Integration by parts formula for killed processes: a point of view from approximation theory

Noufel Frikha*   Arturo Kohatsu-Higa†   Libo Li‡

Abstract

In this paper, we establish a probabilistic representation for two integration by parts formulas, one being of Bismut-Elworthy-Li's type, for the marginal law of a one-dimensional diffusion process killed at a given level. These formulas are established by combining a Markovian perturbation argument with a tailor-made Malliavin calculus for the underlying Markov chain structure involved in the probabilistic representation of the original marginal law. Among other applications, an unbiased Monte Carlo path simulation method for both integration by parts formula stems from the previous probabilistic representations.

1 Introduction

In this article, we consider the following one-dimensional stochastic differential equation (SDE)

$$X_t = x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s, \ x \in \mathbb{R} \tag{1.1}$$

where the coefficients $b, \sigma : \mathbb{R} \to \mathbb{R}$ are smooth and bounded functions and $(W_t)_{t\geq 0}$ stands for a one-dimensional Brownian motion on a given filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

---

*Université de Paris, Laboratoire de Probabilités, Statistiques et Modélisation, F-75013 Paris, France.
E-mail: frikha@math.univ-paris-diderot.fr
†Department of Mathematical Sciences, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu, Shiga, 525-8577, Japan.
E-mail: khts00@fc.ritsumei.ac.jp. This author was supported by grants of the Japanese government KAKENHI 16K05215 and 16H03642.
‡Department of Mathematics and Statistics, University of New South Wales, Sydney, Australia.
E-mail: libo.li@unsw.edu.au

The aim of this paper is to provide a probabilistic representation for two integration by parts (IBP) formulas for the marginal law of the process X killed at a fixed given level L. To be more specific, for a starting point $x \geq L$, let $\tau = \inf\{t \geq 0 : X_t < L\}$ be the first hitting time of the level L by the one-dimensional process X. For a given finite horizon $T > 0$, we are interested in establishing probabilistic representations for IBP formulae related to the following quantities

$$\mathbb{E}[f'(X_T)\mathbb{1}_{\{\tau > T\}}] \quad \text{and} \quad \partial_x \mathbb{E}[f(X_T)\mathbb{1}_{\{\tau > T\}}] \tag{1.2}$$

where f is a real-valued smooth function defined on $[L, \infty)$. Extensions to non-smooth functions or to the transition density of the killed process are also obtained.

In the recent past, IBP formulae have raised a lot of interest as these explicit formulae can be further analyzed to obtain properties of densities or for Monte Carlo simulation among other applications, see Nualart [27] or Malliavin and Thalmaier [25]. The former quantity in (1.2) is commonly considered in the literature on Malliavin calculus, while the latter quantity is referred in the literature to as the Bismut-Elworthy-Li (BEL for short) formula. The BEL formula is of interest for many practical applications such as the numerical computation of price sensitivities in finance for Delta hedging purpose. For a more detailed discussion on this topic, we refer the interested reader to Fournié and al. [13], Fournié and al. [12], Gobet et al. [18] [8] for a short sample.

In particular, in Section 2.6 of [25], the authors propose a continuous time version of the IBP formula for d-dimensional diffusion process X killed when it exits an open sub-domain D of $\mathbb{R}^d$. Denoting by τ the first exit time of X from D, the formula writes $\partial_x \mathbb{E}[f(X_T)\mathbb{1}_{\{\tau > T\}}] = \mathbb{E}[f(X_T)\mathbb{1}_{\{\tau > T\}}H]$ where H has an explicit expression using stochastic integrals. Extensions have also been proposed by Arnaudon and Thalmaier [5], [6]. At this stage, it is important to observe that, from a numerical perspective, these formulae will inevitably involve a time discretization, thus introducing a bias, in order to devise a Monte Carlo simulation method, as it is already the case for the quantity $\mathbb{E}[f(X_T)\mathbb{1}_{\{\tau > T\}}]$, see e.g. Gobet [17], Gobet and Menozzi [19]. As observed in [6], it may also require to compute the solution of a control problem which can be done off-line. One may thus claim as stated at the beginning of Section 2.6 in [25]: "Its Monte-Carlo implementation, which has not yet been done, seems to be relatively expensive in computing time".

Our approach is probabilistic and relies on a perturbation argument of Markov semigroups to derive a probabilistic representation for the marginal law of the killed process based on a simple Markov chain approximation scheme for which we develop an appropriate Malliavin calculus machinery.

The main novelty in comparison with the aforementioned previous works lies in the fact that an unbiased Monte Carlo simulation method directly stems from the integration by parts formulae derived here. One may thus devise an estimator which does not involve any bias but only a statistical error. To the best of our knowledge, this feature appears to be new. As a by product of our analysis, we propose a probabilistic representation for the two derivatives $\partial_x p(T, x, z)$ and $\partial_z p(T, x, z)$ where $(0, \infty) \times [L, \infty)^2 \ni (T, x, z) \mapsto p(T, x, z)$ stands for the transition density evaluated at terminal point z at time T of the process X starting from x and killed at level L. We also point out that devising a Monte Carlo estimator for IBP formulae without introducing a bias from the exact simulation methods of Jenkins [23] or Herrmann and Zucca [21] does not seem to be apparent. An extension of the exact simulation method introduced by Beskos and al. [9] to compute the two first derivatives with respect to the starting point x of the quantity $\mathbb{E}[f(X_T^x)]$, X being a one-dimensional diffusion with constant diffusion coefficient, has been recently proposed by Tanré and Reutenauer [28]. However, it seems difficult to implement from this method a Monte Carlo estimator for the aforementioned derivatives of the transition density of

the diffusion without introducing any bias.

The first step towards obtaining an IBP formula is to prove a probabilistic representation for the marginal law of the killed diffusion process, in the spirit of Bally and Kohatsu-Higa [7] which developed such a formula for multi-dimensional diffusion processes (without stopping) and some Lévy driven SDEs by means of a probabilistic perturbation argument for Markov semigroups. We also refer the reader to Labordère et al. [20] and Agarwal and Gobet [1] for some recent contributions in that direction for multi-dimensional diffusion processes.

Such representation involves a simple Markov chain structure evolving along a time grid given by the jump times of an independent renewal process. A similar representation was derived by the same authors in [14] by means of analytic arguments. However, the representation obtained here is different and more amenable for the implementation of Monte Carlo simulation methods or to establish IBP formulae.

Once such representation is established, we then want to prove suitable IBP formulae using the underlying Markov chain structure and eventually the noise provided by the jump times. For this purpose, we set up a tailor-made Malliavin calculus for this new approximation process and perform a careful propagation argument of spatial derivatives, backward in time for the first quantity in (1.2) and forward in time for the second quantity in (1.2).

These developments are not free of mathematical hurdles. In fact, the proposed methodology leads to the appearance of boundary terms which have to be treated carefully. The key idea that we develop to deal with this issue consists in using the noise provided by the jump times. Contrary to the IBP formulae developed here, we point out that in most cases the explicitness of the IBP formulae for diffusions killed at a boundary demands a number of simplifications and approximations. A technical argument commonly used consists in performing a localisation of the underlying process in order to ensure that it is not close to the boundary. This technique is successful but has some theoretical and practical limitations as shown in Delarue [11], [18] and Nakatsu [26]. This is one of the main reasons for the previously quoted statement in Section 2.6 of [25].

We finally emphasize that the variance of the Monte Carlo estimators associated to the IBP formulae established here tends to be large and even infinite. This feature is not new and appears to be reminiscent of the probabilistic representation originally obtained in [7]. Importance sampling or higher order methods have been proposed to circumvent this issue in the case of multi-dimensional diffusions, see Andersson and Kohatsu-Higa [3]. Following the ideas developed in [3], we show how to achieve finite variance for the Monte Carlo estimators obtained from the probabilistic representation formulas of the marginal law of the killed process and of both IBP formulas by employing an importance sampling scheme on the jump times of the renewal process. We finally provide some numerical tests illustrating our previous analysis.

The article is organized as follows. In Section 2, we provide some basic definitions, assumptions and present a reflection principle based on a simple one step Markov chain that will play a central role in our probabilistic representations for the marginal law of the killed process and for our IBP formulae. In addition, we also construct the adequate Malliavin calculus machinery related to the underlying Markov chain upon which both IBP formulae are made. In Section 3, the probabilistic representation for the marginal law of the killed process, based on the Markov chain of Section 2, is established. The change from the process X to the Markov chain coming from the reflection principle of Section 2 simplifies our analysis as all the irregularities of the process appear as indicator functions. Section 4 is devoted to the main ingredients to obtain our first IBP formula, that is, the IBP formula for the quantity $\mathbb{E}[f'(X_T)\mathbb{1}_{\{\tau>T\}}]$. These ingredients

are put to work in Section 5. Theorem 5.1 is the main result of this section. As a by product, we obtain a probabilistic representation for the first derivative of the transition density of the killed process with respect to its terminal point in Corollary 5.3. In Section 6, we establish the BEL formula for the law of the killed process. The main result of this section is Theorem 6.2 and, as a by product, we obtain a probabilistic representation for the first derivative of the transition density of the killed process with respect to the initial condition x in Corollary 6.4. Many of the proofs of Sections 3 and 4 are technical and postponed to the appendix in Section 10. In Section 7, we show how to achieve finite variance for our unbiased Monte Carlo estimators by an importance sampling technique that we briefly present. Some numerical results are presented in Section 8. Clearly, one needs to study numerical issues in more detail and these are left for later studies.

1.1 Notations

For a fixed given point $z \in \mathbb{R}$, the Dirac measure is denoted by $\delta_z(dx)$. Derivatives may be denoted by $f'(x)$ in the one-dimensional case or by $\partial_i f(x)$ in the multi-dimensional case for the partial derivative with respect to the i-th variable appearing in the multivariate function f or also by $\partial_x^i f(x) \equiv f^{(i)}(x)$ where the latest is used in functions of one variable and the former is used mostly for multivariate functions.

We will often work with continuous or smooth functions defined on $[L, \infty)$. In order to shorten notation, we will consider their extensions[1] on $\mathbb{R}$. This is done just in order to shorten the length of equations and notation. Therefore all statements can be rewritten using the same assumptions but restricted to the domain $[L, \infty)$.

The space of functions which are k-times continuously differentiable on a closed domain D is denoted by $\mathscr{C}^k(D)$. At the boundary points of D, all derivatives are considered as limits taken from the interior of D only. In the case that the derivatives are bounded, the space of corresponding functions is denoted by $\mathscr{C}_b^k(D)$, $k \in \mathbb{N} \cup \{0, \infty\}$. In particular, note that functions in $\mathscr{C}_b^1$ may not be bounded but are at most linearly growing. Finally, $\mathscr{C}_p^k(D)$ denote the class of k-times differentiable functions with at most polynomial growth at infinity. The space of p-integrable random variables (r.v.'s) is denoted by $\mathbb{L}^p$ with its extension $\mathbb{L}^\infty$ for $p = \infty$.

Given a measure space S, the space of real valued Borel measurable functions on S will be denoted by $\mathcal{M}(S)$. We also introduce the simplex $A_n := \{t \in (0, T]^n; 0 < t_1 < \cdots < t_n \leq T\}$, $n \in \mathbb{N}$ where T is fixed throughout the paper.

The transition density function at x of the standard Brownian motion at time t is denoted by $g(t, x) = (2\pi t)^{-1/2} \exp(-x^2/(2t))$. Its associated Hermite polynomials of order i, $i \in \mathbb{N}$, are defined by $\mathcal{H}_i(t, x) = (g(t, x))^{-1} \partial_x^i g(t, x)$.

In order to simplify lengthy equation, we may use the symbol $\overset{\mathbb{E}}{=}$ to mean that two quantities are equal in expectation. Sometimes the same symbol maybe used for equality on conditional expectation. This will be clearly indicated at the point where it is used.

Generic constants are usually denoted by C and are independent of all variables unless otherwise explicitly stated. As usual they may change value from one line to the next.

As we will be using discrete Markov chains in this article, we will often have indexes whose range may be the set of integers. In order to shorten the length of statements, we will use the following notation for the most common set of indexes $\mathbb{N}_n \equiv \{1, ..., n\}$ or $\bar{\mathbb{N}}_n \equiv \{0, ..., n\}$ with $n \in \bar{\mathbb{N}} \equiv \mathbb{N} \cup \{0\}$. In the case that $n \leq 0$, then $\mathbb{N}_n := \emptyset$.

---

[1] In most cases, unless explicitly said we assume that functions are extended using e.g. Whitney's extension theorem.

2 Preliminaries

2.1 Assumptions and basic definitions

Throughout the article, we work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which is assumed to be rich enough to support all r.v's considered in what follows. In addition, we will work under the following assumptions on the coefficients:

Assumption (H)

(i) The coefficients of the SDE (1.1) are smooth and bounded, in particular, $b \in \mathscr{C}_b^\infty(\mathbb{R})$ and $a \in \mathscr{C}_b^\infty(\mathbb{R})$.

(ii) The function σ is bounded and uniformly elliptic, that is, there exist $\underline{a}, \overline{a} > 0$ such that for any $x \in \mathbb{R}$, $\underline{a} \leq a(x) = \sigma^2(x) \leq \overline{a}$. Therefore, without loss of generality, we will assume that $\sigma(x) > 0$.

2.2 A reflection principle

Our probabilistic representation involves the following approximation process

$$\bar{X}_t^{s,x} \equiv \bar{X}_t^{s,x}(\rho) = \rho x + (1-\rho)(2L - x) + \sigma(x)(W_t - W_s), \tag{2.1}$$

where ρ is a r.v. distributed according to a Bernoulli(1/2) law, independent of W, namely $\mathbb{P}(\rho = 1) = \mathbb{P}(\rho = 0) = 1/2$. Under our assumption on the coefficients, the flow derivatives of $\bar{X}_t^{s,x}$ exist. In particular, one has

$$\partial_x \bar{X}_t^{s,x} = 2\rho - 1 + \sigma'(x)(W_t - W_s), \quad \partial_x^2 \bar{X}_t^{s,x} = \sigma''(x)(W_t - W_s). \tag{2.2}$$

In the particular case that $s = 0$, we may use the simplified form $\bar{X}_t = \bar{X}_t^{0,x}$. At this point, we give a brief explanation about how the approximation process $\bar{X}$ appears in the forthcoming probabilistic representation. The proof of the following lemma is straightforward by using the reflection principle, see e.g. Karatzas and Shreve [24].

Lemma 2.1. *Define the following approximation process:*

$$\bar{Y}_t = x + \sigma(x)W_t, \quad x \geq L,$$

together with its associated exit time $\bar{\tau} := \inf\{t, \bar{Y}_t = L\}$. Then, for any bounded measurable function f, the following property is satisfied:

$$\mathbb{E}\left[f(\bar{Y}_T)\mathbb{1}_{\{\bar{\tau} > T\}}\right] = \mathbb{E}\left[f(\bar{Y}_T)\mathbb{1}_{\{\bar{Y}_T \geq L\}}\right] - \mathbb{E}\left[f(2L - \bar{Y}_T)\mathbb{1}_{\{\bar{Y}_T < L\}}\right]$$

$$= 2\mathbb{E}\left[(2\rho - 1)f(\bar{X}_T)\mathbb{1}_{\{\bar{X}_T \geq L\}}\right]. \tag{2.3}$$

2.3 Basic Markov chain framework

We also consider a Poisson process with parameter $\lambda > 0$, independent of the one-dimensional Brownian motion W with jump times T_i, $i \in \mathbb{N}$ and we set $\zeta_i := T_i \wedge T$, $i \in \mathbb{N}$ with the convention that $\zeta_0 = T_0 = 0$.

Define π to be the partition of $[0, T]$ given by $\pi := \{0 =: \zeta_0 < \cdots < \zeta_{N_T} \leq T\}$. Associated with this set, we recall the definition of the simplex $A_n := \{t \in (0, T]^n; 0 < t_1 < \cdots < t_n \leq T\}$. For instance, on the set $\{N_T = n\}$, $n \in \mathbb{N}$, we have $(\zeta_1, \ldots, \zeta_n) \in A_n$ and $\zeta_{n+1} = T$. In particular, for the set $\{N_T = 0\}$ (i.e. $n = 0$), we let $\pi := \{0, T\}$ and $A_0 = \emptyset$. In this sense, we will use throughout the rest of the paper the index $n \in \bar{\mathbb{N}}$ without any further mention of its range of values.

As it is the case in the previous observation, many proofs and definitions will be carried out conditioning on the set $\{N_T = n\}$, $n \in \bar{\mathbb{N}}$.

Let $\bar{X} := (\bar{X}_i)_{i\in\bar{\mathbb{N}}}$ be the discrete time Markov chain starting at time 0 from $\bar{X}_0 = x$ and evolving according to

$$\bar{X}_{i+1} := \rho_{i+1}\bar{X}_i + (1 - \rho_{i+1})(2L - \bar{X}_i) + \sigma_i Z_{i+1}, \quad i \in \bar{\mathbb{N}}_{N_T}, \tag{2.4}$$

where for simplicity we set $\sigma_i := \sigma(\bar{X}_i)$, $Z_{i+1} := W_{\zeta_{i+1}} - W_{\zeta_i}$ (note that one has $Z_{i+1} = \sigma_i^{-1}\left(\bar{X}_{i+1} - \rho_{i+1}\bar{X}_i - (1 - \rho_{i+1})(2L - \bar{X}_i)\right)$) and $\{\rho_i; i \in \mathbb{N}\}$ is an i.i.d. sequence of Bernoulli(1/2) r.v.'s such that W, N and $\{\rho_i; i \in \mathbb{N}\}$ are mutually independent. In what follows we use the notation $h_i \equiv h(\bar{X}_i)$, $i \in \bar{\mathbb{N}}_{n+1}$ for any function $h : \mathbb{R} \to \mathbb{R}$. In particular, the reader may have noticed that we already used this notation in the above formula for $h = \sigma$. We also associate to the Markov chain $\bar{X}$ the following sets

$$D_{i,n} := \{\bar{X}_i \geq L, N_T = n\}, \quad \text{for } i \in \bar{\mathbb{N}}_{n+1}. \tag{2.5}$$

It is important to point out that given $N_T = n$, the conditional distribution of $\bar{X}_{n+1}$ given $\bar{X}_n$ is not the same as the conditional distribution of $\bar{X}_{i+1}$ given $\bar{X}_i$ for $i \in \bar{\mathbb{N}}_{n-1}$. This is due to the fact that the length of the last interval is $T - \zeta_n$, rather than the waiting time between two consecutive Poisson jumps. This remark applies to various definitions and results to be stated in the rest of the article.

We define the filtration $\mathcal{G} := (\mathcal{G}_i)_{i\in\bar{\mathbb{N}}}$ where $\mathcal{G}_i := \sigma(Z^i, \zeta^i, \rho^i)$ with the notation $a^i := (a_1, \ldots, a_i)$ for $a = Z$, ζ, ρ, $i \in \mathbb{N}$ and $\mathcal{G}_0$ defined as the trivial σ-field. We assume that the filtration $\mathcal{G}$ satisfies the usual conditions.

2.4 Simplified Malliavin calculus for the underlying Markov chain

In this section we introduce the required material for our Malliavin calculus computations. Instead of using an infinite dimensional calculus as it is usually done in the literature, the approach developed below is based on a finite dimensional calculus for which the dimension is given by the number of jumps of the underlying Poisson process involved in the Markov chain $\bar{X}$. In what follows, $n \in \bar{\mathbb{N}}$ unless stated otherwise.

We start by defining the following space of smooth r.v.'s.

Definition 2.2. *For $i \in \bar{\mathbb{N}}_n$, we define the set $\mathbb{S}_{i+1,n}(\bar{X})$ as the subset of r.v.'s $H \in \mathbb{L}^0$ such that there exists a measurable function $h : \mathbb{R}^2 \times \{0,1\} \times A_2 \to \mathbb{R}$ satisfying*

1. *$H = h(\bar{X}_i, \bar{X}_{i+1}, \rho_{i+1}, \zeta_i, \zeta_{i+1})$ on the set $\{N_T = n\}$,*

2. *For any $r \in \{0,1\}$ and any $(s,t) \in A_2$, $h(\cdot, \cdot, r, s, t) \in \mathscr{C}_p^\infty(\mathbb{R}^2)$.*

For a r.v. $H \in \mathbb{S}_{i+1,n}(\bar{X})$, $i \in \bar{\mathbb{N}}_n$, we may sometimes abuse the notation and write

$$H \equiv H(\bar{X}_i, \bar{X}_{i+1}, \rho_{i+1}, \zeta_i, \zeta_{i+1}), \tag{2.6}$$

that is the same symbol H may denote the r.v. or the function in the set $\mathbb{S}_{i+1,n}(\bar{X})$. One can easily define the flow derivatives for $H \in \mathbb{S}_{i+1,n}(\bar{X})$ as follows:

$$\begin{aligned}
\partial_{\bar{X}_{i+1}} H &:= \partial_2 h(\bar{X}_i, \bar{X}_{i+1}, \rho_{i+1}, \zeta_i, \zeta_{i+1}), \\
\partial_{\bar{X}_i} H &:= \partial_1 h(\bar{X}_i, \bar{X}_{i+1}, \rho_{i+1}, \zeta_i, \zeta_{i+1}) + \partial_2 h(\bar{X}_i, \bar{X}_{i+1}, \rho_{i+1}, \zeta_i, \zeta_{i+1})\partial_{\bar{X}_i}\bar{X}_{i+1}, \\
\partial_{\bar{X}_i}\bar{X}_{i+1} &:= (2\rho_{i+1} - 1) + \sigma_i' Z_{i+1}.
\end{aligned} \tag{2.7}$$

We now define the derivative and integral operators for $H \in \mathbb{S}_{i+1,n}(\bar{X})$, $i \in \bar{\mathbb{N}}_n$, as

$$\mathcal{I}_{i+1}(H) := H\frac{Z_{i+1}}{\sigma_i(\zeta_{i+1} - \zeta_i)} - \mathcal{D}_{i+1}H, \qquad \mathcal{D}_{i+1}H := \partial_{\bar{X}_{i+1}}H. \tag{2.8}$$

Note that due to the above definitions and assumption (**H**), we also have that $\mathcal{I}_{i+1}(H), \mathcal{D}_{i+1}H \in \mathbb{S}_{i+1,n}(\bar{X})$ so that we can define iterations of the above operators,

namely $\mathcal{I}_{i+1}^{\ell+1}(H) = \mathcal{I}_{i+1}(\mathcal{I}_{i+1}^{\ell}(H))$ and similarly $\mathcal{D}_{i+1}^{\ell+1}(H) = \mathcal{D}_{i+1}(\mathcal{D}_{i+1}^{\ell}H)$, $\ell \in \bar{\mathbb{N}}$, with the convention $\mathcal{I}_{i+1}^{0}(H) = \mathcal{D}_{i+1}^{0}(H) = H$.

Through this article, we will use the following notation for a certain type of conditional expectation that will appear frequently. For any $X \in \mathbb{L}^1$ and any $i \in \bar{\mathbb{N}}_n$,

$$\mathbb{E}_{i,n}[X] := \mathbb{E}[X \,|\, \mathcal{G}_i, T^{n+1}, \rho^{n+1}, N_T = n].$$

With the above definitions, the following duality[2] is satisfied for any $f \in \mathscr{C}_p^1(\mathbb{R})$ and any $(i, \ell) \in \bar{\mathbb{N}}_n \times \mathbb{N}$:

$$\mathbb{E}_{i,n}\left[\mathcal{D}_{i+1}^{\ell}f(\bar{X}_{i+1})H\right] = \mathbb{E}_{i,n}\left[f(\bar{X}_{i+1})\mathcal{I}_{i+1}^{\ell}(H)\right]. \tag{2.9}$$

In order to obtain explicit norm estimates for r.v.'s in $\mathbb{S}_{i+1,n}(\bar{X})$ it is useful to define for $H \in \mathbb{S}_{i+1,n}(\bar{X})$, $i \in \bar{\mathbb{N}}_n$ and $p \geq 1$

$$\|H\|_{p,i,n}^p := \mathbb{E}_{i,n}\left[|H|^p\right].$$

Another useful formula that is used at several places later on is the following *extraction formula* for $H_1, H_2 \in \mathbb{S}_{i+1,n}(\bar{X})$:

$$\mathcal{I}_{i+1}^{\ell}(H_1 H_2) = \sum_{j=0}^{\ell}(-1)^j \binom{\ell}{j}\mathcal{I}_{i+1}^{\ell-j}(H_1)\mathcal{D}_{i+1}^{j}H_2. \tag{2.10}$$

The proof of the above statement is done by induction. Then, by iteration, one obtains that $\mathcal{I}_i^{\ell}(1) \in \mathbb{S}_{i,n}(\bar{X})$ and it satisfies $\mathcal{I}_i^{\ell+1}(1) = \mathcal{I}_i^{\ell}(1)\mathcal{I}_i(1) - \ell\mathcal{I}_i^{\ell-1}(1)/(a_{i-1}(\zeta_i - \zeta_{i-1}))$ which, in particular, implies:

$$\mathcal{I}_i^{\ell}(1) = (-1)^{\ell}\mathcal{H}_{\ell}(a_{i-1}(\zeta_i - \zeta_{i-1}), \sigma_{i-1}Z_i). \tag{2.11}$$

Using (2.10) for $H_1 = 1$ and $H_2 = H$ as well as (2.8), the following norm bound for stochastic integrals is clearly satisfied for $i \in \bar{\mathbb{N}}$ and any measurable set $A \in \mathcal{F}$

$$\left\|\mathbb{1}_A \mathcal{I}_{i+1}^{\ell}(H)\right\|_{p,i,n} \leq C_{\ell,p}\sum_{j=0}^{\ell}(\zeta_{i+1} - \zeta_i)^{\frac{j-\ell}{2}}\|\mathbb{1}_A \mathcal{D}_{i+1}^j H\|_{p,i,n}. \tag{2.12}$$

The following Hölder like inequality for smooth r.v.'s $H_1, H_2 \in \mathbb{S}_{i+1,n}(\bar{X})$ is also frequently used in our computations without any further mentioning for any $i \in \bar{\mathbb{N}}$, for any $p, p_1, p_2 \geq 1$ satisfying $p^{-1} = p_1^{-1} + p_2^{-1}$ and any $A \in \mathcal{F}$:

$$\|\mathbb{1}_A H_1 H_2\|_{p,i,n} \leq C\|\mathbb{1}_A H_1\|_{p_1,i,n}\|\mathbb{1}_A H_2\|_{p_2,i,n}. \tag{2.13}$$

We will also frequently manipulate quantities such as $\mathbb{E}_{i,n}[\delta_L(\bar{X}_{i+1})H]$ for $H \in \mathbb{S}_{i+1,n}(\bar{X})$. The previous expression has a clear meaning due to the IBP formula (2.9) (see the theory in [22], Chapter V.9. for a much more general framework) or in the sense of conditional laws

$$\mathbb{E}_{i,n}[\delta_L(\bar{X}_{i+1})H] = \mathbb{E}_{i,n}[\mathbb{1}_{D_{i+1,n}}\mathcal{I}_{i+1}(H)] = \mathbb{E}_{i,n}[H|\bar{X}_{i+1} = L]g(a_i(\zeta_{i+1} - \zeta_i), L - \bar{X}_i).$$

We finally introduce the following space of r.v.'s with certain time (ir)regularity estimates.

