Unbiased Simulation Estimators for Path Intergals of Diffusions

Guanting Chen* Alex Shkolnik† Kay Giesecke‡

October 5, 2020

Abstract

We develop and analyze Monte Carlo simulation estimators for path integrals of a multivariate diffusion with a general state-dependent drift and volatility. We prove that our estimators are unbiased and have finite variance by extending the regularity conditions of the parametrix method. The performance of our estimators is illustrated on numerical examples that highlight some applied problems for which our estimators apply.

*Institute for Computational and Mathematical Engineering, Stanford University, CA 94305 EMAIL: guanting@stanford.edu.

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

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

1. Introduction

Suppose $Y \in \mathbb{R}^d$ solves the stochastic differential equation (SDE)

$$\mathrm{d}Y_t = \mu(Y_t)\,\mathrm{d}t + \sigma(Y_t)\,\mathrm{d}W_t \tag{1}$$

for a m-dimensional (standard) Brownian motion W and suitable drift and diffusion coefficients $\mu : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$. Exact simulation of multivariate diffusions ($d \geq 2$) with general coefficients is known to be a challenging problem (Blanchet & Zhang 2017). Discretization methods are widely applicable however. For example, consider the Euler scheme

$$\mathrm{d}Y_t^\pi = \mu(Y_{t_k}^\pi)\,\mathrm{d}t + \sigma(Y_{t_k}^\pi)\,\mathrm{d}W_t \qquad t_k \leq t < t_{k+1} \tag{2}$$

defined over a finite, increasing sequence $\{t_k\} \subseteq [0, T]$ of discretization points satisfying $0 = t_0 < T = \max_k t_k$. It is well known that the Euler process Y^π may be used to estimate $\mathrm{E}(f(Y_T))$ with order-one (weak) convergence (i.e., $|\mathrm{E}(f(Y_T)) - \mathrm{E}(f(Y_T^\pi))| \leq C_{T,f} \max_k |t_k - t_{k-1}|$) under various assumptions on the function f and the coefficients. The leading constant $C_{T,f}$, the rate of convergence, or both, may be unsuitably large for applications that involve small estimates or model sensitivities. For these (and other) reasons, unbiased estimators based on the representation

$$\mathrm{E}(f(Y_T)) = e^{\lambda T}\mathrm{E}\left(f(Y_T^\pi)\prod_{k=1}^{N_T} \lambda^{-1}\vartheta_{\tau_k-\tau_{k-1}}(Y_{\tau_{k-1}}^\pi, Y_{\tau_k}^\pi)\right) \tag{3}$$

have recently been developed. In (3), the Y^π is defined analogously to (2), but use random discretization points $\{\tau_k\}$ in lieu of $\{t_k\}$. Precisely, $\{\tau_k\}$ form the arrival times of a counting process N, which in (3) is taken to be Poisson of rate $\lambda > 0$. The weight function ϑ is defined via the coefficients μ, σ and corrects for the bias generated by the Euler iterates $\{Y_{\tau_k}^\pi\}$. Approaches based on (3) are known as parametrix methods and lead to unbiased estimates of $\mathrm{E}(f(Y_T))$. In contrast, the plain Euler estimator $f(Y_T^\pi)$ is biased.

The parametrix method has long been applied in the study of solutions to partial differential equations, but only recently have probabilistic representations of the form (3) been proposed for diffusion simulation (Bally & Kohatsu-Higa 2015). The scheme has tremendous potential, not only in terms of accuracy, but also in terms of run-time. For instance, the rate λ of the Poisson process generating the points $\{\tau_k\}$ can be relatively small, resulting in relatively few iterates $Y_{\tau_k}^\pi$ that are needed. A fine discretization (i.e., more $Y_{t_k}^\pi$) is required to control the bias of the plain Euler method. Parametrix estimators do suffer from a large variance, but this drawback may be mitigated with a greater number of parallelized Monte Carlo trials.

Furthermore, significant progress has been made on controlling the variance of a parametrix estimator (Andersson & Kohatsu-Higa 2017).

A genuine shortcoming of the parametrix method, however, lies in the smoothness and compact support assumptions on the objective function f that are required for (3) to hold. These are often too restrictive for applications (however, see Doumbia, Oudjane & Warin (2017) for a relaxation to Lipschitz continuous f). Moreover, while unbiased estimators for expectations of functions of a skeleton $(Y_{t_1}, \dots, Y_{t_n})$ have been designed (Henry-Labordère, Tan & Touzi 2017), to our knowledge, no unbiased estimators for path integrals $\int_0^T \Lambda(Y_t)\,\mathrm{d}t$ given a function $\Lambda : \mathbb{R}^d \to \mathbb{R}$ appear in the literature. Using recently developed Gaussian upper bounds and existence results for the density of a diffusion with a Lipschitz drift (see Menozzi, Pesce & Zhang (2020)), we are able to treat several important extensions.

- We extend the parametrix method to path integrals by generalizing formula (3) to expectations of the form $\mathrm{E}\big(\int_0^T \Lambda(Y_t)\,\mathrm{d}t\big)$ for a Lipschitz continuous function Λ and diffusion Y in (1). A direct application of the parametrix formula to the joint process $\big(Y_t, \int_0^t \Lambda(Y_s)\,\mathrm{d}s\big)_{t\geq 0}$ leads to assumptions on Λ inherited from the parametrix method. Namely, the drift Λ must be twice continuously differentiable and bounded. Requiring Λ be only Lipschitz continuous is a significant relaxation.

- We extend the parametrix method by relaxing f to have exponential growth and to expectations of the form $\mathrm{E}\big(e^{-\int_0^T \Lambda(Y_t)\,\mathrm{d}t} f(Y_T)\big)$. The latter extension treats a wide array of applications in finance where the term $e^{-\int_0^T \Lambda(Y_t)\,\mathrm{d}t}$ serves as a discount factor. Such expressions are also related to killed diffusions as the discount factor may be replaced by the indicator $1_{\{\eta > T\}}$ where η is a "killing" time of Y of intensity $\Lambda(Y)$ (see Collin-Dufresne, Goldstein & Hugonnier (2004) for example).

We develop these extensions without sacrificing the variance properties of our estimators. In particular, a finite variance of the estimator in (3) is achieved by replacing the Poisson process N by an alternative counting process (e.g., one with Beta distributed interarrival times as in Andersson & Kohatsu-Higa (2017)). The same approach extends to our estimators.

