Simulation of Stochastic Volterra Equations Driven by Space–Time Lévy Noise

Bohan Chen, Carsten Chong and Claudia Klüppelberg

Abstract In this paper we investigate two numerical schemes for the simulation of stochastic Volterra equations driven by space–time Lévy noise of pure-jump type. The first one is based on truncating the small jumps of the noise, while the second one relies on series representation techniques for infinitely divisible random variables. Under reasonable assumptions, we prove for both methods L^p- and almost sure convergence of the approximations to the true solution of the Volterra equation. We give explicit convergence rates in terms of the Volterra kernel and the characteristics of the noise. A simulation study visualizes the most important path properties of the investigated processes.

Keywords Simulation of SPDEs · Simulation of stochastic Volterra equations · Space–time Lévy noise · Stochastic heat equation · Stochastic partial differential equation

1   Introduction

The aim of this paper is to investigate different simulation techniques for stochastic Volterra equations (SVEs) of the form

$$Y(t, x) = Y_0(t, x) + \int_0^t \int_{\mathbb{R}^d} G(t, x; s, y) \sigma(Y(s, y)) \, \Lambda(\mathrm{d}s, \mathrm{d}y) , \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d ,$$

(1)

B. Chen · C. Chong · C. Klüppelberg (✉)
Technische Universität München, BoltzmannstraßE 3, 85748 Garching, Germany
e-mail: cklu@ma.tum.de

C. Chong
e-mail: carsten.chong@tum.de

B. Chen
e-mail: bohan.chen@mytum.de

209

where G is a deterministic kernel function, σ a Lipschitz coefficient and Λ a Lévy basis on $\mathbb{R}_+ \times \mathbb{R}^d$ of pure-jump type with no Gaussian part. In the purely temporal case where no space is involved and the kernel G is sufficiently regular on the diagonal $\{(t; s) \in \mathbb{R}_+ \times \mathbb{R}_+ : t = s\}$, the existence and uniqueness of the solution Y to (1) are established for general semimartingale integrators in [14]. The space–time case (1) is treated in [5] for quite general Lévy bases. In particular, G is allowed to be singular on the diagonal, which typically happens in the context of stochastic partial differential equations (SPDEs) where G is the Green's function of the underlying differential operator. More details on the connection between SPDEs and the SVE (1) are presented in Sect. 2, or can be found in [2, 5, 20].

Since in most cases there exists no explicit solution formula for the SVE (1), it is a natural task to develop appropriate simulation algorithms. For SPDEs driven by Gaussian noise, research on this topic is rather far advanced, see e.g. [7, 9, 21]. However, for SPDEs driven by jump noises such as non-Gaussian Lévy bases, the related literature is considerably smaller, see [3] and the work of Hausenblas and coauthors [8, 10, 11]. The case $\sigma \equiv 1$ has been treated in [4]. The contribution of our paper can be summarized as follows:

- We propose and analyze two approximation schemes for (1), each of which replaces the original noise by a truncated noise that only has finitely many atoms on compact subsets of $\mathbb{R}_+ \times \mathbb{R}^d$. For the first scheme, we simply cut off all jumps whose size is smaller than a constant. For the second scheme, we use series representation techniques for the noise as in [17] such that the jumps to be dropped off are chosen randomly. Both methods have already been applied successfully to the simulation of Lévy processes, cf. [1, 18].
- In the case where G originates from an SPDE, the crucial difference of our numerical schemes to the Euler or finite element methods in the references mentioned before is that we do not simulate small space–time increments of the noise but successively the true jumps of the Lévy basis, which is an easier task given that one usually only knows the underlying Lévy measure. It is important to recognize that this is only possible because the noise Λ is of pure-jump type, and contains neither a Gaussian part nor a drift. We shall point out in Sect. 6 how to relax this assumption.

The remaining article is organized as follows: Sect. 2 gives the necessary background for the SVE (1). In particular, we present sufficient conditions for the existence and uniqueness of solutions, and address the connection between (1) and SPDEs. In Sect. 3 we construct approximations to the solution Y of (1) by truncating the small jumps of the Lévy basis. We prove in Theorem 1 their L^p-convergence, and in some cases also their almost sure (a.s.) convergence to the target process Y. In Sect. 4 we approximate the driving Lévy basis using series representation methods. This leads to an algorithm that produces approximations again converging in the L^p-sense, sometimes also almost surely, to Y, see Theorem 2. In both theorems, we find explicit L^p-convergence rates that only depend on the kernel G and the characteristics of Λ. Section 5 presents a simulation study for the stochastic heat equation

which highlights the typical path behaviour of stochastic Volterra equations. The final Sect. 6 compares the two simulation algorithms developed in this paper and discusses some further directions of the topic.

2 Preliminaries

We start with a summary of notations that will be employed in this paper.

$\mathbb{R}_+$	The set $[0, \infty)$ of *positive* real numbers
$\mathbb{N}$	The natural numbers $\{1, 2, \ldots\}$
$\mathbb{B}$	A stochastic basis $(\Omega, \mathscr{F}, \mathbb{F} = (\mathscr{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})$ satisfying the usual hypotheses of completeness and right-continuity
$\bar{\Omega}, \tilde{\Omega}$	$\bar{\Omega} := \Omega \times \mathbb{R}_+$ and $\tilde{\Omega} := \Omega \times \mathbb{R}_+ \times \mathbb{R}^d$ where $d \in \mathbb{N}$
$\mathscr{B}(\mathbb{R}^d)$	The Borel σ-field on $\mathbb{R}^d$
$\tilde{\mathscr{B}}_{\mathrm{b}}$	The collection of all bounded Borel sets of $\mathbb{R}_+ \times \mathbb{R}^d$
$\mathscr{P}$	The predictable σ-field on $\mathbb{B}$ *or* the collection of all predictable processes $\bar{\Omega} \to \mathbb{R}$
$\tilde{\mathscr{P}}$	The product $\mathscr{P} \otimes \mathscr{B}(\mathbb{R}^d)$ *or* the collection all $\mathscr{P} \otimes \mathscr{B}(\mathbb{R}^d)$-measurable processes $\tilde{\Omega} \to \mathbb{R}$
$\tilde{\mathscr{P}}_{\mathrm{b}}$	The collection of sets in $\tilde{\mathscr{P}}$ which are a subset of $\Omega \times [0, k] \times [-k, k]^d$ for some $k \in \mathbb{N}$
δ_a	The Dirac measure at a
p^*	$p \vee 1$
L^p	The space $L^p(\Omega, \mathscr{F}, \mathbb{P})$, $p \in (0, \infty)$ with $\|X\|_{L^p} := \mathbb{E}[|X|^p]^{1/p^*}$
L^0	The space $L^0(\Omega, \mathscr{F}, \mathbb{P})$ of all random variables on $\mathbb{B}$ endowed with the topology of convergence in probability
B^p_{loc}	The set of all $Y \in \tilde{\mathscr{P}}$ for which $\|Y(t, x)\|_{L^p}$ is uniformly bounded on $[0, T] \times \mathbb{R}^d$ for all $T \in \mathbb{R}_+$ $(p \in (0, \infty))$
A^{c}	The complement of A within the superset it belongs to (which will be clear from the context)
$A - B$	$\{x - y \colon x \in A, y \in B\}$
$-A$	$\{-x \colon x \in A\}$
Leb	The Lebesgue measure on $\mathbb{R}^d$ (d should be clear from the context)
$\| \cdot \|$	The Euclidean norm on $\mathbb{R}^d$
$C, C(T)$	Two generic constants in $\mathbb{R}_+$, one dependent and one independent of T, whose values are irrelevant in this paper and may therefore change from one place to the other

We suppose that the stochastic basis $\mathbb{B}$ supports a *Lévy basis*, that is, a mapping $\Lambda \colon \tilde{\mathscr{P}}_{\mathrm{b}} \to L^0$ with the following properties:

- $\Lambda(\emptyset) = 0$ a.s.
- For all pairwise disjoint sets $(A_i)_{i \in \mathbb{N}} \subset \tilde{\mathscr{P}}_{\mathrm{b}}$ with $\bigcup_{i=1}^{\infty} A_i \in \tilde{\mathscr{P}}_{\mathrm{b}}$ we have

$$\Lambda\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Lambda(A_i) \quad \text{in } L^0 . \tag{2}$$

- $(\Lambda(\Omega \times B_i))_{i \in \mathbb{N}}$ is a sequence of independent random variables if $(B_i)_{i \in \mathbb{N}}$ are pairwise disjoint sets in $\tilde{\mathscr{B}}_{\mathrm{b}}$.
- For every $B \in \mathscr{B}_{\mathrm{b}}$, $\Lambda(\Omega \times B)$ has an infinitely divisible distribution.
- $\Lambda(A)$ is $\mathscr{F}_t$-measurable when $A \in \tilde{\mathscr{P}}_{\mathrm{b}}$ and $A \subseteq \Omega \times [0, t] \times \mathbb{R}^d$ for $t \in \mathbb{R}_+$.
- For every $t \in \mathbb{R}_+$, $A \in \tilde{\mathscr{P}}_{\mathrm{b}}$ and $\Omega_0 \in \mathscr{F}_t$ we have a.s.

$$\Lambda(A \cap (\Omega_0 \times (t, \infty) \times \mathbb{R}^d)) = \mathbb{1}_{\Omega_0} \Lambda(A \cap (\Omega \times (t, \infty) \times \mathbb{R}^d)) .$$

Just as Lévy processes are semimartingales and thus allow for an Itô integration theory, Lévy bases belong to the class of L^0-valued σ-finite random measures. Therefore, it is possible to define the stochastic integral

$$\int_{\mathbb{R}_+ \times \mathbb{R}^d} H(s, y) \, \Lambda(\mathrm{d}s, \mathrm{d}y)$$

for $H \in \tilde{\mathscr{P}}$ that are *integrable* with respect to Λ, see [6] for the details.

