ORIGINAL PAPER

Point process simulation of generalised hyperbolic Lévy processes

Yaman Kındap[1] · **Simon Godsill**[1]

Abstract

Generalised hyperbolic (GH) processes are a class of stochastic processes that are used to model the dynamics of a wide range of complex systems that exhibit heavy-tailed behavior, including systems in finance, economics, biology, and physics. In this paper, we present novel simulation methods based on subordination with a generalised inverse Gaussian (GIG) process and using a generalised shot-noise representation that involves random thinning of infinite series of decreasing jump sizes. Compared with our previous work on GIG processes, we provide tighter bounds for the construction of rejection sampling ratios, leading to improved acceptance probabilities in simulation. Furthermore, we derive methods for the adaptive determination of the number of points required in the associated random series using concentration inequalities. Residual small jumps are then approximated using an appropriately scaled Brownian motion term with drift. Finally the rejection sampling steps are made significantly more computationally efficient through the use of squeezing functions based on lower and upper bounds on the Lévy density. Experimental results are presented illustrating the strong performance under various parameter settings and comparing the marginal distribution of the GH paths with exact simulations of GH random variates. The new simulation methodology is made available to researchers through the publication of a Python code repository.

Keywords Generalised inverse Gaussian process · Lévy process · Series representations · Monte Carlo methods · Stochastic differential equations

1 Introduction

The statistical properties of randomly evolving phenomena in many real-world applications can be studied using stochastic differential equations. The random behaviour in such systems are typically characterised through a Brownian motion term which implicitly assumes that some version of the Central-limit theorem is valid for the observed random behaviour. However, numerous dynamical systems exhibit more heavy-tailed characteristics than the Gaussian, see for example applications in financial modelling (Mandelbrot 1963; Fama 1965; Cont and Tankov 2003), communications (Azzaoui and Clavier 2010; Fahs et al. 2012; de Freitas et al. 2017; Liebeherr et al. 2012; Shevlyakov and Kim 2006; Warren and Thomas 1991), signal processing (Nikias and Shao 1995), image analysis (Achim et al. 2001, 2006), audio process-

ing (Godsill and Rayner 1998; Lombardi and Godsill 2006), climatological sciences (Katz and Brown 1992; Katz et al. 2002), in the medical sciences (Chen et al. 2010) and for the understanding of sparse modelling/compressive sensing (Unser et al. 2014a, b; Unser and Tafti 2014; Amini and Unser 2014; Carrillo et al. 2016; Lopes 2016; Zhou and Yu 2017; Tzagkarakis 2009; Achim et al. 2010). In such cases the stochastic driving term can be characterised by a Lévy process which generalises the sample paths and the marginal distributions of the term to include other parametric families such as the Poisson process and the α-stable process.

In general a continuous-time randomly evolving system may possess a continuous random Brownian motion component as well as sudden discrete random changes at random times ('jumps'). General Lévy processes encompass both of these classes of random evolution such that the increments of the process are independent and stationary (Bertoin 1997; Ken-Iti 1999). Here we focus on a broad class of such processes that are composed purely of jumps.

While there is substantial theoretical and applied interest in simulation of Lévy processes *per se*, in our work we are ultimately concerned with modelling and inference

✉ Yaman Kındap
  yk392@cam.ac.uk

  Simon Godsill
  sjg30@cam.ac.uk

1 SigProC Lab, Department of Engineering, University of Cambridge, Cambridge, UK

for systems driven by non-Gaussian Lévy processes, such as the linear stochastic differential equation (SDE) model (Øksendal 2014)

$$dX(t) = AX(t)dt + hdW(t)$$

where the more standard Brownian motion is replaced by a non-Gaussian Lévy process $\{W(t)\}$ (the so-called Background-Driving Lévy process (BDLP) Barndorff-Nielsen and Halgreen 1977), see for example our earlier work with Stable law Lévy processes (Lemke et al. 2015; Godsill et al. 2019; Riabiz et al. 2017). The work presented here though is focussed purely on the generation of the underlying Lévy processes, with the extension of our point process methods to the SDE case being in principle straightforward, as demonstrated in Lemke et al. (2015), Godsill et al. (2019) for the Stable law case. The convergence results from extending this paper to the linear SDE case are presented in Costa et al. (2023).