---

[2]This duality is obtained using the Gaussian density of $\bar{X}_{i+1}$ while in classical Malliavin calculus it is based on the density of the Wiener process. Therefore the derivative and integral, $\mathcal{D}_{i+1}$ and $\mathcal{I}_{i+1}$, $i \in \bar{\mathbb{N}}_n$ defined in the formula (2.9) are renormalizations of the usual duality principle in Malliavin calculus. In our case this notation simplifies greatly many equations.

Definition 2.3. *For $\ell \in \mathbb{Z}$, $i \in \bar{\mathbb{N}}_n$, we define the space $\mathbb{M}_{i+1,n}(\bar{X}, \ell/2)$ as the set of r.v.'s $H \in \mathbb{L}^0$, satisfying the property (1) in Definition 2.2 and such that*

$$\mathbb{1}_{D_{i,n}} \|\mathbb{1}_{D_{i+1,n}} H\|_{p,i,n} \leq C(\zeta_{i+1} - \zeta_i)^{\ell/2}$$

for some deterministic constant C independent of (p, i, n).

We again remark that since the definition of the space $\mathbb{M}_{i+1,n}(\bar{X}, \ell/2)$ uses the conditional norm $\mathbb{E}_{i,n}[.]$ and property *(1)* in Definition 2.2, when we say that a r.v. $H \in \mathbb{M}_{i+1,n}(\bar{X}, \ell/2)$, this statement is understood on the set $\{N_T = n\}$, $n \in \mathbb{N}$.

A straightforward consequence of equation (2.10) and (2.12) is the following property.

Lemma 2.4. *For $j \in \{0, 1\}$, $i \in \bar{\mathbb{N}}$, and $k \in \mathbb{N}$, if $H_1 \in \mathbb{M}_{i+1,n}(\bar{X}, j/2) \cap \mathbb{S}_{i+1,n}(\bar{X})$ with $\|\mathcal{D}_{i+1}^k H_1\|_{p,i,n} \leq C$ then $\mathcal{I}_{i+1}^k(H_1) \in \mathbb{M}_{i+1,n}(\bar{X}, (j-k)/2)$. Furthermore, if $H_2 \in \mathbb{M}_{i+1,n}(\bar{X}, k/2)$ then the product $H_1 H_2 \in \mathbb{M}_{i+1,n}(\bar{X}, (j+k)/2)$.*

Lemma 2.5. *Let $h \equiv h(\bar{X}_i, \bar{X}_{i+1}, \rho_{i+1}, \zeta_i, \zeta_{i+1}) \in \mathbb{S}_{i+1,n}(\bar{X})$ with $i \in \bar{\mathbb{N}}$, then the following chain rule type formula holds*

$$\partial_{\bar{X}_i} \mathcal{I}_{i+1}(h) = \mathcal{I}_{i+1}(\partial_{\bar{X}_i} h) - \frac{\sigma_i'}{\sigma_i} \mathcal{I}_{i+1}(h).$$

Proof. From the extraction formula (2.10), the usual chain rule and the fact that $\partial_{\bar{X}_i} \mathcal{I}_{i+1}(1) = -\frac{\sigma_i'}{\sigma_i} \mathcal{I}_{i+1}(1)$,

$$\partial_{\bar{X}_i} \mathcal{I}_{i+1}(h) = -\frac{\sigma_i'}{\sigma_i} \mathcal{I}_{i+1}(1) h + \mathcal{I}_{i+1}(1) \partial_{\bar{X}_i} h - \partial_{\bar{X}_i} \mathcal{D}_{i+1} h$$

$$\mathcal{I}_{i+1}(\partial_{\bar{X}_i} h) = \mathcal{I}_{i+1}(1) \partial_{\bar{X}_i} h - \mathcal{D}_{i+1} \partial_{\bar{X}_i} h.$$

Note that $\partial_{\bar{X}_i}$ and $\mathcal{D}_{i+1}$ do not commute. Indeed, by the usual chain rule, one has

$$\partial_{\bar{X}_i} \mathcal{D}_{i+1} h - \mathcal{D}_{i+1} \partial_{\bar{X}_i} h = \partial_1 \partial_2 h + \partial_2^2 h \frac{\partial \bar{X}_{i+1}}{\partial \bar{X}_i} - \mathcal{D}_{i+1}(\partial_1 h + \partial_2 h \frac{\partial \bar{X}_{i+1}}{\partial \bar{X}_i}) = -\partial_2 h \frac{\sigma_i'}{\sigma_i}$$

where in the last equality we used the fact that $\mathcal{D}_{i+1} \frac{\partial \bar{X}_{i+1}}{\partial \bar{X}_i} = \frac{\sigma_i'}{\sigma_i}$. By combining the above computations we obtain

$$\partial_{\bar{X}_i} \mathcal{I}_{i+1}(h) - \mathcal{I}_{i+1}(\partial_{\bar{X}_i} h) = -\frac{\sigma_i'}{\sigma_i} \left(\mathcal{I}_{i+1}(1) - \mathcal{D}_{i+1} h \right) = -\frac{\sigma_i'}{\sigma_i} \mathcal{I}_{i+1}(h). \qquad \square$$

Remark 2.6. Heuristically, the result of Lemma 2.5 can be viewed as a chain rule formula of the type

$$\partial_{\bar{X}_i} (\mathcal{I}_{i+1}(h)) \equiv (\mathcal{I}_{i+1})(\partial_{\bar{X}_i} h) + (\partial_{\bar{X}_i} \mathcal{I}_{i+1})(h).$$

3 Markov chain representation for killed processes

In this section, we establish a probabilistic representation for the marginal law of the killed process, that is, for the law of X_T on the set $\{\tau > T\}$ based on the Markov chain $\bar{X}$ introduced in the previous section. The proof of the following result is postponed to Appendix 10.

Theorem 3.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function with at most polynomial growth at infinity such that $f(L) = 0$. Then, one has*

$$\mathbb{E}\left[f(X_T) \mathbb{1}_{\{\tau > T\}} \right] = \mathbb{E}\left[f(\bar{X}_{N_T+1}) \prod_{i=1}^{N_T+1} \mathbb{1}_{D_{i,N_T}} \bar{\theta}_i \right]. \tag{3.1}$$

Here, for $i \in \mathbb{N}_{N_T+1}$, we let

$$\bar{\theta}_i := \mathbb{1}_{\{N_T > i-1\}} 2(2\rho_i - 1)\lambda^{-1} \left(\mathcal{I}_i(c_1^i) + \mathcal{I}_i^2(c_2^i) \right) + \mathbb{1}_{\{N_T = i-1\}} 2e^{\lambda T}(2\rho_{N_T+1} - 1). \quad (3.2)$$

In the above definitions, we have used $c_j^i \in \mathbb{S}_{i,n}(\bar{X})$, $j = 1, 2$, $i \in \mathbb{N}_{n+1}$, $n \in \bar{\mathbb{N}}$ where

$$c_1^i \equiv c_1^i(\bar{X}_{i-1}, \bar{X}_i) := b_i = b(\bar{X}_i), \quad (3.3)$$

$$c_2^i \equiv c_2^i(\bar{X}_{i-1}, \bar{X}_i) := \frac{1}{2}(a_i - a_{i-1}) = \frac{1}{2}(a(\bar{X}_i) - a(\bar{X}_{i-1})).$$

Furthermore, we have $\bar{\theta}_i \in \mathbb{S}_{i,n}(\bar{X})$, $i \in \mathbb{N}_{n+1}$, $n \in \bar{\mathbb{N}}$ and the following time degeneracy estimate is satisfied: for all $p \geq 1$, there exists some positive constant $C := C(T, a, b, p)$ such that for $i \in \mathbb{N}_n$,

$$\mathbb{1}_{\{i \leq n\}}(\zeta_i - \zeta_{i-1})^{\frac{p}{2}} \mathbb{1}_{D_{i-1,n}} \mathbb{E}_{i-1,n} \left[\mathbb{1}_{D_{i,n}} |\bar{\theta}_i|^p \right] + \mathbb{1}_{\{i=n+1\}} \mathbb{1}_{D_{n,n}} \mathbb{E}_{n,n} \left[\mathbb{1}_{D_{n+1,n}} |\bar{\theta}_{n+1}|^p \right] \leq C. \quad (3.4)$$

As a consequence, for all $p \in [0, 2)$, one has $\mathbb{E}\left[\left| \prod_{i=1}^{N_T+1} \mathbb{1}_{D_{i,N_T}} \bar{\theta}_i \right|^p \right] < \infty$.

We stress that the above probabilistic representation allows one to implement an unbiased Monte-Carlo simulation method since one just has to simulate the Poisson process N, then the Markov chain $\bar{X}$ along the jump times of N and finally to compute the explicit product of weights. Contrary to our previous work [14], the Markov chain $\bar{X}$ is defined from the reflection principle introduced in Lemma 2.1 thus reducing further the number of variables in the problem. This probabilistic representation is thus more workable for the forthcoming IBP formulae.

Remark 3.2. We make several remarks before moving on:

(i) The assumption $f(L) = 0$ can be avoided at the cost of longer formulae as one can obtain a probabilistic representation based on the same Markov chain for the probability $\mathbb{P}(\tau \geq T)$ using the results in Section 3 of [14], but we do not pursue this goal here.

(ii) In the proof of the above result, one uses the following crucial property: For $i \in \bar{\mathbb{N}}_n$, the random variables $\mathbb{1}_{D_{i+1,n}} c_1^{i+1}$, $\mathbb{1}_{D_{i+1,n}} \partial_{\bar{X}_i} c_2^{i+1} \in \mathbb{M}_{i+1,n}(\bar{X}, 0)$ and $\mathbb{1}_{D_{i+1,n}} c_2^{i+1} \in \mathbb{M}_{i+1,n}(\bar{X}, 1/2)$ so that $\mathbb{1}_{D_{i+1,n}} \mathcal{I}_{i+1}(c_1^{i+1})$, $\mathbb{1}_{D_{i+1,n}} \mathcal{I}_{i+1}^2(c_2^{i+1}) \in \mathbb{M}_{i+1,n}(\bar{X}, -1/2)$ and therefore $\mathbb{1}_{D_{i+1,n}} \bar{\theta}_{i+1} \in \mathbb{M}_{i+1,n}(\bar{X}, -1/2)$.

(iii) The power $p/2$ appearing in the time degeneracy estimates in (3.4) is crucial in order to determine the integrability of the r.v. appearing on the right hand side of (3.1). This motivates the definition below.

Definition 3.3. A weight r.v. $H \in \mathbb{S}_{i,n}$ is said to satisfy the time degeneracy estimate if for all $p \geq 1$

$$\mathbb{1}_{D_{i-1,n}} \left\| \mathbb{1}_{D_{i,n}} H \right\|_{p,i-1,n} \leq C(\zeta_i - \zeta_{i-1})^{-\frac{1}{2}} \quad (3.5)$$

in the case that $i \in \mathbb{N}_n$ and $\mathbb{1}_{D_{n,n}} \left\| \mathbb{1}_{D_{n+1,n}} H \right\|_{p,n,n} \leq C$ in the case that $i = n+1$.

At this stage, we find it useful to show graphically the dynamic structure of the Markov chain and the random weights $\bar{\theta}_i$, $i \in \mathbb{N}_{N_T+1}$ for the probabilistic representation (3.1). This will be important in order to understand the structure of the IBP formula.
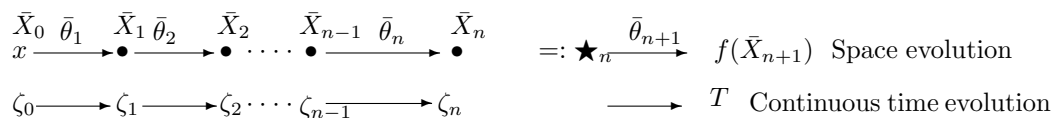
Figure 1: The time evolution of the Markov chain and its weights.

In the above Figure 1, one observes the evolution of the Markov chain $\bar{X}_i$, $i \in \bar{\mathbb{N}}_{n+1}$, together with its associated weight sequence appearing in the probabilistic representation. We also note that at the end of the Markov chain evolution which always happens at time T, the test function f is evaluated at $\bar{X}_{n+1}$. Furthermore, the right hand side of (3.1) is the product of all elements on top of the arrows on the first line of the above figure and the corresponding indicator functions of the sets $D_{i,n}$ defined by (2.5), with $f(\bar{X}_{n+1})$. The second line gives the time evolution followed by the jump times of the Poisson process N which coincide with the times at which transitions of the Markov chain $\bar{X}$ happen.

In other figures that appear later on, we will use the general symbol $\bigstar_k$, $k \in \mathbb{N}_{n+1}$ defined in Figure 1 which stands for product $\bigstar_k := \prod_{i=1}^{k} \mathbb{1}_{D_{i,n}} \bar{\theta}_i$. We remark here that as stated in (3.2), on the set $\{N_T = n\}$, the definition of $(\bar{X}_{n+1}, \bar{\theta}_{n+1})$ differs from all other $(\bar{X}_i, \bar{\theta}_i)$, $i \in \mathbb{N}_n$.

4 Ingredients for an IBP formula

In this section we give the main ingredients in order to establish an integration by parts formula for the marginal law taken at time T of the killed process.

In order to understand the main ingredients to be introduced in this section, one starts by supposing that the problem is to find an IBP formula for $\mathbb{E}\left[f'(X_T)\mathbb{1}_{\{\tau \geq T\}}\right]$. The BEL formula will be easier to handle and is tackled in Section 6.

4.1 Some heuristics

The following heuristic arguments may help the reader to understand the strategy developed in the next sections to prove our IBP formulae.

Step 1: The first step was performed with the probabilistic representation established in Theorem 3.1 involving the Markov chain $\bar{X}$ evolving on a time grid governed by the set of jump times of the Poisson process N.

Step 2: As already explained in the introduction, the central idea is to reduce the infinite dimensional problem of finding an IBP formula for $\mathbb{E}[f'(X_T)\mathbb{1}_{\{\tau > T\}}]$ to a finite dimensional problem, namely finding an IBP formula for $\mathbb{E}\left[f'(\bar{X}_{N_T+1}) \prod_{i=1}^{N_T+1} \mathbb{1}_{D_{i,N_T}} \bar{\theta}_i\right]$, for which the dimension is random and given by the number of jumps of the Poisson process at time T.

At this stage, unfortunately, one cannot perform a standard integration by parts formula as in [27] on the whole time interval $[0, T]$ for various reasons. For example, the Skorokhod integral of the product of weights $\prod_{i=1}^{N_T+1} \mathbb{1}_{D_{i,N_T}} \bar{\theta}_i$ will inevitably involve the Malliavin derivatives of $\bar{\theta}_i$ and the indicator function $\mathbb{1}_{D_{i,N_T}}$, which in turn will raise integrability problems of the resulting Malliavin weight.

The key idea that we use consists in performing IBP formulae locally on each random intervals $[\zeta_i, \zeta_{i+1}]$, for $i \in \bar{\mathbb{N}}_n$ on the set $\{N_T = n\}$, that is, by using the noise of the Markov chain on this specific time interval and then by combining them in an suitable way. The case $i = n$ is easy to consider because, by taking the conditional expectation $\mathbb{E}_{n,n}[.]$ in the original probabilistic representation on the set $\{N_T = n\}$, the IBP reduces to $\mathbb{E}_{n,n}[f'(\bar{X}_{N_T+1})\mathbb{1}_{D_{n+1,n}} \bar{\theta}_n] = \mathbb{E}_{n,n}[f(\bar{X}_{N_T+1})\mathbb{1}_{D_{n+1,n}} \mathcal{I}_{n+1}(\bar{\theta}_{n+1})]$ by (2.9) and the fact that $f(L) = 0$. However, performing the IBP formula on a random interval $[\zeta_i, \zeta_{i+1}]$ for $0 \leq i < n$ is more challenging.

In order to do it, our first ingredient consists in transferring the derivative operator appearing on the test function f backward in time from the last interval to the interval on which we perform the local IBP, say $[\zeta_i, \zeta_{i+1}]$. This operation will unfortunately generate a boundary term each time a transfer is performed, i.e. one for each time interval $[\zeta_j, \zeta_{j+1}]$, $j = i, \cdots, n$. As previously mentioned, the boundary terms will induce integrability

problems. In order to circumvent this issue, we then introduce our second ingredient which consists in performing a *boundary merging procedure* of this term using the time randomness provided by the Poisson process. We refer the reader to Section 4.3 for a more detailed discussion of this issue. Then, an additional ingredient that we use is a time merging lemma about the reduction of jump times of the Poisson process. It is described in Lemma 10.8 and directly employed in the proof of our main result, namely Theorem 5.1.

As already explained above, the last ingredient consists in combining these various local IBP formulae in a suitable way. Roughly speaking, we consider a weighted sum of each integral operator, the weight being the length of the corresponding time interval.

4.2 The transfer of derivative lemma

Lemma 4.1. Let $f \in \mathscr{C}_p^1(\mathbb{R})$ and $n \in \bar{\mathbb{N}}$, the following transfer of derivative formula holds for $i \in \bar{\mathbb{N}}_{n-1}$:

$$\mathbb{E}_{i,n}\big[\partial_{\bar{X}_{i+1}} f(\bar{X}_{i+1}) \mathbb{1}_{D_{i+1,n}} \bar{\theta}_{i+1}\big] = \partial_{\bar{X}_i} \mathbb{E}_{i,n}\big[f(\bar{X}_{i+1}) \mathbb{1}_{D_{i+1,n}} \overleftarrow{\theta}^e_{i+1}\big] \tag{4.1}$$
$$+ \mathbb{E}_{i,n}\big[f(\bar{X}_{i+1})\big(\mathbb{1}_{D_{i+1,n}} \overleftarrow{\theta}^c_{i+1} + \delta_L(\bar{X}_{i+1}) \overleftarrow{\theta}^\partial_{i+1}\big)\big],$$

where the r.v.'s $(\overleftarrow{\theta}^e_{i+1}, \overleftarrow{\theta}^c_{i+1}, \overleftarrow{\theta}^\partial_{i+1}) \in \mathbb{S}_{i+1,n}(\bar{X})$ are defined by

$$\overleftarrow{\theta}^e_{i+1} := 2\lambda^{-1}\left(\mathcal{I}^2_{i+1}(d^{i+1}_2) + \mathcal{I}_{i+1}(d^{i+1}_1)\right),$$
$$\overleftarrow{\theta}^c_{i+1} := \mathcal{I}_{i+1}\left(\bar{\theta}_{i+1} - (2\rho_{i+1}-1)\overleftarrow{\theta}^e_{i+1}\right) - \partial_{\bar{X}_i} \overleftarrow{\theta}^e_{i+1} - \sigma'_i \mathcal{I}_{i+1}\left(Z_{i+1} \overleftarrow{\theta}^e_{i+1}\right), \tag{4.2}$$
$$\overleftarrow{\theta}^\partial_{i+1} := 2(2\rho_{i+1}-1)\lambda^{-1}(a'(L) - b(L))\mathcal{I}_{i+1}(1),$$
$$d^{i+1}_1 := c^{i+1}_1 - (2\rho_{i+1}-1)\partial_{\bar{X}_i} c^{i+1}_2, \tag{4.3}$$
$$d^{i+1}_2 := c^{i+1}_2. \tag{4.4}$$

Assume additionally that $f(L) = 0$. Then, one has

$$\mathbb{E}_{n,n}\big[\partial_{\bar{X}_{n+1}} f(\bar{X}_{n+1}) \mathbb{1}_{D_{n+1,n}} \bar{\theta}_{n+1}\big] = \partial_{\bar{X}_n} \mathbb{E}_{n,n}\big[f(\bar{X}_{n+1}) \mathbb{1}_{D_{n+1,n}} \overleftarrow{\theta}^e_{n+1}\big] \tag{4.5}$$
$$+ \mathbb{E}_{n,n}\big[f(\bar{X}_{n+1}) \mathbb{1}_{D_{n+1,n}} \overleftarrow{\theta}^c_{n+1}\big]$$

with $\overleftarrow{\theta}^\partial_{n+1} := 0$, $\overleftarrow{\theta}^e_{n+1} := 2e^{\lambda T}$ and $\overleftarrow{\theta}^c_{n+1} := -2e^{\lambda T}(\sigma'\sigma)_n(T - \zeta_n)\mathcal{I}^2_{n+1}(1) \in \mathbb{S}_{n+1,n}$. With the above definitions, $x \mapsto \mathbb{E}_{i,n}[f(\bar{X}_{i+1}) \mathbb{1}_{D_{i+1,n}} \overleftarrow{\theta}^e_{i+1} | \bar{X}_i = x] \in \mathscr{C}_p^1(\mathbb{R})$, a.s. for $i \in \bar{\mathbb{N}}_n$. Moreover, one has that $\overleftarrow{\theta}^a_{i+1}$, for $a \in \{e, c, \partial\}$ satisfies the time degeneracy estimates for $i \in \bar{\mathbb{N}}_n$.

The proof of Lemma 4.1 is postponed to Appendix 10.2. The transfer of derivatives procedure starts on the last time interval $[\zeta_n, T]$ according to formula (4.5). It expresses the fact that the derivative operator $\partial_{\bar{X}_{n+1}}$ on the left hand side of the equation is transferred to a flow derivative $\partial_{\bar{X}_n}$ of the conditional expectation on the right hand side of (4.5). Remark that the derivative of f has been written ubiquitously as $\partial_{\bar{X}_{n+1}} f(\bar{X}_{n+1})$ and that exceptionally in the last time interval there is no boundary term due to the assumption $f(L) = 0$.

Then, by the Markov property, the first conditional expectation appearing on the right hand side of (4.1) can be expressed as a function of $\bar{X}_n$ which will be the new test function that will be used in (4.1) for the case $i = n - 1 > 0$ and so on.

The transfer of derivatives for other time intervals is obtained in (4.1). This formula in comparison with (4.5) has a boundary term which is denoted by $\overleftarrow{\theta}^\partial_{i+1}$. In this fashion, various transfer of derivatives formulae can be obtained by transferring successively

the derivative operator through all intervals backward in time. The left pointing arrow appearing on the top of the notation $\overset{\leftarrow}{\theta}{}^{a}_{i+1}$, for $a \in \{e, c, \partial\}$ expresses the fact that we are performing a transfer of derivatives backward in time.
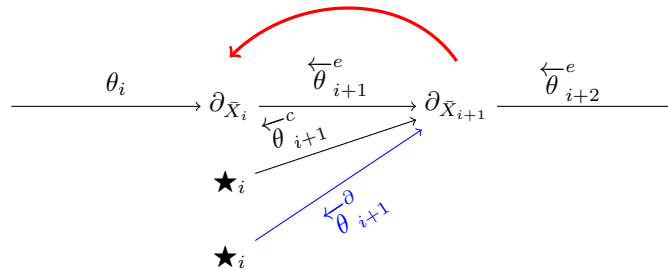

Figure 2: The dynamics of the transfer of derivatives formula.

As explained with the transfer of derivatives for the last time interval, we see that the derivative $\partial_{\bar{X}_{i+1}}$ is transferred (this is the meaning of the left pointing red arrow in the above figure) to become the derivative $\partial_{\bar{X}_i}$, changing the weight $\bar{\theta}_{i+1}$ into $\overset{\leftarrow}{\theta}{}^{e}_{i+1}$ (e for exchange), but at the expense of creating extra terms denoted by $\overset{\leftarrow}{\theta}{}^{c}_{i+1}$ (c for correction) and a boundary term $\overset{\leftarrow}{\theta}{}^{\partial}_{i+1}$ coming from the boundary (denoted by ∂ in what follows). All these weights satisfy the same moment estimates as (3.4).

In the above figure, we can see the structure of the transfer of derivatives will generate two branches of a tree at each time Lemma 4.1 is applied. The blue arrow means that the corresponding term generated by the transfer of derivative is a degenerate boundary term that needs to be addressed because the weight r.v. $\overset{\leftarrow}{\theta}{}^{\partial}_{i+1}$ is multiplied by a Dirac delta distribution. This will be solved using the second ingredient to be introduced in the next section: *the boundary merging lemmas*. The term $\bigstar_i$ standing for the product $\bigstar_i := \prod_{\ell=1}^{i} \mathbb{1}_{D_{\ell,n}} \bar{\theta}_\ell$ means that no more transfer of derivatives is required for these two terms because the derivative operator no longer appears on the test function f for the terms associated to $\overset{\leftarrow}{\theta}{}^{c}_{i+1}$ and $\overset{\leftarrow}{\theta}{}^{\partial}_{i+1}$ in (4.1).

We remark that the last term in (4.1) is singular but well defined if the time variables are fixed which is the case here as there is a conditioning with respect to the Poisson process expressed with the conditional expectations $\mathbb{E}_{i,n}$ in (4.1). Since the boundary weight $\overset{\leftarrow}{\theta}{}^{\partial}_{i+1}$ is multiplied by $\delta_L(\bar{X}_{i+1})$ the time degeneracy estimate for the product $\delta_L(\bar{X}_{i+1}) \overset{\leftarrow}{\theta}{}^{\partial}_{i+1}$ deteriorates so that the resulting probabilistic representation will not be exploitable without the extra ingredient of the boundary merging lemmas. A further study is thus required before proving moment estimates. We will discuss this matter in detail in the next section.

4.3 The boundary merging lemmas

The second ingredient that we need in order to establish our first IBP formula is the boundary merging lemma. Though the sequence of weights $\{\overset{\leftarrow}{\theta}{}^{\partial}_{i}, i \in \mathbb{N}_{N_T}\}$ satisfies the time degeneracy estimate stated in Lemma 4.1, which is similar to the one satisfied by the sequence $\{\bar{\theta}_i, i \in \mathbb{N}_{N_T}\}$ in (3.4), the fact that it is multiplied by $\delta_L(\bar{X}_i)$ increases its time degeneracy by a factor of $(\zeta_i - \zeta_{i-1})^{-1/2}$.

