Karhunen–Loève expansions of Lévy processes

Daniel Hackmann
Institute of Financial Mathematics and Applied Number Theory, Linz, Austria

ABSTRACT
We derive the basis functions and joint distribution of the stochastic coefficients of the Karhunen–Loève expansion of a square-integrable Lévy process. Further, we demonstrate a method for simulating the coefficients via a shot-noise representation.

KEYWORDS
Fourier series; Infinitely divisible; Karhunen-Loève; Lévy process; Simulation; Shot-noise.

MATHEMATICS SUBJECT CLASSIFICATION
Primary: 60G51; Secondary: 60J75, 60G07, 65C05

1. Introduction
The Karhunen–Loève theorem (KLT) allows us to create generalized Fourier series from stochastic processes in an, in some sense, optimal way. Arguably the most famous application of the KLT is to derive the classic sine series expansion of a Wiener process $W$ on $[0, 1]$. Specifically,

$$W_t = \sqrt{2} \sum_{k \geq 1} Z_k \frac{\sin \left( \pi \left( k - \frac{1}{2} t \right) \right)}{\pi \left( k - \frac{1}{2} \right)}$$

where convergence of the series is in $L^2(\Omega, \mathbb{F})$ and uniform in $t \in [0, 1]$, and the \{Z_k\}_{k \geq 1} are i.i.d. standard normal random variables. In this article, we show that a square-integrable Lévy process also admits a Karhunen–Loève expansion (KLE) as a series of sine functions. Although the stochastic coefficients are no longer normal nor independent, they remain in the class of infinitely divisible (ID) distributions; we demonstrate how they may be simulated via a shot-noise representation.

Deriving the KLE of the type (1) for a square-integrable stochastic process $X$ on $[a, b]$ requires two steps: first, one must solve a Fredholm integral equation to obtain the basis functions \{e_k\}_{k \geq 1} (c.f., the sine functions in Equation (1)). Second, one must identify the distribution of the stochastic coefficients

$$Z_k := \int_a^b X_t e_k(t) \, dt, \quad k \in \mathbb{N}.$$
process, the Brownian bridge process and the Anderson–Darling process are two examples. A recent paper by Barczy and Lovas (2016) derives the KLE for an interesting generalization of the Brownian bridge and provides extensive references for Gaussian processes with explicit KLE expansions. For further Gaussian examples with derivation, see also (Jin 2014, Chapter 1). Non Gaussian processes pose an additional challenge and the problem of deriving the KLE is usually left to numerical means (see e.g., Phoon, Huang, and Quek 2005).

To begin, we recall the necessary facts from the theory of Lévy processes and ID random vectors.

2. Facts from the theory of Lévy processes

The Lévy–Khintchine theorem states that every \( d \)-dimensional ID random vector \( \xi \) has a Fourier transform of the form

\[
E[e^{i\langle z, \xi \rangle}] = e^{-\Psi_\xi(z)}, \quad z \in \mathbb{R}^d,
\]

where

\[
\Psi_\xi(z) = \frac{1}{2} z^T Q z - \langle a, z \rangle - \int_{\mathbb{R}^d \setminus \{0\}} e^{i\langle z, x \rangle} - 1 - i\langle z, x \rangle h(x) \nu(dx),
\]

and where \( a \in \mathbb{R}^d \), \( Q \) is a positive semidefinite matrix, and \( \nu(dx) \) is the Lévy measure on \( \mathbb{R}^d \setminus \{0\} \). The function \( h \) is known is the cutoff function; up to a choice of \( h \), the generating triple \( (a, Q, \nu) \) uniquely identifies the distribution of \( \xi \). The Lévy–Khintchine theorem for Lévy processes gives us an analogously powerful result, specifically, for any \( d \)-dimensional Lévy process \( X \) we have

\[
E[e^{i\langle z, X_t \rangle}] = e^{-t\Psi_X(z)}, \quad z \in \mathbb{R}^d, \quad t \geq 0,
\]

where \( \Psi_X \) is as in (3) and \( X \) is uniquely determined, up to identity in distribution, by the triple \( (a, Q, \nu) \). In one dimension, we will write \( (a, \sigma^2, \nu) \) for the generating triple and when \( \nu(dx) = \pi(x)dx \) for some density function \( \pi \), we will write \( (a, \sigma^2, \pi) \). If we wish to be specific regarding the cutoff function, we will write \( (a, Q, \nu)_h \equiv \cdot \) or \( (a, \sigma^2, \nu)_h \equiv \cdot \) for the generating triples.

We will work primarily with one-dimensional Lévy processes from the set \( K \) made up of those processes which satisfy \( E[X_t] = 0 \) and \( E[X_t^2] < \infty \) for every \( t \geq 0 \). For \( X \in K \), we have

\[
\text{Var}(X_t) = E[X_t^2] = at, \quad \text{and} \quad \text{Cov}(X_s, X_t) = E[X_sX_t] = \alpha \min(s, t),
\]

where \( \alpha := \Psi_X''(0) \).

