



**HAL**  
open science

# Branching diffusion representation of semi-linear elliptic PDEs and estimation using Monte Carlo method

Ankush Agarwal, Julien Claisse

► **To cite this version:**

Ankush Agarwal, Julien Claisse. Branching diffusion representation of semi-linear elliptic PDEs and estimation using Monte Carlo method. Stochastic Processes and their Applications, 2020, 10.1016/j.spa.2020.02.009 . hal-01709033

**HAL Id: hal-01709033**

**<https://polytechnique.hal.science/hal-01709033v1>**

Submitted on 14 Feb 2018

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

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

# Branching diffusion representation of semi-linear elliptic PDEs and estimation using Monte Carlo method

Ankush Agarwal<sup>\*1</sup> and Julien Claisse<sup>2</sup>

<sup>1</sup>Adam Smith Business School, University of Glasgow, University Avenue, G128QQ Glasgow, United Kingdom. Email: [ankush.agarwal@glasgow.ac.uk](mailto:ankush.agarwal@glasgow.ac.uk)

<sup>2</sup>Centre de Mathématiques Appliquées (CMAP), École Polytechnique and CNRS, Route de Saclay, 91128 Palaiseau Cedex, France. Email: [julien.claisse@polytechnique.edu](mailto:julien.claisse@polytechnique.edu)

February 14, 2018

## Abstract

We study semi-linear elliptic PDEs with polynomial non-linearity and provide a probabilistic representation of their solution using branching diffusion processes. When the non-linearity involves the unknown function but not its derivatives, we extend previous results in the literature by showing that our probabilistic representation provides a solution to the PDE without assuming its existence. In the general case, we derive a new representation of the solution by using marked branching diffusion processes and automatic differentiation formulas to account for the non-linear gradient term. In both cases, we develop new theoretical tools to provide explicit sufficient conditions under which our probabilistic representations hold. As an application, we consider several examples including multi-dimensional semi-linear elliptic PDEs and estimate their solution by using the Monte Carlo method.

**Key words:** automatic differentiation formula, branching diffusion processes, elliptic, exit time, Monte Carlo method, partial differential equation, semi-linear

**AMS subject classifications (2010):** 35J61, 60H30, 60J85, 65C05

## 1 Introduction

In this paper, we are interested in the following class of semi-linear elliptic partial differential equations (PDEs): given a bounded domain  $\mathcal{O} \subset \mathbb{R}^d$ ,  $f : \mathcal{O} \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $h : \partial\mathcal{O} \rightarrow \mathbb{R}$ ,

$$\mathcal{L}u + f(u, Du) = 0 \text{ in } \mathcal{O}, \quad u = h \text{ on } \partial\mathcal{O}, \quad (1)$$

where  $\mathcal{L}$  is the infinitesimal generator of a diffusion. This class of PDEs arises naturally in many fields of science (see, *e.g.*, [2, 25] and references therein). In most cases, the related PDEs only involve non-linearity in the unknown function but not in its derivatives. For instance, the Poisson–Boltzmann equation or the steady states in nonlinear equations of the Klein–Gordon or Schrödinger type. Besides, the stationary Burgers equation or Hamilton–Jacobi–Bellman equations when the diffusion coefficient is uncontrolled are examples of PDEs where the non-linearity involves both the unknown function and its first order derivatives. In this paper, we provide existence results for this class of PDEs when the generator  $f$  is polynomial and derive a new probabilistic representation of the solution which is well suited for numerical application, especially in high dimensions, using the Monte Carlo method.

The classical probabilistic approach for semi-linear PDEs relies on the theory of backward stochastic differential equations (BSDEs) initiated by Pardoux and Peng [30]. The case of elliptic PDE with Dirichlet condition was first studied by Darling and Pardoux [10] for Lipschitz generator and later more general results were provided by Pardoux [29] and Briand *et al.* [8]. These results rely on the so-called monotonicity assumption, which requires  $y \mapsto f(x, y, z) - \mu y$  to be decreasing for some constant  $\mu \in \mathbb{R}$ , and the assumption that  $f$  has a linear or quadratic growth in  $z$ . We do not put such restrictions in our setting and propose a method to investigate a different class of elliptic PDEs with a generator which is a polynomial in  $y$  and  $z$ .

Our probabilistic representation is based on *branching diffusion processes*. These processes describe the evolution of a population of independent and identical particles moving according to a diffusion process. They were first introduced by Skorokhod [34] and later studied, more thoroughly and systematically, in a

series of papers by Ikeda *et al.* [20, 21, 22]. In particular, these authors, using branching diffusion processes, established a probabilistic representation of semi-linear parabolic PDEs of the form

$$\partial_t u + \mathcal{L}u + \beta \left( \sum_{l \in \mathbb{N}} p_l u^l - u \right) = 0 \text{ in } [0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = g \text{ in } \mathbb{R}^d, \quad (2)$$

where  $g : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $\beta > 0$  and  $(p_l)_{l \in \mathbb{N}}$  is a probability mass function. More precisely, they consider a process where each particle moves according to a diffusion with generator  $\mathcal{L}$  and dies at an exponentially distributed random time with parameter  $\beta$  to give birth to  $l$  offsprings with probability  $p_l$ . Then, the solution of PDE (2) is expressed as  $\mathbb{E}[\prod_{i=1}^{N_T} g(X_T^i)]$  where  $N_T$  and  $(X_T^i)_{i=1, \dots, N_T}$  denote the number of particles and their positions at time  $T$  respectively. The special case  $p_2 = 1$ , which corresponds to the celebrated Fisher–KPP equation, has been particularly well-studied (see, *e.g.*, McKean [26]).

In Henry-Labordère *et al.* [18], the authors extended the probabilistic representation of (path-dependent) PDE (2) where the probability mass function  $(p_l)_{l \in \mathbb{N}}$  is replaced by an arbitrary real-valued sequence of functions. See also Rasulov *et al.* [32]. This result was further extended in Henry-Labordère *et al.* [19] to parabolic semi-linear PDE of the form

$$\partial_t u + \mathcal{L}u + f(u, Du) = 0 \text{ in } [0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = g \text{ in } \mathbb{R}^d, \quad (3)$$

where  $f$  is a  $(d + 1)$ -variate polynomial. The main idea consists of introducing marked particles which carry a weight function given by the Bismuth-Elworthy-Li formula to account for non-linearity in gradient of the solution. Under conditions of “small non-linearity” or small maturity  $T$ , the authors showed that their probabilistic representation provides a continuously differentiable viscosity solution to PDE (3). They also performed numerical simulations to illustrate the accuracy of their method to solve PDEs using the Monte Carlo method.

The aim of this paper is to extend the results in Henry-Labordère *et al.* [18, 19] to the case of elliptic PDEs with Dirichlet condition. The first result on the link between semi-linear elliptic PDE and branching diffusion processes was obtained by Watanabe [36] who derived a criterion for extinction of branching Brownian motion absorbed at the boundary of a domain. Recently, Bossy *et al.* [4] extended this result to derive a probabilistic representation for PDE of the form

$$\mathcal{L}u + \beta \left( \sum_{l \in \mathbb{N}} c_l u^l - u \right) = 0 \text{ in } \mathcal{O}, \quad u = h \text{ on } \partial \mathcal{O}, \quad (4)$$

where  $(c_l)_{l \in \mathbb{N}}$  is a sequence of real-valued functions. They used it to compute a Monte Carlo approximation of the solution to Poisson-Boltzmann equation in dimension three. One of the critical assumptions in their work is the existence of a smooth solution to PDE (4). In this paper, we do not require this assumption as we show directly that the probabilistic representation provides a continuous viscosity solution to PDE (4). The main difficulty compared to [18] is to ensure the continuity and integrability of the probabilistic representation. In particular, the arguments of small maturity used in [18] cannot be exploited here and instead, we resolve this problem by developing new tools which ensure that our probabilistic representation holds in small domain.

In the second part of our paper, we perform a rigorous analysis for the case of semi-linear elliptic PDE involving non-linearity in gradient of the unknown function. To the best of our knowledge, this is the first paper in the literature to provide a representation for this class of PDE using branching diffusion processes. In contrast with Henry-Labordère *et al.* [19], Malliavin calculus cannot be used in this setting since the exit time of a diffusion from a domain is not differentiable in the Malliavin sense. However, Delarue [11] and Gobet [16] have established a suitable automatic differentiation formula, based on the work of Thalmaier [35], which allows us to derive a probabilistic representation analogous to [19]. Under the assumption that the non-linear gradient term vanishes at the boundary, it provides a continuously differentiable solution to PDE (1) when  $f$  is a multivariate polynomial.

Importantly, our probabilistic representation provides a means to evaluate solutions of semi-linear elliptic PDEs by using the Monte Carlo method whose accuracy relies on the dimensionless central limit theorem. For this reason, it is particularly well suited for numerical applications in high dimension where deterministic methods usually fail to provide accurate estimates. Regarding the probabilistic approach, although different numerical methods for BSDEs have been introduced (see, *e.g.*, [6, 37]), to the best of our knowledge, they have never been used in the literature for elliptic PDEs.<sup>1</sup> In addition, the classical approach to solve BSDEs

<sup>1</sup>A related problem has been studied by Bouchard and Menozzi [5] for the case of parabolic PDEs in a cylindrical domain with finite time horizon.

induces a discretization bias and involves a sequence of conditional expectation estimators. The available methods to estimate conditional expectations, such as least-squares regression (see Gobet *et al.* [17]), induce an additional approximation bias and are computationally expensive in high dimension.

We consider several numerical examples, including multi-dimensional elliptic PDEs, to illustrate the accuracy of our method. One of the critical steps in the algorithm consists of simulating the first exit time and position of a diffusion from a domain. There are three classical approaches to perform this task – walk on spheres scheme, walk on parallelepipeds scheme and Euler discretization method (see Bouchard *et al.* [7] and references therein). For instance, the walk on spheres scheme was used by Bossy *et al.* [4] to estimate the solution of Poisson-Boltzmann equation in dimension three. In this paper, we consider numerical examples which involve Brownian motion over a rectangular domain and therefore, we generate unbiased samples for exit time and position using the walk on squares scheme as introduced in Faure [12] and in Milstein and Tretyakov [27], and implemented in the library developed by Lejay [24].

The rest of the paper is organized as follows. In the next section, we provide a precise formulation of the problem and introduce branching diffusion processes used to derive our probabilistic representation. Then, we consider the case of semi-linear PDE with linear gradient term in Section 3 and give explicit sufficient conditions to ensure that the probabilistic representation provides a continuous viscosity solution to the PDE. In Section 4, we perform the analysis for semi-linear PDE with non-linear gradient term. In particular, we establish appropriate automatic differentiation formulas. Finally, in Section 5, we present several numerical examples to illustrate the applicability of our results in different settings. We provide the proof of the automatic differentiation formula in Appendix A.

## 1.1 Notations

Any element  $x \in \mathbb{R}^d$ ,  $d \geq 1$ , is a column vector with  $i$ th component  $x_i$  and Euclidean norm  $|x|$ .  $x \cdot y$  denotes the usual dot product and  $d(x, y)$  denotes the Euclidean distance for any  $x, y \in \mathbb{R}^d$ . Given an open set  $\mathcal{O} \subset \mathbb{R}^d$ ,  $\bar{\mathcal{O}}$  denotes its closure,  $\partial\mathcal{O}$  its boundary and  $\text{diam}(\mathcal{O})$  its diameter. For any  $g : \mathcal{O} \rightarrow \mathbb{R}$ , the supremum norm is defined as  $\|g\|_\infty = \sup\{|g(x)| : x \in \mathcal{O}\}$ .  $\mathcal{C}(\mathcal{O})$  (resp.  $\mathcal{C}(\bar{\mathcal{O}})$ ) denotes the set of continuous functions on  $\mathcal{O}$  (resp.  $\bar{\mathcal{O}}$ ).  $\mathcal{C}^k(\mathcal{O})$  (resp.  $\mathcal{C}^k(\bar{\mathcal{O}})$ ) denotes the set of functions with continuous derivatives of all orders less than or equal to  $k$  (resp. which further have continuous extensions to  $\bar{\mathcal{O}}$ ).  $\mathcal{C}^{k,\alpha}(\mathcal{O})$  (resp.  $\mathcal{C}^{k,\alpha}(\bar{\mathcal{O}})$ ) are subspaces of  $\mathcal{C}^k(\mathcal{O})$  (resp.  $\mathcal{C}^k(\bar{\mathcal{O}})$ ) consisting of functions whose partial derivatives of order less than or equal to  $k$  are locally (resp. globally) Hölder continuous with exponent  $0 < \alpha < 1$ . Given  $T > 0$ ,  $\mathcal{C}^{1,k}((0, T] \times \mathcal{O})$  denotes the set of functions with continuous first derivative with respect to the first component and continuous derivatives of all orders less than or equal to  $k$  with respect to the second component. For a smooth function  $u(t, x)$ ,  $Du$  and  $D^2u$  stand for, respectively, its gradient (as a row vector) and Hessian matrix with respect to its second component. Furthermore, if  $g : \mathbb{R}^m \rightarrow \mathbb{R}^n$  is a differentiable function, its gradient  $Dg = (\partial_{x_1}g(x), \dots, \partial_{x_m}g(x))$  takes values in  $\mathbb{R}^n \otimes \mathbb{R}^m$ . For  $d \geq 1$ ,  $\mathbb{M}^d$  denotes the set of all  $d \times d$  matrices and  $I_d \in \mathbb{M}^d$  the identity matrix.

## 2 Problem Definition

### 2.1 A Class of Semi-Linear Elliptic PDE

Let  $(\mu, \sigma) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{M}^d$  denote the drift and diffusion coefficient. Then for a non-negative integer  $m$ , we consider a subset  $L \subset \mathbb{N}^{m+1}$  and a generator function  $f : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$  defined as

$$f(x, y, z) := \sum_{l=(l_0, l_1, \dots, l_m) \in L} c_l(x) y^{l_0} \prod_{i=1}^m (b_i(x) \cdot z)^{l_i},$$

where  $(c_l)_{l \in L}$ ,  $c_l : \mathbb{R}^d \rightarrow \mathbb{R}$  and  $(b_i)_{i=1, \dots, m}$ ,  $b_i : \mathbb{R}^d \rightarrow \mathbb{R}^d$  are sequences of functions. Furthermore, for every  $l = (l_0, l_1, \dots, l_m)$ , we denote  $|l| := \sum_{i=0}^m l_i$ . Given a bounded domain  $\mathcal{O} \subset \mathbb{R}^d$ , we consider the following semi-linear elliptic PDE:

$$\mathcal{L}u + \beta (f(u, Du) - u) = 0 \quad \text{in } \mathcal{O}, \quad u = h \quad \text{on } \partial\mathcal{O}, \quad (5)$$

where  $\beta$  is a positive constant,  $h : \partial\mathcal{O} \rightarrow \mathbb{R}$  is the Dirichlet boundary condition and  $\mathcal{L}$  is the infinitesimal generator associated to a diffusion process with parameters  $(\mu, \sigma)$ . Under a set of general assumptions, we provide a probabilistic representation of solution  $u$  using the theory of branching diffusion processes. We list the assumptions on parameters of PDE (5) which are needed for our results at the outset.