The key idea in order to circumvent this issue consists in using the time randomness provided by the Poisson process in order to smooth this singularity. This will eventually allow us to retrieve a time degeneracy estimate similar to (3.4). We call this ingredient *the boundary merging lemmas*.

To be more specific, in order to circumvent the time degeneracy problem related to the boundary term on the right hand side of (4.1), one needs to consider two successive random time intervals, say $[\zeta_{i-1}, \zeta_i]$ and $[\zeta_i, \zeta_{i+1}]$, and take expectations with respect to the intermediate time ζ_i in order to regularize the term $\delta_L(\bar{X}_i)$ that appears when one transfers the derivative using (4.1) on the time interval $[\zeta_{i-1}, \zeta_i]$. This operation will remove the time singularity induced by the Dirac delta distribution but will also reduce the number of jumps by one unit.

An important remark is that the boundary merging lemmas are not needed in the case that there are no jumps in the interval $[0, T]$, that is, on the set $\{N_T = 0\}$, because of the condition $f(L) = 0$.

From the application of the boundary merging lemmas, a new Markov chain structure appears which we call the *merged boundary (one-step) Markov chain* $\bar{X}^\partial$ whose transition on the time interval $[\zeta_j, \zeta_{i+1}]$, for $j = i - 1, i$, is given by

$$\bar{X}^\partial_{j,i+1} \equiv \bar{X}^\partial_{j,i+1}(\bar{X}_j) := L\left(1 - \mu(\bar{X}_j)\right) + \bar{X}_j\mu(\bar{X}_j) + \sigma(L)Z_{j,i+1}, \quad Z_{j,i+1} := \sum_{k=j+1}^{i+1} Z_k,$$

(4.6)

with $\mu(x) := -\sigma(L)/\sigma(x)$.

In the above one-step dynamics and in what follows, the index $j = i - 1$ will be used to indicate the boundary merging of the underlying Markov chain $\bar{X}$ on the two consecutive time intervals $[\zeta_{i-1}, \zeta_i]$ and $[\zeta_i, \zeta_{i+1}]$ while the index $j = i$ will be used in a second stage once the reduction of jumps (a.k.a. *time merging*) is performed in Section 5.1. As the proof of time degeneration estimates are similar and in order to simplify the presentation, we will deal with both cases at the same time in the next two Lemmas.

In the same way as it was done in Section 2.4, we define the associated space of smooth r.v.'s.

Definition 4.2. *For $i \in \bar{\mathbb{N}}_n$ and $j \in \{i-1, i\}$, we define the set $\mathbb{S}_{j,i+1,n}(\bar{X}^\partial)$ as the subset of r.v.'s $H \in \mathbb{L}^0$ such that there exists a measurable functions $h : \mathbb{R}^2 \times \{0, 1\} \times A_{i+1-j} \to \mathbb{R}$ satisfying*

1. *the random variable H can be written as $H = h(\bar{X}_j, \bar{X}^\partial_{j,i+1}, \rho_{i+1}, \zeta_j, \zeta_i, \zeta_{i+1})$ on the set $\{N_T = n\}$.*

2. *For any $r \in \{0, 1\}$ and any $s \in A_{i+1-j}$, one has $h(\cdot, \cdot, r, s) \in \mathscr{C}_p^\infty(\mathbb{R}^2)$.*

Similarly to the operators $\mathcal{D}$ and $\mathcal{I}$, defined in Section 2.4, related to the Markov chain $\bar{X}$, we introduce the operators $\bar{\mathcal{D}}$ and $\bar{\mathcal{I}}$ associated to the merged boundary process $\bar{X}^\partial$ defined above. For any (smooth) r.v. $H \in \mathbb{S}_{j,i+1,n}(\bar{X}^\partial)$ and $\ell \geq 1$, we let

$$\bar{\mathcal{I}}_{j,i+1}(H) := H\frac{Z_{j,i+1}}{\sigma(L)(\zeta_{i+1} - \zeta_j)} - \bar{\mathcal{D}}_{i+1}H, \quad \bar{\mathcal{I}}^{\ell+1}_{j,i+1}(H) := \bar{\mathcal{I}}_{j,i+1}(\bar{\mathcal{I}}_{j,i+1}(H))$$

$$\bar{\mathcal{D}}_{i+1}H := \partial_2 h(\bar{X}_j, \bar{X}^\partial_{j,i+1}, \rho_{i+1}, \zeta_j, \zeta_i, \zeta_{i+1}), \quad \bar{\mathcal{D}}^{\ell+1}_{i+1}H := \bar{\mathcal{D}}^{\ell}_{i+1}(\bar{\mathcal{D}}_{i+1}H).$$

Observe that since $Z_{j,i+1} = \sigma^{-1}(L)(\bar{X}^\partial_{j,i+1} - (L\left(1 - \mu(\bar{X}_j)\right) + \bar{X}_j\mu(\bar{X}_j))) \in \mathbb{S}_{j,i+1,n}(\bar{X}^\partial)$, it is clear that the r.v.'s $\bar{\mathcal{I}}_{j,i+1}(1)$ and $\bar{\mathcal{I}}^2_{j,i+1}(1)$ also belong to $\mathbb{S}_{j,i+1,n}(\bar{X}^\partial)$ so that they are explicit functions of the variables $\bar{X}_j$, $\bar{X}^\partial_{j,i+1}$, ρ_{i+1}, ζ_j, ζ_i and ζ_{i+1}.

In the following result, we will also use new weights $\overleftarrow{\theta}^{\partial*e}_{j,i+1}$, for $j = i - 1, i$, obtained by a merging procedure of the two weights $\overleftarrow{\theta}^{\partial}_i$ and $\overleftarrow{\theta}^e_{i+1}$ on the boundary set $\{\bar{X}_i = L\}$

$$\overleftarrow{\theta}^{\partial*e}_{j,i+1} := 4\lambda^{-1}\frac{a'(L) - b(L)}{a_j}\left(\bar{\mathcal{I}}^2_{j,i+1}(\bar{d}^{i+1}_2) + \bar{\mathcal{I}}_{j,i+1}(\bar{d}^{i+1}_1)\right)$$

with

$$\bar{d}_1^{i+1} := (b - a')(\bar{X}_{j,i+1}^{\partial}), \quad \bar{d}_2^{i+1} := \frac{1}{2}(a(\bar{X}_{j,i+1}^{\partial}) - a(L)) \tag{4.7}$$

for $i = 0, \cdots, n-1$. For the last interval, on $\{N_T = n\}$, we have $\overleftarrow{\theta}_{j,n+1}^{\partial *e} := 4e^{\lambda T} \frac{a'(L) - b(L)}{a_j}$, for $j = n-1, n$.

The cases $j = i - 1$ for (4.7) and $j = n - 1$ for $\overleftarrow{\theta}_{n-1,n+1}^{\partial *e}$ are used in the following result where we perform the boundary merging procedure. The remaining cases $j = i$ for $\overleftarrow{\theta}_{i,i+1}^{\partial *e}$ and $j = n$ for $\overleftarrow{\theta}_{n,n+1}^{\partial *e}$ will be used once time merging is performed in Section 5. For this reason, some properties that we will employ later appear in the next lemma for these cases.

The definition of the time degeneracy estimate in the sense of Definition 3.3 is naturally extended to this case if we define

$$D_{j,i+1,n}^{\partial} := \{\bar{X}_{j,i+1}^{\partial} \geq L, N_T = n\}, \quad j = i - 1, i$$

and replace (3.5) for $H \in \mathbb{S}_{j,i+1,n}(\bar{X}^{\partial})$ by

$$\forall p \geq 1, \quad \mathbb{1}_{D_{j,n}} \mathbb{E}\left[\mathbb{1}_{D_{j,i+1,n}^{\partial}} |H|^p \Big| \mathcal{G}_j, \zeta_{i+1}, N_T = n\right] \leq C(\zeta_{i+1} - \zeta_j)^{-\frac{p}{2}}, \ i \in \mathbb{N}_{n-1}, \ j = i - 1, i \tag{4.8}$$

and $\mathbb{1}_{D_{j,n}} \mathbb{E}\left[\mathbb{1}_{D_{j,n+1,n}^{\partial}} |H|^p \Big| \mathcal{G}_j, N_T = n\right] \leq C$ for $j = n - 1, n$.

We are now in position to provide the second ingredient in order to establish our first IBP formula, namely the boundary merging lemma. Its proof is postponed to Section 10.3.2.

Lemma 4.3. Let $f \in \mathscr{C}_p^0(\mathbb{R})$ and $n \in \bar{\mathbb{N}}$. The following property is satisfied for any $i \in \mathbb{N}_{n-1}$

$$\mathbb{E}\big[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}} \overleftarrow{\theta}_{i+1}^e \delta_L(\bar{X}_i) \overleftarrow{\theta}_i^{\partial} \,|\, \mathcal{G}_{i-1}, \zeta_{i+1}, N_T = n\big] \tag{4.9}$$

$$= \frac{\lambda^{-1}}{\zeta_{i+1} - \zeta_{i-1}} \mathbb{E}\big[f(\bar{X}_{i-1,i+1}^{\partial})\mathbb{1}_{D_{i-1,i+1,n}^{\partial}} \overleftarrow{\theta}_{i-1,i+1}^{\partial *e} \,|\, \mathcal{G}_{i-1}, \zeta_{i+1}, N_T = n\big].$$

Similarly, for the last time interval, one has

$$\mathbb{E}\big[f(\bar{X}_{n+1})\mathbb{1}_{D_{n+1,n}} \overleftarrow{\theta}_{n+1}^e \delta_L(\bar{X}_n) \overleftarrow{\theta}_n^{\partial} \,|\, \mathcal{G}_{n-1}, N_T = n\big] \tag{4.10}$$

$$= \frac{\lambda^{-1}}{T - \zeta_{n-1}} \mathbb{E}\big[f(\bar{X}_{n-1,n+1}^{\partial})\mathbb{1}_{D_{n-1,n+1,n}^{\partial}} \overleftarrow{\theta}_{n-1,n+1}^{\partial *e} \,|\, \mathcal{G}_{n-1}, N_T = n\big].$$

Here, $|\overleftarrow{\theta}_{j,n+1}^{\partial *e}| \leq C$ a.s. for $j = n - 1, n$. Moreover, with the above definitions $\overleftarrow{\theta}_{j,i+1}^{\partial *e} \in \mathbb{S}_{j,i+1,n}(\bar{X}^{\partial})$, $j = i - 1, i$ and $\overleftarrow{\theta}_{j,i+1}^{\partial *e}$ satisfies the time degeneracy estimates in the sense of (4.8).

The above lemma can be illustrated using Figure 3 below. The boundary weight r.v. $\overleftarrow{\theta}_i^{\partial}$ marked in blue is merged together with the weight $\overleftarrow{\theta}_{i+1}^e$ appearing in the next time interval, by taking expectations with respect to ζ_i in (4.9). This operation leads to a new transition on the time interval $[\zeta_{i-1}, \zeta_{i+1}]$ for the underlying Markov chain. More precisely, the new r.v. $\bar{X}_{i-1,i+1}^{\partial}$, and the new weight, $\overleftarrow{\theta}_{i-1,i+1}^{\partial *e}$, both marked in red in the above figure, will replace the symbols in blue: $\overleftarrow{\theta}_i^{\partial}$ and $\overleftarrow{\theta}_{i+1}^e$. In this sense, we call $\overleftarrow{\theta}_{i-1,i+1}^{\partial *e}$ the boundary merging of weights $\overleftarrow{\theta}_i^{\partial}$ and $\overleftarrow{\theta}_{i+1}^e$.

Remark that for the branch corresponding to the pair $(\overleftarrow{\theta}_i^e, \overleftarrow{\theta}_{i+1}^e)$, no merging procedure is required so that the respective Markov chain elements $\bar{X}_i$ and $\bar{X}_{i+1}$ remain unchanged. For this reason, we have chosen to leave the weights $\overleftarrow{\theta}_i^e$ and $\overleftarrow{\theta}_{i+1}^e$ in black on the left hand side.

On the right hand side of Figure 3, we see the result of the boundary merging procedure. Clearly the transition for the merged term has changed from $\bar{X}_{i-1}$ to $\bar{X}_{i-1,i+1}^\partial$ and then, for the next time interval $[\zeta_{i+1}, \zeta_{i+2}]$, the transition of the underlying Markov chain remains unchanged and is given by

$$Y_{i+2} := \bar{X}_{i+2}(i+1, \bar{X}_{i-1,i+1}^\partial)$$
$$= \rho_{i+2} \bar{X}_{i-1,i+1}^\partial + (1 - \rho_{i+2})(2L - \bar{X}_{i-1,i+1}^\partial) + \sigma(\bar{X}_{i-1,i+1}^\partial)Z_{i+2},$$

and where $\overleftarrow{\beta}_{i+2}^e := \overleftarrow{\theta}_{i+2}^e(\bar{X}_{i-1,i+1}^\partial, Y_{i+2}, \rho_{i+2}, \zeta_{i+1}, \zeta_{i+2})$ stands for the same weight formula as $\overleftarrow{\theta}_{i+2}^e$ but whose starting point is $\bar{X}_{i-1,i+1}^\partial$ and its end point is Y_{i+2}. We remind the reader that we freely use function notation for smooth r.v.'s on the spaces $\mathbb{S}_{i+2,n}(\bar{X})$ as in (2.6).
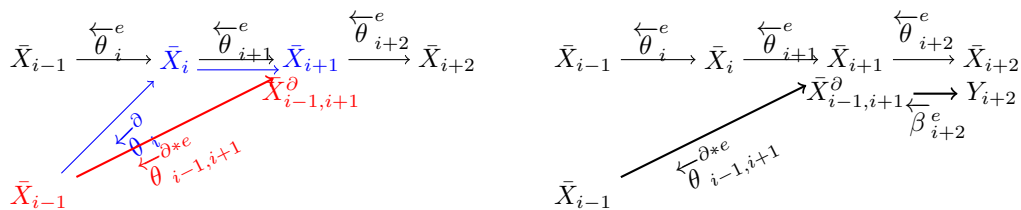

Figure 3: Markov chain structure before merging of boundary terms on the left and after merging on the right.

So far, we have explained how to transfer the derivatives and how to deal with boundary terms. The last step consists in performing a local IBP formula on a fixed time interval, say $[\zeta_i, \zeta_{i+1}]$. This operation will involve the integral operator applied to corresponding weight, namely $\mathcal{I}_i(\mathbb{1}_{D_{i,n}}\bar{\theta}_i)$, and thus will inevitably increase the time singularity in the estimate as stated in (3.5). Moreover, by the extraction formula, the Malliavin derivative of $\mathbb{1}_{D_{i,n}}$ will also generate a boundary term that has to be carefully treated by a merging procedure. This is the purpose of the following second boundary merging Lemma that we now describe.

The proof of the following result is postponed to Section 10.3.3. As the resulting boundary merging weight is different, the notation for the boundary merged weight changes from $\partial * e$ to $\partial \circledast e$. The merged boundary process does not change.

Lemma 4.4. Let $f \in \mathscr{C}_p^0(\mathbb{R})$ and $n \in \bar{\mathbb{N}}$. The following property is satisfied for any $i \in \mathbb{N}_{n-1}$

$$\mathbb{E}\left[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\overleftarrow{\theta}_{i+1}^e \delta_L(\bar{X}_i)(\zeta_i - \zeta_{i-1})\bar{\theta}_i \,|\, \mathcal{G}_{i-1}, \zeta_{i+1}, N_T = n\right] \tag{4.11}$$

$$= \frac{\lambda^{-1}}{\zeta_{i+1} - \zeta_{i-1}}\mathbb{E}\left[f(\bar{X}_{i-1,i+1}^\partial)\mathbb{1}_{D_{i-1,i+1,n}^\partial}\overleftarrow{\theta}_{i-1,i+1}^{\partial \circledast e} \,|\, \mathcal{G}_{i-1}, \zeta_{i+1}, N_T = n\right]$$

where we define the boundary merged weight $\overleftarrow{\theta}_{j,i+1}^{\partial \circledast e}$ for $j = i-1, i$ as

$$\overleftarrow{\theta}_{j,i+1}^{\partial \circledast e} := 4\lambda^{-1}\frac{a'(L) - b(L)}{a_j^{3/2}\sigma(L)}(\bar{X}_j - L)\left(\bar{\mathcal{I}}_{j,i+1}^2(\hat{d}_2^{i+1}) + \bar{\mathcal{I}}_{j,i+1}(\hat{d}_1^{i+1})\right)$$

with coefficients given by

$$\hat{d}_k^{i+1} := \bar{d}_k^{i+1} \times (\bar{\Phi}g^{-1})(a(L)(\zeta_{i+1} - \zeta_j), Z_{j,i+1}), \quad k = 1, 2,$$

$$\bar{\Phi}(t, z) := \int_{|z|}^{\infty} g(t, y) dy.$$

Similarly, for the last time interval, one has

$$\mathbb{E}\big[f(\bar{X}_{n+1}) \mathbb{1}_{D_{n+1,n}} \overset{\leftarrow e}{\theta}_{n+1} \delta_L(\bar{X}_n)(\zeta_n - \zeta_{n-1})\bar{\theta}_n \,|\, \mathcal{G}_{n-1}, N_T = n \big] \qquad (4.12)$$

$$= \frac{\lambda^{-1}}{T - \zeta_{n-1}} \mathbb{E}\big[f(\bar{X}_{n-1,n+1}^{\partial}) \mathbb{1}_{D_{n-1,n+1,n}^{\partial}} \overset{\leftarrow \partial \circledast e}{\theta}_{n-1,n+1} \,|\, \mathcal{G}_{n-1}, N_T = n \big]$$

where $\overset{\leftarrow \partial \circledast e}{\theta}_{j,n+1} := 4e^{\lambda T} \frac{2a'(L) - b(L)}{a_j^{3/2}\sigma(L)}(\bar{X}_j - L)\hat{d}^{n+1}$, $\hat{d}^{n+1} := (\bar{\Phi}g^{-1})(a(L)(\zeta_{n+1} - \zeta_j), Z_{j,n+1})$ *so that* $|\overset{\leftarrow \partial \circledast e}{\theta}_{j,n+1}| \le C$ *a.s., for* $j = n - 1, n$.

Moreover, for any $i \in \bar{\mathbb{N}}_n$, *for any* $j = i - 1, i$, $\overset{\leftarrow \partial \circledast e}{\theta}_{j,i+1} \in \mathbb{S}_{j,i+1,n}(\bar{X}^{\partial})$ *and it satisfies the time degeneracy estimates in the sense of* (4.8).

Remark 4.5. (i) We again emphasize the role of the two indexes $j = i - 1, i$, for $i \in \mathbb{N}_n$, in the definition of the boundary merged weights $\overset{\leftarrow \partial * e}{\theta}_{j,i+1}$ and $\overset{\leftarrow \partial \circledast e}{\theta}_{j,i+1}$ of Lemmas 4.3 and 4.4. The index $j = i - 1$ is used for the boundary merging operation while the index $j = i$ will be used once the reduction of jumps operation (a.k.a. time merging) is performed in Section 5.1.

(ii) The reason why the terms $\bar{\Phi}g^{-1}$ appear in the definitions of the coefficients $\hat{d}^{i+1}$, contrary to the previous Lemma 4.3, is due to the time increments $\zeta_i - \zeta_{i-1}$ and $\zeta_n - \zeta_{n-1}$ on the left hand side of equalities (4.11) and (4.12).

(iii) An important technical remark which follows from the definitions of the merged weights $\overset{\leftarrow \partial * e}{\theta}_{j,i+1}$ and $\overset{\leftarrow \partial \circledast e}{\theta}_{j,i+1}$, $j = i - 1, i$ in Lemmas 4.3 and 4.4 is that they do not depend neither on ρ_i or ρ_{i+1}.

5 A first IBP formula: putting the ingredients to work

As explained at the beginning of Section 4, in order to obtain an IBP formula, one first has to take the conditional expectation w.r.t the Poisson process N inside $\mathbb{E}\left[\partial_{\bar{X}_{N_T+1}} f(\bar{X}_{N_T+1}) \prod_{i=1}^{N_T+1} \mathbb{1}_{D_{i,N_T}} \bar{\theta}_i \right]$. Once the jump times are fixed, one observes that there is a Markov chain structure to which one may apply the transfer of derivatives procedure described in Lemma 4.1. This leads to a tree structure of terms which are combined up to the time interval where we decide to stop the transfer of derivatives and to perform a local IBP formula using only the noise on this specific interval.

Throughout this section we will often denote this interval by the general index k. For example, the tree structure is illustrated in Figure 4 where the IBP is performed on the time interval $[\zeta_{k-1}, \zeta_k]$ for $k = 2$. The application of the IBP formula (2.9) on the time interval $[\zeta_{k-1}, \zeta_k]$ will give after using the extraction formula (2.10) the new weight

$$\mathcal{I}_k(\mathbb{1}_{D_{k,n}} \bar{\theta}_k) = \delta_L(\bar{X}_k)\bar{\theta}_k + \mathbb{1}_{D_{k,n}} \mathcal{I}_k(\bar{\theta}_k).$$

We thus see that a boundary term appears which is treated by the boundary merging procedure described in Lemma 4.4 and thus gives rise to the new weight associated to the symbol $\partial \circledast e$.

The terms which contain boundary weights due to the successive application of the transfer of derivatives formula (4.1) are treated by the boundary merging procedure of

Lemma 4.3. Figure 4 shows these terms before the boundary merging procedure in a simplified algebraic notation that will be introduced in this section.

This boundary merging operation will lead to the recovery of an integrable time degeneracy estimate and to the reduction of one jump unit (after an application of the so-called time merging procedure) in the underlying Poisson process. Finally, a local IBP formula is performed on the interval $[\zeta_{k-1}, \zeta_k]$ where we have stopped the transfer of derivatives. As described above, after using the extraction formula, the new weight given by $\mathcal{I}_k(\mathbb{1}_{D_{k,n}}\bar{\theta}_k)$ will also generate a boundary term to which we apply the boundary merging procedure described by Lemma 4.4. As before, this will in turn reduce the number of jump times by one unit.

We importantly note that when one stops the (backward) transfer of derivatives procedure and decide to perform a local IBP formula on the time interval $[\zeta_{k-1}, \zeta_k]$, the weights $\mathbb{1}_{D_{j,n}}\bar{\theta}_j$ for the preceding time intervals, namely $[\zeta_{j-1}, \zeta_j]$, for $j = 1, \cdots, k-1$, correspond to the original probabilistic representation in Theorem 3.1 and thus remain unchanged. A similar remark applies when a time merging of weights or a correction weight appears on one branch of the tree structure below.

The overall tree diagram before carrying out the boundary merging procedure can be schematically described as follows in the case of $N_T = 4$ jumps of the Poisson process
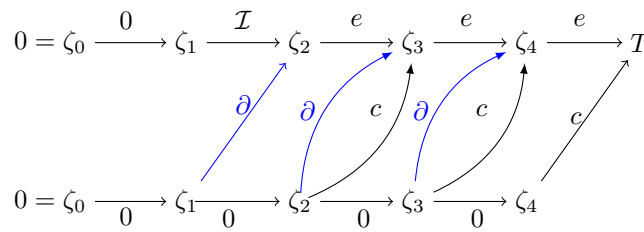

Figure 4: A tree in the case of $N_T = 4$ jump times with an IBP performed on the time interval $[\zeta_1, \zeta_2]$ before any merging.

An important remark in order to understand the tree notation to be introduced in the next section is that the blue curved arrows will lead to the application of the boundary merging Lemma 4.3. Therefore, as the number of jumps will be reduced by the time merging procedure, those time merged weights will no longer be considered as weights associated to the set $\{N_T = 4\}$ but $\{N_T = 3\}$. A similar remark applies when applying Lemma 4.4 to the boundary term generated by $\mathcal{I}_k(\mathbb{1}_{D_{k,n}}\bar{\theta}_k)$ which is denoted by the straight blue arrow in Figure 4.

We believe that the above explanations given before the proof of our main result are important for the reader because they are necessary to understand the following section where the various symbols for the above tree structure are introduced.

5.1 The tree structure in the IBP formula

In this section, we start by describing all the branches of the tree that are generated after performing transfer of derivatives and the two types of boundary merging previously described.

We denote by S_{n+1}, the set of all symbol sequences of length $n+1$ described by the following vectors of length $n+1$. More explicitly, when $n = 0$, we define $S_1 := \{I_1^1, \mathcal{B}_1^1\}$ with $I_1^1 := (\mathcal{I})$ and $\mathcal{B}_1^1 := (\partial \circledast e)$ and for $n \in \bar{\mathbb{N}}$,

$$S_{n+1} := \left(\bigcup_{1 \le k \le n+1} \{I_k^{n+1}, \mathcal{B}_k^{n+1}\} \right) \bigcup \left(\bigcup_{2 \le k \le n+1} \{C_k^{n+1}, B_k^{n+1}\} \right).$$

Therefore in S_{n+1} there are in total $2(n+1) + 2n$ vectors. The above vectors give the description of a branch in the IBP formula and are defined as follows. On the one hand, we let $C_k^{n+1} := (0, ..., 0, c, e, ..., e)$ and $I_k^{n+1} := (0, ..., 0, \mathcal{I}, e, ..., e)$ where the component c or $\mathcal{I}$ appears in the k-th coordinate for $k = 2, ..., n+1$ and we allow $k = 1$ only for the vectors I_k^{n+1}. On the other hand, we denote by $B_k^{n+1} := (0, ..., 0, \partial * e, e, ..., e)$, $k = 2, ..., n+1$ and $\mathcal{B}_k^{n+1} := (0, ..., 0, \partial \circledast e, e, ..., e)$, $k = 1, ..., n+1$, where the element $\partial * e$ or $\partial \circledast e$ appears in the k-th coordinate for $k = 2, ..., n$ and $k = 1, ..., n$ respectively[3].

As mentioned previously, the index k corresponds to the time interval on which we perform the local IBP formula. In other words, if the symbol 0 appears at the j-th coordinate of a vector, for $1 \le j < k$, this means that the weight corresponding to that time interval, namely $[\zeta_{j-1}, \zeta_j]$, remains the same as given in the probabilistic representation of Theorem 3.1, i.e. $\mathbb{1}_{D_{j,n}} \bar{\theta}_j$. The symbol c appearing at the k-th coordinate of the vector C_k^{n+1} means that the new weight (associated to this vector and to the interval $[\zeta_{k-1}, \zeta_k]$) is $\mathbb{1}_{D_{k,n}} \overleftarrow{\theta}_k^c$. The symbol $\mathcal{I}$ appearing at the k-th coordinate of I_k^{n+1} stands for the IBP weight, i.e. $\mathbb{1}_{D_{k,n}} \mathcal{I}_k(\bar{\theta}_k)$. The symbol $\partial * e$ corresponds to the merging between exchange (denoted by e) and boundary (denoted by ∂) r.v.'s according to Lemma 4.3 while the symbol $\partial \circledast e$ corresponds to the same merging but occurring on the same interval as the one where we perform the IBP formula and is computed according to Lemma 4.4.

