



**HAL**  
open science

# Level Crossings and Absorption of an Insurance Model

Romain Azaïs, Alexandre Genadot

► **To cite this version:**

Romain Azaïs, Alexandre Genadot. Level Crossings and Absorption of an Insurance Model. Romain Azaïs; Florian Bouguet. Statistical Inference for Piecewise-deterministic Markov Processes, Wiley, pp.65-105, 2018, 978-1-786-30302-8. 10.1002/9781119507338.ch3 . hal-01862266

**HAL Id: hal-01862266**

**<https://hal.science/hal-01862266v1>**

Submitted on 3 Sep 2019

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

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

# LEVEL CROSSINGS AND ABSORPTION OF AN INSURANCE MODEL

ROMAIN AZAÏS<sup>1</sup> AND ALEXANDRE GENADOT<sup>2</sup>

<sup>1</sup> *Laboratoire Reproduction et Développement des Plantes, Univ Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRA, Inria, F-69342, Lyon, France*

<sup>2</sup> *IMB, Institut de Mathématiques de Bordeaux, Université de Bordeaux, France and INRIA Bordeaux Sud Ouest, Team: CQFD*

In this chapter, we undertake the study of a particular piecewise-deterministic process modeling the evolution of a capital submitted to catastrophic event occurring at random times and with random intensities. When such an event leads the capital to go under a certain threshold, the ruin holds. We are interested in the inference of the probability and times of ruin of the process as well as in the average number of time a given capital is outreached. In this chapter, we revisit the two articles [3] and [4] where additional informations and details can be found, including the implementation of the developed methods in both multi-dimensional and real data frameworks.

When one assumption is stated, it is assumed throughout the chapter, unless explicitly said.

## 1. AN INSURANCE MODEL

We consider the model introduced by [11] describing the evolution of a capital subject to random heavy loss events. We do not attempt to give an exhaustive survey about this model, but refer the reader to the article [11] and the references therein. We consider an individual household whose income  $I_t$  at time  $t$  may be split into

$$I_t = C_t + \sigma_t \tag{1.1}$$

where  $C_t$  denotes the consumption and  $\sigma_t$  is the savings. Consumption is assumed to evolve according to

$$C_t = \begin{cases} I_t & \text{if } I_t \leq I^* \\ I^* + a(I_t - I^*) & \text{else,} \end{cases} \tag{1.2}$$

where  $I^*$  is the critical income level and  $0 < a < 1$ . If the income is smaller than  $I^*$ , the whole income is used for consumption. We denote by  $X_t$  the accumulated capital up to time  $t$ . The capital evolves according to

$$\frac{dX_t}{dt} = c\sigma_t, \tag{1.3}$$

where  $0 < c < 1$ , while the income evolves with

$$I_t = bX_t, \tag{1.4}$$

with  $b > 0$ . Finally, from (1.1), (1.2), (1.3) and (1.4), the capital  $X_t$  satisfies the ODE,

$$\frac{dX_t}{dt} = r(X_t - x^*)^+, \quad (1.5)$$

where  $x^* = I^*/b$  and  $r = (1 - a)bc$ . Now, we assume that the capital  $X_t$  is subject to catastrophic events, which occur in a Poisson-like fashion with homogeneous rate  $\lambda$ . When an event occurs at time  $t$ , the capital  $X_t$  is reduced by a random fraction whose distribution is described by its probability density function  $G$ . After the loss event, the process starts again according to (1.5). We obtain a process for which the interval  $[0, x^*]$  is an absorbing set, called *area of poverty*. Indeed, once the process is below the critical capital  $x^*$ , the next events will reduce the capital and the process will never again reach values above  $x^*$ .

In this chapter, we will be interested in the two following natural questions.

- Can we, from the observation of a trajectory of the process, infer the probability and contingent time of ruin, that is, the absorption features of the process?
- Can we, still from the observation of a trajectory of the process, infer the average number of times the capital overcomes a given threshold, for instance before ruin?

In the sequel, for simplicity, we set the critical capital  $x^*$  at the value 1. The insurance model we just described is a growth-fragmentation model, that is a Markov process  $(X_t)_{t \geq 0}$  defined on some probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  – from a probability density function  $G$  on the interval  $[0, 1]$  and two real numbers  $r, \lambda > 0$  – by its extended generator  $\mathcal{L}$  as follows,

$$\mathcal{L}f(x) = r(x - 1)^+ f'(x) + \lambda \int_0^1 (f(zx) - f(z)) G(z) dz \quad (1.6)$$

with  $\xi^+ = \xi \vee 0$ , and for any smooth function  $f : \mathbf{R} \rightarrow \mathbf{R}$  (see [6] for full details on the domain of  $\mathcal{L}$ ). The generator (1.6) describes a process having an exponential growth at rate  $r$  and which can only decrease by means of downward jumps. The downward jumps occur at rate  $\lambda$  and instantaneously reduce the process of a certain percentage distributed according to the density  $G$ . The positive part present in Equation (1.6) implies that the exponential growth is annihilated below the threshold 1, preventing the process to escape the domain  $[0, 1]$ . In the sequel, we propose to describe the motion of our growth-fragmentation model as the dynamics of a Piecewise-Deterministic Markov Process (PDMP).

In most cases, the dynamics of a one-dimensional real-valued PDMP is described by its three local features  $(\lambda, Q, \Phi)$ .

- $\Phi : \mathbf{R} \times \mathbf{R}_+ \rightarrow \mathbf{R}$  is the deterministic flow. It satisfies,
 
$$\forall \xi \in \mathbf{R}, \forall s, t \geq 0, \Phi(\xi, 0) = \xi \quad \text{and} \quad \Phi(\xi, t + s) = \Phi(\Phi(\xi, t), s).$$
- $\lambda : \mathbf{R} \rightarrow \mathbf{R}_+$  is the jump rate. It satisfies,

$$\forall \xi \in \mathbf{R}, \exists \varepsilon > 0, \int_0^\varepsilon \lambda(\Phi(\xi, s)) ds < \infty.$$

- $Q$  is a Markov kernel on  $\mathbf{R}$  which satisfies,

$$\forall \xi \in \mathbf{R}, Q(\xi, \{\xi\}) = 0.$$

Starting from  $X_0 = x$ , the motion can be described as follows. The first jump time  $T_1$  is a positive random variable whose survival function is,

$$\forall t \geq 0, \mathbf{P}(T_1 > t | X_0 = x) = \exp\left(-\int_0^t \lambda(\Phi(x, s)) ds\right).$$

This jump time occurs in a Poisson-like fashion with nonhomogeneous rate  $\lambda$ . One chooses a real-valued random variable  $Z_1$  according to the distribution  $Q(\Phi(x, T_1), \cdot)$ . Let us remark that the post-jump location  $Z_1$  depends on the interarrival time  $T_1$ , via the deterministic flow starting from  $X_0 = x$ . The trajectory between the times 0 and  $T_1$  is given by

$$X_t = \begin{cases} \Phi(x, t) & \text{for } 0 \leq t < T_1, \\ Z_1 & \text{for } t = T_1. \end{cases}$$

Now, starting from  $X_{T_1}$ , one may choose the interarrival time  $S_2 = T_2 - T_1$  and the post-jump location  $Z_2$  in a similar way as before, and so on. The randomness of such a process is only given by the jump mechanism.

In our particular case, one may easily compute from (1.6) the local features of the growth-fragmentation model  $(X_t)_{t \geq 0}$  that we consider. They are given by

$$\Phi(x, t) = \begin{cases} (x-1)\exp(rt) + 1 & \text{if } x > 1, \\ x & \text{else,} \end{cases} \quad (1.7)$$

$$\lambda(x) = \lambda, \quad \text{and} \quad Q(x, dy) = \frac{1}{x} G\left(\frac{y}{x}\right) dy. \quad (1.8)$$

Notice that the flow  $\Phi$  given by Equation (1.7) arises, according to (1.6), as the solution of the following ODE,

$$y'_t = r(y_t - 1)^+ \quad \text{and} \quad y_0 = x,$$

describing the kinetic of the process between jumps. Moreover, we see in particular from (1.8), that the rate of jump  $\tilde{\lambda}$  is homogeneous and that jumps are downward since  $G$  is a density on  $[0, 1]$ . Therefore, the sequence of the interarrival times  $(S_n)_{n \geq 1}$  is independent and exponentially distributed with rate  $\lambda$ . In addition, the particular form of the transition kernel  $Q$  (1.8) implies that the sequence of the random loss fractions  $(V_n)_{n \geq 1}$  defined from,

$$\forall n \geq 1, Z_n = \Phi(Z_{n-1}, S_n) V_n, \quad (1.9)$$

is independent and independent of  $(S_n)_{n \geq 1}$  with common distribution  $G$  (we refer the reader to Lemma 5.1). As a consequence, the dynamics of the PDMP  $(X_t)_{t \geq 0}$  may be summarized by the observation of the independent sequences  $(S_n)_{n \geq 1}$  and  $(V_n)_{n \geq 1}$ . Two possible trajectories of the process  $(X_t)_{t \geq 0}$  are given in Figure 1.1.

This continuous-time Markov process is called absorbing because the motion may reach the absorbing interval  $\Gamma = [0, 1]$  from any initial value  $X_0 = x$ . In this chapter, we are mainly interested in given statistical answers to the two following questions:

- How to compute the absorption probability and the distribution of the hitting time of  $\Gamma$ ?
- How to compute the average number of crossings of a given level?

Let us notice, for the second section, that since the process is monotone between jumps, the number of crossings of a given level is twice the number of continuous crossings of this level (the continuous crossings are the crossings that are not due to the jumps of the process). This is therefore sufficient to study the average number

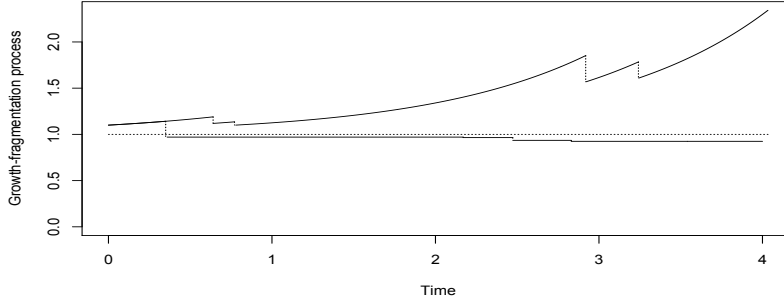


FIGURE 1.1. Two trajectories for the growth-fragmentation model starting from  $X_0 = 1.1$  and whose characteristics are  $\lambda = r = 1$  and  $G(u) = 11u^{10}$ , as in the numerical example investigated in Subsection 3.3. One of them is absorbed at the first jump time, while the other one seems to escape the trapping set.

of continuous crossings. In Section 2, we assess these two above questions from a theoretical point of view.

*Remark 1.1.* The growth-fragmentation process described in this section satisfies the following Poisson driven SDE,

$$dX_t = r(X_t - 1)^+ dt + X_{t-} \int_0^1 (z - 1) \mathcal{N}(dt \times dz), \quad t \geq 0, \quad (1.10)$$

with initial condition  $X_0 = x$ . The measure  $\mathcal{N}$  is a Poisson random measure on  $\mathbf{R}_+ \times [0, 1]$  with intensity  $\nu$ , the measure on  $\mathbf{R}_+ \times [0, 1]$  defined by

$$\nu(dt \times dz) = \lambda dt G(z) dz.$$

Equation (1.10) provides another dynamical representation for the process  $(X_t)_{t \geq 0}$  which may be used, for example, for simulation purposes. A closed, but not very tractable, form for  $(X_t)_{t \geq 0}$  may be derived from (1.10).

*Assumption 1.1.* The initial state  $X_0$  has a density with respect to the Lebesgue measure which is continuous and bounded over  $\mathbf{R}_+^*$ . Moreover, the probability to jump close to zero is not too large in the sense that

$$\int_0^1 \frac{G(u)}{u} du < \infty.$$

Under Assumption 1.1, a cautious study of the closed form associated to (1.10) allows to prove the following proposition.

*Proposition 1.2.* Under Assumption 1.1, for any  $t \in [0, T]$ , the random variable  $X_t$  has a density  $p_{X_t}(x)$  with respect to the Lebesgue measure that is continuous with respect to  $x > 1$ .

## 2. SOME RESULTS ABOUT THE CROSSING AND ABSORPTION FEATURES

**2.1. Transition Density of the Post-jump Locations.** Recall that the sequence of the jump times of the PDMP  $(X_t)_{t \geq 0}$  is  $(T_n)_{n \geq 0}$  and the post-jump

locations are defined, for any  $n \geq 0$ , by  $Z_n = X_{T_n}$ . The Markov kernel  $R$  of  $(Z_n)_{n \geq 0}$  is given, for any  $x \in \mathbf{R}_+^*$  and  $A \in \mathcal{B}(\mathbf{R}_+^*)$ , by

$$R(x, A) = \mathbf{P}(Z_{n+1} \in A \mid Z_n = x).$$

We show that  $R$  can be directly computed from the characteristics  $G$  and  $\lambda$ .

