

STATISTICAL ESTIMATION OF JUMP RATES FOR A PIECEWISE DETERMINISTIC MARKOV PROCESSES WITH DETERMINISTIC INCREASING MOTION AND JUMP MECHANISM *

NATHALIE KRELL¹

Abstract. We consider the class of Piecewise deterministic Markov processes (PDMP), whose state space is $(0, \infty)$, that possess an increasing deterministic motion and with a deterministic jump mechanism. Well known examples for this class of processes are transmission control protocol (TCP) window size process and the processes modeling the size of a “marked” *Escherichia coli* cell. Having observed the PDMP until its n th jump, we construct a nonparametric estimator of the jump rate λ . Our main result is 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 $n^{-s/(2s+1)}$. Simulations illustrate the behavior of our estimator.

Mathematics Subject Classification. 62M05, 62G05, 62G20, 60J25.

Received November 4, 2015. Revised May 3, 2016. Accepted May 11, 2016.

1. INTRODUCTION

The Piecewise deterministic Markov processes were first introduced in the literature by Davis [19, 20], they form a family of càdlàg Markov processes involving a deterministic motion punctuated by random jumps. We refer to the paper [8] and its references for an overview of PDMPs. Let us detail the special case of PDMPs that will be considered in this paper. The motion of the PDMP $(X(t))_{t \geq 0}$ depends on three local characteristics, namely the jump rate λ , the flow ϕ and a deterministic increasing function f which governs the location of the process at the jump time (in the general case it depends on a Markov kernel Q). The process starts from x and follows the flow $\phi(x, t)$ until the first jump time T_1 which occurs spontaneously in a Poisson-like fashion with rate $\lambda(\phi(x, t))$. The location of the process at the jump time T_1 , denoted by $Z_1 = X(T_1)$, is equal to $f(X(T_1^-))$, with f a function such that $0 \leq f(y) \leq \kappa y$ with $0 < \kappa < 1$. The motion restarts from this new point Z_1 as before. This fully describes a piecewise continuous trajectory for $\{X(t)\}$ with jump times $\{T_k\}$ and post jump locations $\{Z_k\}$, and which evolves according to the flow ϕ between two jumps.

This paper analyzes a special case of both Piecewise Deterministic Markov process (PDMPs) and growth-fragmentation model.

Keywords and phrases. Piecewise deterministic markov processes, nonparametric estimation, jump rate estimation, ergodicity of Markov chains.

* *The research of N. Krell is partly supported by the Agence Nationale de la Recherche PIECE 12-JS01-0006-01.*

¹ Université de Rennes 1, Institut de Recherche mathématique de Rennes, CNRS-UMR 6625, Campus de Beaulieu, Bâtiment 22, 35042 Rennes cedex, France. nathalie.krell@univ-rennes1.fr

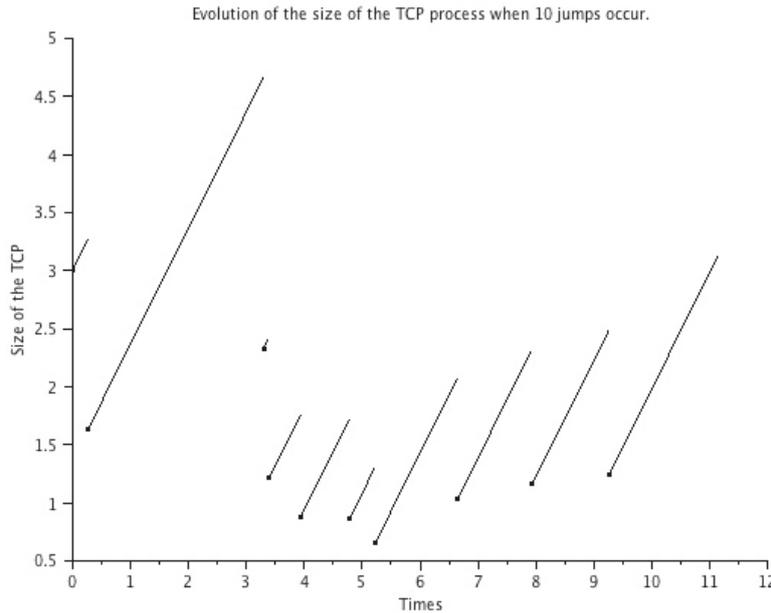


FIGURE 1. Evolution of the TCP process when 10 jumps occur.

The first known example for this class of processes is the TCP window size process (see [11, 17, 22–24]). The TCP protocol is one of the main data transmission protocols of the Internet. It has been designed to adapt to the various traffic conditions of the actual network. For a connection, the maximum number of packets that can be sent at each round is given by a variable W , called the congestion window size. If all the W packets are successfully transmitted, then W is increased by 1, otherwise it is multiplied by $\delta \in (0, 1)$ (detection of a congestion). As shown in [22] a correct scaling of this process leads to a continuous time Markov process, called general TCP window size process. See Figure 1 for a simulation of a TCP process when 10 jumps occur.

The second example is the processes modeling the size of a “marked” *Escherichia coli* cell (see [21, 27]). If by following the evolution of *Escherichia coli* bacterium, we randomly choose a bacteria, and follow his growth, and at his division, we choose randomly, and independently of the process, to follow one of his daughter and so on. We call the bacteria followed at each time, the “marked” bacteria. The size of the “marked” bacteria is a PDMP. More precisely between the jumps, the bacteria grows exponentially with a growth rate which we will refer to later as the instantaneous growth rate, denoted τ . If x is the size of the bacteria at birth, after a time t , his size is $xe^{\tau t}$. The division rate of such process is denoted by $B(\cdot)$ and at each jump its size is divided by two.

This two examples will be call in the rest of the paper us two main examples.

In both cases the value of the process is divided in a deterministically way at each jump.

The purpose of this article is to perform non-parametric estimation of the jump rate based on a single observation of the process over a long time interval. The PDMPs are well studied processes in probability, but few work has been done in statistics.

The assumptions made in this paper ensure that the Markov process which gives the size after each jump of the PDMP is ergodic. The ergodic theorem was already known for some cases of PDMPs and for some one-dimensional jump-diffusions. In [18] Cloez studies a process which is more general than the one considered in this paper. More precisely, his process evolves like a diffusion which satisfies a stochastic differential equation between the jumps, but he requires that λ be bounded below. This paper does not make such assumption. In addition, the well known ergodicity properties of TCP, due to Bardet *et al.* [11], provide quantitative estimates for the exponential convergence to equilibrium, in terms of the total variation and Wasserstein distances. However, such

results cannot be used in the present framework as we need a uniform upper bound for the speed of convergence to the invariant measure over a certain class of functions in order to prove the statistical result.

Our approach is based on the methods used by Doumic *et al.* [21], which were applied to analyse a special case of PDMPs dealing with “marked” bacteria size evolution. The analytical results of this paper can be generalized to more general PDMPs, for example to TCP. In [21] the authors do not merely study the “marked” bacterium which is selected at each division uniformly and independently, but the evolution of all bacteria involved. A dependence structure results when, for example, a bacteria divides and gives birth to two new bacterium of equal size. In a general case, the instantaneous growth rate (constant for a given bacteria) would depend on the bacteria itself. Consequently, the size of the bacterium is no longer a Markov process. However, on the other hand, the size of the process and the instantaneous growth rate, together form a Markov process. The present paper contains the case of a “marked” bacterium, in its simple case, where the instantaneous growth rate is the same for all bacteria.

The aim of this paper is to generalize the study done in [21] to more general PDMP including the TCP process.

This article has two main features that may be useful for future studies. First, it can be used as a tool to verify the PDMP jump rate proposed in the existing literature, for example $\lambda(x) = x$ in the TCP case. Secondly, it can suggest an estimator for some special cases, such as for *Escherichia coli* bacteria, where the jump rate has not been conjectured.

Very few literature exists for estimation about PDMP. As far as I know there is no classical estimator for the jump rate of a PDMP.

In [9] Azaïs *et al.* give an estimator of the conditional distribution of the inter-jump times for a PDMP, which 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 [1–3]. 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 this assumption. As a consequence the tools of their paper and of the present one are different. To the best of my knowledge, [9] is the only work investigating the nonparametric estimation of the conditional distribution of the inter-arrival times for PDMPs. This paper relies on [6] in which the authors focus on the non parametric estimation of the jump rate and the cumulative rate for a class of non homogeneous marked renewal processes. The case where the post-jump locations of the PDMP do not depend on inter-arrival times was considered in [6].

We refer to [9] for an overview of the statistical methods related to this kind of process, as well as to [1–3] for statistical inference related to the multiplicative intensity model. The book of Andersen *et al.* [4] gives a comprehensive account of estimation for jump rates which depend both on time and spatial variable.

As far as I know, the only other paper dealing with general PDMP is the work of Azaïs [7], where the author focuses on a non parametric recursive estimator for the transition kernel of the PDMP.

Other works dealing with specific cases of PDMP can be seen as ruin probabilities, for example, as found in the references of [5]. In addition, the PDMP modeling the quantity of a given food contaminant in the body has been studied in [13–15], assuming that the inter-intake times are i.i.d.. In this paper we do not make this assumption.

The paper of Azaïs and Genadot [10] consider a growth-fragmentation model where λ is constant and the Markov kernel Q is absolutely continuous with respect to the Lebesgue measure. The case considered is totally different from the present one.

The paper is organized as follows. In Section 2 we introduce the class of PDMPs, that will be studied, and we give an explicit construction of the PDMP. Section 3 concerns the statistical estimation of the jump rate. We first define the observation scheme and the class of functions for the 3 parameters of the PDMP concerned, and the assumptions used (in Sect. 3.2). In Section 3.3, some ergodicity results are stated uniformly over the class of functions previously defined. We explicitly construct an estimator $\hat{\lambda}_n$ of λ . In Section 3.5 an upper bound for

the squared-error loss is given in the main Theorem 3.5. In Section 3.6, we illustrate our result with simulations of a TCP process which could not be seen as a “marked” bacteria process, amongst others.

Finally, Section 4 is reserved for the proofs. In Section 4.1 we prove the ergodicity result of Section 3.3. Section 4.2 presents the intermediate results needed in 4.3 to prove the major results given in Section 3.5.

2. PDMP

In general a 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 at the jump time. In this article, we consider a specific class of PDMP which includes the control of congestion TCP/IP used in communication networks [22, 24], for which the transition measure Q is a Dirac mass function, which means that when the process jumps, the size after the jump is a deterministic function of its size before. More precisely,

Assumption 2.1.

