International Journal of Theoretical and Applied Finance
Vol. 10, No. 1 (2007) 51–88

World Scientific
www.worldscientific.com

PRICING PATH-DEPENDENT OPTIONS ON STATE DEPENDENT VOLATILITY MODELS WITH A BESSEL BRIDGE

GIUSEPPE CAMPOLIETI* and ROMAN MAKAROV†
Department of Mathematics, Wilfrid Laurier University
75 University Avenue West, Waterloo, Ontario N2L 3C5, Canada
**gcampoli@wlu.ca*
†*rmakarov@wlu.ca*

Received 18 November 2004
Accepted 27 April 2006

This paper develops bridge sampling path integral algorithms for pricing path-dependent options under a new class of nonlinear state dependent volatility models. Path-dependent option pricing is considered within a new (dual) Bessel family of semimartingale diffusion models, as well as the constant elasticity of variance (CEV) diffusion model, arising as a particular case of these models. The transition p.d.f.s or pricing kernels are mapped onto an underlying simpler squared Bessel process and are expressed analytically in terms of modified Bessel functions. We establish precise links between pricing kernels of such models and the randomized gamma distributions, and thereby demonstrate how a squared Bessel bridge process can be used for exact sampling of the Bessel family of paths. A Bessel bridge algorithm is presented which is based on explicit conditional distributions for the Bessel family of volatility models and is similar in spirit to the Brownian bridge algorithm. A special rearrangement and splitting of the path integral variables allows us to combine the Bessel bridge sampling algorithm with either adaptive Monte Carlo algorithms, or quasi-Monte Carlo techniques for significant numerical efficiency improvement. The algorithms are illustrated by pricing Asian-style and look-back options under the Bessel family of volatility models as well as the CEV diffusion model.

Keywords: Option pricing; hypergeometric, Bessel and CEV diffusion processes; Monte Carlo methods; variance reduction; bridge sampling algorithms; path integration.

1. Introduction

In most works on path-dependent options, the underlying asset price is assumed to follow geometric Brownian motion with constant local volatility, i.e., with linear volatility function. However, the lognormal assumption of asset price returns disagrees with most of the empirical evidence. To improve upon empirical deficiencies of the lognormal model, several alternative models have been proposed in the literature. In most recent years, some new classes of analytically exact transition

probability density functions (p.d.f.s) for nonlinear state dependent volatility models have been derived; for example see [2, 3, 13, 14]. Such new diffusion models provide a greater accuracy and flexibility in calibration, i.e., in fitting computed prices to observed market option prices that in reality exhibit pronounced volatility skews or smiles. This has been demonstrated in part by calibrating theoretical prices obtained under these models for standard European calls and puts to corresponding market prices across a range of constant-maturity strikes. An attractive feature of these state dependent volatility models is that they give a basis for modelling asset prices with various forms of nonlinear local volatility structure while retaining computational tractability.

For much of path-dependent options, there do not exist exact analytical pricing formulas except for a very few exceptions that have appeared, for example, in some works on pricing standard barrier and lookback options under the originally proposed and commonly known *constant elasticity of variance* (CEV) model [12]. Closed-form pricing formulas for such path-dependent options were derived by Davydov and Linetsky [13]. Within a restricted subset of the same CEV model, some related numerical results on pricing barrier and lookback options based on a trinomial lattice approach were presented by Boyle and Tian [5]. However, for more general classes of state dependent volatility models and for other more general Asian style options, exact pricing formulas are not known and, moreover, accurate and efficient numerical pricing algorithms are rather scarce. Thus, the subject of our investigation is to present a useful numerical approach for valuing path-dependent options under new classes of state dependent diffusion models.

In this paper we consider the problem of pricing path-dependent options within certain new subfamilies of a recently developed larger class of multi-parameter volatility models that includes the CEV models as special cases. Generally, these diffusion models give rise to semimartingales, i.e., various subsets include martingale, supermartingale and submartingale processes. The processes are characterized by exact analytical transition p.d.f.s and are referred to as so-called hypergeometric diffusions as their p.d.f.s involve hypergeometric functions. A particular family of such models that we implement in this paper is referred to as the *Bessel family of models*. These models have transition p.d.f.s (or pricing kernels) for the asset price dynamics that reduce to those for the squared Bessel process. Closed-form analytical expressions for the transition p.d.f.s are thereby given in terms of modified Bessel functions [2, 3]. More specifically, we consider the implementation of two new subfamilies of three-parameter volatility models, referred to as the Bessel I-subfamily and Bessel K-subfamily. These two (dual) models have distinctive probabilistic properties that are here classified as strict positive supermartingale (local martingale) and martingale processes, respectively. We show how these Bessel models are closely connected to, but yet form a superset of, the corresponding CEV models, the latter being analytically recovered as limiting cases.

Specializing to discrete-time path-dependent options, we reduce Asian option pricing to an evaluation of a path integral. The stochastic asset path sampling

is done here in a very precise manner since we have exact pricing kernels that map onto the simpler underlying squared Bessel processes. Since we have the problem of evaluating a multi-dimensional integral, we find an efficient solution among adaptive Monte Carlo and quasi-Monte Carlo methods. For numerical integration via randomized quasi-Monte Carlo techniques, there are recent works on the subject of structuring the sampling algorithm so as to concentrate the variance of the integrand to just a few coordinates [9, 19, 22]. For geometric Brownian motion, it is possible to realize this by means of a Brownian bridge sampling algorithm. Recently this approach was extended into the case of the variance gamma model [6, 30]. However, for most nonlinear state dependent volatility models analogous bridge sampling algorithms remain largely unknown to date. In this work we present efficient bridge sampling algorithms for nonlinear state dependent processes. We consider the Bessel family of models and the CEV diffusion model for which sampling of asset prices can be done precisely due to the established connection between the transition p.d.f.s and the randomized gamma distributions. Our computational approach is based on two principles: the use of a Bessel bridge process and a splitting of the path integral into two subspaces. In doing so, the variance of the integrand is concentrated to the first subspace with the smallest dimensionality.

The paper is organized as follows. Section 2 begins with a brief introduction of the Bessel family of volatility models, where derivations of the relevant subfamilies of exact analytical pricing kernels are given and the salient probabilistic properties of the processes are also presented. The mathematical connection between these asset pricing processes, their corresponding CEV counterparts, and the relevant underlying (simpler) squared Bessel process is then also established. In Sec. 3 we show how the Bessel family of kernels is directly connected with the randomized gamma distributions. We describe how these distributions are used to efficiently simulate the squared Bessel process and the standard squared Bessel bridge process. For simulation of the Bessel distribution, we use Devroye's rejection method [16]. Section 4 begins by presenting a discretized path integral framework as well as a general bridge sampling approach for valuing any path-dependent European style option. Sections 4.2 and 4.3 contain exact bridge sampling algorithms, via the randomized gamma distributions, for all models considered within the Bessel family and CEV processes, i.e., for all pricing models that map to a squared Bessel process. In Sec. 4.4 we present both sequential and bridge sampling algorithms for simulating asset price paths under all such models. These algorithms are used within our path integral approach for pricing path-dependent options. Finally, we implement a special splitting (variance reduction) technique for efficiently evaluating path integrals by decomposing the set of integration variables. In Sec. 5, we discuss adaptive Monte Carlo algorithms and randomized quasi-Monte Carlo methods as applied to option pricing by evaluation of path integrals. Section 6 presents numerical results for pricing average Asian style options and lookback options within the Bessel family, including the CEV diffusion model.

2. State Dependent Volatility Models Based on Bessel Process

Of special interest in asset pricing theory are state dependent diffusion models with analytically known transition p.d.f. $U(S, S_0, t)$. The conditional probability of an asset having value (at time t) within an interval D is then

$$\mathsf{P}\{S_t \in D | S_{t=0} = S_0\} = \int_D U(S, S_0, t) dS\,,$$

where $(S_t)_{t\geq 0}$ is assumed to be a positive asset price process. Within such a time-homogeneous model, a standard European-style call option, for example, written on the price S_0 at time zero, struck at K, and maturing in time T, can be simply and exactly priced by evaluating a definite integral:

$$\begin{aligned} C(S_0, K; T) &= e^{-rT}\mathsf{E}^Q[(S_T - K)_+ | S_{t=0} = S_0] \\ &= e^{-rT}\int_0^\infty U(S, S_0, T)(S - K)_+ dS. \end{aligned} \tag{2.1}$$

The transition p.d.f. U in (2.1) is a so-called pricing kernel for an assumed pricing measure $Q = Q(g)$. In this case the money market account $g_t = e^{rt}$ is the numeraire asset; i.e., with expectation $\mathsf{E}^Q[S_T | S_t] = e^{r(T-t)} S_t$, $0 \leq t \leq T$, where r is an assumed constant interest rate. Throughout we denote $(x)_+ = \max(x, 0)$.

In this paper, we are interested in pricing options with path dependent payoffs under a class of diffusion models with nonlinear volatility functions. We shall focus specifically on state dependent models that arise from an underlying λ-dimensional squared Bessel process $(x_t)_{t\geq 0} \in (0, \infty)$ with

$$dx_t = \lambda dt + 2\sqrt{x_t} dW_t, \tag{2.2}$$

where dW_t is a standard Brownian increment. In particular, we consider the CEV model. This commonly known model is related to the process (2.2). As discussed below, by applying appropriate transformations, transition p.d.f.s for the CEV models arise from those for the process in (2.2). Moreover, we also introduce another related, yet new, and more general family of models referred to as the Bessel family of models. Transition p.d.f.s for these models are constructed by means of so-called diffusion canonical transformations — combinations of nonlinear variable transformations and measure changes. In particular, two (dual) Bessel subfamilies of asset price processes with contrasting probabilistic properties are derived and discussed.

2.1. *The Bessel family of volatility models*

Let $(F_t)_{t\geq 0}$ be a diffusion process that formally obeys a stochastic differential equation (SDE) of the form:

$$dF_t = \sigma(F_t) dW_t^Q, \quad F_{t=0} = F_0, \tag{2.3}$$

where dW_t^Q is a standard Brownian increment within an appropriate probability measure Q. In financial terms, F_t may denote the discounted price at calendar

time t of some asset, such as the forward price of a stock under a risk-neutral or forward measure. The state dependent volatility function $\sigma(F)$ is assumed to be generally nonlinear. More precisely, we model asset prices as time-homogeneous Markovian diffusions described by transition p.d.f.s $U = U(F, F_0, t)$ that solve the corresponding Fokker–Planck (forward Kolmogorov) PDE:

$$\frac{\partial U}{\partial t} = \frac{1}{2}\frac{\partial^2}{\partial F^2}\left(\sigma^2(F)\,U\right). \tag{2.4}$$

A solution satisfying the Dirac delta initial condition, $U(F, F_0, 0^+) = \delta(F - F_0)$, and imposed homogeneous boundary conditions at the endpoints of the domain of the process is a transition p.d.f. or solution kernel for the F_t-process. We note that in general situations (2.4) may have multiple solutions.

Exact analytical families of transition p.d.f.s for a class of such F_t-processes can be generated by considering some underlying x_t-process, and in some generally different probability measure, where

$$dx_t = \lambda(x_t)dt + \nu(x_t)dW_t, \tag{2.5}$$

for which the corresponding Kolmogorov PDE, with drift $\lambda(x)$ and volatility $\nu(x)$, can readily be solved analytically. Let $u = u(x, x_0, t)$ be such a transition p.d.f. (or x-space solution kernel). As is shown partly in [2] and in much more detail in [3], by considering a (diffusion canonical) invertible transformation $x = X(F)$ such that

$$\left|\frac{dX}{dF}\right| = \frac{\nu(x)}{\sigma(F)},$$

with volatility function given by

$$\sigma(F) = \frac{\sigma_0 \nu(x) \exp\left(-2\int^x \frac{\lambda(s)}{\nu^2(s)}ds\right)}{[\hat{u}(x,\rho)]^2}, \tag{2.6}$$

then a transition p.d.f. (F-space kernel) U for the above F_t-process is given in terms of an x-space kernel as follows:

$$U(F, F_0, t) = \frac{\nu(X(F))}{\sigma(F)}\frac{\hat{u}\left(X(F), \rho\right)}{\hat{u}\left(X(F_0), \rho\right)}e^{-\rho t}u\left(X(F), X(F_0), t\right). \tag{2.7}$$

The so-called generating function $\hat{u}$ satisfies the second order ODE:

$$\frac{1}{2}\nu^2(x)\frac{d^2}{dx^2}\hat{u}(x,\rho) + \lambda(x)\frac{d}{dx}\hat{u}(x,\rho) - \rho\hat{u}(x,\rho) = 0, \tag{2.8}$$

where the parameter ρ is an arbitrary non-negative real constant, and $\sigma_0 \neq 0$ is an arbitrary real constant. The map $F = F(x)$ is given by

$$F(x) = \bar{F} \pm \sigma_0 \int_{\bar{x}}^{x} \frac{e^{-2\int \frac{\lambda(z)}{\nu^2(z)}dz}}{[\hat{u}(z,\rho)]^2}dz, \tag{2.9}$$

with arbitrary real constant $\bar{F}$, $\bar{x} = X(\bar{F})$. The inverse $x = X(F)$ is given by inverting this relation using either branch ($\pm$ branch for monotonically increasing/decreasing map).

In this paper, we specifically consider models arising from (2.2) as underlying x_t-process where $\lambda(x) = \lambda$ (constant) and $\nu(x) = 2\sqrt{x}$. In this particular subsection we consider $\lambda > 2$. The exact kernel that satisfies homogeneous zero boundary conditions at the endpoints $x = 0$ and ∞ is:

$$u(x, x_0, t) = \frac{1}{2}\left(\frac{x}{x_0}\right)^{\frac{\mu}{2}} \frac{e^{-(x+x_0)/(2t)}}{t} I_\mu(\sqrt{xx_0}/t), \quad t > 0, \quad x, x_0 > 0\,, \tag{2.10}$$

where $\mu \equiv \frac{\lambda}{2} - 1 > 0$. This density satisfies probability conservation: $\int_0^\infty u(x, x_0, t)dx = 1$. The general form for $\hat{u}$ is obtained as the general solution of (2.8),

$$\hat{u}(x, \rho) = x^{-\mu/2}\left[q_1 I_\mu(\sqrt{2\rho x}) + q_2 K_\mu(\sqrt{2\rho x})\right],$$

where $I_\mu(z)$ and $K_\mu(z)$ are the modified Bessel functions (of order μ) of the first and second kind, respectively (see [1]). For arbitrary constants q_1, q_2 such that $\hat{u} > 0$, the volatility function takes the general form

$$\sigma(F) = \frac{2\sigma_0}{\sqrt{X(F)}\left[q_1 I_\mu\left(\sqrt{2\rho X(F)}\right) + q_2 K_\mu\left(\sqrt{2\rho X(F)}\right)\right]^2}. \tag{2.11}$$

By inserting $\hat{u}$ into (2.9) and using certain Bessel integral and Wronskian properties, it is readily shown [3] that one generally obtains a dual family of maps (and for our purposes we simply set $\bar{F} = 0$):

$$F(x) = \begin{cases} \dfrac{2\sigma_0/(q_1 q_2)}{1 + (q_1/q_2) I_\mu\left(\sqrt{2\rho x}\right)/K_\mu\left(\sqrt{2\rho x}\right)} & (q_1 \neq 0), \\ \dfrac{2\sigma_0/(q_1 q_2)}{1 + (q_2/q_1) K_\mu\left(\sqrt{2\rho x}\right)/I_\mu\left(\sqrt{2\rho x}\right)} & (q_2 \neq 0). \end{cases} \tag{2.12}$$

From the small and large argument asymptotics of $I_\mu(z)$ and $K_\mu(z)$ it follows that $x \in (0, \infty)$ is mapped onto $F \in (0, 2\sigma_0/(q_1 q_2))$, where the first mapping in (2.12) is strictly decreasing and the second is strictly increasing. For nonzero q_1, q_2, observe that F is contained in a finite interval which is not useful for our present purposes as we seek asset price processes that lie on the positive half-line as follows.