We remark here that in the last interval there is no boundary term because we always assume that the test function f vanishes at the boundary.

If we fix our attention on all the branches that are generated with the objective of carrying out the local IBP formula on the interval $[\zeta_{k-1}, \zeta_k]$, we find the two following subsets of S_{n+1} which correspond to the branches that finish with a correction weight $\overleftarrow{\theta}^c$ and the branches that finish with a time merging of two time intervals, before leaving unchanged the remaining weights on remaining intervals. For $k = 1, ..., n+1$, we define

$$\bar{S}_{n+1}^k := \bigcup_{k < j \le n+1} \{C_j^{n+1}\}, \quad \dot{S}_{n+1}^k := \{\mathcal{B}_k^{n+1}\} \bigcup \left(\bigcup_{k < j \le n+1} \{B_j^{n+1}\} \right)$$

and in the special case $n = 0$, we let $\dot{S}_1^1 := \{\mathcal{B}_1^1\}$.

For example, the following figure describes the situation in the case when originally there were $n = 4$ jump times. Observe that, on the one hand, due to the merging procedure, all sequences of arrows starting at time 0 and finishing at time T which contain time merged weights are thus considered on the set $\dot{S}_4^2$, i.e. for $n = 3$. On the other hand, the sequences of arrows that contain a correction term c are part of the set $\dot{S}_5^2$.

5.2 The Markov chain and weights associated to the IBP tree branches

For the corresponding time partition $\pi := \{0 = \zeta_0 < ... < \zeta_{n+1} = T\}$ of the underlying Poisson process on the set $\{N_T = n\}$, we now need to define the underlying Markov chain $\bar{X}^{\mathbf{s}}$, for $\mathbf{s} \in S_{n+1}$ or $\mathbf{s} \in \dot{S}_{n+1}^k$, that will be used in the probabilistic representation of our first IBP formula.

It will be defined as the same process as the original Markov chain $\bar{X}$ except that on a certain time interval it may use the one-step transition of the merged boundary Markov chain $\bar{X}^\partial$ given by (4.6). To be more specific, in the case that $\mathbf{s} = B_k^{n+1}$ or $\mathcal{B}_k^{n+1}$,

---

[3]In order to keep index notation short, so as not to have to always consider two cases when k takes as lowest value 1 or 2, we include C_1^{n+1} and B_1^{n+1} as symbols but any statement in this case should be taken as an empty statement or that the symbol corresponds to the zero (empty) element.
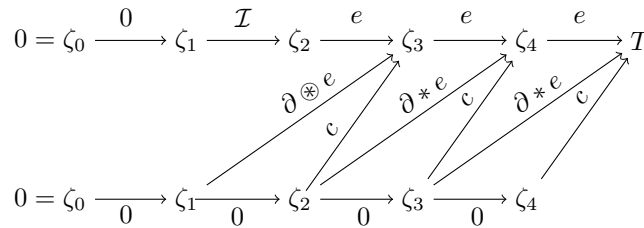
Figure 5: A tree in the case of $N_T = 4$ jump times with an IBP on the interval $[\zeta_1, \zeta_2]$ after time merging. All branches that contain a merged terms are associated to the set $\{N_T = 3\}$.

$k = 1, ..., n+1$, the Markov chain $\bar{X}^{\mathbf{s}}$ is defined by

$$\bar{X}_j^{\mathbf{s}} := \begin{cases} \bar{X}_j, & \text{if } 0 \leq j \leq k-1, \\ \bar{X}_{k-1,k}^{\partial}(\bar{X}_{k-1}) & \text{for } j = k, \\ \bar{X}_j(k, \bar{X}_{k-1,k}^{\partial}) & \text{for } j = k+1, ..., n+1. \end{cases}$$

Here $\bar{X}_{k-1,k}^{\partial}(\bar{X}_{k-1})$ stands for the one-step transition defined in (4.6). Similarly, for $j = k+1, \cdots, n+1$, $\bar{X}_j(k, \bar{X}_{k-1,k}^{\partial})$ is the flow notation for the scheme (2.4) taken at step j starting from the point $\bar{X}_{k-1,k}^{\partial}$ at step k. For any other $\mathbf{s} \in S_{n+1}$ of length $n+1$, the process $\bar{X}^{\mathbf{s}}$ corresponds to the original Markov chain dynamics, that is, we let $\bar{X}^{\mathbf{s}} \equiv \bar{X}$. We also define the associated set

$$D_{i,n}^{\mathbf{s}} := \{\bar{X}_i^{\mathbf{s}} \geq L, N_T = n\}.$$

We now introduce the weights corresponding to the Markov chain $\bar{X}^{\mathbf{s}}$ to be used in the IBP formula. For each $\mathbf{s} \in S_{n+1}$, we will define its associated weights $(\overleftarrow{\theta}_1^{\mathbf{s}}, ..., \overleftarrow{\theta}_{n+1}^{\mathbf{s}})$ and its product[4] as

$$\overleftarrow{\theta}^{\mathbf{s}} := \prod_{j=1}^{n+1} \overleftarrow{\theta}_j^{\mathbf{s}}.$$

We now define the weights for each element $\mathbf{s} = (s_1, ..., s_{n+1}) \in S_{n+1}$ as follows for $i \in \mathbb{N}_{n+1}$:

1. If $s_i = 0$, then $\overleftarrow{\theta}_i^{\mathbf{s}} := \mathbb{1}_{D_{i,n}^{\mathbf{s}}} \bar{\theta}_i(\bar{X}_{i-1}^{\mathbf{s}}, \bar{X}_i^{\mathbf{s}}, \rho_i, \zeta_{i-1}, \zeta_i) = \mathbb{1}_{D_{i,n}} \bar{\theta}_i(\bar{X}_{i-1}, \bar{X}_i, \rho_i, \zeta_{i-1}, \zeta_i)$.

2. If $s_i \in \{c, e\}$, with the notation introduced in (4.2)

$$\overleftarrow{\theta}_i^{\mathbf{s}} := \mathbb{1}_{D_{i,n}^{\mathbf{s}}} \overleftarrow{\theta}_i^{s_i}(\bar{X}_{i-1}^{\mathbf{s}}, \bar{X}_i^{\mathbf{s}}, \rho_i, \zeta_{i-1}, \zeta_i).$$

   As remarked at the beginning of Section 4.2, the weight $\overleftarrow{\theta}_i^{\mathbf{s}}$ is an explicit function of the underlying Markov chain and we will make use of such property in what follows.

3. Similarly, if $s_i \in \{\partial * e, \partial \circledast e, \mathcal{I}\}$, then using (4.2), Lemmas 4.3 and 4.4

$$\overleftarrow{\theta}_i^{\mathbf{s}} := \begin{cases} \overleftarrow{\theta}_{i-1,i}^{\partial * e}(\bar{X}_{i-1}^{\mathbf{s}}, \bar{X}_i^{\mathbf{s}}, \rho_i, \zeta_{i-1}, \zeta_i) \mathbb{1}_{D_{i,n}^{\mathbf{s}}}, & \text{if } s_i = \partial * e, \\ (\zeta_i - \zeta_{i-1})^{-1} \overleftarrow{\theta}_{i-1,i}^{\partial \circledast e}(\bar{X}_{i-1}^{\mathbf{s}}, \bar{X}_i^{\mathbf{s}}, \rho_i, \zeta_{i-1}, \zeta_i) \mathbb{1}_{D_{i,n}^{\mathbf{s}}}, & \text{if } s_i = \partial \circledast e, \\ \mathcal{I}_i(\bar{\theta}_i) \mathbb{1}_{D_{i,n}^{\mathbf{s}}}, & \text{if } s_i = \mathcal{I}. \end{cases}$$

---

[4]Recall the standard convention $\prod_\emptyset = 1$.

In the case of time merging, we importantly refer the reader to Remark 4.5. Due to this, we know that the reduction of one jump does not affect the sequence ρ_i, $i \in \mathbb{N}_{n+1}$ as the boundary merged weight does not depend on it.

5.3 The IBP formula

Theorem 5.1. *Let $f \in \mathscr{C}_b^1(\mathbb{R})$ such that $f(L) = 0$. Under assumption **(H)**, the following IBP formula is satisfied:*

$$T\mathbb{E}[f'(X_T)\mathbb{1}_{\{\tau \geq T\}}]$$

$$= \mathbb{E}\left[\sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1})\left\{f(\bar{X}_{N_T+1})\overleftarrow{\theta}^{I_k^{N_T+1}} + \sum_{\mathbf{s} \in \bar{S}_{N_T+1}^k \cup \dot{S}_{N_T+1}^k} f(\bar{X}_{N_T+1}^{\mathbf{s}})\overleftarrow{\theta}^{\mathbf{s}}\right\}\right].$$

Moreover, the r.v. appearing inside the expectation of the right-hand side of the above equality belongs to $\mathbb{L}^p(\mathbb{P})$ for $p \in [0, 2)$.

Proof. We start from the Markov chain representation given by Theorem 3.1, namely

$$\mathbb{E}[f'(X_T)\mathbb{1}_{\{\tau \geq T\}}] = \mathbb{E}[f'(\bar{X}_{N_T+1})\prod_{i=1}^{N_T+1}\mathbb{1}_{D_{i,N_T}}\bar{\theta}_i]$$

$$= \sum_{n \geq 0}\mathbb{E}\left[\mathbb{E}[f'(\bar{X}_{n+1})\prod_{i=1}^{n+1}\mathbb{1}_{D_{i,n}}\bar{\theta}_i|T^{n+1}]\mathbb{1}_{\{N_T=n\}}\right],$$

where we notice that $\{N_T = n\} = \{T_{n+1} > T\} \cap \{T_n \leq T\}$. We remind the reader that $\mathbb{E}_{i,n}[X]$ is the expectation of X conditional on $\{\mathcal{G}_i, T^{n+1}, \rho^{n+1}, N_T = n\}$ and therefore for the rest of the proof we will work on the set $\{N_T = n\}$.

Step 1: IBP on the last interval. We start by proving the IBP for the last time interval. That is, by the tower property of conditional expectation and the integration by parts formula, (2.9) on the (deterministic) time interval $[\zeta_n, T]$, noting that $f(L) = 0$, one has

$$\mathbb{E}[f'(\bar{X}_{n+1})\mathbb{1}_{D_{i,N_T}}\theta_{n+1}\prod_{i=1}^{n}\mathbb{1}_{D_{i,N_T}}\bar{\theta}_i\,|\,T^{n+1}]$$

$$= \mathbb{E}[\mathbb{E}_{n,n}[\mathcal{D}_{n+1}f(\bar{X}_{n+1})\mathbb{1}_{D_{n+1,n}}\theta_{n+1}]\prod_{i=1}^{n}\mathbb{1}_{D_{i,n}}\bar{\theta}_i\,|\,T^{n+1}]$$

$$= \mathbb{E}[f(\bar{X}_{n+1})\mathbb{1}_{D_{n+1,n}}\mathcal{I}_{n+1}(\bar{\theta}_{n+1})\prod_{i=1}^{n}\mathbb{1}_{D_{i,n}}\bar{\theta}_i\,|\,T^{n+1}]. \tag{5.1}$$

Step 2: The transfer of derivatives. In this step, we will perform the transfer of derivatives from the last time interval $[\zeta_{N_T}, T]$ to the time interval $[\zeta_{k-1}, \zeta_k]$. In order to carry this step, using the Markov property of the process $\bar{X}$, we define for $k \in \mathbb{N}_n$ the functions:

$$F_k(\bar{X}_k) := \mathbb{E}_{k,n}[f(\bar{X}_{n+1})\prod_{i=k+1}^{n+1}\mathbb{1}_{D_{i,n}}\overleftarrow{\theta}_i^e]$$

$$= \mathbb{E}[f(\bar{X}_{n+1})\prod_{i=k+1}^{n+1}\mathbb{1}_{D_{i,n}}\overleftarrow{\theta}_i^e|\bar{X}_k, T^{n+1}, \rho^{n+1}, N_T = n],$$

with the convention that $F_{n+1}(\bar{X}_{n+1}) = f(\bar{X}_{n+1})$ and $\prod_\emptyset \cdots = 1$. Note that the following recursive relation is satisfied for $k \in \mathbb{N}_n$

$$F_k(\bar{X}_k) = \mathbb{E}_{k,n}[F_{k+1}(\bar{X}_{k+1})\mathbb{1}_{D_{k+1,n}}\overleftarrow{\theta}_{k+1}^e]. \tag{5.2}$$

From the transfer of derivatives formula (4.5) of Lemma 4.1, we obtain

$$\mathbb{E}[\partial_{\bar{X}_{n+1}} f(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}] = \mathbb{E}[\partial_{\bar{X}_{n+1}} F_{n+1}(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}]$$

$$= \mathbb{E}[\partial_{\bar{X}_n} F_n(\bar{X}_n) \prod_{i=1}^{n} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}]$$

$$+ \mathbb{E}[f(\bar{X}_n) \mathbb{1}_{D_{n+1,n}} \overleftarrow{\theta}_{n+1}^c \prod_{i=1}^{n} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}].$$

Next, we proceed using a backward induction argument by combining successive applications of the transfer of derivative formula of Lemma 4.1 with the tower property of conditional expectation. To be more specific, from Lemma 4.1, one has for $k \in \mathbb{N}_n$

$$\mathbb{E}_{k,n}[\partial_{\bar{X}_{k+1}} F_{k+1}(\bar{X}_{k+1}) \mathbb{1}_{D_{k+1,n}} \bar{\theta}_{k+1}] = \partial_{\bar{X}_k} F_k(\bar{X}_k) + \mathbb{E}_{k,n}[F_{k+1}(\bar{X}_{k+1}) \mathbb{1}_{D_{k+1,n}} \overleftarrow{\theta}_{k+1}^c] \quad (5.3)$$

$$+ \mathbb{E}_{k,n}[F_{k+1}(\bar{X}_{k+1}) \delta_L(\bar{X}_{k+1}) \overleftarrow{\theta}_{k+1}^\partial]$$

which directly implies by iteration

$$\mathbb{E}[f'(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}] = \mathbb{E}[\partial_{\bar{X}_k} F_k(\bar{X}_k) \prod_{i=1}^{k} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}]$$

$$+ \sum_{j=k+1}^{n} \mathbb{E}[F_j(\bar{X}_j) \mathbb{1}_{D_{j,n}} \overleftarrow{\theta}_j^c \prod_{i=1}^{j-1} \theta_i | T^{n+1}]$$

$$+ \sum_{j=k+1}^{n} \mathbb{E}[F_j(\bar{X}_j) \delta_L(\bar{X}_j) \overleftarrow{\theta}_j^\partial \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}] \quad (5.4)$$

$$+ \mathbb{E}[f(\bar{X}_T) \mathbb{1}_{D_{n+1,n}} \overleftarrow{\theta}_{n+1}^c \prod_{i=1}^{n} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}].$$

for $k = 1, \ldots, n$, with the convention $\sum_\emptyset \cdots = 0$.

Step 3: The local IBP on the interval $[\zeta_{k-1}, \zeta_k]$. To proceed, we first notice that $\bar{\theta}_k \in \mathbb{S}_{k,n}(\bar{X})$, for $k \in \mathbb{N}_{n+1}$, is a smooth r.v.. Then from the tower property of conditional expectation (using $\mathbb{E}_{k-1,n}[\cdot]$), the integration by parts formula (2.9), the extraction formula (2.10) and (10.8) one obtains

$$\mathbb{E}[\partial_{\bar{X}_k} F_k(\bar{X}_k) \theta_k | \mathcal{G}_{k-1}, T^{n+1}]$$
$$= \mathbb{E}[F_k(\bar{X}_k) \mathbb{1}_{D_{k,n}} \mathcal{I}_k(\bar{\theta}_k) | \mathcal{G}_{k-1}, T^{n+1}] - \mathbb{E}[F_k(\bar{X}_k) \delta_L(\bar{X}_k) \bar{\theta}_k | \mathcal{G}_{k-1}, T^{n+1}],$$
$$= \mathbb{E}[F_k(\bar{X}_k) \mathbb{1}_{D_{k,n}} \mathcal{I}_k(\bar{\theta}_k) | \mathcal{G}_{k-1}, T^{n+1}] + \mathbb{E}[F_k(\bar{X}_k) \delta_L(\bar{X}_k) \overleftarrow{\theta}_k^\partial | \mathcal{G}_{k-1}, T^{n+1}].$$

By plugging the previous identity into (5.4) and using the recursive formula (5.2) yield for $k \in \mathbb{N}_n$:

$$\mathbb{E}[f'(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}]$$

$$= \mathbb{E}[F_k(\bar{X}_k) \mathbb{1}_{D_{k,n}} \mathcal{I}_k(\bar{\theta}_k) \prod_{i=1}^{k-1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}] + \sum_{j=k+1}^{n} \mathbb{E}[F_j(\bar{X}_j) \mathbb{1}_{D_{j,n}} \overleftarrow{\theta}_j^c \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}]$$

$$+ \sum_{j=k}^{n} \mathbb{E}[F_{j+1}(\bar{X}_{j+1}) \mathbb{1}_{D_{j+1,n}} \overleftarrow{\theta}_{j+1}^e \delta_L(\bar{X}_j) \overleftarrow{\theta}_j^\partial \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}] \quad (5.5)$$

$$+ \mathbb{E}[f(\bar{X}_{n+1}) \mathbb{1}_{D_{n+1,n}} \overleftarrow{\theta}_{n+1}^c \prod_{i=1}^{n} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}].$$

Step 4: Combining all the local IBP formulae. Multiplying (5.1) by $(T - \zeta_n)$, (5.5) by $(\zeta_k - \zeta_{k-1})$ and summing from $k = 1$ to n the resulting equalities, we obtain

$$T\mathbb{E}[f'(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i |\, T^{n+1}]$$

$$= \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) \mathbb{E}[\partial_{\bar{X}_{n+1}} f(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i |\, T^{n+1}],$$

$$= \mathbb{E}[f(\bar{X}_{n+1}) \mathbb{1}_{D_{n+1,n}} (T - \zeta_n) \mathcal{I}_{n+1}(\bar{\theta}_{n+1}) \prod_{i=1}^{n} \mathbb{1}_{D_{i,n}} \bar{\theta}_i |\, T^{n+1}]$$

$$+ \sum_{k=1}^{n} \mathbb{E}[f(\bar{X}_{n+1}) \prod_{i=k+1}^{n+1} \mathbb{1}_{D_{i,n}} \overleftarrow{\theta}_i^e \times \mathbb{1}_{D_{k,n}} (\zeta_k - \zeta_{k-1}) \mathcal{I}_k(\bar{\theta}_k) \times \prod_{i=1}^{k-1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i |\, T^{n+1}]$$

$$+ \sum_{k=1}^{n} (\zeta_k - \zeta_{k-1}) \sum_{j=k+1}^{n} \mathbb{E}[f(\bar{X}_{n+1}) \prod_{i=j+1}^{n+1} \mathbb{1}_{D_{i,n}} \overleftarrow{\theta}_i^e \times \mathbb{1}_{D_{j,n}} \overleftarrow{\theta}_j^c \times \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i |\, T^{n+1}]$$

$$+ \sum_{k=1}^{n} (\zeta_k - \zeta_{k-1}) \sum_{j=k}^{n} \mathbb{E}[f(\bar{X}_{n+1}) \prod_{i=j+1}^{n+1} \mathbb{1}_{D_{i,n}} \overleftarrow{\theta}_i^e \times \delta_L(\bar{X}_j) \overleftarrow{\theta}_j^\partial \times \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i |\, T^{n+1}]$$

$$+ \sum_{k=1}^{n} (\zeta_k - \zeta_{k-1}) \mathbb{E}[f(\bar{X}_{n+1}) \mathbb{1}_{D_{n+1,n}} \overleftarrow{\theta}_{n+1}^c \prod_{i=1}^{n} \mathbb{1}_{D_{i,n}} \bar{\theta}_i |\, T^{n+1}]$$

where we used the fact that $T - \zeta_n + \sum_{k=1}^{n} (\zeta_k - \zeta_{k-1}) = T - \zeta_0 = T$ in the first equality. The final argument starts by multiplying the above identity by $\mathbb{1}_{\{N_T = n\}}$ and taking the expectation of both hand sides with respect to the Poisson process.

Step 5: The merging procedure. Now, we perform the boundary and time merging procedures for the fourth term appearing in the right-hand side of the above equality. This is done by first conditioning with respect to $\mathcal{G}_{j-1}, \zeta_{j+1}, N_T = n$ in the inside sum and then by applying Lemma 4.3, for $j > k$ and Lemma 4.4. More specifically, for any $n \geq 1$, $k \in \mathbb{N}_n$ and $k < j \leq n$,

$$\mathbb{E}\Big[f(\bar{X}_{n+1}) \prod_{i=j+1}^{n+1} \mathbb{1}_{D_{i,n}} \overleftarrow{\theta}_i^e \times \delta_L(\bar{X}_j) \overleftarrow{\theta}_j^\partial |\mathcal{G}_{j-1}, \zeta_{j+1}, N_T = n\Big]$$

$$= \mathbb{E}\Big[F_{j+1}(\bar{X}_{j+1}) \mathbb{1}_{D_{j+1,n}} \overleftarrow{\theta}_{j+1}^e \times \delta_L(\bar{X}_j) \overleftarrow{\theta}_j^\partial |\mathcal{G}_{j-1}, \zeta_{j+1}, N_T = n\Big]$$

$$= \frac{\lambda^{-1}}{\zeta_{j+1} - \zeta_{j-1}} \mathbb{E}\Big[F_{j+1}(\bar{X}_{j-1,j+1}^\partial) \mathbb{1}_{D_{j-1,j+1,n}^\partial} \overleftarrow{\theta}_{j-1,j+1}^{\partial * e} |\mathcal{G}_{j-1}, \zeta_{j+1}, N_T = n\Big]$$

and for $j = k$,

$$\mathbb{E}\Big[f(\bar{X}_{n+1}) \prod_{i=k+1}^{n+1} \mathbb{1}_{D_{i,n}} \overleftarrow{\theta}_i^e \times \delta_L(\bar{X}_k)(\zeta_k - \zeta_{k-1}) \overleftarrow{\theta}_k^\partial |\mathcal{G}_{k-1}, \zeta_{k+1}, N_T = n\Big]$$

$$= \mathbb{E}\Big[F_{k+1}(\bar{X}_{k+1}) \mathbb{1}_{D_{k+1,n}} \overleftarrow{\theta}_{k+1}^e \times \delta_L(\bar{X}_k) \overleftarrow{\theta}_k^\partial |\mathcal{G}_{k-1}, \zeta_{k+1}, N_T = n\Big]$$

$$= \frac{\lambda^{-1}}{\zeta_{k+1} - \zeta_{k-1}} \mathbb{E}\Big[F_{k+1}(\bar{X}_{k-1,k+1}^\partial) \mathbb{1}_{D_{k-1,k+1,n}^\partial} \overleftarrow{\theta}_{k-1,k+1}^{\partial \circledast e} |\mathcal{G}_{k-1}, \zeta_{k+1}, N_T = n\Big]$$

so that

$$\mathbb{E}\Big[\sum_{k=1}^{n}(\zeta_k-\zeta_{k-1})\sum_{j=k}^{n}\mathbb{E}[f(\bar{X}_{n+1})\prod_{i=j+1}^{n+1}\mathbb{1}_{D_{i,n}}\overleftarrow{\theta}_i^{\,e}\times\delta_L(\bar{X}_j)\overleftarrow{\theta}_j^{\,\partial}\times\prod_{i=1}^{j-1}\mathbb{1}_{D_{i,n}}\bar{\theta}_i|\,T^{n+1}]\mathbb{1}_{\{N_T=n\}}\Big]$$

$$=\mathbb{E}\Big[\sum_{k=1}^{n}(\zeta_k-\zeta_{k-1})\times\sum_{j=k+1}^{n}F_{j+1}(\bar{X}_{j-1,j+1}^{\partial})\lambda^{-1}(\zeta_{j+1}-\zeta_{j-1})^{-1}$$

$$\times\mathbb{1}_{D_{j-1,j+1,n}^{\partial}}\overleftarrow{\theta}_{j-1,j+1}^{\,\partial *e}\times\prod_{i=1}^{j-1}\mathbb{1}_{D_{i,n}}\bar{\theta}_i\mathbb{1}_{\{N_T=n\}}\Big]$$

$$+\mathbb{E}\Big[\sum_{k=1}^{n}(\zeta_k-\zeta_{k-1})F_{k+1}(\bar{X}_{k-1,k+1}^{\partial})\lambda^{-1}(\zeta_{k+1}-\zeta_{k-1})^{-1}$$

$$\times\mathbb{1}_{D_{k-1,k+1,n}^{\partial}}\overleftarrow{\theta}_{k-1,k+1}^{\,\partial \circledast e}\times\prod_{i=1}^{k-1}\mathbb{1}_{D_{i,n}}\bar{\theta}_i\mathbb{1}_{\{N_T=n\}}\Big].$$

We then apply Lemma 10.8 to the above identity. In order to do it, we make use of the following decomposition

$$\mathbb{E}\Big[\sum_{k=1}^{n}(\zeta_k-\zeta_{k-1})\sum_{j=k}^{n}\mathbb{E}[f(\bar{X}_{n+1})\prod_{i=j+1}^{n+1}\mathbb{1}_{D_{i,n}}\overleftarrow{\theta}_i^{\,e}\times\delta_L(\bar{X}_j)\overleftarrow{\theta}_j^{\,\partial}\times\prod_{i=1}^{j-1}\mathbb{1}_{D_{i,n}}\bar{\theta}_i|\,T^{n+1}]\mathbb{1}_{\{N_T=n\}}\Big]$$

$$=\mathbb{E}\Big[\sum_{k=1}^{n}\sum_{j=k+1}^{n}(\zeta_{j+1}-\zeta_{j-1})^{-1}G_1(\zeta_1,\cdots,\zeta_{j-1},\zeta_{j+1},\cdots,\zeta_{N_T+1})\mathbb{1}_{\{N_T=n\}}\Big]$$

$$+\mathbb{E}\Big[\sum_{k=1}^{n}(\zeta_{k+1}-\zeta_{k-1})^{-1}G_2(\zeta_1,\cdots,\zeta_{k-1},\zeta_{k+1},\cdots,\zeta_{N_T+1})\mathbb{1}_{\{N_T=n\}}\Big]$$

where G_1 and G_2 are the measurable functions defined for $j\geq k+1$ by

$$G_1(\zeta_1,\cdots,\zeta_{j-1},\zeta_{j+1},\cdots,\zeta_{N_T+1})$$
$$:=\frac{(\zeta_k-\zeta_{k-1})}{\lambda}\mathbb{E}\Big[F_{j+1}(\bar{X}_{j-1,j+1}^{\partial})\mathbb{1}_{D_{j-1,j+1,n}^{\partial}}\overleftarrow{\theta}_{j-1,j+1}^{\,\partial *e}\times\prod_{i=1}^{j-1}\mathbb{1}_{D_{i,n}}\bar{\theta}_i|\,T^{n+1}\Big],$$

$$G_2(\zeta_1,\cdots,\zeta_{k-1},\zeta_{k+1},\cdots,\zeta_{N_T+1})$$
$$:=\frac{(\zeta_k-\zeta_{k-1})}{\lambda}\mathbb{E}\Big[F_{k+1}(\bar{X}_{k-1,k+1}^{\partial})\mathbb{1}_{D_{k-1,k+1,n}^{\partial}}\overleftarrow{\theta}_{k-1,k+1}^{\,\partial \circledast e}\times\prod_{i=1}^{k-1}\mathbb{1}_{D_{i,n}}\bar{\theta}_i|\,T^{n+1}\Big].$$

Now observe that the weights $\overleftarrow{\theta}_{j-1,j+1}^{\,\partial *e}$, $\overleftarrow{\theta}_{k-1,k+1}^{\,\partial \circledast e}$ satisfy the time degeneracy estimates in the sense of (4.8) and so do the weights $\overleftarrow{\theta}_i^{\,e}$, $\bar{\theta}_i$, so that from the tower property of conditional expectation and Lemma 10.1 with $p=1$, the following estimates hold

$$|G_1|\leq C^n\lambda^{-1}(\zeta_k-\zeta_{k-1})\prod_{i=1,i\neq j,j+1}^{n+1}(\zeta_i-\zeta_{i-1})^{-\frac{1}{2}}(\zeta_{j+1}-\zeta_{j-1})^{-\frac{1}{2}},$$

$$|G_2|\leq C^n\lambda^{-1}(\zeta_k-\zeta_{k-1})\prod_{i=1,i\neq k,k+1}^{n+1}(\zeta_i-\zeta_{i-1})^{-\frac{1}{2}}(\zeta_{k+1}-\zeta_{k-1})^{-\frac{1}{2}}$$

with the convention $\zeta_0=0$, $\zeta_{n+1}=T$ on $\{N_T=n\}$.

