lambda\}} \, d\tilde{\mathcal{N}}_{s,z}^1 - \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq n\mu\}} \, d\tilde{\mathcal{N}}_{s,z}^{-1} \right)$$

Lemma 5.2 ensures that

$$\|\sqrt{n}(X_n - \Lambda) - M_n^\dagger\| \leq \left(\frac{c}{n^{3/4}} \right)$$

Then denoting

$$W = \frac{1}{\sqrt{\lambda + \mu}}(\lambda B_1 + \mu B_2)$$

where B_1 and B_2 are Brownian motions, (W is itself a standard Brownian motion), we can deduce from Corollary 5.9 that

$$\|\sqrt{n}(X_n - \Lambda) - W\| \leq \frac{c}{n^{1/6}} \log(n)$$

The Skorokhod reflection being Lipschitz continuous (see proposition D4 in [103]) we get the corollary. $\square$

2.2 Higher rate of convergence for a smaller class of functions

In [7] we obtained a higher convergence speed for the rescaled $M/M/\infty$ and $M/M/1$ queues by restraining ourselves to a smaller class of functions in the definition of the Wasserstein distance.

The Σ class of functions

Definition 6.3. A function $f : W_n \rightarrow \mathbf{R}$ is said to belong to the class Σ_n whenever it is 1-Lipschitz continuous, twice differentiable in the sense of definition 3.6, and we have

$$\sup_{w \in W_n} \left| \left\langle D_n^{(2)} f_n(w) - D_n^{(2)} f_n(w + g), h \otimes k \right\rangle_{H_n^{\otimes(2)}} \right| \leq \|g\|_W \|h\|_{L^2} \|k\|_{L^2}, \quad (6.5)$$

for any $g \in W_n$, $h, k \in H_n$.

Actually, in the definition of the distance between distributions of processes, the test functions are defined on the whole space W . Hence, we must find a class of functions whose restriction to W_n belong to Σ_n for any $n \geq 1$. This involves the notion of H -differential on W . Let

$$H = \left\{ h, \exists! h' \in L^2([0, T]) \text{ such that } h(t) = \int_0^t h'(s) \, ds \right\}. \quad (6.6)$$

It is an Hilbert space when equipped with the scalar product

$$\langle h, g \rangle_H = \int_0^T h'(s) g'(s) \, ds.$$

A function $f : W \rightarrow \mathbf{R}$ is said to be twice H -differentiable whenever for any $w \in W$, for any $h \in H$, the function

$$\begin{cases} \mathbf{R} & \longrightarrow \mathbf{R} \\ \varepsilon & \longmapsto f(w + \varepsilon h) \end{cases}$$

is twice differentiable in a neighbor of 0. We denote by Df and $D^{(2)}f$ its first and second order gradient, defined by

$$\begin{aligned} \langle Df(x), h \rangle_H &= \left. \frac{d}{d\varepsilon} f(x + \varepsilon h) \right|_{\varepsilon=0}, \\ \left\langle D^{(2)}f(x), h_1 \otimes h_2 \right\rangle_{H^{\otimes(2)}} &= \left. \frac{\partial^2}{\partial \varepsilon_1 \partial \varepsilon_2} f(w + \varepsilon_1 h_1 + \varepsilon_2 h_2) \right|_{\varepsilon_1 = \varepsilon_2 = 0}. \end{aligned}$$

Definition 6.4. The class Σ is the set of 1-Lipschitz continuous, twice H -differentiable functions such that

$$\sup_{w \in W} \left| \left\langle D^{(2)}f(w) - D^{(2)}f(w + g), h \otimes k \right\rangle_{H^{\otimes(2)}} \right| \leq \|g\|_W \|h\|_{L^2} \|k\|_{L^2},$$

for any $g \in W$, $h, k \in H$.

For $f : W \rightarrow \mathbf{R}$, let $f_n = f|_{W_n}$. If f is once H -differentiable, then, we have that for any $w_n \in W_n$, any $j \in \{0, \dots, n-1\}$,

$$\begin{aligned} \langle Df(\mathbf{c}(w_n)), h_n^j \rangle_H &= \left. \frac{d}{dt} f(\mathbf{c}(w_n + \varepsilon h_n^j)) \right|_{\varepsilon=0} = \left. \frac{d}{dt} f_n(w_n + \varepsilon h_n^j) \right|_{\varepsilon=0} \\ &= \langle D_n f_n(w_n), h_n^j \rangle_{H_n}. \end{aligned} \quad (6.7)$$

Thus, it is straightforward that if f belongs to Σ then f_n belongs to Σ_n for any $n \geq 1$.

Remark 6. Let us provide a few instances of functionals which are often encountered in queueing analysis, and which are regular enough to be elements of Σ . This is the case, first, for the function F_f , that is defined for any mild enough function f and $T > 0$, by

$$F_f : \begin{cases} \mathbb{D} & \longrightarrow \mathbf{R} \\ x = (x_t, t \in [0, T]) & \longmapsto \frac{1}{T} \int_0^T f(x_s) ds, \end{cases}$$

observing that $F_f(X)$ goes to $\mathbb{E}_\pi[f]$ for large T whenever the Markov process X is ergodic of invariant probability π . Then, for any $x, y \in W$,

$$|F_f(x) - F_f(y)| \leq \|x - y\|_W$$

provided that f is Lipschitz continuous. Moreover, a classical computation shows that

$$\begin{aligned} \langle D^{(2)} F_f(x + g) - D^{(2)} f(x), h \otimes k \rangle_{H^{\otimes(2)}} \\ = \frac{1}{T} \int_0^T (f''(x_s + g(s)) - f''(x_s)) h_s k_s ds. \end{aligned}$$

Hence F_f belongs to Σ as long as f'' does exist and is Lipschitz continuous. Similarly, for $M \geq 0$ and $p \geq 2$,

$$F_{M,p} : \begin{cases} \mathbb{D} & \longrightarrow \mathbf{R} \\ x & \longmapsto \left(\int_0^T |x_s \wedge M|^p ds \right)^{1/p}. \end{cases}$$

also belongs to the set of admissible test functions. Observe that for M and p large enough, $F_{M,p}(x)$ can be considered as an ersatz to $\sup_{s \leq T} |x_s|$.

For any of these functionals F , if $\text{dist}(\mathbf{P}_{X_n}, \mathbf{P}_X)$ tends to 0 as $n^{-\alpha}$, then the distribution of the random variables $(F(X_n), n \geq 1)$ converges in the sense of a damped Kantorovitch-Rubinstein distance at a rate $n^{-\alpha}$:

$$\sup_{\varphi \in \mathcal{C}_b^3} \left| \mathbf{E} [\varphi(F(X_n))] - \mathbf{E} [\varphi(F(X))] \right| \leq c n^{-\alpha},$$

where $\mathcal{C}_b^3$ is the set of three times differentiable functions from $\mathbf{R}$ to $\mathbf{R}$ with bounded derivatives of any order. Note that this kind of result is inaccessible via the standard Stein method in dimension 1, since we usually cannot achieve the first step of the Stein Method, which consists in devising a functional characterization of the distribution of $F(X)$.

Theorem 5.7 revisited

Using the Σ class of functions we can slightly amend the definition of the Wasserstein distance to

$$\begin{aligned} & \text{dist}_{W_{\eta,p}} \left(\pi_{\varphi(n)}(\pi_n X), \pi_{\varphi(n)}(\pi_n B) \right) \\ &= \sup_{F \in \Sigma(W_{\eta,p})} \left(\mathbf{E} [F_n(\pi_n X)] - \mathbf{E} [F_n(\pi_n B)] \right), \end{aligned}$$

We can then restate theorem 4.5 and get

Theorem 6.5. *Let $(u_a, a \in \mathcal{D}^n)$ be a family of functions satisfying Assumption Orth_p . Then, there exists $c > 0$ such that for all n ,*

$$\text{dist}_{W_{\eta,p}} \left(\pi_{\varphi(n)}(\pi_n X), \pi_{\varphi(n)}(\pi_n B) \right) \leq c \frac{1}{n^{1/2}}.$$

Proof. The newly defined Wasserstein distance, applying 4.10 to F_n is equal to

$$\sup_{F \in \Sigma(W_{\eta,p})} \mathbf{E} \left[\int_0^T L^n P_t^n F_n(\pi_n X) \, dt \right];$$

which can be directly bounded, thanks to lemma 4.4 by

$$\begin{aligned} & c \int_0^T \int_0^1 \sum_{a \in \mathcal{D}^n} \left| \left\langle h_a^n \otimes h_a^n, \nabla^{(2)} P_t^n (F \circ \pi_{\varphi(n)})(x + r h_a^{\varphi(n)}) \right. \right. \\ & \quad \left. \left. - \nabla^{(2)} P_t^n (F \circ \pi_{\varphi(n)})(x) \right\rangle_{I_{1,2}} \right| \, dr \, dt \end{aligned}$$

Using the Mehler formula and definition 6.4, this can be bounded, for $\varphi(n) = n$ by

$$c \int_0^T \int_0^1 n e^{-2t} |r| \|h_a^{\varphi(n)}\|^3 \, dr \, dt$$

and remembering that $\|h_a^{\varphi(n)}\| \leq 1/\sqrt{n}$ we finally get a bound proportional to $1/\sqrt{n}$. The proof is thus complete. $\square$

It is then a simple matter to recalculate the bound in theorem 5.7. If η is small, the dominating term in the triangular inequality becomes $d(S_n^k, \pi_n S_n^k)$, equal to $\frac{\log(n)}{\sqrt{n}}$ (see equation (5.33)) and we get

$$d(S_n^k, B) \leq \frac{\log(n)}{\sqrt{n}}.$$

Chapter 7

Other Applications

We illustrate in this model other uses for our model in the fields of genetics and epidemics.

