# 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 simulation of Lévy semistationary (LSS) processes in the context of power markets. A disadvantage of applying numerical integration to obtain trajectories of LSS processes is that such a scheme is not iterative. We address this problem by introducing and analyzing a Fourier simulation scheme for obtaining trajectories of these processes in an iterative manner. Furthermore, we demonstrate that our proposed scheme is well suited for simulation of a wide range of LSS processes, including, in particular, LSS processes indexed by a kernel function which is steep close to the origin. Finally, we put our simulation scheme to work for simulating the price of path-dependent options to demonstrate the advantages of the proposed Fourier simulation scheme.

**Key words.** Lévy semistationary processes, spot modeling, Fourier inversion, numerical simulation, path-dependent options

**AMS subject classifications.** 60G10, 65C05, 65C20, 65T40, 91B28

**DOI.** 10.1137/130905320

**1. Introduction.** A Lévy semistationary process 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 introduced in the context of modeling electricity spot prices by Barndorff-Nielsen, Benth, and Veraart [3]. The special case of Brownian semistationary processes was studied by Barndorff-Nielsen and Schmiegel [6] to model turbulence in physics.

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 prices, the price is pushed up by increased demand. Modeling commodity spot prices by means of *mean-reverting* processes has been studied by Schwartz [24], and in a more general setting by means of a sum of Lévy process driven Ornstein–Uhlenbeck processes, corresponding to different mean-reversion coefficients, in papers by Benth, Kallsen, and Meyer-Brandis [11] and Klüppelberg, Meyer-Brandis, and

Schmidt [19], respectively. 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 and hence the terminology *semistationary*. Thus a Lévy semistationary processes which is modulated by a 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 studying the corresponding approximation errors. The current paper reintroduces an approximation method presented in that paper and analyzes its convergence. The approximation method in question consists 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 analyze its convergence.

The main purpose of this paper is to analyze the Fourier approximation method and to illustrate its use for simulating prices of derivatives under Lévy semistationary electricity spot dynamics. We show that the Fourier simulation method and numerical integration have a similar convergence behavior. However, an advantage of the Fourier simulation method is that it is an iterative scheme, by which we mean that in order to simulate a point at time $t+\delta$ 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 with the numerical integration approach in which one needs to perform a complete reintegration 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. The Fourier approximation method is also 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. Moreover, we argue that the Fourier method we present here is particularly well suited for simulating Lévy semistationary processes with kernel functions that are steep close to the origin. Finally, we discuss an application of our method where we simulate 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 [25]).

The remaining part of this paper is structured as follows. In the next section we set the stage by introducing Lévy semistationary processes and their basic properties, 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 analyze the error induced by our approximations in the mean square sense. Next, in section 5, we compare our method to more direct alternative methods of simulation and illustrate the advantages of our approach. In section 6, we apply our method to simulate prices for path-dependent options and thereby illustrate our method, as well as demonstrate its relevance for applications. Finally, section 7 concludes.

**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., the $\sigma$-algebras $\mathcal{F}_t$ include all the null sets of $\mathcal{F}$, and 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

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

for $t \in \mathbb{R}$, where $L$ is a two-sided square integrable Lévy process, $g$ is a Borel measurable, real-valued, nonnegative deterministic function on $\mathbb{R}_+$, the positive half line including the origin, such that

$$(2.2) \qquad g \in L^1(\mathbb{R}_+) \cap L^2(\mathbb{R}_+),$$

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 modeled as a *stationary* process *independent* of $L$. Denote, in what follows, the first two moments of the Lévy process $L$ and the volatility process $\sigma$ by

$$(2.3) \qquad m_j := \mathbb{E}[L^j(1)] \quad \text{and} \quad \kappa_j := \mathbb{E}[\sigma^j(0)],$$

respectively, where $j = 1, 2$ and we assume that these moments are finite. We remark that if $g$ is bounded, nonnegative, 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 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 noncentered 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 [21] and Basse-O'Connor, Graversen, and Pedersen [8]).

The characteristic function of $X$ is easily computed by conditioning on the filtration generated by the volatility process $\sigma$:

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

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 *distinguished logarithm* which fulfills $\exp(\psi(\theta)) = \mathbb{E}[\exp(\mathrm{i}\theta L(1))]$ in the sense of p. 33 in Sato [23]. Note, in particular, that if $\sigma = 1$, then the cumulant function of $X(t)$ is

$$\mathrm{Log}\,\mathbb{E}[\exp(\mathrm{i}\theta X(t))] = \int_{0}^{\infty} \psi(\theta g(s))ds,$$

where Log in the formula above should as before be understood as the *distinguished logarithm*. 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)

$$\mathbb{E}[X^2(t)] = -(\psi'(0))^2 \mathbb{E}\left[\left(\int_{-\infty}^t g(t-s)\sigma(s-)ds\right)^2\right] - \psi''(0)\mathbb{E}\left[\int_{-\infty}^t g^2(t-s)\sigma^2(s-)ds\right]$$

$$= m_1^2 \mathbb{E}\left[\left(\int_{-\infty}^t g(t-s)\sigma(s-)ds\right)^2\right] + \mathrm{Var}(L(1))\kappa_2||g||_{L^2(\mathbb{R}_+)}^2$$

(2.6)
$$= m_1^2 \int_0^\infty \int_0^\infty g(u)g(v)\mathbb{E}[\sigma(0)\sigma(|u-v|)]dudv + \mathrm{Var}(L(1))\kappa_2||g||_{L^2(\mathbb{R}_+)}^2.$$

Notice that (2.6) is not time-dependent and $X(t)$ 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 noncentered 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 function $g$ 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, $\sigma$ and $L$, we define a complex-valued LSS process as

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

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 $\sigma$.

*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 $Z(t) = \int_{-\infty}^t g(t-s)\rho(s-)dL(s)$. It holds that*
    *(i)*

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

*where equality is obtained when the Lévy process is centered, i.e., when $m_1 = 0$, and*
    *(ii)*

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

*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 $h > 0$:

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

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

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) we have

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

which concludes the proof. ■

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 $g_\epsilon$ rather than $g$ itself, where $\epsilon$ 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 [6]). Suppose $L$ is a centered Lévy process and $g$ is of the form

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

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 [7]), not a semimartingale process. By Lemma 2.1 we may approximate $X(t)$ with a semimartingale LSS process that has the nonsingular 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 $\nu$ is to $1/2$, the slower the rate is. If we want to simulate from $X(t)$, one could 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 $g_\epsilon$ 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 nonnegative kernel function $g$ consider

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

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

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

By (2.2), the integrals on the right-hand side will tend to zero as $\tau$ increases. This gives the interpretation of LSS processes as limits of moving averages 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. To this end, for a given LSS process and $t, r \in \mathbb{R}$ such that $r \leq t$ consider the sum

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

By the integrability condition (2.2) 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

$$(3.2) \qquad X_r(t) := \int_{r}^{t} g(t - s)\sigma(s-)dL(s)$$

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) in an iterative manner.