The paper is structured as follows. Section 2 describes the construction of the path integral and the killed diffusion estimator. It also supplies the theorems guaranteeing the unbiasedness and the finite variance of both parametrix estimators. Section 3 establishes the proofs of the theoretical results. Section 5 supports our findings with numerical experiments. The Appendix collects the auxiliary results that are required by our proofs.

2. Main results

Fix $T \in (0,\infty)$. Consider the joint process $X = (Y, Z)$ where Y is the diffusion in (1), and Z is a one-dimensional diffusion with drift coefficient Λ and a constant coefficient of diffusion. More precisely, let (Y, Z) solve

$$\begin{aligned} Y_t &= y_0 + \int_0^t \mu(Y_s)\,\mathrm{d}s + \int_0^t \sigma(Y_s)\,\mathrm{d}W_s, \\ Z_t &= \int_0^t \Lambda(Y_s)\,\mathrm{d}s + \nu B_t, \end{aligned} \tag{4}$$

where B is a one-dimensional Brownian motion independent of W, and a constant $\nu > 0$. The nonrandom $y_0 \in \mathbb{R}^d$ denotes the starting point of Y.

Let $C_b^k(\mathscr{D})$ be the space of functions on $\mathscr{D}$ with k bounded derivatives.

Assumption 2.1. *There exist constants $a_1, a_2 > 0$ such that the $a = \sigma\sigma^\top$ satisfies $a_1 I \preceq a(y) \preceq a_2 I$ for all $y \in \mathbb{R}^d$ (and I, the $d \times d$ identity matrix). Moreover, $\mu \in C_b^1(\mathbb{R}^d)$, $\sigma \in C_b^2(\mathbb{R}^d)$ and $\Lambda : \mathbb{R}^d \to \mathbb{R}$ is Lipchitz continuous.*

Weak solutions $X = (Y, Z)$ to the integral SDE (4) exists under the conditions on the coefficient in Assumption 2.1 (Krylov 1974).

We say $G : \mathscr{D} \to \mathbb{R}$ ($|\cdot|$, a Euclidean norm on $\mathscr{D}$) has exponential growth if there are constants $c, C > 0$ with $|G(u)| \le Ce^{c|u|}$ for all $u \in \mathscr{D}$.

Assumption 2.2. *$f : \mathbb{R}^d \to \mathbb{R}$ is measurable and of exponential growth.*

Our parametrix estimators use the following Euler processes,

$$\begin{aligned} \mathrm{d}Y_t^\pi &= \mu(Y_{\tau_k}^\pi)\,\mathrm{d}t + \sigma(Y_{\tau_k}^\pi)\,\mathrm{d}W_t \\ \mathrm{d}Z_t^\pi &= \Lambda(Y_{\tau_k}^\pi)\,\mathrm{d}t + \nu\,\mathrm{d}B_t \end{aligned} \qquad \tau_k \le t < \tau_{k+1}, \tag{5}$$

for an increasing sequence of arrival times $\{\tau_k\}_{k\in\mathbb{N}}$ of a counting process N that is taken to be independent of W and B. We denote by X^π, the joint Euler process (Y^π, Z^π), which starts in $(y_0, 0) \in \mathbb{R}^{d+1}$. When N is nonexplosive (i.e. $N_t < \infty$ almost surely for all $t \ge 0$), the associated discretization spacings π have the form $0 = \tau_0 < \tau_1 < \cdots < \tau_{N_T} < T$ almost surely.

For $x_1 = (y_1, z_1) \in \mathbb{R}^{d+1}$ and $x_2 = (y_2, z_2) \in \mathbb{R}^{d+1}$, define

$$\begin{aligned} \theta_t(x_1, x_2) &= \vartheta_t(y_1, y_2) + (\Lambda(y_2) - \Lambda(y_1))\Big(\frac{z_2 - z_1 - \Lambda(y_1)t}{t\nu^2}\Big) \\ \vartheta_t(y_1, y_2) &= \frac{1}{2}\textstyle\sum_{i,j}\vartheta_t^{i,j}(y_1, y_2) - \sum_i \rho_t^i(y_1, y_2) \\ \vartheta_t^{i,j}(y_1, y_2) &= \partial_{i,j}^2 a^{i,j}(y_2) + \partial_j a^{i,j}(y_2)h_t^i(y_1, y_2) \\ &\quad + \partial_i a^{i,j}(y_2)h_t^j(y_1, y_2) + (a^{i,j}(y_2) - a^{i,j}(y_1))h_t^{i,j}(y_1, y_2) \\ \rho_t^i(y_1, y_2) &= \partial_i \mu^i(y_2) + (\mu^i(y_2) - \mu^i(y_1))h_t^i(y_1, y_2) \\ h_t(y_1, y_2) &= H_{ta(y_1)}(y_2 - y_1 - t\mu(y_1)) \end{aligned} \tag{6}$$

where H denotes the Hermite polynomials. For any matrix m we have 1st-order polynomials $H^i_{\mathrm{m}}(x) = -(\mathrm{m}^{-1}x)_i$ and 2nd-order polynomials $H^{ij}(x) = (\mathrm{m}^{-1}x)_i(\mathrm{m}^{-1}x)_j - (\mathrm{m}^{-1})_{ij}$. See also Bally & Kohatsu-Higa (2015).

For simplicity, we take independent interarrivals $\{\tau_k - \tau_{k-1}\}_{k=1}^{\infty}$ with a common density ψ and survival function $\Psi(t) = \mathrm{P}(\tau_1 > t) = \int_t^{\infty} \psi(s)\,\mathrm{d}s$.

Recall, $X^\pi = (Y^\pi, Z^\pi)$. We adopt the convention, $\prod_{k=1}^{0} c_k = 1$.

Theorem 2.3. *Suppose μ, σ, and Λ in (4) satisfy Assumption 2.1. Then,*

$$\begin{aligned} &\mathrm{E}\Big(\int_0^T \Lambda(Y_t)\,\mathrm{d}t\Big) = \mathrm{E}(\mathrm{U}_T), \\ &\mathrm{U}_T = \frac{Z^\pi_T}{\Psi(T-\tau_{N_T})} \prod_{k=1}^{N_T} \frac{\theta_{\tau_k-\tau_{k-1}}(X^\pi_{\tau_{k-1}}, X^\pi_{\tau_k})}{\psi(\tau_k-\tau_{k-1})}. \end{aligned} \tag{7}$$

Th estimator (7) does not involve the coefficient ν since $\mathrm{E}(B_t) = 0$.