1 The ON/OFF model

Let $(Y_i)_{i \in \mathbf{N}}$ be a sequence of independent and identically distributed pure jump Markov processes taking values in $\{0, 1\}$. For any i , we denote respectively by σ_0 and σ_1 , the transition intensities of Y_i from 0 to 1, and from 1 to 0. We also let π_0 and π_1 be the stationary probability of Y_i being in states 0 and 1 respectively, that is,

$$\pi_0 = \frac{\sigma_1}{\sigma_0 + \sigma_1} \text{ and } \pi_1 = \frac{\sigma_0}{\sigma_0 + \sigma_1}.$$

We are interested in the asymptotic behavior, as n goes large, of the process S_n defined by

$$X_n(t) = \sum_{i=0}^n Y_i(t), \quad t \geq 0.$$

In a networking context, the above can be seen as the process counting the number of 'ON' sources, in a communication system in which there are n sources independently alternating between 'ON' and 'OFF' states.

It is immediate to observe that for any n , the limiting distribution $X_n^*(\infty)$ of the process $(X_n(t))_{t \geq 0}$ is binomial of parameters n , π_1 , and that

$$\sqrt{n} \left(\frac{X_n(\infty)}{n} - \pi_1 \right) \Longrightarrow \mathcal{N}(0, \pi_0 \pi_1).$$

At the process level, it is shown in [25] that under suitable assumptions on the initial conditions, for all $T > 0$,

$$U_n = \sqrt{n} \left(\frac{X_n}{n} - \Lambda \right) \Longrightarrow \Theta(U) \quad \text{in } \mathbb{D}([0, T], \mathbf{R}),$$

where

- For a fixed $\Lambda(0) \in \mathbf{R}_+$,

$$\Lambda(t) = \pi_1 + (\Lambda(0) - \pi_1) \exp(-(\sigma_1 + \sigma_0)t), \quad t \geq 0; \quad (7.1)$$

- Θ is defined by (5.16);
- The process U is defined by

$$U(t) = \sigma_0 \int_0^t (1 - \Lambda(s)) \, dB^1(s) - \sigma_1 \int_0^t \Lambda(s) \, dB^2(s), \quad t \geq 0,$$

for B^1 and B^2 , two independent standard Brownian motions.

Hereafter we provide a bound on the speed of the latter convergence,

Theorem 7.1. *Suppose that condition (i) of the FCLT Assumption is satisfied. Then for all $n \in \mathbf{N}^*$ and all F which are Lipschitz on $\mathbb{D}$ and whose restriction to $W_{\eta,p}$ belongs to $\text{Lip}_1(W_{\eta,p})$, we have that*

$$|\mathbf{E}[F(\Theta(U))] - \mathbf{E}[F(U_n)]| \leq \frac{C}{n^{1/6}} \log(n).$$

Proof. Let for all $n \geq 1$, $\bar{X}_n(t) = X_n(t)/n$, $t \geq 0$. The sequence $(\bar{X}_n)_{n \geq 1}$ naturally takes places in the settings of Section 1 for $d = 1$, $k = 2$, $\alpha_1 = \alpha_2 = 0$ and $\zeta_1 = 1$ and $\zeta_2 = -1$. Indeed, it is easily seen that for any $n \in \mathbf{N}$ and $t \geq 0$ we have the equality in distribution

$$\begin{aligned} \bar{X}_n(t) &\stackrel{(d)}{=} \bar{X}_n(0) + \int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{\{z \leq n\sigma_0(1 - \bar{X}_n(s^-))\}} \, d\mathcal{N}_1(s, z) \\ &\quad - \int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{\{z \leq n\sigma_1 \bar{X}_n(s^-)\}} \, d\mathcal{N}_2(s, z), \end{aligned}$$

where $\mathcal{N}_1$ and $\mathcal{N}_2$ denote two independent marked Poisson measures of common intensity $ds \otimes dz$, representing the overall “up” and “down” jumps, respectively. It is then clear that all conditions of the LLN Assumption are met, for

$$\bar{\rho}_1 : (s, x) \mapsto \sigma_0(1 - x) \quad \text{and} \quad \bar{\rho}_2 : (s, x) \mapsto \sigma_1 x,$$

which are obviously Lipschitz continuous with respect to their second variable. Plainly, Λ defined by (7.1) is the unique solution of (5.5) in the present case. As ρ_1 and ρ_2 are linear in their second variable, condition (ii) in the FCLT Assumption is clearly satisfied for $L_1 = L_2 \equiv 1$ and $E_{n,1} \equiv E_{n,2} = 0$ for all n , and so the corresponding operator Θ defined by (5.16) is linear, continuous and therefore Lipschitz continuous. Thus recalling (5.12), from Theorem 5.2 we get that

$$\mathbf{E}[|U_n - \Theta(M_{n,1,\Lambda} + M_{n,2,\Lambda})|] \leq cn^{-3/4},$$

where in the present case, for any n and t we have that

$$\begin{aligned} M_{n,1,\Lambda}(t) &= \frac{1}{\sqrt{n}} \int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{\{\frac{z}{n} \leq \sigma_0(1 - \Lambda(s))\}} \, d\tilde{\mathcal{N}}^1(s, z); \\ M_{n,2,\Lambda}(t) &= -\frac{1}{\sqrt{n}} \int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{\{\frac{z}{n} \leq \sigma_1 \Lambda(s)\}} \, d\tilde{\mathcal{N}}^2(s, z), \end{aligned}$$

Both $M_{n,1,\Lambda}$ and $M_{n,2,\Lambda}$ clearly respect all the conditions of Theorem 5.7, therefore applying Corollary 5.9 the result follows. $\square$

2 The SIR model

We consider a population of size n in which an epidemics propagates. Individuals can go through three states: susceptible (S), infected (I) and recovered (R). Let for any t , $S(t)$, $I(t)$ and $R(t)$ denote respectively the number of susceptibles, infected and recovered individuals at time t . Each infected individual that is in contact with a susceptible one transmits the disease to the latter after a time that is exponentially distributed of parameter λ/n , independently of everything else. We make the classical mixing assumption of Kermack and Mc Kendrick [68], i.e., the connectivity graph between individuals is complete, i.e., any individual is connected to any other one (contrary to recent models in which the connectivity graph is heterogeneous - see e.g. [31]). For any infectious individual, the duration of the infection is independent of the rest of the model and exponentially distributed with parameter γ . Since $R(t) = n - S(t) - I(t)$ for all t , a Markov description of this system is given by the bivariate process $X_n(t)$, defined by

$$X_n(t) = \begin{pmatrix} S_n(t) \\ I_n(t) \end{pmatrix}, \quad t \geq 0.$$

We perform the classical large graph scaling of X_n , to obtain the scaled process $\bar{X}_n$, defined by

$$\bar{X}_n(t) = \frac{1}{n} X_n(t) = \begin{pmatrix} S_n(t)/n \\ I_n(t)/n \end{pmatrix} =: \begin{pmatrix} \bar{S}_n(t) \\ \bar{I}_n(t) \end{pmatrix}, \quad t \geq 0.$$

Let $\Lambda := \begin{pmatrix} s \\ i \end{pmatrix}$ be the unique solution of the ODE

$$\begin{cases} s' &= -\lambda s i \\ i' &= \lambda s i - \gamma i, \end{cases} \quad (7.2)$$

and for all n ,

$$U_n = \sqrt{n} (\bar{X}_n - \Lambda).$$

If we denote

$$U(t) = \lambda \int_0^t s(u) i(u) \, dB_u^1 \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \gamma \int_0^t i(u) \, dB_u^2 \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

where B^1 and B^2 are two independent Brownian motions, we have the following result

Theorem 7.2. *Suppose that condition (i) of the FCLT Assumption holds. Then for all $n \in \mathbf{N}^*$ and all F which are Lipschitz on $\mathbb{D}$ and whose restriction to $W_{\eta,p}$ belongs to $\text{Lip}_1(W_{\eta,p})$, we have that*

$$|\mathbf{E}[F(\Theta(U))] - \mathbf{E}[F(U_n)]| \leq \frac{C}{n^{1/6}} \log(n),$$

for Θ defined by (5.16).