**Theorem 2.1.** *The transition kernel of  $(Z_n)_{n \geq 0}$  satisfies  $R(x, dy) = R(x, y)dy$ , with*

$$R(x, y) = \begin{cases} \frac{1}{x} G\left(\frac{y}{x}\right) & \text{if } x \leq 1, \\ \frac{\lambda}{r} (x-1)^{\frac{\lambda}{r}} \left[ \int_0^{\frac{y}{x} \wedge 1} G(u) u^{\frac{\lambda}{r}} (y-u)^{-\frac{\lambda}{r}-1} du \right] & \text{else.} \end{cases} \quad (2.1)$$

For convenience, we use in some proofs the following notation. For  $\lambda > 0$ ,  $x \geq 1$ ,  $y \geq 0$  and  $u \in [0, 1]$ , we define

$$\alpha_\lambda(x) = \frac{(x-1)^{\lambda/r}}{r} \quad \text{and} \quad \beta_\lambda(y, u) = u^{\lambda/r} (y-u)^{-\lambda/r-1}. \quad (2.2)$$

*Proof of Theorem 2.1.* In both cases  $x > 1$  and  $x \leq 1$ , we have

$$R(x, dy) = \left[ \int_{\mathbf{R}_+} \frac{1}{z} G\left(\frac{y}{z}\right) \mathcal{S}(x, dz) \right] dy, \quad (2.3)$$

where the conditional distribution  $\mathcal{S}(x, dz)$  is defined from its cumulative version,

$$\mathcal{S}(x, (-\infty, z]) = \mathbf{P}(\Phi(Z_{n-1}, S_n) \leq z \mid Z_{n-1} = x).$$

For  $x \leq 1$ , from (1.7), we have  $\mathcal{S}(x, dz) = \delta_{\{x\}}(dz)$ . This shows (2.1) for  $x \leq 1$ . If  $x > 1$ , for any  $z \geq x$ ,

$$\mathcal{S}(x, (-\infty, z]) = \mathbf{P}\left(S_n \leq \frac{1}{r} \log\left(\frac{z-1}{x-1}\right)\right).$$

As a consequence, we have

$$\mathcal{S}(x, dz) = \frac{\lambda}{r} \frac{(x-1)^{\frac{\lambda}{r}}}{(z-1)^{\frac{\lambda}{r}+1}} \mathbb{1}_{[x, \infty)}(z) dz. \quad (2.4)$$

Since  $G$  is a probability density function on  $[0, 1]$ , together with (2.4), we may re-write (2.3) as

$$R(x, dy) = \frac{\lambda}{r} (x-1)^{\frac{\lambda}{r}} \left[ \int_{x \vee y}^{+\infty} G\left(\frac{y}{z}\right) \frac{(z-1)^{-\frac{\lambda}{r}-1}}{z} dz \right] dy$$

By the change of variable  $u = y/z$ , we obtain

$$R(x, y) = \frac{\lambda}{r} (x-1)^{\frac{\lambda}{r}} \left[ \int_0^{\frac{y}{x} \wedge 1} G(u) \left(\frac{y-u}{u}\right)^{-\frac{\lambda}{r}-1} u^{-1} du \right].$$

This shows the result (2.1) for  $x > 1$ .  $\square$

**2.2. Absorption Time and Probability.** When the trajectory of the PDMP  $(X_t)_{t \geq 0}$  starts from  $X_0 = x > 1$ , the absorption probability is defined by

$$p(x) = \mathbf{P}(X_t \in \Gamma \text{ for some } t \geq 0 \mid X_0 = x),$$

where  $\Gamma = [0, 1]$  is the absorbing set. We state that  $p(x)$  may be found as a solution of an integral equation.

**Theorem 2.2.**  *$p(x)$  is a solution of the following integral equation,*

$$p(x) = \int_0^1 R(x, y) dy + \int_1^{+\infty} p(y) R(x, y) dy. \quad (2.5)$$

*Proof of Theorem 2.2.* First, we propose to rewrite  $p(x)$  from the Markov chain  $(Z_n)_{n \geq 0}$ ,

$$p(x) = \mathbf{P}(Z_n \in \Gamma \text{ for some } n \mid Z_0 = x).$$

In addition, we have

$$p(x) = \mathbf{P}(Z_1 \in \Gamma \mid Z_0 = x) + \mathbf{P}(Z_1 \notin \Gamma, Z_n \in \Gamma \text{ for some } n \geq 2 \mid Z_0 = x).$$

Together with

$$\begin{aligned} & \mathbf{P}(Z_1 \notin \Gamma, Z_n \in \Gamma \text{ for some } n \geq 2 \mid Z_0 = x) \\ &= \int_1^{+\infty} \mathbf{P}(Z_n \in \Gamma \text{ for some } n \geq 2 \mid Z_1 = y) R(x, y) dy, \end{aligned}$$

and the Markov property for  $(Z_n)_{n \geq 1}$ , this shows (2.5).  $\square$

Notice that since  $R$  is a probability kernel, the constant function  $p(x) = 1$  is a solution to Equation (2.5). We will look for a solution going to zero when the starting point  $x$  goes to infinity. For this purpose, we will work in the space  $L^1(1, \infty)$ , leading us to impose the following condition on the probability density  $G$ , which strengthens Assumption 1.1.

*Assumption 2.1.* We assume that the density  $G$  is such that  $\int_0^1 G(u) u^{-1} du < 1 + \frac{r}{\lambda}$ .

**Proposition 2.3.** *Suppose that Assumption 2.1 holds. Then Equation (2.5) has a unique solution in the space  $L^1(1, \infty)$  endowed with its usual norm  $\|\cdot\|$ .*

*Proof of Proposition 2.3.* Equation (2.5) is associated to the following integral operator on  $L^1(1, \infty)$ ,

$$K : p \mapsto \int_1^{+\infty} p(y) R(\cdot, y) dy. \quad (2.6)$$

Using Jensen's and Fubini's theorems, for any  $h \in L^1(1, \infty)$  we have,

$$\begin{aligned} \|Kh\| &= \int_1^{+\infty} \left| \int_1^{+\infty} h(y) R(x, y) dy \right| dx \leq \int_1^{+\infty} \int_1^{+\infty} |h(y)| R(x, y) dy dx \\ &\leq \sup_{y \in [1, \infty[} \int_1^{+\infty} R(x, y) dx \int_1^{+\infty} |h(y)| dy. \end{aligned}$$

According to Lemma 2.4 below, the above inequalities yield

$$\|Kh\| \leq \frac{\lambda}{\lambda + r} \int_0^1 \frac{G(u)}{u} du \|h\|.$$

Therefore, under the condition,  $\frac{\lambda}{\lambda+r} \int_0^1 \frac{G(u)}{u} du < 1$ , we get  $\|K\| < 1$ . One may then rewrite Equation (2.5) as the Fredholm equation of the second kind, see [12],

$$p - Kp = s,$$

where  $s(x) = \int_0^1 R(x, y) dy$ . This equation has obviously a unique solution since  $\|K\| < 1$ . Notice that  $\|p\| < \infty$  since one may write  $p = \sum_{k=0}^{\infty} K^k s$ .  $\square$

**Lemma 2.4.** *The following equality holds,*

$$\sup_{y \in [0, \infty)} \int_1^{\infty} R(x, y) dx = \frac{\lambda}{\lambda+r} \int_0^1 \frac{G(u)}{u} du.$$

*Proof of Lemma 2.4.* By definition of  $R$  one can write

$$\int_1^{\infty} R(x, y) dx = \lambda \int_1^{\infty} \alpha_{\lambda}(x) \int_0^{y/x \wedge 1} \beta_{\lambda}(y, u) G(u) du dx,$$

where for  $\lambda > 0$ ,  $x \geq 1$ ,  $y \geq 0$  and  $u \in [0, 1]$ , we use the following notation,

$$\alpha_{\lambda}(x) = \frac{(x-1)^{\lambda/r}}{r}, \quad \beta_{\lambda}(y, u) = \frac{u^{\lambda/r}}{(y-u)^{\lambda/r+1}}. \quad (2.7)$$

In the above term, one can change the order of integration to integrate with respect to  $x$  first. We obtain

$$\begin{aligned} & \lambda \int_1^{\infty} \alpha_{\lambda}(x) \int_0^{y/x \wedge 1} \beta_{\lambda}(y, u) G(u) du dx \\ &= \lambda \int_0^1 \int_1^{y/u} \alpha_{\lambda}(x) dx \beta_{\lambda}(y, u) G(u) \mathbb{1}_{\{u \leq y\}} du \\ &= \frac{\lambda}{r} \int_0^1 \frac{1}{\lambda/r+1} \left(\frac{y}{u} - 1\right)^{\lambda/r+1} \beta_{\lambda}(y, u) G(u) \mathbb{1}_{\{u \leq y\}} du \\ &= \frac{\lambda}{\lambda+r} \int_0^1 G(u) \mathbb{1}_{\{u \leq y\}} \frac{du}{u}. \end{aligned}$$

The result follows.  $\square$

We now proceed to the computation of  $t_m(x)$ , the probability for the process  $(X_t)_{t \geq 0}$  starting from  $X_0 = x$  to be absorbed at jump  $m$ . For  $x > 1$  we have  $t_1(x) = \mathbf{P}(Z_1 \in \Gamma | Z_0 = x)$  and, as  $\Gamma$  is absorbing, for  $m \geq 2$ ,

$$t_m(x) = \mathbf{P}(Z_m \in \Gamma, Z_{m-1} \notin \Gamma | Z_0 = x).$$

We state in the following result that this sequence satisfies a recurrence relation.

**Theorem 2.5.** *For  $x > 1$ , the functional sequence  $(t_m)_{m \geq 1}$  satisfies  $t_1(x) = \int_0^1 R(x, y) dy$  and the recursion relation:*

$$\forall m \geq 2, t_m(x) = \int_1^{\infty} t_{m-1}(y) R(x, y) dy. \quad (2.8)$$

*Proof of Theorem 2.5.* The proof follows the same reasoning as in the proof of Proposition 2.2.  $\square$



**2.3. Kac-Rice Formula.** In this section, we present a local-time crossing relation for the PDMP  $(X_t)_{t \geq 0}$ . A more general statement, extending the result to a class of one-dimensional piecewise-smooth processes can be found in [4]. Crossing formulas such as (2.9) presented below can also be derived from [9, Theorem 1] where occupation times for some smooth processes are studied.

This local-time crossing relation allows for the proof of a so-called Kac-Rice formula, giving an explicit form for the average number of continuous crossings by the process of a given level. We will use the formula in Section 4 in order to propose an estimator of this the average number of continuous crossings.

**Theorem 2.6** (Local time-crossing relation). *For any level  $x > 1$  and any time horizon  $T$ , we have almost surely*

$$c_x(T) = r(x-1)^+ l_x(T), \quad (2.9)$$

where  $c_x(T)$  is the number of continuous crossings of  $x$  up to time  $T$  and  $l_x(H)$  is the local time spent at level  $x$  by the process  $X$  between times 0 and  $H$ , defined by

$$l_x(H) = \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_0^T \mathbb{1}_{|X_t - x| \leq \delta} dt.$$

*Proof of Theorem 2.6.* We start the proof with the same arguments as in the proof of [5, Theorem 12]. Since  $(x-1)^+ \neq 0$ , we can apply Kac's counting formula to almost all trajectories of  $X$ . Indeed, the continuity of the drift implies that between two successive jumps, the trajectories of  $X$  are  $\mathcal{C}^1$  and such that  $dX_t/dt = (X_t - 1)^+$  for almost every  $t \in [0, T]$ . Moreover, by Proposition 1.2, the level  $x$  is almost surely not reached at  $T$  nor at the jump times of the process. By continuity of this density and the fact that  $(x-1)^+ \neq 0$ , the trajectories of  $X$  have no tangencies at the level  $x$ . Let us also remark that the number of crossings of the level  $x$  is bounded by the number of jumps of  $X$  plus one, and has therefore finite expectation. Thus, we can apply Kac's counting formula [10] and have almost surely

$$c_x(H) = \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_0^T r(X_t - 1)^+ \mathbb{1}_{|X_t - x| \leq \delta} dt.$$

Let us now work differently from the proof of [5, Theorem 12]. For any  $\delta > 0$  we have

$$\begin{aligned} & \frac{r}{2\delta} \int_0^T (X_t - 1)^+ \mathbb{1}_{|X_t - x| \leq \delta} dt \\ &= \frac{r}{2\delta} \int_0^T ((X_t - 1)^+ - (x - 1)^+) \mathbb{1}_{|X_t - x| \leq \delta} dt + r(x - 1)^+ \frac{1}{2\delta} \int_0^T \mathbb{1}_{|X_t - x| \leq \delta} dt. \end{aligned}$$

The left-hand-side converges towards  $c_x(T)$  by Kac's counting formula. The absolute value of the first term of the right hand side is almost surely bounded by

$$\frac{r}{2\delta} \int_0^T |X_t - x| \mathbb{1}_{|X_t - x| \leq \delta} dt \leq \frac{r}{2} \int_0^T \mathbb{1}_{|X_t - x| \leq \delta} dt,$$

The term

$$\int_0^T \mathbb{1}_{|X_t - x| \leq \delta} dt$$

converges almost surely towards zero since  $X_t$  has a density on  $\mathbf{R}$  for all time  $t$ . Finally we almost surely have

$$c_x(T) = r(x-1)^+ \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_0^T \mathbb{1}_{|X_t - x| \leq \delta} dt,$$

which states the expected result.  $\square$

**Corollary 2.7** (Kac-Rice formula). *For any level  $x > 1$  and any time horizon  $T$ , the following Kac-Rice formula holds,*

$$C_x(T) = \mathbf{E}(c_x(T)) = r(x-1)\mathbf{E}(l_x(T)) = r(x-1) \int_0^T p_{X_s}(x) ds. \quad (2.10)$$

*Proof of Corollary 2.7.* First, let us remark that the fact that  $l_x(H)$  has finite expectation is just a consequence of Fatou's lemma and Fubini-Tonelli theorem,

$$\begin{aligned} \mathbf{E}(l_x(T)) &\leq \liminf_{\delta \rightarrow 0} \frac{1}{2\delta} \mathbf{E} \left( \int_0^T \mathbb{1}_{|X_t - x| \leq \delta} dt \right) \\ &= \liminf_{\delta \rightarrow 0} \frac{1}{2\delta} \int_{x-\delta}^{x+\delta} \int_0^T p_{X_t}(y) dt dy = \int_0^T p_{X_t}(x) dt. \end{aligned}$$