Theorem 2.4. *Suppose the μ, σ, and Λ in (4) satisfy Assumption 2.1, the objective function f satisfies Assumption 2.2, and let $\nu_T^2 = \nu^2 T/2$. We have,*

$$\begin{aligned} &\mathrm{E}\big(e^{-\int_0^T \Lambda(Y_t)\,\mathrm{d}t} f(Y_T)\big) = \mathrm{E}(\mathrm{U}_T), \\ &\mathrm{U}_T = \frac{e^{-Z^\pi_T} f(Y^\pi_T)}{e^{\nu_T^2}\Psi(T-\tau_{N_T})} \prod_{k=1}^{N_T} \frac{\theta_{\tau_k-\tau_{k-1}}(X^\pi_{\tau_{k-1}}, X^\pi_{\tau_k})}{\psi(\tau_k-\tau_{k-1})}. \end{aligned} \tag{8}$$

These results may be coupled with those in Chen, Shkolnik & Giesecke (2019) to incorporate a jump-diffusion setting. When $\Lambda = 0$ and the density ψ is exponential with rate λ, we recover (3) ($\theta = \vartheta$), but with significantly relaxed conditions on f than those in the literature. When $f = 1$, we recover the results of Wagner (1987) and Wagner (1988). Note,

$$\mathrm{E}\big(1_{\{\eta > T\}} f(Y_T)\big) = \mathrm{E}(\mathrm{U}_T) \tag{9}$$

for U_T in (8) for η, a stopping time with (nonnegative) intensity $\Lambda(Y)$ (Theorem 3.1 of Giesecke & Shkolnik (2020) establishes this identity under Assumption 2.1). This extends the scope of Theorem 2.4 to killed diffusions.

The variance of the estimators depends heavily on the choice of parameters ν and the choice of the counting process N. We recommend $\nu > 1/2$ to avoid division by small numbers in (6). The choice of a Poisson process for N as in (3) leads to an infinite variance. Choices that lead to a finite variance are discussed in Andersson & Kohatsu-Higa (2017). One example includes Beta distributed interarrivals on $[0, T+\epsilon]$ for a $\epsilon > 0$, i.e.,

$$\psi(\delta) = 1/\sqrt{4\delta(T+\epsilon)}. \tag{10}$$

Theorem 2.5. *Let U_T be the estimator in Theorems 2.3 or 2.4 that satisfies the associated assumptions. Then, $\mathrm{E}(\mathrm{U}_T^2) < \infty$ provided ψ is given by (10).*

The result generalizes to other ψ in Andersson & Kohatsu-Higa (2017).

3. Theorem Proofs

We begin by establishing Theorem 2.5 which guarantees that the variance of our estimators is finite. We leverage following auxiliary result.

Lemma 3.1. *Suppose μ, σ, and Λ satisfy Assumption 2.1. Let $n \in \mathbb{N}$ and set $\Pi_n = \{(t_1, \dots, t_n) \in (0, T)^n : t_{k-1} < t_k\}$. For $1 \le p < \infty$ and measurable $G : \mathbb{R}^{d+1} \to \mathbb{R}$ of exponential growth, define $u_n : \Pi_n \to \mathbb{R}$ as ($t_0 = 0$)*

$$u_n(t_1, \dots, t_n) = \mathrm{E}\left(\left|G(X_T^\pi)\prod_{k=1}^{n}\theta_{t_k - t_{k-1}}(X^\pi_{t_{k-1}}, X^\pi_{t_k})\right|^p\right). \tag{11}$$

There exists a constant C_T such that $u_n(t_1, \dots, t_n) \le C_T^n \prod_{k=1}^n (t_k - t_{k-1})^{-p/2}$.

We defer the proof of Lemma 3.1 to the Appendix.

Remark 3.2. *The constant C_T depends on T, d, p, ν, y as well as the coefficients μ, σ, Λ and the function G, but not on n, nor the times points $\{t_k\}_{k=1}^n$.*

A similar bound appears in Corollary 4.2 of Andersson & Kohatsu-Higa (2017), but for Λ continuously differentiable and bounded and G, also bounded. It is key to establishing finite variance of the estimators involving interarrival densities ψ such as (10). We follow an argument similar to that of Proposition 7.3 of Andersson & Kohatsu-Higa (2017), to show that

$$\begin{aligned} &\mathrm{E}\left(|\mathrm{U}_T|^p\right) < \infty \qquad 1 \le p < \infty, \ \text{and} \\ &\mathrm{U}_T = \frac{G(X_T^\pi)}{\Psi(T - \tau_{N_T})}\prod_{k=1}^{N_T}\frac{\theta_{\tau_k - \tau_{k-1}}(X^\pi_{\tau_{k-1}}, X^\pi_{\tau_k})}{\psi(\tau_k - \tau_{k-1})}. \end{aligned} \tag{12}$$

To see this, observe that by the independence of the $\{\tau_k\}_{k\in\mathbb{N}}$ (hence N_T) and X, almost surely $\phi_{N_T}(\tau_1, \dots, \tau_{N_T}) = \mathrm{E}\left(|\mathrm{U}_T|^p \,|\, \{\tau_k\}_{k\in\mathbb{N}}\right)$ for ϕ given by

$$\phi_n(t_1, \dots, t_n) = \frac{u_n(t_1, \dots, t_n)}{|\Psi(T - t_n)\prod_{k=1}^n \psi(t_k - t_{k-1})|^p}.$$

We have, $\mathrm{E}\left(|\mathrm{U}_T|^p\right) = \sum_{n=0}^\infty \mathrm{E}(1_{\{N_T = n\}}\mathrm{E}(|\mathrm{U}_T|^p \,|\, \{\tau_k\}_{k\in\mathbb{N}})$, and so

$$\begin{aligned} \mathrm{E}\left(|\mathrm{U}_T|^p\right) &= \sum_{n=0}^\infty \mathrm{E}(1_{\{N_T=n\}}\phi_{N_T}(\tau_1, \dots, \tau_{N_T})) \\ &= \sum_{n=0}^\infty \int_{\Pi_n} \frac{u_n(t_1, \dots, t_n)\,\mathrm{d}t_1 \dots \mathrm{d}t_n}{|\Psi(T - t_n)\prod_{k=1}^n \psi(t_k - t_{k-1})|^{p-1}} \end{aligned} \tag{13}$$