Proof. Fix $n \geq 1$. We can use the following representation for the X_n process: for all $t \geq 0$,

$$X_n(t) = \begin{pmatrix} S(0) \\ I(0) \end{pmatrix} + \left(\int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{\{z \leq \frac{\lambda}{n} S_n(u^-) I_n(u^-)\}} d\mathcal{N}^1(s, z) \right) \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \left(\int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{\{z \leq \gamma I_n(u^-)\}} d\mathcal{N}^2(u, z) \right) \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$

for two independent Poisson measures $\mathcal{N}^1$ and $\mathcal{N}^2$ of common intensity $ds \otimes dz$, and thus

$$\bar{X}_n(t) = \begin{pmatrix} \bar{S}^n(0) \\ \bar{I}^n(0) \end{pmatrix} + \frac{1}{n} \left(\int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{\{z \leq \lambda n \bar{S}_n(u^-) \bar{I}_n(u^-)\}} d\mathcal{N}^1(u, z) \right) \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \frac{1}{n} \left(\int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{\{z \leq \gamma n \bar{I}_n(u^-)\}} d\mathcal{N}^2(u, z) \right) \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Consequently, the LLN Assumption is clearly satisfied in dimension $d = 2$, for $\alpha_1 = \alpha_2 = 0$, and for all s and f ,

$$\bar{\rho}_1(s, f) = \lambda f_1 f_2, \quad \text{and} \quad \bar{\rho}_2(s, f) = \gamma f_2.$$

The ODE (5.5) thus becomes

$$\Lambda(t) := \begin{pmatrix} s(t) \\ i(t) \end{pmatrix} = \begin{pmatrix} s(0) \\ i(0) \end{pmatrix} + \left(\int_0^t \lambda s(u^-) i(u^-) du \right) \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \left(\int_0^t \gamma i(u^-) du \right) \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad t \geq 0,$$

or in other words, (7.2).

The functions $\bar{\rho}_1$ and $\bar{\rho}_2$ are obviously Lipschitz continuous with respect to their second variable. Therefore using lemma 5.1, $\begin{pmatrix} s_n(t) \\ i_n(t) \end{pmatrix}$ converges in L^2 to $\Lambda(t) = \begin{pmatrix} s(t) \\ i(t) \end{pmatrix}$ where s and i are the solutions of the system of ordinary differential equations: $\bar{\rho}_2(s, \cdot)$ is linear and $\bar{\rho}_1(s, \cdot)$ is such that

$$\begin{aligned} \rho_1(s, \bar{X}_n) - \rho_1(s, \Lambda(s)) &= \lambda[(s_n - s)i + (i_n - i)s + (i_n - i)(s_n - s)] \\ &= \frac{\lambda}{\sqrt{n}} \begin{pmatrix} i \\ s \end{pmatrix} \cdot U_n + (i_n - i)(s_n - s) \\ &= L_1(s, U_n) + (i_n - i)(s_n - s) \end{aligned}$$

where $U_n = \sqrt{n} \begin{pmatrix} s_n - s \\ i_n - i \end{pmatrix}$.

Condition (ii) in the FCLT Assumption is therefore clearly satisfied for all n , and so the corresponding operator Θ defined by (5.16) is linear, continuous and therefore Lipschitz continuous. Thus recalling (5.12), from Theorem 5.2 we get that

$$\mathbf{E}[|U_n - \Theta(M_{1,n} - M_{2,n})|] \leq cn^{3/4},$$

where

$$M_{1,n} = \frac{1}{\sqrt{n}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{\{z \leq n\lambda\}} d\tilde{\mathcal{N}}_{u,z}^1 \text{ and}$$

$$M_{2,n} = \frac{1}{\sqrt{n}} \begin{pmatrix} 0 \\ -1 \end{pmatrix} \int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{\{z \leq n\gamma\}} d\tilde{\mathcal{N}}_{u,z}^2.$$

Both $M_{1,n}$ and $M_{2,n}$ clearly respect all the conditions of Theorem 5.7, therefore applying Corollary 5.9 the result follows. $\square$

3 The Moran model with selection

We consider in this section the Moran model described in [65]: in a population of size n , each individual bears a gene liable to take two forms : A and B . Each individual has one single parent and its child inherits the genetic form of its parent. At exponential rate a pair of individuals is sampled uniformly at random within the population. One of them dies and the other splits in two. One out of the two forms bears a selection advantage over the other one so that the probability of an increase in the population taking the A form is $rA_n(1 - A_n)/n(n-1)$ whereas the probability of a decrease is $RA_n(1 - A_n)/n(n-1)$. In addition, every individual of type A changes its type independently at rate ν_1 and each individual of type B changes its type independently at rate ν_2 . In [65], Kasprzak uses the ideas of Kurtz ([75]) to write X_n , the proportion of individuals carrying the gene of type A as

$$X_n(t) = \frac{1}{n} P_1(n^2 R_1^n) - \frac{1}{n} P_{-1}(n^2 R_{-1}^n)$$

where R_1^n and R_{-1}^n are time changes given by the equations

$$\begin{aligned} R_1^n &= \int_0^t \left(\frac{1}{2} X_n(s) + \frac{\nu_2}{n} \right) (1 - X_n(s)) ds \\ R_{-1}^n &= \int_0^t \left(\frac{1}{2} (1 - X_n(s)) + \frac{\nu_1}{n} \right) X_n(s) ds \end{aligned}$$

and P_1 and P_{-1} are two independent Poisson processes of intensity 1, which are not independent of R_1^n and R_{-1}^n . He then applies the Stein Method directly, but in order to make his calculations tractable he is forced to consider a similar process instead, where P_1 and P_{-1} are also independent of R_1^n and R_{-1}^n . Our representation of the same Moran Model is different: if $A_n(t)$ denotes the number of individuals of type A , we can write it as:

$$\begin{aligned} A_n(t) &= A_n(0) + \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \frac{rA_n(n-A_n)}{n(n-1)}\}} d\mathcal{N}_{s,z}^1 - \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \frac{RA_n(n-A_n)}{n(n-1)}\}} d\mathcal{N}_{s,z}^{-1} \\ &\quad + \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \nu_2(n-A_n)\}} d\mathcal{N}_{s,z}^{1'} - \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \nu_1 A_n\}} d\mathcal{N}_{s,z}^{-1'} \end{aligned}$$

This process is rescaled by accelerating the time scale of splits and deaths by a factor of $n-1$ and that of mutations by n while dividing the number of individuals by n so that $X_n = A_n/n$ represents the proportion of individuals carrying the gene of type A : For all $t \geq 0$,

$$\begin{aligned} X_n(t) &= X_n(0) + \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq rn \frac{X_n(1-X_n)}{n-1}\}} d\mathcal{N}_{s,z}^1 - \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq Rn \frac{X_n(1-X_n)}{n-1}\}} d\mathcal{N}_{s,z}^{-1} \\ &\quad + \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \nu_2(1-X_n)\}} d\mathcal{N}_{s,z}^{1'} - \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \leq \nu_1 X_n\}} d\mathcal{N}_{s,z}^{-1'} \end{aligned}$$

Using $\alpha_1 = \alpha_{-1} = \alpha_{1'} = \alpha_{-1'} = 1$, we cannot meet all conditions of Lemma 5.1 because the terms in the integrals cannot be made independent of n : the double independence on n and $n-1$ makes this impossible and the Moran model does not fall in the class of Markov models considered in the present study. In order to extend our method to the Moran model, we would need to make the fluid limit Λ appear in the equation and to make the rest negligible.

