Unified moment-based modelling of integrated stochastic processes

Ioannis Kyriakou*, Riccardo Brignone, Gianluca Fusai

Abstract

In this paper we present a new general method for simulating integrals of stochastic processes. We focus on the nontrivial case of time integrals conditional on the state variable levels at the endpoints of a time interval, based on a moment-based probability distribution construction. We present different classes of models with important usages in finance, medicine, epidemiology, climatology, bioeconomics and physics. We highlight the benefits of our method and benchmark its performance against existing schemes.

1. Introduction

Time integrals of stochastic processes and their simulation feature in numerous research problems in finance, medicine, epidemiology, technology, engineering, bioeconomics and physics; the cases mentioned next are certainly non-exhaustive. In financial engineering, time integrals appear in stochastic volatility models, commodity price models (see e-companion Section EC.9), Asian options, volatility options and interest rate derivatives. In turbulent diffusion modelling and related phenomena, the position of a fluid particle at a certain time is given by the integrated velocity (e.g., see Obukhov, 1959). Sums of random variables arise in wireless communications and related areas (see Nadarajah, 2008), portfolio credit risk applications such as portfolio loss process modelling (see Giesecke *et al.*, 2011 and Dassios and Zhao, 2017), but also specialized areas of biomedical engineering involving signal averaging. Within the family of sigmoidal growth models, the stochastic Verhulst population model involves a stochastic integral in its explicit solution that requires accurate simulation for generating probabilistic forecasts in fields including geoscience, oncology to describe tumor growth (e.g., see Laird, 1964, and later research) or epidemic dynamics (e.g., see Shen, 2020, Wu *et al.*, 2020). A similar simulation challenge appears in the randomized Schaefer model, which describes the growth of populations living in a randomly varying environment and being harvested, but also in the Ginzburg–Landau model used originally to describe phase transition for superconductivity, while over the years its use has broadened towards various directions.

A long-lasting concern in the cases described above remains the efficient simulation of stochastic time integrals or the even more involved, as will become clearer later, *conditional* stochastic integrals. Our analysis encompasses stochastic volatility models, such as the Heston and double Heston, stochastic alpha-beta-rho (SABR), Ornstein–Uhlenbeck stochastic volatility (OU-SV), 3/2 and 4/2, but also linear models with multiplicative noise and nonlinear reducible models, such

*Correspondence to: Ioannis Kyriakou, Faculty of Actuarial Science and Insurance, Cass Business School, City, University of London, London EC1Y 8TZ, UK.

Email addresses: ioannis.kyriakou@city.ac.uk (Ioannis Kyriakou), riccardo.brignone@finance.uni-freiburg.de (Riccardo Brignone), gianluca.fusai.1@city.ac.uk & gianluca.fusai@uniupo.it (Gianluca Fusai)

as the stochastic Verhulst, Gordon–Schaefer and Ginzburg–Landau. We extend our application to certain model variants with jumps, such as Bates, Duffie–Pan–Singleton (DPS), time-changed Lévy and self-exciting point processes. The aim of this paper is to unify simulation in a practicable manner that is fast and accurate.

For years, discretization of stochastic differential equations (SDEs) features in the literature (see, for example, Chen *et al.*, 2012 for a review) as a possible way to go round the simulation of integrated processes, inevitably yielding a bias, which can be hard to quantify accurately, besides rendering the procedure particularly tedious. To circumvent this, attempts have been made to simulate *exactly* or, perhaps more precisely phrased, recover the $O(s^{-1/2})$ convergence rate of an unbiased Monte Carlo estimation with a total computational budget s, such as Broadie and Kaya (2006), Cai *et al.* (2017), Kang *et al.* (2017) and Li and Wu (2019), for different models. These approaches have proved to be able to produce accurate results. At their core, they rely on simulating an integrated process over a time interval conditional on its level at the interval endpoints. Although accurate, a serious demerit of them, as pointed out in the seminal work of Broadie and Kaya (2006), is the implicit need to recover the unknown distribution function of the conditional integrated process using numerical inversion of the associated Laplace transform. This can become a heavy load and almost impracticable when generating entire sample trajectories, aside from introducing potential error, and thereby bias, and computational burden increases during numerical integration. Luckily, recent contributions, for example, by Cai *et al.* (2014a) have endowed us with computable error bounds of the Laplace transform inversions guaranteeing their accuracy; nevertheless, computational speed remains an issue in view also of its trade-off with accuracy, especially when integrating into a Monte Carlo simulation application. This still hinders the way between the method and the user, leaving space for further research.

Our contribution is summarized as follows. We present a unified methodological framework for modelling (conditional) integrated processes. To this end, we first employ an adaptively modified moment generating function algorithm from Choudhury and Lucantoni (1996) to compute fast the moments of the conditional integrated process. Thereby, we propose an accurate Pearson curve fit to approximate its probability distribution. This allows us to easily generate random samples from it, bypassing at any stage computational intensive Laplace transform inversion or differentiation for the computation of the moments. We also study the error resulting from this approximation and the bias induced in simulation estimators.

The system of Pearson curves is simple and fast in family selection for varying levels of skewness and kurtosis, in parameter estimation and simulation. As we explain, we approximate the unknown distribution of the time integral, which for each different model is moment-determinate, i.e., it is uniquely determined by its moments, by a *bona fide* distribution with a non-negative density. Although this is based on a first four-moment fit, we do show that the differences between the higher-order moments of a four-moment Pearson curve fit and the true ones are immaterial. Hence, the resulting simulation methodology is convergent and very accurate. Another notable merit of it is that the size Δt of the time interval does not affect the accuracy, as it does not involve any SDE discretization. Therefore, it serves as an ideal substitute of approximations that require a large number of time steps (small Δt) to potentially secure enough accuracy, when actually a single time step of arbitrary size suffices and entire sample trajectories and resulting undesired time increases are unnecessary. It can still be used when access to sample trajectories is intended, offering a speed up over numerical integration- or expansion-based methods.

Our framework makes it possible to investigate important applications in different areas. More specifically, due to increased problem dimensionality, Monte Carlo simulation remains the method of choice for computing expected values of nonlinear functions of driving processes on several occasions, including cases of path-dependence, advanced stochastic volatility models and self-exciting point processes, where Fourier or Laplace-transform solutions are inexistent or slow to compute. Therefore, we start by considering the evaluation of European plain vanilla options and extend to further refinements of the proposed method applicable to path-dependent contracts, such as barriers and lookbacks. In the second main part of the paper, we explore applications in, typically, non-finance models. We cast a spotlight on the stochastic logistic model and present, for the sake of exemplification, a simulation case study of tumor growth having first calibrated the model to growth data of multicellular tumor spheroids. We also revisit the Ginzburg–Landau model and its simulation and demonstrate our method's capability under stressed volatility conditions.

The remainder of the paper is structured as follows. In Section 2 we present the Pearson curve fitting procedure, whereas in Section 3 we study its convergence to the true distribution with increasing shared moments. In Section 4, we introduce the various financial models and present various properties which will serve as the building block for our simulation scheme. In Section 5 we portray our random number generation mechanism using our moment-based probability distribution build-up, whereas in Section 6 we study the bias introduced in simulation estimators. Section 7 presents our numerical study focused on financial applications, whereas in Sections 8 and 9 we extend to other models and unfold the applicability of our simulation method via practical examples in areas such as oncology and bioeconomics. Section 10 concludes the paper. Supplementary results and various additional applications (e.g., early-exercisable contracts) are deferred to the e-companion.

2. Moment-based approximations to probability distributions: the case of Pearson curves

Consider a set Ω equipped with a σ-algebra $\mathcal{F}$. The random variable Ψ on Ω has cumulative distribution function G and moments $\mu_n := \mu_n(G) = \int_{\mathbb{R}} x^n dG(x)$. As said in the introduction, in this paper we focus primarily on Ψ representing (conditional) stochastic time integrals (see later Sections 4, 8 and 9). This is certainly not a restriction and, in fact, in Section EC.8 of the e-companion we present an additional popular application to the Carr–Geman–Madan–Yor (CGMY) model exemplifying this.

As part of this research, we have studied algorithms based on scale mixtures (see Hörmann *et al.*, 2004, p. 325), Cornish–Fisher series expansion (see Abramowitz and Stegun, 1968, p. 935), moment-based approximations using mixtures (Lindsay *et al.*, 2000), and Johnson and Pearson distribution fits by moments (see Devroye, 1986). Among these methods, unreported numerical results have highlighted the superior performance of the Pearson system. Indeed, albeit very fast, algorithms related to scale mixtures and series techniques have been quite inaccurate. The Gram–Charlier series might exhibit unusual behaviour, but also the Cornish–Fisher expansion may diverge as this is the Legendre inversion of the Edgeworth expansion of a distribution, which typically diverges when including many moments (see Fusai and Tagliani, 2002). The Pearson system also takes precedence over the Johnson family of distributions. This is attributed to Pearson family's simpler and faster procedure for member selection and parameter estimation for given moments, as opposed to Johnson which is more complicated based on Hill's algorithm (see Hill *et al.*, 1976);

sometimes the latter even fails to converge for certain parameter settings (see Simonato, 2011). Hence, we adhere to Pearson's system of distributions, which we will generally denote by $\tilde{G}$, as our choice for implementing moment matching and eventually drawing random numbers from.

Let $\tilde{g}(x)$ be the density function associated with $\tilde{G}$ satisfying the differential equation

$$\frac{d\tilde{g}(x)}{dx} = -\frac{c_0 + x}{c_1 + c_2 x + c_3 x^2}\tilde{g}(x). \tag{1}$$

Solving equation (1) yields well-defined density functions with general form

$$\tilde{g}(x) = \mathcal{C}\left(c_1 + c_2 x + c_3 x^2\right)^{-\frac{1}{2c_3}} \exp\left\{\frac{(c_2 - 2c_0c_3)\arctan\left(\frac{c_2+2c_3x}{\sqrt{4c_1c_3-c_2^2}}\right)}{c_3\sqrt{4c_1c_3 - c_2^2}}\right\}, \tag{2}$$

where $\mathcal{C}$ is the normalizing constant and $\{c_0, c_1, c_2, c_3\}$ the parameters that control the shape of the distribution. These are estimated in the distribution fitting using the first four finite raw moments $\{\mu_1, \mu_2, \mu_3, \mu_4\}$ and are given by

$$c_0 = c_2 := \frac{\sqrt{\beta\gamma}(\varepsilon + 3)}{10\varepsilon - 12\gamma - 18}, \quad c_1 := \frac{(4\varepsilon - 3\gamma)\beta}{10\varepsilon - 12\gamma - 18}, \quad c_3 := \frac{2\varepsilon - 3\gamma - 6}{10\varepsilon - 12\gamma - 18}, \tag{3}$$

where

$$\beta := \mu_2 - \mu_1^2, \quad \gamma := \frac{\left(\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3\right)^2}{\left(\mu_2 - \mu_1^2\right)^3}, \quad \varepsilon := \frac{\mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4}{\left(\mu_2 - \mu_1^2\right)^2}$$

are, respectively, the variance, squared skewness and kurtosis of the Pearson random variable. It is worth noting that, although numerical methods exist making possible the Pearson system to fit more than four moments, they inevitably impact the computational effort (see Rose and Smith, 2002, Chapter 5) therefore we do not consider them here. In Section 7, we study the closeness of the higher-order moments of a four-moment Pearson curve fit to the true ones and favourably show that, in practice, the differences remain very small also under challenging model parameterizations. Given knowledge of the first four integer moments, we classify the Pearson distribution family types according to Johnson *et al.* (1994) being standard in the literature. We make a selection based on the η-criterion proposed by Elderton and Johnson (1969): given $\sqrt{\gamma}$ and ε, we compute

$$\eta := \frac{\gamma(\varepsilon + 3)^2}{4(4\varepsilon - 3\gamma)(2\varepsilon - 3\gamma - 6)}.$$

In particular, we get from it the *main* types I ($\eta < 0$), IV ($0 < \eta < 1$) and VI ($\eta > 1$); and the *transition* types, i.e., normal ($\eta = 0, \varepsilon = 3$), II ($\eta = 0, \varepsilon < 3$), III ($\eta = \pm\infty$), V ($\eta = 1$) and VII ($\eta = 0, \varepsilon > 3$).

3. Proximity of distributions with shared moments

A natural question that arises once we obtain the Pearson fit $\tilde{G}$ is how well it approximates the true distribution G with common moments. Our task next is to consider convergence under various notions of distance and provide ways to estimate the closeness in different moment-based metrics.

3.1. Lévy distance

Definition 1. *In the space of distribution functions of random variables, the Lévy distance (Lévy, 1925) between any pair of elements $G(x)$ and $\tilde{G}(x)$ is defined as*

$$L\left(G,\tilde{G}\right)=\inf\left\{\epsilon:\ G(x-\epsilon)-\epsilon\leq\tilde{G}(x)\leq G(x+\epsilon)+\epsilon\ \text{for all}\ x\right\}.$$

To measure the distance between the two distribution functions in terms of their characteristic functions, $\varphi_G(u)$ and $\varphi_{\tilde{G}}(u)$, Zolotarev's metric (Zolotarev and Senatov, 1976) can be employed

$$L\left(G,\tilde{G}\right)=\min_{T>0}\max\left\{\frac{1}{2}\max_{|u|\leq T}\left|\varphi_G(u)-\varphi_{\tilde{G}}(u)\right|,\frac{1}{T}\right\}.$$

The Lévy distance and Zolotarev's metric are equivalent in the sense that they induce the same topology. We quote also the following result from Zolotarev (1970). For any $\mathcal{T}>e$, the Lévy distance satisfies the inequality

$$L\left(G,\tilde{G}\right)\leq\frac{1}{\pi}\int_0^{\mathcal{T}}\left|\frac{\varphi_G(u)-\varphi_{\tilde{G}}(u)}{u}\right|du+\frac{2e\ln\mathcal{T}}{\mathcal{T}}. \tag{4}$$

This useful result allows us to measure the closeness of our Pearson curve fit $\tilde{G}$ and the true distribution G via their characteristic functions. The latter is available from the literature for different models of interest (see later), whereas in the next result we focus on the Pearson characteristic function.

Proposition 1. *Let $\varphi_{\tilde{G}}(u)=\int_{\mathbb{R}}e^{iux}\tilde{g}(x)\,dx$ be the characteristic function of $\tilde{G}$, where $i:=\sqrt{-1}$. This then satisfies the system of ordinary differential equations*

$$\begin{cases}\varphi'_{\tilde{G}}(u)=\vartheta_{\tilde{G}}(u)\\ \vartheta'_{\tilde{G}}(u)=\frac{uc_2+i(1-2c_3)}{iuc_3}\vartheta_{\tilde{G}}(u)+\frac{c_1}{c_3}\varphi_{\tilde{G}}(u)\end{cases}, \tag{5}$$

with set of initial conditions $\left(\varphi_{\tilde{G}}(0),\vartheta_{\tilde{G}}(0)\right)=(1,i\mu_1)$.

Proof. See e-companion Section EC.1. ■

An analytic solution of (5) can be obtained easily with the aid of a software system that allows symbolic computations, such as Mathematica, which, in the interest of space, we omit here.

Next, we present different metrics which give a gauge of the closeness of the target and approximating distributions based on a certain number of moments. Let the distribution functions $G(x)$ and $\tilde{G}(x)$ satisfy

$$\mu_n(G)=\mu_n\left(\tilde{G}\right)=\mu_n,\quad n=0,1,\ldots,2m\quad(m\geq 2). \tag{6}$$

In the first instance, in our case of the four-moment fit, (6) holds for $m=2$. From Klebanov and Mkrtchyan (1986), the estimate of closeness in the L-metric can be expressed in terms of the truncated Carleman's series $\beta_{k,m}:=\sum_{n=k+1}^{m}\left(\mu_{2n}/\mu_{2k}\right)^{-1/(2(n-k))}$. Since divergence of Carleman's series is a sufficient condition for the moment problem to be determinate (Akhiezer, 1965), it is natural to seek an estimate of closeness in terms of β^{-1} which gets smaller for larger m. Based on this and Zolotarev's inequality (4), we have a further inequality.

Theorem 2 (Klebanov and Mkrtchyan, 1986)**.** *Let G and $\tilde{G}$ be two distribution functions for*

which $\beta_m \equiv \beta_{0,m}$ *holds. Then, there exists a constant* C_{μ_2} *dependent on* μ_2 *such that*

$$L\left(G,\tilde{G}\right) \leq \frac{C_{\mu_2}\ln\left(1+\beta_{m-1}\right)}{\beta_{m-1}^{1/4}}. \tag{7}$$

A sharpened version of inequality (7) and a variant when additional moments of the two distributions are not coinciding but are only fairly close can be found in Section EC.2 of the e-companion.

3.2. Uniform distance

In addition to the Lévy distance, we consider here the topology given by the uniform metric (Kolmogorov, 1933, Zolotarev, 1983) between two distributions G and $\tilde{G}$,

$$\rho\left(G,\tilde{G}\right) = \sup_x \left|G\left(x\right) - \tilde{G}\left(x\right)\right|.$$

The Lévy and uniform metrics are linked as follows.

Lemma 3 (Linnik and Ostrovskiĭ, 1977)**.** *For all distribution functions* G *and* $\tilde{G}$*,*

$$L\left(G,\tilde{G}\right) \leq \rho\left(G,\tilde{G}\right) \leq \left(1+\varrho\right) L\left(G,\tilde{G}\right), \tag{8}$$

where $\varrho = \sup_x \tilde{G}'\left(x\right)$ *if* $\tilde{G}$ *is absolutely continuous.*

Lemma 3 applies in the Pearson case upon existence of density $\tilde{g}$ (2) which implies absolute continuity. Combining (7) and (8) gives a bound to the uniform distance between G and $\tilde{G}$. Related to this distance is also Esseen's inequality recorded in Section EC.2 of the e-companion. Next, we present a classical bounding result due to Akhiezer (1965, Corollary 2.5.4), which is revisited by Lindsay and Basak (2000, Theorems 1, 2).

Theorem 4 (Lindsay and Basak, 2000)**.** *Let any two arbitrary distributions* $G\left(x\right)$ *and* $\tilde{G}\left(x\right)$ *satisfy (6). Then, for all the values of* x*,*

$$\left|G\left(x\right) - \tilde{G}\left(x\right)\right| \leq \left\{P_m'\left(x\right) W_m^{-1} P_m\left(x\right)\right\}^{-1} =: \varpi_m\left(x\right), \tag{9}$$

where $P_m\left(x\right) := \left(1, x, x^2, \ldots, x^m\right)'$ *and* $W_m := \left\|\mu_{i+j}\right\|_{i,j=0}^{m}$ *is a Hankel symmetric matrix defined by the first* $2m$ *moments.*

Bound (9) goes to 0 at the rate x^{-2m} as $x \to \infty$ giving relatively sharp tail information. An improvement to (9) is due to Khamis (1954) who introduces a constant non-negative multiplier that is smaller than the unity and is given by $1 + \min\left\{\text{l.u.b}_{c\leq x\leq d}\frac{-G'(x)}{\tilde{G}'(x)}, \text{l.u.b}_{c\leq x\leq d}\frac{-\tilde{G}'(x)}{G'(x)}\right\}$ if it exists.

All the above metrics between probability measures can be verified in our case of the Pearson distribution with some shared moments and admit practical bounding techniques for the approximation error.

3.3. An entropy bound

In problems where the key question is the choice of an approximating distribution, the use of the maximum entropy principle is also popular. Assuming we have some known moments, this

principle suggests to select, among the distributions consistent with such partial information, the one with maximum entropy, that is, the most uncertain.

For a generic continuous distribution function H with support D and associated density function $h(x)$, we define the differential entropy $\mathcal{H}[h] = -\int_D h(x) \ln h(x)\, dx$. Also, in addition to our original approximating Pearson distribution function $\tilde{G}$ with associated density function $\tilde{g}$, we consider the entropy-maximizing distribution $\hat{G}$ with density function

$$\hat{g}(x) := \exp\left(-\sum_{n=0}^{m} \lambda_n x^n\right), \tag{10}$$

where $\{\lambda_n\}$ are the Lagrange multipliers (Kapur and Kesavan, 1992). This is obtained by maximizing the entropy constrained by moments. The resulting distribution $\hat{G}$ shares the same first m moments with the target true distribution G with density g.

Our aim is to derive an entropy bound to the absolute difference of each of $H \in \left\{\hat{G}, \tilde{G}\right\}$ with respect to the true G. To this end, we consider two measures of difference between density functions, that is, the divergence and variation measures

$$\mathcal{I}[g,h] := \int_D g(x) \ln \frac{g(x)}{h(x)} dx \quad \text{and} \quad \mathcal{V}[g,h] := \int_D |h(x) - g(x)|\, dx,$$

respectively. Next, we present a lower bound for the divergence measure, although more complicated ones exist in the literature.

Lemma 5 (Kullback, 1967)**.** *We have that*

$$\mathcal{I} \geq \frac{\mathcal{V}^2}{2} + \frac{\mathcal{V}^4}{36}. \tag{11}$$

Based on the above, we can restate the uniform metric as follows.

Theorem 6. *For the distributions G and $H \in \left\{\hat{G}, \tilde{G}\right\}$,*

$$|H(x) - G(x)| \leq 3\sqrt{-1 + \sqrt{1 + \frac{4\mathcal{I}[g,h]}{9}}}. \tag{12}$$

In addition, $\mathcal{I}[g,\hat{g}] = \sum_{n=0}^{m} \lambda_n \mu_n(G) - \mathcal{H}[g]$.

Proof. Inequality (12) follows from $|H(x) - G(x)| \leq \mathcal{V}[g,h]$ and by solving (11).

In addition, for $H \equiv \hat{G}$, we have

$$\mathcal{I}[g,\hat{g}] = \int_D g(x) \ln \frac{g(x)}{\hat{g}(x)} dx = -\mathcal{H}[g] - \int_D g(x) \ln \hat{g}(x)\, dx = -\mathcal{H}[g] + \sum_{n=0}^{m} \lambda_n \int_D x^n \hat{g}(x)\, dx,$$

where the third equality is due to (10) and from which the result follows. ■

In this section, we have established some estimates of closeness of probability laws, which we will revisit later in Sections 6 and 7. We move now to our first application of the Pearson curve approximation and resulting simulation scheme to stochastic volatility models.

4. Stochastic volatility models

Let $(\Omega, \mathcal{F}, Q, \{\mathcal{F}_t\})$ be a filtered probability space where the filtration satisfies the usual conditions with $\mathcal{F}_0$ trivial. This filtered probability space supports all the processes we encounter in the sequel and Q denotes the risk neutral probability measure. We begin with the general class of affine bivariate stochastic volatility models.

4.1. Bivariate affine models

The Heston model (Heston, 1993) is a classical member of the class of affine stochastic volatility models (Duffie *et al.*, 2000, Duffie *et al.*, 2003):

$$dS(t) = rS(t)dt + \sigma(t)S(t)(\rho dW_2(t) + \sqrt{1-\rho^2}dW_1(t)), \tag{13}$$

$$d\sigma^2(t) = k(\theta - \sigma^2(t))dt + v\sigma(t)dW_2(t), \tag{14}$$

where W_1 and W_2 are two independent standard Brownian motions; r is the continuously compounded risk-free interest rate. $(S(t), t \geq 0)$ denotes the asset price process. $(\sigma(t), t \geq 0)$ is the instantaneous variance process described by a Cox *et al.* (1985) (CIR) square-root diffusion with constant parameters θ, k, v and instantaneous correlation $\rho \in [-1, 1]$ between the two processes. If the Feller condition, $2k\theta \geq v^2$, is satisfied, the zero boundary is unattainable by process (14); otherwise, it is attracting and attainable. At the zero boundary though, the process is immediately reflected into the positive domain. The variance transition is given by

$$\sigma^2(t) \overset{\text{(law)}}{=} \frac{v^2(1-e^{-k(t-u)})}{4k}\chi'^2_d(\lambda), \tag{15}$$

where $\chi'^2_d(\lambda)$ is a noncentral chi-squared random variable with $d := 4\theta k/v^2$ degrees of freedom and noncentrality parameter

$$\lambda := \frac{4ke^{-k(t-u)}\sigma^2(u)}{v^2(1-e^{-k(t-u)})}.$$

By substituting (14) in (13), we get on the log-scale that

$$\begin{aligned}\ln S(t) &= \ln S(u) + r(t-u) - \frac{1}{2}\int_u^t \sigma^2(s)ds \\ &\quad + \frac{\rho}{v}\left(\sigma^2(t) - \sigma^2(u) - k\theta(t-u) + k\int_u^t \sigma^2(s)ds\right) + \widehat{W}_1\left((1-\rho^2)\int_u^t \sigma^2(s)ds\right),\end{aligned}$$

where $\widehat{W}_1$ is a standard Brownian motion independent of $\int_u^t \sigma^2(s)ds$. Therefore, $\widehat{W}_1\left((1-\rho^2)\int_u^t \sigma^2(s)ds\right)$ is an Ocone martingale, is process symmetric (see Rheinländer and Schmutz, 2013) and conditionally strong Markov (see Karatzas and Shreve, 1991). Conditionally on $\sigma^2(t)$ and $\int_u^t \sigma^2(s)ds$, we then have that

$$\left(\ln S(t) \middle| \ln S(u), \sigma^2(u), \sigma^2(t), \int_u^t \sigma^2(s)ds\right) \sim \mathcal{N}(m_{u,t}, s^2_{u,t}),$$

where

$$m_{u,t} := \ln S(u) + r(t-u) - \frac{1}{2}\int_u^t \sigma^2(s)ds + \frac{\rho}{v}\left(\sigma^2(t) - \sigma^2(u) - k\theta(t-u) + k\int_u^t \sigma^2(s)ds\right),$$

$$s^2_{u,t} := (1-\rho^2)\int_u^t \sigma^2(s)ds.$$

Extended model constructions with independent jumps in the asset price process (Bates, 1996), contemporaneous asset price and variance jumps of correlated magnitudes (Duffie *et al.*, 2000) and time-changed Lévy models (Carr *et al.*, 2003) are well known, hence, in the interest of space, we omit their details here and defer to Section 5.3 some more notes on their simulation.

