

# Unbiased Simulation Estimators for Jump-Diffusions

Guanting Chen<sup>\*</sup>   Alex Shkolnik<sup>†</sup>   Kay Giesecke<sup>‡</sup>

June 29, 2019

This Version: September 11, 2019<sup>§</sup>

*Winter Simulation Conference 2019, forthcoming*

## Abstract

We develop and analyze an unbiased Monte Carlo estimator for a functional of a one-dimensional jump-diffusion process with a state-dependent drift, volatility, jump intensity and jump size. The approach combines a change of measure to sample the jumps with the parametrix method to simulate the diffusions. Under regularity conditions on the coefficient functions as well as the functional, we prove the unbiasedness and the finite variance property of the estimator. Numerical experiments illustrate the performance of the scheme.

---

<sup>§</sup>We thank the participants at the 9th Western Conference in Mathematical Finance for the comments and questions that led to improvements of a previous version of this article. We also thank two anonymous referees for detailed comments on the original manuscript.

<sup>\*</sup>Institute for Computational and Mathematical Engineering, Stanford University, CA 94305  
EMAIL: guanting@stanford.edu.

<sup>†</sup>Department of Statistics and Applied Probability, University of California, Santa Barbara, CA 93106. EMAIL: shkolnik@ucsb.edu

<sup>‡</sup>Department of Management Science & Engineering, Stanford University, Stanford, CA 94305. EMAIL: giesecke@stanford.edu.

# 1. Introduction

Jump-diffusion processes are widely used to model the dynamics of prices, interest rates and other financial market quantities. However, only relatively basic jump-diffusion models are analytically tractable. Simulation methods are often employed to treat the computational problems involving richer models.

Discretization methods are broadly applicable, but generate biased simulation estimators (see Shkolnik, Giesecke, Teng & Wei (2019) for an error analysis). Exact sampling schemes for one-dimensional diffusions (Beskos & Roberts 2005, Chen & Huang 2013) and one-dimensional jump-diffusions (Casella & Roberts 2011, Giesecke & Smelov 2013, Gonçalves & Roberts 2014) avoid bias but have a narrower scope. It has proven difficult to extend these exact methods to multivariate processes. For instance, the first exact algorithm for multivariate diffusions has infinite expected run time (Blanchet & Zhang 2017). A less ambitious alternative to exact methods are algorithms that yield unbiased estimators for functionals of the process. Under different sets of assumptions on the coefficients of the process and the functional, Wagner (1989), Rhee & Glynn (2015), Bally & Kohatsu-Higa (2015), Agarwal & Gobet (2017), Andersson & Kohatsu-Higa (2017) and Henry-Labordère, Tan & Touzi (2017) provide such estimators for multivariate diffusions. The jump-diffusion case has not been treated, to our knowledge.

This paper develops an unbiased simulation estimator for functionals of a one-dimensional jump-diffusion process with state-dependent drift, volatility, jump intensity and jump magnitude. We combine a change of measure with an recursive algorithm to extend an unbiased estimator for a diffusion to a process with jumps that arrive with a state-dependent intensity. We exemplify this approach on the “*parametrix method*” estimator for diffusions developed by Bally & Kohatsu-Higa (2015). Here, the diffusion is approximated by an Euler process and the resulting estimator is reweighted to offset the discretization bias. Under the assumption that the intensity function is bounded and has a bounded, continuous derivative, and those inherited from the parametrix method on the diffusion coefficients, we prove that our estimator is unbiased and has finite variance.

While our approach has a somewhat narrower scope than the method of Giesecke & Smelov (2013) for one-dimensional jump-diffusions, our numerical results indicate that it can significantly outperform this exact scheme. Moreover, unlike existing exact algorithms, our approach is extensible to multivariate jump-diffusions. This extension will be pursued in future work, and will illustrate how our change of measure approach harnesses the jumps for alternative unbiased estimators for diffusions, such as that of Henry-Labordère et al. (2017).

The rest of the paper is organized as follows. Section 2 formulates the simulation problem. Section 3 introduces the parametrix method. Section 4 develops and analyzes our approach. Section 5 describes our simulation algorithm and Section 6 supplies the numerical results. An appendix contains the proofs.

## 2. Problem formulation

Fix a time horizon  $T > 0$  and let  $C_b^k(\mathbb{R})$  denote the space of continuous, bounded functions on  $\mathbb{R}$  with  $k$  bounded and continuous derivatives.

**2.1. Jump-diffusion.** We consider an  $\mathbb{R}$ -valued process  $X$  solving the integral SDE

$$(1) \quad X_t = \int_0^t \mu(X_s) ds + \int_0^t \sigma(X_s) dW_s + J_t$$

for  $t \in [0, T]$ , with coefficients  $\mu : \mathbb{R} \rightarrow \mathbb{R}$  and  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$  such that a (unique) weak solution exists for some standard Brownian motion  $W$  and pure jump process  $J$ . Take  $\{\mathcal{F}_t\}_{t \geq 0}$  as the augmented filtration to which both  $W$  and  $J$  (and therefore  $X$ ) are adapted. Specifically, we consider  $J$  of the form,

$$(2) \quad J_t = \sum_{n=0}^{N_t} V_n$$

for a counting process  $N$  with intensity  $\lambda(X)$  for some  $\lambda : \mathbb{R} \rightarrow \mathbb{R}_+$  and  $\{V_n\}_{n \in \mathbb{N}}$ , a sequence of marks (see Brémaud (1981) for definitions). Let  $T_n$  denote the  $n$ th arrival time of  $N$  with  $T_0 = 0$ . Every  $V_n$  is a  $\mathcal{F}_{T_n}$ -measurable random variable with a probability law  $\nu_n$  on  $\mathbb{R}$  that may depend on  $X_{T_n-}$  and  $T_n$ . Observe that,  $X_0 = J_0$  and  $J_0 = V_0$ . Finally, denote by  $\mathbf{P}$  the probability measure on the space on which  $X, W$  and  $J$  are defined, and by  $\mathbf{E}$ , the corresponding expectation.

**Example 2.1.** Taking  $\mu(x) = \mu x$ ,  $\sigma(x) = \sigma x$ ,  $\lambda(x) = \lambda$  and  $V_n = X_{T_n-}(e^{Z_n} - 1)$  with parameters  $\mu, \sigma, \lambda > 0$  and for  $\{Z_n\}_{n \geq 1}$  an i.i.d sequence drawn from the double-exponential distribution, gives the jump-diffusion model of Kou (2002).

**Example 2.2.** Taking  $\mu(x) = (r + \lambda(x))x$ ,  $\sigma(x) = ax^{\beta+1}$  where  $\lambda(x) = b + ca^2x^{2\beta}$  and every  $V_n = -X_{T_n-}$  with parameters  $a, b, r > 0$  while  $c \geq 1/2$  and  $\beta < 0$ , gives the jump-to-default extended CEV model of Carr & Linetsky (2006).

