Generating Sparse Stochastic Processes Using Matched Splines

Leello Dadi, Shayan Aziznejad, and Michael Unser, *Fellow, IEEE*

***Abstract*—We provide an algorithm to generate trajectories of sparse stochastic processes that are solutions of linear ordinary differential equations driven by Lévy white noises. A recent paper showed that these processes are limits in law of generalized compound-Poisson processes. Based on this result, we derive an off-the-grid algorithm that generates arbitrarily close approximations of the target process. Our method relies on a B-spline representation of generalized compound-Poisson processes. We illustrate numerically the validity of our approach.**

***Index Terms*—Sparse stochastic processes, Lévy driven CARMA processes, B-splines, compound-Poisson processes.**

I. Introduction

Motivated by tractability and results such as the central-limit theorem, most of the early work in statistical signal processing has focused on Gaussian models [1]. In particular, the theory of Gaussian stationary processes provided justifications for the use of the discrete cosine transform [2] as an approximation of the Karhunen–Loève transform, and the Kalman filter [3] as an optimal estimator.

However, the analysis of real-world signals has revealed that the Gaussian framework may be insufficient to capture the breadth of the underlying behaviors [4], [5]. An important property that escapes the Gaussian framework is that of sparsity in some transform domains [6]. Sparsity being an essential component of modern signal processing [7]–[9], the authors of [10] proposed a wider stochastic framework that encompasses both Gaussian and sparsity-compatible models. Within this framework, a continuous-time signal is a realization of a stochastic process s that can be whitened by some linear, shift-invariant operator L. The key here is that the resulting white noise, or innovation, is not necessarily Gaussian. Put formally, signals are solutions of

$$\mathrm{L}s = w, \tag{1}$$

where w is a well defined innovation process called a Lévy white noise [11]. The term *Lévy* here comes from the fact that w is an object that can be interpreted as the derivative of a Lévy process in the sense of distributions [12], [13]. Whenever w is non-Gaussian, the realizations of s can be shown to be sparse. Accordingly, they have been named sparse stochastic processes [10]. Specific instances of such processes have been used to model natural signals such as images [14], [15], RF echoes in ultrasound [16], and network traffic in communication systems [17]–[19].

The goal of this paper is to generate realizations of the stochastic process s given its whitening operator L and a statistical characterization of its innovation process w. The computer generation of these signals can be of great interest to practitioners who wish to evaluate their reconstruction algorithms. We are thinking of works such as [20]–[23], where optimal estimators for interpolating and denoising such processes have been derived.

A possible approach to generate realizations of s would be to notice that, if L is a differential operator such as $\mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}t}$ or a polynomial in D, then (1) defines a stochastic differential equation (SDE) [24]. This becomes more apparent when notating w with the alternative notation $\mathrm{d}Z_t$, where $(Z_t)_{t\in\mathbb{R}^+}$ is a Lévy process (Chapter 7.4 in [10]). For example, $(\mathrm{D} - \alpha\mathrm{I})s = w$ can be rewritten as $\mathrm{d}S_t = \alpha S_t \mathrm{d}t + \mathrm{d}Z_t$. A suitable SDE solver, such as the one studied in [25], can then be used to generate an approximation of the signal. In particular, a common method is to solve the linear system of stochastic difference equations that is obtained by considering the discrete counter-part of the operator L (*e.g.* using finite differences instead of the derivative), and by replacing the innovation process w with a discrete white noise (see, for example, [26]).

It turns out that generic SDE solvers do not exploit the linearity of L. Here, the analytic treatment of (1) can be pushed further to obtain an explicit solution. Brockwell shows in [27] that s corresponds to the integral of a deterministic function with respect to a Lévy process. The integral can then be approximated by substituting it with a Riemann sum defined on a partition of the integration interval [28, Theorem 21].

These approaches, although valid, have drawbacks when it comes to the generation of synthetic signals for the evaluation of algorithms. First, they directly depend on the existence of a grid on which the approximation of the continuous process is sampled. This can lead to complication in the context of the multi-resolution algorithms that manipulate grid-free descriptions of signals. Second, the generated approximations are not solutions of an SDE in the form of (1). In other words, the approximations are not mathematical objects of the same nature as s.

In what follows, we propose a method that addresses both issues. It is based on a theoretical result by Fageot *et al.* [26] that states that any solution s of (1) is the limit in law of a sequence of simpler processes s_n. In other words, we have that

$$s_n \xrightarrow{\mathcal{L}} s, \text{ as } n \to \infty.$$

Leello Dadi and Shayan Aziznejad contributed equally to this work.

This work was done at the Biomedical Imaging Group, École polytechnique fédérale de Lausanne, Switzerland (e-mail: leello.tadesse@gmail.com; shayan.aziznejad@epfl.ch; michael.unser@epfl.ch). It was funded by the Swiss National Science Foundation under Grant 200020_184646 / 1.

These simpler processes, called generalized Poisson processes [10], have the advantage of having a grid-free numerical representation despite having a continuously defined domain. They fall within the category of (random) signals with a finite rate of innovation [29], [30]. They also have the desirable property of being whitened by the same operator L as the approximated signal. This implies that they all have the same correlation structure as the target signal (see Proposition 1).

Our method takes a sufficiently large value for n and generates a realization of the process s_n on a chosen interval. To do so, we consider an intermediary process called the generalized increment process. Interestingly, this process can be represented as a weighted sum of shifted B-splines and can be sampled very efficiently [31], [32]. The desired stochastic process s_n is then obtained from the latter by recursive filtering.

The outline of the paper is as follows: In Section II, we provide the necessary mathematical background. In Section III, we give a description of our algorithm: we begin by discussing the simulation of the innovation process in Subsection III-A. We then define the generalized increment process in Subsection III-B and we show how to generate its trajectories in Subsection III-C. Using this, we provide a recipe for generating sparse stochastic processes in Subsection III-D. In Subsection III-E, we show that our generation method perfectly reproduces the correlation structure of the target stochastic process. Finally, we conduct numerical investigations to show the validity of our method in Section IV.

II. Mathematical Foundations

In this section, we give a brief overview of the mathematical concepts that underly our approach. For a more detailed exposition, the reader is referred to [13], [26], [33], and references therein.

The Schwartz space $\mathcal{S}(\mathbb{R})$ is the space of smooth and rapidly decaying test functions. Its continuous dual, denoted by $\mathcal{S}'(\mathbb{R})$, is the space of tempered distributions. It is the space of all continuous linear functionals over $\mathcal{S}(\mathbb{R})$.

We denote by L an operator that is a continuous, linear, shift-invariant mapping from $\mathcal{S}'(\mathbb{R})$ to $\mathcal{S}'(\mathbb{R})$. The operator L is said to be shift-invariant if for any test function φ and any $t_0 \in \mathbb{R}$, we have that

$$\mathrm{L}\{\varphi\}(t-t_0) = \mathrm{L}\{\varphi(\cdot - t_0)\}(t), \quad t \in \mathbb{R},$$

where $\varphi(\cdot - t_0) : t \mapsto \varphi(t-t_0)$ is the shifted version of φ by t_0.

We restrict ourselves to rational operators in $\mathrm{D} = \frac{\mathrm{d}}{\mathrm{dt}}$, written $\mathrm{L} = P(\mathrm{D})Q(\mathrm{D})^{-1}$, where P and Q are polynomials such that $\deg(P) > \deg(Q)$. The latter assumption is crucial to have the minimum required regularity (point-wise definition) for the solution s of (1). The case $\mathrm{L} = \mathrm{D}$ is a typical choice that appears, for example, in the modeling of Brownian motion.

Rational operators are defined through their frequency response

$$\widehat{\mathrm{L}}(\omega) = \frac{P(\mathrm{j}\omega)}{Q(\mathrm{j}\omega)}.$$

They provide a succinct representation of the equation $P(\mathrm{D})s = Q(\mathrm{D})w$ that we can simply rewrite as $\mathrm{L}s = w$.