Taking expectation on both sides of (2.9), we obtain the first equality. We are left to show that

$$\mathbf{E}(l_x(T)) = \int_0^T p_{X_s}(x) ds.$$

For this purpose we first use the dominated convergence theorem

$$\begin{aligned} \mathbf{E}(l_x(H)) &= \mathbf{E} \left( \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_0^T \mathbb{1}_{|X_t - x| \leq \delta} dt \right) \\ &= \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_0^T \mathbf{P}(|X_t - x| \leq \delta) dt. \end{aligned}$$

The domination comes from the fact that the term  $\frac{1}{2\delta} \int_0^T \mathbb{1}_{|X(t) - x| \leq \delta} dt$  is bounded, for some  $\delta_0$  small enough, by the number of jumps of the process between times 0 and  $T$ , which is integrable, times a quantity of order

$$\left( \inf_{\{x' : |x' - x| \leq \delta_0\}} (x' - 1)^+ \right)^{-1} = \frac{1}{x - 1 - \delta_0},$$

that is positive for  $\delta_0$  small enough, since  $x > 1$ . Then, from Fubini-Tonelli theorem, using the fact that  $z \mapsto \int_0^T p_{X_t}(z) dt$  is continuous on a neighborhood of  $x$ , we have

$$\begin{aligned} \mathbf{E}(l_x(T)) &= \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_0^T \mathbf{P}(|X_t - x| \leq \delta) dt = \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_0^T \int_{x-\delta}^{x+\delta} p_{X_t}(y) dy dt \\ &= \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_{x-\delta}^{x+\delta} \int_0^T p_{X_t}(y) dt dy \\ &= \int_0^T p_{X_t}(x) dt, \end{aligned}$$

which ends the proof.  $\square$

*Remark 2.8.* In the stationary case, if the density  $p_{X_0}$  is stationary for the process  $X$ , the formula (2.10) reads

$$C_x(T) = r(x-1)p_{X_0}(x)T. \quad (2.11)$$

### 3. INFERENCE FOR THE ABSORPTION FEATURES OF THE PROCESS

We present here our main results on the estimation of the absorption probability and of the hitting time of  $\Gamma$  for the PDMP  $(X_t)_{t \geq 0}$ . Almost all the proofs are deferred to Section 5.

**3.1. Semi-Parametric Framework.** In this section, we assume that we observe the PDMP  $(X_t)_{t \geq 0}$  within a long time interval. In other words, we observe the  $n$  first terms of the sequence of the interarrival times  $(S_k)_{k \geq 1}$  and of the sequence of the independent random loss fractions  $(V_k)_{k \geq 1}$  defined by (1.9) for large  $n$ .

From these independent observations, we propose to estimate the features  $G$  and  $\lambda$ . In the sequel, we consider an estimator  $\widehat{G}_n$  of the density function  $G$  and an estimator  $\widehat{\lambda}_n$  of the rate  $\lambda$ , computed from the  $n$  first loss events, that is to say from  $S_1, \dots, S_n$  and  $V_1, \dots, V_n$ .

*Assumption 3.1.* In some of our convergence results, we impose a few conditions on the asymptotic behaviors of  $\widehat{G}_n$  and  $\widehat{\lambda}_n$ . When one of the assumptions is used, this is specified in the statement of the result.

- $(C_1^\lambda) : \widehat{\lambda}_n \in [\lambda_*, \lambda^*]$ , with  $\lambda_* > 0$ .
- $(C_2^\lambda) : \widehat{\lambda}_n - \lambda$  tends to 0 in probability.
- $(C_1^G) : \|\widehat{G}_n - G\|_\infty$  tends to 0 in probability.
- $(C_2^G) : \int_0^1 |G(u) - \widehat{G}_n(u)| u^{-1} du$  tends to 0 in probability.

*Remark 3.1.* The condition  $(C_1^\lambda)$  states that the sequence of the estimators  $\widehat{\lambda}_n$  is uniformly bounded on  $(0, \infty)$ . The three other conditions,  $(C_2^\lambda)$  and  $(C_{1,2}^G)$ , concern the convergence in probability of the estimators  $\widehat{\lambda}_n$  and  $\widehat{G}_n$ . Regarding  $\widehat{G}_n$ , this convergence in probability is stated with respect to two different norms: the sup-norm for  $(C_1^G)$  and a weighted  $L^1$ -norm for  $(C_2^G)$ . This latter convergence mode implies that the probability to jump back close to zero should not be too large.

In the present chapter, we are not interested in the demonstration of the asymptotic properties of the estimates  $\widehat{G}_n$  and  $\widehat{\lambda}_n$  but in the estimation of some characteristics of the PDMP  $(X_t)_{t \geq 0}$  from these estimates. In particular, we establish in Theorems 3.5 and 3.6 that the convergences in probability of  $\widehat{G}_n$  and  $\widehat{\lambda}_n$  may be transferred to our estimators of the absorption probability and of the distribution of the hitting time of  $\Gamma$ . Thus, we do not investigate the properties or the choice of the estimators of  $G$  and  $\lambda$ . Nevertheless, the assumptions that we impose on  $\widehat{G}_n$  and  $\widehat{\lambda}_n$  are non restrictive as we proceed to show, see also Subsection 5.6. For the sake of readability, we introduce the following notation:  $\overline{S}_n$  denotes the empirical mean of the  $n$  first interarrival times, while the projection  $\pi_{[a,b]}(x)$  is defined by

$$\pi_{[a,b]}(x) = \begin{cases} a & \text{if } x < a, \\ b & \text{if } x > b, \\ x & \text{else.} \end{cases}$$

In the case where we know two real numbers  $\lambda_*$  and  $\lambda^*$  such that  $0 < \lambda_* < \lambda < \lambda^*$ , the truncated maximum likelihood estimator given by

$$\widehat{\lambda}_n^{ml} = \pi_{[\lambda_*, \lambda^*]}(\overline{S}_n^{-1}), \quad (3.1)$$

obviously satisfies both the conditions  $(C_{1,2}^\lambda)$ . Furthermore, the Parzen-Rosenblatt estimator of a uniformly continuous density satisfies the assumption  $(C_1^G)$  whenever the bandwidth  $(h_n)_{n \geq 1}$  is such that

$$\sum_{n \geq 1} \exp(-\delta n h_n^2) < \infty,$$

for any  $\delta > 0$  (see [14] for instance). Finally, we will establish in Appendix 5.6 that the convergence  $(C_2^G)$  is also satisfied for the Parzen-Rosenblatt estimator under some additional conditions (see Assumption 5.1) on the density of interest  $G$ .

*Remark 3.2.* The situation that we consider in the present chapter may be generalized to estimate the absorption features of a large number of stochastic processes. Indeed, our method may be applied to any absorbing jump process satisfying both the following conditions:

- the absorption occurs only at (downward or upward) jumps;
- the motion of the process depends on an easily estimable quantity  $\Theta$  in a parametric, semi-parametric or nonparametric setting. In our semi-parametric framework,  $\Theta = (\lambda, G)$  is estimated from the independent random variables  $(S_i)_{1 \leq i \leq n}$  and  $(Y_i)_{1 \leq i \leq n}$ . Let us denote by  $\widehat{\Theta}_n$  an estimator of  $\Theta$ .

Indeed, we establish in Theorem 2.2 that the absorption probability is solution of an integral equation without using the specific kinetic of the model but only the fact that the absorption occurs at jumps. This solution depends on the transition kernel of the embedded chain of the jump process, and thus on the quantity  $\Theta$  which governs the dynamics. As a consequence, one may estimate the absorption probability by plugging the estimate  $\widehat{\Theta}_n$  in the expression of this solution. Nevertheless, the transfer of the asymptotic properties of  $\widehat{\Theta}_n$  to the estimator of the absorption features strongly depends on the specific motion of the chosen model.

### 3.2. Estimators and Convergence Results.

*Estimator for the Transition Density of the Post-Jump Locations.* We focus on a procedure for estimating the Markov kernel  $R$  of the post-jump locations of this process. The explicit formula (2.1) leads us to propose to estimate the transition density  $R(x, y)$  by the plug-in estimator

$$\widehat{R}_n(x, y) = \begin{cases} \frac{1}{x} \widehat{G}_n\left(\frac{y}{x}\right) & \text{if } x \leq 1, \\ \frac{\widehat{\lambda}_n}{r} (x-1)^{\frac{\widehat{\lambda}_n}{r}} \int_0^{\frac{y}{x} \wedge 1} \widehat{G}_n(u) u^{\frac{\widehat{\lambda}_n}{r}} (y-u)^{-\frac{\widehat{\lambda}_n}{r}-1} du & \text{else,} \end{cases} \quad (3.2)$$

where  $\widehat{G}_n(\xi)$  and  $\widehat{\lambda}_n$  estimate the quantities  $G(\xi)$  and  $\lambda$  from the observation of the  $n$  first loss events. We establish that the distance between  $R$  and its estimate  $\widehat{R}_n$  is directly related to the estimation error of  $\widehat{G}_n$  and  $\widehat{\lambda}_n$ .

**Proposition 3.3.** *Under Assumption  $(C_1^\lambda)$ , the following inequality holds almost surely*

$$\sup_{(x,y) \in [1,\infty) \times [0,\infty)} |R(x, y) - \widehat{R}_n(x, y)| \leq \|G - \widehat{G}_n\|_\infty + c_* \|\widehat{G}_n\|_\infty |\lambda - \widehat{\lambda}_n|$$

where  $c_* = \frac{1}{\lambda_*} \left( 4e^{-1} \frac{\lambda_*^*}{\lambda_*} + 1 \right)$ .

This result has the following corollary.

**Corollary 3.4.** *Under Assumptions  $(C_{1,2}^\lambda)$  and  $(C_1^G)$ , the estimator  $\widehat{R}_n$  converges towards  $R$  in probability uniformly in  $(x, y) \in [1, \infty) \times [0, \infty)$ ,*

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow \infty} \mathbf{P} \left( \sup_{(x,y) \in [1,\infty) \times [0,\infty)} \left| R(x, y) - \widehat{R}_n(x, y) \right| \geq \varepsilon \right) = 0.$$

In addition, the rate of convergence may be obtained from (5.2).

*Estimator for the Absorption Probability.* The previous results on the estimation of the transition density  $R$  allow us to estimate the absorption probability of the PDMP  $(X_t)_{t \geq 0}$ . We have shown in Theorem 2.2 that the absorption probability  $p(x)$  starting from  $x$  is a solution of an integral equation. Thus, by virtue of (2.5), we propose to estimate  $p(x)$  by the unique solution of the estimated integral equation

$$\widehat{p}_n(x) = \int_0^1 \widehat{R}_n(x, y) dy + \int_1^{+\infty} \widehat{p}_n(y) \widehat{R}_n(x, y) dy, \quad (3.3)$$

which satisfies both conditions

$$\lim_{x \searrow 1} \widehat{p}_n(x) = 1 \quad \text{and} \quad \lim_{x \rightarrow \infty} \widehat{p}_n(x) = 0.$$

Nevertheless, the equation above is not in a proper form to obtain a closed-form expression for  $\widehat{p}_n$ . As a consequence, we propose to solve numerically this estimated equation. In the spirit of Proposition 2.3, on the space  $L^1(1, \infty)$ , endowed with its usual norm denoted by  $\|\cdot\|$ , we define the operator

$$\widehat{K}_n : h \mapsto \int_1^{+\infty} h(y) \widehat{R}_n(x, y) dy, \quad (3.4)$$

and we introduce the following additional notation,

$$\widehat{s}_n(\cdot) = \int_0^1 \widehat{R}_n(\cdot, y) dy. \quad (3.5)$$

Thus, Equation (3.3) can be rewritten as a Fredholm equation of the second kind on the space  $L^1(1, \infty)$ , as we precisely did in the proof of Proposition 2.3,

$$\widehat{p}_n - \widehat{K}_n \widehat{p}_n = \widehat{s}_n. \quad (3.6)$$

It is well known, one may approximate a solution of (3.6) by the quantity

$$\widehat{p}_{n,m} = \sum_{k=0}^m \widehat{K}_n^k \widehat{s}_n, \quad (3.7)$$

as long as  $\|\widehat{K}_n\| < 1$ , this condition being ensured by Assumption 2.1, as stated in the following theorem of convergence of  $\widehat{p}_{n,m}$  towards  $p$ .

**Theorem 3.5.** *Under the conditions  $(C_{1,2}^\lambda)$  and  $(C_2^G)$ , and Assumption 2.1, Equation (2.5) has a unique solution and moreover,  $\|\widehat{p}_{n,m} - p\|$  tends to 0 in probability when  $n$  and  $m$  go to infinity.*

*Estimator for the Distribution of the Hitting Time.* We now proceed to the estimation of  $t_m(x)$ , the probability for the process  $(X_t)_{t \geq 0}$  starting from  $X_0 = x$  to be absorbed at jump  $m$ . Following the same approach as in Subsection 3.2 and in accordance with Theorem 2.5, we propose to estimate the functional sequence  $(t_m)_{m \geq 1}$  by the recursive procedure,

$$\hat{t}_{1,n}(x) = \int_0^1 \hat{R}_n(x, y) dy,$$

and for  $m \geq 2$ ,

$$\hat{t}_{m,n}(x) = \int_1^\infty \hat{t}_{m-1,n}(y) \hat{R}_n(x, y) dy.$$

Using the operator  $\hat{K}_n$  defined in (3.4) and the notation (3.5), this recursion relation is closed to give

$$\hat{t}_{m,n} = \hat{K}_n^{m-1} \hat{s}_n. \quad (3.8)$$

**Theorem 3.6.** *Under the conditions  $(C_{1,2}^\lambda)$  and  $(C_2^G)$ , and Assumption 2.1, for any integer  $m$ ,  $\|\hat{t}_{m,n} - t_m\|$  tends to 0 in probability when  $n$  goes to infinity.*

We give the relation between the functional sequence  $(\hat{t}_{m,n})_{m \geq 1}$  and the estimate  $\hat{p}_{n,m}$  of the absorption probability in the following remark.