Similarly to Lévy processes, there exist two notions of characteristics for Lévy bases: one going back to [15, Proposition 2.1] that is based on the Lévy-Khintchine formula and is independent of $\mathbb{F}$, and a filtration-based one that is useful for stochastic analysis [6, Theorem 3.2]. For the whole paper, we will assume that both notions coincide such that Λ has a canonical decomposition under the filtration $\mathbb{F}$ of the form

$$\Lambda(\mathrm{d}t, \mathrm{d}x) = B(\mathrm{d}t, \mathrm{d}x) + \Lambda^{\mathrm{c}}(\mathrm{d}t, \mathrm{d}x) + \int_{\mathbb{R}} z \mathbb{1}_{\{|z| \leq 1\}} (\mu - \nu)(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z)$$

$$+ \int_{\mathbb{R}} z \mathbb{1}_{\{|z| > 1\}} \mu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) ,$$

where B is a σ-finite signed Borel measure on $\mathbb{R}_+ \times \mathbb{R}^d$, Λ^{c} a Lévy basis such that $\Lambda(\Omega \times B)$ is normally distributed with mean 0 and variance $C(B)$ for all $B \in \tilde{\mathscr{B}}_{\mathrm{b}}$, and μ a Poisson measure on $\mathbb{R}_+ \times \mathbb{R}^d$ relative to $\mathbb{F}$ with intensity measure ν (cf. [12, Definition II.1.20]). There exists also a σ-finite Borel measure λ on $\mathbb{R}_+ \times \mathbb{R}^d$ such that

$$B(\mathrm{d}t, \mathrm{d}x) = b(t, x) \, \lambda(\mathrm{d}t, \mathrm{d}x) , \quad C(\mathrm{d}t, \mathrm{d}x) = c(t, x) \, \lambda(\mathrm{d}t, \mathrm{d}x) \quad \text{and}$$

$$\nu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) = \pi(t, x, \mathrm{d}z) \lambda(\mathrm{d}t, \mathrm{d}x) \tag{3}$$

with two functions $b\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ and $c\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}_+$ as well as a transition kernel π from $(\mathbb{R}_+ \times \mathbb{R}^d, \mathscr{B}(\mathbb{R}_+ \times \mathbb{R}^d))$ to $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ such that $\pi(t, x, \cdot)$ is a Lévy measure for each $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$.

We have already mentioned in the introduction that we will assume

$$C = 0 \tag{4}$$

throughout the paper. For simplicity we will also make two further assumptions: first, that there exist $b \in \mathbb{R}$ and a Lévy measure π such that for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ we have

$$b(t, x) = b \,, \quad \pi(t, x, \cdot) = \pi \quad \text{and} \quad \lambda(\mathrm{d}t, \mathrm{d}x) = \mathrm{d}(t, x) \,; \tag{5}$$

second, that

$$\Lambda \in \mathscr{S} \cup \mathscr{V}_0 \,, \tag{6}$$

where $\mathscr{S}$ is the collection of all *symmetric* Lévy bases and $\mathscr{V}_0$ is the class of Lévy bases with *locally finite variation and no drift*, defined by the property that

$$\int_{\mathbb{R}} |z| \mathbb{1}_{\{|z| \le 1\}} \pi(\mathrm{d}z) < \infty \,, \quad \text{and} \quad b_0 := b - \int_{\mathbb{R}} z \mathbb{1}_{\{|z| \le 1\}} \pi(\mathrm{d}z) = 0 \,.$$

Furthermore, if Λ has a finite first moment, that is,

$$\int_{\mathbb{R}} |z| \mathbb{1}_{\{|z| > 1\}} \pi(\mathrm{d}z) < \infty \,, \tag{7}$$

we define

$$B_1(\mathrm{d}t, \mathrm{d}x) := b_1 \, \mathrm{d}(t, x) \,, \quad b_1 := b + \int_{\mathbb{R}} z \mathbb{1}_{\{|z| > 1\}} \pi(\mathrm{d}z) \,,$$

$$M(\mathrm{d}t, \mathrm{d}x) := \Lambda(\mathrm{d}t, \mathrm{d}x) - B_1(\mathrm{d}t, \mathrm{d}x) = \int_{\mathbb{R}} z \, (\mu - \nu)(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) \,.$$

Next, let us summarize the most important facts regarding the SVE (1). All details that are not explained can be found in [5]. First, many SPDEs of evolution type driven by Lévy noise can be written in terms of (1), where G is the Green's function of the corresponding differential operator. Most prominently, taking G being the heat kernel in $\mathbb{R}^d$, (1) is the so-called *mild formulation* of the stochastic heat equation (with constant coefficients and multiplicative noise). Typically for parabolic equations, the heat kernel is very smooth in general but explodes on the diagonal $t = s$ and $x = y$. In fact, it is only p-fold integrable on $[0, T] \times \mathbb{R}^d$ for $p < 1 + 2/d$. In particular, as soon as $d \ge 2$, it is not square-integrable, and as a consequence, no solution to the stochastic heat equation in the form (1) will exist for Lévy noises with non-zero Gaussian component. This is another reason for including assumption (4) in this paper.

Second, let us address the existence and uniqueness problem for (1). By a *solution* to this equation we mean a predictable process $Y \in \tilde{\mathscr{P}}$ such that for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$, the stochastic integral on the right-hand side of (1) is well defined and the equation itself for each $(t, x) \in [0, T] \times \mathbb{R}^d$ holds a.s. We identify two solutions as soon as they are modifications of each other. Given a number $p \in (0, 2]$, the following conditions A1–A6 guarantee a unique solution to (1) in B_{loc}^p by [5, Theorem 3.1]:

A1. $Y_0 \in B_{\mathrm{loc}}^p$ is independent of Λ.

A2. $\sigma \colon \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous, that is, there exists $C \in \mathbb{R}_+$ such that

$$|\sigma(x) - \sigma(y)| \le C|x - y| , \quad x, y \in \mathbb{R} . \tag{8}$$

A3. $G \colon (\mathbb{R}_+ \times \mathbb{R}^d)^2 \to \mathbb{R}$ is a measurable function with $G(t, \cdot; s, \cdot) \equiv 0$ for $s > t$.

A4. Λ satisfies (3)–(6) and

$$\int_{\mathbb{R}} |z|^p \, \pi(\mathrm{d}z) < \infty . \tag{9}$$

A5.   If we define for $(t, x), (s, y) \in \mathbb{R}_+ \times \mathbb{R}^d$

$$\tilde{G}(t, x; s, y) := |G(t, x; s, y)| \mathbb{1}_{\{p>1, \Lambda \notin \mathscr{S}\}} + |G(t, x; s, y)|^p , \tag{10}$$

then we have for all $T \in \mathbb{R}_+$

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \int_0^T \int_{\mathbb{R}^d} \tilde{G}(t, x; s, y) \, \mathrm{d}(s, y) < \infty . \tag{11}$$

A6.   For all $\varepsilon > 0$ and $T \in \mathbb{R}_+$ there exist $k \in \mathbb{N}$ and a partition $0 = t_0 < \cdots < t_k = T$ such that

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \sup_{i=1,\ldots,k} \int_{t_{i-1}}^{t_i} \int_{\mathbb{R}^d} \tilde{G}(t, x; s, y) \, \mathrm{d}(s, y) < \varepsilon . \tag{12}$$

Apart from A1–A6, we will add another assumption in this paper:

A7.   There exists a sequence $(U^N)_{N \in \mathbb{N}}$ of compact sets increasing to $\mathbb{R}^d$ such that for all $T \in \mathbb{R}_+$ and compact sets $K \subseteq \mathbb{R}^d$ we have, as $N \to \infty$,

$$r_1^N(T, K) := \sup_{(t,x) \in [0,T] \times K} \left(\int_0^t \int_{(U^N)^c} |G(t, x; s, y)| \mathbb{1}_{\{p>1, \Lambda \notin \mathscr{S}\}} \, \mathrm{d}(s, y) \right.$$
$$\left. + \left(\int_0^t \int_{(U^N)^c} |G(t, x; s, y)|^p \, \mathrm{d}(s, y) \right)^{1/p^*} \right) \to 0 . \tag{13}$$

Conditions A6 and A7 are automatically satisfied if $|G(t, x; s, y)| \le g(t - s, x - y)$ for some measurable function g and A5 holds with G replaced by g. For A6 see [5, Remark 3.3(3)]; for A7 choose $U^N := \{x \in \mathbb{R}^d : \|x\| \le N\}$ such that for $p > 1$

$$\sup_{(t,x)\in[0,T]\times K} \int_0^t \int_{(U^N)^c} |G(t,x;s,y)|^p \, d(s,y)$$

$$\leq \sup_{(t,x)\in[0,T]\times K} \int_0^t \int_{(U^N)^c} g^p(t-s,x-y) \, d(s,y)$$

$$= \sup_{x\in K} \int_0^T \int_{x-(U^N)^c} g^p(s,y) \, d(s,y) \leq \int_0^T \int_{K-(U^N)^c} g^p(s,y) \, d(s,y)$$

$$\to 0 \quad \text{as} \quad N \to \infty$$

by the fact that $K - (U^N)^c \downarrow 0$. A similar calculation applies to the case $p \in (0, 1]$ and the first term in $r_1^N(T, K)$.

Example 1 We conclude this section with the stochastic heat equation in $\mathbb{R}^d$, whose mild formulation is given by the SVE (1) with

$$G(t,x;s,y) = g(t-s,x-y) \,, \quad g(t,x) = \frac{\exp(-\|x\|^2/(4t))}{(4\pi t)^{d/2}} \mathbb{1}_{[0,t)}(s) \qquad (14)$$

for $(t, x), (s, y) \in \mathbb{R}_+ \times \mathbb{R}^d$. We assume that Y_0 and σ satisfy conditions A1 and A2, respectively. Furthermore, we suppose that (3)–(6) are valid, and that (9) holds with some $p \in (0, 1 + 2/d)$. It is straightforward to show that then A3–A6 are satisfied with the same p. Let us estimate the rate $r_1^N(T, K)$ for $T \in \mathbb{R}_+$, $K := \{\|x\| \leq R\}$ with $R \in \mathbb{N}$, and $U^N := \{\|x\| \leq N\}$. We first consider the case $p \leq 1$ or $\Lambda \in \mathscr{S}$. Since $K - (U^N)^c = (U^{N-R})^c$ for $N \geq R$, the calculations after A7 yield ($\Gamma(\cdot, \cdot)$ denotes the upper incomplete gamma function and $p(d) := 1 + (1 - p)d/2$)

$$(r_1^N(T, K))^{p^*} \leq \int_0^T \int_{(U^{N-R})^c} g^p(t,x) \, d(t,x) = \int_0^T \int_{N-R}^\infty \frac{\exp(-pr^2/(4t))}{(4\pi t)^{pd/2}} r^{d-1} \, dr \, dt$$

$$= C \int_0^T t^{p(d)-1} \Gamma\left(\frac{d}{2}, \frac{p(N-R)^2}{4t}\right) dt$$

$$= C(T)\left((p(d))^{-1}\Gamma\left(\frac{d}{2}, \frac{p(N-R)^2}{4T}\right) - \left(\frac{p(N-R)^2}{4T}\right)^{p(d)}\right.$$

$$\left. \times \Gamma\left(\frac{d}{2} - p(d), \frac{p(N-R)^2}{4T}\right)\right)$$

$$\leq C(T) \exp\left(-\frac{p(N-R)^2}{4T}\right)(N-R)^{d-2} \,, \qquad (15)$$

which tends to 0 exponentially fast as $N \to \infty$. If $p > 1$ and $\Lambda \notin \mathscr{S}$, it follows from formula (13) that we need an extra summand for $r_1^N(T, K)$, namely (15) with $p = 1$.

3   Truncation of Small Jumps

In this section we approximate Eq. (1) by cutting off the small jumps of Λ. To this end, we first define for each $N \in \mathbb{N}$

$$G^N(t, x; s, y) := G(t, x; s, y)\mathbb{1}_{U^N}(y) , \quad (t, x), (s, y) \in \mathbb{R}_+ \times \mathbb{R}^d , \qquad (16)$$

where the meaning of the sets U^N is explained in A7. Furthermore, we introduce

$$r_2^N := \left(\int_{[-\varepsilon^N, \varepsilon^N]} |z|^p \, \pi(\mathrm{d}z) \right)^{1/p^*} , \quad r_3^N := \left| \int_{[-\varepsilon^N, \varepsilon^N]} z \mathbb{1}_{\{p>1, \Lambda \notin \mathscr{S}\}} \, \pi(\mathrm{d}z) \right| , \quad (17)$$

where $(\varepsilon^N)_{N \in \mathbb{N}} \subseteq (0, 1)$ satisfies $\varepsilon^N \to 0$ as $N \to \infty$. Condition A4 implies that $r_2^N, r_3^N \to 0$ as $N \to \infty$. Next, defining truncations of the Lévy basis Λ by

$$\Lambda^N(\mathrm{d}t, \mathrm{d}x) := \int_{[-\varepsilon^N, \varepsilon^N]^c} z \, \mu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) , \qquad (18)$$

our approximation scheme for the solution Y to (1) is given as:

$$Y^N(t, x) := Y_0(t, x) + \int_0^t \int_{\mathbb{R}^d} G^N(t, x; s, y)\sigma(Y^N(s, y)) \, \Lambda^N(\mathrm{d}s, \mathrm{d}y) \qquad (19)$$

for $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Indeed, Y^N can be simulated exactly because for all $T \in \mathbb{R}_+$ the truncation Λ^N only has a finite (random) number $R^N(T)$ of jumps on $[0, T] \times U^N$, say at the space–time locations (τ_i^N, ξ_i^N) with sizes J_i^N. This implies that we have the following alternative representation of $Y^N(t, x)$ for $(t, x) \in [0, T] \times \mathbb{R}^d$:

$$Y^N(t, x) = Y_0(t, x) + \sum_{i=1}^{R^N(T)} G(t, x; \tau_i^N, \xi_i^N)\sigma(Y^N(\tau_i^N, \xi_i^N))J_i^N \mathbb{1}_{\{\tau_i^N < t\}} . \qquad (20)$$

What remains to do is to simulate $Y^N(\tau_i^N, \xi_i^N)$, $i = 1, \ldots, R^N(T)$, iteratively, from which the values $Y(t, x)$ for all other $(t, x) \in [0, T] \times \mathbb{R}^d$ can be computed.

The following algorithm summarizes up the simulation procedure:

Algorithm 1 Consider a finite grid $\mathscr{G} \subseteq [0, T] \times \mathbb{R}^d$. For each N proceed as follows:

1. Draw a Poisson random variable $R^N(T)$ with parameter

$$R^N(T) := \int_0^T \int_{U^N} \int_{\mathbb{R}} \mathbb{1}_{\{z \in [-\varepsilon^N, \varepsilon^N]^c\}} \, \nu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) = T \mathrm{Leb}(U^N)\pi\left([-\varepsilon^N, \varepsilon^N]^c\right) .$$

2. For $i = 1, \ldots, R^N(T)$:

    a. Draw a pair (τ_i^N, ξ_i^N) with uniform distribution from $[0, T] \times U^N$.

    b. Draw J_i^N from $[-\varepsilon^N, \varepsilon^N]^c$ with distribution $\pi / \pi \left([-\varepsilon^N, \varepsilon^N]^c \right)$.

    c. Rename $(\tau_i^N, \xi_i^N, J_i^N : i = 1, \ldots, R^N(T))$ by sorting $(\tau_i^N : i = 1, \ldots, R^N(T))$ in increasing order.

3. For each $i = 1, \ldots, R^N(T)$ and $(t, x) \in \mathscr{G}$ simulate $Y_0(\tau_i^N, \xi_i^N)$ and $Y_0(t, x)$.

4. For each $i = 1, \ldots, R^N(T)$ set

$$Y^N(\tau_i^N, \xi_i^N) := Y_0(\tau_i^N, \xi_i^N) + \sum_{j=1}^{i-1} G(\tau_i^N, \xi_i^N; \tau_j^N, \xi_j^N) \sigma(Y^N(\tau_j^N, \xi_j^N)) J_j^N .$$

5. For each $(t, x) \in \mathscr{G}$ define $Y^N(t, x)$ via (20).

The next theorem determines the convergence behaviour of (19) to (1).