By considering the above two maps wherein one of the parameters q_1 or q_2 tends to zero (for fixed nonzero the other), the above Bessel family gives rise to two separate 3-parameter subfamilies:

$$F(x) = \begin{cases} a_I \dfrac{K_\mu\left(\sqrt{2\rho x}\right)}{I_\mu\left(\sqrt{2\rho x}\right)}, & \text{Bessel } I\text{-subfamily as } q_2 \to 0, \\ a_K \dfrac{I_\mu\left(\sqrt{2\rho x}\right)}{K_\mu\left(\sqrt{2\rho x}\right)}, & \text{Bessel } K\text{-subfamily as } q_1 \to 0, \end{cases} \tag{2.13}$$

where $a_I = 2\sigma_0/q_1^2$, $a_K = 2\sigma_0/q_2^2$ and ρ, μ are independently adjustable positive parameters. The two functions $F(x)$ (and the respective inverses $x = X(F)$) in

(2.13) now map $x \in (0,\infty)$ and $F \in (0,\infty)$ into one another, where $F(x)$ is monotonic with $dF(x)/dx = \pm\sigma(F(x))/(2\sqrt{x})$ (with respective + (or −) sign for K (or I) subfamily). The transformation (2.13) hence leads to a pair of subfamilies of processes $F_t \in (0,\infty)$ with respective volatility functions

$$\sigma(F) = \begin{cases} \dfrac{a_I}{\sqrt{X(F)}I_\mu^2(\sqrt{2\rho X(F)})}, & I\text{-subfamily}, \\ \dfrac{a_K}{\sqrt{X(F)}K_\mu^2(\sqrt{2\rho X(F)})}, & K\text{-subfamily}. \end{cases} \tag{2.14}$$

Despite the similarity between these two subfamilies, they possess essentially different properties. Of interest are the qualitative differences in the local volatility functions $\sigma_{loc}(F) \equiv \sigma(F)/F$ for both models. Let $\sigma_{loc}^{(K)}$ and $\sigma_{loc}^{(I)}$ denote the respective local volatility functions of the K and I-subfamilies. From (2.13) and (2.14), and the large and small argument asymptotics of the modified Bessel functions, we obtain the local volatility asymptotes:

$$\sigma_{loc}^{(K)}(F) \sim \begin{cases} \mu\sqrt{2\rho}\left[\dfrac{2a_K}{\Gamma(\mu)\Gamma(\mu+1)}\right]^{\frac{1}{2\mu}} F^{-\frac{1}{2\mu}}, & \text{as } F \to 0^+, \\ 2\sqrt{2\rho}, & \text{as } F \to \infty, \end{cases} \tag{2.15}$$

and

$$\sigma_{loc}^{(I)}(F) \sim \begin{cases} 2\sqrt{2\rho}, & \text{as } F \to 0^+, \\ \mu\sqrt{2\rho}\left[\dfrac{2a_I^{-1}}{\Gamma(\mu)\Gamma(\mu+1)}\right]^{\frac{1}{2\mu}} F^{\frac{1}{2\mu}}, & \text{as } F \to \infty\,. \end{cases} \tag{2.16}$$

Figure 1 shows typical plots of $\sigma_{loc}^{(K)}$ and $\sigma_{loc}^{(I)}$. The Bessel K-subfamily hence approaches a constant local volatility in the limit of large asset prices (i.e., approaches a lognormal model), whereas for small asset values the local volatility exhibits a power law singularity. On the other hand, the Bessel I-subfamily approaches a constant local volatility for small asset values, with unbounded power law in the opposing large asset value limit. As explained further below, the union of the two models forms a superset of the CEV model with zero drift function.

A pair of exact transition p.d.f.s for the two models are readily obtained by combining Eq. (2.7) (where $\hat{u}(x,\rho) = x^{-\mu/2}I_\mu(\sqrt{2\rho x})$ for I-subfamily and $\hat{u}(x,\rho) = x^{-\mu/2}K_\mu(\sqrt{2\rho x})$ for K-subfamily) with Eq. (2.10) and (2.14) giving:

$$U(F,F_0,t) = \begin{cases} \dfrac{e^{-\rho t-(x+x_0)/(2t)}}{a_I t}\dfrac{xI_\mu^3(\sqrt{2\rho x})}{I_\mu(\sqrt{2\rho x_0})} I_\mu\left(\dfrac{\sqrt{xx_0}}{t}\right) & (\text{for } I\text{-subfamily}), \\ \dfrac{e^{-\rho t-(x+x_0)/(2t)}}{a_K t}\dfrac{xK_\mu^3(\sqrt{2\rho x})}{K_\mu(\sqrt{2\rho x_0})} I_\mu\left(\dfrac{\sqrt{xx_0}}{t}\right) & (\text{for } K\text{-subfamily}), \end{cases} \tag{2.17}$$

where $x = X(F)$, $x_0 = X(F_0)$ are given by the respective inverses of (2.13).

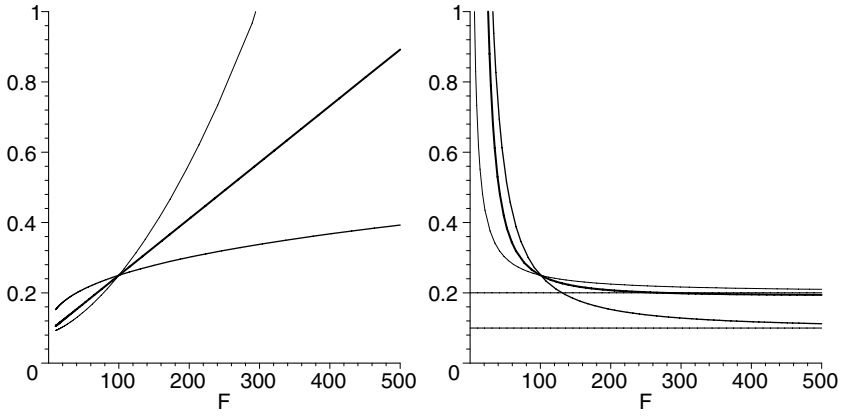

Fig. 1. Local volatility plots for the Bessel subfamilies. The Bessel I-subfamily (left plot) is plotted using (2.14) for the following choices of parameters: $a_I = 24.5302$, $\rho = 0.001$, $\mu = 0.25$ (the thinnest line); $a_I = 17.7323$, $\rho = 0.001$, $\mu = 0.5$ (the thickest line); and $a_I = 5.0574$, $\rho = 0.001$, $\mu = 1.5$. The Bessel K-subfamily (right plot) is plotted using (2.14) for the following choices of parameters: $a_K = 78.5398$, $\rho = 0.005$, $\mu = 0.5$ (the thinnest line); $a_K = 129.4506$, $\rho = 0.005$, $\mu = 0.125$ (the thickest line); and $a_K = 362.2907$, $\rho = 0.00125$, $\mu = 0.25$. The parameters are chosen in such a way that the value of the local volatility functions is fixed at 0.25 for $F = 100$.

We now briefly discuss the salient properties of these kernels and hence the probabilistic character of the corresponding F_t-processes described by them. The various properties follow by applying the large and small argument asymptotics of the modified Bessel functions. The first of these concerns boundary conditions at the endpoints, with fixed F_0, t. For the I-subfamily: $U = O(F^{-2-\frac{1}{\mu}}) \to 0$, as $F \to \infty$, and $U \to 0$, as $F \to 0^+$, where in this latter limit U approaches zero faster than any power law. A second property of interest is whether probability is conserved. That is, given any path with initial value $F_0 \in (0, \infty)$, then

$$R(F_0, t) = \frac{\partial}{\partial t}\int_0^\infty U(F, F_0, t)dF = \frac{1}{2}\frac{\partial}{\partial F}\big(\sigma^2(F)\,U(F, F_0, t)\big)\big|_{F=0^+}^{F=\infty},$$

gives the rate of change of the probability that any such path will have terminal value $F_t \in (0, \infty)$ at time t. This expression follows simply by taking the time derivative of U and using the forward Eq. (2.4). From (2.16) and the asymptotics of U, it follows that $R(F_0, t) = 0$ for the I-subfamily. This is consistent with directly integrating U by using the appropriate Bessel-I integral identity (see [3]). Hence U integrates to unity for all $t \geq 0$ and probability is conserved (i.e., no absorption occurs) for the I-subfamily. The last property of interest concerns whether the process F_t is a martingale (or else a strict super/submartingale). For this purpose we consider the rate of change of the expectation, $M(F_0, t) \equiv \frac{\partial}{\partial t}\mathsf{E}[F_t|F_{t=0} = F_0]$ where

$$M(F_0, t) = \int_0^\infty F\frac{\partial U}{\partial t}\,dF = \frac{1}{2}\left[F\frac{\partial}{\partial F}\big(\sigma^2(F)\,U(F, F_0, t)\big) - \sigma^2(F)U(F, F_0, t)\right]_{F=0^+}^{F=\infty},$$

follows from (2.4) and integrating by parts. From the asymptotics of U and $\sigma(F)$ it readily follows that $M(F_0,t)<0$, for all $t>0$, and hence that the process is a strict supermartingale, i.e., $\mathsf{E}[F_t|F_{t=0}=F_0]<F_0$ for $t>0$. Although we skip the details here, the explicit expression is easily derived and gives the time decay of $M(F_0,t)$ as a function that depends on the parameters of the model besides time t. The I-subfamily of processes with p.d.f. in (2.17) are therefore positive conservative strict supermartingales (not true martingales) for all choices of positive model parameters a_I,ρ,μ. If F_t represents a forward price or discounted asset price process under an assumed risk-neutral Q-measure, then the I-subfamily is a candidate for a so-called "bubble market" model for which F_t is a strict supermartingale, and where the usual textbook results of equivalent martingale arbitrage-free pricing do *not* strictly hold. For discussions on some of the financial implications of using strict supermartingale models for option pricing we refer the reader to [11] and references therein.

Repeating the above analysis for the K-subfamily with respective p.d.f. in (2.17) gives quite different properties, stated as follows. The boundary conditions at the endpoints, for fixed F_0,t, are now: $U=O(F^{\frac{1}{\mu}-1})$, as $F\to 0^+$, and $U\to 0$, as $F\to\infty$, where in this latter limit U approaches zero faster than any power law. Hence, at the lower endpoint limit $F\to 0^+$ we have either $U\to constant>0$, when $\mu=1$, or U approaches an integrable singularity if $1<\mu<\infty$, or $U\to 0$ (zero absorbing boundary condition) when $0<\mu<1$. The analysis of the absorption rate $R(F_0,t)$, for all $\mu>0$, shows that the lower endpoint is an absorbing (exit) boundary. More precisely, let τ denote the time for absorption, or first passage time for paths hitting the origin $F=0$ and starting at $F_0>0$. Then, $p(\tau;F_0)=-R(F_0,\tau)$ gives the first passage time density which is readily shown to have the generalized inverse Gaussian distribution:

$$p(\tau;F_0)=\frac{\big(X(F_0)/2\rho\big)^{\mu/2}}{2K_\mu\big(\sqrt{2\rho X(F_0)}\,\big)}\,\tau^{-\mu-1}e^{-\rho\tau-X(F_0)/2\tau},\ \ \tau>0. \tag{2.18}$$

Using the K-subfamily kernel U into the above expression for $M(F_0,t)$, it follows readily that $M(F_0,t)=0$. Hence, the F_t-processes of the K-subfamily are all martingales in the sense that $\mathsf{E}[F_t|F_{t=0}=F_0]=F_0$. This is in direct constrast to the strict supermartingale property that holds for the I-subfamily.

An interesting property of the above Bessel subfamilies is that they form a superset of the CEV models (with zero drift function). That is, the CEV volatility and the respective transition p.d.f.s are recovered in special limiting cases as $\rho\to 0^+$. The CEV models have a power function volatility $\sigma(F)=\delta F^{1+\beta}$ and hence are specified by two parameters δ,β. As was shown in [3], the I-subfamily recovers the CEV model for $\beta>0$ by considering the limit $\rho\to 0^+$ while keeping $a_I\rho^{-\mu}=constant$. In particular, defining $\beta\equiv(2\mu)^{-1}>0$ and $\delta\equiv C^{-\beta}/\beta$ with constant $C\equiv(a_I/2)(\rho/2)^{-\mu}\Gamma(\mu)\Gamma(\mu+1)$, and using the small argument asymptotics of I_μ and K_μ, for positive order $\mu=(2\beta)^{-1}$, the limiting form of the map in (2.13) simplifies to the explicit form $F(x)=(\delta^2\beta^2x)^{-\frac{1}{2\beta}}$, $X(F)=\frac{F^{-2\beta}}{\delta^2\beta^2}$. Then, by applying the limit to (2.14) and (2.17), and using the Gamma function property $\Gamma(\mu+1)=\mu\Gamma(\mu)$, the

I-subfamily volatility (with originally 3 parameters) reduces to the CEV volatility $\sigma(F) = \delta F^{1+\beta}$, while the transition p.d.f. reduces to the known p.d.f. function U_0 in equation (2.20) below, i.e., $U(F, F_0, t) \to U_0(F, F_0, t)$ where $\vartheta = \beta$. As further discussed below, the (increasing local volatility) CEV process described by such a transition p.d.f. exhibits the same basic probabilitistic properties as the I-subfamily.

The corresponding limiting form of the K-family has never been considered. The procedure is similar, but now $a_K \rho^\mu$ is kept constant as $\rho \to 0^+$. Setting $\beta = -(2\mu)^{-1}$ ($\mu = 1/2|\beta|$), $\delta \equiv C^{|\beta|}/|\beta|$, $C \equiv (2a_K)(\rho/2)^\mu/(\Gamma(\mu)\Gamma(\mu+1))$, the limit again gives $X(F) = F^{-2\beta}/(\delta^2\beta^2)$. The volatility again reduces to $\delta F^{1+\beta}$, but now with the important distinction that $\beta < 0$ (i.e., with decreasing local volatility). Moreover, the limiting form of the K-subfamily p.d.f. is given by (2.20) where $\vartheta = |\beta|$. This is the known density for the case of the martingale CEV process with absorbing boundary at $F = 0$. Taking the limit $\rho \to 0^+$ of the first passage time density in (2.18) also readily recovers the exact known expression for the first passage time density of the corresponding CEV model: $p^{CEV}(\tau; F_0) = ((\tilde{x}_0)^\mu/\Gamma(\mu))\tau^{-\mu-1}e^{-\tilde{x}_0/\tau}$, where $\tilde{x}_0 \equiv F_0^{-2\beta}/(2\delta^2\beta^2)$. Integrating this density from $\tau = 0$ to t then gives the known c.d.f. for absorption up to time t as $P(t; F_0) = \Gamma(\mu, \tilde{x}_0/t)/\Gamma(\mu)$, where $\Gamma(a, x) \equiv \int_x^\infty t^{a-1}e^{-t}dt$ is the complementary incomplete Gamma function.

