Universität Trier

Quantitative Finance and Risk Analysis

Working Paper Series Nº 20 - 4:

Branching Diffusions with Jumps and Valuation with Systemic Counterparties

Christoph Belak

Daniel Hoffmann

Frank T. Seifried

Branching Diffusions with Jumps and Valuation with Systemic Counterparties

Christoph Belak[*] Daniel Hoffmann[†] Frank T. Seifried[‡]

January 14, 2020

We extend the branching diffusion Monte Carlo method of Henry-Labordère e.a. [11] to the case of parabolic PDEs with mixed local-nonlocal analytic nonlinearities. We investigate branching diffusion representations of classical solutions, and we provide sufficient conditions under which the branching diffusion representation solves the PDE in the viscosity sense. Our theoretical setup directly leads to a Monte Carlo algorithm, whose applicability is showcased in a stylized high-dimensional example. As our main application, we demonstrate how the methodology can be used to value financial positions with defaultable, systemically important counterparties.

Mathematics Subject Classification (2010): 60J85; 65C05; 91G80

Keywords: Branching Diffusion; Mixed Local-Nonlocal PDE; Nonlinear Jumps; Monte Carlo Simulation; Credit Valuation Adjustment

1 Introduction

The objective of this article is to derive probabilistic representations of solutions of a certain class of nonlinear parabolic partial differential equations with nonlocal terms in the nonlinearity. The representation is based on a branching diffusion mechanism with jumps at branching times and makes it possible to compute solutions by direct (non-nested) Monte Carlo simulation, leading to a numerical algorithm that does not suffer from the curse of dimensionality.

The class of partial differential equations under consideration in this article takes the form

$$\partial_t u(t,x) + \mathcal{A}[u](t,x) + \int_\Xi f\big(t,x,\xi,\mathcal{J}[u](t,x,\xi)\big)\gamma(\mathrm{d}\xi) = 0, \qquad \text{(PDE)}$$

[*]Technische Universität Berlin, Faculty II — Mathematics and Natural Sciences, Institute of Mathematics, Straße des 17. Juni 136, 10623 Berlin, e-mail: belak@math.tu-berlin.de

[†]University of Trier, Department IV — Mathematics, Universitätsring 19, 54296 Trier, e-mail: d.hoffmann@uni-trier.de

[‡]University of Trier, Department IV — Mathematics, Universitätsring 19, 54296 Trier, e-mail: seifried@uni-trier.de

where $\mathcal{A}$ denotes the infinitesimal generator of an Itō diffusion, i.e. a (possibly degenerate) linear partial differential operator of second order; $\mathcal{J}$ is a nonlocal operator; and the nonlinearity f is analytic in the jump terms.

In recent years, starting with [10], there has been significant progress in the realm of probabilistic representations of partial differential equations with analytic nonlinearities acting on zeroth- and first-order derivatives. We refer in particular to [12] for the zeroth-order case and [11] for the first-order case. We also refer to [4] for an extension of the branching diffusion approach to the case of locally analytic nonlinearities, [3] for the case of Lipschitz nonlinearities, [13] for higher-order partial differential equations, and [1] for an extension to elliptic equations. The main contribution of this article is to extend the branching diffusion approach to the case of nonlocal terms inside the nonlinearity. This extension is achieved by introducing jump marks in the branching diffusion underlying the probabilistic representation result: We consider a branching diffusion similar to the one introduced in [12] with the additional feature that, at each branching time, a subset of offspring particles jumps away from their parent's position. We refer to the resulting object as a *branching diffusion with jumps*.[1]

Our main motivation behind the derivation of a probabilistic representation is to open up the possibility to apply numerical algorithms for nonlocal nonlinear PDEs in high dimensions via Monte Carlo simulation. The effectiveness and efficiency of such algorithms is showcased in an example of a nonlocal nonlinear PDE, which we solve in dimensions up to 100. In addition, we show how our Monte Carlo methodology can be used in the pricing of (equivalently, the computation of credit valuation adjustments for) financial positions where the counterparty is a systemically important financial institution whose default causes a devaluation of the underlying. This jump-at-default model represents a particularly realistic setup for wrong-way risk; see, e.g., [5], [17] and [19]. Our Monte Carlo approach complements existing methods by making it possible to price systemic defaultable positions for settings where the underlying dynamics are not tractable by PDE methods as in, e.g., [6] or [14].

The remainder of the article is organized as follows: Section 2 provides the stochastic construction of the branching diffusion with jumps and derives a probabilistic representation of classical solutions of (PDE) in terms of it. In Section 3 we conversely establish sufficient conditions for the branching diffusion representation to yield a viscosity solution of (PDE). Section 4 illustrates how the branching diffusion representation allows for efficient simulation of solutions via the Monte Carlo method in a stylized high-dimensional example. Finally, in Section 5 we showcase our methodology in the context of pricing with a systemically important, defaultable counterparty.

[1] This is not to be confused with the straightforward notion of branching jump-diffusions, where jump terms appear in the infinitesimal generator $\mathcal{A}$; see also the discussion below Theorem 2.4.

2 Branching Diffusion Representations for Nonlocal PDEs

Throughout this paper, we fix a time horizon $T > 0$, a non-empty set $\mathcal{I} \subseteq \mathbb{N}_0^m$ of multi-indices, and a jump distribution γ on an abstract measurable space $(\Xi, \mathfrak{B})$. The goal of this section is to provide a stochastic representation of a classical solution $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ of a nonlocal partial differential equation of the form

$$\partial_t u(t,x) + \mathcal{A}[u](t,x) + \int_\Xi f\big(t,x,\xi,\mathcal{J}[u](t,x,\xi)\big)\gamma(\mathrm{d}\xi) = 0, \tag{PDE}$$

$$u(T,x) = g(x), \tag{BC}$$

where

$$\mathcal{A}[u](t,x) \triangleq \mu(t,x)^\intercal \, \mathrm{D}_x \, u(t,x) + \frac{1}{2} \operatorname{tr}\Big[\sigma(t,x)\sigma(t,x)^\intercal \, \mathrm{D}_x^2 \, u(t,x)\Big]$$

is the infinitesimal generator of a diffusion process; the nonlinearity

$$f : [0,T] \times \mathbb{R}^d \times \Xi \times \mathbb{R}^m \to \mathbb{R}, \quad f(t,x,\xi,y) \triangleq \sum_{i\in\mathcal{I}} c_i(t,x,\xi) y^i$$

is (multivariate) analytic[2] in y with measurable coefficients $c_i : [0,T] \times \mathbb{R}^d \times \Xi \to \mathbb{R}$ for $i \in \mathcal{I}$; and the jump operator $\mathcal{J}$ is given by

$$\mathcal{J}_\ell[u](t,x,\xi) \triangleq u\big(t, \Gamma_\ell(t,x,\xi)\big) \quad \text{for} \quad \ell \in [1:m]$$

where the jump maps $\Gamma_\ell : [0,T] \times \mathbb{R}^d \times \Xi \to \mathbb{R}^d$, $\ell \in [1:m]$, are measurable.

2.1 Preliminaries

For $n \in \mathbb{N}$, we denote by $\mathbf{N}_n \triangleq \bigcup_{\nu=1}^n \mathbb{N}^\nu$ the set of all $\mathbb{N}$-words with length at most n and by $\mathbf{N} \triangleq \bigcup_{n\in\mathbb{N}} \mathbf{N}_n$ the set of finite $\mathbb{N}$-words. We work on a probability space $(\Omega, \mathfrak{A}, \mathbb{P})$ that supports the following random variables, all of which are taken to be mutually independent:

- A family $\{W^{(k)}\}_{k\in\mathbf{N}}$ of independent $\mathbb{R}^d$-valued Brownian motions that serve as driving noise for the emerging diffusion processes with infinitesimal generator $\mathcal{A}$.
- A family $\{\Delta^{(k)}\}_{k\in\mathbf{N}}$ of i.i.d. $\mathbb{R}^d$-valued random variables with distribution γ.
- A family $\{\tau^{(k)}\}_{k\in\mathbf{N}}$ of i.i.d. $\mathbb{R}_+$-valued random variables serving as lifetimes of the particles underlying the branching mechanism. We assume that the distribution of $\tau^{(k)}$, $k \in \mathbf{N}$, admits a continuous density $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$F(T) \triangleq \int_T^\infty \rho(s)\mathrm{d}s > 0 \qquad \text{for all } T > 0. \tag{2.1}$$

- A family $\{I^{(k)}\}_{k\in\mathbf{N}}$ of i.i.d. $\mathcal{I}$-valued random variables modeling the number of offspring

[2] We use standard multi-index notation and write $y^i \triangleq \prod_{\ell=1}^m y_\ell^{i_\ell}$ for $y \in \mathbb{R}^m$ and $|i| \triangleq \sum_{\ell=1}^m i_\ell$, $i \in \mathbb{N}_0^m$. The case of a multivariate polynomial obtains if $\mathcal{I}$ is finite.

of each particle as well as marks for their initial positions. We assume that

$$p_i \triangleq \mathbb{P}[I^{(k)} = i] > 0 \qquad \text{for all } i \in \mathcal{I},\ k \in \mathbf{N} \qquad \text{and} \qquad \sum_{i \in \mathcal{I}} |i| p_i < \infty.$$

In the following, we describe the branching mechanism and the spatial dynamics separately to finally obtain the branching diffusion representation of (PDE).

2.2 Branching Mechanism

We fix an initial time $t \in [0,T]$. The branching mechanism is defined by recursion on successive generations, for each $\omega \in \Omega$.

Given a particle of generation $n \in \mathbb{N}$ labeled $k = (k_1, \dots, k_n) \in \mathbb{N}^n$, we denote its parent particle by $k- \triangleq (k_1, \dots, k_{n-1}) \in \mathbb{N}^{n-1}$. The branching time of particle k is given by $T_t^{(k)} \triangleq (T_t^{(k-)} + \tau^{(k)}) \wedge T$; on the event $\{T_t^{(k)} < T\}$ the particle k is removed at time $T_t^{(k)}$ and branches into $|I^{(k)}|$ descendants, which are labeled by $(k_1, \dots, k_n, k_{n+1}) \in \mathbb{N}^{n+1}$ for $k_{n+1} \in [1 : |I^{(k)}|]$. We attach the jump mark $J^{(k)} \triangleq 1$ to the first $I_1^{(k)}$ particles, the mark $J^{(k)} \triangleq 2$ to the following $I_2^{(k)}$ particles, etc., so each offspring particle k carries a mark $J^{(k)} \in [1 : m]$. This iteration is well-defined and uniquely determines the branching dynamics if we assume that the mechanism starts with a single particle with label (1) of generation 1 at time t and

$$(1)- \triangleq () \triangleq \emptyset \qquad \text{and} \qquad T_t^{\emptyset} \triangleq t.$$

In the following, we only refer to $k \in \mathbf{N}$ as a *particle* if either $k = (1)$ or if $k-$ is a particle and $k_n \in [1 : |I^{(k-)}|]$. Figure 1 below visualizes this branching mechanism.

We denote the random set of all particles of generation $n \in \mathbb{N}$ alive at time $s \in [t,T]$ by

$$\mathcal{K}_t^n(s) \triangleq \begin{cases} \{k \in \mathbb{N}^n : k \text{ is a particle and } T_t^{(k-)} \le s < T_t^{(k)}\} & \text{if } s \in [t,T), \\ \{k \in \mathbb{N}^n : k \text{ is a particle and } T_t^{(k)} = T\} & \text{if } s = T. \end{cases}$$

The set of all particles of generation n alive before or at time s is given by

$$\overline{\mathcal{K}}_t^n(s) \triangleq \bigcup_{r \in [t,s]} \mathcal{K}_t^n(r).$$

Finally, the set of all particles alive at time s and the set of all particles alive before or at time s are defined as

$$\mathcal{K}_t(s) \triangleq \bigcup_{n \in \mathbb{N}} \mathcal{K}_t^n(s) \qquad \text{and} \qquad \overline{\mathcal{K}}_t(s) \triangleq \bigcup_{n \in \mathbb{N}} \overline{\mathcal{K}}_t^n(s), \qquad \text{respectively.}$$

For ease of notation, we subsequently write

$$\mathcal{K}_t^n \triangleq \mathcal{K}_t^n(T), \qquad \mathcal{K}_t \triangleq \mathcal{K}_t(T), \qquad \overline{\mathcal{K}}_t^n \triangleq \overline{\mathcal{K}}_t^n(T), \qquad \overline{\mathcal{K}}_t \triangleq \overline{\mathcal{K}}_t(T).$$

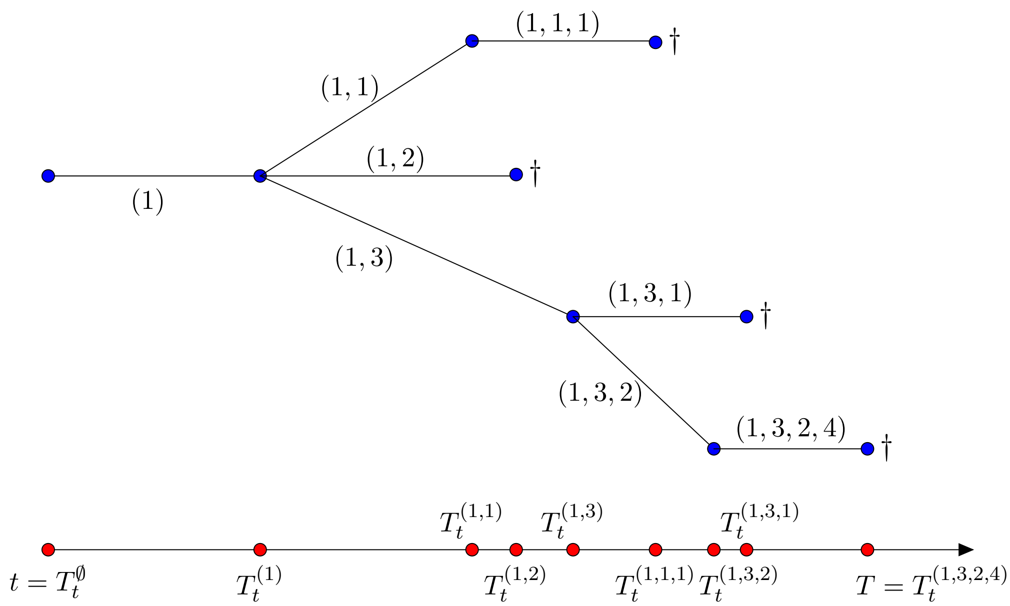

Figure 1 Illustration of the branching mechanism (without jump marks).