We are interested in generalized stochastic processes defined over $\mathcal{S}'(\mathbb{R})$. A generalized stochastic process w can be viewed as a random element of $\mathcal{S}'(\mathbb{R})$ in the sense that, for any $\varphi \in \mathcal{S}(\mathbb{R})$, the linear functional $\varphi \mapsto \langle \varphi, w \rangle \in \mathbb{R}$ is a well defined random variable over $\mathbb{R}$ (See Appendix A for a formal definition).

A. Lévy White Noises

Lévy white noises constitute an important class of generalized stochastic processes, whose specification is essential to our framework. The three important operational properties of Lévy white noises for our purpose are:

1) **Stationarity:** For any $\varphi \in \mathcal{S}(\mathbb{R})$ and $\tau \in \mathbb{R}$, the random variables $\langle \varphi, w \rangle$ and $\langle \varphi(\cdot - \tau), w \rangle$ are identically distributed.
2) **Independence:** For any $\varphi_1, \varphi_2 \in \mathcal{S}(\mathbb{R})$ with disjoint supports, the random variables $\langle \varphi_1, w \rangle$ and $\langle \varphi_2, w \rangle$ are independent.
3) **Characterization of the probability law:** For any Lévy white noises w in $\mathcal{S}'(\mathbb{R})$ and for any test function $\varphi \in \mathcal{S}(\mathbb{R})$, the characteristic function of the random variable $X_\varphi = \langle \varphi, w \rangle$ can be specified as

$$\widehat{\mathcal{P}}_{X_\varphi}(\xi) = \mathbb{E}[\mathrm{e}^{\mathrm{j}\xi\langle \varphi, w\rangle}] = \exp\left(\int_{\mathbb{R}} f(\xi\varphi(r))\mathrm{d}r\right), \quad (2)$$

where the function $f : \mathbb{R} \to \mathbb{C}$ is called the Lévy exponent of w.

Formally, this Lévy exponent can be obtained as

$$f(\xi) = \log\left(\widehat{\mathcal{P}}_{X_{\text{rect}}}(\xi)\right),$$

where $X_{\text{rect}} = \langle \text{rect}_{[0,1]}, w \rangle$[1] is the observation of w through the rectangular window

$$\text{rect}_{[0,1]}(x) = \begin{cases} 1, & 0 < x \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

The distribution of X_{rect} gives us the Lévy exponent f that defines (2), so that we can determine all the statistics of w from the knowledge of X_{rect}.

In particular, the following Proposition from [10] connects the second-order statistics of w to those of X_{rect}.

Proposition 1 ([10], Theorem 4.15)**.** *Let w be a Lévy white noise such that $X_{\text{rect}} = \langle \text{rect}_{[0,1]}, w \rangle$ has zero mean and a finite variance $\sigma_w^2 = \mathbb{E}[X_{\text{rect}}^2]$. Then,*

$$\forall \varphi_1, \varphi_2 \in \mathcal{S}(\mathbb{R}), \quad \mathbb{E}[\langle \varphi_1, w \rangle \langle \varphi_2, w \rangle] = \sigma_w^2 \langle \varphi_1, \varphi_2 \rangle.$$

It turns out that X_{rect} is an infinitely divisible random variable in the sense of Definition 1 [35].

Definition 1. *A real-valued random variable X is said to be infinitely divisible if, for any natural number $M \in \mathbb{N}$, there exist M independent and identically distributed random variables $X_1, ..., X_M$ such that*

$$X = X_1 + \cdots + X_M.$$

[1] Although $\text{rect}_{[0,1]}$ is not in $\mathcal{S}(\mathbb{R})$, the random variable X_{rect} can still be defined. For more details, see [34].

TABLE I
INFINITELY DIVISIBLE DISTRIBUTIONS AND THEIR LÉVY EXPONENTS

Distribution	**Lévy exponent**
Gaussian (μ, σ)	$\mathrm{j}\mu\xi - \sigma^2\xi^2/2$
Symmetric α-stable $(\alpha, c), \alpha \in (0, 2]$	$-\lvert c\xi\rvert^\alpha$
Gamma(α, β)	$-\beta \log(1 - \mathrm{j}\xi/\alpha)$
Laplace (μ, b)	$\mathrm{j}\mu\xi - \log(1 + b^2\xi^2)$

To check the infinite divisibility of X_{rect}, one can note that, for any $M \in \mathbb{N}$, we have that

$$X_{\text{rect}} = \langle \text{rect}_{[0,1]}, w \rangle = \langle \sum_{m=0}^{M-1} \text{rect}_{[\frac{m}{M}, \frac{m+1}{M}]}, w \rangle = \sum_{m=0}^{M-1} \langle \text{rect}_{[\frac{m}{M}, \frac{m+1}{M}]}, w \rangle. \quad (3)$$

The terms in the sum (3) are independent and identically distributed random variables as a consequence of the independence and stationarity properties of white noises, which certifies that $\langle \text{rect}_{[0,1]}, w \rangle$ is infinitely divisible.

The converse is also true: for any regular[2] infinitely divisible random variable X with Lévy exponent $f(\xi) = \log(\mathbb{E}[\mathrm{e}^{\mathrm{j}\xi X}])$, there exists a well defined Lévy white noise w whose statistics are determined by (2) [35]–[37]. This shows that there is a one-to-one correspondence between infinitely divisible distributions and *Lévy* white noises through $\langle \text{rect}_{[0,1]}, w \rangle$.

The Gaussian, gamma, and α-stable distributions are classical examples of infinitely divisible distributions [12]. We can plug in their Lévy exponents in (2) to define their corresponding Lévy white noises. We repeat in Table I some infinitely divisible distributions of interest, along with their Lévy exponents [35].

A case of special interest is when f is the Lévy exponent of a compound-Poisson distribution. A compound-Poisson random variable X, with rate λ and amplitude law ν, is defined as

$$X = \sum_{k=1}^{K} A_k,$$

where the number K is a Poisson random variable with parameter λ and $(A_k)_{k=1}^K$ is an i.i.d. sequence drawn according to ν. We refer to the corresponding Lévy white noise w as a compound-Poisson innovation. It is known to be equal in law to

$$w = \sum_{k \in \mathbb{Z}} A_k \delta(\cdot - \tau_k), \quad (4)$$

where $(\tau_k)_{k \in \mathbb{Z}}$ are the locations of impulses with rate λ [13]. The law of these impulses is as follows: for any interval $[a, b]$, the number of impulses in $[a, b]$ is a Poisson random variable with parameter $\lambda(b - a)$.

On any finite interval, compound-Poisson innovations have a finite representation. They can be stored on a computer with the quantization of real numbers as sole source of information loss. They are therefore well adapted to simulation purposes.

[2]The random variable X is said to be regular, if $\mathbb{E}[|X|^\epsilon] < +\infty$ for some $\epsilon > 0$.

B. Generalized Lévy Processes

The sparse-stochastic-process framework of Unser *et al.* [10] is a comprehensive theory of generalized Lévy Processes. These are stochastic processes that can be whitened by some admissible linear, shift-invariant operator. More precisely, s is a generalized Lévy process if there exists an operator L such that $w = \mathrm{L}s$ is a Lévy white noise. Equivalently, one may view generalized Lévy processes as the solution of the stochastic differential equation

$$\mathrm{L}s = w. \quad (5)$$

It has been shown that, under mild technical assumptions on L and w, a solution s of (5) exists and constitutes a properly defined generalized stochastic process over $\mathcal{S}'(\mathbb{R})$ [36].

