DEPT. OF MATH./CMA UNIVERSITY OF OSLO
PURE MATHEMATICS NO 4
ISSN 0806–2439 APRIL 2013

APPROXIMATING LÉVY SEMISTATIONARY PROCESSES VIA FOURIER METHODS IN THE CONTEXT OF POWER MARKETS

FRED ESPEN BENTH, HEIDAR EYJOLFSSON, AND ALMUT E. D. VERAART

Abstract. The present paper discusses Lévy semistationary processes in the context of power markets. A Fourier simulation scheme for obtaining trajectories of these processes is discussed and its rate of convergence is analysed. Finally we put our simulation scheme to work for simulating the price of path dependent options.

1. Introduction

A Lévy semistationary processes is a continuous time integral process driven by a Lévy process and a deterministic kernel function on the entire real line which is modulated by stochastic volatility. Lévy semistationary processes were initially suggested in the context of modelling turbulence (see Barndorff-Nielsen and Schmiegel [4, 5]) and have also been employed to model tumor growth (see Jónsdóttir et al. [16]). Moreover, in Barndorff-Nielsen, Benth and Veraart [2] Lévy semistationary processes (and more generally volatility modulated Lévy-driven Volterra processes) are introduced as a general modelling framework for electricity spot prices.

It is well known that electricity spot markets display a strong mean reversion effect. This means that at times when the spot price is high, the price is pushed down by lowered demand and conversely during periods of low price, the price is pushed up by increased demand. Modelling commodity spot by means of *mean-reverting* processes has been studied by Schwartz [21] and more generally by means of a a sum of Lévy process driven Ornstein-Uhlenbeck processes, corresponding to different mean-reversion coefficients, has been studied in Benth, Kallsen and Meyer-Brandis [8] and Klüppelberg, Meyer-Brandis and Schmidt [17]. The idea behind employing Lévy semistationary processes in the electricity spot setting is to generalize further to processes which mean revert in the *weak* probabilistic sense, i.e. by being stationary. Whenever a Lévy semistationary process is modulated by a stationary volatility process, the Lévy semistationary process is stationary, hence the terminology semistationary. Thus a Lévy semistationary processes which is

Date: April 15, 2013.

Financial support from the Norwegian Research Council of the eVita project 205328 "Energy Markets: modeling, optimization and simulation" (Emmos) is greatly acknowledged by F. E. Benth. A. E. D. Veraart acknowledges financial support by CREATES and by a Marie Curie FP7 Integration Grant within the 7th European Union Framework Programme.

modulated by stationary volatility process can be thought of as a stationary analogue of Lévy semimartingales of the type $t \mapsto \int_0^t \sigma(s)dL(s)$.

Benth and Eyjolfsson [10] discuss numerical methods for simulating discrete trajectories of Lévy semistationary processes, albeit without going into deep analysis of the methods. The current paper re-introduces an approximation method presented in that paper and analyses its convergence. The approximation method in question consist of considering an integral representation of the deterministic kernel function and considering methods to numerically approximate the corresponding integral. As it turns out, we can under some conditions interpret the approximation of the integral as a Fourier series, thereby giving ourselves some tools to analyse its convergence.

The main purpose of this paper is to analyse the Fourier approximation method and to illustrate its use for simulating prices of derivatives under Lévy semistationary electricity spot dynamics. We show that in terms of rate of convergence the Fourier simulation method and numerical integration are similar. However, an advantage of the Fourier simulation method is that it is an iterative scheme. By which we mean that to simulate a point $t + \delta$ given a value at t we merely need to simulate the increments in the Lévy and volatility processes and numerically evaluate a Fourier integral. This contrasts the corresponding numerical integration approach in which one needs to perform a complete re-integration in order to obtain the same iterative step. It follows that if the coefficients in the numerical Fourier integration are fewer than the number of previous time increments, our method is faster for simulating the incremental value. Moreover the Fourier approximation method is more flexible in the sense that given the Lévy and volatility processes one can easily simulate multiple Lévy semistationary processes driven by different kernel functions but with the same Lévy and volatility processes. Finally we discuss an application of our method which is simulating derivatives based on path-dependent options, such as Asian options which have been traded at the Nordic electricity exchange NordPool as OTC contracts (see Weron [22]).

This paper is structured as follows. In the next section we set the stage by introducing Lévy semistationary processes and some basic results on them, which shall be used throughout the paper. In section 3 we introduce the Fourier approximation scheme discuss methods to obtain trajectories by means of it and put it into context with Fourier series approximation. Following that, in section 4 we analyse the error induced by our approximation, in the mean square sense. Next in section 5 we compare our method to the more direct approach of numerical integration and illustrate the advantages of our approach. Finally in section 6 we apply our method to simulate prices for path dependent options and thereby demonstrate our method before arriving at our conclusive remarks.

2. Preliminaries

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{R}}, \mathbb{P})$ be a complete filtered probability space which satisfies the usual conditions, i.e. that the σ-algebras $\mathcal{F}_t$ include all the null sets of $\mathcal{F}$ and that the filtration $\{\mathcal{F}_t\}_{t\in\mathbb{R}}$ is right-continuous. We define a Lévy semistationary (LSS) process to be a process

of the type

$$X(t) = \int_{-\infty}^{t} g(t-s)\sigma(s-)dL(s) \tag{2.1}$$

for $t \in \mathbb{R}$, where L is a two-sided square integrable Lévy process, g is a real-valued, non-negative deterministic function on $\mathbb{R}_+$ and $\{\sigma(t)\}_{t\in\mathbb{R}}$ is a càdlàg process adapted to the filtration $\{\mathcal{F}_t\}_{t\in\mathbb{R}}$. We restrict our attention to LSS processes with a stochastic volatility process $\sigma(t)$ where $\sigma(t)$ is modelled as a *stationary* process *independent* of L. Denote for the sequel the first two moments of the Lévy process L and the volatility process σ by of $\sigma(t)$ by

$$m_j := \mathbb{E}[L^j(1)] \text{ and } \kappa_j := \mathbb{E}[\sigma^j(0)], \tag{2.2}$$

respectively, where $j = 1, 2$ and we assume that these moments are finite. Furthermore, g is assumed to be a Borel measurable function on $\mathbb{R}_+$, the positive half line including the origin, such that

$$g \in L^1(\mathbb{R}_+) \cap L^2(\mathbb{R}_+). \tag{2.3}$$

We remark that for our purposes if it holds that g is bounded, non-negative and in $L^1(\mathbb{R}_+)$, then it holds that $g \in L^2(\mathbb{R}_+)$. Note also that, as we shall see below, in the case when the Lévy process L is centered, i.e. when $m_1 = 0$ the milder condition $g \in L^2(\mathbb{R}_+)$ is in fact a sufficient to guarantee that the corresponding LSS process is well defined. However, we prefer to present the general conditions here, since we want to include LSS processes driven by non-centered Lévy processes in our analysis. These conditions, which we shall always assume to hold in what follows, ensure that $X(t)$ is well-defined, and square-integrable (see Protter [19] and Basse-O'Connor et al. [7]).

The characteristic function of X is easily computed by conditioning on the volatility process σ:

$$\mathbb{E}[\exp(\mathrm{i}\theta X(t))] = \mathbb{E}[\exp(\int_{-\infty}^{t} \psi(\theta g(t-s)\sigma(s-))ds)], \tag{2.4}$$

where $\psi(\theta)$ is the cumulant (i.e., the log-characteristic function) of $L(1)$. Here and in what follows, the cumulant function $\psi(\theta)$ is the unique real solution of $\exp(\psi(\theta)) = \mathbb{E}[\exp(\mathrm{i}\theta L(1))]$. Note in particular that if $\sigma = 1$, then the cumulant function of $X(t)$ is

$$\log \mathbb{E}[\exp(\mathrm{i}\theta X(t))] = \int_0^\infty \psi(\theta g(s))ds,$$

where $C = \int_0^\infty \psi(\theta g(s))ds$ denotes the unique real solution of $\exp(C) = \mathbb{E}[\exp(\mathrm{i}\theta X(t))]$. We observe that

$$\mathbb{E}[X(t)] = -\mathrm{i}\psi'(0) \int_{-\infty}^{t} g(t-s)\mathbb{E}[\sigma(s-)]ds = m_1\kappa_1||g||_{L^1(\mathbb{R}_+)}.$$

Furthermore, we find that the second moment is

(2.5)

$$\begin{aligned}\mathbb{E}[X^2(t)] &= -(\psi'(0))^2\mathbb{E}\left[\left(\int_{-\infty}^t g(t-s)\sigma(s-)ds\right)^2\right] - \psi''(0)\mathbb{E}\left[\int_{-\infty}^t g^2(t-s)\sigma^2(s-)ds\right] \\ &= m_1^2\mathbb{E}\left[\left(\int_{-\infty}^t g(t-s)\sigma(s-)ds\right)^2\right] + \mathrm{Var}(L(1))\kappa_2||g||^2_{L^2(\mathbb{R}_+)} \\ &= m_1^2\int_0^\infty\int_0^\infty g(u)g(v)\mathbb{E}[\sigma(0)\sigma(|u-v|)]dudv + \mathrm{Var}(L(1))\kappa_2||g||^2_{L^2(\mathbb{R}_+)}. \qquad (2.6)\end{aligned}$$

Notice that (2.6) is not time dependent, hence, $X(t)$ is second order stationary. Notice also, that in the case where the Lévy process is centered, i.e. when $m_1 = 0$, condition (2.6) reduces to a square integrability condition on the kernel function, g. Thus for LSS processes driven by centered Lévy processes the class of kernel functions is observed to be larger than in the case of non-centered Lévy processes.

It will be convenient to generalize the LSS processes introduced above to allow for complex-valued kernel functions. Thus, for a complex-valued kernel functions g such that $\mathrm{Re}\, g, \mathrm{Im}\, g \in L^1(\mathbb{R}_+) \cap L^2(\mathbb{R}_+)$, and real-valued volatility and Lévy processes, σ and L, we define a complex-valued LSS process as

$$X(t) = \int_{-\infty}^t \mathrm{Re}\, g(t-s)\sigma(s-)dL(s) + \mathrm{i}\int_{-\infty}^t \mathrm{Im}\, g(t-s)\sigma(s-)dL(s), \qquad (2.7)$$

and shall denote it by (2.1) as before.

The next Lemma concerns continuity of LSS processes with respect to the kernel function g and the stochastic volatility function σ.