2.2. *CEV diffusion models*

The CEV models with generally nonzero linear drift function formally obey the SDE [12]:

$$dS_t = \nu S_t dt + \delta S_t^{\beta+1} dW_t, \quad t \geq 0, \quad S_0 > 0. \tag{2.19}$$

In the previous section we derived transition p.d.f.s for the CEV diffusion models as limiting cases of the kernels for the Bessel subfamilies for $\nu = 0$. For both theoretical and computational purposes it is useful to give the direct connection between CEV models and squared Bessel processes. For the case of zero drift parameter $\nu = 0$, the change of variables $x_t = X(S_t) \equiv S_t^{-2\beta}/(\delta^2\beta^2)$ directly reduces the CEV process to a squared Bessel process (2.2) with drift parameter $\lambda = 1/\beta + 2$. This is easily seen, for example, by formally taking the stochastic differential of $X(S_t)$ and using (2.19) with $\nu = 0$. More precisely, by specializing the (diffusion canonical) methodology of the previous section for the case of only a variable transformation (and renaming variables $F_t \to S_t$) then one easily obtains transition p.d.f.s that solve the Kolmogorov PDE for CEV diffusions in S-space in terms of the p.d.f.s that solve the corresponding Kolmogorov PDE for the squared Bessel diffusions in x-space. There is no measure change (i.e., the factor $\frac{\hat{u}(X(S))}{\hat{u}(X(S_0))}e^{-\rho t}$ in (2.7) is simply unity in this case) and we have the relation $U_0(S, S_0, t) = |X'(S)|\, u(X(S), X(S_0), t)$. In our notation we use $U_\nu(S, S_0, t)$ to denote transition p.d.f.s for the CEV diffusions with drift parameter ν. The densities U_0 for $\nu = 0$ therefore obtain by simply substituting the p.d.f.s u for the x_t-process. In Sec. 2.1 we considered $\lambda > 2$ ($\mu = \frac{1}{2\beta} > 0$), with u given by (2.10), and thereby generated the Bessel I and K subfamilies. However,

in generating kernels for the CEV models directly from the squared Bessel process, one can also consider the case $\lambda < 2$ ($\mu = \frac{1}{2\beta} < 0$). Moreover, besides (2.10) we also consider another p.d.f. for $\mu < 0$ (i.e., an alternate solution kernel u to the Kolmogorov PDE for the x_t-process) which is given by the expression in (2.10) but with the replacement $I_\mu \to I_{|\mu|}$ (i.e., changing order of Bessel I from μ into $|\mu|$). By combining the different cases, the relevant transition p.d.f.s for the CEV models for $\nu = 0$ take the form:

$$U_0(S, S_0, t) = \frac{S^{-2\beta-3/2} S_0^{1/2}}{\delta^2 |\beta| t} \exp\left(-\frac{S^{-2\beta} + S_0^{-2\beta}}{2\delta^2 \beta^2 t} \right) I_{\frac{1}{2\vartheta}} \left(\frac{S^{-\beta} S_0^{-\beta}}{\delta^2 \beta^2 t} \right), \qquad (2.20)$$

Using the result of Goldberg [21], a CEV process for generally nonzero drift parameter ν, denoted by $S_t^{(\nu)}$, is obtained from a CEV process with $\nu = 0$ by means of a scale and time change:

$$S_t^{(\nu)} = e^{\nu t} S_{\tau(t)}^{(0)}, \quad \tau(t) = \begin{cases} \dfrac{1}{2\nu\beta} \left(e^{2\nu\beta t} - 1 \right), & \nu \neq 0, \\ t, & \nu = 0. \end{cases} \qquad (2.21)$$

From the Kolmogorov PDE for generally nonzero ν, one readily derives transition p.d.f.s as $U_\nu(S, S_0, t) = e^{-\nu t} U_0(e^{-\nu t} S, S_0, \tau(t))$. Observe that any homogeneous endpoint boundary condition satisfied by U_0 then also holds for U_ν. The probabilistic properties of the CEV models depend upon the value of β and which solution kernel is employed. Using (2.20), one generates three different transition p.d.f.s U_ν and hence three types of processes with features briefly summarized as follows:

(1) For $\beta > 0$ and $\vartheta = \beta$: U_ν integrates to unity (over $S \in (0, \infty)$) for all $t \geq 0$ (i.e., the process $S_t^{(\nu)}$ is not absorbed and hence probability is conserved), the process $S_t^{(0)}$ (with $\nu = 0$ and corresponding p.d.f. U_0) is a strict supermartingale, and U_ν obeys zero boundary conditions at $S = 0^+$ and ∞, where $U_\nu = O(S^{-2\beta-2})$ as $S \to \infty$, for any fixed values $S_0, t > 0$.

(2) For $\beta < 0$ and $\vartheta = |\beta|$: U_ν does not integrate to unity for $t > 0$, the lower endpoint $S = 0$ is absorbing and the process $S_t^{(0)}$ obeys the martingale property, i.e., is driftless where $\mathsf{E}[S_t^{(0)} | S_0^{(0)} = S_0] = S_0$. Moreover, $U_\nu \to 0$ as $S \to \infty$. As $S \to 0^+$, $U_\nu = O(S^{2|\beta|-1})$ which goes to zero only when $\beta < -1/2$.

(3) For $\beta < -1/2$ and $\vartheta = \beta$: U_ν integrates to unity (over $S \in (0, \infty)$) for all $t \geq 0$ (giving a process $S_t^{(\nu)}$ that is not absorbed) and the process $S_t^{(0)}$ is a strict submartingale. $U_\nu \to 0$ as $S \to \infty$, and $U_\nu = O(S^{2|\beta|-2})$ as $S \to 0^+$ which is a zero boundary condition when $\beta < -1$.

The above properties follow by use of the same methods described in the previous subsection, where the absorption rate for a process $S_t^{(\nu)}$ is given in terms of the rate for the corresponding $S_t^{(0)}$ process as $R_\nu(S_0, t) = \dot{\tau}(t) R_0(S_0, \tau(t))$. Moreover, for any U_ν one has the identity $\mathsf{E}[S_t^{(\nu)} | S_0^{(\nu)} = S_0] = e^{\nu t} \mathsf{E}[S_{\tau(t)}^{(0)} | S_0^{(0)} = S_0]$. Hence, for the process defined by case (2) above: $\mathsf{E}[S_t^{(\nu)} | S_0^{(\nu)} = S_0] = e^{\nu t} S_0$ so that $S_t^{(\nu)}$ actually drifts at rate ν, i.e., $e^{-\nu t} S_t^{(\nu)}$ is driftless, and under the risk-neutral measure

with constant interest rate r the forward price $F_t = e^{-rt}S_t$ (where $S_t = S_t^{(\nu=r)}$) is a martingale.

3. The Bessel and Randomized Gamma Distributions

A nonnegative integer random variable Y is said to be a Bessel random variable with parameters $\mu > -1$ and $b > 0$ if

$$\mathsf{P}\{Y = n\} = \frac{(b/2)^{2n+\mu}}{I_\mu(b)\, n!\, \Gamma(n+\mu+1)}, \quad n = 0, 1, 2, \ldots.$$

This expression follows from the Taylor series:

$$I_\mu(x) = \sum_{n=0}^{\infty} \frac{(x/2)^{2n+\mu}}{n!\, \Gamma(n+\mu+1)}. \tag{3.1}$$

For simplicity, we use the notation $Bes(\mu, b)$ for the Bessel distribution with parameters μ and b, i.e., $Y \sim Bes(\mu, b)$. This distribution arises in Pitman and Yor [27] and is related to many other distributions, where the Bessel function is involved in the density, including multivariate and randomized gamma distributions and the squared Bessel bridge distribution (see [32]).

Let $G(\nu, \kappa)$ denote the gamma distribution with shape parameter ν and scale parameter κ, where $\nu > 0$ and $\kappa > 0$, and let $P(\alpha)$ denote the Poisson distribution with mean α. The randomized gamma distribution of the *first type* is the mixture distribution $G(\mu + \eta + 1, \kappa)$, $\kappa > 0$, $\mu > -1$, where η has the Poisson distribution. For any positive numbers κ, α and $\mu > -1$, the randomized gamma variable, with distribution $G(\mu + \eta + 1, \kappa)$ where $\eta \sim P(\alpha)$, has the p.d.f.

$$f_1(y) = \kappa y^{\mu/2} e^{-\alpha-\kappa y} I_\mu(\sqrt{4\kappa\alpha y}) \left(\frac{\kappa}{\alpha}\right)^{\mu/2}, \quad y > 0. \tag{3.2}$$

The randomized gamma distribution of the *second type* is the mixture distribution $G(\mu + \eta_1 + 2\eta_2 + 1, \kappa)$, $\kappa > 0$, $\mu > -1$, where η_1 and η_2 are independent variates with the Poisson and Bessel distributions respectively. This distribution can be viewed as a generalization of the first type distribution, which is recovered in the limiting case with $\eta_2 = 0$. For any positive numbers κ, a, b and $\mu > -1$, the randomized gamma distribution $G(\mu + \eta_1 + 2\eta_2 + 1, \kappa)$ with $\eta_1 \sim P((a+b)/(4\kappa))$ and $\eta_2 \sim Bes(\mu, \sqrt{ab}/(2\kappa))$ has the p.d.f.

$$f_2(y) = \frac{\kappa}{I_\mu(\sqrt{ab}/(2\kappa))} e^{-(a+b)/(4\kappa)-\kappa y} I_\mu(\sqrt{ay}) I_\mu(\sqrt{by}), \quad y > 0. \tag{3.3}$$

These two distributions are closely connected with the squared Bessel process and the squared Bessel bridge. For real $\lambda > 0$ (i.e., $\mu > -1$), the λ-dimensional squared Bessel process has paths $x(t)$, $t \geq 0$, conditional on $x(0) = x_0$, with transition p.d.f. given by (2.10), with replacement $I_\mu \to I_{|\mu|}$ for $\mu < 0$. As is evident from

(3.2), the conditional distribution of $x(t)$, given $x(0) = x_0$, is then a randomized gamma distribution of the first type:

$$x(t) \sim G(\mu + \eta + 1, 1/(2t)) \quad \text{with}\, \eta \sim P(x_0/(2t)). \tag{3.4}$$

A standard squared Bessel bridge $y_{x_0,x_1}(t)$ is a stochastic process on $[t_0, t_1]$, $0 \le t_0 < t_1$, generated by $x(t)$ and with $x(t_0)$ and $x(t_1)$ fixed at points x_0 and x_1, respectively. The bridge p.d.f. $u_{x_0,x_1}(y, t, t_0, t_1)$ of $y = y_{x_0,x_1}(t)$ for $t_0 < t < t_1$ is given by

$$u_{x_0,x_1}(y, t, t_0, t_1) = \frac{u(y, x_0, t - t_0)u(x_1, y, t_1 - t)}{u(x_1, x_0, t_1 - t_0)}. \tag{3.5}$$

Regardless of whether or not $u(x, x_0, t - t_0)$ integrates to unity, it is clear that the bridge p.d.f. (3.5) integrates to unity thanks to the Chapman–Kolmogorov equation:

$$u(x_1, x_0, t_1 - t_0) = \int_0^\infty u(x_1, y, t_1 - t)u(y, x_0, t - t_0)dy\,, \quad t \in (t_0, t_1)\,.$$

Upon using (2.10), the bridge p.d.f. in (3.5) reduces to the form in (3.3) by setting $a \equiv x_0/(t - t_0)^2$, $b \equiv x_1/(t_1 - t)^2$, $\kappa \equiv \frac{t_1 - t_0}{2(t-t_0)(t_1-t)}$. Hence, the p.d.f. of $y_{x_0,x_1}(t)$ is a randomized gamma distribution of the second type. Then, $y(t) = y_{x_0,x_1}(t)$, $t_0 < t < t_1$, can be obtained by generating independent random variables

$$\eta_1(t) \sim P\left(\frac{1}{2(t_1 - t_0)}\left[\frac{t_1 - t}{t - t_0}x_0 + \frac{t - t_0}{t_1 - t}x_1\right]\right) \quad \text{and} \quad \eta_2 \sim Bes\left(\mu, \frac{\sqrt{x_0 x_1}}{t_1 - t_0}\right),$$

$$\text{and then } \; y(t) \sim G\left(\mu + \eta_1(t) + 2\eta_2 + 1, \frac{t_1 - t_0}{2(t - t_0)(t_1 - t)}\right). \tag{3.6}$$

To simulate a Bessel distribution, we can use a rejection method (see [16]). One of the available algorithms is based on known properties of the Bessel distribution and a discrete log-concave distribution stated below without proof.

Lemma 3.1. *The Bessel distribution is log-concave, that is, the ratio* p_{n+1}/p_n, *where* $p_n = \mathsf{P}\{Y = n\}$, *and* $Y \sim Bes(\mu, b)$, *is decreasing in* n. *Furthermore, the Bessel distribution is unimodal and has a unique mode or two modes at consecutive integers, and one mode is always located at*

$$m = \left\lfloor \frac{\sqrt{\mu^2 + b^2} - \mu}{2} \right\rfloor.$$

Lemma 3.2. [15] *For any discrete log-concave distributions with mode at* m, *we have, for all* $n \ge 0$:

$$p_n \le p_m \min\left\{1, e^{1 - p_m|n-m|}\right\}.$$

As is seen from the previous lemma, for all $k \ge -m$ and $k - 1/2 \le x \le k + 1/2$, $p_{m+k} \le g(x)$ holds, where $g(x) = p_m \min\left\{1, e^{1-p_m(|x|-1/2)}\right\}$. So, the function g is a dominant function for any discrete log-concave distribution with mode at m and with the probability vector $\{p_n\}_{n\ge 0}$. Observe that g is a mixture

of a rectangular function on $[-w/p_m, w/p_m]$ and two exponential tails outside $[-w/p_m, w/p_m]$, where $w = 1 + p_m/2$. When g is used in the rejection algorithm, the rejection constant (expected number of iterations before succeeding) is $\int_{-\infty}^{\infty} g(x)dx = 2w + 2 = 4 + p_m \leq 5$ [15]. The algorithm is as follows.

```
Let w = 1 + p_m/2;
Repeat
    Generate independent uniform random variables α_1, α_2 ~ U[0,1)
             and a random sign s = ±1;
    If α_1 ≤ w/(1+w)
       then Generate α_3 ~ U[0,1);
            Set X = α_3/p_m
       else Generate e ~ E(1);
            Set X = (w+e)/p_m;
    Set K = s · round(X)
Until K ≥ −m and α_2 min{1, e^{w−p_m X}} ≤ p_{m+K}/p_m;
Return m + K
```

Here, $U[0,1)$ denotes the uniform distribution on the interval $[0,1)$, and $E(1)$ stands for the exponential distribution with mean one. Note that probability p_m need only be computed once, and that p_{m+K}/p_m does not require the evaluation of the modified Bessel function, nor the Gamma function:

$$\frac{p_{m+K}}{p_m} = \frac{b^{2K}}{4^K} \times \begin{cases} \prod_{l=1}^{K} ((m+l)(m+l+\mu))^{-1}, & K \geq 0, \\ \prod_{l=1}^{-K} (m-l+1)(m-l+\mu+1), & K < 0. \end{cases}$$

4. A Bessel Bridge for Path Dependent Option Pricing

4.1. *Path integral approach*

As discussed in [10], the problem of valuing European style options with (discrete-time) path dependent payoffs (e.g., Asian and lookback options) within the path integral approach reduces to the evaluation of an integral in asset price space. Within this approach, the total time interval $[0,T]$, $T > 0$, is discretized with partition $\mathcal{T}$, and time steps $\Delta t_i = t_i - t_{i-1} > 0$, $i = 1, \ldots, N$:

$$\mathcal{T} = \{t_i\}_{i=0,1,\ldots,N}, \quad 0 = t_0 < t_1 < \cdots < t_N = T. \tag{4.1}$$

The values of the price process[1] F_t at time points $t = t_i$, are denoted by F_i, for all $i = 0, 1, \ldots, N$. Within the discrete time path integral framework, the payoff functions Λ depend upon all intermediate path values $(F_i)_{1 \leq i \leq N}$, i.e., $\Lambda = \Lambda(\mathbf{F})$

[1]Note that an analogous formulation applies to the use of p.d.f.s in the asset price (e.g., for the stock price S_t) in place of the p.d.f.s for the discounted asset price F_t.

where $\boldsymbol{F}$ denotes the N-dimensional vector formed by the values F_i, $i = 1, \ldots, N$ along any time discretized path. Within any of the above models, the domain of allowable price path values is the positive real line $R_+ \equiv (0, \infty)$. Using a transition p.d.f., i.e., a pricing kernel U for each time lapse Δt_i, we obtain the joint p.d.f. (conditional on the initial asset value F_0 at time t_0) of any asset path $\boldsymbol{F} \in R_+^N$:

$$\boldsymbol{U}(\boldsymbol{F}; F_0, \mathcal{T}) = \prod_{i=1}^{N} U(F_i, F_{i-1}, \Delta t_i). \tag{4.2}$$

By repeatedly applying the Chapman–Kolmogorov relation, this p.d.f. then integrates to

$$W(F_0, T) \equiv \int_{R_+} \cdots \int_{R_+} \boldsymbol{U}(\boldsymbol{F}; F_0, \mathcal{T}) dF_1 \cdots dF_N = \int_{R_+} U(F_N, F_0, T) dF_N. \tag{4.3}$$

Hence, for kernels U without absorption $W(F_0, T) = 1$. This applies to the Bessel I-subfamily and the super/sub-martingale CEV models. For kernels U with absorption $W(F_0, T) = 1 - P(T; F_0)$, where $P(T; F_0)$ is the c.d.f. for absorption at the lower endpoint of the process up to time T (this applies to the Bessel K-subfamily and the CEV model (defined by case (2) of Sec. 2.2) where forward prices are martingales in a risk-neutral measure).