Using that conditional on the event $\{N_T = n, \zeta_1, \cdots, \zeta_{j-1}, \zeta_{j+1}, \cdots, \zeta_{N_T}\}$, the distribution of ζ_j is uniform on $[\zeta_{j-1}, \zeta_{j+1}]$, we get

$$
\mathbb{E}[(\zeta_{j+1} - \zeta_{j-1})^{-1}|G_1|\mathbb{1}_{\{N_T=n\}}]
$$

$$
\leq C^n \int_0^T \cdots \int_0^{s_2} (s_k - s_{k-1})(s_{j+1} - s_{j-1})^{-\frac{3}{2}} \prod_{i=1, \neq j, j+1}^{n+1} (s_i - s_{i-1})^{-\frac{1}{2}} \, ds_1 \cdots ds_n
$$

$$
\leq C^n \int_0^T \cdots \int_0^{s_2} (s_k - s_{k-1})(s_{j+1} - s_{j-1})^{-\frac{1}{2}}
$$

$$
\times \prod_{i=1, i \neq j, j+1}^{n+1} (s_i - s_{i-1})^{-\frac{1}{2}} \, ds_1 \cdots ds_{j-1} ds_{j+1} \cdots ds_n
$$

$$
< \infty.
$$

The same argument yields $\mathbb{E}[(\zeta_{k+1} - \zeta_{k-1})^{-1}|G_2|\mathbb{1}_{\{N_T=n\}}] < \infty$. From Lemma 10.8, we thus obtain

$$
\mathbb{E}\Big[(\zeta_k - \zeta_{k-1}) \sum_{j=k}^n f(\bar{X}_{n+1}) \prod_{i=j+1}^{n+1} \mathbb{1}_{D_{i,n}} \overleftarrow{\theta}_i^e \times \delta_L(\bar{X}_j) \overleftarrow{\theta}_j^\partial \times \prod_{i=1}^{j-1} \theta_i \mathbb{1}_{\{N_T=n\}}\Big]
$$

$$
= \mathbb{E}\Big[(\zeta_k - \zeta_{k-1}) \sum_{j=k+1}^n f(\bar{X}_n^{B_j^n}) \prod_{i=j+1}^n \mathbb{1}_{D_{i,n}^{B_j^n}} \overleftarrow{\theta}_i^{e, B_j^n}
$$

$$
\times \mathbb{1}_{D_{j,n}^{B_j^n}} \overleftarrow{\theta}_j^{\partial * e, B_j^n} \times \prod_{i=1}^{j-1} \theta_i \mathbb{1}_{\{N_T=n-1\}}\Big]
$$

$$
+ \mathbb{E}\Big[f(\bar{X}_n^{\mathcal{B}_k^n}) \prod_{i=k+1}^n \mathbb{1}_{D_{i,n}^{\mathcal{B}_k^n}} \overleftarrow{\theta}_i^{e, \mathcal{B}_k^n} \mathbb{1}_{D_{k,n}^{\mathcal{B}_k^n}} \overleftarrow{\theta}_k^{\partial \circledast e, \mathcal{B}_k^n} \prod_{i=1}^{k-1} \theta_i \mathbb{1}_{\{N_T=n-1\}}\Big].
$$

Step 6: Integrability properties. The proof of the $\mathbb{L}^p(\mathbb{P})$-integrability, $p \in [0, 2)$, follows from the time degeneracy estimates for each weight in the Lemmas 4.1, 4.3, 4.4 and 10.1 combined with a similar argument to the one employed at the end of the proof in Section 10 (see the discussion following Lemma 10.1). This also implies that the infinite sum over n converges absolutely and therefore after a re-ordering of the different terms one obtains the claimed formula. This concludes the proof. $\qquad \square$

Remark 5.2. (i) The right hand side of the IBP formula may alternatively be written in a longer but maybe more appealing format as

$$
T\mathbb{E}[f'(X_T)\mathbb{1}_{\{\tau \geq T\}}] = \mathbb{E}\left[f(\bar{X}_{N_T+1}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \left\{\overleftarrow{\theta}^{I_k^{N_T+1}} + \sum_{j=k+1}^{N_T+1} \overleftarrow{\theta}^{C_j^{N_T+1}}\right\}\right]
$$

$$
+ \mathbb{E}\left[\sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \sum_{j=k+1}^{N_T+1} f(\bar{X}_{N_T+1}^{B_j^{N_T+1}}) \overleftarrow{\theta}^{B_j^{N_T+1}}\right]
$$

$$
+ \mathbb{E}\left[\sum_{k=1}^{N_T+1} f(\bar{X}_{N_T+1}^{\mathcal{B}_k^{N_T+1}}) \overleftarrow{\theta}^{\mathcal{B}_k^{N_T+1}}\right].
$$

(ii) The restriction $f(L) = 0$, can be easily removed if one considers the test function $\tilde{f}(x) = f(x) - f(L)$ instead of f.

Note that, the previous theorem not only yields an IBP formula which is suitable for Monte Carlo simulation, it also provides a probabilistic representation for the derivative,

with respect to the terminal point, of the transition density of the killed process at time T. To be more specific, under assumption **(H)**, for any bounded measurable f defined on $[L, \infty)$, one deduces $\mathbb{E}[f(X_T)\mathbb{1}_{\{\tau > T\}}] = \int_L^\infty f(z)p(T, x, z)\, dz$ where $(0, \infty) \times [L, \infty)^2 \ni (T, x, z) \mapsto p(T, x, z)$ is the transition density of the killed process at time T starting from x at time 0. Moreover, from Theorem 5.1, by a standard approximation argument that we omit, one deduces that $z \mapsto p(T, x, z)$ is differentiable on $[L, \infty)$.[5] The next result provides a probabilistic representation of this derivative from which directly stems an unbiased Monte Carlo simulation method.

Corollary 5.3. *Under assumption* **(H)**, *the transition density of the killed process at time T is differentiable with respect to its terminal point. Moreover, for all $(T, x, z) \in (0, \infty) \times [L, \infty)^2$ the following probabilistic representation holds*

$$T\partial_z p(T, x, z) = \mathbb{E}\left[\sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1})g(a(\bar{X}_{N_T})(T - \zeta_{N_T}), z - \bar{X}_{N_T}) \overleftarrow{\theta}^{I_k^{N_T+1}} \right]$$

$$+ \mathbb{E}\left[\sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \sum_{\mathbf{s} \in \bar{S}_{N_T+1}^k \cup \dot{S}_{N_T+1}^k} p_{\mathbf{s}}(T - \zeta_{N_T}, \bar{X}_{N_T}^{\mathbf{s}}, z) \overleftarrow{\theta}^{\mathbf{s}} \right].$$

Here $p_{\mathbf{s}}$ denotes the transition density of the Markov chain $\bar{X}^{\mathbf{s}}$, that is, $z \mapsto p_{\mathbf{s}}(T - \zeta_{N_T}, x, z)$ is the density of the r.v. $\bar{X}_{N_T+1}^{\mathbf{s}}$ conditional on $\{\bar{X}_{N_T}^{\mathbf{s}} = x\}$. In particular, for $\mathbf{s} = B_{N_T+1}^{N_T+1}$ or $\mathcal{B}_{N_T+1}^{N_T+1}$, the dynamics (4.6) readily gives $p_{\mathbf{s}}(T - \zeta_{N_T}, \bar{X}_{N_T}^{\mathbf{s}}, z) = g(a(L)(\zeta_{N_T+1} - \zeta_{N_T}), z - (L(1 - \mu(\bar{X}_{N_T})) + \bar{X}_{N_T}\mu(\bar{X}_{N_T}))$ since $\bar{X}_{N_T}^{\mathbf{s}} = \bar{X}_{N_T}$.

6 Bismut-Elworthy-Li type formula

In this section, we briefly derive the IBP formula for $\partial_x \mathbb{E}[f(X_T)\mathbb{1}_{\{\tau \geq T\}}]$ commonly referred in the literature as the Bismut-Elworthy-Li formula. As we will see, obtaining this formula is simpler than the IBP formula derived in Theorem 5.1 as it does not involve boundary weights. Therefore the proof of this formula will not require any merging procedure contrary to the one of Theorem 5.1.

First, we give the transfer of derivatives lemma which is carried out forward in time in comparison with Lemma 4.1 where it is done backward in time. The proof is similar to the one of Lemma 4.1 given in Section 10.2 and therefore we omit it. Note that due to the change of time direction some changes of notation and sign occur among other changes in the formulae for the weights.

Lemma 6.1. *Let $f \in \mathscr{C}_p^1(\mathbb{R})$ and $n \in \bar{\mathbb{N}}$. Then, the following transfer of derivative formula holds for $i \in \bar{\mathbb{N}}_{n-1}$:*

$$\partial_{\bar{X}_i} \mathbb{E}_{i,n}[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}} \bar{\theta}_{i+1}] = \mathbb{E}_{i,n}[\partial_{\bar{X}_{i+1}} f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}} \overrightarrow{\theta}_{i+1}^{e}]$$

$$+ \mathbb{E}_{i,n}[f(\bar{X}_{i+1})\left(\mathbb{1}_{D_{i+1,n}} \overrightarrow{\theta}_{i+1}^{c} + \delta_L(\bar{X}_{i+1})\overrightarrow{\theta}_{i+1}^{\partial}\right)],$$

where the r.v.'s $(\overrightarrow{\theta}_{i+1}^{e}, \overrightarrow{\theta}_{i+1}^{c}, \overrightarrow{\theta}_{i+1}^{\partial}) \in \mathbb{S}_{i+1,n}(\bar{X})$ are defined by

$$\overrightarrow{\theta}_{i+1}^{e} := 2\lambda^{-1}\left(\mathcal{I}_{i+1}^2(d_2^{i+1}) + \mathcal{I}_{i+1}(d_1^{i+1})\right),$$

$$\overrightarrow{\theta}_{i+1}^{c} := \mathcal{I}_{i+1}\left((2\rho_{i+1} - 1)\bar{\theta}_{i+1} - \overrightarrow{\theta}_{i+1}^{e}\right) + \partial_{\bar{X}_i}\bar{\theta}_{i+1} + \sigma_i'\mathcal{I}_{i+1}\left(Z_{i+1}\bar{\theta}_{i+1}\right),$$

$$\overrightarrow{\theta}_{i+1}^{\partial} := \overrightarrow{\theta}_{i+1}^{e},$$

$$d_1^{i+1} := c_1^{i+1} + (2\rho_{i+1} - 1)\partial_{\bar{X}_i} c_2^{i+1},$$

$$d_2^{i+1} := c_2^{i+1}.$$

---

[5]We refer the interested reader to [14] for an alternative proof based on analytic arguments.

In a similar way, let $f \in \mathscr{C}_p^1(\mathbb{R})$ such that $f(L) = 0$. Then, the following transfer of derivative formula is satisfied

$$\partial_{\bar{X}_n} \mathbb{E}_{n,n}\big[f(\bar{X}_{n+1})\mathbb{1}_{D_{n+1,n}}\bar{\theta}_{n+1}\big] = \mathbb{E}_{n,n}\big[\partial_{\bar{X}_{n+1}} f(\bar{X}_{n+1})\mathbb{1}_{D_{n+1,n}} \overrightarrow{\theta}^e_{n+1}\big]$$

with $\overrightarrow{\theta}^e_{n+1} := 2e^{\lambda T}(1 + (2\rho_{n+1} - 1)\sigma'_n Z_{n+1})$ so that $\forall p \geq 1$, $\|\overrightarrow{\theta}^e_{n+1}\|_{p,n,n} \leq C$. We also set $\overrightarrow{\theta}^c_{n+1} := 0$ for notational convenience.

With the above definitions, for any $i \in \bar{\mathbb{N}}_{n-1}$, we have $\mathbb{E}_{i,n}[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\bar{\theta}_{i+1}]\big|_{\bar{X}_i = \cdot} \in \mathscr{C}_p^1(\mathbb{R})$ a.s. Moreover, the weights $\overrightarrow{\theta}^a_{i+1}$, $a \in \{e, c, \partial\}$, $i \in \bar{\mathbb{N}}_n$, satisfy the time degeneracy estimates.

As we did before the statement of Theorem 5.1, we need to define the weights that will be used in the BEL formula. It will be apparent in what follows that no boundary terms will appear so that no merging operation is needed. Similarly to Section 5.1, we consider the vectors of length $n + 1$ defined by $\widehat{C}_k^{n+1} = (e, \ldots, e, c, 0, \ldots, 0)$ where the symbol c appears in the k-th coordinate for $k \in \mathbb{N}_{n+1}$ and $\widehat{I}_k^{n+1} = (e, \ldots, e, \mathcal{I}, 0, \ldots, 0)$ where the symbol $\mathcal{I}$ appears in the k-th coordinate for $k \in \mathbb{N}_{n+1}$. Again in order to keep index notation short, we included $\widehat{C}_{n+1}^{n+1}$ as a symbol but any statement in this case should be taken as an empty statement or that the symbol corresponds to an empty element. We then define the set $\widehat{S}_{n+1}^k = \bigcup_{1 \leq j \leq k}\{\widehat{C}_j^{n+1}\}$ for $k \in \mathbb{N}_{n+1}$. Note that since there are no merging terms, the weights $\overrightarrow{\theta}^e_i$ and $\overrightarrow{\theta}^c_i$ are given by Lemma 6.1 applied on the set $\{N_T = n\}$. Again, we consider the mapping, which to a vector $\mathbf{s} \in \widehat{S}_{n+1}^k$, provides the associated product of weights denoted by $\overrightarrow{\theta}^{\mathbf{s}}$, namely for $i \in \mathbb{N}_{n+1}$

$$\overrightarrow{\theta}^{C_i^{n+1}} := \prod_{\ell=i+1}^{n+1} \theta_\ell \times \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^c_i \times \prod_{j=1}^{i-1} \mathbb{1}_{D_{j,n}} \overrightarrow{\theta}^e_j,$$

$$\overrightarrow{\theta}^{I_i^{n+1}} := \prod_{\ell=i+1}^{n+1} \theta_\ell \times \mathbb{1}_{D_{i,n}} \mathcal{I}_i(\overrightarrow{\theta}^e_i) \times \prod_{j=1}^{i-1} \mathbb{1}_{D_{j,n}} \overrightarrow{\theta}^e_j.$$

Theorem 6.2. *Let $f \in \mathscr{C}_b^1(\mathbb{R})$ such that $f(L) = 0$. Using the weights defined above, the following Bismut-Elworthy-Li formula is satisfied for any initial point $x \in [L, \infty)$:*

$$T\partial_x \mathbb{E}[f(X_T)\mathbb{1}_{\{\tau \geq T\}}] = \mathbb{E}\left[f(\bar{X}_{N_T+1}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1})\left\{\overrightarrow{\theta}^{I_k^{N_T+1}} + \sum_{\mathbf{s} \in \widehat{S}_{N_T+1}^k} \overrightarrow{\theta}^{\mathbf{s}}\right\}\right].$$

Moreover, the r.v. appearing inside the expectation in the right-hand side of the above equality belongs to $\mathbb{L}^p(\mathbb{P})$, for any $p \in [0, 2)$.

Proof. As in the proof of Theorem 5.1, we have

$$\mathbb{E}[f(X_T)\mathbb{1}_{\{\tau \geq T\}}] = \sum_{n \geq 0} \mathbb{E}\Big[\mathbb{E}[f(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \mathbb{1}_{D_{i,n}}\bar{\theta}_i | T^{n+1}]\, \mathbb{1}_{\{N_T = n\}}\Big].$$

In most of the arguments below, we will work on the set $\{N_T = n\}$. In order to perform a forward induction argument through the Markov chain structure, we define for $k \in \mathbb{N}_{n+1}$ the functions

$$\widehat{F}_k(\bar{X}_k) := \mathbb{E}_{k,n}\big[f(\bar{X}_{n+1}) \prod_{i=k+1}^{n+1} \mathbb{1}_{D_{i,n}}\bar{\theta}_i\big]$$

$$= \mathbb{E}\big[f(\bar{X}_{n+1}) \prod_{i=k+1}^{n+1} \mathbb{1}_{D_{i,n}}\bar{\theta}_i | \bar{X}_k, T^{n+1}, \rho^{n+1}, N_T = n\big].$$

We let $\widehat{F}_{n+1}(\bar{X}_{n+1}) := f(\bar{X}_{n+1})$ and the following recursive relation is satisfied for $k \in \mathbb{N}_n$

$$\widehat{F}_k(\bar{X}_k) = \mathbb{E}_{k,n}[\widehat{F}_{k+1}(\bar{X}_{k+1}) \mathbb{1}_{D_{k+1,n}} \theta_{k+1}]. \tag{6.1}$$

Then, iterating the transfer of derivative formula in Lemma 6.1 for $k \in \mathbb{N}_n$, we obtain[6]

$$\partial_x \mathbb{E}[f(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}] = \mathbb{E}[\mathcal{D}_k \widehat{F}_k(\bar{X}_k) \prod_{i=1}^{k} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}]$$

$$+ \sum_{j=1}^{k} \mathbb{E}[\widehat{F}_j(\bar{X}_j) \mathbb{1}_{D_{j,n}} \overrightarrow{\theta}^c_j \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}] \tag{6.2}$$

$$+ \sum_{j=1}^{k} \mathbb{E}[\widehat{F}_j(\bar{X}_j) \delta_L(\bar{X}_j) \overrightarrow{\theta}^\partial_j \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}].$$

To further simplify the first term on the right-hand side of the above equation, we use the tower property of conditional expectation, the integration by parts formula (2.9) and the extraction formula (2.10) to obtain

$$\mathbb{E}[\mathcal{D}_k \widehat{F}_k(\bar{X}_k) \mathbb{1}_{D_{k,n}} \overrightarrow{\theta}^e_k | \mathcal{G}_{k-1}, T^{n+1}] = \mathbb{E}[\widehat{F}_k(\bar{X}_k) \mathbb{1}_{D_{k,n}} \mathcal{I}_k(\overrightarrow{\theta}^e_k) | \mathcal{G}_{k-1}, T^{n+1}]$$

$$- \mathbb{E}[\widehat{F}_k(\bar{X}_k) \delta_L(\bar{X}_k) \overrightarrow{\theta}^e_k | \mathcal{G}_{k-1}, T^{n+1}].$$

In the case $k = n$, using the transfer of derivative formula of Lemma 6.1 on the last time interval and then performing the IBP formula (2.9), noting that $f(L) = 0$, we obtain the representation

$$\partial_x \mathbb{E}[f(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}] = \mathbb{E}[\mathcal{D}_{n+1} f(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}]$$

$$+ \sum_{j=1}^{n} \mathbb{E}[\widehat{F}_j(\bar{X}_j) \mathbb{1}_{D_{j,n}} \overrightarrow{\theta}^c_j \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}]$$

$$+ \sum_{j=1}^{n} \mathbb{E}[\widehat{F}_j(\bar{X}_j) \delta_L(\bar{X}_j) \overrightarrow{\theta}^\partial_j \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}].$$

$$= \mathbb{E}[f(\bar{X}_{n+1}) \mathbb{1}_{D_{n+1,n}} \mathcal{I}_{n+1}(\overrightarrow{\theta}^e_{n+1}) \prod_{i=1}^{n} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}]$$

$$+ \sum_{j=1}^{n+1} \mathbb{E}[\widehat{F}_j(\bar{X}_j) \mathbb{1}_{D_{j,n}} \overrightarrow{\theta}^c_j \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}] \tag{6.3}$$

$$+ \sum_{j=1}^{n} \mathbb{E}[\widehat{F}_j(\bar{X}_j) \delta_L(\bar{X}_j) \overrightarrow{\theta}^\partial_j \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}]$$

where we remind the reader that we previously set $\overrightarrow{\theta}^e_{n+1} = 0$ in Lemma 6.1 for notational convenience. At this stage we emphasize that all the boundary terms in (6.3) vanish. In fact, as $f(L) = 0$, the boundary term in the last interval vanishes as stated in Lemma 6.1. For the other intervals, from (6.1), we claim that for $j \in \mathbb{N}_n$,

$$\mathbb{E}[\widehat{F}_{j+1}(\bar{X}_{j+1}) \mathbb{1}_{D_{j+1,n}} \bar{\theta}_{j+1} \delta_L(\bar{X}_j) \overrightarrow{\theta}^\partial_j | \mathcal{G}_{j-1}, T^{n+1}] = 0. \tag{6.4}$$

---

[6]As before, we use the convention $\sum_\emptyset \cdots = 0$, $\prod_\emptyset \cdots = 1$.

To see this, we note that for $\bar{X}_j = L$, one has $\bar{X}_{j+1} = L + \sigma(L)Z_{j+1}$, which is independent of ρ_{j+1}, and $\bar{\theta}_{j+1}$ given by (3.2) reduces to $\bar{\theta}_{j+1} = 2(2\rho_{j+1} - 1)\lambda^{-1}(\frac{1}{2}(a(\bar{X}_{j+1}) - a(L))\mathcal{I}_{j+1}^2(1) + (b(L) - a'(\bar{X}_{j+1}))\mathcal{I}_{j+1}(1) + \frac{a''(\bar{X}_{j+1})}{2} - b'(\bar{X}_{j+1}))$. Therefore the conclusion follows by conditioning with respect to $\{\bar{X}_j = L\}$ in (6.4) and by noting that $\mathbb{E}[2\rho_{j+1} - 1|\mathcal{G}_{j-1}, T^{n+1}, \bar{X}_j = L] = 0$.

From this property the identity (6.2) becomes

$$\partial_x \mathbb{E}[f(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}] = \mathbb{E}[\widehat{F}_k(\bar{X}_k)\mathbb{1}_{D_{k,n}}\mathcal{I}_k(\overrightarrow{\theta}^e_k) \prod_{i=1}^{k-1} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}]$$
$$+ \sum_{j=1}^{k} \mathbb{E}[\widehat{F}_j(\bar{X}_j)\mathbb{1}_{D_{j,n}} \overrightarrow{\theta}^c_j \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}].$$

Now, for each $k \in \mathbb{N}_n$, one can multiply the above equality by the length of the interval on which the local IBP formula is performed, namely $\zeta_k - \zeta_{k-1}$ and sum them over all k. For the last interval, we multiply (6.3) by $T - \zeta_n$. This gives

$$T\partial_x \mathbb{E}[f(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i | T^{n+1}]$$
$$= \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1})\mathbb{E}[f(\bar{X}_{n+1}) \prod_{i=k+1}^{n+1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i \times \mathbb{1}_{D_{k,n}}\mathcal{I}_k(\overrightarrow{\theta}^e_k) \times \prod_{i=1}^{k-1} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}]$$
$$+ \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) \sum_{j=1}^{k} \mathbb{E}[f(\bar{X}_{n+1}) \prod_{i=j+1}^{n+1} \mathbb{1}_{D_{i,n}} \bar{\theta}_i \times \mathbb{1}_{D_{j,n}} \overrightarrow{\theta}^c_j \times \prod_{i=1}^{j-1} \mathbb{1}_{D_{i,n}} \overrightarrow{\theta}^e_i | T^{n+1}].$$

From the above formula, the $\mathbb{L}^p$-moment estimate, $p \in [0, 2)$, follows by similar arguments as described at the end of the proof of Theorem 3.1. Finally, one concludes by using the Lebesgue differentiation which yields

$$T\partial_x \mathbb{E}[f(X_T)\mathbb{1}_{\{\tau \geq T\}}] = T\partial_x \mathbb{E}[f(\bar{X}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i]$$
$$= \sum_{n \geq 0} \mathbb{E}\Big[T\partial_x \mathbb{E}[f(\bar{X}_{n+1}) \prod_{i=1}^{n+1} \theta_i | T^{n+1}] \mathbb{1}_{\{N_T = n\}}\Big]$$

and summing the previous formula over n. $\qquad\square$

Remark 6.3. (i) The right hand side of the IBP formula may alternatively be written as

$$T\partial_x \mathbb{E}[f(X_T)\mathbb{1}_{\{\tau \geq T\}}] = \mathbb{E}\left[f(\bar{X}_{N_T+1}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \left\{ \overrightarrow{\theta}^{I_k^{N_T+1}} + \sum_{j=1}^{k} \overrightarrow{\theta}^{C_j^{N_T+1}} \right\}\right].$$

(ii) We note that the above formula does not involve any merging procedure. This is due to the fact that only the first derivative is being considered here. In fact, the key property (6.4) would not be satisfied if one considers second order derivatives. Therefore, a merging procedure similar to the one described in Section 4.3 would be necessary.