**Assumption 2.1.** (i) The functions  $b_i, c_l$  are continuous on  $\bar{\mathcal{O}}$ .

(ii) The function  $h$  is continuous on  $\partial\mathcal{O}$ .

(iii) The function  $(x, y, \zeta) \mapsto \sum_{l \in L} c_l(x) y^{l_0} \prod_{i=1}^m \zeta_i^{l_i}$  is continuous on  $\bar{\mathcal{O}} \times \mathbb{R} \times \mathbb{R}^m$ .

Part (iii) above implies that the multivariate power series  $(y, \zeta) \mapsto \sum_{l \in L} c_l(x) y^{l_0} \prod_{i=1}^m \zeta_i^{l_i}$  has an infinite radius of convergence for all  $x \in \bar{\mathcal{O}}$ . It is a simplifying assumption which avoids the use of localization arguments in the rest of the paper.

## 2.2 Marked Branching Diffusion Processes with Absorption

A (age-dependent) marked branching diffusion process is characterized by diffusion parameters  $(\mu, \sigma)$ , a probability density function  $\rho: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , and a probability mass function  $(p_l)_{l \in L}$ . In this process, we start with one particle of mark 0 at position  $x \in \mathbb{R}^d$  which undergoes a diffusion with parameters  $(\mu, \sigma)$  during its lifetime distributed according to  $\rho$ . At the end of its lifetime (arrival time), the particle dies and gives rise to  $|l| = \sum_{i=0}^m l_i$  offsprings with probability  $p_l$ , among which  $l_0$  have mark 0,  $l_1$  have mark 1, and so on. After their birth, each offspring performs the same but an independent branching diffusion process as the parent particle. Additionally, we consider that particles are absorbed at the boundary, *i.e.*, they die without giving rise to any offspring when they leave the domain  $\mathcal{O}$ . In order to construct the above process, we denote by  $\mathbb{K} := \{\emptyset\} \cup \bigcup_{n \geq 1} \mathbb{N}^n$  the set of labels and we consider the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  equipped with

- a sequence of i.i.d. positive random variables  $(\tau^k)_{k \in \mathbb{K}}$  distributed with density function  $\rho$
- a sequence of i.i.d. random elements  $(I^k)_{k \in \mathbb{K}}$  with  $\mathbb{P}(I^k = l) = p_l, l \in L$
- a sequence of independent  $d$ -dimensional Brownian motions  $(W^k)_{k \in \mathbb{K}}$

Furthermore, we consider the sequences  $(\tau^k)_{k \in \mathbb{K}}, (I^k)_{k \in \mathbb{K}}$  and  $(W^k)_{k \in \mathbb{K}}$  to be mutually independent. The age-dependent branching process is constructed as follows:

1. Start from one particle at position  $x \in \mathbb{R}^d$  and index it by label  $\emptyset$ . This is the common ancestor of all particles, the only particle of generation 0. For notational convenience, we simply write  $\tau, I$  and  $W$  instead of  $\tau^\emptyset, I^\emptyset$  and  $W^\emptyset$ . The dynamic of particle  $\emptyset$  is given as follows:

- The position  $X^\emptyset = X^x$  of the particle during its lifetime is given by

$$X_t^x = x + \int_0^t \mu(X_s^x) ds + \int_0^t \sigma(X_s^x) dW_s, \quad \mathbb{P} - \text{a.s.},$$

- The arrival time of the particle is given by

$$T^\emptyset := \tau \wedge \eta^x \quad \text{where } \eta^x := \inf \left\{ t \geq 0; X_t^x \notin \mathcal{O} \right\}.$$

- At the arrival time, if  $\eta^x \leq \tau$ , then the particle  $\emptyset$  dies without giving rise to any offspring, else the particle  $\emptyset$  dies and gives rise to  $|I|$  offsprings which belong to the first generation, and are indexed by label  $i$  for  $i = 0, \dots, |I| - 1$ .
- Given  $I = (I_0, I_1, \dots, I_m)$ , we have  $|I| = \sum_{i=0}^m I_i$  offspring particles, among which the first  $I_0$  have mark 0,  $I_1$  have mark 1, and so on, so that each particle has mark  $i$  for  $i = 0, \dots, m$ .

2. For generation  $n \geq 1$ , let the label for a particle be given as  $k = (k_1, \dots, k_{n-1}, k_n) \in \mathbb{N}^n$ . Furthermore, denote by  $k^- := (k_1, \dots, k_{n-2}, k_{n-1})$  the parent particle of  $k$ . The particle  $k$  starts from  $X_{T_{k^-}}^{k^-}$  at time  $T_{k^-}$ :

- The position  $X^k$  of the particle during its lifetime is given by

$$X_t^k = X_{T_{k^-}}^{k^-} + \int_{T_{k^-}}^t \mu(X_s^k) ds + \int_{T_{k^-}}^t \sigma(X_s^k) dW_{s-T_{k^-}}^k, \quad \mathbb{P} - \text{a.s.}$$

- The arrival time of the particle is given by

$$T_k := (T_{k^-} + \tau^k) \wedge \inf \left\{ t \geq T_{k^-}; X_t^k \notin \mathcal{O} \right\}.$$

- At the arrival time, if  $X_{T_k}^k \notin \mathcal{O}$ , then the particle  $k$  dies without giving rise to any offspring, else the particle  $k$  dies and gives rise to  $|I^k|$  offsprings which belong to the  $(n+1)$ th generation, and are indexed by label  $(k_1, \dots, k_{n-1}, k_n, i)$  for  $i = 0, \dots, |I^k| - 1$ .
- Given  $I^k = (I_0^k, I_1^k, \dots, I_m^k)$ , we have  $|I^k| = \sum_{i=0}^m I_i^k$  offspring particles, among which the first  $I_0^k$  have mark 0,  $I_1^k$  have mark 1, and so on, so that each particle has mark  $i$  for  $i = 0, \dots, m$ .

In addition, we denote by  $\mathcal{K}_n^x$  the collection of particles of the  $n$ th generation and by  $\mathcal{K}^x = \bigcup_{n \in \mathbb{N}} \mathcal{K}_n^x$  the collection of all particles. Finally, we introduce the filtration

$$\mathcal{F}_n := \sigma \left( \tau^k, I^k, W^k, k \in \{\emptyset\} \cup \bigcup_{i=1}^n \mathbb{N}^i \right), \quad n \in \mathbb{N}.$$

We make the following assumptions to ensure that the branching diffusion process above is well-defined and for further developments.

- Assumption 2.2.** (i) *The probability distribution  $(p_l)_{l \in L}$  satisfies  $p_l > 0$  and  $\sum_{l \in L} |l| p_l < \infty$ .*  
(ii) *The probability density  $\rho$  is strictly positive.*  
(iii) *The coefficients  $(\mu, \sigma)$  are Lipschitz on  $\bar{\mathcal{O}}$ .*

Part (i) of the assumption above ensures that the number of particles remains finite in finite time in the underlying branching process, see, *e.g.*, Athreya and Ney [1, Theorem 4.1.1]. In addition, under assertion (iii), there exists a unique solution (up to the boundary) to the stochastic differential equation (SDE) corresponding to  $(\mu, \sigma)$  and so the branching diffusion process is well-defined.

**Assumption 2.3.** *The branching diffusion process goes extinct almost surely.*

It is clearly sufficient to assume that  $\sum_{l \in L} |l| p_l \leq 1$  for Assumption 2.3 to hold. However, since particles are absorbed at the boundary, we can derive much weaker conditions, especially if the domain is small. For instance, in the case of branching Brownian motion with exponential lifetime of parameter  $\beta$ , Assumption 2.3 is equivalent to

$$\beta \left( \sum_{l \in L} |l| p_l - 1 \right) - \frac{\lambda_1}{2} \leq 0, \quad (6)$$

where  $\lambda_1$  is the first positive eigenvalue of the Laplacian operator in the domain  $\mathcal{O}$ , see Sevast'yanov [33] or Watanabe [36]. Also see Remark 3.3 below for further developments.

### 3 Semi-Linear PDEs with Linear Gradient Term

In this section, we are concerned with semi-linear PDEs with polynomial non-linearity involving the unknown function but not its derivatives, *i.e.*, we assume that  $m = 0$  in Section 2 so that all the particles have the same mark 0 and PDE (5) reads as

$$\mathcal{L}u + \beta(f(u) - u) = 0 \quad \text{in } \mathcal{O}, \quad u = h \quad \text{on } \partial\mathcal{O}, \quad (7)$$

where  $f(x, y) = \sum_{l \in L} c_l(x) y^l$ ,  $L \subset \mathbb{N}$ . Throughout this section, we suppose that Assumption 2.1–2.3 remain valid. Additionally, we set  $\rho$  as the density of exponential distribution with parameter  $\beta$ .

#### 3.1 Probabilistic Representation

We consider a branching diffusion process starting from  $x \in \mathcal{O}$  as in Section 2.2 and we introduce the following random variable:

$$\psi^x := \prod_{\substack{k \in \mathcal{K}^x \\ X_{T_k}^k \notin \mathcal{O}}} h(X_{T_k}^k) \prod_{\substack{k \in \mathcal{K}^x \\ X_{T_k}^k \in \mathcal{O}}} \frac{c_{I^k}(X_{T_k}^k)}{p_{I^k}}.$$

**Proposition 3.1.** *Suppose that PDE (7) has a solution  $u \in \mathcal{C}^2(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  such that the sequence  $(\psi_n^x)_{n \in \mathbb{N}}$  given by*

$$\psi_n^x := \prod_{\substack{k \in \bigcup_{i=0}^n \mathcal{K}_i^x \\ X_{T_k}^k \notin \mathcal{O}}} h(X_{T_k}^k) \prod_{\substack{k \in \bigcup_{i=0}^n \mathcal{K}_i^x \\ X_{T_k}^k \in \mathcal{O}}} \frac{c_{I^k}(X_{T_k}^k)}{p_{I^k}} \prod_{k \in \mathcal{K}_{n+1}^x} u(X_{T_k}^k),$$

*is uniformly integrable, then we have  $u(x) = \mathbb{E}[\psi^x]$ .*

*Proof.* By applying Itô's formula to  $(e^{-\beta t}u(X_t^x))_{t \geq 0}$ , we obtain the following Feynman-Kac representation:

$$u(x) = \mathbb{E} \left[ e^{-\beta \eta^x} h(X_{\eta^x}^x) + \int_0^{\eta^x} \beta e^{-\beta s} f(\cdot, u)(X_s^x) ds \right].$$

See, *e.g.*, Freidlin [13, Theorem 2.2.1] for a detailed proof. Then, we can write

$$\begin{aligned} u(x) &= \mathbb{E} \left[ h(X_{\eta^x}^x) \mathbf{1}_{\tau \geq \eta^x} + f(\cdot, u)(X_{\tau}^x) \mathbf{1}_{\tau < \eta^x} \right] \\ &= \mathbb{E} \left[ h(X_{\eta^x}^x) \mathbf{1}_{\tau \geq \eta^x} + \frac{c_I(X_{\tau}^x)}{p_I} u^I(X_{\tau}^x) \mathbf{1}_{\tau < \eta^x} \right]. \end{aligned} \quad (8)$$

Since an empty product is equal to 1 by convention, it follows that  $u(x) = \mathbb{E}[\psi_0^x]$ . Next, since each offspring has the same dynamic as the parent particle, we can repeat the above calculations to write for  $k \in \mathcal{K}_1^x$ ,

$$u(X_{T_k^-}^k) = \mathbb{E} \left[ h(X_{T_k}^k) \mathbf{1}_{X_{T_k}^k \notin \mathcal{O}} + \frac{c_{I^k}(X_{T_k}^k)}{p_{I^k}} u^{I^k}(X_{T_k}^k) \mathbf{1}_{X_{T_k}^k \in \mathcal{O}} \middle| \mathcal{F}_0 \right]. \quad (9)$$

We use the result in (9) and plug it back in (8) to obtain  $u(x) = \mathbb{E}[\psi_1^x]$  by conditional independence of particles in  $\mathcal{K}_1^x$  given  $\mathcal{F}_0$ . Similarly, we can show by iteration that for any  $n \in \mathbb{N}$ , we have  $u(x) = \mathbb{E}[\psi_n^x]$ . To conclude, it remains to observe that  $\psi_n^x$  converges to  $\psi^x$  almost surely in view of Assumption 2.3. Thus, if we suppose that  $(\psi_n^x)_{n \in \mathbb{N}}$  is uniformly integrable, as  $n \rightarrow \infty$ , we get  $u(x) = \mathbb{E}[\psi^x]$ .  $\square$

Proposition 3.1 shows that there is at most one solution to PDE (7) satisfying an appropriate integrability condition and such a solution admits a probabilistic representation using branching diffusion processes. It is not completely satisfactory since one needs to prove first the existence of a classical solution and even then the uniform integrability condition is hard to derive as illustrated by the example of Section 5.1. In order to overcome these limitations, we next establish a result which shows directly that the probabilistic representation provides a (viscosity) solution of PDE (7). This is the main result of this section. It is stated under abstract assumptions for which we will provide explicit sufficient conditions in the subsequent sections.