An ID random vector \( \xi \) or Lévy process \( X \) with associated Lévy measure \( \nu \) has a finite second moment (meaning the component-wise moments) if, and only if,

\[
\int_{|x| > 1} |x|^2 \nu(dx) < \infty \quad \text{(Condition A)}.
\]

We will denote the class of ID random vectors with zero first moment and finite second moment by \( C \). The subset of \( C \) which also satisfies

\[
\int_{|x| \leq 1} |x| \nu(dx) < \infty \quad \text{(Condition B)}.
\]

will be denoted \( CB \) and \( KB \) will denote the analogous subset of \( K \). We remark that any \( \xi \in C \) (resp. \( X \in K \)) necessarily has a representation of the form \( (0, Q, \nu)_{h=1} \) (resp. \( (0, \sigma^2, \nu)_{h=1} \)).
Additionally, any $d$-dimensional $\xi \in CB$ necessarily has representation \((a, Q, v)_{h=0}\) where $a$ has entries
\[
- \int_{\mathbb{R}^d \setminus \{0\}} P_k(x) v(dx), \quad k \in \{1, 2, \ldots d\}
\]
and $P_k$ is the projection onto the $k$th component. Analogously, if $X \in KB$, then we have representation \((a, \sigma^2, v)_{h=0}\) where $a = -\int_{\mathbb{R} \setminus \{0\}} x v(dx)$.

3. The Karhunen–Loève theorem

Given a real-valued continuous time stochastic process $X$ defined on an interval $[a, b]$ and an orthonormal basis $\{\phi_k\}_{k\geq 1}$ for $L^2([a, b])$ we might try to express $X$ as a generalized Fourier series
\[
X_t = \sum_{k=1}^{\infty} Y_k \phi_k(t), \quad \text{where } Y_k := \int_a^b X_t \phi_k(t) dt.
\]

The expansion so formed, which takes as a basis the normalized eigenfunctions $\{\phi_k\}_{k\geq 1}$ corresponding to the non zero eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3, \ldots$ of the integral operator $K_X : L^2([a, b]) \to L^2([a, b])$,
\[
(K_X f)(s) := \int_a^b \text{Cov}(X_s, X_t) f(t) dt,
\]
is called the KLE. We will write the stochastic coefficients for the KLE as $\{Z_k\}_{k\geq 1}$ and the first $d$ of these as $Z^{(d)} := (Z_1, Z_2, \ldots, Z_d)$.

**Theorem 1 (The KLT).** Let $X$ be a real-valued continuous time stochastic process on $[a, b]$ such that $0 \leq a \leq b < \infty$ and let $\mathbb{E}[X_t] = 0$ and $\mathbb{E}[X_t^2] < \infty$ for each $t \in [a, b]$. Further, suppose $\text{Cov}(X_s, X_t)$ is continuous on $[a, b] \times [a, b]$.

(i) Then,
\[
\mathbb{E}\left[ X_t - \sum_{k=1}^{d} Z_k \epsilon_k(t) \right]^2 \to 0, \quad \text{as } d \to \infty
\]
uniformly for $t \in [a, b]$. Additionally, the random variables $\{Z_k\}_{k\geq 1}$ are uncorrelated and satisfy $\mathbb{E}[Z_k] = 0$ and $\mathbb{E}[Z_k^2] = \lambda_k$.

(ii) For any other basis $\{\phi_k\}_{k\geq 1}$ of $L^2([a, b])$, with corresponding stochastic coefficients $\{Y_k\}_{k\geq 1}$, and any $d \in \mathbb{N}$, we have
\[
\int_a^b \mathbb{E}[(\epsilon_d(t))^2] dt \leq \int_a^b \mathbb{E}[(\tilde{\epsilon}_d(t))^2] dt,
\]
where $\epsilon_d$ and $\tilde{\epsilon}_d$ are the remainders $\epsilon_d(t) := \sum_{d+1}^{\infty} Z_k \phi_k(t)$ and $\tilde{\epsilon}_d(t) := \sum_{d+1}^{\infty} Y_k \phi_k(t)$.

According to Ghanem and Spanos (1991), the KLT was proposed independently by Karhunen (1947), Loève (1948), and Kac and Siegert (1947). Modern proofs of the first part of the theorem can be found in Ash and Gardner (1975) and Ghanem and Spanos (1991) and the second part—the optimality of the truncated approximation—is also proven in Ghanem and Spanos (1991). A concise and readable overview of this theory is given in (Luo 2006, Chapter 7.1).

Since the covariance function of a process $X \in \mathcal{K}$ differs from that of a Wiener process only by the scaling factor $\alpha$, the method for determining the eigenfunctions and the eigenvalues...
of $K_X$ is identical to that employed for a Wiener process. Therefore, we omit the proof of the following proposition and direct the reader to (Ash and Gardner 1975, p. 41) where the proof for the Wiener process is given.

**Proposition 1.** For $X \in \mathcal{K}$, the eigenvalues and associated eigenfunctions of the operator $K$ defined on $L^2([0, T])$ are given by

$$
\lambda_k = \frac{\alpha T^2}{\pi^2 (k - \frac{1}{2})^2}, \quad \text{and} \quad e_k(t) = \sqrt{\frac{2}{T}} \sin \left( \frac{\pi}{T} \left( k - \frac{1}{2} \right) t \right), \quad k \in \mathbb{N}, \ t \in [0, T]. \tag{6}
$$

A nice consequence of Proposition 1 and Theorem 1 is that it allows us to estimate the amount of total variance $v(T) := \int_0^T \text{Var}(X_t) dt = \int_0^T \mathbb{E}[X_t^2] dt = \alpha T^2/2$ we capture when we represent our process by a truncated KLE. Since it is straightforward to show that $v(T) = \sum_{k \geq 1} \lambda_k$, the total variance explained by a $d$-term approximation is therefore given by

$$
\sum_{k=1}^d \frac{\lambda_k}{v(T)} = \frac{2}{\pi^2} \sum_{k=1}^d \frac{1}{(k - \frac{1}{2})^2},
$$

and this holds independently of $\alpha$ or $T$. To determine the distribution of $Z^{(d)}$, we need the following, which we remark applies not only to process $X \in \mathcal{K}$.