- The flow $\phi : [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$ is a one-parameter group of homeomorphisms: ϕ is C^1 , $\phi(\cdot, t)$ is an homeomorphism for each $t \in [0, \infty)$, satisfying the semigroup property: $\phi(\cdot, t + s) = \phi(\phi(\cdot, s), t)$ and $\phi_x(\cdot) := \phi(x, \cdot)$ is an C^1 -diffeomorphism.
- The jump rate $\lambda : [0, \infty) \rightarrow \mathbb{R}_+$ is assumed to be a measurable function satisfying

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

- $f : [0, \infty) \rightarrow [0, \infty)$ is an increasing C^1 -diffeomorphism and $Q(u, \{y\}) = \mathbb{1}_{\{f(u)=y\}}$.

Given these three characteristics, it can be shown ([19], pp. 62–66), that there exists a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \{\mathbb{P}_x\})$ such that the motion of the process $\{X(t)\}$ starting from a point $x \in [0, \infty)$ may be constructed as follows. Consider a random variable T_1 such that

$$\mathbb{P}_x(T_1 > t) = e^{-A(x,t)}, \tag{2.1}$$

where for $x \in [0, \infty)$ and $t \in [0, \infty)$

$$A(x, t) = \int_0^t \lambda(\phi(x, s)) ds.$$

If T_1 is equal to infinity, then the process X follows the flow, *i.e.* for $t \in \mathbb{R}_+$, $X(t) = \phi(x, t)$. Otherwise let

$$Z_1 = f(\phi(x, T_1)). \tag{2.2}$$

The trajectory of $\{X(t)\}$ starting at x , for $t \in [0, T_1]$, is given by

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

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 $\{X(t)\}$ with jump times $\{T_k\}_{k \in \mathbb{N}}$ (where $T_0 = 0$). A very natural Markov chain is linked to $\{X(t)\}$, namely the jump chain $(Z_n)_{n \in \mathbb{N}}$.

$\{X(t)\}$ is a Markov process with infinitesimal generator G :

$$Gh(y) = \phi'_x(y)h'(y) + \lambda(y)(h(f(y)) - h(y)) \tag{2.3}$$

for $h : [0, \infty) \rightarrow \mathbb{R}$ a bounded measurable functional.

Thanks to (2.1), we get that

$$\mathbb{P}(T_1 \in dt | Z_0 = x) = \lambda(\phi_x(t)) e^{-\int_0^t \lambda(\phi_x(u)) du} dt.$$

Using (2.2), the monotonicity of $f \circ \phi_x$ and a simple change of variables, we get the transition probability of the Markov chain $(Z_n)_{n \in \mathbb{N}}$:

$$\mathbb{P}(Z_1 \in dy | Z_0 = x) = \lambda(f^{-1}(y)) e^{-\int_{f(x)}^y \lambda(f^{-1}(u)) g_x(u) du} g_x(y) \mathbb{1}_{\{y \geq f(x)\}} dy, \quad (2.4)$$

where

$$g_x(y) := \left[(f \circ \phi_x)' \left((f \circ \phi_x)^{-1}(y) \right) \right]^{-1}.$$

3. STATISTICAL ESTIMATION OF THE JUMP RATE

3.1. The observation scheme

Statistical inference is based on the observation scheme:

$$(X(t), t \leq T_n)$$

and asymptotics are considered when the number of jumps of the process, n , goes to infinity.

Actually the simpler observation scheme:

$$(X(0), X(T_i), 1 \leq i \leq n) = (Z_i, 0 \leq i \leq n)$$

is sufficient, as you remark that for all $n \geq 1$: $T_n = (f \circ \phi_{Z_{n-1}})^{-1}(Z_n)$.

3.2. Class of functions

We suppose that f and ϕ_x are known. Typically in our two main examples $f(x) = x/2$.

We notice that as we suppose that we observe all the trajectory, even if we do not know ϕ_x , we could have a good estimation of it. Nevertheless, if we suppose that ϕ_x is not known and introduce an estimation $\hat{\phi}_x$ in all the rest of the paper it would change everything and a new strategy shall be developed.

We now give some definitions that we will need in the rest of the paper. Let \mathcal{D} be a compact intervals of $(0, \infty)$. 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 $\varphi : \mathcal{D} \rightarrow \mathbb{R}$ possessing a derivative of order $\lfloor s \rfloor$ that satisfies

$$|\varphi^{\lfloor s \rfloor}(y) - \varphi^{\lfloor s \rfloor}(x)| \leq c(\varphi) |x - y|^{\{s\}}. \quad (3.1)$$

The minimal constant $c(\varphi)$ such that (3.1) holds defines a semi-norm $|\varphi|_{\mathcal{H}^s(\mathcal{D})}$. We equip the space $\mathcal{H}^s(\mathcal{D})$ with the norm

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

and the associated Hölder balls

$$\mathcal{H}^s(\mathcal{D}, M_1) = \{\lambda : \|\lambda\|_{\mathcal{H}^s(\mathcal{D})} \leq M_1\}, \quad M_1 > 0.$$

Assumption 3.1. We suppose that for a compact interval of $(0, \infty)$, $f : [0, \infty) \rightarrow [0, \infty)$ and $\phi_x : [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$ are known and follow, for $M_2 > 0$, $0 < \kappa < 1$, $n_0 \in \mathbb{N} \setminus \{0\}$, and two positive continuous functions $m : [0, \infty) \rightarrow (0, \infty)$ and $M : [0, \infty) \rightarrow (0, \infty)$ such that for all $x \in [0, \infty) : M(x) \geq m(x) > 0$,

$$g_x \in \mathcal{H}^{n_0}(\mathcal{D}), \quad \|f\|_{L^\infty(\mathcal{D})} \leq M_2, \quad \text{and } f^{-1} \in \mathcal{H}^{n_0}(\mathcal{D})$$

$$\forall x > 0, \quad 0 < f(x) \leq \kappa x, \quad (3.2)$$

$$\forall y > 0, \quad \forall x \geq 0, \quad m(y) \leq g_x(y) \leq M(y). \quad (3.3)$$

We notice that $f(0) = 0$. Assumption (3.2) is quite reasonable as you want that the size reduce when the process jump. Typically an interesting case would be $f(x) = \kappa x$ with $\kappa \in (0, 1)$ and then (3.3) would simply be

$$M^{-1}(y) \leq \kappa \phi'_x(\phi_x^{-1}(y/\kappa)) \leq m^{-1}(y).$$

This seems quite reasonable because, in view of the definition of the infinitesimal generator of the PDMP defined in (2.3), we would like $\phi'_x(\cdot)$ not to be identically zero.

Also the cases where $\phi_x(t) = xe^{\alpha t}$ and $f(x) = x/2$, we get $g_x(y) = \frac{1}{\alpha x}$ and for $m(x) = \frac{1}{\alpha \max_{y \in \mathcal{D}} y}$ and $M(x) = \frac{1}{\alpha \min_{y \in \mathcal{D}} y}$ we get (3.3). For $\phi_x(t) = x + t$ and $f(x) = x/2$, we get $g_x(y) = 2$ and for $m(x) = 2$ and $M(x) = 2$ we get (3.3).

Each parameter from the last assumption are from now on fixed.

We want to bound from above the squared-loss error of our estimator over compact intervals \mathcal{D} of $(0, \infty)$. We need to specify the local smoothness properties of λ over \mathcal{D} , together with general properties that ensure that the empirical measurements of the PDMP converge toward the invariant probability with an appropriate speed of convergence. So we have to impose technical assumptions on λ in particular near the origin and infinity.

Definition 3.2. For $b > 0$, a vector of positive constants $\mathbf{c} = (r, l, L, a)$, we introduce the class $\mathcal{F}(\mathbf{c}, b)$ of continuous function $\lambda : [0, \infty) \rightarrow [0, \infty)$, such that

$$\int_0^{f(r)} M(u)\lambda(f^{-1}(u))du \leq L, \tag{3.4}$$

$$\int_{f(r)}^\infty m(u)\lambda(f^{-1}(u))du = \infty, \tag{3.5}$$

$$\int_{f(r)}^r M(u)\lambda(f^{-1}(u))du \geq l, \tag{3.6}$$

$$(\kappa^{b+1} - 1) \frac{a}{b+1} (f(r))^{b+1} < \log(1 - \kappa^{b+1}), \tag{3.7}$$

$$\lambda(x) \geq \frac{a(f(x))^b}{m(f(x))} \quad \forall x \geq r. \tag{3.8}$$

The assumptions (3.4) to (3.6) are weak and mostly technical.

Define

$$\delta(\mathbf{c}, b) := \frac{1}{1 - \kappa^{b+1}} \exp\left(- (1 - \kappa^{b+1}) \frac{a}{b+1} (f(r))^{b+1}\right).$$

Therefore thanks to (3.7) we have $\delta(\mathbf{c}, b) < 1$

Typically in the case where $f(x) = x/2$, (3.7) is working for $b = 1$ and $r > \sqrt{-32 \log(3/4)/(3a)}$. Additionally if $a \leq 2 \min_{y \in \mathcal{D}} m(y/2)$ the case $\lambda(x) = x$ follow (3.8) for us two main examples.

Fix a vector of positive constants $\mathbf{c} = (r, l, L, a)$ and a constant b .

3.3. Geometric ergodicity of the discrete model

Let $x \in [0, \infty)$. Introduce the transition kernel

$$\mathcal{P}_\lambda(x, dx') = \mathbb{P}(Z_n \in dx' \mid Z_{n-1} = x)$$

of the size of the process at the n th jump time, given the size of the process at the $(n - 1)$ th jump time. From (2.4), we infer that $\mathbb{P}(Z_1 \in dy \mid Z_0 = x)$ is equal to

$$\lambda(f^{-1}(y)) e^{-\int_{f(x)}^y \lambda(f^{-1}(u))g_x(u)du} g_x(y) \mathbb{1}_{\{y \geq f(x)\}} dy.$$

Thus we obtain an explicit formula for

$$\mathcal{P}_\lambda(x, dy) = \mathcal{P}_\lambda(x, y)dy$$

with

$$\mathcal{P}_\lambda(x, y) = \lambda(f^{-1}(y))e^{-\int_{f(x)}^y \lambda(f^{-1}(u))g_x(u)du} g_x(y) \mathbb{1}_{\{y \geq f(x)\}}. \tag{3.9}$$

Denote the left action of positive measures $\mu(dx)$ on $[0, \infty)$ for the transition kernel \mathcal{P}_λ by

$$\mu \mathcal{P}_\lambda(dy) = \int_{[0, \infty)} \mu(dx) \mathcal{P}_\lambda(x, dy)$$

and the right action of a function ψ on \mathbb{R} for the transition \mathcal{P}_λ by

$$\mathcal{P}_\lambda \psi(x) = \int_{[0, \infty)} \psi(y) \mathcal{P}_\lambda(x, dy).$$

We now give the geometric ergodic theorem that we will need for the statistical part. We need an uniformity on the class of functions $\mathcal{F}(c, b)$ defined in Section 3.2.