From the p.d.f. in (4.2), and assuming a payoff function Λ dependent on the set of all intermediate path values at the discrete times, then the value of a path dependent European style derivative at current time $t_0 = 0$, and maturing at time $T = t_N$, is expressible in terms of an integral, $\Theta = \Theta_N(F_0, T)$, over all attainable intermediate asset price values, where

$$\begin{aligned} \Theta &= \int_{R_+} \cdots \int_{R_+} \boldsymbol{U}(\boldsymbol{F}; F_0, \mathcal{T}) \Lambda(F_1, \ldots, F_N) dF_1 \cdots dF_N \\ &= \int_{R_+} \cdots \int_{R_+} \prod_{i=1}^{N} U(F_i, F_{i-1}, \Delta t_i) \Lambda(F_1, \ldots, F_N) dF_1 \cdots dF_N. \end{aligned} \tag{4.4}$$

It is evident that within this formulation one can work with either the discounted asset price process F_t or with the asset (e.g., stock) price process S_t. Let D_t stand for a discount factor that depends on calendar time t. Then $F_t = D_t S_t$ (or $S_t = D_t^{-1} F_t$). Although we can also use other numeraires, in what follows we will assume a numeraire asset with value $g_t = e^{rt}$ as the domestic money market account with constant interest rate r, so that $D_t = g_0/g_t = e^{-rt}$ and $F_t = e^{-rt} S_t$. Two types of arithmetic average Asian payoffs are: one with time-averaging over F_t and the other with time-averaging over S_t. That is, we consider options with payoffs $\Lambda(F_T, I_T^{(F)})$, $I_T^{(F)} \equiv \frac{1}{T} \int_0^T F_t dt$, or with payoffs $\Lambda(S_T, I_T^{(S)})$, $I_T^{(S)} \equiv \frac{1}{T} \int_0^T S_t dt$. This latter form can also be recast as an average over a time-changed forward price: $I_T^{(S)} = \frac{1}{T} \int_0^{\frac{1}{r}(e^{rT}-1)} F_{\bar{\tau}(t)} dt$, $\bar{\tau}(t) \equiv \frac{1}{r} \log(1 + rt)$. Note that these payoffs are the continuous time limits ($\Delta t_i \to 0$, $N \to \infty$) of the discretized forms. Examples include the arithmetic average forward price call struck at K where

$\Lambda = (I_T^{(F)} - K)_+$, $I_T^{(F)} = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N F_i$, and its analoguous arithmetic average asset (stock) price call with $\Lambda = (I_T^{(S)} - K)_+$, $I_T^{(S)} = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N S_i$. Hence, the time-discretized arithmetic average forward price call has a payoff of the form $\Lambda(F_1, \ldots, F_N) = (\frac{1}{N}\sum_{i=1}^N F_i - K)_+$ where N is the number of discrete monitoring times, and an obvious similar form for the average stock price call.

Within a risk-neutral pricing measure $Q = Q(g)$, with $g_t = e^{rt}$ as numeraire, the current value of the derivative with $\Lambda(F_T, I_T^{(F)})$ as payoff is

$$V(F_0, T) = e^{-rT}\mathsf{E}\big[\Lambda(F_T, I_T^{(F)})|F_{t_0} = F_0\big] = e^{-rT}\lim_{N\to\infty}\Theta_N(F_0, T), \tag{4.5}$$

and with payoff $\Lambda(S_T, I_T^{(S)})$ is

$$V(S_0, T) = e^{-rT}\mathsf{E}\big[\Lambda(S_T, I_T^{(S)})|S_{t_0} = S_0\big] = e^{-rT}\lim_{N\to\infty}\Theta_N(S_0, T). \tag{4.6}$$

Of course, when interest rates are negligible the two values are the same. The conditional expectations $\mathsf{E}\big[\,\cdot\,\big] \equiv \mathsf{E}^Q\big[\,\cdot\,\big]$ are taken w.r.t. the Q-measure using the p.d.f. in (4.2), or using its analogue $\boldsymbol{U}(\boldsymbol{S}; S_0, \mathcal{T}) = \Pi_{i=1}^N U(S_i, S_{i-1}, \Delta t_i)$, where $U(F_i, F_{i-1}, \Delta t_i)$ or $U(S_i, S_{i-1}, \Delta t_i)$ are the respective risk-neutral pricing kernels for time step Δt_i. In particular, for the Bessel family of models we shall use $U(F_i, F_{i-1}, \Delta t_i)$ given by (2.17) in a forward price framework and where our calculations shall be restricted to pricing (4.5) within such models, although the evaluation of (4.6) for nonzero interest rates follows readily with minor modifications as partly depicted above. For the CEV model, we shall value average Asian options of the form in (4.6) and consider only the processes for which the forward price $F_t = e^{-rt}S_t$ is a martingale (defined by case (2) of Sec. 2.2) with respective pricing kernel $U(S_i, S_{i-1}, \Delta t_i) = U_{\nu=r}(S_i, S_{i-1}, \Delta t_i)$.[2]

The value of a path-dependent derivative in the above framework hence involves computing a conditional expectation of the payoff:

$$\Theta = \Theta_N(F_0, T) = \mathsf{E}[\Lambda(F_1, F_2, \ldots, F_N)|F_{t_0} = F_0]. \tag{4.7}$$

In particular, for some N, the valuation of $V(F_0, T)$ in (4.5) is reduced to the problem of evaluating a finite N-dimensional integral as given by (4.4), with an analogous form for (4.6). So, for processes without absorption we can generate sample discretized paths for $\boldsymbol{F}$ (or $\boldsymbol{S}$) within a direct simulation method. The situation is slightly different, if the assumed asset process is absorbing. In this case, we cannot use a direct simulation technique, but rather should normalize the p.d.f.

[2]As shown, the Bessel I-subfamily is a class of strict supermartingale processes. The CEV model also gives rise to subclasses of strict supermartingales or submartingales (cases (1) and (3) of Sec. 2.2). If desired, these models can also be used as precursors to new processes obeying the martingale condition, at least in a piecewise sense as follows. Let $(F_t)_{t\in[t_{i-1},t_i]}$ be any such semimartingale process, then $\tilde{F}_t \equiv F_t\tilde{F}_{t_{i-1}}/\mathsf{E}[F_t|F_{t_{i-1}}=F_{i-1}]$ defines a new forward price process obeying $\mathsf{E}[\tilde{F}_t|\tilde{F}_{t_{i-1}}=\tilde{F}_{i-1}]=\tilde{F}_{i-1}$, for $t \in [t_{i-1}, t_i]$. It also follows by iteration that $\mathsf{E}[\tilde{F}_t|\tilde{F}_0]=\tilde{F}_0$.

and apply a weighted method. Hence, to evaluate the expectation in (4.7) for the case of an absorption we use a weighted estimate:

$$\Theta = W(F_0, T)\,\tilde{\mathsf{E}}[\Lambda(F_1, F_2, \ldots, F_N)|F_{t_0} = F_0], \tag{4.8}$$

with weight $W(F_0, T)$ as given in (4.3), and where the expectation $\tilde{\mathsf{E}}[\,\cdot\,]$ is with respect to the normalized distribution of $\boldsymbol{F}$ defined by $\tilde{\boldsymbol{U}} \equiv \boldsymbol{U}(\boldsymbol{F}; F_0, \mathcal{T})/W(F_0, T)$, which integrates to unity over $\boldsymbol{F} \in R_+^N$.

As the example below for $N = 4$ shows, for bridge sampling it is better, at first, to reorganize the joint p.d.f. with the goal of simplifying the expression of the weight. Consider (4.2) for a discretized path (F_1, F_2, F_3, F_4):

$$\boldsymbol{U}(\boldsymbol{F}; F_0, \mathcal{T}) = U(F_1, F_0, \Delta t_1)U(F_2, F_1, \Delta t_2)U(F_3, F_2, \Delta t_3)U(F_4, F_3, \Delta t_4).$$

By multiplying and dividing by some transition densities and reorganizing terms, the joint p.d.f. can be rewritten as follows:

$$\begin{aligned}\boldsymbol{U}(\boldsymbol{F}; F_0, \mathcal{T}) = {} & \frac{U(F_2, F_1, \Delta t_2)U(F_1, F_0, \Delta t_1)}{U(F_2, F_0, \Delta t_1 + \Delta t_2)} \times \frac{U(F_4, F_3, \Delta t_4)U(F_3, F_2, \Delta t_3)}{U(F_4, F_2, \Delta t_3 + \Delta t_4)} \\ & \times \frac{U(F_4, F_2, \Delta t_3 + \Delta t_4)U(F_2, F_0, \Delta t_1 + \Delta t_2)}{U(F_4, F_0, T)} \times U(F_4, F_0, T).\end{aligned} \tag{4.9}$$

The first term in the right-hand side of (4.9) is the bridge p.d.f. of F_1 conditional on the process having values $F_{t_0} = F_0$ and $F_{t_2} = F_2$. The second term is the bridge p.d.f. of F_3 conditional on $F_{t_2} = F_2$ and $F_{t_4} = F_4$. The third term is the bridge p.d.f. of F_2 conditional on $F_{t_0} = F_0$ and $F_{t_4} = F_4$. All these bridge p.d.f.s integrate to unity (thanks to the Chapman–Kolmogorov equation) and, for the models considered here, can be reduced to the randomized gamma distribution of the second kind (see next two subsections). The only term with absorption is the last transition p.d.f. $U(F_N, F_0, T)$. According to the principle of importance sampling, the sampling probability density needs to be proportional to an absorbing probability density. To estimate (4.4) we use a weighted Monte Carlo estimate in accordance with (4.8). The last variable F_N is simulated using the normalized density $U(F_N, F_0, T)/W(F_0, T)$, and the other variables $F_1, \ldots, F_{N-1}$ are simulated using bridge p.d.f.s.

The next two subsections are devoted to developing simulation methods for sampling with either the direct use of transition p.d.f.s or bridge p.d.f.s. We use the fact that the transition densities for both the Bessel family and CEV diffusion models reduce to the randomized gamma distributions.

4.2. *Bessel bridge for the Bessel family of models*

As discussed in Sec. 2, the state dependent F_t-processes that we consider are reducible to some underlying process (2.5) by applying a nonlinear change of state variables and a measure change. According to (2.7), we can also express the F-space

pricing kernel U in terms of x-space variables:

$$U(F, F_0, t) = \frac{\nu(X(F))}{\sigma(F)} \tilde{u}(X(F), X(F_0), t),$$

where

$$\tilde{u}(x, x_0, t) = e^{-\rho t} \frac{\hat{u}(x)}{\hat{u}(x_0)} u(x, x_0, t).$$

An F_t-process can also be viewed as a direct transformation $F_t = F(x_t)$ of an x_t-process having transition p.d.f. $\tilde{u}(x, x_0, t)$ within the same pricing measure. For our particular purpose, the last equation applied to (2.17) gives:

$$\tilde{u}(x, x_0, t) = \begin{cases} \dfrac{1}{2t} e^{-\rho t - (x+x_0)/(2t)} \dfrac{I_\mu(\sqrt{2\rho x})}{I_\mu(\sqrt{2\rho x_0})} I_\mu\left(\dfrac{\sqrt{x x_0}}{t}\right), & I\text{-subfamily}, \\ \dfrac{1}{2t} e^{-\rho t - (x+x_0)/(2t)} \dfrac{K_\mu(\sqrt{2\rho x})}{K_\mu(\sqrt{2\rho x_0})} I_\mu\left(\dfrac{\sqrt{x x_0}}{t}\right), & K\text{-subfamily}. \end{cases} \tag{4.10}$$

The following result allows us to do exact sampling of paths for the Bessel I-subfamily within a sequential approach. Its implementation is discussed further in Sec. 4.4.

Statement 4.1. The process $(x_t)_{t\geq 0}$ with transition p.d.f. $\tilde{u}(x_t, x_0, t)$ for the I-subfamily in (4.10) follows a randomized gamma distribution of the second type. Let $t > 0$ and $F_0 > 0$ hold, then the random variate $F \equiv F_t$, having the p.d.f. $U(F, F_0, t)$ in (2.17) for the I-subfamily, can be obtained by generating independent random variables

$$\eta_1(t) \sim P\left(\rho t + \frac{X(F_0)}{2t}\right) \quad \text{and} \quad \eta_2 \sim Bes\left(\mu, \sqrt{2\rho X(F_0)}\right),$$

and then

$$F_t = F(x_t), \quad \text{where } x_t \sim G\left(\mu + \eta_1(t) + 2\eta_2 + 1, \frac{1}{2t}\right).$$

Proof. The statement results immediately by considering the formulas (3.3), (3.4), and (4.10). □

This result can also be applied to the Bessel K-subfamily when evaluating the expectation in (4.8). This only requires minor modifications as follows. Let $\tilde{u}_I$ and $\tilde{u}_K$ denote the transition p.d.f.s in (4.10) for I and K-subfamily, respectively. Then, by changing integration variables $F_i = F(x_i)$, $i = 1, \ldots, N$, as defined by (2.13) for the K-subfamily, the path integral in (4.4) can be recast as an N-dimensional integral over the x-space variables $x_i \in R_+$ with integrand $\frac{I_\mu(\sqrt{2\rho X(F_0)})}{K_\mu(\sqrt{2\rho X(F_0)})} \frac{K_\mu(\sqrt{2\rho x_N})}{I_\mu(\sqrt{2\rho x_N})} \Lambda(F(x_1), \ldots, F(x_N)) \prod_{i=1}^N \tilde{u}_I(x_i, x_{i-1}, \Delta t_i)$. This follows from the product of all ratios $\frac{\tilde{u}_K(x_i, x_{i-1}, \Delta t_i)}{\tilde{u}_I(x_i, x_{i-1}, \Delta t_i)}$, which simplifies by repeated cancellations, and by using the Jacobian $\prod_{i=1}^N |F'(x_i)| = \prod_{i=1}^N \sigma(F(x_i))/\nu(x_i)$.