**Theorem 3.1.** *Suppose  $u : x \mapsto \mathbb{E}[\psi^x]$  is well-defined and continuous on  $\bar{\mathcal{O}}$ . Then  $u$  is a viscosity solution of PDE (7).*

*Proof.* We first observe that by definition of  $u$ , it holds

$$u(x) = \mathbb{E} \left[ h(X_{\eta^x}^x) \mathbf{1}_{\tau \geq \eta^x} + \frac{c_I(X_{\tau}^x)}{p_I} \prod_{i=0}^{I-1} \psi_i^{X_{\tau}^x} \mathbf{1}_{\tau < \eta^x} \right].$$

where

$$\psi_i^{X_{\tau}^x} := \prod_{\substack{k=(i,\dots) \in \mathcal{K}^x \\ X_{T_k}^k \notin \mathcal{O}}} h(X_{T_k}^k) \prod_{\substack{k=(i,\dots) \in \mathcal{K}^x \\ X_{T_k}^k \in \mathcal{O}}} \frac{c_{I^k}(X_{T_k}^k)}{p_{I^k}}.$$

Furthermore, the branching property, which says that each offspring starts the same but an independent branching diffusion as the parent particle, yields that conditioned on  $X_{\tau}^x$  and  $I$ ,  $(\psi_0^{X_{\tau}^x}, \dots, \psi_{I-1}^{X_{\tau}^x})$  are independent random variables identical in law to  $\psi^{X_0}$  where  $X_0$  is distributed as  $X_{\tau}^x$  and independent of  $\mathcal{F}_n$  for all  $n \in \mathbb{N}$ . Through this argument, we deduce that

$$\mathbb{E} \left[ \prod_{i=0}^{I-1} \psi_i^{X_{\tau}^x} \middle| X_{\tau}^x, I \right] \mathbf{1}_{\tau < \eta^x} = u^I(X_{\tau}^x) \mathbf{1}_{\tau < \eta^x}.$$

Working backward along the lines of the proof of Proposition 3.1, we deduce that

$$u(x) = \mathbb{E} \left[ e^{-\beta \eta^x} h(X_{\eta^x}^x) + \int_0^{\eta^x} \beta e^{-\beta s} f(\cdot, u)(X_s^x) ds \right]. \quad (10)$$

Next, for any  $\delta > 0$ , it follows from Markov property that

$$u(x) = \mathbb{E} \left[ e^{-\beta(\eta^x \wedge \delta)} u(X_{\eta^x \wedge \delta}^x) + \int_0^{\eta^x \wedge \delta} \beta e^{-\beta s} f(\cdot, u)(X_s^x) ds \right]. \quad (11)$$

The fact that  $u$  is a viscosity solution of PDE (7) now follows from classical arguments. For the sake of completeness, let us show that  $u$  is a viscosity subsolution. Consider  $\varphi \in \mathcal{C}_b^2(\mathcal{O})$  such that  $x \in \mathcal{O}$  is a maximum point of  $u - \varphi$  and  $u(x) = \varphi(x)$ . First, we observe by applying Itô's formula that

$$\mathbb{E} \left[ e^{-\beta(\eta^x \wedge \delta)} \varphi(X_{\eta^x \wedge \delta}^x) \right] = \varphi(x) + \mathbb{E} \left[ \int_0^{\eta^x \wedge \delta} e^{-\beta s} (\mathcal{L}\varphi - \beta\varphi)(X_s^x) ds \right].$$

Since  $u(x) = \varphi(x)$  and  $u \leq \varphi$  otherwise, we deduce by using (11) that

$$\mathbb{E} \left[ \frac{1}{\delta} \int_0^{\eta^x \wedge \delta} e^{-\beta s} (\mathcal{L}\varphi - \beta\varphi + \beta f(\cdot, u))(X_s^x) ds \right] \geq 0.$$

Since  $y \mapsto f(y, u(y))$  is continuous and bounded, it follows from the mean value theorem and the dominated convergence theorem that

$$\mathcal{L}\varphi(x) + \beta (f(x, \varphi(x)) - \varphi(x)) \geq 0.$$

Thus  $u$  is a viscosity subsolution of PDE (7). The fact that  $u$  is a viscosity supersolution results from similar arguments. □

**Remark 3.1.** *It will turn out from the study of the next sections that the explicit conditions we provide to ensure that  $x \mapsto \mathbb{E}[\psi^x]$  is a solution to PDE (7) also entail that  $x \mapsto \mathbb{E}[|\psi^x|]$  is a solution to*

$$\mathcal{L}v + \beta \left( \sum_{l \in L} |c_l| v^l - v \right) = 0 \text{ in } \mathcal{O}, \quad v = |h| \text{ on } \partial\mathcal{O}.$$

*In particular, we observe that the monotonicity assumption is not relevant in our setting. The limiting assumption in our approach is the integrability condition on  $(\psi^x)_{x \in \mathcal{O}}$  which ensures that  $u : x \mapsto \mathbb{E}[\psi^x]$  is well-defined. See Section 3.3 for more details.*

**Remark 3.2.** *We can also work with a general lifetime distribution  $\rho$  and extend the arguments above, as done in Section 4, to derive another probabilistic representation which is given as*

$$u(x) = \mathbb{E} \left[ \prod_{\substack{k \in \mathcal{K}^x \\ X_{T_k}^k \notin \mathcal{O}}} \frac{e^{-\beta \Delta T_k} h(X_{T_k}^k)}{\bar{F}(\Delta T_k)} \prod_{\substack{k \in \mathcal{K}^x \\ X_{T_k}^k \in \mathcal{O}}} \frac{\beta e^{-\beta \Delta T_k} c_{I^k}(X_{T_k}^k)}{p_{I^k} \rho(\Delta T_k)} \right],$$

where  $\Delta T_k := T_k - T_{k-}$  is the lifetime of particle  $k$  and  $\bar{F}(t) := \int_t^\infty \rho(s) ds$ ,  $t \geq 0$ . However, this leads to more stringent assumptions when studying the integrability condition as in Section 3.3.

## 3.2 Continuity Assumption

Let us give explicit sufficient conditions for the continuity assumption in Theorem 3.1 to hold. We use two different approaches leading to slightly different conditions. Both approaches rely on appropriate integrability conditions on  $(\psi^x)_{x \in \mathcal{O}}$  which we discuss in the next section.

### 3.2.1 PDE Approach

**Assumption 3.1.** (i) *The diffusion coefficient  $\sigma$  is uniformly elliptic.*<sup>2</sup>

(ii) *The boundary  $\partial\mathcal{O}$  is of class  $\mathcal{C}^{1,\alpha}$ .*

**Proposition 3.2.** *Suppose Assumption 3.1 holds and  $(\psi^x)_{x \in \mathcal{O}}$  is uniformly bounded in  $L^1$ , then the map  $u : x \mapsto \mathbb{E}[\psi^x]$  is continuous.*

The proof follows immediately from (10) and the following lemma.

**Lemma 3.1.** *Suppose Assumption 3.1 holds, then the following statements are satisfied:*

(i) *The map  $\bar{\mathcal{O}} \ni x \mapsto \mathbb{E}[e^{-\beta \eta^x} h(X_{\eta^x}^x)]$  is continuous.*

(ii) *For any  $g : \mathcal{O} \rightarrow \mathbb{R}$  bounded measurable, the map  $\bar{\mathcal{O}} \ni x \mapsto \mathbb{E}[\int_0^{\eta^x} e^{-\beta s} g(X_s^x) ds]$  is continuous.*

<sup>2</sup>There exists  $\lambda > 0$  such that  $\sigma\sigma^*(x) \geq \lambda I_d$  for all  $x \in \bar{\mathcal{O}}$ .



*Proof.* (i) Let us first study the continuity of  $x \mapsto \mathbb{E}[e^{-\beta\eta^x} h(X_{\eta^x}^x)]$ . Under the assumptions above, it is well-known that there exists a smooth solution  $\varphi \in \mathcal{C}^2(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  to the following PDE:

$$\mathcal{L}\varphi - \beta\varphi = 0 \quad \text{in } \mathcal{O}, \quad \varphi = h \quad \text{on } \partial\mathcal{O}.$$

See, *e.g.*, Gilbarg and Trudinger [15, Theorem 6.13]. By Itô's formula, we deduce that  $\varphi(x) = \mathbb{E}[e^{-\beta\eta^x} h(X_{\eta^x}^x)]$  and the conclusion follows.

(ii) Let us now turn to the continuity of  $x \mapsto \mathbb{E}[\int_0^{\eta^x} e^{-\beta s} g(X_s^x) ds]$ . We start by showing that  $x \mapsto \mathbb{E}[g(X_s^x) \mathbf{1}_{s < \eta^x}]$  is continuous for  $s > 0$ . If  $g$  is continuous and  $g = 0$  on  $\partial\mathcal{O}$ , it is known that the unique smooth solution  $\chi \in \mathcal{C}^{1,2}((0, T] \times \mathcal{O}) \cap \mathcal{C}([0, T] \times \bar{\mathcal{O}})$  of

$$\begin{aligned} \partial_t \chi - \mathcal{L}\chi &= 0, & \text{in } (0, T] \times \mathcal{O}, \\ \chi(0, \cdot) &= g, & \text{on } \mathcal{O}, \\ \chi &= 0, & \text{on } (0, T] \times \partial\mathcal{O}, \end{aligned} \tag{12}$$

is of the form

$$\chi(s, x) = \int_{\mathcal{O}} G(s, x; 0, y) g(y) dy, \quad 0 < s \leq T,$$

where  $G$  is the so-called Green function of PDE (12) (see, *e.g.*, Ladyženskaja *et al.* [23, Theorem 4.16.2]). Then, it follows from Itô's formula that

$$\mathbb{E}[g(X_s^x) \mathbf{1}_{s < \eta^x}] = \int_{\mathcal{O}} G(s, x; 0, y) g(y) dy.$$

In particular,  $y \mapsto G(s, x; 0, y)$  appears as the density of  $X_s^x$  on the event  $\{s < \eta^x\}$ , and so the identity above remains valid for any  $g$  bounded measurable. Furthermore,  $x \mapsto G(s, x; 0, y)$  is continuous and satisfies

$$|G(s, x; 0, y)| \leq C s^{-\frac{d}{2}} e^{-C \frac{|x-y|^2}{s}},$$

for some constant  $C > 0$  depending on  $T$ , see [23, Equation 4.16.16]. Hence the desired result follows from the dominated convergence theorem. To conclude, it remains to observe that

$$\mathbb{E} \left[ \int_0^{\eta^x} e^{-\beta s} g(X_s^x) ds \right] = \int_0^{+\infty} e^{-\beta s} \mathbb{E}[g(X_s^x) \mathbf{1}_{s < \eta^x}] ds$$

and to apply once again the dominated convergence theorem. □

### 3.2.2 Probabilistic Approach

- Assumption 3.2.** (i) *The diffusion coefficient  $\sigma$  is uniformly elliptic on  $\partial\mathcal{O}$ .*<sup>3</sup>  
(ii) *The boundary  $\partial\mathcal{O}$  satisfies an exterior cone condition.*<sup>4</sup>  
(iii) *The stopping time  $\eta^x$  is finite almost surely.*

Sufficient conditions for Part (iii) to hold are provided in Freidlin [13, Lemma 3.3.1]. For instance, it suffices to assume that there exists  $1 \leq i \leq d$  such that  $\sum_{j=1}^d \sigma_{ij}^2(x) > 0$  or  $|\mu_i(x)| > 0$  for all  $x \in \mathcal{O}$ .

**Proposition 3.3.** *Suppose Assumption 3.2 holds and  $(\psi^x)_{x \in \mathcal{O}}$  is uniformly integrable, then  $u : x \mapsto \mathbb{E}[\psi^x]$  is continuous.*

*Proof.* Clearly it suffices to prove that  $\bar{\mathcal{O}} \ni x \mapsto \psi^x$  is almost surely continuous, in the sense that for all sequence  $(x_n)_{n \in \mathbb{N}}$  converging to  $x$ , it holds

$$\mathbb{P} \left( \lim_{n \rightarrow \infty} \psi^{x_n} = \psi^x \right) = 1.$$

This essentially follows from the almost sure continuity of  $x \mapsto \eta^x$  stated in Lemma 3.2 below.

<sup>3</sup>There exists  $\lambda > 0$  such that  $\sigma\sigma^*(x) \geq \lambda I_d$  for all  $x \in \partial\mathcal{O}$ .

<sup>4</sup>See, *e.g.*, Gilbarg and Trudinger [15, Problem 2.12] for a definition.

(i) Let us first show that the contribution of the first particle to  $\psi^x$  is almost surely continuous, *i.e.*,

$$\lim_{n \rightarrow \infty} h(X_{\eta^{x_n}}^x) \mathbf{1}_{\tau \geq \eta^{x_n}} = h(X_{\eta^x}^x) \mathbf{1}_{\tau \geq \eta^x}, \quad \mathbb{P} - \text{a.s.}, \quad (13)$$

$$\lim_{n \rightarrow \infty} \frac{c_I(X_{\tau}^{x_n})}{p_I} \mathbf{1}_{\tau < \eta^{x_n}} = \frac{c_I(X_{\tau}^x)}{p_I} \mathbf{1}_{\tau < \eta^x}, \quad \mathbb{P} - \text{a.s.} \quad (14)$$

To achieve this, let us consider the set

$$\Omega^\theta := \{\eta^x \neq \tau\} \cap \left\{ \lim_{n \rightarrow \infty} \eta^{x_n} = \eta^x \right\} \cap \left\{ \lim_{n \rightarrow \infty} X^{x_n} = X^x \right\}.$$

In view of Lemma 3.2 below, it is clear that  $\mathbb{P}(\Omega^\theta) = 1$ . Furthermore, one easily checks that, for every  $\omega \in \Omega^\theta$ , both (13)–(14) hold.

(ii) Let us show next that the contribution of the particles of the first generation to  $\psi^x$  is almost surely continuous. For every  $i \in \mathbb{N}$ , we denote by  $(X_s^{i,x})_{s \geq 0}$  the unique solution of

$$X_t^{i,x} = x + \int_0^t \mu(X_s^{i,x}) ds + \int_0^t \sigma(X_s^{i,x}) dB_s^i, \quad \mathbb{P} - \text{a.s.},$$

where the Brownian motion  $B^i$  is defined by

$$B_t^i := W_{t \wedge \tau} + W_{t - \tau}^i \mathbf{1}_{t \geq \tau}.$$

Clearly,  $(X_s^{i,x})_{s \geq 0}$  coincides with the trajectory  $X^i$  of particle  $i$  during its lifetime. Let  $\eta^{i,x}$  be the first exit time of  $X^{i,x}$  from  $\mathcal{O}$  and consider

$$\Omega^i := \{\eta^{i,x} \neq \bar{\tau}^i\} \cap \left\{ \lim_{n \rightarrow \infty} \eta^{i,x_n} = \eta^{i,x} \right\} \cap \left\{ \lim_{n \rightarrow \infty} X^{i,x_n} = X^{i,x} \right\}.$$

where  $\bar{\tau}^i := \tau + \tau^i$ . Once again,  $\mathbb{P}(\Omega^i) = 1$  and

$$\begin{aligned} \lim_{n \rightarrow \infty} h\left(X_{\eta^{i,x_n}}^{i,x_n}\right) \mathbf{1}_{\bar{\tau}^i \geq \eta^{i,x_n}} &= h\left(X_{\eta^{i,x}}^{i,x}\right) \mathbf{1}_{\bar{\tau}^i \geq \eta^{i,x}}, \quad \text{on } \Omega^i, \\ \lim_{n \rightarrow \infty} \frac{c_{I^i}\left(X_{\bar{\tau}^i}^{i,x_n}\right)}{p_{I^i}} \mathbf{1}_{\bar{\tau}^i < \eta^{i,x_n}} &= \frac{c_{I^i}\left(X_{\bar{\tau}^i}^{i,x}\right)}{p_{I^i}} \mathbf{1}_{\bar{\tau}^i < \eta^{i,x}}, \quad \text{on } \Omega^i. \end{aligned}$$

Thus, we conclude that on the set  $\Omega^\theta \cap \bigcap_{i \in \mathbb{N}} \Omega^i$  of probability 1, the contributions of the particles of generations 0 and 1 to  $\psi^x$  is almost surely continuous.