4.2. SABR

In this model, proposed by Hagan *et al.* (2002), the forward asset price dynamics follows a constant elasticity of variance (CEV) diffusion process with volatility evolving according to a driftless geometric Brownian motion:

$$\begin{aligned} dS(t) &= \sigma(t) S^{\beta}(t) \left(\rho dW_2(t) + \sqrt{1-\rho^2} dW_1(t) \right), \\ d\sigma(t) &= v\sigma(t) dW_2(t), \end{aligned}$$

where $\beta \in [0,1]$ is constant. In this case, we straightforwardly have that

$$\ln \sigma(t) \sim \mathcal{N}\left(\ln \sigma(u) - \frac{1}{2} v^2 (t-u), v^2 (t-u) \right).$$

Also, conditionally on $\left(\sigma^2(t), \int_u^t \sigma^2(s) ds \right)$, we distinguish among several cases of describing the distribution of $\ln S$. For $\beta = 1$,

$$\left(\ln S(t) \middle| \ln S(u), \sigma(u), \sigma^2(t), \int_u^t \sigma^2(s) ds \right) \sim \mathcal{N}(m_{u,t}, s_{u,t}^2),$$

where

$$m_{u,t} := \ln S(u) - \frac{1}{2} \int_u^t \sigma^2(s) ds + \frac{\rho}{v} (\sigma(t) - \sigma(u)), \quad s_{u,t}^2 := (1-\rho^2) \int_u^t \sigma^2(s) ds.$$

For $\beta \in [0,1)$, $\rho = 0$ and $(S(t), t \geq 0)$ with an absorbing boundary at 0 (Islah, 2009, Cai *et al.*, 2017),

$$\begin{aligned} P\left(S(t) = 0 \middle| \sigma(u), \sigma(t), \int_u^t \sigma^2(s) ds, S(u) \right) &= 1 - Q_{\chi^2}\left(A_0; \frac{1}{1-\beta} \right), \\ P\left(S(t) \leq y \middle| \sigma(u), \sigma(t), \int_u^t \sigma^2(s) ds, S(u) \right) &= 1 - Q_{\chi'^2}\left(A_0; \frac{1}{1-\beta}, C_0(y) \right) \end{aligned}$$

for any $y > 0$, where

$$A_0 := \left(\int_u^t \sigma^2(s) ds \right)^{-1} \left(\frac{S(u)^{1-\beta}}{1-\beta} \right)^2, \quad C_0(y) := \left(\int_u^t \sigma^2(s) ds \right)^{-1} \left(\frac{y^{1-\beta}}{1-\beta} \right)^2,$$

$Q_{\chi^2}(\cdot; d)$ and $Q_{\chi'^2}(\cdot; d, \lambda)$ are, respectively, the chi-squared and noncentral chi-squared cumulative distribution functions. Finally, for $\beta \in [0,1)$, $\rho \neq 0$ and $(S(t), t \geq 0)$ with an absorbing boundary

at 0,

$$P\left(S(t)=0\,\middle|\sigma(u),\sigma(t),\int_u^t\sigma^2(s)ds,S(u)\right)\approx 1-Q_{\chi^2}\left(A;1+\frac{\beta}{(1-\beta)(1-\rho^2)}\right),$$
$$P\left(S(t)\le y\,\middle|\sigma(u),\sigma(t),\int_u^t\sigma^2(s)ds,S(u)\right)\approx 1-Q_{\chi'^2}\left(A;1+\frac{\beta}{(1-\beta)(1-\rho^2)},C(y)\right)$$

for $y>0$, where

$$A:=\frac{\left(\frac{S(u)^{1-\beta}}{1-\beta}+\frac{\varrho}{v}(\sigma(t)-\sigma(u))\right)^2}{(1-\rho^2)\int_u^t\sigma^2(s)ds},\quad C(y):=\frac{\left(\frac{y^{1-\beta}}{1-\beta}\right)^2}{(1-\rho^2)\int_u^t\sigma^2(s)ds}.$$

4.3. Ornstein–Uhlenbeck driven stochastic volatility

This model started becoming known since the early works of Scott (1987) and Stein and Stein (1991). The original assumption of zero correlation between asset price and volatility was relaxed later by Schöbel and Zhu (1999). More specifically, the volatility process $\sigma(t)$ is represented by a Gaussian OU model:

$$dS(t)=rS(t)dt+\sigma(t)S(t)\left(\rho dW_2(t)+\sqrt{1-\rho^2}dW_1(t)\right),$$
$$d\sigma(t)=k(\theta-\sigma(t))dt+vdW_2(t).$$

In this case, we have (e.g., see Li and Wu, 2019) that

$$\left(\sigma(t),\int_u^t\sigma(s)ds\right)\sim\mathcal{N}_2\left(\underline{\mu}_{u,t},\Sigma_{u,t}\right),$$

where $\mathcal{N}_2$ is a bivariate normal distribution with mean vector and covariance matrix

$$\underline{\mu}_{u,t}:=\begin{pmatrix}(\sigma(u)-\theta)(1-\zeta_{u,t})+\theta\\ \theta(t-u)+\frac{(\sigma(u)-\theta)\zeta_{u,t}}{k}\end{pmatrix}\quad\text{and}\quad\Sigma_{u,t}:=\begin{pmatrix}\frac{v^2\zeta_{u,t}(2-\zeta_{u,t})}{2k} & \frac{v^2\zeta_{u,t}^2}{2k^2}\\ \frac{v^2\zeta_{u,t}^2}{2k^2} & \frac{-v^2\zeta_{u,t}^2+2kv^2\left(t-u-\frac{\zeta_{u,t}}{k}\right)^2}{2k^3}\end{pmatrix},$$

respectively, for $\zeta_{u,t}:=1-\exp\left(-k\left(u-t\right)\right)$. Then,

$$\left(\ln S(t)\,\middle|\,\ln S(u),\sigma(u),\int_u^t\sigma(s)ds,\int_u^t\sigma^2(s)ds\right)\sim\mathcal{N}(m_{u,t},s_{u,t}^2),$$

with

$$m_{u,t}:=\ln S(u)+\left(r-\frac{\rho v}{2}\right)(t-u)+\frac{\rho}{2v}(\sigma^2(t)-\sigma^2(u))-\frac{\rho k\theta}{v}\int_u^t\sigma(s)ds+\left(\frac{\rho k}{v}-\frac{1}{2}\right)\int_u^t\sigma^2(s)ds,$$
$$s_{u,t}^2:=(1-\rho^2)\int_u^t\sigma^2(s)ds.$$

4.4. Double Heston model

Aiming to improve the empirical performance of the Heston model, Christoffersen *et al.* (2009) proposed the multivariate extension

$$\begin{aligned} dS(t) &= rS(t)dt + \sum_{j=1}^{2} \sigma_j(t)S(t)\left(\rho_j dW_{j+2}(t) + \sqrt{1-\rho_j^2}dW_j(t)\right), \\ d\sigma_j^2(t) &= k_j(\theta_j - \sigma_j^2(t))dt + v_j\sigma_j(t)dW_{j+2}(t), \quad j = 1, 2, \end{aligned}$$

where the standard Brownian motions $\{W_j\}_{j=1}^4$ are mutually independent. Similarly to the original Heston model, we have that

$$\sigma_j^2(t) \overset{\text{(law)}}{=} \frac{v_j^2(1-e^{-k_j(t-u)})}{4k_j}\chi_{d_j}^2(\lambda_j)$$

and

$$\left(\ln S(t)\middle|\, \ln S(u), \{\sigma_j^2(t)\}_{j=1}^2, \left\{\int_u^t \sigma_j^2(s)ds\right\}_{j=1}^2\right) \sim \mathcal{N}(m_{u,t}, s_{u,t}^2),$$

where

$$m_{u,t} := \ln S(u) + r(t-u) + \sum_{j=1}^{2}\left(\rho_j\int_u^t \sigma_j(s)dW_{j+2}(s) - \frac{1}{2}\int_u^t \sigma_j^2(s)ds\right), \quad s_{u,t}^2 := \sum_{j=1}^{2}(1-\rho_j^2)\int_u^t \sigma_j^2(s)ds.$$

4.5. 4/2 and 3/2 models

Finally, Grasselli (2017) introduced another generalization of the Heston model, the so-called 4/2 stochastic volatility model, where

$$\begin{aligned} dS(t) &= rS(t)dt + S(t)\left(\alpha\sigma(t) + \frac{\beta}{\sigma(t)}\right)(\rho dW_2(t) + \sqrt{1-\rho^2}dW_1(t)), \\ d\sigma^2(t) &= k(\theta - \sigma^2(t))dt + v\sigma(t)dW_2(t). \end{aligned}$$

In this case,

$$\left(\ln S(t)\middle|\, \ln S(u), \sigma^2(u), \sigma^2(t), \int_u^t \sigma^2(s)ds, \int_u^t \frac{ds}{\sigma^2(s)}\right) \sim \mathcal{N}(m_{u,t}, s_{u,t}^2),$$

where

$$\begin{aligned} m_{u,t} &:= \ln S(u) + \left(r - \alpha\beta - \frac{\alpha\rho k\theta}{v} + \frac{\beta\rho k}{v}\right)(t-u) + \frac{\alpha\rho}{v}\left(\sigma^2(t) - \sigma^2(u)\right) + \frac{\beta\rho}{v}\ln\frac{\sigma^2(t)}{\sigma^2(u)} \\ &\quad + \left(\frac{\alpha\rho k}{v} - \frac{\alpha^2}{2}\right)\int_u^t \sigma^2(s)ds + \left[\frac{\beta\rho}{v}\left(\frac{v^2}{2} - k\theta\right) - \frac{\beta^2}{2}\right]\int_u^t \frac{ds}{\sigma^2(s)}, \\ s_{u,t}^2 &:= (1-\rho^2)\left(\alpha^2\int_u^t \sigma^2(s)ds + \beta^2\int_u^t \frac{ds}{\sigma^2(s)} + 2\alpha\beta(t-u)\right). \end{aligned}$$

This model reduces to the original Heston when $\beta = 0$ and $\alpha = 1$. Instead, when $\alpha = 0$, the model coincides with the 3/2 model which was proposed independently by Heston (1997) and Platen (1997).

5. Moment-based random number generator

Key quantity in the models presented in the previous section is the conditional time integral of a function of the variance. To highlight this commonality across the different models, we will denote it by $\Psi(u,t)$ and the relevant conditioning arguments by $\Phi(u,t)$. Their exact forms for the models of Section 4 are presented in Table 1.

[Insert Table 1]

The simulation of the models can then be summarized as follows:

1. Simulate $(\Phi(u,t)|\sigma(u))$
2. Simulate $(\Psi(u,t)|\sigma(u),\Phi(u,t))$
3. Simulate $(\ln S(t)|\ln S(u),\sigma(u),\Phi(u,t),\Psi(u,t))$.

Steps 1 and 3 are trivial as we have seen that the conditional distributions are known; what remains a challenge is the generation of conditional samples from $\Psi(u,t)$ in step 2. (The case of the double Heston model is a mere extension of the original Heston by independence of the variance processes.) Hitherto in the literature, exact simulation of the conditional $\Psi(u,t)$ relies on numerical inversion of its Laplace transform $\mathcal{L}(a) := E\left[\exp\left(-a\Psi(u,t)\right)|\Phi(u,t)\right]$ (refer to the e-companion Section EC.3 for the various models), which is nevertheless the hardest and most time-consuming step of the whole simulation scheme. For this reason, typical exact simulation schemes are very slow when used to simulate entire trajectories due to multiple numerical inversion of the conditional Laplace transform of $\Psi(u,t)$ at each time u. For ease of notation when referring to the conditional distribution of $\Psi(u,t)$, we focus hereafter on the condition on $\Phi(u,t)$, which is the key feature in our problem, with the rest of the information up to time u incorporated.

5.1. Random number generation based on fitted Pearson curves

We aim to circumvent the above stumbling block by proposing a new approach to simulating $(\Psi(u,t)|\Phi(u,t))$ relying on fitting a Pearson curve to the corresponding theoretical distribution introduced in Section 2. A major advantage of this, as we will ascertain in our numerical study in Section 7, is that it gives excellent results for different models and parameter sets. This is explained, as we have shown in Section 3, by convergence of the approximating distribution under various notions of distance. In addition, whilst we are exactly fitting only four moments, in practice higher moments of the two distributions are very close, as we will show in Section 7.1, resulting in even sharpened bounds, hence high precision. This is because the moment problem is determinate, i.e., the distribution of $(\Psi(u,t)|\Phi(u,t))$ is uniquely determined by the sequence of its moments. A criterion for this is the existence of the moment generating function for all $a \in (-c,c)$, where $c>0$ (see Stoyanov, 2013). From Section EC.3 of the e-companion, we have, from (EC.4) for Heston, that $a \geq -k^2/2v^2$; from (EC.5) for SABR, $a \geq -(\sigma(t)-\sigma(u))^2/2v^2$; from (EC.6) for OU-SV, $a \geq -k^2/2v^2$; and for 4/2 from (EC.7), $a \geq -k^2/2v^2$ and $b \geq -(2\theta k - v^2)^2/8v^2$.

Once the Pearson curve type is determined, we can draw random numbers from it. First, we generate random numbers $\hat{Y}$ from the selected standardized family with zero mean and unitary variance, based on the procedure summarized in Section EC.4 of the e-companion (see Johnson

et al., 1994 for more details). Then, we rescale and shift using the true mean, μ_1, and standard deviation, $\beta^{1/2}$, to obtain

$$Y := \mu_1 + \beta^{1/2}\hat{Y} \tag{16}$$

corresponding to a random sample from $(\Psi(u,t)|\Phi(u,t))$.

5.2. Computation of moments

Associated with our moment-based technique is the efficient computation of the integer moments. The n-th conditional moment of $\Psi(u,t)$ is traditionally given by

$$\mu_n = (-1)^n \left.\frac{\partial^n}{\partial a^n}\mathcal{L}(a)\right|_{a=0}. \tag{17}$$

However, computing the moments from (17) may not be practicable, especially for high-order moments, as the Laplace transform may involve special functions as in the majority of the cases examined in this paper. For example, consider the relevant Laplace transform (EC.4) for the Heston model. Evaluating the first four moments of $\Psi(u,t)$ from (17), using, e.g., the symbolic toolbox of Mathematica, requires 307 Bessel function evaluations, which is highly computational intensive but also endangers numerical errors.

In this paper, we bypass such kind of problem by numerical inversion of an adaptively modified moment generating function introduced by Choudhury and Lucantoni (1996). According to this,

$$\begin{aligned}
\mu_n &= \frac{n!}{2nlr_n^n\alpha_n^n}\left\{\mathcal{L}(\alpha_n r_n) + (-1)^n\mathcal{L}(-\alpha_n r_n) + 2\sum_{j=1}^{nl-1}\Re(\mathcal{L}(\alpha_n r_n e^{\pi ij/nl})e^{-\pi ij/l})\right\} - \hat{e}_n, \quad (18)\\
\hat{e}_n &:= \sum_{j=1}^{\infty}\alpha_n^{2ljn}\frac{n!}{(n+2ljn)!}\mu_{n+2ljn}10^{-\gamma j},
\end{aligned}$$

where $\Re(z)$ denotes the real part of z and $\hat{e}_n$ is the error term. As Choudhury and Lucantoni (1996) report, the choice $r_n := 10^{-\gamma/(2nl)}$ is made in order to bound the error and achieve accuracy of the order $10^{-\gamma}$. Algorithm 1 summarizes the procedure for computing any m integer moments as well as the parameter l and the adaptive α_n.

[Insert Algorithm 1]

The proposed method has several merits. First, it is fast: for the first four moments, the Laplace transform is evaluated 14 times = 2 (for μ_1) + 3 (for μ_2) + 4 (for μ_3) + 5 (for μ_4) (for $l = 1$ used in equation 18). Hence, for (EC.4), the Bessel function is evaluated just 28 times, which is a considerable reduction from 307 times in (17). Second, it is very accurate: using $\gamma = 11$ as per the recommendation of Choudhury and Lucantoni (1996), the error appears, consistently with them, only in the eleventh to thirteenth significant place. This is generally easy to verify by calculating just the first few terms in $\hat{e}$ based on the true moments, as the infinite series is heavily damped and the error cannot be significant. Finally, the Laplace transform can be evaluated for the whole sample of $\Phi(u,t)$ draws altogether (as opposed to one by one), which we favour in our application requiring multiple moment computations for different random realizations of $\Phi(u,t)$ due to dependence of the moments of $\Psi(u,t)$ on $\Phi(u,t)$.

5.3. Summary of the simulation method and extensions

Our method still hinges on the Laplace transform, but circumvents its numerical inversion in virtue of the speedy computation of the moments and simulation of the Pearson proxy. In Section 7, it is shown that the computing time is drastically reduced, benefiting the simulation of the asset price process on a set of multiple observation dates. We summarize the methodology in Algorithm 2 based on sampling from (16). (For martingale-correction of the discounted asset price trajectory, see Andersen, 2008, Proposition 7 and Glasserman, 2004, equation 4.60.) If we care to simulate the terminal asset price only, we use $m = 1$.

[Insert Algorithm 2]

The 4/2 model (see also Section 4.5) requires some extra care as it requires sampling from $(\Psi(u,t)|\Phi(u,t)) \equiv \left(\int_u^t \sigma^2(s)ds, \int_u^t \frac{ds}{\sigma^2(s)} \middle| \sigma^2(t) \right)$. In Section EC.5 of the e-companion, we present an extension of the framework described above to the multivariate case, in particular the bivariate case of relevance here. Alternatively, the sampling problem can be simplified by entirely bypassing the time integrals using directly the conditional Laplace transform $E\left[\exp\left(-a \ln S(t)\right)| \ln S(u), \sigma^2(u), \sigma^2(t)\right]$ from Grasselli (2017, Proposition 4.1).

A few more cases of models with jumps ensue. Extending to the Bates and DPS models is straightforward, following Broadie and Kaya (2006, Sections 6.1–6.2), by replacing the time integral with our usual four-moment Pearson fit. Similar logic applies to the simulation of the 4/2 model extension with independent jumps of Lin *et al.* (2017). In addition, we have a couple of cases that deserve some more attention.

5.3.1. Time-changed Lévy models

A Lévy model with integrated CIR time-change is given by

$$\ln S(t) = \ln S(u) + r(t-u) - \psi(1) \int_u^t \sigma^2(s)ds + L\left(\int_u^t \sigma^2(s)ds \right),$$

where $\sigma(t)$ is a square-root diffusion process, L is a Lévy process and $\psi(u) := \ln E\left(e^{uL(1)}\right)$. Having first simulated $\int_u^t \sigma^2(s)ds$ according to the fitted Pearson distribution, we then simulate the Lévy process L on the new (stochastic) time scale. The simulation of standard Lévy processes including the variance gamma or normal inverse Gaussian is trivial (e.g., see Cont and Tankov, 2004); the case of the CGMY process is much harder (e.g., see Ballotta and Kyriakou, 2014) and one can use the new technique we develop in Section EC.8 of the e-companion based on a new Pearson curve fit.

5.3.2. Jump-diffusion model with state-dependent jump intensity

Next, we draw attention to the class of self-exciting point processes, in particular, a jump-diffusion process with state-dependent drift, diffusion coefficient and jump intensity:

$$d\lambda(t) = k_\lambda(\theta_\lambda - \lambda(t))dt + \sigma_\lambda \sqrt{\lambda(t)} dW(t) + dJ(t), \tag{19}$$

where $k_\lambda, \theta_\lambda, \sigma_\lambda$ are constant, $J(t) := \beta N(t)$ is a jump process given by a compound self-exciting point process with CIR intensity, $N(t) := \sum_{i \geq 1} \mathbf{1}_{\{\bar{T}_i < t\}}$ is the total number of jumps and $\bar{T}_i$ is the i-th jump time (see Dassios and Zhao, 2017 for more details). Model (19) accounts for jumps, for example, in the asset price process, that are not uniformly distributed over time, but tend to appear

in clusters (e.g., see Fulop and Li, 2019 and Du and Luo, 2019). For the sake of exemplification, we will assume that

$$\ln S(t) = \ln S(u) + \left(r - \frac{\sigma^2}{2}\right)(t-u) - \omega \int_u^t \lambda(s)ds + \sigma W(t) + \sum_{i=1}^{N(t)} J_i \tag{20}$$

for jump sizes $J \sim \mathcal{N}\left(\mu_J, \sigma_J^2\right)$ and $\omega := \exp\left(\mu_J + \frac{1}{2}\sigma_J^2\right) - 1$. This model can be seen as a self-exciting extension of the one proposed in Wachter (2013) with $\beta = 0$. It is also a special case with constant variance σ^2 of the affine model in Fulop and Li (2019, Model III). (Our method is adaptable to the full model specification; upon assuming stochastic variance with jumps, the simulation task becomes similar to that of the DPS model.)

Giesecke *et al.* (2011) and Dassios and Zhao (2017) show us how to simulate exactly the pairs $\left\{\left(\bar{T}_i, \lambda(\bar{T}_i)\right)\right\}_{i=1}^{N(t)}$ for $u < \bar{T}_i < t \ \forall i > 0$; nevertheless, on several occasions, such as (20), we need to be able to simulate also

$$\left(\int_{\bar{T}_i}^{\bar{T}_{i+1}} \lambda(s)ds \middle| \lambda(\bar{T}_i), \lambda(\bar{T}_{i+1})\right). \tag{21}$$

Between two consecutive jump times $\bar{T}_i$ and $\bar{T}_{i+1}$, the process λ evolves as a CIR diffusion and our method lends itself to the efficient simulation of (21). Then, we have that

$$\int_u^t \lambda(s)ds = \sum_{i=0}^{N(t)} \int_{\bar{T}_i}^{\bar{T}_{i+1}} \lambda(s)ds + \int_{\bar{T}_{N(t)}}^t \lambda(s)ds$$

and

$$\left(\ln S(t) \middle| \ln S(u), \int_u^t \lambda(s)ds, \sum_{i=1}^{N(t)} J_i\right) \sim \mathcal{N}(m_{u,t}, s_{u,t}^2)$$

with

$$m_{u,t} := \ln S(u) + \left(r - \frac{1}{2}\sigma^2\right)(t-u) - \omega \int_u^t \lambda(s)ds + \sum_{i=1}^{N(t)} J_i, \quad s_{u,t}^2 := \sigma^2(t-u).$$

6. Error propagation to target simulation output

Having fitted a Pearson curve type $\tilde{G}$ to the distribution G of $(\Psi|\Phi)$, what we are further interested in is how any error from this distributional approximation, which we have introduced in Section 3, translates to the ultimate simulation estimators for quantities of interest, such as a derivative's price. In this section, we take a step forward in that direction and study how the bias introduced into the simulation estimator is bounded.

To highlight the dependence on Φ with marginal density function f_Φ as required in what follows, we augment accordingly the original notation used in Sections 2–3 for the target and approximating distribution functions of Ψ conditional on Φ, $G(\cdot|\,\Phi)$ and $\tilde{G}(\cdot|\,\Phi)$, respectively. Suppose that we are interested in estimating $E[\Pi]$ based on G using $\tilde{E}[\Pi]$ based on $\tilde{G}$; g and $\tilde{g}$ are the associated conditional density functions. Then, it is always true that

$$E[\Pi] = E[E[\Pi|\,\Psi = x, \Phi = \phi]] = \int f_\Phi(\phi) \int_L^\infty \xi(x;\phi)\, g(x|\,\phi)\, dx d\phi, \tag{22}$$

where (L,∞) is the support of distribution G and

$$\xi(x;\phi) := E[\Pi|\,\Psi = x, \Phi = \phi].$$

The absolute bias induced by the approximation satisfies

$$\left|E[\Pi] - \tilde{E}[\Pi]\right| \leq \int f_\Phi(\phi) \int_L^\infty |\xi(x;\phi)|\,|g(x|\,\phi) - \tilde{g}(x|\,\phi)|\,dxd\phi,$$

where

$$|g(x|\,\phi) - \tilde{g}(x|\,\phi)| \leq \frac{1}{2\pi}\int_{\mathbb{R}} \left|\varphi_G(u|\,\phi) - \varphi_{\tilde{G}}(u|\,\phi)\right| du$$

for integrable characteristic functions φ_G and $\varphi_{\tilde{G}}$; for the latter, refer to Proposition 1.

Next, we aim to derive a more easily accessible and computable upper bound relying on the moment-based metrics of proximity of G and $\tilde{G}$ in Section 3, such as the uniform distance, the entropy bound or the reciprocal polynomial bound, which we generally denote by $\Upsilon(x;\phi)$. We do that in the following theorem and corollary, specifying the necessary conditions for existence of the bound. We then put under scrutiny specific cases.

Theorem 7. *Let*

$$\left|G(x|\,\phi) - \tilde{G}(x|\,\phi)\right| \leq \Upsilon(x;\phi). \tag{23}$$

For $\xi(x;\phi)$ of finite variation for each ϕ, the total variation of ξ is a bounded function of ϕ and

$$\left|E[\Pi] - \tilde{E}[\Pi]\right| \leq \int f_\Phi(\phi) \int_L^\infty \left|\frac{\partial\xi}{\partial x}(x;\phi)\right| \Upsilon(x;\phi)\,dxd\phi \tag{24}$$

is finite.

Proof. From (22),

$$\begin{aligned} E[\Pi|\,\Psi,\Phi] &= \int_L^\infty \xi(x;\phi)\,g(x|\,\phi)\,dx = [\xi(x;\phi)\,G(x|\,\phi)]_L^\infty - \int_L^\infty \frac{\partial\xi}{\partial x}(x;\phi)\,G(x|\,\phi)\,dx \\ &= -\int_L^\infty \frac{\partial\xi}{\partial x}(x;\phi)\,G(x|\,\phi)\,dx \end{aligned}$$

by finite limits of $\xi(x;\phi)\,G(x|\,\Phi)$ as $x \to L$ and $x \to \infty$. Similar result holds for $\tilde{G}$. Then,

$$\left|E[\Pi|\,\Psi = x, \Phi = \phi] - \tilde{E}[\Pi|\,\Psi = x, \Phi = \phi]\right| \leq \int_L^\infty \left|\frac{\partial\xi}{\partial x}(x;\phi)\right| \left|\tilde{G}(x|\,\phi) - G(x|\,\phi)\right| dx$$

from which (24) follows by finite total variation of ξ and by virtue of (23). ■

Corollary 8. *For $\partial\xi(x;\phi)/\partial x > 0$,*

$$\left|E[\Pi] - \tilde{E}[\Pi]\right| \leq \int f_\Phi(\phi) \int_L^\infty \xi(x;\phi) \left|\frac{\partial\Upsilon}{\partial x}(x;\phi)\right| dxd\phi. \tag{25}$$

If, then, $\limsup_{x\to\infty} \xi(x;\phi) < \infty$, the absolute difference $\left|E[\Pi] - \tilde{E}[\Pi]\right|$ is bounded.