We introduce the Lyapunov function

$$\mathbb{V}(x) = \exp\left(\frac{a}{b+1}(f(x))^{b+1}\right) \text{ for } x \in [0, \infty). \tag{3.10}$$

The function \mathbb{V} controls the rate of the geometric ergodicity of the chain with transition \mathcal{P}_λ and appears in the proof of Proposition 3.3.

Proposition 3.3. *Under Assumptions 2.1 and 3.1, for every $\lambda \in \mathcal{F}(c, b)$ there exists a unique invariant probability measure of the form $\nu_\lambda(dx) = \nu_\lambda(x)dx$ on $[0, \infty)$. Moreover there exist $0 < \gamma < 1$, a constant R and a function $\mathbb{V} : [0, \infty) \rightarrow [1, \infty)$ such that*

$$\sup_{\lambda \in \mathcal{F}(c, b)} \sup_{|\psi| \leq V} \left| \mathcal{P}_\lambda^k \psi(x) - \int_{[0, \infty)} \psi(z) \nu_\lambda(z) dz \right| \leq R \mathbb{V}(x) \gamma^k \tag{3.11}$$

for every $x \in [0, \infty)$, $k \geq 0$, where the supremum is taken over all functions $\psi : [0, \infty) \rightarrow \mathbb{R}$ satisfying $|\psi(x)| \leq \mathbb{V}(x)$ for all $x \in [0, \infty)$. The function \mathbb{V} is ν_λ -integrable for every $\lambda \in \mathcal{F}(c, b)$.

For all $y \in [0, \infty)$ we have the relation:

$$\lambda(y) \mathbb{E}_{\nu_\lambda}(g_{Z_0}(f(y)) \mathbb{1}_{\{f(Z_0) \leq f(y)\}} \mathbb{1}_{\{Z_1 \geq f(y)\}}) = \nu_\lambda(f(y)). \tag{3.12}$$

3.4. Construction of a nonparametric estimator

By formula (3.12),

$$\lambda(y) = \frac{\nu_\lambda(f(y))}{\mathbb{E}_{\nu_\lambda}(g_{Z_0}(f(y)) \mathbb{1}_{\{Z_0 \leq y\}} \mathbb{1}_{\{Z_1 \geq f(y)\}})},$$

provided the denominator is positive. This representation suggests an estimation procedure, replacing the marginal density $\nu_\lambda(f(y))$ and the expectation in the denominator by their empirical counterparts. To that end, pick a kernel function

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

and set $K_h(y) = h^{-1}K(h^{-1}y)$ for $y \in \mathbb{R}$ and $h > 0$. Our estimator is defined by

$$\widehat{\lambda}_n(y) = \frac{n^{-1} \sum_{k=1}^n K_{h_n}(Z_k - f(y))}{(n^{-1} \sum_{k=1}^n g_{Z_{k-1}}(f(y)) \mathbb{1}_{\{Z_k \geq f(y), y \geq Z_{k-1}\}}) \vee \varpi_n}, \tag{3.13}$$

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

Assumption 3.4. The function K has compact support, and for some integer $n_0 \geq 1$ (the same n_0 as in Assumption 3.1), we have $\int_{\mathbb{R}} x^k K(x) dx = \mathbf{1}_{\{k=0\}}$ for $0 \leq k \leq n_0$.

3.5. Speed of convergence

We are now ready to state our main result.

Theorem 3.5. *Work under Assumption 2.1 and 3.1. Specify $\widehat{\lambda}_n$ with a kernel K satisfying Assumption 3.4 for some $n_0 > 0$ and*

$$h_n = c_0 n^{-1/(2s+1)}, \quad \varpi_n \text{ such that } \lim_{n \rightarrow \infty} \varpi_n = 0$$

For every $M_1 > 0$, there exist $c_0 = c_0(\mathbf{c}, M_1, M_2)$ and $d(\mathbf{c}) \geq 0$ such that for every $0 < s < n_0$ and every compact interval $\mathcal{D} \subset (d(\mathbf{c}), \infty)$ such that $\inf \mathcal{D} \geq f(r)$, we have

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

where the supremum is taken over

$$\lambda \in \mathcal{F}(\mathbf{c}, b) \cap \mathcal{H}^s(\mathcal{D}, M_1),$$

and $\mathbb{E}_{\mu}[\cdot]$ denotes expectation with respect to any initial distribution $\mu(d\mathbf{x})$ for (Z_0) on $[0, \infty)$ such that we have $\int \mathbb{V}(x)^2 \mu(d\mathbf{x}) < \infty$.

We observe that we recover the result for the marked bacteria of [21]. In this case $\phi(x, t) = xe^{\kappa_0 t}$ with $\kappa_0 \in \mathbb{R}^{+*}$ and $f(x) = x/2$, so that $g_x(y) = \frac{1}{\kappa_0 y}$. We find the same estimator but the speed of convergence is a little bit better, as φ_n need not be $\log(n)^{-1}$; rather we only require that $\lim_{n \rightarrow \infty} \varphi_n^{-1} = 0$.

3.6. Numerical implementation

The goal of this subsection is to illustrate the finite behaviour of our estimator *via* numerical experiments.

At least for the simulation, I proposed an adaptive estimation procedure. I used a Smoothed Cross-validation, SCV, to choose this bandwidth (see for example the paper of Hall, Marron and Park [26]. This idea was first published by Bowman[16] and Rudemo [28]). As the bandwidth interfere only for the estimation of the invariant probability ν_{λ} , I will use a Cross validation for this estimation. I cut the trajectory in three part, I use the first one to estimate $\widehat{\nu}_{\lambda}$ and then for the last part of the trajectory, I use the method of the least squares cross validation to minimize in h

$$CV(h) = \int \widehat{\nu}_{\lambda}(x)^2 dx - 2n^{-1} \sum_{i=2n}^{3n} \widehat{\nu}_{\lambda}(X_j)$$

(where I had approximated the integral term by a Riemann approximation). For this minimizer denoted \widehat{h} , I calculated the estimator $\widehat{\lambda}$ on all the trajectory. In the next figures, I use this method to find the reconstructed λ with h adaptative.

More precisely we first investigate numerical simulations for the TCP. The TCP window-size process appears as the scaling limit of the transmission rate of a server uploading packets on the Internet according to the algorithm used in the TCP (transmission control protocol) in order to avoid congestion (see [22] for details on this scaling limit). This PDMP takes values in \mathbb{R}^+ and the jump rate λ is the identity function. The function f which represents the proportion of the size kept after the jump is $f(x) = x/2$. The flow is $\phi(x, t) = x + t$.

As a consequence the size of the process after the n th jump Z_n , conditional on Z_{n-1} , has the same law as $\sqrt{(1/4)Z_{n-1}^2 + e_n/2}$, where $(e_n)_{n \geq 0}$ is a family of i.i.d. random variables with exponential distribution of parameter 1. The variable e_n is also independent of $(Z_i)_{i \leq n-1}$. As a consequence it is easy to generate the $(Z_n)_{n \geq 0}$

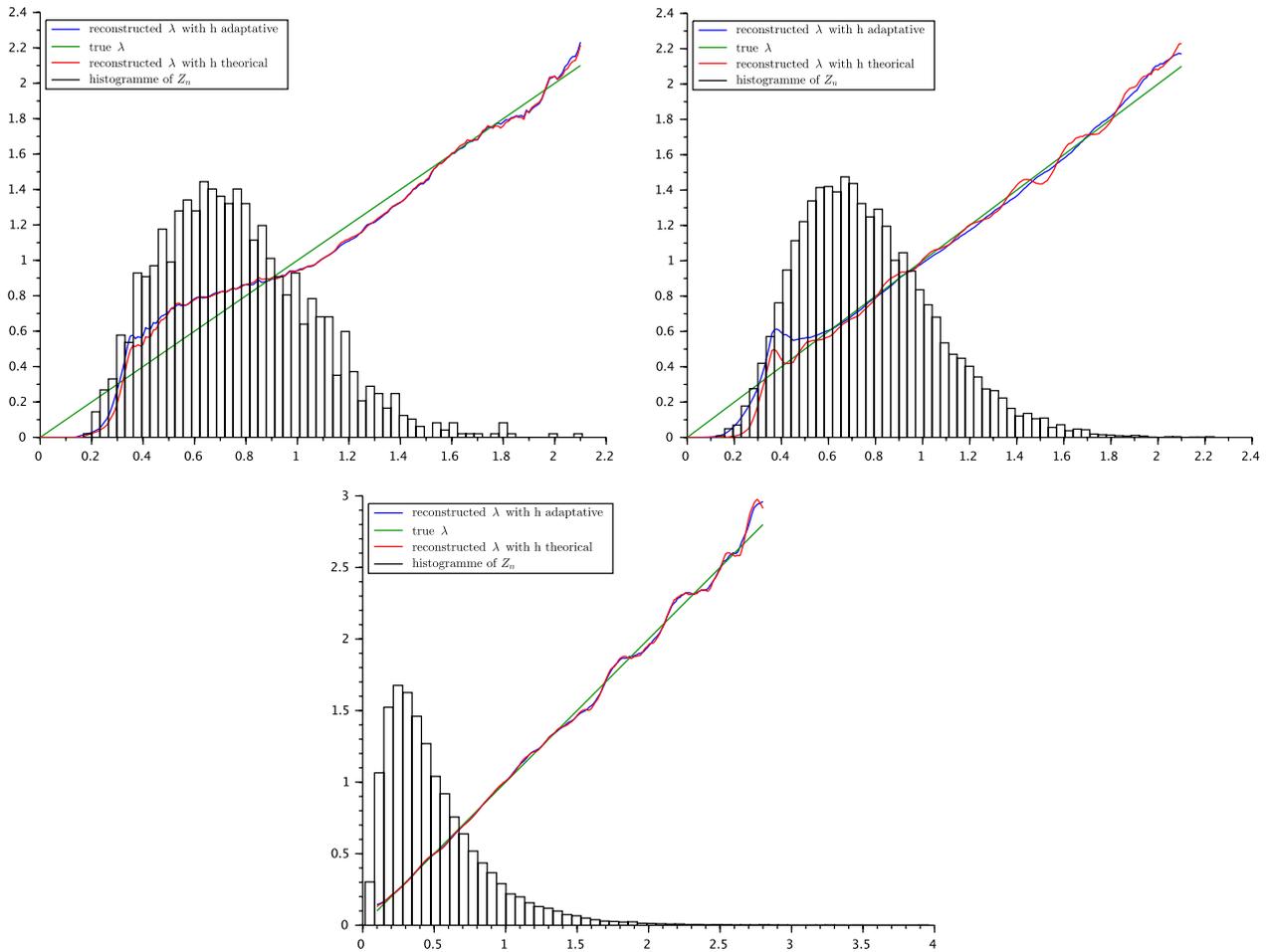


FIGURE 2. Reconstruction of λ for $n = 1500$, $n = 15\,000$ and $n = 45\,000$ in the TCP case.

recursively. A trajectory of such a PDMP is given in Figure 1. These processes satisfy the assumptions required for our theorem, with $\kappa = 1/2$, $m(\cdot) = 2$, $M(\cdot) = 2$, $g(\cdot) = 2$, $r = 1$, $a = 1$, $b = 1/2$, $\varpi_n = 0.1$, for the h theoretical I use $h_n = n^{-1/3}$ and K a Epanechnikov Kernel, for which $n_0 = 1$ for Assumption 3.4. Thus we expect a speed of convergence of order $n^{1/3}$ at best.