Chapter 8

Appendix

1 Besov-Liouville Spaces

. Let f be in $\mathcal{L}^p([0, T]; dt)$, (denoted $\mathcal{L}^p$ in a compact way) the fractional left and right integrals of f are defined respectively by :

$$(I_{0+}^\alpha f)(x) = \frac{1}{\Gamma(\alpha)} \int_0^x f(t)(x-t)^{\alpha-1} dt, \quad x \geq 0,$$

$$(I_{T-}^\alpha f)(x) = \frac{1}{\Gamma(\alpha)} \int_x^T f(t)(t-x)^{\alpha-1} dt, \quad x \leq T,$$

where $\alpha \geq 0$, and Γ is the Euler function. Moreover, $I_{0+}^0 = I_{T-}^0 = \text{Id}$. For all $\alpha \geq 0, p, q \geq 1$, $f \in \mathcal{L}^p$ and $g \in \mathcal{L}^q$ such that $p^{-1} + q^{-1} \leq \alpha + 1$, the following equality, a direct consequence of the Fubini theorem, is verified:

$$\int_0^T f(s)(I_{0+}^\alpha g)(s) ds = \int_0^T (I_{T-}^\alpha f)(s)g(s) ds. \quad (8.1)$$

For all $p \in [1, +\infty]$, the Besov-Liouville space $I_{0+}^\alpha(\mathcal{L}^p) := \mathcal{I}_{\alpha,p}^+$ is usually equipped with the following norm :

$$\|I_{0+}^\alpha f\|_{\mathcal{I}_{\alpha,p}^+} = \|f\|_{\mathcal{L}^p}. \quad (8.2)$$

Similarly, the Besov-Liouville space $I_{1-}^\alpha(\mathcal{L}^p) := \mathcal{I}_{\alpha,p}^-$ is usually equipped with the norm :

$$\|I_{T-}^\alpha f\|_{\mathcal{I}_{\alpha,p}^-} = \|f\|_{\mathcal{L}^p}.$$

The next theorem enables to compare Besov-Liouville spaces with the spaces of Hölderian continuous functions, and to formulate some continuity implications (see [45],[105]) :

Theorem 8.1. *i. If $0 < \alpha < 1$, and $1 < p < 1/\alpha$, I_{0+}^α is a bounded operator of $\mathcal{L}^p$ on $\mathcal{L}^q$ where $q = p(1 - \alpha p)^{-1}$. (see [105], Theorem 3.5)*

ii. For all $0 < \alpha < 1$ and $p \geq 1$, $\mathcal{I}_{\alpha,p}^+$ is continuously embedded in $\text{Hol}_0(\alpha - 1/p)$ if $\alpha - 1/p > 0$. $\text{Hol}_0(\nu)$ denotes the space of α -Hölderian functions, vanishing in $t=0$, equipped with the usual norm. (see [105], Theorem 3.8)

iii. For all $0 < \alpha < \beta < 1$, $\text{Hol}_0(\beta)$ is compactly embedded in $\mathcal{I}_{\alpha,\infty}$ (see [45], Proposition 2)

iv. $I_{0+}^{-\alpha}$, and $I_{T-}^{-\alpha}$, denote the inverse mappings of I_{0+}^{α} , and I_{T-}^{α} respectively. Equations $I_{0+}^{\alpha} I_{0+}^{\beta} f = I_{0+}^{\alpha+\beta} f$, and $I_{T-}^{\alpha} I_{T-}^{\beta} f = I_{T-}^{\alpha+\beta} f$, are verified if $\beta > 0, \alpha + \beta > 0$ and $f \in \mathcal{L}^1$. (see [105], Theorem 2.5)

v. For $\alpha p < 1$, the $\mathcal{I}_{\alpha,p}^+$ and $\mathcal{I}_{\alpha,p}^-$ spaces are canonically isomorphic, which enables to write in an abusive way $\mathcal{I}_{\alpha,p}$ to denote those spaces. (see [105], Corollary 1 of Theorem 11.5)

We now recall the definitions and properties of Besov-Liouville spaces of negative order, which are dual to the positive order spaces (see [26]). Let $\mathcal{D}_+$ be the space of C^∞ functions defined over $[0, 1]^n$ and such that $\phi^{(k)}(0) = 0$, for all $k \in \mathbb{N}$. Similarly let $\mathcal{D}_-$ be the space of C^∞ functions defined over and such that $\phi^{(k)}(1) = 0$, for all $k \in \mathbb{N}$. Those two spaces are equipped with the projective topology induced by the semi norms

$$p_k(\phi) = \sum_{j \leq k} \|\phi^{(j)}\|_\infty, \forall k \in \mathbb{N}.$$

Denote $\mathcal{D}'_+$, resp. $\mathcal{D}'_-$ their strong topological dual. It is clear that $\mathcal{D}_+$ is stable under I_{0+}^β as well as $\mathcal{D}_-$ under I_{1-}^β , for all $\beta \in \mathbb{R}^+$. Therefore thanks to equality (8.1), we can define a fractional integral for all distributions (ie. for all members of $\mathcal{D}'_-$ ou $\mathcal{D}'_+$) :

$$\text{For } \theta \in \mathcal{D}'_-, I_{0+}^\beta \theta : \phi \in \mathcal{D}_- \mapsto \langle \theta, I_{T-}^\beta \phi \rangle_{\mathcal{D}'_-, \mathcal{D}_-},$$

$$\text{For } \theta \in \mathcal{D}'_+, I_{T-}^\beta \theta : \phi \in \mathcal{D}_+ \mapsto \langle \theta, I_{0+}^\beta \phi \rangle_{\mathcal{D}'_+, \mathcal{D}_+}$$

The Besov-Liouville space of negative order is therefore defined in the following way:

Definition 8.2. Let $\beta > 0$ and $r > 1$, $\mathcal{I}_{-\beta,r}^+$ (resp. $\mathcal{I}_{-\beta,r}^-$) is the space of those distributions $\theta \in \mathcal{D}'_-$ (resp. $\theta \in \mathcal{D}'_+$) such that $I_{0+}^\beta \theta$ (resp. $I_{T-}^\beta \theta$) belongs to $\mathcal{L}^r$. The norm of a θ element in this space is the norm of $I_{0+}^\beta \theta$ in $\mathcal{L}^r$ (resp. that of $I_{T-}^\beta \theta$).

Theorem 8.3. For $\beta > 0$ and $r > 1$, the dual space of $\mathcal{I}_{\beta,r}^+$ (resp. $\mathcal{I}_{\beta,r}^-$) is canonically isometric and isomorphic to $I_{T-}^{-\beta}(\mathcal{L}^{r*})$ (resp. $I_{0+}^{-\beta}(\mathcal{L}^{r*})$) where $r^* = r(r-1)^{-1}$. In addition for all $\beta \geq \alpha \geq 0$ and $r > 1$, I_{T-}^β is continuous from $\mathcal{I}_{-\alpha,r}^-$ to $\mathcal{I}_{\beta-\alpha,r}^-$.

Proof. $\mathcal{L}^r$ and $\mathcal{L}^{r*}$ are clearly dual. For $(f, g) \in \mathcal{L}^r \times \mathcal{L}^{r*}$,

$$\begin{aligned} \left\langle I_{0+}^\beta(f), I_{T-}^{-\beta}(g) \right\rangle_{\mathcal{I}_{\beta,r}^+, \mathcal{I}_{T-}^{-\beta}(\mathcal{L}^{r*})} &= \int_{[0;T]} [I_{0+}^\beta(f)](s) [I_{T-}^{-\beta}(g)](s) ds \\ &= \int_{[0;T]} [I_{0+}^{-\beta} \circ I_{0+}^\beta(f)](s) g(s) ds \text{ because of (8.1)} \\ &= \int_{[0;T]} f(s) g(s) ds \text{ by the semi group property} \\ &= \langle f, g \rangle_{\mathcal{L}^r, \mathcal{L}^{r*}}. \end{aligned}$$

□

According to theorem 8.1 i) all functions in $\mathcal{I}_{\alpha,2}^-$ belong to $\subset \mathcal{L}^2$, and we even have

Theorem 8.4. *The canonical embedding κ_α from $\mathcal{I}_{\alpha,2}^-$ to $\mathcal{L}^2$ is Hilbert-Schmidt if and only if $\alpha > 1/2$ (i.e. $\forall i, \alpha_i > \frac{1}{2}$). In addition,*

$$c_\alpha := \|\kappa_\alpha\|_{HS} = \|I_{0+}^\alpha\|_{HS} = \|I_{T-}^\alpha\|_{HS} = \frac{1}{\sqrt{2}} \Gamma(\alpha) \left(\frac{T^{2\alpha}}{\alpha(\alpha - 1/2)} \right)^{1/2}.$$

Proof. Recall that in a Hilbert space, a trace operator is a compact operator for which we can define a trace in the sense of linear algebra, which is finite and independent of the choice for the basis. In other words, an operator A is a trace operator if there exists a complete orthonormal basis $(e_k, k \geq 1)$ such that the following series is converging:

$$\text{Tr}(|A|) = \sum_k |\langle Ae_k, e_k \rangle|.$$

An operator A between two Hilbert spaces H^1 and H^2 equipped with the complete orthonormal bases $(e_k^1)_{k \geq 1}$ and $(e_k^2)_{k \geq 1}$ respectively is said to be Hilbert Schmidt if

$$\|A\|_{HS}^2 := \sum_k \|Ae_k^1\|_{H_2}^2 = \sum_k \sum_n \langle Ae_k^1, e_n^2 \rangle_{H_2}^2 < \infty.$$