In this paper, we study simulation methods for a very broad class of Lévy processes, the generalised hyperbolic (GH) process (Barndorff-Nielsen et al. 2001; Eberlein and Hammerstein 2004) [also known as generalised hyperbolic Lévy motion (Eberlein 2001)], which captures various degrees of semi-heavy- or heavy-tailed behaviour such that the tails may be designed to be lighter than non-Gaussian Stable laws (which possess infinite variance (Samorodnitsky and Taqqu 1994), but heavier than a Gaussian (Borak et al. 2011. Some important special cases include the hyperbolic process (Eberlein and Keller 1995), the normal inverse-Gaussian process (NIG) (Barndorff-Nielsen 1978) and the variance-gamma process (Madan and Seneta 1990), which were introduced in the context of modelling empirical financial returns, and the Student-t process (see also Shah et al. 2014; Solin and Särkkä 2015 which was introduced as an extension to Gaussian processes in Machine Learning, although such processes are not equivalent to the Student-t Lévy processes simulated here). The GH distribution is defined as a normal variance-mean mixture where the required mixing distribution is the generalised inverse-Gaussian (GIG) distribution. Our current work improves a point process simulation framework for GIG processes (Godsill and Kındap 2021), extending it to the variance-mean mixture representation of the GH process, and in addition providing substantial modifications and improvements to the original methods.

The simulation of the sample paths of Lévy processes is a key area of research that enables the use of Lévy processes in inference and decision-making. In Rosiński (2001), Rosiński surveys generalised shot-noise series representations of Lévy processes and their relation with point processes, and this is the general framework adopted for the current paper (see Godsill and Kındap 2021; Lemke and Godsill 2015; Riabiz et al. 2020; Godsill et al. 2019 and references therein, for

our previous studies using this methodology). Other relevant developments include (Barndorff-Nielsen 1997a) which present the theory of NIG processes, Rydberg (1997) which present approximate sampling methods for the NIG case, and Barndorff-Nielsen and Shephard (2001) which give applications of shot-noise series based methods for non-Gaussian Ornstein-Uhlenbeck (OU) processes. Exact simulation methods for the class of tempered stable (TS) processes are studied in Zhang (2011), Qu et al. (2021), Grabchak (2019), Sabino (2022). In addition, approximate simulation methods for GH Lévy fields, which are infinite-dimensional GH Lévy processes, are studied in Barth and Stein (2017).

It is shown in Barndorff-Nielsen and Halgreen (1977) that the GH distribution is infinitely divisible and hence can be the distribution of a Lévy process at time $t = 1$. The GH distribution possesses a five parameter probability density function defined for random variables on the real line as follows (Eberlein 2001; Eberlein and Hammerstein 2004)

$$f_{GH}(x) = a(\lambda, \alpha, \beta, \delta) \left(\delta^2 + (x - \mu)^2\right)^{(\lambda - \frac{1}{2})/2}$$
$$\times K_{\lambda - \frac{1}{2}} \left(\alpha \sqrt{\delta^2 + (x - \mu)^2}\right) \exp(\beta(x - \mu)) \tag{1}$$

where

$$a(\lambda, \alpha, \beta, \delta) = \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi} \alpha^{\lambda - \frac{1}{2}} \delta^\lambda K_\lambda(\delta \sqrt{\alpha^2 - \beta^2})}$$

$K_\nu(\cdot)$ is the modified Bessel function of the second kind with index ν. The parameter $\lambda \in \mathbb{R}$ characterises the tail behaviour, $\alpha > 0$ determines the shape, $0 \le |\beta| < \alpha$ controls the skewness, $\mu \in \mathbb{R}$ is a location parameter and $\delta > 0$ is the scale parameter. Alternative parametrisations of the probability density function in the limiting parameter settings are discussed in Eberlein and Hammerstein (2004).

