Can Piecewise deterministic Markov process help biologist?

Nathalie Krell

Université de Rennes

Habilitation à diriger les recherches de l'Université de Rennes

Spécialité : **Mathématiques**

Par

Nathalie Krell

Can Piecewise deterministic Markov process help biologist?

Rapporteurs :

Mme. Sophie	**Hautphenne**	The University of Melbourne.
M. Florent	**Malrieu**	Institut Denis Poisson, Université de Tours.
Mme. Patricia	**Reynaud-Bouret**	Laboratoire J.A. Dieudonné, Université Côte d'Azur.

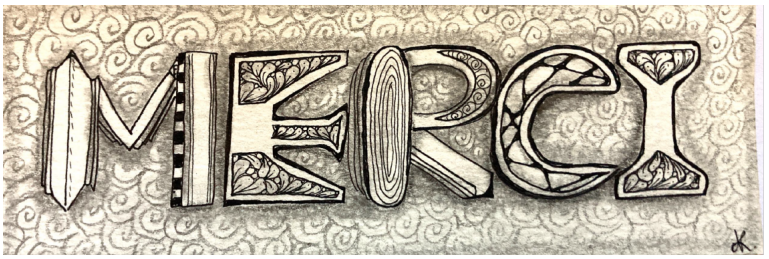
Soutenue publiquement le **2 juin 2023** devant le jury composé de

M. Jean-Christophe	**Breton**	Université de Rennes	Examinateur
Mme. Benoîte	**de Saporta**	Université de Montpellier	Examinatrice
Mme. Anne	**Gegout-Petit**	Université de Lorraine	Examinatrice
Mme. Sophie	**Hautphenne**	The University of Melbourne	Rapportrice
Mme. Eva	**Löcherbach**	Université Paris–1	Examinatrice
M. Florent	**Malrieu**	Université de Tours	Rapporteur
Mme. Patricia	**Reynaud-Bouret**	Université Côte d'Azur	Rapportrice

IRMAR

Université de Rennes

Remerciements



Même si un dessin vaut mille mots, je vais tout de même en écrire quelques-uns.

Je souhaite remercier les membres du jury pour avoir accepté de prendre part à cette soutenance et de s'être déplacé pour l'occasion. À ceci s'ajoute une gratitude particulière pour Eva pour m'avoir permis de découvrir le co-encadrement de thèse. Jean-Christophe pour sa patience infinie à répondre à toutes mes questions dans des domaines très variés. Benoîte pour ses conseils, discussions, parties de jeux,... D'une façon générale, je remercie vivement mes rapporteur/trices pour leur travail malgré leur emploi du temps déjà très chargé. Merci Florent de m'avoir fait découvrir le merveilleux monde des PDMPs et de m'avoir ainsi donné une nouvelle direction de recherche. Merci Sophie d'avoir accepté de rédiger un premier rapport d'HDR. Je remercie Patricia et Anne pour leurs discussions et conseils.

Au fil des ans j'ai eu l'occasion de nouer de fructueuses collaborations, dont certaines m'ont laissée d'excellents souvenirs tant personnels que professionnels. Je tiens donc à remercier tous mes collaborateurs passés, présents et futurs, ... même si certains projets prennent beaucoup plus de temps que prévu et nous confrontent à des difficultés inattendues je suis sûre qu'on finira pas y arriver.

Il serait fastidieux de dresser la liste de tous les collègues de Rennes (en France et dans le monde) à qui je suis redevable : l'ensemble des membres des équipes de probabilités et de statistiques du laboratoire pourrait y trouver sa place mais aussi de nombreux autres collègues dans d'autres équipes. Sans oublier aussi tous les collègues rencontrés en conférences, au CNU, en comité de sélection,...

J'ai aussi une pensée particulière pour ceux qui essaye de faire à leur échelle des choses face aux enjeux climatiques, en particulier, j'espère que le livre que l'on commence à écrire sur des exercices pour sensibiliser les étudiants aux problèmes climatiques verra bien le jour.

Bien sûr il y a aussi tout le personnel administratif que ça soit les secrétaires du labo, des enseignements, les bibliothécaires, les informaticiens,... sans vous la vie aurait été impossible et aussi beaucoup moins sympa.

Mais aussi les étudiants qui comme les collègues m'ont aussi beaucoup appris sur des choses à faire ou au contraire à ne surtout pas faire... Merci à Pierre de m'avoir fait découvrir l'encadrement doctoral.

Je n'oublierais pas non plus mes amis et ma famille et celle qui a donné un tout autre sens à ma vie: Jeanne.

Contents

Introduction and presentation of the links with my thesis

0.1 Introduction

This document is an overview of my different works since my PhD thesis. There are essentially two themes which constitute the main part of my work since my thesis: PDMP (piecewise deterministic Markov process) and the modeling of processes coming from biology. We will see that these two themes are very often linked.

This document consists of an introduction and three chapters.

In the introduction, I will present my main object of study during my thesis: the fragmentations. I will especially highlight the key tool that I used in the work I did on PDMP: the many to-one-formula. This tool is the link between my thesis work about fragmentations and what I did afterwards related to PDMP. I will not make a chronological presentation, because I prefer a more logical presentation of the different results in order to show the underlying links between the different works. Having introduced some notations and definitions, at the end of the introduction I will detail more precisely the contents of the following chapters.

Then, there will be three chapters that make up my three main research themes. The first one is about the modeling of the growth of the size of E. Coli bacterium. In a second chapter, I will be interested in the estimation of the jump rate of PDMP. In the last chapter, I will care about more complex processes because they are processes with values in $\mathbb{R}_+^N$ which have more complex dependency structures. I will go into more details about these three main axes at the end of this introduction.

And at the end of each chapter, I will give some research perspectives that are works in progress or longer term research perspectives.

0.2 The rates of presence in homogeneous fragmentation

There are different types of fragmentation. You can refer to Jean Bertoin's book (Ber06) (see also (Bas06) and (Ber03)). I will make a short overview of this subject. I will start by defining the interval fragmentation point of view because it allows to define naturally the associated marked fragment.

We consider a homogenous fragmentation F of intervals, which is a Markov process in continuous time taking its values in the set $\mathcal{U}$ of open sets of $(0,1)$. Informally, each interval component - or *fragment* - splits as time goes on, independently of the others and with the same law, up to a rescaling. We make the restriction that the fragmentation is conservative, which means that no mass is lost. In this case, the law of the fragmentation F is completely characterized by the so-called dislocation measure ν (which corresponds to the jump-component of the process) which is a measure on $\mathcal{U}$ fulfilling the following conditions

$$\nu((0,1)) = 0,$$

$$\int_{\mathcal{U}} (1-u_1)\nu(dU) < \infty, \tag{0.1}$$

and

$$\sum_{i=1}^{\infty} u_i = 1 \qquad \text{for } \nu - \text{almost every } U \ \in \mathcal{U},$$

where for $U \in \mathcal{U}$,

$$|U|^{\downarrow} := (u_1, u_2, ...)$$

is the decreasing sequence of the lengths of the interval components of U.

It appears quite natural to study the rates of decay of fragments. If we measure the fragments by logarithms of their sizes, a homogeneous fragmentation can be considered as an extension of a classical branching random walk in continuous time. The common feature of many branching models consists in the alternative between exponential growth and extinction. Let us recall some basic facts about a Galton-Watson process $(\zeta_n)_{n\geq 0}$ started from $\zeta_0 = 1$ with finite mean $m = \mathbb{E}\zeta_1$. We have $\lim_{n\to\infty} n^{-1} \log \mathbb{E}(\zeta_n) = \log m$ and

(a) if $m > 1$, and $\mathbb{P}(\zeta_1 \geq 1) = 1$ then $\lim_{n\to\infty} n^{-1} \log Z_n = \log m$ a.s.

(b) if $m < 1$ then $\lim_{n\to\infty} n^{-1} \log \mathbb{P}(Z_n \neq 0) = \log m$.

More generally, in branching random walks, when the local rate of growth of the population in expectation is (exponentially) positive, it is a.s. the effective local rate of growth of the population. When it is negative, it is the local rate of decrease of the probability of presence.

After the defense of my Thesis, Alain Rouault who was one of the reviewer

proposed me to go further in the study of the asymptotic behavior of fragments having an "exponential decay".

The goal of the paper written with Alain Rouault (KR11) is to present results of the second type, i.e. asymptotic study of presence of abnormally large fragments. Let us first explain known results of the first type - exponential growth - and fix some notation.

For $x \in (0,1)$ let $I_x(t)$ be the component of the interval fragmentation $F(t)$ which contains x, and let $|I_x(t)|$ be its length. Jean Bertoin showed in (Ber01) that if V is an uniform random variable on $[0,1]$ independent of the fragmentation, then

$$\xi(t) := -\log |I_V(t)|$$

is a subordinator whose distribution is entirely determined by the characteristics of the fragmentation. Its Laplace exponent is given by

$$\mathbb{E}\mathrm{e}^{-q\xi(t)} = \mathrm{e}^{-t\kappa(q)}$$

where κ is the concave positive function :

$$\kappa(q) := \int_{\mathcal{U}} \left(1 - \sum_{j=1}^{\infty} u_j^{q+1}\right) \nu(dU) \quad \forall q > \underline{p} \tag{0.2}$$

and $\underline{p}$ is the smallest real number for which κ remains finite :

$$\underline{p} := \inf \left\{ p \in \mathbb{R} : \int_{\mathcal{U}} \sum_{j=2}^{\infty} u_j^{p+1} \nu(dU) < \infty \right\} .$$

The strong law of large numbers tells us that a.s. $\lim_{t\to\infty} \xi(t)/t \to \kappa'(0) =: v_{typ}$, so that a.s.

$$\lim_{t\to\infty} -t^{-1} \log |I_V(t)| = v_{typ} .$$

In fact, there is an interval $(v_{\min}, v_{\max})$ straddling v_{typ} of effective asymptotic exponential rates of decreasing of fragments which we describe now. Let $\overline{p}$ be the unique solution of the equation

$$\kappa(q) = (q+1)\kappa'(q), \quad q > \underline{p} .$$

We define $v_{\min} := \kappa'(\overline{p})$ and $v_{\max} := \kappa'(\underline{p}^+)$.

In all the following, we fix a and b such that $0 < a < 1 < b$.

If we set

$$\widetilde{G}_{v,a,b}(t) = \{I_x(t) : x \in (0,1) \text{ and } ae^{-vt} < |I_x(t)| < be^{-vt}\}$$

then it is known ((Ber03), (BR05) Corollary 3) that the asymptotic growth of $\widetilde{G}_{v,a,b}(t)$

is ruled by the concave function C defined for $v < v_{\max}$ by

$$C(v) = \inf_{q>\underline{p}} \left((q+1)v - \kappa(q)\right) , \tag{0.3}$$

or

$$C(v) = (\Upsilon_v + 1)v - \kappa(\Upsilon_v) \quad , \quad \kappa'(\Upsilon_v) = v \, . \tag{0.4}$$

More precisely we have:

• for $v \in (v_{\min}, v_{\max})$, $C(v)$ is strictly positive and

$$\lim_{t\to\infty} \; t^{-1} \log \sharp \widetilde{G}_{v,a,b}(t) = C(v) \text{ a.s.} \tag{0.5}$$

• for $v \leq v_{\min}$, $C(v)$ is strictly negative and the set $\widetilde{G}_{v,a,b}(t)$ is a.s. empty for t large enough.

Let us stress that $C(v)$ depends only on v and not on a, b.

In the sequel, the latter setting referred to *classical*.

In (Kre08) done during my PhD, I studied the more constrained set

$$G_{v,a,b}(t) = \{I_x(t) : x \in (0,1) \text{ and } ae^{-vs} < |I_x(s)| < be^{-vs} \;\; \forall\, s \leq t\} \, ,$$

and proved a result of the same kind. In particular, Proposition 3 (p.908) (Kre08) shows us that it exists a positive number $\rho(v,a,b)$ depending upon v, a, b such that

• for $v > \rho(v,a,b)$, conditionally on $\{\inf\{t : G_{v,a,b}(t) \neq \emptyset\} = \infty\}$

$$\lim_{t\to\infty} \; t^{-1} \log \sharp G_{v,a,b}(t) = v - \rho(v,a,b) \, , \text{ a.s.} \tag{0.6}$$

• for $v \leq \rho(v,a,b)$, $\lim_{t\to\infty} \sharp G_{v,a,b}(t) = 0$ a.s..

This result holds under the following Assumption α, which comes from (Lam00) and (Ber97), and ensures the absolute continuity of the marginals of the underlying Lévy process.

Assumption α**.** The image ν_1 of the measure ν by the mapping $U \mapsto u_1$ satisfies

$$\nu_1^{ac}((1-\epsilon, 1]) = \infty \quad \text{for any} \quad \epsilon > 0 \, , \tag{0.7}$$

where ν_1^{ac} is the absolutely continuous part of ν_1.

If we refer to the above general comments on branching models, we can say that the above assertions (0.5) and (0.6) are of the first type. Our main aim here is to present results of the second type.

For the classical model, an assumption is needed. A fragmentation is called r-lattice with $r > 0$, if $(\xi(t))_{t\geq 0}$ is a compound Poisson process whose jump measure has a support carried by a discrete subgroup of $\mathbb{R}$ and r is the mesh. If there is no such r, the fragmentation is called non-lattice.

Theorem 0.1. *(BM05b) Under either the fragmentation is non-lattice, or it is r-lattice and a, b satisfy $b > ae^r$, if $v < v_{\min}$, then*

$$\lim_{t\to\infty} t^{-1} \log \mathbb{P}(\widetilde{G}_{v,a,b}(t) \neq \emptyset) = C(v). \tag{0.8}$$

In (BM05b), the result of Theorem 5 is more precise since it gives sharp (i.e. non logarithmic) estimates of the latter probability.

For the more constrained set $G_{v,a,b}(t)$, the corresponding result is the following.

Theorem 0.2. *Under Assumption α, if $v - \rho(v, a, b) < 0$, then*

$$\lim_{t\to\infty} t^{-1} \log \mathbb{P}(G_{v,a,b}(t) \neq \emptyset) = v - \rho(v, a, b). \tag{0.9}$$

Let us remark that since $G_{v,a,b} \subset \tilde{G}_{v,a,b}$, the limits (0.5), (0.8), (0.6) and (0.9) are comparable. In fact, we have the following general result

Proposition 0.1. *Under Assumption α, for all $v < v_{max}$*

$$C(v) > v - \rho(v, a, b). \tag{0.10}$$

Theorem 0.2 is the main result of the paper (KR11). The crucial tool consists in first introducing additive martingales to make a change of probability and then using a decomposition according to the spine method.

I decided I will not develop in this part how the many-to-one formula solved the problem because I would have to go in too many details to give the link between partition fragmentations and interval fragmentations. Therefore, a lot of unnecessary notations and definitions would have been introduced in the following. Instead, I refer to the article (KR11) for more details on this subject. On the other hand I will detail further the paper I wrote with Marc Hoffmann on the estimation of the Lévy measure associated to the subordinator, which is the marked fragment. It will indeed allow to understand the link, there will be then with the study of the bacteria and the associated marked bacteria.

0.3 Statistical analysis of self-similar conservative fragmentation chains

0.3.1 The fragmentation chains

In a first time after my thesis I finished two papers that I had started during my thesis. The first paper (FKM10) was about a problem from the mining industry. During my thesis I went to Chile for 6 months. Where a collaboration with Joaquim Fontbona and Servet Martinez started. We worked on minimizing the energetic cost of reducing the size of a fragment of mass x to fragments of size less than or

equal to ϵ. For this purpose it is assumed that the crushing of the ore, using devices can be modeled by fragmentation. Two devices are considered. We represent them as different stochastic fragmentation processes. We followed the self-similar energy model introduced by Jean Bertoin and Servet Martinez (BM05), to calculate the average energy required to reach a ϵ size with this two-device procedure. Then we asymptotically compare, when ϵ goes to 0 or 1, its energy requirement with that of individual fragmentation processes. In particular, it is interesting to observe that for certain ranges of parameters of the fragmentation processes and their energy cost functions, the consecutive use of two devices can be asymptotically more efficient than the use of each of them separately, or vice versa.

One of the key tools for this is to use a many-to-one formula and the properties of the Lévy process associated with the marked fragment.

To do this, we will first introduce the notions and definitions related to mass fragmentation. We can notice that if we consider the size reordered by decreasing order of the interval fragmentation introduced in the previous section, we will obtain a mass fragmentation.

Let $X = \big(X(t), t \geq 0\big)$ be a fragmentation chain with state space

$$\mathcal{S}^{\downarrow} := \big\{\mathbf{s} = (s_1, s_2, \ldots),\ s_1 \geq s_2 \geq \ldots \geq 0,\ \sum_{i=1}^{\infty} s_i \leq 1\big\}.$$

We assume that X has parameter of self-similarity $\alpha \geq 0$. To ensure that everything is well-defined, see *e.g.* Jean Bertoin (Ber06), the following mild assumptions on the dislocation measure $\nu(d\mathbf{s})$ of X are in force in the rest of this section:

Assumption A. *We have* $\nu\big((1,0,\ldots)\big) = 0$ *and* $\nu\big(s_1 \in (0,1)\big) > 0$. *Moreover, for every* $\varepsilon > 0$*:* $\int_{\mathcal{S}^{\downarrow}} \sum_{i=1}^{\infty} 1_{\{s_i > \varepsilon\}} \nu(d\mathbf{s}) < \infty$.

We denote by $\mathbb{P}_m$ the law of X started from the initial configuration $(m, 0, \ldots)$ with $m \in (0,1]$. Under $\mathbb{P}_m$, X is a Markov process and its evolution can be described as follows: a fragment with size x lives for an exponential time with parameter $x^{\alpha}\nu(\mathcal{S}^{\downarrow})$ and then splits and gives rize to a family of smaller fragments distributed as $x\xi$, where ξ is distributed according to $\nu(\cdot)/\nu(\mathcal{S}^{\downarrow})$. Under $\mathbb{P}_m$, the law of X is entirely determined by α and $\nu(\cdot)$.

We will repeatedly use the representation of fragmentation chains as random infinite marked trees. Let

$$\mathcal{V} := \bigcup_{n=0}^{\infty} \mathbb{N}^n$$

denote the infinite genealogical tree (with $\mathbb{N}^0 := \{\emptyset\}$) associated to X as follows: to each node $u \in \mathcal{V}$, we set a mark

$$(\xi_u, a_u, \zeta_u), \tag{0.11}$$

where ξ_u is the size of the fragment labelled by u, a_u is its birthtime and ζ_u is its lifetime. We have the following identity between point measures on $(0, +\infty)$:

$$\sum_{i=1}^{\infty} 1_{\left\{X_i(t)>0\right\}} \delta_{X_i(t)} = \sum_{u\in\mathcal{V}} 1_{\left\{t\in[a_u, a_u+\zeta_u)\right\}} \delta_{\xi_u}, \quad t \geq 0, \tag{0.12}$$

with $X(t) = \big(X_1(t), X_2(t), \dots\big)$, and where δ_x denotes the Dirac mass at x. Finally, X has the following branching property: for every fragment $\mathbf{s} = (s_1, \dots) \in \mathcal{S}^{\downarrow}$ and every $t \geq 0$, the distribution of $X(t)$ given $X(0) = \mathbf{s}$ is the same as the decreasing rearrangement of the terms of independent random sequences $X^{(1)}(t)$, $X^{(2)}(t), \dots$ where, for each i, $X^{(i)}(t)$ is distributed as $X(t)$ under $\mathbb{P}_{s_i}$.

If we can observe the whole fragmentation process as time grows, then the statistical problem is somehow degenerate. In this paper, we postulate a more realistic observation scheme, based on mining industry, where the goal is to separate metal from non valued components in large mineral blocks by a series of operations such as of blasting, crushing and grinding operations. The context is exactly the same as the one used previously in the paper with Joaquim Fontbona and Servet Martinez (FKM10).

0.3.2 The empirical measure

If we keep in mind the motivation of mineral crushing, we consider the fragmentation under $\mathbb{P} := \mathbb{P}_1$, initiated with an unique block of size $m = 1$ and we observe the process stopped at the time when all the fragments become smaller than some given threshold $\varepsilon > 0$, so we have data ξ_u, for every $u \in \mathcal{V}_\varepsilon$, with

$$\mathcal{V}_\varepsilon := \big\{u \in \mathcal{U}, \ \ \xi_{u-} \geq \varepsilon, \ \xi_u < \varepsilon\big\},$$

where we denote by $u-$ the parent of the fragment labelled by u. We will further assume that the total mass of the fragments remains constant through time:

Assumption B. *(Conservative property). We have:* $\nu\big(\sum_{i=1}^{\infty} s_i = 1\big) = 1$.

We next consider the empirical measure integrated against a test function $g(\cdot)$

$$\mathcal{E}_\varepsilon(g) := \sum_{u\in\mathcal{V}_\varepsilon} \xi_u \, g(\xi_u/\varepsilon).$$

Indeed, under Assumption B, we have

$$\sum_{u\in\mathcal{V}_\varepsilon} \xi_u = 1 \ \ \mathbb{P} - \text{almost surely}, \tag{0.13}$$

so $\mathcal{E}_\varepsilon(g)$ appears as a weighted empirical version of $g(\cdot)$. Notice that the empirical

measure $\mathcal{E}_\varepsilon$ only depends on the size of the fragmentation and is thus independent of the self-similarity parameter α. We consider α as a nuisance parameter. Jean Bertoin and Servet Martinez show in Corollary 1 of (BM05) that under mild assumptions on $\nu(\cdot)$, the random variable measure $\mathcal{E}_\varepsilon(g)$ converges to

$$\mathcal{E}(g) := \frac{1}{c(\nu)} \int_0^1 \frac{g(a)}{a} \int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} s_i \, 1_{\{s_i<a\}} \nu(d\mathbf{s}) da$$

in $L^1(\mathbb{P})$, as $\varepsilon \to 0$, with $c(\nu) = -\int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} s_i \log s_i \, \nu(d\mathbf{s})$, tacitly assumed to be well-defined. This suggests a strategy for recovering information about $\nu(\cdot)$ by picking suitable test functions $g(\cdot)$.

In (HK11) we give a rate of convergence for the empirical measure $\mathcal{E}_\varepsilon$ toward its limit, extending former results (under more stringent assumptions) of Jean Bertoin and Servet Martinez (BM05). The rate is of the form $\varepsilon^{1/2-\ell(\pi)}$, where $\ell(\pi) > 0$ can arbitrarily be made small under suitable exponential moment conditions for π. We additionally consider the more realistic framework of observations with limited accuracy, where each fragment is actually known up to a systematical stochastic error of order $\sigma \ll \varepsilon$: $X_{\varepsilon,\sigma} := \big(\xi_u^{(\sigma)},\ u \in \mathcal{V}_{\varepsilon,\sigma}\big)$ with

$$\mathcal{V}_{\varepsilon,\sigma} := \big\{u \in \mathcal{U},\ \xi_{u-}^{(\sigma)} \geq \varepsilon,\ \xi_u^{(\sigma)} < \varepsilon\big\},$$

and

$$\xi_u^{(\sigma)} := \xi_u + \sigma U_u. \tag{0.14}$$

The random variables $(U_u, u \in \mathcal{V})$ are identically distributed, centred and are almost surely bounded in absolute value by 1. They account for a systematic experimental microstructure noise in the measurement of X_ε, independent of X_ε. The noise level $0 \leq \sigma = \sigma(\varepsilon) \ll \varepsilon$ is assumed to be known and represents the accuracy level of the statistician.