Let $(e_k, k \geq 1)$ be an complete orthonormal basis of $\mathcal{L}^2([0; T]^n)$ then $(h_k^\alpha = I_{T-}^\alpha(e_k), k \in \mathbb{N})$ is a complete orthonormal basis $\mathcal{I}_{\alpha,2}^-$ and

$$\|\kappa_\alpha\|_{HS}^2 = \sum_{k \geq 1} \|h_k^\alpha\|_{\mathcal{L}^2}^2 = \sum_k \|I_{T-}^\alpha(e_k)\|_{\mathcal{L}^2}^2 = \|I_{T-}^\alpha\|_{HS}^2.$$

$$\begin{aligned}
\|\kappa_\alpha\|_{\text{HS}}^2 &= \sum_k \|I_{T-}^\alpha(e_k)\|_{\mathcal{L}^2}^2 \\
&= \sum_k \int_0^T |I_{T-}^\alpha(e_k)(s)|^2 ds \\
&= \frac{1}{\Gamma(\alpha)^2} \sum_k \int_0^T |\langle e_k, (\cdot - s)^{\alpha-1} \mathbf{1}_{[s,T]}(\cdot) \rangle|^2 ds \\
&= \frac{1}{\Gamma(\alpha)^2} \int_0^T \sum_k (|\langle e_k, (\cdot - s)^{\alpha-1} \mathbf{1}_{[s,T]}(\cdot) \rangle|^2) ds \\
&= \frac{1}{\Gamma(\alpha)^2} \int_0^T \|(\cdot - s)^{\alpha-1} \mathbf{1}_{[s,T]}(\cdot)\|_{L^2}^2 ds \\
&= \frac{1}{\Gamma(\alpha)^2} \int_0^T \int_0^T \|(t - s)^{2(\alpha-1)} \mathbf{1}_{[s,T]}(t)\|_{L^2}^2 dt ds \\
&= \frac{2}{\Gamma(\alpha)^2} \int_{[0,T]} \int_s^T (t - s)^{2\alpha-2} dt ds \\
&= \frac{2}{\Gamma(\alpha)^2} \int_{[0,T]} \left[\frac{(t - s)^{2\alpha-1}}{2\alpha - 1} \right]_s^T ds \\
&= \frac{2}{\Gamma(\alpha)^2} \int_{[0,T]} \frac{(T - s)^{2\alpha-1}}{2\alpha - 1} ds = \frac{2}{\Gamma(\alpha)^2} \left[\frac{-(T - s)^{2\alpha}}{2\alpha(2\alpha - 1)} \right]_0^T \\
&= \frac{1}{\Gamma(\alpha)^2} \frac{T^{2\alpha}}{2\alpha(\alpha - 1/2)} < \infty.
\end{aligned}$$

By the same line of reasoning, $c_\alpha = \|I_{0+}^\alpha\|_{HS}$.

□

Let $\tau \in [0, T]$, and ϵ_τ the Dirac measure at point τ . According to theorem (8.1), i, ϵ_τ belongs to $(\mathcal{I}_{\alpha,2}^-)'$ for all $\alpha > 1/2$. Let us estimate the norm of ϵ_τ on this space.

Lemma 8.5. *For all $\alpha > 1/2$, $\tau \in [0, T]$, the image of ϵ_τ by j_α , the canonical isometry between $(\mathcal{I}_{\alpha,2}^-)'$ and $\mathcal{I}_{\alpha,2}^-$, is the function*

$$j_\alpha(\epsilon_\tau) : s \mapsto I_{T-}^\alpha((\cdot - \tau)_+^{\alpha-1})(s),$$

and

$$\|j_\alpha(\epsilon_\tau)\|_{\mathcal{I}_{\alpha,2}^-}^2 = \sum_{k \in \mathbb{N}} |h_k^\alpha(\tau)|^2 = \frac{(T - \tau)^{2\alpha-1}}{(2\alpha - 1)\Gamma(\alpha)^2}. \quad (8.3)$$

Proof. By definition of the dual product, for all $h = I_{T-}^\alpha(\dot{h})$ where $\dot{h} \in \mathcal{L}^2$,

$$\begin{aligned}
\langle \epsilon_\tau, h \rangle_{(\mathcal{I}_{\alpha,2}^-)', \mathcal{I}_{\alpha,2}^-} &= h(\tau) = \frac{1}{\Gamma(\alpha)} \int_\tau^T (s - \tau)^{\alpha-1} \dot{h}(s) ds \\
&= \left\langle \frac{1}{\Gamma(\alpha)} (\cdot - \tau)_+^{\alpha-1}, \dot{h} \right\rangle_{\mathcal{L}^2} \\
&= \langle I_{T-}^\alpha((\cdot - \tau)_+^{\alpha-1}), h \rangle_{\mathcal{I}_{\alpha,2}^-, \mathcal{I}_{\alpha,2}^-},
\end{aligned}$$

hence the first claim. Moreover, thanks to the Parseval identity in $\mathcal{L}^2$,

$$\begin{aligned}
\sum_{k \in \mathbb{N}} \|h_k^\alpha(\tau)\|^2 &= \sum_{k \in \mathbb{N}} \langle \epsilon_\tau, h_k^\alpha \rangle_{(\mathcal{I}_{\alpha,2}^-)', \mathcal{I}_{\alpha,2}^-}, \\
&= \sum_{k \in \mathbb{N}} \langle j_\alpha(\epsilon_\tau), h_k^\alpha \rangle_{\mathcal{I}_{\alpha,2}^-, \mathcal{I}_{\alpha,2}^-}, \\
&= \frac{1}{\Gamma(\alpha)^2} \sum_{k \in \mathbb{N}} \langle (\cdot - \tau)_+^{\alpha-1}, e_k \rangle_{\mathcal{L}^2}^2 \\
&= \frac{1}{\Gamma(\alpha)^2} \|(\cdot - \tau)_+^{\alpha-1}\|_{\mathcal{L}^2}^2 \\
&= \|j_\alpha(\epsilon_\tau)\|_{\mathcal{I}_{\alpha,2}^-}^2.
\end{aligned}$$

So that (8.3) follows by quadrature. $\square$

Since standard Brownian motion exhibits Hölderian trajectories of any order strictly smaller than $1/2$, Theorem (8.1) ensures that such trajectories belong to $\mathcal{I}_{\beta,\infty} \subset \mathcal{I}_{\beta,2}$ for all $\beta < 1/2$. In addition, a simple calculation shows that for all $\alpha \in (0, 1)$,

$$\mathbf{1}_{[a,+\infty)} = \Gamma(\alpha) I_{0+}^\alpha ((\cdot - a)_+^{-\alpha}).$$

where $(\cdot - a)_+$ is the function from $\mathbb{R}^n$ into $\mathbb{R}$ which maps $x = (x_1, \dots, x_n)$ to $(x_1 - a)_+$. Hence $\mathbf{1}_{[a,+\infty)}$ belongs to $\mathcal{I}_{1/2-\epsilon,2}$ for all $\epsilon > 0$. Notice that piecewise linear functions also belong to $\mathcal{I}_{\beta,2}$ for all $\beta < 1/2$. We can therefore chose to do all the calculations in a $\mathcal{I}_{\beta,2}$ space with $\beta < 1/2$. The closer β is to $1/2$, the more meaningful is the distance, but the higher the error bound.

2 Moment bound for Poisson variables

By following closely Chapter 2 in [10], we show hereafter a moment bound for the maximum of n Poisson variables. (Notice that, contrary to Exercise 2.18 in [10] we do not assume here that the Poisson variables are independent.)

Proposition 8.6. *Let $(X_i, i = 1 \dots n)$ be Poisson random variables of parameter ν . The Lambert W function is defined over $[-1/e, \infty]$ by the equation $W(x)e^{W(x)} = x$. Then*

$$\mathbf{E} \left[\max_{i=1, \dots, n} X_i \right] \leq \frac{\log n / e^\nu}{W(\log(n/e^\nu)/\nu e)} = \nu e \exp(W(\log(n/e^\nu)/\nu e))$$

Proof. Let's consider $(Z_i, i = 1 \dots n)$ the centered Poisson variables (i.e., for all i , $Z_i = X_i - \nu$). By a straightforward calculation, for all $u \in \mathbf{R}$ and all i ,

$$\mathbf{E} [e^{uZ_i}] = e^{-u\nu} \sum_{k=0}^{\infty} e^{uk} e^{-\nu} \frac{\nu^k}{k!} = e^{-u\nu - \nu} e^{\nu e^u}$$

Therefore the logarithm of the moment generating function of Z_i is $\Psi_{Z_i}(u) = \nu(e^u - u - 1)$

By Jensen's inequality, and because $\exp(\cdot)$ is an increasing function,

$$\exp\left(u\mathbf{E}\left[\max_{i=1,\dots,n} Z_i\right]\right) \leq \mathbf{E}\left[\exp\left(u\max_{i=1,\dots,n} Z_i\right)\right] = \mathbf{E}\left[\max_{i=1,\dots,n} \exp(uZ_i)\right]$$

Because the maximum of a sequence of positive numbers is lower than its sum, the right-hand side of the last equation is lower than $\mathbf{E}\left[\sum_{i=0}^n \exp(uZ_i)\right]$.