The three parameter probability density function $f_{GIG}(\lambda, \delta, \gamma)$ of the GIG distribution may be linked to Eq. (1) via a variance-mean mixture of Gaussians Eberlein (2001). Using the parameterisation $\gamma = \sqrt{\alpha^2 - \beta^2}$, the variance-mean mixture for the GH distribution may be expressed as

$$f_{GH}(x) = \int_0^\infty N(x; \mu + \beta u, u) f_{GIG}\left(u; \lambda, \delta, \sqrt{\alpha^2 - \beta^2}\right) du \tag{2}$$

where u is a GIG distributed random variable. Random variate generation algorithms for a GIG variable are studied in Devroye (2014); Hörmann and Leydold (2013), and their extension to GH distributed random variables are then obtained through the normal variance-mean construction shown in Eq. (2).

GH processes are generally intractable for simulation since the Lévy density associated with the GIG process is expressed as an integral involving certain Bessel functions. Simulation methods based on generalised shot-noise representation of the GIG Lévy process are given in Godsill and Kındap (2021). These methods rely on the construction of dominating point processes that are tractable for simulation, followed by thinning methods derived from upper bounds on the intractable integrand. In the earlier work the series generated must be truncated to a finite number of terms which needs to be tuned by the user, and hence may be inefficient in some parameter regimes.

In this paper we show for the first time a practical method for simulation of paths of the GH process, based on subordination with a GIG process. The paper provides several significant contributions. The first contribution is to provide improved simulation methods for the underlying GIG process based on tighter bounds for the construction of dominating processes and the corresponding thinning method, and a proof of convergence is provided for the first time, based on an earlier result by Rosiński (2001). Secondly, we derive *adaptive* truncation methods for approximating the infinite series involved in our representation which allow for the first time an automatic choice of the truncation level for the jumps of the GIG process. Once truncation has occurred we then approximate the residual error committed by adding an appropriately scaled Brownian motion term with drift, motivated by Central-limit theorem-style results for the residual error. Furthermore, the thinning (rejection sampling) methods are made significantly more computationally efficient through the introduction of 'squeezing' functions that both upper- and lower-bound the acceptance probabilities. Finally, acceptance probability bounds are derived and convergence properties of the novel simulation methods are compared against the methods introduced in Godsill and Kındap (2021). The simulation methodology is made available to researchers through the publication of a Python code repository.[1]

The paper is organised as follows. Section 2 presents the necessary preliminaries for simulation of Lévy processes and their corresponding point processes, using a generalised shot-noise approach. Section 3 introduces the specific form of the GIG Lévy density and derives various bounds on these densities as well as constructing dominating Lévy densities from the related bounds. Section 4 gives simulation algorithms for the GH Lévy process based on the simulation of the previously discussed dominating Lévy processes and associated thinning methods. Section 5, presents an adaptive truncation method for the infinite series involved in generalised shot-noise representations and a method of approximating the residual series. Section 6 gives a practical sampling algorithm based on squeezing functions for increasing the efficiency

of simulation. Section 7 presents example simulations, comparing the distribution of the paths generated with exact simulations of GH random variates.

2 Generalised shot-noise representations

In this section series representations of Lévy processes given in Rosiński (2001), Kallenberg (2002) that enable their simulation are reviewed. Let $W(t)$ be a Lévy process on some time interval of interest $t \in [0, T]$ having no drift or Brownian motion part, and hence containing purely jumps; then the characteristic function (CF) is given by Kallenberg (2002, Corollary 13.8), as

$$E\left[\exp(iuW(t))\right]$$
$$= \exp\left(t\left[\int_{\mathbb{R}\setminus\{0\}} (e^{iuw} - 1 - iw\mathbb{I}(|w| < 1))Q(dw)\right]\right)$$

where Q is a Lévy measure on $\mathbb{R}_0 := \mathbb{R}\setminus\{0\}$ satisfying $\int_{\mathbb{R}_0} \min(1, w^2)Q(dw) < \infty$. Under this definition $W(T)$ is a random variable whose distribution is infinitely divisible.

We will also require a more restricted class of non-negative, non-decreasing Lévy processes $X(t)$, the *subordinator* process, whose CF is given by:

$$E\left[\exp(iuX(t))\right] = \exp\left(t\left[\int_0^\infty (e^{iux} - 1)Q_X(dx)\right]\right)$$

and which has the more restrictive requirement that

$$\int_0^\infty \min(1, x)Q_X(dx) < \infty \tag{3}$$

$Q_X(dx)$ defines the *density* of jumps for $\{X(t)\}$ such that the expected number of jumps of size $x \in [a, b]$ is $\mu_{[a,b]} = \int_a^b Q_X(dx)$ and the number of jumps is a Poisson random variable with mean $\mu_{[a,b]}$. We will be dealing with *infinite activity* processes for which $\int_0^\infty Q_X(dx) = \infty$ and hence there are almost surely an infinite number of jumps in time interval $[0, T]$.