Proof. From (24), we get that

$$\int_L^\infty \frac{\partial\xi}{\partial x}(x;\phi)\,\Upsilon(x;\phi)\,dx = [\xi(x;\phi)\,\Upsilon(x;\phi)]_L^\infty - \int_L^\infty \xi(x;\phi)\,\frac{\partial\Upsilon}{\partial x}(x;\phi)\,dx \leq \int_L^\infty \xi(x;\phi)\left|\frac{\partial\Upsilon}{\partial x}(x;\phi)\right| dx,$$

hence (25) follows. ■

For a further study of the above results, we need to make some explicit assumptions. Consider the plain vanilla call option case with

$$\xi(x;\phi) \equiv E\left[(S-e^{\kappa})^{+}\middle|\,\Psi=x,\Phi=\phi\right],$$

where $y^{+}:=\max(y,0)$ and $\kappa\in\mathbb{R}$, in conjunction, for example, with the bound $\Upsilon(x;\phi)\equiv\varpi_m(x)$ in (9). Also, for the majority of the models in Section 4 (see also Table 1),

$$(\ln S|\,\Psi=x,\Phi=\phi)\sim\mathcal{N}(\gamma+\varepsilon\phi+\lambda x,\eta x),$$

where $\gamma,\varepsilon,\lambda,\eta$ are non-random (we may adapt to other models in the paper, if needed). Then,

$$\xi(x;\phi)=e^{m(x;\phi)+\frac{s^2(x)}{2}}N\left(\frac{m(x;\phi)+s^2(x)-\kappa}{s(x)}\right)-e^{\kappa}N\left(\frac{m(x;\phi)-\kappa}{s(x)}\right),$$

where $N(\cdot)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\cdot}e^{-y^2/2}dy$ and

$$m(x;\phi)=\gamma+\varepsilon\phi+\lambda x,\quad s^2(x)=\eta x\quad(\eta>0).$$

The bound (25) holds, depending on the interplay of λ and η. If $\lambda+\eta/2=0$, then $\xi(x;\phi)$ converges to a finite constant, however $\varpi_m'(x)$ converges to 0 at polynomial rate. If $\lambda+\eta/2<0$, then $\lambda<0$ so $\xi(x;\phi)\sim e^{\left(\lambda+\frac{\eta}{2}\right)x}$ as $x\to\infty$, i.e., ξ decreases exponentially fast in the limit; this is not affected by the convergence rate of $\varpi_m'(x)$. If, contrary to these cases, $\xi(x;\phi)\to\infty$, then we can try to truncate $\xi^{c}(x;\phi)=E\left[\min\left\{(S-e^{\kappa})^{+},c\right\}\middle|\,\Psi=x,\Phi=\phi\right]$ using some $c<\infty$, so that $\lim_{x\to\infty}\xi^{c}(x;\phi)=c$ (instead, a put-type $\xi(x;\phi)\equiv E\left[(e^{\kappa}-S)^{+}\middle|\,\Psi=x,\Phi=\phi\right]$ is automatically bounded). If $\xi(x;\phi)\to\infty$, capping can still be avoided depending on the limiting behaviour of $G(x|\,\phi)$ and $\tilde{G}(x|\,\phi)$.

Theorem 9. *If $\xi(x;\phi)\to\infty$ but not as fast as $G(x|\,\phi),\tilde{G}(x|\,\phi)\to 1$, then $\left|E[\Pi]-\tilde{E}[\Pi]\right|$ is bounded.*

Proof. This follows from

$$E[\Pi|\,\Psi,\Phi]=\int_{L}^{\infty}\xi(x;\phi)\,g(x|\,\phi)\,dx=-\left[\xi(x;\phi)\left(1-G(x|\,\phi)\right)\right]_{L}^{\infty}+\int_{L}^{\infty}\frac{\partial\xi}{\partial x}(x;\phi)\left(1-G(x|\,\phi)\right)dx$$

if $\xi(x;\phi)(1-G(x|\,\phi))$ has limit 0 as $x\to\infty$; similarly for $\tilde{E}[\Pi|\,\Psi,\Phi]$ with respect to $\tilde{G}$. ■

With regard to Theorem 9, we have from (1) when $c_3=0$ that

$$\ln\tilde{g}(x|\,\phi)=-\frac{x}{c_2}+\left(\frac{c_1}{c_2^2}-1\right)\ln(c_1+c_2x)+\mathcal{C}.$$

The support of this density can be (L,∞) or $(-\infty,L)$, where $L=-c_1/c_2$, depending on $c_2>0$ or $c_2<0$. Accordingly, we have that

$$\tilde{g}(x|\,\phi)\propto(\pm x\mp L)^{c_1/c_2^2-1}\exp\left(-\frac{x}{c_2}\right).$$

Hence, Theorem 9 will hold when $1/c_2>\lambda+\eta/2$. Notice that we require also $c_1/c_2^2-1>-1$, i.e., $c_1>0$ for the integral to converge as $x\to L$; this has to be the case if the support is to

contain 0. Also, this case corresponds to a gamma distribution with shape parameter c_1/c_2^2 and scale parameter $\pm 1/c_2 > 0$.

In addition to the plain vanilla call option with payoff

$$\Pi(T) = (S(T) - K)^+$$

at some maturity time $T > 0$ for fixed strike price K, other functions we consider in this paper include

$$\Pi(T) = (S(T) - K)^+ \mathbf{1}_{\left\{\max_{0\le t\le T} S(t)\le U\right\}} \tag{26}$$

corresponding to the payoff of a type of barrier option (up-and-out call) with $\mathbf{1}_{\{\cdot\}}$ denoting the indicator of the event $\{\cdot\}$ and U the fixed barrier level;

$$\Pi(T) = \left(K - \min_{0\le t\le T} S(t)\right)^+ \tag{27}$$

corresponding to a lookback put; and

$$\Pi(T) = \left(\bar{S}(T) - K\right)^+ \tag{28}$$

equalling the payoff of an Asian call option with $\bar{S}$ the arithmetic average of the asset price recordings up to time T. For some of these cases and some models, explicit expressions for ξ exist and we present them in Section EC.7 of the e-companion.

7. Numerical study

The following section is dedicated to a numerical analysis of the proposed simulation strategies in the Heston (one- and two-factor), SABR, OU-SV, 3/2 and 4/2, Bates, DPS, NIGCIR (NIG with integrated CIR time-change) and SECIRJD (self-exciting point process with CIR intensity) models. We use parameter sets from the relevant literature that are practically relevant and representative of different markets and market conditions and are not benign (including high correlations, high volatility of variance, and long maturities). All parameter values and their sources are reported in Table 2. All numerical experiments are run in Matlab R2019b in Microsoft Windows 10 on a machine with an Intel(R) Core(TM) i7-9750HQ CPU @2.60GHz and 16 GB of RAM. A code is made available from https://openaccess.city.ac.uk/id/eprint/26427/.

[Insert Table 2]

7.1. Analysis of error and computing time

Before moving to the actual application, we study the core of our method, that is, the ability of Algorithm 1 to compute fast and accurately the moments of the conditional time integral as well as the quality of the subsequent Pearson distribution fit. In the interest of space, we present results relating only to the Heston model (parameter sets H1 and H5, being common in the literature), which is the most computationally complicated with associated Laplace transform (EC.4) that involves Bessel functions whose calculation generally slows down the simulation. We break down the procedure in Algorithm 2 for simulating over the one-period $[0, T]$ into the following four pieces:

Step 1 Generate a sample of $\left(\sigma^2(T)\middle|\,\sigma^2(0)\right)$

Step 2 Compute the first four integer moments of $\left(\int_0^T \sigma^2(s)ds \middle| \sigma^2(T)\right)$ using Algorithm 1

Step 3 Generate a sample of $\left(\int_0^T \sigma^2(s)ds \middle| \sigma^2(T)\right)$ from the Pearson curve fitted by moments

Step 4 Generate a sample of $\left(S(T) \middle| \int_0^T \sigma^2(s)ds, \sigma^2(T)\right)$.

We have trialed the procedure described above several times; the mean execution times corresponding to each step are reported in Section EC.6 of the e-companion. Increasing the number of simulations results in linear increases of the computing time of each step, but with constant percentages of each with respect to the total time. Steps 1, 3 and 4 account for 4%, 11% and 8% of the total time and can be easily implemented using standard routines in numerical computing environments such as Matlab (`ncx2rnd` for step 1; `pearsrnd` for step 3; `randn` for step 4). The moments' evaluation, that is, step 2, is the dominant element occupying 77% of the total execution time. In addition to this, we compare evaluation done using Algorithm 1 and via analytical moments from equation (17). Results reported in Section EC.6 of the e-companion for the first four integer moments indicate a considerable speed-up using Algorithm 1, which in the case of the Heston model translates to 28 Bessel function evaluations versus 307 via equation (17).

Next, we study the potential sources of error from our moment-based approximation. First, we compare moments evaluated using Algorithm 1 and analytically from equation (17) for three different random realizations of the terminal variance corresponding to the 25, 50 and 75th percentiles of that distribution, for convenience. (Similar performances were observed for different terminal variances and parameter sets.) The absolute discrepancies are extremely small, as expected based on the discussion in Section 5.2 (see also Section EC.6 of the e-companion).

Second, we investigate the ability of the Pearson fit to represent the true distribution of the conditional integrated variance. To this end, we perform two tests. In the first one, we aim to assess the closeness of the Pearson moments from the true ones, i.e., having matched the first four integer moments, how far are the higher-order moments from the true ones? Adhering to the same 25, 50 and 75th percentiles of the terminal variance distribution as before for convenience, we present on a log-scale in Figure 1 the first eight integer moments. Indeed, the moments are very close and, for example, for parameters H1 and conditional on the 50th percentile of the terminal variance, we experience absolute differences of the fifth to eighth moments of 0.0003, 0.0014, 0.0034 and 0.0064. Other parameter sets yield similar results insinuating that we are implicitly able to almost fit more than four moments, hence corroborating the high accuracy of our method.

[Insert Figure 1]

Having replaced the original steps 2 and 3 of Broadie and Kaya (2006) by our approach, it also comes natural to evaluate the potential impact on accuracy. For this, we compute the true and fitted Pearson cumulative distribution functions, G and $\tilde{G}$, of $\left(\int_0^T \sigma^2(s)ds \middle| \sigma^2(T)\right)$ for different terminal variances and present both in Figure 2 for terminal variance corresponding to the 50th percentile only in the interest of space. The two plots are practically almost identical. Putting under the microscope the left side of the distribution (see bottom plots in Figure 2) reveals maximum absolute differences between the two of 0.001 and 0.007 for parameters H1 and H5 respectively, whereas for larger values of the integrated variance we observe even smaller differences. These also fall within the theoretical upper bound (7)–(8) computed values. Finally, an implementation of a two-sample Kolmogorov–Smirnov test leads to non-rejection of the null hypothesis that the

samples are drawn from the same distribution with very significantly high p-values for the different parameter sets (smallest p-value being 0.6 for H1).

[Insert Figure 2]

In Figure 3, we show the outcome from the implementation of the entropy approach in Section 3.3. More specifically, we compare the entropy-maximizing distribution based on m moments and our (formally four-moment) fitted Pearson distribution against the true distribution. We report the evaluated bound (12) to the absolute differences between each approximating distribution $H \in \left\{\hat{G}, \tilde{G}\right\}$ and the true distribution G (top plots) as well as the individual entropies (bottom plots). It is obvious that the distances are small and are reducing with increasing m. The fitted Pearson is very close in terms of entropy to the true distribution, while the entropy-maximizing distribution converges to the true distribution with increasing m, remaining, though, even with twelve moments matched, behind the Pearson fit.

[Insert Figure 3]

We conclude this part with a short note of some additional attempts, thanks to a suggestion of one of our reviewers, of fitting of Pearson curves based, instead, on exponential, negative (see Cressie *et al.*, 1981) or even fractional (see Cressie and Borkent, 1986) moments, i.e., $E(e^{-m\alpha\Psi(0,T)}|\Phi(0,T))$ for $m = 1, \dots 4$ and given α, $E(\Psi(0,T)^{-m}|\Phi(0,T))$, $E(\Psi(0,T)^{\gamma}|\Phi(0,T))$ for $\gamma \in \mathbb{Q}$, consistently with the notation in Table 1. More specifically, the computation of the latter proved particularly slow; the application of negative moments produced some accurate simulation estimates but was not sufficiently fast. On the contrary, the use of exponential moments did speed up the computations. However, the resulting error bounds were found to be substantially wider, hence we did not consider this any further. More detailed results can be provided upon request.

7.2. Path-independent derivatives

In this section, we aim to assess the accuracy and speed of our proposed Algorithm 2 in the context of pricing path-independent options, in particular, European plain vanilla options, by comparing with various methods starring in the literature. We study simulation schemes such as those of Broadie and Kaya (2006), Glasserman and Kim (2011), Giles (2008), Giles and Szpruch (2014), Cai *et al.* (2017) and Li and Wu (2019), but also the method of Fang and Oosterlee (2008) as a very accurate proxy for the true option prices. We consider the one-factor and two-factor Heston, SABR, OU-SV, 3/2 and 4/2 models as well as models with jumps including Bates, DPS, NIGCIR and SECIRJD. We compute root mean square errors, RMSE $= \sqrt{\text{bias}^2 + \text{standard error}^2}$, where the bias, as defined in Section 6, is given by the difference between the simulated and true option price. The bias is bounded and can be computed based on Theorem 7. Alternatively, this can be estimated accurately using the hybrid analytical-Monte Carlo valuation described in Sections EC.7.1.1–EC.7.1.2 of the e-companion for variance reduction and $\mathcal{M} = 10^8$ simulation trials (true prices in the SABR model can be obtained as in Cai *et al.*, 2017). We report on the top panel of Table 3 the estimated biases (and standard errors) and true prices for at-the-money plain vanilla options in the different model settings. Obviously these are consistently very small across all the different parameter sets in Table 2. (More, similar results for in-the-money and out-of-the-money options can be made available upon request.)

[Insert Table 3]

We compute RMSEs for varying number of simulations $\mathcal{M}$ for the different models and build the speed-accuracy portrayals of the most competing methods. Numerical results are displayed in Table 4, whereas Figures 4–6 present a few relevant illustrations on a log-log scale. As $\mathcal{M}$ increases, the standard error reduces and the RMSE depends eventually on any residual error. Nevertheless, in the case of our Algorithm 2 this is usually of the order 10^{-4}, for example, for Heston, double Heston, 3/2, Bates, DPS, NIGCIR and SECIRJD, or even 10^{-6} for the SABR model (see Table 3) and, thus, from Figures 4 and 6 we can see how close the RMSEs are to those from methods with optimal convergence like Broadie and Kaya (2006), Giles (2008), Giles and Szpruch (2014), Cai *et al.* (2017) and Li and Wu (2019). Our scheme exhibits the same convergence rate but more importantly reduces the computing time, hence the parallelly shifted plots to the left.

In particular, from Table 4, the computing times compared to Broadie and Kaya (2006), Li and Wu (2019) and Cai *et al.* (2017) decrease, respectively, by an approx. factor of 800, 190 and 8. In the case of the SABR and OU-SV models, the Laplace transforms (EC.5) and (EC.6) do not involve special functions like Heston's (EC.4), therefore the resulting speed-up gain in the Li and Wu (2019) and Cai *et al.* (2017) approaches is magnified in our method which gathers momentum becoming impressively faster when computing the required moments using Algorithm 1. Similarly in relation to the expansion approach of Glasserman and Kim (2011) involving infinite summations requiring truncation (here we have used 10 terms), we achieve power saving by a factor of 3–4. In the case of models with jumps, Bates and NIGCIR exhibit similar computing times to Heston. The simulation of DPS and SECIRJD is slower as the conditional time integral must be simulated more times according to the (random) number of jumps (which, for the particular parameter values, are expected to be more for the latter).

[Insert Table 4]
[Insert Figures 4, 5, 6]

7.3. Path-dependent derivatives

In what follows, we study the efficiency of Algorithm 2 in generating sample paths by turning our attention to the evaluation of path-dependent contracts. Although our method is applicable to general payoff structures and models, we focus here on barrier (up-and-out) and arithmetic Asian call options with payoffs (26) and (28) and monthly monitoring in the Heston model for which we have easily accessible true prices. In particular, these are given for Asians by the accurate lower bound of Fusai and Kyriakou (2016) used in a control variate Monte Carlo setting to achieve extra accuracy; for barriers, we use the Fourier-based technique of Fang and Oosterlee (2011) and barrier level $U = 120$. Accessing these numbers smooths the way for the estimation of biases which we report on the remaining panels of Table 3 for $\mathcal{M} = 10^8$ simulations. Similarly to the case of path-independent options the biases are still very small, which is not surprising as the size of the time interval does not affect the error.

Furthermore, we compare our simulation results based on Algorithm 2 with the time-discretization scheme of Andersen (2008). For a fair comparison, we have tried different number of time steps and chosen the minimum of 250 time steps per year to have an as similar as possible RMSE (for $\mathcal{M} = 10^4$) based on the pre-computed biases. Results are exhibited in Table 5 and Figure 7. We find that Andersen's scheme is faster for a small number of replications. However, as it is obvious from the graphs in Figure 7, for increasing number of simulations, the (large) bias becomes

dominant and the RMSE decay is severely slowed down. This contrasts our method which results in moderate to high levels of precision as $\mathcal{M}$ increases. This is welcoming news as our method constitutes a valid new methodology for applications where the entire sample paths are needed.

[Insert Table 5]
[Insert Figure 7]

We devote the final part of this section to the SABR model given its rising popularity in the literature. First, we apply Algorithm 2 to pricing barrier options with payoff (26) and compare with the mSABR method of Leitao *et al.* (2017), which, even if non-exact, represents a satisfactory benchmark, and the low-bias simulation scheme of Chen *et al.* (2012) using the formers' choice of parameters. Based on the reports in Table 6, our Algorithm 2 and the mSABR method agree at 2–3 decimal places, whereas the discrepancies with Chen *et al.* (2012) are higher. This can be attributed to the larger bias induced by the small-disturbance expansion approximation of moments of the integrated variance and the potentially restrictive lognormal fit they employ. The observed increases in computing time are due to the varying maturity times (see Table 2) subject to quarterly monitoring per annum.

[Insert Table 6]

In what follows, we consider further applications entailing accurate simulation solutions. This is especially important in super-linear growth settings where it is well known that, for standard SDEs, the explicit Euler scheme runs into difficulties, performs poorly or fails to converge (see Hutzenthaler *et al.*, 2010).

8. Linear SDEs: multiplicative noise

The constant-coefficients inhomogeneous model with general form

$$dX(t) = (aX(t) + c)\,dt + (bX(t) + q)\,dW(t) \tag{29}$$

belongs to this class of SDEs. From Kloeden and Platen (1992), (29) has solution

$$X(t) = Y(t; a, b, 1)\left(X(0) + (c - bq)\int_0^t Y(s; a, b, -1)ds + q\int_0^t Y(s; a, b, -1)dW(s)\right),$$

where

$$Y(t; a, b, \gamma) := \exp\left(\gamma\left(a - \frac{1}{2}b^2\right)t + \gamma bW(t)\right). \tag{30}$$

The above solution can be extended to the more general case of variable coefficients.

A well-known special case of (29) in finance is the Brennan and Schwartz (1980) process

$$dX(t) = \kappa\left(\theta - X(t)\right)dt + \sigma X(t)dW(t) \tag{31}$$

with $\kappa := -a > 0$, $\theta := -c/a \in \mathbb{R}$, $\sigma := b > 0$ and $q := 0$, in consistency with a more standard parameterization of the model (e.g., see Cai *et al.*, 2014b). For instance, (31) has been used to model interest rate uncertainty, but has also appeared in the energy marketplace with the spot price reverting towards an equilibrium price level (which might be stochastic, that is, the so-called

Pilipovic model). In Li *et al.* (2018), (31) is referred to as a GARCH linear SDE and is used to model stochastic default intensity, with additional references given therein for uses in modelling the volatility or variance of asset returns. The model (31) can be simulated using our Algorithm 3 for $n = 0$, $a = -\kappa$, $b = \sigma$ and $c = \kappa\theta$; for more details, refer to the next section.

9. Reducible SDEs

9.1. Stochastic generalized logistic (Richards) growth model

This is an autonomous, nonlinear reducible SDE model with polynomial drift of degree n of the general form

$$dX(t) = (cX(t)^n + aX(t))\,dt + bX(t)dW(t). \tag{32}$$

The substitution x^{1-n} reduces (32) to a linear SDE with multiplicative noise, from which its explicit solution is

$$X(t) = Y(t;a,b,1)\left(X(0)^{1-n} + c\,(1-n)\int_0^t Y(s;a,b,n-1)ds\right)^{\frac{1}{1-n}}, \tag{33}$$

where Y is given by (30).

[Insert Algorithm 3]

We summarize the simulation methodology for the model (33) in Algorithm 3. If we are interested in simulating X at a terminal time $T > 0$ only, we use $m = 1$. Simulation steps 3 and 6 of Algorithm 3 are trivial; for step 5, the conditional distribution of $\int_0^t Y(s;a,b,\gamma)ds$ plays a crucial role in the implementation of the methodology developed in this paper. To this end, let $W^{(\mu)}(t) = \mu t + W(t)$ be a Brownian motion with constant drift $\mu \in \mathbb{R}$. We also recall from Matsumoto and Yor (2005) the additive functional $A^{(\mu)}(t) = \int_0^t \exp\left(2W^{(\mu)}(s)\right) ds$. Then, they show (see also Cai *et al.*, 2017) that

$$E\left[\exp\left(-\frac{u}{A^{(\mu)}(t)}\right)\middle| W^{(\mu)}(t) = w\right] = \exp\left\{-\frac{g(w,u)^2 - w^2}{2t}\right\} \tag{34}$$

for any $t > 0$ and $g(w,u) := \operatorname{arcosh}(ue^{-w} + \cosh w)$.

Proposition 2. *The Laplace transform of* $\left(\int_0^t Y(s;a,b,\gamma)ds\right)^{-1}$ *conditional on* $Y(t;a,b,\gamma)$ *is given by*

$$E\left[\exp\left(-\frac{u}{\int_0^t Y(s;a,b,\gamma)ds}\right)\middle| Y(t;a,b,\gamma)\right] = \exp\left\{-\frac{g\left(\frac{\ln Y(t;a,b,\gamma)}{2}, \frac{\gamma^2 b^2 u}{4}\right)^2 - \left(\frac{\ln Y(t;a,b,\gamma)}{2}\right)^2}{\frac{1}{2}\gamma^2 b^2 t}\right\}. \tag{35}$$

Proof. See e-companion Section EC.1. ■

The range of admissible values $u \geq -\frac{2}{\gamma^2 b^2}\left(\sqrt{Y(t;a,b,\gamma)} - 1\right)^2$ ensures that the moment problem is determinate. Special cases of great practical importance are the stochastic Verhulst and Gordon–Schaefer models for $n = 2$ and the stochastic Ginzburg–Landau equation for $n = 3$ which we explore further in the following sections.

9.2. Stochastic logistic (Verhulst) growth model

The Verhulst (1838) model, also referred to usually as logistic or $\mathcal{S}$-shaped, belongs to the Richards family of sigmoidal growth models. The deterministic model is described by

$$\frac{d\tilde{X}(t)}{dt} = \lambda \left(K - \tilde{X}(t)\right) \tilde{X}(t), \tag{36}$$

where λ is the Malthusian growth coefficient and $K > 0$ a finite supportable carrying capacity in an environment of finite resources. To account for seasonal variations, (36) is augmented by a random component which results in the autonomous stochastic linear-quadratic Verhulst equation

$$d\tilde{X}(t) = \tilde{\lambda} \left(1 - \frac{\tilde{X}(t)}{K}\right) \tilde{X}(t)dt + \sigma \tilde{X}(t)dW(t),$$

where W is a standard Brownian motion, $X(0) \in (0, \infty)$, $\tilde{\lambda} := \lambda K$ and $\lambda, \sigma > 0$ are constant. The extinction boundary 0 is non-attractive if $\tilde{\lambda} > \sigma^2/2$ and is attractive if $\tilde{\lambda} < \sigma^2/2$. Also, there is no explosion as the boundary $+\infty$ is non-attractive. Equivalently, we can write

$$dX(t) = \left(\tilde{\lambda} X(t) - X^2(t)\right) dt + \sigma X(t)dW(t), \tag{37}$$

where $X := \lambda \tilde{X}$. In this model, an approximately exponential growth is succeeded by a slowed down linear growth, as saturation begins, and ceases at maturity. The solution exists and is unique for all $t \geq 0$, and from (32)–(33) for $n = 2$

$$X(t) = Y(t; \tilde{\lambda}, \sigma, 1) \left(X(0)^{-1} + \int_0^t Y(s; \tilde{\lambda}, \sigma, 1)ds\right)^{-1} = \frac{X(0)Y(t; \tilde{\lambda}, \sigma, 1)}{1 + X(0) \int_0^t Y(s; \tilde{\lambda}, \sigma, 1)ds}.$$

This is an environmental stochasticity model meaning that random fluctuations in the environment, such as weather, epidemics, natural disasters, crop failures, can affect the entire population. Originally, the model (36) aimed at portraying the self-limiting growth of a biological population. According to this, the rate of reproduction is proportional to the existing population and the amount of available resources. A multi-dimensional version allows for affection of co-existing species by coupling together separate single species models. The logistic model appears also in branches of medicine such as oncology. In what follows, we focus the spotlight on the logistic model for tumor growth, we parameterize this and use it to exemplify the application of our simulation method.