The observations $\xi_u + \sigma U_u$ are further discarded below a threshold $\sigma \leq t_\varepsilon \leq \varepsilon$ beyond which they become irrelevant, leading to the modified empirical measure

$$\mathcal{E}_{\varepsilon,\sigma}(g) := \sum_{u \in \mathcal{V}_{\varepsilon,\sigma}} 1_{\{\xi_u^{(\sigma)} \geq t_\varepsilon\}} \, \xi_u^{(\sigma)} \, g\big(\xi_u^{(\sigma)}/\varepsilon\big).$$

In the sequel, we take $t_\varepsilon = \gamma_0 \varepsilon$ for some (arbitrary) $0 < \gamma_0 < 1$ and assume further that $\sigma \leq \frac{1}{2} t_\varepsilon$.

Assumptions A and B are in force. At this stage, we can relate $\mathcal{E}(g)$ to a more appropriate quantity by means of the so-called *tagged fragment* approach.

0.3.3 The randomly tagged fragment

Let us first consider the homogeneous case $\alpha = 0$. Assume we can randomly "tag" a point –according to an uniform distribution– on the initial fragment and imagine we can follow the evolution of the fragment that contains this point.

Let us denote by $\big(\chi(t), t \geq 0\big)$ the process of the size of the fragment that contains the randomly chosen point. This fragment is a typical observation in our data set X_ε, and it appears at time

$$T_\varepsilon := \inf\big\{t \geq 0, \;\; \chi(t) < \varepsilon\big\}.$$

Jean Bertoin (Ber06) shows that the process $\zeta(t) := -\log \chi(t)$ is a subordinator, with Lévy measure:

$$\pi(dx) := e^{-x} \sum_{i=1}^{\infty} \nu(-\log s_i \in dx). \tag{0.15}$$

We can anticipate that the information we get from X_ε is actually information about the Lévy measure $\pi(dx)$ of $\zeta(t)$ obtained via $\zeta(T_\varepsilon)$. The dislocation measure $\nu(d\mathbf{s})$ and $\pi(dx)$ are related by (0.15) which reads

$$\int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} s_i f(s_i) \nu(d\mathbf{s}) = \int_{(0,+\infty)} f(e^{-x}) \pi(dx), \tag{0.16}$$

for any suitable $f(\cdot) : [0,1] \to [0,+\infty)$. In particular, by Assumption B and the fact that $\nu(\mathcal{S}^\downarrow) = 1$, $\pi(dx)$ is a probability measure hence $\zeta(t)$ is a compound Poisson process. Informally, a typical observation takes the form $\zeta(T_\varepsilon)$, which is the value of a subordinator with Lévy measure $\pi(dx)$ at its first passage time strictly above $-\log \varepsilon$. The case $\alpha \neq 0$ is a bit more involved and reduces to the homogeneous case by a time change.

In terms of the limit of the empirical measure $\mathcal{E}_\varepsilon(g)$, we equivalently have

$$\mathcal{E}(g) = \frac{1}{c(\pi)} \int_0^1 \frac{g(a)}{a} \pi(-\log a, +\infty)\, da = \frac{1}{c(\pi)} \int_0^{+\infty} g(e^{-x}) \pi(x, +\infty)\, dx,$$

with $c(\pi) = \int_{(0,+\infty)} x\, \pi(dx)$, where both representations will be found useful. Except in the binary case, knowledge of $\pi(\cdot)$ does not, in general, allow us to recover $\nu(\cdot)$.

More precisely, the many-to-one formula described in the following lemma allows us to see the link between the evolution of the fragmentation and of the marked fragment. Thus, with the help of the marked fragment and thus of the associated subordinator, we can obtain the information we need about the fragmentation.

Lemma 0.1. *Let $f(\cdot) : [0, +\infty) \to [0, +\infty)$. Then for $\varepsilon > 0$,*

$$\mathbb{E}\Big[\sum_{v \in \mathcal{U}_\varepsilon} \xi_v \, f(\xi_v) \Big] = \mathbb{E}^\star \big[f\big(\chi(T_\varepsilon)\big) \big], \tag{0.17}$$

where $\chi(t) = \exp\big(-\zeta(t)\big)$ and $\big(\zeta(t), t \geq 0\big)$ is a subordinator with Lévy measure $\pi(\cdot)$ defined on an appropriate probability space $(\Omega^\star, \mathbb{P}^\star)$, and

$$T_\varepsilon := \inf\big\{ t \geq 0, \ \zeta(t) > -\log \varepsilon \big\}.$$

It is important to specify here that, the process of the associated many-to-one is explicit according to the starting process and that one does not have only the existence in law of a process which is present in the many-to-one formula, as it is often the case (Clo17; BT11).

We can also note that in the demonstration we will need a special case of Sgibnev's result (Sgi02) on uniform rates of convergence in the key renewal theorem.

For statistical purposes, our main tool is the empirical measure $\mathcal{E}_\varepsilon$ of the size of fragments when they reach a size smaller than a threshold ε in the limit $\varepsilon \to 0$. We highlight the fact that $\mathcal{E}_\varepsilon$ captures information about the dislocation measure through the Lévy measure π of a randomly tagged fragment associated to the fragmentation process.

0.3.4 The statistical inference

Definition 0.1. *For $\kappa > 0$, we say that a non-lattice probability measure $\pi(dx)$ defined on $[0, +\infty)$ belongs to $\Pi(\kappa)$ if*

$$\int_{[0,+\infty)} e^{\kappa x} \, \pi(dx) < +\infty,$$

Assumption C. *The probability $\pi(dx)$ is absolutely continuous w.r.t. the Lebesgue measure: $\pi(dx) = \pi(x)dx$. Moreover, its density function $x \rightsquigarrow \pi(x)$ is continuous on $(0, +\infty)$ and satisfies $\limsup_{x \to +\infty} e^{\vartheta x} \pi(x) < +\infty$ for some $\vartheta \geq 1$.*

We distinguish two cases: the *parametric case*, where we estimate a linear functional of $\pi(\cdot)$ of the form

$$m_k(\pi) := \int_0^{+\infty} x^k \pi(x) dx, \quad k = 1, 2, \ldots$$

and the *non-parametric case*, where we estimate the function $x \rightsquigarrow \pi(x)$ pointwise. In that latter case, it will prove convenient to assess the local smoothness properties

of $\pi(\cdot)$ on a logarithmic scale. Henceforth, we consider the mapping

$$a \rightsquigarrow \beta(a) := a^{-1}\pi(-\log a), \quad a \in (0,1). \tag{0.18}$$

In the non-parametric case, we estimate $\beta(a)$ for every $a \in (0,1)$.

The parametric case

Preliminaries. For $k \geq 1$, we estimate

$$m_k(\pi) := \int_0^{+\infty} x^k \, \pi(x)dx = \int_0^1 \log(1/a)^k \beta(a)da$$

by the correspondence (0.18) and implicitly assumed it is well-defined. We first focus on the case $k = 1$. Pick a sufficiently smooth test function $f(\cdot) : [0,1] \to \mathbb{R}$ such that $f(1) = 0$ and let $g(a) := -af'(a)$. Plainly

$$\begin{aligned}\mathcal{E}(g) &= \frac{1}{c(\pi)} \int_0^1 \frac{g(a)}{a} \pi(-\log a, +\infty)\, da \\ &= -\frac{1}{m_1(\pi)} \int_0^1 f'(a) \int_0^a \beta(u)du\, da = \frac{1}{m_1(\pi)} \int_0^1 f(a)\beta(a)da. \end{aligned} \tag{0.19}$$

Formally, taking $f(\cdot) \equiv 1$ would identify $1/m_1(\pi)$ since $\beta(\cdot)$ integrates to one, but this choice is forbidden by the boundary condition $f(1) = 0$. We shall then consider instead a family of regular functions which are close to the constant function 1 while it satisfies the boundary condition $f(1) = 0$.

Construction of the approximating functions. Let $f_\gamma : [0,1] \to \mathbb{R}$ with $0 < \gamma < 1$ be a family of smooth functions satisfying the following conditions.

- $f_\gamma(a) = 1$ for $a \leq 1 - \gamma$ and $f_\gamma(1) = 0$.
- $$\sup_{\gamma>0} \left(\|f_\gamma\|_\infty + \gamma \|f'_\gamma\|_\infty + \gamma^2 \|f''_\gamma\|_\infty \right) < +\infty. \tag{0.20}$$
- For some $\delta > 0$, we have

$$\limsup_{a\to 0} \left(a^{-1-\delta} \sup_{\gamma>0} \gamma^{1+\delta} |f_\gamma(1-a)| + a^{-1} \sup_{\gamma>0} \gamma^2 |f'_\gamma(1-a)| \right) < +\infty. \tag{0.21}$$

The family $(f_\gamma, \gamma > 0)$ imitates the behaviour of the target function $f_0(a) = 1$ for $0 \leq a < 1$ and $f_0(1) = 0$ and is close to f_0 as $\gamma \to 0$.

Construction of an estimator. We are now ready to give an estimator of the first moment $m_1(\pi)$ of π, and more generally, of any moment $m_k(\pi)$, $k \geq 1$. For a

parametrization $\gamma := \gamma_\varepsilon \to 0$ to be specified later, we set

$$g_{\gamma_\varepsilon}(a) := -af'_{\gamma_\varepsilon}(a), \quad a \in (0,1).$$

By Theorem 1 in (HK11), we exhibit explicit rates in the convergence of $\mathcal{E}_{\sigma,\varepsilon}(g_{\gamma_\varepsilon})$ to $\mathcal{E}(g_{\gamma_\varepsilon})$ which in turn is equal to $m_1(\pi)^{-1}\int_0^1 f_{\gamma_\varepsilon}(a)\beta(a)da$ by (0.19). Since $f_{\gamma_\varepsilon} \approx 1$ and $\beta(\cdot)$ is a density function, by appropriate regularity assumptions on π, we may further expect this last quantity to be close to $1/m_1(\pi)$. We therefore set

$$\widehat{m}_{1,\varepsilon} := \frac{1}{\mathcal{E}_{\varepsilon,\sigma}\big(g_{\gamma_\varepsilon}\big)} \tag{0.22}$$

for an estimator of $m_1(\pi)$. More generally, for $k > 1$, we define successive moment estimators as follows. Set $h_{\gamma_\varepsilon}(a) := f_{\gamma_\varepsilon}(1-a)\log(1/a)^k$ and $\widetilde{g}_{\gamma_\varepsilon}(a) := -ah'_{\gamma_\varepsilon}(a)$. The same heuristics as before lead to the following estimator

$$\widehat{m}_{k,\varepsilon} := \frac{\mathcal{E}_{\varepsilon,\sigma}\big(\widetilde{g}_{\gamma_\varepsilon}\big)}{\mathcal{E}_{\varepsilon,\sigma}\big(g_{\gamma_\varepsilon}\big)}.$$

In the parametric case (Theorem 2.3), we establish that the best achievable rate is $\varepsilon^{1/2}$ in the particular case of binary fragmentations, where a particle splits exactly in two blocks at each step.

The non-parametric case

Definition 0.2. *For $\kappa > 0$, we say that the probability $\pi(\cdot)$ belongs to the class $\mathcal{R}(\kappa)$ if*

$$\limsup_{x\to 0} x^{-\kappa+1}\pi(x) < +\infty$$

appended with $\mathcal{R}(\infty) := \bigcap_{\kappa>0}\mathcal{R}(\kappa)$.

Given $s > 0$, we say that $\beta(\cdot)$ belongs to the Hölder class $\Sigma(s)$ if there exists a constant $c > 0$ such that

$$\big|\beta^{(n)}(y) - \beta^{(n)}(x)\big| \le c|y-x|^{\{s\}},$$

with $s = n + \{s\}$, where n is a non-negative integer and $\{s\} \in (0,1]$. We also need to relate $\beta(\cdot)$ to the decay of its corresponding Lévy measure $\pi(\cdot)$. Abusing again notation, we identify $\Pi(\kappa)$ with the set of $\beta(\cdot)$ such that $e^x\beta(e^{-x})dx \in \Pi(\kappa)$, thanks to the inverse of (0.18). Likewise for $\mathcal{R}(\kappa)$.

We construct an estimator of $\beta(\cdot)$ in the same way as for the parametric case: for $a \in (0,1)$ and a normalizing factor $0 < \gamma_\varepsilon \to 0$, set

$$\varphi_{\gamma_\varepsilon,a}(x) := \gamma_\varepsilon^{-1}\varphi\big((x-a)/\gamma_\varepsilon\big),$$

where φ is a smooth function with support in $(0,1)$ that satisfies the following oscillating property: for some integer $N \geq 1$,

$$\int_0^1 \varphi(a)da = 1, \quad \int_0^1 a^k \varphi(a)da = 0, \;\; k = 1, \ldots, N. \tag{0.23}$$

So the function $\varphi_{\gamma_\varepsilon,a}$ plays the role of a kernel centred around a. Set

$$h_{a,\varepsilon}(x) = -x\,\varphi'_{\gamma_\varepsilon,a}(x), \quad x \in (0,1).$$

We have

$$\mathcal{E}(h_{a,\varepsilon}) = \frac{1}{m_1(\pi)} \int_0^1 \varphi_{\gamma_\varepsilon,a}(x)\beta(x)dx$$

by (0.19). By letting $h_\varepsilon \to 0$ with an appropriate rate as $\varepsilon \to 0$, we expect this term to be close to $\beta(a)/m_1(\pi)$. We can eventually get rid of the denominator by our preliminary estimator $\widehat{m}_{1,\varepsilon}$. Our non-parametric estimator of $\beta(a)$ takes thus the form

$$\widehat{\beta}_\varepsilon(a) := \widehat{m}_{1,\varepsilon}\, \mathcal{E}_{\varepsilon,\sigma}\big(h_{a,\varepsilon}\big), \quad a \in (0,1),$$

where $\widehat{m}_{1,\varepsilon}$ is the estimator of $m_1(\pi)$ defined in (0.22).

Theorem 0.3. *Work under Assumptions A, B and C. Let $\kappa_1 \geq 4$ and $\kappa_2 > 1$. For any $1 \leq \mu < \kappa_1$, let $\widehat{\beta}_\varepsilon(\cdot)$ be specified by $\gamma_\varepsilon := \varepsilon^{\mu/(\mu+1)(2s+3)}$. For every $a \in (0,1)$, the family*

$$\big(\varepsilon^{-\mu/(\mu+1)}\big)^{s/(2s+3)}\big(\widehat{\beta}_\varepsilon(a) - \beta(a)\big)$$

is tight, as soon as

$$\beta \in \Sigma(s) \cap \Pi(\kappa_1) \cap \mathcal{R}(\kappa_2)$$

for $0 < s < \min\{N, 3\kappa_2\}$ and $\sigma\varepsilon^{-3}$ remains bounded.

0.4 The use of the tools from the fragmentation

The marked fragment and the many-to-one formula were keytools for the paper with Marc Hofmann (HK11) about estimation on fragmentation and also in the paper with Alain Rouault (KR11) and Joaquim Fontbona et Servet Martinez (FKM10).

It gave me the idea of using it in an other context of the growth of bacteria. It will be the first chapter of this HDR. This new direction has been an important part of my research since my PhD. Thanks to a lot of interesting discussions about the behavior of a growing bacteria with Lydia Robert who is a biologist with incredible knowledge in probability, a new world was open to me. The first question was to try to understand the factor that plays a primordial role in the division of the bacteria. Lots of different factors (size, age, elongation,...) can intervene and of course there can be combinations of several factors. To answer this question, a collaboration

with Marie Doumic, Marc Hoffmann and Lydia Robert began which gave rise to two publications (DHK$^+$14; DHKR15). What is the factor which determinates the fact that the bacteria decides to divide? The two most common possibilities were the fact that the bacteria was too old or was too tall. In the first paper, we constructed two models, the first was based on length as key element of the division and the second one was based on age. The model taking into account the age led to simulations where there would be an accumulation of bacteria of small sizes, this did not correspond at all to the experimental data which existed on this subject. In a second paper (DHKR15), we decided to explore the model where the length would determinate the division. The first difficulty was to build a model which allowed the growth of bacteria to be modeled. As I was familiar with fragmentations, this gave me the idea to make a construction analogous to that of fragmentations using an indexed tree of the classical Ulam-Harris-Neveu notation. The strength of this tool is that we could put as many parameters as we wanted in the nodes of our tree and then build our process. Like in the fragmentation case it exists a many-to-one formula which makes an explicit link between the whole process and a "marked" bacterium. Thus we will simplify the study of a process with value in $\mathbb{R}_+^{\mathbb{N}}$ to a process with value in $\mathbb{R}_+$. In addition, this marked process turns out to be a PDMP.

Then, with Bernard Delyon, Benoîte de Saporta, and Lydia Robert we were interested in understanding how the growth rate parameter evolves particulary if there was a difference between these old bacteria and young bacteria. A young bacteria had inherited from youngest part of their mother. A old bacteria had inherited the oldest part. This difference between the kinds of bacteria is an issue that interested a lot Lydia Robert and we decided to look at it for growth rates which gave birth to the (DSKR18) paper. At this moment with Bertrand Cloez, Benoîte de Saporta and Tristan Roget we were interested in looking at the influence of this difference between the old pole and the young pole and how this also intervenes in the division rate and in the factor of division in the size of the bacterium after its division. We also tried to compare the division model of the bacteria because of size using a new model, the Adder model. It is a work in progress.

Thanks to the PDMP, which appeared for the modeling of the growth of the size of a marked bacterium, I started to get interested in PDMPs. In a first paper (Kre16), I was interested in a particular class of PDMP, which includes the case of the marked bacteria. Thus, I obtained an upper bound for the speed of convergence of my estimator. Then, in a second step I wanted to go further in this study by making the estimator adaptive and by showing that the speed of convergence what min-max. For this, I started a collaboration with Emeline Schmisser who is a specialist in this kind of tool. For that we had to change our estimator, in (Kre16) I used a kernel estimator and there we took an estimator by projection. We also succeeded in showing that the estimator did indeed have a min-max speed and we also generalized the class of PDMP studied. The results can be found in the publication (KS21). This part will be the second part of this HDR.

As I studied problems coming from biological issue such as neuronal process, where the interaction between the element are more much complicated as the one for the bacterium. We will need other kinds of tools. In fact this process was no longer a branching process and the many-to-one formula was lost. This will lead to a publication with Eva Löcherbach and Pierre Hodara (whom I co-supervised with her) (HKL18). This direction will be detailed in the last chapter.

0.5 List of publications

I have published 11 works, in international peer-reviewed journals, as well as a book chapter and a proceedings. They are all available on my web page `https://perso.univ-rennes1.fr/nathalie.krell/Recherche.html`

List of publications

[DSKR18] Bernard Delyon, Benoîte de Saporta, Nathalie Krell, Lydia Robert.
Investigation of asymmetry in E. coli growth rate.
Case Studies in Business, Industry and Government Statistics, Société Française de Statistique , 7 (1), pp.1-13, 2018.

[DHKARR14] Marie Doumic, Marc Hoffmann, Nathalie Krell, Stéphane Aymerich, Jérome Robert, and Lydia Robert.
Division control in escherichia coli is based on a size-sensing rather than timing mechanism.
BMC Biology, 12:17, 2014.

[DHKR15] Marie Doumic, Marc Hoffmann, Nathalie Krell, and Lydia Robert.
Statistical estimation of a growth-fragmentation model observed on a genealogical tree.
Bernoulli, 21(3):1760–1799, 2015.

[FKM10] Joaquín Fontbona, Nathalie Krell, and Servet Martínez.
Energy efficiency of consecutive fragmentation processes.
J. Appl. Probab., 47(2):543–561, 2010.

[HK11] Marc Hoffmann and Nathalie Krell.
Statistical analysis of self-similar conservative fragmentation chains.
Bernoulli, 17(1):395–423, 2011.

[HKL18] Pierre Hodara, Nathalie Krell and Eva Löcherbach.
Non-parametric estimation of the spiking rate in systems of interacting neurons.
Stat. Inference Stoch. Process., 21(1):81–111, 2018.

[KR11] Nathalie Krell and Alain Rouault. Martingales and rates of presence in homogeneous fragmentations.
Stochastic Process. Appl., 121(1):135–154, 2011.

[Kre08] Nathalie Krell.
Multifractal spectra and precise rates of decay in homogeneous fragmentations.
Stochastic Process. Appl., 118(6):897–916, 2008.

[Kre09] Nathalie Krell.
Self-similar branching Markov chains.
Séminaire de Probabilités XLII, 261-280, Lecture Notes in Math. 979, Springer, Berlin, 2009.

[Kre16] Nathalie Krell.
Statistical estimation of jump rates for a piecewise deterministic Markov processes with deterministic increasing motion and jump mechanism.
ESAIM: PS 20 196-216, 2016.

[KS21] Nathalie Krell and Émeline Schmisser.
Nonparametric estimation of jump rates for a specific class of piecewise deterministic Markov processes.
Bernoulli, 27(4):2362–2388, 2021.

Book chapter

[HKL18b] Pierre Hodara, Nathalie Krell and Eva Löcherbach.
Regularity of the invariant measure and non-parametric estimation of the jump rate.
Statistical inference for piecewise-deterministic Markov processes, Romain Azais ; Florian Bouguet. Wiley, Chapitre 2 - p. 39-64, 2018.

Proceedings

[ABGKZ14] Romain Azaïs, Jean-Baptiste Bardet, Alexandre Génadot, Nathalie Krell, and Pierre-André Zitt.
Piecewise deterministic Markov process-recent results.
Journées MAS 2012, volume 44 of *ESAIM Proc.*, pages 276–290. EDP Sci., Les Ulis, 2014.

Chapter 1

Cell Division structured Models

1.1 Why do bacteria divide?

1.1.1 Two dataset

Following the discussion with Lydia Robert the goal was to try to model the growth of *Escherichia coli* (denoted *E. coli*) bacteria.

The reported results in the modeling of the growth of bacteria were obtained from the analysis of two different datasets which were obtained through microscopic time-lapse imaging of *E. coli* single-cells growing in rich medium, by Eric Stewart et al. (SMPT05) and Robert Wang, et al. (WRPDTWS10). We will refer to these two models. Stewart et al. followed *E. coli* single cells growing into microcolonies on LB-agarose pads at 300C. The cell lineages were followed and the length of each cell in the microcolony was measured every 2 minutes. In the data from Robert Wang, et al., the cells were grown in LB medium at 370C in a micro fuid setup (WRPDTWS10) and the length of the cells was measured every minute. Due to the micro fuid device structure, only one daughter cell is followed at each division (data s_i: sparse tree), in contrast to the experiment of Eric Stewart et al. where all the individuals of a genealogical tree are followed (data f_i: full tree). It is worth noting that this different structure of the data f_i and s_i leads to different PDE models, and the statistical analysis was adapted to each situation. From each dataset (f_i and s_i) we extracted the results of three experiments (experiments f_1; f_2 and f_3 and s_1; s_2 and s_3). Each experiment f_i corresponds to the growth of 6 microcolonies up to 600 cells and each experiment s_i to the growth of bacteria in a hundred microchannels for 40 generations.

Given the accuracy of image analysis, we do not take into account variations of cell width within the population, because they are negligible compared to cell-cycle induced length variations. Thus, in the present study we do not distinguish between length, volume and mass and use the term cell size as a catch-all descriptor.