When L is an operator with a trivial null space, such as $\mathrm{L} = (\mathrm{D} - \alpha\mathrm{I})$ with $\Re(\alpha) \neq 0$, we can write that

$$s = \mathrm{L}^{-1} w,$$

where L^{-1} is the inverse of L. However, when the null space is nontrivial, for instance when L corresponds to an unstable ordinary differential equation, the specification of the boundary conditions become necessary to uniquely identify the solution. The boundary conditions take the form

$$\phi_\ell(s) = c_\ell, \quad \ell = 1, \ldots, N_0, \quad (6)$$

where $\phi_\ell : s \mapsto \phi(s) \in \mathbb{R}$ are appropriate linear functionals, $c_\ell \in \mathbb{R}$, and N_0 is the dimension of the null space of L. For instance, one can impose that the process s takes fixed values at reference locations $t_1 < \ldots < t_{N_0}$; that is, $\phi_\ell(s) = s(t_\ell) = c_\ell$ for $\ell = 1, \ldots, N_0$. Such boundary conditions appear in the classical definition of Lévy processes (including Brownian motion), where we have that $\phi(s) = s(0) = 0$ (Chapter 7 of [10]). We formally write

$$s = \mathrm{L}_{\boldsymbol{\phi}}^{-1} w,$$

where $\mathrm{L}_{\boldsymbol{\phi}}^{-1}$ is the right inverse of L. It incorporates the boundary conditions (6) (Chapter 5.4 of [10]).

When w is a compound-Poisson innovation of the form (4), the process $s = \mathrm{L}_{\boldsymbol{\phi}}^{-1} w$ ($\mathrm{L}^{-1}s$, respectively, when the null space of L is trivial) is called a generalized Poisson process.

The fundamental property for this work is that any sparse stochastic process s that is the solution of (5) can be specified as the limit in law of a sequence $\{s_n\}_{n \in \mathbb{N}}$ of generalized Poisson processes [26]. The corresponding driving processes $w_n = \mathrm{L}s_n$ are compound-Poisson innovations of the form

$$w_n = \sum_{k \in \mathbb{Z}} A_{k,n} \delta(\cdot - \tau_{k,n}) \quad (7)$$

with rates $\lambda_n = n$ and with i.i.d. amplitudes $A_{k,n}$ that are infinitely divisible random variables with Lévy exponent $f_n = \frac{1}{n} f$, where f is the Lévy exponent of w.

C. Green's Functions

The Green's function of a differential operator L is a tempered distribution $\rho_{\mathrm{L}} \in \mathcal{S}'(\mathbb{R})$ that satisfies

$$\mathrm{L}\rho_{\mathrm{L}} = \delta.$$

It can be viewed as the impulse response of the inverse of L. The canonical Green's function is

$$\rho_{\mathrm{L}} = \mathcal{F}^{-1}\left\{\frac{1}{\widehat{\mathrm{L}}(\omega)}\right\},$$

where $\widehat{\mathrm{L}}$ is the frequency response of L (Chapter 5.2 of [10]). This definition can be made to stay valid even when $\widehat{\mathrm{L}}$ vanishes at some points, as long as $\frac{1}{\widehat{\mathrm{L}}(\omega)}$ is in $\mathcal{S}'(\mathbb{R})$. For details on how to compute Green's functions, the reader is referred to Appendix B. We have plotted the Green's function of several operators in Figure 1 to highlight their variety and their dependence on L.

III. Method

In this section, we introduce our method for generating (approximate) trajectories of a sparse stochastic process s that is whitened by an operator L and whose innovation noise is w. When necessary, we assume general boundary conditions of the form $\phi_\ell(s) = 0$ for $\ell = 1, ..., N_0$, where N_0 is the dimension of the null space of L.

As mentioned earlier, the process s is the limit of generalized compound-Poisson processes s_n driven by $w_n = \mathrm{L}s_n$, a compound-Poisson innovation of the form (7). The process s_n can therefore be written

$$s_n = \sum_{k\in\mathbb{Z}} A_{k,n}\rho_{\mathrm{L}}(\cdot - \tau_{k,n}) + p_{0,n},$$

where ρ_{L} is a Green's function of L and $p_{0,n}$ is an element of the null space of L determined by boundary conditions (it vanishes when L is invertible). Indeed, we have that

$$\begin{aligned}\mathrm{L}\left\{\sum_{k\in\mathbb{Z}} A_{k,n}\rho_{\mathrm{L}}(\cdot - \tau_{k,n}) + p_{0,n}\right\} &= \sum_{k\in\mathbb{Z}} A_{k,n}\mathrm{L}\{\rho_{\mathrm{L}}(\cdot - \tau_{k,n})\}\\ &= \sum_{k\in\mathbb{Z}} A_{k,n}\delta(\cdot - \tau_{k,n})\\ &= w_n.\end{aligned}$$

For large values of n, the process s_n is assumed to be a good approximation of s. So, our goal is to generate samples of s_n on any uniform grid over any interval $[0, T]$. More precisely, once an interval $[0, T]$ is specified and a regular grid with step size h is provided, our aim is to obtain the vector $\mathbf{s}_n$ whose components are $[\mathbf{s}_n]_i = s_n(ih)$, for $i = 0, ..., \left(\lceil \frac{T}{h} \rceil - 1\right)$.

A. Simulating the Innovation Process

We begin by obtaining a realization of the driving innovation w_n. It consists of a sequence of impulse locations $(\tau_{k,n})$ and a corresponding sequence of amplitudes $(A_{k,n})$.

The sequence $(\tau_{k,n})$ is a point Poisson process. Its realization on the interval $[0, T]$ is simulated in two steps. First, a Poisson random variable K with parameter $\lambda = nT$ is generated. Then, K impulse locations $(\tau_{k,n})_{k\in\{1,...,K\}}$ are sampled uniformly on $[0, T]$.

The next step is to simulate the K corresponding amplitudes $(A_{k,n})_{k\in\{1,...,K\}}$. The characteristic function of the amplitudes variable A is

$$\xi \mapsto \exp\left(\frac{1}{n} f(\xi)\right).$$

TABLE II
The nth root of infinitely divisible distributions

Distribution	**nth Root**
Gaussian (μ, σ)	Gaussian $(\frac{\mu}{n}, \frac{\sigma}{\sqrt{n}})$
α-Stable (α, β, μ, c)	If $\alpha \neq 1$, $(\alpha, \beta, \frac{\mu}{n}, \frac{c}{n^{\frac{1}{\alpha}}})$,
	If $\alpha = 1$, $(\alpha, \beta, \frac{\mu}{n} - \frac{2}{\pi}c\beta\frac{\log(n)}{n}, \frac{c}{n})$
Gamma(α, β)	Gamma$(\frac{\alpha}{n}, \beta)$
Compound-Poisson of intensity λ	Compound-Poisson of intensity $\frac{\lambda}{n}$
Laplace (μ, b)	$X_n = \frac{\mu}{n} + b(G_1^{(n)} - G_2^{(n)})$
	with $G_1^{(n)}, G_1^{(n)} \sim$ Gamma$(\frac{1}{n}, 1)$

We refer to it as the *n*th root of the law of $\langle \mathrm{rect}_{[0,1]}, w\rangle$. Our assumption in this paper is that there exists, for any $n \in \mathbb{N}$, a known method[3] to generate infinitely divisible variables with Lévy exponent $\frac{1}{n}f(\xi)$. For common parametric distributions such as α-stable, Laplace, and gamma distributions, such sampling methods [39] are well known and implemented in scientific computing libraries[4]. Simulating from their *n*th root is a simple matter of rescaling their parameters, as summarized in Table II. By applying the correct rescaling, we simulate K independent amplitudes and thus obtain the sequence $(A_{k,n})_{k\in\{1,...,K\}}$.

B. Generalized Increment Process

With the impulse locations $(\tau_{k,n})_{k\in\{1,...,K\}}$ and amplitudes $(A_{k,n})_{k\in\{1,...,K\}}$ in hand, we can compute samples of

$$s_n(\cdot) = \sum_{k=1}^{K} A_{k,n}\rho_{\mathrm{L}}(\cdot - \tau_{k,n}) + p_{0,n} \tag{8}$$

on a grid.