(iii) Similarly to the previous section, the above theorem yields a probabilistic representation for the derivative of the transition density of the killed process at time T from which stems an unbiased Monte Carlo simulation method. In contrast with Corollary 5.3, the derivative is taken with respect to the starting point. This can be seen by formally taking the Dirac mass at point z as a test function in Theorem 6.2.[7]

---

[7]We again refer the interested reader to [15] or to [14] for an analytical proof of the differentiability of the map $[L, \infty) \ni x \mapsto p(T, x, z)$.

Corollary 6.4. *Under assumption **(H)**, the transition density of the killed process at time T is differentiable with respect to its starting point. Moreover, for all $(T, x, z) \in (0, \infty) \times [L, \infty)^2$ the following probabilistic representation holds*

$$T\partial_x p(T, x, z)$$

$$= \mathbb{E}\left[g(a(\bar{X}_{N_T})(T - \zeta_{N_T}), z - \bar{X}_{N_T}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \left\{\overrightarrow{\theta}^{I_k^{N_T+1}} + \sum_{\mathbf{s} \in \widehat{S}_{N_T+1}^k} \overrightarrow{\theta}^{\mathbf{s}}\right\}\right].$$

7 Achieving finite variance by importance sampling

The previous probabilistic representations of Theorems 3.1, 5.1 and 6.2 as well as Corollaries 5.3 and 6.4 allows one to devise an unbiased Monte Carlo simulation. However, in general, its use is hampered by the fact that the variance is infinite as suggested for instance by the moment estimate of Lemma 10.2. The main tool that we develop here in order to circumvent this issue consists in employing an importance sampling scheme on the jump times of the Poisson process N as originally proposed by Andersson and Kohatsu-Higa [3]. Since the arguments developed below follow similar lines of reasonings as those employed in [3], we will omit some technical details.

Let us first introduce a renewal process in the following sense:

Definition 7.1. *Let $(T_n)_{n \geq 1}$ be a sequence of random variables such that $(T_n - T_{n-1})_{n \geq 1}$, with the convention $T_0 = 0$, are i.i.d. with density f and c.d.f: $t \mapsto F(t) = \int_{-\infty}^t f(s)\,ds$. Then, the renewal process $J := (J_t)_{t \geq 0}$ with jump times $(T_n)_{n \geq 1}$ is defined by $J_t := \sum_{n \geq 1} \mathbb{1}_{\{T_n \leq t\}}$.*

As previously done, we assume that J is independent of the Brownian motion W. It is readily seen that $\{J_t = n\} = \{T_n \leq t < T_{n+1}\}$ and by an induction argument that we omit, one may prove that the joint distribution of $(T_1, \cdots, T_n)$ is given by

$$\mathbb{P}(T_1 \in ds_1, \cdots, T_n \in ds_n) = \prod_{j=0}^{n-1} f(s_{i+1} - s_i)\mathbb{1}_{\{0 < s_1 < \cdots < s_n\}}$$

which in turn implies

$$\mathbb{E}[\mathbb{1}_{\{J_t = n\}}\Phi(T_1, \cdots, T_n)] = \mathbb{E}[\mathbb{1}_{\{T_n \leq t < T_{n+1}\}}\Phi(T_1, \cdots, T_n)] \tag{7.1}$$

$$= \int_t^\infty \int_{\Delta_n(t)} \Phi(s_1, \cdots, s_n) \prod_{j=0}^n f(s_{j+1} - s_j)\,d\mathbf{s}_{n+1} \tag{7.2}$$

so that, by Fubini's theorem

$$\mathbb{E}[\mathbb{1}_{\{J_t = n\}}\Phi(T_1, \cdots, T_n)] = \int_{\Delta_n(t)} \Phi(s_1, \cdots, s_n)(1 - F(t - s_n)) \prod_{j=0}^{n-1} f(s_{j+1} - s_j)\,d\mathbf{s}_n \tag{7.3}$$

for any map $\Phi : \Delta_n(t) \to \mathbb{R}$ satisfying $\mathbb{E}[\mathbb{1}_{\{J_t = n\}}|\Phi(T_1, \cdots, T_n)|] < \infty$. Usual choices that we will consider are the followings:

Examples:

1. If the density function f is given by $f(t) = \lambda e^{-\lambda t}\mathbb{1}_{[0,\infty)}(t)$ so that $F(t) = 1 - e^{-\lambda t}$, $t \geq 0$, for some positive parameter λ, then J is a Poisson process with intensity λ.

2. If the density function f is given by $f(t) = \frac{1-\alpha}{\bar{\tau}^{1-\alpha}} \frac{1}{t^\alpha}\mathbb{1}_{[0,\bar{\tau}]}(t)$, so that $F(t) = (t/\bar{\tau})^{1-\alpha}$, $t \in [0, \bar{\tau}]$, for some parameters $(\alpha, \bar{\tau}) \in (0, 1) \times (0, \infty)$, then J is a renewal process with $[0, \bar{\tau}]$-valued $Beta(1 - \alpha, 1)$ jump times.

3. More generally, if the density function f is given by $f(t) = \frac{\bar{\tau}^{1-\alpha-\beta}B(\alpha,\beta)^{-1}}{t^{1-\alpha}(\bar{\tau}-t)^{1-\beta}}\mathbb{1}_{[0,\bar{\tau}]}(t)$, so that $F(t) = B(t/\bar{\tau}, \alpha, \beta)/B(\alpha, \beta)$, $[0,1] \ni x \mapsto B(x, \alpha, \beta)$ being the incomplete Beta function, for some parameters $(\alpha, \beta, \bar{\tau}) \in (0,1)^2 \times (0,\infty)$, then J is a renewal process with $[0,\bar{\tau}]$-valued $Beta(\alpha,\beta)$ jump times.

Having these definitions at hand, as done before, we define the partition π of $[0,T]$ given by $\pi := \{0 =: \tau_0 < \cdots < \tau_{J_T} \leq T\}$ with $\tau_i := T_i \wedge T$. The probabilistic representation of Theorem 3.1 then becomes

$$\mathbb{E}\left[f(X_T)\mathbb{1}_{\{\tau>T\}}\right] = \mathbb{E}\left[f(\bar{X}_{J_T+1})\prod_{i=1}^{J_T+1}\mathbb{1}_{D_{i,J_T}}\bar{\theta}_i\right], \tag{7.4}$$

where the dynamics of the Markov chain $\bar{X}$ is given by (2.4) with $(Z_{i+1} = W_{\tau_{i+1}} - W_{\tau_i})_{0 \leq i \leq J_T}$, the set $D_{i,n} := \{\bar{X}_i \geq L, J_T = n\}$ for $i \in \bar{\mathbb{N}}_{n+1}$, and for $i \in \bar{\mathbb{N}}_{J_T}$

$$\bar{\theta}_i := \mathbb{1}_{\{J_T > i-1\}}2(2\rho_i - 1)(f(\tau_i - \tau_{i-1}))^{-1}\left(\mathcal{I}_i(c_1^i) + \mathcal{I}_i^2(c_2^i)\right) \tag{7.5}$$
$$+ \mathbb{1}_{\{J_T=i-1\}}(1 - F(T - \tau_{J_T}))^{-1}(2\rho_{J_T+1} - 1).$$

Moreover, the following estimates holds on the set $\{J_T = n\}$

$$\mathbb{1}_{\{i \leq n\}}(\tau_i - \tau_{i-1})^{\frac{p}{2}}f(\tau_i - \tau_{i-1})^p\mathbb{1}_{D_{i-1,n}}\mathbb{E}_{i-1,n}\left[\mathbb{1}_{D_{i,n}}|\bar{\theta}_i|^p\right]$$
$$+ \mathbb{1}_{\{i=n+1\}}\mathbb{1}_{D_{n,n}}(1 - F(T - \tau_n))^p\mathbb{E}_{n,n}\left[\mathbb{1}_{D_{n+1,n}}|\bar{\theta}_{n+1}|^p\right] \leq C \tag{7.6}$$

where C is positive constant independent of n. Hence, using (7.6) and then (7.3), for all $p \geq 1$

$$\mathbb{E}\left[\left|\prod_{i=1}^{J_T+1}\mathbb{1}_{D_{i,J_T}}\bar{\theta}_i\right|^p\right]$$
$$= \sum_{n\geq 0}\mathbb{E}\left[\left|\prod_{i=1}^{n+1}\mathbb{1}_{D_{i,n}}\bar{\theta}_i\right|^p\mathbb{1}_{\{J_T=n\}}\right]$$
$$\leq C\sum_{n\geq 0}C^{n+1}\int_{A_n}(1 - F(T - t_n))^{-p+1}\prod_{i=1}^n(t_i - t_{i-1})^{-\frac{p}{2}}f(t_i - t_{i-1})^{-p+1}\,dt_1\cdots dt_n.$$

As already mentioned before, the previous estimate is actually quite sharp and the series appearing in the right-hand side is finite for $p \in [0,2)$ if J is a Poisson process. In order to achieve a finite variance, one has to select the law of the jump times suitably. Indeed, if for instance J is a renewal process with $[0,\bar{\tau}]$-valued $Beta(1-\alpha, 1)$ jump times, $\alpha \in (0,1)$, $\bar{\tau} > T$, a simple computation shows that the above series is finite as soon as $-\frac{p}{2} + \alpha(p-1) > -1$, that is, $p(\frac{1}{2} - \alpha) < 1 - \alpha$. In particular, taking $\alpha = 1/2$, it is readily seen that the moment of order p of the random variable appearing inside the expectation in the right-hand side of (7.4) is finite for all $p \geq 1$. Similarly, if J is a renewal process with $[0,\bar{\tau}]$-valued $Beta(1-\alpha, 1-\beta)$ jump times, $(\alpha,\beta) \in (0,1)^2$, $\bar{\tau} > T$, the integral appearing in the right-hand side of the above inequality is finite as soon as $p(\frac{1}{2} - \alpha) < 1 - \alpha$ and $-p\beta < 1 - \beta$, the later condition being always satisfied. In particular, taking $\alpha = 1/2$ and any $\beta \in (0,1)$, any moment of order p for $p \geq 1$ is finite.

We now provide the probabilistic representation for the two IBP formulas using the above importance sampling technique. We thus redefine the weights appearing in the first IBP formula keeping in mind that the sequence $(\bar{\theta}_i)_{1 \leq i \leq J_T+1}$ is now given by (7.5). The new random variables $(\overleftarrow{\theta}_{i+1}^e, \overleftarrow{\theta}_{i+1}^c, \overleftarrow{\theta}_{i+1}^\partial) \in \mathbb{S}_{i+1,n}(\bar{X})$ of Lemma 4.1 are now

defined by: for $i \in \bar{\mathbb{N}}_{J_T-1}$

$$\overleftarrow{\theta}^e_{i+1} := 2(f(\tau_{i+1} - \tau_i))^{-1}\left(\mathcal{I}^2_{i+1}(d^{i+1}_2) + \mathcal{I}_{i+1}(d^{i+1}_1)\right),$$

$$\overleftarrow{\theta}^c_{i+1} := \mathcal{I}_{i+1}\left(\bar{\theta}_{i+1} - (2\rho_{i+1} - 1)\overleftarrow{\theta}^e_{i+1}\right) - \partial_{\bar{X}_i}\overleftarrow{\theta}^e_{i+1} - \sigma'_i\mathcal{I}_{i+1}\left(Z_{i+1}\overleftarrow{\theta}^e_{i+1}\right), \qquad (7.7)$$

$$\overleftarrow{\theta}^\partial_{i+1} := 2(2\rho_{i+1} - 1)(f(\tau_{i+1} - \tau_i))^{-1}(a'(L) - b(L))\mathcal{I}_{i+1}(1)$$

and $\overleftarrow{\theta}^\partial_{J_T+1} := 0$, $\overleftarrow{\theta}^e_{J_T+1} := 2(1 - F(T - \tau_{J_T}))^{-1}$, $\overleftarrow{\theta}^c_{J_T+1} := -2(1 - F(T - \tau_{J_T}))^{-1}(\sigma'\sigma)_{J_T}(T - \tau_{J_T})\mathcal{I}^2_{J_T+1}(1)$. For the boundary merging weights of Lemmas 4.3 and 4.4, we first redefine the Markov chain $\bar{X}^\partial$ with dynamics (4.6) by modifying the corresponding increments $Z_{j,i+1}$ as done previously. We then set $D^\partial_{j,i+1,n} := \{\bar{X}^\partial_{j,i+1} \geq L, J_T = n\}$, $j = i - 1, i$. Finally, on the set $\{J_T = n\}$,

$$\overleftarrow{\theta}^{\partial *e}_{j,i+1} := 4(f(\tau_{i+1} - \tau_j))^{-1}\frac{a'(L) - b(L)}{a_j}\left(\bar{\mathcal{I}}^2_{j,i+1}(\bar{d}^{i+1}_2) + \bar{\mathcal{I}}_{j,i+1}(\bar{d}^{i+1}_1)\right), \qquad (7.8)$$

$$\overleftarrow{\theta}^{\partial \circledast e}_{j,i+1} := 4(f(\tau_{i+1} - \tau_j))^{-1}\frac{a'(L) - b(L)}{a_j^{3/2}\sigma(L)}(\bar{X}_j - L)\left(\bar{\mathcal{I}}^2_{j,i+1}(\hat{d}^{i+1}_2) + \bar{\mathcal{I}}_{j,i+1}(\hat{d}^{i+1}_1)\right) \qquad (7.9)$$

for $j = i - 1, i$ with coefficients given by

$$\hat{d}^{i+1}_k := \bar{d}^{i+1}_k \times (\bar{\Phi}g^{-1})(a(L)(\tau_{i+1} - \tau_j), Z_{j,i+1}), \quad k = 1, 2,$$

and $\overleftarrow{\theta}^{\partial *e}_{j,n+1} := 4(1 - F(T - \tau_j))^{-1}\frac{a'(L)-b(L)}{a_j}$, $\overleftarrow{\theta}^{\partial \circledast e}_{j,n+1} := 4(1 - F(T - \tau_j))^{-1}\frac{2a'(L)-b(L)}{a_j^{3/2}\sigma(L)}(\bar{X}_j - L)\hat{d}^{n+1}$ for $j = n - 1, n$ with $\hat{d}^{n+1} := (\bar{\Phi}g^{-1})(a(L)(\tau_{n+1} - \tau_j), Z_{j,n+1})$.

At this stage, it is important to remark that the above new weights satisfy a time degeneracy estimate, with a slight modification of Definition 3.3. The notation $\mathbb{E}_{i,n}[X]$ now is used for the expectation of X conditional on $\{\mathcal{G}_i, T^{n+1}, \rho^{n+1}, J_T = n\}$ and one considers the corresponding norm $\|.\|_{p,i,n}$. We now say that a weight $H \in \mathbb{S}_{i,n}$ satisfies the time degeneracy estimate if for all $p \geq 1$

$$\mathbb{1}_{D_{i-1,n}}\left\|\mathbb{1}_{D_{i,n}}H\right\|_{p,i-1,n} \leq Cf(\tau_i - \tau_{i-1})^{-1}(\tau_i - \tau_{i-1})^{-\frac{1}{2}} \qquad (7.10)$$

in the case that $i \in \mathbb{N}_n$ and $\mathbb{1}_{D_{n,n}}\left\|\mathbb{1}_{D_{n+1,n}}H\right\|_{p,n,n} \leq C$ in the case that $i = n + 1$. In a completely analogous manner as done in Lemma 4.1, the weights $\overleftarrow{\theta}^a_i$ for $a \in \{e, c, \partial\}$ satisfies the time degeneracy estimate (7.10). For the boundary merging weights, we replace (7.10) for $H \in \mathbb{S}_{j,i+1,n}(\bar{X}^\partial)$ by

$$\forall p \geq 1, \quad \mathbb{1}_{D_{j,n}}\mathbb{E}\left[\mathbb{1}_{D^\partial_{j,i+1,n}}|H|^p\Big|\mathcal{G}_{j-1}, \tau_{i+1}, J_T = n\right] \leq C(f(\tau_{i+1} - \tau_j))^{-p}(\zeta_{i+1} - \zeta_j)^{-\frac{p}{2}}, \qquad (7.11)$$

for $i \in \mathbb{N}_{n-1}$, $j = i - 1, i$ and $\mathbb{1}_{D_{j,n}}\mathbb{E}\left[\mathbb{1}_{D^\partial_{j,n+1,n}}|H|^p\Big|\mathcal{G}_{j-1}, J_T = n\right] \leq C$ for $j = n - 1, n$. Doing so, the new boundary merging weights $\overleftarrow{\theta}^{\partial *e}_{j,i+1}$ and $\overleftarrow{\theta}^{\partial \circledast e}_{j,i+1}$ defined respectively by (7.8) and (7.9) both satisfy the time degeneracy estimate (7.11).

With the above new definitions and properties, we finally redefine the corresponding weights $\overleftarrow{\theta}^{\mathbf{s}}$ with the related Markov chain $\bar{X}^{\mathbf{s}}$ for each $\mathbf{s} \in S_{n+1}$ or $\mathbf{s} \in \dot{S}^k_{n+1}$ on the time partition $\pi := \{0 =: \tau_0 < \cdots < \tau_{n+1} := T\}$ of the underlying renewal process on the set $\{J_T = n\}$. This is done in a completely analogous manner as presented in the subsection 5.2.

We can now restate Theorem 5.1 as follows. For any function $f \in \mathscr{C}_b^1(\mathbb{R})$ satisfying $f(L) = 0$,

$$T\mathbb{E}[f'(X_T)\mathbb{1}_{\{\tau \geq T\}}]$$

$$= \mathbb{E}\left[\sum_{k=1}^{J_T+1}(\tau_k - \tau_{k-1})\left\{f(\bar{X}_{J_T+1})\overleftarrow{\theta}^{I_k^{J_T+1}} + \sum_{\mathbf{s} \in \bar{S}_{J_T+1}^k \cup \dot{S}_{J_T+1}^k} f(\bar{X}_{J_T+1}^{\mathbf{s}})\overleftarrow{\theta}^{\mathbf{s}}\right\}\right].$$

Notethat the corresponding formulation of Remark 5.2 also holds. Moreover, if J is a renewal process with $[0, \bar{\tau}]$-valued $Beta(1-\alpha, 1)$ jump times, $\bar{\tau} > T$, with α satisfying $p(\frac{1}{2}-\alpha) < 1-\alpha$ or if J is a renewal process with $[0, \bar{\tau}]$-valued $Beta(1-\alpha, 1-\beta)$ jump times, $\bar{\tau} > T$, with α and β such that $p(\frac{1}{2}-\alpha) < 1-\alpha$ and $\beta \in (0, 1)$, then the r.v. appearing inside the expectation of the right-hand side of the above equality belongs to $\mathbb{L}^p(\mathbb{P})$ for any $p \geq 1$. The proof follows similar lines of reasoning as those employed above in order to deal with the new probabilistic representation (7.4) and is thus omitted.

We proceed similarly for the BEL formula of Theorem 6.2. Namely, we redefine the weights $\overrightarrow{\theta}^a$, for $a \in \{e, c, \partial\}$ as follows

$$\overrightarrow{\theta}_i^e := 2(f(\tau_i - \tau_{i-1}))^{-1}\left(\mathcal{I}_i^2(d_2^i) + \mathcal{I}_i(d_1^i)\right),$$

$$\overrightarrow{\theta}_i^c := \mathcal{I}_i\left((2\rho_i - 1)\bar{\theta}_i - \overrightarrow{\theta}_i^e\right) + \partial_{\bar{X}_{i-1}}\bar{\theta}_i + \sigma_i'\mathcal{I}_i\left(Z_i\bar{\theta}_i\right),$$

$$\overrightarrow{\theta}_i^\partial := \overrightarrow{\theta}_i^e$$

and also set $\overrightarrow{\theta}_{n+1}^e := 2(1 - F(T - \zeta_n))^{-1}(1 + (2\rho_{n+1} - 1)\sigma_n' Z_{n+1})$, $\overrightarrow{\theta}_{n+1}^c := 0$. These new weights satisfy the time degeneracy estimate (7.10). Then, the following BEL formula is satisfied for any $f \in \mathcal{C}_b^1(\mathbb{R})$ satisfying $f(L) = 0$ and any initial point $x \in [L, \infty)$:

$$T\partial_x\mathbb{E}[f(X_T)\mathbb{1}_{\{\tau \geq T\}}] = \mathbb{E}\left[f(\bar{X}_{J_T+1})\sum_{k=1}^{J_T+1}(\zeta_k - \zeta_{k-1})\left\{\overrightarrow{\theta}^{I_k^{J_T+1}} + \sum_{\mathbf{s} \in \widehat{S}_{J_T+1}^k} \overrightarrow{\theta}^{\mathbf{s}}\right\}\right].$$

Moreover, the r.v. appearing inside the expectation in the right-hand side of the above equality belongs to $\mathbb{L}^p(\mathbb{P})$, for any $p \geq 1$ in the case of $[0, \bar{\tau}]$-valued $Beta(1-\alpha, 1)$ jump times or $[0, \bar{\tau}]$-valued $Beta(1-\alpha, 1-\beta)$ jump times under the condition $p(\frac{1}{2}-\alpha) < 1-\alpha$ and $\beta \in (0, 1)$. In particular, choosing $\alpha = 1/2$, the $L^p(\mathbb{P})$ moment is finite for any $p \geq 1$.

8 Numerical tests

In this section, we provide some numerical results for the unbiased Monte Carlo simulation method based on the probabilistic representation formula established in Theorem 3.1 for the marginal law of the killed process and the Bismut-Elworthy-Li (BEL for short) formula of Theorem 6.2. A similar numerical analysis could be done for the IBP formula established in Theorem 5.1 but we restrict to the two aforementioned case for sake of simplicity. For a one dimensional Brownian motion W, we thus consider the following one-dimensional SDE with dynamics

$$X_t = x_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s, \ x \in \mathbb{R} \tag{8.1}$$

and we choose the coefficients b, σ and the test function f as follows

$$\sigma(x) = \bar{\sigma} \times (\sin(\omega x) + 2), \quad b(x) = -\frac{x}{x^2 + \frac{c_1}{3c_3}}\sigma(x), \quad f(x) = c_3 x^3 + c_1 x + c_0$$

for some positive constant $\bar{\sigma} > 0$. With this particular choice, we first observe that assumption **(H)** is clearly satisfied and that a direct computation yields $\mathcal{L}f(x) := b(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x) = 0$. The process $(f(X_t))_{t \geq 0}$ is thus a martingale and by Doob's stopping theorem one gets $\mathbb{E}[f(X_{\tau \wedge T})] = f(x_0)$ where $\tau = \inf\{t \geq 0 : X_t \leq L\}$. We now shift the function f by considering $h(x) = f(x) - f(L)$ satisfying $h(L) = 0$ instead of f. It is readily seen that h satisfies $\mathcal{L}h(x) = 0$ so that $\mathbb{E}[h(X_{\tau \wedge T})] = \mathbb{E}[h(X_T)\mathbb{1}_{\{\tau > T\}}] = h(x_0)$. Note that since h is not bounded but of polynomial growth, only Theorem 3.1 directly applies. However, the extension of Theorem 6.2 (and also of Theorem 5.1) to polynomially growing function can be performed by a standard approximation argument noting that all moments of $\bar{X}_{N_T+1}$ and X_T are bounded.

Having this extension in mind, by Theorem 6.2, one has

$$\forall x \geq L, \quad Th'(x) = \partial_x \mathbb{E}[h(X_T)\mathbb{1}_{\{\tau > T\}}]$$

$$= \mathbb{E}\left[h(\bar{X}_{N_T+1}) \sum_{k=1}^{N_T+1}(\zeta_k - \zeta_{k-1})\left\{\overrightarrow{\theta}^{I_k^{N_T+1}} + \sum_{\mathbf{s} \in \widehat{S}_{N_T+1}^k} \overrightarrow{\theta}^{\mathbf{s}}\right\}\right].$$

We select the following parameters: $T = 0.5$, $L = 0$, $c_0 = 0$, $c_1 = 1$, $c_3 = 1$ and $x_0 = 1$ so that $h(x_0) = 2$ and $Th'(x_0) = 2$. We use three different parameters sets for $\bar{\sigma} = \omega = 0.1, 0.2, 0.3$. We examine the performance of the proposed Monte Carlo estimator with respect to the previous sets of parameters when one uses the Exponential sampling (the distribution of the jump times is exponential with parameter λ) as it is written in Theorems 3.1 and 6.2 and when one uses an importance sampling technique with Beta distribution with parameters $(\gamma, \bar{\tau})$ for the jump times of the renewal process, as exposed in Section 7, which allows to achieve finite variance for our estimators. We note that though the variance is not finite in the case of Exponential time sampling, we include it here in order to compare its performance with the Beta sampling scheme. In both cases, we first select the optimal parameters which minimize the variance of the estimator using few samples, that is, the optimal λ in the case of Exponential sampling and the optimal $(\gamma, \bar{\tau})$ in the case of Beta sampling. We then use $M = 4 \times 10^6$ i.i.d. samples to estimate the considered quantities.

The numerical results are summarized in the two tables below. The first column of Table 1 and Table 2 provides the value of the two parameters $\bar{\sigma} = \omega$. The second column (resp. third column) of Table 1 provides the estimated value of the quantity $\mathbb{E}[h(X_T)\mathbb{1}_{\{\tau > T\}}]$ with its associated variance, $L^1(\mathbb{P})$-error and 95%-confidence interval in the case of Exponential sampling (resp. Beta sampling). The second column (resp. third column) of Table 2 provides the estimated value of the quantity $\partial_x \mathbb{E}[h(X_T)\mathbb{1}_{\{\tau > T\}}]$ with its associated variance, $L^1(\mathbb{P})$-error and 95%-confidence interval in the case of Exponential sampling (resp. Beta sampling).

Most notably, we observe that in both tables the performance of our estimators quickly deteriorates as $\bar{\sigma} = \omega$ increases. Actually, for large values of $\bar{\sigma}$, ω, say greater than 0.4, the variance becomes difficult to estimate from the simulations and the obtained estimates become unreliable. We also see that the Beta time sampling method outperforms the Exponential time sampling for all values of the considered parameters especially for large values of $\bar{\sigma}$. This behavior was already observed in [3] and is reminiscent of unbiased simulation methods for multidimensional diffusion processes. It was thus expected here since there is no hope that our estimator will overcome this problem. To circumvent this issue, one may resort to more sophisticated method such as the second order approximation method developed by Andersson et al. [4].