as $\mathrm{E}(1_{\{N_T=n\}}\phi_n(\tau_1,\ldots,\tau_n))=\int_{\Pi_n}\phi_n(t_1,\ldots,t_n)\Psi(T-t_n)\prod_{k=1}^n\psi(t_k-t_{k-1})\,\mathrm{d}t_k$ by Andersson & Kohatsu-Higa (2017, Lemma 7.1), provided the right side converges. Applying the bound $\Psi(T-t_n)\geq\Psi(T)=\int_T^{T+\epsilon}\psi(s)\,\mathrm{d}s>0$ and the ψ in (10), we indeed have by Lemma 3.1 that the series converges, and

$$\mathrm{E}(|\mathrm{U}_T|^p)\leq\sum_{n=0}^{\infty}K_T^n\int_{\Pi_n}\prod_{k=1}^{n}(t_k-t_{k-1})^{-1/2}\mathrm{d}t_k<\infty \tag{14}$$

for a constant $K_T\geq C_T$ and all $1\leq p<\infty$. Having established (12), by considering only $1\leq p\leq 2$, we deduce the claim in Theorem 2.5, as the estimator U_T in both (7) and (8) has a G of exponential growth. By the same argument, alternative interarrival densities ψ may be employed.

Next, we prove our estimators are unbiased.

Lemma 3.3. *Suppose $X=(Y,Z)$ in (4) has coefficients μ, σ and Λ satisfying Assumption 2.1. For $G:\mathbb{R}^{d+1}\to\mathbb{R}$, measurable and of exponential growth,*

$$\mathrm{E}(G(X_T))=\mathrm{E}\left(\frac{G(X_T^\pi)}{\Psi(T-\tau_{N_T})}\prod_{k=1}^{N_T}\frac{\theta_{\tau_k-\tau_{k-1}}(X_{\tau_{k-1}}^\pi,X_{\tau_k}^\pi)}{\psi(\tau_k-\tau_{k-1})}\right). \tag{15}$$

We defer the proof of Lemma 3.3 to the Appendix. The lemma extends the parametrix method to objective functions G of exponential growth and to one of the components of the diffusion (namely, Z) to a Lipschitz continuous drift. Here, we illustrate its application to Theorems 2.3 and 2.4.

Since B is independent of W (hence, independent of Y),

$$\begin{aligned}\mathrm{E}(Z_T)&=\mathrm{E}\Big(\int_0^T\Lambda(Y_t)\,\mathrm{d}t\Big),\\ \mathrm{E}\big(e^{-Z_T}f(Y_T)\big)&=e^{\nu_T^2}\,\mathrm{E}\Big(e^{-\int_0^T\Lambda(Y_t)\,\mathrm{d}t}f(Y_T)\Big).\end{aligned} \tag{16}$$

We apply Lemma 3.3 with $G(x)=z$ and $G(x)=e^{-z}f(y)$ for $(y,z)=x$ and f measurable and of exponential growth to the left sides of (16) to obtain parmetrix formulas like (15). Finally, adjusting for the term ν_T^2 in (8) (but not in (7)) yields the claims of Theorems (2.3) and 2.4.

We remark that the above argument leads to the design of additional estimators, provided the adjustment for the Brownian motion B may be performed. For example, the quadratic $G(x)=z^2$ may be addressed using

$$\mathrm{E}(Z_T^2)-\nu^2T=\mathrm{E}\bigg(\Big(\int_0^T\Lambda(Y_t)\,\mathrm{d}t\Big)^2\bigg) \tag{17}$$

again appealing to the independence of Y and B, and that $\mathrm{E}(B_T)=0$ and $\mathrm{E}(B_T^2)=T$. Again applying Lemma 3.3 to the left side of (17), yields the appropriate paramtetix formula for estimating the right side of (17) without bias. The parameter ν cannot be set to zero as it is key to establishing (15).

4. Numerical example

We provide numerical results to demonstrate the performance of our estimator. Based on the Assumptions 2.1 and 2.2, we consider Y solving

$$\mathrm{d}Y_t = \sin(Y_t)\,\mathrm{d}t + \sqrt{0.4 + 0.2\sin(Y_t)}\,\mathrm{d}W_t. \tag{18}$$

We test estimators for $\int_0^{T_1} \Lambda(X_t)\,\mathrm{d}t$ and $e^{-\int_0^{T_2} \Lambda(Y_t)\,\mathrm{d}t} f(Y_{T_2})$, where $T_1 = 1$, $T_2 = 0.2$, $\Lambda(x) = 0.2 + x$ and $f(x) = e^x$. Different T_1 and T_2 ensure the two objective function values are roughly of the same order of magnitude. The drift and volatility of Y are smooth and bounded, but the functions Λ and f both extend the typical scope where the parametrix estimator applies.

We compare the parametrix estimators with the Euler method, to test for a bias in the Euler method and for its absence in the parametrix estimator. A nearly exact expectation is computed with a very large number of Monte Carlo trials and a very fine discretization. For the benchmark Euler method, the computation budget is allocated according to the trade-off rule in Duffie & Glynn (1995). Specifically, letting p denote the number of (uniformly spaced) discretization points t_k in (2), we run $M = p^2$ Monte Carlo trials. This approach balances the bias of the Euler scheme with the statistical error of Monte Carlo. For $p = 2^7$, the Euler and parametrix scheme have a nearly identical run-time of 5×10^{-4} seconds per similation trial.

Tables 1 and 2 summarize our numerical experiments. Table 1 reports estimates of the expectaton of a path intergral of the diffusion Y. The error of each estimate falls within the 99% confidence interval, indicating no bias in either method. Hence, the trade-off rule between bias and computional budget masks the bias of the Euler scheme; (alternatively, this effect may be due to cancellation errors facilitated by the linear choice of Λ). The variance of the Euler samples is roughly 0.22 for each discretization. In contrast, the parametrix estimator exhibits a sample variance that is hundreds of times larger; yet, it appears bounded, confirming our theoretical findings. Significantly longer running times were then required to match the performance of the Euler scheme. While many of the computations to assemble the parametrix estimator may be parallelized, more work to reduce its variance is needed to outperform the Euler method on this example.