1.1.2 By what factor do you model the division of the bacteria?

Many organisms coordinate cell growth and division through size control mechanisms: cells must reach a critical size to trigger some cell cycle event. Bacterial division is often assumed to be controlled in this way, but we miss experimental evidence to support this assumption. Theoretical arguments show that size control is required to maintain size homeostasis in the case of exponential growth of individual cells. Nevertheless, if the growth law deviates slightly from exponential in very small cells, homeostasis can be maintained with a simple "timer" triggering division.Therefore, to decide if division control in bacteria relies on a "timer" or "sizer" mechanism requires quantitative comparisons between models and data. This is the question we were interested in in (DHK$^+$14) where we give reference about the result existing before.

The "timer" and "sizer" hypotheses are easily expressed in mathematical terms: two different PDE models are commonly used to describe bacterial growth, using a division rate (i.e. the instantaneous probability of division) depending either on cell age or cell size. In the age-structured model (Age model) the division rate B_1 is a function only of the age a of the cell. The density $n(t;a)$ of cells of age a at time t is given as a solution to the Mc-Kendrick Von Forster equation ((Per07) and references therein).

We have (in a weak sense):

$$\partial_t n(t,a) + \partial_a n(t,a) = -B_1(a)n(t,a), \tag{1.1}$$

with the boundary condition

$$n(t,a=0) = 2\int_0^\infty B_1(a)n(t,a)da.$$

The methods relie on tagged fragment approach ((Ber06), (Haa03)) and many-to-one formula ((Ban09), (BT11), (Kre08) (DHKR15) and (Clo17)) .

In this model, a cell of age a at time t has the probability $B_1(a)dt$ of dividing between time t and $t+dt$.

In the size-structured model (Size model), the division rate B is a function only of the size x of the cell. Assuming that the size of a single cell grows with a rate $v(x)$, the density $n(t;x)$ of cells of size x at time t is given as a solution to the size-structured cell division equation (Per07).

We have (in a weak sense) if we keep the 2 daugthers at each generation:

$$\boxed{\partial_t n(t,x) + \partial_x\big(v(x)\,n(t,x)\big) + B(x)n(t,x) = 4B(2x)n(t,2x).} \tag{1.2}$$

Therefore the mean empirical distribution of $X(t)$ satisfies the deterministic transport-fragmentation equation. In the Size Model, a cell of size x at time t has the probability $B(x)dt$ of dividing between time t and $t + dt$. This model is related to the so-called "sloppy size control" model (Whe82) describing division in S. pombe.

The PDE given by Eq. (1.1) and (1.2) can be embedded into a two-dimensional age-and-size-structured equation (Age and Size Model), describing the temporal evolution of the density n(t; a; x) of cells of age a and size x at time t, with a division rate $B_{a,x}$ a priori depending on both age and size:

$$(\partial_t + \partial_a)n(t, a, x) + \partial_x(v(x)n(t, a, x)) = -B_{a,x}(a)n(t, a, x), \tag{1.3}$$

with the boundary condition

$$n(t, a = 0, x) = 4 \int_0^\infty B_{a,s}(a, x)n(t, a, 2x)da.$$

In the paper (DHK$^+$14) we confront these models with recent data on *E. coli* single cell growth, using a rigorous statistical methodology. In particular, we develop the estimation part with the dependence on height in the different part and for the model of the dependence on age we will be able to refer to the works of Adélaïde Olivier and Marc Hoffmann (HO16). We demonstrate that a size-independent "timer" mechanism for division control, though theoretically possible, is quantitatively incompatible with the data and extremely sensitive to slight variations in the growth law. In contrast, a "sizer" model is robust and fits the data well. In addition, we tested the effect of variability in individual growth rates and noise in septum positioning and found that size control is robust to this phenotypic noise.

In the following figures 1.1, we will show the experimental and reconstructed age-size distributions for representative experiments from Eric Stewart et al.(SMPT05) (f_1) and Robert Wang, et al.(WRPDTWS10) (s_1). In the Figure 1.1 A and B: It is the experimental age-size distribution for a representative experiment f_1 (A) and s_1 (B). The frequency of cells of age a and size s in the population is represented by the color of the figure at the point of coordinate a on the x-axis and s on the y-axis, according to the scale indicated on the right of the figure. C and D: It is the reconstruction of the distributions using the Age model (C: reconstruction of the data f_1 shown in panel A; D: reconstruction of the data s_1 shown in panel B). These reconstructed distributions are obtained by simulations of the Age model using a division rate estimated from the data (C: from f_1, D: from s_1); the growth functions used for the simulations are detailed in the Methods section in (DHK$^+$14). E and F: It is the reconstruction of the distributions using the Size model (E: reconstruction of the data f_1 shown in panel A; F: reconstruction of the data s_1 shown in panel B); these distributions are obtained by simulations of the Size model using a division rate estimated from the data (E: from f_1, F: from s_1) with an exponential growth

function.

Experimental and reconstructed age-size distributions

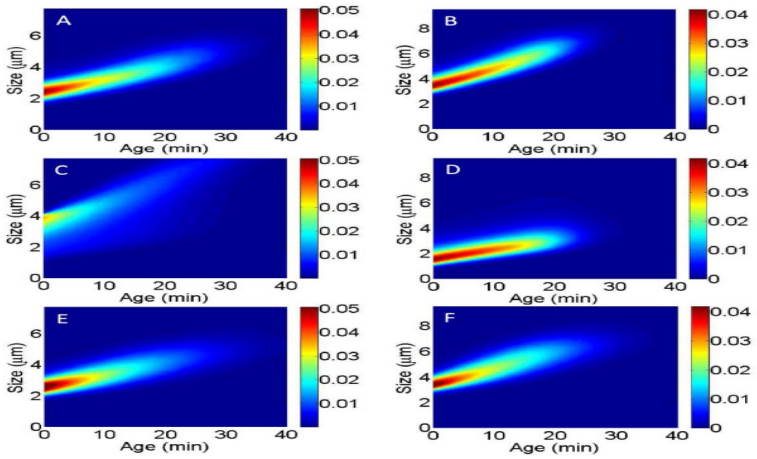


Figure 1.1: Age Size Distribution for all cells - whole tree data

We can conclude that the confrontations between cell cycle models and data usually suffer from a lack of high-quality data and suitable statistical estimation techniques. Here we overcome these limitations by using high precision measurements of tens of thousands of single bacterial cells combined with recent statistical inference methods to estimate the division rate within the models. We, therefore, provide the first precise quantitative assessment of different cell cycle models.

1.2 The estimation of the division rate

After the first work with Lydia Robert, Marc Hoffmann and Marie Doumic (DHK$^+$14), we decided to study more precisely the evolution of the size of the bacteria when the division factor was induced by the size of the bacteria.

The underlying biological problem concerns the division of *Escherichia coli* (*E. coli*) which is a single cell bacterium. Single cells grow and divide to give birth to two daughter cells, that grow and divide and so on. So a colony of cells from a single ancestor is structured as a binary genealogical tree.

E. coli is a rod-shaped bacterium with constant width and elongating length, hence its length (or size) is representative of its biomass or volume. It is commonly admitted, that starting from size x at birth, the bacterium size growths exponentially fast with time at constant rate until its division. This goes back to Monod (1942). More specifically, if T is the age of the bacterium at division, there exists a constant τ, which will be called the growth rate, such that the size of the bacterium at time $0 \leq t \leq T$ equals $xe^{\tau t}$.

The mother cell gives rize to two offsprings, at a rate $B(x)$ that depend on its size x. The two offsprings have initial size $x_1/2$, where x_1 is the size of the mother at

division and start independent growth according to the rate τ and divide according to the rate $B(x)$.

We could also notice that the variability of the growth rate from one cell to another comes from exogeneous and endogeneous factors. Using the dataset of Stewart (f_i) consists of 88 microcolonies followed for a few hours (average time of division is of order 20 minutes): approximately 5 microcolonies are followed everyday, for 16 days. We can see that the variability in growth rate may vary from one day to the next (exogeneous factor). And there are also variability in growth rate may vary within a microcolony if specific factors are transmitted from parents to offsprings. (endogeneous factor).

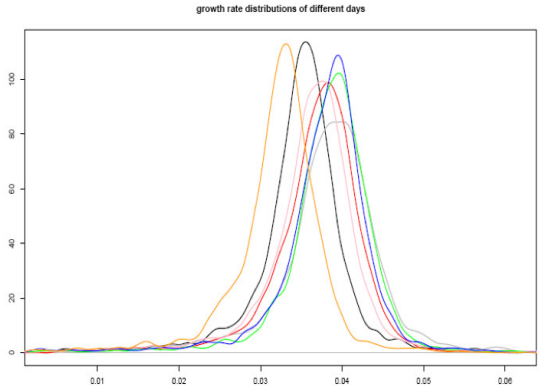


Figure 1.2: one curve = 1 day

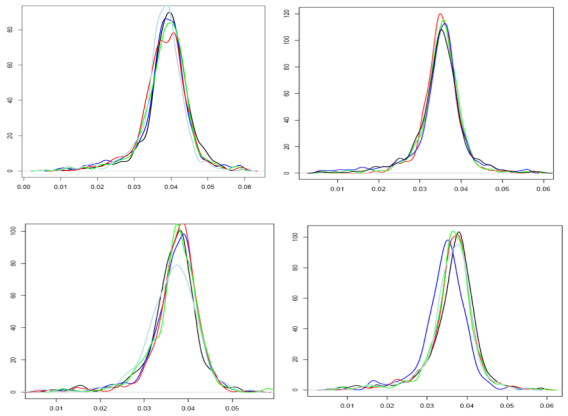


Figure 1.3: one curve = 1 microcolony

1.2.1 The genealogical construction

As we did for the fragmentations we will use a genealogical tree with the Ulam-Harris-Neveu numbering. In each node we will be able to store all information we need to build our process of growth of the size of our bacteria.

We can note that we are in a simpler case because each division of a bacterium gives birth to two bacteria.

Let $\mathcal{U} := \bigcup_{n=0}^{\infty}\{0,1\}^n$ (with $\{0,1\}^0 := \{\emptyset\}$) denotes the infinite binary genealogical tree. Each node $u \in \mathcal{U}$ is identified with a cell of the population and has a mark

$$(\xi_u, b_u, \zeta_u, \tau_u),$$

where ξ_u is the size at birth, τ_u the growth rate, b_u the birthtime and ζ_u the lifetime of u. The evolution $\big(\xi_t^u, t \in [b_u, b_u + \zeta_u)\big)$ of the size of u during its lifetime is governed by

$$\xi_t^u = \xi_u \exp\big(\tau_u(t - b_u)\big) \text{ for } t \in [b_u, b_u + \zeta_u). \tag{1.4}$$

Each cell splits into two offsprings of the same size according to a division rate $B(x)$ for $x \in (0, \infty)$. Equivalently

$$\mathbb{P}\big(\zeta_u \in [t, t + dt] \,\big|\, \zeta_u \geq t, \xi_u = x, \tau_u = v\big) = B\big(x \exp(vt)\big)dt. \tag{1.5}$$

At division, a cell splits into two offsprings of the same size. If u^- denotes the parent of u, we thus have

$$2\,\xi_u = \xi_{u^-} \exp\big(\tau_{u^-}\zeta_{u^-}\big). \tag{1.6}$$

Finally, the growth rate τ_u of u is inherited from its parent τ_{u^-} according to a Markov kernel

$$\rho(v, dv') = \mathbb{P}(\tau_u \in dv' \,|\, \tau_{u^-} = v), \tag{1.7}$$

where $v > 0$ and $\rho(v, dv')$ is a probability measure on $(0, \infty)$ for each $v > 0$. Eq. (1.4), (1.5), (1.6) and (1.7) completely determine the dynamics of the model $\big((\xi_u, \tau_u), u \in \mathcal{U}\big)$, as a Markov chain on a tree, given an additional initial condition $(\xi_\emptyset, \tau_\emptyset)$ on the root. The chain is embedded into a piecewise deterministic continuous Markov process thanks to (1.4) by setting

$$(\xi_t^u, \tau_t^u) = \big(\xi_u \exp\big(\tau_u(t - b_u)\big), \tau_u\big) \text{ for } t \in [b_u, b_u + \zeta_u)$$

and $(0, 0)$ otherwise. Define

$$\big(X(t), V(t)\big) = \Big(\big(X_1(t), V_1(t)\big), \big(X_2(t), V_2(t)\big), \dots\Big)$$

as the process of sizes and growth rates of the living particles in the system at time t. As for the fragmentation process we have an identity between point measures

$$\sum_{i=1}^{\infty} \mathbf{1}_{\{X_i(t)>0\}}\delta_{(X_i(t),V_i(t))} = \sum_{u\in\mathcal{U}} \mathbf{1}_{\{b_u \leq t < b_u+\zeta_u\}}\delta_{(\xi_t^u,\tau_t^u)} \tag{1.8}$$

where δ denotes the Dirac mass.

If μ is a probability measure on the state space $\mathcal{S} = [0, \infty) \times \mathcal{E}$, we shall denote

indifferently by $\mathbb{P}_\mu$ the law of any of the three processes above where the root $(\xi_\emptyset, \tau_\emptyset)$ has distribution μ. The construction is classical (see for instance (Ber06) and the references therein).

1.2.2 The behaviour of the mean empirical measure

Denote by $\mathcal{C}_0^1(\mathcal{S})$ the set of real-valued test functions with compact support in the interior of $\mathcal{S}$.

Theorem 1.1. *Work under Assumption 1 in (DHKR15). Let μ be a probability distribution on $\mathcal{S}$. Define the distribution $n(t, dx, dv)$ by*

$$\langle n(t,\cdot), \varphi\rangle = \mathbb{E}_\mu\Big[\sum_{i=1}^{\infty} \varphi\big(X_i(t), V_i(t)\big)\Big] \text{ for every } \varphi \in \mathcal{C}_0^1(\mathcal{S}).$$

Then $n(t,\cdot)$ solves (in a weak sense)

$$\begin{cases} \partial_t n(t,x,v) + v\,\partial_x\big(xn(t,x,v)\big) + B(x)n(t,x,v) \\ \quad = \; 4B(2x)\int_{\mathcal{E}} \rho(v',v)n(t,2x,dv'), \\ n(0,x,v) = n^{(0)}(x,v), x \geq 0 \end{cases}$$

with initial condition $n^{(0)}(dx, dv) = \mu(dx, dv)$.

Theorem 1.1 somehow legitimates our methodology: by enabling each cell to have its own growth rate and by building-up new statistical estimators in this context, we still have a translation in terms of the approach in (DPZ09). In particular, we will be able to compare our estimation results with (DHRBR12). Our proof is based on fragmentation techniques, inspired by Jean Bertoin (Ber06) and Bénédicte Haas (Haa03). Alternative approaches to the same kind of questions include the probabilistic studies of Brigitte Chauvin et al. (CRW91), Vincent Bansaye et al. (BDMT11) or Simon Harris and Matthew Roberts (HR17) and the references therein.

1.2.3 A many-to-one formula via a tagged branch

For $u \in \mathcal{U}$, we set $m^i u$ for the i-th parent along the genealogy of u. Define

$$\overline{\tau_t^u} = \sum_{i=1}^{|u|} \tau_{m^i u}\zeta_{m^i u} + \tau_t^u(t - b_u) \text{ for } t \in [b_u, b_u + \zeta_u)$$

and 0 otherwise for the cumulated growth rate along its ancestors up to time t. In the same spirit as tagged fragments in fragmentation processes (see the book by Jean Bertoin (Ber06) for instance) we pick a branch at random along the genealogical tree

at random: for every $k \geq 1$, if ϑ_k denotes the node of the tagged branch at the k-th generation, we have

$$\mathbb{P}(\vartheta_k = u) = 2^{-k} \text{ for every } u \in \mathcal{U} \text{ such that } |u| = k,$$

and 0 otherwise. For $t \geq 0$, the relationship

$$b_{\vartheta_{C_t}} \leq t < b_{\vartheta_{C_t}} + \zeta_{\vartheta_{C_t}}$$

uniquely defines a counting process $(C_t, t \geq 0)$ with $C_0 = 0$. The process C_t enables in turn to define a tagged process of size, growth rate and cumulated growth rate via

$$\big(\chi(t), \mathcal{V}(t), \overline{\mathcal{V}}(t)\big) = \Big(\xi_t^{\vartheta_{C_t}}, \tau_t^{\vartheta_{C_t}}, \overline{\tau_t^{\vartheta_{C_t}}}\Big) \text{ for } t \in [b_{\vartheta_{C_t}}, b_{\vartheta_{C_t}} + \zeta_{\vartheta_{C_t}})$$

and 0 otherwise. We have the representation

$$\chi(t) = \frac{x e^{\overline{\mathcal{V}}(t)}}{2^{C_t}} \tag{1.9}$$

and since $\mathcal{V}(t) \in [e_{\min}, e_{\max}]$, we note that

$$e_{\min} t \leq \overline{\mathcal{V}}(t) \leq e_{\max} t. \tag{1.10}$$

The behaviour of $\big(\chi(t), \mathcal{V}(t), \overline{\mathcal{V}}(t)\big)$ can be related to certain functionals of the whole particle system via a so-called many-to-one formula. This is the key tool to obtain Theorem 1.1.

Proposition 1.1 (A many-to-one formula)**.** *Work under Assumption 1 in (DHKR15). For $x \in (0, \infty)$, let $\mathbb{P}_x$ be defined as in Lemma 1 in (DHKR15). For every $t \geq 0$, we have*

$$\mathbb{E}_x\big[\phi\big(\chi(t), \mathcal{V}(t), \overline{\mathcal{V}}(t)\big)\big] = \mathbb{E}_x\Big[\sum_{u \in \mathcal{U}} \xi_t^u \frac{e^{-\overline{\tau_t^u}}}{x} \phi\big(\xi_t^u, \tau_t^u, \overline{\tau_t^u}\big)\Big] \tag{1.11}$$

for every $\phi : \mathcal{S} \times [0, \infty) \to [0, \infty)$.

1.2.4 Statistical estimation of the division rate

Two observation schemes

Let $\mathcal{U}_n \subset \mathcal{U}$ denote a subset of size n of connected nodes: if u belongs to $\mathcal{U}_n$, so does its parent u^-. We look for a nonparametric estimator of the division rate

$$y \rightsquigarrow B(y) \text{ for } y \in (0, \infty).$$

Statistical inference is based on the observation scheme

$$\big((\xi_u, \tau_u), u \in \mathcal{U}_n\big)$$

and asymptotic study is undertaken as the population size of the sample $n \to \infty$. We are interested in two specific observation schemes.

The full tree case. We observe every pair (ξ_u, τ_u) over the first N_n generations of the tree:

$$\mathcal{U}_n = \{u \in \mathcal{U},\ \ |u| \leq N_n\}$$

with the notation $|u| = n$ if $u = (u_0, u_1, \ldots, u_n) \in \mathcal{U}$, and N_n is chosen such that that 2^{N_n} has order n. This model corresponds to the database of the publication Eric Stewart et al. (SMPT05) described in the section 1.1.1. □

The sparse tree case. We follow the first n offsprings of a single cell, along a fixed line of descendants. This means that for some $u \in \mathcal{U}$ with $|u| = n$, we observe every size ξ_u and growth rate τ_u of each node (u_0), (u_0, u_1), (u_0, u_1, u_2) and so on up to a final node $u = (u_0, u_1, \ldots, u_n)$.This model corresponds to the database of the publication Robert Wang et al. (WRPDTWS10) described in the Section 1.1.1. □

Remark 1.1. *For every $n \geq 1$, we tacitly assume that there exists a (random) time $T_n < \infty$ almost surely, such that for $t \geq T_n$, the observation scheme $\mathcal{U}_n$ is well-defined. This is a consequence of the behaviour of B near infinity that we impose later on in* (1.19) *below.*

Estimation of the division rate

We denote by $\boldsymbol{x} = (x, v)$ an element of the state space $\mathcal{S} = [0, \infty) \times \mathcal{E}$. Introduce the transition kernel

$$\mathcal{P}_B(\boldsymbol{x}, d\boldsymbol{x'}) = \mathbb{P}\big((\xi_u, \tau_u) \in d\boldsymbol{x'} \,\big|\, (\xi_{u^-}, \tau_{u^-}) = \boldsymbol{x}\big)$$

of the size and growth rate distribution (ξ_u, τ_u) at the birth of a descendant $u \in \mathcal{U}$, given the size at birth and growth rate of its parent (ξ_{u^-}, τ_{u^-}). From (1.5), we infer that $\mathbb{P}(\zeta_{u^-} \in dt \,\big|\, \xi_{u^-} = x, \tau_{u^-} = v)$ is equal to

$$B\big(x \exp(vt)\big) \exp\Big(- \int_0^t B\big(x \exp(vs)\big) ds\Big) dt.$$

Using formula (1.6), by a simple change of variables

$$\mathbb{P}\big(\xi_u \in dx' \big|\, \xi_{u^-} = x, \tau_{u^-} = v\big) = \frac{B(2x')}{vx'} \mathbf{1}_{\{x' \geq x/2\}} \exp\big(- \int_{x/2}^{x'} \tfrac{B(2s)}{vs} ds\big) dx'.$$

Incorporating (1.7), we obtain an explicit formula for

$$\mathcal{P}_B(\boldsymbol{x}, d\boldsymbol{x'}) = \mathcal{P}_B\big((x,v), x', dv'\big)dx',$$

with

$$\mathcal{P}_B\big((x,v), x', dv'\big) = \frac{B(2x')}{vx'}\mathbf{1}_{\{x'\geq x/2\}}\exp\big(-\int_{x/2}^{x'}\tfrac{B(2s)}{vs}ds\big)\rho(v,dv'). \tag{1.12}$$

Assume further that $\mathcal{P}_B$ admits an invariant probability measure $\nu_B(d\boldsymbol{x})$, *i.e.* a solution to

$$\nu_B\mathcal{P}_B = \nu_B, \tag{1.13}$$

where

$$\mu\mathcal{P}_B(d\boldsymbol{y}) = \int_{\mathcal{S}}\mu(d\boldsymbol{x})\mathcal{P}_B(\boldsymbol{x}, d\boldsymbol{y})$$

denotes the left action of positive measures $\mu(d\boldsymbol{x})$ on $\mathcal{S}$ for the transition $\mathcal{P}_B$.

Proposition 1.2. *Work under Assumption 1 in (DHKR15). Then $\mathcal{P}_B$ admits an invariant probability measure ν_B of the form $\nu_B(d\boldsymbol{x}) = \nu_B(x,dv)dx$ and we have*

$$\nu_B(y) = \frac{B(2y)}{y}\mathbb{E}_{\nu_B}\Big[\frac{1}{\tau_{u^-}}\mathbf{1}_{\{\xi_{u^-}\leq 2y,\,\xi_u\geq y\}}\Big] \tag{1.14}$$

where $\mathbb{E}_{\nu_B}[\cdot]$ denotes expectation when the initial condition $(\xi_\emptyset, \tau_\emptyset)$ has distribution ν_B and where we have set $\nu_B(y) = \int_{\mathcal{E}}\nu_B(y,dv')$ in (1.14) for the marginal density of the invariant probability measure ν_B with respect to y.