A direct approach to generate $\mathbf{s}_n$ is to use the expansion (8) and represent the process as a sum of shifted Green's functions. However in this case, the determination of $s_n(t)$ at any point $t \in [0, T]$ may require nontrivial computation of each and every term in (8). This stems from the fact that Green's functions are infinitely supported in general. There are therefore potential drawbacks to expansions in the basis of shifted Green's functions like (8). To overcome these issues, we propose instead an alternative method based on B-splines.

Recall that L is a rational operator of the form $P(\mathrm{D})Q(\mathrm{D})^{-1}$, where we take $\{\alpha_1, ..., \alpha_{\deg(P)}\}$ to be the roots of P, with possible repetitions. Its discrete counterpart $\mathrm{L}_{\mathrm{d}}^h$ is defined as

$$\mathrm{L}_{\mathrm{d}}^h\{f\} = \sum_{m=0}^{\deg(P)} r[m]f(\cdot - mh),$$

[3]Workarounds exists for when a sampling method for the nth root $\frac{1}{n}f(\xi)$ is unavailable. For instance, one can opt for an approximate sampling scheme such as in [38].

[4]E. Jones, et al., "SciPy: Open source scientific tools for Python," 2001.

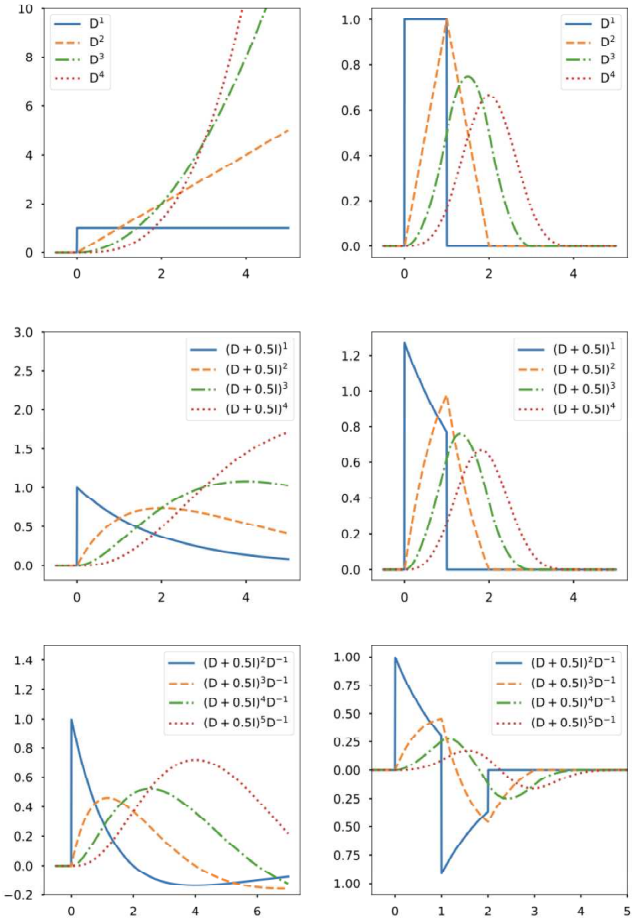

Fig. 1. Green's functions (left) and B-splines (right) associated with specific operators L.

where the sequence r is determined through its Fourier transform

$$R(\mathrm{e}^{\mathrm{j}\omega}) = \sum_{m=0}^{\deg(P)} r[m]\mathrm{e}^{-\mathrm{j}\omega m} = \prod_{m=1}^{\deg(P)} (1 - \mathrm{e}^{\alpha_m h}\mathrm{e}^{-\mathrm{j}\omega h}).$$

It is a finite impulse-response filter (FIR). Its null space contains the null space of L [31]. The function $\beta_{\mathrm{L}}^h := \mathrm{L}_{\mathrm{d}}^h\{\rho_{\mathrm{L}}\}$ is called the B-spline corresponding to L [40]. The B-spline has the fundamental property of being the shortest possible function within the space of cardinal L-splines (its support is included in $[0, \deg(P) \times h]$) [41], [42]. This will turn out to be crucial for the numerical efficiency of our method. Moreover, they reproduce both the Green's function and elements in the null space of their corresponding operator L [10, Section 6.4.]. Examples of relevant generalized B-splines are shown in Figure 1 (right figures). Note how they contrast with the corresponding infinitely supported Green's functions (left figures).

The application of $\mathrm{L}_{\mathrm{d}}^h$ to s_n yields

$$u_n(t) = \mathrm{L}_{\mathrm{d}}^h\{s_n\}(t) = \sum_{m=0}^{\deg(P)} r[m]s_n(t - mh). \tag{9}$$

The process u_n in (9) is called the generalized increment process. Interestingly, it can be written as a sum of compactly supported terms, like

$$\begin{aligned} u_n(t) &= \mathrm{L}_{\mathrm{d}}^h\{s_n\}(t) \\ &= \sum_{k=1}^{K} A_{k,n}\mathrm{L}_{\mathrm{d}}^h\{\rho_{\mathrm{L}}(\cdot - \tau_{k,n})\}(t) + \mathrm{L}_{\mathrm{d}}^h\{p_{0,n}\}(t) \\ &= \sum_{k=1}^{K} A_{k,n}\beta_{\mathrm{L}}^h(t - \tau_{k,n}) + \quad 0. \end{aligned}$$

The process u_n, along with boundary conditions, is our alternate representation of s_n. Now, let $\mathbf{u}_n$ be the vector whose components are $[\mathbf{u}_n]_i = u(ih)$, for $i = 1, ..., \left(\lceil \frac{T}{h} \rceil - 1\right)$. This vector can be computed more efficiently than $\mathbf{s}_n$ since the process u_n admits a representation with compactly supported terms. Moreover, $\mathbf{u}_n$ is linearly related to the vector $\mathbf{s}_n$ via a discrete system of difference equations. Indeed, we have that

$$[\mathbf{u}_n]_i = \sum_{m=0}^{\deg(P)} r[m][\mathbf{s}_n]_{i-m}, \tag{10}$$

for $\deg(P) \leq i \leq \left(\lceil \frac{T}{h} \rceil - 1\right)$. For $0 < i < \deg(P)$, we have that

$$[\mathbf{u}_n]_i = \sum_{m=0}^{\deg(P)} r[m]\mathbf{s}_n((i-m)h),$$

where the values $\mathbf{s}_n(-mh)$ for $m = 0, ..., (\deg(P) - 1)$ provide the boundary values. These relations are established by writing (9) with $t = ih$. The boundary values are determined by the null-space term $p_{0,n}$, which is itself determined by the boundary conditions.

Thus, once we have evaluated $\mathbf{u}_n$, we can obtain $\mathbf{s}_n$ by solving (10), which is accomplished by applying a recursive reverse filter to $\mathbf{u}_n$. This is performed by rewriting (10) as

$$[\mathbf{s}_n]_i = \frac{1}{r[0]}\left([\mathbf{u}_n]_i - \sum_{m=1}^{\deg(P)} r[m][\mathbf{s}_n]_{i-m}\right). \tag{11}$$

By substitution of the boundary values when necessary (*i.e.*, taking $\mathbf{s}_n((i-m)h)$ instead of $[\mathbf{s}_n]_{i-m}$ when $(i-m) \leq 0$), (11) allows one to recursively compute the components of $\mathbf{s}_n$.

C. Computing the Generalized Increment Process

We now describe an efficient procedure to compute the generalized increment process. The components of $\mathbf{u}_n$ are given by

$$[\mathbf{u}_n]_i = \sum_{k=0}^{K} A_{k,n}\beta_{\mathrm{L}}^h(ih - \tau_{k,n}).$$

The naive approach here would be to iterate through each grid point i independently and compute $[\mathbf{u}_n]_i$. Doing so would require one to read the entire sequence of impulse locations (τ_k) for each i. This cannot be avoided since there is no information on the sequence (τ_k), aside from its inclusion in $[0, T]$. We simply would not know *which* B-spline terms are inactive, so we would have to iterate through them all. A

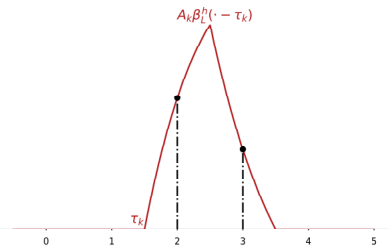

Fig. 2. For a single B-spline term, it is only the grid points that sit within the support of the B-spline that are incremented (black stems).

more efficient approach is to iterate through the list of impulses instead of the grid points.