**2.2. Objective.** Our goal is to derive an unbiased estimator of  $\mathbf{E}(f(X_T))$  for Borel  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Precisely, we construct a random variable  $U$  such that

$$(3) \quad \mathbf{E}(f(X_T)) = \mathbf{E}(U).$$

We also insist the estimator  $U$  have finite variance (i.e.,  $\mathbf{Var}(U) < \infty$ ). This leads to a natural Monte Carlo algorithm for estimating  $\mathbf{E}(f(X_T))$ , i.e., average a sufficient number of i.i.d. samples of  $U$  to ensure that the estimated confidence intervals provide a desired level of precision. The finite variance of the estimator ensures that the CLT may be used to construct the desired confidence intervals.

**2.3. Preliminaries.** Define a process  $Y$  on  $[0, T]$  as a (weak) solution to the SDE

$$(4) \quad dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t$$

for coefficients  $\mu, \sigma$  and  $W$  in (1). A (weak) solution of (1) may be constructed from i.i.d. copies  $\{W^n\}_{n \in \mathbb{N}}$  of the  $W$  and a sequence of standard exponential random variables  $\{\mathcal{E}_n\}_{n \in \mathbb{N}}$ . To this end, for the intensity function  $\lambda$ , define  $A$  as

$$(5) \quad A_t = \int_0^t \lambda(Y_s)ds.$$

We take  $(Y^n, A^n)$  to be defined via (4)–(5) but with respect to  $W^n$ . This pair corresponds to the interval  $[T_n, T_{n+1})$  with the right endpoint given by the relation  $T_{n+1} = T_n + (A^n)_{\mathcal{E}_n}^{-1}$ . Now, starting at  $T_0 = 0$  and  $X_{0-} = 0$ , we proceed as,

$$(6) \quad \begin{aligned} X_{T_n} &= X_{T_n-} + V_n \\ X_t &= Y_{t-T_n}^n; \quad t \in (T_n, T_{n+1}), \quad Y_0^n = X_{T_n}. \end{aligned}$$

The  $(N, J)$  may be constructed from  $(T_n, V_n)_{n \in \mathbb{N}}$ . A solution  $X$  that follows the above recipe is càdlàg and enjoys the strong Markov property at each stopping time  $T_n$ . For further detail, we refer the reader to Shkolnik et al. (2019). This construction may be used to simulate the paths of  $X$  but the implied procedure for sampling the jump times is, in general, subject to simulation bias. In general, these errors arise in sampling of the  $Y$ , but also in the inversion of the  $A$ .

Throughout,  $\mathbf{E}_x$  denotes the expectation given  $X$  (or  $Y$ ) is started in  $x \in \mathbb{R}$ .

**Assumption 2.3.** *The function  $\mu$  is  $C_b^1(\mathbb{R})$  and Lipschitz. The function  $\sigma$  is  $C_b^2(\mathbb{R})$  and Lipschitz. There are constants  $b, c > 0$  such that  $\sigma(x) \in [c, b]$  for all  $x \in \mathbb{R}$ .*

**Assumption 2.4.** *The function  $\lambda$  is  $C_b^1(\mathbb{R})$  and is further strictly positive and Lipschitz. Every  $V_n$  has a finite second moment and may be sampled without bias.*

**Assumption 2.5.** *The function  $f$  is bounded.*

Assumption 2.3 guarantees the existence and uniqueness of a weak solution of (4). Assumption 2.4 guarantees further that  $X$  solving (1) is non-explosive. Bias free samples and a finite second moment of  $V_n$  are required for an unbiased estimator that also has finite variance. The requirement that  $f$  be bounded (in conjunction with 2.4) guarantees that a “*parametrix formula*” holds for  $Y$ .

### 3. Parametrix method

Bally & Kohatsu-Higa (2015) develop unbiased estimator for a diffusion via the parametrix method. This estimator relies on a formula that, under Assumption 2.3, states that the diffusion  $Y$  solving (4) and its Euler process  $Y^\pi$  satisfy

$$(7) \quad \mathbf{E}_y(g(Y_t)) = \mathbf{E}_y(g(Y_t^\pi) \Theta_\pi(y, t)) \quad y \in \mathbb{R},$$

for  $t \geq 0$ , any smooth function  $g$  on  $\mathbb{R}$  with compact support and a random weight  $\Theta_\pi(y, t)$  that corrects for the bias due to  $Y^\pi$ . The Euler process  $Y^\pi$  is defined over a random time discretization  $\pi = \{\tau_k\}_{k \in \mathbb{N}}$  with  $0 = \tau_0 \leq \tau_k \leq \tau_{k+1}$ . More precisely, for every  $k \in \mathbb{N}$ , we take  $Y^\pi$  to be the solution of the SDE,

$$(8) \quad dY_t^\pi = \mu(Y_{\tau_k}^\pi) dt + \sigma(Y_{\tau_k}^\pi) dW_t, \quad \tau_k \leq t \leq \tau_{k+1}.$$

We assume the  $\{\eta_k\}_{k \geq 1}$ , with  $\eta_k = \tau_k - \tau_{k-1}$ , is an i.i.d. sequence and that each  $\tau_k$  is the  $k$ th arrival of  $K$ , some counting process independent of  $W$ . Let  $\phi$  be the density function of  $\eta_1$  and let  $\Psi$  be its survival function, i.e.,  $\Psi(x) = \mathbf{P}(\tau_1 > x)$ . Then, given that  $Y^\pi$  starts in  $y \in \mathbb{R}$ , the weight  $\Theta_t(y, t)$  in (7) takes the form

$$(9) \quad \Theta_\pi(y, t) = \left( \prod_{k=1}^{K_t} \frac{\theta_{\eta_k}(Y_{\tau_{k-1}}^\pi, Y_{\tau_k}^\pi)}{\phi(\eta_k)} \right) / \Phi(t - \tau_{K_t}).$$

The function  $\theta_t : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  for  $t \in \mathbb{R}_+$  depends on the coefficients  $\mu$  and  $\sigma$  of the diffusion. We omit these specifics here (but, see Remark 4.1 for details).

When  $K$  is a Poisson process, the variance of the estimator  $g(Y_t^\pi) \Theta_\pi(y, t)$  of  $\mathbf{E}_y(g(Y_t))$  may be infinite. Andersson & Kohatsu-Higa (2017) provide such an example and propose taking  $\eta_1$  to have the Beta distribution. With this choice, the unbiased estimator  $g(Y_t^\pi) \Theta_\pi(y, t)$  has finite variance (see Section 4.1).

To accommodate jumps we extend the class of functions for which (7) holds.