As in Proposition 2.4 in [11], the total number of particles is almost surely finite, i.e.

$$\#\overline{\mathcal{K}}_t < \infty;$$

see also Theorem IV.1.1 in [2] and Chapter VI §§12f. in [9].

2.3 Branching Diffusion Dynamics

The next step is to specify the dynamics of the individual particles. We first impose some standard regularity assumptions on the coefficient functions in the infinitesimal generator $\mathcal{A}$.

Assumption 1. *The functions*

$$\mu : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d \qquad \textit{and} \qquad \sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times d}$$

are measurable and satisfy the following Lipschitz and linear growth conditions: There exists a constant $L > 0$ such that

$$\begin{aligned} |\mu(t, x) - \mu(t, y)| + |\sigma(t, x) - \sigma(t, y)| &\leq L|x - y|, && t \in [0, T],\ x, y \in \mathbb{R}^d, \\ |\mu(t, x)|^2 + |\sigma(t, x)|^2 &\leq L^2\big(1 + |x|^2\big), && t \in [0, T],\ x \in \mathbb{R}^d. \end{aligned} \qquad \diamond$$

Under this assumption, classical results such as Theorem 3.21 in [18] imply that the stochastic differential equation

$$\begin{aligned} \overline{X}_s^{t,x} &= x, && s \in [0,t], \\ \mathrm{d}\overline{X}_s^{t,x} &= \mu\big(s, \overline{X}_s^{t,x}\big)\mathrm{d}s + \sigma\big(s, \overline{X}_s^{t,x}\big)\mathrm{d}\overline{W}_s, && s \in [t,T], \end{aligned} \tag{2.2}$$

admits a pathwise unique strong solution for each starting configuration $(t,x) \in [0,T] \times \mathbb{R}^d$. Here, $\overline{W}$ is an $\mathbb{R}^d$-valued Brownian motion on $(\Omega, \mathfrak{A}, \mathbb{P})$ that is independent of all other random variables that occurred so far. The natural filtration of $\overline{X}^{t,x}$ augmented by all $\mathbb{P}$-nullsets is denoted by $\overline{\mathfrak{F}}^{t,x} = \{\overline{\mathfrak{F}}_s^{t,x}\}_{s\in[0,T]}$. An application of the results in Chapter 3.7 of [18] yields the following:

Lemma 2.1 (Properties of the Diffusion)**.** *The random field* $\{\overline{X}_s^{t,x}\}_{s,t\in[0,T],x\in\mathbb{R}^d}$ *defined via* (2.2) *can be chosen such that it satisfies the following conditions:*

(i) Continuity with respect to initial data: *The map*

$$\overline{X} : [0,T] \times \mathbb{R}^d \times [0,T] \to \mathbb{R}^d, \qquad (t,x,s) \mapsto \overline{X}_s^{t,x},$$

is almost surely continuous.

(ii) Flow property: *For all* $(t,x) \in [0,T] \times \mathbb{R}^d$ *and any* $[t,T]$*-valued* $\overline{\mathfrak{F}}^{t,x}$*-stopping time* τ*,*

$$\overline{X}_{\tau+s}^{\tau,\overline{X}_\tau^{t,x}} \mathbb{1}_{\{\tau+s\le T\}} = \overline{X}_{\tau+s}^{t,x} \mathbb{1}_{\{\tau+s\le T\}}, \qquad s \in [0,\infty).$$

(iii) Strong Markov property: *For all* $(t,x) \in [0,T] \times \mathbb{R}^d$ *and* $s \in [0,\infty)$*, for any* $[t,T]$*-valued* $\overline{\mathfrak{F}}^{t,x}$*-stopping time* τ*, and for every bounded measurable function* $h : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$*, we have*

$$\mathbb{E}\Big[h\big(\tau+s, \overline{X}_{\tau+s}^{t,x}\big)\mathbb{1}_{\{\tau+s\le T\}}\Big|\overline{\mathfrak{F}}_\tau^{t,x}\Big] = \mathbb{E}\Big[h\big(\tau+s, \overline{X}_{\tau+s}^{t,x}\big)\mathbb{1}_{\{\tau+s\le T\}}\Big|\big(\tau, \overline{X}_\tau^{t,x}\big)\Big]. \qquad \diamond$$

The branching diffusion is constructed by attaching to each particle in the branching mechanism a diffusion with the same dynamics as $\overline{X}$, but with a different driving noise and a suitable initial condition. To make this precise, we fix $x \in \mathbb{R}^d$ and define for each $k \in \overline{\mathcal{K}}_t$ with $k = (1, k_2, \ldots, k_n) \in \mathbb{N}^n$ an associated diffusion $X^{(k)} = X^{k,t,x} = \{X_s^{k,t,x}\}_{s\in[T_t^{(k-)}, T_t^{(k)}]}$ as the unique strong solution of

$$\begin{aligned} X^{(k)}_{T_t^{(k-)}} &= \Gamma_{J^{(k)}}\big(T_t^{(k-)}, X^{(k-)}_{T_t^{(k-)}}, \Delta^{(k-)}\big) \\ \mathrm{d}X_s^{(k)} &= \mu\big(s, X_s^{(k)}\big)\mathrm{d}s + \sigma\big(s, X_s^{(k)}\big)\mathrm{d}W_s^{(k)}, && s \in \big[T_t^{(k-)}, T_t^{(k)}\big]. \end{aligned}$$

Note that this iteration is well-defined starting from $X_t^{(1)} = x$. It follows that $X^{(k)}$ has the same dynamics as $\overline{X}$, but with the different, independent, driving noise $W^{(k)}$. The lifetime of $X^{(k)}$ coincides with the lifetime of the particle k. Moreover, the initial value of $X^{(k)}$ is the terminal value of the diffusion $X^{(k-)}$ associated with its parent particle $k-$, plus an additional jump whose size is given by the jump map $\Gamma_{J^{(k)}}$ corresponding to its mark $J^{(k)}$ and the jump parameter $\Delta^{(k-)}$. For later reference, we encode all information available up to generation $n \in \mathbb{N}$ by setting

$$\mathfrak{F}^n \triangleq \sigma\big(W^{(k)}, \tau^{(k)}, \Delta^{(k)}, I^{(k)} : k \in \mathbf{N}_n\big).$$

For notational convenience, we furthermore write $\mathfrak{F}^0 \triangleq \{\emptyset, \Omega\}$ for the trivial σ-algebra. Finally, for $n \in \mathbb{N}_0$, we enlarge these σ-algebras by the branching time information of one future generation, i.e. we set

$$\mathfrak{G}^n \triangleq \mathfrak{F}^n \vee \sigma\big(\tau^{(k)} : k \in \mathbf{N}_{n+1}\big).$$

For $n \in \mathbb{N}$ and $k \in \overline{\mathcal{K}}_t^n$, we observe that

$$\text{conditional on } \mathfrak{G}^{n-1}, \text{ the laws of } X^{(k)} \text{ and } \overline{X}^{T_t^{(k-)}, X^{(k)}_{T_t^{(k-)}}} \text{ on } \big[T_t^{(k-)}, T_t^{(k)}\big] \text{ are identical.}$$

We stress that $X^{(k)}$ and $\overline{X}^{T_t^{(k-)}, X^{(k)}_{T_t^{(k-)}}}$ do not coincide pathwise since the dynamics of $X^{(k)}$ are driven by $W^{(k)}$, while those of $\overline{X}^{T_t^{(k-)}, X^{(k)}_{T_t^{(k-)}}}$ are driven by $\overline{W}$. Replacing the driving Brownian motion with a new, independent one for each offspring particle – without changing its distribution – will be the key step in the branching representation below.

2.4 Branching Diffusion Representation

We now address the branching diffusion representation of classical solutions of (PDE). To begin with, we specify suitable boundedness assumptions on the coefficient functions c_i and the terminal condition g.

Assumption 2. *The functions $g : \mathbb{R}^d \to \mathbb{R}$ and $c_i : [0,T] \times \mathbb{R}^d \times \Xi \to \mathbb{R}$, $i \in \mathcal{I}$, in* (PDE) *and* (BC) *are bounded and measurable.*

In order for the possibly infinite series in the nonlinearity of (PDE) to be defined unambiguously, we subsequently agree on the following definition of classical solutions.

Definition 2.2 (Classical Solution)**.** *Under Assumption 2, a continuous function*

$$u : [0,T] \times \mathbb{R}^d \to \mathbb{R}, \qquad (t,x) \mapsto u(t,x)$$

is said to be a classical solution *of* (PDE) *with terminal condition* (BC) *if*

(i) $u \in \mathcal{C}^{1,2}([0,T) \times \mathbb{R}^d)$;

(ii) for each $(t,x) \in [0,T) \times \mathbb{R}^d$*, it holds that*

$$\sum_{i \in \mathcal{I}} \int_{\Xi} \big| c_i(t,x,\xi) \mathcal{J}[u](t,x,\xi)^i \big| \gamma(\mathrm{d}\xi) < +\infty;$$

(iii) and u satisfies (PDE) *for each* $(t,x) \in [0,T) \times \mathbb{R}^d$ *and* (BC) *for each* $x \in \mathbb{R}^d$. ◇

We are now in a position to establish the main result of this section, which allows us to represent a classical solution of (PDE) by means of a functional of the branching diffusion. The key idea is to introduce *randomization* across subsequent generations and subsequently exploit the conditional independence structure.

Remark 2.3 (Randomization)**.** We fix a particle $k \in \overline{\mathcal{K}}_t^n$ of generation $n \in \mathbb{N}$. By a slight abuse of notation, we drop any indices pertaining to the initial position (t, x) and write

$$\overline{X} \triangleq \overline{X}^{T^{(k-)}, X^{(k)}_{T^{(k-)}}} \qquad \text{and} \qquad (\Delta, \tau, I) \triangleq (\Delta^{(k)}, \tau^{(k)}, I^{(k)}).$$

Under suitable regularity and integrability assumptions, any classical solution u of (PDE) admits a Feynman-Kac representation[3] of the form

$$\begin{aligned} u\big(T^{(k-)}, X^{(k)}_{T^{(k-)}}\big) &= \mathbb{E}\Big[g\big(\overline{X}_T\big) + \int_{T^{(k-)}}^T f\big(r, \overline{X}_r, \Delta, \mathcal{J}[u]\big(r, \overline{X}_r, \Delta\big)\big)\mathrm{d}r\Big|\mathfrak{F}^{n-1}\Big] \\ &= \mathbb{E}\Big[g\big(\overline{X}_T\big) + \int_{T^{(k-)}}^T \sum_{i \in \mathcal{I}} c_i\big(r, \overline{X}_r, \Delta\big)\mathcal{J}[u]\big(r, \overline{X}_r, \Delta\big)^i \mathrm{d}r\Big|\mathfrak{F}^{n-1}\Big]. \end{aligned} \tag{2.3}$$

The key idea underlying the branching diffusion representation is to represent the right hand side recursively in terms of the branching diffusion $X^{(k)}$, thus eliminating the integral, sum and nonlinearity within the conditional expectation. More precisely, we claim that

$$\begin{aligned} &u\big(T^{(k-)}, X^{(k)}_{T^{(k-)}}\big) \\ &= \mathbb{E}\Big[\mathbb{1}_{\{T^{(k)}=T\}}\frac{g\big(X^{(k)}_T\big)}{F(T - T^{(k-)})} + \mathbb{1}_{\{T^{(k)}<T\}}\frac{c_I\big(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta\big)}{\rho(T^{(k)} - T^{(k-)})p_I}\mathcal{J}[u]\big(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta\big)^I\Big|\mathfrak{F}^{n-1}\Big]. \end{aligned} \tag{2.4}$$

Let us start by considering the first summand in (2.4). Since $T^{(k)} = (T^{(k-)} + \tau) \wedge T$ and $\overline{X}$ and $X^{(k)}$ have the same conditional distribution given $\mathfrak{G}^{n-1}$, we have

$$\mathbb{E}\Big[\mathbb{1}_{\{T^{(k)}=T\}}\frac{g\big(X^{(k)}_T\big)}{F(T - T^{(k-)})}\Big|\mathfrak{F}^{n-1}\Big] = \mathbb{E}\Big[\mathbb{1}_{\{\tau \geq T - T^{(k-)}\}}\frac{g\big(\overline{X}_T\big)}{F(T - T^{(k-)})}\Big|\mathfrak{F}^{n-1}\Big].$$

But then, since τ is independent of $\mathfrak{F}^{n-1}$ and $\overline{X}$, we can simply integrate with respect to the density of τ and use the definition of F in (2.1) to obtain

$$\mathbb{E}\Big[\mathbb{1}_{\{T^{(k)}=T\}}\frac{g\big(X^{(k)}_T\big)}{F(T - T^{(k-)})}\Big|\mathfrak{F}^{n-1}\Big] = \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T - T^{(k-)})}F(T - T^{(k-)})\Big|\mathfrak{F}^{n-1}\Big] = \mathbb{E}\big[g\big(\overline{X}_T\big)\big|\mathfrak{F}^{n-1}\big]$$

as in (2.3). The second term in (2.4) is slightly more involved, but can be handled similarly:

[3] As demonstrated in its proof, this holds in particular under the conditions of Theorem 2.4.