By applying the analogue of formula (3.5), for an F_t-process with p.d.f. $U(F_t, F_{t_0}, t-t_0)$ in place of the p.d.f. $u(x_t, x_{t_0}, t-t_0)$, and using the representation (2.7), we have the following expression for the bridge p.d.f. of an F_t-process:

$$\begin{aligned} &U_{F_1,F_2}(F,t,t_1,t_2) \\ &\quad = \frac{\nu(X(F))}{\sigma(F)} \frac{\tilde{u}(X(F),X(F_1),t-t_1)\tilde{u}(X(F_2),X(F),t_2-t)}{\tilde{u}(X(F_2),X(F_1),t_2-t_1)} \\ &\quad = \frac{\nu(X(F))}{\sigma(F)} \frac{u(X(F),X(F_1),t-t_1)u(X(F_2),X(F),t_2-t)}{u(X(F_2),X(F_1),t_2-t_1)}. \end{aligned} \tag{4.11}$$

Thus, in the particular case when the underlying process x_t is a squared Bessel process, a bridge process generated by F_t is just a diffusion transformation of the standard Bessel bridge $y_t = y_{x_1,x_2}(t)$, where $x_1 = X(F_1)$ and $x_2 = X(F_2)$. Also, the distribution of the bridge process generated by the underlying x_t-process coincides with that of the standard squared Bessel bridge. So, we can again use the randomized gamma distribution of the second type for modelling the bridge process in F-space for either Bessel I or K-subfamilies as stated just below.

Statement 4.2. Let $0 \le t_1 < t < t_2$ hold, then for all positive values F_1 and F_2, the random variate $F = F_t$, having the bridge p.d.f. $U_{F_1,F_2}(F,t,t_1,t_2)$ can be obtained by generating independent random variables

$$\eta_1(t) \sim P\left(\frac{1}{2(t_2-t_1)}\left[\frac{t_2-t}{t-t_1}X(F_1) + \frac{t-t_1}{t_2-t}X(F_2)\right]\right) \quad \text{and}$$

$$\eta_2 \sim Bes\left(\mu, \frac{\sqrt{X(F_1)X(F_2)}}{t_2-t_1}\right),$$

and then

$$F = F(y_t), \quad \text{where } y_t \sim G\left(\mu + \eta_1(t) + 2\eta_2 + 1, \frac{t_2-t_1}{2(t-t_1)(t_2-t)}\right).$$

Proof. The statement results by considering formulas (3.6) and (4.11). □

4.3. *A Bessel bridge for the CEV model*

As discussed in Sec. 2.2, the CEV process with nonzero drift arises from its respective process with zero drift by means of a scale and time change. Moreover, the latter process reduces to the squared Bessel process, which follows the randomized gamma distribution. Using the samplings in (3.4) and (3.6), it follows that a CEV process without absorption can be simulated with the result just below.

Statement 4.3. For all $S_t^{(\nu)} > 0$, $t \ge 0$, and $\Delta t > 0$, the value $S_{t+\Delta t}^{(\nu)}$ of a CEV process with drift parameter ν and without absorption is obtained by generating random variables

$$\eta \sim P\left(\frac{x_t}{2\Delta\tau}\right), \quad \gamma \sim G\left(\frac{1}{2\beta} + \eta + 1, \frac{1}{2\Delta\tau}\right),$$

and then

$$S^{(\nu)}_{t+\Delta t} = e^{\nu(t+\Delta t)}(\gamma\delta^2\beta^2)^{-1/(2\beta)},$$

where $x_t = \frac{\left(e^{-\nu t}S^{(\nu)}_t\right)^{-2\beta}}{\delta^2\beta^2}$ and $\Delta\tau = \tau(t+\Delta t) - \tau(t)$.

On the other hand, all bridge CEV processes, whether with or without absorption, are simulated as follows.

Statement 4.4. Let $0 \le t_1 < t < t_2$ hold, then for all positive values S_1 and S_2, the value $S^{(\nu)}_t$ of the bridge CEV process with drift ν, conditional on $S^{(\nu)}_{t_1} = S_1$ and $S^{(\nu)}_{t_2} = S_2$, is obtained by generating independent random variables

$$\eta_1(t) \sim P\left(\frac{1}{2(\tau_2-\tau_1)}\left[\frac{\tau_2-\tau(t)}{\tau(t)-\tau_1}x_1 + \frac{\tau(t)-\tau_1}{\tau_2-\tau(t)}x_2\right]\right), \quad \eta_2 \sim Bes\left(\frac{1}{2\vartheta}, \frac{\sqrt{x_1x_2}}{\tau_2-\tau_1}\right),$$

and then

$$S^{(\nu)}_t = e^{\nu t}(x_t\delta^2\beta^2)^{-1/(2\beta)},$$

where

$$x_t \sim G\left(\frac{1}{2\vartheta} + \eta_1(t) + 2\eta_2 + 1, \frac{\tau_2-\tau_1}{2(\tau(t)-\tau_1)(\tau_2-\tau(t))}\right),$$

$$\tau_i = \tau(t_i),\ x_i = \frac{\left(e^{-\nu t_i}S_i\right)^{-2\beta}}{\delta^2\beta^2}, \quad i = 1, 2.$$

We emphasize that this bridge sampling result applies to all viable cases of the CEV model discussed in Sec. 2.2 where either $\vartheta = \beta$ or $\vartheta = |\beta|$.

When computing path integrals for option valuation, our methods use a combination of sequential and bridge sampling of asset paths, as described in the next section. Moreover, we exploit the use of the underlying x_t-processes and thereby compute integrals in the corresponding x-space. As part of the integration procedure, some paths are then sampled directly with the use of transition p.d.f.s for the x_t-process. Hence, within the CEV model with absorption (where $\beta < 0$ and $\vartheta = |\beta|$) a similar direct sequential type of simulation still applies by slightly modifying the procedure in Statement 4.3 as follows. Recall from Sec. 2.2 that in this case the transition p.d.f. for the x_t-process is given by (2.10) but with the change $I_\mu \to I_{|\mu|}$, where $\mu = 1/2\beta < 0$. The point $x = 0$ is an exit in this case, and in particular we must first normalize transition p.d.f.s that are used in a direct sequential sampling of the x-space paths. In particular, given a p.d.f. $u(x, x_0, t)$ we compute $\int_0^\infty u(x, x_0, t)\,dx = \frac{\gamma\left(\frac{x_0}{2t},\theta\right)}{\Gamma(\theta)}$, where $\theta = |\mu|$ and $\gamma(a, x) = \int_0^x t^{a-1}e^{-t}\,dt$ is the lower incomplete Gamma function. By using (3.1), the normalized p.d.f. becomes:

$$\frac{\Gamma(\theta)}{\gamma(\frac{x_0}{2t},\theta)}u(x, x_0, t) = \sum_{n=0}^{\infty} e^{-\frac{x_0}{2t}}\frac{(\frac{x_0}{2t})^{n+\theta}}{\Gamma(n+\theta+1)}\frac{\Gamma(\theta)}{\gamma(\frac{x_0}{2t},\theta)}\,e^{-x/2t}\frac{x^n}{n!\,(2t)^{n+1}}. \tag{4.12}$$

The last part of the general term of this series is the p.d.f. of the Gamma distribution $G(n+1, 1/(2t))$. The first parts for $n = 0, 1, 2, \ldots$ form a sequence of probabilities

of a discrete random variate. By analogy with the previous section, this random variate γ, taking on nonnegative integer values with probabilities

$$\mathsf{P}\{\gamma = n\} = e^{-a}\frac{a^{n+\theta}}{\Gamma(n+\theta+1)}\frac{\Gamma(\theta)}{\gamma(a,\theta)}, \quad n = 0, 1, 2, \ldots, \tag{4.13}$$

is said to follow a "*incomplete Gamma*" discrete distribution which we simply denote by $\mathrm{I}\Gamma(a,\theta)$ with parameters $a > 0$ and $\theta > -1$. Thus, the above normalized transition p.d.f. for the squared Bessel process with absorption follows a mixture Gamma distribution $G(\gamma + 1, 1/(2t))$ where $\gamma \sim \mathrm{I}\Gamma(x_0/(2t), \theta)$.

Lemma 4.1. *The incomplete Gamma distribution is log-concave, that is, the ratio* $\mathsf{P}\{\gamma = n+1\}/\mathsf{P}\{\gamma = n\}$, *where* $\gamma \sim \mathrm{I}\Gamma(a,\theta)$, *is decreasing in* n. *Furthermore, the incomplete Gamma distribution is unimodal and has a unique mode or two modes at consecutive integers.*

Proof. By using (4.13), we obtain that $\frac{\mathsf{P}\{\gamma=n+1\}}{\mathsf{P}\{\gamma=n\}} = \frac{a}{(n+\theta+1)}$ is a monotonically decreasing function of n. Hence the incomplete Gamma distribution is log-concave. To prove the second statement it is enough to show that $f(x;a,\theta) = a^x/\Gamma(x+\theta+1)$ is monotonically decreasing or has only one maximum point with respect to x. The sign of its derivative

$$\frac{df}{dx}(x;a,\theta) = \frac{a^x}{\Gamma(x+\theta+1)}\left(\ln(a) - \Psi(x+\theta+1)\right),$$

is determined by $\ln(a) - \Psi(x+\theta+1)$, where $\Psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ is the digamma function. $\Psi(x)$ is a monotonically increasing function for $x > 0$, hence $\frac{df}{dx}(x;a,\theta)$ is negative for all $x > 0$ or might change its sign at only a root. Thus, the incomplete Gamma distribution has a unique mode at $n = 0$ if $\Psi(\theta+1) \leq \ln(a)$ or has one or two modes at consecutive integers n and $n+1$ such that $\Psi(n+\theta+1) \leq \ln(a)$ and $\Psi(n+\theta+2) \geq \ln(a)$. □

Based on this result, Lemma 3.2 and the algorithm given below it are then also valid for the incomplete Gamma distribution.

4.4. *Sequential and bridge sampling approaches*

Based on the integral representation (4.4) and the fact that F_t-processes are generated by an underlying squared Bessel process, we develop two types of path generation algorithms: sequential sampling and bridge sampling (see Fig. 2). (Note that here we refer explicitly to discounted asset price process F_t without absorption, but the same algorithms hold for the case of absorption and/or for S_t-processes after employing slight modifications as explained previously for simulating processes with absorption, like the martingale Bessel K-subfamily and the CEV model in case (2) of Sec. 2.2)

The first algorithm is rather natural and is based on the Markov property of the process and on Statement 4.1. Using the randomized gamma distribution of the second type, we simulate the underlying x_t-process with transition p.d.f. $\tilde{u}(x_i, x_{i-1}, \Delta t_i)$ at a discrete set of time points and obtain the values of F_t by applying the mapping $F(x)$. A pseudocode of the algorithm is as follows:

```
Let x(t_0) = X(F_0);
For i = 1 to N
    Generate:
        η_1^i ~ P( ρΔt_i + x(t_{i-1})/(2Δt_i) ),
        η_2^i ~ Bes( μ, sqrt(2ρx(t_{i-1})) ),
        x(t_i) ~ G( μ + η_1^i + 2η_2^i + 1, 1/(2Δt_i) );
    Let F_i = F(x(t_i))
Next i
```

The other method is analogous to the Brownian bridge sampling algorithm with the only exception that more complicated (yet still tractable) distributions are involved. Assume that the time-interval partition has $N = 2^k$ points, $k \geq 1$, and let a trajectory of the process be sampled at the time points in the following order of generation:

$$\begin{aligned} &\underbrace{t_N}, \underbrace{t_{N/2}}, \underbrace{t_{N/4}, t_{3N/4}}, \underbrace{t_{N/8}, t_{3N/8}, t_{5N/8}, t_{7N/8}}, \dots \\ &\dots, \underbrace{t_2, t_6, t_{10}, \dots, t_{N-2}}, \underbrace{t_1, t_3, \dots, t_{N-1}} \end{aligned} \tag{4.14}$$

where N is large enough such that all indices are integers. Note that this corresponds to the bridge sampling in Fig. 2.

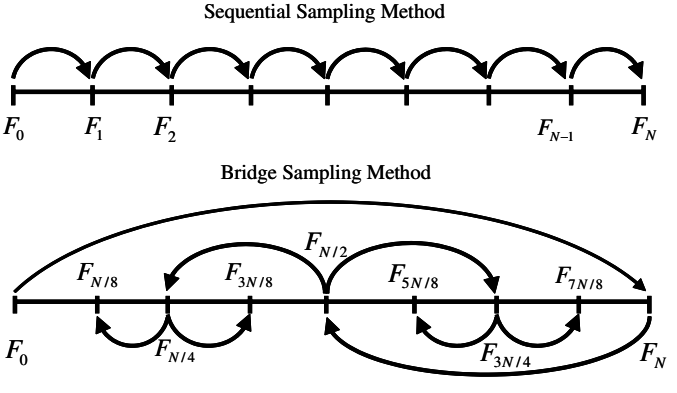

Fig. 2. Comparing two sampling algorithms for simulating asset price values.

Based on Statements 4.1 and 4.2, we have the following bridge sampling algorithm.

```
Let x(t_0) = X(F_0);
Generate:
    η_1^N ~ P( ρT + x(t_0)/(2T) ),
    η_2^N ~ Bes( μ, sqrt(2ρx(t_0)) ),
    x(t_N) ~ G( μ + η_1^N + 2η_2^N + 1, 1/(2T) );
Let F_N = F(x(t_N));
For l = 1 to k
    For m = 1 to 2^{l-1}
        Let i = (2m-1)N/2^l;
        Let τ_1 = t_{i-N/2^l}, τ = t_i, τ_2 = t_{i+N/2^l};
        Generate:
            η_1^i ~ P( 1/(2(τ_2-τ_1)) [ (τ_2-τ)/(τ-τ_1) x(τ_1) + (τ-τ_1)/(τ_2-τ) x(τ_2) ] ),
            η_2^i ~ Bes( μ, sqrt(x(τ_1)x(τ_2)) / (τ_2-τ_1) ),
            x(t_i) ~ G( μ + η_1^i + 2η_2^i + 1, (τ_2-τ_1)/(2(τ-τ_1)(τ_2-τ)) );
        Let F_i = F(x(t_i))
    Next m
Next l
```

In the case of uniform partition of the time interval, i.e., $\Delta t_i = \Delta t = T/N$, $i = 1, \ldots, N$, the formulas are simplified:

$$\eta_1^i \sim P\left(\frac{1}{4\Delta t}\left(x(\tau_1) + x(\tau_2)\right)\right),$$

$$\eta_2^i \sim Bes\left(\mu, \frac{\sqrt{x(\tau_1)x(\tau_2)}}{2\Delta t}\right),$$

$$x(t_i) \sim G\left(\mu + \eta_1^i + 2\eta_2^i + 1, \frac{1}{\Delta t}\right).$$

The bridge sampling algorithm allows us to simulate variables F_i in descending order of contribution to the variance of an estimate of the path integral. All other integration variables will add just slight details to the final path. Hence, we can concentrate on the distribution of the first set of variables.