**Lemma 3.1.** *Under Assumption 2.3, for any  $g$  Borel and bounded, (7) holds.*

The proof relies on standard dominated convergence type arguments (see Appendix A). The conditions may be relaxed further but we do pursue this here.

**Remark 3.2.** *Lemma 3.1 generalizes to  $\mathbb{R}^d$  in an obvious way (Andersson & Kohatsu-Higa 2017). In particular, we apply the parametrix formula in  $\mathbb{R}^2$  to the joint process  $(Y, A)$  to accommodate the jumps in  $X$  (see Remark 4.1 and Appendix A).*

## 4. Unbiased estimator

Derivations of the parametrix formula rely on the structure of the infinitesimal generator of a diffusion. They do not extend in a straightforward manner to jump-diffusions. To address this, we apply a change of measure that transforms the  $\mathcal{F}_{T_n}$ -conditional distribution the waiting time (at  $T_n$ ) until  $T_{n+1}$  to be exponentially distributed with rate  $\lambda(X_{T_n})$ . The waiting time may then be sampled exactly. The parametrix formula is then applied to the diffusion on each interval  $[T_n, T_{n+1})$  to form an estimator  $U$  such that (3) holds under our assumptions.

**4.1. Main Results.** For  $(x, t) \in \mathbb{R} \times \mathbb{R}_+$  and  $(Y, A)$  in (4)–(5), define the variables

$$(10) \quad G(y, t) = e^{-A_t} f(Y_t) \quad \text{and} \quad H(y, t) = \lambda(Y_t) e^{-A_t}.$$

Recall the Euler process  $Y^\pi$  of  $Y$  in (8) and define  $A^\pi$  as the Euler process of  $A$  in (5) so that for each  $k \in \mathbb{N}$  given  $A_{\tau_k}^\pi$ , the process  $A^\pi$  satisfies

$$dA_t^\pi = \lambda(Y_{\tau_k}^\pi) dt, \quad \tau_k \leq t < \tau_{k+1},$$

starting from  $A_0^\pi = 0$ . Define the associated random variables  $G^\pi$  and  $H^\pi$  via the Euler processes  $Y^\pi$  and  $A^\pi$  and the weight  $\Theta_\pi$  in (9), i.e., let

$$(11) \quad G^\pi(y, t) = e^{-A_t^\pi} f(Y_t^\pi) \Theta_\pi(y, t) \quad \text{and} \quad H^\pi(y, t) = \lambda(Y_t^\pi) e^{-A_t^\pi} \Theta_\pi(y, t).$$

To compute  $\Theta_\pi(y, t)$ , we take (9) with  $\phi$ , the Beta( $\alpha, \beta$ ) density with parameters  $\alpha \in (\frac{1}{2}, \frac{2}{3})$  and  $\beta = 1$  rescaled to the domain  $[0, t + \epsilon]$  for some  $\epsilon > 0$ . A practical choice for  $\epsilon$  is proportional to  $t$ . Finally, the function  $\theta$  in (9) is given by

$$(12) \quad \begin{aligned} \theta_t(x, y) &= \frac{1}{2} q_t(x, y) - \rho_t(x, y) - (\lambda(y) - \lambda(x)) \\ &\quad \text{where, for } a(x) = \sigma(x)^2 \text{ we have,} \\ \rho_t(x, y) &= \mu'(y) + (\mu(y) - \mu(x)) h_t(x, y) \\ q_t(x, y) &= a''(y) + 2a'(y) h_t(x, y) + (a(y) - a(x)) b_t(x, y) \\ b_t(x, y) &= \frac{h_t(x, y)^2 t a(x) - 1}{t a(x)} \quad \text{and} \quad h_t(x, y) = \frac{t \mu(x) + x - y}{t a(x)}. \end{aligned}$$

**Remark 4.1.** *The parametrix method in our context is being applied to the joint process  $(Y, A)$ . Not surprisingly, setting  $\lambda = 0$  recovers the  $\Theta_\pi(y, t)$  in the parametrix formula (7) for  $Y$  only. This setting corresponds to the case when  $A$  is constant.*

**Remark 4.2.** *The restriction to the parameter  $\alpha$  in the density  $\phi$  to the interval  $(\frac{1}{2}, \frac{2}{3})$  is motivated by Proposition 7.3 in Andersson & Kohatsu-Higa (2017). This ensures that a sufficient number of moments of the parametrix estimator are finite.*

Equipped with the above parametrix variables we state the unbiased estimator  $U$  for  $f(X_T)$  as in (3). Let  $\xi_x$  be exponentially distributed with rate  $\lambda(x)$ , and given the exponential density function  $\ell(y, t) = \lambda(y) e^{-\lambda(y)t}$ , we further define

$$(13) \quad L(y, t) = \frac{H(y, t)}{\ell(y, t)} \quad \text{and} \quad L^\pi(y, t) = \frac{H^\pi(y, t)}{\ell(y, t)}.$$

Take  $\{G_n^\pi\}_{n \in \mathbb{N}}$  and  $\{L_n^\pi\}_{n \in \mathbb{N}}$  to be mutually i.i.d. sequences of the random variables with laws of  $G^\pi$  and  $L^\pi$  respectively. Here, both  $G_n^\pi(y, t)$  and  $H_n^\pi(y, t)$  are independent samples under  $\mathbf{P}$ , conditional on a corresponding sample of  $Y^\pi$  started in  $y$ . Taking  $Y^{n, \pi}$  to denote the sample of  $Y^\pi$  that is associated with the sample  $L_n^\pi(y, t)$ , and  $\{V_n\}_{n \in \mathbb{N}}$  as the marks in (2), we recursively define  $U_n$  as

$$(14) \quad \begin{aligned} U_n(x, t) &= 0 \text{ for all } t < 0, \text{ and otherwise,} \\ U_n(x, t) &= L_n^\pi(x, \xi_x) U_{n+1}(Y_{\xi_x}^{n, \pi} + V_{n+1}, t - \xi_x) + G_n^\pi(x, t). \end{aligned}$$

The estimator  $U$  in (3) is evaluated by sampling  $x = V_0$  and setting

$$(15) \quad U = U_0(x, T).$$

Some remarks are in order. The time  $t$  in  $U_n(x, t)$  corresponds to the time remaining until horizon  $T$ . The subscript  $n$  denotes the number of jumps with  $t$  time remaining. When  $t$  is negative all jumps have been realized, i.e.,  $N_T = n$ . At each  $T_n$ , we have  $X_{T_n} = x$ , also the starting point of  $Y^{n,\pi}$ . Each exponential  $\xi_x$  facilitates the sampling of the jump times without bias. It corresponds to the time increment  $T_n - T_{n-1}$  under a change of measure as explained in Section 4.2.

The next result supplies a certificate of correctness for our estimator.