Table 2 reports estimates of $\mathrm{E}(1_{\{\eta > T_2\}} f(Y_{T_2}))$ where η is a stopping time with intensity $\Lambda(Y)$, or equivalently the expectation of a discounted payoff $f(Y_{T_2})$. In this example, we begin to observe the bias in the Euler scheme, as the error of its estimates begins to exceed the size of the confidence interval. In contrast, the parametrix estimate error always remains well within the bounds of the statistical uncertainty. The sample variance of the estimates produced by the Euler scheme is roughly 0.73, increasing relative to the path-integral example. In contrast, the sample variance of the

Table 1. *Estimation of* $\mathrm{E}\big(\int_0^{T_1} \Lambda(Y_t)\,\mathrm{d}t\big)$ *for model* (18) *with the parametrix and Euler methods. "Error" reports the absolute value between the (nearly) exact value and the Monte Calro estimate based on* M *trials. Normal confidence intervals (CI) accompany each estimate.*

Parametrix				Euler			
M	Error	Variance	99% CI	M	p	Error	99% CI
10^6	0.0013	62.06	0.0200	2^{14}	2^7	0.0057	0.0094
10^7	0.0009	65.23	0.0065	2^{16}	2^8	0.0031	0.0047
10^8	0.0014	73.21	0.0021	2^{18}	2^9	0.0006	0.0024
10^9	0.0003	69.78	0.0007	2^{20}	2^{10}	0.0005	0.0012

Table 2. *Estimation of* $\mathrm{E}\big(e^{-\int_0^{T_2} \Lambda(Y_t)\,\mathrm{d}t} f(Y_{T_2})\big)$ *for model* (18) *with the parametrix and Euler methods. "Error" reports the absolute value between the (nearly) exact value and the Monte Calro estimate based on* M *trials. Normal confidence intervals (CI) accompany each estimate.*

Parametrix				Euler			
M	Error	Variance	99% CI	N	h	Error	99% CI
10^6	0.0072	57.35	0.0193	2^{14}	2^7	0.0134	0.0172
10^7	0.0016	56.56	0.0061	2^{16}	2^8	0.0015	0.0086
10^8	0.0003	56.61	0.0019	2^{18}	2^9	0.0040	0.0043
10^9	0.0001	56.54	0.0003	2^{20}	2^{10}	0.0033	0.0022

parametrix estimator decreased. For this reason, the errors made by the parametrix method are significantly lower than those of the Euler method. This highlights the advantages of the paramterix family of estimators.

5. Conclusion

The parametrix method allows for unbiased simulation estimators of diffusions. While the approach has great potential for fast and accurate simulation, current theoretical results limit the scope of its application. We extend the parametrix method to accommodate several classes of problems encountered in practice. In particular, we extend the method to path integrals and to killed diffusions which arise frequently in modeling. Our theoretical results establish the unbiasedness and finite variance properties of the estimators. Numerical examples illustrates the performance of the estimators relative to a biased, Euler discretization scheme.

A. Auxiliary Results and Notation

Let φ_c denote the multivariate Gaussian density with a zero mean and variance (matrix) c. Denote by q the transition kernel of the Euler process $X^\pi = (Y^\pi, Z^\pi)$ defined in (5) on the interval $[\tau_k, \tau_{k+1})$, given $(X^\pi_{\tau_k}, \tau_k, \tau_{k+1})$. The law of (Y^π, Z^π) given $X^\pi_{\tau_k} = (y_1, z_1) \in \mathbb{R}^{d+1}$ is Gaussian with covariances $a(y_1) = (\sigma\sigma^\top)(y_1) \in \mathbb{R}^{d\times d}$ and ν^2 for Y^π and Z^π respectively with means $y_1 + \mu(y_1)$ and $z_1 + \Lambda(z_1)$. For initial and final points $x_1 = (y_1, z_1)$ and $x_2 = (y_2, z_2) \in \mathbb{R}^{d+1}$, the density $q_t(x_1, x_2)$ decomposes as

$$q_t(x_1,x_2) = \varphi_{ta(y_1)}(y_2 - y_1 - \mu(y_1)t)\varphi_{t\nu^2}(z_2 - z_1 - \Lambda(y_1)t) \tag{19}$$

as the Brownian motions W and B driving Y^π and Z^π are independent.

Lemma A.1. *For any symmetric $c \in \mathbb{R}^{d\times d}$ for which there exist constants $c_1, c_2 > 0$ satisfying $c_1 I \preceq c \preceq c_2 I$, we have that for $0 < t < T$,*

$$\varphi_{tc}(y_2 - y_1 - b(y_1)t) \le C_T\,\varphi_{2tc_2I}(y_2 - y_1) \tag{20}$$

where $C_T = 2^{d/2}e^{T\|b\|_\infty/(2c_1)}$ and for all $y_1, y_2 \in \mathbb{R}^d$.

Proof. See Lemma A.1 of Andersson & Kohatsu-Higa (2017). ■

We provide a bound on the bias correcting weights θ in (6). For a function f, let $\|f\|_\infty = \sup_x |f(x)|$ where $|\cdot|$ is the Eucledian norm.

Lemma A.2. *Suppose that Assumption 2.1 holds for the coefficients μ, σ and Λ determining θ in (6). Then for any $1 \le p < \infty$ and $0 < t \le T$, we have*

$$|\theta_t(x_1,x_2)^p q_t(x_1,x_2)| \le \frac{K_T}{t^{p/2}}\varphi_{4ta_2I}(y_2 - y_1)\varphi_{2t\nu^2}(z_2 - z_1 - \Lambda(y_1)t) \tag{21}$$

for all $x_1 = (y_1, z_1), x_2 = (y_2, z_2) \in \mathbb{R}^{d+1}$ and a constant K_T that depends on $T, d, p, c_1, c_2, \|\mu\|_\infty$ and the Lipschitz constant of Λ (per Assumption 2.1).

Proof. We have $\theta(x_1,x_2) = \vartheta_t(y_1,y_2) + (\Lambda(y_2) - \Lambda(y_1))\left(\frac{z_2 - z_1 - \Lambda(y_1)t}{t\nu^2}\right)$, where ϑ satisfies, by Lemma 4.1 in Andersson & Kohatsu-Higa (2017),

$$|\vartheta_t(y_1,y_2)^p \varphi_{ta(y_1)}(y_2 - y_1 - \mu(y_1)t)| \le \frac{C'_T}{t^{p/2}}\varphi_{4ta_2I}(y_2 - y_1) \tag{22}$$

for a constant C'_T that depends on $d, a_1, a_2, \|\mu\|_\infty$ and T.