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) evaluates the kernel function $g$ only 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 $\tau_0, \tau$ such that $0 < \tau_0 < \tau$ and $T - r \leq \tau_0$ consider the function $h : \mathbb{R} \to \mathbb{R}$, defined by

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

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

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

Now let the Fourier transform of $h_\lambda$ be (see Folland [16])

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

and suppose, furthermore, that

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

Then the inverse Fourier transform exists, and we have (see Folland [16])

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

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 $h_\lambda$ 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 nonnegative real numbers and may be approximated as

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

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 nonperiodic 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 $2\tau$. 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

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

for $j, k = 0, \ldots, N$. Given $\tau > 0$, the orthogonality condition (3.7) is satisfied if $\xi_n = n\pi/\tau$ for $n = 0, \ldots, N$, which yields 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

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

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

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

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 what follows we shall work under the orthogonality assumption (3.7) together with

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

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

Furthermore, the approximation (3.6) may be written as

$$(3.9) \quad h(x) \approx e^{-\lambda|x|} \left( \frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(n\pi x/\tau) \right), \quad \text{with } a_n = \frac{\widehat{h}_\lambda(n\pi/\tau)}{\tau} \quad \text{for } n = 0, \ldots, N,$$

which 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 $N$th partial Fourier series of $h_\lambda$ times $e^{-\lambda|x|}$ if it is extended as a periodic function from $[-\tau, \tau]$ to the entire

real line (see, e.g., Folland [16]). Notice here that, since we have defined $h_\lambda$ 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

$$\int_r^t g(t-s)\sigma(s-)dL(s) \approx \int_r^t e^{-\lambda(t-s)} \left( \frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(n\pi(t-s)/\tau) \right) \sigma(s-)dL(s)$$

$$(3.10) \qquad = \frac{a_0}{2}\widehat{X}_{\lambda,r}(t,0) + \mathrm{Re} \sum_{n=1}^{N} a_n \widehat{X}_{\lambda,r}(t, n\pi/\tau),$$

where

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

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 $\sigma$ than we have assumed here), whereas for any $\lambda > 0$ and $r < T$ we have by Lemma 2.1 that

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

$$(3.11) \qquad = \frac{2m_1^2 \kappa_2}{\lambda^2} + \frac{\mathrm{Var}(L(1))\kappa_2}{2\lambda}.$$

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. Indeed, we can obtain a trajectory stepwise as follows.

Fix $\delta > 0$, and we find

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

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

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

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

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 $\delta$. In principle, we could simulate $\widehat{X}_{\lambda,r}(t,y)$ exactly.

For example, if $\sigma(s) = 1$, then the residual is an independent outcome of a random variable $Z$ with cumulant function given by

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

Thus, the error is from the numerical integration in the 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).

We employ the residual approximation (3.13) together with (3.12) to derive an algorithm for obtaining discrete trajectories of arbitrary LSS processes as follows. Notice that in order to use the approximation (3.13) in practice one needs to know the value of the volatility process $\sigma$ in the left limit of the point $t$. Hence, employing (3.12) and approximation (3.13) recursively on a given discrete time grid $t_0 < t_1 < \cdots < t_M$, with a given initial value $\widehat{X}_{\lambda,r}(t_0, y) = \hat{x}_0(y)$ for some $y \in \mathbb{R}$, requires that the values $\sigma(t_0-), \sigma(t_1-), \ldots, \sigma(t_{M-1}-)$ be known. We shall thus in what follows, when discussing how to obtain a trajectory of $\widehat{X}_{\lambda,r}(\cdot, y)$ for a given $y \in \mathbb{R}$, assume that the corresponding volatility process discrete trajectory $\sigma$ is known. As a motivation for this we remark that in the setting of Barndorff-Nielsen, Benth, and Veraart [3] the authors employ LSS processes (and slightly more general processes) as a general modeling framework for electricity spot prices, with the volatility process being modeled as a square root of an LSS process driven by a Lévy subordinator process where the volatility of the LSS volatility process is constant. Thus in cases where the volatility process $\sigma$ is unknown, but modeled as the square root of an LSS process with constant volatility, one can simply simulate the volatility process on the given time grid, e.g., by using the method proposed here.

Now let us introduce an Euler–Maruyama-type scheme for simulating trajectories of $\widehat{X}_{\lambda,r}(\cdot, y)$ inspired by (3.12) and the approximation (3.13). Let $y \in \mathbb{R}$, and consider the equidistant time grid $t_0 < t_1 < \cdots < t_M$ with step size $\Delta t > 0$. Suppose that the values $\{\sigma(t_j-)\}_{j=0}^{M-1}$ of the volatility process on the discrete time grid are known. Given an initial value $\widehat{X}_{\lambda,r}(t_0, y) = \hat{x}_0(y)$, consider the time series $\{\eta_j(y)\}_{j=0}^M$ defined by the iterative scheme $\eta_0(y) := \hat{x}_0(y)$ and

$$(3.14) \qquad \eta_j(y) := \mathrm{e}^{(-\lambda+\mathrm{i}y)\Delta t}(\eta_{j-1}(y) + \sigma(t_{j-1}-)\Delta L(t_{j-1}))$$

for $j = 1, \ldots, M$, where $\Delta L(t_{j-1}) := L(t_j) - L(t_{j-1})$. A notable advantage of the above scheme is that we have the same residual term for every $y$, apart from a deterministic scaling by a complex exponential. Thus, given a simulated trajectory $\{\sigma(t_{j-1}-)\Delta L(t_{j-1})\}_{j=1}^M$ and a discretization $\{y_n\}_{n=0}^N$ in the Fourier domain, we can easily obtain a field of simulated values indexed by $(j, n)$. In particular, in light of the discussion in the current section, we can take $y_n = n\pi/\tau$ for an appropriately selected $\tau > 0$. Such a field can in turn be employed to simulate trajectories of general (truncated) LSS processes as follows.

To simulate a discrete trajectory $X_r(t_0), X_r(t_1), \ldots, X_r(t_M)$ given the discrete volatility process trajectory $\{\sigma(t_j-)\}_{j=0}^{M-1}$, we apply the following algorithm.

Algorithm 1. *For each $t_j$, where $j = 1, \ldots, M$, do the following:*
   1. *Simulate $\Delta L(t_{j-1})$.*

2. *For each* $n = 0, \ldots, N$, *simulate* $\eta_j(n\pi/\tau)$ *from* $\eta_{j-1}(n\pi/\tau)$ *and* $\Delta L(t_{j-1})$ *using* (3.14).