Key idea for the proof of (1.14)
As $\nu_B(d\boldsymbol{x}) = \nu_B(x,dv)dx$, it follows that for every $y\in(0,\infty)$,

$$\begin{aligned}\nu_B(y,dv') &= \int_{\mathcal{S}}\nu_B(x,dv)dx\,\mathcal{P}_B\big((x,v),y,dv'\big)\\ &= \frac{B(2y)}{y}\int_{\mathcal{E}}\int_0^{2y}\nu_B(x,dv)\exp\big(-\int_{x/2}^{y}\tfrac{B(2s)}{vs}ds\big)\frac{\rho(v,dv')}{v}dx.\end{aligned}$$

By Assumption 1 in (DHKR15), we have $\int_{x/2}^{\infty}\frac{B(2s)}{s}ds = \infty$ hence

$$\exp\big(-\int_{x/2}^{y}\tfrac{B(2s)}{vs}ds\big) = \int_y^{\infty}\tfrac{B(2s)}{vs}\exp\big(-\int_{x/2}^{s}\tfrac{B(2s')}{vs'}ds'\big)ds.$$

It follows that $\nu_B(y,dv')$ is equal to

$$\begin{aligned}&\frac{B(2y)}{y}\int_{\mathcal{E}}\int_0^{2y}\nu_B(x,dv)dx\int_y^{\infty}\tfrac{B(2s)}{vs}\exp\big(-\int_{x/2}^{s}\tfrac{B(2s')}{vs'}ds'\big)ds\frac{\rho(v,dv')}{v}\\ =\ &\frac{B(2y)}{y}\int_{\mathcal{S}}\int_{[0,\infty)}\mathbf{1}\{x\leq 2y, s\geq y\}v^{-1}\nu_B(x,dv)dx\,\mathcal{P}_B\big((x,v),s,dv'\big)ds.\end{aligned}$$

Integrating with respect to dv', we obtain the result.

Construction of a nonparametric estimator

Inverting (1.14) and applying an appropriate change of variables, we obtain

$$B(y) = \frac{y}{2} \frac{\nu_B(y/2)}{\mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y,\, \xi_u \geq y/2\}}\right]}, \tag{1.15}$$

provided the denominator is positive. Representation (1.15) suggests an estimation procedure, replacing the marginal density $\nu_B(y/2)$ and the expectation in the denominator by their empirical counterparts. To that end, pick a kernel function

$$K : [0, \infty) \to \mathbb{R}, \quad \int_{[0,\infty)} K(y)dy = 1,$$

and set $K_h(y) = h^{-1}K\big(h^{-1}y\big)$ for $y \in [0, \infty)$ and $h > 0$. Our estimator is defined as

$$\widehat{B}_n(y) = \frac{y}{2} \frac{n^{-1} \sum_{u \in \mathcal{U}_n} K_h(\xi_u - y/2)}{n^{-1} \sum_{u \in \mathcal{U}_n} \frac{1}{\tau_{u^-}} \mathbf{1}\{\xi_{u^-} \leq y, \xi_u \geq y/2\} \bigvee \varpi}, \tag{1.16}$$

where $\varpi > 0$ is a threshold that ensures that the estimator is well defined in all cases and $x \bigvee y = \max\{x, y\}$. Thus $(\widehat{B}_n(y), y \in \mathcal{D})$ is specified by the choice of the kernel K, the bandwidth $h > 0$ and the threshold $\varpi > 0$.

Assumption. *The function K has compact support, and for some integer $n_0 \geq 1$, we have $\int_{[0,\infty)} x^k K(x)dx = \mathbf{1}_{\{k=0\}}$ for $0 \leq k \leq n_0$.*

Rate of convergence

We are ready to state our main result. For $s > 0$, with $s = \lfloor s \rfloor + \{s\}$, $0 < \{s\} \leq 1$ and $\lfloor s \rfloor$ an integer, introduce the Hölder space $\mathcal{H}^s(\mathcal{D})$ of functions $f : \mathcal{D} \to \mathbb{R}$ possessing a derivative of order $\lfloor s \rfloor$ that satisfies

$$|f^{\lfloor s \rfloor}(y) - f^{\lfloor s \rfloor}(x)| \leq c(f)|x - y|^{\{s\}}. \tag{1.17}$$

The minimal constant $c(f)$ such that (1.17) holds defines a semi-norm $|f|_{\mathcal{H}^s(\mathcal{D})}$. We equip the space $\mathcal{H}^s(\mathcal{D})$ with the norm

$$\|f\|_{\mathcal{H}^s(\mathcal{D})} = \|f\|_{L^\infty(\mathcal{D})} + |f|_{\mathcal{H}^s(\mathcal{D})}$$

and the Hölder balls

$$\mathcal{H}^s(\mathcal{D}, M) = \{B,\ \|B\|_{\mathcal{H}^s(\mathcal{D})} \leq M\},\ M > 0.$$

For $\lambda > 0$ and a vector of positive constants $\mathfrak{c} = (r, m, \ell, L)$, introduce the class $\mathcal{F}^\lambda(\mathfrak{c})$ of continuous functions $B : [0, \infty) \to [0, \infty)$ such that

$$\int_0^{r/2} x^{-1} B(2x) dx \le L, \quad \int_{r/2}^{r} x^{-1} B(2x) dx \ge \ell, \tag{1.18}$$

and

$$B(x) \ge m\, x^\lambda \ \text{ for } \ x \ge r. \tag{1.19}$$

Theorem 1.2. *Work under Assumption 3 in (DHKR15) in the sparse tree case and Assumption 4 in (DHKR15) in the full tree case. Specify $\widehat{B}$ with a kernel K satisfying Assumption 2 in (DHKR15) for some $n_0 > 0$ and*

$$h = c_0 n^{-1/(2s+1)}, \ \ \varpi_n = (\log n)^{-1}.$$

For every $M > 0$ there exist $c_0 = c_0(\mathfrak{c}, M)$ and $d(\mathfrak{c}) \ge 0$ such that for every $0 < s < n_0$ and every compact interval $\mathcal{D} \subset (d(\mathfrak{c}), \infty)$ such that $\inf \mathcal{D} \ge r/2$, we have

$$\sup_{\rho, B} \mathbb{E}_\mu \big[\| \widehat{B}_n - B \|^2_{L^2(\mathcal{D})} \big]^{1/2} \lesssim (\log n) n^{-s/(2s+1)},$$

where the supremum is taken over

$$\rho \in \mathcal{M}(\rho_{\min}, \rho_{\max}) \ \text{ and } \ B \in \mathcal{F}^\lambda(\mathfrak{c}) \cap \mathcal{H}^s(\mathcal{D}, M),$$

and $\mathbb{E}_\mu[\cdot]$ denotes expectation with respect to any initial distribution $\mu(d\boldsymbol{x})$ for $(\xi_\emptyset, \tau_\emptyset)$ on $\mathcal{S}$ such that $\int_{\mathcal{S}} \mathbb{V}(\boldsymbol{x})^2 \mu(d\boldsymbol{x}) < \infty$.

Remarks 1.1.

1. *We obtain the classical rate $n^{-s/(2s+1)}$ (up to a log term) which is optimal in a minimax sense for density estimation. It is presumably optimal in our context, using for instance classical techniques for nonparametric estimation lower bounds on functions of transition densities of Markov chains, see for instance (GHR04).*

2. *The extra logarithmic term is due to technical reasons: we need it in order to control the decay of correlations of the observations over the full tree structure.*

3. *The knowledge of the smoothness s that is needed for the construction of $\widehat{B}_n$ is not realistic in practice. An adaptive estimator could be obtained by using a data-driven bandwidth in the estimation of the invariant density $\nu_B(y/2)$ in* (1.16). *The Goldenschluger-Lepski bandwidth selection method (GL11), see also (DHRBR12) would presumably yield adaptation, but checking the assumptions still requires a proof in our setting. We implement data-driven bandwidth in the numerical Section 1.2.5 below.*

1.2.5 Numerical implementation

For the simulations, we first generated simulated data in order to verify that our protocol did work. For more details we can refer to (DHKR15) and that gave the figure 1.4.

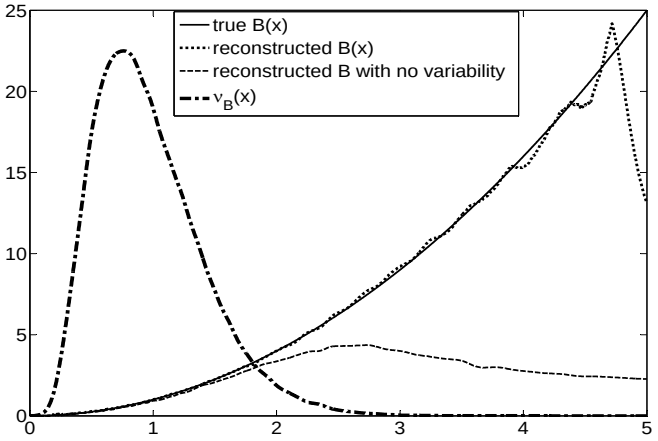


Figure 1.4: *Reconstruction for $n = 2^{17}$ and $\varpi = n^{-1/2}$. When the variability in the growth rate is ignored, the estimate reveals unsatisfactory. The parameter values are the reference ones.*

We proceed as in the above protocol. Figure 1.5 shows the reconstructed B and ν_B for a sample of $n = 2335$ cells. Though much more precise and reliable, thanks both to the experimental device and the reconstruction method, our results are qualitatively in accordance with previous indirect reconstructions carried out in (DMZ10) on old datasets published in (Kub69) back in 1969. The reconstruction of the division rate is prominent here since it appears to be the last component needed for a full calibration of the model. Thus, our method provides the biologists with a complete understanding of the size dependence of the biological system. Phenotypic variability between genetically identical cells has recently received growing attention with the recognition that it can be genetically controlled and subject to selection pressures (KEBC05). Our mathematical framework allows the incorporation of this variability at the level of individual growth rates. It should allow the study of the impact of variability on the population fitness and should be of particular importance to describe the growth of populations of cells exhibiting high variability of growth rates. Several examples of high variability have been described, both in genetically engineered or natural bacterial populations (SMPT05; TMY09).

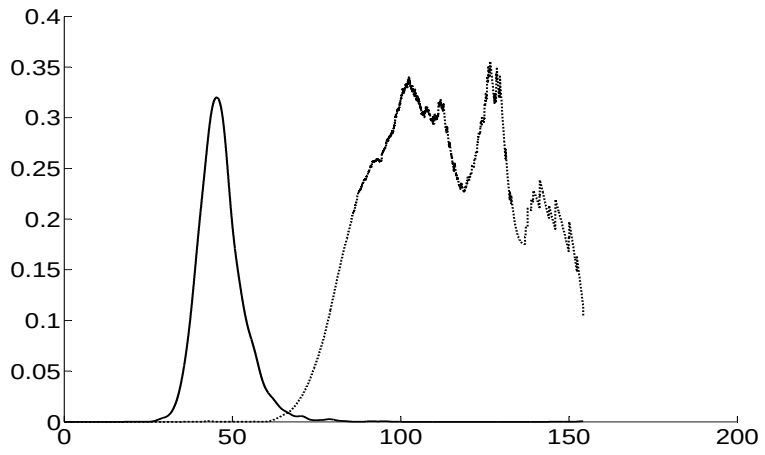


Figure 1.5: *Estimation of B (dotted line) and ν_B (solid line) on experimental data of E. coli dividing cells, $n = 2335$. In abscissae, the bacterial length is in arbitrary unit.*

1.3 The old and the new pole

1.3.1 The growth rate

One of the many questions that interested Lydia Robert, in addition to better understanding the growth of *E. coli* bacteria, was to understand a little about the influence for a bacterium to be old poles or not. In this section, we will interest us on the paper (DSKR18). In this paper written with Benoîte de Saporta, Bernard Delyon and Lydia Robert, we tried to respond to this question.

Although two sister cells are clones with identical genetic material, asymmetry in *E. coli* division makes sense biologically as *E. coli* grows and reproduces by dividing roughly at its middle. Each cell has thus a new *pole* (created at the division of its mother) and an old one (one of the two original poles of its mother), see in the next Figure 1.6. The cell that inherits the old pole of its mother is called the *old pole* cell, the other one is called the *new pole* cell. It is suspected that both cells inherit different material or material of different quality from their mother cell. Therefore, each cell has a *type*: old pole (O) or new pole (N) cell. On experimental data, we usually do not know the type of the original cell and its two daughters at the root of the genealogy, but from generation 2 on, the type of each cell is known. For the cells of unknown types we will note them UT in the Figure 1.6 and we will write it in black. The cells of young type will be written in blue and the old one in red. For further generations, we can associate to one cell not only its type, but also the sequence of types of its ancestors, see Figure 1.6. The original ancestor is labelled 1 and the two daughters of cell n are labelled $2n$ for the new pole one and $2n + 1$ for the old pole one. Therefore, even-labelled cells are type N and odd-labelled cells are

type O and the whole sequence of types of their ancestors can be retrieved from the decomposition of their label in base 2 (with 0 coding for N and 1 coding for O). For instance, cell number 19 is type NOO which means, it is type O, its mother is type O and its grand-mother is type N.

One of the difficulties comes from the fact that we have access to two different data sets structured as binary genealogical trees. For the statistician, this special structure is hard to take into account rigorously because of the intricate dependence structure within a tree. The data sets come from two different biological experiments. One set corresponds to small complete trees (the Steward sets), whereas the other one corresponds to long specific sub-trees (the Wang sets). In the previous figure Wang's observations correspond to the bacteria with the orange circle. You can refer to the Section 1.1.1 for more details on these two data sets. Our aim is to compare both sets, which is especially complicated as they have very different tree structures.

The starting point of the present work is that the latter questions have seemingly opposite answers in the biological literature: in (SMPT05), the growth rate of older cells is significantly slowed down, whereas in (WRPDTWS10) it is stable. We provide the data sets from both of these papers, and our aim is to conduct a new statistical study of both data sets to investigate the behavior of the growth rate of *E. coli* and try to decide whether both experiments yield contradictory results or not.

For the study of the data there was a big work of treatment of the data because the measurements are extremely noisy. We can refer to (DSKR18) for more details on this point.

The main difficulty to analyze these data sets lies on the special dependence structure coming from the genealogical trees. To take this into account, we may use the BAR model from (GBP+05; Guy07; dSGPM11; dSGPM12; dSGPM14).

Let $X_{j,k}$ be the growth rate of cell number k in tree number j. The asymmetric BAR model is an autoregressive model defined as follows: $X_{j,1}$ is arbitrary and for $k \geq 1$, one has

$$X_{j,2k} \quad = \quad a_0 + b_0 X_{j,k} + \varepsilon_{j,2k},$$

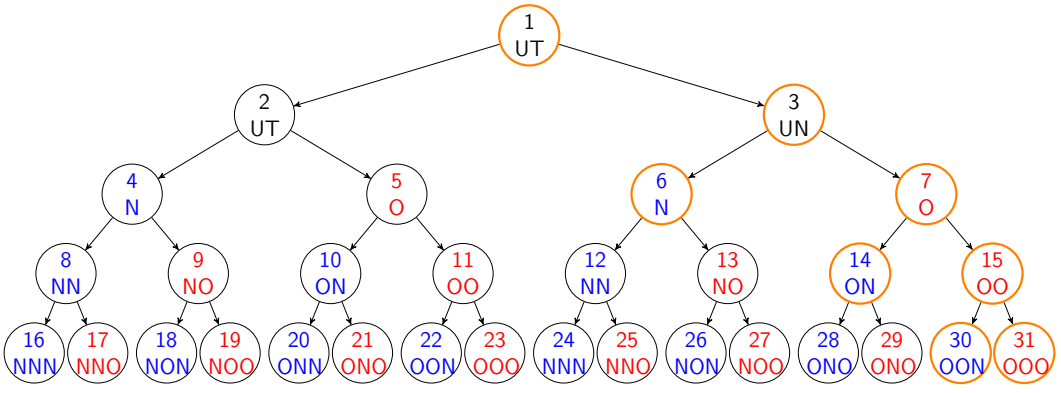


Figure 1.6: Cell division binary tree with the type of each cell

$$X_{j,2k+1} \quad = \quad a_1 + b_1 X_{j,k} + \varepsilon_{j,2k+1},$$

where $(\varepsilon_{j,k})$ is a noise sequence and $\theta = (a_0, b_0, a_1, b_1)$ parameters to be estimated.

By adapting the techniques of (dSGPM14), we obtain the estimation results given in Figure 1.7.

1.3.2 The influence of being old or new

As it is not possible to compare the BAR model for both data sets, we turned to more basic tools to compare the influence of the mother and higher ancestors on the growth rate of a given cell.

We averaged the growth rates of cells within the same generation of the same tree (without taking care of the border distance), and normalized the growth rate of each cell with the suitable average. Then we computed the mean growth rate over all normalized cells that have cumulated n new poles or n old poles (for $1 \leq n \leq 7$). The results are given on Figure 1.8 (a), circles are cumulated new-pole cells and stars cumulated old-pole cells. This figure corresponds to Figure 3 in (SMPT05). Then we compared the mean of all new-pole cells which mother cumulated n old poles, and old-pole cells which mother cumulated n new poles (for $1 \leq n \leq 6$), see Figure 1.8 (b), circles are new-pole cells with cumulated old-pole mother and stars old-pole cells with cumulated new-pole mother. The scales of both figures are the same to facilitate comparison.

The linear regression slope coefficients are respectively 4.4% for the new pole cells and -1.1% in Figure 1.8 (a), 0.1% for the new pole cells and -0.5% in Figure 1.8 (b).

We can conclude that one new pole is enough to *forget* an accumulation of old poles and similarly one old pole is enough to forget an accumulation of new poles.

The influence of the mother and the grand-mother

For each tree, we selected the old cell branch (rightmost branch in Figure 1.6) and we fit an additive regression model explaining the growth rate of a cell with the one

	Estimation	95% confidence interval
$\widehat{a}_{0,n}$	0.0304	$[0.0200; 0.0410]$
$\widehat{b}_{0,n}$	0.0664	$[-0.4652; 0.5980]$
$\widehat{a}_{1,n}$	0.0281	$[0.0178; 0.0385]$
$\widehat{b}_{1,n}$	0.0994	$[-0.3194; 0.5182]$

Figure 1.7: Estimated parameters for the BAR model, Wang data, $n = 302$, $m = 224$.

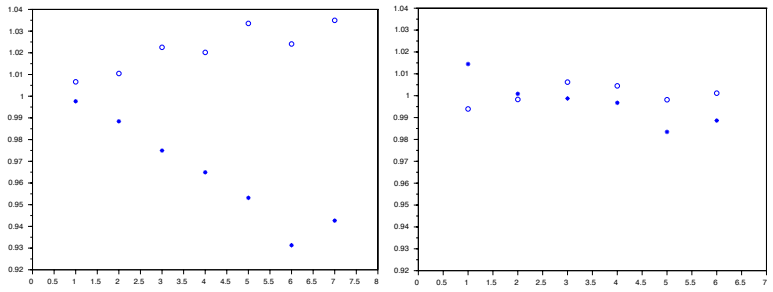


Figure 1.8: Mean normalized growth rate within generations and trees for cells that have cumulated (a) n consecutive new poles (circles) or n consecutive old poles (stars) for $1 \leq n \leq 7$; (b) 1 new pole after n consecutive old poles (circles), 1 old pole after n consecutive new poles (stars), for $1 \leq n \leq 6$, Stewart data set.

of its mother and the one of its grand mother

$$r_n = \beta_m m_n + \beta_g g_n + \beta_0 + e_n \tag{1.20}$$

where

- r_n is the growth rate of the n-th generation cell ($X_{2^{n+1}-1}$ with previous notation),
- m_n is the growth rate of its mother (X_{2^n-1}),
- g_n is the growth rate of its grand mother ($X_{2^{n-1}-1}$)
- e_n the prediction error.

The triple $(\beta_0, \beta_m, \beta_g)$ depends on the tree. The R command is `lm(rate~ratemo+rategdmo)`. Histograms of p-values for the significance of the mother coefficient β_m and for the grand mother coefficient β_g are plotted in Figure 1.9 .

We conclude that the effect of the grand mother is not significant. The coefficient β_m is significantly positive with a value around 0.3.

Looking at Wang's data, we compared new pole and old pole cells means as well as mother-daughter correlation. More specifically,

1. Student test for comparison of the mean of the growth rate of old pole cells and of new pole cells yields a p-value $< 10^{-16}$, and 1% confidence intervals for

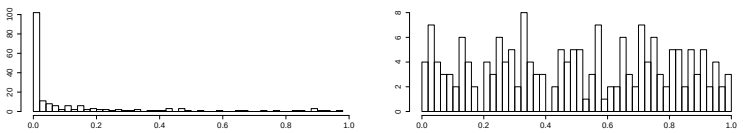


Figure 1.9: Histogram of p-value for significance of the mother coefficient β_m (a) and for the grand mother coefficient β_g (b), Wang data set.

mean growth rates are: $[0.0309, 0.031]$ for old pole cells and $[0.03186, 0.03195]$ for new pole cells;

2. Correlation daughter mother. We have computed one confidence interval for overall correlation between old pole daughters and their mother, and another for new pole daughters. 1% confidence intervals for correlation between growth rates of new pole daughters and that of their mother is: $[0.085, 0.123]$, the same for old pole cells is $[0.125, 0.16]$.

A significant difference thus holds for the mean as well as for the correlation with the mother cell.

The stationary of the process

Both experiments were not conducted at the same stage in the life of *E. coli* cells. In (SMPT05), they selected a random cell from a previous colony and let it grow and divide in a new medium. Thus, the first generations of observed cells are stressed, leading to a reduced growth rate, see Figure 1.11. This corresponds to a transient phase.

We see on Figure 1.10 an uniform distribution of the p-value, which is characteristic of the non-significance of the hypothesis of different distributions.

Conclusion

In these two data sets, we made efforts to take in account the tree structure of the data. We tried different statistical procedures that can be summed up as follows.

Wang data: Because of the simple structure of this data set, each tree is here just the orange subtree in Figure 1.6. We have tried dynamical models in which the growth rate of a cell may have a multi-generation memory, with coefficients possibly

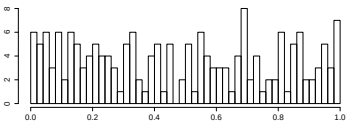


Figure 1.10: P-values for the Kolmogorov test of stationarity, Wang data set.

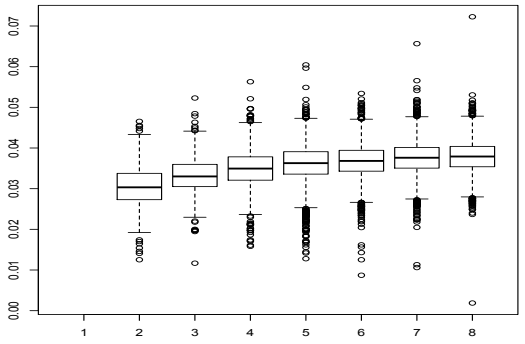


Figure 1.11: Box plots of growth rates for cells in generations 2 to 8, Stewart data set.

dependent on the tree (mixed effects). We did not find a significant improvement over the simplest model where the rate of a cell depends only on the one of its mother, and that the grand mother has no significant influence. We found that

1. The old pole cell growth rate is significantly more correlated to its mother than the new pole cell.
2. The mean old pole cell growth rate is significantly smaller than the mean new pole cell growth rate.
3. The stationarity cannot be rejected.

Stewart data: The tree structure induces dependency in the data which we have taken into account in our testing procedures...