**Lemma 1.** Let $X$ be a real-valued Lévy process and let $\{f_k\}_{k=1}^d$ be a collection of functions which are in $L^1([0, T])$. Then, the vector $\xi$ consisting of elements

$$
\xi_k = \int_0^T X_t f_k(s) ds, \quad k \in \{1, 2, \ldots, d\},
$$

has an ID distribution with characteristic exponent

$$
\Psi_{\xi}(z) = \int_0^T \Psi_X((z, u(t))) dt, \quad z \in \mathbb{R}^d, \tag{7}
$$

where $u : [0, T] \rightarrow \mathbb{R}^d$ is the function with kth component $u_k(t) := \int_0^T f_k(s) ds$, $k \in \{1, 2, \ldots, d\}$.

**Remark 1.** A similar identity to (7) is known, see p. 128 in Bertoin (2001). In the proof of Lemma 1, we borrow some ideas from there. Since the proof is rather tedious, we relegate it to the Appendix. We note in passing that an alternative proof (to that given in the Appendix) could also be established using the identity $\xi_k = \int_0^T u_k(t) dX_t$, which holds for continuous functions $f_k$ and follows from integration by parts. While this identity is interesting, using it does not shorten the proof considerably, and a rigorous treatment would require additional concepts and notation from the theory of stochastic calculus, which the current proof avoids.

With Lemma 1 and Proposition 1 in hand, we come to our first main result. Going forward, we use the notation $\mathcal{B}_S$ for the Borel sigma algebra on the topological space $S$.

**Theorem 2.** If $X \in \mathcal{K}$ with generating triple $(0, \sigma^2, \nu)_{k=1}$, then $Z^{(d)} \in \mathcal{C}$ with generating triple $(0, Q, \Pi)_{k=1}$ where $Q$ is a diagonal $d \times d$ matrix with entries

$$
q_{k,k} := \frac{\sigma^2}{2} \frac{T^2}{\pi^2 (k - \frac{1}{2})^2}, \quad k \in \{1, 2, \ldots, d\}, \tag{8}
$$
and $\Pi$ is the measure,

$$\Pi(B) := \int_{\mathbb{R} \setminus \{0\} \times [0, T]} \mathbb{I}(f(v) \in B)(v \times \lambda)(dv), \quad B \in \mathcal{B}_{\mathbb{R}^d \setminus \{0\}}.$$  

(9)

where $\lambda$ is the Lebesgue measure on $[0, T]$ and $f : \mathbb{R} \times [0, T] \to \mathbb{R}^d$ is the function

$$(x, t) \mapsto \frac{\sqrt{2T}x}{\pi} \left( \cos \left( \frac{T}{\pi} \frac{1 - \frac{1}{2}t}{t} \right), \cos \left( \frac{T}{\pi} \frac{2 - \frac{1}{2}t}{t} \right), \ldots, \cos \left( \frac{T}{\pi} \frac{d - \frac{1}{2}t}{t} \right) \right)^T.$$  

(10)

Proof. We substitute the formula for the characteristic exponent (Formula 3 with $a = 0$ and $h = 1$) and the eigenfunctions (Formula 6) into (7) and carry out the integration. Then, (8) follows from the fact that $X$ is a Lévy measure that satisfies Condition $A$ since

$$\text{and that the } \{u_k\}_{k \geq 1} \text{ are therefore also orthogonal on } [0, T].$$

Next, we note that $f$ is a continuous function from $\mathbb{R} \setminus \{0\} \times [0, T]$ to $\mathbb{R}^d$ and is therefore $(\mathcal{B}_{\mathbb{R} \setminus \{0\} \times [0, T]}, \mathcal{B}_{\mathbb{R}^d \setminus \{0\}})$ measurable. Therefore, $\Pi$ is nothing other than the push-forward measure obtained from $(v \times \lambda)$ and $f$; in particular, it is a well-defined measure on $\mathcal{B}_{\mathbb{R}^d \setminus \{0\}}$. It is also a Lévy measure that satisfies Condition $A$ since

$$\int_{|x| > 1} |x|^2 \Pi(dx) \leq \int_{\mathbb{R}^d \setminus \{0\}} |x|^2 \Pi(dx) = \frac{2T}{\pi^2} \int_0^T \left( \sum_{k=1}^d u_k^2(t) \right) dt \int_{\mathbb{R} \setminus \{0\}} x^2 \nu(dx) < \infty,$$  

(11)

where the final inequality follows from the fact that $X \in K$. Applying Fubini's theorem and a change of variables, i.e.,

$$\int_0^T \int_{\mathbb{R} \setminus \{0\}} e^{ix(z, u(t))} - 1 - ix(z, u(t)) \nu(dx) dt = \int_{\mathbb{R} \setminus \{0\} \times [0, T]} e^{i(x, f(v))} - 1 - i(x, f(v)) (v \times \lambda)(dv)$$

$$= \int_{\mathbb{R}^d \setminus \{0\}} e^{i(z, x)} - 1 - i(z, x) \Pi(dx),$$

concludes the proof. \hfill \Box

Remark 2. Note, that if we set $\sigma = 1$, $\nu \equiv 0$, and $T = 1$, we may easily recover the KLE of the Wiener process, i.e., (1), from Theorem 2.

We gather some fairly obvious but important consequences of Theorem 2 in the following corollary, whose proof we leave to the reader.