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: In step 2 above, we just need to have stored the $N + 1$ values of $\eta_{j-1}(n\pi/\tau)$ from the previous time $t_{j-1}$ along with the simulated $\Delta L(t_{j-1})$, 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 the step size $\Delta t$ in the time domain, whereas with the Fourier technique the accuracy is connected to the number of sampling points, $N$, in the Fourier domain. That is, when employing numerical integration to obtain a trajectory of an LSS process driven by a "nasty" kernel function, e.g., one having a singularity at the origin, one can increase the precision by making the time partition finer in a neighborhood around the singularity. By contrast, for any $y \in \mathbb{R}$, the complex LSS process $t \mapsto \widehat{X}_{\lambda,r}(t, y)$ is driven by an exponential function with a negative real part, which is a well-behaved continuous function and thus well suited for simulation on an equidistant time grid. Thus by using our method, the precision of the method for ill-behaved functions is connected to the number of sampling points, $N$, in the Fourier domain.

**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 then approximating the resulting integral. Furthermore, we have seen that by splitting a general LSS process into a sum of two integrals (3.1), 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 analyze the error induced by our approximation.

For a fixed $\lambda > 0$ and $N \geq 1$ we denote by

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

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

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

(see Folland [16]). We may employ the sum representation (3.1) to get

$$
\mathbb{E}\left[ \left| X(t) - \int_{t_0}^{t} h_N(t - s)\sigma(s-)dL(s) \right|^2 \right]
$$

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

By Lemma 2.1 it holds that

$$\mathbb{E}\left[\left|\int_{-\infty}^{t_0} g(t-s)\sigma(s-)dL(s)\right|^2\right]$$

$$(4.3) \qquad \leq 2m_1^2\kappa_2||g(t-t_0+\cdot)||_{L^1(\mathbb{R}_+)}^2 + \mathrm{Var}(L(1))\kappa_2||g(t-t_0+\cdot)||_{L^2(\mathbb{R}_+)}^2$$

and

$$\mathbb{E}\left[\left|\int_{t_0}^{t} (g(t-s) - h_N(t-s))\sigma(s-)dL(s)\right|^2\right]$$

$$(4.4) \qquad \leq 2m_1^2\kappa_2||g - h_N||_{L^1([0,t-t_0])}^2 + \mathrm{Var}(L(1))\kappa_2||g - h_N||_{L^2([0,t-t_0])}^2.$$

Clearly, by (2.2), 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

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

$$\leq 2\int_0^{t-t_0} e^{-2\lambda x} \left| \sum_{n=N+1}^{\infty} a_n \cos(n\pi x/\tau) \right|^2 dx$$

$$\leq 2\int_0^{t-t_0} e^{-2\lambda x} \left( \sum_{n=N+1}^{\infty} |a_n \cos(n\pi x/\tau)| \right)^2 dx$$

$$\leq \frac{1 - e^{-2\lambda(t-t_0)}}{\lambda} \left( \sum_{n=N+1}^{\infty} |a_n| \right)^2.$$

Similarly, by using the Cauchy–Schwarz inequality we furthermore find that

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

$$\leq (t-t_0)\frac{1 - e^{-2\lambda(t-t_0)}}{\lambda} \left( \sum_{n=N+1}^{\infty} |a_n| \right)^2.$$

Combining these findings with (4.4) we thus find that

$$\mathbb{E}\left[\left|\int_{t_0}^{t} (g(t-s) - h_N(t-s))\sigma(s-)dL(s)\right|^2\right]$$

$$(4.5) \qquad \leq \left(2m_1^2(t - t_0) + \mathrm{Var}(L(1))\right)\kappa_2 \frac{1 - \mathrm{e}^{-2\lambda(t-t_0)}}{\lambda}\left(\sum_{n=N+1}^{\infty}|a_n|\right)^2.$$

Thus, unsurprisingly, the Fourier series kernel function approximation (3.9) is most dependent on the choice of $\lambda > 0$ and on the number of 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)? More precisely, 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

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

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

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

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

$$(4.8) \qquad\qquad\qquad h_\lambda'(\tau) = 0,$$

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 (4.7) and (4.8) combined yield

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

Now, for our purposes, recalling the definition (3.3), this means that the interpolating function $\phi$ 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*

$$\mathbb{E}\left[\left|\int_{t_0}^{t}(g(t - s) - h_N(t - s))\sigma(s-)dL(s)\right|^2\right]$$

$$\leq \left(2m_1^2(t - t_0) + \mathrm{Var}(L(1))\right)\kappa_2\frac{1 - \mathrm{e}^{-2\lambda(t-t_0)}}{\lambda}\frac{\tau^{4k-2}\|h_\lambda^{(2k)}\|_{L^1([-\tau,\tau])}^2}{\pi^{4k}}\left(\sum_{n=N+1}^{\infty}\frac{1}{n^{2k}}\right)^2.$$
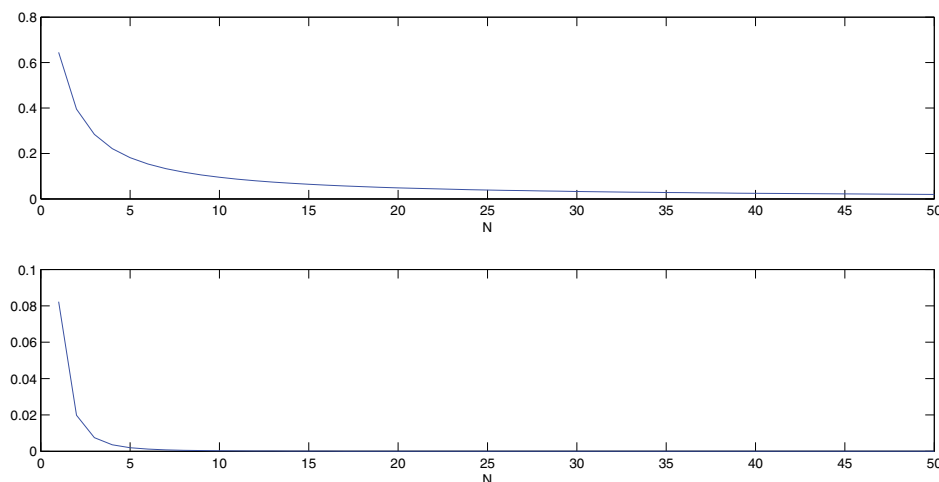
**Figure 1.** *The convergence rate of the series* (4.10) *for* $k = 1, 2$. *The upper graph depicts the function* $N \mapsto \pi^2/6 - \sum_{n=1}^{N} n^{-2}$. *The lower graph shows the function* $N \mapsto \pi^4/90 - \sum_{n=1}^{N} n^{-4}$.