*Remark 3.7.* The estimation procedures for  $p$  and  $(t_m)_{m \geq 1}$  may be carried out at the same time. In light of (3.5), (3.7) and (3.8), we have

$$\hat{p}_{n,m} = \hat{s}_n + \sum_{k=1}^m \hat{t}_{k,n}.$$

As a consequence, the estimation of the absorption probability  $p$  from the estimated sequence  $(\hat{t}_{m,n})_{m \geq 1}$  does not require extra computations.

**3.3. Numerical Illustration.** This part of the chapter is dedicated to some numerical illustrations of our main convergence results stated in the previous section. All the simulations have been implemented in the **R** language, which is commonly used in the statistical community, with an extensive use of the **integrate** function (numerical integration routine with adaptive quadrature of functions). As an example in our simulations, we choose for the probability density function  $G$  the following power function,  $G(u) = 11u^{10}$  for any  $u \in [0, 1]$ . This density function charges the interval  $[0.8, 1]$  at more than 90%. This means that the process is weakly affected by a fragmentation event. For the jump rate we choose  $\lambda = 1$  and for the growth rate  $r = 1$ . Then,

$$\frac{\lambda}{\lambda + r} \int_0^1 G(u) u^{-1} du = 0.55 < 1,$$

so that we are in the scope of application of Theorems 3.5 and 3.6. We propose to illustrate the theoretical Theorems 3.5 and 3.6 from the observation of different numbers of data ( $n = 50, 75$  and  $100$ ). In addition, we always present the distribution of our estimates from a fixed number of data over 100 replicates of the numerical experiment.

For these simulation experiments, we choose to estimate the density  $G(x)$  by the Parzen-Rosenblatt estimator  $\widehat{G}_n^{PR}(x)$  defined by

$$\widehat{G}_n^{PR}(x) = \frac{1}{nh_n} \sum_{i=1}^n \mathbb{K} \left( \frac{Y_i - x}{h_n} \right),$$

where  $\mathbb{K}$  is the Gaussian kernel and the parameter  $h_n$  is the bandwidth. The estimator is computed from the R function `density` with an optimal choice of the bandwidth parameter. In addition,  $\lambda$  is estimated from the observations  $S_i$ 's by the truncated maximum likelihood estimator  $\widehat{\lambda}_n^{ml}$  defined in (3.1). These estimates satisfy the conditions imposed in the chapter.

First, we present some simulation results for the transition kernel  $R(x, y)$  (see Figures 3.2, 3.3 and 3.4). The transition kernel is not really a quantity of interest in the model, contrary to the rate and measure of jumps  $\lambda$  and  $G(u)du$ . Nevertheless, the kernel appears when we want to compute the probability of hitting  $\Gamma$ , or the hitting time of  $\Gamma$ . This is therefore required to be able to estimate  $R(x, y)$  in our approach. Recall the definition (3.2) of the estimator of  $R$  from  $\widehat{\lambda}_n$  and  $\widehat{G}_n^{PR}$ . In Figure 3.2 are displayed the trajectory of  $R(x, 2)$  and its estimates for  $1 \leq x \leq 4$  from  $n = 50, 75$  and  $100$  data as well as the pointwise error with boxplots over 100 replications between  $R(\cdot, 2)$  and  $\widehat{R}_{100}(\cdot, 2)$  within the interval  $[1, 4]$ . Figure 3.3 presents the same numerical results for the estimation of  $R(2, y)$ ,  $1 \leq y \leq 4$ . Notice that, according to Figures 3.2 and 3.3, the pointwise error in the estimation of  $R(\cdot, 2)$  and  $R(2, \cdot)$  is maximum around 2. This may be explained by the presence of a singularity at 2 for both functions. The corresponding integrated square errors are given in Figure 3.4. In both cases, we observe a decrease in the error when the number of data grows. However, this is not very surprising here since the transition kernel is estimated from its exact expression (see Proposition 2.1), substituting  $\lambda$  by  $\widehat{\lambda}_n$  and  $G$  by  $\widehat{G}_n^{PR}$ .

Now, we proceed to the simulation of the estimation of  $p(x)$ , the probability for the process  $(X_t)_{t \geq 0}$  to be absorbed by  $\Gamma = [0, 1]$  starting from  $x > 1$ . This is, with the time of absorption, one of the two main quantities of interest in the model. Indeed, for the ruin model of Section 1,  $p(x)$  corresponds to the probability to be ruined starting from some capital  $x$ . Nevertheless, we cannot compute directly the function of interest  $p$ . As a consequence, we propose to compare  $\widehat{p}_{n,m}$  and the numerical approximation  $p_m = \sum_{k=0}^m K^k s$  of  $p$ , where the operator  $K$  is defined in (2.6) and  $s(\cdot) = \int_0^1 R(\cdot, y) dy$ . Roughly speaking,  $K$  and  $s$  are the deterministic limits of the estimates  $\widehat{K}_n$  and  $\widehat{s}_n$  presented in (3.4) and (3.5). The error in  $L^1(1, \infty)$ -norm between  $p$  and  $p_m$  satisfies

$$\|p - p_m\| \leq \|s\| \frac{\|K\|^{m+1}}{1 - \|K\|}.$$

Together with the chosen numerical values and  $m = 10$ , we have  $\|p - p_m\| \leq 1.6 \times 10^{-4}$ . Consequently, the numerical error due to the approximation of  $p$  is very low and does not affect the comparison results presented in the sequel.

Recall that our approximation of  $p$  is given by  $\widehat{p}_{n,m} = \sum_{k=0}^m \widehat{K}_n^k \widehat{s}_n$ . In the simulations, we compare  $\widehat{p}_{n,m}$  with  $p_m$  for  $m = 10$  and  $n = 50, 70$  and  $100$  data. Figure 3.5 displays the shape of  $p_m$  and  $\widehat{p}_{n,m}$  as well as the boxplots of the punctual error between the curves  $p_m$  and  $p_{100,m}$ . The corresponding integrated square error is presented in Figure 3.6. A decrease in the error is observed when  $n$  grows. Note

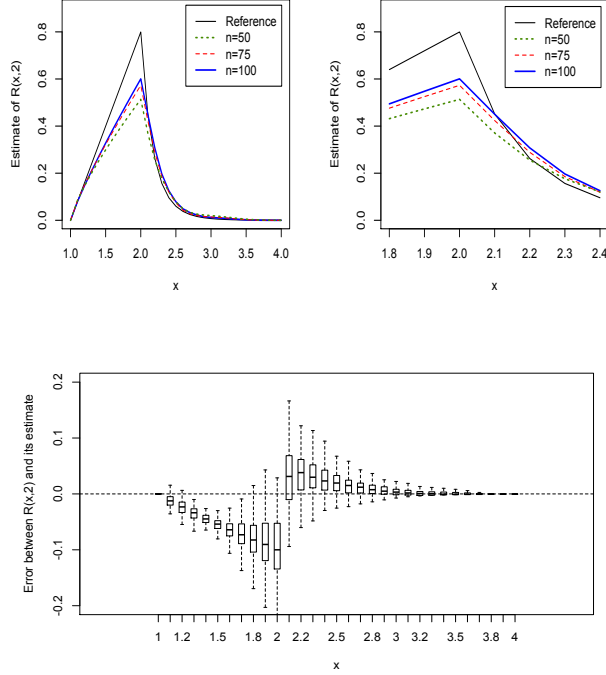


FIGURE 3.2. The figure displays the reference curve  $R(\cdot, 2)$  and its estimates from the observation of  $n = 50, 75$  or  $100$  random loss events (top left) with a zoom around  $R(2, 2)$  (top right), and the pointwise error on 100 replicates between  $R(\cdot, 2)$  and  $\hat{R}_{100}(\cdot, 2)$  (bottom).

that the error is already small for  $n = 50$  and seems to behave quite well despite the successive applications of the kernel  $\hat{K}_n$ .

Finally, we go on with the estimation of  $t_m(x)$ , the probability for the process  $(X_t)_{t \geq 0}$  starting from  $x$ , to be absorbed at jump  $m$ . The quantity  $t_m(x)$  is an important feature of the model and provides additional information to that given by  $p(x)$ . Remark that according to Proposition 2.5,  $t_m$  may be computed in an exact way contrary to  $p(x)$ . There is therefore no numerical error in this case (if we do not consider the numerical errors introduced by the computation of the kernel integrals). Thus, we compare directly  $t_m(x)$  with its estimator  $\hat{t}_{m,n}(x)$  given by Equation (3.8). At first, we notice that the estimation of the probability of absorption  $\hat{p}_{n,m}$  and the estimation of the times at which an absorption occurs  $\hat{t}_{m,n}(x)$  are related through the following formula:

$$\hat{p}_{n,m} = \sum_{k=0}^m \hat{t}_{k+1,n}.$$

Therefore, in our previous computations of  $\hat{p}_{n,m}$ , we already have computed the quantities  $\hat{t}_{m,n}$  and no further computations are required. In Figure 3.7, we present



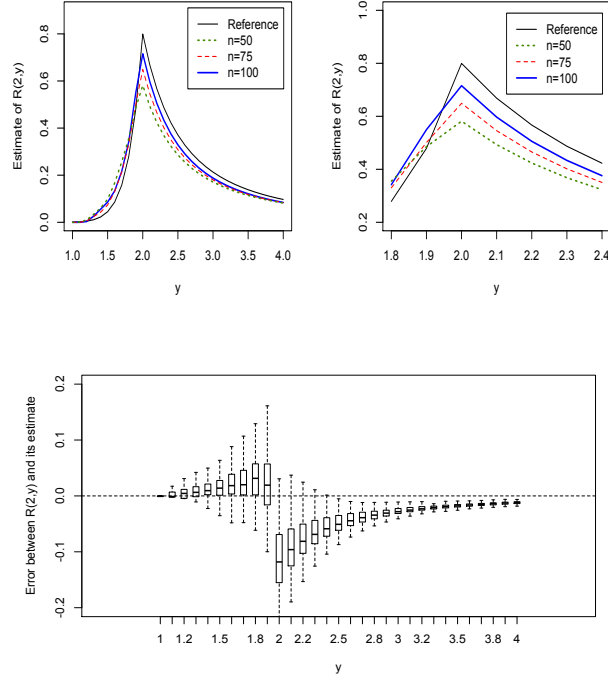


FIGURE 3.3. The figure displays the reference curve  $R(2, \cdot)$  and its estimates from the observation of  $n = 50, 75$  or  $100$  random loss events (top left) with a zoom around  $R(2, 2)$  (top right), and the pointwise error on 100 replicates between  $R(2, \cdot)$  and  $\hat{R}_{100}(2, \cdot)$  (bottom).

the integrated square error between  $t_m$  and its estimate  $\hat{t}_{m,n}$  from the observation of  $n = 50, 75$  or  $100$  random loss events and for  $m = 1, 2, 3$  and  $4$ , that is for the four first absorption times. There is a decrease of the error when  $n$  grows for each value of  $m$ . Quantitatively, this does not make sense to compare the error for  $m = 2$  and  $m = 4$  since, as displayed in Figure 3.8, the order of magnitude of the estimated probabilities are really different. Figure 3.8 presents the distribution of the hitting time of  $\Gamma$ ,  $t_m(x)$ , for  $x = 1.1$  and  $m = 1, \dots, 6$ , and also the distribution of its estimates  $\hat{t}_{m,n}$  from the observation of  $n = 50, 75$  or  $100$  random loss events. More precisely, in this figure is represented the mean of the estimators together with the first and third quartiles, over 100 replications. Once again, a decrease in the error is observed when  $n$  grows showing that the law of the hitting times of  $\Gamma$  is well estimated. These results, coupled with the estimate of  $p(x)$ , give all the interesting information in the study of this model. In all the procedure, the estimates are of high quality despite the low number of data used, in particular with respect to the sample sizes used in [1, 2, 7, 8].

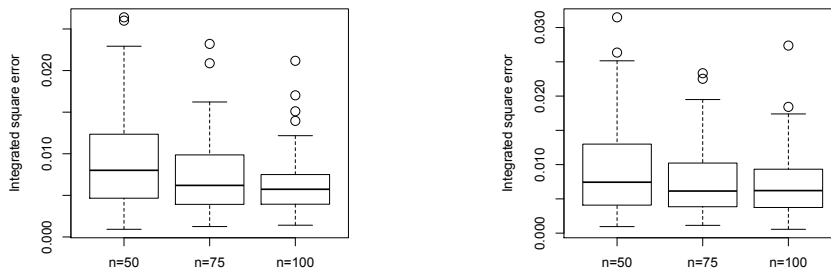


FIGURE 3.4. The figure displays the integrated square error on 100 replicates between  $R(\cdot, 2)$  and its estimate (left), and between  $R(2, \cdot)$  and its estimate (right), from the observation of  $n = 50, 75$  or 100 random loss events.

#### 4. INFERENCE FOR THE AVERAGE NUMBER OF CROSSINGS

We present in this section an innovative procedure to estimate the average number of continuous crossings of a piecewise-deterministic process based on the results of Subsection 2.3. A description and analysis of this procedure in various situations (from simulations as well as from real data) is also proposed in [4]. We begin here with the presentation of the estimators and then test their performance thanks to the use of simulated data.