Corollary 1. Suppose $X \in K$, then:

(i) $X \in KB$ with generating triple $(a, \sigma^2, \nu)_{h=0}$ if, and only if, $Z^{(d)} \in CB$ with generating triple $(a, Q, \Pi)_{h=0}$, where $Q$ and $\Pi$ are as defined in (8) and (9) and $a$ is the vector with entries

$$a_k := a \frac{(-1)^{k+1} \sqrt{2T}^2}{\pi^2 (k - \frac{1}{2})^2}, \quad k \in \{1, 2, \ldots, d\}.$$  

(12)

(ii) $X$ has finite Lévy measure $\nu$ if, and only if, $Z^{(d)}$ has finite Lévy measure $\Pi$.

Also intuitively obvious, but slightly more difficult to establish rigorously, is the fact that the entries of $Z^{(d)}$ are dependent unless $\nu \equiv 0$. 
Corollary 2. If \( X \in \mathcal{K} \), then \( Z^{(d)} \) has independent entries if, and only if, \( \nu \) is the zero measure.

To prove Corollary 2, we use the fact that a \( d \)-dimensional ID random vector with generating triple \((a, Q, \nu)\) has independent entries if, and only if, \( \nu \) is supported on the union of the coordinate axes and \( Q \) is diagonal (see E 12.10 on p. 67 in Sato 1999). For this purpose, we define for a vector \( \mathbf{x} = (x_1, x_2, \ldots, x_d)^T \in \mathbb{R}^d \) such that \( x_k > 0 \), \( k \in \{1, 2, \ldots, d\} \), the sets
\[
\mathcal{I}^+(\mathbf{x}) := \prod_{k=1}^d (x_k, \infty), \quad \text{and} \quad \mathcal{I}^-(\mathbf{x}) := \prod_{k=1}^d (-\infty, -x_k),
\]
where we caution the reader that the symbol \( \prod \) indicates the Cartesian product and not the Lévy measure of \( Z^{(d)} \).

In the proof below, and throughout the remainder of the article, \( f \) will always refer to the function defined in (10), and \( f_k \) to the \( k \)th coordinate of \( f \).

**Proof of Corollary 2.** We assume that \( \nu \) is not identically zero and show that there exists \( \mathbf{x} \) such that either \( \Pi(\mathcal{I}^+(\mathbf{x})) \) or \( \Pi(\mathcal{I}^-(\mathbf{x})) \) is strictly greater than zero.

Since \( \nu(\mathbb{R}\setminus\{0\}) > 0 \), there must exist \( \delta > 0 \) such that one of \( \nu((\infty, -\delta)) \) and \( \nu((\delta, \infty)) \) is strictly greater than zero; we will initially assume the latter. We observe that for \( d \in \mathbb{N}, \ d \geq 2 \), the zeros of the function \( h_d : [0, T] \rightarrow \mathbb{R} \) defined by
\[
t \mapsto \cos \left( \frac{\pi}{T} \left( d - \frac{1}{2} \right) t \right) \left( d - \frac{1}{2} \right)
\]
occur at points \( \{nT/(2d - 1)\}_{n=1}^{2d-1} \), and therefore the smallest zero is \( t_d := T/(2d - 1) \). From the fact that the cosine function is positive and decreasing on \( [0, \pi/2] \), we may conclude that
\[
\frac{\cos \left( \frac{\pi}{T} \left( k - \frac{1}{2} \right) t \right)}{\left( k - \frac{1}{2} \right)} > \epsilon, \quad k \in \{1, 2, \ldots, d\}, \ t \in [0, t_d/2],
\]
where \( \epsilon = h_d(t_d/2) > 0 \). Now, let \( \mathbf{x} \) be the vector with entries \( x_k = \delta \epsilon \sqrt{2T}/\pi \) for \( k \in \{1, 2, \ldots, d\} \).

Then,
\[
(\delta, \infty) \times [0, t_d/2] \subset f^{-1}(\mathcal{I}^+(\mathbf{x})),
\]
since for \( (x, t) \in (\delta, \infty) \times [0, t_d/2] \) we have
\[
f_k(x, t) = \frac{\sqrt{2T}}{\pi} x \cos \left( \frac{\pi}{T} \left( k - \frac{1}{2} \right) t \right) > \delta \epsilon \frac{\sqrt{2T}}{\pi} = x_k, \quad k \in \{1, 2, \ldots, d\}.
\]
But then,
\[
\Pi(\mathcal{I}^+(\mathbf{x})) \geq \nu((\delta, \infty)) \times \lambda([0, t_d/2]) > 0.
\]
If we had initially assumed that \( \nu((\infty, -\delta)) > 0 \), we would have reached the same conclusion by using the interval \( (-\infty, -\delta) \) and \( \mathcal{I}^-(\mathbf{x}) \). We conclude that \( \Pi \) is not supported on the union of the coordinate axes, and so \( Z^{(d)} \) does not have independent entries. The proof of the converse statement is straightforward. \( \square \)

4. **Shot-noise representation of \( Z^{(d)} \)**

In general, simulating a random vector with dependent entries with only the knowledge of the characteristic function seems to be a difficult problem; even generating random variables from the characteristic function is not straightforward (see e.g., Devroye 1986). In our case, thanks
to Theorem 2 we know that $Z^{(d)}$ is ID and that the Lévy measure $\Pi$ has a special disintegrated form. This will help us build the connection with the so-called shot-noise representation of our vector $Z^{(d)}$. The goal is to represent $Z^{(d)}$ as an almost surely convergent series of random vectors.

To explain this theory—nicely developed and explained in Rosiński (1990, 2001)—we assume that we have two random sequences $\{V_i\}_{i \geq 1}$ and $\{\Gamma_i\}_{i \geq 1}$ which are independent of each other and defined on a common probability space. We assume that each $\Gamma_i$ is distributed like a sum of $i$ independent exponential random variables with mean 1, and that the $\{V_i\}_{i \geq 1}$ take values in a measurable space $D$, and are i.i.d. with common distribution $F$.