From the above proposition it is clear that the $L^2$-rate of convergence of the approximation scheme (3.10) towards the quantity to simulate is controlled by three main factors. The most apparent factor is the infinite series

$$(4.10) \qquad \sum_{n=1}^{\infty} \frac{1}{n^{2k}},$$

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 $\zeta$ 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 $h_\lambda$ function, which in turn is determined by the original kernel function $g$ and, perhaps to a bigger extent, the interpolation function $\phi$. Note that the conclusion of Proposition 4.1 holds if $h$ is a $C^{2k}$-function and $h_\lambda^{(2n-1)}(\tau) = 0$ 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 $\phi$ be an appropriately chosen polynomial. Let us outline how this can be achieved. Suppose that $g$ is a $C^{2k}$-function, and let $\phi$ be the $4k + 1$ degree polynomial

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

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

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

Equipped with this interpolation function $h$ (and thus $h_\lambda$) 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 [16, section 8.4]). But, on the other hand, a high value of $\lambda > 0$ usually leads to high values of $||h_\lambda^{(2k)}||^2_{L^1([-\tau, \tau])}$. It follows that letting $\lambda$ be "too high" can in fact increase the error, so one needs to pay some attention to the parameter selection.

Having quantified the error induced from employing a kernel function approximation of the type (3.6) by using the methodology developed in the previous section, let us move on to quantifying the error induced from employing the scheme (3.14) to obtain trajectories of $\widehat{X}_{\lambda,r}(\cdot, y)$ for $y \in \mathbb{R}$. Consider the following auxiliary result.

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

$$\left(\mathbb{E}[|\widehat{X}_{\lambda,r}(t_j, y) - \eta_j(y)|^2]\right)^{1/2} \le \left((2m_1^2(t_j - t_0) + \mathrm{Var}(L(1)))(t_j - t_0)\right)^{1/2} \left\{ \sqrt{\frac{\kappa_2(\lambda^2 + y^2)}{3}} \Delta t \right.$$

$$\left. + \left( \sup_{1 \le k \le j, s \in [t_{k-1}, t_k)} \mathbb{E}\left[|\sigma(t_{k-1}-) - \sigma(s-)|^2\right] \right)^{1/2} \right\},$$

*where $\eta_j$ is obtained by (3.14) and $1 \le j \le M$.*

*Proof.* By iterating (3.14) it holds that

$$\eta_j(y) = e^{(-\lambda+iy)j\Delta t}\hat{x}_0(y) + \sum_{k=1}^{j} e^{(-\lambda+iy)(j+1-k)\Delta t}\sigma(t_{k-1}-)\Delta L(t_{k-1})$$

$$= e^{(-\lambda+iy)j\Delta t}\hat{x}_0(y) + \int_{t_0}^{t_j} \sum_{k=1}^{j} e^{(-\lambda+iy)(j+1-k)\Delta t} 1_{[t_{k-1}, t_k)}(s)\sigma(s-)dL(s)$$

$$+ \int_{t_0}^{t_j} \sum_{k=1}^{j} e^{(-\lambda+iy)(j+1-k)\Delta t}(\sigma(t_{k-1}-) - \sigma(s-))1_{[t_{k-1}, t_k)}(s)dL(s).$$

Now observe that by Lemma 2.1, the Cauchy–Schwarz inequality, and the inequality

$$|e^{zx} - e^{zy}|^2 = \left|\int_{x \wedge y}^{x \vee y} z e^{zu} du\right|^2 \le |z|^2|x - y|^2,$$

where $x, y \in \mathbb{R}$ are nonnegative, $z \in \mathbb{C}$, and $\mathrm{Re}\, z \le 0$, it holds that

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

$$\le 2m_1^2\kappa_2 \left(\int_{t_0}^{t_j} \left|e^{(-\lambda+iy)(t_j-s)} - \sum_{k=1}^{j} e^{(-\lambda+iy)(j+1-k)\Delta t} 1_{[t_{k-1}, t_k)}(s)\right| ds\right)^2$$

$$+ \operatorname{Var}(L(1))\kappa_2 \int_{t_0}^{t_j} \left| \mathrm{e}^{(-\lambda+\mathrm{i}y)(t_j-s)} - \sum_{k=1}^{j} \mathrm{e}^{(-\lambda+\mathrm{i}y)(j+1-k)\Delta t} 1_{[t_{k-1},t_k)}(s) \right|^2 ds$$

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

$$\leq (2m_1^2(t_j-t_0) + \operatorname{Var}(L(1)))\kappa_2(\lambda^2+y^2) \int_{t_0}^{t_j} \left| s - \sum_{k=1}^{j} t_{k-1} 1_{[t_{k-1},t_k)}(s) \right|^2 ds$$

$$= (2m_1^2(t_j-t_0) + \operatorname{Var}(L(1)))\kappa_2(\lambda^2+y^2)j\frac{(\Delta t)^3}{3}$$

$$= (2m_1^2(t_j-t_0) + \operatorname{Var}(L(1)))\kappa_2(\lambda^2+y^2)(t_j-t_0)\frac{(\Delta t)^2}{3}.$$

Furthermore, it holds by Lemma 2.1 that

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

$$\leq (2m_1^2(t_j-t_0) + \operatorname{Var}(L(1)))(t_j-t_0) \sup_{1\leq k\leq j, s\in[t_{k-1},t_k)} \mathbb{E}\left[|\sigma(t_{k-1}-) - \sigma(s-)|^2\right].$$

Putting these results together and applying Minkowski's inequality completes the proof. ∎

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 $r \leq t_0 < t_1 < \cdots < t_M$, with step size $\Delta t > 0$, it holds that*

$$\mathbb{E}\left[\left|\int_r^{t_j} h_N(t_j-s)\sigma(s-)dL(s) - \frac{a_0}{2}\eta_j(0) - \operatorname{Re}\sum_{n=1}^{N} a_n \eta_j(n\pi/\tau)\right|^2\right]$$

$$\leq (2m_1^2(t_j-t_0) + \operatorname{Var}(L(1)))(t_j-t_0)\left\{\kappa_2\left(\lambda^2\left(\frac{|a_0|}{2} + \sum_{n=1}^{N}|a_n|\right)^2 + \frac{\pi^2}{\tau^2}\left(\sum_{n=1}^{N}|a_n|n\right)^2\right)(\Delta t)^2\right.$$

$$\left. + 3\left(\frac{|a_0|}{2} + \sum_{n=1}^{N}|a_n|\right)^2 \sup_{1\leq k\leq j, s\in[t_{k-1},t_k)} \mathbb{E}\left[|\sigma(t_{k-1}-) - \sigma(s-)|^2\right]\right\},$$