**Theorem 4.3.** *Under Assumptions 2.3, 2.4 & 2.5, we have  $\mathbf{E}(U) = \mathbf{E}(f(X_T))$ .*

We also guarantee that the Monte Carlo confidence intervals constructed for  $U$  may be trusted. Recall that the samples  $G_n(y, t)$  and  $L_n(y, t)$  of the variables in (11) are independent (conditional on the starting point  $y$ ). This design reduces variance and is key to the proof of the following finite variance result.

**Theorem 4.4.** *Under Assumptions 2.3, 2.4 & 2.5, we have  $\mathbf{Var}(U) < \infty$ .*

The proofs of both theorems are deferred to Appendix A.

**4.2. Estimator Derivation.** The estimator in (14)–(15) consists of two main ingredients. The first is a change of measure that ensures the waiting times in between jumps are (conditionally) exponential. The second is the parametrix formula (7) which provides unbiased estimation for the diffusions in between the jumps.

For  $t \geq 0$  and a change of time  $\gamma(t) = T_{N_t}$ , define a càdlàg process  $Z$  by

$$Z_t = \exp\left(\int_0^t (\lambda(X_s) - \lambda(X_{\gamma(s)})) ds\right) \prod_{n=1}^{N_t} \lambda(X_{T_{n-1}}) / \lambda(X_{T_n}).$$

Giesecke & Shkolnik (2019) Theorem 3.1 guarantees the existence of a probability  $\mathbf{Q}$  via the Radon-Nikodym derivative  $Z_T$  of  $\mathbf{Q}$  with respect to  $\mathbf{P}$ . This result addresses the case when  $N_T$  is bounded. The extension to an infinite but countable number of jumps is facilitated by Kolmogorov's extension theorem. We omit the proof for brevity. Here, we require  $\lambda > 0$  per Assumption 2.4.

The probability  $\mathbf{Q}$  has the property that the  $\mathbf{P}$ -Brownian motion  $W$  in the construction of  $X$  is a  $\mathbf{Q}$ -Brownian motion. However, the intensity of  $N$  under  $\mathbf{Q}$  is now given by  $\lambda(X_\gamma)$ . It follows that the waiting time  $T_{n+1} - T_n$  has the  $\mathcal{F}_{T_n}$ -conditional  $\mathbf{Q}$ -distribution that is exponential with rate  $\lambda(X_{T_n})$ . This is not the case under  $\mathbf{P}$  (see Section 4 of Giesecke & Shkolnik (2019) for further detail).

Denoting by  $\mathbf{E}_x^{\mathbf{Q}}$  the expectation with respect to  $\mathbf{Q}$  with  $X$  started in  $x$ , let

$$u(x, T) = \mathbf{E}_x(f(X_T)) = \mathbf{E}_x^{\mathbf{Q}}(f(X_T)/Z_T).$$

By splitting the right side on the event  $\{T_1 > T\}$  and its complement, conditioning on  $\mathcal{F}_{T_1}$  and applying the strong Markov property of  $X$  at  $T_1$ , we obtain

$$(16) \quad u(x, T) = \mathbf{E}_x^{\mathbf{Q}}(\mathbf{1}_{\{T_1 > T\}} f(X_T)/Z_T) + \mathbf{E}_x^{\mathbf{Q}}(\mathbf{1}_{\{T_1 \leq T\}} u(X_{T_1}, T - T_1)/Z_{T_1}).$$

Note that on the event  $\{T_1 > T\}$ , given  $X$  and  $Y$  start in  $x$ , the laws of the random variables  $f(X_T)/Z_T$  and  $G(x, T)e^{\lambda(x)T}$  coincide. Here, the random variable  $G$  is defined in terms of the diffusion  $(Y, A)$  in (10). Recall that under  $\mathbf{Q}$ , almost surely we have  $\xi_x = T_1$ , which is exponentially distributed with rate  $\lambda(x)$ . It follows that  $\mathbf{Q}(\xi_x > T | \mathcal{F}_0) = e^{-\lambda(x)T}$ , and by the tower property of expectations,

$$(17) \quad \mathbf{E}_x^{\mathbf{Q}}(\mathbf{1}_{\{T_1 > T\}} f(X_T)/Z_T) = \mathbf{E}_x^{\mathbf{Q}}(G(x, T)) = \mathbf{E}_x(G(x, T)).$$

In the second equality, we used the fact that a  $\mathbf{P}$ -Brownian motion is a  $\mathbf{Q}$ -Brownian motion and that  $\xi_x$  has the same exponential distribution under both probabilities (note,  $\xi_x \neq T_1$  under  $\mathbf{P}$ ). Similarly, on the event  $\{T_1 \leq T\}$ , the laws of  $u(X_{T_1}, T - T_1)/Z_{T_1}$  and  $u(Y_{\xi_x} + V_1, T - \xi_x)L(x, \xi_x)$  coincide. Here,  $L(x, t)$  is defined in (13) and corresponds to a likelihood ratio of the probability densities of  $T_1$  under  $\mathbf{P}$  and  $\mathbf{Q}$ . Adopting the convention  $u(\cdot, t) = 0$  for  $t < 0$ , we write

$$(18) \quad \mathbf{E}_x^{\mathbf{Q}}(\mathbf{1}_{\{T_1 \leq T\}} u(X_{T_1}, T - T_1)/Z_{T_1}) = \mathbf{E}_x(u(Y_{\xi_x} + V_1, T - \xi_x)L(x, \xi_x)).$$

Combining (17)–(18) with (16) we obtain the following identity.

$$u(x, T) = \mathbf{E}_x(L(x, \xi_x)u(Y_{\xi_x} + V_1, T - \xi_x)) + \mathbf{E}_x(G(x, T)).$$

With the law of the jump diffusion  $X$  absent from expectations on the right side, we may apply the parametrix formula to the diffusion  $(Y, A)$  to obtain

$$(19) \quad u(x, T) = \mathbf{E}_x(L^\pi(x, \xi_x)u(Y_{\xi_x}^\pi + V_1, T - \xi_x)) + \mathbf{E}_x(G^\pi(x, T)).$$

The estimator  $U(x, T)$  of  $f(X_T)$  in (14) is based on formula (19). The independence of the samples  $G_n^\pi$  and  $L_n^\pi$  corresponds to evaluating the two expectations in (19) separately. The proof of Theorem 4.3 makes the application of the parametrix formula for the joint process  $(Y, A)$  defining  $G$  and  $L$  rigorous.

## 5. Algorithm design

The estimator  $U$  in (14)–(15) leads to a straightforward recursive implementation. However, such an implementation can exhibit poor performance. Here, we design an iterative scheme which is preferable computationally. It is based on unrolling the recursion in (14). This iterative estimator is sampled in Algorithm 1. It is important to note that while the values  $(n, x_n)$  in Algorithm 1 represent those of  $(N, X)$  at the jump-times, they are not samples of  $(N, X)$  at these times. The change of measure and the parametrix formula preclude this interpretation. Algorithm 1 returns a sample of an unbiased estimator of  $f(X_T)$ .