We display our numerical results as specified above in Figures 1, 2 and 3.

Figure 2 displays the reconstruction of λ for different simulated samples, for $n = 1500$, $n = 15\,000$ and $n = 45\,000$. As expected, the estimation is better for larger n . The estimator performs worse for small x as these sizes are rarely reached by the TCP process and above all as we have used a threshold ϖ_n in us estimator, which introduce a bias. Indeed, thanks to the simulation, we notice that the threshold play an important role for the small value of x for which $(n^{-1} \sum_{k=1}^n g_{Z_{k-1}}(f(y)) \mathbb{1}_{\{Z_k \geq f(y), f(y) \geq f(Z_{k-1})\}})$ could be zero. As a consequence I have choosed it equal to 0.1.

In Figure 3, we plot the empirical mean error of our estimation procedure on a log-log scale. The numerical results agree with the theory.

We now consider an other example, for which the size of the process are divided by 3 at division. Thus we have $f(x) = x/3$. We choice $\phi(x, t) = xe^t$, $\lambda(x) = x$ and therefore $g_x(y) = 1/y$. Let $(e_i)_{i \in \mathbb{N}^*}$ a i.i.d family of continuous

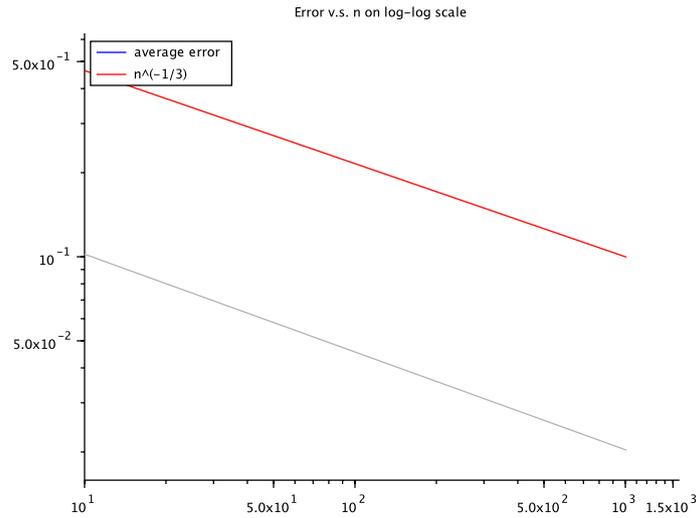


FIGURE 3. The empirical mean error of the estimation procedure *vs.* the theoretical rate on a log-log scale in the TCP case.

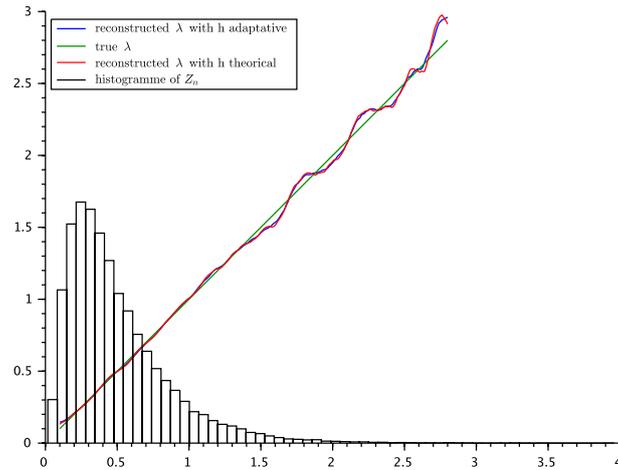


FIGURE 4. Reconstruction of λ for $n = 45\,000$ in the marked bacteria case.

uniform distribution on $[0, 1]$. Conditionally on $Z_n = x$ the law of Z_{n+1} is equal in law to $\frac{x}{3} - \frac{1}{3} \log(e_{n+1})$. We also use for the h theoretical I use $h_n = n^{-1/3}$, K a Epanechnikov Kernel, $\varpi_n = 0.1$.

As before Figure 4 displays the reconstruction of λ for a simulated sample with $n = 45\,000$.

We also give an example with a non monotonic function $\lambda(x) = 1 + 0.5x(x/4 - 1)$, the f and the ϕ are the same as for the TCP case. As I do not find an explicit formula for the law of the Z_n , it was simulate with a rejection method, the process having a size z at time t will jump between time t and $t + H$ with a probability $H\lambda(z)$. We will use $H = 1/50$ in the simulation. We also use for the h theoretical I use $h_n = n^{-1/3}$, K a Epanechnikov Kernel, $\varpi_n = 0.1$.

As before Figure 5 displays the reconstruction of λ for a simulated sample with $n = 15\,000$.

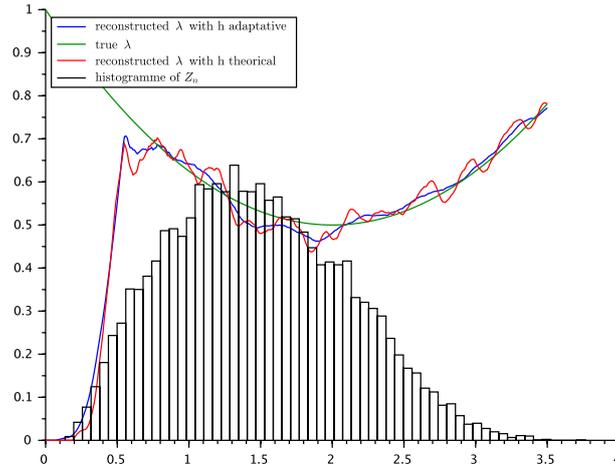


FIGURE 5. Reconstruction of $\lambda = 1 + 0.5x(x/4 - 1)$ for $n = 15\,000$ in the marked bacteria case.

4. PROOF

4.1. Proof of Proposition 3.3

We will follow the same idea as in [21]. We prove a minorisation condition, strong aperiodicity and a drift condition for the transition operator \mathcal{P}_λ in order to use Theorem 1.1 of [12].

We could also use for example the paper of Hairer and Mattingly [25] to get the existence of a unique invariant measure, the assumption could be easily prove and the bound for γ is more explicit but we would have inequality: there exist $0 < \gamma < 1$, a constant R and a function $\mathbb{V} : [0, \infty) \rightarrow [1, \infty)$ such that

$$\sup_{\lambda \in \mathcal{F}(c,b)} \sup_{\|\psi\| < \infty} |\mathcal{P}_\lambda^k \psi(x) - \int_{[0,\infty)} \psi(z) \nu_\lambda(z) dz| \leq R \|\psi - \int_{[0,\infty)} \psi(z) \nu_\lambda(z) dz\|_V \gamma^k$$

with

$$\|\psi\|_V = \sup_x \frac{|\psi(x)|}{1 + V(x)}.$$

Minorisation condition. Let $\lambda \in \mathcal{F}(c, b)$ and set $\mathcal{C} = (0, r)$ where r is specified by c . Fix a measurable $A \in \mathcal{F}$ and $x \in \mathcal{C}$; thanks to (3.3), we have

$$\mathcal{P}_\lambda(x, A) \geq \int_A \lambda(f^{-1}(y)) e^{-\int_{f(x)}^y \lambda(f^{-1}(u)) M(u) du} m(y) \mathbb{1}_{\{y \geq f(x)\}} dy.$$

We introduce the function φ_λ

$$\varphi_\lambda(y) := \lambda(f^{-1}(y)) e^{-\int_0^y \lambda(f^{-1}(u)) M(u) du} m(y) \quad \forall y \in [0, \infty), \tag{4.1}$$

and the measure μ_λ

$$\mu_\lambda(dy) := \frac{\varphi_\lambda(y)}{c_\lambda} \mathbb{1}_{\{y > f(r)\}} dy,$$

where $c_\lambda = \int_{f(r)}^\infty \varphi_\lambda(u) du$. Thus, we have

$$\mathcal{P}_\lambda(x, A) \geq \mu_\lambda(A) c_\lambda.$$

By using (3.5) and (3.4), we get that

$$c_\lambda = \int_{f(r)}^{\infty} \varphi_\lambda(u) du = \left[\frac{-m(y)e^{-\int_0^y \lambda(f^{-1}(u))M(u)du}}{M(y)} \right]_{f(r)}^{\infty} \geq \frac{m(f(r))e^{-L}}{M(f(r))} := \tilde{\beta} > 0.$$

This shows that the following minorisation condition holds for every $x \in \mathcal{C}$ and $A \in \mathcal{F}$ uniformly in $\lambda \in \mathcal{F}(\mathbf{c}, b)$:

$$\mathcal{P}_\lambda(x, A) \geq \mu_\lambda(A)\tilde{\beta}. \quad (4.2)$$

□

Strong aperiodicity condition. We have

$$\begin{aligned} \mu_\lambda(\mathcal{C})\tilde{\beta} &= c_\lambda^{-1}\tilde{\beta} \int_{f(r)}^r \varphi_\lambda(y) dy = c_\lambda^{-1}\tilde{\beta} \left[\frac{-m(y)e^{-\int_0^y \lambda(f^{-1}(u))M(u)du}}{M(y)} \right]_{f(r)}^r \\ &= \tilde{\beta} \left(1 - \frac{m(r)M(f(r))}{M(r)m(f(r))} e^{-\int_{f(r)}^r \lambda(f^{-1}(u))M(u)du} \right), \end{aligned} \quad (4.3)$$

using the computation we just did for c_λ .