*where $\eta_j$ is obtained by (3.14) and $1 \leq j \leq M$.*

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

$$\left(\mathbb{E}\left[\left|\int_r^{t_j} h_N(t_j - s)\sigma(s-)dL(s) - \frac{a_0}{2}\eta_j(0) - \mathrm{Re}\sum_{n=1}^N a_n\eta_j(n\pi/\tau)\right|^2\right]\right)^{1/2}$$

$$= \left(\mathbb{E}\left[\left|\frac{a_0}{2}(\widehat{X}_{\lambda,r}(t_j,0) - \eta_j(0)) - \mathrm{Re}\sum_{n=1}^N a_n(\widehat{X}_{\lambda,r}(t_j,n\pi/\tau) - \eta_j(n\pi/\tau))\right|^2\right]\right)^{1/2}$$

$$\leq \frac{|a_0|}{2}\left(\mathbb{E}\left[\left|\widehat{X}_{\lambda,r}(t_j,0) - \eta_j(0)\right|^2\right]\right)^{1/2} + \sum_{n=1}^N |a_n|\left(\mathbb{E}\left[\left|\widehat{X}_{\lambda,r}(t_j,n\pi/\tau) - \eta_j(n\pi/\tau)\right|^2\right]\right)^{1/2}$$

$$\leq \left((2m_1^2(t_j - t_0) + \mathrm{Var}(L(1)))(t_j - t_0)\right)^{1/2}\left\{\sqrt{\frac{\kappa_2}{3}}\left(\frac{|a_0|}{2}\lambda + \sum_{n=1}^N |a_n|\sqrt{\lambda^2 + \left(\frac{n\pi}{\tau}\right)^2}\right)\Delta t\right.$$

$$\left. + \left(\frac{|a_0|}{2} + \sum_{n=1}^N |a_n|\right)\left(\sup_{1\leq k\leq j, s\in[t_{k-1},t_k)} \mathbb{E}\left[|\sigma(t_{k-1}-) - \sigma(s-)|^2\right]\right)^{1/2}\right\}.$$

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

$$\frac{|a_0|}{2}\lambda + \sum_{n=1}^N |a_n|\sqrt{\lambda^2 + \left(\frac{n\pi}{\tau}\right)^2} \leq \lambda\left(\frac{|a_0|}{2} + \sum_{n=1}^N |a_n|\right) + \frac{\pi}{\tau}\sum_{n=1}^N |a_n|n.$$

Thus by the elementary inequality $(x + y + z)^2 \leq 3(x^2 + y^2 + z^2)$, it follows that

$$\mathbb{E}\left[\left|\int_r^{t_j} h_N(t_j - s)\sigma(s-)dL(s) - \frac{a_0}{2}\eta_j(0) - \mathrm{Re}\sum_{n=1}^N a_n\eta_j(n\pi/\tau)\right|^2\right]$$

$$\leq \left((2m_1^2(t_j - t_0) + \mathrm{Var}(L(1)))(t_j - t_0)\right)\left\{\sqrt{\frac{\kappa_2}{3}}\lambda\left(\frac{|a_0|}{2} + \sum_{n=1}^N |a_n|\right)\Delta t\right.$$

$$\left. + \sqrt{\frac{\kappa_2}{3}}\frac{\pi}{\tau}\sum_{n=1}^N |a_n|n\Delta t + \left(\frac{|a_0|}{2} + \sum_{n=1}^N |a_n|\right)\left(\sup_{1\leq k\leq j, s\in[t_{k-1},t_k)} \mathbb{E}\left[|\sigma(t_{k-1}-) - \sigma(s-)|^2\right]\right)^{1/2}\right\}^2$$

$$\leq (2m_1^2(t_j - t_0) + \mathrm{Var}(L(1)))(t_j - t_0)\left\{\kappa_2\left(\lambda^2\left(\frac{|a_0|}{2} + \sum_{n=1}^N |a_n|\right)^2 + \frac{\pi^2}{\tau^2}\left(\sum_{n=1}^N |a_n|n\right)^2\right)(\Delta t)^2\right.$$

$$\left. + 3\left(\frac{|a_0|}{2} + \sum_{n=1}^N |a_n|\right)^2 \sup_{1\leq k\leq j, s\in[t_{k-1},t_k)} \mathbb{E}\left[|\sigma(t_{k-1}-) - \sigma(s-)|^2\right]\right\}. \quad \blacksquare$$

The above proposition characterizes the error of the approximation of the trajectories $t \mapsto \widehat{X}_{\lambda,r}(t,y)$ in terms of the step size, $\Delta t$, in the time domain. Thus it gives us the complete

picture of the error induced by our approximations in the sense that together with Proposition 4.1 it bounds the error made by our approximations of the kernel function and the increments of the $\widehat{X}_{\lambda,r}(\cdot, y)$ processes in the time domain.

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 $r \leq t_0 < t_1 < \cdots < t_M$, with step size $\Delta t > 0$, it holds that*

$$\mathbb{E}\left[\left|\left|\int_r^{t_j} h_N(t_j - s)\sigma(s-)dL(s) - \frac{a_0}{2}\eta_j(0) - \text{Re}\sum_{n=1}^N a_n\eta_j(n\pi/\tau)\right|\right|^2\right]$$

$$\leq (2m_1^2(t_j - t_0) + \text{Var}(L(1)))(t_j - t_0)\left\{\kappa_2\left(\lambda^2\left(|a_0|^2 + \frac{\tau^2}{3}||h'_\lambda||_2^2\right) + \frac{\tau^2}{6}||h''_\lambda||_2^2\right)(\Delta t)^2\right.$$

$$\left. + 3\left(|a_0|^2 + \frac{\tau^2}{3}||h'_\lambda||_2^2\right)\sup_{1\leq k\leq j, s\in[t_{k-1}, t_k)}\mathbb{E}\left[|\sigma(t_{k-1}-) - \sigma(s-)|^2\right]\right\},$$

*where $\eta_j$ is obtained by* (3.14), $1 \leq j \leq M$, *and $||\cdot||_2$ denotes the norm on $L^2([-\tau, \tau])$.*

*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

$$\left(\frac{|a_0|}{2} + \sum_{n=1}^N |a_n|\right)^2 \leq |a_0|^2 + 2\left(\sum_{n=1}^N |a_n|\right)^2$$

$$\leq |a_0|^2 + 2\left(\sum_{n=1}^N \frac{1}{n^2}\right)\left(\sum_{n=1}^N |a_n n|^2\right)$$

$$= |a_0|^2 + \frac{2}{\pi^2}\left(\sum_{n=1}^N \frac{1}{n^2}\right)\left(\sum_{n=1}^N |\widehat{h'_\lambda}(n\pi/\tau)|^2\right)$$

$$\leq |a_0|^2 + \frac{\tau^2}{3}||h'_\lambda||_{L^2([-\tau,\tau])}^2.$$

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

$$\left(\sum_{n=1}^N |a_n|n\right)^2 \leq \left(\sum_{n=1}^N \frac{1}{n^2}\right)\left(\sum_{n=1}^N |a_n n^2|^2\right)$$

$$= \frac{\tau^2}{\pi^4}\left(\sum_{n=1}^N \frac{1}{n^2}\right)\left(\sum_{n=1}^N |\widehat{h''_\lambda}(n\pi/\tau)|^2\right)$$

$$\leq \frac{1}{6}\frac{\tau^4}{\pi^2}||h''_\lambda||_{L^2([-\tau,\tau])}^2.$$

The proof is concluded by Proposition 4.3. ∎

**5. Comparison to other methods.** In the present section we will present the advantages of employing the Fourier method to simulate paths of LSS processes. Recall from our preliminary section that, in order for an LSS process to be well defined, we impose that $g \in L^1(\mathbb{R}_+) \cap L^2(\mathbb{R}_+)$. Moreover, when describing our simulation method in section 3 we imposed that the kernel function be continuous on the bounded interval it is evaluated on. For functions that do not fulfill this condition we assume that the corresponding LSS process can be approximated in an $L^2$-sense (using Lemma 2.1) by an LSS process having a continuous kernel function. Thus, we effectively restrict our attention to continuous functions but treat LSS processes driven by discontinuous kernels by means of approximations. Indeed, one can always find such an approximation, since the space of continuous functions with compact support is dense in any $L^p$ space where $1 \le p < \infty$ (see Proposition 7.9 in Folland [16]). Now, having imposed such a continuity condition, let us point out some of the advantages of employing our method for obtaining trajectories of LSS processes, as opposed to numerical integration or the Euler method discussed by Benth and Eyjolfsson in [10].