Theorem 1 *Grant assumptions A1–A7 under which the SVE (1) has a unique solution in B_{loc}^p. Then Y^N as defined in (19) belongs to B_{loc}^p for all $N \in \mathbb{N}$, and for all $T \in \mathbb{R}_+$ and compact sets $K \subseteq \mathbb{R}^d$ there exists a constant $C(T) \in \mathbb{R}_+$ independent of N and K such that*

$$\sup_{(t,x) \in [0,T] \times K} \| Y(t, x) - Y^N(t, x) \|_{L^p} \leq C(T)(r_1^N(T, K) + r_2^N + r_3^N) . \quad (21)$$

Furthermore, if $\sum_{N=1}^\infty (r_1^N(T, K) + r_2^N + r_3^N)^{p^*} < \infty$ is fulfilled, then we also have for all $(t, x) \in [0, T] \times K$ that $Y^N(t, x) \to Y(t, x)$ a.s. as $N \to \infty$.

Proof It is obvious that $|G^N| \leq |G|$ pointwise and that we have $\nu^N \leq \nu$ for the third characteristic ν^N of Λ^N. Thus, A1–A6 are still satisfied when G and ν are replaced by G^N and ν^N (if $\Lambda \in \mathscr{S}$, also $\Lambda^N \in \mathscr{S}$). So Y^N as a solution to (1) with G^N and Λ^N instead of G and Λ belongs to B_{loc}^p as well. Moreover, for all $T \in \mathbb{R}_+$ there exists $C(T) \in \mathbb{R}_+$ independent of $N \in \mathbb{N}$ such that

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \| Y^N(t, x) \|_{L^p} \leq C(T) , \quad N \in \mathbb{N} . \quad (22)$$

We only sketch the proof for this statement. In fact, using [5, Lemma 6.1(1)] it can be shown that the left-hand side of (22) satisfies an inequality of the same type as in Lemma 6.4(3) of the same paper. In particular, it is bounded by a constant $C^N(T)$ that depends on N only through $|G^N|$ and ν^N, and that this constant is only increased if we replace $|G^N|$ and ν^N by the larger $|G|$ and ν. In this way, we obtain an upper bound $C(T)$ that does not depend on N.

Next, we prove the convergence of Y^N to Y as stated in (21). We have

$$
\begin{aligned}
Y(t, x) - Y^N(t, x) &= \int_0^t \int_{\mathbb{R}^d} [G(t, x; s, y) - G^N(t, x; s, y)] \sigma(Y(s, y)) \, \Lambda(\mathrm{d}s, \mathrm{d}y) \\
&\quad + \int_0^t \int_{\mathbb{R}^d} G^N(t, x; s, y) [\sigma(Y(s, y)) - \sigma(Y^N(s, y))] \, \Lambda(\mathrm{d}s, \mathrm{d}y) \\
&\quad + \int_0^t \int_{\mathbb{R}^d} G^N(t, x; s, y) \sigma(Y^N(s, y)) \, (\Lambda - \Lambda^N)(\mathrm{d}s, \mathrm{d}y) \\
&=: I_1^N(t, x) + I_2^N(t, x) + I_3^N(t, x), \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d . \quad (23)
\end{aligned}
$$

If $p > 1$, we have by (8), Hölder's inequality and the Burkholder-Davis-Gundy-inequality

$$
\begin{aligned}
\|I_2^N(t, x)\|_{L^p} &\leq \left\| \int_0^t \int_{\mathbb{R}^d} G^N(t, x; s, y) [\sigma(Y(s, y)) - \sigma(Y^N(s, y))] B_1(\mathrm{d}s, \mathrm{d}y) \right\|_{L^p} \\
&\quad + \left\| \int_0^t \int_{\mathbb{R}^d} G^N(t, x; s, y) [\sigma(Y(s, y)) - \sigma(Y^N(s, y))] M(\mathrm{d}s, \mathrm{d}y) \right\|_{L^p} \\
&\leq C \Bigg(\left(\int_0^t \int_{\mathbb{R}^d} |G(t, x; s, y)| \, |B_1|(\mathrm{d}s, \mathrm{d}y) \right)^{p-1} \\
&\quad \times \int_0^t \int_{\mathbb{R}^d} |G(t, x; s, y)| \|Y(s, y) - Y^N(s, y)\|_{L^p}^p \, |B_1|(\mathrm{d}s, \mathrm{d}y) \Bigg)^{1/p} \\
&\quad + C \left(\int_0^t \int_{\mathbb{R}^d} |G(t, x; s, y)|^p \|Y^N(s, y) - Y^N(s, y)\|_{L^p}^p \, \mathrm{d}(s, y) \right)^{1/p} . \\
&\hspace{12cm} (24)
\end{aligned}
$$

If $p \in (0, 1]$, we have $\Lambda \in \mathscr{V}_0$ by (6) and (9), and thus Jensen's inequality gives

$$
\begin{aligned}
\|I_2^N(t, x)\|_{L^p} &= \mathbb{E} \left[\left(\int_0^t \int_{\mathbb{R}^d} G^N(t, x; s, y) [\sigma(Y(s, y)) - \sigma(Y^N(s, y))] z \, \mu(\mathrm{d}s, \mathrm{d}y, \mathrm{d}z) \right)^p \right] \\
&\leq \mathbb{E} \left[\int_0^t \int_{\mathbb{R}^d} |G^N(t, x; s, y) [\sigma(Y(s, y)) - \sigma(Y^N(s, y))] z|^p \, \nu(\mathrm{d}s, \mathrm{d}y, \mathrm{d}z) \right] \\
&\leq C \int_0^t \int_{\mathbb{R}^d} |G(t, x; s, y)|^p \|Y(s, y) - Y^N(s, y)\|_{L^p} \, \mathrm{d}(s, y) . \quad (25)
\end{aligned}
$$

Inserting (24) and (25) back into (23), we have for $w^N(t, x) := \|Y(t, x) - Y^N(t, x)\|_{L^p}$

$$
\begin{aligned}
w^N(t, x) &\leq C(T) \Bigg(\left(\int_0^t \int_{\mathbb{R}^d} |G(t, x; s, y)| \mathbb{1}_{\{p>1, \Lambda \notin \mathscr{S}\}} (w^N(s, y))^p \, \mathrm{d}(s, y) \right)^{1/p} \\
&\quad + \left(\int_0^t \int_{\mathbb{R}^d} |G(t, x; s, y)|^p (w^N(s, y))^{p^*} \, \mathrm{d}(s, y) \right)^{1/p^*} \Bigg) \\
&\quad + \|I_1^N(t, x) + I_3^N(t, x)\|_{L^p} , \quad (t, x) \in [0, T] \times \mathbb{R}^d .
\end{aligned}
$$

By a Gronwall-type estimate, which is possible because of A5 (see the proof of [5, Theorem 4.7(3)] for an elaboration of an argument of this type), we conclude

$$\sup_{(t,x)\in[0,T]\times K} w^N(t,x) \leq C(T) \sup_{(t,x)\in[0,T]\times K} \|I_1^N(t,x) + I_3^N(t,x)\|_{L^p} .$$

where $C(T)$ does not depend on K because of (11). For $I_1^N(t,x)$ we have for $p > 1$

$$
\begin{aligned}
\|I_1^N(t,x)\|_{L^p} &\leq \left\| \int_0^t \int_{\mathbb{R}^d} [G(t,x;s,y) - G^N(t,x;s,y)]\sigma(Y(s,y)) B_1(ds,dy) \right\|_{L^p} \\
&+ \left\| \int_0^t \int_{\mathbb{R}^d} [G(t,x;s,y) - G^N(t,x;s,y)]\sigma(Y(s,y)) M(ds,dy) \right\|_{L^p} \\
&\leq C\left(1 + \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \|Y(t,x)\|_{L^p}\right)\left(\int_0^t \int_{(U^N)^c} |G(t,x;s,y)| \, |B_1|(ds,dy) \right. \\
&+ \left. \left(\int_0^t \int_{(U^N)^c} |G(t,x;s,y)z|^p \, v(ds,dy,dz) \right)^{1/p} \right) \\
&\leq C(T) r_1^N(T,K) ,
\end{aligned}
\tag{26}
$$

uniformly in $(t,x) \in [0,T] \times K$. In similar fashion one proves the estimate (26) for $p \in (0,1]$, perhaps with a different $C(T)$. Next, when $p > 1$, (22) implies

$$
\begin{aligned}
\|I_3^N(t,x)\|_{L^p} &= \left\| \int_0^t \int_{\mathbb{R}^d} \int_{[-\varepsilon^N,\varepsilon^N]} G^N(t,x;s,y)\sigma(Y^N(s,y))z\,(\mu - v)(ds,dy,dz) \right\|_{L^p} \\
&+ \left\| \int_0^t \int_{\mathbb{R}^d} \int_{[-\varepsilon^N,\varepsilon^N]} G^N(t,x;s,y)\sigma(Y^N(s,y))z\mathbb{1}_{\{\Lambda\notin\mathscr{S}\}} v(ds,dy,dz) \right\|_{L^p} \\
&\leq C(T)\left(\left(\int_0^t \int_{\mathbb{R}^d} \int_{[-\varepsilon^N,\varepsilon^N]} |G(t,x;s,y)z|^p \, \pi(dz)\,d(s,y) \right)^{1/p} \right. \\
&+ \left. \left| \int_{[-\varepsilon^N,\varepsilon^N]} z\mathbb{1}_{\{\Lambda\notin\mathscr{S}\}} \pi(dz) \right| \int_0^t \int_{\mathbb{R}^d} |G(t,x;s,y)|\mathbb{1}_{\{\Lambda\notin\mathscr{S}\}} d(s,y) \right) \\
&\leq C(T)(r_2^N + r_3^N) .
\end{aligned}
$$