(iii) The desired result follows by iteration. We denote for all  $n \geq 2$ ,  $k \in \mathbb{N}^n$ ,

$$\Omega^k := \{\eta^{k,x} \neq \bar{\tau}^k\} \cap \left\{ \lim_{n \rightarrow \infty} \eta^{k,x_n} = \eta^{k,x} \right\} \cap \left\{ \lim_{n \rightarrow \infty} X^{k,x_n} = X^{k,x} \right\},$$

where  $\bar{\tau}^k := \bar{\tau}^{k^-} + \tau$  and  $(X_s^{k,x})_{s \geq 0}$  is the unique solution of

$$X_t^{k,x} = x + \int_0^t \mu(X_s^{k,x}) ds + \int_0^t \sigma(X_s^{k,x}) dB_s^k, \quad \mathbb{P} - \text{a.s.},$$

with

$$B_t^k := B_{t \wedge \bar{\tau}^{k^-}}^{k^-} + W_{t - \bar{\tau}^{k^-}}^k \mathbf{1}_{t \geq \bar{\tau}^{k^-}}.$$

Then on the set  $\bigcap_{k \in \mathbb{K}} \Omega^k$  of probability 1, it holds that  $\lim_{n \rightarrow \infty} \psi^{x_n} = \psi^x$ .  $\square$

**Lemma 3.2.** *Suppose Assumption 3.2 holds, then the map  $\bar{\mathcal{O}} \ni x \mapsto \eta^x$  is almost surely continuous, in the sense that for all sequence  $(x_n)_{n \in \mathbb{N}}$  converging to  $x$ , it holds*

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \eta^{x_n} = \eta^x\right) = 1.$$

*Proof.* The proof is due to Darling and Pardoux [10], see also [28, Proposition 4.4]. More precisely, the authors show that if the stopping time

$$\tau^x := \{s \geq 0, X_s^x \notin \bar{\mathcal{O}}\},$$

is finite almost surely and  $\mathbb{P}(\tau^x > 0) = 0$  for all  $x \in \partial\mathcal{O}$ , then  $x \mapsto \tau^x$  is almost surely continuous. It remains to observe that, under Assumption 3.2 (i)–(ii),  $\mathbb{P}(\tau^x = 0) = 1$  for all  $x \in \partial\mathcal{O}$  (see Bass [3, Corollary 3.3.2] or Pinsky [31, Theorem 2.3.3]) and thus  $\tau^x = \eta^x$  for all  $x \in \bar{\mathcal{O}}$ .  $\square$

### 3.3 Integrability Condition

In this section, we provide explicit sufficient conditions to verify the integrability conditions on  $(\psi_x)_{x \in \mathcal{O}}$  required for Propositions 3.2 and 3.3. Actually we study boundedness of  $(\psi_x)_{x \in \mathcal{O}}$  in  $L^q$  for  $q \geq 1$ . In particular, the case  $q = 2$  ensures that the corresponding Monte Carlo estimator has finite variance. Let us introduce the constant

$$C_0 := \max \left( \|h\|_\infty, \sup_{l \in L} \left\{ \frac{\|c_l\|_\infty}{p_l} \right\} \right).$$

Clearly, it holds  $|\psi^x| \leq C_0^{|\mathcal{K}^x|}$  where  $|\mathcal{K}^x|$  denotes the cardinality of the set  $\mathcal{K}^x$ , *i.e.*, the total number of particles. In particular, if  $C_0 \leq 1$ , then  $|\psi^x| \leq 1$ . To the best of our knowledge, this is the only condition that has been used so far in the literature to ensure the integrability of  $\psi^x$ .

In the rest of this section, we provide more technical but weaker conditions. First, we establish the desired result under minimal assumptions.

**Proposition 3.4.** *Suppose that there exists a non-negative function  $v \in \mathcal{C}^2(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  satisfying*

$$\mathcal{L}v + \beta \left( \sum_{l \in L} \frac{|c_l|^q}{p_l^{q-1}} v^l - v \right) \leq 0 \quad \text{in } \mathcal{O}, \quad v \geq |h|^q \quad \text{on } \partial\mathcal{O},$$

*then we have  $\mathbb{E}[|\psi^x|^q] \leq v(x)$ . In particular,  $(\psi^x)_{x \in \mathcal{O}}$  is uniformly bounded in  $L^q$ .*

*Proof.* First, we observe that Itô's formula yields

$$v(x) \geq \mathbb{E} \left[ e^{-\beta\eta^x} |h(X_{\eta^x}^x)|^q + \int_0^{\eta^x} \beta e^{-\beta s} \sum_{l \in L} \frac{|c_l(X_s^x)|^q}{p_l^{q-1}} v^l(X_s^x) ds \right].$$

Next, by repeating the arguments of Proposition 3.1, we get

$$v(x) \geq \mathbb{E} \left[ \prod_{\substack{k \in \cup_{i=0}^n \mathcal{K}_i^x \\ X_{T_k}^k \notin \mathcal{O}}} |h(X_{T_k}^k)|^q \prod_{\substack{k \in \cup_{i=0}^n \mathcal{K}_i^x \\ X_{T_k}^k \in \mathcal{O}}} \frac{|c_{I^k}(X_{T_k}^k)|^q}{p_{I^k}^q} \prod_{k \in \mathcal{K}_{n+1}^x} v(X_{T_k}^k) \right].$$

Then the conclusion follows easily from Fatou's lemma.  $\square$

In practice, one can look for a constant supersolution in order to apply Proposition 3.4. For instance, for the case  $q = 1$ , such a solution exists if and only if  $\sum_{l \in L} |c_l| \|h\|_\infty^l \leq \|h\|_\infty$ . Otherwise, finding a supersolution has to be done on a case by case basis and might turn out to be a difficult task. For this reason, we provide an alternative result for which conditions are easier to verify than Proposition 3.4 and less stringent than assuming  $C_0 \leq 1$ . Essentially, it signifies that the integrability condition is satisfied if the domain is sufficiently small.

The idea is to dispose of the spatial parameter  $x$  by introducing a branching process which stochastically dominates the branching diffusion process. Denote

$$\delta := 1 - \inf_{x \in \mathcal{O}} \left\{ \mathbb{E} \left[ e^{-\beta\eta^x} \right] \right\}.$$

Let us introduce a new probability mass function  $(\tilde{p}_l)_{l \in \tilde{L}}$  with  $\tilde{L} := L \cup \{0\}$  as follows:

$$\tilde{p}_0 := 1 - \delta + \delta p_0 \quad \text{and} \quad \tilde{p}_l := \delta p_l \quad \text{for all } l \geq 1,$$

and the corresponding transition matrix  $\tilde{P} = (\tilde{P}_{i,j})_{i,j \geq 0}$  given by

$$\tilde{P}_{0,0} = 1 \quad \text{and} \quad \tilde{P}_{i,i+l-1} = \tilde{p}_l \quad \text{for all } i \geq 1, l \in \tilde{L}.$$

**Proposition 3.5.** *Denote by  $R$  the common radius of convergence of  $f(s) := \sum_{l \in L} p_l s^l$  and  $\tilde{f}(s) := \sum_{l \in \tilde{L}} \tilde{p}_l s^l$ . If  $R > 1$  and  $\sum_{l \in \tilde{L}} l \tilde{p}_l < 1$ , then  $(\psi^x)_{x \in \mathcal{O}}$  is uniformly bounded in  $L^q$  provided that  $C_0^q \leq \gamma$  where*

$$\gamma := \frac{s^*}{\tilde{f}(s^*)} = \frac{s^*}{1 - \delta + \delta f(s^*)}, \quad (15)$$

*where  $s^*$  is the solution (if any) to  $s\tilde{f}'(s) = \tilde{f}(s)$  and  $s^* = R$  otherwise. In addition, if  $R = \infty$ , then  $\gamma$  goes to infinity as  $\text{diam}(\mathcal{O})$  goes to zero.*

*Proof.* (i) Denote by  $(N_n^x)_{n \in \mathbb{N}}$  and  $(\tilde{N}_n)_{n \in \mathbb{N}}$  the number of particles at the  $n$ th jump time (arrival time) in the branching diffusion process and in a branching process with offspring distribution  $(\tilde{p}_l)_{l \in \tilde{L}}$  respectively. In particular,  $(\tilde{N}_n)_{n \in \mathbb{N}}$  is a Markov chain with transition matrix  $\tilde{P}$ . Let us first show that  $(N_n^x)_{n \in \mathbb{N}}$  is stochastically dominated by  $(\tilde{N}_n)$ , *i.e.*,  $\mathbb{P}(N_n^x \geq l) \leq \mathbb{P}(\tilde{N}_n \geq l)$  for all  $n, l \geq 1$ . We observe that

$$\tilde{p}_0 = \inf_{x \in \mathcal{O}} \{\mathbb{P}(N_1^x = 0)\} \quad \text{and} \quad \tilde{p}_l = \sup_{x \in \mathcal{O}} \{\mathbb{P}(N_1^x = l)\} \quad \text{for } l \in \tilde{L} \setminus \{0\}.$$

It follows that  $\mathbb{P}(N_1^x \geq l) \leq \mathbb{P}(\tilde{N}_1 \geq l)$  for all  $l \geq 1$ . For the incremental step, we first observe that for all  $l \geq 1$ ,

$$\begin{aligned} \mathbb{P}(N_{n+1}^x \geq l, N_n^x \geq l) &\leq \mathbb{P}(N_n^x \geq l) - \mathbb{P}(N_n^x = l, N_{n+1}^x = l-1) \\ &\leq \mathbb{P}(N_n^x \geq l) - \tilde{p}_0 \mathbb{P}(N_n^x = l) \\ &\leq (1 - \tilde{p}_0) \mathbb{P}(N_n^x \geq l) + \tilde{p}_0 \mathbb{P}(N_n^x \geq l+1), \end{aligned}$$

where the second inequality follows from Markov property. In addition, it holds

$$\begin{aligned} \mathbb{P}(N_{n+1}^x \geq l, N_n^x \leq l-1) &\leq \sum_{j=1}^{l-1} \mathbb{P}(N_n^x = j, N_{n+1}^x - N_n^x \geq l-j) \\ &\leq \sum_{j=1}^{l-1} \sum_{i=l-j+1}^{\infty} \tilde{p}_i \mathbb{P}(N_n^x = j) \\ &\leq \sum_{i=l}^{\infty} \tilde{p}_i \mathbb{P}(N_n^x \geq 1) - \sum_{i=2}^{\infty} \tilde{p}_i \mathbb{P}(N_n^x \geq l) + \sum_{j=2}^{l-1} \tilde{p}_{l-j+1} \mathbb{P}(N_n^x \geq j), \end{aligned}$$

where the second inequality follows once again from Markov property. Combining both computations above, we deduce that

$$\mathbb{P}(N_{n+1}^x \geq l) \leq \sum_{i=l}^{\infty} \tilde{p}_i \mathbb{P}(N_n^x \geq 1) + \sum_{j=2}^{l-1} \tilde{p}_{l-j+1} \mathbb{P}(N_n^x \geq j) + \tilde{p}_1 \mathbb{P}(N_n^x \geq l) + \tilde{p}_0 \mathbb{P}(N_n^x \geq l+1).$$

Hence the desired result follows easily by induction and Markov property.

(ii) Let us show next that  $\mathbb{E}[|\psi^x|] \leq \mathbb{E}[C_0^\zeta]$  where  $\zeta$  the extinction time of  $(\tilde{N}_n)_{n \in \mathbb{N}}$ . First we recall that  $|\psi^x| \leq C_0^{|\mathcal{K}^x|}$  where  $|\mathcal{K}^x|$  denotes the cardinality of the set  $\mathcal{K}^x$ . We observe further that  $|\mathcal{K}^x|$  coincides with the extinction time of  $(N_n^x)_{n \in \mathbb{N}}$ . Clearly, Step (i) above yields that  $\mathbb{P}(|\mathcal{K}^x| \geq n) \leq \mathbb{P}(\zeta \geq n)$  for all  $n \geq 1$ , and thus

$$\mathbb{E}[|\psi^x|] \leq \mathbb{E}[C_0^{|\mathcal{K}^x|}] \leq \mathbb{E}[C_0^\zeta].$$

(iii) We are now in a position to conclude the proof. It follows from Daley [9, Theorem 2] that the power series  $\mathbb{E}[s^\zeta]$  converges on its radius of convergence  $\gamma = \frac{s^*}{\tilde{f}(s^*)}$ . In addition, if  $R = \infty$ , then there exists a solution  $s^*$  to  $s\tilde{f}'(s) = \tilde{f}(s)$ . Using further  $\tilde{f}(s) = 1 - \delta + \delta f(s)$ , we deduce that

$$s^* f'(s^*) - f(s^*) = \frac{1 - \delta}{\delta}.$$

It follows that  $s^*$  goes to infinity as  $\delta$  goes to zero, or equivalently,  $\text{diam}(\mathcal{O})$  goes to zero. Then we have

$$\gamma = \frac{1}{\tilde{f}'(s^*)} = \frac{1}{\delta f'(s^*)} = \frac{s^*}{1 - \delta} \left( 1 - \frac{f(s^*)}{s^* f'(s^*)} \right).$$

To conclude, it remains to observe that  $\frac{f(s^*)}{s^* f'(s^*)}$  is bounded away from 1 since  $s \mapsto \frac{f(s)}{s f'(s)}$  is decreasing.  $\square$

**Remark 3.3.** *It follows immediately from Step (i) of the proof of Proposition 3.5 that if  $\sum_{l \in \tilde{L}} l \tilde{p}_l \leq 1$ , then the branching diffusion process goes extinct almost surely, *i.e.*, Assumption 2.3 holds.*

## 4 Semi-Linear PDEs with Non-Linear Gradient Term

In this section we study the case of semi-linear PDE with non-linearity in gradient of the solution, *i.e.*, we assume that  $m \geq 1$  in Section 2 so that the particles in the branching diffusion process carry different marks to account for it. Throughout this section, we suppose that Assumption 2.1–2.3 remain valid.

## 4.1 Probabilistic Representation

Our next assumption is the key automatic differentiation condition on the underlying diffusion process  $X^x$ . We will provide explicit conditions and formulas for it in the next sections.