The samples  $G_n^\pi$  and  $L_n^\pi$  in Steps 1 and 4 of the algorithm involve two independent samples of the Euler process  $Y^\pi$ , each with its own time discretization  $\pi$  and weight  $\Theta_\pi$ . Each  $x_n$  is constructed from the sample  $Y^{n, \pi}$  from Step 4.



---

**Algorithm 1** (Generates a sample  $U$  for which  $\mathbf{E}(U) = \mathbf{E}(f(X_T))$ ).

---

0. Initialize  $t = T$ ,  $n = 0$ ,  $x_0 = V_0$ ,  $\Pi = 1$  and  $U = 0$ .
  1. Generate a sample  $G_n^\pi(x_n, t)$  of  $G^\pi(x_n, t)$  in (11) by simulating an Euler process  $(Y^\pi, A^\pi)$  started at  $(x_n, 0)$  for  $\pi$  sampled over the interval  $[0, t]$ .
  2. Update  $U \leftarrow U + \Pi \times G_n^\pi(x_n, t)$ .
  3. Generate an exponential  $\xi_n$  with rate  $\lambda(x_n)$  and **return**  $U$  if  $\xi_n > t$ .
  4. Generate a sample  $L_n^\pi(x_n, \xi_n)$  of  $L^\pi(x_n, \xi_n)$  in (11) by simulating an Euler process  $(Y^\pi, A^\pi)$  started at  $(x_n, 0)$  for  $\pi$  sampled over the interval  $[0, \xi_n]$ .
  5. Update  $\Pi \leftarrow \Pi \times L_n^\pi(x_n, \xi_n)$ .
  6. Sample  $V_{n+1}$  and set  $x_{n+1} = Y_{\xi_n}^{n,\pi} + V_{n+1}$  for the sample  $Y^{n,\pi}$  in Step 4.
  7. Update  $t \leftarrow t - \xi_n$  and  $n \leftarrow n + 1$ . Go to Step 1.
- 

## 6. Numerical results

We provide numerical results to demonstrate the proposed estimator is unbiased. Consider a jump-diffusion  $X$  solving on  $[0, T]$  the SDE

$$(20) \quad dX_t = \kappa(\mu_0 - X_t)dt + \sqrt{\sigma_0 + \sigma_1 X_t}dW_t + dJ_t$$

where  $\kappa, \mu_0, \sigma_0, \sigma_1 > 0$  are constant parameters, and  $J$  a marked point process with mark sequence  $\{V_n\}$  and intensity  $\lambda(X) = \lambda_0 + \lambda_1 X$  for constants  $\lambda_1 \geq 0$  and  $\lambda_0 \geq \lambda_1 \sigma_0 / \sigma_1$ . The Feller condition applied to the process  $\sigma_0 + \sigma_1 X$  guarantees that  $\lambda(X) \geq 0$  almost surely, as required. Model (20) is an example of an affine jump-diffusion, a class of models widely utilized in financial applications (Duffie, Pan & Singleton (2000); Errais, Giesecke & Goldberg (2010)).

We estimate the expected (undiscounted) price  $\mathbf{E}((X_T - \mathcal{K})_+)$  of an interest rate cap with strike  $\mathcal{K}$  and maturity  $T$ , written on the short rate  $X$  (here,  $\mathbf{P}$  is the risk-neutral measure). For comparison, we also supply results for  $\mathbf{E}(X_T)$ .

Model (20) and the functions  $f(x) = (x - \mathcal{K})_+$  and  $f(x) = x$  above violate several of the assumptions of Theorems 4.3 and 4.4. Note that the drift  $\mu(x) = \kappa(\mu_0 - x)$  and the volatility  $\sigma(x) = \sqrt{\sigma_0 + \sigma_1 x}$  are not bounded. Furthermore, the variance function  $\sigma^2$  is not uniformly elliptic, and the intensity function  $\lambda$  is not bounded. All of these are in violation of Assumptions 2.3–2.5. However, the numerical results below suggest that our estimator remains unbiased even when many of the required conditions do not hold. This indicates a broader applicability of our scheme and points to extensions in future work. An

additional motivation for choosing this setting is that it is tractable (yet, nontrivial). This allows us to test if our algorithm indeed outputs unbiased samples.

**Table 1.** *Numerical simulation results for  $\mathbf{E}(X_T - \mathcal{K})_+$ . The Error and Variance values represent sample quantities over  $m$  Monte Carlo trials.*

$m$	Error $\times 10^{-3}$	Variance $\times 10^3$	99% CI $\times 10^{-3}$	Time (sec)
$2^{26}$	6.86	1.32404	11.5	339.381
$2^{28}$	5.66	1.86632	6.8	1357.52
$2^{30}$	0.21	1.68057	3.2	5430.10
$2^{32}$	0.06	1.54682	1.5	21720.4
$2^{34}$	0.31	2.04001	0.9	86881.6

**Table 2.** *Numerical simulation results for  $\mathbf{E}(X_T)$ . The Error and Variance values represent sample quantities over  $m$  Monte Carlo trials.*

$m$	Error $\times 10^{-3}$	Variance $\times 10^3$	99% CI $\times 10^{-3}$	Time (sec)
$2^{26}$	8.42992	4.40456	20.9	338.180
$2^{28}$	9.41392	6.78829	13.0	1352.72
$2^{30}$	1.60108	6.13765	6.2	5410.88
$2^{32}$	0.24092	5.52924	2.9	21643.5
$2^{34}$	0.97092	6.87710	1.6	86574.2

To illustrate the performance of our estimator, we report the sample error, given by  $|\mathbf{U}_m - \mathbf{E}(f(X_T))|$ , where  $\mathbf{U}_m$  is the unbiased estimate of  $\mathbf{E}(f(X_T))$  generated using  $m$  Monte Carlo trials. The exact value  $\mathbf{E}(f(X_T))$  is computed semi-analytically (see below). We also include the sample variance of the estimate  $\mathbf{U}_m$  along with 99% confidence intervals in Tables 1 and 2. We report the estimated values for an increasing number of trials, the largest being,  $2^{34} \geq 17 \times 10^9$ , to test the behavior over as many scenarios as possible. We observe that the sample error stays within the 99% confidence interval for each number of trials used. While the variance estimates are large, no indication of outliers was detected.