First, we use the conditional identity in law given $\mathfrak{G}^{n-1}$ of $\overline{X}$ and $X^{(k)}$ to obtain

$$\mathbb{E}\Big[\mathbb{1}_{\{T^{(k)}<T\}}\frac{c_I\big(T^{(k)},X^{(k)}_{T^{(k)}},\Delta\big)}{\rho(T^{(k)}-T^{(k-)})p_I}\mathcal{J}[u]\big(T^{(k)},X^{(k)}_{T^{(k)}},\Delta\big)^I\Big|\mathfrak{F}^{n-1}\Big]$$
$$=\mathbb{E}\Big[\mathbb{1}_{\{T^{(k)}<T\}}\frac{c_I\big(T^{(k)},\overline{X}_{T^{(k)}},\Delta\big)}{\rho(T^{(k)}-T^{(k-)})p_I}\mathcal{J}[u]\big(T^{(k)},\overline{X}_{T^{(k)}},\Delta\big)^I\Big|\mathfrak{F}^{n-1}\Big].$$

Next, independence of I and τ from all other objects involved allows us to integrate with respect to the associated probability mass function and density, and we obtain

$$\mathbb{E}\Big[\mathbb{1}_{\{T^{(k)}<T\}}\frac{c_I\big(T^{(k)},\overline{X}_{T^{(k)}},\Delta\big)}{\rho(T^{(k)}-T^{(k-)})p_I}\mathcal{J}[u]\big(T^{(k)},\overline{X}_{T^{(k)}},\Delta\big)^I\Big|\mathfrak{F}^{n-1}\Big]$$
$$=\mathbb{E}\Big[\int_{T^{(k-)}}^{T}\sum_{i\in\mathcal{I}}c_i\big(r,\overline{X}_r,\Delta\big)\mathcal{J}[u]\big(r,\overline{X}_r,\Delta\big)^i\mathrm{d}r\Big|\mathfrak{F}^{n-1}\Big].$$

Combining the above equations, we have therefore argued that

$$\mathbb{E}\Big[\mathbb{1}_{\{T^{(k)}<T\}}\frac{c_I\big(T^{(k)},X^{(k)}_{T^{(k)}},\Delta\big)}{\rho(T^{(k)}-T^{(k-)})p_I}\mathcal{J}[u]\big(T^{(k)},X^{(k)}_{T^{(k)}},\Delta\big)^I\Big|\mathfrak{F}^{n-1}\Big]$$
$$=\mathbb{E}\Big[\int_{T^{(k-)}}^{T}\sum_{i\in\mathcal{I}}c_i\big(r,\overline{X}_r,\Delta\big)\mathcal{J}[u]\big(r,\overline{X}_r,\Delta\big)^i\mathrm{d}r\Big|\mathfrak{F}^{n-1}\Big]$$

and thus (2.4) holds. The key advantage of (2.4) is that the nonlinearity is represented via

$$\mathcal{J}[u]\big(T^{(k)},X^{(k)}_{T^{(k)}},\Delta\big)^I=\prod_{\ell=1}^{m}u\big(T^{(k)},\Gamma_\ell\big(T^{(k)},X^{(k)}_{T^{(k)}},\Delta\big)\big)^{I_\ell},$$

where it is possible to iterate over all generations of particles, using conditional independence across generations: Indeed, the terms $\Gamma_\ell(T^{(k)},X^{(k)}_{T^{(k)}},\Delta^{(k)})$ appearing as arguments in the function u correspond to the initial positions of the particles of generation $n+1$. This leads to the branching diffusion representation made rigorous in Theorem 2.4 below. $\diamond$

We next state the first main result of this paper: A branching diffusion representation of classical solutions of (PDE).

Theorem 2.4 (Branching Representation of Classical Solutions)**.** *Suppose Assumptions 1 and 2 hold, let* $u:[0,T]\times\mathbb{R}^d\to\mathbb{R}$ *be a classical solution of* (PDE) *satisfying* (BC) *and fix* $(t,x)\in[0,T]\times\mathbb{R}^d$*. For each* $n\in\mathbb{N}_0$*, iteratively define the random variables*[4] $\mathcal{G}_0^{t,x}\triangleq\mathcal{C}_0^{t,x}\triangleq 1$,

$$\mathcal{G}_n^{t,x}\triangleq\mathcal{G}_{n-1}^{t,x}\prod_{k\in\mathcal{K}^n}\frac{g(X_T^{(k)})}{F(T-T^{(k-)})}\qquad\textit{and}\qquad\mathcal{C}_n^{t,x}\triangleq\mathcal{C}_{n-1}^{t,x}\prod_{k\in\overline{\mathcal{K}}^n\setminus\mathcal{K}^n}\frac{c_{I^{(k)}}(T^{(k)},X^{(k)}_{T^{(k)}},\Delta^{(k)})}{\rho(T^{(k)}-T^{(k-)})p_{I^{(k)}}}$$

[4]For notational convenience, we suppress the dependence on (t,x) on the right hand sides of these definitions but highlight here that $X^{(k)}=X^{k,t,x}$ and $T^{(k)}=T_t^{(k)}$ as well as $\overline{\mathcal{K}}^n=\overline{\mathcal{K}}_t^n$ and $\mathcal{K}^n=\mathcal{K}_t^n$.

and

$$\mathcal{R}_n^{t,x} \triangleq \prod_{k\in\mathcal{K}^{n+1}} \frac{g(X_T^{(k)})}{F(T-T^{(k-)})}$$
$$\times \prod_{k\in\overline{\mathcal{K}}^{n+1}\setminus\mathcal{K}^{n+1}} \frac{c_{I^{(k)}}(T^{(k)}, X_{T^{(k)}}^{(k)}, \Delta^{(k)})\mathcal{J}[u]\big(T^{(k)}, X_{T^{(k)}}^{(k)}, \Delta^{(k)}\big)^{I^{(k)}}}{\rho(T^{(k)}-T^{(k-)})p_{I^{(k)}}},$$

and set

$$\Psi_n^{t,x} \triangleq \mathcal{G}_n^{t,x}\mathcal{C}_n^{t,x}\mathcal{R}_n^{t,x} \qquad \textit{and} \qquad \Psi^{t,x} \triangleq \lim_{n\to\infty} \Psi_n^{t,x}. \tag{2.5}$$

Suppose that

(i) *the family* $\{\Psi_n^{t,x}\}_{n\in\mathbb{N}_0}$ *is uniformly integrable;*

(ii) *for every* $(s,y)\in[t,T]\times\mathbb{R}^d$*, it holds that*

$$\sum_{i\in\mathcal{I}} \mathbb{E}\Big[\int_s^T \int_\Xi \big|c_i(r,\overline{X}_r^{s,y},\xi)\mathcal{J}[u]\big(r,\overline{X}_r^{s,y},\xi\big)^i\big|\gamma(\mathrm{d}\xi)\mathrm{d}r\Big] < +\infty;$$

(iii) *for any* $(s,y)\in[t,T]\times\mathbb{R}^d$*, the local martingale*

$$M^{s,y} \triangleq \int_s^\cdot \mathrm{D}_x\, u(r,\overline{X}_r^{s,y})\sigma(r,\overline{X}_r^{s,y})\mathrm{d}\overline{W}_r$$

is a martingale.

Then $\Psi^{t,x}$ *is integrable and* u *admits the branching diffusion representation*

$$u(t,x) = \mathbb{E}[\Psi^{t,x}]. \qquad \diamond$$

Before we turn to the proof, note that unwinding the definitions we have $\Psi^{t,x} = \mathcal{G}^{t,x}\mathcal{C}^{t,x}$ where

$$\mathcal{G}^{t,x} = \prod_{k\in\mathcal{K}} \frac{g(X_T^{(k)})}{F(T-T^{(k-)})} \qquad \text{and} \qquad \mathcal{C}^{t,x} = \prod_{k\in\overline{\mathcal{K}}\setminus\mathcal{K}} \frac{c_{I^{(k)}}(T^{(k)}, X_{T^{(k)}}^{(k)}, \Delta^{(k)})}{\rho(T^{(k)}-T^{(k-)})p_{I^{(k)}}}$$

so the branching diffusion representation in Theorem 2.4 can be written more explicitly as

$$u(t,x) = \mathbb{E}\Big[\prod_{k\in\mathcal{K}} \frac{g(X_T^{(k)})}{F(T-T^{(k-)})} \prod_{k\in\overline{\mathcal{K}}\setminus\mathcal{K}} \frac{c_{I^{(k)}}(T^{(k)}, X_{T^{(k)}}^{(k)}, \Delta^{(k)})}{\rho(T^{(k)}-T^{(k-)})p_{I^{(k)}}}\Big].$$

Proof of Theorem 2.4. Since $\lim_{n\to\infty}\Psi_n^{t,x} = \Psi^{t,x}$ and $\{\Psi_n^{t,x}\}_{n\in\mathbb{N}_0}$ is uniformly integrable, it follows from Vitali's theorem that $\Psi^{t,x}$ is integrable and $\lim_{n\to\infty}\mathbb{E}[\Psi_n^{t,x}] = \mathbb{E}[\Psi^{t,x}]$. Hence, to prove the result, it suffices to show that

$$u(t,x) = \mathbb{E}[\Psi_n^{t,x}] \quad \text{for each } n\in\mathbb{N}_0. \tag{2.6}$$

Step 1. Using Itō's lemma and the fact that u solves (PDE) subject to (BC), we have

$$\begin{aligned} u(s,y) = {} & u(T,\overline{X}_T^{s,y}) - \int_s^T \Big[\partial_t u(r,\overline{X}_r^{s,y}) + \mathcal{A}[u](r,\overline{X}_r^{s,y})\Big]\mathrm{d}r - M_T^{s,y} \\ = {} & g(\overline{X}_T^{s,y}) + \int_s^T \int_\Xi f\big(r,\overline{X}_r^{s,y},\xi,\mathcal{J}[u]\big(r,\overline{X}_r^{s,y},\xi\big)\big)\gamma(\mathrm{d}\xi)\mathrm{d}r - M_T^{s,y} \end{aligned}$$

for any $(s,y) \in [t,T] \times \mathbb{R}^d$. Fix a particle $k \in \overline{\mathcal{K}}^n$ of generation $n \in \mathbb{N}$, write $\overline{X}^{(k)} \triangleq \overline{X}^{T^{(k-)},X^{(k)}_{T^{(k-)}}}$, and note that $(T^{(k-)}, X^{(k)}_{T^{(k-)}})$ is $\mathfrak{F}^{n-1}$-measurable. Choosing $(s,y) = (T^{(k-)}, X^{(k)}_{T^{(k-)}})$ in the Itō representation above, taking conditional expectations and using (ii) and (iii) we obtain

$$u(T^{(k-)}, X^{(k)}_{T^{(k-)}}) = \mathbb{E}\Big[g(\overline{X}_T^{(k)}) + \int_{T^{(k-)}}^T \int_\Xi f\big(r,\overline{X}_r^{(k)},\xi,\mathcal{J}[u]\big(r,\overline{X}_r^{(k)},\xi\big)\big)\gamma(\mathrm{d}\xi)\mathrm{d}r\Big|\mathfrak{F}^{n-1}\Big].$$

Since $\Delta^{(k)}$ has distribution γ, is independent of $\mathfrak{F}^{n-1}$ and $\overline{X}^{(k)}$ and $\Delta^{(k)}$ are independent, it follows that

$$u(T^{(k-)}, X^{(k)}_{T^{(k-)}}) = \mathbb{E}\Big[g(\overline{X}_T^{(k)}) + \int_{T^{(k-)}}^T f\big(r,\overline{X}_r^{(k)},\Delta^{(k)},\mathcal{J}[u]\big(r,\overline{X}_r^{(k)},\Delta^{(k)}\big)\big)\mathrm{d}r\Big|\mathfrak{F}^{n-1}\Big],$$

and by the argument in Remark 2.3 this can be further rewritten as

$$\begin{aligned} u(T^{(k-)}, X^{(k)}_{T^{(k-)}}) = \mathbb{E}\Big[& \mathbb{1}_{\{T^{(k)}=T\}} \frac{g\big(X_T^{(k)}\big)}{F(T-T^{(k-)})} \\ & + \mathbb{1}_{\{T^{(k)}<T\}} \frac{c_{I^{(k)}}\big(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)}\big)}{\rho(T^{(k)}-T^{(k-)})p_{I^{(k)}}} \mathcal{J}[u]\big(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)}\big)^{I^{(k)}} \Big|\mathfrak{F}^{n-1}\Big]. \end{aligned} \tag{2.7}$$