We observe that

1. The old pole cell growth rate is significantly more correlated to its mother than the new pole cell.
2. The mean old pole cell growth rate is significantly smaller than the mean new pole cell growth rate.
3. There is no stationarity of the growth rate across generations. This means that the initial stress of the experiment has not enough time to vanish during only 10 generations.
4. An important factor is the number of generations since the last change of pole type, for example, cell 17 (NNO) in figure 1.6 should behave similarly as cell 21 (ONO), or NONOONN as ONONONN.

To conclude, in both data sets, we recover a statistically significant difference between the growth rate of sister cells. Therefore, asymmetry is present in the division of the *E. coli*, even after hundreds of generations.

The apparent conflict between both data sets may simply come from observations at different phases: Stewart's data are still in a transient phase whereas Wang's data are stationary. From this point of view, the two data sets are not contradictory. To our best knowledge, there is no available data set of *E. coli* division with both transient and steady states. It would be interesting to design an experiment where both the transient and the stationary phase could be observed on the same colonies.

1.4 To go further

After having tried to take into account the influence of the old pole and the young pole of a bacterium for the growth rates in the previous section, we could ask the question more generally for all the other parameters of the bacterium. This is what motivated the current collaboration with Bertrand Cloez, Benoîte de Saporta and Tristan Roget. We also wanted to generalize the model studied in the Section 1.2 by incorporating the fact that this time the size of the bacteria would not be divided exactly in two, but that there would be a proportion θ_i of the size of the mother for young bacteria and $1-\theta_i$ for old bacteria with the θ_i depending on the type of mother. The growth rate τ_i will also depend on the type, as well as the division rate B_i. One of the first question is to estimate the new parameter θ_i. A second step is to make the same work that was done in (DHKR15) to estimate the new division rate by incorporating the adaptive method that we developed in (KS21).

Let (τ_0, τ_1) describes the growth rates of old and young cells and the parameters (θ_0, θ_1) describe the size proportions inherited by the old and young cell.

In the same way as in the Subsection 1.2.1, we can construct the genealogical tree and the underlying process. We can make the convention that the labels ending with 1 will be of young type and those ending with 0 will be of old type. The difference is that for the equation (1.6) we will obtain instead

$$\xi_{uij} = \theta_i \xi_{ui} \exp\left(\tau_i \zeta_i\right) \tag{1.21}$$

for $u \in \mathcal{U}$ and $(i,j) \in \{0,1\}^2$.

Similarly the equation (1.12) for the transition kernel of the size at death and the type (0 for the old cell and 1 for the young) will become:

$$Q_B\big((x,i),(x',i')\big) = \frac{B(x',i')}{2\alpha_{i'}x'}\mathbf{1}_{\{x' \geq x\theta_{i'}\}} \exp\Big(-\int_{x\theta_{i'}}^{x'} \frac{B(s,i')}{\alpha_{i'}s} ds\Big). \tag{1.22}$$

With methods similar to (DHKR15), one could obtain that under good assumptions:

Proposition 1.3. *$\mathcal{Q}_B$ admits an invariant probability measure with density $\mu_B(x,i)$ and we have*

$$\mu_B(y,j) = \frac{B(y,j)}{y}\mathbb{E}_{\mu_B}\Big[\frac{1}{\tau_j}\mathbf{1}_{\{\theta_j d_{u-} \leq y,\, d_u \geq y,\, j_u = j\}}\Big]. \tag{1.23}$$

We could then combine the methods used in (DHKR15) with the adaptive estimators developed in (KS21) in order to have an adaptive estimate of B.

In this work in progress, we have obtained first simulations concerning the estimation of the division rate for young bacteria and another for old bacteria that is visibly significantly different, as you can observe in the figure 1.12.

We could also extend the question that was asked in (DHK$^+$14) in order to find the best possible modeling of the division rate. We could add a possible modeling

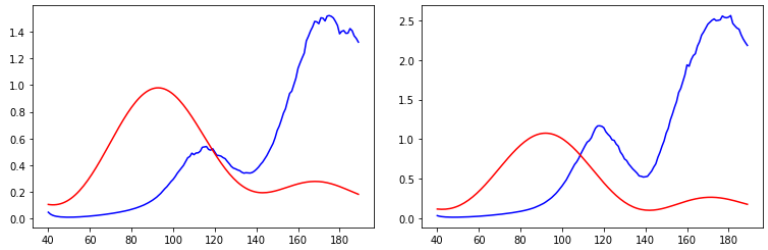


Figure 1.12: In red the invariant probability and in blue the division rate. On the left for the old cell and on the right for the young.

where the factor of the division, would also be the elongation and thus propose a model Adder.

Parameter estimation in branching processes has received significant attention; we refer the reader to the survey by (Yan08) , and to (Gut91) for a book-length treatment and also to the articles (DHK21) for the case of birth-and-death processes and (BHM22) for the case of branching processes with almost sure extinction. Another approach would be to study this literature to see what can be done in different branching process settings.

Chapter 2

A specific class of PDMP

2.1 Introduction

During the study of the evolution of bacterial growth (see previous chapter), there is a PDMP that appeared naturally, it was the process modeling the size of a marked bacterium. And at that time, Florent Malrieu was in Rennes and created his ANR on PDMP. It was a chance for me, it allowed me to know this processes better.

PDMPs were first introduced in the literature by Davis ((Dav84) and (Dav93)). Already at this time, the theory of diffusions had such powerful tools as the theory of Itô calculus and stochastic differential equations at its disposal. Davis's goal was to endow the PDMP with rather general tools. The main reason for that was to provide a general framework. Until this work only particular cases had been dealt with, which turned out not to be easily generalizable.

PDMPs form a family of càdlàg Markov processes involving a deterministic motion punctuated by random jumps. The motion of the PDMP $(X(t))_{t\geq 0}$ depends on three local characteristics, namely the jump rate λ, the flow ϕ and the transition measure Q according to which the location of the process at the jump time is chosen. The process starts from x and follows the flow $\phi(x,t)$ until the first jump time T_1 which occurs either spontaneously in a Poisson-like fashion with rate $\lambda(\phi(x,t))$ or when the flow $\phi(x,t)$ hits the boundary of the state-space. In both cases, the location of the process at the jump time T_1, denoted by $Z_1 = X(T_1)$, is selected by the transition measure $Q(\phi(x,T_1),\cdot)$ and the motion restarts from this new point as before. This fully describes a piecewise continuous trajectory for $\{X(t)\}_{t\geq 0}$ with jump times $\{T_k\}_{k\geq 1}$ and post jump locations $\{Z_k\}_{k\geq 1}$, and which evolves according to the flow ϕ between two jumps.

These processes have been heavily studied from both a theoretical and an applied perspective. For example in communication networks with the control of congestion TCP/IP (V. Dumas and al (DGR02), V. Guillemin et al. (GRZ04)), for molecular biology (BLM15), for the model of Hodgkin-Huxley concerning the neuronal activity (K. Pakdaman et al. (PTW10)), for bacterial chemotaxis (FGM16) in reliability

(F. Dufour and Y. Dutuit (DD02)) and for the movement of a population of bacteria (H.G. Othmer et al.(ODA88) as well as R. Erban and H.G. Othmer (EO05)). You can also refer to the survey (BT21) for applications in biology.

With Romain Azaïs, Jean-Baptiste Bardet, Alexandre Génadot and Pierre-André Zitt we wrote a proceedings (ABGKZ14) for the MAS days. In a first part, we give a precise definition and some general properties of the PDMPs. Then, we illustrate the state of the art regarding PDMPs through three specific examples: a model of switched vector fields, the TCP process, and a modelization of neuronal activity. Finally, we briefly review some results about a non-parametric statistical method to get an estimation of the conditional density associated with the jumps of a PDMP defined on a separable metric space and we end with a survey of numerical methods.

2.2 Nonparametric estimation of jump rates for a specific class of PDMP

2.2.1 The background

After the writing of the paper (DHKR15) on the estimation of the division rate of bacteria and the study of the specific PDMP that appeared in this study, I wanted to try to see how I could generalize the estimation of the division rate of a more general PDMP. Therefore, I wrote a first paper on the estimation of a class of PDMP (Kre16) which included 2 important examples, the TCP and the evolution of the marked bacteria.

The TCP (transmission control protocol) (see (DGR02), (GRZ04) for instance) is one of the main data transmission protocol on Internet. It is a piecewise deterministic Markov process $(X_t)_{t\geq 0} with flow \phi(x,t) = x+ct$ and deterministic transition measure $Q(x,y) = \mathbb{1}_{\{y=\kappa x\}}$. The data transmitted by the network grows in a linear way, until an error occurs then we restrict the quantity of transmitted data, it is only a proportion κ of the previous data which are transmitted. This process grows linearly (by construction) and the constant κ can be configured in the server implementation (so that is also known), but the moment when the transmission fails is of course unknown. In the literature, it is usually supposed that the jump rate satisfies $\lambda(x) = x$, but with this work we can check if it is a realistic assumption or not.

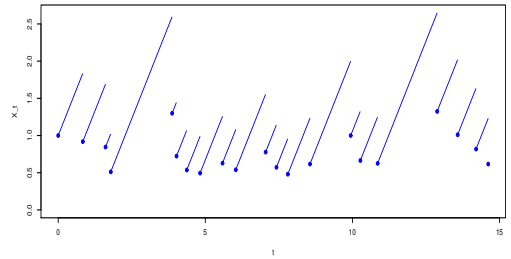
Another example of PDMP is the size of a marked bacteria (see (DHKR15), (LP09)). We randomly choose a bacteria, and follow its growth, until it divides in two. Then we randomly choose one of its daughters, and so on. Between the jumps, the bacteria grows exponentially: $\phi(x,t) = xe^{ct}$. The size of the bacteria after the division is random, as the bacteria does not divide itself in two equal parts.

Simulations for these 2 types of processes (with non-standard parameter choices ($\lambda(x) = \sqrt{x}$ for the TCP and a beta transition probability for the labeled bacterium) are given in Figure 2.1.

At that time, there were only few studies about estimation on PDMP. The paper (ADGP14) Azaïs et al. is an exception, which gives an estimator of the conditional distribution of the inter-jump times for a PDMP. The estimator is uniformly consistent when only one observation of the process within a long time is available. They deal with PDMPs which jump when they hit the boundary (this case is not considered in our paper). Their method relies on a generalization of Aalen's multiplicative intensity model (Aal75; Aal77; Aal78). But they only prove the uniform consistency of their estimator. They also have to assume that the process $(X(t))_{t\geq 0}$ evolves in a bounded space. Here we do not make such assumption. As a consequence the tools of their paper and of (Kre16) are different. To the best of my knowledge at this time, (ADGP14) was the only work investigating the nonparametric estimation of the conditional distribution of the inter-arrival times for PDMPs.

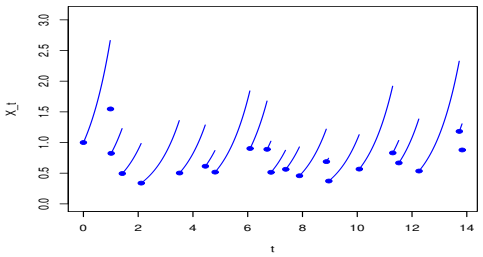
Figure 2.1: Examples of simulations of processes $\{X_t\}_{t\geq 0}$ and $\{Z_k\}_{k\in\mathbb{N}}$

TCP protocol



$\phi(x,t) = x + t$, $Z_k = Y_k/2$,
$\lambda(x) = \sqrt{x}$

Bacterial growth



$\phi(x,t) = xe^t$, $Z_k = Y_k U$, $U \sim \beta(20,20)$,
$\lambda(x) = x^2$

• : process $\{Z_k\}_{k\in\mathbb{N}}$ — : process $\{X(t)\}_{t\geq 0}$

I was interested in the study of the class of PDMP, whose state space is $(0,\infty)$, that possess an increasing deterministic motion and with a deterministic jump mechanism. As we observed the PDMP until its nth jump, I construct a nonparametric estimator of the jump rate λ. The main result was that for $\mathcal{D}$ a compact subset of $(0,\infty)$, if λ is in the Hölder space $\mathcal{H}^s(\mathcal{D})$, the squared-loss error of the estimator is asymptotically close to the speed of $n^{-s/(2s+1)}$. I decided not to detail it, in the presentation of the work (KS21) that generalizes it, instead I will insist on the differences with (Kre16).

Indeed, I was interested in generalizing this first work. There were two motivations, the first one was to have an adaptive estimator and the second one was to show that the speed was indeed min-max. For that I asked Emeline Schmisser to work with me because she was much more familiar with these notions than me, so a new collaboration was born.

At the time of the work (KS21), there was a bit more literature on the subject in particular (AMG16) but it was in a different framework than ours because it supposed that the distribution transition was absolutely continuous relative to the Lebesgue measure and that there was a boundary. But these were two things that we did not want, if we wanted to include our two key examples. We also refere on (Fuj13) and (AG18) but which only showed the consistency of their estimator. For a more detailed description of the existing works, we can refer to the papers cited in (KS21).

2.2.2 Presentation of the model

In the second paper (KS21) on estimation about PDMP, the framework is more general than the one I had supposed at the beginning, in particular the transition distribution can contain in addition to the deterministic part an absolutely continuous part with respect to the Lebesgue measure, which moreover will lead to an estimator

different from the one I had proposed and more complicated to study.

We consider a filtered PDMP $(X_t)_{t\geq 0}$ taking values in $\mathbb{R}^+$, with flow ϕ, transition measure $Q(x, dy)$ and homogeneous jump rate λ. Starting from initial value x_0, the process follows the flow ϕ until the first jump time T_1 which occurs spontaneously in a Poisson-like fashion with rate $\lambda(\phi(x,t))$. The post-jump location of the process at time T_1 is governed by the transition distribution $Q(\phi(x_0, T_1), dy)$ and the motion restarts from this new point as before.

A piecewise deterministic Markov process (PDMP) is defined by its local characteristics, namely, the jump rate λ, the flow ϕ and the transition measure Q according to which the location of the process is chosen after the jump. In this article, we consider an one-dimensional PDMP $\{X(t)\}_{t\geq 0}$. More precisely:

Assumption (A1)**.**

1. *The flow $\phi : \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}^+$ is a one-parameter group of homeomorphisms: ϕ is $\mathcal{C}^1$, for each $t \in \mathbb{R}^+$, $\phi(., t)$ is an homeomorphism satisfying the semigroup property: $\phi(., t+s) = \phi(\phi(., s), t)$ and for each $x \in \mathbb{R}^+$, $\phi_x(.) := \phi(x, .)$ is an increasing $\mathcal{C}^1$-diffeormorphism. In particular, $\phi(x, 0) = x$.*

2. *The jump rate $\lambda : \mathbb{R}^+ \to \mathbb{R}^+$ is a measurable function satisfying*

$$\forall x \in \mathbb{R}^+, \ \exists\, \varepsilon' > 0 \ \text{ such that } \ \int_0^{\varepsilon'} \lambda(\phi(x, s))ds < \infty$$

   that is, the jump rate does not explode.

3. $\forall x \in \mathbb{R}^+$, $Q(x, \mathbb{R}^+ \setminus \{x\}) = 1$.

For instance, we can take $\phi(x, t) = x+ct$ (linear flow) or $\phi(x, t) = xe^{ct}$ (exponential flow). The transition measure may be continuous with respect to the Lebesgue measure or deterministic (for example $Q(x, \{y\}) = \mathbb{1}_{\{y=f(x)\}}$).

Given these three characteristics, it can be shown ((Dav93, p62-66)), that it exists a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \{\mathbb{P}_x\}_{x\in\mathbb{R}^+})$ such that the motion of the process $\{X(t)\}_{t\geq 0}$ starting from a point $x_0 \in \mathbb{R}^+$ may be constructed as follows: consider a random variable T_1 with survival function

$$\mathbb{P}\left(T_1 > t \middle| X_0 = x_0\right) = e^{-\Lambda(x_0,t)}, \text{ where } \Lambda(x, t) = \int_0^t \lambda(\phi(x, s))ds. \tag{2.1}$$

If T_1 is equal to infinity, then the process $\{X(t)\}_{t\geq 0}$ follows the flow, i.e. for $t \in \mathbb{R}^+$, $X(t) = \phi(x_0, t)$. Otherwise let $Y_1 = \phi(x_0, T_1^-)$ the pre-jump location and Z_1 the post-jump location. Z_1 is defined through the transition kernel Q: $\mathbb{P}\left(Z_1 \in A \middle| Y_1 = y\right) = \int_A Q(y, dz)$. The trajectory of $\{X(t)\}_{t\geq 0}$ starting at x_0, for $t \in [0, T_1]$, is given by

$$X(t) = \begin{cases} \phi(x_0, t) & \text{for } t < T_1, \\ Z_1 & \text{for } t = T_1. \end{cases}$$

Inductively starting from $X(T_n) = Z_n$, we now select the next inter-jump time $T_{n+1} - T_n$ and post-jump location $X(T_{n+1}) = Z_{n+1}$ in a similar way. This construction properly defines a strong Markov process $\{X(t)\}_{t\geq 0}$ with jump times $\{T_k\}_{k\in\mathbb{N}}$ (where $T_0 = 0$). A very natural Markov chain is linked to $\{X(t)\}_{t\geq 0}$, namely the jump chain $\{Y_n, Z_n\}_{n\in\mathbb{N}}$ (or, equivalently, $\{T_n, Z_n\}_{n\in\mathbb{N}}$).

To simplify the notations, let us set $\phi_x(t) = \phi(x,t)$ and $z_0 = x_0$. By (2.1),

$$\begin{aligned}\mathbb{P}\left(Y_1 > y \middle| Z_0 = z_0\right) &= \mathbb{P}\left(T_1 > (\phi_{z_0})^{-1}(y) \middle| Z_0 = z_0\right) \\ &= \exp\left(-\int_0^{(\phi_{z_0})^{-1}(y)} \lambda(\phi_{z_0}(s))ds\right) \mathbb{1}_{\{y\geq z_0\}}\end{aligned}$$

and by the change of variable $u = \phi_{z_0}(s)$ (we recall that for any $z \in \mathbb{R}^+$, ϕ_z is a monotonic function), we get

$$\mathbb{P}\left(Y_1 > y \middle| Z_0 = z_0\right) = \exp\left(-\int_{z_0}^{y} \lambda(u)\left(\phi_{z_0}^{-1}\right)'(u)du\right) \mathbb{1}_{\{y\geq z_0\}}. \tag{2.2}$$

If the function $\lambda(y)(\phi_{z_0}^{-1})'(y)$ is finite, we obtain the conditional density:

$$\mathcal{P}(z_0, y) := \lambda(y)(\phi_{z_0}^{-1})'(y)e^{-\int_{z_0}^{y} \lambda(u)(\phi_{z_0}^{-1})'(u)du} \mathbb{1}_{\{y\geq z_0\}}. \tag{2.3}$$

Estimating directly λ is difficult, but we can construct a quotient estimator. By (2.2) and (2.3), we get that, for any $y \in \mathcal{I}$,

$$\begin{aligned}\lambda(y)(\phi_{z_0}^{-1})'(y)\mathbb{1}_{\{z_0\leq y\}}\mathbb{P}\left(Y_1 > y \middle| Z_0 = z_0\right) &= \mathcal{P}(z_0, y) \\ \lambda(y)\mathbb{E}\left(\mathbb{1}_{\{Z_0\leq y<Y_1\}}(\phi_{Z_0}^{-1})'(y) \middle| Z_0 = z_0\right) &= \mathcal{P}(z_0, y)\end{aligned}$$

and we integrate with respect to the stationary distribution μ of Z_0

$$\lambda(y)\mathbb{E}_\xi\left(\left(\phi_{Z_0}^{-1}\right)'(y)\mathbb{1}_{\{Z_0\leq y<Y_1\}}\right) = \int \mathcal{P}(z, y)\mu(dz) = \nu(y)$$

recalling that ξ is the stationary measure of the couple (Z_0, Y_1). Let us set

$$\mathbf{D}(y) := \mathbb{E}_\xi\left((\phi_{Z_0}^{-1})'(y)\mathbb{1}_{\{Z_0\leq y<Y_1\}}\right). \tag{2.4}$$

Then, if $\mathbf{D}(y) > 0$, we get:

$$\lambda(y) = \frac{\nu(y)}{\mathbf{D}(y)}. \tag{2.5}$$

Our aim is to estimate the jump rate λ on the compact interval $\mathcal{I} := [i1, i2] \subset (0, \infty)$. For that purpose, we assume:

Assumption (S)**.**

1. *The transition kernel is a contraction mapping: there exists $\kappa < 1$, such that*

$\mathbb{P}(Z_1 \leq \kappa Y_1) = 1$.

2. *The flow is bounded: there exist two functions* $\mathbf{m}$ *and* $\mathbf{M}$ *such that,* $\forall x, y \in (\mathbb{R}^+)^2$*:*

$$0 < \mathbf{m}(y) \leq (\phi_x^{-1})'(y) \leq \mathbf{M}(y).$$

3. *The jump rate is positive on* $[i_1, \infty[$ *and there exists* $\mathbf{a} > 0$*,* $b > -1$ *such that*

$$\forall y \geq i_1, \quad \lambda(y)\mathbf{m}(y) \geq \mathbf{a}\frac{y^b}{b+1}.$$

Then $\forall y \geq z$*,* $\mathbb{P}_z(Y_1 \geq y) \leq \exp(-\mathbf{a}(y^{b+1} - z^{b+1}))$ *and* $\lim_{y\to\infty} \mathbb{P}_z(Y_1 \geq y) = 0$.

4. *The jump rate does not explode too soon: there exist two positive constants* $\mathbf{L}, \mathbf{l}$*, such that* $\|\lambda\|_{L^\infty([i_1, i_2'])} \leq \mathbf{L}$ *and* $\int_0^{i_1} \lambda(u)\mathbf{M}(u)du \leq \mathbf{l}$ *where*

$$i_2' = \max\left(i_2, (i_2 - i_1) + \left(\frac{1}{\mathbf{a}(1-\kappa^{b+1})}\ln\left(\frac{2\kappa^{b+1}}{1-\kappa^{b+1}}\right)\right)^{1/(b+1)} \mathbb{1}_{\{\kappa^{b+1} \geq 1/3\}}\right).$$

These conditions ensure that $\mathbf{D}(y) > 0$ *and that the Markov chain* (Y_k, Z_k) *is geometrically* β*-mixing. The following two assumptions allow us to control the regularity of* ν *(the rate of convergence of the estimator* $\hat{\lambda}_n$ *depends on the regularity of* ν*, not on the regularity of* λ*).*

1. *For any* $y \in \mathbb{R}^+$*,* $\lambda(y) < \infty$*. This ensures that* ν *and* $\mathcal{P}$ *are continuous with respect to the Lebesgue measure on* $\mathbb{R}^+$*.*

2. *There exists* $\alpha > 0$ *such that:*

   - $\forall K \subset \mathbb{R}^{+*}$ *compact,* $\forall z \in \mathbb{R}^{+*}$*, the function* $(\phi_\cdot^{-1})'(.)$ *belongs to* $H^\alpha([0, z] \times K)$*.*
   - $\forall K \subset \mathbb{R}^{+*}$ *compact,* $\lambda \in H^\alpha(K)$*.*
   - *The transition measure* Q *can be written*

$$Q(x, dy) = Q_1(x, y)dy + p_0(x)\delta_0(dy) + \sum_{i=1}^{j_Q} p_i(x)\delta_{f_i(x)}(dy)$$

   with, for any compact K*,* Q_1 *and* $(p_i)_{0\leq i\leq j_Q}$ *in* $H^{\alpha-1}(K)$*, and* $(f_i)_{1\leq i\leq j_Q}$ *invertible functions such that* $(f_i^{-1})_{1\leq i\leq j_Q} \in H^\alpha(K)$*.*

Remark 2.1. *The Assumption (S) is quite technical but it has the advantage of being directly linked to the function we want to estimate and of not having to suppose that there is an invariant probability and to make assumptions about it, when we have no idea of what it is worth. This is often done in articles.*

If Assumption (S) is satisfied, for fixed flow ϕ and transition measure Q, we can introduce the class of functions

$$\mathcal{E}(\mathfrak{s}, b, \alpha) = \left\{ \lambda \in H^\alpha(\mathcal{J}), \forall y \geq i_1, \lambda(y)\mathbf{m}(y) \geq \frac{\mathbf{a}y^b}{b+1},\ \int_0^{i_1} \lambda(u)\mathbf{M}(u) \leq \mathbf{l},\ \|\lambda\|_{H^\alpha(\mathcal{J})} \leq \mathbf{L} \right\}$$

with $\mathfrak{s} = (\mathbf{a}, \mathbf{l}, \mathbf{L}) \in (\mathbb{R}^+)^3$ and the convex set

$$\mathcal{J} = \mathcal{J}_{\lfloor \alpha \rfloor} \cup [i_1, i_2'] := [j_1, j_2] \tag{2.6}$$

is defined by the recurrence:

$$\mathcal{J}_0 = \mathcal{I} \quad \text{and} \quad \mathcal{J}_{k+1} = \text{Conv}\left(\mathcal{I} \cup \bigcup_{i=1}^{j_Q} f_i^{-1}(\mathcal{J}_k) \right).$$

Note that we want to do the estimation on $\mathcal{I}$ but we need more technical assumptions on $\mathcal{J}$, so that the results are valid.