Further, we assume we have a measurable function $H : (0, \infty) \times D \rightarrow \mathbb{R}^d$ which we use to define the random sum

$$S_n := \sum_{i=1}^{n} H(\Gamma_i, V_i), \quad n \in \mathbb{N},$$

and the measure

$$\mu(B) := \int_0^\infty \int_D \mathbb{1}(H(r, v) \in B) F(dv) dr, \quad B \in \mathcal{B}_{\mathbb{R}^d \setminus \{0\}}.$$  \hspace{1cm} (15)

The function $C : (0, \infty) \rightarrow \mathbb{R}^d$ is defined by

$$C_k(s) := \int_0^s \int_D P_k(H(r, v)) F(dv) dr, \quad k \in \{1, 2, \ldots, d\},$$

where, as before, $P_k$ is the projection onto the $k$th component. The connection between (14) and ID random vectors is then explained in the following theorem whose results can be obtained by restricting Theorems 3.1, 3.2, and 3.4 in Rosiński (1990) from a general Banach space setting to $\mathbb{R}^d$, and by making a slight modification to the proof of Theorem 3.4.

**Theorem 3 (Theorems 3.1, 3.2, and 3.4 in Rosiński 1990).** Suppose $\mu$ is a Lévy measure, then:

(i) If Condition B holds then $S_n$ converges almost surely to an ID random vector with generating triple $(0, 0, \mu)_{n=0}$ as $n \rightarrow \infty$.

(ii) If Condition A holds, and for each $v \in D$, the function $r \rightarrow |H(r, v)|$ is non increasing, then

$$M_n := S_n - C(n), \quad n \in \mathbb{N}$$  \hspace{1cm} (17)

converges almost surely to an ID random vector with generating triple $(0, 0, \mu)_{n=1}$.

The name “shot-noise representation” comes from the idea that $|H|$ can be interpreted as a model for the volume of the noise of a shot $V_i$ which occurred $\Gamma_i$ seconds ago. If $|H|$ is non increasing in the first variable, as we assume in case (ii) in Theorem 3, then the volume decreases as the elapsed time grows. The series $\lim_{n \rightarrow \infty} S_n$ can be interpreted as the total noise at the present time of all previous shots.

The goal is to show that for any process in $K$ whose Lévy measure admits a strictly positive density $\pi$, the vector $Z^{(d)}$ has a shot-noise representation of the form (14) or (17). To simplify notation, we make some elementary but necessary observations/assumptions: First, we assume that $X$ has no Gaussian component $\sigma^2$. There is no loss of generality to this assumption, since if $X$ does have a Gaussian component, then $Z^{(d)}$ changes by the addition of a vector of independent Gaussian random variables. This poses no issue from a simulation standpoint.
Second, from (3) we see that any Lévy process $X$ with representation $(0, 0, \pi)_{h=1}, j \in \{0, 1\}$ can be decomposed into the difference of two independent Lévy processes, each having only positive jumps. Indeed, splitting the integral and making a change of variable $x \mapsto -x$ gives

$$\Psi_X(z) = -\int_{\mathbb{R} \setminus \{0\}} e^{izx} - 1 - izx j\pi(x)dx$$

$$= -\int_0^\infty e^{izx} - 1 - izx j\pi(x)dx - \int_0^\infty e^{iz(-x)} - 1 - iz(-x) j\pi(-x)dx$$

$$= \Psi_{X^+}(z) + \Psi_{X^-}(z)$$

(18)

where $X^+$ (resp. $X^-$) has Lévy density $\pi(\cdot)$ (resp. $\pi(-\cdot)$) restricted to $(0, \infty)$. In light of this observation, the results of Theorem 4 are limited to Lévy processes with positive jumps. It should be understood that for a general process we can obtain $Z^{(d)}$ by simulating $Z_+^{(d)}$ and $Z_-^{(d)}$—corresponding to $X^+$ and $X^-$, respectively—and then subtracting the second from the first to obtain a realization of $Z^{(d)}$.

Last, for a Lévy process with positive jumps and strictly positive Lévy density $\pi$, we define the function

$$g(x) := \int_x^\infty \pi(s)ds.$$  

(19)

which is just the tail integral of the Lévy measure. We see that $g$ is strictly monotonically decreasing to zero, and so admits a strictly monotonically decreasing inverse $g^{-1}$ on the domain $(0, g(0))$.

**Theorem 4.** Let $\pi$ be a strictly positive Lévy density on $(0, \infty)$ and identically zero elsewhere.

(i) If $X \in \mathcal{KB}$ with generating triple $(a, 0, \pi)_h=0$, then $Z^{(d)}$ has a shot-noise representation

$$Z^{(d)} \overset{d}{=} a + \sum_{i \geq 1} H(\Gamma_i, U_i)$$

(20)

where $a$ is as defined in (12), $\{U_i\}_{i \geq 1}$ is an i.i.d. sequence of uniform random variables on $[0, 1]$, and

$$H(r, v) := f(g^{-1}(r/T)\mathbb{1}(0 < r < g(0)), Tv).$$

(21)

where $f$ is the function defined in (10).

(ii) If $X \in \mathcal{K}$ with generating triple $(0, 0, \pi)_h=1$, then $Z^{(d)}$ has a shot-noise representation

$$Z^{(d)} \overset{d}{=} \lim_{n \to \infty} \sum_{i=1}^n H(\Gamma_i, U_i) - C(n),$$

(22)

where $H$ and $\{U_i\}_{i \geq 1}$ are as in Part (i) and $C$ is defined as in (16).