Step 2. We establish (2.6) by induction on n. For $n = 0$, let $k = (1)$ be the only particle of generation 1, recall that $\mathfrak{F}^0$ is trivial, and note that (2.7) rewrites as

$$\begin{aligned} u(t,x) &= \mathbb{E}\Big[\mathbb{1}_{\{T^{(k)}=T\}} \frac{g\big(X_T^{(k)}\big)}{F(T-T^{(k-)})} \\ &\qquad + \mathbb{1}_{\{T^{(k)}<T\}} \frac{c_{I^{(k)}}\big(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)}\big)\mathcal{J}[u]\big(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)}\big)^{I^{(k)}}}{\rho(T^{(k)}-T^{(k-)})p_{I^{(k)}}}\Big] \\ &= \mathbb{E}\big[\mathcal{R}_0^{t,x}\big] = \mathbb{E}\big[\Psi_0^{t,x}\big]. \end{aligned}$$

Now, let $n \in \mathbb{N}$ and suppose that the claim is true for $n-1$, i.e.

$$u(t,x) = \mathbb{E}\big[\Psi_{n-1}^{t,x}\big] = \mathbb{E}\big[\mathcal{G}_{n-1}^{t,x}\mathcal{C}_{n-1}^{t,x}\mathcal{R}_{n-1}^{t,x}\big]. \tag{2.8}$$

Let $k \in \overline{\mathcal{K}}^n$ be an arbitrary particle of generation n. On the event $\{k \in \overline{\mathcal{K}}^n \backslash \mathcal{K}^n\} = \{T^{(k)} < T\}$, the particle k branches into $|I^{(k)}|$ offspring particles (k, k_{n+1}), $k_{n+1} \in [1 : |I^{(k)}|]$, of which the first $I_1^{(k)}$ have mark 1, i.e. jump to $\Gamma_1\big(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)}\big)$, the next $I_2^{(k)}$ have mark 2, i.e. jump to $\Gamma_2\big(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)}\big)$, and so forth. Thus, on the event $\{k \in \overline{\mathcal{K}}^n \setminus \mathcal{K}^n\}$, we have

$$\begin{aligned}
&\mathcal{J}[u]\big(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)}\big)^{I^{(k)}} = \prod_{\ell=1}^{m} u\big(T^{(k)}, \Gamma_\ell\big(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)}\big)\big)^{I_\ell^{(k)}} \\
&= \prod_{k_{n+1}=1}^{|I^{(k)}|} u\big(T^{(k)}, X^{(k,k_{n+1})}_{T^{(k)}}\big) = \prod_{\overline{k} \in \overline{\mathcal{K}}^{n+1}, \overline{k}-=k} u\big(T^{(\overline{k}-)}, X^{(\overline{k})}_{T^{(\overline{k}-)}}\big) \\
&= \prod_{\overline{k} \in \overline{\mathcal{K}}^{n+1}, \overline{k}-=k} \mathbb{E}\Big[\mathbb{1}_{\{T^{(\overline{k})}=T\}} \frac{g\big(X_T^{(\overline{k})}\big)}{F(T - T^{(\overline{k}-)})} \\
&\qquad + \mathbb{1}_{\{T^{(\overline{k})}<T\}} \frac{c_{I^{(\overline{k})}}\big(T^{(\overline{k})}, X^{(\overline{k})}_{T^{(\overline{k})}}, \Delta^{(\overline{k})}\big)\mathcal{J}[u]\big(T^{(\overline{k})}, X^{(\overline{k})}_{T^{(\overline{k})}}, \Delta^{(\overline{k})}\big)^{I^{(\overline{k})}}}{\rho(T^{(\overline{k})} - T^{(\overline{k}-)})p_{I^{(\overline{k})}}} \Big| \mathfrak{F}^n\Big],
\end{aligned}$$

where the final identity is due to (2.7). Thus using $\mathfrak{F}^n$-conditional independence of individual offspring particles, we have

$$\begin{aligned}
&\prod_{k \in \overline{\mathcal{K}}^n \setminus \mathcal{K}^n} \mathcal{J}[u]\big(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)}\big)^{I^{(k)}} \\
&= \mathbb{E}\Big[\prod_{\overline{k} \in \overline{\mathcal{K}}^{n+1}} \Big(\mathbb{1}_{\{T^{(\overline{k})}=T\}} \frac{g\big(X_T^{(\overline{k})}\big)}{F(T - T^{(\overline{k}-)})} \\
&\qquad + \mathbb{1}_{\{T^{(\overline{k})}<T\}} \frac{c_{I^{(\overline{k})}}\big(T^{(\overline{k})}, X^{(\overline{k})}_{T^{(\overline{k})}}, \Delta^{(\overline{k})}\big)\mathcal{J}[u]\big(T^{(\overline{k})}, X^{(\overline{k})}_{T^{(\overline{k})}}, \Delta^{(\overline{k})}\big)^{I^{(\overline{k})}}}{\rho(T^{(\overline{k})} - T^{(\overline{k}-)})p_{I^{(\overline{k})}}} \Big) \Big| \mathfrak{F}^n\Big] \\
&= \mathbb{E}\big[\mathcal{R}_n^{t,x} \big| \mathfrak{F}^n\big].
\end{aligned} \tag{2.9}$$

But then, using (2.9) and the definition of $\mathcal{R}_{n-1}^{t,x}$, we obtain

$$\mathcal{R}_{n-1}^{t,x} = \frac{\mathcal{G}_n^{t,x}}{\mathcal{G}_{n-1}^{t,x}} \frac{\mathcal{C}_n^{t,x}}{\mathcal{C}_{n-1}^{t,x}} \prod_{k \in \overline{\mathcal{K}}^n \setminus \mathcal{K}^n} \mathcal{J}[u]\big(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)}\big)^{I^{(k)}} = \frac{\mathcal{G}_n^{t,x}}{\mathcal{G}_{n-1}^{t,x}} \frac{\mathcal{C}_n^{t,x}}{\mathcal{C}_{n-1}^{t,x}} \mathbb{E}\big[\mathcal{R}_n^{t,x} \big| \mathfrak{F}^n\big].$$

Plugging this into (2.8) yields the claim since $(\mathcal{G}_n^{t,x}, \mathcal{C}_n^{t,x})$ is $\mathfrak{F}^n$-measurable and thus

$$u(t,x) = \mathbb{E}\big[\mathcal{G}_{n-1}^{t,x}\mathcal{C}_{n-1}^{t,x}\mathcal{R}_{n-1}^{t,x}\big] = \mathbb{E}\Big[\mathcal{G}_n^{t,x}\mathcal{C}_n^{t,x}\mathbb{E}\big[\mathcal{R}_n^{t,x} \big| \mathfrak{F}^n\big]\Big] = \mathbb{E}\big[\Psi_n^{t,x}\big]. \qquad \square$$

Remark 2.5. A sufficient condition for (ii) in Theorem 2.4 to hold is that $\|u\|_\infty \leq 1$ and $\sum_{i \in \mathcal{I}} \|c_i\|_\infty < +\infty$; (iii) holds if $\mathrm{D}_x\, u$ and σ are bounded. If u is bounded, similar arguments as in Section 3.2 can be used to obtain sufficient conditions for (i). ⋄

The branching diffusion representation in Theorem 2.4 can be extended in several ways: Upon combining our approach with that in [11], one can also treat mixed local-nonlocal nonlinearities that include first-order derivatives. Moreover, the notion of branching diffusions with jumps can also be extended to that of branching jump-diffusions with jumps, allowing for an additional (linear) nonlocal term in the infinitesimal generator.

3 Viscosity Solutions and Branching Diffusion Representations

In the previous section, our point of view was to start with a classical solution of (PDE) and derive its branching diffusion representation. This result was achieved under the assumption that a classical solution exists. In this section, we study the converse question: Can the branching diffusion representation be used to define a solution of the PDE? It is clear that the branching diffusion representation does in general not yield a sufficiently regular solution to qualify as a classical solution, hence we subsequently work with the weaker concept of viscosity solutions.

The main result of this section gives sufficient conditions under which the branching diffusion representation defines a viscosity solution of (PDE). We then derive conditions under which the family $\{\Psi^{t,x}\}_{(t,x)\in[0,T]\times\mathbb{R}^d}$ is uniformly integrable, which implies one of the key assumptions needed to obtain the viscosity property of the branching diffusion representation.

3.1 Viscosity Solutions of Nonlocal Nonlinear PDEs

We first provide a definition of viscosity solutions of (PDE) appropriate to deal with the nonlocal terms in the nonlinearity.

Definition 3.1 (Viscosity Solution)**.** *Suppose Assumption 2 holds and let $u : [0,T]\times\mathbb{R}^d \to \mathbb{R}$ be a continuous function such that*

$$\sum_{i\in\mathcal{I}} \int_{\Xi} \left| c_i(t,x,\xi)\mathcal{J}[u](t,x,\xi)^i \right| \gamma(\mathrm{d}\xi) < +\infty \qquad \textit{for all } (t,x)\in[0,T)\times\mathbb{R}^d.$$

We say that

1. u *is a* viscosity subsolution *of* (PDE) *if for all* $(t,x)\in[0,T)\times\mathbb{R}^d$ *and all test functions* $\varphi\in\mathcal{C}^{1,2}([0,T]\times\mathbb{R}^d)$ *with* $\varphi(t,x)=u(t,x)$ *and* $\varphi\geq u$ *we have*

$$-\partial_t\varphi(t,x) - \mathcal{A}[\varphi](t,x) - \int_{\Xi} f(t,x,\xi,\mathcal{J}[u](t,x,\xi))\gamma(\mathrm{d}\xi) \leq 0;$$

2. u *is a* viscosity supersolution *of* (PDE) *if for all* $(t,x)\in[0,T)\times\mathbb{R}^d$ *and all test functions* $\varphi\in\mathcal{C}^{1,2}([0,T]\times\mathbb{R}^d)$ *with* $\varphi(t,x)=u(t,x)$ *and* $\varphi\leq u$ *we have*

$$-\partial_t\varphi(t,x) - \mathcal{A}[\varphi](t,x) - \int_{\Xi} f(t,x,\xi,\mathcal{J}[u](t,x,\xi))\gamma(\mathrm{d}\xi) \geq 0;$$

3. u *is a* viscosity solution *of* (PDE) *if it is both a viscosity sub- and supersolution.* ◇

With this definition in place, we can state our second main result.

Theorem 3.2 (Viscosity Property of the Branching Representation). *For all $(t,x) \in [0,T]\times\mathbb{R}^d$, let $\Psi^{t,x}$ be given as in* (2.5) *and define*

$$u : [0,T] \times \mathbb{R}^d \to \mathbb{R}, \qquad (t,x) \mapsto u(t,x) \triangleq \mathbb{E}\big[\Psi^{t,x}\big]. \tag{3.1}$$

In addition to Assumptions 1 and 2, assume that the SDE coefficients μ, σ and the PDE coefficients c_i, $i \in \mathcal{I}$, and g are continuous. Moreover, suppose that the following conditions are satisfied for all $(t,x) \in [0,T] \times \mathbb{R}^d$:

(i) There exists $\varepsilon > 0$ such that the family $\{\Psi^{s,y}\}_{(s,y)\in\mathcal{B}_\varepsilon(t,x)}$ is uniformly integrable.

(ii) There exist $\delta > 0$ and a measurable function $\zeta : \mathcal{I} \times \mathbb{R}^d \to \mathbb{R}$ such that

$$\big|c_i(s,y,\xi)\mathcal{J}[u](s,y,\xi)^i\big| \leq \zeta(i,\xi) \qquad \textit{for all } (s,y) \in \overline{\mathcal{B}}_\delta(t,x)$$

with

$$\sum_{i\in\mathcal{I}} \int_\Xi |\zeta(i,\xi)|\gamma(\mathrm{d}\xi) < +\infty.$$

(iii) It holds that

$$\sum_{i\in\mathcal{I}} \mathbb{E}\Big[\int_t^T \int_\Xi \Big|c_i\big(s,\overline{X}_s^{t,x},\xi\big)\mathcal{J}[u]\big(s,\overline{X}_s^{t,x},\xi\big)^i\Big|\gamma(\mathrm{d}\xi)\mathrm{d}s\Big] < +\infty.$$

Then u is a viscosity solution of (PDE). ◇

Note that if $\sum_{i\in\mathcal{I}} \|c_i\|_\infty < +\infty$ and $\|u\|_\infty \leq 1$ (as in Remark 2.5), then conditions (ii) and (iii) are satisfied; sufficient conditions for (i), which simultaneously imply boundedness of u, are given in Section 3.2 below.