9.3. Tumor growth: a simulation case study

Cancer cells have an inflated rate of proliferation leading to a rapid tumor growth. A survey of the relevant literature reveals that, generally, small tumors grow exponentially, however the growth slows down when they get larger. This decelerated growth has the consequence that the diameter (if a solid tumor) typically eventually remains constant in time. Therefore, a model that is able to accommodate these features can be a sensible choice (e.g., see Laird, 1964, and later research in the field). Note that, while growth curves are relevant for modelling the untreated tumor, they cannot be applied to the treated tumor as it may decrease and regrow after a treatment such as radiation (e.g., see Demidenko, 2013, Chapter 10); studying post-treatment tumor is beyond the scope of this exercise.

Here, we adopt the stochastic logistic model (37) for the growth of tumors, where we denote by $X(t)$ the tumor volume at time t. For the purposes of the illustration, we calibrate the model to growth data of multicellular tumor spheroids, three-dimensional aggregates of cancer cells, that have been grown under controlled experimental conditions, as described in Chignola *et al.* (1999), that were made available by Demidenko (2013). The nonlinear least-squares parameter estimates of the model are $\hat{\lambda} = 0.00109$, $\hat{K} = 203.10135$ and $\hat{X}(0) = 9.63691$; in addition, the estimated daily volatility is $\hat{\sigma} = 10.6539\%$.

[Insert Figure 8]

Figure 8 exhibits the historical daily tumor growth and the fitted deterministic Verhulst equation (red solid line). In addition, our simulation method serves a useful tool for the analysis of the tumor growth based on the model (37). First, it can be used to generate probabilistic forecasts: the medium (green) line shows the expected value forecasts and the upper and lower (orange dotted) lines are confidence bands such that there is a 90% chance that the tumor volume will not exceed these bands. Second, it allow us to study the state of growth at certain points in time. In particular, we document three critical time points (see vertical lines in Figure 8) corresponding to different phases of the tumor growth *in vivo*, therefore signifying the timing of relevant guideline for treatment: T_1, when tumor slow growth (tumor vasculature) is completed since commencement of cell division; T_2, when aggressive tumor growth occurs and the maximum rate of growth is reached; T_3, when growth is slowed again (due to limited supply of oxygen, nutrients, and space). For the given model, these critical points are given by

$$\hat{T}_{1,3} = \frac{1}{\hat{\lambda}} \ln\left(\left(\pm\sqrt{3} - 2\right)\left(1 - \frac{\hat{K}}{\hat{X}(0)}\right)\right)$$

and

$$\hat{T}_2 = \frac{1}{\hat{\lambda}} \ln\left(\frac{\hat{K}}{\hat{X}(0)} - 1\right)$$

(inflection point). Finally, Figure 9 shows the tumor volume (for the 95th percentile simulated radius) at the three different phases and the corresponding simulated distributions of the tumor radii and volumes as percentages of the estimated maximum limit, $100X(\hat{T}_i)/\hat{K}\%$. At time $\hat{T}_1$, there is a large concentration of probability mass at lower volume levels but there is also a nonzero probability of volume expansion as implied by the left tail. The distribution of the relative volume becomes more symmetric and the kurtosis reduces as we approach the terminal phase at $\hat{T}_3$.

[Insert Figure 9]

9.4. Stochastic Gordon–Schaefer model

The model (32) is popular, for example, in bioeconomics where it can be used to represent the growth of a population living in a randomly varying environment and being harvested, such as a fish population under fishing, or a wildlife population under hunting, or even a tree population in forestry.

Early important contributions on specific stochastic models for fisheries are due to Beddington and May (1977), Braumann (1985) and Hanson and Ryan (1998). The stochastic Gordon–Schaefer

population growth SDE model

$$dX(t) = rX(t)\left(1 - \frac{X(t)}{K}\right)dt - H(t)dt + \sigma X(t)dW(t) \tag{38}$$

includes the term $H(t) := qEX(t)$ which represents the (e.g., fishing) harvesting rate, where $E \geq 0$ is the (fishing) effort (e.g., hours trawled or number of hooks sets per day) and $q > 0$ is the fraction of the population harvested per unit of effort, so that the difference

$$r\left(1 - \frac{X(t)}{K}\right) - qE$$

is the natural growth rate adjusted for mortality due to harvesting. Under a basic profit structure, the profit per unit time is

$$\Pi(t) = pH(t) - cE, \tag{39}$$

where the revenues (first term) depend on the harvesting rate and p is the price per unit sold, and the costs (second term) appreciate subject to c cost per unit effort per unit time. p and c can, respectively, be functions of the harvesting rate (higher rate implying lower selling price) and the effort E (increasing effort leading, for example, to overtime and therefore higher cost). Finally, from (32)–(33)

$$X(t) = Y(t; r - qE, \sigma, 1)\left(X(0)^{-1} + \frac{r}{K}\int_0^t Y(s; r - qE, \sigma, 1)ds\right)^{-1}.$$

[Insert Figure 10]

In Figure 10, we revisit the bioeconomic resource model framework of Hanson and Ryan (1998) for the Pacific halibut. One can determine the optimal sustainable constant effort E^* that maximizes the expected value of the asymptotic profit. By adopting their parameter estimates based on a realistic set of data, we use our method to generate population size (in kilograms) sample paths. The top panel shows the impact of increasing fishing effort starting from $E = 0$ (top-left), which reduces the model (38) to (37), implying a significantly overestimated population size under a misspecified model with harvesting mortality ignored, for example, by 25%, 50% and 75% when $E = 50\%E^*$, E^* and $150\%E^*$ (not explicitly reported in the plot), respectively. In particular, the central and bottom panels focus on the impact of the fishing effort on the profit process (39). Along these panels, the middle plot corresponds to the optimal sustainable effort E^*, whilst as we diverge from it the profit drops. This mirrors the position of the mode of the estimated density of the terminal profit in the bottom panels, from which we observe that as E reduces below or increases above the optimal level the mode shifts to left; the standard deviation, skewness and kurtosis also increase with E.

9.5. Ginzburg–Landau equation

Our last application is devoted to the Ginzburg–Landau equation. In its deterministic version, this was introduced by Ginzburg and Landau (1950) to describe phase transition for superconductivity. Over the years, this model has been used in bistable systems, chemical turbulence, phase transitions in non-equilibrium systems, optics with dissipation, thermodynamics and hydrodynamics, etc. It has also played an important role as a modulation equation and served as a simple model for the transition from regular to turbulent behaviour (see Mielke, 2002).

Because random noise is often unavoidable, taking into consideration stochastic disturbances is needed. A stochastic version of it is provided by Kloeden and Platen (1992) and is given by

$$dX(t) = \left(\left(\alpha + \frac{1}{2}\sigma^2\right) X(t) - \beta X^3(t)\right) dt + \sigma X(t) dW(t), \tag{40}$$

where $X(0) \in (0, \infty)$ and $\alpha \geq 0$ and $\beta, \sigma > 0$ are constant. From (33), its solution is given explicitly by

$$X(t) = \frac{X(0) Y(t; \alpha + \frac{1}{2}\sigma^2, \sigma, 1)}{\sqrt{1 + 2X^2(0)\beta \int_0^t Y(s; \alpha + \frac{1}{2}\sigma^2, \sigma, 2) ds}}.$$

Several variants of (40) with a colored noise or regime switching also exist.

Here, we focus on (40) and run a simulation experiment borrowed from Hutzenthaler *et al.* (2010) in order to demonstrate the efficiency of our proposed scheme. Table 7 shows, for different values of the parameter σ, Monte Carlo estimates of $E\left(X^2(3)\right)$ using our method as well as estimates based on different implementations of the Euler approximation. As σ increases, the bias of the Euler scheme increases and the resulting estimates become perceptibly inaccurate. For very large $\sigma = 6$ or 7, in most of the runs the Euler scheme explodes returning 'NaN' ('not-a-number'), whilst our method remains robust, even under such extremely volatile conditions, and is remarkably faster by a factor of 20.

[Insert Table 7]

10. Conclusion

In this paper, we propose a novel method for the simulation of integrals of general stochastic processes. We focus our analysis on especially hard cases of dependence on the terminal value of the process. We manage to relax the most time-consuming parts of other methods: i) the repeatedly recovered conditional distribution by numerical transform inversion; ii) the large number of time steps intermediating the actual monitoring dates for the purpose of reducing the bias of time-discretization techniques; iii) the number of terms in expansion approaches; iv) the attainment of the required moments by direct differentiation of a Laplace transform.

Our method builds on efficient computation of the moments of the integrated process. We then propose a moment-based random number generator which is accurate, with very small and bounded bias, and fast, improving on the complexity and speed of pre-existing techniques. Bypassing the most prolonged parts of others, we are able to generate entire paths of stochastic processes of interest in reasonable time. We explore the flexibility and robustness of our method in different practical problem structures and show that it can efficiently contribute to the interface of simulation with various areas in financial engineering, medicine, bioeconomics and physics and, thereby, lead to useful implications within the application domain.

References

Abramowitz, M. and Stegun, I. A. (1968) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover Books on Mathematics, 2 edn.

Akhiezer, N. I. (1965) *The Classical Moment Problem and Some Related Questions in Analysis*. University Mathematical Monographs. Oliver & Boyd.

Andersen, L. (2008) Simple and efficient simulation of the Heston stochastic volatility model. *Journal of Computational Finance*, **11**, 1–42.

Ballotta, L. and Kyriakou, I. (2014) Monte Carlo simulation of the CGMY process and option pricing. *Journal of Futures Markets*, **34**, 1095–1121.

Bates, D. S. (1996) Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options. *Review of Financial Studies*, **9**, 69–107.

Beddington, J. R. and May, R. M. (1977) Harvesting natural populations in a randomly fluctuating environment. *Science*, **197**, 463–465.

Braumann, C. A. (1985) Stochastic differential equation models of fisheries in an uncertain world: Extinction probabilities, optimal fishing effort, and parameter estimation. In *Mathematics in Biology and Medicine* (eds. V. Capasso, E. Grosso and S. L. Paveri-Fontana), 201–206. New York: Springer.

Brennan, M. J. and Schwartz, E. S. (1980) Analyzing convertible bonds. *The Journal of Financial and Quantitative Analysis*, **15**, 907–929.

Broadie, M. and Kaya, O. (2006) Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, **54**, 217–231.

Cai, N., Kou, S. G. and Liu, Z. (2014a) A two-sided Laplace inversion algorithm with computable error bounds and its applications in financial engineering. *Advances in Applied Probability*, **46**, 766–789.

Cai, N., Li, C. and Shi, C. (2014b) Closed-form expansions of discretely monitored Asian options in diffusion models. *Mathematics of Operations Research*, **39**, 789–822.

Cai, N., Song, Y. and Chen, N. (2017) Exact simulation of the SABR model. *Operations Research*, **65**, 931–951.

Callegaro, G., Fiorin, L. and Grasselli, M. (2019) Quantization meets Fourier: a new technology for pricing options. *Annals of Operations Research*, **282**, 59–86.

Carr, P., Geman, H., Madan, D. and Yor, M. (2003) Stochastic volatility for Lévy processes. *Mathematical Finance*, **13**, 345–382.

Chen, B., Oosterlee, C. W. and van der Weide, H. (2012) A low-bias simulation scheme for the SABR stochastic volatility model. *International Journal of Theoretical and Applied Finance*, **15**, 125–161.

Chignola, R., Schenetti, A., Chiesa, E., Foroni, R., Sartoris, S., Brendolan, A., Tridente, G., Andrighetto, G. and Liberati, D. (1999) Oscillating growth patterns of multicellular tumour spheroids. *Cell Proliferation*, **32**, 39–48.

Choudhury, G. L. and Lucantoni, D. M. (1996) Numerical computation of the moments of a probability distribution from its transform. *Operations Research*, **44**, 368–381.

Christoffersen, P., Heston, S. and Jacobs, K. (2009) The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science*, **55**, 1914–1932.

Cont, R. and Tankov, P. (2004) *Financial Modelling With Jump Processes*. Financial Mathematics Series. Boca Raton: Chapman & Hall/CRC.

Corsaro, S., Kyriakou, I., Marazzina, D. and Marino, Z. (2019) A general framework for pricing Asian options under stochastic volatility on parallel architectures. *European Journal of Operational Research*, **272**, 1082–1095.

Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985) A theory of the term structure of interest rates. *Econometrica*, **53**, 385–407.

Cressie, N. and Borkent, M. (1986) The moment-generating function has its moments. *Journal of Statistical Planning and Inference*, **13**, 337–344.

Cressie, N., Davis, A. S., Folks, J. L. and Policello, G. E. (1981) The moment-generating function and negative integer moments. *The American Statistician*, **35**, 148–150.

Dassios, A. and Zhao, H. (2017) Efficient simulation of clustering jumps with CIR intensity. *Operations Research*, **65**, 1494–1515.

Demidenko, E. (2013) *Mixed Models: Theory and Applications with R*. Wiley Series in Probability and Statistics. New York: Wiley, 2 edn.

Devroye, L. (1986) *Non-Uniform Random Variate Generation*. New York: Springer.

Du, D. and Luo, D. (2019) The pricing of jump propagation: Evidence from spot and options markets. *Management Science*, **65**, 2360–2387.

Duffie, D., Filipović, D. and Schachermayer, W. (2003) Affine processes and applications in finance. *The Annals of Applied Probability*, **13**, 984–1053.

Duffie, D., Pan, J. and Singleton, K. (2000) Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, **68**, 1343–1376.

Elderton, W. P. and Johnson, N. L. (1969) *Systems of Frequency Curves*. Cambridge: Cambridge University Press.

Fang, F. and Oosterlee, C. W. (2008) A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, **31**, 826–848.

Fang, F. and Oosterlee, C. W. (2011) A Fourier-based valuation method for Bermudan and barrier options under Heston's model. *SIAM Journal on Financial Mathematics*, **2**, 439–463.

Fulop, A. and Li, J. (2019) Bayesian estimation of dynamic asset pricing models with informative observations. *Journal of Econometrics*, **209**, 114–138.

Fusai, G. and Kyriakou, I. (2016) General optimized lower and upper bounds for discrete and continuous arithmetic Asian options. *Mathematics of Operations Research*, **41**, 531–559.

Fusai, G. and Tagliani, A. (2002) An accurate valuation of Asian options using moments. *International Journal of Theoretical and Applied Finance*, **5**, 147–169.

Gauthier, P. and Possamaï, D. (2010) Efficient simulation of the double Heston model. Working Paper.

Giesecke, K., Kakavand, H. and Mousavi, M. (2011) Exact simulation of point processes with stochastic intensities. *Operations*

Research, **59**, 1233–1245.

Giles, M. B. (2008) Multilevel Monte Carlo path simulation. *Operations Research*, **56**, 607–617.

Giles, M. B. and Szpruch, L. (2014) Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation. *Annals of Applied Probability*, **24**, 1585–1620.

Ginzburg, V. L. and Landau, L. D. (1950) On the theory of superconductivity. *Zh. Eksp. Teor. Fiz.*, **20**, 1064–1082.

Glasserman, P. (2004) *Monte Carlo Methods in Financial Engineering*. Stochastic Modelling and Applied Probability. New York: Springer.

Glasserman, P. and Kim, K. K. (2011) Gamma expansion of the Heston stochastic volatility model. *Finance and Stochastics*, **15**, 267–296.

Gnoatto, A., Grasselli, M. and Platen, E. (2016) A penny saved is a penny earned: Less expensive zero coupon bonds. Working Paper.

Grasselli, M. (2017) The 4/2 stochastic volatility model: A unified approach for the Heston and the 3/2 model. *Mathematical Finance*, **27**, 1013–1034.

Hagan, P. S., Kumar, D., Lesniewski, A. S. and Woodward, D. E. (2002) Managing smile risk. *Wilmott Magazine*, 84–108.

Hanson, F. B. and Ryan, D. (1998) Optimal harvesting with both population and price dynamics. *Mathematical Biosciences*, **148**, 129–146.

Heston, S. (1997) A simple new formula for options with stochastic volatility. Tech. rep., Washington University of St. Louis.

Heston, S. L. (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, **6**, 327–343.

Hill, I. D., Hill, R. and Holder, R. L. (1976) Algorithm as 99: Fitting Johnson curves by moments. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **25**, 180–189.

Hörmann, W., Leydold, J. and Derflinger, G. (2004) *Automatic Nonuniform Random Variate Generation*. Statistics and Computing. New York: Springer.

Hutzenthaler, M., Jentzen, A. and Kloeden, P. E. (2010) Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proceedings of the Royal Society A*, **467**, 1563–1576.

Islah, O. (2009) Solving SABR in exact form and unifying it with LIBOR market model. Working Paper.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994) *Continuous Univariate Distributions*. New York: Wiley, 2 edn.

Kang, C., Kang, W. and Lee, J. M. (2017) Exact simulation of the Wishart multidimensional stochastic volatility model. *Operations Research*, **65**, 1190–1206.

Kapur, J. N. and Kesavan, H. K. (1992) Entropy optimization principles and their applications. In *Entropy and Energy Dissipation in Water Resources* (eds. V. P. Singh and M. Fiorentino), 3–20. Dordrecht: Springer Netherlands.

Karatzas, I. and Shreve, S. E. (1991) *Brownian Motion and Stochastic Calculus*. New York: Springer, 2 edn.

Khamis, S. H. (1954) On the reduced moment problem. *The Annals of Mathematical Statistics*, **25**, 113–122.

Klebanov, L. B. and Mkrtchyan, S. T. (1986) Estimation of the closeness of distributions in terms of identical moments. *Journal of Soviet Mathematics*, **32**, 54–60.

Kloeden, P. E. and Platen, E. (1992) *Numerical Solution of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. New York: Springer.

Kolmogorov, A. N. (1933) Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, **4**, 83–91.

Kullback, S. (1967) A lower bound for discrimination information in terms of variation. *IEEE Transactions on Information Theory*, **13**, 126–127.

Laird, A. K. (1964) Dynamics of tumor growth. *British Journal of Cancer*, **13**, 490–502.

Leitao, Á., Grzelak, L. A. and Oosterlee, C. W. (2017) On an efficient multiple time step Monte Carlo simulation of the SABR model. *Quantitative Finance*, **17**, 1549–1565.

Lévy, P. (1925) *Calcul des Probabilités*. Paris: Gauthier-Villars.

Li, C. and Wu, L. (2019) Exact simulation of the Ornstein–Uhlenbeck driven stochastic volatility model. *European Journal of Operational Research*, **275**, 768–779.

Li, M., Mercurio, F. and Resnick, S. (2018) The Garch linear SDE: explicit formulas and the pricing of a quanto CDS. Available online at https://www.risk.net/media/download/1001271/download.

Lin, W., Li, S., Luo, A. and Chern, S. (2017) Consistent pricing of VIX and equity derivatives with the 4/2 stochastic volatility plus jumps model. *Journal of Mathematical Analysis and Applications*, **447**, 778–797.

Lindsay, B. G. and Basak, P. (2000) Moments determine the tail of a distribution (but not much else). *The American Statistician*, **54**, 248–251.

Lindsay, B. G., Pilla, R. S. and Basak, P. (2000) Moment-based approximations of distributions using mixtures: Theory and applications. *Annals of the Institute of Statistical Mathematics*, **52**, 215–230.

Linnik, J. V. and Ostrovskiĭ, I. V. (1977) *Decomposition of Random Variables and Vectors*, vol. 48 of *Translations of Mathematical Monographs*. Providence, R. I.: American Mathematical Society.

Matsumoto, H. and Yor, M. (2005) Exponential functionals of Brownian motion, I: Probability laws at fixed time. *Correspondance mathématique et physique*, **2**, 312–347.

Mielke, A. (2002) The Ginzburg-Landau equation in its role as a modulation equation. In *Handbook of Dynamical Systems* (ed. B. Fiedler), vol. 2 of *Handbook of Dynamical Systems*, 759–834. Elsevier Science.

Nadarajah, S. (2008) A review of results on sums of random variables. *Acta Applicandae Mathematicae*, **103**, 131–140.

Obukhov, A. M. (1959) Description of turbulence in terms of Lagrangian variables. vol. 6 of *Advances in Geophysics*, 113–116. Elsevier.

Platen, E. (1997) A non-linear stochastic volatility model. *Financial Mathematics Research Report No. FMRR 005-97, Center for Financial Mathematics, Australian National University, Canberra*.

Rheinländer, T. and Schmutz, M. (2013) Self-dual continuous processes. *Stochastic Processes and their Applications*, **123**, 1765–1779.

Rose, C. and Smith, M. D. (2002) *Mathematical Statistics with Mathematica*. New York: Springer.

Schöbel, R. and Zhu, J. (1999) Stochastic volatility with an Ornstein–Uhlenbeck process: An extension. *European Finance Review*, **3**, 23–46.

Scott, L. (1987) Option pricing when the variance changes randomly: Theory, estimation, and an application. *Journal of Financial Quantitative Analysis*, **22**, 419–438.

Shen, C. Y. (2020) Logistic growth modelling of COVID-19 proliferation in China and its international implications. *International Journal of Infectious Diseases*, **96**, 582–589.

Simonato, J. G. (2011) The performance of Johnson distributions for value at risk and expected shortfall computation. *Journal of Derivatives*, **19**, 7–24.

Stein, E. and Stein, J. (1991) Stock price distributions with stochastic volatility: An analytic approach. *Review of Financial Studies*, **4**, 727–752.

Stoyanov, J. M. (2013) *Counterexamples in Probability*. New York: Dover Publications, 3 edn.

Verhulst, P. F. (1838) Notice sur la loi que la population poursuit dans son accroissement. *Correspondance mathématique et physique*, **10**, 113–121.

Wachter, J. A. (2013) Can time-varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance*, **68**, 987–1035.

Wu, K., Darcet, D., Wang, Q. and Sornette, D. (2020) Generalized logistic growth modeling of the COVID-19 outbreak: comparing the dynamics in the 29 provinces in China and in the rest of the world. *Nonlinear Dynamics*, **101**, 1561–1581.

Zhang, S. M. and Feng, Y. (2019) American option pricing under the double Heston model based on asymptotic expansion. *Quantitative Finance*, **19**, 211–226.

Zolotarev, V. M. (1970) Several new probabilistic inequalities connected with the Lévy metric. *Dokl. Akad. Nauk SSSR*, **190**, 1019–1021.

Zolotarev, V. M. (1983) Probability metrics. *Theory of Probability & Its Applications*, **28**, 278–302.

Zolotarev, V. M. and Senatov, V. V. (1976) Two-sided estimates of Levy's metric. *Theory of Probability & Its Applications*, **20**, 234–245.