In order to generate sample paths from the GH process, we will use the so-called variance-mean mixture representation of its Lévy measure,

$$Q_{GH}(dw) = \int_0^\infty N(dw; \mu + \beta x, x)Q_{GIG}(dx) \tag{4}$$

which is the normal mixture representation of the GH Lévy measure, analogous to the normal mixture representation of its probability density (2), and where Q_{GIG} is the Lévy measure of a Generalised inverse Gaussian (GIG) subordinator process (see Barndorff-Nielsen 1997b; Wolpert and Ickstadt 1998b and the following section for further detail). Hence

through standard subordination techniques the GH process can be expressed as $W(t) = \mu_W X(t) + \sigma_W B(X(t))$ where B is a standard Brownian motion (Simon 1999; Veraart and Winkel 2010; Barndorff-Nielsen and Shephard 2012).

It is first required to simulate a realisation $\{x_i\}_{i=1}^{\infty}$ of the jumps from the underlying GIG subordinator process and to use a generalised shot-noise representation (Rosiński 2001) to simulate from Q_{GH}:

$$W(t) = \sum_{i=1}^{\infty} W_i \mathbb{I}_{V_i \leq t} - t c_i \tag{5}$$

where $\{V_i \in [0, T]\}_{i=1}^{\infty}$ are i.i.d. uniform random variables representing the arrival time of jumps, and independent of the jump sizes $\{W_i\}_{i=1}^{\infty}$, which are independently distributed as:

$$W_i \sim N(\mu + \beta x_i, x_i)$$

Note that Rosiński (2001) proves the almost sure convergence of such series to $W(t)$ for x_i non-increasing, i.e. jumps of $X(t)$ are generated in order of decreasing size. The terms c_i are centering terms which we may take as zero for the GH class of processes as a result of the condition in Eq. (3).

The task remaining is to generate ordered realisations of the jumps in the subordinator, $\{x_i\}_{i=1}^{\infty}$. Here the Lévy-Itô integral representation of $X(t)$ may be invoked:

$$X(t) = \int_{(0,\infty)} x N([0, t], dx) \tag{6}$$

where N is a bivariate point process with mean measure $Leb. \times Q$ on $[0, T] \times \mathbb{R}_0$ which may be represented with Dirac functions as

$$N = \sum_{i=1}^{\infty} \delta_{V_i, X_i} \tag{7}$$

where again $\{V_i \in [0, T]\}_{i=1}^{\infty}$ are i.i.d. uniform random variables independent of $\{X_i\}$ which represent the arrival time of jumps, $\{X_i\}_{i=1}^{\infty}$ are the jump sizes, and T is the duration of the time interval considered. If we substitute N into Eq. (6) we obtain a series representation of $X(t)$ as:

$$X(t) = \sum_{i=1}^{\infty} X_i \mathbb{I}_{V_i \leq t} \tag{8}$$

The classical method to generate such a subordinator process, with Lévy measure Q (Ferguson and Klass 1972; Rosiński 2001; Wolpert and Ickstadt 1998a, b) simulates jumps of decreasing size by an appropriate transformation of the epochs of a unit rate Poisson process. Briefly, an arbitrarily large number of epochs $\{\Gamma_i\}_{i=1,2,...}$ is randomly

simulated from a unit rate Poisson process. These terms may be transformed into the jump magnitudes of the corresponding subordinator process by calculating the upper tail probability of the Lévy measure $Q^+(x) = Q([x, \infty)) < \infty$. A corresponding non-increasing function $h(\cdot)$ is then defined as the inverse tail probability, $h(\gamma) = \inf_x \{x; \ Q^+(x) = \gamma\}$ which assigns a non-increasing jump value to each of the ordered Poisson epochs, $\{X_i = h(\Gamma_i)\}$. Thus, small Γ_i values correspond to large jumps $h(\Gamma_i)$ and vice versa. It can be seen from this definition that $\mathbb{E}[\#\{X_i; \ X_i \geq x\}] = Q^+(x)$: this procedure is essentially following an analogous formulation to the standard inverse CDF method for random variate generation, but applied here to a point process intensity function instead of a probability distribution. Formally, the mapping theorem (Kingman 1992) ensures that the resulting transformed process is a Poisson process having the correct Lévy density $Q(x)$.