Lemma 2.1. *Consider the complex LSS processes $X(t) = \int_{-\infty}^t g(t-s)\sigma(s-)dL(s)$, $Y(t) = \int_{-\infty}^t h(t-s)\sigma(s-)dL(s)$ and $Z(t) = \int_{-\infty}^t g(t-s)\rho(s-)dL(s)$. Then,*

i)

$$\mathbb{E}\left[|X(t)-Y(t)|^2\right] \le 2m_1^2\kappa_2||g-h||^2_{L^1(\mathbb{R}_+)} + \mathrm{Var}(L(1))\kappa_2||g-h||^2_{L^2(\mathbb{R}_+)},$$

where equality is obtained when the Lévy process is centered, i.e. when $m_1 = 0$.

ii)

$$\mathbb{E}\left[|X(t)-Z(t)|^2\right] \le \left(2m_1^2||g||^2_{L^1(\mathbb{R}_+)} + \mathrm{Var}(L(1))||g||^2_{L^2(\mathbb{R}_+)}\right)\sup_{s\in(-\infty,t]}\mathbb{E}\left[|\sigma(s-)-\rho(s-)|^2\right].$$

Proof. The first part of the proof goes by a straightforward calculation using (2.6), (2.7) and $\mathbb{E}[\sigma(0)\sigma(h)] < \kappa_2$ for any $h > 0$:

$$\begin{aligned}&\mathbb{E}\left[|X(t)-Y(t)|^2\right] \\ &= \mathbb{E}\left[\left|\mathrm{Re}\int_{-\infty}^t (g(t-s)-h(t-s))\sigma(s-)dL(s)\right|^2 + \left|\mathrm{Im}\int_{-\infty}^t (g(t-s)-h(t-s))\sigma(s-)dL(s)\right|^2\right]\end{aligned}$$

$$\begin{aligned}
&\leq m_1^2\kappa_2(||\operatorname{Re} g - \operatorname{Re} h||^2_{L^1(\mathbb{R}_+)} + ||\operatorname{Im} g - \operatorname{Im} h||^2_{L^1(\mathbb{R}_+)}) \\
&\quad + \operatorname{Var}(L(1))\kappa_2(||\operatorname{Re} g - \operatorname{Re} h||^2_{L^2(\mathbb{R}_+)} + ||\operatorname{Im} g - \operatorname{Im} h||^2_{L^2(\mathbb{R}_+)}) \\
&\leq 2m_1^2\kappa_2||g-h||^2_{L^1(\mathbb{R}_+)} + \operatorname{Var}(L(1))\kappa_2||g-h||^2_{L^2(\mathbb{R}_+)}.
\end{aligned}$$

Similarly for the second part, notice first that for a real-valued g it holds that

$$\mathbb{E}\left[\left(\int_{-\infty}^t g(t-s)(\sigma(s-)-\rho(s-))ds\right)^2\right] \leq \left(\int_{-\infty}^t g(t-s)ds\right)^2 \sup_{s\in(-\infty,t]}\mathbb{E}\left[|\sigma(s-)-\rho(s-)|^2\right].$$

Thus by (2.5) it holds that

$$\begin{aligned}
&\mathbb{E}\left[|X(t)-Z(t)|^2\right] \\
&= \mathbb{E}\left[\left|\operatorname{Re}\int_{-\infty}^t g(t-s)(\sigma(s-)-\rho(s-))dL(s)\right|^2 + \left|\operatorname{Im}\int_{-\infty}^t g(t-s)(\sigma(s-)-\rho(s-))dL(s)\right|^2\right] \\
&\leq \left(m_1^2\left(||\operatorname{Re} g||^2_{L^1(\mathbb{R}_+)} + ||\operatorname{Im} g||^2_{L^1(\mathbb{R}_+)}\right)\right. \\
&\quad \left. + \operatorname{Var}(L(1))\left(||\operatorname{Re} g||^2_{L^2(\mathbb{R}_+)} + ||\operatorname{Im} g||^2_{L^2(\mathbb{R}_+)}\right)\right) \sup_{s\in(-\infty,t]}\mathbb{E}\left[|\sigma(s-)-\rho(s-)|^2\right] \\
&\leq \left(2m_1^2||g||^2_{L^1(\mathbb{R}_+)} + \operatorname{Var}(L(1))||g||^2_{L^2(\mathbb{R}_+)}\right) \sup_{s\in(-\infty,t]}\mathbb{E}\left[|\sigma(s-)-\rho(s-)|^2\right].
\end{aligned}$$

Hence, the Lemma follows. □

In practice, for a given LSS spot price model, we would estimate the kernel function g from observed price data in the market. Such estimates are prone to statistical error, and hence we find g_ϵ rather than g itself, where ϵ is the error induced from statistical estimation, being a function of the number of data n at hand. The above result shows that the variance of $X(t)$ is robust towards this estimation error.

Let us consider an example of an approximation of a singular kernel g coming from applications to turbulence (see Barndorff-Nielsen and Schmiegel [5]). Suppose L is a centered Lévy process and g is of the form

$$g(x) = x^{\nu-1}\mathrm{e}^{-\alpha x}, \tag{2.8}$$

where $1/2 < \nu < 1$ and $\alpha > 0$. Note that g is singular at the origin, and $X(t)$ is thus, in general (unless L has bounded variation, see Basse and Pedersen [6]), not a semimartingale process. By Lemma 2.1 we may approximate $X(t)$ with a semimartingale LSS process that has the non-singular kernel function

$$g_\epsilon(x) = \begin{cases} g(x) & \text{if } x \geq \epsilon \\ g(\epsilon) & \text{if } x \in [0,\epsilon]. \end{cases}$$

We easily find that

$$\int_0^\infty (g(x)-g_\epsilon(x))^2\,dx \leq 2\int_0^\epsilon x^{2\nu-2}\mathrm{e}^{-2\alpha x}\,dx + 2\epsilon^{2\nu-1}\mathrm{e}^{-2\alpha\epsilon}$$

$$\leq \frac{2\epsilon^{2\nu-1}}{2\nu-1} + 2\epsilon^{2\nu-1} = \frac{4\nu\epsilon^{2\nu-1}}{2\nu-1}.$$

Thus we have the rate

$$||g - g_\epsilon||^2_{L^2(\mathbb{R}_+)} \leq \frac{4\nu\epsilon^{2\nu-1}}{2\nu-1},$$

from which we may observe that the closer ν is to $1/2$, the slower the rate is. If we want to simulate from $X(t)$, one would do numerical integration of $g(t-s)$ with respect to the paths of $L(s)$ and $\sigma(s)$ for $s \leq t$. To avoid problems around the singularity $s = t$, we can use h_ϵ rather than g in the numerical integration, with an error that we can control.

Another application of Lemma 2.1 is to view the LSS process $X(t)$ as a sliding window. To this end, fix $\tau > 0$, and for a real-valued non-negative kernel function g consider

$$X_\tau(t) := \int_{t-\tau}^{t} g(t-s)\sigma(s-)\, dL(s)\,. \tag{2.9}$$

Since,

$$X_\tau(t) = \int_{-\infty}^{t} g(t-s)1(\tau \geq t-s)\sigma(s-)\, dL(s)$$

we find from the Lemma that

$$\mathbb{E}\left[|X(t) - X_\tau(t)|^2\right] \leq m_1^2\kappa_2 \int_\tau^\infty g(x)dx + \mathrm{Var}(L(1))\kappa_2 \int_\tau^\infty g^2(x)dx. \tag{2.10}$$

By (2.3), the integrals on the right hand side will tend to zero as τ increases. This gives the interpretation of LSS processes as limits of moving average over a sliding window.

3. Fourier methods

In this section we discuss an alternative way of representing the kernel function g in order to allow for easy simulation of LSS process trajectories. Indeed, since the kernel function g in (2.1) is t dependent, i.e. changes with t, simulating a trajectory of X on a predefined set of points given a starting point, the Lévy and volatility processes using numerical integration requires that we evaluate a new integral at each point, and not just an increment.

To this end, for a given LSS process and $t, r \in \mathbb{R}$ such that $r \leq t$ consider the sum

$$X(t) = \int_{-\infty}^{r} g(t-s)\sigma(s-)dL(s) + \int_r^t g(t-s)\sigma(s-)dL(s). \tag{3.1}$$

By the integrability condition (2.3) it holds that

$$\lim_{r\to-\infty} \mathbb{E}\left[\left|\int_{-\infty}^{r} g(t-s)\sigma(s-)dL(s)\right|^2\right] = 0.$$

It follows that at a fixed time $t \in \mathbb{R}$ what happened in the past at time $r < t$ becomes less and less relevant for the present as $t - r$ becomes larger. Thus suppose we are interested

in simulating a discrete trajectory $X(t_0), X(t_1), \ldots, X(t_M)$ of a particular LSS process X. Then it follows that simulating a discrete trajectory $X_r(t_0), X_r(t_1), \ldots, X_r(t_M)$ where

$$X_r(t) := \int_r^t g(t-s)\sigma(s-)dL(s), \tag{3.2}$$

and $r \leq t$, yields an adequate approximation with an error we can make arbitrarily small. In what follows we introduce a simulation algorithm for efficiently simulating trajectories of (3.2).

Suppose that we fix the time horizon of X_r, i.e. we consider X_r on a bounded interval $[t_0, T]$ where $T < \infty$. The key observation here is that the truncated LSS process (3.2) only evaluates the kernel function g on the *bounded* interval $[0, T-r]$. It follows that we may adjust the kernel function as we wish outside the interval $[0, T-r]$. To that end, for fixed τ_0, τ such that $0 < \tau_0 < \tau$ and $T - r \leq \tau_0$ consider the function $h : \mathbb{R} \to \mathbb{R}$, defined by

$$h(x) = \begin{cases} g(|x|) & \text{if } |x| \leq \tau_0 \\ \phi(|x|) & \text{if } |x| \in (\tau_0, \tau] \\ 0 & \text{if } |x| > \tau \end{cases} \tag{3.3}$$

where $\phi : [\tau_0, \tau] \to \mathbb{R}$ is a continuous interpolating function such that $\phi(\tau_0) = g(\tau_0)$ and $\phi(\tau) = 0$. Notice in particular that $g = h$ on $[0, T-r]$ so $X_r(t) = \int_r^t h(t-s)\sigma(s-)dL(s)$ on $[t_0, T]$. Further, since h has a bounded support it holds for any given $\lambda > 0$ that

$$h_\lambda(x) := h(x)\mathrm{e}^{\lambda|x|} \in L^1(\mathbb{R}). \tag{3.4}$$

Now, let the Fourier transform of h_λ be (see Folland [14])

$$\widehat{h}_\lambda(y) = \int_{\mathbb{R}} h_\lambda(x)\mathrm{e}^{-\mathrm{i}xy}dx$$

and suppose furthermore that

$$\widehat{h}_\lambda \in L^1(\mathbb{R}).$$

Then the inverse Fourier transform exists, and we have (see Folland [14])

$$h(x) = \frac{\mathrm{e}^{-\lambda|x|}}{2\pi} \int_{\mathbb{R}} \widehat{h}_\lambda(y)\mathrm{e}^{\mathrm{i}yx}dy. \tag{3.5}$$

Note however, that since the Fourier transform maps integrable functions to continuous functions that vanish at infinity, i.e. $\mathcal{F}(L^1(\mathbb{R})) \subset C_0(\mathbb{R})$, we shall require that $h \in C_0(\mathbb{R})$. We remark that in the case when the kernel function g of interest is discontinuous, e.g. if it has a singularity, we need to approximate it, in the L^2 sense, with a continuous kernel function. We shall illustrate this with an example later. Now that we have an integral representation (3.5), we investigate to what extent we can estimate it, given the Fourier transform $\widehat{h}_\lambda$. By construction for a given $\lambda > 0$ the function h_λ is even which implies that the resulting Fourier transform, $\widehat{h}_\lambda$ is also even and real-valued. Thus the domain of

integration in the integral representation of h is reduced to the non-negative real numbers and may be approximated as

$$h(x) = \frac{\mathrm{e}^{-\lambda|x|}}{\pi} \int_0^\infty \widehat{h}_\lambda(y) \cos(yx) dy \approx \frac{\mathrm{e}^{-\lambda|x|}}{\pi} \sum_{n=0}^{N} \widehat{h}_\lambda(\xi_n) \cos(\xi_n x) \Delta y_n, \tag{3.6}$$

where $0 = y_0 < y_1 < \cdots < y_{N+1} < \infty$, $\Delta y_n = y_{n+1} - y_n$ and $\xi_n \in [y_n, y_{n+1}]$ for all $n = 0, \ldots, N$. As a function on the domain $[-\tau, \tau]$ the function h is continuous and even around the origin. Observe that in (3.6) the non-periodic function h is approximated by a finite sum of periodic functions. But since we are merely interested in approximating h on a bounded interval we may think of h as a periodic function on the entire real line with period 2τ. Now let us investigate which parameters in the finite sum numerical integration (3.6) constitute a good choice for the purpose of approximating h adequately on $[0, T-r]$.