Algorithm 1 Numerical inversion of adaptively modified moment generating function

Input: m, γ, $\mathcal{L}(\cdot)$
Output: $\{\mu_n\}_{n=1}^{m}$

Set $l = \alpha_1 = 1$ and compute μ_1 from (18)
Compute $\alpha_2 = 1/\mu_1$ and μ_2
Set $l = 1 \vee 2$ and $\alpha_1 = \alpha_2 = 2\mu_1/\mu_2$ and compute new values for μ_1 and μ_2 (from 18)
Set $n = 3$
while $n \leq m$ **do**
    $l = 1 \vee 2$, compute $\alpha_n = (n-1)\mu_{n-2}/\mu_{n-1}$
    Compute μ_n from (18)
    $n = n + 1$
end

Algorithm 2 Moment-matched conditional sampling scheme: asset price process under stochastic volatility

Input: Model parameters, terminal time T, number of monitoring dates N
Output: Asset price path $\{S(t)\}$ for $t = \{0, \Delta, 2\Delta, 3\Delta, ..., T\}$

Set $\Delta = \frac{T}{N}$
for $t = 0 : \Delta : T - \Delta$ **do**
    Given $\sigma(t)$, generate $\Phi(t, t+\Delta)$
    Compute the moments of $\Psi(t, t+\Delta)$ conditional on $\sigma(t)$ and $\Phi(t, t+\Delta)$ using Algorithm 1
    Sample from the conditional $\Psi(t, t+\Delta)$ from (16) given the moments
    Sample $S(t+\Delta)$ given $S(t), \sigma(t), \Phi(t, t+\Delta)$ and $\Psi(t, t+\Delta)$
end for
return $\{S(t)\}$ for $t = \{0, \Delta, 2\Delta, 3\Delta, ..., T\}$

Algorithm 3 Moment-matched conditional sampling scheme: linear & reducible SDEs

Input: Model parameters, terminal time T, number of monitoring dates N
Output: Sample path $\{X(t)\}$ for $t = \{0, \Delta, 2\Delta, 3\Delta, ..., T\}$

Set $\Delta = \frac{T}{N}$
for $t = 0 : \Delta : T - \Delta$ **do**
    Given $Y(t; a, b, 1)$, simulate $Y(t+\Delta; a, b, 1)$
    Compute the moments of $\int_t^{t+\Delta} Y(s; a, b, n-1)ds$ conditional on $Y(t+\Delta; a, b, 1)$ using Algorithm 1
    Simulate the conditional $\int_t^{t+\Delta} Y(s; a, b, n-1)ds$ from (16) given the moments
    Simulate $X(t+\Delta)$ from (33) given $X(t)$
end for
return $\{X(t)\}$ for $t = \{0, \Delta, 2\Delta, 3\Delta, ..., T\}$

Table 1: Key quantities in simulation of different models

Model	$\Phi(u,t)$	$\Psi(u,t)$
Heston	$\sigma^2(t)$	$\int_u^t \sigma^2(s)ds$
SABR	$\sigma^2(t)$	$\frac{1}{\int_u^t \sigma^2(s)ds}$
OU-SV	$\left(\sigma(t), \int_u^t \sigma(s)ds\right)$	$\int_u^t \sigma^2(s)ds$
Double Heston	$(\sigma_1^2(t), \sigma_2^2(t))$	$\left(\int_u^t \sigma_1^2(s)ds, \int_u^t \sigma_2^2(s)ds\right)$
3/2	$\sigma^2(t)$	$\int_u^t \frac{ds}{\sigma^2(s)}$
4/2	$\sigma^2(t)$	$\left(\int_u^t \sigma^2(s)ds, \int_u^t \frac{ds}{\sigma^2(s)}\right)$

Table 2: Model parameter sets

Heston model												
	$S(0)$	k	θ	v	$\sigma^2(0)$	ρ	r	T				
H1	100	6.21	0.019	0.61	0.010201	-0.7	3.19%	1				
H2	100	2	0.09	1	0.09	-0.3	5%	5				
H3	100	0.5	0.04	1	0.04	-0.9	3%	1				
H4	100	0.3	0.04	0.9	0.04	-0.5	3%	1				
H5	100	1	0.09	1	0.09	-0.3	3%	1				
H6	100	6.2	0.02	0.6	0.02	-0.7	3%	1				
SABR model												
	$S(0)$	β	v	$\sigma(0)$	ρ	r	T					
SABR1	0.05	0.3	0.6	0.4	0	0	1					
SABR2	0.05	0.3	0.6	0.4	0	0	3					
SABR3	0.05	0.3	0.6	0.4	0	0	5					
SABR4	0.5	0.5	0.4	0.5	0	0	4					
SABR5	0.04	1	0.3	0.2	-0.5	0	5					
SABR6	1	0.6	0.3	0.25	-0.5	0	20					
OU-SV model												
	$S(0)$	k	θ	v	$\sigma^2(0)$	ρ	r	T				
OU-SV1	100	4	0.02	0.1	0.04	-0.7	9.53%	1				
OU-SV2	100	4	0.02	0.1	0.04	-0.7	9.53%	3				
OU-SV3	100	4	0.02	0.1	0.04	-0.7	9.53%	5				
Double Heston model												
	$S(0)$	$k_{1,2}$	$\theta_{1,2}$	$v_{1,2}$	$\sigma^2_{1,2}(0)$	$\rho_{1,2}$	r	T				
DH1	100	0.9 1.2	0.1 0.15	0.36 0.2	0.36 0.2	-0.5 -0.5	3.00%	1				
DH2	100	1.0738 0.0326	0.1026 0.7078	0.826 1.5355	0.0028 0.0059	-0.2819 -0.687	3.00%	1				
3/2 model												
	$S(0)$	k	θ	v	$\sigma^2(0)$	ρ	β	r	T			
3/2-1	100	1.8	0.04	0.2	0.04	-0.7	0.025	2%	1			
3/2-2	100	1.1705	0.6853	0.398	0.8992	-0.8637	0.0192	2%	1			
4/2 model												
	$S(0)$	k	θ	v	$\sigma^2(0)$	ρ	β	α	r	T		
4/2-1	100	1.8	0.04	0.2	0.04	-0.7	0.025	0.3	2%	1		
4/2-2	100	1.1705	0.6853	0.398	0.8992	-0.8637	0.0192	0.0218	2%	1		
Bates model												
	$S(0)$	k	θ	v	$\sigma^2(0)$	ρ	λ	$\bar{\mu}$	σ_s	r	T	
B1	100	3.99	0.014	0.2700	0.008836	-0.79	0.11	-0.12	0.15	3.19%	5	
DPS model												
	$S(0)$	k	θ	v	$\sigma^2(0)$	ρ	λ	$\bar{\mu}$	σ_s	μ_v	ρ_J	r	T
DPS1	100	3.46	0.008	0.1400	0.007569	-0.82	0.47	-0.1	0.0001	0.05	-0.38	3.19%	1
NIGCIR model													
	$S(0)$	k	θ	v	$\sigma^2(0)$	$\underline{\theta}$	$\underline{k}$	$\underline{\sigma}$	r	T			
NIGCIR1	100	3.99	0.014	0.27	0.008836	-11.00604	0.00294	0.84059	4%	1			
SECIRJD model													
	$S(0)$	$\lambda(0)$	k_λ	θ_λ	σ_λ	β	σ^2	μ_J	σ_J	r	T		
SECIRJD1	100	2.935	2.266	2.935	0.585	1.782	0.041	-0.018	0.002	3.19%	1		

Notes. Parameter sets H1–2, B1, DPS1 and H3–6 are from Broadie and Kaya (2006) and Glasserman and Kim (2011), DH1–2 from Gauthier and Possamaï (2010) and Zhang and Feng (2019), SABR1–3 and SABR4–6 from Cai *et al.* (2017) and Leitao *et al.* (2017), OU-SV1–3 from Li and Wu (2019), 3/2-1 and 4/2-1 from Grasselli (2017) and Callegaro *et al.* (2019), 3/2-2 and 4/2-2 from Gnoatto *et al.* (2016), NIGCIR1 from Corsaro *et al.* (2019, Table 2) and SECIRJD1 from Fulop and Li (2019, Table 8).

Table 3: True option prices and biases (with standard errors) of Monte Carlo price estimates for each model and different parameter sets for at-the-money European plain vanilla, arithmetic Asian and barrier call options

			Plain vanilla option			
	H1	H2	H3	H4	H5	H6
True option price	6.8061	34.9998	6.7304	7.0972	11.3743	7.0737
Bias	0.0001	-0.0029	-0.0002	0.0007	-0.0006	-0.0014
Standard error	0.0007	0.0058	0.0005	0.0011	0.0020	0.0008
	SABR1	SABR2	SABR3	OU-SV1	OU-SV2	OU-SV3
True option price	0.0394	0.0436	0.0447	13.2149	40.7977	62.7631
Bias	1.72e-06	-2.3e-06	1.16e-07	0.0013	-0.0014	0.0012
Standard error	4.11e-07	3.99e-07	3.91e-07	0.0015	0.0043	0.0067
	DH1	DH2	3/2-1	3/2-2	4/2-1	4/2-2
True option price	26.9504	9.3663	6.3991	2.1537	8.6592	9.8629
Bias	-0.0002	-0.0005	-0.0003	0.0002	0.0060	0.0010
Standard error	0.0020	0.0002	0.0008	0.0002	0.0015	0.0017
	B1	DPS1	NIGCIR1	SECIRJD1		
True option price	20.1645	6.8619	6.6425	9.7639		
Bias	0.0001	0.0027	-3.87E-05	0.0004		
Standard error	7.09E-04	2.32E-04	2.35E-04	4.63E-04		
			Asian option			
	H1	H2	H3	H4	H5	H6
True option price	3.5665	18.1576	4.1061	4.3222	6.6513	3.8590
Bias	-0.0001	-0.0028	0.0003	0.0003	0.0005	-0.0007
Standard error	0.0004	0.0027	0.0003	0.0006	0.0011	0.0004
			Barrier option			
	H1	H2	H3	H4	H5	H6
True option price	4.9142	0.1803	6.3748	4.5714	2.6489	4.2925
Bias	-0.0004	1.34e-05	-0.0004	-0.0004	-7.1e-05	0.0008
Standard error	0.0006	0.0001	0.0004	0.0005	0.0005	0.0006

Table 4: Speed-accuracy profiles of Algorithm 2 and some competent benchmarks (Broadie and Kaya, 2006, Glasserman and Kim, 2011, Cai *et al.*, 2017, Li and Wu, 2019) for different models and parameter sets: the case of European plain vanilla option

	Broadie–Kaya		Algorithm 2		Broadie–Kaya		Algorithm 2		Glasserman–Kim		Algorithm 2	
$\mathcal{M} \times 10^4$	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time
	H1				H2				H3			
4	0.0373	273.45	0.0372	0.21	0.2904	160.85	0.3000	0.23	0.0246	0.49	0.0244	0.25
16	0.0186	1076.72	0.0186	0.82	0.1464	643.97	0.1440	0.84	0.0123	1.95	0.0122	0.81
64	0.0093	4028.62	0.0093	3.29	0.0726	2458.06	0.0734	3.15	0.0061	7.74	0.0061	3.33
256	0.0046	16884.00	0.0046	13.21	0.0362	10467.25	0.0367	14.28	0.0031	31.96	0.0031	12.95
	Glasserman–Kim		Algorithm 2		Glasserman–Kim		Algorithm 2		Glasserman–Kim		Algorithm 2	
$\mathcal{M} \times 10^4$	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time
	H4				H5				H6			
4	0.0507	0.56	0.0513	0.25	0.0973	0.86	0.1018	0.28	0.0394	0.82	0.0411	0.31
16	0.0257	2.30	0.0262	0.94	0.0489	3.44	0.0489	0.92	0.0198	3.28	0.0205	0.80
64	0.0132	9.38	0.0132	3.14	0.0248	13.86	0.0249	3.39	0.0099	13.11	0.0103	3.14
256	0.0068	35.94	0.0066	13.11	0.0124	55.43	0.0123	13.09	0.0050	52.94	0.0053	12.28
	Cai–Song–Chen		Algorithm 2		Cai–Song–Chen		Algorithm 2		Cai–Song–Chen		Algorithm 2	
$\mathcal{M} \times 10^4$	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time
	SABR1				SABR2				SABR3			
4	2.04e-05	1.65	2.07e-05	0.23	1.98e-05	1.55	2.00e-05	0.2	1.93e-05	1.53	1.97e-05	0.19
16	1.03e-05	6.34	1.05e-05	0.87	9.96e-06	6.28	1.02e-05	0.79	9.76e-06	6.3	9.83e-06	0.78
64	5.14e-06	25.39	5.42e-06	3.35	4.99e-06	25.12	5.50e-06	3.04	4.89e-06	25.07	4.90e-06	2.95
256	2.57e-06	100.7	3.09e-06	13.7	2.50e-06	99.64	3.39e-06	12.41	2.45e-06	99.41	2.45e-06	12.29
	Li–Wu		Algorithm 2		Li–Wu		Algorithm 2		Li–Wu		Algorithm 2	
$\mathcal{M} \times 10^4$	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time
	OU-SV1				OU-SV2				OU-SV3			
4	0.0764	15.60	0.0757	0.11	0.2345	21.66	0.2002	0.11	0.3774	21.64	0.3382	0.10
16	0.0382	63.20	0.0377	0.38	0.1176	82.87	0.1024	0.39	0.1891	85.25	0.1687	0.40
64	0.0192	253.55	0.0188	1.49	0.0591	332.22	0.0550	1.51	0.0951	341.73	0.0839	1.49
256	0.0096	1013.31	0.0095	6.47	0.0295	1328.01	0.0339	6.22	0.0475	1366.03	0.0421	6.14
	Algorithm 2		Algorithm 2		Algorithm 2		Algorithm 2		Algorithm 2		Algorithm 2	
$\mathcal{M} \times 10^4$	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time
	DH1		DH2		3/2-1		3/2-2		4/2-1		4/2-2	
4	0.0989	0.83	0.0107	0.74	0.0392	0.70	0.0989	0.75	0.0740	2.12	0.0872	2.18
16	0.0497	3.01	0.0053	2.96	0.0197	2.93	0.0497	3.11	0.0375	9.17	0.0432	9.75
64	0.0248	12.29	0.0027	12.11	0.0098	12.09	0.0248	12.39	0.0195	36.85	0.0216	37.29
256	0.0124	48.15	0.0014	47.89	0.0049	50.75	0.0124	51.42	0.0110	147.47	0.0108	149.12
	Algorithm 2		Algorithm 2		Algorithm 2		Algorithm 2					
$\mathcal{M} \times 10^4$	RMSE	time	RMSE	time	RMSE	time	RMSE	time				
	B1		DPS1		NIGCIR1		SECIRJD1					
4	0.1117	0.18	0.0368	7.96	0.0372	0.26	0.0731	12.22				
16	0.0559	0.58	0.0186	32.23	0.0186	0.81	0.0366	51.30				
64	0.0281	2.26	0.0096	138.43	0.0093	2.84	0.0183	203.79				
256	0.0141	8.82	0.0054	527.43	0.0046	11.38	0.0092	849.69				

Notes. All computing times are in seconds. For convenience, out of scale Monte Carlo benchmarks for the DH, 3/2, 4/2, Bates, DPS, NIGCIR and SECIRJD models are not reported.

Table 5: Speed-accuracy profiles of Algorithm 2 and Andersen's (2008) method in the Heston model and different parameter sets: the case of path-dependent derivatives

Asian option

	Andersen		Algorithm 2		Andersen		Algorithm 2		Andersen		Algorithm 2	
$\mathcal{M} \times 10^4$	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time
	H1				H2				H3			
4	0.0261	0.98	0.0192	3.42	0.9372	5.03	0.9213	15.78	0.0239	0.86	0.0158	3.01
16	0.0201	5.06	0.0096	12.48	0.5350	27.89	0.5066	59.80	0.0195	3.91	0.0078	12.08
64	0.0183	34.74	0.0048	50.45	0.2908	331.26	0.2343	238.44	0.0183	17.69	0.0039	56.55
256	0.0179	213.71	0.0024	201.43	0.2083	3312.59	0.1172	953.92	0.0180	192.28	0.0020	272.12
	Andersen		Algorithm 2		Andersen		Algorithm 2		Andersen		Algorithm 2	
$\mathcal{M} \times 10^4$	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time
	H4				H5				H6			
4	0.0334	0.86	0.0282	3.27	0.0595	1.07	0.0523	3.51	0.0284	0.95	0.0220	3.94
16	0.0231	3.93	0.0145	12.50	0.0386	4.90	0.0260	14.62	0.0211	5.12	0.0111	15.50
64	0.0194	18.22	0.0072	53.84	0.0314	21.62	0.0132	66.83	0.0188	20.08	0.0056	59.86
256	0.0184	198.61	0.0036	279.79	0.0292	186.54	0.0066	355.53	0.0182	196.87	0.0029	235.37

Barrier option

	Andersen		Algorithm 2		Andersen		Algorithm 2		Andersen		Algorithm 2	
$\mathcal{M} \times 10^4$	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time
	H1				H2				H3			
4	0.0338	0.98	0.0289	3.42	0.0054	5.03	0.0023	15.78	0.0220	0.86	0.0222	3.01
16	0.0233	5.07	0.0145	12.48	0.0031	27.90	0.0012	59.80	0.0111	3.92	0.0110	12.08
64	0.0199	34.76	0.0072	50.46	0.0022	331.17	0.0006	238.46	0.0058	17.71	0.0055	56.56
256	0.0189	213.74	0.0036	201.46	0.0019	3311.68	0.0003	954.00	0.0033	192.29	0.0028	272.15
	Andersen		Algorithm 2		Andersen		Algorithm 2		Andersen		Algorithm 2	
$\mathcal{M} \times 10^4$	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time	RMSE	time
	H4				H5				H6			
4	0.0242	0.87	0.0243	3.27	0.0274	1.07	0.0243	3.51	0.0336	0.95	0.0284	3.95
16	0.0128	3.94	0.0121	12.51	0.0194	4.90	0.0122	14.62	0.0241	5.13	0.0142	15.50
64	0.0076	18.23	0.0060	53.85	0.0169	21.64	0.0061	66.83	0.0210	20.11	0.0071	59.87
256	0.0056	198.62	0.0030	279.82	0.0161	186.63	0.0030	355.57	0.0202	197.01	0.0036	235.40

Notes. All computing times are in seconds.

Table 6: Barrier option prices (with standard errors, s.e.) in the SABR model

Method	price	s.e.	time	price	s.e.	time	price	s.e.	time
	SABR4			SABR5			SABR6		
mSABR	0.0384	–	–	0.00523	–	–	0.1078	–	–
Chen et al.	0.0380	1.12e-02	3.77	0.00520	8.43e-04	4.47	0.1022	2.12e-02	21.81
Algorithm 2	0.0385	1.13e-02	21.91	0.00523	8.41e-04	25.72	0.1079	2.12e-02	102.24

Notes. All computing times are in seconds and correspond to $\mathcal{M} = 10^6$ simulations and quarterly monitoring (as in Leitao *et al.*, 2017, Section 4). Benchmarks: mSABR (Leitao *et al.*, 2017); Chen *et al.* (2012). Barrier levels: $U = 1.2$ (SABR4); $U = 0.08$ (SABR5); $U = 2$ (SABR6).

Table 7: Simulation of $E\left(X^2(3)\right)$ (with standard errors, s.e.) in the stochastic Ginzburg–Landau model

	$\sigma = 2$		$\sigma = 4$		$\sigma = 5$		$\sigma = 6$		$\sigma = 7$	
	$E\left(X^2(3)\right)$	s.e.	$E\left(X^2(3)\right)$	s.e.	$E\left(X^2(3)\right)$	s.e.	$E\left(X^2(3)\right)$	s.e.	$E\left(X^2(3)\right)$	s.e.
Algorithm 3	0.4689	4.12e-04	0.9138	1.18e-03	1.1455	1.66e-03	1.3693	2.18e-03	1.5989	2.75e-03
	Euler approximation									
Batch 1	0.4556	4.10e-03	0.7553	1.10e-02	0.7106	1.33e-02	NaN	NaN	NaN	NaN
Batch 2	0.4514	4.09e-03	0.7551	1.12e-02	0.6964	1.32e-02	NaN	NaN	NaN	NaN
Batch 3	0.4576	4.16e-03	0.7582	1.10e-02	0.7065	1.33e-02	0.5191	1.43e-02	NaN	NaN
Batch 4	0.4572	4.08e-03	0.7417	1.08e-02	0.7126	1.39e-02	0.5353	1.43e-02	NaN	NaN
Batch 5	0.4608	4.09e-03	0.7274	1.05e-02	0.7043	1.36e-02	NaN	NaN	NaN	NaN
Batch 6	0.4491	4.07e-03	0.7308	1.05e-02	0.7224	1.36e-02	0.5415	1.47e-02	NaN	NaN
Batch 7	0.4595	4.08e-03	0.7393	1.08e-02	0.7310	1.42e-02	NaN	NaN	NaN	NaN
Batch 8	0.4592	4.03e-03	0.7333	1.07e-02	0.7344	1.43e-02	NaN	NaN	NaN	NaN
Batch 9	0.4589	4.10e-03	0.7539	1.11e-02	0.7108	1.34e-02	NaN	NaN	NaN	NaN
Batch 10	0.4626	4.12e-03	0.7555	1.08e-02	0.7096	1.34e-02	0.5366	1.41e-02	NaN	NaN

Notes. The number of Monte Carlo simulations used for each estimate is 10^5 for the Euler approximation (based on $1,000$ time steps) and 10^7 for Algorithm 3. Parameters are from Hutzenthaler *et al.* (2010): $X(0) = 1$, $\alpha = 0$, $\beta = 1$, $\sigma = \{2, 4, 5, 6, 7\}$.

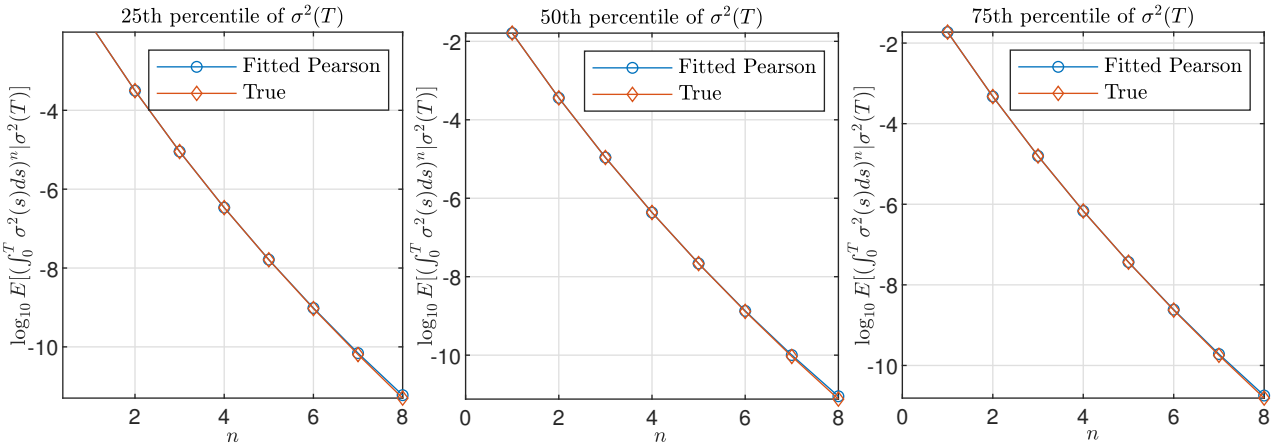


Figure 1: True moments of the integrated variance in the Heston model (parameter set H1) conditional on the 25, 50, 75th percentiles of the terminal variance and corresponding moments based on fitted Pearson distribution

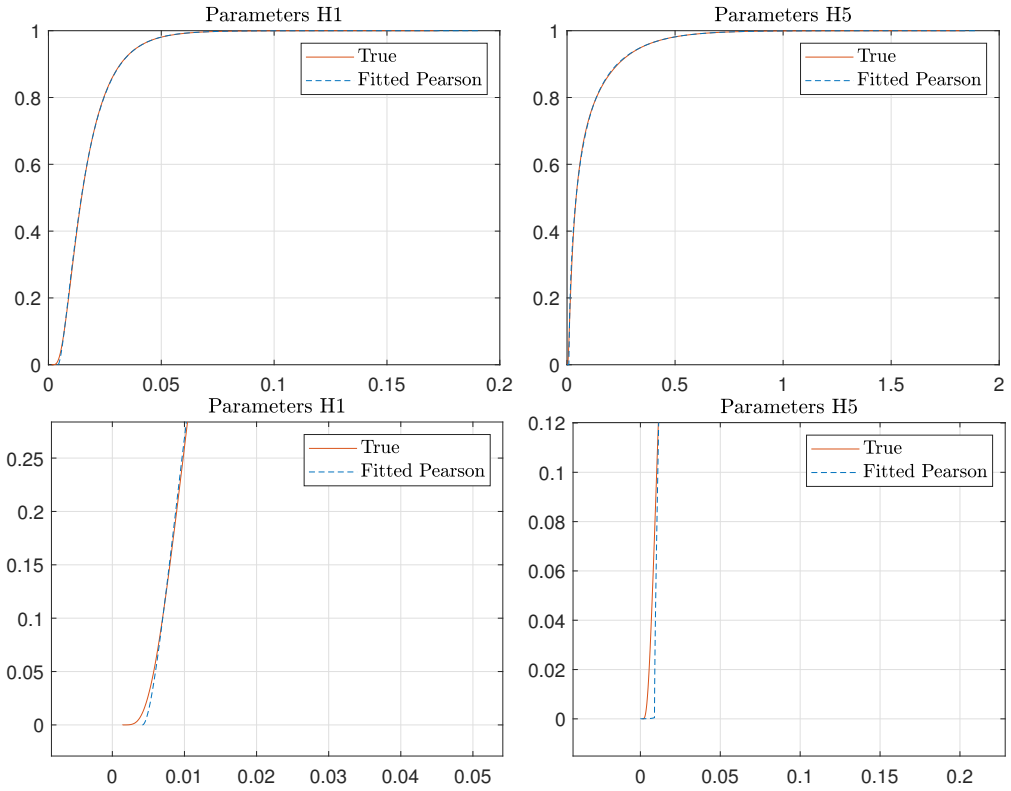
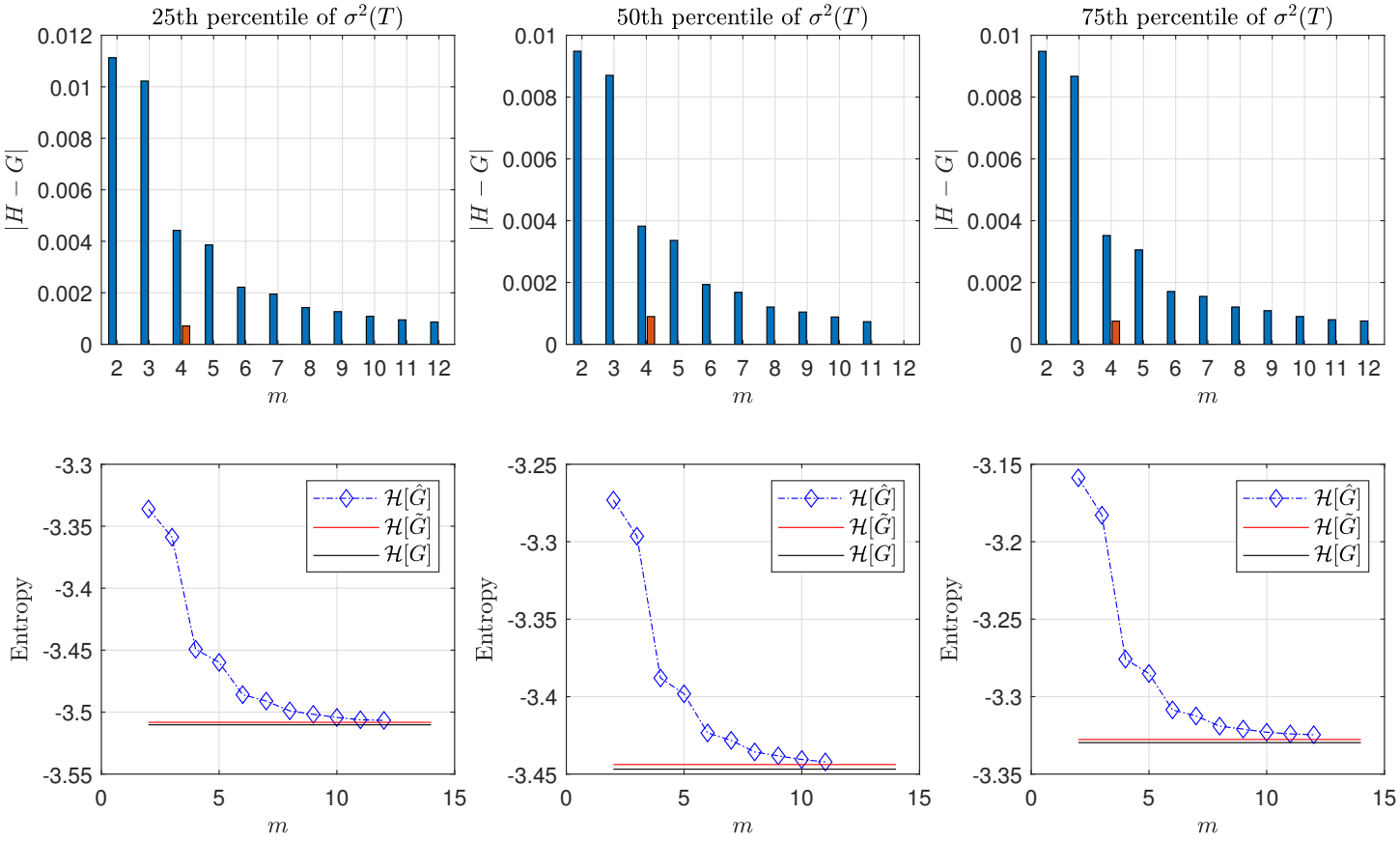