Consider a decomposition of the set of integration variables $\{F_i\}_{i=1}^N$, arranged in the order (4.14), into two subsets containing $D = 2^l$, $1 \le l \le k$, and $N - D$ variables correspondingly:

$$\boldsymbol{F} = \boldsymbol{F}^1 \cup \boldsymbol{F}^2 = \{F_{N/D}, F_{2N/D}, \ldots, F_N\} \cup \{F_i \,|\, i \neq jN/D,\, 1 \le i \le N,\, j \in \mathbb{N}^+\}.$$

According to this partition, the N-dimensional integral (4.4) can be written as an integral with respect to $\boldsymbol{F}^1$ and $\boldsymbol{F}^2$:

$$\Theta = \int_{R_+^D} d\boldsymbol{F}^1 \boldsymbol{U}^1(\boldsymbol{F}^1; F_0, \mathcal{T}^1) \int_{R_+^{N-D}} d\boldsymbol{F}^2 \boldsymbol{U}^2(\boldsymbol{F}^2; \boldsymbol{F}^1, F_0, \mathcal{T}^2) \Lambda(\boldsymbol{F}^1, \boldsymbol{F}^2). \tag{4.15}$$

Multidimensional p.d.f.s $\boldsymbol{U}^1$ and $\boldsymbol{U}^2$ can be represented as products of one-dimensional p.d.f.s. The final expressions of $\boldsymbol{U}^1$ and $\boldsymbol{U}^2$ depend on what type of sampling algorithm is chosen. For example, sampling $\boldsymbol{F}^1$ sequentially and applying the bridge sampling method for $\boldsymbol{F}^2$, we have

$$\begin{aligned}
\boldsymbol{U}^1(\boldsymbol{F}^1; F_0, \mathcal{T}^1) &= \prod_{i=1}^{2^l} U(F_{iN/D}, F_{(i-1)N/D}, t_{iN/D}), \\
\boldsymbol{U}^2(\boldsymbol{F}^2; \boldsymbol{F}^1, F_0, \mathcal{T}^2) &= \prod_{j=l+1}^{k} \prod_{m=1}^{2^{j-1}} U_{F_{i_1(m)}, F_{i_2(m)}}\left(F_{i(m)}, t_{i(m)}, t_{i_1(m)}, t_{i_2(m)}\right),
\end{aligned}$$

where $i_1(m) = \frac{(2m-2)N}{2^j}$, $i(m) = \frac{(2m-1)N}{2^j}$, and $i_2(m) = \frac{2mN}{2^j}$. Note that one can also use different combinations of sequential and/or bridge sampling in the sets of first and second variables.

Consider some Monte Carlo estimate of the above integral and assume that the variables from $\boldsymbol{F}^1$ give much larger contributions to the variance of the random estimator in comparison with the variables from $\boldsymbol{F}^2$. If, in this case, the first coordinates account for at least $100 \cdot p_D\%$ of the variance, we say that *the integral has effective dimension D in proportion p_D in the truncation sense* [9, 22]. Clearly, p_D tends to one as D tends to N. As we show numerically below, the situation, where p_D is close to one even for $D \ll N$, is fairly typical for pricing path-dependent options. Usually, to obtain good accuracy, the time interval $[0, T]$ is discretized with high frequency, as for instance a one year interval might be divided into 250–260 parts accordingly to the number of business days per year. So, without the decomposition of variables it is rather difficult to improve the efficiency of evaluating the integral. Among the methods that can be used to improve the plain Monte Carlo method for high-dimensional integration we can distinguish stratified sampling, importance sampling, and quasi-Monte Carlo methods. All these techniques depend dramatically on the cardinality of the set of integration variables.

Before applying one of the techniques mentioned above, we change the region of integration of the first integral in (4.15). Let $\boldsymbol{\Phi}(\boldsymbol{F}^1; F_0, \mathcal{T}^1)$ be the joint c.d.f. corresponding to the p.d.f. $\boldsymbol{U}^1(\boldsymbol{F}^1; F_0, \mathcal{T}^1)$. By inverting this c.d.f. with respect to

$\boldsymbol{F}^1$ and changing variables $\boldsymbol{F}^1 = \boldsymbol{\Phi}^{-1}(\boldsymbol{\alpha}; F_0, \mathcal{T}^1)$, where $\boldsymbol{\alpha} \in [0,1)^D$, the region of integration with respect to $\boldsymbol{F}^1$ reduces to the D-dimensional unit hypercube:

$$\Theta = \int_{[0,1)^D} d\boldsymbol{\alpha} \int_{R_+^{N-D}} d\boldsymbol{F}^2 \boldsymbol{U}^2(\boldsymbol{F}^2; \boldsymbol{F}^1, F_0, \mathcal{T}^2) \Lambda(\boldsymbol{F}^1, \boldsymbol{F}^2). \tag{4.16}$$

where $\boldsymbol{F}^1 = \boldsymbol{\Phi}^{-1}(\boldsymbol{\alpha}; F_0, \mathcal{T}^1)$. This is best illustrated by a concrete example. Let $D = 4$ with bridge sampling used on $\boldsymbol{F}^1 = (F_{N/4}, F_{N/2}, F_{3N/4}, F_N)$. Fix $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) \in [0,1)^D$. According to the bridge sampling method, first obtain F_N by solving the equation $\alpha_1 = \int_0^{F_N} U(F, F_0, T) dF$. Then, by solving $\alpha_2 = \int_0^{F_{N/2}} U_{F_0, F_N}(F, T/2, 0, T) dF$, derive $F_{N/2}$. In the same manner, using the corresponding bridge p.d.f.s, obtain $F_{N/4}$ and $F_{3N/4}$ as functions of $\alpha_3, F_0, F_{N/2}$ and $\alpha_4, F_{N/2}, F_N$, correspondingly.

Since we have a problem of numerical integration over a unit high-dimensional hypercube, we have more familiar choices of methods for improving the efficiency of the Monte Carlo simulations.

5. Adaptive Monte Carlo Algorithms and Randomized Quasi Monte Carlo Methods

5.1. *Monte Carlo methods and generating pseudorandom numbers*

The Monte Carlo method is based on the Strong Law of Large Numbers. The Monte Carlo approximation of the integral (4.4) is obtained by the formula

$$\Theta \approx \widehat{\Theta}_n^{MC} = \frac{1}{n} \sum_{i=1}^{n} \Lambda(\boldsymbol{F}_{(i)}),$$

with $\boldsymbol{F}_{(i)} \in R_+^N$ denoting independent samples drawn from the joint p.d.f. $\boldsymbol{U}(\boldsymbol{F}; F_0, T)$. These samples can be obtained by using sequential and/or bridge sampling algorithms. The Strong Law of Large Numbers guarantees an almost certain convergence of the approximation as n tends to infinity. In addition to that, the Central Limit Theorem provides a rate of convergence:

$$\mathsf{P}\left\{ |\Theta - \widehat{\Theta}_n| \le \frac{\sigma}{\sqrt{n}} \gamma \right\} \approx \int_0^{\gamma} \frac{\sqrt{2}}{\sqrt{\pi}} e^{-x^2/2} dx,$$

with σ denoting the standard deviation of the random estimate $\Lambda(\boldsymbol{F})$ and γ is a positive real number. Choosing the parameter γ equal, for example, to one, two, or three, gives confidence intervals $\left(\widehat{\Theta}_n - \frac{\sigma}{\sqrt{n}}\gamma, \widehat{\Theta}_n + \frac{\sigma}{\sqrt{n}}\gamma \right)$, which cover the value Θ with probabilities 84.1%, 97.7%, and 99.9% respectively.

As a practical generator of pseudorandom numbers, we consider one of the congruential generators, based on a residue method of the form:

$$v_0 = 1, \quad v_n \equiv v_{n-1} M (\mathrm{mod}\, 2^r), \quad \alpha_n = v_n 2^{-r}. \tag{5.1}$$

Here $r = 128$ is a number of binary bits, used to represent a real number, $M = 5^{100109} (\mathrm{mod}\, 2^{128})$ is a sufficiently large number mutually disjoint from 2^r. This

generator was tested in [17, 24] and has period length $L_p = 2^{126} \approx 10^{38}$. One can distribute pseudorandom numbers among K processors. The sequence v_n is first split into subsequences of sufficiently large length L, which begin with the numbers v_{kL}, where $k = 1, 2, \ldots, K, \ldots$ is the identifying number of a processor in a cluster.

The initial values v_{kL} of subsequences can be obtained by the formula

$$v_{(k+1)L} \equiv v_{kL} M^L (\mathrm{mod}\, 2^r). \tag{5.2}$$

Thus, to run a Monte Carlo simulation on the kth processor, we use the subsequence of the residue method, which begins with $\alpha_{kL} = v_{kL} 2^{-r}$. The chosen value of the "jump" length for the auxiliary generator (5.2) is $L = 10^{26}$. This allows one to distribute the initial sequence by equal parts of length 10^{26} among about 10^{12} processors.

5.2. *Randomization of low discrepancy sequences*

Suppose that the integral (4.4) has been reduced to an integral over unit hypecube $[0,1)^d$. This is done, for example, by inverting the multi-dimensional cumulative distribution function $\Phi(\boldsymbol{F}; F_0, \mathcal{T})$ with respect to $\boldsymbol{F}$. In this case, the dimension $d = N$, and the integral (4.4) is written as follows:

$$\Theta = \int_{[0,1)^d} \Lambda\left(\boldsymbol{\Phi}^{-1}(\boldsymbol{\alpha}; F_0, \mathcal{T})\right) d\boldsymbol{\alpha}.$$

Another way consists of using some simulation routine $\boldsymbol{F} = \boldsymbol{f}(\boldsymbol{\alpha})$, which allows us to transform a bounded sequence of independent uniformly distributed (on $[0,1)$) variates $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_d)$ onto a random vector $\boldsymbol{F}$ having p.d.f. $\boldsymbol{U}(\boldsymbol{F}; F_0, \mathcal{T})$. Hence, the evaluation of Θ reduces to computing the expectation

$$\Theta = \mathsf{E}_0[\Lambda(\boldsymbol{f}(\boldsymbol{\alpha}))] = \int_{[0,1)^d} \Lambda(\boldsymbol{f}(\boldsymbol{\alpha})) d\boldsymbol{\alpha}.$$

Quasi-Monte Carlo methods approximate integrals over unit hypercubes using

$$\Theta \approx \widehat{\Theta}_n^{QMC} = \frac{1}{n} \sum_{i=1}^{n} \Lambda(\boldsymbol{f}(\boldsymbol{\alpha}^{(i)})),$$

for carefully and deterministically chosen points $\boldsymbol{\alpha}^{(1)}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(n)}$. Thus, to sample continuous uniform distributions on $[0,1)^d$, a low discrepancy sequence is used (see, for example, [19, 20, 25] for additional details). To obtain sample-based error estimates, a randomized version of quasi-Monte Carlo is considered (see, for instance, a survey in [18]). A procedure of randomization starts with a low discrepancy sequence $\boldsymbol{\alpha}^{(1)}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(n)} \in [0,1)^d$. A mapping from $[0,1)^d$ onto $[0,1)^d$ is chosen at random and applied to $\boldsymbol{\alpha}^{(i)}$. The result, denoted by $\boldsymbol{\alpha}_1^{(i)}, \boldsymbol{\alpha}_2^{(i)}, \ldots, \boldsymbol{\alpha}_m^{(i)}$, is the sequence of independent randomized versions of $\boldsymbol{\alpha}^{(i)}$. Using the randomized low

discrepancy sequence $\boldsymbol{\alpha}_j^{(1)}, \boldsymbol{\alpha}_j^{(2)}, \ldots, \boldsymbol{\alpha}_j^{(n)} \in [0,1)^d$, where $j = 1, 2, \ldots, m$, the quasi-Monte Carlo estimate $\widehat{\Theta}_n^{QMC}$ is evaluated. Then Θ is estimated as in the Monte Carlo method:

$$\Theta \approx \frac{1}{m} \sum_{j=1}^{m} \left(\frac{1}{n} \sum_{i=1}^{n} \Lambda \left(\boldsymbol{f}(\boldsymbol{\alpha}_j^{(i)}) \right) \right).$$

For low discrepancy sampling we use a Sobol's sequence based on [7]. Two types of randomizations are used here: random shift modulo one [18] and a linear matrix scrambling [26]. Random shift modulo 1 of $\boldsymbol{\alpha}$ takes the form $\boldsymbol{\alpha} \to \boldsymbol{\alpha} + \boldsymbol{\beta} (\mathrm{mod}\, 1)$, where $\boldsymbol{\beta} \in [0,1)^d$ is formed from independent uniformly distributed variates.

Digital scrambles of $[0,1)$ obtain by randomizing the digits of α in an integer base $b \geq 2$. For $\alpha \in [0,1)$ write $\alpha = \sum_{k=1}^{\infty} a_k b^{-k}$. Scrambling of α results in a number $\beta = \sum_{k=1}^{\infty} b_k b^{-k}$, where the digits $b_k \in \{0, 1, \ldots, b-1\}$ are obtained by some permutation scheme applied to the digits of α. An affine matrix scrambling of $\alpha = \sum_{k=1}^{\infty} a_k b^{-k}$ takes the form $\sum_{k=1}^{\infty} a_k b^{-k} \to \sum_{k=1}^{\infty} b_k b^{-k}$, where each b_k is given by $b_k = \sum_{j=1}^{k} M_{kj} a_j + C_k (\mathrm{mod}\, b)$ for random elements M_{kj} and C_k from $\{0, 1, \ldots, b-1\}$. The structure of the scrambling can be described by the structure of the matrix M. In particular, we use Tezuka's [31] i-binomial scrambling, which corresponds to the matrix

$$\begin{pmatrix} h_1 & 0 & 0 & 0 & \cdots \\ g_2 & h_1 & 0 & 0 & \cdots \\ g_3 & g_2 & h_1 & 0 & \cdots \\ g_4 & g_3 & g_2 & h_1 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix},$$

where a matrix element, uniformly distributed in $\{0, 1, \ldots, b-1\}$, is denoted by a symbol g, and a matrix element, uniformly distributed in $\{1, \ldots, b-1\}$, is represented by a symbol h. All elements h_1, g_2, g_3, and so forth, are independent and uniformly distributed over their ranges. Notice that for $b = 2$ the scrambling matrix has units on the principal diagonal, and zeros and units below the diagonal.

Owen [26] showed that for any $\alpha \in [0,1)$ the scrambled version of α in the base b under affine matrix scrambling is uniformly distributed in $[0,1)$. Applying independent digital scrambles to each coordinate of $\boldsymbol{\alpha} \in [0,1)^d$, we obtain a uniformly distributed point in $[0,1)^d$. Thus, the use of scrambling provides unbiased estimates of a multivariate integral.

5.3. *Adaptive algorithms MISER and VEGAS*

Consider again the problem of multidimensional integration over a unit hypercube. The MISER integration routine [28, 29] is based on the adaptive stratified sampling. This method smooths down many irregularities in the distribution of random points.

Suppose that we partition the region of integration into K non-overlapping strata $A_1, A_2, \ldots, A_K$. So, the expectation can be computed as follows:

$$\mathsf{E}[\Lambda(\boldsymbol{f}(\boldsymbol{\alpha}))] = \sum_{i=1}^{K} \mathsf{P}\{\boldsymbol{\alpha}^{(i)} \in A_i\} \mathsf{E}[\Lambda(\boldsymbol{f}(\boldsymbol{\alpha}^{(i)}))|\boldsymbol{\alpha}^{(i)} \in A_i].$$

In random sampling we generate uniformly distributed $\boldsymbol{\alpha}^{(1)}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(K)}$ in the corresponding stratum $A_1, A_2, \ldots, A_K$. Moreover, we decide in advance what fraction of the samples should be drawn from each stratum A_i. It is possible to derive optimal allocation in order to minimize the variance of the estimator (see, for example, [20]).

The main problem with stratified sampling is to avoid exponential explosion of the number of stratum with growing dimensionality of the region. A technique called recursive stratified sampling [28, 29] attempts to do this by successive bisection of a region, not along all d dimensions, but rather along only one dimension at a time.