Since there is an infinite number of jumps in the series representation (8), the simulation is in practice truncated at a finite number of terms and the remaining small jumps are ignored or approximated somehow (Asmussen and Rosiński 2001), a topic that is addressed in Sect. 5 of this paper.

The generic method reviewed here requires the explicit evaluation of the inverse tail measure $h(\gamma)$ which is not tractable for the GIG process. An alternative approach was devised in Godsill and Kındap (2021), simulating from a tractable dominating point process N_0 having Lévy measure Q_0 such that $dQ_0(x)/dQ(x) \geq 1, \ \forall x \in (0, \infty)$ for which $h_0(\gamma)$ is directly available. The resulting samples from N_0 are then thinned with probability $dQ(x)/dQ_0(x)$ as in (Lewis and Shedler 1979; Rosiński 2001) to obtain the desired jump magnitudes $\{x_i\}$ of the subordinator process. The generic procedure is given in Algorithm 1 for a point process $Q(x)$ having dominating density $Q_0(x) \geq Q(x)$ and $h_0(\gamma) = \inf_x \{x; \ Q_0^+(x) = \gamma\}$.

Algorithm 1 Generation of the jumps of a point process having Lévy density $Q(x)$ and dominating process $Q_0(x) \geq Q(x)$.

---

1. Assign $N = \emptyset$,
2. Generate the epochs of a unit rate Poisson process, $\{\Gamma_i; \ i = 1, 2, 3...\}$,
3. For $i = 1, 2, 3...$,

   – Compute $x_i = h_0(\Gamma_i)$
   – With probability $Q(x_i)/Q_0(x_i)$, accept x_i and assign $N = N \cup x_i$.

Note that our work here will later require partial simulation of such point processes on measurable sets A on jump magnitudes, i.e. $Q_A(x) = \mathbb{I}_A(x) Q(x)$, and typically A will simply be an interval $(a, b], b \leq \infty$. This partial simulation

is straightforwardly achieved by replacing Step 2) in Algorithm 1 with the steps provided in Algorithm 2.

Algorithm 2 Generation of Poisson process epochs corresponding to jump magnitudes $x_i \in (a, b]$ where $a > b$.

- $i = 1, \Gamma_0 = Q_0^+(a)$
- While $\Gamma_{i-1} < Q_0^+(b)$
  1. $E_i \sim Exp(1)$
  2. $\Gamma_i = \Gamma_{i-1} + E_i$
  3. $i = i + 1$
- Return $\{\Gamma_j\}_{j=1}^{i-1}$

As before $Q_0^+(x) = Q_0([x, \infty))$ and $Exp(1)$ is the unit mean exponential distribution, and noting that the While loop in Algorithm 2 may in practice be replaced with a draw from $M \sim Poisson(Q_0^+(b) - Q_0^+(a))$ followed by M i.i.d. draws for the (unordered) Γ_i terms from a uniform distribution $U(Q_0^+(b) - Q_0^+(a))$.

A rejection sampling procedure such as Algorithm 1 may be viewed within the generalised shot-noise framework of Rosiński (2001) in which the process is expressed as a random function of the underlying Poisson epochs $\{\Gamma_i\}$ as follows

$$X(t) = \sum_i H(\Gamma_i, e_i) \mathbb{I}(V_i \leq t)$$

where $H(\gamma, .)$ is a non-increasing function of γ, and e_i are random variables or vectors drawn independently across i. Rosiński (2001) Th. 4.1 proves the almost sure convergence of such series under mild conditions. In particular the conditions of Th. 4.1 (A) are satisfied. First the distribution of $H(\cdot, \cdot)$ is expressed as a probability kernel $\sigma(\gamma, A)$ for measurable sets A:

$$\mathbb{P}\{H(\gamma, e) \in A\} = \sigma(\gamma, A)$$

and it follows from the Marking Theorem (Kingman 1992) that the resulting point process has Lévy measure