The idea is as follows: First, initialize the vector $\mathbf{u}_n$ to zeros. Then, read the list of impulse locations one by one. For each impulse at τ_k, find the grid points that lie within the support of the B-spline at τ_k. Then, increment the value of $\mathbf{u}_n$ on those grid points by the contribution of the considered B-spline (see Figure 2). In one pass over the list of impulses, this method computes the values $[\mathbf{u}_n]_i = \mathbf{u}_n(ih)$.

This intermediate computation of the generalized-increment process provides a considerable gain in terms of efficiency. Instead of having a number of operations that scales with $\lceil \frac{T}{h} \rceil \times K$ for the Green's function representation, we have one that scales with $\deg(P) \times \big(\lceil \frac{T}{h} \rceil + K\big)$.

D. Recipe to Generate Trajectories

Here is a summary of the procedure that generates trajectories of $\mathrm{L}s_n = w_n$.

First, fix the infinitely divisible distribution[5] that corresponds to w and define the operator L by identifying the polynomials P and Q.

Pick a sufficiently large value for n. Intuitively, n should be large enough to ensure the occurrence of several jumps in each bin. In other words, we expect n to be of the same order as h^{-1}. This has been validated with our numerical experiments as well, where we show that it provides a good approximation of the underlying statistics of the process (see Subsection IV-C and Figures 5, 6, and 7).

Pick a simulation interval $[0, T]$ and generate w_n as described in section III-A. Determine an explicit form for ρ_{L}. At this point, the grid-free approximation s_n (expressed as in (8)) is available and can be stored.

Fix a grid on $[0, T]$ by choosing a step size h. Then determine the vector $\mathbf{s}_n$ with component $[\mathbf{s}_n]_i = s_n(ih)$. Compute the FIR filter $\mathrm{L}_{\mathrm{d}}^h$ and obtain $\beta_{\mathrm{L}}^h = \mathrm{L}_{\mathrm{d}}^h\{\rho_{\mathrm{L}}\}$. Then, compute the generalized increment vector $\mathbf{u}_n$ as described in Section III-C.

To obtain $\mathbf{s}_n$, apply the reverse filter to $\mathbf{u}_n$ following (11). Take the values $\mathbf{s}_n(-mh)$ for $m = 0, ..., (\deg(P) - 1)$ to be zero for most cases except when L has a nontrivial null space, in which case it is derived from boundary conditions. The pseudocode of our method is provided in Algorithm 1.

[5]The choice here is restricted to parametric families we can rescale and simulate.

Input : Coefficients of P and Q, approximation level n, interval size T, step size h
Output: Vector $\mathbf{s}_n$
Compute ρ_{L} and the FIR filter $r[m]$
Compute $\beta_L^h = \mathrm{L}_d^h\{\rho_{\mathrm{L}}\}$
Generate $[(\tau_1, A_1), ..., (\tau_K, A_K)]$
Initialize $\mathbf{u}_n$ with zeros as an array of size $\lceil \frac{T}{h} \rceil$
foreach (τ_k, A_k) **do**
  Find closest grid point $i_{\mathrm{grid}} = \lfloor \frac{\tau_k}{h} \rfloor$
  foreach *i in* $\{i_{\mathrm{grid}}, \ldots, i_{\mathrm{grid}} + \deg(P)\}$ **do**
    $[\mathbf{u}_n]_i \leftarrow [\mathbf{u}_n]_i + A_k \times \beta_L^h(ih - \tau_k)$
  end
end
Recursively apply a reverse filter to $\mathbf{u}_n$ following (11)

Algorithm 1: Procedure to obtain $\mathbf{s}_n$.

E. Correlation Structure

In this section, we show a merit of our method by proving that the generated approximations preserve the correlation structure of the target process.

First note that for any white Lévy noise w, we have that

$$w_n \xrightarrow{\mathcal{L}} w,$$

where the sequence of compound-Poisson innovations $(w_n)_{n\in\mathbb{N}}$ is defined in (7). We refer to this approximating sequence in Proposition 2.

Proposition 2. *Let w be a Lévy white noise such that $X_{\mathrm{rect}} = \langle \mathrm{rect}_{[0,1]}, w \rangle$ has zero mean and the finite variance $\sigma_w^2 = \mathbb{E}[X_{\mathrm{rect}}^2]$. Let $n \in \mathbb{N}$ and let w_n be a compound-Poisson innovation that approximates w as defined in* (7)*. Denoting $X_{\mathrm{rect},n} = \langle \mathrm{rect}_{[0,1]}, w_n \rangle$, we have that*

$$\mathbb{E}[X_{\mathrm{rect},n}] = \mathbb{E}[X_{\mathrm{rect}}] = 0$$

and

$$\sigma_{w_n}^2 = \mathbb{E}[X_{\mathrm{rect},n}^2] = \mathbb{E}[X_{\mathrm{rect}}^2] = \sigma_w^2.$$

The proof can be found in Appendix C. Now, if $s_n = \mathrm{L}^{-1}w_n$ is a generalized Poisson process that approximates $s = \mathrm{L}^{-1}w$, then

$$\begin{aligned}
\mathbb{E}[\langle \varphi_1, s_n \rangle \langle \varphi_2, s_n \rangle] &= \mathbb{E}[\langle \varphi_1, \mathrm{L}^{-1} w_n \rangle \langle \varphi_2, \mathrm{L}^{-1} w_n \rangle] \\
&= \mathbb{E}[\langle \mathrm{L}^{-1*} \varphi_1, w_n \rangle \langle \mathrm{L}^{-1*} \varphi_2, w_n \rangle] \\
&= \sigma_{w_n}^2 \langle \mathrm{L}^{-1*} \varphi_1, \mathrm{L}^{-1*} \varphi_2 \rangle \\
&= \sigma_w^2 \langle \mathrm{L}^{-1*} \varphi_1, \mathrm{L}^{-1*} \varphi_2 \rangle \\
&= \mathbb{E}[\langle \varphi_1, s \rangle \langle \varphi_2, s \rangle]. \qquad (12)
\end{aligned}$$

From (12), we concluded that, more than just approximated, the correlation structure is preserved *exactly* in our method.

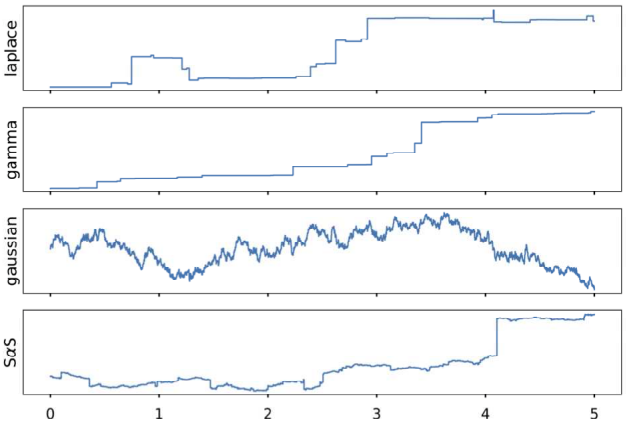

Fig. 3. Trajectories of Lévy processes (L = D) with different innovations. From top to bottom: Laplace(0, 1), gamma(1, 1), Gaussian(0,1), and symmetric-α-stable with $\alpha = 1.23$.