Proof. Note that the uniform integrability assumption (i) implies that $\Psi^{t,x} \in \mathrm{L}^1(\mathbb{P})$ and hence u is well-defined. Moreover, Lemma 2.1 and continuity of c_i, $i \in \mathcal{I}$, and g guarantee that $(t,x) \mapsto \Psi^{t,x}$ is continuous, so u is continuous by Vitali's theorem. Finally, condition (ii) guarantees that u satisfies the integrability conditions required on viscosity solutions in Definition 3.1.

Step 1. We fix $(t,x) \in [0,T) \times \mathbb{R}^d$. As before, we drop the index (t,x) in some of the random variables and processes to simplify notation. Moreover, we set

$$\big(X, I, \Delta\big) \triangleq \big(X^{(1)}, I^{(1)}, \Delta^{(1)}\big).$$

We first observe that

$$\mathbb{1}_{\{T^{(1)}=T\}}\Psi^{t,x} = \mathbb{1}_{\{T^{(1)}=T\}}\frac{g(X_T)}{F(T-t)}$$

as well as

$$\mathbb{1}_{\{T^{(1)}<T\}}\Psi^{t,x} = \mathbb{1}_{\{T^{(1)}<T\}}\frac{c_I(T^{(1)}, X_{T^{(1)}}, \Delta)}{\rho(T^{(1)}-t)p_I}$$

$$\times \prod_{k\in\mathcal{K}\setminus\{(1)\}} \frac{g\big(X_T^{(k)}\big)}{F(T-T^{(k-)})} \prod_{k\in\overline{\mathcal{K}}\setminus(\mathcal{K}\cup\{(1)\})} \frac{c_{I^{(k)}}\big(T^{(k)}, X_{T^{(k)}}^{(k)}, \Delta^{(k)}\big)}{\rho(T^{(k)}-T^{(k-)})p_{I^{(k)}}}$$

$$= \mathbb{1}_{\{T^{(1)}<T\}} \frac{c_I(T^{(1)}, X_{T^{(1)}}, \Delta)}{\rho(T^{(1)}-t)p_I} \prod_{k_2=1}^{|I|} \Psi^{T^{(1)}, X_{T^{(1)}}^{(1,k_2)}}.$$

Using the definition of the branching diffusion and conditional independence structure, we find that

$$\mathbb{E}\Big[\prod_{k_2=1}^{|I|} \Psi^{T^{(1)}, X_{T^{(1)}}^{(1,k_2)}} \Big| \mathfrak{F}^1\Big] = \prod_{k_2=1}^{|I|} \sum_{\ell=1}^{m} \mathbb{1}_{\{J^{(1,k_2)}=\ell\}} \mathbb{E}\big[\Psi^{s,y}\big]\big|_{(s,y)=(T^{(1)}, \Gamma_\ell(T^{(1)}, X_{T^{(1)}}, \Delta))}$$
$$= \mathcal{J}[u]\big(T^{(1)}, X_{T^{(1)}}, \Delta\big)^I.$$

Putting these equations together and using the tower property of conditional expectation, it follows as in Remark 2.3 that

$$\begin{aligned} u(t,x) &= \mathbb{E}\big[\mathbb{E}[\Psi^{t,x}|\mathfrak{F}^1]\big] \\ &= \mathbb{E}\Big[\mathbb{1}_{\{T^{(1)}=T\}} \frac{g(X_T)}{F(T-t)} + \mathbb{1}_{\{T^{(1)}<T\}} \frac{c_I\big(T^{(1)}, X_{T^{(1)}}, \Delta\big)\mathcal{J}[u]\big(T^{(1)}, X_{T^{(1)}}, \Delta\big)^I}{\rho(T^{(1)}-t)p_I}\Big] \\ &= \mathbb{E}\Big[g\big(\overline{X}_T\big) + \int_t^T \int_\Xi f\big(s, \overline{X}_s, \xi, \mathcal{J}[u]\big(s, \overline{X}_s, \xi\big)\big)\gamma(\mathrm{d}\xi)\mathrm{d}s\Big]. \end{aligned} \quad (3.2)$$

For $\varepsilon > 0$, we now introduce the stopping time

$$\tau_\varepsilon \triangleq \inf\big\{s \geq t : |\overline{X}_s - x| \geq \varepsilon\big\} \wedge (t+\varepsilon) \wedge T.$$

From (3.2), the flow property and the strong Markov property of $\overline{X}$ as noted in Lemma 2.1, in combination with the conditional version of Fubini's theorem, which is applicable by (iii), it follows that u satisfies the dynamic programming representation

$$u(t,x) = \mathbb{E}\Big[u\big(\tau_\varepsilon, \overline{X}_{\tau_\varepsilon}\big) + \int_t^{\tau_\varepsilon} \int_\Xi f\big(s, \overline{X}_s, \xi, \mathcal{J}[u]\big(s, \overline{X}_s, \xi\big)\big)\gamma(\mathrm{d}\xi)\mathrm{d}s\Big]. \quad (3.3)$$

Step 2. From the dynamic programming representation (3.3), the viscosity property of u follows by standard arguments. To keep the presentation self-contained, we provide a proof of the subsolution property (the supersolution property is established analogously). We fix a test function $\varphi \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R}^d)$ with $\varphi(t,x) = u(t,x)$ and $\varphi \geq u$. By the dynamic programming representation and Itō's lemma, we find that

$$\begin{aligned} \varphi(t,x) &= u(t,x) \\ &= \mathbb{E}\Big[u\big(\tau_\varepsilon, \overline{X}_{\tau_\varepsilon}\big) + \int_t^{\tau_\varepsilon} \int_\Xi f\big(s, \overline{X}_s, \xi, \mathcal{J}[u]\big(s, \overline{X}_s, \xi\big)\big)\gamma(\mathrm{d}\xi)\mathrm{d}s\Big] \end{aligned}$$

$$
\begin{aligned}
&\leq \mathbb{E}\Big[\varphi\big(\tau_\varepsilon, \overline{X}_{\tau_\varepsilon}\big) + \int_t^{\tau_\varepsilon} \int_\Xi f\big(s, \overline{X}_s, \xi, \mathcal{J}[u]\big(s, \overline{X}_s, \xi\big)\big)\gamma(\mathrm{d}\xi)\mathrm{d}s\Big] \\
&= \mathbb{E}\Big[\varphi(t,x) + \int_t^{\tau_\epsilon} \partial_t\varphi(s, \overline{X}_s) + \mathcal{A}[\varphi](s, \overline{X}_s) \\
&\qquad\qquad\qquad + \int_\Xi f\big(s, \overline{X}_s, \xi, \mathcal{J}[u]\big(s, \overline{X}_s, \xi\big)\big)\gamma(\mathrm{d}\xi)\mathrm{d}s\Big].
\end{aligned}
$$

For $(s,y) \in [0,T] \times \mathbb{R}^d$, we now define

$$
I_\varphi(s,y) \triangleq \partial_t\varphi(s,y) + \mathcal{A}[\varphi](s,y) + \int_\Xi f(s,y,\xi,\mathcal{J}[u](s,y,\xi))\gamma(\mathrm{d}\xi)
$$

to arrive at

$$
\mathbb{E}\Big[\int_t^{\tau_\epsilon} I_\varphi(s, \overline{X}_s)\mathrm{d}s\Big] \geq 0.
$$

From condition (ii) and dominated convergence, it follows that I is continuous. But then

$$
0 \leq \mathbb{E}\Big[\int_t^{\tau_\epsilon} I_\varphi(s, \overline{X}_s)\mathrm{d}s\Big] \leq \mathbb{E}[\tau_\varepsilon - t] \max_{(s,y)\in\overline{\mathcal{B}}_\varepsilon(t,x)} I_\varphi(s,y)
$$

and thus, since $\mathbb{E}[\tau_\varepsilon - t] > 0$,

$$
\max_{(s,y)\in\overline{\mathcal{B}}_\varepsilon(t,x)} I_\varphi(s,y) \geq 0.
$$

Letting $\varepsilon \downarrow 0$ allows us to conclude that $I_\varphi(t,x) \geq 0$, so u is a viscosity subsolution in the sense of Definition 3.1. □

3.2 Sufficient Conditions for Uniform Integrability of $\{\Psi^{t,x}\}$

In this section, we provide a ramification of the results in [11] for branching diffusions with jumps to give sufficient conditions for uniform integrability of $\{\Psi^{t,x}\}$ as required in Theorem 3.2.

Theorem 3.3 (Integrability Conditions)**.** *Let $\kappa \in (1,\infty)$ and define*

$$
C_1 \triangleq \frac{\|g\|_\infty^\kappa}{F(T)^{\kappa-1}} \qquad \textit{and} \qquad C_2 \triangleq \sup_{i\in\mathcal{I},\, t\in[0,T]} \left(\frac{\|c_i\|_\infty}{\rho(t)p_i}\right)^{\kappa-1}.
$$

Then the family $\{\Psi^{t,x}\}_{(t,x)\in[0,T]\times\mathbb{R}^d}$ is bounded in $\mathrm{L}^\kappa(\mathbb{P})$, and in particular uniformly integrable, in either of the following two cases:

(i) It holds that

$$
\frac{C_1}{F(T)} \vee C_2^{\frac{\kappa}{\kappa-1}} \leq 1.
$$

(ii) The power series $\sum_{i\in\mathcal{I}} \|c_i\|_\infty x^{|i|}$ is nonzero[5] *and has an infinite radius of convergence, and*

[5] Otherwise (PDE) becomes linear; for that case the branching approach becomes unnecessary.

the terminal time $T > 0$ is sufficiently small in that

$$T < \int_{C_1}^{\infty} \Big(C_2 \sum_{i \in \mathcal{I}} \|c_i\|_\infty x^{|i|} \Big)^{-1} \mathrm{d}x. \qquad \diamond$$

Proof. Fix some $(t, x) \in [0, T] \times \mathbb{R}^d$. By definition of $\Psi^{t,x}$, we have

$$\big|\Psi^{t,x}\big|^\kappa = \prod_{k \in \mathcal{K}} \frac{\big|g\big(X_T^{(k)}\big)\big|^\kappa}{F(T - T^{(k-)})^\kappa} \prod_{k \in \overline{\mathcal{K}} \setminus \mathcal{K}} \frac{\big|c_{I^{(k)}}\big(T^{(k)}, X_{T^{(k)}}^{(k)}, \Delta^{(k)}\big)\big|^\kappa}{\big|\rho(T^{(k)} - T^{(k-)}) p_{I^{(k)}}\big|^\kappa}.$$

With this, under condition (i), and since F is decreasing we immediately find that

$$\mathbb{E}\big[|\Psi^{t,x}|^\kappa\big] \leq \mathbb{E}\Big[\prod_{k \in \mathcal{K}} \frac{C_1}{F(T)} \prod_{k \in \overline{\mathcal{K}} \setminus \mathcal{K}} C_2^{\frac{\kappa}{\kappa-1}}\Big] \leq \mathbb{E}\Big[\prod_{k \in \overline{\mathcal{K}}} \frac{C_1}{F(T)} \vee C_2^{\frac{\kappa}{\kappa-1}}\Big] \leq 1$$

and the proof is complete. Let us therefore subsequently assume that we are in case (ii); we follow the approach of [11, Theorem 3.12]. The basic idea is to identify an upper bound χ_∞ for $|\Psi^{t,x}|^\kappa$ that can itself be regarded as a branching estimator for an ODE admitting a global solution. First note that

$$\mathbb{E}\big[|\Psi^{t,x}|^\kappa\big] \leq \mathbb{E}\Big[\prod_{k \in \mathcal{K}} \frac{C_1}{F(T - T_t^{(k-)})} \prod_{k \in \overline{\mathcal{K}} \setminus \mathcal{K}} \frac{C_2 \|c_{I^{(k)}}\|_\infty}{\rho(T^{(k)} - T^{(k-)}) p_{I^{(k)}}}\Big]. \tag{3.4}$$

Now consider the the following ODE to be solved backwards in time:

$$\dot{\eta}(t) + C_2 \sum_{i \in \mathcal{I}} \|c_i\|_\infty \eta(t)^{|i|} = 0, \quad t \in [0, T]; \qquad \eta(T) = C_1. \tag{3.5}$$

Define the map

$$\varphi : [C_1, \infty) \to [0, \infty], \qquad y \mapsto \varphi(y) \triangleq \int_{C_1}^{y} \Big(C_2 \sum_{i \in \mathcal{I}} \|c_i\|_\infty x^{|i|} \Big)^{-1} \mathrm{d}x.$$

Since the power series is non-degenerate, φ is a continuous mapping that is strictly increasing where finite. Upon rearranging and integrating the ODE, note that $\eta \in \mathcal{C}^1([0, T])$ is a solution of (3.5) if and only if

$$-\int_t^T \Big(C_2 \sum_{i \in \mathcal{I}} \|c_i\|_\infty \eta(t)^{|i|} \Big)^{-1} \dot{\eta}(s) \mathrm{d}s = \int_t^T 1 \mathrm{d}s, \quad t \in [0, T]; \qquad \eta(T) = C_1,$$

and the substitution $x \triangleq \eta(s)$ shows that this is the case if and only if

$$\varphi(\eta(t)) = \int_{C_1}^{\eta(t)} \Big(C_2 \sum_{i \in \mathcal{I}} \|c_i\|_\infty x^{|i|} \Big)^{-1} \mathrm{d}x = T - t, \quad t \in [0, T].$$