The implementation of the scheme follows Algorithm 1. To evaluate the weight  $\Theta_\pi$  for the samples  $G_n^\pi$  and  $L_n^\pi$  with time discretization  $\pi$  over, say  $[0, \Delta]$ , we take  $\phi(x) = \alpha(2\Delta)^{-\alpha}/(x^{1-\alpha})$  defined on the domain  $[0, 2\Delta]$  with  $\alpha = 0.5$ . The function  $\theta_t(x, y)$  in (12) for the model in (20) takes the form

$$\begin{aligned} \theta_t(x, y) = & \frac{\sigma_1}{2}(y-x) \left( \frac{y-x-\kappa(\mu_0-x)t}{t(\sigma_0+\sigma_1x)} \right)^2 - \sigma_1 \left( \frac{y-x-\kappa(\mu_0-x)t}{t(\sigma_0+\sigma_1x)} \right) \\ & - \frac{\kappa(x-y)(y-x-\kappa(\mu_0-x)t)}{t(\sigma_0+\sigma_1x)} + \kappa - \lambda_1(y-x). \end{aligned}$$

The parameters for  $X$  are selected as  $\sigma_1 = 0.2$  and  $\kappa = \mu_0 = \sigma_0 = \lambda_0 = \lambda_1 = 1$ . We start  $X$  at 2 and draw each  $V_n$  for  $n \geq 1$  uniformly from  $\{0.5, 1.0\}$ . The maturity and strike price are set at  $T = 1$  and  $\mathcal{K} = 3$  for all simulations.

We obtain the exact value for  $\mathbf{E}_x((X_T - \mathcal{K})_+)$  using Fourier inversion. This is possible for model (20) as the transform  $\psi_x(u) = \mathbf{E}_x(e^{uX_T})$  may be computed semi-analytically. Indeed, we have  $\psi_x(u) = e^{a(0)+b(0)x}$  for  $u \in \mathbb{C}$  and  $(a, b)$  solve

$$\begin{aligned} \dot{a} &= -\kappa\mu_0b - \frac{1}{2}\sigma_0b^2 - q(b)\lambda_0 & a(T) &= 0, \\ \dot{b} &= \mu_0b - \frac{1}{2}\sigma_1b^2 - q(b)\lambda_1 & b(T) &= u, \end{aligned}$$

a system of ODEs on  $[0, T]$  where  $q(b) = \frac{1}{2}(e^{b/2} + e^b - 2)$ . Then,

$$\mathbf{E}_x((X_T - \mathcal{K})_+) = \frac{e^{-z\mathcal{K}}}{\pi} \int_0^\infty \operatorname{Re} \left\{ e^{-iu\mathcal{K}} \frac{\psi_x(u - zi)}{(z + ui)^2} \right\} du$$

for any  $z > 0$  for which the above integral converges. We take  $z = 0.4$ . For more details we refer the reader to Duffie et al. (2000) and Lee (2004).

To compute the exact value for  $\mathbf{E}_x(X_T)$ , let  $\mu_x(t) = \mathbf{E}_x(X_t)$  which using (20) may be shown to satisfy the following ODE for  $v = \mathbf{E}(V_1)$ .

$$\dot{\mu}_x = (\kappa\mu_0 + v\lambda_0) + (v\lambda_1 - \kappa)\mu_x \quad \mu_x(0) = x.$$

We observe that the algorithm is extremely fast in practice because relatively few (random) discretization points are sampled for the Euler processes involved. This is unlike traditional Euler methods that require a fine discretization to maintain accuracy. Despite a relatively high sample variance, in our simulations, the scheme significantly outperforms the one proposed in Giesecke & Smelov (2013). Utilizing a single core of a Intel(R) Xeon(R) CPU (E7-8890 v3 at 2.50GHz), Algorithm 1 takes roughly 24 hours to achieve a 99% confidence interval of length 0.002. It would take 130 times longer for the algorithm of Giesecke & Smelov (2013) to achieve the same result. If an effective variance reduction technique could be implemented, our approach could be made even more efficient.

## A. Proofs

The proof of Lemma 3.1 below accomodates the  $\mathbb{R}^d$  case. As discussed in Remark 4.1, we apply the parametrix formula in  $\mathbb{R}^2$  to the process  $(Y, A)$ . It should be noted that  $(Y, A)$  does not satisfy Assumption 2.3, and we make the proper modifications in the proofs to address this difficulty.

**PROOF OF LEMMA 3.1.** By Corollary 4.2 and Proposition 7.3 of Andersson & Kohatsu-Higa (2017), (take  $\Phi(x) = 1$  and  $p = 1$  therein),  $\mathbf{E}_y(|\Theta_\pi(y, t)|) < \infty$ .

For  $M \in \mathbb{N}$  and  $g$  Borel and bounded on  $\mathbb{R}^d$ , consider  $g_M = g \mathbf{1}_{[-M, M]}$ . Then, there exists a sequence of compactly supported, smooth functions  $\{g_M^i\}_{i=0}^\infty$  such that  $g_M^i \rightarrow g_M$  in  $L_1(\mathbb{R}^d)$  (Tao 2011, Theorem 1.3.20). Furthermore, there is a subsequence  $\{g_M^{i_k}\}_{k \in \mathbb{N}}$  that is uniformly bounded (as  $g$  is bounded) and converges pointwise almost everywhere to  $g_M(y)$  (Folland 2007, Corollary 2.32)). Consequently, the parametrix formula holds for each  $g_M^{i_k}$ , i.e.,  $\mathbf{E}_y(g_M^{i_k}(Y_t)) = \mathbf{E}_y(g_M^{i_k}(Y_t^\pi) \Theta_\pi(y, t))$ . Since  $g$  is bounded, the law of  $Y_t$  is non-atomic and therefore no point has positive probability mass, and  $\mathbf{E}_y(|\Theta_\pi(x, t)|) < \infty$ , applying dominated convergence with  $k \uparrow \infty$  yields  $\mathbf{E}_y(g_M(Y_t)) = \mathbf{E}_y(g_M(Y_t^\pi) \Theta_\pi(y, t))$ . Applying dominated convergence again with  $M \uparrow \infty$  yields the desired result.  $\blacksquare$

Before proceeding to the proofs of Theorems 4.3 and 4.4 we make some preliminary definitions. The estimator  $U = U_0(X_0, T)$  constructed in (14)–(15) has a  $\mathbf{P}$ -law that coincides with the  $\mathbf{Q}$ -law. It implies the existence of a counting process  $\hat{N}$  that is constructed from the exponential waiting times generated at each arrival (e.g., see Step 3 of Algorithm 1). This process should not be confused with the original counting process  $N$ , the law of which depends on the probability measure. The auxiliary process  $\hat{N}$  represents  $N$  and is involved in the left side of

$$(21) \quad \mathbf{E}(U) = \mathbf{E}(f(X_T)).$$

Define  $\mathcal{B}_n(t) = \{\hat{N}_t = n\}$  and note that this event satisfies

$$(22) \quad \mathbf{1}_{\{T_1 \leq T\}} \mathcal{B}_{n+1}(t) = \mathbf{1}_{\{T_1 \leq T\}} \mathcal{B}_n(t - T_1).$$

We proceed to prove the main results of the paper.