**4.1. Estimation Procedures.** We observe a dataset  $\mathcal{T} = \{\mathcal{T}_i : 1 \leq i \leq n\}$  made of  $n$  discrete trajectories within the time window  $[0, T]$ . Each trajectory  $\mathcal{T}_i$  is itself composed of  $n_T$  points in  $\mathbf{R}$ ,  $\mathcal{T}_i = \{z_i^j : 1 \leq j \leq n_T\}$  given on the regular temporal grid

$$\left\{ \frac{T(j-1)}{n_T-1} : 1 \leq j \leq n_T \right\},$$

which is common to all the trajectories. We assume that the  $n$  trajectories come from the same underlying PDMP  $X$  and are independent. Our goal is to estimate the average number of continuous crossings  $C_x(T)$  of the level  $x$  (with  $x > 1$ ) by the process  $X$  from the dataset  $\mathcal{T}$ .

In the sequel, we introduce two estimation procedures of the average number of crossings  $C_x(T)$ . The Monte Carlo method only returns the empirical mean of the number of observed crossings. This technique does not exploit the characteristics of the underlying model. The Kac-Rice estimator consists in plugging an estimator of the distribution of  $X$  in the Kac-Rice formula (2.10), taking into account the knowledge of the velocity  $r$ .

*Monte Carlo method.* The most naive estimator of the average number of crossings is the Monte Carlo method. For each trajectory  $\mathcal{T}_i$ ,  $1 \leq i \leq n$ , we count the number of crossings of the given level  $x$ . We then average over the  $n$  trajectories and denote

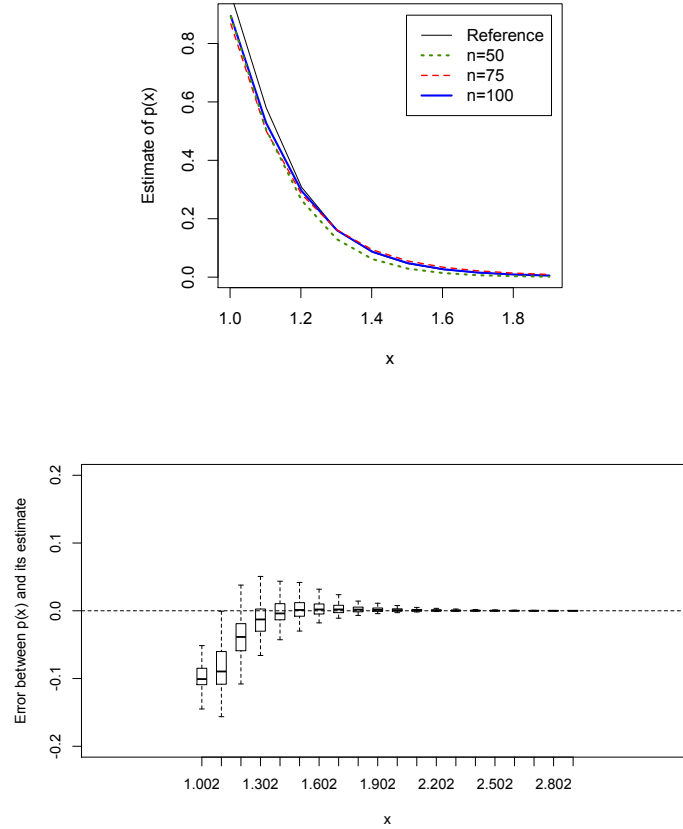


FIGURE 3.5. The figure displays the absorption probability  $p$  (approximated by  $p_m$ ) and its estimates  $\hat{p}_{n,m}$  from the observation of  $n = 50, 75$  or  $100$  random loss events and for  $m = 10$  iterations of the estimated kernel  $\hat{K}_n$  (top), and the associated pointwise error on 100 replicates from  $n = 100$  random loss events (bottom).

the result by  $\check{C}_x(T)$ . The latter quantity reads

$$\check{C}_x(T) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n_T-1} \mathbb{1}_{(-\infty, 0)}((z_i^j - x)(z_{i+1}^j - x)).$$

As mentioned in the introduction, this estimator has a natural inclination to be biased. Indeed,  $\check{C}_S(T)$  might miss one or several continuous crossings between two consecutive temporal steps given the discrete nature of the temporal grid, especially when the size of the time step is large.

*Kac-Rice method.* This technique consists in replacing the distribution of  $X$  appearing in (2.10) by some non-parametric estimator. More precisely, we estimate the density  $p_{X_{h_j}}$  at time  $h_j = T(j-1)/(n_H-1)$  by kernel smoothing methods

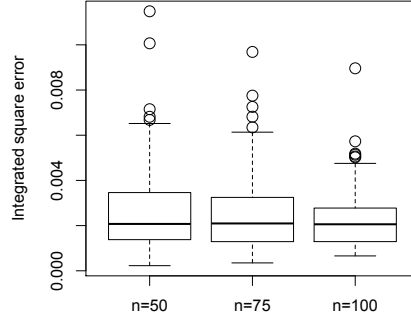


FIGURE 3.6. The figure displays the integrated square error on 100 replicates between the absorption probability  $p$  (approximated by  $p_m$ ) and its estimates  $\hat{p}_{n,m}$  from the observation of  $n = 50, 75$  or  $100$  random loss events and for  $m = 10$  iterations of the estimated kernel  $\hat{K}_n$ .

using the whole dataset, i.e., all the observed values  $z_i^j$  of  $X_{h_j}$ ,

$$\forall x \in S, \hat{p}_{X_{h_j}}(x) = \frac{1}{n\sqrt{\det(B_n)}} \sum_{i=1}^n \mathbb{K}\left(B_n^{-1/2}(x - z_i^j)\right),$$

with  $\mathbb{K}$  being some kernel function and  $B_n$  the bandwidth, i.e., a positive real number. We then compute the plug-in estimator (see Theorem 2.7), with  $x > 1$  :

$$\hat{C}_x(T) = r(x-1) \frac{T}{n_T - 1} \sum_{j=1}^{n_T} \hat{p}_{X_{h_j}}(x). \quad (4.1)$$

*Remark 4.1.* The theoretical consistency of the non-stationary Kac-Rice estimator may be established from the asymptotic properties of the kernel estimator of the distribution  $p_{X_t}$  of  $X$  at time  $t$  from i.i.d. observed values. In this chapter, we do not focus on this question. Nevertheless, we would like to emphasize that the convergence of the kernel estimator  $\hat{p}_{X_t}$  can be directly reported on the rectangle approximation of the estimated Kac-Rice formula (4.1), which only consists in a linear combination of the latter estimators.

**4.2. Numerical Application.** In the sequel, we observe  $n = 100$  trajectories of the process within the time window  $[0, 10]$  and we consider two different time step sizes:  $h = 0.01$ , i.e.,  $nH = 1001$ ,  $h = 0.1$ , i.e.,  $nH = 101$  and  $h = 0.2$ , i.e.,  $nH = 51$ . Thinner the time step is, better the kernel estimator of the stationary measure should behave (a large number of data is available) and smaller the bias of the Monte Carlo estimator should be (few crossings are missed). We denote by MC and KR the Monte-Carlo and Kac-Rice estimators defined in Subsection 4.1.

In Figure 4.9, for the threshold 1.1 and 1.15, we observe as expected the increase of the bias of the MC estimator when the time step grows. Moreover, for these two thresholds, the behavior of the KR estimator remains good when the time steps

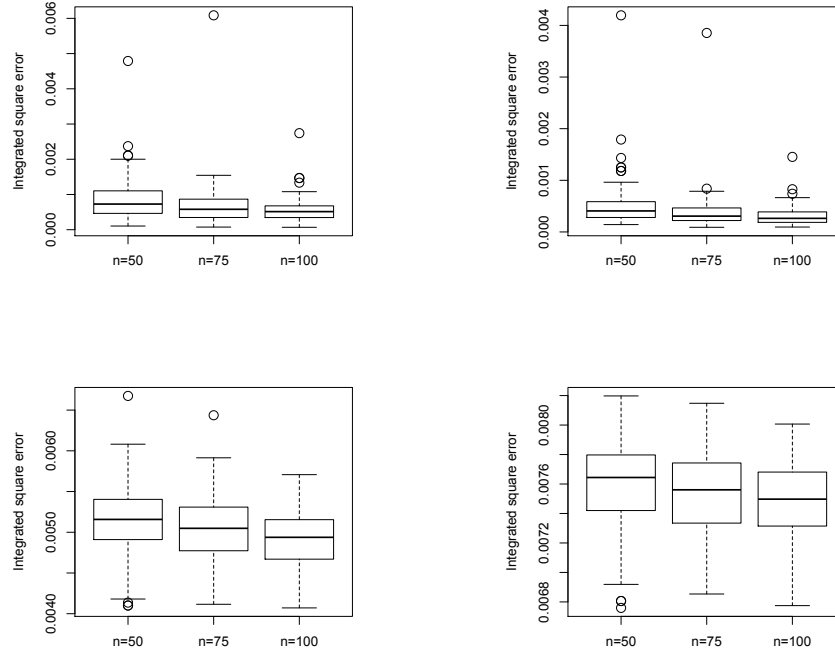


FIGURE 3.7. The figure displays the integrated square error on 100 replicates between  $t_m$  and its estimate  $\hat{t}_{m,n}$  from the observation of  $n = 50, 75$  or  $100$  random loss events and for  $m = 1$  (top left),  $m = 2$  (top right),  $m = 3$  (bottom left) and  $m = 4$  (bottom right).

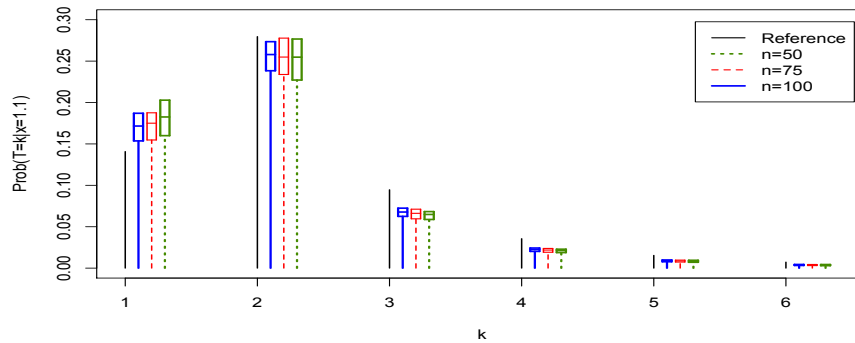


FIGURE 3.8. The figure displays the distribution of the hitting time of  $\Gamma t_m(x)$  for  $x = 1.1$  and  $m = 1, \dots, 6$  and its estimates  $\hat{t}_{m,n}$  from the observation of  $n = 50, 75$  or  $100$  random loss events.

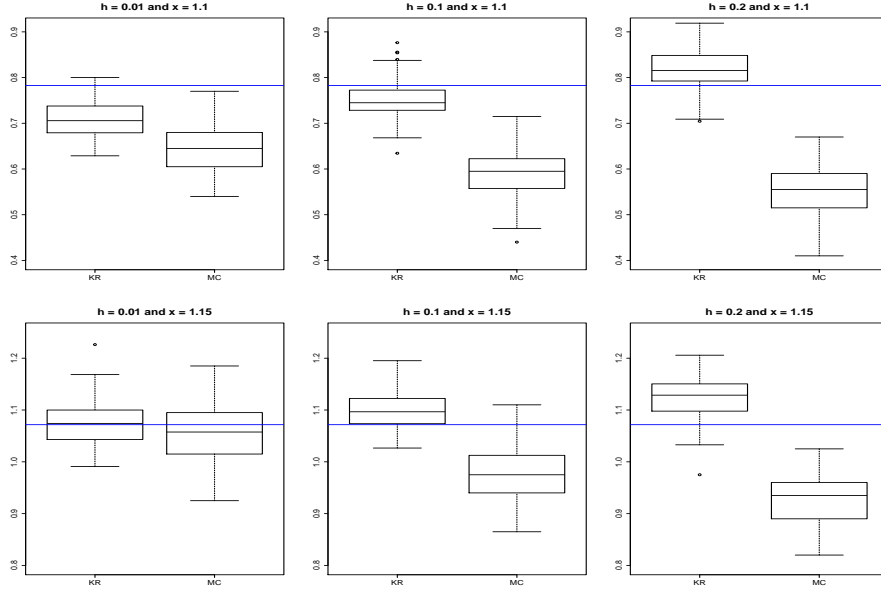


FIGURE 4.9. Boxplots of the average number of continuous crossings for 100 replications. For each figure, the MC estimator is at the right of the KR estimator. Top: the level is  $x = 1.1$  and the time step is increasing from left to right  $h = 0.01, 0.1, 0.2$ . Bottom: the level is  $x = 1.15$  and the time step is increasing from left to right  $h = 0.01, 0.1, 0.2$ .

grow. The KR estimator always gives, for these two thresholds, a more accurate result than the MC estimator.

In Figure 4.10, for the threshold 1, even if the KR estimator gives better results than the MC estimator, the mean value and entire boxplots are quite far from the real value (in blue) of the average number of continuous crossings. The KR estimator does not seem to be affected by the increase of the time step. The latter remark also applies for the average number of continuous crossings of the threshold 1.2 presented in Figure 4.10. But for this threshold, the MC estimator is more accurate than the KR estimator.