Hence, because of the definition of Ψ_{Z_i} ,

$$\exp\left(u\mathbf{E}\left[\max_{i=1,\dots,n} Z_i\right]\right) \leq \sum_{i=1}^n \mathbf{E}[\exp(uZ_i)] \leq n \exp(\Psi_{Z_i}(u)) = n \exp(\nu(e^u - u - 1))$$

Taking the log, for any u in $\mathbf{R}$,

$$u\mathbf{E}\left[\max_{i=1,\dots,n} Z_i\right] - \nu(e^u - u - 1) \leq \log(n)$$

so that

$$\mathbf{E}\left[\max_{i=1,\dots,n} Z_i\right] \leq \inf_{u \in \mathbf{R}} \left(\frac{\log(n) + \nu(e^u - u - 1)}{u} \right)$$

Taking the derivative on the right-hand side, it is easy to check that the infimum is reached when

$$\nu u e^u - \nu e^u + \nu = \log(n) \quad (8.4)$$

which means that it is reached when $(u - 1)e^{u-1} = \frac{\log(n/e^\nu)}{e^\nu}$ i.e. $u = 1 + W(a)$ where $a = \frac{\log(n/e^\nu)}{e^\nu}$. Then the infimum is equal to

$$\frac{\log n + \nu(e^{1+W(a)} - 1 - W(a) - 1)}{1 + W(a)} \quad (8.5)$$

But we know from (8.4) that $\nu(1 + W(a))e^{1+W(a)} - \log(n) = \nu e^{1+W(a)} - \nu$ so that (8.5) is equal to

$$\nu e^{1+W(a)} - \nu = \nu e e W(a) - \nu = \nu e \frac{a}{W(a)} - \nu$$

Remembering that the Z_i are the centered X_i we thus obtain that

$$\mathbf{E}\left[\max_{i=1,\dots,n} X_i\right] \leq \nu e \frac{a}{W(a)} - \nu + \nu = \frac{\log(n/e^\nu)}{W(\log(n/e^\nu)/\nu e)}$$

which completes the proof. $\square$

Remark 7. We conclude by observing that $W(z) \geq \log(z) - \log \log(z)$ for all $z > e$. Therefore for $n \geq \exp(e^{\nu+1} + \nu)$,

So that if ν is small relative to n , using the second expression for the bound of the expectation of the maximum in Proposition 8.6,

$$\mathbf{E}\left[\max_{i=1,\dots,n} X_i\right] \leq \frac{\log(n/e^\nu)}{\log\left(\frac{\log(n/e^\nu)}{\nu e}\right)}$$

Bibliography

- [1] Robert A. Adams and John J. F. Fournier, *Sobolev Spaces*, Academic Press, jun 2003 (en).
- [2] Bremaud Baccelli, *Elements of queueing theory*, 01 2003.
- [3] A. D. Barbour, *Stein's Method and Poisson Process Convergence*, Journal of Applied Probability **25** (1988), 175–184.
- [4] ———, *Stein's method for diffusion approximations*, Probability Theory and Related Fields **84** (1990), no. 3, 297–322 (en).
- [5] M. S. Bartlett, *Some evolutionary stochastic processes*, Journal of the Royal Statistical Society. Series B (Methodological) **11** (1949), no. 2, 211–229.
- [6] Forest Baskett, K. Mani Chandy, Richard R. Muntz, and Fernando G. Palacios, *Open, closed, and mixed networks of queues with different classes of customers*, J. ACM **22** (1975), no. 2, 248–260.
- [7] Eustache Besançon, Laurent Decreusefond, and Pascal Moyal, *Stein's method for diffusive limits of queueing processes*, Queueing Systems: Theory and Applications **95** (2020), no. 3, 173–201.
- [8] Patrick Billingsley, *Convergence of Probability Measures*, 2 edition ed., Wiley-Interscience, New York, jul 1999 (English).
- [9] A. A. Borovkov, *Limit Laws for queueing processes in multichannel systems*, Sibirsk Mat **8** (1967), no. 5, 983–1004.
- [10] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart, *Concentration inequalities*, Oxford University Press, 2013.
- [11] Anton Braverman, *Stein's method for steady-state diffusion approximations*, 2017.
- [12] Anton Braverman and J. G. Dai, *Stein's method for steady-state diffusion approximations of $m/Ph/n + m$ systems*, Ann. Appl. Probab. **27** (2017), no. 1, 550–581.
- [13] H. Brézis, *Analyse fonctionnelle*, masson ed., 1987.
- [14] Krzysztof Burdzy and Ellen Toby, *A skorohod-type lemma and a decomposition of reflected brownian motion*, Ann. Probab. **23** (1995), no. 2, 586–604.

- [15] Duncan S Callaway, Mark EJ Newman, Steven H Strogatz, and Duncan J Watts, *Network robustness and fragility: Percolation on random graphs*, Physical review letters **85** (2000), no. 25, 5468.
- [16] Louis H. Y. Chen, *On the Convergence of Poisson Binomial to Poisson Distributions*, The Annals of Probability **2** (1974), no. 1, 178–180 (EN).
- [17] Sylvie Roelly- Coppoletta, *A criterion of convergence of measure-valued processes: application to measure branching processes*, Stochastics **17** (1986), no. 1-2, 43–65.
- [18] Cristina Costantini, *The skorohod oblique reflection problem in domains with corners and application to stochastic differential equations*, Probability Theory and Related Fields **91** (1992), 43–70.
- [19] L. Coutin and L. Decreusefond, *Stein’s method for Brownian approximations*, Communications on Stochastic Analysis **7** (2013), no. 3, 349–372, 00000.
- [20] L. Coutin and Laurent Decreusefond, *Convergence rate in the rough donsker theorem*, arXiv:1707.01269 [math] (2017), arXiv: 1707.01269.
- [21] ———, *Donsker’s theorem in Wasserstein-1 distance*, working paper or preprint, April 2019.
- [22] Laure Coutin and Laurent Decreusefond, *Stein’s method for rough paths*, Potential Analysis (2019).
- [23] Donald Dawson, *Measure-valued markov processes*, Ecole d’Eté de Probabilités de Saint-Flour XXI - 1991 (Berlin, Heidelberg) (Paul-Louis Hennequin, ed.), Springer Berlin Heidelberg, 1993, pp. 1–260.
- [24] R. Dante Deblasse and Ellen H. Toby, *Reflecting brownian motion in a cusp*, Transactions of the American Mathematical Society **339** (1993), no. 1, 297–321.
- [25] Decreusefond, Etienne, Lang, Robin, and Vallois, *Longest excursion of the Ornstein Uhlenbeck Process: applications to genomics and telecom*, 2015.
- [26] L. Decreusefond, *Stochastic integration with respect to Volterra processes*, Annales de l’Institut Henri Poincare (B) Probability and Statistics **41** (2005), no. 2, 123–149.
- [27] L. Decreusefond and P. Moyal, *Fluid limit of the $m/m/1/1+gi$ -edf queue*, IEEE Proceedings Symposium on Applications and the Internet Workshops 05, 2005, pp. 248–251.
- [28] L. Decreusefond and P. Moyal, *Fluid limit of a heavily loaded EDF queue with impatient customers.*, Markov Process. Relat. Fields **14** (2008), no. 1, 131–158 (English).
- [29] L. Decreusefond and P. Moyal, *Stochastic modeling and analysis of telecom networks*, ISTE Ltd and John Wiley & Sons Inc, 2012, 00000.

- [30] Laurent Decreusefond, *The Stein-Dirichlet-Malliavin method*, ESAIM: Proceedings (2015), 11 (en).
- [31] Laurent Decreusefond, Jean-Stéphane Dhersin, Pascal Moyal, and Viet Chi Tran, *Large graph limit for an sir process in random network with heterogeneous connectivity*, Ann. Appl. Probab. **22** (2012), no. 2, 541–575.
- [32] Laurent Decreusefond and Pascal Moyal, *A functional central limit theorem for the $m/gi/\infty$ queue*, Ann. Appl. Probab. **18** (2008), no. 6, 2156–2178.
- [33] O Diekmann and JAP Heesterbeek, *Mathematical epidemiology of infectious diseases*, Model Building, Analysis (1989).
- [34] Peter Donnelly and Thomas G. Kurtz, *A countable representation of the fleming-viot measure-valued diffusion*, Ann. Probab. **24** (1996), no. 2, 698–742.
- [35] Monroe D. Donsker, *Justification and extension of doob’s heuristic approach to the kolmogorov-smirnov theorems*, Ann. Math. Statist. **23** (1952), no. 2, 277–281.
- [36] Bogdan Doytchinov, John Lehoczky, and Steven Shreve, *Real-time queues in heavy traffic with earliest-deadline-first queue discipline*, Ann. Appl. Probab. **11** (2001), no. 2, 332–378.
- [37] P. Dupuis and K. Ramanan, *Convex duality and the skorokhod problem. i*, Probability Theory and Related Fields **115** (1999), 153–195.
- [38] ———, *Convex duality and the skorokhod problem. ii*, Probability Theory and Related Fields **115** (1999), 197–236.
- [39] Paul Dupuis and Hitoshi Ishii, *On lipschitz continuity of the solution mapping to the skorokhod problem, with applications*, Stochastics and Stochastic Reports **35** (1991), no. 1, 31–62.
- [40] Paul Dupuis and Hitoshi Ishii, *Sdes with oblique reflection on nonsmooth domains*, Ann. Probab. **21** (1993), no. 1, 554–580.
- [41] Paul Erdős and Alfréd Rényi, *On the evolution of random graphs*, Publ. Math. Inst. Hung. Acad. Sci **5** (1960), no. 1, 17–60.
- [42] Agner Krarup Erlang, *The theory of probabilities and telephone conversations*.
- [43] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes: characterization and convergence*, Wiley series in probability and statistics, Wiley Interscience, Hoboken, 2005, OCLC: ocm62724166.
- [44] Denis Feyel and Arnaud de La Pradelle, *Fractional integrals and brownian processes*, 1996.
- [45] Denis Feyel and Arnaud de la Pradelle, *On Fractional Brownian Processes*, Potential Analysis **10** (1999), no. 3, 273–288 (en).