**Proof.** Rewriting (9) to suit our assumptions and making a change of variables $t = Tv$ gives, for any $B \in \mathcal{B}_{\mathbb{R}^d \setminus \{0\}}$

$$\Pi(B) = \int_0^T \int_0^\infty \mathbb{1}(f(x, t) \in B)\pi(x)dxdt = \int_0^1 \int_0^\infty \mathbb{1}(f(x, Tv) \in B)T\pi(x)dxdv.$$ 

Making a further change of variables $r = Tg(x)$ gives

$$\Pi(B) = \int_0^1 \int_0^{g(0)} \mathbb{1}(f(g^{-1}(r/T), Tv) \in B)drdv.$$
Since \( 0 \notin B \), so that \( \mathbb{1}(0 \in B) = 0 \), we may conclude that

\[
\Pi(B) = \int_0^\infty \int_0^1 \mathbb{1}(f(g^{-1}(r/T)\mathbb{1}(0 < r < g(0)), Tv) \in B) dv dr.
\]

From the definition of the function \( f \) (Formula 10), and that of \( g^{-1} \), it is clear that

\[
(r, v) \mapsto f(g^{-1}(r/T)\mathbb{1}(0 < r < g(0)), Tv)
\]

is measurable and non increasing in absolute value for any fixed \( v \). Therefore, we can identify (23) with the function \( H \), the uniform distribution on \([0, 1] \) with \( F \), and \( \Pi \) with \( \mu \). The results then follow by applying the results of Theorems 2 and 3 and Corollary 1.

Going forward, we will write simply \( H(r, v) = f(g^{-1}(r/T), Tv) \) where it is understood that \( g^{-1} \) vanishes outside the interval \((0, g(0))\).

It is clear that the efficacy of simulating \( Z^{(d)} \) in this way is tied to the decay of \( g^{-1} \) (in those cases where \( g(0) = \infty \)) and our ability to compute \( g^{-1} \). Rosiński (2001) and Rosiński (2007) demonstrate several related series techniques, which avoid computation of \( g^{-1} \) entirely. A potentially interesting area of future research would be to adapt these to the present case. However, if \( g \) is computable and well-behaved enough, we can also apply Theorem 4 directly via numerical inversion of \( g \).

**Example 1.** Consider the variance gamma (VG) process which was first introduced in Madan and Seneta (1990). The process can be constructed as the difference of two independent Gamma processes, i.e., processes with Lévy measures of the form

\[
\nu(dx) = c e^{-\rho x} dx, \quad x > 0,
\]

where \( c, \rho > 0 \). For this example, we use a gamma process \( X^+ \) with parameters \( c \) and \( \rho \) and subtract a gamma process \( X^- \) with parameters \( c \) and \( \hat{\rho} \) to yield a VG process \( X \). We observe that \( X^+, X^- \notin \mathbb{K} \) since

\[
\mathbb{E}[X^+_1] = ut \Psi'_{X^+}(0) = t \frac{c}{\rho} \neq 0 \quad \text{and} \quad \mathbb{E}[X^-_1] = ut \Psi'_{X^-}(0) = t \frac{c}{\hat{\rho}} \neq 0.
\]

However, this is not a problem, since we can always construct processes \( \tilde{X}^+, \tilde{X}^- \in \mathbb{K} \) by subtracting \( t c \rho^{-1} \) and \( t c \hat{\rho}^{-1} \) from \( X^+ \) and \( X^- \), respectively. We then generate the KLE of \( \tilde{X}^+ \) and add back \( t c \rho^{-1} \) to the result, and apply the analogous procedure for \( X^- \). From (24), we see that the function \( g \) will have the form

\[
g(x) = c \int_x^\infty e^{-\rho s} \frac{ds}{s} = cE_1(\rho x),
\]

where \( E_1(x) := \int_x^\infty s^{-1} e^{-s} ds \) is the exponential integral function. Therefore,

\[
g^{-1}\left(\frac{T}{r}\right) = \frac{1}{\rho}E_1^{-1}\left(\frac{r}{Te}\right).
\]

There are many routines available to compute \( E_1 \); we choose a Fortran implementation to create a lookup table for \( E_1^{-1} \) with domain \([6.226 \times 10^{-22}, 45.47]\). We discretize this domain into 200,000 unevenly spaced points, such that the distance between two adjacent points is no more than 0.00231. Then, we use polynomial interpolation between points.
Now we fix $c = 1$, $\rho = 2$, $\hat{\rho} = 5$, and $T = 3$ and simulate some sample paths of
\[ S^{(d)} := \sum_{k=1}^{d} Z_k e_k(t), \]
for $d \in \{5, 10, 15, 20, 25, 100, 250, 500, 3000\}$. When simulating $Z_+^{(d)}$, we truncate the series (20) when $(Tc)^{-1} \Gamma_j > 45.47$; at this point, we have $g^{-1}(T^{-1} \Gamma_j) < \rho^{-1} 10^{-19}$. Using the fact that the $\{\Gamma_j\}_{j \geq 1}$ are nothing other than the waiting times of a Poisson process with intensity one, we estimate that we need to generate on average $45 Tc = 135$ random variables to simulate $Z_+^{(d)}$ and similarly for $Z_-^{(d)}$. We remark that for the chosen process both the decay and computation of $g^{-1}$ are manageable.