It is essential for our approach that we select parameters which allow us to represent the finite sum approximation of the kernel function, h, as an orthogonal expansion. Meaning that $\{\cos(\xi_n x)\}_{n=0}^N$ is such that

$$\int_{-\tau}^{\tau} \cos(\xi_j x) \cos(\xi_k x) dx = \begin{cases} 0 & \text{if } j \neq k \\ 2\tau & \text{if } j = k = 0 \\ \tau & \text{if } j = k \geq 1 \end{cases} \tag{3.7}$$

for $j, k = 0, \ldots, N$. Given $\tau > 0$ the orthogonality condition (3.7) is satisfied if $\xi_n = n\pi/\tau$ for $n = 0, \ldots, N$. Thus yielding an equidistant evaluation grid $0 = \xi_0 < \xi_1 < \cdots < \xi_N$ with step size $\Delta\xi = \pi/\tau$. Furthermore, under the assumption (3.7), choosing parameters $c_0, c_1, \ldots, c_N$ to minimize the least squares integral

$$\int_{-\tau}^{\tau} \left(h_\lambda(x) - \sum_{n=0}^{N} c_n \cos(\xi_n x) \right)^2 dx \tag{3.8}$$

by differentiating with respect to c_n and using the orthogonality relation (3.7) for $n = 0, \ldots, N$ yields

$$c_0 = \frac{\widehat{h}_\lambda(\xi_0)}{2\tau} \quad \text{and} \quad c_n = \frac{\widehat{h}_\lambda(\xi_n)}{\tau} \text{ for } \quad n = 1, \ldots, N.$$

Together with (3.6) this suggests that if we take $\Delta y_0 = \pi/(2\tau)$ and $\Delta y_n = \pi/\tau$, for $n = 1, \ldots, N$, then under (3.7) the approximation (3.6) is optimal in the least squares sense. On the other hand it is easy to see that if the orthogonality condition (3.7) is not fulfilled, then the least squares sum (3.8) is no longer minimized by taking $c_n = c\widehat{h}_\lambda(\xi_n)$ for a constant c. Indeed, c_n will in general be a function of $\widehat{h}_\lambda(\xi_0), \widehat{h}_\lambda(\xi_1), \ldots, \widehat{h}_\lambda(\xi_N)$, which is inconsistent with the approximation (3.6), which we have made. Therefore in the sequel we shall work under the orthogonality assumption (3.7) together with

$$\xi_n = n\pi/\tau \text{ for } n = 0, \ldots, N,$$

$$\Delta y_0 = \pi/(2\tau) \quad \text{and} \quad \Delta y_n = \pi/\tau \text{ for } n = 1, \ldots, N.$$

Furthermore denoting by

$$a_n = \widehat{h}_\lambda(n\pi/\tau)/\tau \text{ for } n = 0, \dots, N$$

the approximation (3.6) may be written as

$$h(x) \approx \mathrm{e}^{-\lambda|x|}\left(\frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(n\pi x/\tau)\right) \tag{3.9}$$

and is in the least squares sense the optimal orthogonal expansion of the family $\{\cos(\xi_n x\}_{n=0}^N$ on its domain. Furthermore the expression on the right hand side is the Nth partial Fourier series of h_λ times $\mathrm{e}^{-\lambda|x|}$, if it is extended as a periodic function from $[-\tau, \tau]$ to the entire real line (see e.g. Folland [14]). Notice here that, since we have defined h_λ as a continuous function we avoid the Gibbs phenomenon in the partial Fourier series, which would cause increased error in our approximations.

Now applying what we have just found to simulate a trajectory of the integral (3.2) for a fixed $r \in \mathbb{R}$, we find that

$$\begin{aligned}\int_r^t g(t-s)\sigma(s-)dL(s) &\approx \int_r^t \mathrm{e}^{-\lambda(t-s)}\left(\frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(n\pi(t-s)/\tau)\right)\sigma(s-)dL(s) \\ &= \frac{a_0}{2}\widehat{X}_{\lambda,r}(t,0) + \mathrm{Re}\sum_{n=1}^{N} a_n \widehat{X}_{\lambda,r}(t, n\pi/\tau), \end{aligned} \tag{3.10}$$

where

$$\widehat{X}_{\lambda,r}(t,y) = \int_r^t \mathrm{e}^{(-\lambda+\mathrm{i}y)(t-s)}\sigma(s-)dL(s).$$

Note that since $\lambda > 0$, $\widehat{X}_{\lambda,r}(t,y)$ is a (complex-valued) LSS process for each $y \in \mathbb{R}$. We observe that for $\lambda = 0$, the definition of $\widehat{X}_0(t,y)$ fails since the complex exponential has norm 1 (except under stronger conditions on σ than we have assumed here). Whereas for any $\lambda > 0$ and $r < T$ we have by Lemma 2.1 that

$$\begin{aligned}\mathbb{E}\left[\left|\widehat{X}_{\lambda,r}(t,y)\right|^2\right] &\leq 2m_1^2\kappa_2\left(\int_0^\infty \mathrm{e}^{-\lambda x}dx\right)^2 + \mathrm{Var}(L(1))\kappa_2\int_0^\infty \mathrm{e}^{-2\lambda x}dx \\ &= \frac{2m_1^2\kappa_2}{\lambda^2} + \frac{\mathrm{Var}(L(1))\kappa_2}{2\lambda}. \end{aligned} \tag{3.11}$$

Thus for an arbitrary kernel function g (under our conditions) we represent the LSS process (3.2) as an approximation of an integral in the Fourier domain, where the integrand is a deterministic Fourier transform times a complex valued LSS process driven by an exponential kernel function that varies with t. Now, assuming the Fourier transform $\widehat{h}_\lambda$ is relatively easy to evaluate, this presents us with the advantage of employing the nice properties of the exponential function, when simulating a trajectory of an LSS process, such as being able to evaluate a trajectory stepwise, as follows.

Fix $\delta > 0$, and we find

$$\widehat{X}_{\lambda,r}(t+\delta, y) = \int_r^{t+\delta} \mathrm{e}^{(-\lambda+\mathrm{i}y)(t+\delta-s)}\sigma(s-)dL(s)$$

$$= \mathrm{e}^{(-\lambda+\mathrm{i}y)\delta}\widehat{X}_{\lambda,r}(t,y) + \mathrm{e}^{(-\lambda+\mathrm{i}y)\delta}\int_t^{t+\delta} \mathrm{e}^{(-\lambda+\mathrm{i}y)(t-s)}\sigma(s-)dL(s). \tag{3.12}$$

Now, the residuals can e.g. be simulated by the approximation

$$\int_t^{t+\delta} \mathrm{e}^{(-\lambda+\mathrm{i}y)(t-s)}\sigma(s-)\, dL(s) \approx \sigma(t-)\Delta L(t), \tag{3.13}$$

where $\Delta L(t) = L(t+\delta) - L(t)$. One can show that the variance of the error in this approximation is independent of y, and is of order δ. In principle, we could simulate $\widehat{X}_\lambda(t,y)$ exactly. For example, if $\sigma(s) = 1$, we have that the residual is an independent outcome of a random variable Z with cumulant

$$\log \mathbb{E}[\exp(i\theta \int_t^{t+\delta} \exp((-\lambda+\mathrm{i}y)(t-s))dL(s))] = \int_0^\delta \psi(\theta \mathrm{e}^{(-\lambda+\mathrm{i}y)u})du.$$

Thus, error is from numerical integration in Fourier domain only, and not connected to the simulations which are in principle exact. Notice furthermore that by approximating the above cumulant with $\delta\psi(\theta)$ we get the cumulant function of $\Delta L(t)$, which can be used to warrant the residual approximation (3.13).

Hence, to simulate a discrete trajectory $X_r(t_0), X_r(t_1), \ldots, X_r(t_M)$ given all information available at time t_0, we do the following: For each t_j where $j = 1, \ldots, M$.

(1) Simulate $\Delta L(t_j)$
(2) For each $n = 0, \ldots, N$, simulate $\widehat{X}_{\lambda,r}(t_j, n\pi/\tau)$ from $\widehat{X}_{\lambda,r}(t_{j-1}, n\pi/\tau)$ and $\Delta L(t_j)$
(3) Compute numerically the inverse Fourier transform in (3.10) where $a_n = \widehat{h}_\lambda(n\pi/\tau)/\tau$ for $n = 0, \ldots, N$.

Note the advantages here: We have the same residual term for every y_n, except from a deterministic scaling by a complex exponential. This means that to simulate $\widehat{X}_\lambda(t_j, y)$, we simulate the outcome of *one* random variable Z, and then compute

$$\widehat{X}_{\lambda,r}(t_j, y) = \exp((-\lambda+\mathrm{i}y)\delta)\left\{\widehat{X}_{\lambda,r}(t_{j-1}, y) + Z\right\}.$$

Hence, in step 2 above, we just need to have stored the $N+1$ values of $\widehat{X}_\lambda(t_{j-1}, y_n)$ from the previous time t_{j-1} along with the simulated Z, in order to compute the next iterative step. Notice also that the number of sampling points N depends on the damping properties of $\widehat{h}_\lambda$. The faster $\widehat{h}_\lambda(y)$ decays to zero for large values of y, the smaller N can be chosen. Further we have seen that the points of evaluation $\{\xi_n\}_{n=0}^N$ are optimally chosen under our orthogonality condition (3.7). We can also easily change the kernel function g, without having to redo the whole simulation algorithm, since the first steps are independent of g. This may prove advantageous in estimation studies, where one may want to simulate over parametric g's in order to find the optimal one. Finally, another advantage compared to direct numerical integration is that with the latter, the accuracy is linked to how many

sample points we simulate the Lévy process in time, whereas with the Fourier technique this is converted into sampling an integral over space instead.