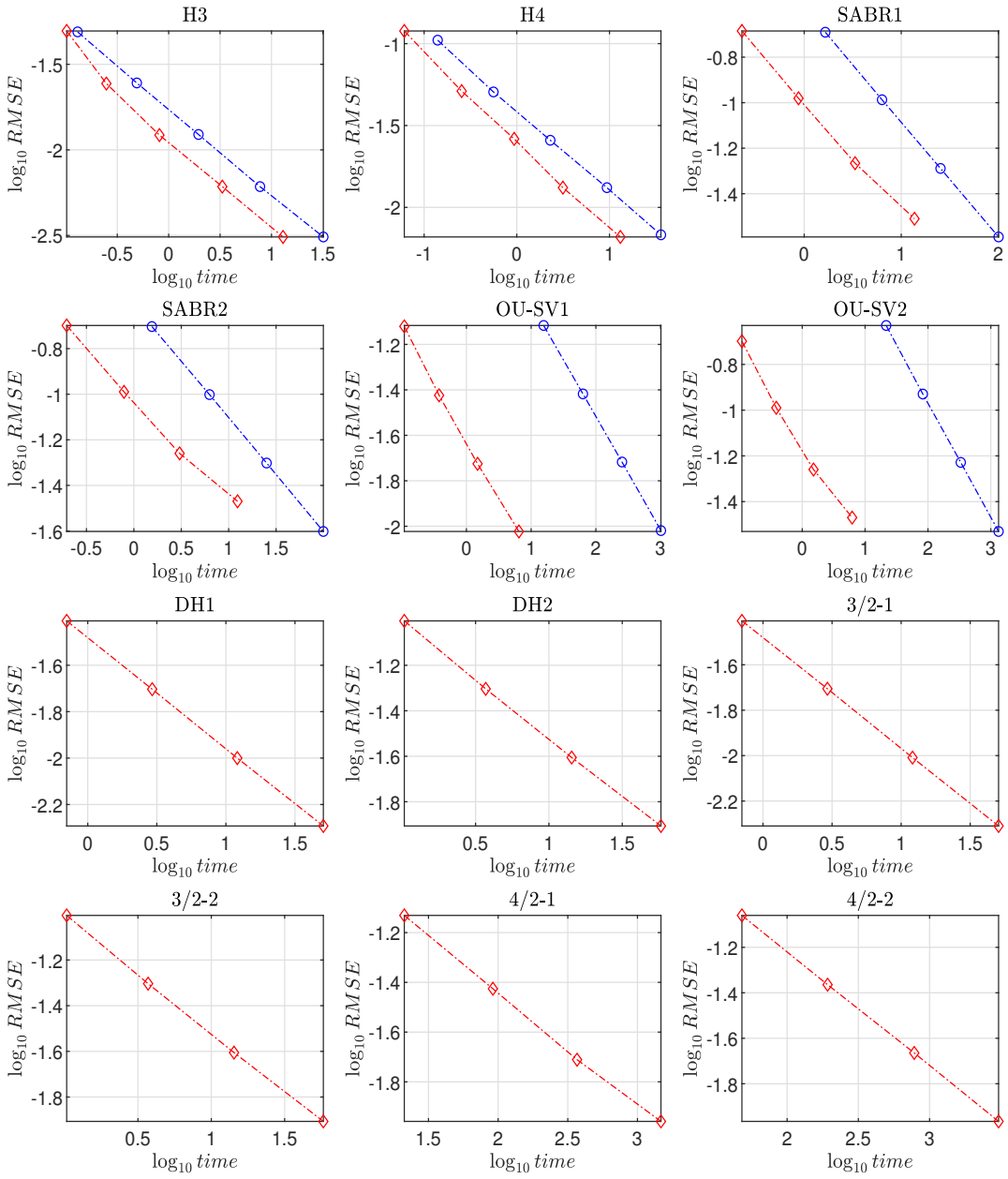
Figure 2: True and fitted Pearson cumulative distribution functions, G and $\tilde{G}$, of the integrated variance in the Heston model (parameter sets H1 and H5) conditional on the 50th percentile of the terminal variance

Figure 3: Entropy bound



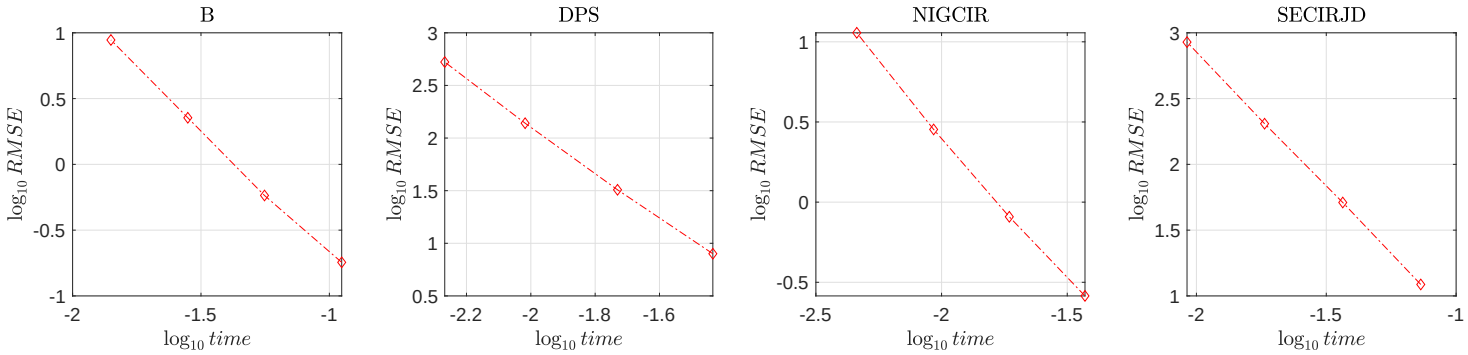
Notes. Top plots: absolute difference $|H(x) - G(x)|$ bounds (12) for true distribution G and approximation $H \in \left\{\hat{G}, \tilde{G}\right\}$, where $\hat{G}$ corresponds to the entropy-maximizing distribution with m moments and $\tilde{G}$ to the four-moment fitted Pearson distribution, corresponding to the integrated variance in the Heston model (parameter set H1) conditional on the 25, 50, 75th percentiles of the terminal variance. Bottom plots: corresponding entropies $\mathcal{H}[\hat{G}]$, $\mathcal{H}[\tilde{G}]$ and $\mathcal{H}[G]$.

Figure 4: Speed-accuracy comparisons of Algorithm 2 and competent benchmarks for different models and parameter sets: the case of European plain vanilla option



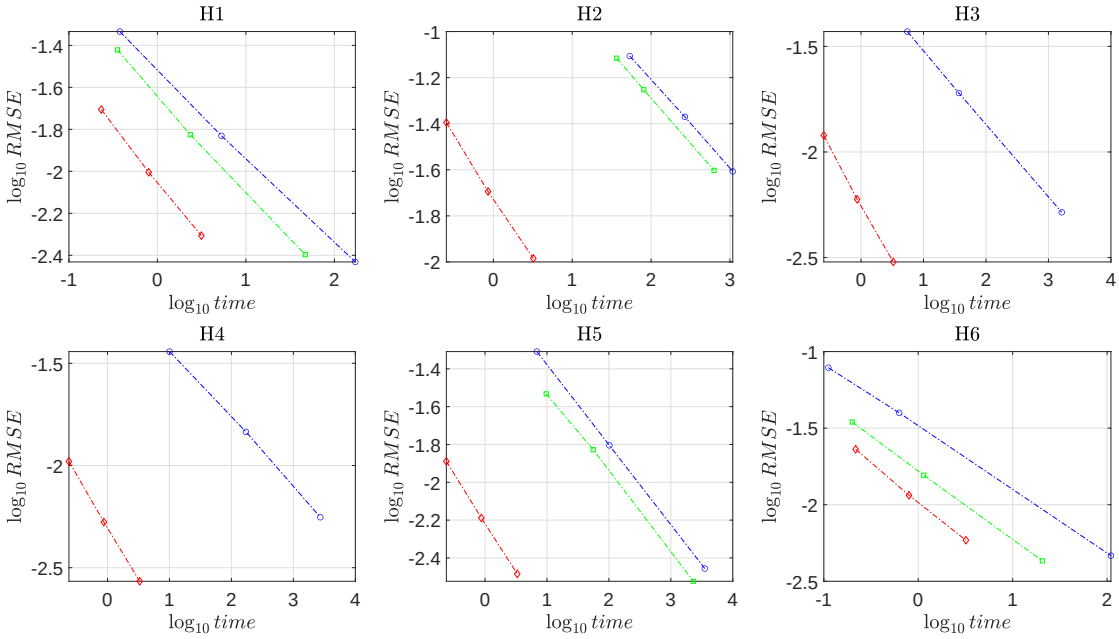
Notes. Algorithm 2: plots with diamond markers; benchmarks: plots with circle markers. Benchmarks: Glasserman and Kim (2011) (H3–4); Cai *et al.* (2017) (SABR1–2); Li and Wu (2019) (OU-SV1–2). For convenience, out of scale Monte Carlo benchmarks for the DH, 3/2, 4/2 models are not reported. All computing times are in seconds.

Figure 5: Speed-accuracy profiles of Algorithm 2 in the Bates, DPS, NIGCIR and SECIRJD models: the case of European plain vanilla option



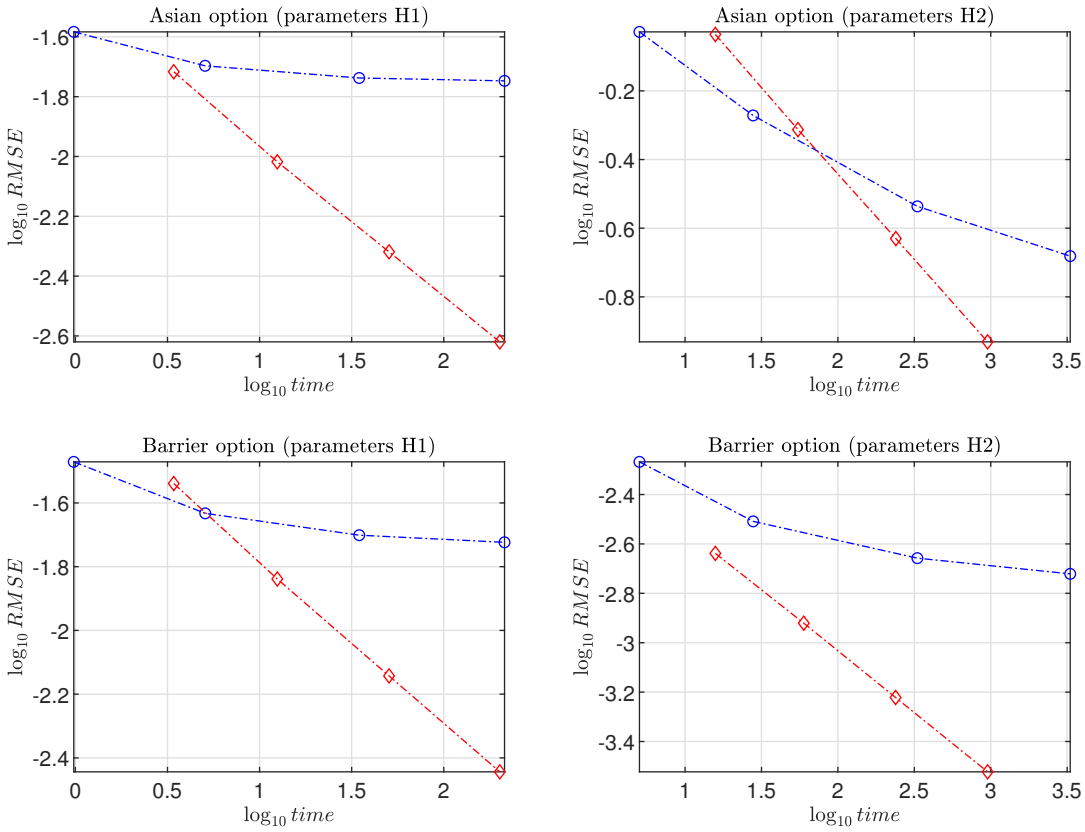
Notes. All computing times are in seconds.

Figure 6: Speed-accuracy comparisons of Algorithm 2 and multi-level Monte Carlo (MLMC) methods for different parameter sets in the Heston model: the case of European plain vanilla option.



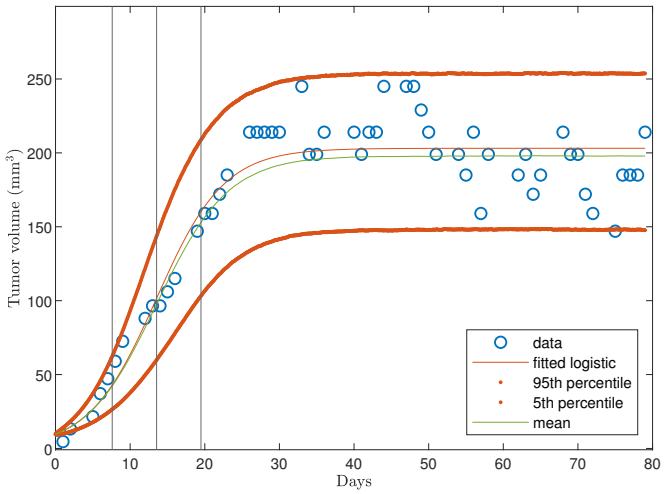
Notes. Algorithm 2: plots with red diamond markers; benchmarks: plots with blue circle (green square) markers correspond to MLMC (antithetic MLMC). MLMC: geometric sequence of time step sizes $h_l = M^{-l}T$, $l = 0, 1, \ldots, L$, for $M = 4$ and $L = 6$, optimal initial number of simulations $\mathcal{M}_0 = 10^4$ (see Giles, 2008), target RMSE $= \{0.05, 0.02, 0.005\}$ and simulation performed based on Euler scheme. Antithetic MLMC: based on a Milstein numerical approximation of the Heston SDE with use of antithetic variables (see Giles and Szpruch, 2014). All computing times are in seconds.

Figure 7: Speed-accuracy profiles of Algorithm 2 and Andersen's (2008) method in the Heston model and different parameter sets: the case of path-dependent derivatives



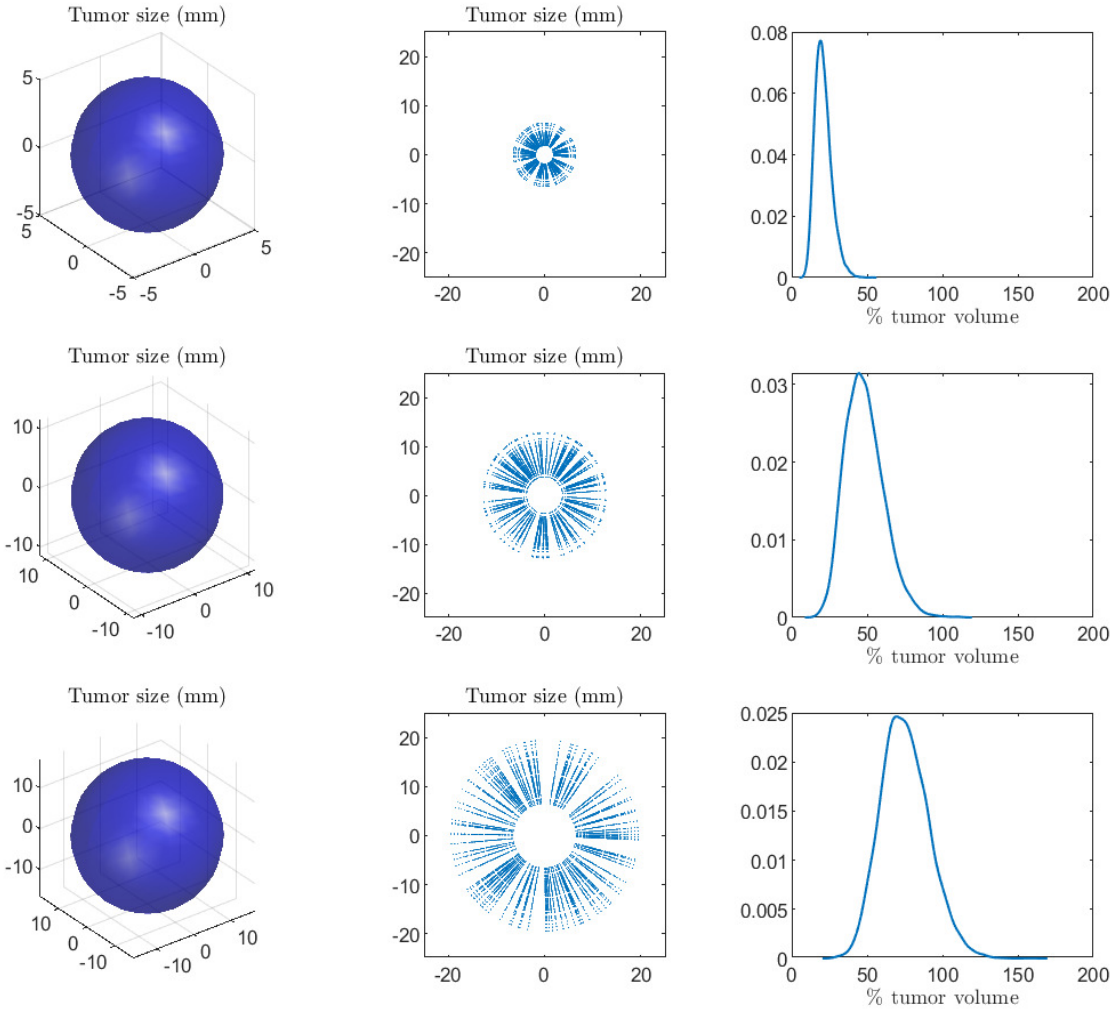
Notes. Algorithm 2: plots with diamond markers; benchmark: plots with circle markers. Benchmark: Andersen (2008) (H1–2). All computing times are in seconds.

Figure 8: Stochastic logistic model for tumor volume



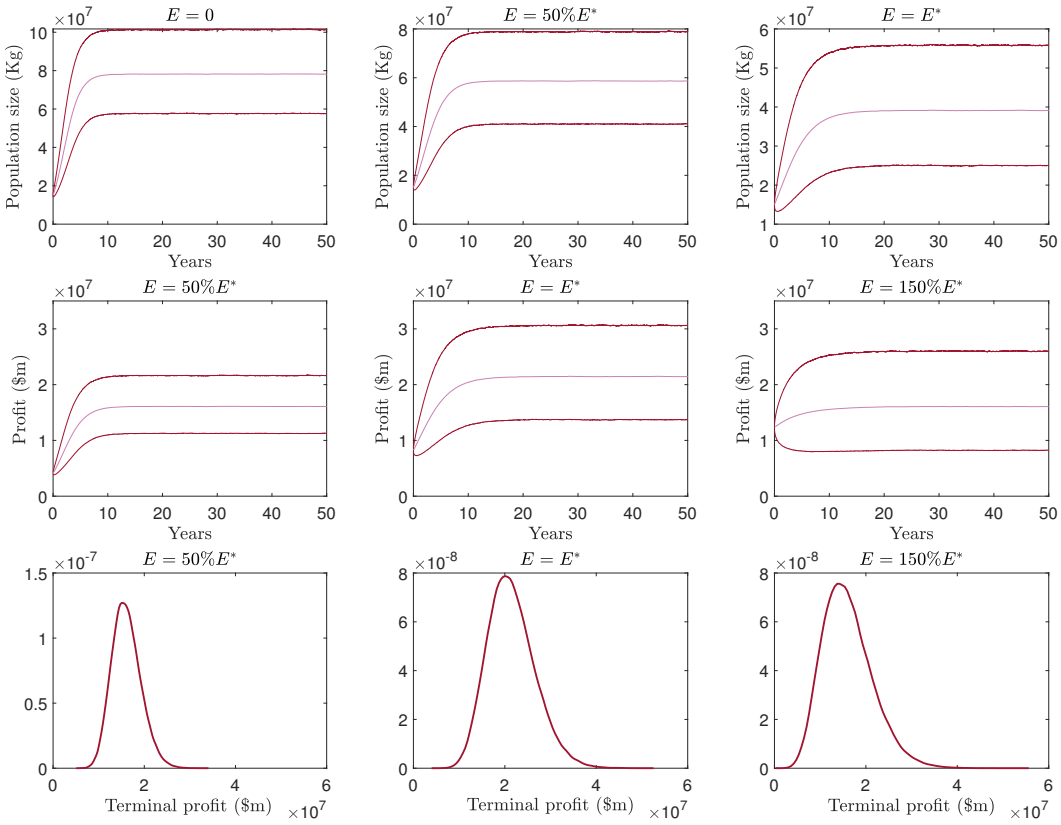
Notes. Historical daily tumor growth; fitted deterministic Verhulst equation; probabilistic forecasts for tumor growth generated by the model (37): 95% and 5% confidence bands (upper and lower lines) based on 10^5 simulated trajectories; mean value estimate (medium line).

Figure 9: Tumor growth



Notes. Tumor size (3D & 2D) (left & central plots) and probability density estimates of % tumor volume with respect to estimated maximum limit (right plots) at the three different phases (top to bottom): T_1 (completion of initial growth), T_2 (aggressive tumor growth) and T_3 (terminal growth).

Figure 10: Impact of varying fishing effort E on population size of Pacific halibut and fishing profit



Notes. Population size (top plots) and profit (central plots): 95% and 5% confidence bands (upper and lower lines) and mean value estimate (medium line) based on 10^5 simulated trajectories. Probability density estimates of terminal profit (bottom plots). Parameter estimates are from Hanson and Ryan (1998): $r = 0.71$, $K = 8.05 \times 10^7$ kg, $q = 3.30 \times 10^{-6}$, $p = 1.59$, $c = 96 \times 10^{-6} + 0.10 \times 10^{-6} E$. In addition, $X(0) = 1.5 \times 10^7$ kg, $\sigma = 0.2$ and $E^* = 104,540$ (optimal sustainable effort).

E-companion to "Unified moment-based modelling of integrated stochastic processes"

EC.1. Proofs

Proof of Proposition 1. From (1),

$$\begin{aligned}\int_{\mathbb{R}} e^{iux}\tilde{g}(x)\,dx &= \left[e^{iux}\left(c_1+c_2x+c_3x^2\right)\tilde{g}(x)\right]_{-\infty}^{\infty} \\ &\quad -\int_{\mathbb{R}} e^{iux}\left\{iu\left(c_1+c_2x+c_3x^2\right)+(c_2+2c_3x)\right\}\tilde{g}(x)\,dx \\ &= -(iuc_1+c_2)\int_{\mathbb{R}} e^{iux}\tilde{g}(x)\,dx-(iuc_2+2c_3)\int_{\mathbb{R}} e^{iux}x\tilde{g}(x)\,dx \quad \text{(EC.1)} \\ &\quad -iuc_3\int_{\mathbb{R}} e^{iux}x^2\tilde{g}(x)\,dx,\end{aligned}$$

where in the first equality $\left[e^{iux}\left(c_1+c_2x+c_3x^2\right)\tilde{g}(x)\right]_{-\infty}^{\infty}=0$ by $\mu_2<\infty$. Differentiation of $\varphi_{\tilde{G}}(u)$ under the integral sign is allowed because $e^{iux}\tilde{g}(x)$ is twice continuously differentiable in u and once in x. Hence, we obtain $\varphi'_{\tilde{G}}(u)=i\int_{\mathbb{R}} e^{iux}x\tilde{g}(x)\,dx$ and $\varphi''_{\tilde{G}}(u)=-\int_{\mathbb{R}} e^{iux}x^2\tilde{g}(x)\,dx$, based on which (EC.1) can be written as

$$iuc_3\varphi''_{\tilde{G}}(u)+(-uc_2+2ic_3-i)\,\varphi'_{\tilde{G}}(u)+(c_0-iuc_1-c_2)\,\varphi_{\tilde{G}}(u)=0.$$

From this, (5) follows for $c_0-c_2=0$ by construction. ■

Proof of Proposition 2. We have that

$$\begin{aligned}\int_0^t Y(s;a,b,\gamma)ds &= \int_0^t \exp\left(\gamma\left(a-\frac{1}{2}b^2\right)s+\gamma bW(s)\right)ds \\ &\overset{\text{(law)}}{=} \frac{4}{\gamma^2b^2}\int_0^{\gamma^2b^2t/4}\exp\left(\frac{4}{\gamma}\left(\frac{a}{b^2}-\frac{1}{2}\right)s+2W(s)\right)ds \\ &= \frac{4}{\gamma^2b^2}\int_0^{\gamma^2b^2t/4}\exp\left(2W^{(\mu)}(s)\right)ds,\end{aligned}$$

where the second equality follows from the scaling property of the Brownian motion for $\gamma,b>0$ and $\mu:=2a/(\gamma b^2)-1/\gamma$. In addition, by considering the condition in (34), we get

$$w=W^{(\mu)}\left(\frac{\gamma^2b^2t}{4}\right)\overset{\text{(law)}}{=}\mu\frac{\gamma^2b^2t}{4}+\frac{\gamma b}{2}W(t)=\frac{\ln Y(t;a,b,\gamma)}{2}.$$

Expression (35) then follows and the proposition is proved. ■

EC.2. More on moment-based metrics

A sharpened version of Klebanov and Mkrtchyan's (1986) inequality is

$$L\left(G,\tilde{G}\right)\leq C\frac{\gamma_{k,m-1}^{1/(2k+2)}}{(k+1)^2}\left|\ln\frac{(2k)!}{\gamma_{k,m-1}}\right|$$

for $0 \leq k \leq m-1$, where

$$\begin{cases} \gamma_{k,m} = \min\left\{2\mu_{2k}, \mu_{2k}\frac{2+\sqrt{\frac{\mu_{2k+2}}{\mu_{2k}}}}{\sqrt{\beta_{k,m}}}\right\} \text{ for } k \leq m-1 \\ \gamma_{m,m} = \mu_{2m} \end{cases}.$$

Although by construction we are exactly matching the first four integer moments, Rachev *et al.* (2013, Theorem 10.3.4) allows us to restate Klebanov and Mkrtchyan's (1986) inequality when additional moments of the two distributions are not coinciding but are only fairly close. Therefore, for G and $\tilde{G}$ with finite moments up to order $2m$ and

$$\left|\mu_n(G) - \tilde{\mu}_n\left(\tilde{G}\right)\right| \leq \delta_{G,\tilde{G}}, \quad n = 1, \ldots, 2m,$$

where $0 < \delta_{G,\tilde{G}} \leq 1$, we have that

$$L\left(G, \tilde{G}\right) \leq \frac{2}{\ln\left(1 + \delta_{G,\tilde{G}}^{-1/2}\right)} + \left(\frac{2\mu_{2m}}{(2m)!}\right)^{\frac{1}{2m+1}}.$$

Finally, related to the distance between distribution functions is also Esseen's inequality recorded next. Consider first some non-negative function $H(x)$ and $h(u) := \int_{\mathbb{R}} e^{iux} H(x)\, dx$ satisfying

$$\begin{cases} \int_{\mathbb{R}} H(x)\, dx = 1;\ b = \int_{\mathbb{R}} |x|\, H(x)\, dx < \infty \\ h(0) = 1;\ h(u) = 0 \text{ for } |u| \geq 1;\ 0 \leq |h(u)| \leq 1 \text{ for } |u| \leq 1 \end{cases}. \tag{EC.2}$$

A possible choice for H, as suggested by Esseen (1945), is $H(x) = (3/(8\pi))(4\sin(x/4)/x)^4$. Then, for the Pearson distribution function $\tilde{G}(x)$ which is differentiable almost everywhere, has bounded variation and finite ϱ (see Lemma 3 in the paper), Theorem 2a from Esseen (1945) applies. More specifically, to every arbitrary $k > 1$, there correspond $0 < \bar{m}(k) < 1$ and $\alpha(k)$ so that

$$(2 - \bar{m}(k)) \int_{-\bar{m}(k)\alpha(k)}^{(1-\bar{m}(k))\alpha(k)} H(x)\, dx = \frac{k+1}{k}, \tag{EC.3}$$

where H satisfies the conditions in (EC.2), and

$$\left|G(x) - \tilde{G}(x)\right| \leq \max\left\{\frac{k\epsilon}{2\pi} + \frac{kb\varrho}{\mathcal{T}}, \frac{\alpha(k)\,\varrho}{\mathcal{T}}\right\} \leq \frac{k\epsilon}{2\pi} + \frac{(kb + \alpha(k))\,\varrho}{\mathcal{T}}$$

with $(\mathcal{T}, \epsilon)$ satisfying

$$\int_{-\mathcal{T}}^{\mathcal{T}} \left|\frac{\varphi_G(u) - \varphi_{\tilde{G}}(u)}{u}\right| du = \epsilon.$$

We can always choose $\bar{m}(k)$ sufficiently small and $\alpha(k)$ sufficiently large so that (EC.3) is fulfilled.