## 5. SOME ADDITIONAL PROOFS

**5.1. Technical Lemmas.** This part is dedicated to the presentation of some technical results which will be useful in the proofs of the results presented in Sections 2 and 3. For convenience, we use in the sequel the following notation. For  $\lambda > 0$ ,  $x \geq 1$ ,  $y \geq 0$  and  $u \in [0, 1]$ , we define

$$f_\lambda(x, y) = \alpha_\lambda(x) \int_0^{y/x \wedge 1} \beta_\lambda(y, u) du, \quad (5.1)$$

where  $\alpha_\lambda$  and  $\beta_\lambda$  are defined in (2.2).

**Lemma 5.1.** *The sequence  $(Y_n)_{n \geq 1}$  has  $G$  as common distribution, is independent and independent of the sequence of interarrival times  $(S_n)_{n \geq 1}$ .*

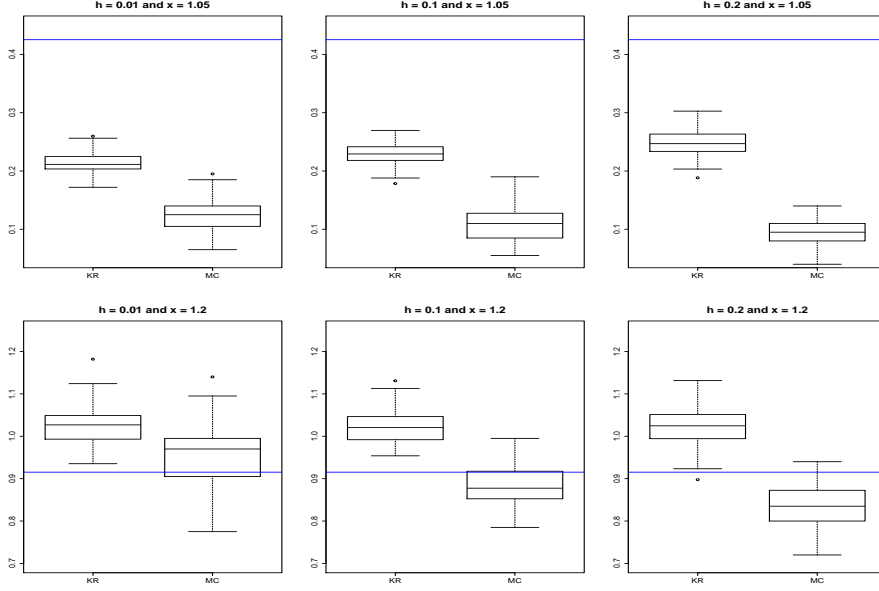


FIGURE 4.10. Boxplots of the average number of continuous crossings for 100 replications. For each figure, the MC estimator is at the right of the KR estimator. Top: the level is  $x = 1.05$  and the time step is increasing from left to right  $h = 0.01, 0.1, 0.2$ . Bottom: the level is  $x = 1.2$  and the time step is increasing from left to right  $h = 0.01, 0.1, 0.2$ .

*Proof of Lemma 5.1.* For any integer  $n$ , the  $\sigma$ -algebra  $\sigma(X_0, Y_1, \dots, Y_{n-1}, S_1, \dots, S_n)$  is denoted  $\mathcal{F}_{n-1}$ . First, let us notice that the post-jump location  $Z_{n-1}$  is  $\mathcal{F}_{n-1}$ -measurable. By the expression of the transition kernel  $R$  given in (2.1), for any measurable function  $\varphi$ , we have

$$\begin{aligned}
 & \mathbf{E} [\varphi(Y_n) | \mathcal{F}_{n-1}] \\
 &= \mathbf{E} \left[ \varphi \left( \frac{Z_n}{\Phi(Z_{n-1}, S_n)} \right) \middle| \mathcal{F}_{n-1} \right] \\
 &= \int \varphi \left( \frac{\zeta}{\Phi(Z_{n-1}, S_n)} \right) \Phi(Z_{n-1}, S_n)^{-1} G \left( \frac{\zeta}{\Phi(Z_{n-1}, S_n)} \right) d\zeta \\
 &= \int \varphi(y) G(y) dy,
 \end{aligned}$$

by the change of variables  $\zeta = y \Phi(Z_{n-1}, S_n)$ . This yields the expected result.  $\square$

**Lemma 5.2.** For any  $\lambda, \lambda_1, \lambda_2 \in [\lambda_*, \lambda^*]$ ,  $x \geq 1$ ,  $y \geq 0$  and  $u \in (0, y/x)$ , we have

$$\begin{aligned}
(i) \quad & f_\lambda(x, y) \leq \\
& \frac{1}{\lambda} \mathbb{1}_{\{y < x\}} \left[ 1 - \left( \frac{x-1}{x} \right)^{\lambda/r} \right] + \frac{1}{\lambda} \mathbb{1}_{\{y \geq x\}} \left[ \left( \frac{x-1}{y-1} \right)^{\lambda/r} - \left( \frac{x-1}{y} \right)^{\lambda/r} \right] \leq \frac{1}{\lambda}, \\
(ii) \quad & \left| \alpha_{\lambda_1}(x) \beta_{\lambda_1}(y, u) - \alpha_{\lambda_2}(x) \beta_{\lambda_2}(y, u) \right| \leq \\
& \frac{1}{r^2} \frac{1}{y-u} \left( \frac{u(x-1)}{y-u} \right)^{\lambda_*/r} \log \left( \frac{u(x-1)}{y-u} \right) |\lambda_1 - \lambda_2|, \\
(iii) \quad & \sup_{x \geq 1, y \geq 0} \int_0^{y/x \wedge 1} \left| \alpha_{\lambda_1}(x) \beta_{\lambda_1}(y, u) - \alpha_{\lambda_2}(x) \beta_{\lambda_2}(y, u) \right| du \leq \frac{4e^{-1}}{\lambda_*^2} |\lambda_1 - \lambda_2|.
\end{aligned}$$

*Proof of Lemma 5.2.* We begin with the inequality (i). Let  $(x, y)$  be in  $[1, \infty) \times [0, \infty)$ . For any  $u \in [0, y/x \wedge 1]$  we have

$$0 \leq u^{\lambda/r} (y-u)^{-\lambda/r-1} \leq (y/x \wedge 1)^{\lambda/r} (y-u)^{-\lambda/r-1}.$$

Therefore,

$$f_\lambda(x, y) \leq \frac{1}{r} (y/x \wedge 1)^{\lambda/r} (x-1)^{\lambda/r} \int_0^{y/x \wedge 1} (y-u)^{-\lambda/r-1} du.$$

Computing the integral leads to

$$f_\lambda(x, y) \leq \frac{1}{\lambda} (y/x \wedge 1)^{\lambda/r} (x-1)^{\lambda/r} \left[ (y - y/x \wedge 1)^{-\lambda/r} - y^{-\lambda/r} \right].$$

We now split the latter term in two using the elementary fact that  $1 = \mathbb{1}_{\{y < x\}} + \mathbb{1}_{\{y \geq x\}}$ . This yields

$$\begin{aligned}
f_\lambda(x, y) & \leq \frac{1}{\lambda} (y/x)^{\lambda/r} (x-1)^{\lambda/r} \mathbb{1}_{\{y < x\}} \left[ (y - y/x)^{-\lambda/r} - y^{-\lambda/r} \right] \\
& \quad + \frac{1}{\lambda} (x-1)^{\lambda/r} \mathbb{1}_{\{y \geq x\}} \left[ (y-1)^{-\lambda/r} - y^{-\lambda/r} \right] \\
& = \frac{1}{\lambda} \mathbb{1}_{\{y < x\}} \left[ 1 - \left( \frac{x-1}{x} \right)^{\lambda/r} \right] \\
& \quad + \frac{1}{\lambda} \mathbb{1}_{\{y \geq x\}} \left[ \left( \frac{x-1}{y-1} \right)^{\lambda/r} - \left( \frac{x-1}{y} \right)^{\lambda/r} \right].
\end{aligned}$$

Using the fact that when  $y \geq x$ ,  $\frac{x-1}{y-1} \leq 1$  and noticing that the two terms  $\frac{x-1}{y}$  and  $\frac{x-1}{x}$  are non negative, we obtain

$$f_\lambda(x, y) \leq \frac{1}{\lambda} \mathbb{1}_{\{y < x\}} + \frac{1}{\lambda} \mathbb{1}_{\{y \geq x\}} = \frac{1}{\lambda}.$$

We go on with the second part of the lemma. One may differentiate with respect to  $\lambda$  to obtain

$$\partial_\lambda \alpha_\lambda(x) \beta_\lambda(y, u) = \frac{1}{r^2} \frac{1}{y-u} \left( \frac{u(x-1)}{y-u} \right)^{\lambda/r} \log \left( \frac{u(x-1)}{y-u} \right).$$



Then, we use that for  $x \geq 1$ ,  $y \geq 0$  and  $u \in [0, y/x \wedge 1]$ , one has  $\Xi = \frac{u(x-1)}{y-u} \in (0, 1]$  such that  $\Xi^{\lambda/2r} |\log \Xi| \leq \frac{2r}{\lambda} e^{-1}$ . This yields

$$|\partial_\lambda \alpha_\lambda(x) \beta_\lambda(y, u)| \leq \frac{2e^{-1}}{r\lambda} \frac{1}{y-u} \left( \frac{u(x-1)}{y-u} \right)^{\frac{\lambda}{2r}} \leq \frac{2e^{-1}}{r\lambda_*} \frac{1}{y-u} \left( \frac{u(x-1)}{y-u} \right)^{\frac{\lambda_*}{2r}}.$$

Notice that the last inequality is uniform in  $\lambda \in [\lambda_*, \lambda^*]$ . This proves the second assertion (ii). For the third one, using the mean value theorem and similar computations as above, we obtain that

$$\begin{aligned} & \int_0^{y/x \wedge 1} \left| \alpha_{\lambda_1}(x) \beta_{\lambda_1}(y, u) - \alpha_{\lambda_2}(x) \beta_{\lambda_2}(y, u) \right| du \\ & \leq \frac{4e^{-1}}{\lambda_*^2} \mathbb{1}_{\{y < x\}} \left[ 1 - \left( \frac{x-1}{x} \right)^{\frac{\lambda_*}{2r}} \right] |\lambda_1 - \lambda_2| \\ & \quad + \frac{4e^{-1}}{\lambda_*^2} \mathbb{1}_{\{y \geq x\}} \left[ \left( \frac{x-1}{y-1} \right)^{\frac{\lambda_*}{2r}} - \left( \frac{x-1}{y} \right)^{\frac{\lambda_*}{2r}} \right] |\lambda_1 - \lambda_2|. \end{aligned}$$

We obtain once again

$$\begin{aligned} & \int_0^{y/x \wedge 1} \left| \alpha_{\lambda_1}(x) \beta_{\lambda_1}(y, u) - \alpha_{\lambda_2}(x) \beta_{\lambda_2}(y, u) \right| du \\ & \leq \mathbb{1}_{\{y < x\}} \frac{4e^{-1}}{\lambda_*^2} |\lambda_1 - \lambda_2| + \mathbb{1}_{\{y \geq x\}} \frac{4e^{-1}}{\lambda_*^2} |\lambda_1 - \lambda_2| \\ & = \frac{4e^{-1}}{\lambda_*^2} |\lambda_1 - \lambda_2|. \end{aligned}$$

The result follows.  $\square$

**Lemma 5.3.** *Under Assumption  $(C_1^\lambda)$ , the following inequality holds almost surely:*

$$\begin{aligned} & \sup_{y \in [0, \infty)} \int_1^\infty |R(x, y) - \widehat{R}_n(x, y)| dx \\ & \leq \frac{\lambda}{\lambda + r} \int_0^1 \left| G(u) - \widehat{G}_n(u) \right| \frac{du}{u} + \lambda^* \left[ 4e^{-1} \frac{1}{\lambda_*^2} + \frac{1}{\lambda_* + r} \right] \int_0^1 |\widehat{G}_n(u)| \frac{du}{u} |\lambda - \widehat{\lambda}_n|. \end{aligned}$$

*Proof of Lemma 5.3.* One may write

$$\begin{aligned} R(x, y) - \widehat{R}_n(x, y) & = \lambda \alpha_\lambda(x) \int_0^{y/x \wedge 1} \beta_\lambda(y, u) \left[ G(u) - \widehat{G}_n(u) \right] du \\ & \quad + \lambda \int_0^{y/x \wedge 1} \widehat{G}_n(u) \left[ \alpha_\lambda(x) \beta_\lambda(y, u) - \alpha_{\widehat{\lambda}_n}(x) \beta_{\widehat{\lambda}_n}(y, u) \right] du \\ & \quad + (\lambda - \widehat{\lambda}_n) \alpha_{\widehat{\lambda}_n}(x) \int_0^{y/x \wedge 1} \beta_{\widehat{\lambda}_n}(y, u) \widehat{G}_n(u) du. \end{aligned}$$

Thus, for any  $y \geq 0$ , we have

$$\begin{aligned} & \int_1^\infty \left| R(x, y) - \widehat{R}_n(x, y) \right| dx \\ & \leq \lambda \int_1^\infty \alpha_\lambda(x) \int_0^{\frac{y}{x} \wedge 1} \beta_\lambda(y, u) \left| G(u) - \widehat{G}_n(u) \right| du dx \\ & \quad + \lambda \int_1^\infty \int_0^{\frac{y}{x} \wedge 1} \left| \widehat{G}_n(u) \right| \left| \alpha_\lambda(x) \beta_\lambda(y, u) - \alpha_{\widehat{\lambda}_n}(x) \beta_{\widehat{\lambda}_n}(y, u) \right| du dx \\ & \quad + \left| \lambda - \widehat{\lambda}_n \right| \int_1^\infty \alpha_{\widehat{\lambda}_n}(x) \int_0^{\frac{y}{x} \wedge 1} \beta_{\widehat{\lambda}_n}(y, u) \left| \widehat{G}_n(u) \right| du dx. \end{aligned}$$