The case $p \in (0,1]$ can be treated similarly, cf. the estimation of $I_2^N(t,x)$ above.

It remains to prove that for each $(t,x) \in [0,T] \times K$ the convergence of $Y^N(t,x)$ to $Y(t,x)$ is almost sure when $r_1^N(T,K)$, r_2^N and r_3^N are p^*-summable. To this end, choose an arbitrary sequence $(a_N)_{N\in\mathbb{N}} \subseteq (0,1)$ converging to 0 such that

$$\sum_{N=1}^{\infty} A_N < \infty \quad \text{with} \quad A_N := \frac{(r_1^N(T,K) + r_2^N + r_3^N)^{p^*}}{a_N^p} .$$

Such a sequence always exists, see [13, Theorem 175.4], for example. So by (21) and Chebyshev's inequality we derive

$$\mathbb{P}\left[|Y(t, x) - Y^N(t, x)| \geq a_N\right] \leq \frac{\|Y(t, x) - Y^N(t, x)\|_{L^p}^{p^*}}{a_N^p} \leq C(T)A_N \ .$$

Our assertion now follows from the Borel-Cantelli lemma. □

Example 2 The rates r_2^N and r_3^N from (17) only depend on the underlying Lévy measure π. Let $p, q \in (0, 2]$ with $q < p$, and assume that $\int_{[-1,1]} |z|^q \, \pi(\mathrm{d}z) < \infty$. If $\Lambda \in \mathcal{V}_0$, assume that $q < 1$. Then

$$r_2^N = \left(\int_{[-\varepsilon^N, \varepsilon^N]} |z|^p \, \pi(\mathrm{d}z)\right)^{1/p^*} \leq \left(\int_{[-1,1]} |z|^q \, \nu(\mathrm{d}z)(\varepsilon^N)^{p-q}\right)^{1/p^*}$$
$$= \mathcal{O}\left((\varepsilon^N)^{(p-q)/p^*}\right) \ ,$$
$$r_3^N = \left|\int_{[-\varepsilon^N, \varepsilon^N]} z \mathbb{1}_{\{p>1, \Lambda \notin \mathscr{S}\}} \, \pi(\mathrm{d}z)\right| \leq \mathcal{O}\left((\varepsilon^N)^{1-q} \mathbb{1}_{\{p>1, \Lambda \notin \mathscr{S}\}}\right) \ .$$

For instance, if $\varepsilon^N = 1/N^k$, then the sequence $(r_2^N)^{p^*} = \mathcal{O}(N^{-k(p-q)})$ is summable for all $k > (p-q)^{-1}$. So in order to obtain a.s. convergence of $Y^N(t, x) \to Y(t, x)$, a sufficient condition is to choose the truncation rates ε^N small enough. Similar conclusions are valid for the other two rates $r_1^N(T, K)$ and r_3^N.

4 Truncation via Series Representations

Another approach to the simulation of (1) uses series representations for the Lévy basis. The idea, going back to [16, 17] and already applied to the simulation of Lévy processes [18], is to choose the jumps to be simulated in a random order. Instead of selecting the big jumps first and the smaller jumps later as in Sect. 3, we only choose the big jumps first more *likely*. The details are as follows: we fix a finite time horizon $T \in \mathbb{R}_+$ and, recalling A7, a partition $(Q^i)_{i \in \mathbb{N}}$ of $\mathbb{R}^d$ into pairwise disjoint compact sets such that $U^N = \bigcup_{i=1}^N Q^i$. We now assume that the jump measure μ of Λ on the strip $[0, T] \times \mathbb{R}^d \times \mathbb{R}$ can be represented in the form

$$\mu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) = \sum_{i=1}^{\infty} \mu_i(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) \ ,$$
$$\mu_i(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) = \sum_{j=1}^{\infty} \delta_{(\tau_j^i, \xi_j^i, H(\Gamma_j^i, V_j^i))}(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) \quad \text{a.s.} \ , \tag{27}$$

where $H \colon (0, \infty) \times \mathbb{R} \to \mathbb{R}$ is a measurable function, satisfying $H(\cdot, v) = -H(\cdot, -v)$ for all $v \in \mathbb{R}$ when $\Lambda \in \mathscr{S}$, and the random variables involved have the following properties for each $i \in \mathbb{N}$:

- $(\tau_j^i\colon j \in \mathbb{N})$ and $(\xi_j^i\colon j \in \mathbb{N})$ are i.i.d. sequences with uniform distribution on $[0, T]$ and Q^i, respectively.
- $(\Gamma_j^i\colon j \in \mathbb{N})$ is a random walk whose increments are exponentially distributed with mean $1/T$.
- $(V_j^i\colon j \in \mathbb{N})$ is an i.i.d. sequence with distribution F on $\mathbb{R}$, which we should be able to simulate from. We assume that F is symmetric when $\Lambda \in \mathscr{S}$.
- The sequences τ^i, ξ^i, Γ^i and V^i are independent from each other.
- $(\tau^i, \xi^i, \Gamma^i, V^i)$ is independent from $(\tau^k, \xi^k, \Gamma^k, V^k\colon k \neq i)$.

Because of (6), μ can always be written in the form (27) whenever the underlying stochastic basis is rich enough. We give three examples of such series representations.

Example 3 The proofs that the following choices are valid can be found in [18, Sect. 3], where also more examples are discussed.

1. *LePage's method*: we set $F := (\delta_{-1} + \delta_1)/2$ and $H(r, \pm 1) := \pm \rho^\leftarrow(r, \pm 1)$, where $\rho^\leftarrow(r, \pm 1) = \inf\{x \in (0, \infty)\colon \pi(\pm[x, \infty)) < r\}$ for $r \in (0, \infty)$.
2. *Bondesson's method*: we assume that $\pi(A) = \int_0^\infty F(A/g(t))\,\mathrm{d}t$ for $A \in \mathscr{B}(\mathbb{R}^d)$ with some non-increasing $g\colon \mathbb{R}_+ \to \mathbb{R}_+$. Then we define $H(r, v) := g(r)v$.
3. *Thinning method*: we choose F in such a way that Q is absolutely continuous with respect to F with density q, and define $H(r, v) := v\mathbb{1}_{\{q(v) \geq r\}}$.

Our approximation scheme is basically the same as in Sect. 3: we define G^N by (16) and Y^N by (19), with the difference that Λ^N on $[0, T] \times \mathbb{R}^d$ is now defined as

$$\Lambda^N(\mathrm{d}t, \mathrm{d}x) := \int_{\mathbb{R}} z\, \mu^N(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z)\,,$$

$$\mu^N(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) := \sum_{i=1}^{\infty} \sum_{j\,\colon\, \Gamma_j^i \leq N} \delta_{(\tau_j^i, \xi_j^i, H(\Gamma_j^i, V_j^i))}(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z)\,. \tag{28}$$

We can therefore rewrite $Y^N(t, x)$ for $(t, x) \in [0, T] \times \mathbb{R}^d$ as

$$Y^N(t, x) = Y_0(t, x) + \sum_{i=1}^{N} \sum_{j\,\colon\, \Gamma_j^i \leq N} G(t, x; \tau_j^i, \xi_j^i)\sigma(Y^N(\tau_j^i, \xi_j^i))H(\Gamma_j^i, V_j^i)\mathbb{1}_{\{\tau_j^i < t\}}\,. \tag{29}$$

This yields the following simulation algorithm:

Algorithm 2 Let $\mathscr{G}$ be a finite grid in $[0, T] \times \mathbb{R}^d$ and $N \in \mathbb{N}$.