Since φ is strictly increasing where it is finite, the latter statement is equivalent to

$$T \in \mathrm{range}(\varphi) \cap \mathbb{R} = \Big\{ \xi \in \mathbb{R} : 0 \leq \xi \leq \int_{C_1}^{\infty} \Big(C_2 \sum_{i \in \mathcal{I}} \|c_i\|_\infty z^{|i|} \Big)^{-1} \mathrm{d}z \Big\}.$$

Thus condition (ii) on T is both necessary and sufficient to guarantee the existence of a solution η of ODE (3.5) on $[0, T]$. With this, we are now able to define

$$\chi_n \triangleq \prod_{k \in \bigcup_{\nu=1}^{n} \mathcal{K}^\nu} \frac{C_1}{F(T - T^{(k-)})} \prod_{k \in \bigcup_{\nu=1}^{n} \overline{\mathcal{K}}^\nu \setminus \mathcal{K}^\nu} \frac{C_2 \|c_{I^{(k)}}\|_\infty}{\rho(T^{(k)} - T^{(k-)}) p_{I^{(k)}}} \prod_{k \in \overline{\mathcal{K}}^{n+1}} \eta(T^{(k-)})$$

as well as

$$\chi_\infty \triangleq \lim_{n \to \infty} \chi_n = \prod_{k \in \mathcal{K}} \frac{C_1}{F(T - T^{(k-)})} \prod_{k \in \overline{\mathcal{K}} \setminus \mathcal{K}} \frac{C_2 \|c_{I^{(k)}}\|_\infty}{\rho(T^{(k)} - T^{(k-)}) p_{I^{(k)}}}.$$

By analogous arguments as in Remark 2.3, we obtain

$$\eta(t) = \eta(T) + \int_t^T C_2 \sum_{i \in \mathcal{I}} \|c_i\|_\infty \eta(s)^{|i|} \mathrm{d}s = \mathbb{E}[\chi_1] = \ldots = \mathbb{E}[\chi_n], \ t \in [0, T]; \quad n \in \mathbb{N}.$$

But then, thanks to (3.4) and Fatou's lemma, it follows that

$$\mathbb{E}\big[|\Psi^{t,x}|^\kappa\big] \leq \mathbb{E}[\chi_\infty] \leq \liminf_{n \to \infty} \mathbb{E}[\chi_n] = \eta(t) \leq \sup_{t \in [0,T]} \eta(t) < \infty,$$

and the proof is complete. □

Under the conditions of Theorem 3.3, it follows in particular that $\{\Psi^{t,x}\}_{(t,x) \in [0,T] \times \mathbb{R}^d}$ is bounded in $\mathrm{L}^1(\mathbb{P})$, so the function $u : [0, T] \times \mathbb{R}^d \to \mathbb{R}$, $u(t, x) \triangleq \mathbb{E}[\Psi^{t,x}]$ is in fact bounded.

4 Monte Carlo Simulation: A High-Dimensional Example

The representation (3.1) derived in Theorem 3.2 makes it possible to compute solutions of (PDE) by direct (non-nested, plain vanilla) Monte Carlo simulation. To illustrate the effectiveness and efficiency of the branching Monte Carlo algorithm, we consider the following stylized problem in $d \geq 1$ dimensions:

$$\partial_t u(t,x) + \frac{1}{2d^2}\Delta u(t,x) + \int_{\mathbb{R}^d} \sum_{i \in \mathcal{I}} c_i(t,x) u(t,x)^{i_1} u(t,x+\xi)^{i_2} \gamma(\mathrm{d}\xi) = 0$$

$$u(T,x) = \cos(1_d^\intercal x).$$

Here Δ denotes the Laplace operator in $\mathbb{R}^d$; the time horizon is $T = 1$; γ is the discrete uniform distribution supported on[6] $\{-(\pi/2)e_i \in \mathbb{R}^d : i \in [1:d]\}$; and the set of possible descendants is given by $\mathcal{I} \triangleq \{i \in \mathbb{N}_0^2 : |i| \leq 2\}$ with coefficients

$$\begin{aligned}
c_{(0,0)}(t,x) &= (\alpha + 1/(2d)) \cos(1_d^\intercal x) \exp\{\alpha(T-t)\} + \cos(1_d^\intercal x)^2/d - 1/(2d), \\
c_{(1,0)}(t,x) &= (-1) \cdot \cos(1_d^\intercal x) \exp\{-\alpha(T-t)\}/d, \\
c_{(0,1)}(t,x) &= (-1) \cdot c_{(1,0)}(t,x), \\
c_{(2,0)}(t,x) &= c_{(0,2)}(t,x) = \exp\{-2\alpha(T-t)\}/(2d), \\
c_{(1,1)}(t,x) &= (-2) \cdot c_{(2,0)}(t,x),
\end{aligned}$$

where $\alpha = 0.2$. It is not hard to verify that the solution is given in closed form by

$$u(t,x) = \cos(1_d^\intercal x) e^{\alpha(T-t)} = \cos\big(\textstyle\sum_{i=1}^d x_i\big) e^{\alpha(T-t)}, \qquad (t,x) \in [0,T] \times \mathbb{R}^d.$$

We refer to u as the *exact solution* and use it as a benchmark to quantify the error of the estimates produced by our algorithm. All subsequent simulation results correspond to the initial configuration $(t,x) = (0, 1_d)$, i.e. we determine $u^* \triangleq u(0, 1_d)$.

The choice of parameters for the branching diffusion is reported in Table 1.

Parameter	Value
$\mathsf{Law}(\tau^{(1)})$	$\Gamma(\kappa,\theta)$ with $\kappa = 0.5$ and $\theta = 2.5$
$p_{(0,0)}$	$1/3$
$p_{(1,0)}, p_{(0,1)}, p_{(1,1)}$	$1/6$
$p_{(2,0)}, p_{(0,2)}$	$1/12$

Table 1 Parameters of the branching diffusion.

Our simulation study is conducted as follows: Given a spatial dimension $d \in \mathbb{N}$ and a number of Monte Carlo simulations N, a simulation run consists of computing the estimator $\hat{u}_{d,N}$ of u^* as the average of N i.i.d. copies $\{\Psi_n^{0,1_d}\}_{n\in[1:N]}$ of $\Psi^{0,1_d}$. All numerical computations are implemented in Matlab.[7] Table 2 presents the simulation results for the estimator $\hat{u}_{d,N}$ for

[6] Here e_i denotes the i^{th} unit vector in $\mathbb{R}^d$.

[7] Hardware: Intel Core i7-6700 (3.40 GHz) with 31.2 GiB RAM; Linux Mint OS (18.1 Serena); no parallelization.

several values of d and N.

N \ d	3	5	10	20	50	100
10^2	-1.6463 (0.255)	0.3150 (0.130)	-0.9049 (0.131)	0.4225 (0.068)	1.2471 (0.122)	1.3105 (0.109)
10^3	-1.2254 (0.124)	0.3240 (0.043)	-1.0134 (0.037)	0.4909 (0.019)	1.1319 (0.038)	1.0323 (0.034)
10^4	-1.2643 (0.054)	0.3573 (0.014)	-1.0311 (0.012)	0.5013 (0.006)	1.1668 (0.012)	1.0488 (0.011)
10^5	-1.1758 (0.017)	0.3492 (0.005)	-1.0260 (0.004)	0.4956 (0.002)	1.1797 (0.004)	1.0549 (0.003)
10^6	-1.2064 (0.007)	0.3485 (0.001)	-1.0262 (0.001)	0.4969 (0.001)	1.1791 (0.001)	1.0536 (0.001)
u^*	-1.2092	0.3465	-1.0248	0.4984	1.1786	1.0532

Table 2 Simulation results for $\hat{u}_{d,N}$, standard deviation in brackets.

To quantify the accuracy of the Monte Carlo algorithm, we further denote the relative error compared to the exact solution by

$$\mathsf{err}_{\mathsf{rel}}(\hat{u}_{d,N}) \triangleq \frac{|\hat{u}_{d,N} - u^*|}{|u^*|}$$

and report our simulation results in Table 3.

N \ d	3	5	10	20	50	100
10^2	33.6% (36.1%)	30.1% (26.7%)	8.9% (7.2%)	9.9% (7.8%)	7.8% (5.7%)	7.9% (6.3%)
10^3	13.5% (14.5%)	10.9% (8.8%)	3.1% (2.4%)	3.3% (2.4%)	2.5% (2.0%)	2.6% (2.1%)
10^4	4.4% (4.3%)	3.5% (2.6%)	0.8% (0.6%)	1.0% (0.8%)	0.8% (0.7%)	0.7% (0.6%)
10^5	1.6% (1.4%)	1.0% (0.8%)	0.3% (0.2%)	0.3% (0.2%)	0.3% (0.2%)	0.2% (0.2%)
10^6	0.6% (0.4%)	0.3% (0.3%)	0.1% (0.1%)	0.1% (0.1%)	0.1% (0.1%)	0.1% (0.1%)

Table 3 Relative error $\mathsf{err}_{\mathsf{rel}}(\hat{u}_{d,N})$, standard deviation in brackets.[8]

The algorithm is able to achieve high levels of accuracy even for high dimensions of $d = 50$ and $d = 100$, provided N is sufficiently large. Note that in this example the precision of the results actually increases with the dimension d; this is in line with Theorem 3.3, as in the above specification the norm $\|c_i\|_\infty$ appearing in the constant C_2 decreases with d.

At the same time, the running times required to achieve these levels of accuracy are rather modest; the exact values are reported in Table 4.

[8] Estimates and standard deviations are based on 100 mutually independent simulation runs of $\hat{u}_{d,N}$.

N \ d	3	5	10	20	50	100
10^2	0.1 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
10^3	0.1 (0.0)	0.2 (0.0)	0.2 (0.0)	0.1 (0.0)	0.2 (0.0)	0.2 (0.0)
10^4	1.4 (0.1)	1.5 (0.1)	1.5 (0.1)	1.5 (0.1)	1.6 (0.1)	3.2 (0.1)
10^5	14.3 (0.5)	14.4 (0.3)	14.4 (0.3)	14.6 (0.3)	15.2 (0.3)	16.7 (0.3)
10^6	140.8 (1.1)	140.5 (1.4)	141.2 (1.3)	142.5 (1.0)	162.2 (4.6)	185.5 (14.9)

Table 4 Running time (in seconds) to compute $\hat{u}_{d,N}$, standard deviation in brackets.[8]

Finally, Figure 2 displays the relative error, the standard deviation, and the running time as functions of the number of Monte Carlo samples N. As expected from the Central Limit Theorem, the slope in the logarithmic plot of the standard deviation is approximately $-1/2$. The running times also demonstrate that there is no curse of dimensionality effect.

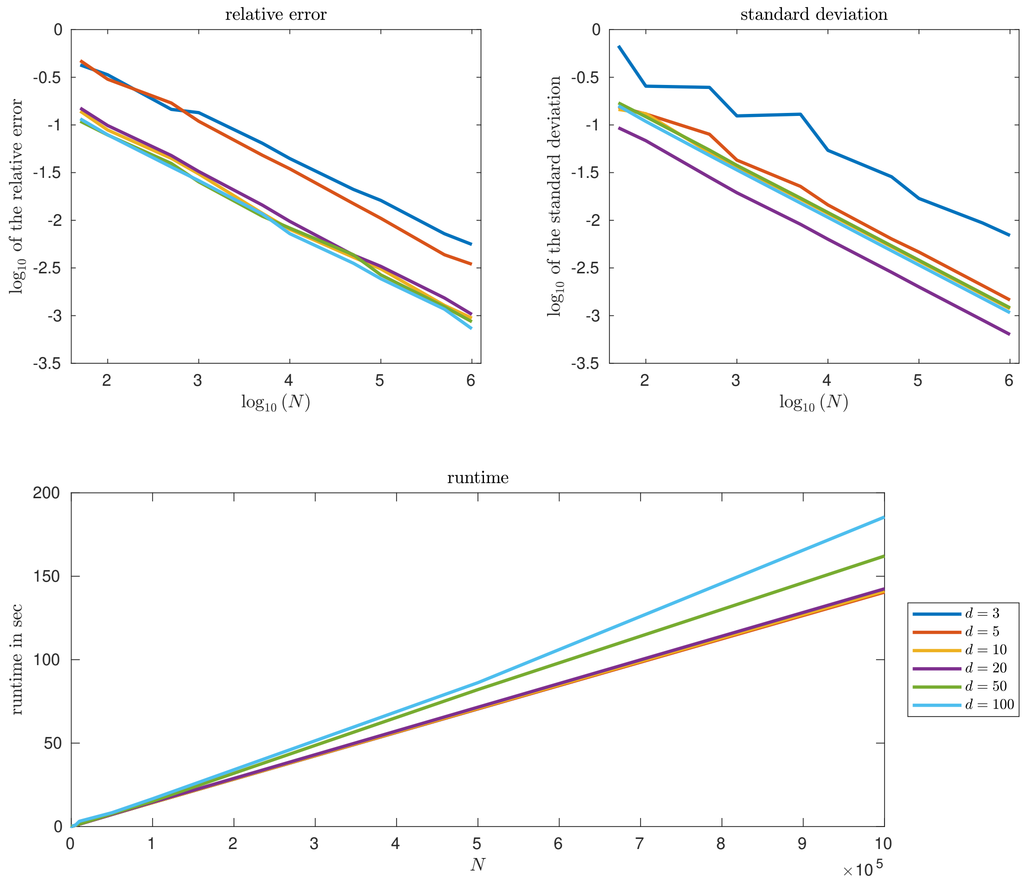

Figure 2 Simulation results.