The VEGAS algorithm, invented by Lepage [23, 29], is a Monte Carlo integration routine applying adaptive importance sampling. Due to the identity

$$\int_{[0,1)^d} \Lambda(\boldsymbol{f}(\boldsymbol{\alpha}))d\boldsymbol{\alpha} = \int_{[0,1)^d} \frac{\Lambda(\boldsymbol{f}(\boldsymbol{\alpha}))}{p(\boldsymbol{\alpha})} dP(\boldsymbol{\alpha}),$$

with P denoting some continuous probabilistic measure in $[0,1)^d$ with the probability density function $p(\cdot)$, the integral can be estimated by the sample mean

$$\frac{1}{n} \sum_{i=1}^{n} \frac{\Lambda(\boldsymbol{f}(\boldsymbol{\alpha}^{(i)}))}{p(\boldsymbol{\alpha}^{(i)})},$$

with $\boldsymbol{\alpha}^{(i)}$ denoting n independent P-distributed sample variates.

The basic idea in the VEGAS algorithm is to construct, adaptively, a multidimensional p.d.f. as a product of one-dimensional step functions, i.e.,

$$p(\boldsymbol{\alpha}) = p(\alpha_1, \alpha_2, \ldots, \alpha_d) = p_1(\alpha_1)p_2(\alpha_2)\cdots p_d(\alpha_d).$$

Each p_i, $i = 1, \ldots, d$, has a fixed number K of steps with the area below each step equal to $1/K$. The optimal separable density function can be shown to be

$$p_1(\alpha_1) \propto \sqrt{\int d\alpha_2 \cdots \int d\alpha_d \frac{(\Lambda(\boldsymbol{f}(\alpha_1, \alpha_2, \ldots, \alpha_d)))^2}{p_2(\alpha_2)\cdots p_d(\alpha_d)}},$$

and $p_2(\alpha_2), \ldots, p_d(\alpha_d)$ are defined similarly.

6. Numerical Results

In this section we report various numerical results on pricing Asian and lookback options.

6.1. *Test setup*

In the following numerical tests, uniformly distributed variates are obtained using the congruential generator (5.1). Normal, gamma and Poisson random variates are generated using the Alan J. Miller library, which is downloadable from `http://users.bigpond.net.au/amiller`. To compute the inverse of the cumulative randomized gamma distribution function, we solve the algebraic-differential system of equations:

$$\begin{cases} \Phi_i'(y) = f_i(y), \\ \Phi_i(0) = 0, \\ \Phi_i(y) = \Phi_i^*, \end{cases}$$

where f_i, $i = 1, 2$, are the p.d.f.s of the randomized gamma distributions of the first or second types respectively. $\Phi_i(y) = \int_0^y f_i(x)dx$, $i = 1, 2$, are the corresponding c.d.f.s and Φ_i^* is a given value from $[0, 1)$. To solve ordinary differential equations we use a method of the sixth order of accuracy with variable integration step and controlled accuracy. This particular procedure is relatively slow, in comparison, for instance, with the computation of the inverse of the normal c.d.f. Using some faster routine could also give greater efficiency gain to the algorithms than what we find below. For low discrepancy sampling we use a Sobol's sequence based on [7], and the generating algorithm is downloadable from `http://www.netlib.org/toms/659`.

6.2. *Path-dependent options*

We start with a common example of an Asian-style call option. Here, a payoff of an Asian-style option depends on the arithmetic average of the underlying asset values: $A_N = \frac{1}{N}\sum_{i=1}^N F_i$. For an *average price call option*, the option holder receives at expiration time T the payoff $\Lambda_{AP}^C = (A_N - K)_+$, where K is the strike price of the option. This type of Asian option is also called fixed strike Asian call.

For pricing this type of option under the Bessel family of state dependent volatility models, an effective control variate method was suggested in [10], where the average of European-style call options with the payoff

$$\Lambda_{AE}^C = \frac{1}{N}\sum_{i=1}^N (F_i - K)_+,$$

is used as a control variate. The second type of Asian call option that we consider is an *average strike call option*, or a floating-strike Asian call with payoff $\Lambda_{AS}^C = (F_N - A_N)_+$.

Lookback options are another example of path-dependent options that we consider. Their payoffs depend on the maximum M_T or minimum m_T values of underlying asset prices attained during the option's life $[t_0, T]$. For discrete-time versions of these options, the following quantities are used: $M_N = \max\limits_{i=0,\ldots,N} F_i$ and

$m_N = \min_{i=0,\ldots,N} F_i$. A *standard lookback call* gives the right to buy at the lowest price recorded during the option's life, so the payoff to a holder is $\Lambda^C_{LB} = F_N - m_N$. A *standard lookback put* gives the right to sell at the highest price recorded during the option's life, consequently $\Lambda^P_{LB} = M_N - F_N$.

6.3. *Valuing path-dependent options on the Bessel family and CEV diffusion models*

We value average price, average strike, and lookback call options using a bridge Monte Carlo algorithm based on generating randomized gamma distributions. As competitive approaches, we test hybrid adaptive algorithms, using the randomized quasi-Monte Carlo (RQMC) method and adaptive Monte Carlo routines such as VEGAS and MISER.

As described in Sec. 4.4, we split the path integral into two subspaces. For integration over $\boldsymbol{F}^1$ we use RQMC or adaptive routines, whereas for integration over $\boldsymbol{F}^2$, the bridge MC technique is applied. Thus, we have a bunch of hybrid algorithms. MCMB is the bridge Monte Carlo method as it was described above. RQMCM is the randomized quasi-Monte Carlo method which uses digital scrambling for the randomization. VEGAS and RQMCVEG are adaptive routines, which use pseudo-random numbers and randomized Sobol's sequences correspondingly. Finally, there is the MISER adaptive routine, which utilizes pseudorandom numbers only.

Suppose, for some type of option, the bridge MC method gives an estimate of the path integral with standard error σ_M for the time $T_M = nt_M$. For an alternative method we have σ_A and $T_A = nt_A$ accordingly, where t_M and t_A denote average times, required to simulate one path, and n is the number of sample paths. The efficiency gain E of the alternative method to the plain MC method is then:

$$E = \frac{\sigma_M^2 T_M}{\sigma_A^2 T_A} = \frac{\sigma_M^2 t_M}{\sigma_A^2 t_A} \approx \frac{\hat{\sigma}_M^2 t_M}{\hat{\sigma}_A^2 t_A}. \tag{6.1}$$

In actual computations, we use a sample estimation $\hat{\sigma}$ instead of an exact value of the standard error σ. Note that a slightly more accurate approach involves the statistical F-distribution to compute a confidence interval on the ratio of variances.

All the above algorithms, except MCMB, involve numerical inversion of the randomized gamma c.d.f. Hence, they demand considerably more time for computations compared to the bridge MC method. Splitting the path integral into D-dimensional and $(N - D)$-dimensional subspaces reduces this shortage, since the algorithm of numerical inversion of a c.d.f. is used for only $D \ll N$ variables.

To compare competitive approaches, we use the same number of sample paths. If, in the case of randomized QMC, we generate m randomized series of Sobol's sequences with n points, then for MC methods we simulate nm independent paths. In what follows, we simulate $m = 100$ independent randomized Sobol's sequences.

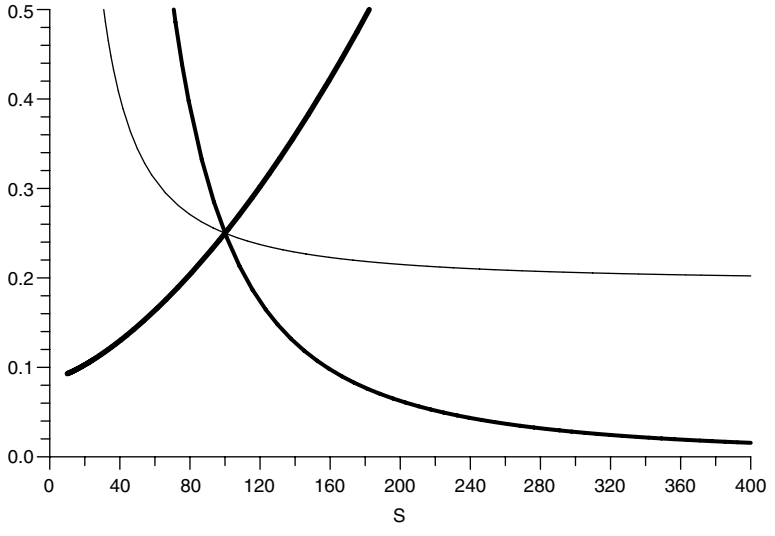

Fig. 3. Plots of local volatility $\sigma_{loc}(S)$ versus S computed for three state-dependent models: the Bessel I-subfamily using (2.14) with $a_I = 24.5302$, $\rho = 0.001$, and $\mu = 0.5$ (the thickest line); the Bessel K-subfamily using (2.14) with $a_K = 111.4760$, $\rho = 0.005$, and $\mu = 0.5$ (the thinnest line); and the CEV model (case 2 of Sec. 2.2) with $\sigma_{loc}(S) = \delta S^{\beta}$, $\delta = 2500$, $\beta = -2$.

In computational tests, three models with different volatility functions are considered (see Fig. 3). We fixed the value of the annual local volatility function $\sigma_{loc}(S_0) = 0.25$ at the initial asset price $S_0 = F_0 = 100$. The strike price is $K = 100$.

Example 6.1. Comparison of adaptive MC and RQMC methods. We start with the Bessel I-subfamily model. As discussed in Secs. 2.1 and 2.2, this is a model for a strict supermartingale in the forward price process whereas for the other types of models considered below (i.e., the Bessel K-subfamily and CEV (case 2) models) the forward price is a martingale under the assumed risk-neutral measure.[3] By varying the Bessel order parameter μ, different shapes of the local volatility function are obtained (see Fig. 1). For example, taking μ as 0.25, 0.5 or 1.5 we obtain, respectively, concave, straight (quadratic model) or convex plots (see [10]). In our consideration, we fixed $\rho = 0.001$ and varied a_I and μ: $\mu = 0.25$, $a_I = 24.5302$ (model 1), $\mu = 0.5$, $a_I = 17.7323$ (model 2), and $\mu = 1.5$, $a_I = 5.0574$ (model 3). For all three cases we obtained a similar numerical convergence. Hence, to give a relative comparison of the methods as applied to the Bessel I-subfamily, we present an expanded outcome for model 1 only. The number of asset price observations is $N = 128$, the interest rate $r = 0$, and all options have one year to expiration: $T = 1.0$.

[3]The reader should carefully note that the risk-neutral pricing formula is formally used here to compute the expectation integrals for the Bessel I-subfamily for the sole purpose of demonstrating the relative performance of the simulation methods.

Table 1. Comparison of the hybrid methods.

	Dimension					
	$D = 4$		$D = 8$		$D = 16$	
Average Price Call						
MCM			5.7515 ± 0.0121	(1.0)		
MCM+CV			5.7761 ± 0.0013	(92.9)		
MISER	5.7817 ± 0.0039	(6.8)	5.7820 ± 0.0042	(4.8)	5.7726 ± 0.0043	(3.1)
VEGAS	5.7750 ± 0.0035	(8.6)	5.7771 ± 0.0038	(5.7)	5.7716 ± 0.0043	(3.2)
RQMCVEG	5.7754 ± 0.0016	(39.1)	5.7785 ± 0.0019	(24.0)	5.7775 ± 0.0024	(10.4)
RQMCM	5.7793 ± 0.0020	(27.3)	5.7761 ± 0.0012	(56.7)	5.7805 ± 0.0010	(60.0)
RQMCM+CV	5.7792 ± 0.0006	(251.9)	5.7786 ± 0.0004	(566.8)	5.7786 ± 0.0006	(270.8)
Average Strike Call						
MCM			5.8389 ± 0.0159	(1.0)		
MISER	5.8222 ± 0.0042	(10.4)	5.8211 ± 0.0044	(7.5)	5.8274 ± 0.0045	(5.0)
VEGAS	5.8253 ± 0.0035	(15.1)	5.8254 ± 0.0037	(10.5)	5.8246 ± 0.0042	(5.7)
RQMCVEG	5.8254 ± 0.0016	(69.2)	5.8232 ± 0.0015	(66.6)	5.8257 ± 0.0025	(16.1)
RQMCM	5.8264 ± 0.0030	(19.9)	5.8274 ± 0.0023	(27.8)	5.8238 ± 0.0024	(17.5)
Standard Lookback Call						
MCM			16.5754 ± 0.0247	(1.0)		
MISER	16.5308 ± 0.0098	(4.6)	16.5254 ± 0.0091	(4.2)	16.5264 ± 0.0091	(3.0)
VEGAS	16.5253 ± 0.0084	(6.2)	16.5234 ± 0.0093	(4.0)	16.5203 ± 0.0109	(2.1)
RQMCVEG	16.5134 ± 0.0049	(18.6)	16.5174 ± 0.0046	(16.7)	16.5146 ± 0.0074	(4.5)
RQMCM	16.5217 ± 0.0038	(31.2)	16.5243 ± 0.0036	(27.6)	16.5228 ± 0.0029	(29.6)
Standard Lookback Put						
MCM			22.9171 ± 0.0202	(1.0)		
MISER	22.9788 ± 0.0181	(0.9)	22.9733 ± 0.0170	(0.8)	22.9899 ± 0.0185	(0.5)
VEGAS	22.9647 ± 0.0134	(1.7)	22.9541 ± 0.0143	(1.1)	22.9596 ± 0.0156	(0.7)
RQMCVEG	22.9456 ± 0.0086	(4.1)	22.9474 ± 0.0075	(4.2)	22.9464 ± 0.0084	(2.3)
RQMCM	22.9661 ± 0.0124	(1.9)	22.9425 ± 0.0093	(2.7)	22.9735 ± 0.0088	(2.1)

Note: Average price call, average strike call, standard lookback call and put, and average of standard European calls are evaluated under the Bessel I-subfamily with $\rho = 0.001$, $a_I = 24.5302$, $\mu = 0.25$. The value of sample standard error is given after $\pm$ sign. The value of achieved speedup in comparison with the standard Monte Carlo estimate is given in brackets.

Table 1 compares results of different adaptive MC and RQMC methods. We employed three different values of the dimension D for adaptive methods: $D = 4, 8, 16$. For the quasi-Monte Carlo method, we used a Sobol's sequence containing $n = 8191$ points. Since we generated $m = 100$ independent replicates of randomization, for Monte Carlo algorithms, $mn = 819100$ simulations were run. Each table entry gives the sample value, the sample standard error (after $\pm$ sign), and the efficiency improvement factor (in round brackets) using (6.1).

Example 6.2. Valuing Path-Dependent Options on the Bessel *K*-subfamily. Here, we consider the Bessel K-subfamily with parameters: $\rho = 0.005$, $a_K = 111.4761$, and $\mu = 0.25$ (or $\lambda = 2.5$). When $F \to \infty$, $\sigma_{loc}(F) \to 0.2$. As $F \to 0$, then $\sigma_{loc}(F) \to (constant)F^{-2}$. The number of asset price observations is set to $N = 128$, the interest rate is simply set to zero, $r = 0$, and all options have six months to expiration: $T = 0.5$.

Table 2. Comparison of the RQMC algorithm over MC for the Bessel K-subfamily.