Now we use (3.6) to get that

$$\mu_\lambda(\mathcal{C})\tilde{\beta} \geq \tilde{\beta} \left(1 - \frac{m(r)M(f(r))}{M(r)m(f(r))} e^{-l} \right) := \beta > 0. \quad (4.4)$$

□

Drift condition. Let $\lambda \in \mathcal{F}(\mathbf{c}, b)$ and recall that $\mathbb{V} : [0, \infty) \rightarrow [1, \infty)$, which is defined in (3.10), is continuously differentiable and satisfies

$$\lim_{y \rightarrow \infty} \mathbb{V}(y) \exp\left(-\frac{a}{b+1}y^{b+1}\right) = 0. \quad (4.5)$$

For $x \geq r$, using (3.3), (3.8) and (4.5) and integration by parts with the boundary condition (3.5), we have,

$$\begin{aligned} \mathcal{P}_\lambda \mathbb{V}(x) &= \int_{f(x)}^{\infty} \mathbb{V}(y) \lambda(f^{-1}(y)) e^{-\int_{f(x)}^y \lambda(f^{-1}(u))g_x(u)du} g_x(y) dy \\ &\leq \int_{f(x)}^{\infty} \mathbb{V}'(y) e^{-\int_{f(x)}^y \lambda(f^{-1}(u))m(u)du} dy. \end{aligned}$$

Thanks to (3.8), we get that

$$\mathcal{P}_\lambda \mathbb{V}(x) \leq \int_{f(x)}^{\infty} \mathbb{V}'(y) e^{-\int_{f(x)}^y a u^b du} dy \leq e^{a \frac{f(x)^{b+1}}{b+1}} \int_{f(x)}^{\infty} \mathbb{V}'(y) e^{-a \frac{y^{b+1}}{b+1}} dy.$$

Integrating again by parts and using (4.5), we obtain that

$$\mathcal{P}_\lambda \mathbb{V}(x) \leq e^{a \frac{f(x)^{b+1}}{b+1}} \int_{f(x)}^{\infty} \mathbb{V}(y) a y^b e^{-a \frac{y^{b+1}}{b+1}} dy.$$

Now use the change of variable $z = a \frac{y^{b+1}}{b+1}$ and the definition of $\mathbb{V}(x)$. As (4.5) is satisfied, we get

$$\mathcal{P}_\lambda \mathbb{V}(x) \leq \mathbb{V}(x) \int_{a \frac{f(x)^{b+1}}{b+1}}^{\infty} e^{\frac{a}{b+1} \left(f \left(\left(\frac{z(b+1)}{a} \right)^{1/(b+1)} \right) \right)^{b+1} - z} dz.$$

By using (3.2), we obtain, for $x \geq r$

$$\mathcal{P}_\lambda \mathbb{V}(x) \leq \mathbb{V}(x) \int_a^{\frac{f(x)b+1}{b+1}} e^{(\kappa^{b+1}-1)z} dz.$$

Therefore,

$$\mathcal{P}_\lambda \mathbb{V}(x) \leq \mathbb{V}(x) \delta(\mathbf{c}, b), \quad (4.6)$$

with

$$\delta(\mathbf{c}, b) = \frac{1}{1 - \kappa^{b+1}} \exp\left(- (1 - \kappa^{b+1}) \frac{a}{b+1} (f(r))^{b+1}\right),$$

and we have $\delta(\mathbf{c}, b) < 1$ by (3.7).

We next need to control $\mathcal{P}_\lambda \mathbb{V}$ outside $x \in [r, \infty)$, that is on the small set \mathcal{C} . For every $x \in \mathcal{C}$, we have

$$\begin{aligned} \mathcal{P}_\lambda \mathbb{V}(x) &\leq \left(\int_{f(x)}^{f(r)} \mathbb{V}(y) \lambda(f^{-1}(y)) g_x(y) dy + \int_{f(x)}^{\infty} \mathbb{V}(y) \lambda(f^{-1}(y)) e^{-\int_{f(x)}^y \lambda(f^{-1}(u)) g_x(u) du} g_x(y) dy \right) \\ &\leq M \sup_{y \in [0, f(r)]} \mathbb{V}(y) L + \delta(\mathbf{c}, b) \mathbb{V}(r) =: K_1 < \infty, \end{aligned} \quad (4.7)$$

where we used (3.4), (3.3), (4.6) for $x = r$ and the fact that $\lambda \in \mathcal{F}(\mathbf{c}, b)$. Combining (4.6) and (4.7), we conclude that

$$\mathcal{P}_\lambda \mathbb{V}(\mathbf{x}) \leq \delta(\mathbf{c}, b) \mathbb{V}(\mathbf{x}) \mathbb{1}_{\{\mathbf{x} \notin \mathcal{C}\}} + K_1 \mathbb{1}_{\{\mathbf{x} \in \mathcal{C}\}}. \quad (4.8)$$

□

End of the proof of Proposition 3.3. By Theorem 1.1 in Baxendale [12] the minorisation condition (4.2) together with the strong aperiodicity condition (4.4) and the drift condition (4.8) imply inequality (3.11), with $R > 0$ and $0 < \gamma < 1$ that explicitly depend on $\delta(\mathbf{c}, b)$, β , $\tilde{\beta}$, \mathbb{V} and K . As a, b, κ, \mathbb{V} and f are fix, the R and the γ are the same for each $\lambda \in \mathcal{F}(\mathbf{c}, b)$. Therefore, we have (3.11).

It remains to prove equality (3.12). As $\mathcal{P}_\lambda(x, dy) = \mathcal{P}_\lambda(x, y) dy$ and

$$\nu_\lambda \mathcal{P}_\lambda = \nu_\lambda,$$

we have that $\nu_\lambda(dy) = \nu_\lambda(y) dy$ and

$$\begin{aligned} \nu_\lambda(y) &= \int_{\mathbb{R}_+} \nu_\lambda(x) \mathcal{P}_\lambda(x, y) dx \\ &= \int_{\mathbb{R}_+} \nu_\lambda(x) \lambda(f^{-1}(y)) e^{-\int_{f(x)}^y \lambda(f^{-1}(u)) g_x(u) du} g_x(y) \mathbb{1}_{\{f(x) \leq y\}} dx. \end{aligned}$$

Thanks to (3.5), we get that

$$e^{-\int_{f(x)}^y \lambda(f^{-1}(u)) g_x(u) du} = \int_y^\infty \lambda(f^{-1}(u)) g_x(u) e^{-\int_{f(x)}^u \lambda(f^{-1}(u')) g_x(u') du'} du.$$

Therefore,

$$\begin{aligned} \nu_\lambda(y) &= \lambda(f^{-1}(y)) \int_{[0, \infty)} \nu_\lambda(x) g_x(y) \mathbb{1}_{\{f(x) \leq y\}} \int_y^\infty \lambda(f^{-1}(u)) g_x(u) e^{-\int_{f(x)}^u \lambda(f^{-1}(u')) g_x(u') du'} du' dx \\ &= \lambda(f^{-1}(y)) \int_{[0, \infty)} \nu_\lambda(x) g_x(y) \mathbb{1}_{\{f(x) \leq y\}} \int_y^\infty \mathbb{1}_{\{u \geq y\}} \mathcal{P}_\lambda(x, u) du dx \\ &= \lambda(f^{-1}(y)) \mathbb{E} \nu_\lambda(g_{Z_0}(y) \mathbb{1}_{\{f(Z_0) \leq y\}} \mathbb{1}_{\{Z_1 \geq y\}}). \end{aligned} \quad \square$$

4.2. Speed of convergence for the empirical measure

We now give a few results that we will need for the proof of Theorem 3.5 in the next subsection. In fact, we decompose the square loss error into a sum of three terms that we will study in the following Propositions.

The notation \lesssim means inequality up to a constant that not depend on n .

Lemma 4.1. *For any \mathfrak{c} such that Assumptions 3.1 and 2.1 are satisfied, there exists a constant $d(\mathfrak{c}) \geq 0$ such that for any compact interval $\mathcal{D} \subset (d(\mathfrak{c}), \infty)$, we have*

$$\inf_{\lambda \in \mathcal{F}(\mathfrak{c}, b)} \inf_{x \in \mathcal{D}} \varphi_\lambda(x)^{-1} \nu_\lambda(x) > 0,$$

where $\varphi_\lambda(x)$ is defined in (4.1).

Proof. Recall that $\mathbb{V}(x) = \exp\left(\frac{a}{b+1}(f(x))^{b+1}\right)$ for every $x \in [0, \infty)$. By Proposition 3.3 (and, more precisely, equation (3.11)) we have

$$\sup_{\lambda \in \mathcal{F}(\mathfrak{c}, b)} \int_{[0, \infty)} \mathbb{V}(x) \nu_\lambda(x) dx < \infty, \quad (4.9)$$

additionally from (4.8) in the proof of Proposition 3.3, we have that $\sup_{\lambda \in \mathcal{F}(\mathfrak{c}, b)} \mathcal{P}_\lambda \mathbb{V}(x) < \infty$ for every $x \in [0, \infty)$. As a consequence, for every $x \in (0, \infty)$, we have

$$\int_{f^{-1}(x)}^{\infty} \nu_\lambda(y) dy \leq \exp\left(-\frac{a}{b+1}(f(x))^{b+1}\right) \int_{[0, \infty)} \mathbb{V}(y) \nu_\lambda(y) dy,$$

and this bound is uniform in λ such that $\lambda \in \mathcal{F}(\mathfrak{c}, b)$ by (4.9). Therefore, for every $x \in (0, \infty)$, we have

$$\sup_{\lambda \in \mathcal{F}(\mathfrak{c}, b)} \int_{f^{-1}(x)}^{\infty} \nu_\lambda(y) dy \leq c(\mathfrak{c}) \exp\left(-\frac{a}{b+1}(f(x))^{b+1}\right), \quad (4.10)$$

for some $c(\mathfrak{c}) > 0$. Let

$$d(\mathfrak{c}) > f^{-1}\left(\left(\frac{b+1}{a} \log(c(\mathfrak{c}))\right)^{1/(b+1)}\right). \quad (4.11)$$

By the definition of ν_λ and using (3.3), for every $y \in (0, \infty)$, we now have

$$\begin{aligned} \nu_\lambda(y) &= \int_0^\infty \nu_\lambda(x) \lambda(f^{-1}(y)) e^{-\int_{f(x)}^y \lambda(f^{-1}(u)) g_x(u) du} g_x(y) \mathbb{1}_{\{f(x) \leq y\}} dx \\ &\geq e^{-\int_0^y \lambda(f^{-1}(u)) M(u) du} \lambda(f^{-1}(y)) m(y) \int_0^{f^{-1}(y)} \nu_\lambda(x) dx \\ &\geq e^{-\int_0^y \lambda(f^{-1}(u)) M(u) du} \lambda(f^{-1}(y)) m(y) \left(1 - \int_{f^{-1}(y)}^\infty \nu_\lambda(x) dx\right) \\ &\geq e^{-\int_0^y \lambda(f^{-1}(u)) M(u) du} \lambda(f^{-1}(y)) m(y) \left(1 - c(\mathfrak{c}) \exp\left(-\frac{a}{b+1}(f(y))^{b+1}\right)\right) \end{aligned}$$

where we used (4.10) for the last inequality. By (4.11), for $y \geq d(\mathfrak{c})$ we have

$$\left(1 - c(\mathfrak{c}) \exp\left(-\frac{a}{b+1}(f(y))^{b+1}\right)\right) > 0,$$

and the conclusion follows readily by the definition of φ_λ . \square

For every $y \in (0, \infty)$, define

$$D(y) = \mathbb{E}_{\nu_\lambda} [g_{Z_0}(f(y)) \mathbb{1}_{\{Z_1 \geq f(y), y \geq Z_0\}}], \quad (4.12)$$

$$D_n(y) = n^{-1} \sum_{k=1}^n g_{Z_{k-1}}(f(y)) \mathbb{1}_{\{Z_k \geq f(y), y \geq Z_{k-1}\}}, \quad (4.13)$$

and

$$D_n(y)_{\varpi_n} = \left(n^{-1} \sum_{k=1}^n g_{Z_{k-1}}(f(y)) \mathbb{1}_{\{Z_k \geq f(y), y \geq Z_{k-1}\}} \right) \bigvee \varpi_n. \quad (4.14)$$

Proposition 4.2. *Work under Assumptions 2.1 and 3.1. Let μ be a probability measure on $[0, \infty)$ such that $\int_{[0, \infty)} \mathbb{V}(\mathbf{x})^2 \mu(d\mathbf{x}) < \infty$. If $1 \geq \varpi_n \rightarrow 0$ as $n \rightarrow \infty$, we have*

$$\sup_{y \in \mathcal{D}} \mathbb{E}_\mu [(D_n(y)_{\varpi_n} - D(y))^2] \lesssim n^{-1} \quad (4.15)$$

uniformly in $\lambda \in \mathcal{F}(\mathbf{c}, b) \cap \mathcal{H}^s(f(\mathcal{D}), M_1)$.