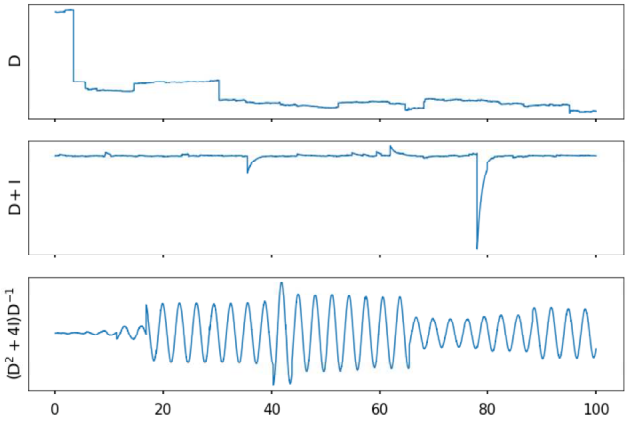

Fig. 4. Trajectories of the solution s of $\mathrm{L}s = w$ for different operators L. In all cases, we considered a symmetric-α-stable white noise w with $\alpha = 1.23$.

IV. Numerical Experiments

In this section, we validate our approach by conducting several numerical experiments. Let us also mention that a Python library that implements our algorithm can be found online[6]. Moreover, an accompanying web interface is also designed and is available [7].

A. Generating Lévy Processes

Among all processes we can generate, those that are solutions to $\mathrm{D}s = w$ are called Lévy processes when the boundary condition is $s(0) = 0$. We showcase in Figure 3 different Lévy processes that correspond to several infinitely divisible distributions. For all four simulations, we took $n = 1{,}000$ and $h = 0.001$. As we demonstrate in Section IV-C, a reasonable choice for these parameter is to set nh to be a small integer (here, $nh = 1$). The visual appearance of the trajectories matches our expectations: The trajectory driven by a Gaussian innovation has the appearance of Brownian motion; the gamma Lévy process is nondecreasing.

B. Choice of the Operator

Our framework allows for any rational operator of the form $P(\mathrm{D})Q(\mathrm{D})^{-1}$, so long as $\deg(P) > \deg(Q)$. In Figure 4, we generate trajectories of s that are solution of $\mathrm{L}s = w$, where w is a symmetric-α-stable innovation with $\alpha = 1.23$. Here we took $n = 200$ and $h = 0.001$. We see that, for various choices of L, the characteristics of the signal are markedly different, which exhibits the breadth of the modeling framework proposed in [10].

C. Convergence as n Grows

In Figure 5, we illustrate how an increase in n improves the approximation. In addition, we have depicted the convergence of moments in Figure 6. While the two figures emphasize the effect of n, they are insufficient to provide a quantitative way to choose n.

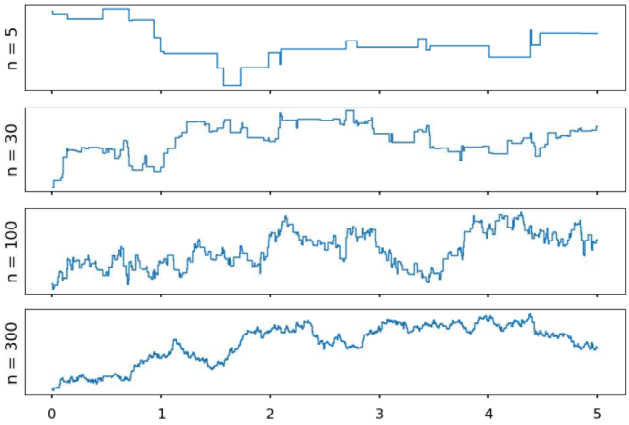

Fig. 5. Approximations of Brownian motion (solution to $\mathrm{D}s = w$, with w a Gaussian white Lévy noise) as n increases.

Here, we propose a measure that is based on the statistics of the generalized increment process. Since the process u_n is maximally decoupled, we can estimate the distribution of $U_n = \langle \beta_{\mathrm{L}}^{h\vee}, w_n \rangle$ from the samples $\{[\mathbf{u}_n]_i\}_i$ of the generalized increment process on the grid and obtain the empirical cumulative distribution function (CDF) $\tilde{F}_n(\cdot)$ of U_n. We then compare this empirical function to the reference CDF $F(\cdot)$ of $U = \langle \beta_{\mathrm{L}}^{h\vee}, w \rangle$. For the comparison, we use the Kolmogorov-Smirnov (KS) divergence [43] defined as

$$\mathcal{KS}(\tilde{F}_n, F) = \max_{x \in \mathbb{R}} |\tilde{F}_n(x) - F(x)|.$$

We then select n such that the KS-divergence is smaller than a certain threshold (*e.g.*, smaller than 0.1). The choice of the threshold is conditioned by the desired numerical precision: The lower the threshold, the more faithful the trajectories, but the higher the computational cost of the algorithm.

[6]https://github.com/Biomedical-Imaging-Group/Generating-Sparse-Processes

[7]https://saturdaygenfo.pythonanywhere.com

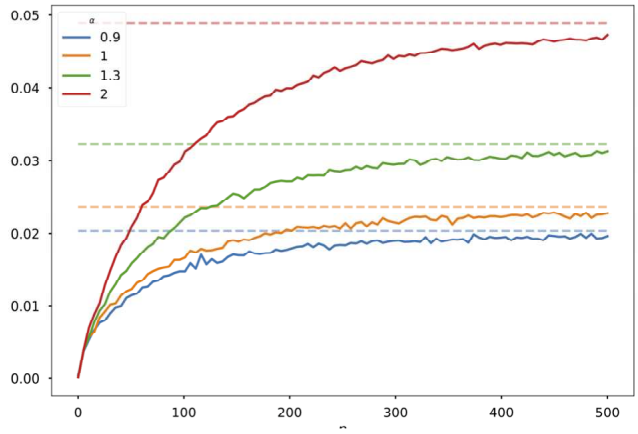

Fig. 6. Convergence of $\mathbb{E}[|\langle \text{rect}_{[0,h]}, s_n \rangle|^p]$ to $\mathbb{E}[|\langle \text{rect}_{[0,h]}, s \rangle|^p]$ for $p = 0.4$, $h = 0.01$, and several symmetric-α-stable Lévy white noises w. The expectations are estimated with 10,000 trajectories for each n.

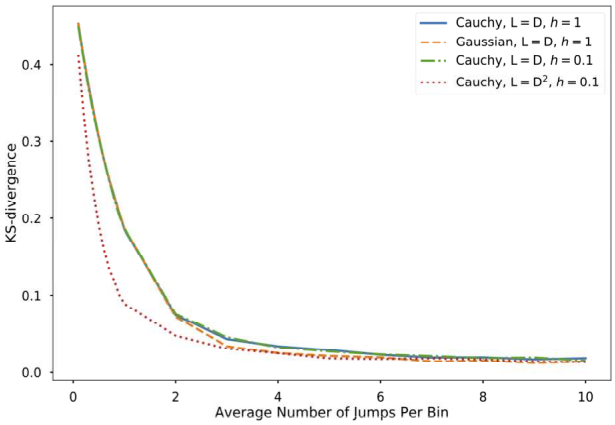

Fig. 7. Kolmogorov-Smirnov (KS) divergence versus the average number of jumps per bin ($N_{\text{jumps}} = nh$).

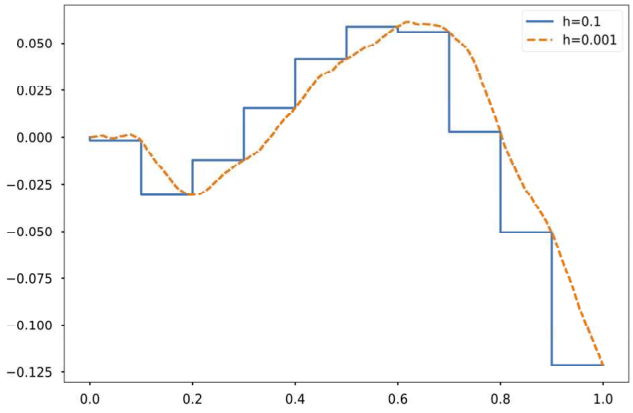

Fig. 8. Single grid-free approximation sampled on grids that differ by their step size.