**Assumption 4.1.** (i) *The map  $x \mapsto \mathbb{E}[e^{-\beta\eta^x} h(X_{\eta^x}^x)]$  belongs to  $\mathcal{C}^1(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  and there exists a measurable function  $\mathcal{W}_{\partial\mathcal{O}}(x, W) = \mathcal{W}_{\partial\mathcal{O}}(x, (W_r)_{r \in [0, \eta^x]})$  such that*

$$D\mathbb{E} \left[ e^{-\beta\eta^x} h(X_{\eta^x}^x) \right] = \mathbb{E} \left[ e^{-\beta\eta^x} h(X_{\eta^x}^x) \mathcal{W}_{\partial\mathcal{O}}(x, W) \right].$$

(ii) *For any  $g : \mathcal{O} \rightarrow \mathbb{R}$  bounded measurable, the map  $x \mapsto \mathbb{E}[\int_0^{\eta^x} e^{-\beta s} g(X_s^x) ds]$  belongs to  $\mathcal{C}^1(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  and there exists a measurable function  $\mathcal{W}_{\mathcal{O}}(s, x, W) = \mathcal{W}_{\mathcal{O}}(s, x, (W_r)_{r \in [0, s]})$  such that*

$$D\mathbb{E} \left[ \int_0^{\eta^x} e^{-\beta s} g(X_s^x) ds \right] = \mathbb{E} \left[ \int_0^{\eta^x} e^{-\beta s} g(X_s^x) \mathcal{W}_{\mathcal{O}}(s, x, W) ds \right].$$

Let us define  $\mathcal{W}(s, x, W) = \mathcal{W}(s, x, (W_r)_{r \in [0, s]})$  as follows:

$$\mathcal{W}(s, x, W) := \mathcal{W}_{\partial\mathcal{O}}(x, W) \mathbf{1}_{X_s^x \notin \mathcal{O}} + \mathcal{W}_{\mathcal{O}}(s, x, W) \mathbf{1}_{X_s^x \in \mathcal{O}}.$$

We consider a marked branching diffusion process starting from  $x \in \mathcal{O}$  as in Section 2.2 and denote

$$\mathcal{W}_k := \mathbf{1}_{m_k=0} + \mathbf{1}_{m_k \neq 0} b_{m_k}(X_{T_k^-}^k) \cdot \mathcal{W}(\Delta T_k, X_{T_k^-}^k, W^k),$$

where  $m_k$  and  $\Delta T_k := T_k - T_k^-$  stand for the mark and the lifetime of particle  $k$  respectively. We next introduce the following random variable:

$$\psi^x := \prod_{\substack{k \in \mathcal{K}^x \\ X_{T_k}^k \notin \mathcal{O}}} \frac{e^{-\beta \Delta T_k} h(X_{T_k}^k)}{\bar{F}(\Delta T_k)} \mathcal{W}_k \prod_{\substack{k \in \mathcal{K}^x \\ X_{T_k}^k \in \mathcal{O}}} \frac{\beta e^{-\beta \Delta T_k} c_{I^k}(X_{T_k}^k)}{p_{I^k} \rho(\Delta T_k)} \mathcal{W}_k.$$

where  $\bar{F}(t) := \int_t^\infty \rho(s) ds$ ,  $t \geq 0$ .

**Proposition 4.1.** *Suppose Assumption 4.1 holds. Assume further that PDE (5) has a solution  $u \in \mathcal{C}^2(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  such that the functions  $(b_i \cdot Du)_{i=1, \dots, m}$  are bounded and the sequence  $(\psi_n^x)_{n \in \mathbb{N}}$  defined by*

$$\psi_n^x := \prod_{\substack{k \in \cup_{i=1}^n \mathcal{K}_i^x \\ X_{T_k}^k \notin \mathcal{O}}} \frac{e^{-\beta \Delta T_k} h(X_{T_k}^k)}{\bar{F}(\Delta T_k)} \mathcal{W}_k \prod_{\substack{k \in \cup_{i=1}^n \mathcal{K}_i^x \\ X_{T_k}^k \in \mathcal{O}}} \frac{\beta e^{-\beta \Delta T_k} c_{I^k}(X_{T_k}^k)}{p_{I^k} \rho(\Delta T_k)} \mathcal{W}_k \prod_{\substack{k \in \mathcal{K}_{n+1}^x \\ m_k=0}} u(X_{T_k}^k) \prod_{\substack{k \in \mathcal{K}_{n+1}^x \\ m_k \neq 0}} (b_{m_k} \cdot Du)(X_{T_k}^k),$$

*is uniformly integrable. Then, we have  $u(x) = \mathbb{E}[\psi^x]$ .*

*Proof.* The proof follows similar arguments as the proof of Proposition 3.1. Using Itô's formula, we have the following Feynman-Kac representation:

$$u(x) = \mathbb{E} \left[ e^{-\beta\eta^x} h(X_{\eta^x}^x) + \int_0^{\eta^x} \beta e^{-\beta s} f(\cdot, u, Du)(X_s^x) ds \right].$$

Then, we can write

$$\begin{aligned} u(x) &= \mathbb{E} \left[ \frac{e^{-\beta\eta^x} h(X_{\eta^x}^x)}{\bar{F}(\eta^x)} \mathbf{1}_{\tau \geq \eta^x} + \frac{\beta e^{-\beta\tau} f(\cdot, u, Du)(X_\tau^x)}{\rho(\tau)} \mathbf{1}_{\tau < \eta^x} \right] \\ &= \mathbb{E} \left[ \frac{e^{-\beta\eta^x} h(X_{\eta^x}^x)}{\bar{F}(\eta^x)} \mathbf{1}_{\tau \geq \eta^x} + \frac{\beta e^{-\beta\tau} c_I(X_\tau^x)}{p_I \rho(\tau)} u^{I_0}(X_\tau^x) \prod_{i=1}^m (b_i \cdot Du)^{I_i}(X_\tau^x) \mathbf{1}_{\tau < \eta^x} \right]. \end{aligned} \quad (16)$$

In other words, we have  $u(x) = \mathbb{E}[\psi_0^x]$ . Furthermore, in the original Feynman-Kac formula, we obtain by differentiating and using Assumption 4.1,

$$\begin{aligned} Du(x) &= \mathbb{E} \left[ e^{-\beta\eta^x} h(X_{\eta^x}^x) \mathcal{W}_{\partial\mathcal{O}}(x, W) + \int_0^{\eta^x} \beta e^{-\beta s} f(\cdot, u, Du)(X_s^x) \mathcal{W}_{\mathcal{O}}(s, x, W) ds \right] \\ &= \mathbb{E} \left[ \psi_0^x \mathcal{W}(T_\emptyset, x, W) \right]. \end{aligned}$$

Since each offspring has the same dynamic as the parent particle, we can repeat the above calculations for  $k \in \mathcal{K}_1^x$  and plug the results back in (16) to obtain  $u(x) = \mathbb{E}[\psi_1^x]$  by conditional independence of particles in  $\mathcal{K}_1^x$  given  $\mathcal{F}_0$ . We conclude by iteration that, for any  $n \in \mathbb{N}$ ,  $u(x) = \mathbb{E}[\psi_n^x]$  and, as  $n \rightarrow \infty$ ,  $u(x) = \mathbb{E}[\psi^x]$ .  $\square$

Similar to Section 3.1, Proposition 4.1 provides a result of uniqueness for a class of semi-linear PDEs given an appropriate uniform integrability condition is satisfied. Next we establish a result of existence by showing that the probabilistic representation is a viscosity solution of PDE (5). This is the main result of this section.

**Theorem 4.1.** *Suppose Assumption 4.1 holds. If we further assume that  $(\psi^x)_{x \in \mathcal{O}}$  and, for  $i = 1, \dots, m$ ,  $(\psi^x b_i(x) \cdot \mathcal{W}(T_\emptyset, x, W))_{x \in \mathcal{O}}$  are uniformly bounded in  $L^1$ , then  $u : x \mapsto \mathbb{E}[\psi^x]$  belongs to  $\mathcal{C}^1(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  and solves PDE (5) in the viscosity sense.*

*Proof.* The proof follows similar arguments as the proof of Theorem 3.1. We observe first that, by definition of  $u$ , it holds

$$u(x) = \mathbb{E} \left[ \frac{e^{-\beta \eta^x} h(X_{\eta^x}^x)}{\bar{F}(\eta^x)} \mathbf{1}_{\tau \geq \eta^x} + \frac{\beta e^{-\beta \tau} c_I(X_\tau^x)}{p_I \rho(\tau)} \prod_{i=0}^{|\mathcal{I}|-1} \psi_i^{X_\tau^x} \mathbf{1}_{\tau < \eta^x} \right].$$

where

$$\psi_i^{X_\tau^x} := \prod_{\substack{k=(i,\dots) \in \mathcal{K}^x \\ X_{T_k}^k \notin \mathcal{O}}} \frac{e^{-\beta \Delta T_k} h(X_{T_k}^k)}{\bar{F}(\Delta T_k)} \mathcal{W}_k \prod_{\substack{k=(i,\dots) \in \mathcal{K}^x \\ X_{T_k}^k \in \mathcal{O}}} \frac{\beta e^{-\beta \Delta T_k} c_{I^k}(X_{T_k}^k)}{p_{I^k} \rho(\Delta T_k)} \mathcal{W}_k.$$

It further follows from the branching property that, conditioned on  $\mathcal{F}_0$ ,  $(\psi_i^{X_\tau^x})_{i=0,\dots,|\mathcal{I}|-1}$  are independent random variables, among which the first  $I_0$  are identical in law to  $\psi^{X_0}$ , the next  $I_1$  are identical in law to  $\psi^{X_0} b_1(x) \cdot \mathcal{W}(T_\emptyset, x, W)$  and so on, where  $X_0$  is distributed as  $X_\tau^x$  and independent of  $\mathcal{F}_n$  for all  $n \in \mathbb{N}$ . Through this argument, we deduce that

$$\mathbb{E} \left[ \prod_{i=0}^{|\mathcal{I}|-1} \psi_i^{X_\tau^x} \mid \mathcal{F}_0 \right] \mathbf{1}_{\tau < \eta^x} = u^{I_0}(X_\tau^x) \prod_{i=1}^m v_i^{I_i}(X_\tau^x) \mathbf{1}_{\tau < \eta^x}.$$

where  $(v_i)_{i=1,\dots,m}$ ,  $v_i : \mathcal{O} \mapsto \mathbb{R}$  are defined as

$$v_i(x) := \mathbb{E}[\psi^x b_i(x) \cdot \mathcal{W}(T_\emptyset, x, W)].$$

Working backward along the lines of the proof of Proposition 4.1, we deduce that

$$u(x) = \mathbb{E} \left[ e^{-\beta \eta^x} h(X_{\eta^x}^x) + \int_0^{\eta^x} \beta e^{-\beta s} \left( \sum_{l \in L} c_l u^{l_0} \prod_{i=1}^m v_i^{l_i} \right) (X_s^x) ds \right]. \quad (17)$$

In particular, since  $u$  and  $(v_i)_{i=1,\dots,m}$  are bounded by assumption, it follows from Assumption 4.1 that  $u$  belongs to  $\mathcal{C}^1(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  and

$$\begin{aligned} Du(x) &= \mathbb{E} \left[ e^{-\beta \eta^x} h(X_{\eta^x}^x) \mathcal{W}_{\partial \mathcal{O}}(x, W) + \int_0^{\eta^x} \beta e^{-\beta s} \left( \sum_{l \in L} c_l u^{l_0} \prod_{i=1}^m v_i^{l_i} \right) (X_s^x) \mathcal{W}_{\mathcal{O}}(s, x, W) ds \right] \\ &= \mathbb{E}[\psi^x \mathcal{W}(T_\emptyset, x, W)]. \end{aligned}$$

Thus, for all  $i = 1, \dots, m$ ,  $v_i$  coincides with  $b_i \cdot Du$  and (17) reads as

$$u(x) = \mathbb{E} \left[ e^{-\beta \eta^x} h(X_{\eta^x}^x) + \int_0^{\eta^x} \beta e^{-\beta s} f(\cdot, u, Du)(X_s^x) ds \right].$$

The fact that  $u$  is a viscosity solution of PDE (5) now follows by classical arguments.  $\square$

## 4.2 Automatic Differentiation Formula: the General Case

The aim of this section is to provide sufficient conditions to ensure that Assumption 4.1 holds and to derive explicit formula for  $\mathcal{W}$ . The automatic differentiation formula discussed in the following originates from Thalmaier [35] and was subsequently developed by Delarue [11] and Gobet [16].

- Assumption 4.2.** (i) The coefficients  $(\mu, \sigma)$  belong to  $\mathcal{C}^{1,\alpha}(\bar{\mathcal{O}})$ .  
(ii) The diffusion coefficient  $\sigma$  is uniformly elliptic.  
(iii) The boundary  $\partial\mathcal{O}$  is of class  $\mathcal{C}^2$ .  
(iv) The function  $h$  can be extended to a function of class  $\mathcal{C}^{1,\alpha}$  on  $\bar{\mathcal{O}}$ .

We start by establishing a technical lemma. Fix a finite horizon  $T > 0$  and denote for any  $s > 0$ ,

$$\theta_s(r, y) := \frac{1}{d(y, \partial\mathcal{O})^2 (s-r)}, \quad \text{for all } y \in \mathcal{O}, r \in [0, s].$$

**Lemma 4.1.** Under Assumption 4.2, it holds for all  $x \in \mathcal{O}$  and  $s > 0$ ,

$$\int_0^{\eta^x \wedge s} \theta_s(r, X_r^x) dr = \infty, \quad \mathbb{P} - a.s. \quad (18)$$

In addition, if we denote

$$\zeta_s := \inf \left\{ t > 0 : \int_0^t \theta_s(r, X_r^x) dr = 1 \right\},$$

then there exists  $t < s$  such that  $\zeta_s \leq \eta^x \wedge t$  and for all  $q \geq 1$ ,

$$\mathbb{E} \left[ \left( \int_0^{\zeta_s \wedge T} \theta_{s \wedge T}^2(r, X_r^x) dr \right)^q \right] \leq \frac{C}{d(x, \partial\mathcal{O})^{4q-2} (s \wedge T)^q}, \quad (19)$$

where  $C > 0$  depends on  $q$  and  $T$  but not on  $x$  or  $s$ .