We first need the following estimate

Lemma 4.3. *Work under Assumptions 2.1 and 3.1. Let $d(\mathbf{c})$ be defined as in Lemma 4.1. For every compact interval $\mathcal{D} \subset (d(\mathbf{c}), \infty)$ such that $\inf \mathcal{D} \leq f(r)$, we have*

$$\inf_{\lambda \in \mathcal{F}(\mathbf{c}, b) \cap \mathcal{H}^s(f(\mathcal{D}), M_1)} \inf_{y \in \mathcal{D}} D(y) > 0.$$

Proof. By (3.12) and the definition of φ_B in (4.1), we readily have that

$$D(y) = \frac{\nu_\lambda(f(y))}{\lambda(y)} = \frac{\nu_\lambda(f(y))}{\varphi_\lambda(f(y))} e^{-\int_0^{f(y)} \lambda(f^{-1}(u)) M(u) du} m(y).$$

Since

$$\lambda \in \mathcal{F}(\mathbf{c}, b) \cap \mathcal{H}^s(f(\mathcal{D}), M_1),$$

by applying (3.4), we obtain

$$\begin{aligned} \int_0^{f(y)} \lambda(f^{-1}(u)) M(u) du &\leq \int_0^{\sup \mathcal{D}} \lambda(f^{-1}(u)) M(u) du \\ &\leq L + \int_{f(r)}^{\sup \mathcal{D}} \lambda(f^{-1}(u)) \sup_{y \in \mathcal{D}} |M(y)| du \\ &\leq \left(L + \sup_{y \in \mathcal{D}} |M(y)| M_1 \sup \mathcal{D} \right) < \infty, \end{aligned}$$

where we used that $\inf \mathcal{D} \leq f(r)$. It follows that

$$\inf_{y \in \mathcal{D}} \exp \left(- \int_0^{f(y)} \lambda(f^{-1}(u)) M(u) du \right) \geq \exp \left(- (L + \sup_{y \in \mathcal{D}} |M(y)| M_1 \sup \mathcal{D}) \right) > 0$$

and Lemma 4.3 follows by applying Lemma 4.1. \square

Proof of Proposition 4.2. Since $D_n(y)$ and $D(y)$ are bounded by $\sup_{y \in f(\mathcal{D})} M(y)$, we have

$$(D_n(y)_{\varpi_n} - D(y))^2 \lesssim (D_n(y) - D(y))^2 + \mathbf{1}_{\{D_n(y) < \varpi_n\}}. \quad (4.16)$$

Thus, by Lemma 4.3 we may choose n sufficiently large that

$$0 < \varpi_n \leq q = \frac{1}{2} \inf_{\lambda \in \mathcal{F}(\epsilon, b) \cap \mathcal{H}^s(f(\mathcal{D}), M_1)} \inf_{y \in \mathcal{D}} D(y).$$

Since

$$\{D_n(y) < \varpi_n\} \subset \{D_n(y) - D(y) < -q\},$$

by integrating (4.16), we have that $\mathbb{E}_\mu[(D_n(y)_{\varpi_n} - D(y))^2]$ is less than a constant times

$$\mathbb{E}_\mu[(D_n(y) - D(y))^2] + \mathbb{P}_\mu(|D_n(y) - D(y)| \geq q).$$

By the Bienaymé–Tchebychev inequality, this quantity is less than a constant times

$$\mathbb{E}_\mu[(D_n(y) - D(y))^2].$$

Set $G(x, z, y) = g_x(f(y))\mathbb{1}_{\{z \geq f(y), y \geq x\}}$ and note that $G(\cdot, \cdot, \cdot)$ is bounded on $[0, \infty)$ by $\sup_{y \in f(\mathcal{D})} |M(y)|$. It follows that

$$D_n(y) - D(y) = n^{-1} \sum_{k=1}^n \left(G(Z_{k-1}, Z_k, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k-1}, Z_k, y)] \right).$$

Therefore,

$$\begin{aligned} \mathbb{E}_\mu[(D_n(y) - D(y))^2] &= \frac{1}{n^2} \sum_{k, k' \in \{1, \dots, n\}} \mathbb{E}_\mu \left[\left(G(Z_{k-1}, Z_k, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k-1}, Z_k, y)] \right) \right. \\ &\quad \left. \left(G(Z_{k'-1}, Z_{k'}, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k'-1}, Z_{k'}, y)] \right) \right]. \end{aligned} \quad (4.17)$$

For $|k - k'| \geq 2$, applying Markov's property, we get that

$$\begin{aligned} &\mathbb{E}_\mu \left[\left(G(Z_{k-1}, Z_k, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k-1}, Z_k, y)] \right) \right. \\ &\quad \left. \left(G(Z_{k'-1}, Z_{k'}, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k'-1}, Z_{k'}, y)] \right) \middle| Z_i \forall i \leq k \wedge k' \right] \\ &= \int \int \mathcal{P}_\lambda^{k \vee k' - k \wedge k'}(Z_{k \wedge k'}, dz) \mathcal{P}_\lambda(z, dz') \left(G(z, z', y) - \mathbb{E}_{\nu_\lambda} [G(Z_0, Z_1, y)] \right) \\ &\quad \left[G(Z_{k \wedge k' - 1}, Z_{k \wedge k'}, y) - \mathbb{E}_{\nu_\lambda} [G(Z_0, Z_1, y)] \right], \end{aligned}$$

with $k \wedge k' = \min\{k, k'\}$.

Applying Proposition 3.3 with $h(z) = \int \mathcal{P}_\lambda(z, dz') G(z, z', y)$, we get

$$\begin{aligned} &\mathbb{E}_\mu \left[\left(G(Z_{k-1}, Z_k, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k-1}, Z_k, y)] \right) \left(G(Z_{k'-1}, Z_{k'}, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k'-1}, Z_{k'}, y)] \right) \right] \\ &\leq R \mathbb{E}_\mu \left[\mathbb{V}(Z_{k \wedge k'}) \left(G(Z_{k \wedge k' - 1}, Z_{k \wedge k'}, y) - \mathbb{E}_{\nu_\lambda} [G(Z_0, Z_1, y)] \right) \right] \gamma^{k \vee k' - k \wedge k'} \\ &\lesssim \int_E \mathcal{P}_\lambda^{k \wedge k'} \mathbb{V}(x) \mu(dx) \gamma^{k \vee k' - k \wedge k'}, \end{aligned}$$

as the function G is bounded by $\sup_{y \in \mathcal{D}} |M(y)|$.

For $|k - k'| = 1$, we suppose for example that $k' = k - 1$. Applying the Markov property, we get that

$$\begin{aligned} & \mathbb{E}_\mu \left[\left(G(Z_{k-1}, Z_k, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k-1}, Z_k, y)] \right) \right. \\ & \quad \left. \left(G(Z_{k'-1}, Z_{k'}, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k'-1}, Z_{k'}, y)] \right) \middle| Z_i \forall i \leq k-1 \right] \\ &= \int \mathcal{P}_\lambda(Z_{k-1}, dz) \left(G(Z_{k-1}, z, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k'-1}, Z_{k'}, y)] \right) \\ & \quad \left[G(Z_{k-2}, Z_{k-1}, y) - \mathbb{E}_{\nu_\lambda} [G(Z_0, Z_1, y)] \right]. \end{aligned}$$

Applying Proposition 3.3 again, we get

$$\begin{aligned} & \mathbb{E}_\mu \left[\left(G(Z_{k-1}, Z_k, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k-1}, Z_k, y)] \right) \left(G(Z_{k'-1}, Z_{k'}, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k'-1}, Z_{k'}, y)] \right) \right] \\ & \leq R \mathbb{E}_\mu [\mathbb{V}(Z_{k \wedge k'})] \left(G(Z_{k-2}, Z_{k-1}, y) - \mathbb{E}_{\nu_\lambda} [G(Z_0, Z_1, y)] \right) \gamma^{k \vee k' - k \wedge k'} \\ & \lesssim \int_E \mathcal{P}_\lambda^{k \wedge k'} \mathbb{V}(x) \mu(dx) \gamma^{k \vee k' - k \wedge k'}, \end{aligned}$$

as the function G is bounded by $\sup_{y \in f(\mathcal{D})} |M(y)|$.

For $k = k'$,

$$\mathbb{E}_\mu \left[\left(G(Z_{k-1}, Z_k, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k-1}, Z_k, y)] \right)^2 \middle| Z_i \forall i \leq k-1 \right] \lesssim \gamma^{k \vee k' - k \wedge k'},$$

as the function G is bounded by $\sup_{y \in f(\mathcal{D})} |M(y)|$.