1. For each $i = 1, \ldots, N$ set $j := 1$ and repeat the following:

   a. Draw E_j^i from an exponential distribution with mean $1/T$.
   b. Define $\Gamma_j^i := \Gamma_{j-1}^i + E_j^i$ $(\Gamma_0^i := 0)$.
   c. If $\Gamma_j^i > N$, set $J_i := j - 1$ and leave the loop; otherwise set $j := j + 1$.

2. For each $i = 1, \ldots, N$ and $j = 1, \ldots, J_i$ simulate independently

   a. a pair (τ_j^i, ξ_j^i) with uniform distribution on $[0, T] \times Q^i$;
   b. a random variable V_j^i with distribution F;
   c. the random variable $Y_0(\tau_j^i, \xi_j^i)$.

3. Sort the sequence $(\tau_j^i : i = 1, \ldots, N, j = 1, \ldots, J_i)$ in increasing order, yielding sequences $(\tau_i, \xi_i, \Gamma_i, V_i : i = 1, \ldots, \sum_{j=1}^N J_j)$. Now define

$$Y^N(\tau_i, \xi_i) := Y_0(\tau_i, \xi_i) + \sum_{j=1}^{i-1} G(\tau_i, \xi_i; \tau_j, \xi_j) \sigma(Y^N(\tau_j, \xi_j)) H(\Gamma_j, V_j) .$$

4. For each $(t, x) \in \mathscr{G}$ simulate $Y_0(t, x)$ and define $Y(t, x)$ by (29).

We can now prove a convergence theorem for Y^N to Y, similar to Theorem 1. Define

$$r_2^N := \left(\int_N^\infty \int_{\mathbb{R}} |H(r, v)|^p F(\mathrm{d}v) \, \mathrm{d}r \right)^{1/p^*} ,$$

$$r_3^N := \left| \int_N^\infty \int_{\mathbb{R}} H(r, v) \mathbb{1}_{\{p>1, \Lambda \notin \mathscr{S}\}} F(\mathrm{d}v) \, \mathrm{d}r \right| . \tag{30}$$

Theorem 2 *Grant assumptions A1–A7 under which the SVE (1) has a unique solution in B_{loc}^p. Further suppose that the jump measure μ of Λ has a representation in form of (27). Then Y^N as defined in (29) belongs to B_{loc}^p for all $N \in \mathbb{N}$, and for all $T \in \mathbb{R}_+$ and compact sets $K \subseteq \mathbb{R}^d$ there exists a constant $C(T) \in \mathbb{R}_+$ independent of N and K such that*

$$\sup_{(t,x) \in [0,T] \times K} \|Y(t, x) - Y^N(t, x)\|_{L^p} \leq C(T)(r_1^N(T, K) + r_2^N + r_3^N) . \tag{31}$$

If $\sum_{N=1}^\infty (r_1^N(T, K) + r_2^N + r_3^N)^{p^*} < \infty$, then we also have for all $(t, x) \in [0, T] \times K$ that $Y^N(t, x) \to Y(t, x)$ a.s. as $N \to \infty$.

Proof We start with some preliminaries. It follows from (27) and [18, Proposition 2.1] that on $[0, T] \times \mathbb{R}^d \times \mathbb{R}$ we have $\nu = \bar{\nu} \circ h^{-1}$ where $\bar{\nu}(\mathrm{d}t, \mathrm{d}x, \mathrm{d}r, \mathrm{d}v) = \mathrm{d}t \, \mathrm{d}x \, \mathrm{d}r \, F(\mathrm{d}v)$ and $h(t, x, r, v) = (t, x, H(r, v))$. Therefore, conditions (6) and (9) imply that

$$\int_0^\infty \int_{\mathbb{R}} |H(r, v)|^p F(\mathrm{d}v) \, \mathrm{d}r = \int_{\mathbb{R}} |z|^p \pi(\mathrm{d}z) < \infty ,$$

and $\quad \int_0^\infty \int_{\mathbb{R}} |H(r, v)| \mathbb{1}_{\{p>1, \Lambda \notin \mathscr{S}\}} F(\mathrm{d}v) \, \mathrm{d}r = \int_{\mathbb{R}} |z| \mathbb{1}_{\{p>1, \Lambda \notin \mathscr{S}\}} \pi(\mathrm{d}z) < \infty .$

Consequently, r_2^N and r_3^N are well defined and converge to 0 when $N \to \infty$. Similarly, the compensator ν_N of the measure $\mu - \mu^N$ is given by $\nu_N(dt, dx, dz) = dt\, dx\, \pi_N(dz)$, where $\pi_N = (\text{Leb} \otimes F) \circ H_N^{-1}$ and $H_N(r, v) = H(r, v) \mathbb{1}_{(N, \infty)}(r)$.

For the actual proof of Theorem 2 one can basically follow the proof of Theorem 1. Only the estimation of $I_3^N(t, x)$ as defined in (23) is different, which we shall carry out now. In the case of $p > 1$, we again use the Burkholder-Davis-Gundy inequality and obtain for $(t, x) \in [0, T] \times \mathbb{R}^d$

$$
\begin{aligned}
\|I_3^N(t, x)\|_{L^p} &= \left\| \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}} G^N(t, x; s, y) \sigma(Y^N(s, y)) z\, (\mu_N - \nu_N)(ds, dy, dz) \right\|_{L^p} \\
&\quad + \left\| \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}} G^N(t, x; s, y) \sigma(Y^N(s, y)) z \mathbb{1}_{\{\Lambda \notin \mathscr{S}\}} \nu_N(ds, dy, dz) \right\|_{L^p} \\
&\leq C(T) \left(\left(\int_0^t \int_{\mathbb{R}^d} \int_N^\infty \int_{\mathbb{R}} |G(t, x; s, y) H(r, v)|^p F(dv)\, dr\, d(s, y) \right)^{1/p} \right. \\
&\quad \left. + \int_0^t \int_{\mathbb{R}^d} |G(t, x; s, y)| \left| \int_N^\infty \int_{\mathbb{R}} H(r, v) \mathbb{1}_{\{\Lambda \notin \mathscr{S}\}} F(dv)\, dr \right| d(s, y) \right) \\
&\leq C(T)(r_2^N + r_3^N)\,.
\end{aligned}
$$

The case $p \in (0, 1]$ is treated analogously. One only needs to replace $\mu_N - \nu_N$ by μ_N and estimate via Jensen's inequality. $\square$

Example 4 (*Continuation of Example 3*) We calculate the rates r_2^N and r_3^N from (30) for the series representations given in Example 3. We assume that $p, q \in (0, 2]$ with $q < p$ are chosen such that $\int_{[-1,1]} |z|^q \pi(dz) < \infty$, and $q < 1$ if $\Lambda \in \mathscr{V}_0$. For all three examples we use the fact that $\pi = (\text{Leb} \otimes F) \circ H^{-1}$ and that $r > N$ implies $|H(r, v)| \leq |H(N, v)|$ for all $v \in \mathbb{R}$.

1. *LePage's method*: We have

$$
\begin{aligned}
(r_2^N)^{p^*} &= \int_N^\infty \frac{|H(r, 1)|^p + |H(r, -1)|^p}{2}\, dr \leq \frac{1}{2} \int_{[H(N, -1), H(N, 1)]} |z|^p\, \pi(dz) \\
&\leq \frac{1}{2} \int_{[H(1, -1), H(1, 1)]} |z|^q\, \pi(dz) (|H(N, -1)| \vee |H(N, 1)|)^{p-q}\,,
\end{aligned}
$$

and therefore

$$
\begin{aligned}
r_2^N &= \mathscr{O}\left((\rho^{\leftarrow}(N, 1) \vee \rho^{\leftarrow}(N, -1))^{(p-q)/p^*} \right)\,, \\
r_3^N &= \mathscr{O}\left((\rho^{\leftarrow}(N, 1) \vee \rho^{\leftarrow}(N, -1))^{1-q} \mathbb{1}_{\{p>1, \Lambda \notin \mathscr{S}\}} \right)\,.
\end{aligned}
$$

2. *Bondesson's method*: Since $H(r, v) = g(r)v$ and g is non-increasing, we obtain

$$
(r_2^N)^{p^*} = \int_N^\infty \int_{\mathbb{R}} |g(r)v|^p\, F(dv)\, dr \leq (g(N))^{p-q} \int_0^\infty g^q(r)\, dr \int_{\mathbb{R}} |v|^p\, F(dv)\,,
$$

and consequently

$$r_2^N = \mathcal{O}\left(g(N)^{(p-q)/p^*}\right) , \quad r_3^N = \mathcal{O}\left(g(N)^{1-q}\mathbb{1}_{\{p>1,\Lambda\notin\mathscr{S}\}}\right) .$$

3. *Thinning method*: Here we have

$$r_2^N = \left(\int_{\mathbb{R}}\int_N^{q(v)\vee N}|v|^p \, \mathrm{d}r \, F(\mathrm{d}v)\right)^{1/p^*} = \left(\int_{\mathbb{R}}|v|^p\,\frac{(q(v)-N)\vee 0}{q(v)}\,\pi(\mathrm{d}v)\right)^{1/p^*}$$

$$\le \left(\int_{\mathbb{R}}|z|^p\mathbb{1}_{\{q(v)\ge N\}}\,\pi(\mathrm{d}z)\right)^{1/p^*} ,$$

$$r_3^N \le \int_{\mathbb{R}}|z|\mathbb{1}_{\{q(v)\ge N\}}\mathbb{1}_{\{p>1,\Lambda\notin\mathscr{S}\}}\,\pi(\mathrm{d}z) .$$

In most situations, there exist $(\varepsilon^N)_{N\in\mathbb{N}} \subseteq \mathbb{R}_+$ with $\varepsilon^N \to 0$ as $N \to \infty$ such that $\{q(v) \ge N\} \subseteq [-\varepsilon^N, \varepsilon^N]$. In this case, one can apply the estimates in Example 2.