Next, applying Lemma A.1 with constant C_T and $y = y_2 - y_1 - \mu(y_1)t$, and letting L denote the Lischitz constant of Λ, we obtain

$$\begin{aligned} |\Lambda(y_1) - \Lambda(y_2)|^p \varphi_{ta(y_1)}(y) &\le \frac{C_T L}{t^{p/2}}|y_2 - y_1|^p \varphi_{2ta_2I}(y_2 - y_1) \\ &\le \frac{L_T}{t^{p/2}}\varphi_{4ta_2I}(y_2 - y_1) \end{aligned} \tag{23}$$

for a constant L_T, where the Gaussian decay of φ is used to bound $|y_2 - y_1|^p$. The latter argument also yields that for $z = z_2 - z_1 - \Lambda(y_1)t$,

$$\left|\frac{z}{t\nu^2}\right|^p \varphi_{t\nu^2}(z) \leq \left(8p\sqrt{T/t}\right)^p \varphi_{2\nu^2 t}(z).$$

The claim now follows by using the identity $|a+b|^p \leq 2^{p-1}(|a|^p + |b|^p)$ for the two terms of θ and applying the bounds above to arrive at (21).

∎

B. Proof of Lemma 3.1

Using the transition law q of X^π (Appendix A) to write $u_n(t_1, \ldots, t_n) = \mathrm{E}\left(\left|G(X_T^\pi)\prod_{k=1}^n \theta_{t_k - t_{k-1}}(X_{t_{k-1}}^\pi, X_{t_k}^\pi)\right|^p\right)$, yields the Lebesgue itegrand

$$\Big(\prod_{k=1}^n |\theta_{t_k - t_{k-1}}(x_{k-1}, x_k)|^p q_{t_k - t_{k-1}}(x_{k-1}, x_k)\Big)|G(x_{n+1})|^p q_{T-t_n}(x_n, x_{n+1})$$

which is integrated over all $(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{(d+1)\times(n+1)}$. By Lemma A.2, for a constant K_T', this integrand is bounded above by

$$\Big(\prod_{k=1}^n \frac{K_T'}{(t_k - t_{k-1})^{p/2}} \widehat{q}_{t_k - t_{k-1}}(x_{k-1}, x_k)\Big)|G(x_{n+1})|^p q(x_n, x_{n+1})$$

where $\widehat{q}$ is the transition law (c.f., the transition law q of X^π in (19)),

$$\hat{q}_t(x_{k-1}, x_k) = \varphi_{4ta_2I}(y_k - y_{k-1})\varphi_{2t\nu^2}(z_k - z_{k-1} - \Lambda(y_{k-1})t) \tag{24}$$

with $(y_k, z_k) = x_k$ of a process $\widehat{X} = (\widehat{Y}, \widehat{Z})$ that starts in $(y_0, 0) \in \mathbb{R}^{d+1}$. Observe that $\widehat{Y}$ is mean-zero Gaussian (w.l.o.g., driven by W) with independent components and covariance matrix $4a_2 I$. But, $\widehat{Y}$ depends on $\widehat{Z}$, defined by $\widehat{Z}_{t_\ell} = \sum_{k=1}^{\ell} \Lambda(\widehat{Y}_{t_{k-1}})(t_k - t_{k-1}) + \sqrt{2\nu}B_{t_\ell}$ (w.l.o.g., B in (5)).

By Lemma A.1 and (19), for a constant C_T, we have the upper bound

$$q_t(x_n, x_{n+1}) \leq C_T \varphi_{4ta_2I}(y_2 - y_1)\varphi_{2t\nu^2}(z_2 - z_1 - \Lambda(y_1)t) = C_T\widehat{q}_t(x_n, x_{n+1})$$

and assembling the above estimates, we deduce the following bound.

$$u_n(t_1, \ldots, t_n) \leq C_T \mathrm{E}\left(|G(\widehat{X}_T)|^p\right) \prod_{k=1}^n \frac{K_T'}{(t_k - t_{k-1})^{p/2}} \tag{25}$$

It remains to bound $\mathrm{E}\left(|G(\widehat{X}_T)|^p\right)$. For constants $c, C > 0$ we have

$$|G(\widehat{X}_T)|^p \leq C\exp(cp|Z_T| + cp\textstyle\sum_{i=1}^d |\widehat{Y}_T^i|)$$

by the exponential growth of G and $|x| = \sqrt{\sum_{i=1}^{d+1} x_i^2} \le \sum_{i=1}^{d+1} |x_i|$. Further,

$$\mathrm{E}(|G(\widehat{X}_T)|^p)^2 \le C\,\mathrm{E}(e^{2cp|Z_T|}) \prod_{i=1}^d \mathrm{E}(e^{2cp|\widehat{Y}_T^i|}) \tag{26}$$

by the independence of the $\{\widehat{Y}_T^i\}$. Applying the Lipschitz property of Λ, the triangle inequality and that $\sum_{k=1}^n \sum_{\ell=1}^k |b_\ell - b_{\ell-1}| = \sum_{k=1}^n k|b_{n-k+1} - b_{n-k}|$,

$$\begin{aligned} |\widehat{Z}_T| &\le |\Lambda(y)|T + TL \sum_{k=1}^n |\widehat{Y}_{t_k} - y| + \sqrt{2\nu^2}|B_T| \\ &\le |\Lambda(y)|T + TL \sum_{k=1}^n k|\widehat{Y}_{t_{n-k+1}} - \widehat{Y}_{t_{n-k}}| + \sqrt{2\nu^2}|B_T| \end{aligned}$$

for L the Lipschitz constant of Λ. Applying the independence of increments property of $\widehat{Y}$ and its independence with the Brownian motion B yields,

$$\mathrm{E}(e^{2cp|Z_T|}) \le e^{2|\Lambda(y)|T} \mathrm{E}(e^{\sqrt{8(cp\nu)^2}|B_T|}) \prod_{k=1}^n \mathrm{E}(e^{2cpTL|\widehat{Y}_{t_{n-1+1}} - \widehat{Y}_{t_{n-k}}|}) \tag{27}$$

To conclude, note that for any zero-mean normal random variable $Q \in \mathbb{R}^d$ with independent components each of variance $4a_2$ and $\delta_k > 0$, we have

$$\mathrm{E}(e^{\delta_k|Q|}) \le \prod_{i=1}^d \mathrm{E}(e^{\delta_k|Q^i|}) \le \prod_{i=1}^d 2\mathrm{E}(e^{\delta_k Q^i}) = e^{d(1+2a_2\delta_k^2)}.$$

As $|\widehat{Y}_{t_k} - \widehat{Y}_{t_{k-1}}| \overset{\mathscr{D}}{=} (t_k - t_{k-1})Q$, assembling, (25), (26) and (27), we see $C_T\mathrm{E}(|G(\widehat{X}_T|^p)$ is bounded above by a constant of the form b^{nd} for $b > 0$ that depends on a_2, T, L, c, p, y, ν. This constant does not depend on the points $\{t_k\}$ as $\sum_{k=1}^n \delta_k^2 \le (\sum_{k=1}^n |\delta_k|)^2 \le T^2$ where we set $\delta_k = t_k - t_{k-1}$, as $t_n < T$. The latter point addresses Remark 3.2 and concludes the proof.