Moreover as \mathbb{V} satisfies (4.8), we get

$$\sup_{\lambda \in \mathcal{F}(\mathbf{c}, b)} \mathcal{P}_\lambda^{k \wedge k'} \mathbb{V}(x) \lesssim 1 + \mathbb{V}(x) \quad (4.18)$$

and, thus, for any k and k' ,

$$\begin{aligned} & \mathbb{E}_\mu \left[\left(G(Z_{k-1}, Z_k, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k-1}, Z_k, y)] \right) \left(G(Z_{k'-1}, Z_{k'}, y) - \mathbb{E}_{\nu_\lambda} [G(Z_{k'-1}, Z_{k'}, y)] \right) \right] \\ & \lesssim \int_{\mathbb{R}_+} (1 + \mathbb{V}(x)) \mu(dx) \gamma^{k \vee k' - k \wedge k'}. \end{aligned}$$

Since \mathbb{V} is μ -integrable by assumption, thanks to (4.17), we have

$$\mathbb{E}_\mu \left[(D_n(y) - D(y))^2 \right] \lesssim \frac{1}{n^2} \sum_{k, k' \in \{1, 2, \dots, n\}} \gamma^{k \vee k' - k \wedge k'} \lesssim n^{-1}, \quad (4.19)$$

uniformly in $y \in \mathcal{D}$ and λ such that $\lambda \in \mathcal{F}(\mathbf{c}, b)$. □

Proposition 4.4. *Work under Assumptions 2.1 and 3.4. Let μ be a probability measure on $[0, \infty)$ such that $\int_{[0, \infty)} \mathbb{V}(x)^2 \mu(dx) < \infty$. Then we have*

$$\sup_{y \in \mathcal{D}} \mathbb{E}_\mu \left[(K_{h_n} \star \widehat{\nu}_n(y) - K_{h_n} \star \nu_\lambda(y))^2 \right] \lesssim (nh_n)^{-1} \quad (4.20)$$

uniformly in $\lambda \in \mathcal{F}(\mathbf{c}, b)$ and $\|f\|_{L^\infty(\mathcal{D})} \leq M_2$ with $\widehat{\nu}_n(\cdot) = \frac{1}{n} \sum_{k \in \{1, \dots, n\}} \mathbb{1}_{\{Z_k\}}(\cdot)$

Proof. By definition,

$$\begin{aligned} \mathbb{E}_\mu \left[\left(K_{h_n} \star \widehat{\nu}_n(y) - K_{h_n} \star \nu_\lambda(y) \right)^2 \right] &= (nh_n)^{-2} \mathbb{E}_\mu \left[\left(\sum_{k \in \{1, \dots, n\}} K \left(\frac{Z_k - y}{h_n} \right) - \mathbb{E}_{\nu_\lambda} \left[K \left(\frac{Z_0 - y}{h_n} \right) \right] \right)^2 \right] \\ &= (nh_n)^{-2} \sum_{k, k' \in \{1, \dots, n\}} \mathbb{E}_\mu \left[\widetilde{K} \left(\frac{Z_k - y}{h_n} \right) \widetilde{K} \left(\frac{Z_{k'} - y}{h_n} \right) \right], \end{aligned}$$

with $\widetilde{K} \left(\frac{Z_k - y}{h_n} \right) = K \left(\frac{Z_k - y}{h_n} \right) - \mathbb{E}_{\nu_\lambda} \left[K \left(\frac{Z_0 - y}{h_n} \right) \right]$.

As in the proof of Proposition 4.2, thanks to the Markov property we obtain

$$\mathbb{E}_\mu \left(\widetilde{K} \left(\frac{Z_k - y}{h_n} \right) \widetilde{K} \left(\frac{Z_{k'} - y}{h_n} \right) \middle| Z_{k \wedge k'} \right) = \left[\mathcal{P}_\lambda^{k \vee k' - k \wedge k'} J(Z_{k \wedge k'}) - \mathbb{E}_{\nu_\lambda} (K \left(\frac{Z_0 - y}{h_n} \right)) \right] \left(K \left(\frac{Z_{k \wedge k'} - y}{h_n} \right) - \mathbb{E}_{\nu_\lambda} (K \left(\frac{Z_0 - y}{h_n} \right)) \right),$$

with $J(\cdot) = K \left(\frac{\cdot - y}{h_n} \right)$. First, as K has bounded support, $K \lesssim \mathbb{V}$ and so we can apply (3.11) from Proposition 3.3. We obtain

$$\left| \left[\mathcal{P}_\lambda^{k \vee k' - k \wedge k'} J(Z_{k \wedge k'}) - \mathbb{E}_{\nu_\lambda} (K \left(\frac{Z_0 - y}{h_n} \right)) \right] \right| \leq R \mathbb{V}(Z_{k \wedge k'}) \gamma^{k \vee k' - k \wedge k'}. \quad (4.21)$$

Moreover, thanks to (3.12) and (3.3) and the fact that $\lambda \in \mathcal{H}^s(\mathcal{D}, M_1)$, we have that

$$\sup_{x \in f(\mathcal{D})} \nu_\lambda(x) \leq M_1 \sup_{y \in \mathcal{D}} M(y), \quad (4.22)$$

so that

$$\left| \mathbb{E}_{\nu_\lambda} (K \left(\frac{Z_0 - y}{h_n} \right)) \right| \leq \int_{[0, \infty)} |K \left(\frac{x - y}{h_n} \right)| \nu_\lambda(x) dx \lesssim h_n. \quad (4.23)$$

Putting together (4.21) and (4.23) we derive

$$\begin{aligned} &\mathbb{E}_\mu \left[\mathbb{E}_\mu \left(\widetilde{K} \left(\frac{Z_k - y}{h_n} \right) \widetilde{K} \left(\frac{Z_{k'} - y}{h_n} \right) \middle| Z_{k \wedge k'} \right) \right] \\ &\lesssim \mathbb{E}_\mu (R J(Z_{k \wedge k'}) \mathbb{V}(Z_{k \wedge k'})) \gamma^{k \vee k' - k \wedge k'} + \mathbb{E}_\mu (R \mathbb{V}(Z_{k \wedge k'})) \gamma^{k \vee k' - k \wedge k'} h_n. \end{aligned}$$

On the one hand, by using the Markov property, the fact that

$$e^{-\int_{f(Z_{k \wedge k' - 1})}^u \lambda(f^{-1}(v)) g_{Z_{k \wedge k' - 1}}(v) dv} \leq 1,$$

(3.3), that $\lambda \in \mathcal{H}^s(f(\mathcal{D}), M_1)$ and as f is increasing, we can bound $\mathbb{V}(u) \mathbb{1}_{\{u \in \mathcal{D}\}}$ by $e^{\frac{\alpha}{b+1} M_2^{b+1}}$. Thus we get that

$$\begin{aligned} &\mathbb{E}_\mu \left[\left| J(Z_{k \wedge k'}) \mathbb{V}(Z_{k \wedge k'}) \right| \right] \\ &\leq \mathbb{E}_\mu \left[\left| \int_{f(Z_{k \wedge k' - 1})}^\infty K \left(\frac{u - y}{h_n} \right) \lambda(f^{-1}(u)) e^{-\int_{f(Z_{k \wedge k' - 1})}^u \lambda(f^{-1}(v)) g_{Z_{k \wedge k' - 1}}(v) dv} g_{Z_{k \wedge k' - 1}}(u) \mathbb{V}(u) du \right| \right] \\ &\leq \left(\sup_{x \in \mathcal{D}} |M(x)| \right) M_1 \int_{[0, \infty)} \mathbb{E}_\mu |K(h_n^{-1}(u - y))| du \lesssim h_n \end{aligned}$$

as K has compact support. On the other hand, as \mathbb{V}^2 and \mathbb{V} are μ integrable by assumption we obtained that

$$\mathbb{E}_\mu \left[\mathbb{E}_\mu \left(\widetilde{K} \left(\frac{Z_k - y}{h_n} \right) \widetilde{K} \left(\frac{Z_{k'} - y}{h_n} \right) \middle| Z_{k \wedge k'} \right) \right] \lesssim h_n \gamma^{k \vee k' - k \wedge k'}.$$

Therefore,

$$\mathbb{E}_\mu \left[\left(K_{h_n} \star \widehat{\nu}_n(y) - K_{h_n} \star \nu_\lambda(y) \right)^2 \right] \lesssim (nh_n)^{-2} \sum_{k, k' \in \{1, \dots, n\}} h_n \gamma^{k \vee k' - k \wedge k'} \lesssim (nh_n)^{-1}. \quad \square$$

4.3. Proof of Theorem 3.5

Recall that

$$\widehat{\lambda}_n(y) = \frac{n^{-1} \sum_{k=1}^n K_{h_n}(Z_k - f(y))}{(n^{-1} \sum_{k=1}^n g_{Z_{k-1}}(f(y)) \mathbb{1}_{\{Z_k \geq f(y), y \geq Z_{k-1}\}}) \vee \varpi_n}$$

and

$$\lambda(y) = \frac{\nu_\lambda(f(y))}{\mathbb{E}_{\nu_\lambda}(g_{Z_0}(f(y)) \mathbb{1}_{\{Z_0 \leq y\}}) \mathbb{1}_{\{Z_1 \geq f(y)\}}}.$$

We will use the decomposition

$$\widehat{\lambda}_n(y) - \lambda(y) = (I + II + III),$$

where

$$\begin{aligned} I &= \frac{K_{h_n} \star \nu_\lambda(f(y)) - \nu_\lambda(f(y))}{D(y)}, \\ II &= \frac{K_{h_n} \star \widehat{\nu}_n(f(y)) - K_{h_n} \star \nu_\lambda(f(y))}{D_n(y) \varpi_n}, \\ III &= \frac{K_{h_n} \star \nu_\lambda(f(y))}{D_n(y) \varpi_n D(y)} (D(y) - D_n(y) \varpi_n), \end{aligned}$$

and where $D(y)$ and $D_n(y) \varpi_n$ are defined in (4.12) and (4.14) respectively. It follows that

$$\|\widehat{\lambda}_n - \lambda\|_{L^2(\mathcal{D})}^2 = \int_{\mathcal{D}} (\widehat{\lambda}_n(y) - \lambda(y))^2 dy \lesssim IV + V + VI,$$

where

$$\begin{aligned} IV &= \int_{\mathcal{D}} (K_{h_n} \star \nu_\lambda(f(y)) - \nu_\lambda(f(y)))^2 \frac{1}{D(y)^2} dy \\ V &= \int_{\mathcal{D}} (K_{h_n} \star \widehat{\nu}_n(f(y)) - K_{h_n} \star \nu_\lambda(f(y)))^2 D_n(y) \varpi_n^{-2} dy \\ VI &= \int_{\mathcal{D}} (D_n(y) \varpi_n - D(y))^2 (K_{h_n} \star \nu_\lambda(f(y)))^2 (D_n(y) \varpi_n D(y))^{-2} dy. \end{aligned}$$

The term IV. We get rid of the term $\frac{1}{D(y)^2}$ using Lemma 4.3. By Assumption 3.4 and classical kernel approximation, we have for every $0 < s \leq n_0$

$$IV \lesssim \|K_{h_n} \star \nu_\lambda - \nu_\lambda\|_{L^2(f(\mathcal{D}))}^2 \lesssim |\nu_\lambda|_{\mathcal{H}^s(f(\mathcal{D}))}^2 h_n^{2s}. \quad (4.24)$$

Lemma 4.5. *We work under Assumptions 2.1 and 3.1. Let $\mathcal{D} \subset (0, \infty)$ be a compact interval, $\lambda \in \mathcal{F}(\mathbf{c}, b)$ for some \mathbf{c} satisfying Assumption 3.1, $g_x \in \mathcal{H}^s(\mathcal{D})$ and $f^{-1} \in \mathcal{H}^s(\mathcal{D})$. Then we have*

$$\|\nu_\lambda\|_{\mathcal{H}^s(\mathcal{D})} \leq \psi(\mathcal{D}, \|\lambda\|_{\mathcal{H}^s(\mathcal{D})}, \|g_x\|_{\mathcal{H}^s(\mathcal{D})}, \|f^{-1}\|_{\mathcal{H}^s(\mathcal{D})})$$

for some continuous function ψ .