2.2.3 Estimation

(Kre16) and (AMG16) construct a pointwise kernel estimator of ν before deriving an estimator of λ. Indeed, densities are often approximated by kernels methods (see (Tsy04) for instance). If the kernel is positive, the estimator is also a density. However, we want to control the L^2 risk of our estimator (not the pointwise risk), and also to construct an adaptive estimator. Estimators by projection are well adapted for L^2 estimation: if they are longer to compute at a single point than pointwise estimators, it is sufficient to know the estimated coefficients to construct the whole function. Furthermore, to find an adaptive estimator, we minimize a function of the norm of our estimator, that is the sum of the square of the coefficients, and the dimension. That is the reason why we choose an estimation by projection.

We first aim at estimating ν on the compact set $\mathcal{I}$. We construct a sequence of L^2 estimators by projection on an orthonormal basis. As usual in nonparametric estimation, their risks can be decomposed in a variance term and a bias term which depends of the regularity of the density function ν. We choose to use the Besov spaces to characterize the regularity, which are well adapted to L^2 estimation (particularly for the wavelet decomposition). The "best" estimator is then selected by penalization. To construct the sequence of estimators, we introduce a sequence of vectorial subspaces S_m. We construct an estimator $\hat{\nu}_m$ of ν on each subspace and then select the best estimator $\hat{\nu}_{\hat{m}}$.

Assumption (4)**.**

1. *The subspaces S_m are increasing and have finite dimension D_m.*

2. *The L^2-norm and the L^∞-norm are connected:*

$$\exists \psi_1 > 0, \forall m \in \mathbb{N}, \forall s \in S_m, \quad \|s\|_\infty^2 \leq \psi_1 D_m \|s\|_{L^2}^2.$$

This implies that, for any orthonormal basis $(\varphi_l)_{1\leq l\leq D_m}$ of S_m,

$$\|\sum_{l=1}^{D_m} \varphi_l^2\|_\infty \leq \psi_1 D_m.$$

3. *There exists a constant $\psi_2 > 0$ such that, for any $m \in \mathbb{N}$, there exists an orthonormal basis $(\varphi_l)_{1\leq l\leq D_m}$ such that:*

$$\|\sum_{l=1}^{D_m} \|\varphi_l\|_\infty |\varphi_l(x)|\|_\infty \leq \psi_2 D_m.$$

4. *There exists $\mathbf{r} \in \mathbb{N}$, called the regularity of the decomposition, such that:*

$$\exists C > 0, \forall \alpha \leq \mathbf{r}, \forall s \in B_{2,\infty}^\alpha, \quad \|s - s_m\|_{L^2} \leq C D_m^{-\alpha} \|s\|_{B_{2,\infty}^\alpha}$$

where s_m is the orthogonal projection of s on S_m and $B_{2,\infty}^\alpha$ is a Besov space.

Conditions 1, 2 and 4 are usual (see (CGCR07, Section 2.3) for instance). They are satisfied for subspaces generated by wavelets, piecewise polynomials or trigonometric polynomials (see (DL93) for trigonometric polynomials and piecewise polynomials and (Mey90) for wavelets). Condition 3 is necessary because we are not in the stationary case: it helps us to control some covariance terms. It is obviously satisfied for bounded bases (trigonometric polynomials), and localized bases (piecewise polynomials). Let us prove it for a wavelet basis. Let φ be a father wavelet function, then $D_m = 2^m$ and $\varphi_l(x) = 2^{m/2}\varphi(2^m x - l)$. We get that $\|\sum_{l=1}^{D_m} \|\varphi_l\|_\infty |\varphi_l(x)|\|_\infty \leq 2^m \|\varphi\|_\infty \|\sum_{l\in\mathbb{Z}} |\varphi(x - l)|\|_\infty$. As φ is at least 0-regular, for $m = 2$, there exists a constant C such that $|\varphi(x)| \leq C(1 + |x|^{-2})$. Then $\sup_x \sum_{l\in\mathbb{Z}} |\varphi(x - l)| \leq C \sup_x \sum_{l\in\mathbb{Z}} (1 + |x - l|^{-2}) < \infty$ and condition 3 is satisfied.

Estimation of the stationary density

Let us now construct an estimator $\hat{\nu}_m$ of ν on the subspace S_m. We consider an orthonormal basis $(\varphi_l)_{1\leq l\leq D_m}$ of S_m satisfying Assumption A4. Let us set

$$a_l = \langle \varphi_l, \nu \rangle = \int_{\mathcal{I}} \varphi_l(x)\nu(x)dx \quad \text{and} \quad \nu_m(x) = \sum_{l=1}^{D_m} a_l \varphi_l(x).$$

The function ν_m is the orthogonal projection of ν on $L^2(\mathcal{I})$. We consider the estimator

$$\hat{\nu}_m(x) = \sum_{l=1}^{D_m} \hat{a}_l \varphi_l(x) \quad \text{with} \quad \hat{a}_l = \frac{1}{n}\sum_{k=1}^{n} \varphi_l(Y_k).$$

Proposition 2.1. *If $D_m^2 \leq n$, under Assumptions A1-A2 and A4, we have*

$$\mathbb{E}_{z_0}\left(\|\hat{\nu}_m - \nu\|^2_{L^2(\mathcal{I})}\right) \leq \|\nu_m - \nu\|^2_{L^2(\mathcal{I})} + (\psi_1 + C_\lambda \psi_2)\frac{D_m}{n} + \frac{c}{n}$$

where $C_\lambda = \frac{2R}{1-\gamma}\int \mathbf{V}_\lambda(z)\mu(dz)$ and c depends explicitly on $\mathbf{V}_\lambda$, γ, R.

When m increases, the bias term decreases whereas the variance term increases. It is important to find a good bias-variance compromise. If ν belongs to the Besov space $B^\alpha_{2,\infty}(\mathcal{I})$ (defined in the Section A.4. from the Supplemental Content of (KS21), then $\|\nu_m - \nu\|^2_{L^2(\mathcal{I})} \leq C\|\nu\|_{B^\alpha_{2,\infty}(\mathcal{I})} D_m^{-2\alpha}$ (see Assumption A4). We choose to use the Besov spaces to characterize the regularity, which are well adapted to L^2 estimation (particularly for the wavelet decomposition). If $\alpha \geq 1/2$, the risk is then minimum for $D_{m_{opt}} \propto n^{1/(2\alpha+1)}$ and we have, for some continuous function ψ:

$$\mathbb{E}_{z_0}\left(\|\hat{\nu}_{m_{opt}} - \nu\|^2_{L^2(\mathcal{I})}\right) \leq \psi\left(\|\nu\|_{B^\alpha_{2,\infty}(\mathcal{I})}, \mathbf{V}_\lambda, R, \gamma\right) n^{-2\alpha/(2\alpha+1)}.$$

This is the usual nonparametric convergence rate (see (Tsy04)). If $\alpha < 1/2$, then the risk is minimum for $D_m = n^{1/2}$ and the bias term is greater than the variance term. We can remark that a piecewise continuous function belongs to $B^{1/2}_{2,\infty}$.

Let us now construct the adaptive estimator. We compute $(\hat{\nu}_0, \ldots, \hat{\nu}_m, \ldots)$ for $m \in \mathcal{M}_n = \{m, D_m^2 \leq n\}$. Our aim is to select automatically m, without knowing the regularity of the stationary density ν. Let us introduce the contrast function $\gamma_n(s) = \|s\|^2_{L^2} - \frac{2}{n}\sum_{k=1}^n s(Y_k)$. If $s \in S_m$, then we can write $s = \sum_{l=1}^{D_m} b_l \varphi_l$ and

$$\gamma_n(s) = \sum_{l=1}^{D_m} b_l^2 - \sum_{l=1}^{D_m} b_l \frac{2}{n}\sum_{k=1}^{n} \varphi_l(Y_k).$$

The minimum is obtained for $b_l = \hat{a}_l = \frac{1}{n}\sum_{k=1}^n \varphi_l(Y_k)$. Therefore

$$\hat{\nu}_m = \arg\min_{s \in S_m} \gamma_n(s). \tag{2.7}$$

As the subspaces S_m are increasing, the function $\gamma_n(\hat{\nu}_m)$ decreases when m increases. To find an adaptive estimator, we need to add a penalty term $pen(m)$. Let us set $pen(m) = \frac{48(\psi_1 + C_\lambda\psi_2)D_m}{n} + \frac{48 c_\lambda \psi_1}{n}$ (or more generally $pen(m) = \frac{\sigma D_m}{n} + \frac{\sigma'}{n}$, with $\sigma \geq 48(\psi_1 + C_\lambda\psi_2)$, $\sigma' \geq 48 c_\lambda \psi_1$) and choose

$$\hat{m} = \arg\min_{m \in \mathcal{M}_n} \gamma_n(\hat{\nu}_m) + pen(m). \tag{2.8}$$

We obtain an adaptive estimator $\hat{\nu}_{\hat{m}}$.

Theorem 2.1 (Risk of the adaptive estimator)**.** *Under Assumptions A1-A2 and A4,* $\forall \sigma \geq 48(\psi_1 + C_\lambda \psi_2)$, $\sigma' \geq 48 c_\lambda \psi_1$, $pen(m) = \frac{\sigma D_m}{n} + \frac{\sigma'}{n}$,

$$\mathbb{E}_{z_0}\left(\|\nu - \hat{\nu}_{\hat{m}}\|^2_{L^2(\mathcal{I})}\right) \leq \min_{m \in \mathcal{M}_n}\left(3\|\nu_m - \nu\|^2_{L^2(\mathcal{I})} + 4pen(m)\right) + \frac{c'}{n}.$$

where c' *is a function of* $(\mathbf{V}_\lambda, R, \gamma, \|\nu\|_{L^2(\mathcal{I})})$. *We recall that* $\mathcal{M}_n = \{m, D_m^2 \leq n\}$.

The estimator is adaptive: it realizes the best bias-variance compromise, up to a multiplicative constant. We have an explicit rate of convergence if ν belongs to some (unknown) Besov space $B^\alpha_{2,\infty}$: in that case,

$$\|\nu - \nu_m\|^2_{L^2(\mathcal{I})} \leq 3\|\nu_{m_{\text{opt}}} - \nu\|^2_{L^2(\mathcal{I})} + 4pen(m_{\text{opt}}) + \frac{c}{n} \leq C\|\nu\|_{B^\alpha_{2,\infty}} D_m^{-2\alpha}$$

and if $\alpha \geq 1/2$,

$$\mathbb{E}_{z_0}\left(\|\nu - \hat{\nu}_{\hat{m}}\|^2_{L^2(\mathcal{I})}\right) \leq \psi\left(\|\nu\|_{B^\alpha_{2,\infty}(\mathcal{I})}, \mathbf{V}_\lambda, R, \gamma\right) n^{-2\alpha/(2\alpha+1)} \tag{2.9}$$

for some continuous function ψ.

Estimation of λ.

Remark 2.2. *1. We notice that this formula is different as the one used in (Kre16)*

$$\lambda(y) = \frac{f(\nu(y))}{\tilde{\mathbf{D}}(y)}$$

where

$$\tilde{\mathbf{D}}(y) := \mathbb{E}_\nu\left(((f \circ \phi_{Z_0})^{-1})'(f(y)) \mathbb{1}_{\{f(Z_0) \leq f(y)\}} \mathbb{1}_{\{Z_1 \geq f(y)\}}\right).$$

As in (Kre16), the author works under the assumption that $Q(x, \{y\}) = \mathbb{1}_{\{y = f(x)\}}$, *the study was easier, here we need to consider the Markov chain* $(Y_k, Z_k)_{k \in \mathbb{N}}$.

2. It is interesting to see that

$$\lambda(y) = \frac{\nu(y)}{\mathbf{D}(y)}$$

is a similar as the key representation for the bacterium case (1.14). *In both case, we manage to obtain a rewriting which allows us to find an estimator of* λ.

To estimate the jump rate, we construct a quotient estimator. Let us consider the estimator

$$\hat{\lambda}_n(y) = \frac{\hat{\nu}_{\hat{m}}(y)}{\hat{\mathbf{D}}_n(y)} \mathbb{1}_{\{\hat{\nu}_{\hat{m}}(y) \geq 0\}} \mathbb{1}_{\{\hat{\mathbf{D}}_n(y) \geq \ln(n)^{-1}\}} \tag{2.10}$$

where

$$\hat{\mathbf{D}}_n(y) := \frac{1}{n}\sum_{k=1}^{n}(\phi^{-1}_{Z_{k-1}})'(y)\mathbb{1}_{\{Z_{k-1}\le y\le Y_k\}}.$$

Remark 2.3. *As the process $\{X(t)\}_{t\ge 0}$ is observed continuously without errors, ϕ^{-1} (and therefore $(\phi^{-1})'$) is known on $\cup_{k\in\mathbb{N}^*}[Z_{k-1}, Y_k]$ so $\hat{\mathbf{D}}_n(y)$ is computable.*

The estimator $\hat{\lambda}_n$ converges with nearly the same rate of convergence as $\hat{\nu}$.

The process $(X_t)_{t\ge 0}$ is observed continuously without errors (so the flow ϕ is known). Another major difference with existing papers on the estimation of the jump rate of a PDMP is that assumptions are made to ensure the process to be ergodic, with fast convergence toward the stationary measure, and exponentially β-mixing. We denote by $(T_1, \dots, T_n)$ the jump times and consider the Markov chain $(Z_0 = x_0, (Y_k = X_{T_k^-}, Z_k = X_{T_k})_{k\in\mathbf{N}})$. Our aim is to construct a non-parametric adaptive estimator of the jump rate λ on a compact interval.

We find the same speed of convergence as in (Kre16) in $n^{-2\alpha/(\alpha+1)}$ up to a factor $\ln(n)^2$. We show this result uniformly on a good class of function. We refer to (KS21) for the definitions of spaces. We must specify that here contrary to (Kre16) the estimator does not depend on the regularity class of our λ, even if the speed depends on it. Moreover we show that this speed is indeed min-max.

Theorem 2.2. *Under A1, (S) and A4, as soon as $\ln(n)^{-1} \le D_0/2$, for any $\alpha \ge 1/2$,*

$$\sup_{\lambda\in\mathcal{E}(\mathfrak{s},b,\alpha)} \mathbb{E}_{z_0}\left(\|\hat{\lambda}_n - \lambda\|^2_{L^2(\mathcal{I})}\right) \lesssim \ln^2(n) n^{-2\alpha/(2\alpha+1)}.$$

This Theorem is Corollary 9 in (KS21).

In (KS21), it is well on to specify how we choose the penalization and there are results, specify on the increase of the speed of convergence of the estimator of the probability invariance, as well as that of λ.

We have proved that, under assumptions A1, (S) and A4,

$$\sup_{\lambda\in\mathcal{E}(\mathfrak{s},b,\alpha)} \mathbb{E}_{z_0}\left(\|\hat{\lambda}_n - \lambda\|^2_{L^2(\mathcal{I})}\right) \lesssim \ln^2(n) n^{-2\alpha/(2\alpha+1)}.$$

We would like to verify that our estimator converges with the minimax rate of convergence, i.e:

$$\inf_{\hat{\lambda}_n}\sup_{\lambda\in\mathcal{E}(\mathfrak{s},b,\alpha)} \mathbb{E}_{z_0}\left(\|\hat{\lambda}_n - \lambda\|^2_{L^2(\mathcal{I})}\right) \ge C\ln^2(n) n^{-2\alpha/(2\alpha+1)}.$$

The $\ln^2(n)$ factor comes from the quotient estimator, we can not expect it will stay in the minimax bound. Indeed, it is clear that we could replace $\ln^{-1}(n)$ in (2.10) by any function $w(n)$ greater than $D_0/2$. The best estimator will be obtained of course by taking $w(n) = D_0/2$ and the risk of this estimator (unreachable as D_0 is unknown) will be proportional to $n^{-2\alpha/(2\alpha+1)}$.

Theorem 2.3 (Minimax bound)**.** *If A1, (S) and A4 are satisfied, then*

$$\inf_{\hat{\lambda}_n} \sup_{\lambda \in \mathcal{E}(\mathfrak{s},b,\alpha)} \mathbb{E}_{z_0}\left(\|\hat{\lambda}_n - \lambda\|^2_{L^2(\mathcal{I})}\right) \geq C n^{-2\alpha/(2\alpha+1)}$$

where the infimum is taken among all estimators.

Some simulations for the TCP protocol and the bacterial growth are provided in (KS21), with various functions λ. The outcomes are consistent with the theoretical results.

2.3 To go further

After this generalization with Emeline Schmisser, a natural question arises: how can we generalize this to n dimensions? There will be two particular cases, which will be studied in the next chapter. They are particular models coming from biological problems, with complicated dependency structures. One of my research track would be to see how to study the growth rate of a PDMP in the framework of a PDMP with value in $\mathbb{R}_+^n$, and having a good branching structure in order to have the existence of a many-to-one formula. The goal is to have the existence of an equation of the type (1.11) for the framework of the modeling of the growth of bacteria or (0.17) within the framework of the fragmentations (even if in this case it is not a PDMP which appears but a subordinator). For that, the first step would be to see which family class for PDMPs I would like to consider. Indeed, the structure using the generation tree and the fact that at each node, we can put all the i.i.d. information we need is very powerful. For example, for the modeling of the growth of bacteria, it allowed to take into account the variability in the growth rate and that it is modeled by a Markov process. The model studied in the perspective of the previous chapter (in the joint work in progress with Bertrand Cloez, Benoîte de Saporta and Tristan Roget) would also be an example to include in my more general case study. A first step would be to define such a process, then show that it has the right properties and then estimate it.

Chapter 3

More dependency

One day Eva Löcherbach asked me to co-supervise Pierre Hodora who was doing a PhD thesis with her. They wanted to estimate the spiking rate for an interacting neural network.

Thus I had the chance to co-supervise Pierre Hodara on a part of his thesis and to discover the role of co-supervisor. It is a role that I really appreciated and that I would like to renew in the future.

The good thing is that we can model the "activity of a biological neural network" by a PDMP. On the other hand, the difficulty is that contrary to the case we had looked to model the growth of bacteria here we did not have that conditioning to their size at birth, the bacteria had independent behaviors. Neither have we a nice many-to-one formula. We could not therefore reduce ourselves to the study of a problem in dimension 1, and we had to really study the process with values in $\mathbb{R}^N$.

Building a model for the activity of a neural network that can fit biological considerations is crucial in order to understand the mechanics of the brain. Many papers in the literature use Hawkes processes in order to describe the spatio-temporal dependencies which are typical for huge systems of interacting neurons, see (GL13), (HRBR15) and (HL17) for example. Our model can be interpreted as Hawkes process with memory of variable length (see (GL16)); it is close to the model presented in (DO16). It is of crucial interest for modern neuro-mathematics to be able to statistically identify the basic parameters defining the dynamics of a model for neural networks. The most relevant mechanisms to study are the way the neurons are connected to each other and the way that a neuron deals with the information it receives. In (DGLO19) and in (HRBR15), the authors build an estimator for the interaction graph, in discrete or in continuous time. In the present work, we assume that we observe a subsystem of neurons which are all interconnected and behave in a similar way. We then focus on the estimation of the firing rate of a neuron within this system. This rate depends on the membrane potential of the neuron, which is influenced by the activity of the other neurons.

3.1 Interacting neurons

3.1.1 The dynamics

Let $N > 1$ be fixed and $(N^i(ds, dz))_{i=1,\dots,N}$ be a family of *i.i.d.* Poisson random measures on $\mathbb{R}_+ \times \mathbb{R}_+$ having intensity measure $dsdz$. We study the Markov process $(X_t)_{t\geq 0} = (X_t^1, \dots, X_t^N)_{t\geq 0}$ taking values in $[0, K]^N$ and solving, for $i = 1, \dots, N$, for $t \geq 0$,

$$\begin{aligned} X_t^i &= X_0^i - \lambda \int_0^t (X_s^i - m)ds - \int_0^t \int_0^\infty X_{s-}^i 1_{\{z\leq f(X_{s-}^i)\}} N^i(ds, dz) \\ &\quad + \sum_{j\neq i} \int_0^t \int_0^\infty a_K(X_{s-}^i) 1_{\{z\leq f(X_{s-}^j)\}} N^j(ds, dz). \end{aligned} \tag{3.1}$$

In the above equation, $\lambda > 0$ is a positive number, m is the equilibrium potential value such that $0 < m < K$. Moreover, we will always assume that $K \geq \frac{2}{N}$. Finally, the functions $a_K : [0, K] \to [0, K]$ and $f : \mathbb{R}_+ \mapsto \mathbb{R}_+$ satisfy (at least) the following assumption.

Assumption (B1)**.**
1. $a_K : [0, K] \to [0, \frac{1}{N}]$ is non-increasing and smooth, $a_K(x) = \frac{1}{N}$, for all $x < K - \frac{2}{N}$ and $a_K(x) < K - x$ for all $x \geq K - \frac{2}{N}$.
2. $f \in C^1(\mathbb{R}_+)$, f is non-decreasing, $f(0) = 0$, and there exists $f_{min} : \mathbb{R}_+ \mapsto \mathbb{R}_+$, non-decreasing, such that $f(x) \geq f_{min}(x) > 0$ for all $x > 0$.

K is the maximal height of the membrane potential of a single neuron. λ gives the speed of attraction of the potential value of each single neuron to an equilibrium value m. The function a_K denotes the increment of membrane potential received by a neuron when an other neuron fires. For neurons with membrane potential away from the bound K, this increment is equal to $\frac{1}{N}$. However, for neurons with membrane potential close to K, this increment may bring their membrane potential above the bound K. This is why we impose this dynamic close to the bound K.