In the right-hand terms above, one may change the order of integration to integrate with respect to  $x$  first. Let us deal with these three terms separately. For the first term, we have

$$\begin{aligned} & \lambda \int_1^\infty \alpha_\lambda(x) \int_0^{y/x \wedge 1} \beta_\lambda(y, u) \left| G(u) - \widehat{G}_n(u) \right| du dx \\ & = \lambda \int_0^1 \int_1^{y/u} \alpha_\lambda(x) dx \beta_\lambda(y, u) \left| G(u) - \widehat{G}_n(u) \right| \mathbb{1}_{\{u \leq y\}} du \\ & = \frac{\lambda}{r} \int_0^1 \frac{1}{\lambda/r + 1} \left( \frac{y}{u} - 1 \right)^{\lambda/r + 1} \beta_\lambda(y, u) \left| G(u) - \widehat{G}_n(u) \right| \mathbb{1}_{\{u \leq y\}} du \\ & = \frac{\lambda}{\lambda + r} \int_0^1 \left| G(u) - \widehat{G}_n(u) \right| \mathbb{1}_{\{u \leq y\}} \frac{du}{u}. \end{aligned}$$

For the third term, a similar computation gives

$$\begin{aligned} & \left| \lambda - \widehat{\lambda}_n \right| \int_1^\infty \alpha_{\widehat{\lambda}_n}(x) \int_0^{\frac{y}{x} \wedge 1} \beta_{\widehat{\lambda}_n}(y, u) \left| \widehat{G}_n(u) \right| du dx \\ & = \left| \lambda - \widehat{\lambda}_n \right| \frac{\widehat{\lambda}_n}{\widehat{\lambda}_n + r} \int_0^1 \left| \widehat{G}_n(u) \right| \mathbb{1}_{\{u \leq y\}} \frac{du}{u}. \end{aligned}$$

The most intricate term is the second. Using Lemma 5.2, we have

$$\begin{aligned} & \lambda \int_1^\infty \int_0^{y/x \wedge 1} \widehat{G}_n(u) \left| \alpha_\lambda(x) \beta_\lambda(y, u) - \alpha_{\widehat{\lambda}_n}(x) \beta_{\widehat{\lambda}_n}(y, u) \right| du dx \\ & \leq \lambda \left| \lambda - \widehat{\lambda}_n \right| \int_1^\infty \int_0^{y/x \wedge 1} \frac{2e^{-1}}{r\lambda_*} \frac{1}{y-u} \left( \frac{u(x-1)}{y-u} \right)^{\lambda_*/2r} \left| \widehat{G}_n(u) \right| du dx \\ & = \frac{\lambda 2e^{-1}}{r\lambda_*} \left| \lambda - \widehat{\lambda}_n \right| \int_0^1 \frac{1}{y-u} \left( \frac{u}{y-u} \right)^{\frac{\lambda_*}{2r}} \left| \widehat{G}_n(u) \right| \int_1^{\frac{y}{u}} (x-1)^{\frac{\lambda_*}{2r}} dx \mathbb{1}_{\{u \leq y\}} du \\ & = \frac{4\lambda_* e^{-1}}{\lambda_*^2} \left| \lambda - \widehat{\lambda}_n \right| \int_0^1 \left| \widehat{G}_n(u) \right| \mathbb{1}_{\{u \leq y\}} \frac{du}{u}. \end{aligned}$$

The result follows by aggregation of the three above estimates.  $\square$

**5.2. Proof of Proposition 3.3.** Let  $n \in \mathbf{N}$ ,  $\lambda \in [\lambda_*, \lambda^*]$ ,  $x \geq 1$  and  $y \geq 0$ . We work  $\omega$  by  $\omega$  so that the desired almost sure inequality will follow. Recall that, with the notation (5.1),

$$R(x, y) = \lambda \alpha_\lambda(x) \int_0^{y/x \wedge 1} \beta_\lambda(y, u) G(u) du$$

and

$$\widehat{R}_n(x, y) = \widehat{\lambda}_n \alpha_{\widehat{\lambda}_n} \int_0^{y/x^{\wedge 1}} \beta_{\widehat{\lambda}_n}(y, u) \widehat{G}_n(u) du.$$

By an elementary rearranging, one may write

$$\begin{aligned} R(x, y) - \widehat{R}_n(x, y) &= \lambda \alpha_\lambda(x) \int_0^{y/x^{\wedge 1}} \beta_\lambda(y, u) [G(u) - \widehat{G}_n(u)] du \\ &\quad + \lambda \int_0^{y/x^{\wedge 1}} \widehat{G}_n(u) [\alpha_\lambda(x) \beta_\lambda(y, u) - \alpha_{\widehat{\lambda}_n}(x) \beta_{\widehat{\lambda}_n}(y, u)] du \\ &\quad + (\lambda - \widehat{\lambda}_n) \alpha_{\widehat{\lambda}_n}(x) \int_0^{y/x^{\wedge 1}} \beta_{\widehat{\lambda}_n}(y, u) \widehat{G}_n(u) du. \end{aligned}$$

We deal with the three above terms separately. For the first term we have

$$\begin{aligned} &\left| \lambda \alpha_\lambda(x) \int_0^{y/x^{\wedge 1}} \beta_\lambda(y, u) [G(u) - \widehat{G}_n(u)] du \right| \\ &\leq \lambda \alpha_\lambda(x) \int_0^{y/x^{\wedge 1}} \beta_\lambda(y, u) du \|G - \widehat{G}_n\|_\infty. \end{aligned}$$

Thus, by the first part of Lemma 5.2, we obtain,

$$\begin{aligned} &\left| \lambda \alpha_\lambda(x) \int_0^{y/x^{\wedge 1}} \beta_\lambda(y, u) [G(u) - \widehat{G}_n(u)] du \right| \\ &\leq \lambda f_\lambda(x, y) \|G - \widehat{G}_n\|_\infty \leq \|G - \widehat{G}_n\|_\infty. \end{aligned}$$

Now for the second term, using this time the third part of Lemma 5.2,

$$\begin{aligned} &\lambda \left| \int_0^{y/x^{\wedge 1}} \widehat{G}_n(u) [\alpha_\lambda(x) \beta_\lambda(y, u) - \alpha_{\widehat{\lambda}_n}(x) \beta_{\widehat{\lambda}_n}(y, u)] du \right| \\ &\leq \lambda \|\widehat{G}_n(u)\|_\infty \int_0^{y/x^{\wedge 1}} |\alpha_\lambda(x) \beta_\lambda(y, u) - \alpha_{\widehat{\lambda}_n}(x) \beta_{\widehat{\lambda}_n}(y, u)| du \\ &\leq \lambda \|\widehat{G}_n(u)\|_\infty \frac{4e^{-1}}{\lambda_*^2} |\lambda - \widehat{\lambda}_n| \\ &\leq \frac{4\lambda^* e^{-1}}{\lambda_*^2} \|\widehat{G}_n(u)\|_\infty |\lambda - \widehat{\lambda}_n|. \end{aligned}$$

For the last term we have, using again the first part of Lemma 5.2,

$$\begin{aligned} &|\lambda - \widehat{\lambda}_n| \alpha_{\widehat{\lambda}_n}(x) \int_0^{y/x^{\wedge 1}} \beta_{\widehat{\lambda}_n}(y, u) \widehat{G}_n(u) du \\ &\leq |\lambda - \widehat{\lambda}_n| \|\widehat{G}_n\|_\infty f_{\widehat{\lambda}_n}(x, y) \leq |\lambda - \widehat{\lambda}_n| \frac{\|\widehat{G}_n\|_\infty}{\lambda_*}. \end{aligned}$$

This ends the proof.

**5.3. Proof of Corollary 3.4.** Let us introduce the notation

$$\begin{aligned} I_n &= \sup_{(x, y) \in [1, \infty) \times [0, +\infty)} |R(x, y) - \widehat{R}_n(x, y)|, \\ C &= \frac{1}{\lambda_*} \left( 4e^{-1} \frac{\lambda^*}{\lambda_*} + 1 \right). \end{aligned}$$

For any  $\varepsilon > 0$ , according to Proposition 3.3, we have

$$\begin{aligned} \mathbf{P}(I_n \geq \varepsilon) &\leq \mathbf{P}\left(\|G - \widehat{G}_n\|_\infty + C\|\widehat{G}_n\|_\infty \mid \lambda - \widehat{\lambda}_n \mid \geq \varepsilon\right) \\ &\leq \mathbf{P}\left(\|G - \widehat{G}_n\|_\infty \geq \frac{\varepsilon}{2}\right) + \mathbf{P}\left(C\|\widehat{G}_n\|_\infty \mid \lambda - \widehat{\lambda}_n \mid \geq \frac{\varepsilon}{2}\right). \end{aligned}$$

Let  $\eta > 0$ . Using the elementary inequality satisfied for any real numbers  $a$  and  $b$ ,

$$ab \leq \frac{1}{4\eta}a^2 + \eta b^2,$$

we have

$$\begin{aligned} \mathbf{P}\left(\|\widehat{G}_n\|_\infty \mid \lambda - \widehat{\lambda}_n \mid \geq \frac{\varepsilon}{2C}\right) &\leq \mathbf{P}\left(\eta\|\widehat{G}_n\|_\infty^2 + \frac{1}{4\eta} \mid \lambda - \widehat{\lambda}_n \mid^2 \geq \frac{\varepsilon}{2C}\right) \\ &\leq \mathbf{P}\left(\eta\|\widehat{G}_n\|_\infty^2 \geq \frac{\varepsilon}{4C}\right) + \mathbf{P}\left(\frac{1}{4\eta} \mid \lambda - \widehat{\lambda}_n \mid^2 \geq \frac{\varepsilon}{4C}\right). \end{aligned}$$

Notice that

$$\begin{aligned} \mathbf{P}\left(\|\widehat{G}_n\|_\infty^2 \geq \frac{\varepsilon}{4C\eta}\right) &= \mathbf{P}\left(\|\widehat{G}_n\|_\infty \geq \sqrt{\frac{\varepsilon}{4C\eta}}\right) \\ &\leq \mathbf{P}\left(\|G\|_\infty + \|\widehat{G}_n - G\|_\infty \geq \sqrt{\frac{\varepsilon}{4C\eta}}\right) \\ &\leq \mathbf{P}\left(\|G\|_\infty \geq \frac{1}{2}\sqrt{\frac{\varepsilon}{4C\eta}}\right) + \mathbf{P}\left(\|\widehat{G}_n - G\|_\infty \geq \frac{1}{2}\sqrt{\frac{\varepsilon}{4C\eta}}\right). \end{aligned}$$

With  $\eta^* = \frac{\varepsilon}{16C(\|G\|_\infty + 1)^2}$ , we have

$$\mathbf{P}\left(\|G\|_\infty \geq \frac{1}{2}\sqrt{\frac{\varepsilon}{4C\eta}}\right) = 0.$$

Thus, we obtain

$$\mathbf{P}\left(\|\widehat{G}_n\|_\infty^2 \geq \frac{\varepsilon}{4C\eta^*}\right) \leq \mathbf{P}\left(\|\widehat{G}_n - G\|_\infty \geq \frac{1}{2}\sqrt{\frac{\varepsilon}{4C\eta^*}}\right).$$

To sum up,

$$\begin{aligned} \mathbf{P}(I_n \geq \varepsilon) &\leq \mathbf{P}\left(\|\widehat{G}_n - G\|_\infty \geq \frac{\varepsilon}{2}\right) \\ &\quad + \mathbf{P}\left(\|\widehat{G}_n - G\|_\infty \geq \|G\|_\infty + 1\right) \\ &\quad + \mathbf{P}\left(\mid \lambda - \widehat{\lambda}_n \mid \geq \frac{\varepsilon}{4C(1 + \|G\|_\infty)}\right). \end{aligned} \tag{5.2}$$

The result follows.