EC.3. Laplace transforms

The conditional Laplace transforms for the Heston, SABR, OU-SV and 4/2 models follow, respectively, from Broadie and Kaya (2006), Cai *et al.* (2017), Li and Wu (2019) and Grasselli (2017, Proposition A.4).

EC.3.1. Heston model

The Laplace transform of the conditional integrated variance is

$$\begin{aligned}\mathcal{L}(a) &= E\left[\exp\left(-a\int_u^t \sigma^2(s)ds\right)\middle|\,\sigma(u),\sigma(t)\right] = \frac{\gamma(a)e^{-(\gamma(a)-k)(t-u)/2}(1-e^{-k(t-u)})}{k(1-e^{-\gamma(a)(t-u)})}\times \\ &\exp\left\{\frac{\sigma^2(u)+\sigma^2(t)}{v^2}\left(\frac{k(1+e^{-k(t-u)})}{1-e^{-k(t-u)}}-\frac{\gamma(a)(1+e^{-\gamma(a)(t-u)})}{1-e^{-\gamma(a)(t-u)}}\right)\right\}\times \\ &\frac{I_{d/2-1}\left(\sigma(u)\sigma(t)\frac{4\gamma(a)e^{-\gamma(a)(t-u)/2}}{v^2(1-e^{-\gamma(a)(t-u)})}\right)}{I_{d/2-1}\left(\sigma(u)\sigma(t)\frac{4ke^{-k(t-u)/2}}{v^2(1-e^{-k(t-u)})}\right)},\end{aligned} \tag{EC.4}$$

where $\gamma(a) := \sqrt{k^2+2v^2a}$ and $I_\nu(\cdot)$ denotes the modified Bessel function of the first kind.

EC.3.2. SABR model

The Laplace transform of the conditional reciprocal of the integrated variance is

$$\begin{aligned}\mathcal{L}(a) &= E\left[\exp\left\{-a\left(\int_u^t \sigma^2(s)ds\right)^{-1}\right\}\middle|\,\sigma(u),\sigma(t)\right] \\ &= \exp\left\{-\frac{g\left(\ln\frac{\sigma(t)}{\sigma(u)},\frac{av^2}{\sigma^2(u)}\right)^2-\ln\left(\frac{\sigma(t)}{\sigma(u)}\right)^2}{2v^2(t-u)}\right\},\end{aligned} \tag{EC.5}$$

where $g(x,\lambda) := \operatorname{arcosh}(\lambda e^{-x}+\cosh x)$.

EC.3.3. OU-SV model

The Laplace transform of the conditional integrated variance is

$$\mathcal{L}(a) = E\left[\exp\left(-a\int_u^t \sigma^2(s)ds\right)\middle|\,\sigma(u),\sigma(t),\int_u^t\sigma(s)ds\right] = \frac{f(\gamma(a))}{f(k)}, \tag{EC.6}$$

where

$$\begin{aligned}f(x) := \frac{x^2}{2\pi\sqrt{\eta(x)}}\exp\left\{-\frac{1}{2\eta(x)v^2}\left[2x^2\left((\sigma(t)+\sigma(u))\int_u^t\sigma(s)ds-\sigma(u)\sigma(t)(t-u)\right)+\right.\right. \\ x^2\left((\sigma^2(t)+\sigma^2(u))(t-u)-2(\sigma(t)+\sigma(u))\int_u^t\sigma(s)ds\right)\cosh(x(t-u))+ \\ \left.\left.x\left(x^2\left(\int_u^t\sigma(s)ds\right)^2-(\sigma(t)-\sigma(u))^2\right)\sinh(x(t-u))\right]\right\}\end{aligned}$$

and

$$\eta(x) := 2-2\cosh(x(t-u))+x(t-u)\sinh(x(t-u)).$$

EC.3.4. 4/2 and 3/2 models

The joint Laplace transform of $\left(\int_u^t\sigma^2(s)ds,\int_u^t\frac{ds}{\sigma^2(s)}\middle|\,\sigma^2(u),\sigma^2(t)\right)$ is

$$
\begin{aligned}
&E\left[\exp\left(-a\int_u^t \sigma^2(s)ds - b\int_u^t \frac{ds}{\sigma^2(s)}\right)\middle|\, \sigma^2(u), \sigma^2(t)\right] \\
&= \frac{\sqrt{\gamma(a)}\sinh\left(\frac{k(t-u)}{2}\right)}{k\sinh\left(\frac{\sqrt{\gamma(a)}(t-u)}{2}\right)} e^{\frac{\sigma^2(u)+\sigma^2(t)}{v^2}\left(k\coth\left(\frac{k(t-u)}{2}\right)-\sqrt{\gamma(a)}\coth\left(\frac{\sqrt{\gamma(a)}(t-u)}{2}\right)\right)} \\
&\times \frac{I_{\frac{\sqrt{(2\theta k-v^2)^2+8v^2b}}{v^2}}\left(\frac{2\sqrt{\gamma(a)\sigma^2(u)\sigma^2(u)}}{v^2\sinh\left(\frac{\sqrt{\gamma(a)}(t-u)}{2}\right)}\right)}{I_{\frac{2\theta k}{v^2}-1}\left(\frac{2k\sqrt{\sigma^2(u)\sigma^2(t)}}{v^2\sinh\left(\frac{k(t-u)}{2}\right)}\right)}.
\end{aligned}
\tag{EC.7}
$$

The Laplace transform of $\left(\int_u^t \frac{ds}{\sigma^2(s)}\middle|\, \sigma^2(u), \sigma^2(t)\right)$ in the 3/2 model follows as special case by setting $a = 0$ in (EC.7).

EC.4. Sampling from the Pearson family of distributions

In what follows, we present the different Pearson distribution types and corresponding generators of random numbers $\hat{Y}$ with zero mean, unitary variance $\beta = 1$, skewness $\sqrt{\gamma}$ and kurtosis ε. To this end, we consider first a few quantities that will be used next:

$$
z := -\frac{\sqrt{\gamma}(\varepsilon+3) + \operatorname{sgn}(\sqrt{\gamma}(\varepsilon+3))\sqrt{\gamma^{1/2}(\varepsilon+3)^2 - 4(4\varepsilon-3\gamma)(2\varepsilon-3\gamma-6)}}{2}, \quad a_1 := \frac{z}{2\varepsilon-3\gamma-6},
$$

$$
a_2 := \frac{4\varepsilon-3\gamma}{z}, \quad c_0 := \frac{4\varepsilon-3\gamma}{10\varepsilon-12\gamma-18}, \quad c_1 := \frac{\sqrt{\gamma}(\varepsilon+3)}{10\varepsilon-12\gamma-18}, \quad c_2 := \frac{2\varepsilon-3\gamma-6}{10\varepsilon-12\gamma-18},
$$

where $\operatorname{sgn}(x) := x/|x|$ for $x \neq 0$ and $\operatorname{sgn}(0) := 0$. Then:

Type I (four-parameter beta)

$$
\hat{Y} \equiv a_1 + (a_2 - a_1)\mathcal{B}\left(\frac{c_1+a_1}{c_2(a_2-a_1)} + 1, -\frac{c_1+a_2}{c_2(a_2-a_1)} + 1\right),
$$

where $\mathcal{B}(a,b)$ denotes a random number generator from a beta distribution with parameters a and b (see Devroye, 1986);

Type II (symmetric four-parameter beta)

$$
\hat{Y} \equiv a_1 + 2|a_1|\mathcal{B}\left(\frac{c_1+a_1}{2c_2|a_1|} + 1, \frac{c_1+a_1}{2c_2|a_1|} + 1\right);
$$

Type III (three-parameter gamma)

$$
\hat{Y} \equiv c_1\Gamma\left(\frac{\frac{c_0}{c_1} - c_1}{c_1} + 1, 1\right) + a_1,
$$

where $\Gamma(a,b)$ denotes a random number generator from a gamma distribution with shape parameter a and scale parameter b (see Marsaglia and Tsang, 2000);

Type IV (density proportional to $\exp\left(-\nu \arctan(\frac{x-\lambda}{a})\right) / \left(1+(\frac{x-\lambda}{a})^2\right)^m$)

$$\hat{Y} \equiv \mathcal{P}_{IV}(m, \nu, a, \lambda),$$

where $\mathcal{P}_{IV}(m, \nu, a, \lambda)$ is a random number generator from a Pearson Type IV based on the exponential rejection method for log-concave densities of Devroye (1986, Section 7.2) and adapted to this case by Heinrich (2004), with

$$m := \frac{1}{2c_2}, \quad \nu := \frac{2c_1(1-m)}{\sqrt{4c_0c_2 - c_1^2}}, \quad b := 2(m-1), \quad a := \sqrt{\frac{b^2(b-1)}{b^2+\nu^2}}, \quad \lambda := \frac{a\nu}{b};$$

Type V (inverse gamma location-scale)

$$\hat{Y} \equiv -\frac{c_1 - \frac{c_1}{2c_2}}{c_2\Gamma(\frac{1}{c_2} - 1, 1)} - \frac{c_1}{2c_2};$$

Type VI (F location-scale)

$$\begin{aligned}\hat{Y} \quad \equiv \quad & \left(a_2 + \frac{2(m_2+1)(a_2-a_1)}{-2(m_1+m_2+1)}\mathcal{F}(2(m_2+1), -2(m_1+m_2+1))\right)\mathbf{1}_{\{a_2<0\}} \\ & + \left(a_2 + \frac{2(m_1+1)(a_2-a_1)}{-2(m_1+m_2+1)}\mathcal{F}(2(m_1+1), -2(m_1+m_2+1))\right)\mathbf{1}_{\{a_2\geq 0\}},\end{aligned}$$

where $\mathbf{1}_{\{.\}}$ denotes the indicator of the event $\{\cdot\}$,

$$m_1 := \frac{a_1 + c_1}{c_2(a_2 - a_1)}, \quad m_2 := -\frac{a_2 + c_1}{c_2(a_2 - a_1)}$$

and $\mathcal{F}(a, b)$ is a random number generator from a Snedecor F distribution with numerator and denominator degrees of freedom a and b respectively (see Devroye, 1986); and

Type VII (t location-scale)

$$\hat{Y} \equiv \sqrt{\frac{c_0}{1-c_2}}\mathcal{T}\left(\frac{1}{c_2} - 1\right),$$

where $\mathcal{T}(a)$ is a random number generator from a Student's t distribution with a degrees of freedom (see Devroye, 1986).

EC.5. Sampling from $\left(\int_u^t \boldsymbol{\sigma^2(s)ds}, \int_u^t \frac{\boldsymbol{ds}}{\boldsymbol{\sigma^2(s)}}\right)$ in the 4/2 model using a bivariate Pearson distribution approach

In this section, we focus on the special case of sampling from the couple $\left(\int_u^t \sigma^2(s)ds, \int_u^t \frac{ds}{\sigma^2(s)}\middle| \sigma^2(t)\right) \equiv (\Psi(u,t)|\Phi(u,t))$ in the 4/2 model (see also Section 4.5 of the paper).

Parrish (1987, 1990) propose a conditional nested factorization approach to simulating from a multivariate Pearson distribution; for our purposes, we consider here the bivariate case and the couple $\left(\int_u^t \sigma^2(s)ds, \int_u^t \frac{ds}{\sigma^2(s)}\right)$ in the 4/2 model. To this end, define the cross-moments

$$\mu_{r_1,r_2} = E\left[\left(\int_u^t \sigma^2(s)ds\right)^{r_1}\left(\int_u^t \frac{ds}{\sigma^2(s)}\right)^{r_2}\middle|\sigma^2(t)\right] = \left.\frac{\partial^{r_1+r_2}\mathcal{L}(-a,-b)}{\partial a^{r_1}\partial b^{r_2}}\right|_{a=b=0},$$

where $r_1, r_2 = \{0, 1, 2, 3, 4\}$, $r_1 + r_2 \leq 4$ and

$$\mathcal{L}(a, b) = E\left[e^{-a\int_u^t \sigma^2(s)ds - b\int_u^t \frac{ds}{\sigma^2(s)}} \middle| \sigma^2(t)\right]$$

is given by (EC.7). The simulation is then summarized in the following steps:

1. Simulate $\left(\sigma^2(t) \middle| \sigma^2(u)\right)$ based on $\sigma^2(t) \overset{\text{(law)}}{=} \chi'^2_d(\lambda)\, v^2(1 - e^{-k(t-u)})/4k$, where $\chi'^2_d(\lambda)$ is the noncentral chi-squared random variable with $d = 4\theta k/v^2$ degrees of freedom and noncentrality parameter $\lambda = 4kv^{-2}e^{-k(t-u)}\sigma^2(u)/(1 - e^{-k(t-u)})$
2. Simulate $\left(\int_u^t \frac{ds}{\sigma^2(s)} \middle| \sigma^2(t)\right)$ having first fitted a Pearson curve by moments $\mu_{0,1}, \mu_{0,2}, \mu_{0,3}, \mu_{0,4}$
3. Compute $\mu_{r_1} := E\left[\left(\int_u^t \sigma^2(s)ds\right)^{r_1} \middle| \int_u^t \frac{ds}{\sigma^2(s)}, \sigma^2(t)\right]$ via equations (EC.8)–(EC.9)
4. Simulate $\left(\int_u^t \sigma^2(s)ds \middle| \int_u^t \frac{ds}{\sigma^2(s)}, \sigma^2(t)\right)$ which follows a Pearson distribution law, if $\left(\int_u^t \sigma^2(s)ds, \int_u^t \frac{ds}{\sigma^2(s)} \middle| \sigma^2(t)\right)$ has a bivariate Pearson distribution (see Parrish, 1987 for a proof of this).

Focusing on step 2, we have from Parrish (1990, equations 4–5) that

$$\mu_1 := \frac{b_1^\star - a_0^\star}{1 - 2b_{11}^\star}, \quad \mu_2 := \frac{(2b_1^\star - a_0^\star)\,\mu_1 + b_0^\star}{1 - 3b_{11}^\star}, \tag{EC.8}$$

$$\mu_3 := \frac{(3b_1^\star - a_0^\star)\,\mu_2 + 2b_0^\star\mu_1}{1 - 4b_{11}^\star}, \quad \mu_4 := \frac{(4b_1^\star - a_0^\star)\,\mu_3 + 3b_0^\star\mu_2}{1 - 5b_{11}^\star} \tag{EC.9}$$

for

$$a_0^\star := a_0 + a_1x_1, \quad b_0^\star := -\left(b_0 + b_1x_1 + b_{11}x_1^2\right),$$
$$b_1^\star := -\left(b_2 + b_{12}x_1\right), \quad b_{11}^\star := -b_{22}$$

with x_1 a random sample from $\left(\int_u^t \frac{ds}{\sigma^2(s)} \middle| \sigma^2(t)\right)$ and $a_0, a_1, b_0, b_1, b_2, b_{11}, b_{12}, b_{22}$ satisfying the system of linear equations (see Parrish, 1987, Figure 3)

$$\begin{bmatrix}
\mu_{0,0} & \mu_{1,0} & 0 & 0 & \mu_{0,0} & 0 & \mu_{1,0} & 2\mu_{0,1} \\
\mu_{1,0} & \mu_{2,0} & 0 & 0 & \mu_{1,0} & 0 & \mu_{2,0} & 2\mu_{1,1} \\
\mu_{0,1} & \mu_{1,1} & \mu_{0,0} & \mu_{1,0} & 2\mu_{0,1} & \mu_{2,0} & 2\mu_{1,1} & 3\mu_{0,2} \\
\mu_{2,0} & \mu_{3,0} & 0 & 0 & \mu_{2,0} & 0 & \mu_{3,0} & 2\mu_{2,1} \\
\mu_{1,1} & \mu_{2,1} & \mu_{1,0} & \mu_{2,0} & 2\mu_{1,1} & \mu_{3,0} & 2\mu_{2,1} & 3\mu_{1,2} \\
\mu_{0,2} & \mu_{1,2} & 2\mu_{0,1} & 2\mu_{1,1} & 3\mu_{0,2} & 2\mu_{2,1} & 3\mu_{1,2} & 4\mu_{0,3} \\
\mu_{3,0} & \mu_{4,0} & 0 & 0 & \mu_{3,0} & 0 & \mu_{4,0} & 2\mu_{3,1} \\
\mu_{2,1} & \mu_{3,1} & \mu_{2,0} & \mu_{3,0} & 2\mu_{2,1} & \mu_{4,0} & 2\mu_{3,1} & 3\mu_{2,2} \\
\mu_{1,2} & \mu_{2,2} & 2\mu_{1,1} & 2\mu_{2,1} & 3\mu_{1,2} & 2\mu_{3,1} & 3\mu_{2,2} & 4\mu_{1,3} \\
\mu_{0,3} & \mu_{1,3} & 3\mu_{0,2} & 3\mu_{1,2} & 4\mu_{0,3} & 3\mu_{2,2} & 4\mu_{1,3} & 5\mu_{0,4}
\end{bmatrix}
\begin{bmatrix} a_0 \\ a_1 \\ b_0 \\ b_1 \\ b_2 \\ b_{11} \\ b_{12} \\ b_{22} \end{bmatrix}
= -
\begin{bmatrix} \mu_{0,1} \\ \mu_{1,1} \\ \mu_{0,2} \\ \mu_{2,1} \\ \mu_{1,2} \\ \mu_{0,3} \\ \mu_{3,1} \\ \mu_{2,2} \\ \mu_{1,3} \\ \mu_{0,4} \end{bmatrix}.$$

EC.6 More experimental results

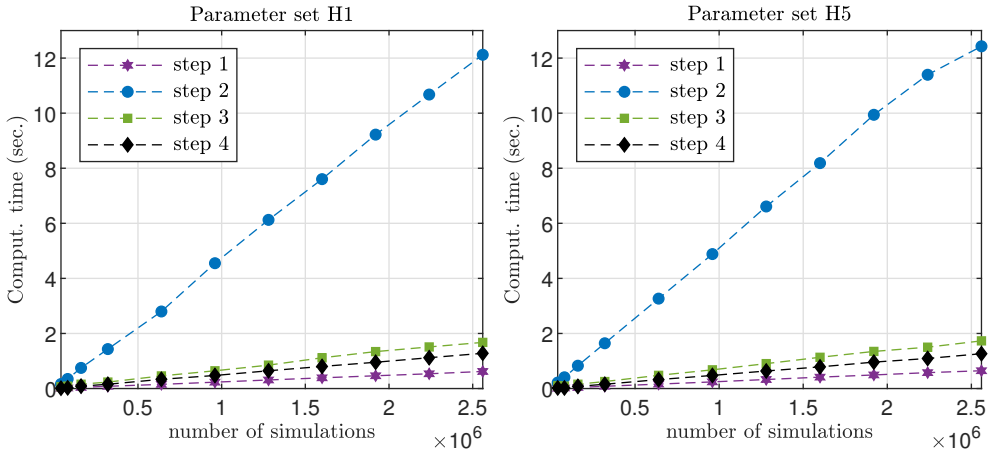


Figure 1: Computing times (in seconds) corresponding to each implementation step of Algorithm 2 as described in Section 7.1 of the paper.

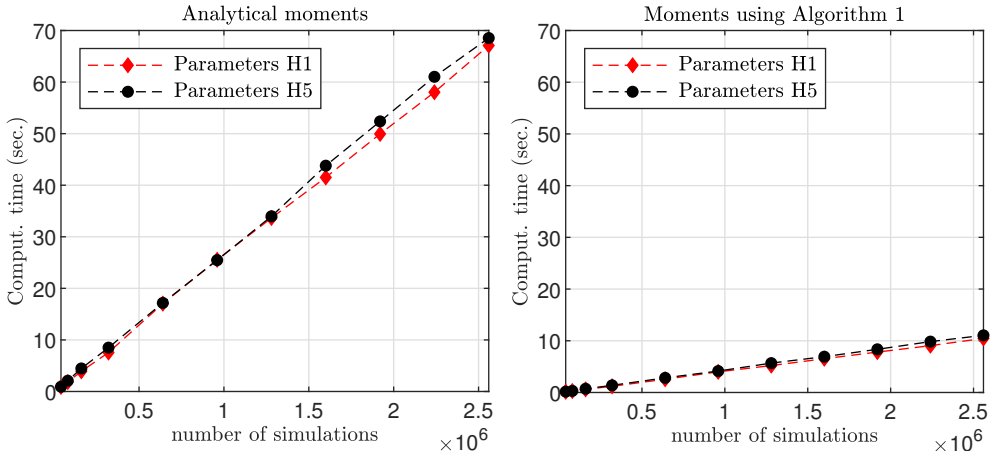


Figure 2: Computing total times (in seconds) for the first four moments of the conditional integrated variance in the Heston model

Table 1: Examples of differences between n-th analytical moments and moments of $\left(\int_0^T \sigma^2(s)ds \middle| \sigma^2(T)\right)$ computed using Algorithm 1 conditional on the 25, 50, 75th percentiles of the terminal variance

	H1			H5		
n	$\sigma^2(T) = 0.003$	$\sigma^2(T) = 0.010$	$\sigma^2(T) = 0.026$	$\sigma^2(T) = 0.001$	$\sigma^2(T) = 0.008$	$\sigma^2(T) = 0.081$
1	3.20e-11	-5.10e-11	2.27e-11	2.16e-11	3.68e-11	3.41e-12
2	2.41e-15	-2.00e-14	-1.30e-14	1.32e-13	-1.40e-13	-2.60e-13
3	6.18e-16	-5.20e-17	-7.30e-16	-2.80e-14	8.77e-14	-1.10e-14
4	-1.10e-17	-2.10e-17	1.54e-17	1.21e-14	3.15e-14	5.41e-15

EC.7. Advancing conditional Monte Carlo valuation

In this section, we present conditional Monte Carlo methods to improve the efficiency of the simulation estimators for the prices of both path-independent options and path-dependent options with payoffs dependent on the maximum (or minimum) of the trajectory of the underlying asset's price, such as barriers and lookbacks. Conditional option prices often have closed or

semi-closed-form representations and contribute to reducing the estimation error by eliminating sources of noise via the conditioning (justified by Jensen's inequality). By further developing a line of work initiated by Willard (1997) for path-independent options in the Heston model, we pave the way to a more general application which comprises different stochastic volatility models but also path-dependent and early-exercisable contracts. Although the principle is general, the application is problem-specific as we show in the following sections.

EC.7.1. European path-independent options

Aiming to improve the efficiency of the simulation estimators under stochastic volatility models, Willard (1997) proposed a conditional Monte Carlo approach. Let $\Pi(T)$ be a generic payoff function which depends on the price of the underlying asset at some terminal time $T > 0$. By conditioning on $\mathcal{F}_T^\sigma$, that is, the σ-algebra generated by $\{\underline{\sigma}(t)\}_{0\le t\le T}$, the derivative's price at time 0 is given by

$$E[e^{-rT}\Pi(T)] = E[e^{-rT}E[\Pi(T)|\mathcal{F}_T^\sigma]]$$

(e.g., see Billingsley, 1995, Theorem 33.3). Thereby, the volatility can be treated as a deterministic quantity and the conditional derivative's price then boils down to that in the simplified model framework.