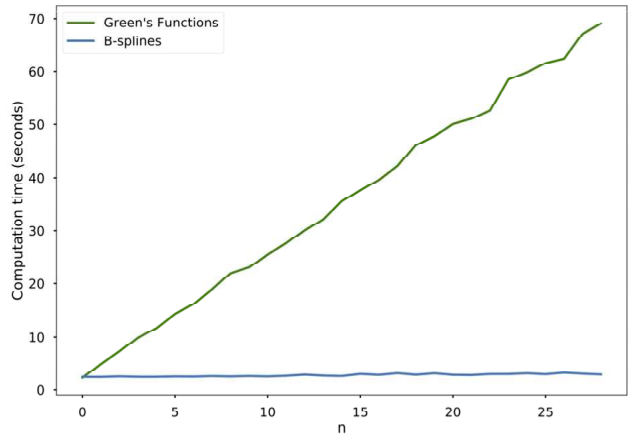

Fig. 9. Average computation time for a trajectory of the solution of $(\mathrm{D} - 0.5\mathrm{I})s = w$, where w a Gaussian white noise. The simulation interval is $[0, 1]$ with step size $h = 0.001$.

Intuitively, we expect that it is necessary to have several jumps in each bin in order to properly approximate the statistics of the process. The average number of jumps in each bin of length h is $N_{\text{jumps}} = nh$, so we expect n to be in the order of h^{-1}.

In Figure 7, we have validated this intuition by plotting the KS-divergence for different values of N_{jumps} in various settings. In all cases, as N_{jumps} increases, the KS-divergence decreases to a baseline error value, due to the finite-sample estimation of the underlying distribution.

D. Benefits of Grid-Free Approximations

Recall that a main motivation for our algorithm was to make it compatible with multi-grid methods. In our approach, the approximation s_n lives off the grid. It is only after the specification of the step size h that s_n is sampled on a grid. The generation of the random variables to determine s_n and the sampling on a grid are completely decoupled. This means that the same approximation s_n can be viewed through different grids, which we illustrate in Figure 8. The solution to $(\mathrm{D} + 1)^2 s = w$, where w is a Gaussian white Lévy noise, is first approximated by s_{1000}. Then, it is viewed on different regular grids on $[0, 1]$.

E. Computational Efficiency

A crucial component of our approach is the computation of the generalized increment $\mathbf{u}$ in order to obtain the values of s_n on a grid. This provides a gain in numerical efficiency that can be felt even on moderately sized simulations. As can be seen in Figure 9, using a Green's function representation requires significantly more time than using an intermediate B-spline representation.

V. Conclusion

We have described a novel approach for generating sparse stochastic processes. Our method leverages the properties of B-splines to guarantee good numerical efficiency. A possible direction for future work is to provide theoretical guidance

on how one should choose the parameter n in terms of a prescribed tolerance on the approximation error.

ACKNOWLEDGMENT

The authors would like to thank Dr. Julien Fageot, Pakshal Bohra, and Thomas Debarre for enlightening discussions.

APPENDIX A
GENERALIZED STOCHASTIC PROCESSES

Generalized stochastic processes are random elements of $\mathcal{S}'(\mathbb{R})$ that can be fully specified by their characteristic functionals. Those are infinite-dimensional generalizations of the characteristic functions of real random variables.

Definition 2. *The characteristic functional of the generalized stochastic process s is the functional $\widehat{\mathcal{P}}_s : \mathcal{S}(\mathbb{R}) \to \mathbb{C}$ such that*

$$\widehat{\mathcal{P}}_s(\varphi) = \mathbb{E}[\mathrm{e}^{\mathrm{j}\langle \varphi, s\rangle}], \text{ for all } \varphi \in \mathcal{S}(\mathbb{R}).$$

It is a continuous, positive-definite functional and $\widehat{\mathcal{P}}_s(0) = 1$.

Just as in finite dimensions, $\widehat{\mathcal{P}}_s$ contains all the statistical information of s. In particular, for any test function $\varphi \in \mathcal{S}(\mathbb{R})$, the distribution of the real random variable $\langle \varphi, s\rangle$ is entirely determined by $\widehat{\mathcal{P}}_s$ as its probability density function p is given by

$$p(t) \propto \mathcal{F}^{-1}\left\{\mathbb{E}[\mathrm{e}^{\mathrm{j}\omega\langle \varphi, s\rangle}]\right\}(t) = \mathcal{F}^{-1}\left\{\widehat{\mathcal{P}}_s(\omega\varphi)\right\}(t),$$

where $\mathcal{F}^{-1}$ is the inverse Fourrier transform. The construction of such objects was initiated in [44]. Their use for modeling sparse signals was introduced in [10].

APPENDIX B
COMPUTING GREEN'S FUNCTIONS

Here, we describe a method to compute Green's functions of rational operators. We begin with the intermediate computation of the Green's function of $\mathrm{L} = (\mathrm{D} - \alpha \mathrm{I})^k$. We have that

$$\begin{aligned}\rho_{\alpha,k}(t) &= \mathcal{F}^{-1}\left\{\frac{1}{(\mathrm{j}\omega - \alpha)^k}\right\}(t) \\ &= \begin{cases} +(t)\frac{t^{k-1}}{(k-1)!}\mathrm{e}^{\alpha t}, & \Re(\alpha) \leq 0 \\ -\mathbb{1}_+(-t)\frac{t^{k-1}}{(k-1)!}\mathrm{e}^{\alpha t}, & \text{otherwise} \end{cases} \end{aligned} \tag{13}$$

is a Green's function of L.

Now, recall that rational operators are of the form $\mathrm{L} = P(\mathrm{D})Q(\mathrm{D})^{-1}$, where P and Q are polynomials. Taking $\{\alpha_1, ..., \alpha_m\}$ to be the roots of P with multiplicity $\{\gamma_1, ..., \gamma_m\}$, the inverse of the frequency response is given by

$$\frac{1}{\widehat{\mathrm{L}}(\omega)} = \frac{Q(\mathrm{j}\omega)}{\prod_{i=1}^m (\mathrm{j}\omega - \alpha_i)^{\gamma_i}}.$$

This inverse is known to admit a partial-fraction decomposition of the form

$$\frac{1}{\widehat{\mathrm{L}}(\omega)} = \sum_{i=1}^m \sum_{k=1}^{\gamma_i} \frac{c_{ik}}{(\mathrm{j}\omega - \alpha_i)^k}$$

for some constants $c_{ik} \in \mathbb{C}$. The corresponding Green's function is then given by:

$$\begin{aligned}\rho_{\mathrm{L}}(t) &= \mathcal{F}^{-1}\left\{\frac{1}{\widehat{\mathrm{L}}(\omega)}\right\}(t) \\ &= \sum_{i=1}^m \sum_{k=1}^{\gamma_i} c_{ik}\mathcal{F}^{-1}\left\{\frac{1}{(\mathrm{j}\omega - \alpha_i)^k}\right\}(t) \\ &= \sum_{i=1}^m \sum_{k=1}^{\gamma_i} c_{ik}\rho_{\alpha_i,k}(t)\end{aligned}$$

The Green's function of L is then be obtained by summing the Green's function of the partial fractions given in (13).

APPENDIX C

Proof of Proposition 1: Since w_n is a compound-Poisson innovation, $X_{\text{rect, n}}$ is a compound-Poisson random variable. It can be written

$$X_{\text{rect,n}} = \sum_{i=1}^N A_i$$

where N is a Poisson random variable with rate $\lambda = n$ and the A_i are independent identically distributed infinitely divisible random variables with Lévy exponent $\frac{1}{n}f$ independent from N. We have by independence of the (A_i), that

$$X_{\text{rect,n}} = \sum_{i=1}^n A_i =_d X_{\text{rect}}$$

because the characteristic function of $\sum_{i=1}^n A_i$ is $(\mathrm{e}^{\frac{1}{n}})^n = \mathrm{e}^f$. This directly implies that $X_{\text{rect,n}}$ and X_{rect} have the same moments.