- [46] Sir Fisher, Ronald Aylmer, *The genetical theory of natural selection.*, OxfordClarendon Press, 1930.
- [47] P. K. Friz and N. B. Victoir, *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*, Cambridge University Press, feb 2010 (en), Google-Books-ID: CVgwLatxfGsC.
- [48] Erol Gelenbe, *G-networks with triggered customer movement*, Journal of Applied Probability **30** (1993), no. 3, 742–748.
- [49] William J. Gordon and Gordon F. Newell, *Closed queuing systems with exponential servers*, Operations Research **15** (1967), no. 2, 254–265.
- [50] R.C. Griffiths, *Lines of descent in the diffusion approximation of neutral wright-fisher models*, Theoretical Population Biology **17** (1980), no. 1, 37 – 50.
- [51] H. Gromoll, Philippe Robert, Bert Zwart, and Richard Bakker, *The impact of reneging in processor sharing queues*, vol. 34, 06 2006, pp. 87–96.
- [52] H. Christian Gromoll, Amber L. Puha, and Ruth J. Williams, *The fluid limit of a heavily loaded processor sharing queue*, Ann. Appl. Probab. **12** (2002), no. 3, 797–859.
- [53] Leonard Gross, *Abstract wiener spaces*, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Contributions to Probability Theory, Part 1 (Berkeley, Calif.), University of California Press, 1967, pp. 31–42.
- [54] J. Michael Harrison and Martin I. Reiman, *Reflected brownian motion on an orthant*, Ann. Probab. **9** (1981), no. 2, 302–308.
- [55] Herbert W Hethcote, *Qualitative analyses of communicable disease models*, Mathematical Biosciences **28** (1976), no. 3-4, 335–356.
- [56] ———, *An immunization model for a heterogeneous population*, Theoretical population biology **14** (1978), no. 3, 338–349.
- [57] ———, *An age-structured model for pertussis transmission*, Mathematical biosciences **145** (1997), no. 2, 89–136.
- [58] ———, *The mathematics of infectious diseases*, SIAM review **42** (2000), no. 4, 599–653.
- [59] HW Hethcote, L Gross, TG Hallam, and SA Levin, *Applied mathematical ecology*, 1989, p. 119.
- [60] Kiyosi Itô and Makiko Nisio, *On the convergence of sums of independent Banach space valued random variables*, Osaka Journal of Mathematics **5** (1968), no. 1, 35–48 (EN).
- [61] James R. Jackson, *Networks of waiting lines*, Operations Research **5** (1957), no. 4, 518–521.
- [62] ———, *Jobshop-like queueing systems*, Management Science **10** (1963), no. 1, 131–142.

- [63] Weining Kang and Kavita Ramanan, *Characterization of stationary distributions of reflected diffusions*, Ann. Appl. Probab. **24** (2014), no. 4, 1329–1374.
- [64] ———, *On the submartingale problem for reflected diffusions in domains with piecewise smooth boundaries*, Ann. Probab. **45** (2017), no. 1, 404–468.
- [65] Mikolaj J. Kasprzak, *Diffusion approximations via stein’s method and time changes*, 2017.
- [66] F. P. Kelly, *Networks of queues with customers of different types*, Journal of Applied Probability **12** (1975), no. 3, 542–554.
- [67] David G. Kendall, *Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded markov chain*, Ann. Math. Statist. **24** (1953), no. 3, 338–354.
- [68] William Ogilvy Kermack, A. G. McKendrick, and Gilbert Thomas Walker, *A contribution to the mathematical theory of epidemics*, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **115** (1927), no. 772, 700–721.
- [69] Aleksandr Iakovlevich Khinchin, D. M. Andrews, M. H. (Maurice Henry) Quenouille, and E. Wolman, *Mathematical methods in the theory of queueing*, 2nd ed. with additional notes by Eric Wolman ed., London : Griffin, 1969 (English), "Translated from the Russian text first published under the title Matematicheskie metody teorii massovogo obsluzhivaniia, as Vol. 49 (1955) of the series Trudy Matematicheskogo Instituta Steklov, Akad. Nauk, U.S.S.R.".
- [70] J. F. C. Kingman, *The single server queue in heavy traffic*, Mathematical Proceedings of the Cambridge Philosophical Society **57** (1961), no. 4, 902–904.
- [71] J. F. C. Kingman, *On the genealogy of large populations*, 1982.
- [72] J.F.C. Kingman, *The first erlang century—and the next*, Queueing Systems **63** (2009), 3–12.
- [73] Leonard Kleinrock, *Information flow in large communication nets*, *ph.d. thesis proposal*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 1961.
- [74] Hui-Hsiung Kuo and Y. J. Lee, *Integration by parts formula and the stein lemma on abstract wiener space*, 2011.
- [75] Thomas G. Kurtz, *Representations of markov processes as multiparameter time changes*, Ann. Probab. **8** (1980), no. 4, 682–715.
- [76] Y. Kwon and R. J. Williams, *Reflected brownian motion in a cone with radially homogeneous reflection field*, Transactions of the American Mathematical Society **327** (1991), no. 2, 739–780.

- [77] Youngmee Kwon, *The submartingal problem for brownian motion in a cone with non-constant oblique reflection*, Probability Theory and Related Fields **92** (1992), 351–391.
- [78] John Lamperti, *Probability: A survey of the mathematical theory*, 2 edition ed., Wiley-Interscience, New York, aug 1996 (English).
- [79] Steve Hung-Mou Lu and P. R. Kumar, *Distributed scheduling based on due dates and buffer priorities*, IEEE Transactions on Automatic Control **36** (1991), 1406–1416.
- [80] S. Méléard and S. Roelly, *Sur les convergences étroite ou vague de processus à valeurs mesures*, C. R. Acad. Sci. Paris, Série I **317**, 1993, pp. 785–788.
- [81] Lauren Meyers, *Contact network epidemiology: Bond percolation applied to infectious disease prediction and control*, Bulletin of the American Mathematical Society **44** (2007), no. 1, 63–86.
- [82] Lauren Ancel Meyers, MEJ Newman, and Babak Pourbohloul, *Predicting epidemics on directed contact networks*, Journal of theoretical biology **240** (2006), no. 3, 400–418.
- [83] JC Miller and E Volz, *Simple rules govern epidemic dynamics in complex networks*, to appear (2011).
- [84] Joel C Miller, *A note on a paper by erik volz: Sir dynamics in random networks*, Journal of mathematical biology **62** (2011), no. 3, 349–358.
- [85] P. A. P. Moran, *Random processes in genetics*, Mathematical Proceedings of the Cambridge Philosophical Society **54** (1958), no. 1, 60–71.
- [86] Pascal Moyal, *Stationarity of measure-valued stochastic recursions: Applications to the pure delay system and the SRPT queue.*, C. R., Math., Acad. Sci. Paris **345** (2007), no. 4, 233–237 (English).
- [87] Mark EJ Newman, *Spread of epidemic disease on networks*, Physical review E **66** (2002), no. 1, 016128.
- [88] Mark EJ Newman, Steven H Strogatz, and Duncan J Watts, *Random graphs with arbitrary degree distributions and their applications*, Physical review E **64** (2001), no. 2, 026118.
- [89] Ivan Nourdin and Giovanni Peccati, *Stein’s method on Wiener chaos*, Probability Theory and Related Fields **145** (2009), no. 1-2, 75–118.
- [90] S. Novak, *Extreme Value Methods with Applications to Finance*. <https://www.crcpress.com/Extreme-Value-Methods-with-Applications-to-Finance/Novak/9781439835746>, 2011.
- [91] David Nualart and Giovanni Peccati, *Central limit theorems for sequences of multiple stochastic integrals*, Ann. Probab. **33** (2005), no. 1, 177–193.
- [92] Esa Nummelin, *General irreducible markov chains and non-negative operators*, Cambridge Tracts in Mathematics, Cambridge University Press, 1984.