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

$$X_r(t_j) \approx \sum_{k=0}^{j-1} g(t_j - t_k)\sigma(t_k-)\Delta L(t_k)$$

(5.1)
$$= \int_{t_0}^{t_j} \sum_{k=0}^{j-1} g(t_j - t_k)\sigma(t_k-)1_{[t_k,t_{k+1})}(s)dL(s).$$

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

(5.2)
$$\mathbb{E}\left[\left|\int_{t_0}^{t_j} \left(g(t-s) - \sum_{k=0}^{j-1} g(t_j - t_k)1_{[t_k,t_{k+1})}(s)\right)\sigma(s-)dL(s)\right|^2\right]$$

$$\le 2m_1^2\kappa_2 \left(\int_{t_0}^{t_j}\left|g(t-s) - \sum_{k=0}^{j-1} g(t_j - t_k)1_{[t_k,t_{k+1})}(s)\right|ds\right)^2$$

$$+ \mathrm{Var}(L(1))\kappa_2 \int_{t_0}^{t_j}\left|g(t-s) - \sum_{k=0}^{j-1} g(t_j - t_k)1_{[t_k,t_{k+1})}(s)\right|^2 ds$$

$$\le (2m_1^2(t_j - t_0) + \mathrm{Var}(L(1)))\kappa_2 \int_{t_0}^{t_j}\left|g(t-s) - \sum_{k=0}^{j-1} g(t_j - t_k)1_{[t_k,t_{k+1})}(s)\right|^2 ds$$

and furthermore that

$$\mathbb{E}\left[\left|\int_{t_0}^{t_j} \sum_{k=0}^{j-1} g(t_j - t_k)1_{[t_k,t_{k+1})}(s)(\sigma(t_k-) - \sigma(s-))dL(s)\right|^2\right] \le \left\{2m_1^2\left(\sum_{k=0}^{j-1} g(t_j - t_k)\Delta t\right)^2\right.$$

(5.3)
$$\left. + \mathrm{Var}(L(1))\sum_{k=0}^{j-1} g^2(t_j - t_k)\Delta t\right\} \sup_{1\le k\le j, s\in[t_{k-1},t_k)} \mathbb{E}\left[|\sigma(t_{k-1}-) - \sigma(s-)|^2\right].$$

So by joining the above manipulations with

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

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

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

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 $L^2$-approximation of the volatility process; cf. (5.3).

So, for instance, in the case when the kernel function $g$ is Lipschitz continuous with a Lipschitz constant $C > 0$, the right-hand side of (5.2) is bounded by

$$(2m_1^2(t_j - t_0) + \text{Var}(L(1)))\kappa_2 C^2(t_j - t_0)(\Delta t)^2.$$

This yields a convergence rate of $(\Delta t)^2$, the same as for our Fourier method from the previous section.

Furthermore, as we have mentioned, in [10] the authors present an iterative Euler-type scheme for simulating trajectories of LSS processes. The basic idea behind the scheme is to assume that there exists a positive function $h : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$|g(u + \delta) - h(\delta)g(u)| \leq c(u, \delta)$$

holds for all $u \geq 0$ and $\delta > 0$, where $u \mapsto c(u, \delta)$ is a square-integrable function for all $\delta > 0$. Next, for a given equidistant time grid $t_0 < t_1 < \cdots < t_M$ with step size $\Delta t > 0$ and an initial value $X(t_0)$, one considers the time series defined by $\widetilde{X}(t_0) := X(t_0)$, and

$$\widetilde{X}(t_j) := h(\Delta t)\widetilde{X}(t_{j-1}) + \int_{t_{j-1}}^{t_j} g(t_j - s)\sigma(s-)dL(s).$$

Given a uniform step size $\Delta t = 1/N$ on the unit interval $[0, 1]$, Benth and Eyjolfsson [10] find by iterating that in the case when $\sigma = 1$ and $h(\Delta t) < 1$ it holds that

$$\left(\mathbb{E}[|\widetilde{X}(1) - X(1)|^2]\right)^{1/2} \leq \frac{||c(\cdot, \Delta t)||_{L^2(\mathbb{R}_+)}}{1 - h(\Delta t)}.$$

However, as argued in the paper by Benth and Eyjolfsson [10], the above expression does not do a good job of bounding the error in many cases. Indeed, by examining the scheme it is evident that $h(\Delta t) \to 1$ as $\Delta t \to 0^+$ needs to hold, and it is thus not clear how one should choose $h$ which guarantees an arbitrarily small error in the general case.

We would like to make the following points, which we believe make our method more feasible than the standard numerical integration technique and the Euler method. First, 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. Due to 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. The Euler method, on the other hand, is indeed an iterative scheme, but as we have argued it is not clear how to bound its error in all cases.

Second, the Fourier simulation scheme offers greater flexibility in that it first uses the increments of the Lévy and volatility processes to simulate the field of complex LSS processes which have exponential kernel functions

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

This field can in turn be employed to simulate LSS process trajectories for multiple choices of kernel functions. 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 $\{g_\theta\}_{\theta \in \Theta}$, where $\Theta$ is a given set of parameters.

Third, we believe that our method is particularly well suited for simulating LSS processes with kernel functions that are steep close to the origin (as we demonstrate in the following section). Indeed, in the iterative part of the Fourier method the kernel function we evaluate on the time grid is $t \mapsto e^{(-\lambda + iy)t}$, which is a complex exponential function with a negative real part, i.e., a well-behaved function with the same total variation on any two intervals of equal lengths. By contrast, using the more direct numerical integration and Euler methods, the total variation of the kernel function on an interval close to the origin is larger than the total variation of the same kernel function on an interval of equal length further from the origin. For a given equidistant time grid $0 = t_1 < t_2 < \cdots < t_M$ this means that a discrete trajectory $g(t_0), g(t_1), \ldots, g(t_M)$ may miss some information close to the origin. This may be avoided by making the grid finer close to the origin, which in turn increases the computational cost. However, as we have argued, such an effect is not observed using the Fourier method, since we have encoded the information on the kernel function into the $a_0, a_1, \ldots, a_N$ coefficients.