In what follows, we are interested in the estimation of the intensity function f, assuming that the parameters K, f_{min} and a_K are known and that the function f belongs to a certain Hölder class of functions. The parameters of this class of functions are also supposed to be known. The assumption $f(0) = 0$ comes from biological considerations and expresses the fact that a neuron, once it has fired, has a refractory period during which it is not likely to fire.

The generator of the process X is given for any smooth test function $\varphi : [0, K]^N \to \mathbb{R}$ and $x \in [0, K]^N$ by

$$L\varphi(x) = \sum_{i=1}^N f(x_i) \left[\varphi(\Delta_i(x)) - \varphi(x)\right] - \lambda \sum_{i=1}^N \left(\frac{\partial \varphi}{\partial x_i}(x) \left[x_i - m\right]\right), \tag{3.2}$$

where

$$(\Delta_i(x))_j = \left\{ \begin{array}{ll} x_j + a_K(x_j) & j \neq i \\ 0 & j = i \end{array} \right\}. \tag{3.3}$$

For more details on the existence of such a process X, we can refer to (HKL18) and Theorem 9.1 in chapter IV of (IW81).

We denote by P_x the probability measure under which the solution $(X_t)_{t\geq 0}$ of (3.1) starts from $X_0 = x \in [0, K]^N$. Moreover, $P_\nu = \int_{[0,K]^N} \nu(dx) P_x$ denotes the probability measure under which the process starts from $X_0 \sim \nu$. Figure 3.1 is an example of trajectory for $N = 5$ neurons, choosing $f = Id$, $\lambda = 1$, $m = 1$, and $K = 2$.

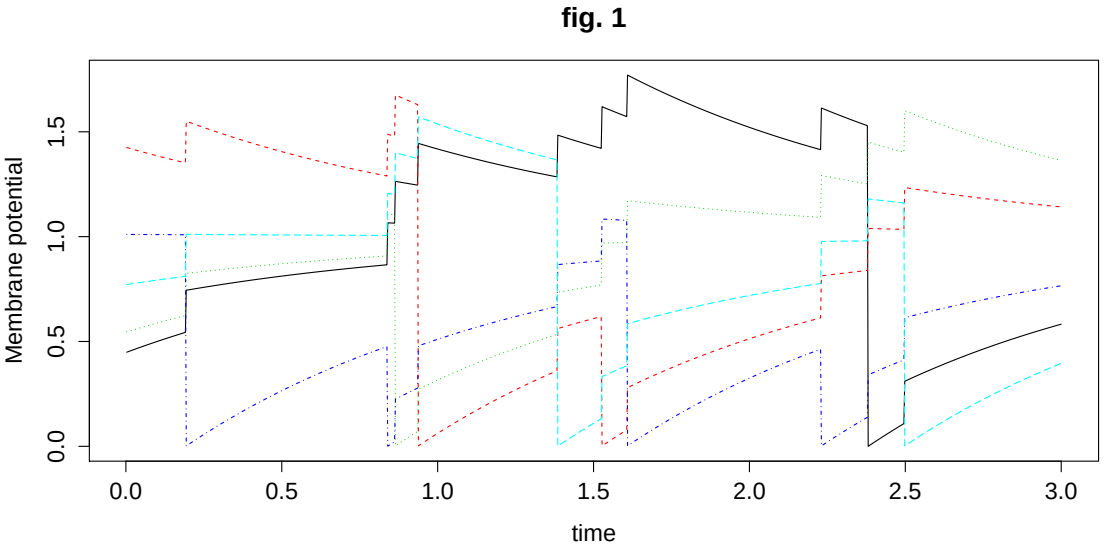


Figure 3.1: Trajectory of 5 neurons

The aim of this work is to estimate the unknown firing rate function f based on an observation of X continuously in time. Notice that for all $1 \leq i \leq N$, X^i reaches the value 0 only through jumps. Therefore, the following definition gives the successive spike times of the $i-$th neuron, $1 \leq i \leq N$. We set

$$T_0^i = 0, T_n^i = \inf\{t > T_{n-1}^i : X_{t-}^i > 0, X_t^i = 0\}, n \geq 1,$$

and introduce the jump measures

$$\mu^i(ds, dy) = \sum_{n\geq 1} 1_{\{T_n^i < \infty\}} \delta_{(T_n^i, X^i_{T_n^i -})}(dt, dy), \quad \mu(dt, dx) = \sum_{i=1}^N \mu^i(ds, dx).$$

By our assumptions, μ^i is compensated by $\hat{\mu}^i(ds, dy) = f(X_s^i) ds \delta_{X_s^i}(dy)$, and therefore

the compensator $\hat{\mu}$ of μ is given by

$$\hat{\mu}(dt, dy) = f(y)\eta(dt, dy), \text{ where } \eta(A \times B) = \int_A \left(\sum_{i=1}^{N} 1_B(X_s^i) \right) ds$$

is the total occupation time measure of the process X.

We will also write $T_n, n \geq 0$, for the successive jump times of the process X, *i.e.*

$$T_0 = 0, T_n = \inf\{T_k^i : T_k^i > T_{n-1}, k \geq 1, 1 \leq i \leq N\}, n \geq 1.$$

To estimate the jump rate f in a position a, we propose a Nadaraya-Watson type kernel estimator which is roughly speaking of the form

$$\hat{f}_t(a) = \frac{\sharp \text{ spikes in positions in } B_h(a) \text{ during } [0,t]}{\text{occupation time of } B_h(a) \text{ during } [0,t]},$$

where $B_h(a)$ is a neighborhood of size h of the position a where we estimate the jump rate function f.

More precisely, for some kernel function Q such that

$$Q \in C_c(\mathbb{R}), \int_{\mathbb{R}} Q(y)dy = 1, \tag{3.4}$$

we define the kernel estimator for the unknown function f at a point a with bandwidth h, based on observation of X up to time t by

$$\hat{f}_{t,h}(a) = \frac{\int_0^t \int_{\mathbb{R}} Q_h(y-a)\mu(ds, dy)}{\int_0^t \int_{\mathbb{R}} Q_h(y-a)\eta(ds, dy)}, \text{ where } Q_h(y) := \frac{1}{h} Q\left(\frac{y}{h}\right) \text{ and } \frac{0}{0} := 0. \tag{3.5}$$

For h small, $\hat{f}_{t,h}(a)$ is a natural estimator for $f(a)$. Indeed, this expression as a ratio follows the intuitive idea to count the number of jumps that occurred with a position close to a and to divide by the occupation time of a neighborhood of a, which is natural to estimate an intensity function depending on the position a. More precisely, by the martingale convergence theorem, the numerator $\int_0^t \int_{\mathbb{R}} Q_h(y-a)\mu(ds, dy)$ should behave, for t large, as $\int_0^t \int_{\mathbb{R}} Q_h(y-a)f(y)\eta(ds, dy)$. But by the ergodic theorem,

$$\frac{\int_0^t \int_{\mathbb{R}} Q_h(y-a)f(y)\eta(ds, dy)}{\int_0^t \int_{\mathbb{R}} Q_h(y-a)\eta(ds, dy)} \to \frac{\pi_1(Q_h(\cdot - a)f)}{\pi_1(Q_h(\cdot - a))}$$

as $t \to \infty$, where π_1 is the stationary measure of each neuron $(X_t^i)_{t \geq 0}$. Finally, if the invariant measure π_1 is sufficiently regular, then

$$\frac{\pi_1(Q_h(\cdot - a)f)}{\pi_1(Q_h(\cdot - a))} \to f(a)$$

as $h \to 0$.

We restrict our study to fixed Hölder classes of rate functions f. For that sake, we introduce the notation $\beta = k + \alpha$ for $k = \lfloor \beta \rfloor \in \mathbb{N}$ and $0 \leq \alpha < 1$. We consider the following Hölder class for arbitrary constants $F, L > 0$, and a function f_{min} as in Assumption B1.

$$H(\beta, F, L, f_{min}) = \{f \in C^k(\mathbb{R}_+) : |\frac{d^l}{dx^l} f(x)| \leq F, \text{ for all } 0 \leq l \leq k, x \in [0, K], \\ f(x) \geq f_{min}(x) \text{ for all } x \in [0, K], \ |f^{(k)}(x) - f^{(k)}(y)| \leq L|x-y|^\alpha \text{ for all } x, y \in [0, K]\}. \tag{3.6}$$

3.1.2 Probabilistic results

In this section, we collect important probabilistic results. We first establish that the process $(X_t)_{t \geq 0}$ is recurrent in the sense of Harris.

Theorem 3.1. *Grant Assumption B1. Then the process X is positive Harris recurrent having unique invariant probability measure π, i.e. for all $B \in \mathcal{B}([0, K]^N)$,*

$$\pi(B) > 0 \ \textit{ implies } P_x \left(\int_0^\infty 1_B(X_s) ds = \infty \right) = 1 \tag{3.7}$$

for all $x \in [0, K]^N$. Moreover, there exist constants $C > 0$ and $\kappa > 1$ which do only depend on the class $H(\beta, F, L, f_{min})$, but not on f, such that

$$\sup_{f \in H(\beta, F, L, f_{min})} \|P_t(x, \cdot) - \pi\|_{TV} \leq C\kappa^{-t}. \tag{3.8}$$

It is well-known that the behavior of a kernel estimator such as the one introduced in (3.5) depends heavily on the regularity properties of the invariant probability measure of the system. Our system is however very degenerate. Firstly, it is a PDMP in dimension N, with interactions between particles. Hence, no Brownian noise is present to smoothen the dynamic. Moreover, the transition kernels associated to the jumps of system (3.1) are highly degenerate (recall (3.3)). The transition kernel

$$K(x, dy) = \mathcal{L}(X_{T_1} | X_{T_1 -} = x)(dy) = \sum_{i=1}^N \frac{f(x^i)}{\bar{f}(x)} \delta_{\Delta^i(x)}(dy)$$

with $\bar{f}(x) := \sum_{i=1}^N f(x^i)$ puts one particle (the one which is just spiking) to the level 0. As a consequence, the above transition does not create density – and it even destroys smoothness due to the reset to 0 of the spiking neuron. Finally, the only way that "smoothness" is generated by such process is the smoothness which is present in the "noise of the jump times" (which are basically of exponential density). For this reason, we have to stay away from the point $x = m$, where the drift of the flow vanishes. Moreover, the reset-to-0 of the spiking particles implies that we are not

able to say anything about the behavior of the invariant density of a single particle in 0 (actually, near to 0) neither. Finally, we also have to stay strictly below the upper bound of the state space K. This is the reason for introducing the following open set $S_{d,\beta}$ given by

$$S_{d,\beta} := \{w \in [0,K] : \frac{\lfloor\beta\rfloor}{N} < w < K - \frac{\lfloor\beta\rfloor}{N}, |w-m| > d\}, \tag{3.9}$$

where β is the smoothness of the fixed class $H(\beta, F, L, f_{min})$ that we consider and where d is fixed such that $d > \frac{\lfloor\beta\rfloor+2}{N}$. Notice that $S_{d,\beta}$ also depends on K, m and N which are supposed to be known. We are able to obtain a control of the invariant measure only on this set $S_{d,\beta}$. The dependence in β is due to the fact that the regularity of f is transmitted to the invariant measure by the means of successive integration by parts (see (Löc18) for more details).

We quote the following theorem from (Löc18).

Theorem 3.2. *(Theorem 5 of (Löc18))*
Suppose that $f \in H(\beta, F, L, f_{min})$. Let

$$\pi_1 := \mathcal{L}_\pi(X_t^1)$$

be the invariant measure of a single neuron, i.e. $\int g d\pi_1 = E_\pi(g(X_t^1))$. Then π_1 possesses a bounded continuous Lebesgue density π^1 on $S_{d,\beta}$ for any d such that $d > (\lfloor\beta\rfloor + 2)/N$, which is bounded on $S_{d,\beta}$, uniformly in $f \in H(\beta, F, L, f_{min})$. Moreover, $\pi^1 \in C^k(S_{d,\beta})$ and

$$\sup_{\ell \le \lfloor\beta\rfloor, w \in S_{d,\beta}} |\pi_1^{(\ell)}(w)| + \sup_{w \neq w', w, w' \in S_{d,\beta}} \frac{\pi_1^{(\lfloor\beta\rfloor)}(w) - \pi_1^{(\lfloor\beta\rfloor)}(w')}{|w-w'|^\alpha} \le C_F, \tag{3.10}$$

where the constant C_F depends on d and on the smoothness class $H(\beta, F, L, f_{min})$, but on nothing else.

3.1.3 Statistical results

We can now state the main theorem of (HKL18) which describes the quality of our estimator in the minimax theory. We assume that m and λ are known and that f is the only parameter of interest of our model. We shall always write P_x^f and E_x^f in order to emphasize the dependence on the unknown f. Fix some $r > 0$ and some suitable point $a \in S_{d,\beta}$. For any possible rate of convergence $(r_t)_{t\ge 0}$ increasing to ∞ and for any process of $\mathcal{F}_t-$measurable estimators $\hat{f}_t$ we shall consider pointwise square risks of the type

$$\sup_{f \in H(\beta,F,L,f_{min})} r_t^2 E_x^f \left[|\hat{f}_t(a) - f(a)|^2 | A_{t,r} \right],$$

where

$$A_{t,r} := \left\{ \frac{1}{Nt} \int_0^t \int_{\mathbb{R}} Q_h(y-a)\eta(ds,dy) \geq r \right\}$$

is roughly the event ensuring that sufficiently many observations have been made near a, during the time interval $[0,t]$. We are able to choose r small enough such that

$$\liminf_{t\to\infty} \inf_{f\in H(\beta,F,L,f_{min})} P_x^f(A_{t,r^*}) = 1, \tag{3.11}$$

see Proposition 8 in (HKL18).

Recall that the kernel Q is chosen to be of compact support. Let us write R for the diameter of the support of Q, therefore $Q(x) = 0$ if $|x| \geq R$. For any fixed $a \in S_{d,\beta}$, write $h_0 := h_0(a,R,\beta,d) := \sup\{h > 0 : B_{hR}(a) \subset S_{d/2,\beta}\}$. Here, $B_{hR}(a) = \{y \in \mathbb{R}_+ : |y-a| < hR\}$.

Theorem 3.3. *Let $f \in H(\beta,F,L,f_{min})$ and choose $Q \in C_c(\mathbb{R})$ such that $\int_{\mathbb{R}} Q(y)y^j dy = 0$ for all $1 \leq j \leq \lfloor\beta\rfloor$, and $\int_{\mathbb{R}} |y|^\beta Q(y)dy < \infty$. Then there exists $r^* > 0$ such that the following holds for any $a \in S_{d,\beta}, r \leq r^*$ and for any $h_t \leq h_0$.*

(i) For the kernel estimate (3.5) with bandwidth $h_t = t^{-\frac{1}{2\beta+1}}$, for all $x \in [0,K]$,

$$\limsup_{t\to\infty} \sup_{f\in H(\beta,F,L,f_{min})} t^{\frac{2\beta}{2\beta+1}} E_x^f \left[|\hat{f}_{t,h_t}(a) - f(a)|^2 | A_{t,r} \right] < \infty.$$

(ii) Moreover, for $h_t = o(t^{-1/(1+2\beta)})$, for every $f \in H(\beta,F,L,f_{min})$ and $a \in S_{d,\beta}$

$$\sqrt{th_t}\left(\hat{f}_{t,h_t}(a) - f(a)\right) \to \mathcal{N}(0,\Sigma(a))$$

weakly under P_x^f, where $\Sigma(a) = \frac{f(a)}{N\pi_1(a)} \int Q^2(y)dy$.

The next theorem shows that the rate of convergence achieved by the kernel estimate $\hat{f}_{t,t^{-1/(2\beta+1)}}$ is indeed optimal.

Theorem 3.4. *Let $a \in S_{d,\beta}$ and $x \in [0,K]$ be any starting point. Then we have*

$$\liminf_{t\to\infty} \inf_{\hat{f}_t} \sup_{f\in H(\beta,F,L,f_{min})} t^{\frac{2\beta}{1+2\beta}} E_x^f[|\hat{f}_t(a) - f(a)|^2] > 0, \tag{3.12}$$

where the infimum is taken over the class of all possible estimators $\hat{f}_t(a)$ of $f(a)$.

3.1.4 Simulation results

In this subsection, we present some results on simulations, for different jump rates f. The other parameters are fixed: $N = 100,\ \lambda = 1,\ K = 2$ and $m = 1$. The dynamics of the system are the same when λ and f have the same ratio. In other words, variations of λ and f keeping the same ratio between the two parameters lead to the

same law for the process rescaled in time. This is why we fix $\lambda = 1$ and propose different choices for f. The kernel Q used here is a truncated Gaussian kernel with standard deviation 1.

We present for each choice of a jump rate function f the associated estimated function $\hat{f}$ and the observed distribution of X or more precisely of $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X^i$. Figures 2, 3 and 4 correspond respectively to the following definitions of f : $f(x) = x$, $f(x) = \log(x+1)$ and $f(x) = \exp(x) - 1$.

For Figures 3.2, 3.3 and 3.4, we fixed the length of the time interval for observations respectively to $t = 200$, 300 and 150. This allows us to obtain a similar number of jump for each simulation, respectively equal to 17324, 18579 and 21214. These simulations are realized with the software R.

The optimal bandwidth $h_t = t^{-\frac{1}{2\beta+1}}$ depends on the regularity of f given by the parameter β. Therefore, we propose a data-driven bandwidth chosen according to a Cross-validation procedure. For that sake, we define the sequence $(Z_k)_{k\in\mathbb{N}^*}$ by $Z_k^i = X^i_{T_k^i -}$ for all $1 \leq i \leq N$. For each $a \in [0, K]$ and each sample $Z = (Z_1, ..., Z_n)$, for $1 \leq \ell \leq n$ we define the random variable $\hat{\pi}_1^{\ell,n,h}(a)$ by

$$\hat{\pi}_1^{\ell,n,h}(a) = \frac{1}{(n-\ell)N} \sum_{k=\ell+1}^{n} \sum_{i=1}^{N} Q_h(Z_k^i - a).$$

$\hat{\pi}_1^{\ell,n,h}(a)$ can be seen as an estimator of the invariant measure π_1^Z of the discrete Markov chain.

We propose an adaptive estimation procedure at least, for this simulation part. We use a Smoothed Cross-validation (SCV) to choose the bandwidth (see for example the paper of Hall, Marron and Park (HMP92)), based on ideas which were first published by Bowmann (Bow84) and Rudemo (Rud82). As the bandwidth is mainly important for the estimation of the invariant probability π_1^Z, we use a Cross validation procedure for this estimation. More precisely, we use a first part of the trajectory to estimate $\widehat{\pi}_1^{\ell,n,h}$ and then another part of the trajectory to minimize the Cross validation $SCV(h)$ in h. In order to be closer to the stationary regime, we chose the two parts of the trajectory far from the starting time. Moreover we chose two parts of the trajectory sufficiently distant from each other. This is why we consider m_1, m_2 and ℓ such that $1 \langle m_1 \leq m_2 \langle \ell \leq n$.

We use the method of the least squares Cross validation and minimize

$$SCV(h) = \int \left(\widehat{\pi}_1^{\ell,n,h}(x)\right)^2 dx - \frac{2}{N(m_2 - m_1)} \sum_{k=m_1+1}^{m_2} \sum_{i=1}^{N} \widehat{\pi}_1^{\ell,n,h}(Z_k^i)$$

(where we have approximated the integral term by a Riemann approximation), giving rise to a minimizer $\hat{h}$. We then calculate the estimator $\hat{f}$ along the trajectory. In the next figure, we use this method to find the reconstructed f with an adaptive choice of h.

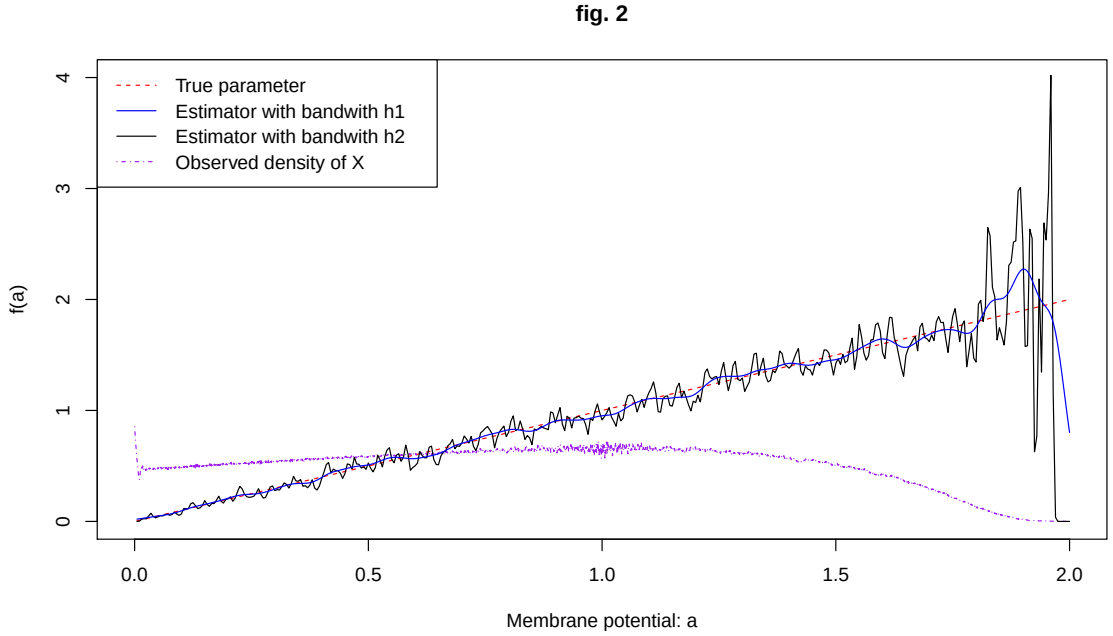


Figure 3.2: Estimation of the intensity function $f(x) = x$

On Figure 3.2, it is interesting to see, that the two ways of finding the bandwith give differents results. Minimizing the log-likelihood gives a smoother estimator.

As expected, we can see that the less observations we have, the worse is our estimator. Note that close to 0 the observed density of X explodes. This was also expectable due to the reset to 0 of the jumping neurons, and this may explain why our estimator is less performing close to 0.

Moreover, the simulations show a lack of regularity of the observed density close to m, which is consistent with our results, but this does not seem to affect the quality of the estimator.

For the sake of readability, the observed density of X on Figure 3.4 has been multiplied by 3.