$$Q(A) = \int_0^\infty \sigma(\gamma, A) d\gamma$$

Applying this to verify Algorithm 1, take $H(\gamma, e) = h(\gamma)e$ and $e \in \{0, 1\}$ binomial with $\mathbb{P}\{e = 1\} = Q(h_0(\gamma))/Q_0(h_0(\gamma))$. We will consider only non-zero jump sizes, since *jumps* of size zero do not impact the point process, and indeed Lévy measures $Q(dx)$ are not defined for $x = 0$. Then it follows for measurable sets $A_0 = A \backslash 0$ that

$$\sigma(\gamma, A_0) = \mathbb{I}(h_0(\gamma) \in A_0) Q(h_0(\gamma))/Q_0(h_0(\gamma))$$

and hence the resulting Lévy measure is

$$
\begin{aligned}
Q_1(A_0) &= \int_0^\infty \mathbb{I}(h_0(\gamma) \in A_0) Q(h_0(\gamma))/Q_0(h_0(\gamma)) d\gamma \\
&= \int_{x \in A_0} Q(x)/Q_0(x)(Q_0(x)dx) = \int_{x \in A_0} Q(x)dx,
\end{aligned}
$$

as required. Here we have made the substitution $x = h_0(\gamma)$, so $\gamma = Q_0^+(x)$ and $d\gamma = Q_0(x)dx$. While the procedure of Algorithm 1 is well known to be valid, see e.g. Rosiński (2001), we include the sketch proof here since we will use more sophisticated versions of it to prove validity of our own algorithms for GIG and GH process simulation in subsequent sections of the paper.

Simple and well known examples of the procedures in Algorithms 1 and 2 are the *tempered stable* and *gamma* processes, which we will require as part our sampling procedures for the GIG process later in the paper. The corresponding Lévy densities and thinning probabilities for these cases are given in Godsill and Kındap (2021) (Section 2.1 and 2.2). The associated sampling algorithms are repeated here for reference purposes in Algorithms 3 and 4.

Algorithm 3 Generation of the jumps of a tempered stable process with Lévy density $Q_{TS}(x) = Cx^{-1-\alpha}e^{-\beta x}$ ($x \geq 0$) where $0 < \alpha < 1$ is the tail parameter and $\beta \geq 0$ is the tempering parameter.

---

1. Assign $N_{TS} = \emptyset$,
2. Generate the epochs of a unit rate Poisson process, $\{\Gamma_i; i = 1, 2, 3...\}$,
3. For $i = 1, 2, 3...$,
   - Compute $x_i = \left(\frac{\alpha \Gamma_i}{C}\right)^{-1/\alpha}$,
   - With probability $e^{-\beta x_i}$, accept x_i and assign $N_{TS} = N_{TS} \cup x_i$.

Algorithm 4 Generation of the jumps of a gamma process with Lévy density $Q_{Ga}(x) = Cx^{-1}e^{-\beta x}$ ($x \geq 0$) where $C > 0$ is the shape parameter and $\beta > 0$ is the rate parameter.

---

1. Assign $N_{Ga} = \emptyset$,
2. Generate the epochs of a unit rate Poisson process, $\{\Gamma_i; i = 1, 2, 3...\}$,
3. For $i = 1, 2, 3...$,
   - Compute $x_i = \frac{1}{\beta(\exp(\Gamma_i/C)-1)}$,
   - With probability $(1 + \beta x) \exp(-\beta x_i)$, accept x_i and assign $N_{Ga} = N_{Ga} \cup x_i$.

---

3 The generalised inverse Gaussian Lévy process

In this section, the GIG Lévy process and its Lévy measure are defined. Tractable bounds on this Lévy measure are required in order to simulate the GIG (and hence the GH) process, and in this section we provide improved bounds compared with those in Godsill and Kındap (2021). These improved bounds are proven in the following section to have higher acceptance rates for the rejection sampling procedures that underlie the algorithms.