5 Valuation with Systemically Important Counterparties

In this section we illustrate the usefulness of the theory developed in this article in the valuation of financial positions with systemic counterparty credit risk. Specifically, we assume that the counterparty in a given financial position is a systemically important bank (SIB);[9] its systemic importance is captured by jumps in the underlying risk factors, or equivalently devaluations in risky asset prices, that occur upon the SIB's default. This model setup was first proposed by [19]; see also [5] and [17]. We wish to stress that our focus here is on situations where finite-difference or fixed point methods for the pricing PDE (as developed by, e.g., [6] or [14]) are not applicable.

5.1 Valuation with Systemic Risk

We consider a financial market that is free of arbitrage; the underlying probability measure (denoted by $\mathbb{P}$ in the preceding sections) is taken as the relevant risk-neutral pricing measure $\mathbb{Q}$. The financial market consists of a locally riskless money market account $B = \{B_t\}_{t\in[0,T]}$ and $d \in \mathbb{N}$ dynamically traded risky assets with prices $X = \{X_t\}_{t\in[0,T]}$ given by an $\mathbb{R}^d_+$-valued semimartingale such that X/B is a local $\mathbb{Q}$-martingale. The financial position whose price is to be determined promises a time-T payoff $g(X_T)$ where g is a bounded measurable function of the underlyings X, provided the counterparty does not default before time T.

Crucially, the counterparty in this financial position is a defaultable, systemically important bank (SIB). This means that (i) the financial position is subject to credit risk; and (ii) if and when the SIB counterparty defaults, there is a negative impact on risky asset prices X, which simultaneously affects the value of the financial position $g(X_T)$. Assuming fractional recovery of post-default mark-to-market value as in [7], risk-neutral pricing yields

$$\frac{V_t}{B_t} = \mathbb{E}^{\mathbb{Q}}_t\Big[\mathbb{1}_{\{\tau>T\}}\frac{g(X_T)}{B_T} + \mathbb{1}_{\{\tau\le T\}}\frac{h(V_\tau)}{B_\tau}\Big] \qquad \text{on } \{\tau > t\} \tag{5.1}$$

where V represents the value of the financial position *provided the counterparty has not defaulted*; τ is the (original) counterparty's default time; and the recovery value the investor retrieves is given by $h(V_\tau)$, where $h(v) \triangleq Rv^+ - v^-$ with a recovery rate $R \in [0, 1]$. Note that in (5.1), V_τ represents the time-τ mark-to-market price of an identical financial position with an SIB counterparty that has *not* defaulted, but is otherwise identical to the original one, immediately after the original counterparty's default. In particular, V_τ is based on *post-default* risky asset prices $X_\tau = X_{\tau-} + \Delta X_\tau$.

Specifically, we assume that default events are modeled within a classical reduced-form framework, and that the risk-neutral dynamics of X are given by an $\mathbb{R}^d_+$-valued jump diffusion. Thus the SIB counterparty's default time is given by

$$\tau \triangleq \inf\{t \ge 0 \,:\, Y_t \ne 0\}$$

[9] A list of banks classified as globally systemically important by the Financial Stability Board is available at http://www.fsb.org/work-of-the-fsb/policy-development/addressing-sifis/global-systemically-important-financial-institutions-g-sifis.

where $Y = \{Y_t\}_{t\in[0,T]}$ is a Cox process with intensity $\{\lambda(t, X_t)\}_{t\in[0,T]}$, and the risk-neutral dynamics of B and X are given by

$$
\begin{aligned}
\mathrm{d}B_t &= r(t, X_t)B_t\mathrm{d}t \\
\mathrm{d}X_t &= \mathrm{diag}(X_{t-})\big[\mu(t, X_{t-})\mathrm{d}t + \sigma(t, X_{t-})\mathrm{d}W_t + \Delta_{Y_t}\mathrm{d}Y_t\big], \qquad X_0 = x
\end{aligned}
$$

where $\{\Delta_n\}_{n\in\mathbb{N}}$ are i.i.d. $E \triangleq (-1, \infty)^d$-valued and independent of Y and W, and

$$\mu(t, x) \triangleq r(t, x)1_d - \lambda(t, x)\mathbb{E}[\Delta_1].$$

In particular, the SIB counterparty's default triggers a simultaneous devaluation in the financial position's underlyings of size Δ_1. The corresponding pricing PDE is given by

$$\partial_t u(t, x)+\mathcal{A}^\star[u](t, x)-r(t, x)u(t, x)+\lambda(t, x)\int_E \Big(h\circ u\big(t, x+\mathrm{diag}(\xi)x\big)-u(t, x)\Big)\nu(\mathrm{d}\xi) = 0$$

subject to the boundary condition $u(T, x) = g(x)$, where ν denotes the distribution of Δ_n, $n \in \mathbb{N}$, and the operator $\mathcal{A}^\star$ is given by

$$\mathcal{A}^\star[u](t, x) \triangleq \mu(t, x)^\mathsf{T}\,\mathrm{diag}(x)\,\mathrm{D}_x\, u(t, x) + \frac{1}{2}\,\mathrm{tr}\Big[\mathrm{diag}(x)\sigma(t, x)\sigma(t, x)^\mathsf{T}\,\mathrm{diag}(x)\,\mathrm{D}_x^2\, u(t, x)\Big].$$

Note that, similarly as in other credit risk valuation problems with recovery of mark-to-market value, the pricing formula (5.1) and the corresponding pricing equation are inherently implicit, with the price to be determined appearing inside the nonlinearity that represents the recovery value (see, e.g., [6] or [14] and the references therein); with the additional complication that jumps at default imply that this term is also non-local.

5.2 Branching Diffusion with Jumps Approach

In order to solve the pricing equation, we follow §5 in [10] and approximate the recovery function h by a polynomial. Thus we assume without loss that $\|g\|_{\mathrm{L}^\infty} \le 1$ (since g is bounded, this can always be achieved by appropriate rescaling) and approximate the recovery value function using $v^\pm \approx \sum_{m=0}^M \alpha_m^\pm v^m$, $v \in [-1, 1]$; we obtain

$$
\begin{aligned}
\partial_t u(t, x) + \mathcal{A}^\star[u](t, x) - \big[r(t, x) + \lambda(t, x)\big]u(t, x)& \\
+ \int_E \sum_{m=0}^{M} \big[R\alpha_m^+ - \alpha_m^-\big]\lambda(t, x)u\big(t, x + \mathrm{diag}(\xi)x\big)^m\nu(\mathrm{d}\xi) = 0.& \qquad (5.2)
\end{aligned}
$$

Hence, setting $\mathcal{I} \triangleq \{(\ell, m) \in [0 : M]^2 \,:\, \ell = 0 \text{ or } m = 0\}$ we can rewrite (5.2) as

$$
\begin{aligned}
\partial_t u(t, x) + \mathcal{A}^\star[u](t, x) + \int_E \sum_{i\in\mathcal{I}} c_i(t, x)u\big(t, x + \mathrm{diag}(\xi)x\big)^i\nu(\mathrm{d}\xi) = 0,& \qquad (5.3) \\
u(T, x) = g(x),&
\end{aligned}
$$

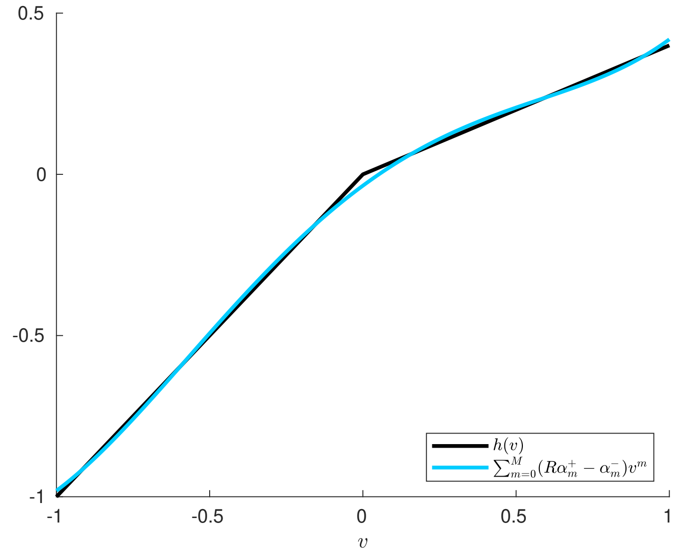

Figure 3 Approximation of h by a polynomial.

where the coefficients $c_{(\ell,m)}$, $(\ell, m) \in \mathcal{I}$, are given by

$$c_{(0,m)}(t,x) \triangleq [R\alpha_m^+ - \alpha_m^-]\lambda(t,x) \quad \text{for } m \in [0:M]$$
$$c_{(1,0)}(t,x) \triangleq -[r(t,x) + \lambda(t,x)] \qquad \text{and} \qquad c_{(\ell,0)}(t,x) \triangleq 0 \quad \text{for } \ell \in [2:M].$$

Since (5.3) is a special case of (PDE), the branching diffusion with jumps approach developed in this paper can be applied to compute the solution via $u(0,x) = \mathbb{E}[\Psi^{0,x}]$ with

$$\Psi^{0,x} \triangleq \prod_{k\in\mathcal{K}} \frac{g(X_T^{(k)})}{F(T - T^{(k-)})} \quad \times \prod_{k\in\overline{\mathcal{K}}\setminus\mathcal{K},\, I^{(k)}=(1,0)} -\frac{r(T^{(k)}, X_{T^{(k)}}^{(k)}) + \lambda(T^{(k)}, X_{T^{(k)}}^{(k)})}{\rho(T^{(k)} - T^{(k-)})p_{(1,0)}}$$
$$\times \prod_{m=0}^{M} \prod_{k\in\overline{\mathcal{K}}\setminus\mathcal{K},\, I^{(k)}=(0,m)} \frac{[R\alpha_m^+ - \alpha_m^-]\lambda(T^{(k)}, X_{T^{(k)}}^{(k)})}{\rho(T^{(k)} - T^{(k-)})p_{(0,m)}}.$$

Here $X^{(k)} \triangleq X^{k,0,x}$, $k \in \overline{\mathcal{K}} \triangleq \overline{\mathcal{K}}^0$, is a branching diffusion with jumps as specified in Section 2, where the dynamics of each individual particle $k \in \overline{\mathcal{K}}$ are given by

$$\mathrm{d}X_t^{(k)} = \mathrm{diag}\big(X_{t-}^{(k)}\big)\Big[\mu\big(t, X_{t-}^{(k)}\big)\mathrm{d}t + \sigma\big(t, X_{t-}^{(k)}\big)\mathrm{d}W_t\Big], \quad t \in [T^{(k-)}, T^{(k)}],$$

with initial conditions

$$X_0^{(1)} = x \qquad \text{and} \qquad X_{T^{(k-)}}^{(k)} \triangleq \begin{cases} X_{T^{(k-)}}^{(k-)} & \text{if } I^{(k)} = (1,0), \\ X_{T^{(k-)}}^{(k-)} + \mathrm{diag}\big(\Delta^{(k-)}\big)X_{T^{(k-)}}^{(k-)} & \text{else.} \end{cases}$$

5.3 Numerical Illustration

While our setup and Monte Carlo methodology allow for general diffusion dynamics and arbitrary dimensionality in the underlying risky assets, for illustration we use a baseline Black-Scholes-Merton model, enhanced by the SIB's credit risk with constant default intensity and devaluations captured by the model of Kou [16]. Thus we take Y as a Poisson process with intensity $\lambda \geq 0$; the riskless rate $r \in \mathbb{R}$ is constant; and g has a single underlying ($d = 1$) with volatility $\sigma > 0$. Jumps at default are such that $-\log(1 + \Delta)$ is exponentially distributed with parameter $\eta \geq 0$. We consider two financial positions representing a shifted put and shifted discount call, respectively, i.e.

$$g_{\text{short}}(X_T) \triangleq (K - X_T)^+ - L \qquad \text{and} \qquad g_{\text{long}}(X_T) = L - (K - X_T)^+.$$

Figure 4 provides an illustration of the corresponding payoffs.

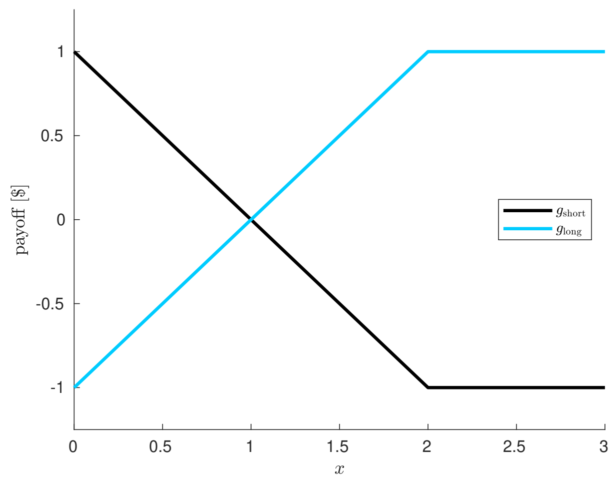

Figure 4 Payoff profiles g_{short} and g_{long} with $K = 2$ and $L = 1$.