- [93] WP Peterson, *Diffusion approximations for networks of queues with multiple customer types*, Mathematics of Operations Research **16** (1991), 90–118.
- [94] M. Ann Piech, *The ornstein-uhlenbeck semigroup in an infinite dimensional l_2 setting*, Journal of Functional Analysis **18** (1975), no. 3, 271 – 285.
- [95] F. Pollaczek, *Über eine aufgabe der wahrscheinlichkeitstheorie. ii (mitteilung aus dem telegraphentechnischen reichsamt)*, Mathematische Zeitschrift **32** (1930), 729–750.
- [96] Félix Pollaczek, *Problèmes stochastiques posés par le phénomène de formation d’une queue d’attente à un guichet et par des phénomènes apparentés*, Gauthier-Villars, 1957 (fre).
- [97] Nicolas Privault, *Stochastic analysis in discrete and continuous settings with normal martingales*, Lecture Notes in Mathematics, vol. 1982, Springer-Verlag, Berlin, 2009.
- [98] Yu. V. Prokhorov, *Convergence of random processes and limit theorems in probability theory*, Theory of Probability & Its Applications **1** (1956), no. 2, 157–214.
- [99] Kavita Ramanan and Martin I. Reiman, *Fluid and heavy traffic diffusion limits for a generalized processor sharing model*, Ann. Appl. Probab. **13** (2003), no. 1, 100–139.
- [100] V. Ramaswami, *A stable recursion for the steady state vector in markov chains of m/g/1 type*, Communications in Statistics. Stochastic Models **4** (1988), no. 1, 183–188.
- [101] Martin I. Reiman, *Open queueing networks in heavy traffic*, Mathematics of Operations Research **9** (1984), no. 3, 441–458.
- [102] Erdos RENYI, *On random graph*, Publicationes Mathematicae **6** (1959), 290–297.
- [103] P. Robert, *Stochastic networks and queues*, french ed., Applications of Mathematics, vol. 52, Springer-Verlag, Berlin, 2003, Stochastic Modelling and Applied Probability.
- [104] Philippe Robert, *Réseaux et files d’attente: méthodes probabilistes | Philippe Robert | Springer*, Springer, 2000.
- [105] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional integrals and derivatives: theory and applications*, Gordon and Breach Science Publishers, Switzerland ; Philadelphia, Pa., USA, 1993 (eng).
- [106] Laurent Schwartz, *Seminaire Laurent Schwartz : 1969-1970: Applications radonifiantes*, Ecole Polytechnique, 1969 (fr), Google-Books-ID: 1GjmnAAACA AJ.

- [107] ———, *Produits tensoriels gp et dp , applications p -sommantes, applications p -radonifiantes*, Séminaire Bourbaki vol. 1970/71 Exposés 382–399 (Berlin, Heidelberg) (A. Donald Heidelberg and B. Eckmann Zürich, eds.), Springer Berlin Heidelberg, 1972, pp. 63–88.
- [108] Thomas Seidman, ‘*first come, first served*’ can be unstable!, Automatic Control, IEEE Transactions on **39** (1994), 2166 – 2171.
- [109] Hsin-Hung Shih, *On Stein’s method for infinite-dimensional Gaussian approximation in abstract Wiener spaces*, Journal of Functional Analysis **261** (2011), no. 5, 1236–1283.
- [110] A. V. Skorokhod, *Limit theorems for stochastic processes*, Theory of Probability & Its Applications **1** (1956), no. 3, 261–290.
- [111] ———, *Stochastic equations for diffusion processes in a bounded region*, Theory of Probability & Its Applications **6** (1961), no. 3, 264–274.
- [112] Charles Stein, *A bound for the error in the normal approximation to the distribution of a sum of dependent random variables*, The Regents of the University of California, 1972 (EN).
- [113] Daniel W. Stroock and S. R.S. Varadhan, *Diffusion processes with boundary conditions*, Communications on Pure and Applied Mathematics **24** (1971), no. 2, 147–225 (English (US)).
- [114] Hiroshi Tanaka, *Stochastic differential equations with reflecting boundary condition in convex regions*, Hiroshima Math. J. **9** (1979), no. 1, 163–177.
- [115] Lee M. Taylor and Ruth J. Williams, *Existence and uniqueness of semi-martingale reflecting brownian motions in an orthant*, Probability Theory and Related Fields **96** (1993), 283–317.
- [116] S. R. S. Varadhan and R. J. Williams, *Brownian motion in a wedge with oblique reflection*, Communications on Pure and Applied Mathematics **38** (1985), no. 4, 405–443.
- [117] Cedric Villani, *Topics in Optimal Transportation*, American Mathematical Society, Providence, RI, mar 2003 (English).
- [118] ———, *Optimal Transport - Old and New* | Cedric Villani | Springer, 2009.
- [119] Ruth J. Williams, *Reflected brownian motion with skew symmetric data in a polyhedral domain*, Probability Theory and Related Fields **75** (1987), 459–485.
- [120] Sewall Wright, *Evolution in mendelian populations*, Genetics **16** (1931), no. 2, 97.

Titre : Vitesse de convergence des approximations diffusions

Mots clés : Calcul de Malliavin, Méthode de Stein, Approximations diffusions

Résumé Dans de nombreux champs d'applications, les processus de Markov sont un outil privilégié de modélisation de processus aléatoires. Malheureusement, il est souvent nécessaire d'avoir recours à des espaces d'états très grands voire infinis, rendant l'analyse exacte des différentes caractéristiques (stabilité, loi stationnaire, temps d'atteinte de certains domaines, etc.) du processus délicate ou impossible. Depuis longtemps, grâce notamment à la théorie des martingales, on procède à des approximations par des diffusions browniennes. Celles-ci permettent souvent une analyse approchée du modèle d'origine. Le principal défaut de cette approche est que l'on ne connaît pas l'erreur commise dans cette approximation. Il s'agit donc ici de développer une théorie du calcul d'erreur dans les approximations diffusion. Depuis quelques temps, le développement de la méthode de Stein-Malliavin a permis de préciser les vitesses de convergence dans les théorèmes classiques comme le théorème de Donsker (convergence fonctionnelle d'une marche aléatoire vers un mouvement brownien) ou la généralisation trajectorielle de l'approximation binomiale-Poisson.

Il s'agit dans ce travail de poursuivre le développement de cette théorie pour des processus de Markov comme ceux que l'on rencontre en théorie des files d'attente ou en épidémiologie et dans bien d'autres domaines appliqués. Partant de la représentation des processus de Markov comme mesures de Poisson, on étend la méthode développée par Laurent Decreusefond et Laure Coutin pour estimer la vitesse de convergence dans les approximations diffusion. Pour ce faire, on étend la méthode de Stein-Malliavin à des vecteurs de processus dépendants plutôt qu'à un seul processus. La limite est un processus gaussien changé de temps. La méthode de Stein Malliavin étant développée surtout pour montrer la convergence vers le mouvement Brownien standard, on l'adapte à la convergence vers un processus changé de temps à travers des méthodes d'approximations linéaires. On fait donc appel à l'analyse gaussienne pour caractériser les dépendances entre intervalles de temps et à l'analyse fonctionnelle pour déterminer les bons espaces probabilisés.

Title : Speed of Convergence of Diffusion Approximations

Keywords : Malliavin Calculus, Stein Method, Diffusion Approximations

Abstract : In many fields of interest, Markov processes are a primary modelisation tool for random processes. Unfortunately it is often necessary to use very large or even infinite dimension state spaces, making the exact analysis of the various characteristics of interest (stability, stationary law, hitting times of certain domains, etc.) of the process difficult or even impossible. For quite a time, thanks in particular to martingale theory, it has been possible to make use of approximations by brownian diffusions. This enables an approximate analysis of the initial problem.

The main drawback of this approach is that it does not measure the error made in this approximation. The purpose is to develop a theory of error calculation for diffusion approximations.

For some time, the development of the Stein-Malliavin method has enabled to get some precision over speed of convergence in classical theorems such as the Donsker theorem (functional convergence of a random walk towards the Brownian motion) or in the generalisation of the Binomial Poisson approximation path by path.

In this work we intend to extend the development of this theory for Markovian processes such as those than can be found in queueing theory, in epidemiology or in other fields of application.

Starting from the representation of Markov processes as Poisson measures, we extend the method developed by Laurent Decreusefond and Laure Coutin to assess the speed of convergence in diffusion approximations. To do so, we extend the Stein-Malliavin method to vectors of processes rather than a single process. The limit is a gaussian process changed in time. The Stein Malliavin method being mainly developed to calculate convergence towards the standard Brownian motion, it is adapted to the problem of convergence towards a time changed process using linear approximation methods. We therefore make use of Gaussian analysis to assess the dependency between the various time periods and to functional analysis to elect the right probabilistic spaces.