	Number of Quasi-Random Points					
	8191		16383		32767	
Average Price Call (Control Variate Method)						
$D = 4$	4.095175 ± 0.000435	(3.6)	4.094867 ± 0.000290	(4.0)	4.095025 ± 0.000199	(4.3)
$D = 8$	4.095210 ± 0.000288	(4.5)	4.094554 ± 0.000244	(3.1)	4.094985 ± 0.000136	(5.0)
$D = 16$	4.094942 ± 0.000233	(3.1)	4.094542 ± 0.000164	(3.1)	4.094711 ± 0.000078	(6.8)
$D = 32$	4.094873 ± 0.000294	(0.6)	4.094635 ± 0.000220	(0.5)	4.094861 ± 0.000229	(0.3)
MCM	4.093448 ± 0.000879	(1.0)	4.093875 ± 0.000621	(1.0)	4.094296 ± 0.000439	(1.0)
Average of European Calls (exact value is 4.727415)						
$D = 4$	4.726402 ± 0.001002	(40)	4.728040 ± 0.000666	(45)	4.727929 ± 0.000468	(46)
$D = 8$	4.726769 ± 0.000480	(95)	4.727742 ± 0.000345	(92)	4.727100 ± 0.000223	(110)
$D = 16$	4.727071 ± 0.000255	(152)	4.727742 ± 0.000223	(99)	4.727652 ± 0.000131	(144)
$D = 32$	4.727878 ± 0.000170	(108)	4.727462 ± 0.000125	(100)	4.727559 ± 0.000107	(68)
MCM	4.742780 ± 0.006766	(1)	4.732532 ± 0.004782	(1)	4.732841 ± 0.003383	(1)
Average Strike Call						
$D = 4$	4.052490 ± 0.000960	(37)	4.050282 ± 0.000537	(66)	4.050092 ± 0.000462	(46)
$D = 8$	4.050777 ± 0.000542	(80)	4.050840 ± 0.000391	(86)	4.050383 ± 0.000249	(106)
$D = 16$	4.051245 ± 0.000300	(117)	4.050282 ± 0.000234	(98)	4.049959 ± 0.000131	(160)
$D = 32$	4.050647 ± 0.000327	(30)	4.050237 ± 0.000303	(17)	4.050028 ± 0.000277	(11)
MCM	4.051054 ± 0.007081	(1)	4.048529 ± 0.005001	(1)	4.049127 ± 0.003541	(1)
Standard Lookback Call						
$D = 4$	12.740585 ± 0.002442	(18)	12.742700 ± 0.001826	(18)	12.738696 ± 0.001105	(25)
$D = 8$	12.738265 ± 0.001685	(26)	12.741012 ± 0.001096	(34)	12.739561 ± 0.000959	(23)
$D = 16$	12.740134 ± 0.001319	(19)	12.740942 ± 0.000922	(20)	12.739400 ± 0.000616	(23)
$D = 32$	12.741110 ± 0.001362	(5)	12.739901 ± 0.000974	(5)	12.740564 ± 0.000721	(5)
MCM	12.750070 ± 0.012553	(1)	12.742361 ± 0.008871	(1)	12.742163 ± 0.006277	(1)

Note: Average price call, average strike call, standard lookback call, and average of standard European calls are evaluated under the Bessel K-subfamily with $\rho = 0.005$, $a_K = 111.4761$, $\mu = 0.25$. The value of sample standard error is given after $\pm$ sign. The value of achieved speedup in comparison with the standard Monte Carlo estimate is given in brackets.

Table 2 reports the comparison of results for the randomized QMC method only, which were obtained by varying the number of quasi-random points n and integral dimension D. The notations are consistent with those mentioned above.

Example 6.3. Valuing Path-Dependent Options on the CEV Model. For the numerical illustration of the case with the CEV model, we take similar data as for the previous test. The interest rate is now chosen as $r = 0.1$ per annum and all options on the underlying have six months to expiration: $T = 0.5$. As was pointed out in [13], typical values of the CEV elasticity parameter β implicit in the S&P 500 stock index option prices are strongly negative. So we take $\beta = -2$. The last parameter δ is taken in such a way as to satisfy the condition $\delta S_0^{\beta} = 0.25$. For $\beta = -2$, we have $\delta = 2500$.

As is known, a discretely monitored exotic option depending on the maximal or/and minimal underlying asset price reached during the lifetime of the option

Table 3. Lookback-style option prices for the CEV model: comparison of RQMC estimates of discretely monitored option prices for different numbers N of observations (with standard errors given in brackets) and extrapolated estimates of continuously monitored option prices with the exact values.

Option Type	$N = 256$	$N = 512$	$N = 1024$	$N = 2048$	Extrapolation	Exact
Lookback Call	16.2613	16.4849	16.6381	16.7473	17.0026	17.0049
	(0.0012)	(0.0012)	(0.0013)	(0.0012)		
Call on Max	14.8597	15.0097	15.1173	15.1935	15.3800	15.3807
	(0.0009)	(0.0009)	(0.0009)	(0.0008)		
Lookback Put	9.8754	10.0333	10.1452	10.2234	10.4129	10.5036
	(0.0009)	(0.0009)	(0.0009)	(0.0008)		
Put on Min	11.3343	11.5373	11.6803	11.7844	12.0349	12.1278
	(0.0012)	(0.0012)	(0.0013)	(0.0013)		

approximates its continuously monitored analogue with order $O(1/\sqrt{N})$ (see [8]) where N is the number of observations. This is because when we simulate the process at discrete points we lose some information about the part of the path between observation dates. For the lognormal model this discretization error is corrected by simulating maximum (or minimum) values of the underlying asset price over each time interval $[t_i, t_{i+1}]$, $i = 0, 1, \ldots, N-1$, conditional upon S_i and S_{i+1} (see [4]). The other, more general but less accurate, approach is based on extrapolating the option price. Here, our aim is to compare estimates of discretely monitored path-dependent options with, wherever possible, exact values of their continuous analogues. As a reference we use [13], where a variety of closed form expressions for values of different exotic options under the CEV model were obtained. We price path-dependent options for different values of N and use the regression method for extrapolation. We follow [8] in assuming that the price of a discrete lookback-style option V_N with N observations can be successfully fitted by the model function $f(\Delta t) = a + b\sqrt{\Delta t} + c\Delta t$ where $\Delta t = T/N$. Hence the value $f(0) \equiv a$ can be used as an approximation to the price V_∞ of a continuously monitored option.

We made $m = 200$ independent randomizations for the QMC method with $n = 8{,}191$ quasi-random points, using dimension $D = 16$. Table 3 provides estimates of discrete path-dependent option prices for different values of N, and compares our extrapolated values for continuous option prices with the exact values provided in [13]. The current minimum $S^0_{\min}$ and maximum $S^0_{\max}$ values of an asset price is assumed to be equal to spot $S_0 = 100$.

6.4. *Further analysis and concluding remarks*

As observed in Tables 1 and 2, the numerical results clearly show an efficiency improvement for all hybrid adaptive algorithms as compared with the plain MC method. This means that the bridge sampling algorithm allows us to transfer a large part of the variance to the first few random variables when evaluating path integrals for Asian style options. As we can see, further enlargement of dimension

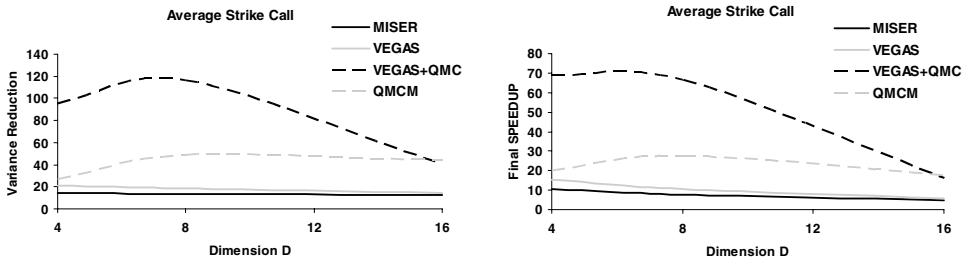

Fig. 4. Comparing rates of variance reduction and of speedup achieved with the use of different computational methods for pricing average strike call options under the Bessel I-subfamily.

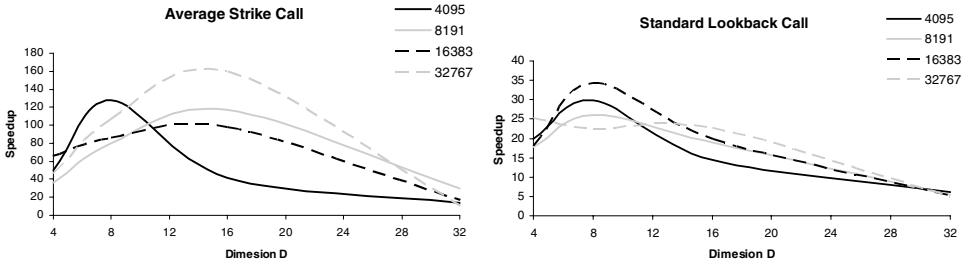

Fig. 5. Comparing rates of computational cost reduction achieved with the use of the randomized quasi-Monte Carlo method with different sample volumes for pricing average strike and lookback call options under the Bessel K-subfamily.

D in adaptive procedures has no significant effect on the variance reduction and, moreover, can impair the efficiency of algorithms because of the increase in computational time.

Randomized QMC adaptive methods excel MC adaptive algorithms such as MISER and VEGAS, as observed in Fig. 4. However, in some tests for smaller values of D, RQMCVEG, which combines importance sampling and randomized QMC, surpasses plain RQMC, but as D increases the closeness of the sampling p.d.f. to the integrand is decreased. If there is an effective control variate method, as in the case of the average price call option, the combination of this method with randomized QMC gives considerable efficiency improvement. Nevertheless, we can suppose that further development of our approach consists in more substantial combination of randomized QMC with classical methods of variance reduction. In the case of importance sampling, we might use piecewise constant or piecewise linear density functions, which can be more close to the integrand. Unfortunately, such adaptive methods as importance sampling and stratified sampling depend on the type of option very considerably, and usually cannot be used for the efficient simultaneous evaluation of different options in contrast to usual MC and RQMC methods.

Table 4. Efficiency improvement factors for RQMC algorithm, averaged over all options considered.

	Number of Quasi-Random Points			
	8191	16383	32767	65535
$D = 4$	0.71	0.97	0.71	0.87
$D = 8$	1.22	1.28	1.37	1.77
$D = 16$	1.20	0.96	1.47	3.05
$D = 32$	0.99	0.94	0.64	1.04

As observed in Fig. 5, comparative results for the RQMC algorithm obtained for different numbers of quasi-random points n and values of dimension D show no essential efficiency improvement with increasing D. Changing the dimension D profoundly influences the gain. Hence, the factors of efficiency improvement for all models reached their maximal values at $D = 8, 16$ and decreased with the following subsequent rise in D. Table 4 summarizes the efficiency improvements. Here, we calculated final improvement factors for the RQMC algorithm by averaging the achieved speedup over all path-dependent options considered.

Acknowledgments

The first author acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) for a discovery research grant as well as SHARCNET (the Shared Hierarchical Academic Research Computing Network) in providing support for a research chair in financial mathematics as well as for a postdoctoral fellowships grant. The second author is a SHARCNET postdoctoral fellow.

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).
[2] C. Albanese, G. Campolieti, P. Carr and A. Lipton, Black-Scholes goes hypergeometric, *Risk* **14** (2001) 99–103.
[3] C. Albanese and G. Campolieti, *Advanced Derivatives Pricing and Risk Management* (Academic Press, Elsevier Science, USA, 2005).
[4] L. Andersen and R. Brotherton-Ratcliffe, Exact exotics, *Risk* **9**(10) (1995) 85–89.
[5] P. P. Boyle and Y. Tian, Pricing lookback and barrier options under the CEV process, *Journal of Financial and Quantitative Analysis* **34** (1999) 241–264. (Correction: P. P. Boyle, Y. Tian and J. Imai, Lookback options under the CEV process: A correction, JFQA website, http://depts.washington.edu/jfqa/ in Notes, Comments, and Corrections).
[6] A. N. Avramidis, P. L'Ecuyer and P. A. Tremblay, Efficient simulation of gamma and variance-gamma processes, *Proceedings of the* 2003 *Winter Simulation Conference* (2003) 319–326.
[7] P. Bratley and B. L. Fox, Algorithm 659: Implementing Sobol's quasirandom sequence generator, *ACM Transaction on Mathematical Software* **14** (1988) 88–100.

[8] M. Broadie, P. Glasserman and S. G. Kou, Connecting discrete and continuous path-dependent options, *Finance and Stochastic* **3** (1999) 55–82.
[9] R. E. Caflisch, W. Morokoff and A. Owen, Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimension, *The Journal of Computational Finance* **1** (1997) 27–46.
[10] G. Campolieti and R. Makarov, Path Integral pricing of Asian options on state dependent volatility models, to appear in *Quantitative Finance.*
[11] A. M. G. Cox and D. G. Hobson, Local Martingales, bubbles and option prices, *Finance and Stochastics* **9** (2005) 477–492.
[12] J. Cox, Notes on option pricing I: Constant elasticity of variance diffusions, Working paper, Stanford University (1975). (Reprinted in *Journal of Portfolio Management* **22** (1996) 15–17).
[13] D. Davydov and V. Linetsky, Pricing and hedging path-dependent options under the CEV process, *Management Science* **47** (2001) 949–965.
[14] V. Linetsky, The spectral decomposition of the option value, *International Journal of Theoretical Applied Finance* **7**(3) (2004) 337–384.
[15] L. Devroye, A simple generator for discrete log-concave distributions, *Computing* **39** (1987) 87–91.
[16] L. Devroye, Simulating Bessel random variables, *Statistics and Probability Letters* **57** (2002) 249–257.
[17] I. G. Dyadkin and K. G. Hamilton, A study of 128-bit multipliers for congruential pseudorandom number generators, *Computer Physics Communications* **125** (2000) 239–258.
[18] P. L'Ecuyer and C. Lemieux, Recent advances in randomized quasi-Monte Carlo methods, in *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, eds. M. Dror, P. L'Ecuyer and F. Szidarovszki (Kluwer Academic Publishers, Boston, 2002), pp. 419–474.
[19] B. L. Fox, *Strategies for Quasi-Monte Carlo* (Kluwer Academic, Boston, 1999).
[20] P. Glasserman, *Monte Carlo Methods in Financial Engineering* (Springer-Verlag, New York, 2004).
[21] D. Goldenberg, A unified method for pricing options on diffusion processes, *J. Financial Econom.* **29** (1991) 3–34.
[22] F. J. Hickernell, A generalized discrepancy and quadrature error bound, *Mathematics of Computation* **67** (1998) 299–322.
[23] G. P. Lepage, A new algorithm for adaptive multidimensional integration, *Journal of Computational Physics* **27** (1978) 192–203.
[24] G. A. Mikhailov and M. A. Marchenko, Parallel realization of statistical simulation and random number generators, *Russian Journal of Numerical Analysis and Mathematical Modelling* **17**(1) (2002) 113–124.
[25] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods* (S.I.A.M., Philadelphia, 1992).
[26] A. B. Owen, Variance and discrepancy with alternative scramblings, *ACM Transactions of Modeling and Computer Simulation* **13** (2003) 363–378.
[27] J. Pitman and M. Yor, A decomposition of Bessel bridges, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **59** (1982) 425–457.
[28] W. H. Press and G. R. Farrar, Recursive stratified sampling for multidimensional Monte Carlo integration, *Computers in Physics* **4** (1990) 190–195.
[29] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in Fortran 77* (Cambridge University Press, 1992).

[30] C. Ribeiro and N. Webber, Valuing path-dependent options in the variance-gamma model by Monte Carlo with a gamma bridge, *Journal of Computational Finance* **7**(2) (Winter 2003/04) 81–100.
[31] S. Tezuka and H. Faure, I-binomial scrambling of digital nets and sequences, *Journal of Complexity* **19** (2003) 744–757.
[32] L. Yuan and J. D. Kalbfleisch, On the Bessel distribution and related problems, *Annals of the Institute of Statistical Mathematics* **52**(3) (2000) 438–477.