*Proof.* The proof essentially follows from Delarue [11]. Indeed, if  $\mathcal{O}$  is a ball, both identities (18) and (19) are easily obtained by repeating the arguments of Propositions 2.3 and 2.4 in [11] while working with  $(r, y) \mapsto d(y, \partial\mathcal{O})\sqrt{s-r}$  instead of  $(r, y) \mapsto d(y, \partial\mathcal{O})(s-r)$ . For an arbitrary domain, it suffices to work with a  $\mathcal{C}^2$ -extension of the distance to the boundary (see, e.g., Gilbarg and Trudinger [15, Lemma 14.16]). Additionally, it follows from (18) that  $\zeta_s \leq \eta^x \wedge s$ . Furthermore, it holds for all  $t < s$ ,

$$\int_0^t \theta_s(r, X_r^x) dr \mathbf{1}_{t \leq \eta^x} \geq -C^{-1} \log \left( 1 - \frac{t}{s} \right) \mathbf{1}_{t \leq \eta^x},$$

where  $C := \text{diam}(\mathcal{O})^2/4$ . Thus for  $t = (1 - e^{-C})s$ , we have  $\zeta_s \leq \eta^x \wedge t$ .  $\square$

**Proposition 4.2.** Under Assumption 4.2, the assertions of Assumption 4.1 are satisfied with

$$\begin{aligned} \mathcal{W}_{\partial\mathcal{O}}^\top(x, W) &= \int_0^{\zeta_s} \theta_T(r, X_r^x) (\sigma^{-1}(X_r^x) Y_r^x)^\top dW_r, \\ \mathcal{W}_{\mathcal{O}}^\top(s, x, W) &= \int_0^{\zeta_s \wedge T} \theta_{s \wedge T}(r, X_r^x) (\sigma^{-1}(X_r^x) Y_r^x)^\top dW_r, \end{aligned}$$

where  $Y^x$  is the tangent process given by

$$Y_s^x = I_d + \int_0^s D\mu(X_r^x) Y_r^x dr + \sum_{i=1}^d \int_0^s D\sigma_i(X_r^x) Y_r^x dW_r^i.$$

and  $\sigma_i$  denotes the  $i$ th column of  $\sigma$ . In addition, it holds for all  $q \geq 1$ ,

$$\mathbb{E} \left[ \left| \int_0^{\zeta_s \wedge T} \theta_{s \wedge T}(r, X_r^x) (\sigma^{-1}(X_r^x) Y_r^x)^\top dW_r \right|^q \right] \leq \frac{C}{d(x, \partial\mathcal{O})^{2q-1} (s \wedge T)^{\frac{q}{2}}}. \quad (20)$$

This result is a slight extension of the automatic differentiation formula obtained by Delarue [11] and Gobet [16]. We postpone the proof to Appendix A. Note that estimate (20) is an easy consequence of Lemma 4.1. Indeed, successively using the Burkholder-Davis-Gundy inequality, the boundedness of  $\sigma^{-1}$  and the fact that  $\sup_{0 \leq r \leq T} \{|Y_r^x|\}$  has finite moments, we derive for all  $s \leq T$ ,

$$\begin{aligned} \mathbb{E} \left[ \left| \int_0^{\zeta_s} \theta_s(r, X_r^x) (\sigma^{-1}(X_r^x) Y_r^x)^\top dW_r \right|^q \right] &\leq C \mathbb{E} \left[ \sup_{0 \leq r \leq T} \{|Y_r|^q\} \left( \int_0^{\zeta_s} \theta_s^2(r, X_r^x) dr \right)^{\frac{q}{2}} \right] \\ &\leq C \mathbb{E} \left[ \left( \int_0^{\zeta_s} \theta_s^2(r, X_r^x) dr \right)^q \right]^{1/2} \\ &\leq \frac{C}{d(x, \partial\mathcal{O})^{2q-1} s^{\frac{q}{2}}}. \end{aligned}$$

**Remark 4.1.** *The choice of weight function in the automatic differentiation formula is not unique. Actually we can replace  $(\theta_s(r, X_r^x))_{r \geq 0}$  by any predictable process  $(\tilde{\theta}_s(r))_{r \geq 0}$  satisfying appropriate integrability conditions and such that*

$$\tilde{\theta}_s(r) = 0 \text{ for all } r \geq \eta^x \wedge s \quad \text{and} \quad \int_0^{\eta^x \wedge s} \tilde{\theta}_s(r) dr = 1.$$

We refer to Delarue [11] for more examples.

We now provide sufficient conditions to ensure that  $(\psi^x)_{x \in \mathcal{O}}$  and  $(\psi^x b_i(x) \cdot \mathcal{W}(T_\emptyset, x, W))_{x \in \mathcal{O}}$ ,  $i = 1, \dots, m$ , are uniformly bounded in  $L^q$  for  $q \geq 1$ . Let us define

$$\begin{aligned} C_{1,q} &:= \sup_{t \geq 0} \left\{ \frac{e^{-\beta t}}{\bar{F}(t)} \right\} \sup_{x \in \partial\mathcal{O}} \{|h(x)|\} \sup_{x \in \mathcal{O}, i=1, \dots, m} \left\{ \mathbb{E} [|b_i(x) \cdot \mathcal{W}_{\partial\mathcal{O}}(x, W)|^q]^{\frac{1}{q}} \vee 1 \right\}, \\ C_{2,q} &:= \sup_{x \in \mathcal{O}, l \in L} \left\{ \frac{|c_l(x)|}{p_l} \right\} \sup_{t \geq 0, x \in \mathcal{O}, i=1, \dots, m} \left\{ \frac{\beta e^{-\beta t}}{\rho(t)} \left( \mathbb{E} [|b_i(x) \cdot \mathcal{W}_{\mathcal{O}}(t, x, W)|^q]^{\frac{1}{q}} \vee 1 \right) \right\}. \end{aligned}$$

In view of (20),  $C_{1,q}$  and  $C_{2,q}$  are clearly finite if the following condition is satisfied:

$$\sup_{t \geq 0} \left\{ \frac{e^{-\beta t}}{\bar{F}(t)} \right\} \vee \sup_{t \geq 0} \left\{ \frac{\beta e^{-\beta t}}{\sqrt{t} \rho(t)} \right\} \vee \sup_{l \in L} \left\{ \frac{\|c_l\|_\infty}{p_l} \right\} \vee \sup_{x \in \mathcal{O}, i=1, \dots, m} \left\{ \frac{|b_i(x)|}{d(x, \partial\mathcal{O})^{2-\frac{1}{q}}} \right\} < \infty.$$

Unlike in the case of linear gradient term, we need to chose a lifetime distribution  $\rho(t) \neq \beta e^{-\beta t}$  to ensure that the first two terms in the above expression are finite. For instance, we can take a Gamma distribution with shape parameter 0.5 and rate parameter  $\beta' < \beta$ , i.e.,  $\rho(t) = \sqrt{\beta'}/(\pi t) e^{-\beta' t}$ . Another choice consists in taking a generalized gamma distribution of the form  $\rho(t) = \beta'/(2\sqrt{t}) e^{-\beta' \sqrt{t}}$  for any  $\beta' > 0$ .

Let us denote

$$\delta := 1 - \inf_{x \in \mathcal{O}} \left\{ \mathbb{E} [\bar{F}(\eta^x)] \right\}.$$

Similar to Section 3.3, we then introduce a new probability mass function  $(\tilde{p}_l)_{l \in \tilde{L}}$  with  $\tilde{L} := L \cup \{0_{m+1}\}$  where  $0_{m+1}$  is the zero element of  $\mathbb{N}^{m+1}$  as follows:

$$\tilde{p}_{0_{m+1}} := 1 - \delta + \delta p_{0_{m+1}} \quad \text{and} \quad \tilde{p}_l := \delta p_l \text{ for all } l \neq 0_{m+1},$$

and the corresponding transition matrix  $\tilde{P} = (\tilde{P}_{i,j})_{i,j \geq 0}$  given by

$$\tilde{P}_{0,0} = 1 \quad \text{and} \quad \tilde{P}_{i,i+j-1} = \sum_{|l|=j} \tilde{p}_l \text{ for all } i \geq 1, j \geq 0.$$

**Proposition 4.3.** *Denote  $C_{0,q} := C_{1,q} \vee C_{2,q}$ .*

- (i) *If  $C_{0,q} \leq 1$ , then  $(\psi^x)_{x \in \mathcal{O}}$  and  $(\psi^x b_i(x) \cdot \mathcal{W}(T_\emptyset, x, W))_{x \in \mathcal{O}}$ ,  $i = 1, \dots, m$ , are bounded by 1 in  $L^q$ .*
- (ii) *Denote by  $R$  the common radius of convergence of  $f(s) := \sum_{l \in L} p_l s^{|l|}$  and  $\tilde{f}(s) := \sum_{l \in \tilde{L}} \tilde{p}_l s^{|l|}$ . If  $R > 1$  and  $\sum_{l \in \tilde{L}} |l| \tilde{p}_l < 1$ , then  $(\psi^x)_{x \in \mathcal{O}}$  and  $(\psi^x b_i(x) \cdot \mathcal{W}(T_\emptyset, x, W))_{x \in \mathcal{O}}$ ,  $i = 1, \dots, m$ , are uniformly bounded in  $L^q$  provided that  $C_{0,q}^q \leq \gamma$  where*

$$\gamma := \frac{s^*}{\tilde{f}(s^*)} = \frac{s^*}{1 - \delta + \delta f(s^*)},$$

where  $s^*$  is the solution (if any) to  $s \tilde{f}'(s) = \tilde{f}(s)$  and  $s^* = R$  otherwise. In addition, if  $R = +\infty$ , then  $\gamma$  goes to infinity as  $\text{diam}(\mathcal{O})$  goes to zero.



*Proof.* Let us denote by  $\mathcal{G}$  the following  $\sigma$ -algebra:

$$\mathcal{G} := \sigma(\tau^k, I^k, k \in \mathbb{K}).$$

Conditioning by  $\mathcal{G}$ , we obtain for all  $x \in \mathcal{O}$ ,

$$\mathbb{E}[|\psi^x|^q] \leq \mathbb{E}\left[C_{0,q}^{g|\mathcal{K}^x|}\right],$$

and the same inequality holds if we consider  $\psi^x b_i(x) \cdot \mathcal{W}(T_\emptyset, x, W)$  instead of  $\psi^x$ . Part (i) of Proposition 4.3 follows immediately. As for Part (ii), the desired result follows by the same arguments as in Proposition 3.5.  $\square$

### 4.3 Automatic Differentiation Formula: the One-Dimensional Case

The automatic differentiation formula given by Proposition 4.2, though quite general, has several drawbacks for numerical applications. Indeed one needs to compute a stochastic integral for each particle that holds a non-zero mark. In addition, the weight function  $\mathcal{W}(s, x, W)$  explodes as  $s$  goes to 0 or  $x$  approaches the boundary. In this section, we provide a simpler automatic differentiation formula that is satisfied for a one-dimensional Brownian motion exiting from an interval.

Let us assume throughout this section that  $d = 1$ ,  $b = 0$ ,  $\sigma = 1$  and  $\mathcal{O} = (-r, r)$  for some  $r > 0$ . We denote

$$W_s^x := x + W_s, \quad s \geq 0, x \in \mathcal{O}.$$

We also introduce the function  $\mathfrak{W} : \mathcal{O} \times \bar{\mathcal{O}} \rightarrow \mathbb{R}$  as

$$\mathfrak{W}(x, y) := \begin{cases} \frac{\sqrt{2\beta}}{\tanh(\sqrt{2\beta}(x+r))}, & \text{if } y > x, \\ \frac{\sqrt{2\beta}}{\tanh(\sqrt{2\beta}(x-r))}, & \text{if } y \leq x. \end{cases}$$

Both lemmas below show that Assumption 4.1 is satisfied in this setting.

**Lemma 4.2.** *For any  $g : \mathcal{O} \rightarrow \mathbb{R}$  bounded measurable, the map  $x \mapsto \mathbb{E}[\int_0^{\eta^x} e^{-\beta s} g(W_s^x) ds]$  belongs to  $\mathcal{C}^1(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  and*

$$DE \left[ \int_0^{\eta^x} e^{-\beta s} g(W_s^x) ds \right] = \mathbb{E} \left[ \int_0^{\eta^x} e^{-\beta s} g(W_s^x) \mathfrak{W}(x, W_s^x) ds \right].$$

*Proof.* Let us denote

$$\chi(x) := \mathbb{E} \left[ \int_0^{\eta^x} e^{-\beta s} g(W_s^x) ds \right].$$

We first assume that  $g$  is continuous. Then  $u$  satisfies the following ODE:

$$\frac{1}{2}\chi'' - \beta\chi + g = 0 \quad \text{in } \mathcal{O}, \quad \chi = 0 \quad \text{on } \partial\mathcal{O}. \quad (21)$$

By standard arguments, including variation of parameters, we deduce that

$$\chi(x) = \int_{-r}^r G(x, y) g(y) dy, \quad (22)$$

where  $G : \bar{\mathcal{O}} \times \mathcal{O} \rightarrow \mathbb{R}$  is given by

$$G(x, y) := \frac{-2}{\sqrt{2\beta}} \left( \sinh(\sqrt{2\beta}(x-y)^+) - \frac{\sinh(\sqrt{2\beta}(r+x))}{\sinh(2\sqrt{2\beta}r)} \sinh(\sqrt{2\beta}(r-y)) \right).$$

By a monotone class argument, (22) remains valid if  $g$  is only assumed bounded measurable. Then the continuity of  $\chi$  follows immediately from the dominated convergence theorem. Further, a direct calculation yields that

$$\frac{\partial_x G(x, y)}{G(x, y)} = \begin{cases} \frac{\sqrt{2\beta}}{\tanh(\sqrt{2\beta}(x+r))}, & \text{if } y > x, \\ \frac{\sqrt{2\beta}}{\tanh(\sqrt{2\beta}(x-r))}, & \text{if } y < x. \end{cases}$$

The desired result then follows by differentiation under the integral sign in (22).  $\square$

**Lemma 4.3.** *The map  $x \mapsto \mathbb{E}[e^{-\beta\eta^x} h(W_{\eta^x}^x)]$  belongs to  $\mathcal{C}^1(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  and*

$$D\mathbb{E} \left[ e^{-\beta\eta^x} h(W_{\eta^x}^x) \right] = \mathbb{E} \left[ e^{-\beta\eta^x} h(W_{\eta^x}^x) \mathfrak{W}(x, W_{\eta^x}^x) \right].$$

*Proof.* Let us denote

$$\varphi(x) := \mathbb{E} \left[ e^{-\beta\eta^x} h(W_{\eta^x}^x) \right].$$

It satisfies the following ODE:

$$\frac{1}{2}\varphi'' - \beta\varphi = 0 \quad \text{in } \mathcal{O}, \quad \varphi = h \quad \text{on } \partial\mathcal{O}.$$

It follows that

$$\varphi(x) = H(x, r)h(r) + H(x, -r)h(-r),$$

where  $H : \mathcal{O} \times \partial\mathcal{O} \rightarrow \mathbb{R}$  is given by

$$H(x, y) := \frac{\sinh(\sqrt{2\beta}(x+y))}{\sinh(2\sqrt{2\beta}y)}.$$

In particular,  $\varphi$  clearly belongs to  $\mathcal{C}^1(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$ . To conclude, it remains to observe that

$$\frac{\partial_x H(x, y)}{H(x, y)} = \frac{\sqrt{2\beta}}{\tanh(\sqrt{2\beta}(x+y))}.$$

□

**Remark 4.2.** *We observe that the weight function still explodes when  $x$  approaches the boundary as in Section 4.2. In view of the proof of Lemma 4.2, this feature cannot be avoided since  $\chi'$  satisfies an ODE of type (21) with non-zero boundary conditions. As a consequence, we need to assume once again that the non-linear gradient term vanishes at the boundary of the domain in order to control the explosion of the weight function.*

Since the weight function does not explode in time anymore, we can work with  $\rho(t) = \beta e^{-\beta t}$  as in Section 3. Thus we have

$$\psi^x := \prod_{\substack{k \in \mathcal{K}^x \\ X_{T_k}^k \notin \mathcal{O}}} h(X_{T_k}^k) \mathcal{W}_k \prod_{\substack{k \in \mathcal{K}^x \\ X_{T_k}^k \in \mathcal{O}}} \frac{c_{I^k}(X_{T_k}^k)}{p_{I^k}} \mathcal{W}_k,$$

where  $\mathcal{W}_k$  are constructed from the formulas of Lemmas 4.2–4.3. We now provide sufficient conditions to ensure that  $(\psi^x)_{x \in \mathcal{O}}$  and  $(\psi^x b_i(x) \cdot \mathcal{W}(T_\emptyset, x, W))_{x \in \mathcal{O}}$ ,  $i = 1, \dots, m$ , are uniformly bounded in  $L^q$ . Let us define

$$C_1 := \sup_{x \in \mathcal{O}, y \in \bar{\mathcal{O}}, i=1, \dots, m} \{|b_i(x) \mathfrak{W}(x, y)|\} \vee 1.$$

Clearly, the constant  $C_1$  is finite if the following condition is satisfied:

$$\sup_{x \in \mathcal{O}, i=1, \dots, m} \left\{ \frac{|b_i(x)|}{d(x, \partial\mathcal{O})} \right\} < \infty.$$

Now we consider the family  $(\bar{\psi}^x)_{x \in \mathcal{O}}$  given by

$$\bar{\psi}^x := \prod_{\substack{k \in \mathcal{K}^x \\ X_{T_k}^k \notin \mathcal{O}}} C_1 h(X_{T_k}^k) \prod_{\substack{k \in \mathcal{K}^x \\ X_{T_k}^k \in \mathcal{O}}} C_1 \frac{c_{I^k}(X_{T_k}^k)}{p_{I^k}}.$$

The tools developed in Section 3.3 allow to derive sufficient conditions to ensure that the family  $(\bar{\psi}^x)_{x \in \mathcal{O}}$  is bounded in  $L^q$ . Since  $|\psi^x| \leq |\bar{\psi}^x|$ , the boundedness on  $(\psi^x)_{x \in \mathcal{O}}$  in  $L^q$  follows immediately. The same holds for  $(\psi^x b_i(x) \cdot \mathcal{W}(T_\emptyset, x, W))_{x \in \mathcal{O}}$  for  $i = 1, \dots, m$ .