**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

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

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 written on the power spot price have been traded at the Nordic electricity exchange NordPool for some time around the year 2000 (see Weron [25]). Here, $X(t)$ denotes deseasonalized power spot price and we

shall assume arithmetic spot dynamics, that is, that the spot is modeled as an LSS process

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

Under a given risk-neutral measure $Q$ such that $P(T) \in L^1(Q)$ and a risk-free asset (a bank account) yielding a continuously compounded rate of return $r > 0$, Benth, Benth, and Koekebakker [9] define the option price $P(T)$ at time $t \le T$ as

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

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*

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

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

$$\mathbb{E}\left[ \left| f\left( \int_0^T X(t)dt \right) - f\left( \int_0^T \widetilde{X}(t)dt \right) \right| \right] \le C\mathbb{E}\left[ \left| \int_0^T (X(t) - \widetilde{X}(t))dt \right| \right]$$

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

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

The above proposition together with Propositions 4.1 and 4.3 gives us a convergence rate for employing our method to price path-dependent options. So, in particular we are able to 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

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

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

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

Thus, plugging this into (6.4) yields

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

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

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

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

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

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 compare with our proposed Fourier method.

In order to illustrate that let us apply 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

$$
\mathbb{E}\left[\max\left(\frac{1}{T}\int_0^T X(t)dt - K, 0\right)\bigg|\mathcal{F}_t\right]
$$

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

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 [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 $\mu$ and variance $\rho^2$, it holds that

$$\mathbb{E}[\max(Z,0)] = \frac{1}{\sqrt{2\pi}}\int_{-\mu/\rho}^\infty (\mu + \rho x)e^{-\frac{x^2}{2}}dx$$

$$= \mu\Phi(\mu/\rho) + \frac{\rho}{\sqrt{2\pi}}\exp(-\mu^2/(2\rho^2)),$$

where $\Phi$ denotes the cumulative distribution function of the standard normal distribution. Thus we conclude that

$$\mathbb{E}\left[\max\left(\frac{1}{T}\int_0^T X(t)dt - K, 0\right)\Big| \mathcal{F}_t\right]$$

$$= \mathbb{E}\left[\max\left(\frac{1}{T}\int_t^T G(T,s)dW(s) - \left(K - \frac{1}{T}\int_{-\infty}^t G(T,s)dW(s)\right), 0\right)\Big| \mathcal{F}_t\right]$$

$$(6.7) \qquad = \mu_{t,T}\Phi(\mu_{t,T}/\rho_{t,T}) + \frac{\rho_{t,T}}{\sqrt{2\pi}}\exp(-\mu_{t,T}^2/(2\rho_{t,T}^2)),$$

where

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

denote the mean and variance processes, respectively. Hence, we have an explicit pricing formula in the Gaussian case.

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

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

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 $\gamma(\nu, x) = \int_0^x u^{\nu-1}e^{-u}du$ 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 to 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

$$g_\epsilon(x) = \begin{cases} \phi_0(x) & \text{if } x \in [0,\epsilon], \\ g(x) & \text{if } x \geq \epsilon, \end{cases}$$
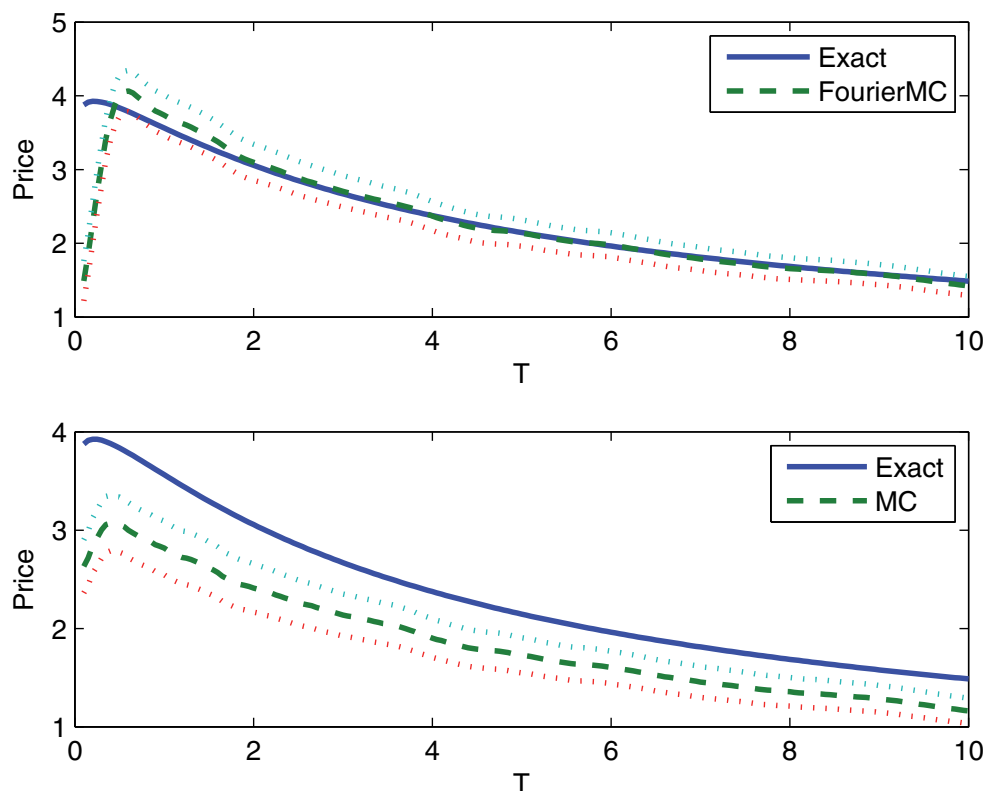
**Figure 2.** *The price curve $T \mapsto \mathbb{E}[\max(A(T) - K, 0)|\mathcal{F}_t]$ on $[0, 10]$, where $A(T)$ is given by (6.6), $K = 1$, $t = 0$, $L$ is standard Brownian motion on $[0, 10]$, and $L = 0$ on $(-\infty, 0)$. Here $\sigma = 1$ and $g(x) = Cx^{\nu-1}e^{-\alpha x}$, with $C = 10$, $\alpha = 1$, and $\nu = 0.55$. The first picture shows the exact price curve obtained by (6.7), and the approximation (6.4), with $\lambda = 1.9$, $\Delta t = 0.05$, and $N = 30$, together with 95% dotted confidence bounds. The second picture depicts the exact price curve obtained by (6.7), and numerical integration to evaluate $X(t)$, together with 95% dotted confidence bounds. In both cases the expectations and confidence bounds are obtained by averaging over 1000 Brownian motion paths.*

where $\phi_0$ 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 (4.12) with $k = 1$ we interpolate $g_\epsilon$ to zero in the tail from $\tau_0 = 10$ to $\tau = 11$.