5   Simulation Study

In this section we visualize the sample path behaviour of the stochastic heat equation from Example 2 via a simulation study, using MATLAB programs from [4]. We take Λ to be a Lévy basis without drift, whose Lévy measure π is that of a gamma process, i.e.

$$\pi(\mathrm{d}z) = \gamma z^{-1}\exp(-\lambda z)\mathbb{1}_{\{z>0\}}\,\mathrm{d}z$$

with two parameters $\gamma, \lambda > 0$. In the figures below their values are always $\gamma = 10$ and $\lambda = 0.1$. Furthermore, we set $Y_0 \equiv 0$ and $\sigma \equiv 1$. Especially the latter choice simplifies the subsequent discussion a lot, but none of the issues we address below relies on this assumption. Thus, the process we would like to simulate is

$$Y(t, x) = \int_0^t\int_{\mathbb{R}^d}g(t - s, x - y)\,\Lambda(\mathrm{d}s, \mathrm{d}y) , \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d , \qquad (32)$$

with g being the heat kernel given in (14). In order to understand the path properties of Y, it is important to notice that g is smooth on the whole $\mathbb{R}_+ \times \mathbb{R}^d$ except at the origin where it explodes. More precisely, for every $t \in (0, \infty)$ the function $x \mapsto g(t, x)$ is the Gaussian density with mean 0 and variance $2t$, which is smooth and assumes its maximum at 0. Also, for every $x \ne 0$, the function $t \mapsto g(t, x)$ is smooth (also at $t = 0$), with maximum at $t = \|x\|^2/(2d)$. However, if $x = 0$, then $g(t, 0) = (4\pi t)^{-d/2}$ has a singularity at $t = 0$.

These analytical properties have direct consequences on the sample paths of Y. When Λ is of compound Poisson type, that is, has only finitely many atoms on compact sets, it can be readily seen from (32) that the evolution of Y after a jump J

at (τ, ξ) follows the shape of the heat kernel until a next jump arrives. In particular, for $x = \xi$, $Y(t, x)$ jumps to infinity at τ, and decays in t like $J(4\pi(t - \tau))^{-d/2}$ afterwards. But for every $x \neq \xi$, the evolution $t \mapsto Y(t, x)$ is *smooth* at $t = \tau$. In fact, it first starts to increase until $t = \tau + \|x - \xi\|^2/(2d)$ and then decays again. As a consequence, in space dimension 1, the space–time plot of Y shows a basically smoothly evolving path, only interrupted with slim poles at the jump locations of Λ; see the case $N = 50$ in Fig. 1. However, when Λ has infinite activity, that is, has infinitely many jumps on any non-empty open set, then it is known from [16, Theorem 4] that on any such set Y is unbounded, at least with positive probability. Therefore, the space–time plots of the approximations of Y with finitely many jumps must be treated with caution: in the limiting situation, no smooth area exists any more, but there will be a dense subset of singularities on the plane, which is in line with Fig. 1.

Another interesting observation, however, is the following: if we consider a countable number of x- or t-sections of Y (for $x \in \mathbb{R}^d$, the x-section of Y is given by the function $t \mapsto Y(t, x)$; for $t \in \mathbb{R}_+$, the t-section of Y is the function $x \mapsto Y(t, x)$), then it is shown in [19, Sect. 2] that these are continuous with probability one. Intuitively, this is possible because a.s. the sections never hit a jump (although they are arbitrarily close). For instance, Figs. 2 and 3 show t-sections of a realization of (32) in one, respectively two space dimensions. So as long as we only take countably many "measurements", we do not observe the space–time singularities of Y but only its relatively regular sections. In theory, this also includes the x-sections of the process Y. But if we plot them for one space dimension as in Fig. 4, one would conjecture from the simulation that they exhibit jumps in time. However, this is *not* true: the jump-like appearance of the x-sections are due to the fact that $g(\cdot, x)$ resembles a
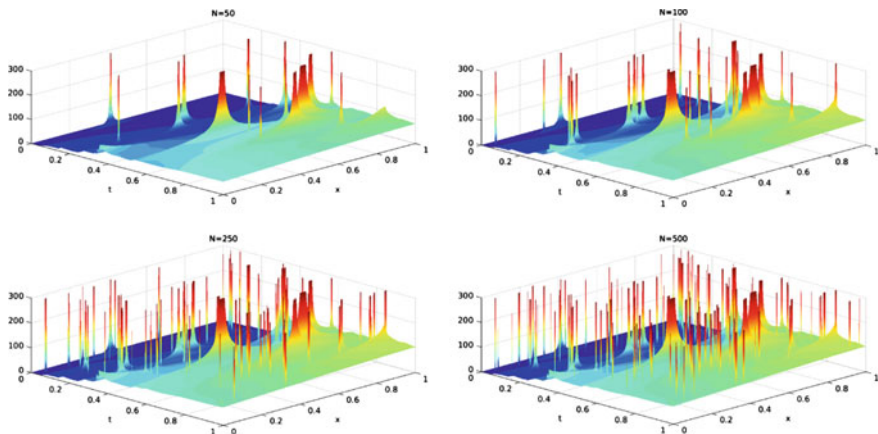

Fig. 1 Successive approximations of Y as given in (32) via Bondesson's method in dimension 1 for $(t, x) \in [0, 1] \times [0, 1]$ with $N \in \{50, 100, 250, 500\}$ jumps in the region $[0, 1] \times [-1, 2]$
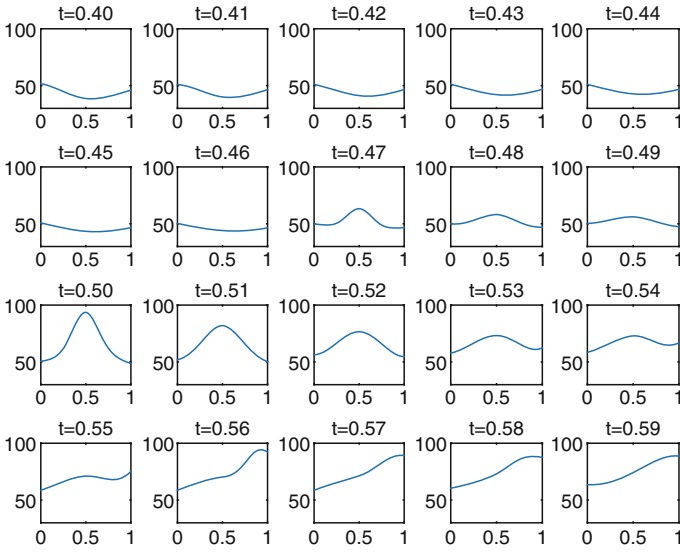
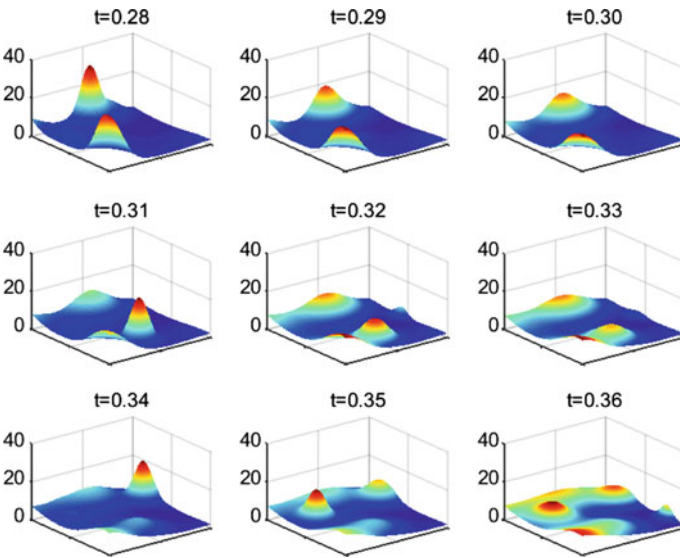**Fig. 2** Several t-sections of the realization of Y shown in Fig. 1 with $N = 500$