4. Error estimation

In the previous section we introduced a method for simulating LSS processes by means of representing the kernel function as an integral and approximating the resulting integral. Furthermore we have seen that by splitting a general LSS process into the sum (3.1) and focusing on the latter integral there is an optimal way of approximating the latter integral in mean square under our assumptions. In the present section we shall employ the assumptions we made in the previous section together with Lemma 2.1 to analyse the error induced by our approximation and compare it to the error induced by employing a step function to approximate the kernel function. To that end for a fixed $\lambda > 0$ and $N \geq 1$ we denote by

$$h_N(x) = \mathrm{e}^{-\lambda|x|}\left(\frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(n\pi x/\tau)\right) \tag{4.1}$$

the Nth partial Fourier series of h_λ times $\mathrm{e}^{-\lambda|x|}$, as introduced in the previous section. Since h_λ is integrable, it is represented by $\mathrm{e}^{\lambda|x|}h_N(x)$ on $[-\tau,\tau]$ in the sense that

$$\lim_{N\to\infty}\int_{-\tau}^{\tau}\left|h_\lambda(x) - \frac{a_0}{2} - \sum_{n=1}^{N} a_n\cos(n\pi x/\tau)\right|^2 dx = 0 \tag{4.2}$$

(see Folland [14]). We may employ the sum representation (3.1), and Lemma 2.1 to get

$$\begin{aligned}&\mathbb{E}\left[\left|X(t) - \int_{t_0}^{t} h_N(t-s)\sigma(s-)dL(s)\right|^2\right]\\ &\leq 2\mathbb{E}\left[\left|\int_{-\infty}^{t_0} g(t-s)\sigma(s-)dL(s)\right|^2\right] + 2\mathbb{E}\left[\left|\int_{t_0}^{t}(g(t-s)-h_N(t-s))\sigma(s-)dL(s)\right|^2\right].\end{aligned}$$

By Lemma 2.1 it holds that

$$\begin{aligned}&\mathbb{E}\left[\left|\int_{-\infty}^{t_0} g(t-s)\sigma(s-)dL(s)\right|^2\right]\\ &\leq 2m_1^2\kappa_2||g(t-t_0+\cdot)||^2_{L^1(\mathbb{R}_+)} + \mathrm{Var}(L(1))\kappa_2||g(t-t_0+\cdot)||^2_{L^2(\mathbb{R}_+)}\end{aligned} \tag{4.3}$$

and

$$\begin{aligned}&\mathbb{E}\left[\left|\int_{t_0}^{t}(g(t-s)-h_N(t-s))\sigma(s-)dL(s)\right|^2\right]\\ &\leq 2m_1^2\kappa_2||g-h_N||^2_{L^1([0,t-t_0])} + \mathrm{Var}(L(1))\kappa_2||g-h_N||^2_{L^2([0,t-t_0])}.\end{aligned} \tag{4.4}$$

Clearly by (2.3), it holds that (4.3) tends to zero as $t - t_0$ increases, with the rate of convergence being controlled by the decay of the corresponding kernel function g, whereas the error (4.4) is controlled by the L^1 and L^2 convergence of Fourier series approximating the corresponding kernel function. By employing the L^2 convergence (4.2) and the elementary inequality $|a+b|^2 \leq 2(|a|^2 + |b|^2)$ we find that

$$\begin{aligned}
||g - h_N||^2_{L^2([0,t-t_0])} &= \int_0^{t-t_0} \mathrm{e}^{-2\lambda x} \left| h_\lambda(x) - \frac{a_0}{2} - \sum_{n=1}^{N} a_n \cos(n\pi x/\tau) \right|^2 dx \\
&\leq 2 \int_0^{t-t_0} \mathrm{e}^{-2\lambda x} \left| \sum_{n=N+1}^{\infty} a_n \cos(n\pi x/\tau) \right|^2 dx \\
&\leq 2 \int_0^{t-t_0} \mathrm{e}^{-2\lambda x} \left(\sum_{n=N+1}^{\infty} |a_n \cos(n\pi x/\tau)| \right)^2 dx \\
&\leq \frac{1 - \mathrm{e}^{-2\lambda(t-t_0)}}{\lambda} \left(\sum_{n=N+1}^{\infty} |a_n| \right)^2 .
\end{aligned}$$

Similarly by using the Cauchy-Schwarz inequality we furthermore find that

$$\begin{aligned}
||g - h_N||^2_{L^1([0,t-t_0])} &\leq \frac{1 - \mathrm{e}^{-2\lambda(t-t_0)}}{2\lambda} \int_0^{t-t_0} \left| h_\lambda(x) - \frac{a_0}{2} - \sum_{n=1}^{N} a_n \cos(n\pi x/\tau) \right|^2 dx \\
&\leq (t-t_0) \frac{1 - \mathrm{e}^{-2\lambda(t-t_0)}}{\lambda} \left(\sum_{n=N+1}^{\infty} |a_n| \right)^2 .
\end{aligned}$$

Combining these findings with (4.4) we thus find that

$$\begin{aligned}
&\mathbb{E}\left[\left| \int_{t_0}^{t} (g(t-s) - h_N(t-s)) \sigma(s-) dL(s) \right|^2 \right] \\
&\leq \left(2m_1^2(t-t_0) + \mathrm{Var}(L(1)) \right) \kappa_2 \frac{1 - \mathrm{e}^{-2\lambda(t-t_0)}}{\lambda} \left(\sum_{n=N+1}^{\infty} |a_n| \right)^2 . \qquad (4.5)
\end{aligned}$$

Thus, unsurprisingly, the Fourier series kernel function approximation (3.9) is most dependent on the choice of $\lambda > 0$ and how many terms N we use to approximate h with its corresponding partial Fourier series. The following question arises: How many terms in the numerical integration (3.9) yield an adequate approximation in terms of the error bound (4.5). Or to rephrase, what is the order of a_n, $n \geq 1$? Let us take a closer look at the coefficients a_n, $n = 1, \ldots, N$. Assuming that the kernel function h is sufficiently smooth, integration by parts yields

$$\tau a_n = \int_{-\tau}^{\tau} h_\lambda(x) \cos(n\pi x/\tau) dx$$

$$= 0 - \frac{\tau}{n\pi}\int_{-\tau}^{\tau} h'_\lambda(x)\sin(n\pi x/\tau)dx \tag{4.6}$$

$$= \left(\frac{\tau}{n\pi}\right)^2 2(-1)^n h'_\lambda(\tau) - \left(\frac{\tau}{n\pi}\right)^2 \int_{-\tau}^{\tau} h''_\lambda(x)\cos(n\pi x/\tau)dx. \tag{4.7}$$

Thus if $h'_\lambda(\tau) \neq 0$ the first surviving term of a_n will be of the order n^{-2}, whereas if

$$h'_\lambda(\tau) = 0 \tag{4.8}$$

then the first term will be of the order n^{-4}. Thus if (4.8) holds then the $|a_n|$ terms decrease at a faster rate, speeding up the convergence of our method, since (4.8) and (4.7) combined yield

$$|a_n| \leq \frac{\tau||h''_\lambda||_{L^1([-\tau,\tau])}}{\pi^2}\frac{1}{n^2}. \tag{4.9}$$

Now for our purposes, recalling the definition (3.3), this means that the interpolating function ϕ should fulfill

$$\phi'_\lambda(\tau) = 0.$$

Thus, selecting a smooth interpolation function is generally a good idea, since it will speed up the convergence of the corresponding Fourier series and thus reduce the computational burden in approximating the kernel function by numerical integration in the Fourier domain. Our findings are summarized, and slightly generalized, in the following proposition.

Proposition 4.1. *Suppose that h defined by (3.3) is a C^{2k} function for some $k \geq 1$, such that $h_\lambda^{(2n-1)}(\tau) = 0$ for $n = 1, \ldots, k$. Then it holds that*

$$\mathbb{E}\left[\left|\int_{t_0}^{t}(g(t-s) - h_N(t-s))\sigma(s-)dL(s)\right|^2\right]$$
$$\leq \left(2m_1^2(t-t_0) + \mathrm{Var}(L(1))\right)\kappa_2\frac{1-\mathrm{e}^{-2\lambda(t-t_0)}}{\lambda}\frac{\tau^{4k-2}||h_\lambda^{(2k)}||^2_{L^1([-\tau,\tau])}}{\pi^{4k}}\left(\sum_{n=N+1}^{\infty}\frac{1}{n^{2k}}\right)^2.$$

From the above Proposition it is clear that the rate of convergence of is controlled by three main factors. The most apparent factor is the infinite series

$$\sum_{n=1}^{\infty}\frac{1}{n^{2k}}, \tag{4.10}$$

where $k \geq 1$. This series is convergent for all $k \geq 1$ and is in fact equal to $\zeta(2k)$ for a given $k \geq 1$, where ζ denotes the Riemann zeta function. Thus for instance it holds that $\zeta(2) = \pi^2/6$ and $\zeta(4) = \pi^4/90$. In figure 1 we illustrate the convergence rates of the series (4.10) for $k = 1, 2$ respectively. Clearly the convergence of the series is considerably faster for $k = 2$ than for $k = 1$.

The size of the parameter $k \geq 1$ is determined by the h_λ function, which in turn is determined by the original kernel function g and, perhaps to a bigger extent, the interpolation function ϕ. Indeed since the conclusion of Proposition 4.1 holds if h is a C^{2k} function and

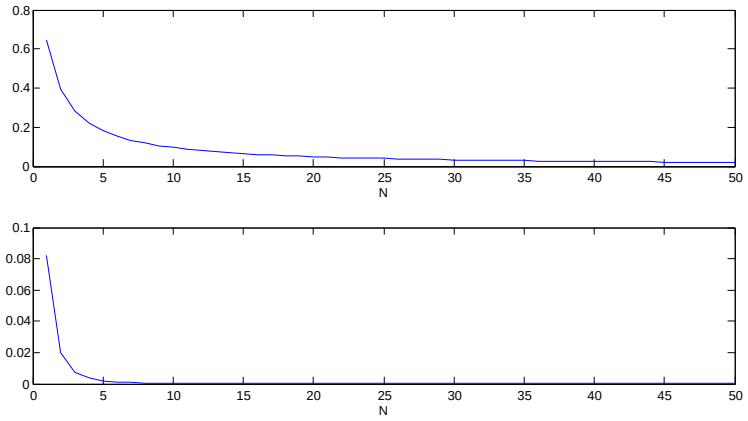

FIGURE 1. The convergence rate of the series (4.10) for $k = 1, 2$. Above: $N \mapsto \pi^2/6 - \sum_{n=1}^{N} n^{-2}$. Below: $N \mapsto \pi^4/90 - \sum_{n=1}^{N} n^{-4}$.