To quantify the impact of systemic risk, we compare the valuations of the financial positions g_{short} and g_{long} in three benchmark scenarios that are identical in all respects, except the choice of counterparty:

- *SIB counterparty:* The counterparty is systemically important, and their default triggers devaluations in risky asset prices (see above).
- *Non-SIB counterparty:* The counterparty is non-systemic, but otherwise identical to the SIB; in particular, it is defaultable with the same credit risk characteristics as the SIB.
- *Default-free counterparty:* There is no counterparty credit risk.[10]

In all three scenarios, the SIB is part of the model and will cause devaluations upon default; the scenarios thus differ only in the choice of counterparty and the resulting wrong- or right-way

[10]In this scenario, the pricing PDE becomes linear and can be solved in closed form; see, e.g., [15] or [16].

risk. The implementation of the branching diffusion with jumps is based on the parameters specified in Table 6. In contrast to Section 4, where we deliberately employ a non-accelerated standard Monte Carlo method for performance analysis, here we exploit standard techniques for variance reduction such as control variates and parallelization, see [8] and [15].[11] The relevant model and simulation parameters are reported in Tables 5 and 6; note in particular that the expected devaluation upon the SIB's default is -50%.

Coefficient	T	r	σ	R	η	x	K	L
Value	1	0.5%	25%	40%	1	1	2	1

Table 5 Market coefficients.

Parameter	Value
N	$8.8 \cdot 10^6$
$\mathsf{Law}(\tau)$	$\Gamma(\kappa, \theta)$ with $\kappa = 0.5$ and $\theta = 2.5$
M	4
$(\alpha_0^\pm, \ldots, \alpha_4^\pm)$	$(0.06, \pm 0.50, 0.82, 0.00, -0.41)$
$p_{(0,0)}$	q_0
$p_{(\ell,0)}$	$q_{\mathsf{loc}} \cdot q_\ell$ for $\ell \in [1:M]$
$p_{(0,m)}$	$(1 - q_{\mathsf{loc}}) \cdot q_m$ for $m \in [1:M]$
q_{loc}	$\frac{\lvert r+\lambda \rvert}{\lvert r+\lambda \rvert + \lambda \cdot \sum_{m=0}^M \lvert R\alpha_m^+ - \alpha_m^- \rvert}$
q_m	$\frac{\lvert c_{(1,0)} \rvert^2 \cdot \mathbb{1}_{\{m=1\}} + \lvert c_{(0,m)} \rvert \cdot \sum_{m=0}^M \lvert c_{(0,m)} \rvert}{\lvert c_{(1,0)} \rvert^2 + (\sum_{m=0}^M \lvert c_{(0,m)} \rvert)^2}$ for $m \in [0:M]$

Table 6 Simulation parameters.

Figures 5 through 7 display our simulation results for different specifications of the default intensity. Figure 5 illustrates the impact of systemic interaction for the financial position g_{short}: Devaluations imply a positive correlation between the SIB's default events and underperforming risky asset prices, causing significant wrong-way risk for short positions. While this is qualitatively apparent, the quantitative size of this effect, in particular relative to the non-SIB counterparty, is remarkable.

In Figure 6 we thus decompose the implied SIB spread (i.e., the difference between the value of an otherwise identical financial position with a default-free counterparty and that with an SIB counterparty) into (i) a pure credit risk component (green), which we identify with the spread between the non-systemic defaultable counterparty and the default-free one, and (ii) a systemic risk component (red), which is present only due to the counterparty's systemic importance. It is apparent that systemic risk is the main driver of the spread, accounting for $80 - 85\%$ of the total spread across realistic default intensities.

Finally, Figure 7 demonstrates that the effect is reversed for long positions; we might refer to this as right-way risk, i.e. negative correlation between the counterparty's default and the value of the financial position. Technically, the systemic risk component turns negative, and the spread becomes significantly smaller than for a non-systemic counterparty.

[11] The control variate is the default-free price.

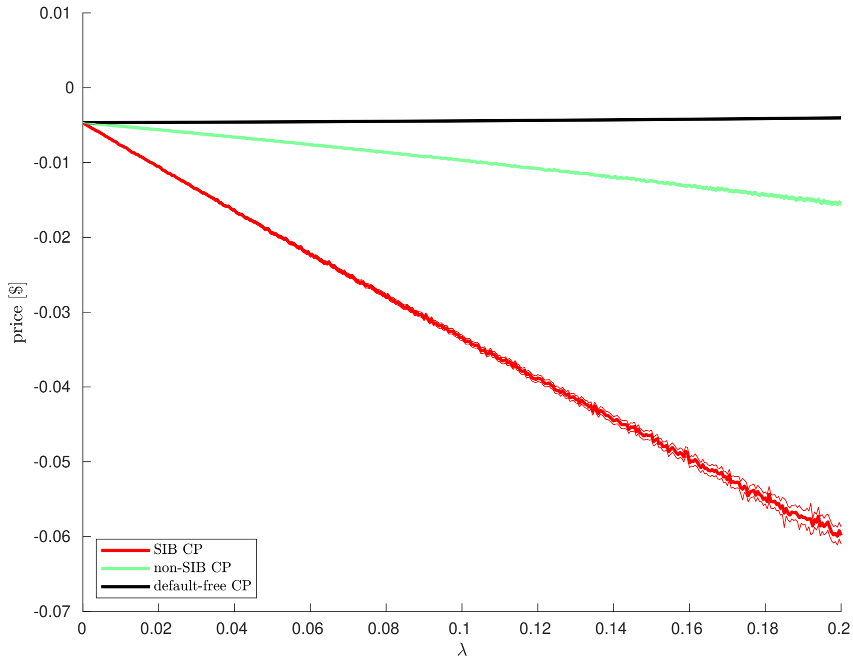

Figure 5 Valuation of g_{short} as a function of the counterparty's default intensity (with 99%-confidence bands).

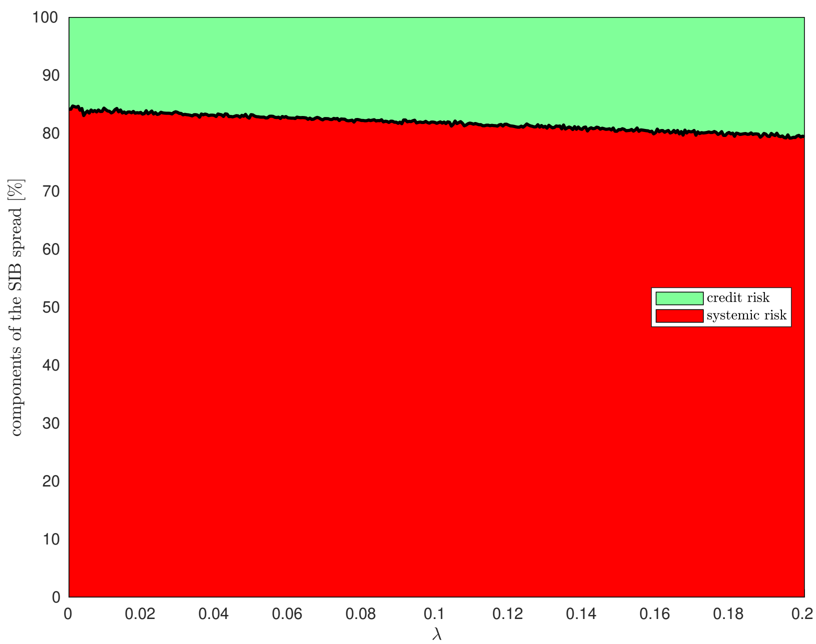

Figure 6 Decomposition of the SIB spread into credit risk (green) and systemic risk (red).

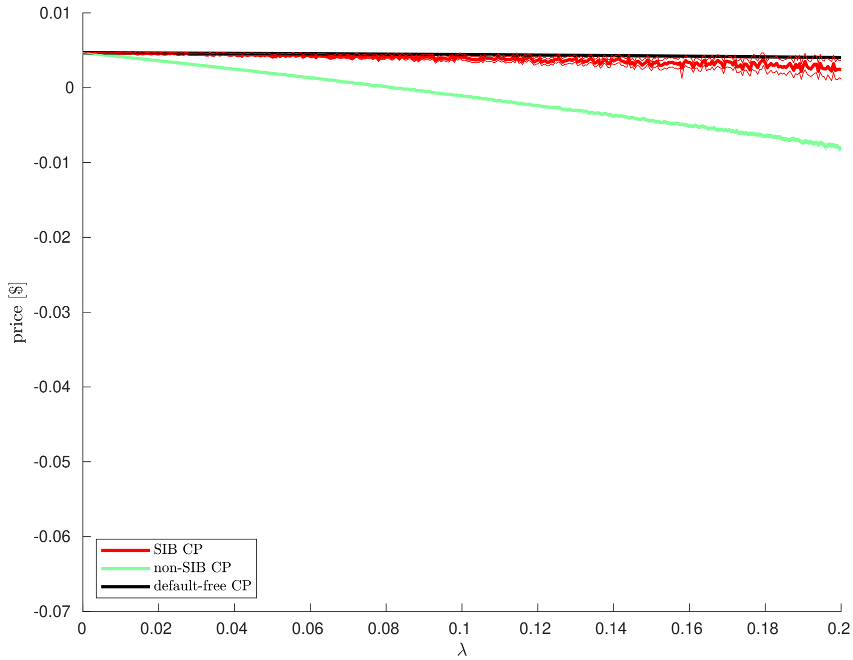

Figure 7 Valuation of g_2 as a function of the counterparty's default intensity (with 99%-confidence bands).

To conclude, this paper has developed a branching diffusion with jumps approach to solving parabolic PDEs with nonlocal nonlinearities. We have showcased the performance of the resulting non-nested Monte Carlo methodology, and we have demonstrated how it applies to the valuation of financial positions with systemic counterparties and mark-to-market recovery. Several extensions to the above pricing model are conceivable within the setup of this paper, including stochastic recovery rates, shocks to state variables (e.g., volatilities), idiosyncratic jumps in asset prices, and margin periods of risk.

Declarations of Interest. Daniel Hoffmann and Frank Seifried gratefully acknowledge financial support from Deutsche Forschungsgemeinschaft (DFG) within the Research Training Group 2126 (Algorithmic Optimization).

References

[1] A. Agarwal and J. Claisse. Branching diffusion representation of semi-linear elliptic PDEs and estimation using Monte Carlo method. https://arxiv.org/abs/1704.00328, 2018.

[2] K. B. Athreya and P. E. Ney. *Branching Processes*. Springer, 1972.

[3] B. Bouchard, X. Tan, and X. Warin. Numerical approximation of general Lipschitz BSDEs with branching processes. *ESAIM: Proceedings and Surveys*, 65:309–329, 2019.

[4] B. Bouchard, X. Tan, X. Warin, and Y. Zou. Numerical approximation of BSDEs using local polynomial drivers and branching processes. *Monte Carlo Methods and Applications*, 23(4):241–263, 2017.

[5] D. Brigo, N. Pede, and A. Petrelli. Multi currency credit default swaps: Quanto effects and FX devaluation jumps. https://arxiv.org/abs/1512.07256, 2018.

[6] P. Carr and S. Ghamami. Derivatives pricing under bilateral counterparty risk. *Journal of Risk*, 20(1):77–107, 2017.

[7] D. Duffie and K. Singleton. Modeling term structure of defaultable bonds. *Review of Financial Studies*, 12:687–720, 1999.

[8] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2003.

[9] T. E. Harris. *The Theory of Branching Processes*. Springer, 1963.

[10] P. Henry-Labordère. Counterparty risk valuation: A marked branching diffusion approach. *Risk Magazine*, July 2012.

[11] P. Henry-Labordère, N. Oudjane, X. Tan, N. Touzi, and X. Warin. Branching diffusion representation of semilinear PDEs and Monte Carlo approximation. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 55(1):184–210, 2019.

[12] P. Henry-Labordère, 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.

[13] P. Henry-Labordère and N. Touzi. Branching diffusion representation for nonlinear Cauchy problems and Monte Carlo approximation. https://arxiv.org/abs/1801.08794, 2018.

[14] J. Kim and T. Leung. Pricing derivatives with counterparty risk and collateralization: A fixed point approach. *European Journal of Operational Research*, 249(2):525–539, 2016.

[15] R. Korn, E. Korn, and G. Kroisandt. *Monte Carlo Methods and Models in Finance and Insurance*. CRC Press, 2010.

[16] S. G. Kou. A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101, 2002.

[17] F. Mercurio and M. Li. Jumping with default: Wrong-way risk modeling for CVA. *Risk*, pages 58–63, December 2015.

[18] E. Pardoux and A. Răşcanu. *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations*. Springer, 2014.

[19] M. Pykhtin and A. Sokol. Exposure under systemic risk impact. *Risk*, pages 88–93, September 2013.