C. Proof of Lemma 3.3

Let $\{\Lambda_\ell\}_{\ell\in\mathbb{N}}$ be a sequence of functions in $C_b^1(\mathbb{R})$ such that $\Lambda_\ell \to \Lambda$ pointwise (this is guaranteed by first truncating Λ so that it is bounded and then using the fact that smooth, bounded functions are dense in the space of bounded, continuous functions; in norm and hence pointwise). Define,

$$Z_t^\ell = \int_0^t \Lambda(Y_s)\,\mathrm{d}s + \nu B_t \tag{28}$$

and let $X^\ell = (Y, Z^\ell)$. Denote by $X^{\ell,\pi} = (Y^\pi, Z^{\ell,\pi})$ the associated process, defined identically to (5) by with Λ_ℓ replacing Λ. It follows that $X^{\ell,\pi}$ is a $(d+1)$-dimensional diffusion that satisfies the assumptions of the plain parametrix method (i.e., C_b^1 drift and uniformly elliptic, C_b^2 volatility) for a weight function θ^ℓ defined identically to θ in (6) but with Λ_ℓ replacing Λ.

We accomplish the proof in two steps. First, we prove the statement for G smooth and of compact support. In the second step, we extend the result to G measurable and of exponential growth.

Step 1. Let G be smooth and of compact support. The G is bounded, with bound $\|G\|_\infty < \infty$. We then have the standard parametrix formula,

$$
\begin{aligned}
&\mathrm{E}\,(G(X_T^\ell)) = \mathrm{E}\,(\mathrm{U}_T^\ell)\\
(29)\qquad &\mathrm{U}_T^\ell = \frac{G(X_T^{\ell,\pi})}{\Psi(T-\tau_{N_T})}\prod_{k=1}^{N_T}\frac{\theta^\ell_{\tau_k-\tau_{k-1}}(X^{\ell,\pi}_{\tau_{k-1}}, X^{\ell,\pi}_{\tau_k})}{\psi(\tau_k-\tau_{k-1})}.
\end{aligned}
$$

for every $\ell \in \mathbb{N}$. Since $X_T^\ell \to X_T$ almost surely, $G(x) \leq \|G\|_\infty$ and G is continuous, by dominated convergence, $\lim_{\ell\uparrow\infty} \mathrm{E}G(X_T^\ell) = \mathrm{E}G(X_T)$. Similarly, as $\mathrm{U}_T^\ell \to \mathrm{U}_T$ (the integrand on the right side of (15)) almost surely,

$$
\begin{aligned}
&\mathrm{E}\,(G(X_T)) = \lim_{\ell\uparrow\infty}\mathrm{E}(\mathrm{U}_T^\ell) = \mathrm{E}(\mathrm{U}_T)\\
(30)\qquad &\mathrm{U}_T = \frac{G(X_T^{\pi})}{\Psi(T-\tau_{N_T})}\prod_{k=1}^{N_T}\frac{\theta_{\tau_k-\tau_{k-1}}(X^{\pi}_{\tau_{k-1}}, X^{\pi}_{\tau_k})}{\psi(\tau_k-\tau_{k-1})}
\end{aligned}
$$

provided there exists an integrable random variable $\mathrm{V}_T \geq \sup_\ell |\mathrm{U}_T^\ell|$.

We show that the required random variable $\mathrm{V} = |\mathrm{V}_T|$ is given by

$$
\mathrm{V}_T = \frac{\|G\|_\infty}{\Psi(T-\tau_{N_T})}\prod_{k=1}^{N_T}\frac{|\vartheta_{t_k-t_{k-1}}(Y^\pi_{t_{k-1}}, Y^\pi_{t_k})| + L|Y^\pi_{t_k} - Y^\pi_{t_{k-1}}||Q_k|}{\psi(\tau_k-\tau_{k-1})},
$$

where $Q_k = Z^{\ell,\pi}_{t_k} - Z^{\ell,\pi}_{t_{k-1}} - \Lambda(Y^\pi_{t_{k-1}})(t_k - t_{k-1}) = B_{t_k} - B_{t_{k-1}}$ is has the standard Normal distribution and is independent of Y^π.

To show that $\mathrm{E}(\mathrm{V}_T) < \infty$, observe that (see (6)),

$$
\theta^\ell_{t_k-t_{k-1}}(X^{\ell,\pi}_{t_{k-1}}, X^{\ell,\pi}_{t_k}) = \vartheta_{t_k-t_{k-1}}(Y^\pi_{t_{k-1}}, Y^\pi_{t_k}) + (\Lambda_\ell(Y^\pi_{t_k}) - \Lambda_\ell(Y^\pi_{t_{k-1}}))Q_k
$$

and due to the fact that each Λ_ℓ is Lipschitz with constant L_ℓ,

$$
|\theta^\ell_{t_k-t_{k-1}}(X^{\ell,\pi}_{t_{k-1}}, X^{\ell,\pi}_{t_k})| = |\vartheta_{t_k-t_{k-1}}(Y^\pi_{t_{k-1}}, Y^\pi_{t_k})| + L_*|Y^\pi_{t_k} - Y^\pi_{t_{k-1}}||Q_k|
$$

where $L_* = \sup_\ell L_\ell < \infty$ (since the limit Λ is Lipschitz).