REFERENCES

[1] R. Gray and L. Davisson, *An Introduction to Statistical Signal Processing*. Cambridge University Press, 2004.
[2] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete Cosine Transfom," *IEEE Transactions on Computers*, vol. 23, no. 1, pp. 90–93, Jan. 1974.
[3] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, Mar. 1960.
[4] D. Mumford and A. Desolneux, *Pattern Theory: The Stochastic Analysis of Real-World Signals*. A. K. Peters/CRC Press, Aug. 2010.
[5] A. Srivastava, A. B. Lee, E. P. Simoncelli, and S.-C. Zhu, "On advances in statistical modeling of natural images," *Journal of Mathematical Imaging and Vision*, vol. 18, no. 1, pp. 17–33, 2003.
[6] A. Amini, M. Unser, and F. Marvasti, "Compressibility of Deterministic and Random Infinite Sequences," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5193–5201, Nov. 2011.
[7] S. Mallat, *A Wavelet Tour of Signal Processing*. Elsevier, Sep. 1999.
[8] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[9] J.-L. Starck, F. Murtagh, and J. M. Fadili, *Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity*. Cambridge University Press, May 2010.
[10] M. Unser and P. D. Tafti, *An Introduction to Sparse Stochastic Processes*. Cambridge University Press, Aug. 2014.
[11] T. Kailath, "The innovations approach to detection and estimation theory," *Proceedings of the IEEE*, vol. 58, no. 5, pp. 680–695, 1970.
[12] K.-i. Sato, S. Ken-Iti, and A. Katok, *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Nov. 1999.
[13] M. Unser, P. D. Tafti, and Q. Sun, "A Unified Formulation of Gaussian Versus Sparse Stochastic Processes—Part I: Continuous-Domain Theory," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1945–1962, Mar. 2014.

[14] D. Mumford and B. Gidas, "Stochastic models for generic images," *Quarterly of Applied Mathematics*, vol. 59, no. 1, pp. 85–111, 2001.

[15] E. Bostan, U. S. Kamilov, M. Nilchian, and M. Unser, "Sparse Stochastic Processes and Discretization of Linear Inverse Problems," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2699–2710, Jul. 2013.

[16] M. A. Kutay, A. P. Petropulu, and C. W. Piccoli, "On modeling biomedical ultrasound RF echoes using a power-law shot-noise model," *IEEE Transactions on Iltrasonics, Ferroelectrics, and Frequency Control*, vol. 48, no. 4, pp. 953–968, 2001.

[17] S. M. Kogon and D. G. Manolakis, "Signal modeling with self-similar α-stable processes: The fractional Lévy stable motion model," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 1006–1010, 1996.

[18] J. R. Gallardo, D. Makrakis, and L. Orozco-Barbosa, "Use of α-stable self-similar stochastic processes for modeling traffic in broadband networks," *Performance Evaluation*, vol. 40, no. 1-3, pp. 71–98, 2000.

[19] N. Laskin, I. Lambadaris, F. C. Harmantzis, and M. Devetsikiotis, "Fractional lévy motion and its application to network traffic modeling," *Computer Networks*, vol. 40, no. 3, pp. 363–375, 2002.

[20] A. Amini, P. Thévenaz, J. Ward, and M. Unser, "On the linearity of Bayesian interpolators for non-Gaussian continuous-time AR(1) processes," *IEEE Transactions on Information Theory*, vol. 59, no. 8, pp. 5063–5074, August 2013.

[21] A. Amini, U. S. Kamilov, E. Bostan, and M. Unser, "Bayesian Estimation for Continuous-Time Sparse Stochastic Processes," *IEEE Transactions on Signal Processing*, vol. 61, no. 4, pp. 907–920, Feb. 2013.

[22] U. S. Kamilov, P. Pad, A. Amini, and M. Unser, "MMSE Estimation of Sparse Lévy Processes," *IEEE Transactions on Signal Processing*, vol. 61, no. 1, pp. 137–147, Jan. 2013.

[23] S. J. Godsill and G. Yang, "Bayesian inference for continuous-time ARMA models driven by non-Gaussian Lévy processes," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 5. IEEE, 2006, pp. V–V.

[24] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, 2nd ed. Berlin: Springer, 1989.

[25] S. Rubenthaler, "Numerical simulation of the solution of a stochastic differential equation driven by a Lévy process," *Stochastic Processes and their Applications*, vol. 103, no. 2, pp. 311–349, Feb. 2003.

[26] J. Fageot, V. Uhlmann, and M. Unser, "Gaussian and sparse processes are limits of generalized Poisson processes," *Applied and Computational Harmonic Analysis*, 2018.

[27] P. J. Brockwell, "Lévy-Driven CARMA Processes," *Annals of the Institute of Statistical Mathematics*, vol. 53, no. 1, pp. 113–124, Mar. 2001.

[28] P. E. Protter, *Stochastic Integration and Differential Equations*, 2nd ed., ser. Stochastic Modelling and Applied Probability. Berlin Heidelberg: Springer-Verlag, 2005.

[29] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.

[30] M. Unser and P. D. Tafti, "Stochastic models for sparse and piecewise-smooth signals," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 989–1006, 2010.

[31] M. Unser and T. Blu, "Cardinal exponential splines: Part I - theory and filtering algorithms," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1425–1438, Apr. 2005.

[32] M. Unser, "Cardinal exponential splines: Part II - think analog, act digital," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1439–1449, Apr. 2005.

[33] M. Unser, P. D. Tafti, A. Amini, and H. Kirshner, "A Unified Formulation of Gaussian *versus* Sparse Stochastic Processes—Part II: Discrete-Domain Theory," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 3036–3051, May 2014.

[34] J. Fageot and T. Humeau, "Unified View on Lévy White Noises: General Integrability Conditions and Applications to Linear SPDE," *arXiv:1708.02500 [math]*, Aug. 2017, arXiv: 1708.02500.

[35] A. Amini and M. Unser, "Sparsity and Infinite Divisibility," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2346–2358, Apr. 2014.

[36] J. Fageot, A. Amini, and M. Unser, "On the Continuity of Characteristic Functionals and Sparse Stochastic Modeling," *Journal of Fourier Analysis and Applications*, vol. 20, no. 6, pp. 1179–1211, Dec. 2014.

[37] R. C. Dalang, T. Humeau *et al.*, "Lévy processes and lévy white noise as tempered distributions," *The Annals of Probability*, vol. 45, no. 6B, pp. 4389–4418, 2017.

[38] L. Bondesson, "On Simulation from Infinitely Divisible Distributions," *Advances in Applied Probability*, vol. 14, no. 4, pp. 855–869, 1982.

[39] L. Devroye, "Nonuniform Random Variate Generation," in *Handbooks in Operations Research and Management Science*. Elsevier, Jan. 2006, vol. 13, pp. 83–121.

[40] C. De Boor, "On calculating with B-splines," *Journal of Approximation Theory*, vol. 6, no. 1, pp. 50–62, 1972.

[41] I. J. Schoenberg, "Contributions to the problem of approximation of equidistant data by analytic functions," in *IJ Schoenberg Selected Papers*. Springer, 1988, pp. 3–57.

[42] A. Ron, "Factorization theorems for univariate splines on regular grids," *Israel Journal of Mathematics*, vol. 70, no. 1, pp. 48–68, 1990.

[43] F. Massey Jr, "The Kolmogorov-Smirnov test for goodness of fit," *Journal of the American Statistical Association*, vol. 46, no. 253, pp. 68–78, 1951.

[44] I. M. Gel'fand and N. Y. Vilenkin, *Generalized Functions: Applications of Harmonic Analysis*. Academic Press, May 2014.