Fig. 3 Several t-sections in the region $[-1, 1]^2$ of a realization of Y in dimension 2 by Bondesson's method with $N = 500$ jumps within $[0, 1] \times [-2, 2]^2$
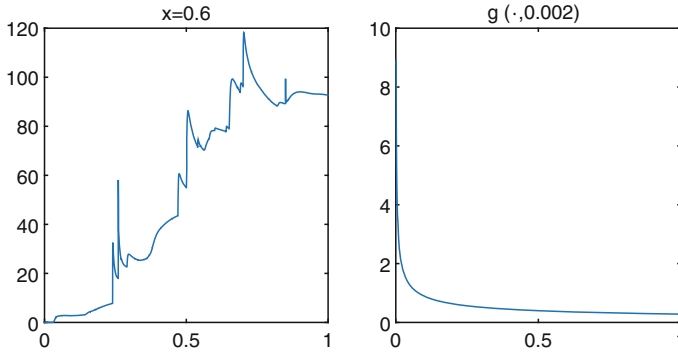
Fig. 4 The x-section of the realization of Y as in Fig. 1 with $N = 500$ at $x = 0.6$ and the heat kernel $g(\cdot, x)$ at $x = 0.002$

discontinuous function at $t = 0$ for small x. Of course, it follows right from the definition (14) that all x-sections of g are smooth everywhere.

Remark 1 Throughout this paper we have taken the stochastic heat equation as our key example for illustration purposes. Obviously, all theory developed in Sects. 3 and 4 holds under much more general assumptions, namely under A1–A7. However, the path behaviour of SVEs as examined through a simulation in this section heavily depends on the chosen kernel. While the path properties above are typical of kernels that originate from parabolic type SPDEs, or more generally, that are smooth everywhere except for an explosive singularity at the origin, a completely different picture arises when we consider bounded kernels. We refer to [2] and the references therein for examples of such kernels in different fields of application. In the case of a simple Ornstein-Uhlenbeck type kernel, a simulation study is carried out in [4, Chap. 5].

6   Conclusion and Outlook

In Sects. 3 and 4 we have presented two simulation algorithms for the SVE (1): Algorithms 1 and 2. In Theorems 1 and 2 we have determined the rate of convergence of the approximations Y^N to Y in the L^p-sense. If these rates are small enough, we have also proved a.s. convergence. Although the theoretical analysis of both schemes lead to quite similar results regarding their convergence behaviour, there are important differences which will decide on whether the one or the other method is preferable in concrete situations. For the first method of truncating the small jumps to work, one must be able to efficiently simulate from the truncated Lévy measure $\pi/\pi([-\varepsilon, \varepsilon]^c)$ for small ε. For the second method, which relies on series representations, the main challenge is to choose H and F in a way such that H is explicitly known *and* F can be easily simulated from. For instance, if one uses LePage's method (see Example 3), then $F = (\delta_{-1} + \delta_1)/2$ is easily simulated, but for

H, which is given by the generalized inverse tails of the underlying Lévy measure, maybe no tractable expression exists.

Finally, let us comment on further generalizations of our results. Throughout this paper, we have assumed that the driving noise Λ is a homogeneous Lévy basis, i.e. satisfies (5). In fact, we have introduced this condition only for the sake of simplicity: with a straightforward adjustment, all results obtained in this paper also hold for time- and space-varying (but deterministic) characteristics. Another issue is the finite time perspective which we have taken up for our analysis. An interesting question would be under which conditions (1) has a stationary solution, and in this case, whether one can simulate from it. Sufficient conditions for the existence and uniqueness of stationary solutions to (1) are determined in [5, Theorem 4.8]. Under these conditions, the methods used to derive Theorems 1 and 2 can indeed be extended to the case of infinite time horizon. We leave the details to the reader at this point.

At last, also the hypothesis that Λ is of pure-jump type can be weakened. If Λ has an additional drift (including the case where Λ has locally infinite variation and is not symmetric) but still no Gaussian part, the approximations Y^N in (19) or (29) will contain a further term that is a Volterra integral with respect to the Lebesgue measure. So each time in between two simulated jumps, a deterministic Volterra equation has to be solved numerically, which boils down to a deterministic PDE in the case where G comes from an SPDE. For this subject, there exists a huge literature, which is, of course, also very different to the stochastic case as considered above. If Λ also contains a Gaussian part, then one has to apply techniques from the papers cited in Sect. 1 and ours simultaneously. We content ourselves with referring to [22], who numerically analyzes a Volterra equation driven by a drift plus a Brownian motion. Finally, let us remark that if $p = 2$ (in particular, G must be square-integrable), it is possible for some Lévy bases to improve the results of Sect. 3 if we do not neglect the small jumps completely but approximate them via a Gaussian noise with the same variance, cf. [1] in the case of Lévy processes.

References

1. Asmussen, S., Rosiński, J.: Approximations of small jumps of Lévy processes with a view towards simulation. J. Appl. Probab. **38**(2), 482–493 (2001)
2. Barndorff-Nielsen, O.E., Benth, F.E., Veraart, A.E.D.: Ambit processes and stochastic partial differential equations. In: Nunno, G.D., Øksendal, B. (eds.) Advanced Mathematical Methods for Finance, pp. 35–74. Springer, Berlin (2011)
3. Barth, A., Lang, A.: Simulation of stochastic partial differential equations using finite element methods. Stochastics **84**(2–3), 217–231 (2012)
4. Chen, B.: Simulation of stochastic partial differential equations. Master's thesis, Technische Universität München (2014). https://mediatum.ub.tum.de/node?id=1225387

5. Chong, C.: Lévy-driven Volterra equations in space–time (2014). Preprint under arXiv:1407.8092 [math.PR]
6. Chong, C., Klüppelberg, C.: Integrability conditions for space–time stochastic integrals: theory and applications. Bernoulli **21**(4), 2190–2216 (2015)
7. Davie, A.M., Gaines, J.G.: Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations. Math. Comput. **70**(233), 121–134 (2001)
8. Dunst, T., Hausenblas, E., Prohl, A.: Approximate Euler method for parabolic stochastic partial differential equations driven by space-time Lévy noise. SIAM J. Numer. Anal. **50**(6), 2873–2896 (2012)
9. Gyöngy, I.: Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise II. Potential Anal. **11**(1), 1–37 (1999)
10. Hausenblas, E.: Finite element approximation of stochastic partial differential equations driven by Poisson random measures of jump type. SIAM J. Numer. Anal. **46**(1), 437–471 (2008)
11. Hausenblas, E., Marchis, I.: A numerical approximation of parabolic stochastic differential equations driven by a Poisson random measure. BIT Numer. Math. **46**(4), 773–811 (2006)
12. Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes, 2nd edn. Springer, Berlin (2003)
13. Knopp, K.: Theory and Application of Infinite Series. Dover, New York (1990)
14. Protter, P.: Volterra equations driven by semimartingales. Ann. Probab. **13**(2), 519–530 (1985)
15. Rajput, B.S., Rosiński, J.: Spectral representations of infinitely divisible processes. Probab. Theory Relat. Fields **82**(3), 451–487 (1989)
16. Rosiński, J.: On path properties of certain infinitely divisible processes. Stoch. Process. Appl. **33**(1), 73–87 (1989)
17. Rosiński, J.: On series representations of infinitely divisible random vectors. Ann. Probab. **18**(1), 405–430 (1990)
18. Rosiński, J.: Series representations of Lévy processes from the perspective of point processes. In: Barndorff-Nielsen, O.E., Mikosch, T., Resnick, S.I. (eds.) Lévy Processes, pp. 401–415. Birkhäuser, Boston (2001)
19. Saint Loubert Bié: E.: Étude d'une EDPS conduite par un bruit poissonnien. Probab. Theory Relat. Fields **111**(2), 287–321 (1998)
20. Walsh, J.B.: An introduction to stochastic partial differential equations. In: Hennequin, P.L. (ed.) École d'Été de Probabilités de Saint Flour XIV - 1984. Lecture Notes in Mathematics, vol. 1180, pp. 265–439. Springer, Berlin (1986)
21. Walsh, J.B.: Finite elment methods for parabolic stochastic PDE's. Potential Anal. **23**(1), 1–43 (2005)
22. Zhang, X.: Euler schemes and large deviations for stochastic Volterra equations with singular kernels. J. Differ. Equ. **244**(9), 2226–2250 (2008)