## 5 Numerical Examples with Related Technical Discussions

In this section we demonstrate the application of our theoretical results with the help of several examples. We restrict our attention to semi-linear elliptic PDEs over a rectangular domain where the underlying diffusion is a Brownian motion. This allows us to generate unbiased samples for exit time and position using the walk on squares scheme as implemented in the numerical library `exitbm` developed by Lejay [24]. We note that our theoretical results are applicable more widely but the numerical estimates for semi-linear PDEs driven by Brownian motion over rectangular domain do not involve any bias. We performed all the numerical computations on a machine with 2,5 GHz Intel Core i5 processor and 4GB RAM.

### 5.1 Example 1

We begin with the following semi-linear PDE in one-dimension (ODE)

$$u'' - u + u^3 = 0,$$

with explicit solution  $u(x) = \frac{\sqrt{2}}{\cosh(x)}$ . To test the validity of our probabilistic representation in Theorem 3.1, we reformulate the above PDE, in the form (7), as following

$$\frac{1}{2}u'' + \left(\frac{1}{2}u^3 + \frac{1}{2}u - u\right) = 0 \text{ in } (-r, r), \quad u(x) = \frac{\sqrt{2}}{\cosh(x)}, \quad x \in \{-r, r\}, \quad (23)$$

For domain  $\mathcal{O} = (-r, r)$ , we know from standard literature that the first eigenvalue of Laplacian is given as  $\lambda_1 = \frac{\pi^2}{4r^2}$ . Then, by choosing  $\beta = 1$ ,  $p_1 = \frac{1}{2}$  and  $p_3 = \frac{1}{2}$ , the criteria for almost sure extinction of the branching Brownian motion (6) holds for any  $r \leq \frac{\pi}{\sqrt{8}}$ . In addition, the integrability condition as presented in Proposition 3.4 is satisfied for  $q = 1$  by  $v(x) = \frac{\sqrt{2}}{\cosh(x)}$ . Therefore, our probabilistic representation of the solution given as

$$u(x) = \mathbb{E} \left[ \prod_{\substack{k \in \mathcal{K}^x \\ X_{T_k}^k \notin \mathcal{O}}} h(X_{T_k}^k) \right],$$

is valid for any  $r \leq \frac{\pi}{\sqrt{8}}$ . Once more, we make use of Proposition 3.4 and see that for  $q = 2$ , it is satisfied by  $v(x) = \sqrt{1 + \lambda} \cos(\sqrt{\lambda}x)$  with  $\lambda = 6$  and  $r \leq 0.338$ . This implies that  $(\psi^x)_{x \in \mathcal{O}}$  is uniformly bounded in  $L^2$  and we are certain that it has finite variance. In Table 5.1, for  $r = 0.3$ , we illustrate numerical results for different values of starting position  $x$  with  $1 \times 10^6$  Monte Carlo sample paths which exhibits the accuracy of our estimator.

$x$	Estimate	99% conf. interval	Std. Dev./Mean	Relative error	Run time (secs)
0	1.4144	[1.4134, 1.4153]	0.2644	0.0112%	13
-0.2	1.3859	[1.3852, 1.3866]	0.1872	0.0358%	26

Table 5.1: Numerical results for the unbiased estimator of probabilistic representation of one-dimensional semi-linear PDE (23).

We observe that even for larger values of  $r$ , our probabilistic representation provides an accurate estimator but with a standard deviation estimate which converges slowly. In Table 5.2, we vary the size of the interval to test the accuracy and notice that as more branching particles are generated for larger intervals, the estimator has higher variance and run time.

$r$	99% conf. interval	Std. Dev./Mean	Run time (secs)
0.4	[1.413654, 1.417886]	0.5793	17
0.5	[1.410199, 1.417666]	0.9134	23

Table 5.2: Numerical results for the unbiased estimator for different values of  $r$  and  $x = 0$ .

**Remark 5.1.** If  $r \geq \operatorname{arcosh}(\sqrt{2})$ , the solution provided by the probabilistic representation takes values in  $[0, 1]$  and thus it does not coincide with  $u(x) = \frac{\sqrt{2}}{\cosh(x)}$ . In particular, in view of Proposition 3.1, it turns out that the sequence  $(\psi_n^x)_{n \in \mathbb{N}}$  is not uniformly integrable for  $\operatorname{arcosh}(\sqrt{2}) \leq r \leq \frac{\pi}{\sqrt{8}}$ . In Table 5.3, we present the results which support our observation.

$x$	Estimate	99% conf. interval	Std. Dev./Mean	Run time (secs)
0	0.9597	[0.9595, 0.9600]	0.0746	123
-0.2	0.9612	[0.9611, 0.9614]	0.0722	121

Table 5.3: Numerical results for the unbiased estimator of probabilistic representation for domain  $(-0.9, 0.9)$ .

## 5.2 Example 2

We consider another semi-linear PDE in one dimension. It is given as

$$\frac{1}{2}u'' + \left(\frac{1}{2} - \frac{3}{2}u^2 - u\right) = 0, \quad (24)$$

with solution  $u(x) = 1 + 2 \tan^2(x)$ . Then, for domain  $\mathcal{O} = (-r, r)$ ,  $\beta = 1$  and probability mass function as  $p_0 = 0.25$  and  $p_2 = 0.75$ , the branching Brownian motion goes extinct almost surely for  $r \leq \frac{\pi}{2}$ . Next, to verify that our probabilistic representation holds, we work with the conditions in Proposition 3.5. We first compute

$$\tilde{p}_0 := p_0 + (1 - p_0) \inf_{x \in \mathcal{O}} \left\{ \mathbb{E} \left[ e^{-\eta^x} \right] \right\} \quad \text{and} \quad \tilde{p}_2 := p_2 \left( 1 - \inf_{x \in \mathcal{O}} \left\{ \mathbb{E} \left[ e^{-\eta^x} \right] \right\} \right).$$

From standard calculations, we get

$$\inf_{x \in \mathcal{O}} \left\{ \mathbb{E} \left[ e^{-\eta^x} \right] \right\} = \inf_{x \in \mathcal{O}} \left\{ \frac{\cosh(\sqrt{2}x)}{\cosh(\sqrt{2}r)} \right\} = \frac{1}{\cosh(\sqrt{2}r)}.$$

This gives us  $\tilde{p}_0 = 0.25 + \frac{0.75}{\cosh(\sqrt{2}r)}$  and  $\tilde{p}_2 = 0.75 \left( 1 - \frac{1}{\cosh(\sqrt{2}r)} \right)$ . In addition, the integrability constant  $C_0 = \max\{2, 1 + 2 \tan^2(r)\}$ . Next, in view of (15), the choice of threshold is given as  $\gamma = \frac{1}{\sqrt{4\tilde{p}_2(1-\tilde{p}_2)}}$ . Then, for our probabilistic representation of the solution to be valid, we need that  $C_0 < \gamma$ , which gives us that  $r \leq 0.31$ . As in the previous example, we would like to compute a confidence interval for the estimate of the solution. Once again, by checking the conditions in Proposition 3.5, we get that for our estimator  $(\psi_x)_{x \in \mathcal{O}}$  to be uniformly bounded in  $L^2$ , we must choose  $r \leq 0.146$ . We set  $r = 0.14$ , and obtain the following results for  $1 \times 10^6$  Monte Carlo sample paths in Table 5.4 which exhibits the accuracy of our estimator.

$x$	Estimate	99% conf. interval	Std. Dev./Mean	Relative error	Run time (secs)
0	0.9999	[0.9989, 1.0001]	0.3931	0.0097%	10
-0.1	1.0198	[1.0190, 1.0205]	0.2745	0.0351%	12

Table 5.4: Numerical results for the unbiased estimator of probabilistic representation of one-dimensional semi-linear PDE (24).

In the course of our experiments, we chose different values of  $r$  and observed that even for higher values, we are able to obtain accurate estimates of the solution. For  $r = 0.3$ , we present the results in Table 5.5.

$x$	Estimate	99% conf. interval	Std. Dev./Mean	Relative error	Run time (secs)
0.0	1.0006	[0.9971, 1.0040]	1.3378	0.0575%	12
-0.1	1.0203	[1.0168, 1.0234]	1.3641	0.0209%	17

Table 5.5: Numerical results for the unbiased estimator of probabilistic representation of one-dimensional semi-linear PDE (24).

### 5.3 Example 3

Next, we consider a multidimensional semi-linear PDE

$$\Delta u = 2d(u^3 + u), \quad (25)$$

with an explicit solution  $u(x) = \tan(\sum_{i=1}^d x_i)$  where  $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ . We reformulate the above PDE as

$$\frac{1}{2}\Delta u + d(-u^3 - u) = 0 \text{ in } \mathcal{O}, \quad u = \tan\left(\sum_{i=1}^d x_i\right), x \in \partial\mathcal{O},$$

with  $\mathcal{O} = (-r, r)^d$  for  $r > 0$  and set  $\beta = d$  and  $p_3 = 1$ . With  $\lambda_1 = \frac{d\pi^2}{4r^2}$ , the branching Brownian motion goes extinct for  $r \leq \frac{\pi}{4}$ .

We first consider  $d = 2$ . For  $r < \frac{\pi}{8}$ , the integrability constant  $C_0 < 1$  and thus our estimator will have finite moments. In order to check if our probabilistic representation is valid for even a larger domain, we verify the condition in Proposition 3.5 for  $q = 1$ . We need to compute

$$\tilde{p}_0 := \inf_{x \in \mathcal{O}} \left\{ \mathbb{E}[e^{-d\eta^x}] \right\} \text{ and } \tilde{p}_3 := 1 - \tilde{p}_0.$$

In the absence of an explicit formula for the Laplace transform of exit time of Brownian motion from a multidimensional rectangular domain, we estimate it by sampling the exit time using the numerical library developed by Lejay [24]. The integrability constant needs to be less than the threshold  $\gamma$  in Proposition 3.5 for our probabilistic representation to be valid. In view of (15), the choice of threshold is given as  $\gamma = \sqrt[3]{\frac{4}{27\tilde{p}_0^2(1-\tilde{p}_0)}}$ . From our estimation procedures, we deduce that  $C_0 < \gamma$  for  $r < 0.484$ . Here, we exhibit our numerical results for  $5 \times 10^5$  Monte Carlo sample paths with  $r = 0.48$  and different values of  $x$  in Table 5.6. We observe that our estimator remains accurate but with an increased variance which is an expected effect due to an increased variance in the estimation of exit time and position of multi-dimensional Brownian motion using the library `exitbm`.

$x$	Estimate	99% conf. interval	Std. Dev./Mean	Relative error	Run time (secs)
(0.0,0.0)	-0.0001	[-0.0021, 0.0019]	–	–	30
(0.1,0.0)	0.0998	[0.0978, 0.1018]	5.4883	0.5726%	34
(0.2,0.1)	0.3095	[0.3075, 0.3114]	1.7342	0.0426%	30
(0.2,0.2)	0.4245	[0.4226, 0.4265]	1.2646	0.4144%	25

Table 5.6: Numerical results for the unbiased estimator of probabilistic representation of semi-linear PDE (25) in  $d = 2$  with  $\beta = 2$ .

Typically, finite difference methods are quite commonly used to numerically solve PDEs. However, such methods are practically implementable and provide stable estimates only when  $d \leq 3$ . Thus, we study PDE (25) in the case of  $d = 4$  to illustrate the broader applicability of our method. We extend the modules in library `exitbm` for our purpose using the existing functions. For  $r < \frac{\pi}{16}$ , the integrability constant  $C_0 < 1$  and thus the estimator will have finite moments. Furthermore, through numerical estimation procedures, we have that  $C_0 < \gamma$  for  $r < 0.242$ . In Table 5.7, we exhibit our results with  $5 \times 10^5$  Monte Carlo sample paths for  $r = 0.24$  which shows that our probabilistic representation provides an accurate estimator of the true solution.

$x$	Estimate	99% conf. interval	Std. Dev./Mean	Relative error	Run time (secs)
(0,0,0,0)	0.0000	[-0.0010, 0.0012]	–	–	190
(0.1,0,0,0)	0.1009	[0.0999, 0.1019]	2.6185	0.5584%	309
(0.1,0.1,0,0)	0.2021	[0.2012, 0.2031]	1.2986	0.2924%	413
(0.1,0.1,0.1,0)	0.3097	[0.3087, 0.31072]	0.8750	0.1305%	510

Table 5.7: Numerical results for the unbiased estimator of probabilistic representation of semi-linear PDE (25) in  $d = 4$  with  $\beta = 4$ .

## 6 Conclusions

In this work, we introduced a probabilistic representation for the solution of semi-linear elliptic PDEs with polynomial non-linearity by using the theory of branching diffusion processes. We performed a detailed analysis to derive explicit conditions under which our representations remain valid. In the linear gradient case, we essentially established that it holds provided that the domain is small enough. In the general case, we need to assume further that the nonlinear gradient term vanishes at the boundary of the domain to balance the explosion of the weight functions. As illustrated by our analysis of the one-dimensional case, this feature is inherent to any automatic differentiation formula. This restricts the choice of semi-linear elliptic PDEs for which we could use our probabilistic representation to obtain numerical estimates. For semi-linear elliptic PDEs of form (7), we illustrated the applicability of our theoretical results to obtain numerical estimates with the help of several examples including multi-dimensional. Finally, in the numerical implementation of semi-linear elliptic PDEs driven by general diffusion processes, the error analysis due to the discretization of the process has been left for future research.

## Acknowledgement

The authors would like to thank Pierre Henry-Labordère, Zhenjie Ren, Xiaolu Tan and Nizar Touzi for their valuable comments and suggestions. The first author research was conducted while at CMAP, École Polytechnique and is part of the Chair *Financial Risks* of the *Risk Foundation*. The second author acknowledges the financial support of ERC 321111 Rofirm.

## A Proof of Proposition 4.2

For the purpose of clarity, we split the proof in two lemmas.