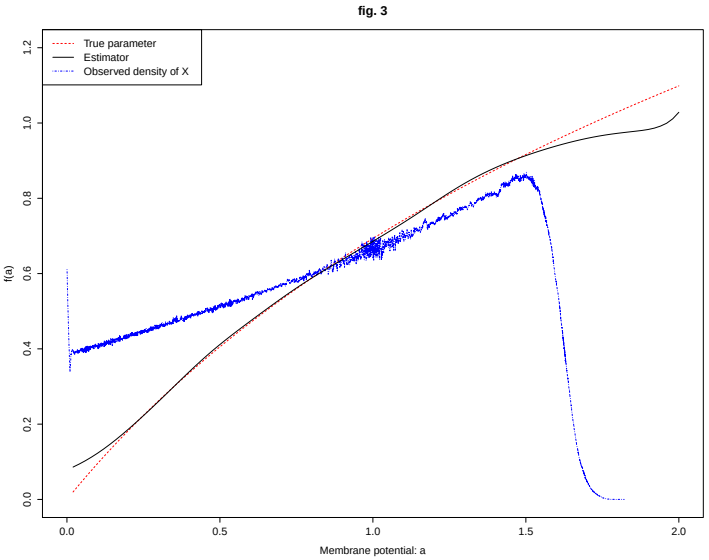


Figure 3.3: Estimation of the intensity function $f(x) = \log(x + 1)$

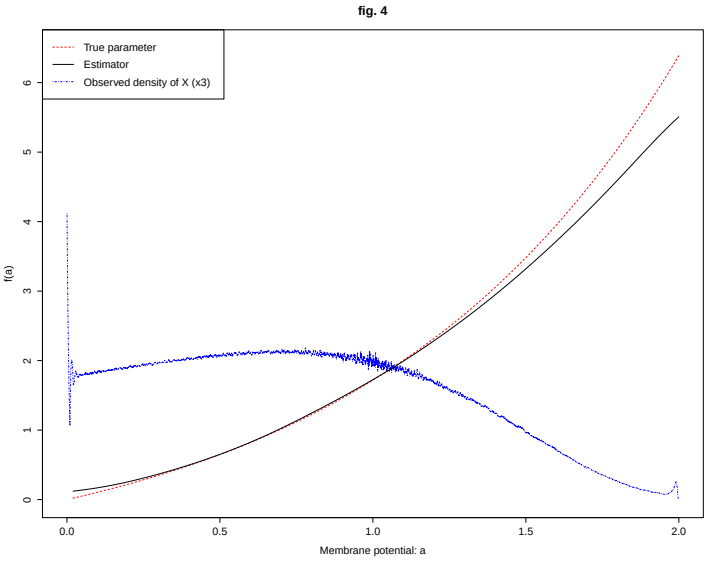


Figure 3.4: Estimation of the intensity function $f(x) = \exp(x) - 1$

3.2 To go further: Stochastic differential equation depending on the rank

During a workshop, I saw the presentation by Fabrice Mahé of a model of the growth of rainbow trout in rearing. I thought that this could be very interesting to model. So we had discussions in order to better understand their growth. During the discussions it appears that there is competition between fish and in particular that the most fearful animals could have less access to food and thus have a lesser growth and conversely. A priori this dominant character of a fish could change over time. With

Hélène Guérin, we thought that it could be very interesting to try to model this. So we were interested in a model in which the growth factors of the fish would depend on its rank within the population.

Let $N \geq 1$. We consider a ranked based interacting diffusion model defined by

$$\mathrm{d}X_t^{i,N} = \sum_{k=1}^{N} b_k^N(t, X_t^{i,N}) \mathbb{1}_{\{X_t^{i,N}=X_t^{(k)}\}} \mathrm{d}t + \sum_{k=1}^{N} \sigma_k^N \mathbb{1}_{\{X_t^{i,N}=X_t^{(k)}\}} \mathrm{d}W_t^i \tag{3.13}$$

where $b_k^N : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ are the growth rate coefficients, σ_k^N are the diffusive coefficients, $\{W^i\}_{1\leq i\leq N}$ are N independent Brownian motions, and $X^{(1)}, X^{(2)}, \ldots, X^{(N)}$ are the order statistics of the N-uplet $(X^{1,N}, X^{2,N}, \ldots, X^{N,N})$: i.e.for all $t \geq 0$,

$$\min_{1\leq i\leq N} X_t^{i,N} = X_t^{(1)} \leq X_t^{(2)} \leq \ldots \leq X_t^{(N)} = \max_{1\leq i\leq N} X_t^{i,N}.$$

When the coefficients $(b_k^N)_{k\in\{1,\ldots,N}$ are constants, we recover the Atlas model, which has been first introduced to model equity markets (see e.g. (Fer02; BFK05)): for $1 \leq i \leq N$

$$\mathrm{d}X_t^{i,N} = \sum_{k=1}^{N} \mathbb{1}_{\{X_t^{i,N}=X_t^{(k)}\}} b_k^N \mathrm{d}t + \sum_{k=1}^{N} \mathbb{1}_{\{X_t^{i,N}=X_t^{(k)}\}} \sigma_k^N \mathrm{d}W_t^i. \tag{3.14}$$

At the beginning we wanted to make b and σ as general as possible and quickly we were confronted with a lot of difficulties. Indeed, the discontinuity of the terms poses important problems in the demonstrations. In the end, we have the feeling that there are very few cases that have been studied in the literature and that could generalize the Atlas model as we would have liked. So for the moment, we are interested in a simplified model where there are only two fishes in interaction. The goal is of course to show that the equation (3.14) admits a weak solution, and it is unique. We also want to make sure that the solution is positive. Indeed, a negative fish size could be a problem and in a second time we would like to estimate the different parameters of the problem. This is a work in progress.

Bibliography

[Aal75] Odd Olai Aalen. *Statistical inference for a family of counting porcesses. ProQuest LLC, Ann Arbor, MI, 1975. Thesis (Ph.D.)–University of California, Berkeley.*

[Aal77] Odd Olai Aalen. Weak convergence of stochastic integrals related to counting processes. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 38(4):261–277, 1977.*

[Aal78] Odd Olai Aalen. Nonparametric inference for a family of counting processes. Ann. Statist.*, 6(4):701–726, 1978.*

[ABGKZ14] Romain Azaïs, Jean-Baptiste Bardet, Alexandre Génadot, Nathalie Krell, and Pierre-André Zitt. Piecewise deterministic Markov process—recent results. In Journées MAS 2012*, volume 44 of* ESAIM Proc.*, pages 276–290. EDP Sci., Les Ulis, 2014.*

[ADGP14] Romain Azaïs, François Dufour, and Anne Gégout-Petit. Nonparametric estimation of the conditional distribution of the interjumping times for piecewise-deterministic Markov processes. Scand. J. Stat.*, 41(4):950–969, 2014.*

[AG18] Romain Azaïs and Alexandre Genadot. A new characterization of the jump rate for piecewise-deterministic Markov processes with discrete transitions. Comm. Statist. Theory Methods*, 47(8):1812–1829, 2018.*

[AMG16] Romain Azaïs and Aurélie Muller-Gueudin. Optimal choice among a class of nonparametric estimators of the jump rate for piecewise-deterministic Markov processes. Electron. J. Stat.*, 10(2):3648–3692, 2016.*

[Ban09] Vincent Bansaye. Cell contamination and branching processes in a random environment with immigration. Adv. in Appl. Probab.*, 41(4):1059–1081, 2009.*

[Bas06] Anne-Laure Basdevant. Fragmentation of ordered partitions and intervals. Electron. J. Probab.*, 11:no. 16, 394–417, 2006.*

[BDMT11] *Vincent Bansaye, Jean-François Delmas, Laurence Marsalle, and Viet Chi Tran. Limit theorems for Markov processes indexed by continuous time Galton-Watson trees.* Ann. Appl. Probab., *21(6):2263–2314, 2011.*

[Ber97] *Jean Bertoin. Exponential decay and ergodicity of completely asymmetric Lévy processes in a finite interval.* Ann. Appl. Probab., *7(1):156–169, 1997.*

[Ber01] *Jean Bertoin. Homogeneous fragmentation processes.* Probab. Theory Related Fields, *121(3):301–318, 2001.*

[Ber03] *Julien Berestycki. Multifractal spectra of fragmentation processes.* J. Statist. Phys., *113(3-4):411–430, 2003.*

[Ber06] *Jean Bertoin.* Random fragmentation and coagulation processes, *volume 102 of* Cambridge Studies in Advanced Mathematics. *Cambridge University Press, Cambridge, 2006.*

[BFK05] *Adrian D. Banner, Robert Fernholz, and Ioannis Karatzas. Atlas models of equity markets.* Ann. Appl. Probab., *15(4):2296–2330, 2005.*

[BLM15] *Michel Benaïm, Stéphane Le Borgne, Florent Malrieu, and Pierre-André Zitt. Qualitative properties of certain piecewise deterministic Markov processes,* Ann. Inst. Henri Poincaré Probab. Stat., **51**, *3, 1040-1075, 2015.*

[BHM22] *Peter Braunsteins, Sophie Hautphenne, and Carmen Minuesa. Parameter estimation in branching processes with almost sure extinction.* Bernoulli, *28(1):33–63, 2022.*

[BM05] *Jean Bertoin and Servet Martínez. Fragmentation energy.* Adv. in Appl. Probab., *37(2):553–570, 2005.*

[BM05b] *Jean Bertoin and Alain Rouault. Asymptotical behaviour of the presence probability in branching random walks and fragmentations.* https://hal.archives-ouvertes.fr/hal-00002955, *2005.*

[Bow84] *Adrian W. Bowman. An alternative method of cross-validation for the smoothing of density estimates.* Biometrika, *71(2):353–360, 1984.*

[BR05] *Jean Bertoin and Alain Rouault. Discretization methods for homogeneous fragmentations.* J. London Math. Soc. (2), *72(1):91–109, 2005.*

[BT11] *Vincent Bansaye and Viet Chi Tran. Branching Feller diffusion for cell division with parasite infection.* ALEA Lat. Am. J. Probab. Math. Stat., *8:95–127, 2011.*

[BT21] *Elena Bandini and Michèle Thieullen. Optimal control of infinite-dimensional piecewise deterministic Markov processes: a BSDE approach. Application to the control of an excitable cell membrane.* Appl. Math. Optim., *84(2):1549–1603, 2021.*

[BT21] *Bertrand Cloez, Renaud Dessalles, Alexandre Genadot, Florent Malrieu, Aline Marguet and Romain Yvinec. Probabilistic and piecewise deterministic models in biology. In Journées MAS 2016 de la SMAI - Phénomènes Complexes et Hétérogènes.* ESAIM Proc. Surveys *60 225-245. Les Ulis: EDP Sci. 2017.*

[CGCR07] *Fabienne Comte, Valentine Genon-Catalot, and Yves Rozenholc. Penalized nonparametric mean square estimation of the coefficients of diffusion processes.* Bernoulli, *13(2):514–543, 2007.*

[Clo17] *Bertrand Cloez. Limit theorems for some branching measure-valued processes.* Adv. in Appl. Probab., *49(2):549–580, 2017.*

[CRW91] *Brigitte Chauvin, Alain Rouault, and Anton Wakolbinger. Growing conditioned trees.* Stochastic Process. Appl., *39(1):117–130, 1991.*

[Dav84] *M. H. A. Davis. Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models.* J. Roy. Statist. Soc. Ser. B, *46(3):353–388, 1984. With discussion.*

[Dav93] *M. H. A. Davis.* Markov models and optimization, *volume 49 of* Monographs on Statistics and Applied Probability. *Chapman & Hall, London, 1993.*

[DD02] *F. Dufour and Y. Dutuit,* Dynamic reliability: A new model, *Proceedings of ESREL 2002 Lambda-Mu 13 Conference, 2002, pp. 350–353.*

[DGLO19] *Aline Duarte, Antonio Galves, Eva Löcherbach, and Guilherme Ost. Estimating the interaction graph of stochastic neural dynamics.* Bernoulli, *25(1):771–792, 2019.*

[DGR02] *Vincent Dumas, Fabrice Guillemin, and Philippe Robert. A Markovian analysis of additive-increase multiplicative-decrease algorithms.* Adv. in Appl. Probab., *34(1):85–111, 2002.*

[DHK+14] *Marie Doumic, Marc Hoffmann, Nathalie Krell, Stéphane Aymerich, Jérome Robert, and Lydia Robert. Division control in escherichia coli*

is based on a size-sensing rather than timing mechanism. BMC Biology*, 12:17, 2014.*

[DHK21] A. C. Davison, S. Hautphenne, and A. Kraus. Parameter estimation for discretely observed linear birth-and-death processes. Biometrics*, 77(1):186–196, 2021.*

[DHKR15] Marie Doumic, Marc Hoffmann, Nathalie Krell, and Lydia Robert. Statistical estimation of a growth-fragmentation model observed on a genealogical tree. Bernoulli*, 21(3):1760–1799, 2015.*

[DHRBR12] Marie Doumic, Marc Hoffmann, Patricia Reynaud-Bouret, and Vincent Rivoirard. Nonparametric estimation of the division rate of a size-structured population. SIAM J. Numer. Anal.*, 50(2):925–950, 2012.*

[DL93] Ronald A. DeVore and George G. Lorentz. Constructive approximation*, volume 303 of* Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1993.*

[DMZ10] M. Doumic, P. Maia, and J.P. Zubelli. On the calibration of a size-structured population model from experimental data. Acta Biotheoretica*, 2010.*

[DO16] A. Duarte and G. Ost. A model for neural activity in the absence of external stimuli. Markov Process. Related Fields*, 22(1):37–52, 2016.*

[DPZ09] Marie Doumic, Benoît Perthame, and Jorge P. Zubelli. Numerical solution of an inverse problem in size-structured population dynamics. Inverse Problems*, 25(4):045008, 25, 2009.*

[dSGPM11] Benoîte de Saporta, Anne Gégout-Petit, and Laurence Marsalle. Parameters estimation for asymmetric bifurcating autoregressive processes with missing data. Electron. J. Stat.*, 5:1313–1353, 2011.*

[dSGPM12] Benoîte de Saporta, Anne Gégout-Petit, and Laurence Marsalle. Asymmetry tests for bifurcating auto-regressive processes with missing data. Statist. Probab. Lett.*, 82(7):1439–1444, 2012.*

[dSGPM14] Benoîte de Saporta, Anne Gégout-Petit, and Laurence Marsalle. Statistical study of asymmetry in cell lineage data. Comput. Statist. Data Anal.*, 69:15–39, 2014.*

[DSKR18] Bernard Delyon, Benoîte de Saporta, Nathalie Krell, Lydia Robert. Investigation of asymmetry in E. coli growth rate Case Studies in Business, Industry and Government Statistics, Société Française de Statistique*, 7 (1), pp.1-13, 2018.*

[EO05] *R. Erban and H. G. Othmer, From individual to collective behavior in bacterial chemotaxis,* SIAM J. Appl. Math. ***65*** *, no. 2, 361-391 (electronic) (2004/05).*

[Fer02] *E. Robert Fernholz. Stochastic portfolio theory, volume 48 of* Applications of Mathematics (New York). Springer-Verlag, *New York, 2002. Stochastic Modelling and Applied Probability.*

[FGM16] *Joaquín Fontbona, Hélène Guérin, and Florent Malrieu. Long time behavior of telegraph processes under convex potentials.* Stochastic Process. Appl., ***126****, 10, 3077-3101, 2016.*

[FKM10] *Joaquín Fontbona, Nathalie Krell, and Servet Martínez. Energy efficiency of consecutive fragmentation processes.* J. Appl. Probab., *47(2):543–561, 2010.*

[Fuj13] *Takayuki Fujii. Nonparametric estimation for a class of piecewise-deterministic Markov processes.* J. Appl. Probab., *50(4):931–942, 2013.*

[GBP+05] *Julien Guyon, Ariane Bize, Grégory Paul, Eric Stewart, Jean-Francois Delmas, and Francois Taddéi. Statistical study of cellular aging. In* CEMRACS 2004—mathematics and applications to biology and medicine*, volume 14 of* ESAIM Proc.*, pages 100–114. EDP Sci., Les Ulis, 2005.*

[GHR04] *Emmanuel Gobet, Marc Hoffmann, and Markus Reiß. Nonparametric estimation of scalar diffusions based on low frequency data.* Ann. Statist., *32(5):2223–2253, 2004.*

[GL11] *Alexander Goldenshluger and Oleg Lepski. Bandwidth selection in kernel density estimation: oracle inequalities and adaptive minimax optimality.* Ann. Statist., *39(3):1608–1632, 2011.*

[GL13] *A. Galves and E. Löcherbach. Infinite systems of interacting chains with memory of variable length—a stochastic model for biological neural nets.* J. Stat. Phys., *151(5):896–921, 2013.*

[GL16] *Antonio Galves and Eva Löcherbach. Modeling networks of spiking neurons as interacting processes with memory of variable length.* J. SFdS, *157(1):17–32, 2016.*

[GRZ04] *Fabrice Guillemin, Philippe Robert, and Bert Zwart. AIMD algorithms and exponential functionals.* Ann. Appl. Probab., *14(1):90–117, 2004.*

[Gut91] *Peter Guttorp.* Statistical inference for branching processes. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1991. A Wiley-Interscience Publication.*

[Guy07] *Julien Guyon. Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging.* Ann. Appl. Probab.*, 17(5-6):1538–1569, 2007.*

[Haa03] *Bénédicte Haas. Loss of mass in deterministic and random fragmentations.* Stochastic Process. Appl.*, 106(2):245–277, 2003.*

[HK11] *Marc Hoffmann and Nathalie Krell. Statistical analysis of self-similar conservative fragmentation chains.* Bernoulli*, 17(1):395–423, 2011.*

[HKL18] *Pierre Hodara, Nathalie Krell and Eva Löcherbach. Non-parametric estimation of the spiking rate in systems of interacting neurons.* Stat. Inference Stoch. Process.*, 21(1):81–111, 2018.*

[HKL18b] *Pierre Hodara, Nathalie Krell and Eva Löcherbach. Regularity of the invariant measure and non-parametric estimation of the jump rate* Statistical inference for piecewise-deterministic Markov processes, Romain Azais ; Florian Bouguet. Wiley*, Chapitre 2 - p. 39-64, 2018.*

[HL17] *Pierre Hodara and Eva Löcherbach. Hawkes processes with variable length memory and an infinite number of components.* Adv. in Appl. Probab.*, 49(1):84–107, 2017.*

[HMP92] *Peter Hall, J. S. Marron, and Byeong U. Park. Smoothed cross-validation.* Probab. Theory Related Fields*, 92(1):1–20, 1992.*

[HO16] *Marc Hoffmann and Adélaïde Olivier. Nonparametric estimation of the division rate of an age dependent branching process.* Stochastic Process. Appl.*, 126(5):1433–1471, 2016.*

[HR17] *Simon C. Harris and Matthew I. Roberts. The many-to-few lemma and multiple spines.* Ann. Inst. Henri Poincaré Probab. Stat.*, 53(1):226–242, 2017.*

[HRBR15] *Niels Richard Hansen, Patricia Reynaud-Bouret, and Vincent Rivoirard. Lasso and probabilistic inequalities for multivariate point processes.* Bernoulli*, 21(1):83–143, 2015.*

[IW81] *N. Ikeda and S. Watanabe. Stochastic differential equations and diffusion processes.* North-Holland Mathematical Library 24 Periodical. Elsevier, Academic Press City*, 1981.*

[KEBC05] *Mads Kaern, Timothy C Elston, William J Blake, and James J Collins. Stochasticity in gene expression: from theories to phenotypes.* Nature Reviews Genetics*, 6(6):451–64, 2005.*

[KR11] *Nathalie Krell and Alain Rouault. Martingales and rates of presence in homogeneous fragmentations.* Stochastic Process. Appl.*, 121(1):135-154, 2011.*

[Kre08] *Nathalie Krell. Multifractal spectra and precise rates of decay in homogeneous fragmentations.* Stochastic Process. Appl.*, 118(6):897-916, 2008.*

[Kre09] *Nathalie Krell. Self-similar branching Markov chains.* Séminaire de Probabilités XLII*, 261-280, Lecture Notes in Math. 979, Springer, Berlin, 2009.*

[Kre16] *Nathalie Krell. Statistical estimation of jump rates for a piecewise deterministic Markov processes with deterministic increasing motion and jump mechanism* ESAIM: PS *20 196-216, 2016.*

[KS21] *Nathalie Krell and Émeline Schmisser. Nonparametric estimation of jump rates for a specific class of piecewise deterministic Markov processes.* Bernoulli*, 27(4): 2362-2388, 2021.*

[Kub69] *H. E. Kubitschek. Growth during the bacterial cell cycle: Analysis of cell size distribution.* Biophysical Journal*, 9(6):792-809, 1969.*

[Löc18] *E. Löcherbach. Absolute continuity of the invariant measure in piecewise deterministic Markov processes having degenerate jumps.* Stochastic Process. Appl.*, 128(6):1797-1829, 2018.*

[Lam00] *A. Lambert. Completely asymmetric Lévy processes confined in a finite interval.* Ann. Inst. H. Poincaré Probab. Statist.*, 36(2):251-274, 2000.*

[LP09] *Philippe Laurençot and Benoit Perthame. Exponential decay for the growth-fragmentation/cell-division equation.* Commun. Math. Sci.*, 7(2):503–510, 2009.*

[Mey90] *Yves Meyer.* Ondelettes et opérateurs. I. *Actualités Mathématiques. [Current Mathematical Topics]. Hermann, Paris, 1990. Ondelettes. [Wavelets].*

[ODA88] *H. G. Othmer, S. R. Dunbar, and W. Alt,* Models of dispersal in biological systems*, J. Math. Biol.* ***26*** *, no. 3, 263-298. 1988.*

[Per07] *Benoît Perthame.* Transport equations in biology*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2007.*

[PTW10] *K. Pakdaman, M. Thieullen, and G. Wainrib. Fluid limit theorems for stochastic hybrid systems with application to neuron models.* Adv. in Appl. Probab., *42(3):761–794, 2010.*

[RT15] *Martin G. Riedler and Michèle Thieullen. Spatio-temporal hybrid (PDMP) models: central limit theorem and Langevin approximation for global fluctuations. Application to electrophysiology.* Bernoulli, *21(2):647–696, 2015.*

[RTT17] *Vincent Renault, Michèle Thieullen, and Emmanuel Trélat. Optimal control of infinite-dimensional piecewise deterministic Markov processes and application to the control of neuronal dynamics via optogenetics.* Netw. Heterog. Media, *12(3):417–459, 2017.*

[Rud82] *Mats Rudemo. Empirical choice of histograms and kernel density estimators.* Scand. J. Statist., *9(2):65–78, 1982.*

[Sgi02] *M. S. Sgibnev. Stone's decomposition of the renewal measure via Banach-algebraic techniques.* Proc. Amer. Math. Soc., *130(8):2425–2430, 2002.*

[SMPT05] *Eric J. Stewart, Richard Madden, Gregory Paul and François Taddei. Aging and death in an organism that reproduces by morphologically symmetric division.* PLoS Biology *3(2), 45, 2005.*

[SMPT05] *Alexander Sturm, Matthias Heinemann, Markus Arnoldini, Arndt Benecke, Martin Ackermann, Matthias Benz, Jasmine Dormann, and Wolf-Dietrich Hardt. The cost of virulence: Retarded growth of salmonella typhimurium cells expressing type iii secretion system 1.* PLoS Pathogens, *7(7):10, 2011.*

[TMY09] *C. Tan, P. Marguet, and L. You. Emergent bistability by a growth-modulating positive feedback circuit.* Nature Chemical Biology, *5(11):842–848, 2009.*

[Tsy04] *Alexandre B. Tsybakov.* Introduction à l'estimation non-paramétrique, *volume 41 of* Mathématiques & Applications (Berlin) [Mathematics & Applications]. *Springer-Verlag, Berlin, 2004.*

[Whe82] *A.E. Wheals. Size control models of saccharomyces cerevisiae cell proliferation.* Molecular and Cellular Biology **2(4)**, *361-368, 1982.*

[WRPDTWS10] *P. Wang, L. Robert, J. Pelletier, W.L. Dang, F. Taddei, A. Wright and S. Jun Robust growth of escherichia coli.* Current Biology *20(12), 1099-1103, 2010.*

[Yan08] N. M. Yanev. Statistical inference for branching processes. In Records and Branching Processes (M. Ahsanullah and G.P. Yanev, eds.) *143-168. Nova Science Publishers, 2008.*