**PROOF OF THEOREM 4.3.** To prove (21) it suffices, for to show that

$$(23) \quad \mathbf{E}_x^{\mathbf{Q}}(\mathbf{1}_{\mathcal{A}_n(t)} U_0(x, t)) = \mathbf{E}_x(\mathbf{1}_{\{N_t \leq n\}} f(X_t)), \quad \mathcal{A}_n(t) = \cup_{k \leq n} \mathcal{B}_k(t), \quad t \leq T.$$

Then, taking  $n \uparrow \infty$ , so that  $\mathbf{1}_{\mathcal{A}_n(t)} \uparrow 1$  almost surely,  $\mathbf{E}_x^{\mathbf{Q}}(U_0(x, t)) = \mathbf{E}_x(f(X_t))$  for  $t \leq T$ . The right side of (23) follows by the bounded convergence theorem ( $f$  is bounded). The left side of (23) follows by dominated convergence with

$|\mathbf{1}_{\mathcal{A}_n(t)} U_0(x, t)| \leq |U_0(x, t)|$  since  $\mathbf{E}_x^Q(|U_0(x, t)|) < \infty$  which holds by the fact that  $\mathbf{E}_x^Q(U_0(x, t)^2) < \infty$  per Theorem 4.4 and Jensen's inequality.

The proof follows induction on  $n$ . To this end, we assume (23) holds for some  $n$  and prove (23) for all  $n + 1$ . Using the definition in (14) we have,

$$\begin{aligned} \mathbf{E}_x^Q(\mathbf{1}_{\mathcal{A}_{n+1}(t)} U_0(x, t)) &= \mathbf{E}_x^Q(\mathbf{1}_{\mathcal{A}_{n+1}(t)} L_0^\pi(x, \xi_x) U_1(Y_{\xi_x}^{0,\pi} + V_1, t - \xi_x)) \\ &\quad + \mathbf{E}_x^Q(\mathbf{1}_{\mathcal{A}_{n+1}(t)} G_0^\pi(x, t)) \end{aligned}$$

We treat the two terms in the sum separately.

The following steps follow from the arguments in Section 4.2. Here, we address the application of the parametrix formula.

$$\begin{aligned} \mathbf{E}_x^Q(\mathbf{1}_{\mathcal{A}_{n+1}(t)} G_0^\pi(x, t)) &= \mathbf{E}_x^Q(\mathbf{1}_{\mathcal{A}_{n+1}(t)} \Theta_\pi(x, t) e^{-A_t^\pi} f(Y_t^\pi)) \\ &= \mathbf{E}_x^Q(\mathbf{1}_{\mathcal{A}_{n+1}(t)} \mathbf{1}_{\{T_1 > t\}} \Theta_\pi(x, t) e^{-A_t^\pi + t\lambda(x)} f(Y_t^\pi)) \\ &= \mathbf{E}_x(\Theta_\pi(x, t) e^{-A_t^\pi} f(Y_t^\pi)) \\ &= \mathbf{E}_x(e^{-A_t} f(Y_t)) \\ &= \mathbf{E}_x(\mathbf{1}_{\{T_1 > t\}} f(X_t)). \end{aligned}$$

In the above steps, we applied the parametrix formula to  $(Y, A)$  starting in  $(x, 0)$  and the map  $(z_1, z_2) = e^{-(z_2 \vee 0)} f(z_1)$ , which is bounded, ensures Lemma 3.1 is applicable. It should be noted that the diffusion  $(Y, A)$  is degenerate which would in theory violate the conditions on its diffusion matrix. This is easily remedied by introducing a second Brownian motion  $W'$  independent of  $W$  and considering  $(Y, A + cW')$  for any  $c > 0$ . Establishing the formula above for this modified process and using bounded convergence with  $c \downarrow 0$  yields the result.

The second term relies on the inductive hypothesis. We define,

$$u_n(x, s) = \mathbf{E}_x^Q(\mathbf{1}_{\mathcal{A}_n(s)} U_1(x, s)) = \mathbf{E}_x^Q(\mathbf{1}_{\mathcal{A}_n(s)} U_0(x, s)) = \mathbf{E}_x(\mathbf{1}_{\{N_t \leq n\}} f(X_s)).$$

Noting that  $\mathbf{1}_{\mathcal{A}_{n+1}(t)} \mathbf{1}_{\{T_1 \leq t\}} = \mathbf{1}_{\{T_1 \leq t\}} \mathbf{1}_{\mathcal{A}_n(t-T_1)}$  upon summing (22), again adopting the convention  $u_n(x, s) = 0$  for  $s < 0$ , and applying the arguments used in the derivation of (18) in Section 4.2, for  $V_1^\pi = Y_{\xi_x}^{0,\pi} + V_1$ , leads to

$$\begin{aligned} \mathbf{E}_x^Q(\mathbf{1}_{\mathcal{A}_{n+1}(t)} L_0^\pi(x, \xi_x) U_1(V_1^\pi, t - \xi_x)) &= \mathbf{E}_x(L_0^\pi(x, t) u_n(V_1^\pi, t - \xi_x)) \\ &= \mathbf{E}_x^Q(L_0(x, t) \mathbf{1}_{\{T_1 \leq t\}} u_n(X_{T_1}, t - T_1)) \\ &= \mathbf{E}_x(\mathbf{1}_{\{T_1 \leq t\}} u_n(X_{T_1}, t - T_1)) \\ &= \mathbf{E}_x(\mathbf{1}_{\{T_1 \leq t\}} \mathbf{1}_{\{N_t \leq n\}} f(X_t)). \end{aligned}$$

The parametrix formula is applied above to  $(Y, A)$  with  $u_n$  is bounded and the  $L_0^\pi$  treated analogously to  $G_0^\pi$  above. Adding the two terms completes the proof. ■

PROOF OF THEOREM 4.4. It suffices for a finite  $C \geq 2$  and  $\gamma \in (0, 1)$  to show

$$(24) \quad \sup_{x \in \mathbb{R}} \mathbf{E}_x^Q(\mathbf{1}_{\mathcal{B}_n(t)} U_0(x, t)^2) \leq C \gamma^n \quad t \leq T.$$

Once (24) holds, we obtain  $\mathbf{E}(U^2) \leq \sup_{x \in \mathbb{R}} \sum_{n \in \mathbb{N}} \mathbf{E}_x^{\mathbf{Q}}(\mathbf{1}_{\mathcal{B}_n(t)} U_0(x, T)^2)$  where the sum and the expectation are exchanged by applying monotone convergence.