$h_\lambda^{(2n-1)}(\tau) = 0$ for $n = 1, \ldots, k$, which translates into $\phi_\lambda^{(2n-1)}(\tau) = 0$ for $n = 1, \ldots, k$. Thus the selection of the interpolation function is paramount for our approach. Clearly it may be selected in a number of different ways, but perhaps the simplest way would be to let ϕ be an appropriately selected polynomial. Let us outline how this can be achieved. Suppose that g is a C^{2k} function and let ϕ be the $4k+1$ degree polynomial

$$\phi(x) = \sum_{j=0}^{4k+1} c_j x^j, \tag{4.11}$$

where the coefficients c_j, $j = 0, \ldots, 4k+1$ are determined by the conditions

$$\phi^{(j)}(\tau_0) = g^{(j)}(\tau_0) \text{ and } \phi^{(j)}(\tau) = 0, \text{ for } j = 0, \ldots, 2k. \tag{4.12}$$

Equipped with this interpolation function h (and thus h_λ) is a C^{2k} function such that $h_\lambda^{(2n-1)}(\tau) = 0$ for $n = 1, \ldots, k$ and $h_\lambda^{(2k)} \in L^1([-\tau, \tau])$.

Finally the parameter $\lambda > 0$ clearly has an impact on the convergence rate in Proposition 4.1. However, the set of admissible $\lambda > 0$ is limited by the condition $\widehat{h}_\lambda \in L^1(\mathbb{R})$, which should be fulfilled in each case as we have assumed in the previous section. In fact it always holds if $h_\lambda \in C^2$ and $h'_\lambda, h''_\lambda \in L^1$ (see Folland [14], section 8.4). But on the other hand a high $\lambda > 0$ usually leads to high values of $||h_\lambda^{(2k)}||^2_{L^1([-\tau,\tau])}$. It follows that letting λ be "too high" can in fact increase the error, so one needs to pay some attention to the parameter selection.

Now, let us turn our attention to the estimation

$$\int_t^{t+\delta} \mathrm{e}^{(\mathrm{i}y-\lambda)(t-s)} \sigma(s-) dL(s) \approx \sigma(t-) \Delta L(t),$$

which we make since in practice we may not be able to simulate $\widehat{X}_\lambda$ exactly. We find that

$$\begin{aligned}\widehat{X}_{\lambda,r}(t+\Delta t,y) &= \mathrm{e}^{(-\lambda+\mathrm{i}y)\Delta t}\left(\widehat{X}_{\lambda,r}(t,y)+\int_t^{t+\Delta t}\mathrm{e}^{(-\lambda+\mathrm{i}y)(t-s)}\sigma(s-)dL(s)\right)\\ &\approx \mathrm{e}^{(-\lambda+\mathrm{i}y)\Delta t}\left(\widehat{X}_{\lambda,r}(t,y)+\epsilon(t,\Delta t,y)\right), \qquad (4.13)\end{aligned}$$

where

$$\epsilon(t,\Delta t,y)\approx\int_t^{t+\Delta t}\mathrm{e}^{(-\lambda+\mathrm{i}y)(t-s)}\sigma(s-)dL(s)$$

denotes the residual approximation, which we assume to be dependent on σ and L on the incremental interval $[t,t+\Delta t]$. Now let us attempt to estimate the mean square error which arises from making this assumption, i.e. by making the estimation (4.13). To this end, let us assume a equidistant time grid $r=t_0<t_1<\cdots<t_M$, with constant step size $\Delta t>0$. By iterating (4.13) we observe that for a given $1\leq j\leq M$ and any $y\in\mathbb{R}$ it holds that

$$\begin{aligned}\widehat{X}_{\lambda,r}(t_j,y) &= \sum_{k=1}^{j}\mathrm{e}^{(-\lambda+\mathrm{i}y)(j+1-k)\Delta t}\int_{t_{k-1}}^{t_k}\mathrm{e}^{(-\lambda+\mathrm{i}y)(t_{k-1}-s)}\sigma(s-)dL(s)\\ &\approx \sum_{k=1}^{j}\mathrm{e}^{(-\lambda+\mathrm{i}y)(j+1-k)\Delta t}\epsilon(t_{k-1},\Delta t,y).\end{aligned}$$

Now suppose that

$$\epsilon(t,\Delta t,y)=\sigma(t-)\Delta L(t), \qquad (4.14)$$

then it furthermore holds that

$$\begin{aligned}\widehat{X}_{\lambda,r}(t_j,y) &\approx \int_{t_0}^{t_j}\sum_{k=1}^{j}\mathrm{e}^{(-\lambda+\mathrm{i}y)(j+1-k)\Delta t}\sigma(t_{k-1}-)1_{[t_{k-1},t_k)}(s)dL(s)\\ &= \int_{t_0}^{t_j}\sum_{k=1}^{j}\mathrm{e}^{(-\lambda+\mathrm{i}y)(j+1-k)\Delta t}1_{[t_{k-1},t_k)}(s)\sigma(s-)dL(s)\\ &\quad+\int_{t_0}^{t_j}\sum_{k=1}^{j}\mathrm{e}^{(-\lambda+\mathrm{i}y)(j+1-k)\Delta t}(\sigma(t_{k-1}-)-\sigma(s-))1_{[t_{k-1},t_k)}(s)dL(s).\end{aligned}$$

Observe that by Lemma 2.1, the Cauchy-Schwarz inequality and Lipschitz continuity of the exponential function it holds that

$$\mathbb{E}\left[\left|\int_{t_0}^{t_j}\left(\mathrm{e}^{(-\lambda+\mathrm{i}y)(t_j-s)}-\sum_{k=1}^{j}\mathrm{e}^{(-\lambda+\mathrm{i}y)(j+1-k)\Delta t}1_{[t_{k-1},t_k)}(s)\right)\sigma(s-)dL(s)\right|^2\right]$$

$$
\begin{aligned}
&\leq 2m_1^2\kappa_2\left(\int_{t_0}^{t_j}\left|\mathrm{e}^{(-\lambda+\mathrm{i}y)(t_j-s)}-\sum_{k=1}^{j}\mathrm{e}^{(-\lambda+\mathrm{i}y)(j+1-k)\Delta t}1_{[t_{k-1},t_k)}(s)\right|ds\right)^2\\
&+\mathrm{Var}(L(1))\kappa_2\int_{t_0}^{t_j}\left|\mathrm{e}^{(-\lambda+\mathrm{i}y)(t_j-s)}-\sum_{k=1}^{j}\mathrm{e}^{(-\lambda+\mathrm{i}y)(j+1-k)\Delta t}1_{[t_{k-1},t_k)}(s)\right|^2ds\\
&\leq (2m_1^2(t_j-t_0)+\mathrm{Var}(L(1)))\kappa_2\int_{t_0}^{t_j}\left|\mathrm{e}^{(-\lambda+\mathrm{i}y)(t_j-s)}-\sum_{k=1}^{j}\mathrm{e}^{(-\lambda+\mathrm{i}y)(j+1-k)\Delta t}1_{[t_{k-1},t_k)}(s)\right|^2ds\\
&\leq (2m_1^2(t_j-t_0)+\mathrm{Var}(L(1)))\kappa_2(\lambda^2+y^2)\int_{t_0}^{t_j}\left|s-\sum_{k=1}^{j}t_{k-1}1_{[t_{k-1},t_k)}(s)\right|^2ds\\
&=(2m_1^2(t_j-t_0)+\mathrm{Var}(L(1)))\kappa_2(\lambda^2+y^2)j\frac{(\Delta t)^3}{3}\\
&=(2m_1^2(t_j-t_0)+\mathrm{Var}(L(1)))\kappa_2(\lambda^2+y^2)(t_j-t_0)\frac{(\Delta t)^2}{3}.
\end{aligned}
$$

Furthermore it holds by Lemma 2.1 that

$$
\begin{aligned}
&\mathbb{E}\left[\left|\int_{t_0}^{t_j}\sum_{k=1}^{j}\mathrm{e}^{(-\lambda+\mathrm{i}y)(j+1-k)\Delta t}(\sigma(t_{k-1}-)-\sigma(s-))1_{[t_{k-1},t_k)}(s)dL(s)\right|^2\right]\\
&\leq (2m_1^2(t_j-t_0)+\mathrm{Var}(L(1)))(t_j-t_0)\mathbb{E}\left[|\sigma(t_0-)-\sigma(t_1-)|^2\right].
\end{aligned}
$$

In the sequel, for a given $\lambda>0$ and an equidistant time grid $r=t_0<t_1<\cdots<t_M$ with constant step size $\Delta t>0$, let us denote by

$$
\eta_j(y):=\sum_{k=1}^{j}\mathrm{e}^{(-\lambda+\mathrm{i}y)(j+1-k)\Delta t}\sigma(t_{k-1}-)\Delta L(t_{k-1}) \tag{4.15}
$$

the time series which approximates $\widehat{X}_{\lambda,r}(t_j,y)$, for $j=1,\ldots,M$ for any $y\in\mathbb{R}$ by means of the assumption (4.14). We summarize our findings in the following Lemma.

Lemma 4.2. *Given an equidistant grid $r=t_0<t_1<\cdots<t_M$, with constant step size $\Delta t>0$, it holds that*

$$
\begin{aligned}
&\mathbb{E}[|\widehat{X}_\lambda(t_j,y)-\eta_j(y)|^2]\\
&\leq (2m_1^2(t_j-t_0)+\mathrm{Var}(L(1)))(t_j-t_0)\left(\kappa_2(\lambda^2+y^2)\frac{(\Delta t)^2}{3}+\mathbb{E}\left[|\sigma(t_0-)-\sigma(t_1-)|^2\right]\right),
\end{aligned}
$$

where η is defined by (4.15).

In Proposition 4.1 we quantified the error induced by the approximation (3.10). Moreover with Lemma 4.2 we have studied the error caused by employing the residual approximation

(3.13). The next result employs what we have just found to complement Proposition 4.1, in the sense that together with Proposition 4.1 it characterizes the error induced by employing our Fourier approximation method which we introduced in the previous section.

Proposition 4.3. *Given an equidistant grid $r = t_0 < t_1 < \cdots < t_M$, with constant step size $\Delta t > 0$, it holds that*

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{t_0}^{t_j} h_N(t_j - s)\sigma(s-)dL(s) - \frac{a_0}{2}\eta_j(0) - \operatorname{Re}\sum_{n=1}^{N} a_n\eta_j(n\pi/\tau)\right|^2\right] \\
&\leq (2m_1^2(t_j - t_0) + \operatorname{Var}(L(1)))(t_j - t_0)\left\{\kappa_2\lambda^2\left(\frac{|a_0|}{2} + \sum_{n=1}^{N}|a_n|\right)^2(\Delta t)^2\right. \\
&\left.+\kappa_2\left(\frac{\pi}{\tau}\right)^2\left(\sum_{n=1}^{N}|a_n|n\right)^2(\Delta t)^2 + 3\left(\frac{|a_0|}{2} + \sum_{n=1}^{N}|a_n|\right)^2\mathbb{E}\left[|\sigma(t_0-) - \sigma(t_1-)|^2\right]\right\},
\end{aligned}$$

where η is defined by (4.15).