In Figure 2 we have plotted the theoretical price curve given by (6.7) (with a solid line), the price curve obtained by means of applying our method with $\lambda = 1.9$, $\Delta t = 0.05$, and $N = 30$, and the price curve obtained by means of numerical integration with $\Delta t = 0.05$. We have, moreover, plotted 95% asymptotic confidence intervals at each time point (dotted lines) using $R = 1000$ Brownian motion paths for the Fourier and numerical integration approximations, respectively. The 95% asymptotic confidence intervals are obtained as follows. Suppose that for given $R$ Brownian motion paths and each $t$ on the discrete time grid

$$\widetilde{A}(t) = (\widetilde{A}_1(t), \widetilde{A}_2(t), \ldots, \widetilde{A}_R(t))$$

denotes the vector of approximations of the average (6.6) as obtained by either the Fourier method or numerical integration, using the $R$ different simulated Brownian motion paths. Then, the 95% asymptotic confidence intervals around the mean prices at each time step $t$ are given by the random interval

$$\left( \overline{\max(\widetilde{A}(t) - K, 0)} - z_{0.975} \frac{s_t}{\sqrt{R}}, \ \overline{\max(\widetilde{A}(t) - K, 0)} - z_{0.025} \frac{s_t}{\sqrt{R}} \right),$$

where

$$\overline{\max(\widetilde{A}(t) - K, 0)} = \frac{1}{R} \sum_{k=1}^{R} \max(\widetilde{A}_k(t) - K, 0),$$

$$s_t = \sqrt{\frac{1}{R-1} \sum_{k=1}^{R} \left( \max(\widetilde{A}_k(t) - K, 0) - \overline{\max(\widetilde{A}(t) - K, 0)} \right)^2},$$

and $z_{0.025}, z_{0.975}$ denote the 0.025 and 0.975 quantiles of the standard normal distribution, respectively (see, e.g., Chapter III in Asmussen and Glynn [2] for more details).

The most striking feature of Figure 2 is how far away the numerical integration estimate, including the confidence bound, is from the theoretical price curve (and the price curve obtained by the Fourier method). The most apparent explanation for this is the steepness of the kernel function close to the origin. Indeed, as we discussed in the previous section, applying a numerical integration technique to simulate an LSS process with such a kernel function appears to be nonoptimal, at least if one wishes to use an equidistant time grid.

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 an 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 by 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 by means of numerical integration.

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

Finally, we remark that in a paper by Eyjolfsson [15], the Fourier method introduced in the current paper is generalized to the tempo-spatial case of so-called ambit fields, and an application to electricity forward pricing is studied.

**Acknowledgment.** We wish to thank an anonymous referee for a number of constructive comments which have improved the paper.

## REFERENCES

[1] R. B. Ash, *Probability and Measure Theory*, 2nd ed., Harcourt/Academic Press, Burlington, MA, 2000.

[2] S. Asmussen and P. W. Glynn, *Stochastic Simulation: Algorithms and Analysis*, Stoch. Model. Appl. Probab. 57, Springer, New York, 2007.

[3] O. E. Barndorff-Nielsen, F. E. Benth, and A. E. D. Veraart, *Modelling energy spot prices by volatility modulated Lévy-driven Volterra processes*, Bernoulli, 19 (2013), pp. 803–845.

[4] O. E. Barndorff-Nielsen, F. E. Benth, and A. E. D. Veraart, *Modelling electricity futures by ambit fields*, Adv. in Appl. Probab., to appear.

[5] O. E. Barndorff-Nielsen and J. Schmiegel, *Lévy-based tempo-spatial modelling; with applications to turbulence*, Uspekhi Mat. Nauk, 59 (2004), pp. 65–91.

[6] O. E. Barndorff-Nielsen and J. Schmiegel, *Brownian semistationary processes and volatility/intermittency*, in Advanced Financial Modelling, Radon Ser. Comput. Appl. Math. 8, H. Albrecher, W. J. Runggaldier, and W. Schachermayer, eds., Walter de Gruyter, Berlin, 2009, pp. 1–25.

[7] A. Basse and J. Pedersen, *Lévy driven moving averages and semimartingales*, Stochastic Process. Appl., 119 (2009), pp. 2970–2991.

[8] A. Basse-O'Connor, S.-E. Graversen, and J. Pedersen, *A unified approach to stochastic integration on the real line*, Theory Probab. Appl., 58 (2013+), to appear.

[9] F. E. Benth, J. Š. Benth, and S. Koekebakker, *Stochastic Modelling of Electricity and Related Markets*, Adv. Ser. Stat. Sci. Appl. Probab. 11, World Scientific, Hackensack, NJ, 2008.

[10] F. E. Benth and H. Eyjolfsson, *Stochastic modeling of power markets using stationary processes*, in Seminar on Stochastic Analysis, Random Fields and Applications VII, Progr. Probab. 67, R. C. Dalang, M. Dozzi, and F. Russo, eds., Springer, Basel, 2013, pp. 261–284.

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

[12] F. E. Benth, R. Kiesel, and A. Nazarova, *A critical empirical study of three electricity spot price models*, Energy Economics, 34 (2012), pp. 1589–1616.

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

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

[15] H. Eyjolfsson, *Ambit Fields via Fourier Methods in the Context of Power Markets*, DUO, University of Oslo, Oslo, Norway, 2013.

[16] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, Pure Appl. Math. (N. Y.), John Wiley & Sons, New York, 1984.

[17] I. García, C. Klüppelberg, and G. Müller, *Estimation of stable CARMA models with an application to electricity spot prices*, Stat. Model., 11 (2011), pp. 447–470.

[18] K. Ýr Jónsdóttir, J. Schmiegel, and E. B. Vedel Jensen, *Lévy-based growth models*, Bernoulli, 14 (2008), pp. 62–90.

[19] C. Klüppelberg, T. Meyer-Brandis, and A. Schmidt, *Electricity spot price modelling with a view towards extreme spike risk*, Quant. Finance, 10 (2010), pp. 963–974.

[20] J. J. Lucia and E. S. Schwartz, *Electricity prices and power derivatives: Evidence from the Nordic power exchange*, Rev. Deriv. Res., 5 (2002), pp. 5–50.

[21] P. E. Protter, *Stochastic Integration and Differential Equations*, 2nd ed., Stoch. Model. Appl. Probab. 21, Springer-Verlag, Berlin, 2005.

[22] P. A. Samuelson, *Proof that properly anticipated prices fluctuate randomly*, Industrial Management Rev., 6 (1965), pp. 41–49.

[23] K. SATO, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge Stud. Adv. Math. 68, Cambridge University Press, Cambridge, UK, 1999.

[24] E. S. SCHWARTZ, *The stochastic behavior of commodity prices: Implications for valuation and hedging*, J. Finance, 52 (1997), pp. 923–973.

[25] R. WERON, *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*, John Wiley & Sons, Chichester, UK, 2006.