**Lemma A.1.** *For any  $g$  bounded measurable, the map  $x \mapsto \mathbb{E}[\int_0^{\eta^x} e^{-\beta s} g(X_s^x) ds]$  belongs to  $\mathcal{C}^1(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  and*

$$D\mathbb{E} \left[ \int_0^{\eta^x} e^{-\beta s} g(X_s^x) ds \right] = \mathbb{E} \left[ \int_0^{\eta^x} e^{-\beta s} g(X_s^x) \mathcal{W}_{\mathcal{O}}(s, x, W) ds \right].$$

*Proof.* The proof consists of four steps. The first three parts essentially follow by repeating the arguments in Gobet [16]. Denote

$$\chi(s, x) := \mathbb{E} [g(X_s^x) \mathbf{1}_{s < \eta^x}].$$

*First step.* Let us collect first preliminary results on the regularity of the function  $\chi$ . Recall that, as established in the proof of Lemma 3.1, we have the following representation

$$\chi(s, x) = \int_{\mathcal{O}} G(s, x; 0, y) g(y) dy,$$

where  $G$  is the Green function of PDE (12) (see, *e.g.*, Ladyženskaja *et al.* [23, Theorem 4.16.2] or Friedman [14, Theorem 3.16]). Furthermore, in view of Theorem 4.16.3 in [23], for any  $0 < t < T$ , we can differentiate under the integral sign and thus  $\chi$  belongs to  $\mathcal{C}_b^{1,2}([t, T] \times \mathcal{O})$  and satisfies

$$\partial_t \chi - \mathcal{L}\chi = 0, \quad \text{in } [t, T] \times \mathcal{O}.$$

In addition, under Assumption 4.2,  $\chi$  belongs to  $\mathcal{C}^{1,3}([t, T] \times \mathcal{O})$  (see, *e.g.*, Friedman [14, Theorem 3.10]).

*Second step.* Next we show that  $N_\theta := D\chi(s - \theta, X_\theta^x) Y_\theta^x$  is a martingale up to time  $\zeta_s$ , *i.e.*,  $(N_{\zeta_s \wedge \theta})_{\theta \geq 0}$  is a martingale. Using Itô's formula, we obtain

$$N_{\zeta_s \wedge \theta} = D\chi(s, x) + \sum_{i=1}^d \int_0^{\zeta_s \wedge \theta} (D\chi(s-r, X_r^x) D\sigma_i(X_r^x) + \sigma_i^\top(X_r^x) D^2\chi(s-r, X_r^x)) Y_r^x dW_r^i.$$

In addition, due to Lemma 4.1, there exists  $t < s$  such that  $\zeta_s \leq t$ . As the maps  $D\chi$  and  $D^2\chi$  are bounded in  $[s-t, s] \times \mathcal{O}$ , we conclude that  $(N_{\zeta_s \wedge \theta})_{\theta \geq 0}$  is a martingale.

*Third step.* Then let us show that

$$D\chi(s, x) = \mathbb{E}\left[g(X_s^x)\mathbf{1}_{s < \eta^x}\mathcal{W}_{\mathcal{O}}(s, x, W)\right].$$

Indeed, Itô's formula yields that

$$\chi(s - \zeta_s, X_{\zeta_s}^x) = \chi(s, x) + \int_0^{\zeta_s} D\chi(s - r, X_r^x)\sigma(X_r^x) dW_r.$$

Successively using Markov property and Itô's isometry, we obtain

$$\begin{aligned} & \mathbb{E}\left[g(X_s^x)\mathbf{1}_{s \leq \eta^x}\left(\int_0^{\zeta_s} \theta_s(r, X_r^x)(\sigma^{-1}(X_r^x)Y_r^x)^\top dW_r\right)^\top\right] \\ &= \mathbb{E}\left[\chi(s - \zeta_s, X_{\zeta_s}^x)\left(\int_0^{\zeta_s} \theta_s(r, X_r^x)(\sigma^{-1}(X_r^x)Y_r^x)^\top dW_r\right)^\top\right] \\ &= \mathbb{E}\left[\int_0^{\zeta_s} \theta_s(r, X_r^x)N_r dr\right]. \end{aligned}$$

Once more, successively using the results that  $N_{\zeta_s \wedge r} = \mathbb{E}[N_{\zeta_s} | \mathcal{F}_r]$  and  $\int_0^{\zeta_s} \theta_s(r, X_r^x) dr = 1$ , we deduce that

$$\mathbb{E}\left[\int_0^{\zeta_s} \theta_s(r, X_r^x)N_r dr\right] = \mathbb{E}\left[N_{\zeta_s} \int_0^{\zeta_s} \theta_s(r, X_r^x) dr\right] = \mathbb{E}[N_{\zeta_s}] = D\chi(s, x).$$

*Fourth step.* We are now in a position to complete the proof. From Fubini's theorem we obtain

$$\mathbb{E}\left[\int_0^{\eta^x} e^{-\beta s} g(X_s^x) ds\right] = \int_0^\infty e^{-\beta s} \mathbb{E}\left[g(X_s^x)\mathbf{1}_{s < \eta^x}\right] ds. \quad (26)$$

Finally, in view of the estimates in Proposition 4.2, we can apply differentiation under the integral sign in (26). This completes the proof.  $\square$

**Lemma A.2.** *The map  $x \mapsto \mathbb{E}[e^{-\beta\eta^x} h(X_{\eta^x}^x)]$  belongs to  $\mathcal{C}^1(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  and*

$$D\mathbb{E}\left[e^{-\beta\eta^x} h(X_{\eta^x}^x)\right] = \mathbb{E}\left[e^{-\beta\eta^x} h(X_{\eta^x}^x)\mathcal{W}_{\partial\mathcal{O}}(x, W)\right].$$

*Proof.* This proof is similar to the proof of Lemma A.1 and we split it in three parts. Denote

$$\varphi(x) := \mathbb{E}\left[e^{-\beta\eta^x} h(X_{\eta^x}^x)\right].$$

*First step.* Let us collect first preliminary results on the regularity of the function  $\varphi$ . Recall that  $\varphi \in \mathcal{C}^2(\mathcal{O}) \cap \mathcal{C}(\bar{\mathcal{O}})$  satisfies the following PDE:

$$\mathcal{L}\varphi - \beta\varphi = 0 \quad \text{in } \mathcal{O}, \quad \varphi = h \quad \text{on } \partial\mathcal{O}.$$

In addition, since  $h \in \mathcal{C}^{1,\alpha}(\bar{\mathcal{O}})$  by assumption, it is known that  $\varphi \in \mathcal{C}^{1,\alpha}(\bar{\mathcal{O}})$  (see, e.g., Gilbarg and Trudinger [15, Theorem 8.34]). Furthermore, under Assumption 4.2, the function  $\varphi$  belongs to  $\mathcal{C}^3(\mathcal{O})$  (see, e.g., Gilbarg and Trudinger [15, Theorem 6.17]).

*Second step.* Next we show that  $N_s := e^{-\beta s} D\varphi(X_s^x)Y_s^x$  is a martingale up to time  $\eta^x$ , i.e.,  $(N_{\eta^x \wedge s})_{s \geq 0}$  is a martingale. Using Itô's formula, we obtain

$$N_{\eta^x \wedge s} = D\varphi(x) + \sum_{i=1}^d \int_0^{\eta^x \wedge s} e^{-\beta r} (D\varphi(X_r^x)D\sigma_i(X_r^x) + \sigma_i^\top(X_r^x)D^2\varphi(X_r^x)) Y_r^x dW_r^i.$$

Thus  $(N_{\eta^x \wedge s})_{s \geq 0}$  is a local martingale. To conclude, it remains to observe that for any  $s \geq 0$ ,

$$\mathbb{E}\left[\sup_{0 \leq r \leq s} \{N_{\eta^x \wedge r}\}\right] < \infty,$$

since  $D\varphi$  is bounded and  $\sup_{0 \leq r \leq s} \{|Y_r^x|\}$  admits finite moments.

*Third step.* We are now in a position to complete the proof. Indeed, Itô's formula yields for all  $s \geq 0$ ,

$$h(X_{\eta^x}^x) = \varphi(x) + \int_0^{\eta^x} D\varphi(X_r^x) \sigma(X_r^x) dW_r.$$

Using Itô's isometry, we obtain

$$\mathbb{E} \left[ h(X_{\eta^x}^x) \left( \int_0^{\zeta^T} \theta_T(r, X_r^x) (\sigma^{-1}(X_r^x) Y_r^x)^\top dW_r \right)^\top \right] = \mathbb{E} \left[ \int_0^{\zeta^T} \theta_T(r, X_r^x) N_r dr \right].$$

Finally, successively using the results that  $N_{\zeta^T \wedge r} = \mathbb{E}[N_{\zeta^T} | \mathcal{F}_r]$  and  $\int_0^{\zeta^T} \theta_T(r, X_r^x) dr = 1$ , we deduce that

$$\mathbb{E} \left[ \int_0^{\zeta^T} \theta_T(r, X_r^x) N_r dr \right] = \mathbb{E} \left[ N_{\zeta^T} \int_0^{\zeta^T} \theta_T(r, X_r^x) dr \right] = \mathbb{E} [N_{\zeta^T}] = D\varphi(x).$$

□

## References

- [1] K. B. Athreya and P. E. Ney. *Branching processes*, volume 196. Springer Science & Business Media, 2012.
- [2] M. Badiale and E. Serra. *Semilinear elliptic equations for beginners*. Universitext. Springer, London, 2011. ISBN 978-0-85729-226-1.
- [3] R. F. Bass. *Diffusions and elliptic operators*. Probability and its Applications (New York). Springer-Verlag, New York, 1998.
- [4] M. Bossy, N. Champagnat, H. Leman, S. Maire, L. Violeau, and M. Yvinec. Monte Carlo methods for linear and non-linear Poisson-Boltzmann equation. *CEMRACS 2013—modelling and simulation of complex systems: stochastic and deterministic approaches, ESAIM Proc. Surveys*, 48:420–446, 2015.
- [5] B. Bouchard and S. Menozzi. Strong approximations of bsdes in a domain. *Bernoulli*, 15(4):1117–1147, 2009.
- [6] B. Bouchard and N. Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their Applications*, 111(2):175–206, 2004.
- [7] B. Bouchard, S. Geiss, and E. Gobet. First time to exit of a continuous Itô process: General moment estimates and L1-convergence rate for discrete time approximations. *ArXiv e-prints*, July 2013.
- [8] P. Briand, B. Delyon, Y. Hu, E. Pardoux, and L. Stoica. Lp solutions of backward stochastic differential equations. *Stochastic Processes and their Applications*, 108(1):109 – 129, 2003.
- [9] D. J. Daley. Quasi-stationary behaviour of a left-continuous random walk. *Annals of Mathematical Statistics*, 40:532–539, 1969.
- [10] R. W. R. Darling and E. Pardoux. Backwards sde with random terminal time and applications to semilinear elliptic pde. *The Annals of Probability*, 25(3):1135–1159, 1997.
- [11] F. Delarue. Estimates of the solutions of a system of quasi-linear PDEs. A probabilistic scheme. In *Séminaire de Probabilités XXXVII*, pages 290–332. Springer, 2003.
- [12] O. Faure. *Simulation du mouvement brownien et des diffusions*. PhD thesis, Ecole Nationale des Ponts et Chaussées, 1992.
- [13] M. Freidlin. *Functional integration and partial differential equations*, volume 109 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1985.
- [14] A. Friedman. *Partial differential equations of parabolic type*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964.



- [15] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Springer, 2015.
- [16] E. Gobet. Revisiting the Greeks for European and American options. In *Proceedings of the International Symposium on Stochastic Processes and Mathematical Finance, Ritsumeikan University, Kusatsu, Japan*, pages 53–71, 2004.
- [17] E. Gobet, J.-P. Lemor, and X. Warin. A regression-based Monte Carlo method to solve backward stochastic differential equations. *The Annals of Applied Probability*, 15(3):2172–2202, 2005.
- [18] P. Henry-Labordere, X. Tan, and N. Touzi. A numerical algorithm for a class of BSDEs via the branching process. *Stochastic Processes and their Applications*, 124(2):1112–1140, 2014.
- [19] P. Henry-Labordere, N. Oudjane, X. Tan, N. Touzi, and X. Warin. Branching diffusion representation of semilinear PDEs and Monte Carlo approximation. *ArXiv 1603.01727*, 2016.
- [20] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. I. *J. Math. Kyoto Univ.*, 8: 233–278, 1968. ISSN 0023-608X.
- [21] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. II. *J. Math. Kyoto Univ.*, 8: 365–410, 1968. ISSN 0023-608X.
- [22] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. III. *J. Math. Kyoto Univ.*, 9: 95–160, 1969. ISSN 0023-608X.
- [23] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva. *Linear and quasilinear equations of parabolic type*. Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., 1968.
- [24] A. Lejay. exitbm: a library for simulating Brownian motion's exit times and positions from simple domains. Technical Report RR-7523, Feb. 2011. URL [HTTPS://HAL.INRIA.FR/INRIA-00561409](https://hal.inria.fr/inria-00561409).
- [25] P. L. Lions. On the existence of positive solutions of semilinear elliptic equations. *SIAM Review*, 24(4):441–467, 1982.
- [26] H. P. McKean. Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov. *Communications on Pure and Applied Mathematics*, 28(3):323–331, 1975.
- [27] G. N. Milstein and M. V. Tretyakov. Simulation of a space-time bounded diffusion. *Annals of Applied Probability*, pages 732–779, 1999.
- [28] É. Pardoux. Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order. In *Stochastic analysis and related topics, VI (Geilo, 1996)*, volume 42 of *Progr. Probab.*, pages 79–127. Birkhäuser Boston, Boston, MA, 1998.
- [29] É. Pardoux. *BSDEs, weak convergence and homogenization of semilinear PDEs*, pages 503–549. Springer Netherlands, 1999.
- [30] É. Pardoux and S. G. Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.
- [31] R. G. Pinsky. *Positive harmonic functions and diffusion*, volume 45 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.
- [32] A. Rasulov, G. Raimova, and M. Mascagni. Monte carlo solution of cauchy problem for a nonlinear parabolic equation. *Mathematics and Computers in Simulation*, 80(6):1118–1123, 2010.
- [33] B. A. Sevast'janov. The extinction conditions for branching processes with diffusion. *Akademija Nauk SSSR. Teorija Veroyatnostej i ee Primenenija*, 6:276–286, 1961.
- [34] A. V. Skorohod. Branching diffusion processes. *Akademija Nauk SSSR. Teorija Veroyatnostej i ee Primenenija*, 9:492–497, 1964.
- [35] A. Thalmaier. On the differentiation of heat semigroups and Poisson integrals. *Stochastics and Stochastics Reports*, 61(3-4):297–321, 1997.
- [36] S. Watanabe. On the branching process for Brownian particles with an absorbing boundary. *Journal of Mathematics of Kyoto University*, 4:385–398, 1965.
- [37] J. Zhang. A numerical scheme for BSDEs. *The Annals of Applied Probability*, 14(1):459–488, 2004.