Table 1: Unbiased Monte Carlo estimation for the quantity $\mathbb{E}[h(X_T)\mathbb{1}_{\{\tau>T\}}]$ based on Theorem 3.1 by Exponential and Beta sampling with its associated 95%-confidence interval.

$\bar{\sigma}=\omega$	Exponential sampling	Beta sampling
0.1	2.0; 26.3; 3.2; (+/−) 0.005	2.0; 14.9; 2.9; (+/−) 0.004
0.2	1.99; 213.2; 4.7; (+/−) 0.014	1.99; 77.2; 4.5; (+/−) 0.009
0.3	2.0; 3064.1; 7.8; (+/−) 0.054	1.98; 681.2; 7.6; (+/−) 0.025

Table 2: Unbiased Monte Carlo estimation for the quantity $\partial_x\mathbb{E}[h(X_T)\mathbb{1}_{\{\tau>T\}}]$ based on Theorem 6.2 by Exponential and Beta sampling with its associated variance and 95%-confidence interval.

$\bar{\sigma}=\omega$	Exponential sampling	Beta sampling
0.1	1.99; 379.7; 8.4; (+/−) 0.019	2.00; 295.0; 8.4; (+/−) 0.017
0.2	1.98; 1008.7; 7.1; (+/−) 0.035	1.98; 467.7; 7.1; (+/−) 0.021
0.3	1.97; 5411.8; 8.9; (+/−) 0.072	1.97; 2358.4; 8.7; (+/−) 0.047

9 Some conclusions

In the present work, we presented a probabilistic representation formula for the marginal law of a killed process based on a basic Markov chain which is obtained using the reflection principle. From this representation, we established two IBP formulae, one being of BEL's type, from which directly stem an unbiased Monte Carlo method. The main element used in this construction is a suitable tailor-made Malliavin calculus for the underlying Markov chain. For this reason, we do not need to use the full fledge power of Malliavin calculus by closing the derivative operator but just the concepts for simple discrete time Markov chain.

The methodology developed here seems to follow a general pattern that could be used to obtain IBP formulae for some other irregular functionals of the Wiener process for which boundary problems may appear such as the exit time, the local time, the running maximum or the occupation time of a multi-dimensional diffusion process.

Although the problem investigated here focuses in the one dimensional case, we believe that the approach developed here also extends to some multi-dimensional cases for which the reflection principle is well understood and densities for basic approximation processes are known, see e.g. [2], [10] and the references therein.

On the other hand, it seems difficult at this moment to generalize the methods in [23] and [21] to the multi-dimensional case or even to obtain an amenable integration by parts formula based on a Markov chain using these formulations or a Lamperti like transform. We will discuss this extension in future works as well as their implementation for simulation purposes.

10 Appendix

10.1 Proof of the probabilistic representation in Theorem 3.1

Let X be the solution to (1.1). Let P denote the semigroup operator associated with the killed process. That is, for a measurable and bounded function f, one defines $P_t f(x) = \mathbb{E}\left[f(X_t)\mathbb{1}_{\{\tau>t\}}\right]$. We remark that due to the indicator function, this semigroup is not conservative. The heuristic argument in order to obtain the probabilistic repre-

sentation is to use Itô's formula on an approximation process obtained from (1.1) by removing the drift and freezing the diffusion coefficient at the starting point. From Itô's formula, one obtains a one step expansion of the law of X around the law of the Markov chain $\bar{X}$. Then, an IBP formula based on the Markov chain $\bar{X}$ has to be used to obtain the probabilistic representation of this first step expansion using one jump of the Poisson process N. Then, one just needs to iterate the first step expansion in order to obtain the full probabilistic representation.

In order to do this rigorously, one needs to use the regularity properties of the semigroup P which can be found in [15], Chapter VI and/or [16]. In particular, under assumption (**H**) on the coefficients, one obtains that if f is a smooth function such that $\lim_{x \downarrow L} f(x) = f(L) = 0$ then $Pf \in \mathscr{C}^{1,2}((0,T] \times [L, \infty))$ and it satisfies $\partial_t P_t f = \mathcal{L} P_t f$ on $[L, \infty)$, $t > 0$, where $\mathcal{L} = \frac{1}{2} a \partial_x^2 + b \partial_x$ is the infinitesimal generator of X. Moreover, one has $\sup_{0 \le t \le T} |\partial_x^\ell P_t f|_\infty \le C$, for $\ell = 1, 2$, for some positive constant $C := C(T, a, b)$.

Under the condition that the test function f vanishes at L, we obtain that P satisfies its corresponding Dirichlet boundary condition $P_t f(L) = f(L) = 0$ together with $P_0 f(x) = f(x)$ for all $x \ge L$.

The argument used to obtain the probabilistic representation starts by applying Itô's formula to $(P_{T-t} f(\bar{Y}_{t \wedge \bar{\tau}}))_{t \in [0,T]}$ where the process $\bar{Y}$ is defined in Lemma 2.1 on the interval $[0, T]$ and $\bar{\tau}$ is its associated exit time. We also recall the definition of the approximation process obtained from the reflection principle of Lemma 2.1, namely $\bar{X}_t^{s,x} = \rho x + (1 - \rho)(2L - x) + \sigma(x)(W_t - W_s)$ (with the shorten notation $\bar{X}_t = \bar{X}_t^{0,x}$):

$$
f(\bar{Y}_T) \mathbb{1}_{\{\bar{\tau} > T\}} \overset{\mathbb{E}}{=} P_T f(x) + \int_0^T \left(-\partial_u P_u f(\bar{Y}_s) \Big|_{u = T-s} + \frac{1}{2} a(x) \partial_x^2 P_{T-s} f(\bar{Y}_s) \right) \mathbb{1}_{\{\bar{\tau} > s\}} ds
$$

$$
\overset{\mathbb{E}}{=} P_T f(x)
$$
$$
+ \int_0^T \left(\frac{1}{2} \left(a(x) - a(\bar{Y}_s) \right) \partial_x^2 P_{T-s} f(\bar{Y}_s) - b(\bar{Y}_s) \partial_x P_{T-s} f(\bar{Y}_s) \right) \mathbb{1}_{\{\bar{\tau} > s\}} ds
$$

$$
\overset{\mathbb{E}}{=} P_T f(x) + 2(2\rho - 1)
$$
$$
\times \int_0^T \left(\frac{1}{2} \left(a(x) - a(\bar{X}_s) \right) \partial_x^2 P_{T-s} f(\bar{X}_s) - b(\bar{X}_s) \partial_x P_{T-s} f(\bar{X}_s) \right) \mathbb{1}_{\{\bar{X}_s \ge L\}} ds.
$$

We now rewrite the previous representation using the Markov chain $(\bar{X}_i)_{0 \le i \le N_T + 1}$ defined by (2.4) together with the Poisson process N. From the previous identity, we get

$$
P_T f(x) \overset{\mathbb{E}}{=} f(\bar{X}_T) 2(2\rho - 1)
$$
$$
+ 2(2\rho - 1) \int_0^T \left(\frac{1}{2} \left(a(\bar{X}_s) - a(x) \right) \partial_x^2 P_{T-s} f(\bar{X}_s) + b(\bar{X}_s) \partial_x P_{T-s} f(\bar{X}_s) \right) \mathbb{1}_{\{\bar{X}_s \ge L\}} ds
$$

(10.1)

$$
\overset{\mathbb{E}}{=} f(\bar{X}_{N_T+1}) \theta_{N_T+1} \mathbb{1}_{\{N_T = 0\}} + e^{\lambda T} 2 \lambda^{-1} (2\rho_{N_T} - 1)
$$
$$
\times \left\{ \frac{1}{2} (a(\bar{X}_1) - a(x)) \partial_x^2 P_{-\zeta_1} f(\bar{X}_1) + b(\bar{X}_1) \partial_x P_{-\zeta_1} f(\bar{X}_1) \right\} \mathbb{1}_{\{\bar{X}_1 \ge L\}} \mathbb{1}_{\{N_T = 1\}}.
$$

Next, we apply the IBP formula (2.9) with respect to the r.v. $\bar{X}_1$ in the above expression. The formula is applied once to the terms associated with the drift coefficient b and two times with respect to the terms related to the diffusion coefficient a. In order to do that one first has to take the conditional expectation $\mathbb{E}_{0,1}[.]$ in the second term of the above equality. As these IBPs involve the indicator function $\mathbb{1}_{\{\bar{X}_1 \ge L\}}$, the correct

calculation has to be done using the theory in [22], Chapter V.9.[8] This yields

$$P_T f(x) \overset{\mathbb{E}}{=} f(\bar{X}_{N_T+1})\theta_{N_T+1}\mathbb{1}_{\{N_T=0\}} + e^{\lambda T}2\lambda^{-1}(2\rho_{N_T}-1)$$
$$\times \left(\frac{1}{2}\mathcal{I}_1\left(\left(a(\bar{X}_1)-a(x)\right)\mathbb{1}_{\{\bar{X}_1\geq L\}}\right)\partial_x P_{T-\zeta_1}f(\bar{X}_1)\right.$$
$$\left. + \mathcal{I}_1\left(b(\bar{X}_1)\mathbb{1}_{\{\bar{X}_1\geq L\}}\right)P_{T-\zeta_1}f(\bar{X}_1)\right)\mathbb{1}_{\{N_T=1\}}$$
$$\overset{\mathbb{E}}{=} f(\bar{X}_{N_T+1})\theta_{N_T+1}\mathbb{1}_{\{N_T=0\}} + e^{\lambda T}2\lambda^{-1}(2\rho_{N_T}-1)$$
$$\times \left(\frac{1}{2}\mathcal{I}_1\left(\left(a(\bar{X}_1)-a(x)\right)\right)\partial_x P_{T-\zeta_1}f(\bar{X}_1)\right.$$
$$\left. + \mathcal{I}_1\left(b(\bar{X}_1)\right)P_{T-\zeta_1}f(\bar{X}_1)\right)\mathbb{1}_{\{\bar{X}_1\geq L\}}\mathbb{1}_{\{N_T=1\}}$$

where we used the extraction formula (2.10) applied to the r.v. $\mathbb{1}_{\{\bar{X}_s\geq L\}}$ and Lemma 10.5 (taking first the conditional expectation w.r.t ζ_1) for $\ell = 1$, $k = 0$ for the last equality.

Now we carry out again the same procedure for the second term in the above integrand using again the extraction formula (2.10) and the fact that $P_t f(L) = f(L) = 0$. We obtain

$$P_{T-t}f(\bar{Y}_t)\mathbb{1}_{\{\bar{\tau}>t\}}$$
$$\overset{\mathbb{E}}{=} f(\bar{X}_{N_T+1})\theta_{N_T+1}\mathbb{1}_{\{N_T=0\}} + e^{\lambda T}2\lambda^{-1}(2\rho_{N_T}-1)$$
$$\times \left(\frac{1}{2}\mathcal{I}_1^2\left(a(\bar{X}_1)-a(x)\right)P_{T-\zeta_1}f(\bar{X}_1) + \mathcal{I}_1\left(b(\bar{X}_1)\right)P_{T-\zeta_1}f(\bar{X}_1)\right)\mathbb{1}_{\{\bar{X}_1\geq L\}}\mathbb{1}_{\{N_T=1\}}$$

(10.2)

We emphasize that, under assumption **(H)**, the following estimates hold: for all $p \geq 1$, there exists $C := C(a,b,T,p)$ such that

$$\|\mathbb{1}_{\{\bar{X}_1\geq L\}}\mathcal{I}_1(a(\bar{X}_1)-a(x))\|_{p,0,1} \leq C, \qquad (10.3)$$
$$\|\mathbb{1}_{\{\bar{X}_1\geq L\}}\mathcal{I}_1^2(a(\bar{X}_1)-a(x))\|_{p,0,1} + \|\mathbb{1}_{\{\bar{X}_1\geq L\}}\mathcal{I}_1(b(\bar{X}_1))\|_{p,0,1} \leq \zeta_1^{-1/2}$$

which in turn, by using the fact that on $\mathbb{P}(N_T = 1, \zeta_1 \in dt) = \lambda e^{-\lambda T}dt$ on $[0,T]$, lead to the integrability of the second term appearing in (10.2).

We now prove the three estimates (10.3). Since they are obtained using the same technique, we only briefly explain how to obtain the first inequality. Using the definition of $\mathcal{I}_1$, one has that

$$\mathcal{I}_1(a(\bar{X}_1)-a(x)) = (a(\bar{X}_1)-a(x))\frac{Z_1}{\sigma(x)\zeta_1} + a'(\bar{X}_1).$$

Therefore in order to bound the above expression, we need to find an upper bound for $|a(\bar{X}_1)-a(x)|$ and then use classical estimates for the moments of Gaussian r.v.'s. In order to bound $|a(\bar{X}_1)-a(x)|$, one uses the Lipschitz property of a as follows

$$\mathbb{1}_{\{\bar{X}_1\geq L\}}|a(\bar{X}_1)-a(x)| \leq C\mathbb{1}_{\{\bar{X}_1\geq L\}}|\bar{X}_1-x| \leq C|\bar{X}_1-x\rho_1-(1-\rho_1)(2L-x)| \leq C|Z_1|.$$

Here we have used the fact that $\bar{X}_1 \geq L$ and $x \geq L$ implies $|\bar{X}_1 - x| \leq |\bar{X}_1 - x\rho - (1-\rho)(2L-x)|$. Finally the estimate (10.3) follows from the application of Lemma 2.4 with $H_1 = a(\bar{X}_1) - a(x)$.

---

[8]An alternative but longer approach can also be achieved using smooth approximations for the indicator function.

With the notations introduced in Section 3, (2.5) and (3.2), the identity (10.2) can be rewritten as

$$P_T f(x) = \mathbb{E}\left[f(\bar{X}_{N_T+1}) \mathbb{1}_{D_{N_T+1,N_T}} \bar{\theta}_{N_T+1} \mathbb{1}_{\{N_T=0\}} \right]$$
$$+ e^{\lambda T} \mathbb{E}\left[P_{T-\zeta_1} f(\bar{X}_1) \mathbb{1}_{D_{N_T,N_T}} \bar{\theta}_1 \mathbb{1}_{\{N_T=1\}} \right]. \tag{10.4}$$

Using again (10.1), applying Lemma 10.7 and by finally performing IBPs as before, we obtain by induction the following formula, $\forall n \in \mathbb{N}$, .

$$P_T f(x) = \sum_{j=0}^{n-1} \mathbb{E}\left[f(\bar{X}_{N_T+1}) \prod_{i=1}^{N_T+1} \mathbb{1}_{D_{i,N_T}} \bar{\theta}_i \mathbb{1}_{\{N_T=j\}} \right]$$
$$+ e^{\lambda T} \mathbb{E}\left[P_{T-\zeta_n} f(\bar{X}_n) \prod_{i=1}^{n} \mathbb{1}_{D_{i,N_T}} \bar{\theta}_i \mathbb{1}_{\{N_T=n\}} \right]. \tag{10.5}$$

To conclude it remains to prove the absolute convergence of the first sum and the convergence to zero of the last term[9]. These two results follow directly from the boundedness of f and the following general estimates on the product of weights.

Lemma 10.1. *The r.v.'s $\bar{\theta}_i \in \mathbb{S}_{i,n}$, $i \in \mathbb{N}_{n+1}$ satisfy the time degeneracy estimates in Definition 3.3. More precisely, for any $p \geq 1$, there exists a (deterministic) constant $C > 0$ (which may depend on λ) such that for $i \in \mathbb{N}_n$*

$$\mathbb{1}_{D_{i-1,n}} \mathbb{E}_{i-1,n}\left[\mathbb{1}_{D_{i,n}} |\bar{\theta}_i|^p \right] \leq C(\zeta_i - \zeta_{i-1})^{-\frac{p}{2}}. \tag{10.6}$$

Furthermore, one has $\mathbb{1}_{D_{n-1,n}} \mathbb{E}_{n,n}\left[\mathbb{1}_{D_{n,n}} |\bar{\theta}_{n+1}|^p \right] \leq C$.

Proof. The proof for $i \leq N_T$ follows from Lemma 2.4 after noting that $c_2^i = \frac{1}{2}(a(\bar{X}_i) - a(\bar{X}_{i-1})) \in \mathbb{M}_i(\bar{X}, 1/2) \cap \mathbb{S}_i(\bar{X})$ and is based on the same arguments as in the proof of (10.3). In the case $i = N_T + 1$ the conclusion follows directly from the boundedness of the r.v. $\bar{\theta}_{N_T+1}$. This argument will be used repeatedly in order to obtain the so-called time degeneracy estimates. $\square$

Lemma 10.2. *Assume that the weights $(\bar{\theta}_i, \mathbb{S}_{i,n})$, $i \in \mathbb{N}_{n+1}$ satisfy the time degeneracy estimates in Definition 3.3. Then for any $p \in [0, 2)$, we have the following moment estimate:*

$$\mathbb{E}\left[\left| \prod_{i=1}^{N_T+1} \mathbb{1}_{D_{i,N_T}} \bar{\theta}_i \right|^p \right] \leq E_{1-\frac{p}{2},1}(CT^{-\frac{p}{2}+1}) < \infty.$$

Here $E_{1-\frac{p}{2},1}$ stands for the Mittag-Leffler function $E_{\alpha,\beta}(z) := \sum_{n\geq 0} \frac{z^k}{\Gamma(\beta+k\alpha)}$ with parameters $\alpha = 1 - \frac{p}{2}$, $\beta = 1$.

Proof. For the proof, it is enough to use the Markov property of the Markov chain $\bar{X}$ together with the time degeneracy estimates and the fact that given $N_T = n$, the jump times of the Poisson process are distributed as the order statistics of n i.i.d. uniform $[0, T]$-valued r.v.'s satisfying $\mathbb{P}(N_T = n, \zeta_1 \in dt_1, \cdots, \zeta_n \in dt_n) = \lambda^n e^{-\lambda T} dt_1, \cdots, dt_n$ on the set $A_n = \{(t_1, \cdots, t_n) \in [0, T]^n : 0 < t_1 < \cdots < t_n < T\}$. This gives

$$\mathbb{E}\left[\left| \prod_{i=1}^{N_T+1} \mathbb{1}_{D_{i,N_T}} \bar{\theta}_i \right|^p \mathbb{1}_{\{N_T=n\}} \right] \leq C^{n+1} \int_{A_n} \prod_{i=1}^{n} (t_i - t_{i-1})^{-\frac{p}{2}} dt_1 \cdots dt_n$$
$$= C^{n+1} T^{n(1-\frac{p}{2})} \frac{\Gamma^n(1-\frac{p}{2})}{\Gamma(1+n(1-\frac{p}{2}))}$$

---

[9]An analytical proof of this fact can be found in [14], Lemma 5.2.

for some positive constant $C := C(T, a, b, p, \lambda)$. One concludes the proof by adding the previous inequality from $n = 0$ to infinity. $\qquad\square$

As a consequence, (3.1) holds for bounded smooth functions. The extension to bounded continuous functions follows from an approximation argument while the extension to bounded measurable function follows from a monotone class argument. The extension to polynomially growing functions is performed by a limit argument noting that all $L^p(\mathbb{P})$ moments of $\bar{X}_{N_T+1}$ and X_T are bounded for $1 \le p < \infty$.

Remark 10.3. (i) From the above proof, it should be clear that inequality (10.6) (or equivalently (3.5)) is strongly tied with the restriction $p \in [0, 2)$ in Lemma 10.1. This is why it is important to always have weights which satisfy a time degeneracy estimate as stated in Definition 3.3.

(ii) On the other hand, the estimate in (10.6), as remarked in [7], also suggests that the variance of the Monte Carlo estimator may be infinite. In order to achieve finite variance, one can resort to an importance sampling technique on the jump times as proposed by Anderson and Kohatsu-Higa [3].

10.2 Proof of the transfer of derivatives of Lemma 4.1

First, we remark that from the explanation given after the definition of the operators $\mathcal{D}$ and $\mathcal{I}$ in (2.8), the r.v.'s $(\overset{\leftarrow}{\theta}{}^e_{i+1}, \overset{\leftarrow}{\theta}{}^c_{i+1}, \overset{\leftarrow}{\theta}{}^\partial_{i+1})$ defined by (4.2) belongs to $\mathbb{S}_{i+1}(\bar{X})$. Then, it is clear that since the coefficients b and σ are smooth as stated in (**H**) and the transition of the Markov chain $\bar{X}$ has a smooth Gaussian transition density, the map $x \mapsto \mathbb{E}_{i,n}[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\overset{\leftarrow}{\theta}{}^e_{i+1}]\big|_{\bar{X}_i = x}$ is in $\mathscr{C}^1_p(\mathbb{R})$ a.s.

In order to prove (4.1), we consider the difference between the term appearing on the left-hand side and the first term appearing on the right-hand side of equation (4.1). Using the integration by parts formula (2.9) and then the extraction formula (2.10), we get

$$
\begin{aligned}
\mathbb{E}_{i,n}[\partial_{\bar{X}_{i+1}} f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\bar{\theta}_{i+1}] ={}& \mathbb{E}_{i,n}[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\mathcal{I}_{i+1}(\bar{\theta}_{i+1})] \\
& - \mathbb{E}_{i,n}[f(\bar{X}_{i+1})\delta_L(\bar{X}_{i+1})\bar{\theta}_{i+1}].
\end{aligned}
$$

From (2.7) and the integration by parts formula (2.9), we also obtain

$$
\begin{aligned}
\partial_{\bar{X}_i}\mathbb{E}_{i,n}[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\overset{\leftarrow}{\theta}{}^e_{i+1}] ={}& \mathbb{E}_{i,n}[\partial_{\bar{X}_i}(f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}})\overset{\leftarrow}{\theta}{}^e_{i+1}] \\
& + \mathbb{E}_{i,n}[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\partial_{\bar{X}_i}\overset{\leftarrow}{\theta}{}^e_{i+1}] \\
={}& \mathbb{E}_{i,n}[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\mathcal{I}_{i+1}([(2\rho_{i+1} - 1) + \sigma'_i Z_{i+1}]\overset{\leftarrow}{\theta}{}^e_{i+1})] \\
& + \mathbb{E}_{i,n}[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\partial_{\bar{X}_i}\overset{\leftarrow}{\theta}{}^e_{i+1}].
\end{aligned}
$$

By combining the two previous computations, we see that the difference

$$
\mathbb{E}_{i,n}[\partial_{\bar{X}_{i+1}} f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\bar{\theta}_{i+1}] - \partial_{\bar{X}_i}\mathbb{E}_{i,n}[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\overset{\leftarrow}{\theta}{}^e_{i+1}]
$$

can be expressed as

$$
\begin{aligned}
& \mathbb{E}_{i,n}[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\mathcal{I}_{i+1}(\bar{\theta}_{i+1} - (2\rho_{i+1} - 1)\overset{\leftarrow}{\theta}{}^e_{i+1})] - \mathbb{E}_{i,n}[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\partial_{\bar{X}_i}\overset{\leftarrow}{\theta}{}^e_{i+1}] \\
& - \mathbb{E}_{i,n}[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\mathcal{I}_{i+1}(\sigma'_i Z_{i+1}\overset{\leftarrow}{\theta}{}^e_{i+1})] - \mathbb{E}_{i,n}[f(\bar{X}_{i+1})\delta_L(\bar{X}_{i+1})\bar{\theta}_{i+1}]. \qquad (10.7)
\end{aligned}
$$

Note that by using the relation $d_1^{i+1} = c_1^{i+1} - (2\rho_{i+1} - 1)\partial_{\bar{X}_i}c_2^{i+1}$ in (4.3) and Lemma

2.5 we obtain

$$\partial_{\bar{X}_i} \overleftarrow{\theta}^e_{i+1} = 2\lambda^{-1} \left(\mathcal{I}^2_{i+1}(\partial_{\bar{X}_i} c^{i+1}_2) + \mathcal{I}_{i+1}(\partial_{\bar{X}_i} d^{i+1}_1) - (\sigma'\sigma^{-1})_i \left(\mathcal{I}_{i+1}(d^{i+1}_1) + 2\mathcal{I}^2_{i+1}(c^{i+1}_2) \right) \right),$$
$$\partial_{\bar{X}_i} c^{i+1}_2 = a'_{i+1} \partial_{\bar{X}_i} \bar{X}_{i+1} - a'_i,$$
$$\partial^2_{\bar{X}_i} c^{i+1}_2 = a''_{i+1} (\partial_{\bar{X}_i} \bar{X}_{i+1})^2 + a'_{i+1} \partial^2_{\bar{X}_i} \bar{X}_{i+1} - a''_i,$$
$$\partial_{\bar{X}_i} d^{i+1}_1 = b'_{i+1} \partial_{\bar{X}_i} \bar{X}_{i+1} - (2\rho_{i+1} - 1)\partial^2_{\bar{X}_i} c^{i+1}_2,$$

which in turn, after some algebraic simplifications, yields the formula for $\overleftarrow{\theta}^c_{i+1}$ appearing in (4.2). In order to conclude the proof of (4.1), it remains to notice that for $f \in \mathscr{C}^1_p([L, \infty))$, one has

$$-\mathbb{E}_{i,n} \left[f(\bar{X}_{i+1})\delta_L(\bar{X}_{i+1})\bar{\theta}_{i+1} \right] = \mathbb{E}_{i,n} \left[f(\bar{X}_{i+1})\delta_L(\bar{X}_{i+1})\overleftarrow{\theta}^\partial_{i+1} \right] \qquad (10.8)$$

which is a direct consequence of Corollary 10.6.

The time degeneracy estimates of the r.v.'s $(\overleftarrow{\theta}^e_{i+1}, \overleftarrow{\theta}^c_{i+1}, \overleftarrow{\theta}^\partial_{i+1})$ are straightforward applications of Lemma 2.4 using the same arguments as in (10.3). The proof of the transfer of derivative formula (4.5) as well as the time degeneracy estimates of the r.v.'s for $\overleftarrow{\theta}^a_{n+1}$, $a \in \{e, c\}$ on the set $\{N_T = n\}$ follow from similar arguments and we omit the remaining technical details. $\qquad \square$

10.3 Proof of the boundary merging lemmas

10.3.1 A time convolution result for Gaussians

In this section, we describe a series of explicit calculations of convolutions of Gaussian densities with respect to the time variable. These are used when merging weights and building the boundary process.