In Figure 1a, we show the sample paths resulting from a simulation of $S^{(d)}$. We notice that the numerical results correspond with the discussion of Section 3: the large movements of the sample path are already captured by the five-term approximation. We also notice peaks resulting from rapid oscillations before the bigger “jumps” in the higher term approximations. In classical Fourier analysis, this is referred to as the Gibbs phenomenon; the solution in that setting is to replace the partial sums by Cesàro sums. We can employ the same technique here, replacing $S^{(d)}$ with $C^{(d)}$, which is defined by
\[ C^{(d)} := \frac{1}{d} \sum_{k=1}^{d} S^{(d)}(k) . \]

It is relatively straightforward to show that $C^{(d)}$ converges to $X$ in the same manner as $S^{(d)}$ (as described in Theorem 1 (i)). In Figure 1b, we show the effect of replacing $S^{(d)}$ with $C^{(d)}$ on all sample paths; now the Gibbs phenomenon is no longer apparent.

References

Ash, R. B., and M. F. Gardner. 1975. Topics in stochastic process. New York–San Francisco–London: Academic Press.

Barczy, M., and R. L. Lovas. 2016. Karhunen–Loève expansion for a generalization of the Wiener bridge. preprint: arXiv:1602.05084v2.

Bertoin, J. 2001. Some elements on Lévy processes. In Stochastic processes: theory and methods, eds. C. R. Rao and D. N. Shanghag. Amsterdam, The Netherlands: Elsevier Science B.V.

Böttcher, B., J. Wang, and R. Schilling. 2013. Lévy matters III. Lévy-type processes: Construction, approximation and sample path properties. Berlin–Heidelberg–New York–London–Paris–Tokyo–Hong Kong–Barcelona–Budapest: Springer.
Devroye, L. 1986. An automated method for generating random variates with a given characteristic function. *SIAM Journal on Applied Mathematics* 46 (4):698–719.

Ghanem, R. G., and P. D. Spanos. 1991. *Stochastic finite elements: A spectral approach*. New York–Berlin–Heidelberg–London–Paris–Tokyo–Hong Kong–Barcelona: Springer–Verlag.

Jin, S. 2014. Gaussian processes: Karhunen–Loève expansion, small ball estimates and applications in time series modes. Unpublished doctoral diss., University of Delaware.

Kac, M., and A. Siegert. 1947. An explicit representation of a stationary Gaussian process. *Annals of Mathematical Statistics* 18:438–42.

Karhunen, K. 1947. Über lineare Methoden in der Wahrscheinlichkeitsrechnung. *American Academiae Scientiarum Fennicae. Series A* 1:37:3–79.

Loéve, M. 1948. Fonctions aleatoires du second ordre. In *Processus stochastique et mouvement Brownien*, ed. P. Lévy. Paris: Gauthier Villars.

Luo, W. 2006. Wiener chaos expansion and numerical solutions of stochastic partial differential equations. Unpublished doctoral diss., California Institute of Technology.

Madan, D. B., and E. Seneta. 1990. The Variance Gamma (V.G.) model for share market returns. *The Journal of Business* 63(4):511–24. URL: http://www.jstor.org/stable/2353303.

Phoon, K. K., H. W. Huang, and Quek, S. T. 2005. Simulation of strongly non-Gaussian processes using Karhunen–Loève expansion. *Probabilistic Engineering Mechanics* 20:188–98.

Rosiński, J. 1990. On series representations of infinitely divisible random vectors. *The Annals of Probability* 18(1):405–30.

Rosiński, J. 2001. Series representations of Lévy processes from the perspective of point processes. In *Lévy processes: theory and applications*, eds. O. E. Barndorff-Nielsen, T. Mikosh, and S. Resnick. Boston–Basel–Berlin: Birkhäuser.

Rosiński, J. 2007. Tempering stable processes. *Stochastic Processes and their Applications* 117 (6):677–707.

Sato, K. 1999. *Lévy processes and infinitely divisible distributions*. Cambridge–New York–Melbourne–Cape Town–Singapore–São Paulo: Cambridge University Press.

### A. Additional proof

**Proof of Lemma 1.** We give a proof for continuously differentiable \( \{ f_k \}_{k=1}^d \) only. The general statement follows fairly directly from the density of \( C^1([0, T]) \) in \( L^1([0, T]) \). Accordingly, we fix \( z \in \mathbb{R}^d \), a collection of continuously differentiable \( \{ f_k \}_{k=1}^d \) defined on \([0, T]\), and a Lévy process \( X \) with state space \( \mathbb{R} \). Instead of proving identity (7) directly for \( X \), we will prove that

\[
\Psi^{(b)}(z) = \int_0^T \Psi_X^{(b)}((z, u(t))) \, dt = b \int_0^T \Psi_X((z, u(t))) \, dt, \quad z \in \mathbb{R}^d, \quad b > 0, \tag{A1}
\]

where \( X^{(b)} \) is the process defined by \( X^{(b)}_t := X_{bt} \) and \( \xi^{(b)} \) is the vector with entries

\[
\xi^{(b)}_k := \int_0^T X_{bt} f_k(t) \, dt, \quad k \in \{1, 2, \ldots, k\}.
\]