Proof. This follows from Lemma 4.2. Since by an application of Minkowski's inequality it holds that

$$\begin{aligned}
&\left(\mathbb{E}\left[\left|\int_{t_0}^{t_j} h_N(t_j - s)\sigma(s-)dL(s) - \frac{a_0}{2}\eta_j(0) - \operatorname{Re}\sum_{n=1}^{N} a_n\eta_j(n\pi/\tau)\right|^2\right]\right)^{1/2} \\
&= \left(\mathbb{E}\left[\left|\frac{a_0}{2}(\widehat{X}_{\lambda,r}(t_j,0) - \eta_j(0)) - \operatorname{Re}\sum_{n=1}^{N} a_n(\widehat{X}_{\lambda,r}(t_j,n\pi/\tau) - \eta_j(n\pi/\tau))\right|^2\right]\right)^{1/2} \\
&\leq \frac{|a_0|}{2}\left(\mathbb{E}\left[\left|\widehat{X}_{\lambda,r}(t_j,0) - \eta_j(0)\right|^2\right]\right)^{1/2} + \sum_{n=1}^{N}|a_n|\left(\mathbb{E}\left[\left|\widehat{X}_{\lambda,r}(t_j,n\pi/\tau) - \eta_j(n\pi/\tau)\right|^2\right]\right)^{1/2} \\
&\leq \left((2m_1^2(t_j - t_0) + \operatorname{Var}(L(1)))(t_j - t_0)\right)^{1/2}\left\{\frac{|a_0|}{2}\left(\kappa_2\lambda^2\frac{(\Delta t)^2}{3} + \mathbb{E}\left[|\sigma(t_0-) - \sigma(t_1-)|^2\right]\right)^{1/2}\right. \\
&\left.+\sum_{n=1}^{N}|a_n|\left(\kappa_2(\lambda^2 + (n\pi/\tau)^2)\frac{(\Delta t)^2}{3} + \mathbb{E}\left[|\sigma(t_0-) - \sigma(t_1-)|^2\right]\right)^{1/2}\right\}.
\end{aligned}$$

Now by noticing that $\sqrt{x^2 + y^2} \leq x + y$ for any non-negative real numbers $x, y \geq 0$ it furthermore holds that

$$\begin{aligned}
&\left(\kappa_2(\lambda^2 + (n\pi/\tau)^2)\frac{(\Delta t)^2}{3} + \mathbb{E}\left[|\sigma(t_0-) - \sigma(t_1-)|^2\right]\right)^{1/2} \\
&\leq \left(\frac{\kappa_2}{3}\right)^{1/2}\left(\lambda^2 + (n\pi/\tau)^2\right)^{1/2}\Delta t + \left(\mathbb{E}\left[|\sigma(t_0-) - \sigma(t_1-)|^2\right]\right)^{1/2}
\end{aligned}$$

$$\leq \left(\frac{\kappa_2}{3}\right)^{1/2} (\lambda + n\pi/\tau)\,\Delta t + \left(\mathbb{E}\left[|\sigma(t_0-) - \sigma(t_1-)|^2\right]\right)^{1/2},$$

for all $n = 0, \ldots, N$. Thus by the elementary inequality $(x+y+z)^2 \leq 3(x^2+y^2+z^2)$ it holds that

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{t_0}^{t_j} h_N(t_j - s)\sigma(s-)dL(s) - \frac{a_0}{2}\eta_j(0) - \mathrm{Re}\sum_{n=1}^{N} a_n\eta_j(n\pi/\tau)\right|^2\right] \\
&\leq \left((2m_1^2(t_j - t_0) + \mathrm{Var}(L(1)))(t_j - t_0)\right)\left\{\left(\frac{\kappa_2}{3}\right)^{1/2}\lambda\left(\frac{|a_0|}{2} + \sum_{n=1}^{N}|a_n|\right)\Delta t\right. \\
&\left. + \left(\frac{\kappa_2}{3}\right)^{1/2}\frac{\pi}{\tau}\sum_{n=1}^{N}|a_n| n\Delta t + \left(\frac{|a_0|}{2} + \sum_{n=1}^{N}|a_n|\right)\left(\mathbb{E}\left[|\sigma(t_0-) - \sigma(t_1-)|^2\right]\right)^{1/2}\right\}^2 \\
&\leq \left((2m_1^2(t_j - t_0) + \mathrm{Var}(L(1)))(t_j - t_0)\right)\left\{\kappa_2\lambda^2\left(\frac{|a_0|}{2} + \sum_{n=1}^{N}|a_n|\right)^2(\Delta t)^2\right. \\
&\left. + \kappa_2\left(\frac{\pi}{\tau}\right)^2\left(\sum_{n=1}^{N}|a_n| n\right)^2(\Delta t)^2 + 3\left(\frac{|a_0|}{2} + \sum_{n=1}^{N}|a_n|\right)^2\mathbb{E}\left[|\sigma(t_0-) - \sigma(t_1-)|^2\right]\right\}
\end{aligned}$$

□

The above Proposition is the final piece in our puzzle in the sense that it characterizes the error of our Fourier approximation in terms of the step size in the time domain. Whereas Proposition 4.1 characterizes the error induced by making the Fourier series approximation. Thus giving us the complete picture. Moreover an application of Bessel's inequality gives us the following result.

Corollary 4.4. *Suppose that $h_\lambda \in C^2$ is such that $h'_\lambda(\tau) = 0$, then given an equidistant grid $r = t_0 < t_1 < \cdots < t_M$, with constant step size $\Delta t > 0$, it holds that*

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{t_0}^{t_j} h_N(t_j - s)\sigma(s-)dL(s) - \frac{a_0}{2}\eta_j(0) - \mathrm{Re}\sum_{n=1}^{N} a_n\eta_j(n\pi/\tau)\right|^2\right] \\
&\leq (2m_1^2(t_j - t_0) + \mathrm{Var}(L(1)))(t_j - t_0)\left\{\kappa_2\lambda^2\left(|a_0|^2 + \frac{\tau^2}{3}||h'_\lambda||^2_{L^2([-\tau,\tau])}\right)(\Delta t)^2\right. \\
&\left. + \kappa_2\frac{\tau^2}{6}||h''_\lambda||^2_{L^2([-\tau,\tau])}(\Delta t)^2 + 3\left(|a_0|^2 + \frac{\tau^2}{3}||h'_\lambda||^2_{L^2([-\tau,\tau])}\right)\mathbb{E}\left[|\sigma(t_0-) - \sigma(t_1-)|^2\right]\right\},
\end{aligned}$$

where η is defined by (4.15).

Proof. By the Cauchy-Schwarz inequality, the identity (4.6), Bessel's inequality and $\sum_{n=1}^{\infty} n^{-2} = \pi^2/6$ it follows that

$$\begin{aligned}\left(\frac{|a_0|}{2} + \sum_{n=1}^{N} |a_n|\right)^2 &\leq |a_0|^2 + 2\left(\sum_{n=1}^{N} |a_n|\right)^2 \\ &\leq |a_0|^2 + 2\left(\sum_{n=1}^{N} \frac{1}{n^2}\right)\left(\sum_{n=1}^{N} |a_n n|^2\right) \\ &= |a_0|^2 + \frac{2}{\pi^2}\left(\sum_{n=1}^{N} \frac{1}{n^2}\right)\left(\sum_{n=1}^{N} |\widehat{h'_\lambda}(n\pi/\tau)|^2\right) \\ &\leq |a_0|^2 + \frac{\tau^2}{3}||h'_\lambda||^2_{L^2([-\tau,\tau])}.\end{aligned}$$

Similarly by the Cauchy-Schwarz inequality, the identity (4.7), Bessel's inequality and $\sum_{n=1}^{\infty} n^{-2} = \pi^2/6$ it holds that

$$\begin{aligned}\left(\sum_{n=1}^{N} |a_n| n\right)^2 &\leq \left(\sum_{n=1}^{N} \frac{1}{n^2}\right)\left(\sum_{n=1}^{N} |a_n n^2|^2\right) \\ &= \frac{\tau^2}{\pi^4}\left(\sum_{n=1}^{N} \frac{1}{n^2}\right)\left(\sum_{n=1}^{N} |\widehat{h''_\lambda}(n\pi/\tau)|^2\right) \\ &\leq \frac{1}{6}\frac{\tau^4}{\pi^2}||h''_\lambda||^2_{L^2([-\tau,\tau])}.\end{aligned}$$

The proof is concluded by Proposition 4.3. □

5. Comparison to numerical integration

In the present section we will present the advantages of employing the Fourier method to simulate paths of LSS processes. Let us begin by recalling the more standard approach of numerical integration given an equidistant grid $r = t_0 < t_1 < \cdots < t_M$ with a constant step size $\Delta t > 0$ which consists of the approximation

$$\begin{aligned}X_r(t_j) &\approx \sum_{k=0}^{j-1} g(t_j - t_k)\sigma(t_k)\Delta L(t_k) \\ &= \int_{t_0}^{t_j} \sum_{k=0}^{j-1} g(t_j - t_k)\sigma(t_k) 1_{[t_k, t_{k+1})}(s) dL(s). \qquad (5.1)\end{aligned}$$

Then by Lemma 2.1 and the Cauchy-Schwarz inequality it holds that

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{t_0}^{t_j}\left(g(t-s)-\sum_{k=0}^{j-1}g(t_j-t_k)1_{[t_k,t_{k+1})}(s)\right)\sigma(s-)dL(s)\right|^2\right]\\
&\leq 2m_1^2\kappa_2\left(\int_{t_0}^{t_j}\left|g(t-s)-\sum_{k=0}^{j-1}g(t_j-t_k)1_{[t_k,t_{k+1})}(s)\right|ds\right)^2\\
&\quad+\mathrm{Var}(L(1))\kappa_2\int_{t_0}^{t_j}\left|g(t-s)-\sum_{k=0}^{j-1}g(t_j-t_k)1_{[t_k,t_{k+1})}(s)\right|^2ds\\
&\leq(2m_1^2(t_j-t_0)+\mathrm{Var}(L(1)))\kappa_2\int_{t_0}^{t_j}\left|g(t-s)-\sum_{k=0}^{j-1}g(t_j-t_k)1_{[t_k,t_{k+1})}(s)\right|^2ds \qquad (5.2)
\end{aligned}$$

and furthermore that

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{t_0}^{t_j}\sum_{k=0}^{j-1}g(t_j-t_k)1_{[t_k,t_{k+1})}(s)(\sigma(t_k)-\sigma(s-))dL(s)\right|^2\right]\\
&\leq\left(2m_1^2\left(\sum_{k=0}^{j-1}g(t_j-t_k)\Delta t\right)^2+\mathrm{Var}(L(1))\sum_{k=0}^{j-1}g^2(t_j-t_k)\Delta t\right)\mathbb{E}\left[|\sigma(t_0-)-\sigma(t_1-)|^2\right]. \qquad (5.3)
\end{aligned}$$

So by joining the above manipulations with

$$\begin{aligned}
&\mathbb{E}\left[\left|X_r(t)-\int_{t_0}^{t_j}\sum_{k=0}^{j-1}g(t_j-t_k)\sigma(t_k)1_{[t_k,t_{k+1})}(s)dL(s)\right|^2\right]\\
&\leq 2\mathbb{E}\left[\left|\int_{t_0}^{t_j}\left(g(t-s)-\sum_{k=0}^{j-1}g(t_j-t_k)1_{[t_k,t_{k+1})}(s)\right)\sigma(s-)dL(s)\right|^2\right]\\
&\quad+2\mathbb{E}\left[\left|\int_{t_0}^{t_j}\sum_{k=0}^{j-1}g(t_j-t_k)1_{[t_k,t_{k+1})}(s)(\sigma(t_k)-\sigma(s-))dL(s)\right|^2\right]
\end{aligned}$$

one obtains a bound for the error induced by numerical integration. Thus, the rate of convergence for numerical integration is determined by the L^2 kernel function rate (5.2) and the L^2 volatility function continuity (5.3). So for instance in the case when the kernel function g is Lipschitz continuous with a Lipschitz constant $C>0$ then the right hand side of (5.2) is bounded by

$$(2m_1^2(t_j-t_0)+\mathrm{Var}(L(1)))\kappa_2C^2(t_j-t_0)(\Delta t)^2.$$

This yields a convergence rate of $(\Delta t)^2$, the same as for our method from the previous section. However, we would like to make the following points, which we believe make our method more feasible.