Lemma 10.4. *For $x, y, \alpha, \beta \in \mathbb{R}_+$ and $\ell \in \{0, 1, 2\}$,*

$$\int_0^t \partial_x g(\alpha^2 s, x) \partial^\ell_y (g(\beta^2(t-s), y)) ds = -\alpha^{-2} \partial^\ell_y (g(\beta^2 t, y + \frac{\beta}{\alpha} x)). \qquad (10.9)$$

Furthermore

$$\int_0^t s\partial_x g(\alpha^2 s, x) \partial^\ell_y (g(\beta^2(t-s), y)) ds = -\alpha^{-3}\beta^{-1} x \partial^\ell_y (\bar{\Phi}(\beta^2 t, y + \frac{\beta}{\alpha} x)) \qquad (10.10)$$

where $\bar{\Phi}(t, z) := \int_{|z|}^\infty g(t, y) dy$.

Proof. The results follows from simple manipulations using the following Laplace transform for $x, \eta > 0$:

$$\mathcal{L}(g_{-1})(\eta) := \int_0^\infty e^{-\eta s} g_{-1}(s)\, ds = x^{-1} e^{-\sqrt{2\eta x^2}} \text{ with } g_{-1}(s) := s^{-1} g(s, x). \qquad (10.11)$$

In particular, one uses the fact that $s\partial_x g(a^2 s, x) = -\frac{x}{a^2} g(a^2 s, x)$ in proving (10.10). $\qquad \square$

10.3.2 Proof of Lemma 4.3

The key ingredient to prove (4.9), that is, the boundary merging of $\overleftarrow{\theta}^e_{i+1}$ and $\overleftarrow{\theta}^\partial_i$, is Lemma 10.4. We divide the proof into several steps.

Step 1: Simplify the statement. First, one observes that it is easier to verify (4.9) for each term appearing in $\overleftarrow{\theta}^e_{i+1}$ and $\overleftarrow{\theta}^\partial_i$ in (4.2), after proper extraction and localization as

in Corollary 10.6, and then add them together. Therefore proving the following general statement will suffice: For any $j = 0, 1, 2$ and real-valued $f \in \mathscr{C}_p^0(\mathbb{R}^2)$,

$$\mathbb{E}\big[f(\bar{X}_i, \bar{X}_{i+1}) \mathbb{1}_{D_{i+1,n}} \mathcal{I}_{i+1}^j(1) \delta_L(\bar{X}_i)(2\rho_i - 1)\mathcal{I}_i(1) \,\big|\, \mathcal{G}_{i-1}, \zeta_{i+1}, N_T = n\big] \qquad (10.12)$$

$$= \frac{\lambda^{-1}}{\zeta_{i+1} - \zeta_{i-1}} \mathbb{E}\big[f(L, \bar{X}_{i-1,i+1}^\partial) \mathbb{1}_{D_{i-1,i+1,n}^\partial} a_{i-1}^{-1} \bar{\mathcal{I}}_{i-1,i+1}^j(1) \,\big|\, \mathcal{G}_{i-1}, \zeta_{i+1}, N_T = n\big].$$

Step 2: Proof of (10.12). *Case* $j = 0$. In order to foster the understanding of the proof, let us first consider the case $j = 0$. The following arguments also gives rise to the definition of the merged boundary process.

First, observe that as $(2\rho_i - 1)\sigma_{i-1}Z_i = L - \bar{X}_{i-1} := Y_{i-1}$ on the set $\{\bar{X}_i = L\}$, one has

$$(2\rho_i - 1)\delta_L(\bar{X}_i)\mathcal{I}_i(1) = \delta_L(\bar{X}_i)\frac{Y_{i-1}}{a_{i-1}(\zeta_i - \zeta_{i-1})}.$$

Next, we rewrite the right hand side of (10.12), namely, using the independence of W and N, we have

$$\mathbb{E}\big[f(\bar{X}_i, \bar{X}_{i+1}) \mathbb{1}_{D_{i+1,n}} \delta_L(\bar{X}_i)(2\rho_i - 1)\mathcal{I}_i(1) \,\big|\, \mathcal{G}_{i-1}, \zeta_{i+1}, \zeta_i, \rho_i, N_T = n\big]$$

$$= \int_L^\infty f(L, z) \frac{Y_{i-1}}{a_{i-1}(\zeta_i - \zeta_{i-1})} g(a_{i-1}(\zeta_i - \zeta_{i-1}), Y_{i-1}) \, g(a(L)(\zeta_{i+1} - \zeta_i), z - L) \, dz.$$

We now use the following key property: for a real-valued bounded $\mathcal{G}_{i-1} \times \mathcal{B}(\mathbb{R}_+^2)$-measurable function $h : \Omega \times \mathbb{R}_+^2 \to \mathbb{R}$, one has

$$\mathbb{E}[h(\zeta_i, \zeta_{i+1}) \,|\, \mathcal{G}_{i-1}, \zeta_{i+1}, N_T = n] = \mathbb{E}[h(s + U, t) \,|\, \mathcal{G}_{i-1}, \zeta_{i+1}, N_T = n]\Big|_{s = \zeta_{i-1}, t = \zeta_{i+1}},$$
$$(10.13)$$

where $U \sim \mathcal{U}(0, t - s)$ is independent of $\mathcal{G}_{i-1}$ and N. More precisely, by using the tower property for conditional expectation and (10.13), we obtain

$$\mathbb{E}\big[f(\bar{X}_i, \bar{X}_{i+1}) \mathbb{1}_{D_{i+1,n}} \delta_L(\bar{X}_i)(2\rho_i - 1)\mathcal{I}_i(1) \,\big|\, \mathcal{G}_{i-1}, \zeta_{i+1}, \rho_i, N_T = n\big] = \frac{1}{\zeta_{i+1} - \zeta_{i-1}} \times$$

$$\int_L^\infty dz \int_0^{\zeta_{i+1} - \zeta_{i-1}} f(L, z) \frac{Y_{i-1}}{a_{i-1}s} g(a_{i-1}s, Y_{i-1}) g(a(L)((\zeta_{i+1} - \zeta_{i-1}) - s), z - L) \, ds.$$

In order to compute the above time convolution and show that the above conditional expectation can be rewritten using the boundary process $\bar{X}^\partial$, we apply Lemma 10.4, in particular (10.9), noticing that $\frac{Y_{i-1}}{a_{i-1}s} g(a_{i-1}s, Y_{i-1}) = \partial_2 g(a_{i-1}s, |Y_{i-1}|)$, to obtain

$$\mathbb{E}\big[f(\bar{X}_i, \bar{X}_{i+1}) \mathbb{1}_{D_{i+1,n}} \delta_L(\bar{X}_i)(2\rho_i - 1)\mathcal{I}_i(1) \,\big|\, \mathcal{G}_{i-1}, \zeta_{i+1}, \rho_i, N_T = n\big]$$

$$= \frac{(\zeta_{i+1} - \zeta_{i-1})^{-1}}{a_{i-1}} \int_L^\infty f(L, z) g(a(L)(\zeta_{i+1} - \zeta_{i-1}), z - (L(1 - \mu(\bar{X}_{i-1})) + \bar{X}_{i-1}\mu(\bar{X}_{i-1}))) dz$$

$$= \frac{(\zeta_{i+1} - \zeta_{i-1})^{-1}}{a_{i-1}} \mathbb{E}\big[f(L, \bar{X}_{i-1,i+1}^\partial) \mathbb{1}_{D_{i-1,i+1,n}^\partial} \,\big|\, \mathcal{G}_{i-1}, \zeta_{i+1}, N_T = n\big].$$

This finishes the proof for the case $j = 0$.

Step 3: The general case: $j = 1, 2$. Using equality (10.9), one can similarly prove (10.12) in the case $j = 1, 2$. In fact, a useful formula to prove this general case directly from (10.9) is the following property for $i \in \mathbb{N}_n$ which was stated in (2.11).

$$\mathcal{I}_{i+1}^j(1) = (-1)^j \left(g^{-1}\partial_y^j g\right)\left(a_i(\zeta_{i+1} - \zeta_i), y - \rho_{i+1}\bar{X}_i - (1 - \rho_{i+1})(2L - \bar{X}_i)\right)\Big|_{y = \bar{X}_{i+1}}.$$

Step 4: The general case follows by linearity. Note that in the general case of (4.9), we have that $\overleftarrow{\theta}^e_{i+1}$ is given by (4.2). This expression can be rewritten using the extraction formula (2.10) as linear combinations of terms of the type $f_j(\bar{X}_i, \bar{X}_{i+1})\mathcal{I}^j_{i+1}(1)$ for some particular functions $f_j \in \mathscr{C}(\mathbb{R}^2)$. Therefore the result in (10.12) applies.

One uses again (2.10) for $\bar{\mathcal{I}}$ in order to rewrite the resulting expressions in a compact form. For this, note that we have conveniently defined $\bar{\mathcal{D}}$ as the adjoint of $\bar{\mathcal{I}}$ for which the extraction formulae are satisfied.

Therefore towards obtaining the formulae which appear in the definition of the weights $\overleftarrow{\theta}^{\partial*e}$ one combines linearly (10.12) to obtain::

$$\mathbb{E}\big[f(\bar{X}_{i+1})\mathbb{1}_{D_{i+1,n}}\overleftarrow{\theta}^e_{i+1}\delta_L(\bar{X}_i)(2\rho_i - 1)\mathcal{I}_i(1)\big|\mathcal{G}_{i-1},\zeta_{i+1},N_T = n\big] = \tag{10.14}$$
$$\frac{2}{\lambda}\mathbb{E}\Big[\ f(\bar{X}^\partial_{i-1,i+1})\mathbb{1}_{D^\partial_{i-1,i+1,n}}$$
$$\times\ \Big\{\frac{1}{a_{i-1}(\zeta_{i+1} - \zeta_{i-1})}\big(\bar{\mathcal{I}}^2_{i-1,i+1}(\bar{d}^{i+1}_2) + \bar{\mathcal{I}}_{i-1,i+1}(\bar{c}^{i+1}_3)\big)\Big\}\Big|\mathcal{G}_{i-1},\zeta_{i+1},N_T = n\Big].$$

Here, $\bar{c}^{i+1}_3$ takes the form

$$\bar{c}^{i+1}_3 := b(\bar{X}^\partial_{i-1,i+1}) - (2\rho_{i+1} - 1)(a'(\bar{X}^\partial_{i-1,i+1})((2\rho_{i+1} - 1)$$
$$+ \sigma'(\bar{X}^\partial_{i-1,i+1})(Z_{i+1} + \Delta Z_i)) - a'(L))$$

and belongs to $\mathbb{S}_{i-1,i+1}(\bar{X}^\partial)$.

To continue simplifying the above expression and to obtain $\bar{d}^{i+1}_1$ from $\bar{c}^{i+1}_3$, note that except for the factor $(2\rho_{i+1} - 1)$ in the coefficient $\bar{c}^{i+1}_3$ all expressions on the right hand side of the above equality do not depend on ρ_{i+1}. In particular, note that the r.v. $\bar{X}^\partial_{i-1,i+1}$ does not depend on ρ_{i+1}. Therefore by the symmetry of the Bernoulli r.v. $2\rho_{i+1} - 1$, the term $(a'\sigma')(\bar{X}^\partial_{i-1,i+1})(Z_{i+1} + Z_i) - a'(L)$ in $\bar{c}^{i+1}_3$ vanishes, leading to the definition of $\bar{d}^{i+1}_1$. From here, (4.9) follows by multiplying (10.14) by $2\lambda^{-1}(a'(L) - b(L))$ in order to obtain $\overleftarrow{\theta}^\partial$ on the left side of the equation.

Finally, the equality (4.10) follows by arguments similar to those employed in (10.12) for $j = 0$ and we omit its proof.

Step 5: The time degeneracy estimates. Using the basic properties for Skorokhod integral, obtained mutatis mutandis from Lemma 2.4, one obtains the time degeneracy inequalities as it was done in Lemma 10.1 using the estimate (10.3). On the set $\{\bar{X}_{i-1} \geq L\}$, one obtains that $|\bar{d}^{i+1}_2(L,\bar{X}_{i-1,i+1})| \leq C|Z_{i+1} + Z_i|$ so that

$$\mathbb{1}_{D_{i-1,n}}\|\mathbb{1}_{D^\partial_{i-1,i+1,n}}\overleftarrow{\theta}^{\partial*j}_{i-1,i+1}\|_{p,i-1} \leq C(\zeta_{i+1} - \zeta_{i-1})^{-1/2}.$$

This ends the proof of Lemma 4.3. □

One may heuristically understand why the above merged structure appears. Actually, the Dirac delta function $\delta_L(\bar{X}_i)$ appearing inside the conditional expectation (10.12) imposes the approximation process to go from $\bar{X}_{i-1}$ to L and then from L to $\bar{X}_i$. The two corresponding Gaussian laws on each interval are convolved in time in the conditional expectation through the random time ζ_i. Therefore, using the branching property of Gaussian kernels, stated in Lemma 10.4 in its analytic form[10], one obtains the results after renormalising the variances. This is the argument used in the above proof.

10.3.3 Proof of Lemma 4.4

As this proof has many intersections with the proof of Lemma 4.3, we only indicate the main points. Essentially we need to perform two steps. First, one simplifies (using the

---

[10]Clearly, this property is connected with a similar one for Bessel processes as stated in Chapter XI of [29].

reduction formula of Corollary 10.6) the expression of $\bar{\theta}_i$ using the fact that one has $\delta_L(\bar{X}_i)$ in the expression (4.11) as was done in the proof of (10.8). Then one follows the lines of reasoning as in the proof of Lemma 4.3, except that one needs to use (10.10) instead of (10.9).

We remark here that the time degeneration estimate in Lemma 4.4 improves in comparison with its corresponding estimate in Lemma 4.3 because of the inequalities

$$|\bar{X}_{i-1} - L| \vee |\bar{X}^{\partial}_{i-1,i+1} - L| \le C|\bar{X}^{\partial}_{i-1,i+1} - L(1 - \mu(\bar{X}_{i-1})) - \bar{X}_{i-1}\mu(\bar{X}_{i-1})| \le C|Z_{i+1} + Z_i|$$

and

$$(\bar{\Phi}g^{-1})(a(L)(\zeta_{i+1} - \zeta_{i-1}), Z_{i+1} + Z_i) \le C(\zeta_{i+1} - \zeta_{i-1})^{1/2},$$

where the last inequality follows from change of variables and Komatsu's inequality and in the first we assume that $\bar{X}_{i-1}, \bar{X}^{\partial}_{i-1,i+1} \ge L$. The proofs corresponding to the last interval are easier and follow similar arguments. This concludes the proof of Lemma 4.4. $\qquad\square$

10.4 Localization and reduction lemmas

The following lemma is the basic result which explains that the expectation of a weight can be simplified thanks to the symmetry of the law of $\bar{X}_i$, for $1 \le i \le N_T$. For this reason, we call the following result a reduction lemma. This section only uses the results and setting of Section 2.

Lemma 10.5. *Let $f \in \mathscr{C}^1_p(\mathbb{R})$ and $(\ell, k, n) \in \bar{\mathbb{N}} \times \mathbb{N}^2$. Then, for any $i \in \mathbb{N}_n$, one has*

$$\mathbb{E}\left[f(\bar{X}_i)\delta_L(\bar{X}_i)(2\rho_i - 1)^{\ell}\mathcal{I}^k_i(1)|\mathcal{G}_{i-1}, T^{n+1}, N_T = n\right]$$
$$= \begin{cases} 0, & \text{if } \ell + k \text{ is odd,} \\ \mathbb{E}\left[f(\bar{X}_i)\delta_L(\bar{X}_i)(2\rho_i - 1)^k\mathcal{I}^k_i(1) \,|\, \mathcal{G}_{i-1}, T^{n+1}, N_T = n\right], & \text{if } \ell + k \text{ is even.} \end{cases}$$

Proof. The proof follows by using (2.11) and noting that the Hermite polynomials of even degree are even functions (as functions of the variable Z_i) and the fact that, conditional on $\mathcal{G}_{i-1}$, T^{n+1} and $\bar{X}_i = L$, the law of the random vector $(Z_i, 2\rho_i - 1)$ is a Bernoulli(1/2) r.v. with symmetric values $\pm(\sigma^{-1}_{i-1}(L - \bar{X}_{i-1}), 1)$ for $i \in \mathbb{N}_n$. $\qquad\square$

The above lemma states that any weight for which its leading term satisfies that $\ell + k$ is odd (in the sense of highest order of time degeneration) will have a reduction in its time degeneration. Combining the extraction formula (2.10) with Lemma 2.10, one obtains a result which encompasses extraction, localization (at L) and reduction in time degeneration.

Corollary 10.6. *Let $f \in \mathscr{C}^1_p(\mathbb{R})$ and $(\ell, k, n) \in \mathbb{N}_0 \times \mathbb{N}^2$. Let $c \in \mathscr{C}^k_b(\mathbb{R})$. Then, for any $i \in \mathbb{N}_n$, one has*

$$\mathbb{E}\left[f(\bar{X}_i)\delta_L(\bar{X}_i)(2\rho_i - 1)^{\ell}\mathcal{I}^{\ell}_i(c(\bar{X}_i))|\mathcal{G}_{i-1}, T^{n+1}, N_T = n\right]$$
$$= \sum_{\substack{j=0 \\ \ell+k-j=even}}^{k} (-1)^j \binom{k}{j} c^{(j)}(L)\mathbb{E}\left[f(\bar{X}_t)\delta_L(\bar{X}_t)(2\rho - 1)^{k-j}\mathcal{I}^{k-j}_i(1)|\mathcal{G}_{i-1}, T^{n+1}, N_T = n\right].$$

10.5 Appendix: Results about emergence and reduction of jumps

The first result is used in the proof of the probabilistic representation in Theorem 3.1 and is used to express that time integrals add jumps to the Poisson process.

Lemma 10.7. *Let $n \in \mathbb{N}$ and $G : \{(t_1, \ldots, t_{n+2}) : 0 < t_1 < \cdots < t_{n+1} < t_{n+2} := T\} \to \mathbb{R}$ be a measurable function such that $\mathbb{E}[\int_{\zeta_n}^T |G(\zeta_1, \ldots, \zeta_n, s, T)| \mathbb{1}_{\{N_T = n\}} ds] < \infty$. Then*

$$\mathbb{E}[\int_{\zeta_n}^T G(\zeta_1, \ldots, \zeta_n, s, T) \mathbb{1}_{\{N_T = n\}} ds] = \lambda^{-1} \mathbb{E}[G(\zeta_1, \ldots, \zeta_n, \zeta_{n+1}, T) \mathbb{1}_{\{N_T = n+1\}}].$$

The proof is straightforward and follows by rewriting the above expectations using that the conditional law of the jump times of a Poisson process given the number of jumps follows the same law as the order statistics for i.i.d. uniformly distributed r.v.'s.

The next result is used for the reduction of jumps after the boundary merging procedure. For this reason we called it the time merging lemma.

Lemma 10.8. *Let $n \in \mathbb{N}$. Let $G : \{(t_1, ..., t_n) : 0 < t_1 < \cdots < t_n < T\} \to \mathbb{R}$ be a measurable function such that for all $i \in \mathbb{N}_n$,*

$$\mathbb{E}[(\zeta_{i+1} - \zeta_{i-1})^{-1} |G(\zeta_1, \ldots, \zeta_{i-1}, \zeta_{i+1}, \ldots, \zeta_{N_T+1})| \mathbb{1}_{\{N_T = n\}}] < \infty.$$

Then, for any $i \in \mathbb{N}_n$, one has

$$\mathbb{E}[(\zeta_{i+1} - \zeta_{i-1})^{-1} G(\zeta_1, \cdots, \zeta_{i-1}, \zeta_{i+1}, \ldots, \zeta_{N_T+1}) \mathbb{1}_{\{N_T = n\}}]$$
$$= \lambda \mathbb{E}[G(\zeta_1, \ldots, \zeta_{N_T+1}) \mathbb{1}_{\{N_T = n-1\}}].$$

Proof. The proof follows from standard computations having at hand the two following important facts: conditionally on the event $\{N_T = n, \zeta_1, \ldots, \zeta_{i-1}, \zeta_{i+1}, \ldots, \zeta_{N_T}\}$, the distribution of ζ_i is uniform on $[\zeta_{i-1}, \zeta_{i+1}]$ and $n\mathbb{P}(N_T = n) = \lambda T \mathbb{P}(N_T = n-1)$. $\square$

Remark 10.9. The above result will be used in the proof of Theorem 5.1 for $i \in \mathbb{N}_n$ on the set $\{N_T = n\}$ for some measurable functions G of the following form

$$G(\zeta_1, \ldots, \zeta_{i-1}, \zeta_{i+1}, \ldots, \zeta_{N_T}) = F(\bar{X}_{i-1,i+1}^\partial, \zeta_{i+1}, \cdots, \zeta_{n+1}) \mathbb{1}_{D_{i-1,i+1,n}^\partial} \overleftarrow{\theta}_{i-1,i+1}^{a_i} \prod_{j=1}^{i-1} \mathbb{1}_{D_{j,n}} \bar{\theta}_j$$

for some symbols $a_i \in \{\partial * e, \partial \circledast e\}$ and for some measurable function $F(x, s_{i+1}, \cdots, T)$.

References

[1] A. Agarwal and E. Gobet, *Finite variance unbiased estimation of stochastic differential equations*, Proceedings of the 2017 Winter Simulation Conference (2017), 1950–1961.

[2] J. Akahori and Y. Imamura, *On a symmetrization of diffusion processes*, Quantitative Finance **14** (2014), no. 7, 1211–1216. MR-3219709

[3] P. Andersson and A. Kohatsu-Higa, *Unbiased simulation of stochastic differential equations using parametrix expansions*, Bernoulli **23** (2017), no. 3, 2028–2057. MR-3624885

[4] P. Andersson, A. Kohatsu-Higa, and T. Yuasa, *Second order probabilistic parametrix method for unbiased simulation of stochastic differential equations*, Tech. report, 2018.

[5] M. Arnaudon and A. Thalmaier, *Gradient estimates for positive harmonic functions by stochastic analysis*, Stochastic Processes and their Applications **117** (2007), no. 2, 202–220. MR-2290193

[6] M. Arnaudon and A. Thalmaier, *The differentiation of hypoelliptic diffusion semigroups*, Illinois J. Math. **54** (2010), no. 4, 1285–1311. MR-2981848

[7] V. Bally and A. Kohatsu-Higa, *A probabilistic interpretation of the parametrix method*, Ann. Appl. Probab. **25** (2015), no. 6, 3095–3138. MR-3404632

[8] G. Bernis, E. Gobet, and A. Kohatsu-Higa, *Monte Carlo evaluation of Greeks for multidimensional barrier and lookback options*, Math. Finance **13** (2003), no. 1, 99–113, Conference on Applications of Malliavin Calculus in Finance (Rocquencourt, 2001). MR-1968099

[9] A. Beskos, O. Papaspiliopoulos, and G. O. Roberts, *Retrospective exact simulation of diffusion sample paths with applications*, Bernoulli **12** (2006), no. 6, 1077–1098. MR-2274855

[10] M. Defosseux, *Affine lie algebras and conditioned space-time brownian motions in affine weyl chambers*, Probab. Theory Relat. Fields (2016), 649–665. MR-3520015

[11] F. Delarue, *Estimates of the solutions of a system of quasi-linear PDEs. A probabilistic scheme*, Séminaire de Probabilités XXXVII, Lecture Notes in Math., vol. 1832, Springer, Berlin, 2003, pp. 290–332. MR-2053051

[12] E. Fournié, J.-M. Lasry, J. Lebuchoux, and P.-L. Lions, *Applications of Malliavin calculus to Monte-Carlo methods in finance. II*, Finance Stoch. **5** (2001), no. 2, 201–236. MR-1841717

[13] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, and N. Touzi, *Applications of Malliavin calculus to Monte Carlo methods in finance*, Finance Stoch. **3** (1999), no. 4, 391–412. MR-1842285

[14] N. Frikha, A. Kohatsu-Higa, and L. Li, *On the first hitting times of one dimensional elliptic diffusions*, arXiv:1609.09327 (2016).

[15] M.G. Garroni and J.-L. Menaldi, *Green functions for second order parabolic integro-differential problems*, Pitman Research Notes in Mathematics Series, vol. 275, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1992. MR-1202037

[16] M.G. Garroni and J.L. Menaldi, *Second order elliptic integro-differential problems*, Chapman & Hall/CRC Research Notes in Mathematics, vol. 430, Chapman & Hall/CRC, Boca Raton, FL, 2002. MR-1911531

[17] E. Gobet, *Weak approximation of killed diffusion using Euler schemes*, Stochastic Process. Appl. **87** (2000), no. 2, 167–197. MR-1757112

[18] E. Gobet and A. Kohatsu-Higa, *Computation of Greeks for barrier and look-back options using Malliavin calculus*, Electron. Comm. Probab. **8** (2003), 51–62. MR-1987094

[19] E. Gobet and S. Menozzi, *Exact approximation rate of killed hypoelliptic diffusions using the discrete Euler scheme*, Stoch. Proc. and Appl. **112** (2004), 210–223. MR-2073411

[20] P. Henry-Labordère, X. Tan, and N. Touzi, *Unbiased simulation of stochastic differential equations*, Ann. Appl. Probab. **27** (2017), no. 6, 3305–3341. MR-3737926

[21] S. Herrmann and C. Zucca, *Exact simulation of the first-passage time of diffusions*, Journal of Scientific Computing **79** (2019), no. 3, 1477–1504. MR-3946466

[22] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, second ed., North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989. MR-1011252

[23] P.A. Jenkins, *Exact simulation of the sample paths of a diffusion with a finite entrance boundary*, arXiv:1311.5777, 2013.

[24] I. Karatzas and S.E. Shreve, *Brownian motion and stochastic calculus*, 2 ed., Springer, 1991. MR-1121940

[25] P. Malliavin and A. Thalmaier, *Stochastic calculus of variations in mathematical finance*, Springer Finance, Springer Berlin Heidelberg, 2005. MR-2189710

[26] T. Nakatsu, *Integration by parts formulas concerning maxima of some SDEs with applications to study on density functions*, Stoch. Anal. Appl. **34** (2016), no. 2, 293–317. MR-3462138

[27] D. Nualart, *The Malliavin calculus and related topics*, second ed., Probability and its Applications (New York), Springer-Verlag, Berlin, 2006. MR-2200233

[28] V. Reutenauer and E. Tanré, *An unbiased Monte Carlo estimator for derivatives. Application to CIR*, Tech. report, arXiv:1609.07431, 2016.

[29] D. Revuz and M. Yor, *Continuous martingales and brownian motion*, Grundlehren der Mathematischen Wissensch Series, Springer, 1991. MR-1083357