Proposition 7.3 in Andersson & Kohatsu-Higa (2017) guarantees there exists a constant  $B \geq 2$  such that for  $G^\pi$  and  $L^\pi$  in (11),

$$\sup_{x \in \mathbb{R}} \mathbf{E}_x^{\mathbf{Q}}(G^\pi(x, t)^4) \leq B^2 \quad \text{and} \quad \sup_{x \in \mathbb{R}} \mathbf{E}_x^{\mathbf{Q}}(L^\pi(x, \xi_x)^4) \leq B^2.$$

An examination of the proof of the result in Andersson & Kohatsu-Higa (2017) reveals that the bound is independent of the starting point  $x \in \mathbb{R}$ . Here, we crucially rely on the value  $\alpha$  in the range specified in Remark 4.2 and that under Assumptions 2.4 and 2.5 both  $\lambda$  and  $f$  in  $G^\pi$  and  $L^\pi$  are bounded.

We proceed by induction on  $n \in \mathbb{N}$ . The base case follows by noting that  $\mathbf{E}_x^{\mathbf{Q}}(\mathbf{1}_{\mathcal{B}_0(t)} U_0(x, t)^2) = \mathbf{E}_x^{\mathbf{Q}}(G_0^\pi(x, t)^2) \leq B$  by Jensen's inequality.

We take  $C \geq B$ . Assuming that (24) holds, it suffices to exhibit the same bound for  $n + 1$  and for  $n$  sufficiently large (taking  $C$  large enough for other  $n$ ). To this end, for  $t \leq T$ , consider  $\{p_n(t)\}_{n \in \mathbb{N}}$  defined by  $p_n^2(t) = \mathbf{Q}(\mathcal{B}_n(t))$ . Assumption 2.4 (bounded  $\lambda$ ) guarantees that for  $n$  sufficiently large,

$$B^2 C p_n(t) \leq \gamma^n / 2.$$

Expanding  $\mathbf{1}_{\mathcal{B}_{n+1}(t)} U_0(x, t)^2$  using the definition (14) yields a sum of three terms. Using (22) and applying Cauchy-Schwarz and the strong Markov property at  $\xi_x = T_1$ , yields the following bounds for these terms.

$$\begin{aligned} \mathbf{E}_x^{\mathbf{Q}}(\mathbf{1}_{\mathcal{B}_{n+1}(t)} G_0^\pi(x, t)^2) &\leq p_{n+1}(t) B^2 C^2 \\ \mathbf{E}_x^{\mathbf{Q}}(\mathbf{1}_{\mathcal{B}_{n+1}(t)} G_0^\pi(x, t) L_0^\pi(x, \xi_x) U_0(Y_{\xi_x}^{0, \pi} + V_1, t - \xi_x)) &\leq p_{n+1}(t) B^2 C^2 \gamma^n / 2 \\ \mathbf{E}_x^{\mathbf{Q}}(\mathbf{1}_{\mathcal{B}_{n+1}(t)} L_0^\pi(x, \xi_x)^2 U_0(Y_{\xi_x}^{0, \pi} + V_1, t - \xi_x)^2) &\leq p_{n+1}(t) B^2 C^2 \gamma^n / 2 \end{aligned}$$

Combining these bounds yields the inductive step, concluding the proof, i.e.,

$$\sup_{x \in \mathbb{R}} \mathbf{E}_x^{\mathbf{Q}}(\mathbf{1}_{\mathcal{B}_{n+1}(t)} U_0(x, t)^2) \leq B^2 C^2 p_{n+1}(t) (1 + \gamma^n) \leq C \gamma^{n+1}.$$

■

## References

- Agarwal, A. & Gobet, E. (2017). Finite variance unbiased estimation of stochastic differential equations, *2017 Winter Simulation Conference*, INFORMS.
- Andersson, P. & Kohatsu-Higa, A. (2017). Unbiased simulation of stochastic differential equations using parametrix expansions, *Bernoulli* **23**(3): 2028–2057.
- Bally, V. & Kohatsu-Higa, A. (2015). A probabilistic interpretation of the parametrix method, *The Annals of Applied Probability* **25**(6): 3095–3138.

- Beskos, A. & Roberts, G. O. (2005). Exact simulation of diffusions, *The Annals of Applied Probability* **15**(4): 2422–2444.
- Blanchet, J. & Zhang, F. (2017). Exact simulation for multivariate ito diffusions, arXiv:1706.05124.
- Brémaud, P. (1981). *Point Processes and Queues – Martingale Dynamics*, Springer-Verlag, New York.
- Carr, P. & Linetsky, V. (2006). A jump to default extended cev model: an application of bessel processes, *Finance and Stochastics* **10**(3): 303–330.
- Casella, B. & Roberts, G. O. (2011). Exact simulation of jump-diffusion processes with monte carlo applications, *Methodology and Computing in Applied Probability* **13**(3): 449–473.
- Chen, N. & Huang, Z. (2013). Localization and exact simulation of brownian motion- driven stochastic differential equations, *Mathematics of Operations Research* **38**(3): 591–616.
- Duffie, D., Pan, J. & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* **68**(6): 1343–1376.
- Errais, E., Giesecke, K. & Goldberg, L. (2010). Affine point processes and portfolio credit risk, *SIAM Journal of Financial Mathematics* **15**: 642–665.
- Folland, G. (2007). *Real Analysis: Modern Techniques and Their Applications*, Wiley.
- Giesecke, K. & Shkolnik, A. (2019). Reducing bias in event time simulations via measure changes. *Mathematics of Operations Research*, forthcoming.
- Giesecke, K. & Smelov, D. (2013). Exact sampling of jump diffusions, *Operations Research* **61**(4): 894–907.
- Gonçalves, F. B. & Roberts, G. O. (2014). Exact simulation problems for jump-diffusions, *Methodology and Computing in Applied Probability* **16**(4): 907–930.
- Henry-Labordère, P., Tan, X. & Touzi, N. (2017). Unbiased simulation of stochastic differential equations, *The Annals of Applied Probability* **27**(6): 3305–3341.
- Kou, S. (2002). A jump-diffusion model for option pricing, *Management Science* **48**(8): 1086–1101.
- Lee, R. W. (2004). Option pricing by transform methods: Extensions, unification, and error control, *Journal of Computational Finance* **7**(3): 51–86.

- Rhee, C.-H. & Glynn, P. (2015). Unbiased estimation with square root convergence for sde models, *Operations Research* **63**(5): 1026–1043.
- Shkolnik, A., Giesecke, K., Teng, G. & Wei, Y. (2019). Numerical solution of jump-diffusion SDEs. *Operations Research*, *forthcoming*.
- Tao, T. (2011). *An Introduction to Measure Theory*, American Mathematical Society.
- Wagner, W. (1989). Unbiased monte carlo estimators for functionals of weak solutions of stochastic differential equations, *Stochastics An International Journal of Probability and Stochastic Processes* **28**(1): 1–20.