Proof of Lemma 4.5. We first recall that

$$\nu_\lambda(y) = \lambda(f^{-1}(y)) \int_E \nu_\lambda(x) e^{-\int_{f(x)}^y \lambda(f^{-1}(u)) g_x(u) du} g_x(y) \mathbb{1}_{\{f(x) \leq y\}} dx.$$

Define

$$A_\lambda(x, y) = e^{-\int_{f(x)}^y \lambda(f^{-1}(u))g_x(u)du} g_x(y).$$

If $\lambda \in \mathcal{F}(\mathbf{c})$, then $A_\lambda(x, \cdot) \in \mathcal{H}^s(\mathcal{D})$ for every $y \in [0, \infty)$, and we have

$$\|A_\lambda(x, \cdot)\|_{\mathcal{H}^s(\mathcal{D})} \leq \psi_1(\|\lambda\|_{\mathcal{H}^s(\mathcal{D})}, \|g_x\|_{\mathcal{H}^s(\mathcal{D})}, \|f^{-1}\|_{\mathcal{H}^s(\mathcal{D})})$$

for some continuous function ψ_1 .

The result is then a consequence of the representation $\nu_\lambda(y) = \lambda(f^{-1}(y)) \int_0^{f^{-1}(y)} \nu_\lambda(x) A_\lambda(x, y) dx$. \square

Returning to (4.24) we deduce from Lemma 4.5 that $\|\nu_\lambda\|_{\mathcal{H}^s(f(\mathcal{D}))}$ is bounded above by a constant that depends on $\mathcal{D}, \|g_x\|_{\mathcal{H}^s(f(\mathcal{D}))}, \|f^{-1}\|_{\mathcal{H}^s(f(\mathcal{D}))}$ and $\|\lambda\|_{\mathcal{H}^s(f(\mathcal{D}))}$ only. It follows that

$$IV \lesssim h_n^{2s} \quad (4.25)$$

uniformly in $\lambda \in \mathcal{H}^s(\mathcal{D}, M_1)$. \square

The term V. We have

$$\mathbb{E}_\mu[V] \leq \varpi_n^{-2} |\mathcal{D}| \sup_{y \in \mathcal{D}} \mathbb{E}_\mu[(K_{h_n} \star \widehat{\nu}_n(f(y)) - K_{h_n} \star \nu_\lambda(f(y)))^2].$$

By (4.20) of Proposition 4.4 we derive that

$$\mathbb{E}_\mu[V] \lesssim \varpi_n^{-2} (nh_n)^{-1} \quad (4.26)$$

uniformly in $\lambda \in \mathcal{F}(\mathbf{c}, b)$. \square

The term VI. First, thanks to Lemma 4.3, we get that

$$\inf_{\lambda \in \mathcal{F}(\mathbf{c}, b) \cap \mathcal{H}^s(f(\mathcal{D}), M_1)} \inf_{y \in f(\mathcal{D})} D_n(y) \varpi D(y) \gtrsim \varpi_n.$$

Next,

$$\sup_{y \in f(\mathcal{D})} |K_{h_n} \star \nu_\lambda(y)| = \sup_{y \in f(\mathcal{D})} \left| \int_{[0, \infty)} K_{h_n}(z - y) \nu_\lambda(z) dz \right| \leq \sup_{y \in \mathcal{D}_1} \nu_\lambda(y) \|K\|_{L^1([0, \infty))}, \quad (4.27)$$

where $\mathcal{D}_1 = \{y + z, y \in f(\mathcal{D}), z \in \text{supp}(K_{h_n})\} \subset \widetilde{\mathcal{D}}$, for some compact interval $\widetilde{\mathcal{D}}$ since K has compact support by Assumption 3.4. Thanks to (4.22), we see that (4.27) holds uniformly in $\lambda \in \mathcal{F}(\mathbf{c}, b)$. We derive that

$$\mathbb{E}_\mu[VI] \lesssim \varpi_n^{-2} \sup_{y \in f(\mathcal{D})} \mathbb{E}_\mu[(D_n(y) \varpi_n - D(y))^2].$$

Applying (4.15) of Proposition 4.2, we conclude that

$$\mathbb{E}_\mu[VI] \lesssim \varpi_n^{-2} n^{-1}, \quad (4.28)$$

uniformly in $\lambda \in \mathcal{F}(\mathbf{c}, b)$. \square

End of the proof of Theorem 3.2. We put together the three estimates (4.25), (4.26) and (4.28) to obtain

$$\mathbb{E}_\mu[\|\widehat{\lambda}_n - \lambda\|_{L^2(\mathcal{D})}^2] \lesssim h_n^{2s} + \varpi_n^{-2} (nh_n)^{-1} + \varpi_n^{-2} n^{-1}$$

uniformly in $\lambda \in \mathcal{H}^s(\mathcal{D}, M_1) \cap \mathcal{F}(\mathbf{c}, b)$. The choice $h_n \sim n^{-1/(2s+1)}$ yields the rate $\varpi_n^{-2} n^{-2s/(2s+1)}$. \square

Acknowledgements. The careful reading and comments of two referees helped to improve a former version of the paper. I also thank Bernard Delyon for the useful discussion about the simulations.

REFERENCES

- [1] O.O. Aalen, *Statistical inference for a family of counting processes*. ProQuest LLC, Ann Arbor, MI. Ph.D. thesis, University of California, Berkeley (1975).
- [2] O.O. Aalen, Weak convergence of stochastic integrals related to counting processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **38** (1977) 261–277.
- [3] O.O. Aalen, Nonparametric inference for a family of counting processes. *Ann. Statist.* **6** (1978) 701–726.
- [4] P.K. Andersen, Ø. Borgan, R.D. Gill and N. Keiding, Statistical models based on counting processes. *Springer Series in Statistics*. Springer-Verlag, New York (1993).
- [5] S. Asmussen and H. Albrecher, Ruin probabilities. Vol. 14 of *Adv. Ser. Statist. Sci. Appl. Probab.* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2nd edition (2010).
- [6] R. Azaïs, F. Dufour and A. Gégout-Petit, Nonparametric estimation of the jump rate for non-homogeneous marked renewal processes. *Ann. Inst. Henri Poincaré, Probab. Stat.* **49** (2013) 1204–1231.
- [7] R. Azaïs, A recursive nonparametric estimator for the transition kernel of a piecewise-deterministic Markov process. *ESAIM: PS* **18** (2014) 726–749.
- [8] R. Azaïs, J.B. Bardet, A. Genadot, N. Krell and P.-A. Zitt, Piecewise deterministic Markov process (pdmps). Recent results. *ESAIM: Proc. Survey* **44** (2014) 276–290
- [9] R. Azaïs, F. Dufour and A. Gégout-Petit, Nonparametric estimation of the conditional distribution of the inter-jumping times for piecewise-deterministic Markov processes. *Scandinavian J. Statist.* **41** (2014) 950–969.
- [10] R. Azaïs and A. Genadot, Semi-parametric inference for the absorption features of a growth-fragmentation model. *TEST* **24** (2015) 341–360.
- [11] J.-B. Bardet, A. Christen and A. Guillin, F. Malrieu and P.-A. Zitt, Total variation estimates for the TCP process. *Electron. J. Probab.* **18** (2013) 10–21.
- [12] P.H. Baxendale, Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Probab.* **15** (2005) 700–738.
- [13] P. Bertail, S. Cléménçon and J. Tressou, A storage model with random release rate for modeling exposure to food contaminants. *Math. Biosci. Eng.* **5** (2008) 35–60.
- [14] P. Bertail, S. Cléménçon and J. Tressou, Statistical analysis of a dynamic model for dietary contaminant exposure. *J. Biol. Dyn.* **4** (2010) 212–234.
- [15] F. Bouguet, Quantitative speeds of convergence for exposure to food contaminants. *ESAIM: PS* **19** (2015) 482–501.
- [16] A.W. Bowman, An alternative method of cross-validation for the smoothing of density estimates. *Biometrika* **71** (1984) 353–360.
- [17] D. Chafaï, F. Malrieu and K. Paroux, On the long time behavior of the TCP window size process. *Stoch. Process. Appl.* **20** (2010) 1518–1534.
- [18] B. Cloez, Wasserstein decay of one dimensional jump-diffusions. Preprint [HAL-00666720](https://hal.archives-ouvertes.fr/00666720), [arXiv:1202.1259](https://arxiv.org/abs/1202.1259) (2012).
- [19] M.H.A. Davis, Markov models and optimization, Vol. 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London (1993).
- [20] M.H.A. Davis, Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models. *J. Roy. Statist. Soc. Ser. B* **46** (1984) 353–388. With discussion.
- [21] M. Doumic, M. Hoffmann, N. Krell and L. Robert, Statistical estimation of a growth-fragmentation model observed on a genealogical tree. *Bernoulli* **21** (2015) 1760–1799.
- [22] V. Dumas, F. Guillemin and Ph. Robert, A Markovian analysis of additive-increase multiplicative-decrease algorithms. *Adv. Appl. Probab.* **34** (2002) 85–111.
- [23] I. Grigorescu and M. Kang, Recurrence and ergodicity for a continuous AIMD model. Preprint, available at http://www.math.miami.edu/~igrigore/pp/b_alpha_0.pdf (2009).
- [24] F. Guillemin, P. Robert and B. Zwart, AIMD algorithms and exponential functionals. *Ann. Appl. Probab.* **14** (2004) 90–117.
- [25] M. Hairer, Martin and J.F. Mattingly, Yet another look at Harris’ ergodic theorem for Markov chains. Seminar on Stochastic Analysis, Random Fields and Applications VI. *Progr. Probab.* **63** (2011) 109–117.
- [26] P. Hall, J.S. Marron and B.U. Park, Smoothed cross-validation. *Probab. Theory Related Fields* **92** (1992) 1–20.
- [27] P. Laurennot and B. Perthame, Exponential decay for the growth-fragmentation/cell-division equation. *Commun. Math. Sci.* **7** (2009) 503–510.
- [28] M. Rudemo, Empirical choice of histograms and kernel density estimators. *Scand. J. Statist.* **9** (1982) 65–78.