Firstly the Fourier simulation scheme is an iterative method, by which we mean that at each step we can simulate the next step by means of simulating the increments in the Lévy and volatility processes and evaluating a Fourier sum. Because of the t dependence of the kernel function g this is however not the case for the numerical integration scheme. Indeed for each time increment one needs to employ all the increments of the Lévy and volatility processes to compute the next step.

Secondly the Fourier simulation scheme offers greater flexibility in that it first uses the increments of the Lévy and volatility processes to simulate the

$$\{\widehat{X}_\lambda(t_j, n\pi/\tau)\}_{j=0,n=0}^{M,N} \tag{5.4}$$

field of complex LSS processes which have exponential kernel functions. This field can in turn be employed to simulate LSS process trajectories for multiple choices of kernel functions. Since by means of calculating different $a_0, a_1, \ldots, a_N$ coefficients corresponding to multiple choices of kernel functions one may reuse the field (5.4) to simulate trajectories of multiple LSS processes. For instance one can easily simulate a parametric family of LSS processes driven by a joint Lévy process and modulated by a joint volatility process but with different kernel functions $\{g_\theta\}_{\theta\in\Theta}$.

6. Pricing path dependent options

In this section we aim to employ our results for path dependent option pricing. More specifically, we shall be concerned with pricing options of the type

$$P(T) = f\left(\int_0^T X(t)dt\right), \tag{6.1}$$

where f is a Lipschitz continuous function with Lipschitz constant $C > 0$. We remark that, selecting $f(x) = \max(x/T - K, 0)$ and $f(x) = \max(K - x/T, 0)$ respectively yield so-called Asian call and put options with strike price $K > 0$. Options of this type have been traded at the Nordic electricity exchange NordPool for some time around the year 2000 (see Weron [22]). Here, $X(t)$ denotes deseasonalised spot price and we shall assume arithmetic spot dynamics, that is, that the spot is modelled as an LSS process

$$X(t) = \int_{-\infty}^t g(t-s)\sigma(s-)dL(s). \tag{6.2}$$

Under a given risk-neutral measure Q such that $P(T) \in L^1(Q)$, and given a risk-free asset (a bank account) yielding a continuously compounded rate of return $r > 0$ Benth et al. [12] define the option price $P(T)$ at time $t \leq T$ as

$$P(T) = \mathrm{e}^{-r(T-t)}\mathbb{E}_Q\left[f\left(\int_0^T X(t)dt\right)|\mathcal{F}_t\right]. \tag{6.3}$$

Suppose that $\widetilde{X}(t)$ is an approximation of the spot price dynamics $X(t)$, e.g. obtained by the method proposed in the previous section. We have the following result.

Proposition 6.1. *Given two LSS processes $X(t)$ and $\widetilde{X}(t)$ and a Lipschitz continuous function f, with Lipschitz constant $C > 0$, it holds that*

$$\mathbb{E}\left[\left|f\left(\int_0^T X(t)dt\right) - f\left(\int_0^T \widetilde{X}(t)dt\right)\right|\right] \leq CT^{1/2}\left(\int_0^T \mathbb{E}\left[|X(t) - \widetilde{X}(t)|^2\right]dt\right)^{1/2}.$$

Proof. By the Lipschitz continuity of f, the Cauchy-Schwarz inequality, Jensen's inequality and Fubini it follows that

$$\begin{aligned}\mathbb{E}\left[\left|f\left(\int_0^T X(t)dt\right) - f\left(\int_0^T \widetilde{X}(t)dt\right)\right|\right] &\leq C\mathbb{E}\left[\left|\int_0^T (X(t) - \widetilde{X}(t))dt\right|\right] \\ &\leq CT^{1/2}\mathbb{E}\left[\left(\int_0^T |X(t) - \widetilde{X}(t)|^2 dt\right)^{1/2}\right] \\ &\leq CT^{1/2}\left(\int_0^T \mathbb{E}\left[|X(t) - \widetilde{X}(t)|^2\right]dt\right)^{1/2}.\end{aligned}$$

□

The above Proposition together with Propositions 4.1 and 4.3 give us a convergence rate for employing our method to price path dependent options. So, in particular we are able approximate the options (6.1) in L^1 by means of simulating the spot price dynamics using the method presented in the previous section. Now, recalling our LSS process approximation (3.10) and using the notation of section 4, we approximate the spot price $X(t)$ by means of

$$X(t) \approx \frac{a_0}{2}\widehat{X}_{\lambda,r}(t,0) + \mathrm{Re}\sum_{n=1}^{N} a_n \widehat{X}_{\lambda,r}(t, n\pi/\tau),$$

for any $t \in \mathbb{R}$. This yields

$$\int_0^T X(t)dt \approx \frac{a_0}{2}\int_0^T \widehat{X}_{\lambda,r}(t,0)dt + \mathrm{Re}\sum_{n=1}^{N} a_n \int_0^T \widehat{X}_{\lambda,r}(t, n\pi/\tau)dt, \tag{6.4}$$

for any $T > 0$. Now furthermore for any $y \in \mathbb{R}$ we may employ Fubini to conclude that

$$\begin{aligned}\int_0^T \widehat{X}_{\lambda,r}(t,y)dt &= \int_0^T \int_r^t \mathrm{e}^{(\mathrm{i}y-\lambda)(t-s)}\sigma(s-)dL(s)dt \\ &= \int_r^T \int_{s\vee 0}^T \mathrm{e}^{(\mathrm{i}y-\lambda)(t-s)}dt\sigma(s-)dL(s) \\ &= \frac{1}{\mathrm{i}y-\lambda}\left(\widehat{X}_{\lambda,r}(T,y) - \widehat{X}_{\lambda,r}(0,y) - \int_0^T \sigma(s-)dL(s)\right).\end{aligned}$$

Thus, plugging this into (6.4) yields

$$\int_0^T X(t)dt \approx \frac{a_0}{2}\frac{\widehat{X}_{\lambda,r}(T,0)-\widehat{X}_{\lambda,r}(0,0)-\int_0^T \sigma(s-)dL(s)}{-\lambda} + \mathrm{Re}\sum_{n=1}^{N} a_n \frac{\widehat{X}_{\lambda,r}(T,n\pi/\tau)-\widehat{X}_{\lambda,r}(0,n\pi/\tau)-\int_0^T \sigma(s-)dL(s)}{\mathrm{i}n\pi/\tau-\lambda},$$

for any $T>0$. Now let us illustrate the error induced by this estimation by considering an example. Note that we may employ Fubini to conclude that

$$\begin{aligned}\int_0^T X(t)dt &= \int_0^T \int_{-\infty}^t g(t-s)\sigma(s-)dL(s)dt \\ &= \int_{-\infty}^T \int_0^T 1_{[0,t]}(s)g(t-s)\sigma(s-)dtdL(s) \\ &= \int_{-\infty}^T \int_{s\vee 0}^T g(t-s)dt\sigma(s-)dL(s) =: \int_{-\infty}^T G(T,s)\sigma(s-)dL(s), \end{aligned} \tag{6.5}$$

thus in general yielding a volatility modulated Volterra (VMV) process with kernel function $G(T,s)=\int_{s\vee 0}^T g(t-s)dt$. For the sequel note in particular we note that

$$A(T)=\frac{1}{T}\int_0^T X(t)dt = \frac{1}{T}\int_{-\infty}^T G(T,s)\sigma(s-)dL(s), \tag{6.6}$$

where $A(T)$ denotes the average of the spot over $[0,T]$. We shall use the above calculations as a benchmark in simulation studies below, when we apply our own method. To illustrate that let us consider applying our proposed simulation scheme from the previous sections to simulate the spot price dynamics. As a benchmark, let us consider the case when $L=W$ is standard Brownian motion, $\sigma=1$ and $f(x)=\max(x/T-K,0)$. Then it holds that $\int_0^T X(t)dt=\int_{-\infty}^T G(T,s)dW(s)$ is a Gaussian process and the option we want to price is an Asian call option. It holds that

$$\begin{aligned}&\mathbb{E}\left[\max\left(\frac{1}{T}\int_0^T X(t)dt-K,0\right)|\mathcal{F}_t\right] \\ &=\mathbb{E}\left[\max\left(\frac{1}{T}\int_t^T G(T,s)dW(s)-\left(K-\frac{1}{T}\int_{-\infty}^t G(T,s)dW(s)\right),0\right)|\mathcal{F}_t\right].\end{aligned}$$

Now let us assume that the conditional probability measure $\mathbb{P}(\cdot|\mathcal{F}_t)=\mathbb{E}[1.|\mathcal{F}_t]$ generated by the conditional expectation operator is regular (see e.g. Ash and Doléans-Dade [1]). Then under $\mathbb{P}(\cdot|\mathcal{F}_t)$ for each $\omega\in\Omega$ it holds that the random variable $T^{-1}\int_t^T G(T,s)dW(s)-(K-T^{-1}\int_{-\infty}^t G(T,s)dW(s))$ is normally distributed with mean $K-T^{-1}\int_{-\infty}^t G(T,s)dW(s,\omega)$ and variance $T^{-2}\int_t^T G^2(T,s)ds$ under the probability measure $\mathbb{P}(\cdot|\mathcal{F}_t)(\omega)$. Thus for our purposes, note that if Z is a normally distributed random variable with mean μ and variance

ρ^2 it holds that

$$\begin{aligned}\mathbb{E}[\max(Z,0)] &= \frac{1}{\sqrt{2\pi}\rho}\int_0^\infty u\mathrm{e}^{-\frac{(u-\mu)^2}{2\rho^2}}du \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\mu/\rho}^\infty (\mu+\rho x)\mathrm{e}^{-\frac{x^2}{2}}dx \\ &= \mu(1-\Phi(-\mu/\rho)) + \frac{\rho}{\sqrt{2\pi}}\int_{-\mu/\rho}^\infty x\mathrm{e}^{-\frac{x^2}{2}}dx \\ &= \mu\Phi(\mu/\rho) + \frac{\rho}{\sqrt{2\pi}}\exp(-\mu^2/(2\rho^2)),\end{aligned}$$

where Φ denotes the cumulative distribution function of the standard normal distribution. Thus we conclude that

$$\begin{aligned}&\mathbb{E}\left[\max\left(\frac{1}{T}\int_0^T X(t)dt - K, 0\right)|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[\max\left(\frac{1}{T}\int_t^T G(T,s)dW(s) - \left(K - \frac{1}{T}\int_{-\infty}^t G(T,s)dW(s)\right), 0\right)|\mathcal{F}_t\right] \\ &= \mu_{t,T}\Phi(\mu_{t,T}/\rho_{t,T}) + \frac{\rho_{t,T}}{\sqrt{2\pi}}\exp(-\mu_{t,T}^2/(2\rho_{t,T}^2)), \qquad (6.7)\end{aligned}$$

where

$$\mu_{t,T} = K - \frac{1}{T}\int_{-\infty}^t G(T,s)dW(s) \quad \text{and} \quad \rho_{t,T}^2 = \frac{1}{T^2}\int_t^T G^2(T,s)ds \qquad (6.8)$$

denote the mean and variance processes respectively. Thus yielding an explicit pricing formula in the Gaussian case.