By steps identical to (13) with $p = 1$, we deduce that

$$
(31)\qquad \mathrm{E}\,(|\mathrm{V}_T|) = \sum_{n=0}^{\infty}\int_{\Pi_n} v_n(t_1,\ldots,t_n)\,\mathrm{d}t_1\ldots\mathrm{d}t_n\,.
$$

where, after setting $\delta_k = t_k - t_{k-1}$, we have

$$
\begin{aligned}
v_n(t_1, \dots, t_n) &= \|G\|_\infty \mathrm{E}\left(\prod_{k=1}^n |\vartheta_{\delta_k}(Y^\pi_{t_{k-1}}, Y^\pi_{t_k})| + L|Y^\pi_{t_k} - Y^\pi_{t_{k-1}}||Q_k|\right) \\
&= \|G\|_\infty \mathrm{E}\left(\prod_{k=1}^n |\vartheta_{\delta_k}(Y^\pi_{t_{k-1}}, Y^\pi_{t_k})| + \sqrt{\frac{2L_*^2}{\pi}}|Y^\pi_{t_k} - Y^\pi_{t_{k-1}}|\right)
\end{aligned}
$$

and the last equality follows by the (mutual) independence of the $\{Q_k\}_{k=1}^n$, their independence of Y^π and the conditioning on the $\{Y^\pi_{t_k}\}_{k=1}^n$ that uses the fact that $\mathrm{E}(|Q_k|\,|\{Y^\pi_{t_k}\}_{k=1}^n) = \mathrm{E}(|Q_k|) = \sqrt{2/\pi}$ for every $1 \le k \le n$.

It now only remains to verify that the series in (31) converges. Denoting by q^Y, the transition law of Y^π (i.e., $q_t^Y(y_1, y_2) = \varphi_{ta(y_1)}(y_2 - y_1 - \mu(y_1)t)$),

$$
\begin{aligned}
|\vartheta_{t_k - t_{k-1}}(y_{k-1}, y_k)| q^Y_{\delta_k}(y_{k-1}, y_k) &\le \frac{C_T}{\sqrt{t}} \varphi_{4\delta_k a_2 I}(y_k - y_{k-1}) \\
\sqrt{\frac{2L_*^2}{\pi}} |y_k - y_{k-1}| q^Y_{\delta_k}(y_{k-1}, y_k) &\le \frac{L_T}{\sqrt{t}} \varphi_{4\delta_k a_2 I}(y_k - y_{k-1})
\end{aligned}
$$

for constants $C_T, L_T > 0$ by using (22) and (23) (taking $p = 1$). It follows that $v_n(t_1, \dots, t_n) \le (C_T + L_T)^n/\sqrt{t}$. As desired (31) converges as in (14).

This concludes the proof of (30) for G smooth and of compact support and Λ in the definition of θ taken to be only Lipschitz continuous.

Step 2. In this step, we extend (30) to G measurable and of exponential growth. First, we observe that by an argument identical to that of Lemma 3.1 of Chen et al. (2019) (see also Remark 3.2 in that reference) (30) holds for G measurable and bounded. Therefore, taking G measurable and of exponential growth, and defining $G_K(x) = (G(x) \wedge K) \vee (-K)$ for $K > 0$, formula (30) holds for each G_K replacing G. The extension to G follows again by applying dominated convergence twice, as done below.

First, to conclude that $\lim_{K\uparrow\infty} \mathrm{E}(G_K(X_T)) = \mathrm{E}(G(X_T))$, we note that $G_K(X_T) \to G(X_T)$ almost surely and observe that $|G_K(X_T)| \le |G(X_T)| \le Ce^{c|X_T|} = Ce^{c|Z_T| + c\sum_{i=1}^d |Y_T^i|}$, denoted by A_T, for some constants $C, c > 0$ per the exponential growth of G. By the Gaussian density bound for the random variable $X_T = (Y_T, Z_T)$ from Menozzi et al. (2020, Theorem 1.2), the random variable A_T is integrable, and the stated claim follows.

Lastly, we prove $\lim_{K\uparrow\infty} \mathrm{E}(\mathrm{U}_T^K) = \mathrm{E}(\mathrm{U}_T)$ with U_T in (30) with G measurable and of exponential growth, and for G_K above, the U_T^K given by

$$
\mathrm{U}_T^K = \frac{G_K(X_T^\pi)}{\Psi(T - \tau_{N_T})} \prod_{k=1}^{N_T} \frac{\theta_{\tau_k - \tau_{k-1}}(X^\pi_{\tau_{k-1}}, X^\pi_{\tau_k})}{\psi(\tau_k - \tau_{k-1})}
$$

We have $|\mathrm{U}_T^K| \le |\mathrm{U}_T|$ with U_T in (30) and $\mathrm{E}(|\mathrm{U}_T|) < \infty$ per (12) (and Lemma 3.1). This concludes the proof; $\mathrm{E}(G(X_T)) = \mathrm{E}(\mathrm{U}_T)$ as required.

References

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.

Blanchet, J. & Zhang, F. (2017). Exact simulation for multivariate ito diffusions. Working Paper.

Chen, G., Shkolnik, A. & Giesecke, K. (2019). Unbiased simulation estimators for jump-diffusions, *in* N. Mustafee, K.-H. G. Bae, S. Lazarova-Molnar, M. Rabe, & C. Szabo (eds), *Proceedings of the 2019 Winter Simulation Conference*, Institute of Electrical and Electronics Engineers, Inc., Piscataway, New Jersey, pp. 890–901.

Collin-Dufresne, P., Goldstein, R. & Hugonnier, J. (2004). A general formula for valuing defaultable securities, *Econometrica* **72**(5): 1377–1407.

Doumbia, M., Oudjane, N. & Warin, X. (2017). Unbiased monte carlo estimate of stochastic differential equations expectations, *ESAIM: Probability and Statistics* **21**: 56–87.

Duffie, D. & Glynn, P. (1995). Efficient monte carlo simulation of security prices, *The Annals of Applied Probability* **5**(4): 897–905.

Giesecke, K. & Shkolnik, A. (2020). Reducing bias in event time simulations via measure changes. *Mathematics of Operations Research*, forthcoming.

Henry-Labordère, P., Tan, X. & Touzi, N. (2017). Unbiased simulation of stochastic differential equations, *The Annals of Applied Probability* **27**(6): 3305–3341.

Krylov, N. V. (1974). Some estimates of the probability density of a stochastic integral, *Mathematics of the USSR-Izvestiya* **8**(1): 233.

Menozzi, S., Pesce, A. & Zhang, X. (2020). Density and gradient estimates for non-degenerate brownian sdes with unbounded measurable drift.

Wagner, W. (1987). Unbiased monte carlo evaluation of certain functional integrals, *Journal of Computational Physics* **71**(1): 21–33.

Wagner, W. (1988). Unbiased multi-step estimators for the monte carlo evaluation of certain functional integrals, *Journal of Computational Physics* **79**(2): 336–352.