EC.7.1.1. Conditional Black–Scholes framework

For any of the Heston, SABR ($\beta = 1$), OU-SV, double Heston, 3/2 and 4/2 models, we have shown in Section 4 of the paper that, under $\mathcal{F}_T^\sigma$, the asset price process is governed by a lognormal model. Thereby, the conditional price of a European plain vanilla call option with strike price K and maturity time T has the following Black–Scholes representation

$$E\left[e^{-rT}(S(T)-K)^+\middle|\,\mathcal{F}_T^\sigma\right] = e^{-rT}\left[e^{m_{0,T}+\frac{1}{2}s_{0,T}^2}N\left(\frac{m_{0,T}+s_{0,T}^2-\ln K}{s_{0,T}}\right) - KN\left(\frac{m_{0,T}-\ln K}{s_{0,T}}\right)\right], \tag{EC.10}$$

where $N(\cdot) := (2\pi)^{-1/2}\int_{-\infty}^{\cdot}\exp\left(-z^2/2\right)dz$ and

$$m_{0,T} := E\left[\ln S(T)\middle|\,\mathcal{F}_T^\sigma\right] \text{ and } s_{0,T}^2 := \mathrm{Var}\left[\ln S(T)\middle|\,\mathcal{F}_T^\sigma\right] \tag{EC.11}$$

are given for the different models in Section 4 of the paper.

EC.7.1.2. SABR model ($\beta \neq 1$)

Conditional lognormality does not hold in this case; instead, we avail ourselves of conditioning arguments and results from Cai *et al.* (2017, Theorem 5.1). Assuming that $(S(t), t \ge 0)$ has an absorbing boundary at 0 and $\beta \neq 1$, we get

$$E\left[(S(T)-K)^+\middle|\,\mathcal{F}_T^\sigma\right] \tag{EC.12}$$
$$\begin{cases} = C_0(1)^{-\frac{1}{2(1-\beta)}}A_0^{\frac{1+\gamma_0}{2}}\left(1-Q_{\chi'^2}(C_0(K);3+\gamma_0,A_0)\right) - KQ_{\chi'^2}(A_0;1+\gamma_0,C_0(K)), & \rho = 0 \\ \approx C(1)^{-\frac{1}{2(1-\beta)}}A^{\frac{1+\gamma_\rho}{2}}\Xi\left(-\frac{\rho^2\gamma_\rho}{2},C(K),3+\gamma_\rho,A\right) - KQ_{\chi'^2}(A;1+\gamma_\rho,C(K)), & \rho \neq 0 \end{cases},$$

where $Q_{\chi'^2}(\cdot; d, \lambda)$ is the noncentral chi-squared cumulative distribution function,

$$A_0 := \left(\int_u^t \sigma^2(s)ds\right)^{-1} \left(\frac{S(u)^{1-\beta}}{1-\beta}\right)^2, \quad C_0(y) := \left(\int_u^t \sigma^2(s)ds\right)^{-1} \left(\frac{y^{1-\beta}}{1-\beta}\right)^2,$$
$$A := \frac{\left(\frac{S(u)^{1-\beta}}{1-\beta} + \frac{\rho}{v}(\sigma(t)-\sigma(u))\right)^2}{(1-\rho^2)\int_u^t \sigma^2(s)ds}, \quad C(y) := \frac{\left(\frac{y^{1-\beta}}{1-\beta}\right)^2}{(1-\rho^2)\int_u^t \sigma^2(s)ds}$$

for any $y > 0$,

$$\gamma_\rho := \frac{\beta}{(1-\beta)(1-\rho^2)}, \quad \Xi(p,k,\delta,\alpha) := 2^p \sum_{n=0}^{\infty} e^{-\frac{\alpha}{2}} \left(\frac{\alpha}{2}\right)^n \frac{\bar{\Gamma}(\delta/2+p+n, k/2)}{n!\tilde{\Gamma}(\delta/2+n)},$$

$\tilde{\Gamma}$ is the gamma function, and $\bar{\Gamma}$ the complementary incomplete gamma function.

EC.7.2. Option payoffs dependent on extreme asset values

Next, we show that the original methodology for plain vanilla options applies more generally to exotic payoff structures, which are sensitive to extremal values of the underlying asset, and obtain pricing expressions that can be related to their counterparts in the Black–Scholes model. This is important as it removes a layer of complexity originating from the more sophisticated driving model and the induced curse of dimensionality when considering, for example, alternative transform techniques, and offers a speed advantage and increased precision by reducing the sampling uncertainty. This application highlights the significance of Monte Carlo valuation in covering notably more complex models and payouts.

EC.7.2.1. Conditional Black–Scholes framework

Conditional on the information of the volatility path generated up until time T, the derivative's (forward) price is given by

$$E[\Pi(T)|\mathcal{F}_T^\sigma], \tag{EC.13}$$

where $\Pi(T)$ depends on the maximum or minimum value of the underlying asset price up to T, like an up-and-out barrier call or lookback put (see Section 6 of the paper for the payoffs). As in the path-independent case, for any of the Heston, SABR ($\beta = 1$), OU-SV, double Heston, 3/2 and 4/2 models, the computation of (EC.13) reduces to a more standard Black–Scholes type problem. To this end, define

$$M(T) = \max_{0\le t\le T} \ln \frac{S(t)}{S(0)} \text{ and } \overline{M}(T) = \min_{0\le t\le T} \ln \frac{S(t)}{S(0)}.$$

To compute (EC.13) under $\mathcal{F}_T^\sigma$, we require, depending on the exact payoff specification, quantities (e.g., from De Gennaro Aquino and Bernard, 2019) including the conditional joint probability density function of $\left(\ln\left(S(t)/S(0)\right), M(T)\right)$ and $\left(\ln\left(S(t)/S(0)\right), \overline{M}(T)\right)$ given by

$$f^{\pm}(y,x) = \frac{\pm\sqrt{2}(2x-y)}{\sqrt{\pi}s_{0,T}^3} e^{-\frac{(2x-y)^2 - 2m_{0,T}y + m_{0,T}^2}{2s_{0,T}^2}},$$

where in $\pm$ the top case applies for $M(T)$ and the bottom for $\overline{M}(T)$, and $m_{0,T}$ and $s^2_{0,T}$ are as in (EC.11). In addition, the conditional marginal density of $M(T)$ and $\overline{M}(T)$ is

$$f^{\pm}(x) = \frac{\sqrt{2}}{\sqrt{\pi}s_{0,T}} e^{-\frac{(m_{0,T}-x)^2}{2s^2_{0,T}}} \mp \frac{2m_{0,T}}{s^2_{0,T}} e^{\frac{2m_{0,T}x}{s^2_{0,T}}} N\left(\mp\frac{x+m_{0,T}}{s_{0,T}}\right),$$

where, in $\pm$ and $\mp$, the top case applies for $M(T)$ and the bottom for $\overline{M}(T)$. For example, for an up-and-out barrier call option with a fixed strike K and barrier U we have

$$E[\Pi(T)|\mathcal{F}^{\sigma}_T] = \int_{-\infty}^{\ln\frac{U}{S(0)}} \int_{\ln\frac{K}{S(0)}}^{x} (S(0)e^x - K) f^+(y,x)\, dydx \tag{EC.14}$$

or, for a lookback put with fixed strike K,

$$E[\Pi(T)|\mathcal{F}^{\sigma}_T] = \int_{-\infty}^{\ln\frac{K}{S(0)}} (K - S(0)e^x) f^-(x)\, dx, \tag{EC.15}$$

where $\underline{K} := \min(K, S_0)$.

In the interest of space, we refrain from presenting here analytical solutions to (EC.14) and (EC.15), which can be obtained easily with the aid of a symbolic toolbox such as Mathematica. We also do not present expressions for other contracts, such as down-and-out barrier or floating-strike lookback options, as they follow trivially based on the same arguments, but we can provide these upon request.

EC.7.2.2. SABR ($\beta \neq 1, \rho = 0$) model

As explained earlier, the case of the SABR model requires a more special treatment. By conditioning with respect to the variance path, the SABR model with $\rho = 0$ reduces to a CEV diffusion, therefore conditional Monte Carlo methods for lookback and barrier options are still feasible. For example, for a lookback put option we get by conditional CEV diffusion from Davydov and Linetsky (2001, Lemma 1) that

$$E[\Pi(T)|\mathcal{F}^{\sigma}_T] = \begin{cases} \int_0^K F\left(y; S(0), \int_0^T \sigma^2(s)ds\right) dy, & S(0) \geq K \\ K - S(0) + \int_0^{S(0)} F\left(y; S(0), \int_0^T \sigma^2(s)ds\right) dy, & S(0) < K \end{cases}, \tag{EC.16}$$

where $F(\cdot;\cdot,\cdot)$ satisfies for any $a > 0$

$$\int_0^{\infty} e^{-aw} F(y;x,w)dw = \frac{1}{a}\sqrt{\frac{x}{y}} \frac{K_v\left(\frac{\sqrt{2a}x^{1-\beta}}{|\beta-1|}\right)}{K_v\left(\frac{\sqrt{2a}y^{1-\beta}}{|\beta-1|}\right)}, \quad 0 < y \leq x \tag{EC.17}$$

and $K_v(\cdot)$ is the modified Bessel function of the second kind with $v := 1/(2|\beta-1|)$ degrees of freedom. Relevant results for knock-out barrier options are available in Davydov and Linetsky (2001, Proposition 3). It is worth noting that (EC.16) depends just on the *unconditional* $\int_0^T \sigma^2(s)ds$, which can still be simulated using the method prescribed in this paper. Finally, the integrals in (EC.16) can be computed by quadrature and $F(\cdot;\cdot,\cdot)$ using, for example, the Euler inversion algorithm of Abate and Whitt (1995).

We conclude this part by applying the mixing Monte Carlo moment-based approach with the semi-analytical formula (EC.16) to price a fixed-strike continuously monitored lookback put option in the SABR model. This way we also bypass the simulation of the intermediate values of the process. We compare with the method of Chen *et al.* (2012) implemented over a refined time grid. From Table 2, the prices from the two methods are comparable but the variance reduction brought in by the mixing approach is substantial. Despite the seemingly high computing time, it is worth noting that our mixing approach is considerably faster than a continuous-monitoring approximation based on a very dense time grid for an especially power-demanding simulation such as that of the SABR model (see Section 4.2 of the paper). In addition, the path-dependent SABR case is the most intensive one (see equation EC.17), whilst any other model case in Section EC.7.2 is significantly simpler and faster.

Table 2: Lookback option prices (with standard errors, s.e.) in the SABR model

Method	price	s.e.	time	price	s.e.	time	price	s.e.	time
		SABR1			SABR2			SABR4	
Chen et al.	0.0446	1.22e-05	490.72	0.04685	9.70e-06	1605.72	0.3542	1.59e-04	2421.91
Hybrid	0.0448	2.10e-06	345.31	0.04688	2.00e-06	347.17	0.3543	6.56e-05	296.11

Notes. Hybrid analytical (formula EC.16)-conditional Monte Carlo (based on fitted Pearson to the integrated variance). Benchmark: Chen *et al.* (2012) with 10^4 time steps per year. All computing times are in seconds and correspond to $\mathcal{M} = 10^6$ simulations.

EC.8. The CGMY model

As explained, for example, in Ballotta and Kyriakou (2014), similarly to other popular Lévy models, like the variance gamma or normal inverse Gaussian, the Carr–Geman–Madan–Yor (CGMY) model can be represented as a subordinated arithmetic Brownian motion; unlike those though, the subordinator is only known via its Laplace transform.

More specifically, the CGMY process is given by

$$X(t) = \theta Z(t) + W(Z(t)), \tag{EC.18}$$

where W is a standard Brownian motion, $\theta := (G - M)/2$ and Z is a subordinator independent of W with Laplace transform

$$E\left[e^{-aZ(t)}\right] = \exp\left((t-u)\, C\Gamma(-Y)\left(2\left(2a + GM\right)^{Y/2}\cos\left(\xi\left(a;G,M\right)Y\right) - M^Y - G^Y\right)\right), \quad u < t, \tag{EC.19}$$

where $\xi(a;G,M) := \arctan\left(\tilde{\theta}^{-1}\sqrt{2a-\theta^2}\right)$ and $\tilde{\theta} := (G+M)/2$.

In view of the above, our moment-based random number generator can be adapted to the efficient simulation of the CGMY model trajectories as follows:

1. Compute the moments of $Z(t)$ (via, for example, Choudhury and Lucantoni, 1996) based on (EC.19)
2. Simulate $Z(t)$ using a fitted a Pearson curve to the corresponding theoretical distribution
3. Simulate W subject to time-change $Z(t)$ and, consequently, $X(t)$ according to (EC.18).

In addition, in financial modelling, the related asset price process under the risk neutral measure is

$$S(t) = S(u) \exp\left((r+\omega)(t-u) + X(t)\right),$$

where $\omega := -C\Gamma(-Y)\left((G+1)^Y - G^Y + (M-1)^Y - M^Y\right)$. Conditionally on $Z(t)$, we also have that

$$(\ln S(t) | \ln S(u), Z(t)) \sim \mathcal{N}(m, s^2),$$

where $m := \ln S(u) + (r+\omega)(t-u) + ((G-M)/2)Z(t)$ and $s^2 := Z(t)$. By analogy to the exposition in Section EC.7.1.1, we get, based on the relevant conditioning argument on this occasion, that

$$E\left[e^{-r(t-u)}(S(t)-K)^+ \middle| Z(t)\right] = e^{-r(t-u)}\left[e^{m+\frac{1}{2}s^2} N\left(\frac{m+s^2-\ln K}{s}\right) - KN\left(\frac{m-\ln K}{s}\right)\right]. \tag{EC.20}$$

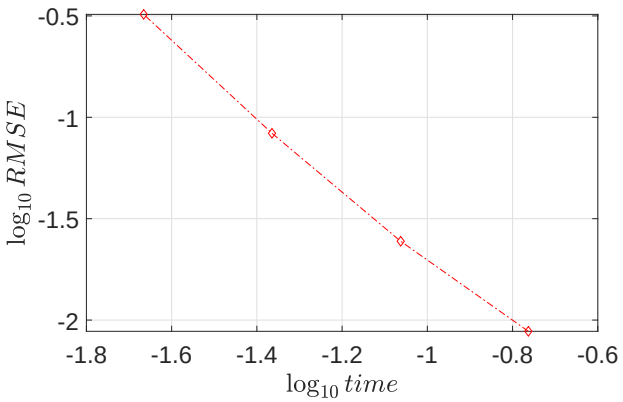
Some numerical results are reported in Table 3. The simulation of the CGMY process is remarkably fast as its increments are independent and we no longer face a conditional Laplace transform of Z, which also does not involve special functions that can severely slow down the overall execution, simplifying considerably the simulation task. Similarly to the other models in the paper, our scheme clearly converges also in this case as shown in Figure 3.

Table 3: European plain vanilla call option prices (with standard errors, s.e.) in the CGMY model

True option price	20.1965	Bias (s.e.)	0.0005	(0.001)
$\mathcal{M} \times 10^4$	4	16	64	256
RMSE	0.1726	0.0866	0.0432	0.0216
Time	0.01	0.02	0.08	0.32

Notes. True (reference) prices via the method of Fang and Oosterlee (2008) based on parameters $S(0) = K = 100$, $C = 0.9795$, $G = 3.512$, $M = 10.96$, $Y = 0.8$, $r = 4\%$, $T = 1$ (Černý and Kyriakou, 2011, Table 2). Hybrid analytical (formula EC.20)-conditional Monte Carlo (based on fitted Pearson to $Z(T)$) based on $\mathcal{M}$ simulation trials for use in RMSE and bias computed as in Section 7.2 of the paper. All computing times are in seconds.

Figure 3: Convergence of our hybrid-conditional Monte Carlo method in the CGMY model: the case of European plain vanilla call option



EC.9. Mean-reverting normal model with Heston stochastic volatility

Evidence of mean-reversion is abundant in various markets including commodities, currencies, energy and temperature. In particular, a mean-reverting model for commodity (such as gas and electricity) log-spot prices with stochastic volatility is postulated by Geman (2005) but also more recent researches such as Kyriakou *et al.* (2016):

$$dS(t) = \alpha S(t)(\beta - \ln S(t))dt + \sigma(t)S(t)\left(\rho dW_2(t) + \sqrt{1-\rho^2}dW_1(t)\right), \quad \text{(EC.21)}$$

$$d\sigma^2(t) = k(\theta - \sigma^2(t))dt + v\sigma(t)dW_2(t), \quad \text{(EC.22)}$$

where α and k are the speeds at which each of processes S and σ revert to their equilibrium levels. The correlation ρ allows for possible inverse leverage effect, i.e., high prices associated with high volatility translating to $\rho > 0$. Substituting (EC.22) in (EC.21) yields on the log-scale

$$d\ln S(t) = \alpha\left(\beta - \frac{\sigma^2(t)}{2\alpha} - \ln S(t)\right)dt + \frac{\rho}{v}d\sigma^2(t) - \frac{\rho k}{v}(\theta - \sigma^2(t))dt + \sigma(t)\sqrt{1-\rho^2}dW_1(t), \quad \text{(EC.23)}$$

from which

$$\begin{aligned}\ln S(t) &= \ln S(u) + \left(\alpha\beta - \frac{\rho k\theta}{v}\right)(t-u) + \left(\frac{\rho k}{v} - \frac{1}{2}\right)\int_u^t \sigma^2(s)ds \\ &\quad + \frac{\rho}{v}\left(\sigma^2(t) - \sigma^2(u)\right) - \alpha\int_u^t \ln S(s)ds + \sqrt{1-\rho^2}\int_u^t \sigma(s)dW_1(s).\end{aligned}$$

In its form (EC.23), for $\ln S \equiv r$ the short-term interest rate, the model reflects the well-known Fong and Vasicek (1992) two-factor model of the term structure, which explicitly recognizes the interest rate volatility as a stochastic factor.

The main challenge here is the loss of analytical tractability due to the mean-reversion in the $\ln S$ dynamics which leads to the extra $\ln S$ time integral in the expression above. In this case, we find that a combinatory solution of our Pearson fit for the integrated variance and a central discretization of the $\ln S$ time integral along the lines of Andersen (2008) is so far the best compromise between accuracy, complexity and computational speed. The proposed method also improves upon the Monte Carlo valuation based on SDE discretization of Clewlow and Strickland (1997) for interest rate derivatives.

EC.10. Simple corrections for early exercise and discrete path-dependence

EC.10.1. Early-exercisable contracts

In this paper, we also consider a simple, yet efficient, way for pricing American-style contracts based on standard Monte Carlo valuation, bypassing the use of a bridge for sampling in reverse time or least-squares Monte Carlo. By suitable conditioning arguments based on the generated path of the volatility process, we wind up in more simplified frameworks, such as a conditional lognormal, which endow us with easily computable, even if approximate, solutions for American options, such as the basic Barone-Adesi and Whaley (1987). In Table 4, we present some results from this application. Better solutions in a model setting with conditionally deterministic parameters can be potentially obtained based on transform techniques (e.g., see Fang

and Oosterlee, 2009), but combined with Monte Carlo valuation they can be quite prolonged and sensitive to grid parameters. An approximate solution is likely to lead to a biased price estimate, where any bias is due to the basal American option price solution, still it offers an adequate runtime–accuracy balance to be an effective counterforce to a two-dimensional transform technique.

Table 4: American and European plain vanilla option prices in the Heston model

$S(0)$	8	9	10	11	12
European (hybrid method)	1.8388	1.0483	0.5015	0.2080	0.0804
American (hybrid method)	1.9999	1.1096	0.5200	0.2149	0.0834
American ref. value	2.0000	1.1076	0.5200	0.2137	0.0820
Error	1.19e-04	2.02e-03	1.02e-05	1.22e-03	1.34e-03

Notes. Hybrid analytical approach (formula EC.10 for Europeans; Barone-Adesi and Whaley, 1987 for Americans)-conditional Monte Carlo (based on fitted Pearson distribution to the integrated variance). American reference values: Fang and Oosterlee (2011) based on parameters $S(0) = \{8, 9, 10, 11, 12\}, K = 10, T = 0.25, r = 0.1, k = 5, \theta = 0.9, v = 0.16, \sigma^2(0) = 0.0625, \rho = 0.1$.

EC.10.2. Connecting discrete and continuous path-dependence

As discussed in Section EC.7, apart from simplifying the valuation process by reducing to a more basic underlying model, conditional Monte Carlo valuation brings in considerable variance reduction. Therefore, having in hand an accurate price estimate for a continuous path-dependent option, such as a lookback or a barrier, as we studied in Section EC.7, enables us to obtain the discrete counterpart using relevant corrections as in Broadie *et al.* (1999). This is very convenient as the accuracy of the continuous price estimate is transferred to the discrete counterpart, subject to the error of the correction, and eventually gives us access to option prices under more sophisticated stochastic volatility models which otherwise would be hard, if not impossible, to compute.

EC.11. Greeks

Unbiased estimation of option Greeks using Monte Carlo simulation is not as straightforward as the computation of option prices due to potential discontinuities in the option payoff function, such as, for example, in the case of barrier and digital options. In general, the standard two methods for the sensitivities are the pathwise (PW) and likelihood ratio (LR) method introduced by Broadie and Glasserman (1996). Both rely on an interchange of integration and differentiation under certain regularity conditions which are easier to justify for the LR method, as density functions are usually smooth functions of their parameters whereas payoff functions are not. Conditioning arguments simplify the problem as they offer us direct access to standard densities. For more details, refer to Broadie and Kaya (2004, 2006).

References

Abate, J. and Whitt, W. (1995) Numerical inversion of Laplace transforms of probability distributions. *ORSA Journal on Computing*, **7**, 36–43.

Andersen, L. (2008) Simple and efficient simulation of the Heston stochastic volatility model. *Journal of Computational Finance*, **11**, 1–42.

Ballotta, L. and Kyriakou, I. (2014) Monte Carlo simulation of the CGMY process and option pricing. *Journal of Futures Markets*, **34**, 1095–1121.

Barone-Adesi, G. and Whaley, R. E. (1987) Efficient analytic approximation of American option values. *The Journal of Finance*, **42**, 301–320.

Billingsley, P. (1995) *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics. New York: Wiley, 3 edn.

Broadie, M. and Glasserman, P. (1996) Estimating security price derivatives using simulation. *Management Science*, **42**, 269–285.

Broadie, M., Glasserman, P. and Kou, S. G. (1999) Connecting discrete and continuous path-dependent options. *Finance and Stochastics*, **3**, 55–82.

Broadie, M. and Kaya, O. (2004) Exact simulation of option Greeks under stochastic volatility and jump diffusion models. In *Proceedings of the 2004 Winter Simulation Conference* (eds. R. G. Ingalls, M. D. Rossetti, J. S. Smith and B. A. Peters), vol. 2, 1607–1615.

Broadie, M. and Kaya, O. (2006) Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, **54**, 217–231.

Cai, N., Song, Y. and Chen, N. (2017) Exact simulation of the SABR model. *Operations Research*, **65**, 931–951.

Černý, A. and Kyriakou, I. (2011) An improved convolution algorithm for discretely sampled Asian options. *Quantitative Finance*, **11**, 381–389.

Chen, B., Oosterlee, C. W. and van der Weide, H. (2012) A low-bias simulation scheme for the SABR stochastic volatility model. *International Journal of Theoretical and Applied Finance*, **15**, 125–161.

Choudhury, G. L. and Lucantoni, D. M. (1996) Numerical computation of the moments of a probability distribution from its transform. *Operations Research*, **44**, 368–381.

Clewlow, L. and Strickland, C. (1997) Monte Carlo valuation of interest rate derivatives under stochastic volatility. *The Journal of Fixed Income*, **7**, 35–45.

Davydov, D. and Linetsky, V. (2001) Pricing and hedging path-dependent options under the CEV process. *Management Science*, **47**, 949–965.

De Gennaro Aquino, L. and Bernard, C. (2019) Semi-analytical prices for lookback and barrier options under the Heston model. *Decisions in Economics and Finance*, **42**, 715–741.

Devroye, L. (1986) *Non-Uniform Random Variate Generation*. New York: Springer.

Esseen, C.-G. (1945) Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. *Acta Mathematica*, **77**, 1–125.

Fang, F. and Oosterlee, C. W. (2008) A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, **31**, 826–848.

Fang, F. and Oosterlee, C. W. (2009) Pricing early-exercise and discrete barrier options by fourier-cosine series expansions. *Numerische Mathematik*, **114**, 27–62.

Fang, F. and Oosterlee, C. W. (2011) A Fourier-based valuation method for Bermudan and barrier options under Heston's model. *SIAM Journal on Financial Mathematics*, **2**, 439–463.

Fong, H. G. and Vasicek, O. A. (1992) Interest rate volatility as a stochastic factor. Tech. rep., Gifford Fong Associates.

Geman, H. (2005) *Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy*. New York: Wiley.

Grasselli, M. (2017) The 4/2 stochastic volatility model: A unified approach for the Heston and the 3/2 model. *Mathematical Finance*, **27**, 1013–1034.

Heinrich, J. (2004) A guide to the Pearson Type IV distribution. Available online at https://www-cdf.fnal.gov/physics/statistics/notes/cdf6820_pearson4.pdf.

Klebanov, L. B. and Mkrtchyan, S. T. (1986) Estimation of the closeness of distributions in terms of identical moments. *Journal of Soviet Mathematics*, **32**, 54–60.

Kyriakou, I., Nomikos, N. K., Papapostolou, N. C. and Pouliasis, P. K. (2016) Affine-structure models and the pricing of energy commodity derivatives. *European Financial Management*, **22**, 853–881.

Li, C. and Wu, L. (2019) Exact simulation of the Ornstein–Uhlenbeck driven stochastic volatility model. *European Journal of Operational Research*, **275**, 768–779.

Marsaglia, G. and Tsang, W. W. (2000) A simple method for generating gamma variables. *ACM Transactions on Mathematical Software*, **26**, 363–372.

Parrish, R. S. (1987) Evaluation and approximation of multivariate cumulative joint probabilities. *Journal of Statistical Computation and Simulation*, **27**, 1–33.

Parrish, R. S. (1990) Generating random deviates from multivariate Pearson distributions. *Computational Statistics & Data Analysis*, **9**, 283–295.

Rachev, S. T., Klebanov, L., Stoyanov, S. V. and Fabozzi, F. (2013) *The Methods of Distances in the Theory of Probability and Statistics*. New York: Springer-Verlag.

Willard, G. A. (1997) Calculating prices and sensitivities for path-independent derivative securities in multifactor models. *Journal of Derivatives*, **5**, 54–61.