It is clear that \( X^{(b)} \) is a Lévy process, that \( \Psi_X^{(b)} = b \Psi_X \), and that (7) corresponds to the special case \( b = 1 \). We focus on this more general result because it will lead directly to a proof of infinite divisibility. We begin by defining

\[
R^{(k)}_N := \frac{T}{N} \sum_{n=0}^{N-1} f_k \left( \frac{(n + 1)T}{N} \right) X_{b(n+1)T/N}, \quad k \in \{1, 2, \ldots, k\}, \quad N \in \mathbb{N}, \tag{A2}
\]

which are \( N \)-point, right-endpoint Riemann sum approximations of the random variables \( \xi_k^{(b)} \). By the usual telescoping sum technique for Lévy processes, we can write
\[ \frac{X_{[n+1]T}}{N} = \left( \frac{X_{[n+1]T}}{N} - \frac{X_{nT}}{N} \right) + \left( \frac{X_{nT}}{N} - \frac{X_{[n-1]T}}{N} \right) + \cdots + \left( \frac{X_{1T}}{N} - \frac{X_{0T}}{N} \right) + \frac{X_{0T}}{N} \]

\[ d \overset{\triangle}{=} \zeta^{(1)} + \zeta^{(2)} + \cdots + \zeta^{(n+1)}, \]

where the random variables \( \zeta^{(i)} \) are independent and each distributed like \( X_{\theta T/N} \). This allows us to rearrange the sum \( R_N^{(k)} \) according to the random variables \( \zeta^{(i)} \), gathering together those with the same index. Therefore, we have

\[ R_N^{(k)} = \sum_{n=0}^{N-1} \zeta^{(n+1)} \left( \frac{T}{N} \sum_{j=n}^{N-1} f_k \left( \frac{(j+1)T}{N} \right) \right). \]

We notice that the term in brackets on the right-hand side is a \( (N - n) \)-point, right-endpoint Riemann sum approximation for the integral of \( f_k \) over the interval \([nT/N, T] \). Let us therefore define

\[ t_{n,N}^{(k)} := \frac{T}{N} \sum_{j=n}^{N-1} f_k \left( \frac{(j+1)T}{N} \right), \quad \text{and} \quad s_{n,N}^{(k)} := \int_{\frac{nT}{N}}^{\frac{T}{N}} f_k(s) \, ds, \quad (A3) \]

as well as the \( d \)-dimensional vectors \( t_{n,N} \) and \( s_{n,N} \) consisting of entries \( t_{n,N}^{(k)} \) and \( s_{n,N}^{(k)} \), respectively. We observe that

\[ \mathbb{E}[\exp(i\langle \mathbf{z}, \xi^{(b)} \rangle)] = \lim_{N \to \infty} \mathbb{E} \left[ \exp \left( \sum_{n=0}^{N-1} i\zeta^{(n+1)} \langle \mathbf{z}, t_{n,N} \rangle \right) \right] = \lim_{N \to \infty} \exp \left( -\frac{bT}{N} \sum_{n=0}^{N-1} \Psi_X (\langle \mathbf{z}, t_{n,N} \rangle) \right), \quad (A4) \]

where we have used the dominated convergence theorem to obtain the first equality and the independence of the \( \zeta^{(i)} \) to obtain the final equality. Further, we get

\[ \exp \left( -\int_0^T \Psi_X^{(b)} (\langle \mathbf{z}, u(t) \rangle) \, dt \right) = \lim_{N \to \infty} \exp \left( -\frac{bT}{N} \sum_{n=0}^{N-1} \Psi_X (\langle \mathbf{z}, s_{n,N} \rangle) \right), \]

by using the left-endpoint Riemann sums. We note that \( |\langle \mathbf{z}, t_{n,N} \rangle - \langle \mathbf{z}, s_{n,N} \rangle| \to 0 \) uniformly in \( n \) since

\[ |\langle \mathbf{z}, t_{n,N} \rangle - \langle \mathbf{z}, s_{n,N} \rangle| \leq \sum_{k=1}^{d} |z_k| |t_{n,N}^{(k)} - s_{n,N}^{(k)}| \leq \frac{dT^2}{N} \max_{1 \leq k \leq d} \left\{ |z_k| \sup_{x \in [0,T]} |f_k'(x)| \right\}, \quad (A5) \]

where the last estimate follows from the well-known error bound \((c - a)^2 \sup_{x \in [a,c]} |g'(x)|)/N \) for the absolute difference between an \( N \)-point, right end-point Riemann sum and the integral of a \( C^1 \) function \( g \) over \([a, c] \). Then, by the continuity of \( \Psi_X \), for any \( \epsilon > 0 \), we may choose an appropriately large \( N \) such that

\[ \left| \frac{1}{N} \sum_{n=0}^{N-1} \Psi_X (\langle \mathbf{z}, t_{n,N} \rangle) - \frac{1}{N} \sum_{n=0}^{N-1} \Psi_X (\langle \mathbf{z}, s_{n,N} \rangle) \right| \leq \frac{1}{N} \sum_{n=0}^{N-1} \left| \Psi_X (\langle \mathbf{z}, t_{n,N} \rangle) - \Psi_X (\langle \mathbf{z}, s_{n,N} \rangle) \right| \leq \epsilon. \]

This proves (A1) and therefore also (7) for \( C^1 \) functions.
To establish the infinite divisibility of $\xi$, we note that (A1) shows that $\Psi_{\xi(b)} = b\Psi_{\xi(1)} = b\Psi_{\xi}$ and that $e^{-b\Psi_{\xi}}$ is therefore a positive definite function for every $b$ since it is the characteristic function of the random vector $\xi^{(b)}$. Positive definiteness follows from Bochner’s theorem (see e.g., Theorem 2.13 in Böttcher, Wang, and Schilling (2013)). Also, we clearly have $\Psi_{\xi}(0) = 0$ since $\Psi_{X}(0) = 0$. By Theorem 2.15 in Böttcher, Wang, and Schilling (2013), these two points combined show that $\Psi_{\xi}$ is the characteristic exponent of an ID probability distribution, and hence $\xi$ is an ID random vector. □