To illustrate this let us consider a Brownian motion driven LSS process, with $\sigma = 1$, driven by a scaled gamma kernel (2.8), on the form

$$g(x) = Cx^{\nu-1}\mathrm{e}^{-\alpha x},$$

where $C > 0$ is a constant, $1/2 < \nu < 1$ and $\alpha > 0$. In this case we find that if $s \geq 0$

$$G(T,s) = \frac{C}{\alpha^\nu}\gamma(\nu, \alpha(T-s)),$$

where $\gamma(\nu,x) = \int_0^x u^{\nu-1}\mathrm{e}^{-u}du$ denotes the lower incomplete gamma function. From which we may easily (numerically) evaluate the variance $\rho_{t,T}^2$ for any given $t \leq T$. Notice that due the singularity of the gamma kernel function at zero, we need to amend it in a neighborhood close to the origin, to make our method applicable. We achieve this for a given $\epsilon > 0$ by means of considering the function

$$g_\epsilon(x) = \begin{cases} \phi_0(x) & \text{if } x \in [0,\epsilon] \\ g(x) & \text{if } x \geq \epsilon \end{cases}$$

where ϕ_0 is a 5th degree interpolating polynomial with coefficients determined by $\phi_0^{(j)}(0) = g^{(j)}(\epsilon)$ and $\phi_0^{(j)}(\epsilon) = g^{(j)}(\epsilon)$ for $j = 0,1,2$. Furthermore, as described above by (4.11) and

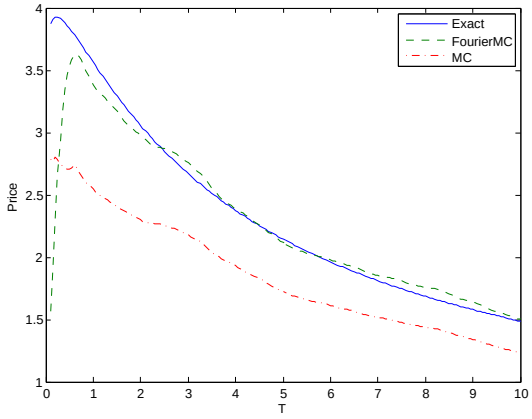

FIGURE 2. The price curve $T \mapsto \mathbb{E}[\max(A(T) - K, 0)|\mathcal{F}_t]$, where $A(T)$ is given by (6.6), $K = 1$, $t = 0$, $L = W$ on $[0, 10]$ and $L = 0$ on $(-\infty, 0)$, $\sigma = 1$ and $g(x) = Cx^{\nu-1}\mathrm{e}^{-\alpha x}$, where $C = 10$, $\alpha = 1$ and $\nu = 0.55$. As obtained by the explicit Gaussian formula (6.7), numerically integration to evaluate $X(t)$ and the approximation (6.4), with $\lambda = 1.9$, $\Delta t = 0.05$ and $N = 30$, where the expectations are obtained by averaging over 1000 Brownian motion paths.

(4.12) with $k = 1$ we interpolate g_ϵ to zero in the tail from $\tau_0 = 10$ to $\tau = 11$. In figure 2 we see that by means of applying our method with $\lambda = 1.9$, $\Delta t = 0.05$ and $N = 30$ and comparing to the explicit path one can get a better result by using the Fourier method than by using numerical integration. The reason for the difference on this scale is the singularity in the kernel function, but one must however take some care when selecting $\lambda > 0$, since not every choice of it will yield a good approximation. To this end one could use the convergence rate results of the current paper as guidelines in the search for the optimal λ.

Obtaining the path in Matlab by means of numerical integration using a convolution routine took 0.1182 seconds, whereas obtaining the path by means of the Fourier method using a two dimensional convolution took 2.7791 seconds. The calculations were performed on a standard laptop computer. However, as pointed out in the previous section one can re-use the field (5.4) to simulate the price for different kernel function. Thus if one is interested in the price for a LSS process where the kernel function depends on a parameter, then the Fourier method has an advantage. Furthermore if one wants to simulate the next time step, then this is more easily accomplished by our method than numerical integration. Finally, as we have seen, by selecting an appropriate $\lambda > 0$ one may get a better approximation by means of the Fourier method than the by means of numerical integration.

7. Conclusion

We have introduced numerical simulation scheme for LSS processes which is based on considering a Fourier integral representation of the corresponding kernel function and approximating it by means of a sum. We have analysed the convergence rate of the method and compared it to numerical integration. Finally we have applied our method to price path dependent options and benchmarked our method and numerical integration against the explicit price.

In future studies we aim to apply our findings to study how this can be applied to simulating forward price fields (as derived by Barndorff-Nielsen et al. [2]), by means of a parametric family of LSS processes, with varying kernel functions.

References

[1] Ash, R. B. and Doléans-Dade, C. *Probability and Measure Theory*, Academic Press, (2000)

[2] Barndorff-Nielsen, O.E., Benth, F. E., and Veraart, A. E. D. Modeling energy spot prices by volatility modulated Lévy-driven Volterra processes. Accepted for publication in *Bernoulli*, (2012).

[3] Barndorff-Nielsen, O. E., Benth, F. E., and Veraart, A. E. D. Modelling electricity forward markets by ambit fields. Available at: http://ssrn.com/abstract=1938704, (2011).

[4] Barndorff-Nielsen, O. E., and Schmiegel, J. Lévy-based Tempo-Spatial Modelling; with Applications to Turbulence. *Uspekhi Mat. NAUK* **59**, 65–91, (2004)

[5] Barndorff-Nielsen, O. E., and Schmiegel, J. Brownian semistationary processes and volatility/intermittency. In *Advanced Financial Modeling*, H. Albrecher, W. Runggaldier and W. Schachermayer (eds.), Radon Series on Computational and Applied Mathematics 8, W. de Gruyter, Berlin, 1–26, (2009).

[6] Basse, A. and Pedersen, J. Lévy driven moving averages and semimartingales. *Stochastic processes and their applications*, 119, 2970–2991, (2009).

[7] Basse-O'Connor, A., Graversen, S.-E. and Pedersen, J. A unified approach to stochastic integration on the real line. To appear in *Theory of Probability and Its Applications*, (2012)

[8] Benth, F. E., Kallsen, J., and Meyer-Brandis, T. A non-Gaussian Ornstein-Uhlenbeck process for electricity spot price modelling and derivatives pricing. *Applied Math. Finance*, 14(2), 153–169, (2007).

[9] Benth, F. E., Kiesel, R., and Nazarova, A. A critical empirical study of three electricity spot price models. *Energy Economics*, Volume 34, Issue 5, 1589-1616, (2012).

[10] Benth, F. E., Eyjolfsson, H. Stochastic Modelling of Power Markets Using Stationary Processes. To appear in the *Proceedings of the Ascona Conference 2011*, R. Dalang, M. Dozzi, and F. Russo (eds.). Springer, (2012)

[11] Benth, F.E., Lempa, J. and Nilssen, T.K., On optimal exercise of swing options in electricity markets. *J. Energy Markets*, 4(4), 3–28, (2012).

[12] Benth, F.E., Šaltytė Benth, J. and Koekebakker, J., *Stochastic Modelling of Electricity and Related Markets*, World Scientific, 2008.

[13] Bjerksund, P., Rasmussen, H., and Stensland, G. Valuation and risk management in the Norwegian electricity market, in *Energy, Natural Resources and Environmental Economics*, pp. 167–185, E. Bjørndal, M. Bjørndal, P. M. Pardalos and M. Rönnqvist (eds.), Springer Verlag, (2010).

[14] Folland, G. B. *Real Analysis – Modern Techniques and their Applications*, J. Wiley& Sons, (1984).

[15] Garcia, I., Klüppelberg, C., and Müller, G. Estimation of stable CARMA models with an application to electricity spot prices.*Statistical Modelling*, 11(5), 447–470, (2011).

[16] Jónsdóttir, K., Ý., Schmiegel, J., and Vedel Jensen, E.,B. Lévy based growth models, *Bernoulli* **14**(1), 62–90, (2008).

[17] Klüppelberg, C., Meyer-Brandis, T., and Schmidt, A. Electricity spot price modelling with a view towards extreme spike risk. *Quantitative Finance* **10** 963–974.
[18] Lucia, J., and Schwartz, E. S. Electricity prices and power derivatives: evidence from the Nordic Power Exchange. *Rev. Derivatives Research*, 5(1), 5–50, (2002).
[19] Protter, Ph. *Stochastic Intergation and Differential Equations*, 2nd Edition Version 2.1, Springer Verlag, (2005).
[20] Samuelson, P. Proof that properly anticipated prices fluctuate randomly. *Industrial Manag. Review*, 6, 41–44, (1965).
[21] Schwartz, E. The Stochastic Behaviour of Commodity Prices: Implications for Valuation and Hedging. *The Journal of Finance* **52**(3), 923–973, (1997).
[22] Weron, R. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. John Wiley & Sons, Chichester, (2006).

Fred Espen Benth, Centre of Mathematics for Applications (CMA), University of Oslo, P.O. Box 1053 Blindern, N-0316 Oslo, Norway
E-mail address: fredb@math.uio.no

Heidar Eyjolfsson, Centre of Mathematics for Applications (CMA), University of Oslo, P.O. Box 1053 Blindern, N-0316 Oslo, Norway
E-mail address: heidar.eyjolfsson@cma.uio.no

Almut E. D. Veraart, Department of Mathematics, Imperial College London, & CREATES, 180 Queen's Gate, London, SW7 2AZ, UK
E-mail address: a.veraart@imperial.ac.uk