#### 5.4. Proof of Theorem 3.5.

**Proposition 5.4.** *The estimators  $\widehat{K}_n$  and  $\widehat{s}_n$  converge in probability towards  $K$  and  $s$  respectively. For any  $\varepsilon > 0$ ,*

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\|\widehat{K}_n - K\| \geq \varepsilon\right) = 0, \quad \text{and} \quad \lim_{n \rightarrow \infty} \mathbf{P}\left(\|\widehat{s}_n - s\| \geq \varepsilon\right) = 0.$$

*Proof of Proposition 5.4.* First, let us notice that for any  $h \in L^1(1, \infty)$ , we have

$$\begin{aligned} \|(K - \widehat{K}_n)h\| &= \int_1^\infty \left| \int_1^\infty h(y)(R(x, y) - \widehat{R}_n(x, y)) dy \right| dx \\ &\leq \int_1^\infty \int_1^\infty |h(y)| |R(x, y) - \widehat{R}_n(x, y)| dy dx \\ &\leq \|h\| \sup_{y \geq 1} \int_1^\infty |R(x, y) - \widehat{R}_n(x, y)| dx. \end{aligned}$$

Therefore,

$$\|K - \widehat{K}_n\| \leq \sup_{y \geq 1} \int_1^\infty |R(x, y) - \widehat{R}_n(x, y)| dx$$

$\mathbf{P}$ -almost surely. Using Lemma 5.3,  $\|K - \widehat{K}_n\|$  is almost surely bounded by

$$\begin{aligned} &\frac{\lambda}{\lambda + r} \int_0^1 |G(u) - \widehat{G}_n(u)| u^{-1} du \\ &+ \lambda^* \left( 4e^{-1} \frac{1}{\lambda_*^2} + \frac{1}{\lambda_* + r} \right) \int_0^1 |\widehat{G}_n(u)| u^{-1} du |\lambda - \widehat{\lambda}_n|. \end{aligned}$$

Then, using Assumption 3.1,  $(C_{1,2}^\lambda)$  and  $(C_2^G)$  the convergence in probability of  $\widehat{K}_n$  towards  $K$  follows. The proof of the convergence in probability of  $\widehat{s}_n$  towards  $s$  is similar.  $\square$

We deduce easily from the above proposition that for any  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \mathbf{P} \left( \|(\widehat{K}_n - K)r\| \geq \varepsilon \right) = 0.$$

Let us choose  $\eta > 0$  and  $\varepsilon > 0$  such that  $\varepsilon < 1 - \|K\|$ . We define

$$\Omega_n = \left\{ \omega \in \Omega ; \|\widehat{K}_n(\omega)\| < 1 - \varepsilon, \|(K - \widehat{K}_n(\omega))r\| \leq \frac{\varepsilon^2}{4}, \|s - \widehat{s}_n(\omega)\| \leq \frac{\varepsilon^2}{4} \right\}.$$

According to Proposition 5.4, there exists  $N$  such that for all  $n \geq N$ ,

$$\begin{aligned} \mathbf{P}(\Omega \setminus \Omega_n) &\leq \mathbf{P} \left( \|\widehat{K}_n\| \geq 1 - \varepsilon \right) + \mathbf{P} \left( \|(K - \widehat{K}_n)r\| > \frac{\varepsilon^2}{4} \right) + \mathbf{P} \left( \|s - \widehat{s}_n\| > \frac{\varepsilon^2}{4} \right) \\ &\leq \frac{\eta}{3} + \frac{\eta}{3} + \frac{\eta}{3} = \eta. \end{aligned}$$

From Equation (3.3),  $\widehat{p}_n$  satisfies almost surely the equation :

$$\widehat{p}_n = \widehat{s}_n + \widehat{K}_n \widehat{p}_n.$$

Therefore, we also have  $\widehat{p}_n = \sum_{k=0}^\infty \widehat{K}_n^k \widehat{s}_n$ . We split the difference  $p - \widehat{p}_{n,m}$  using the quantity  $\widehat{p}_n$ ,

$$p - \widehat{p}_{n,m} = p - \widehat{p}_n + \widehat{p}_n - \widehat{p}_{n,m}.$$

We begin to bound  $p - \widehat{p}_n$  on  $\Omega_n$ . For  $n \geq N$ , on  $\Omega_n$ ,

$$\begin{aligned} \|p - \widehat{p}_n\| &\leq \|(s - \widehat{s}_n)p\| + \|(K - \widehat{K}_n)p\| + \|\widehat{K}_n\| \|p - \widehat{p}_n\| \\ &\leq \frac{\varepsilon^2}{4} + \frac{\varepsilon^2}{4} + (1 - \varepsilon) \|p - \widehat{p}_n\|. \end{aligned}$$

Arranging the terms yields  $\|p - \widehat{p}_n\| \leq \frac{\varepsilon}{2}$  on  $\Omega_n$ . It remains to consider the difference  $\widehat{p}_n - \widehat{p}_{n,m}$ . By definition, we have

$$\widehat{p}_n - \widehat{p}_{n,m} = \sum_{k=m+1}^{\infty} \widehat{K}_n^k \widehat{s}_n.$$

Therefore, for  $n \geq N$ , on  $\Omega_n$ , we have

$$\|\widehat{p}_n - \widehat{p}_{n,m}\| \leq \|\widehat{s}_n\| \frac{\|\widehat{K}_n\|^{m+1}}{1 - \|\widehat{K}_n\|} \leq \left( \frac{\varepsilon^2}{2} + \|s\| \right) \frac{(1 - \varepsilon)^{m+1}}{\varepsilon} \leq \frac{\varepsilon}{2}$$

for  $m \geq M$  with  $M$  large enough. Therefore, for  $n \geq N$  and  $m \geq M$ , on  $\Omega_n$ ,

$$\|p - \widehat{p}_{n,m}\| \leq \|p - \widehat{p}_n\| + \|\widehat{p}_n - \widehat{p}_{n,m}\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This concludes the proof.

**5.5. Proof of Theorem 3.6.** All the ingredients for this proof are in fact already present in the proof of Theorem 3.5. Nevertheless, let us give some details. As in the previous section, let us choose  $\eta > 0$  and  $\varepsilon > 0$  such that  $\varepsilon < 1 - \|K\|$ . We define

$$\Omega_n = \left\{ \omega \in \Omega ; \|\widehat{K}_n(\omega)\| < 1 - \varepsilon, \|(K - \widehat{K}_n(\omega))\| \leq \frac{\varepsilon^2}{2\|s\|}, \|s - \widehat{s}_n(\omega)\| \leq \frac{\varepsilon}{2} \right\}.$$

According to Proposition 5.4, there exists  $N$  such that, for all  $n \geq N$ ,

$$\mathbf{P}(\Omega \setminus \Omega_n) \leq \eta.$$

For  $m = 1$  we have

$$\|\widehat{t}_{1,n} - t_1\| = \|\widehat{s}_n - s\| \leq \frac{\varepsilon}{2}$$

on  $\Omega_n$  and the result follows. Now for  $m \geq 2$ , one may write

$$\widehat{t}_{m,n} - t_m = \widehat{K}_n(\widehat{t}_{m-1,n} - t_{m-1}) + (\widehat{K}_n - K)t_{m-1}.$$

Notice that on  $\Omega_n$ , for  $m \geq 1$ ,

$$\|t_m\| = \|K^{m-1}s\| \leq \|K\|^{m-1}\|s\| \leq (1 - \varepsilon)^{m-1}\|s\|.$$

Then, for  $m \geq 2$ , we write

$$\begin{aligned} \|\widehat{t}_{m,n} - t_m\| &\leq \|\widehat{K}_n(\widehat{t}_{m-1,n} - t_{m-1})\| + \|(\widehat{K}_n - K)t_{m-1}\| \\ &\leq \|\widehat{K}_n\| \|\widehat{t}_{m-1,n} - t_{m-1}\| + \|(\widehat{K}_n - K)\| \|t_{m-1}\| \\ &\leq (1 - \varepsilon) \|\widehat{t}_{m-1,n} - t_{m-1}\| + \frac{\varepsilon^2}{2\|s\|} (1 - \varepsilon)^{m-2} \|s\|. \end{aligned}$$

A straightforward recursion gives, still for  $m \geq 2$  and on  $\Omega_n$ ,

$$\|\widehat{t}_{m,n} - t_m\| \leq (1 - \varepsilon)^{m-1} \|\widehat{t}_{1,n} - t_1\| + \frac{\varepsilon^2}{2} (1 - \varepsilon)^{m-2} \sum_{k=0}^{m-2} (1 - \varepsilon)^k.$$

Therefore, for any  $m \geq 2$ , on  $\Omega_n$ , we obtain

$$\|\widehat{t}_{m,n} - t_m\| \leq (1 - \varepsilon)^{m-1} \frac{\varepsilon}{2} + \frac{\varepsilon^2}{2} (1 - \varepsilon)^{m-2} \frac{1 - (1 - \varepsilon)^{m-1}}{\varepsilon} \leq \varepsilon.$$

The result follows.

5.6. **Discussion on the Condition**  $(C_2^G)$ . Here, we propose to show that the Parzen-Rosenblatt estimator  $\widehat{G}_n^{PR}$  of the density  $G$ , defined by

$$\forall x \in [0, 1], \widehat{G}_n^{PR}(x) = \frac{1}{nh_n} \sum_{i=1}^n \mathbb{K}\left(\frac{Y_i - x}{h_n}\right),$$

where  $\mathbb{K}$  is a kernel function and the bandwidth sequence  $(h_n)_{n \geq 1}$  tends to 0 as  $n$  goes to infinity, satisfies the condition  $(C_2^G)$  given in Assumption 3.1 under the following assumption on the density of interest.

*Assumption 5.1.* We assume that there exists  $\epsilon_1 > 0$  such that, for any  $0 \leq x < \epsilon_1$ ,  $G(x) = 0$ . In addition, we suppose that  $G$  is in the Hölder class  $\Sigma(\beta, L)$  [13, Definition 1.2].

*Remark 5.5.* Notice that Assumption 5.1 does not hold for the example in Subsection 3.3 where  $G(u) = 11u^{10}$  for  $u \in [0, 1]$ . However, the numerical illustrations show that our results still apply in this case. It means that Assumption 5.1 is certainly non optimal and that some weaker assumptions on  $G$ , as being close enough to zero near zero, may be sufficient for our results to apply.

For any  $x$ , we define the mean squared error of  $\widehat{G}_n^{PR}(x)$  by

$$MSE(x) = \mathbf{E} \left[ \left( \widehat{G}_n^{PR}(x) - G(x) \right)^2 \right].$$

By [13, Equation (1.4)], we have the following bias-variance decomposition

$$MSE(x) = b^2(x) + V(x),$$

where, with [13, Equation (1.6)],

$$b(x) = \mathbf{E} \left[ \widehat{G}_n^{PR}(x) \right] - G(x) \quad \text{and} \quad V(x) = \frac{1}{nh_n^2} \mathbf{E} \left[ \mathbb{K}^2 \left( \frac{Y_1 - x}{h_n} \right) \right].$$

In the sequel, we assume that the chosen kernel function  $\mathbb{K}$  has a bounded support. As a consequence, for  $n$  large enough and some  $\epsilon_2 > 0$ ,  $\mathbb{K}\left(\frac{y-x}{h_n}\right) = 0$  for any  $x < \epsilon_2$  and  $y \geq \epsilon_1$ . Thus,

$$\begin{aligned} \int_0^1 \frac{\sqrt{V(x)}}{x} dx &= \frac{1}{\sqrt{nh_n}} \int_{\epsilon_1}^1 \frac{1}{x} \left[ \int_{\epsilon_2}^1 G(y) \mathbb{K}^2 \left( \frac{y-x}{h_n} \right) \right]^{1/2} dx dy \\ &\leq \frac{\|\mathbb{K}\|_\infty \log(\epsilon_1)}{\sqrt{nh_n}}. \end{aligned} \quad (5.3)$$

In addition,  $b(x) = 0$  for any  $x < \epsilon_1 \wedge \epsilon_2$ . Therefore, by virtue of [13, Proposition 1.2],

$$\int_0^1 \frac{b(x)}{x} dx = \int_{\epsilon_1 \wedge \epsilon_2}^1 \frac{b(x)}{x} dx \leq C_1 h_n^\beta, \quad (5.4)$$

for some  $C_1 > 0$ , whenever  $\mathbb{K}$  is a kernel of order  $l = \lfloor \beta \rfloor$  (see [13, Definition 1.3]) satisfying

$$\int |u|^\beta \mathbb{K}(u) du < \infty.$$

Finally, by (5.3) and (5.4), we have

$$\begin{aligned} \mathbf{E} \left[ \int_0^1 \frac{|\widehat{G}_n^{PR}(x) - G(x)|}{x} dx \right] &\leq \int_0^1 \frac{\sqrt{MSE(x)}}{x} dx \\ &\leq \int_0^1 \frac{b(x)}{x} dx + \int_0^1 \frac{\sqrt{V(x)}}{x} dx \\ &\leq C_2 \left( \frac{1}{\sqrt{nh_n}} + h_n^\beta \right), \end{aligned}$$

for some constant  $C_2$ . We conclude that the  $L^1$ -norm vanishes when  $n$  tends to infinity if the bandwidth is such that  $\sqrt{nh_n} \rightarrow 0$ . Therefore, the convergence in probability ( $C_2^G$ ) holds under this condition.

#### REFERENCES

- [1] Romain Azaïs. A recursive nonparametric estimator for the transition kernel of a piecewise-deterministic Markov process. *ESAIM: PS*, 18:726–749, 2014.
- [2] Romain Azaïs, François Dufour, and Anne Gégout-Petit. Non-parametric estimation of the conditional distribution of the interjumping times for piecewise-deterministic Markov processes. *Scandinavian Journal of Statistics*, 41(4):950–969, 2014.
- [3] Romain Azaïs and Alexandre Genadot. Semi-parametric inference for the absorption features of a growth-fragmentation model. *TEST*, 24(2):341–360, Jun 2015.
- [4] Romain Azaïs and Alexandre Genadot. Estimation of the average number of continuous crossings for non-stationary non-diffusion processes. *arXiv preprint arXiv:1703.10192*, 2017.
- [5] Federico Dalmao and Ernesto Mordecki. Rice formula for processes with jumps and applications. *Extremes*, 18(1):15–35, 2015.
- [6] Mark H. A. Davis. *Markov models and optimization*, volume 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.
- [7] Marie Doumic, Marc Hoffmann, Nathalie Krell, and Lydia Robert. Statistical estimation of a growth-fragmentation model observed on a genealogical tree. *To appear in Bernoulli*, 2014.
- [8] 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, 2011.
- [9] D Geman, J Horowitz, et al. Occupation times for smooth stationary processes. *The Annals of Probability*, 1(1):131–137, 1973.
- [10] M. Kac. On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.*, 49(4):314–320, 04 1943.
- [11] R. M. Kovacevic and G. C. Pflug. Does insurance help to escape the poverty trap? –a ruin theoretic approach. *Journal of Risk and Insurance*, 78(4):1003–1028, 2011.
- [12] Wilhelm Niethammer and Walter Schempp. On the construction of iteration methods for linear equations in banach spaces by summation methods. *Aequationes Math*, 5(1):124–125, 1970.
- [13] Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Publishing Company, Incorporated, 1st edition, 2008.
- [14] Dominik Wied and Rafael Weißbach. Consistency of the kernel density estimator: a survey. *Statistical Papers*, 53(1):1–21, 2012.