The density of the Lévy measure of a GIG process (Eberlein and Hammerstein 2004, Eq. 74), following a change of variables as in Godsill and Kındap (2021), is given by

$$\frac{e^{-x\gamma^2/2}}{x} \left[\frac{2}{\pi^2} \int_0^\infty \frac{e^{-\frac{z^2 x}{2\delta^2}}}{z|H_{|\lambda|}(z)|^2} dz + \max(0, \lambda) \right], \quad x > 0$$

where $H_\lambda(z) = J_\lambda(z) + i Y_\lambda(z)$ is the Bessel function of the third kind, also known as the Hankel function of the first kind, which is defined in terms of $J_\lambda(z)$, the Bessel function of the first kind, and $Y_\lambda(z)$, the Bessel function of the second kind. The presence of an integral involving the Bessel function makes the simulation of such processes intractable except for certain edge cases.

Naturally, the GIG Lévy density can be divided into two terms as

$$Q_{GIG}(x) = \frac{2e^{-x\gamma^2/2}}{\pi^2 x} \int_0^\infty \frac{e^{-\frac{z^2 x}{2\delta^2}}}{z|H_{|\lambda|}(z)|^2} dz$$

and a second term, present only for $\lambda > 0$ as

$$\frac{\lambda e^{-x\gamma^2/2}}{x}, \quad x > 0 \tag{9}$$

which is the Lévy density of a gamma process with shape parameter λ and rate $\gamma^2/2$. It is straightforward to simulate from this second term using Algorithm 4, thus our attention is directed towards simulation of the point process with Lévy density $Q_{GIG}(x)$.

In order to avoid any direct calculation of the integral in $Q_{GIG}(x)$, the general approach proposed in Godsill and Kındap (2021) is to consider a bivariate point process $Q_{GIG}(x, z)$ on $(0, \infty) \times (0, \infty)$ which has, by construction, the GIG Lévy density as its marginal, i.e. $Q_{GIG}(x) = \int_0^\infty Q_{GIG}(x, z)dz$ such that

$$Q_{GIG}(x, z) = \frac{2e^{-x\gamma^2/2}}{\pi^2 x} \frac{e^{-\frac{z^2 x}{2\delta^2}}}{z|H_{|\lambda|}(z)|^2} \tag{10}$$

Thus, joint samples $\{x_i, z_i\}$ are simulated from the point process with intensity function $Q_{GIG}(x, z)$, from which the samples $\{x_i\}$ are retained as samples from $Q_{GIG}(x)$. However, simulation from $Q_{GIG}(x, z)$ is still intractable because of the presence of the Bessel function. This is overcome by constructing tractable bivariate dominating point processes with intensity function $Q^0_{GIG}(x, z)$ and thinning with probability $Q_{GIG}(x, z)/Q^0_{GIG}(x, z)$ to yield samples from the desired process with Lévy density Q_{GIG}.

The generic approach proposed here will involve a marginal-conditional factorisation of both point processes:

$$Q^0_{GIG}(x, z) = Q^0_{GIG}(x) Q^0_{GIG}(z|x),$$
$$Q_{GIG}(x, z) = Q_{GIG}(x) Q_{GIG}(z|x),$$

where $Q^0_{GIG}(z|x)$ and $Q_{GIG}(z|x)$ are proper probability densities, i.e. $\int_0^\infty Q^0_{GIG}(z|x)dz = 1$ and $\int_0^\infty Q_{GIG}(z|x)dz = 1$. Thus z may be interpreted as a *marking* variable and $(x, z) \in (0, \infty) \times (0, \infty)$ form a bivariate Poisson process (Kingman 1992). The generic algorithm for sampling $Q_{GIG}(x)$ is then given below, followed by its proof of validity under the generalised shot noise approach.

Algorithm 5 Generation of the jumps of a point process having Lévy density $Q(x) = \int_0^\infty Q(x)Q(z|x)dz$ and dominating process $Q_0(x, z) = Q_0(x)Q_0(z|x)$ such that $Q_0(x, z) \geq Q(x, z)$.

---
1. Assign $N = \emptyset$,
2. Generate Poisson epochs $\{\Gamma_i\}$ and corresponding ordered jump sizes $\{x_i = h_0(\Gamma_i)\}$ from the marginal process $Q_0(x)$ using Algorithm 1
3. For $i = 1, 2, 3...$,
    - Simulate $z_i \sim Q_0(z|x_i)$
    - With probability $Q(x_i, z_i)/Q_0(x_i, z_i)$, accept x_i and assign $N = N \cup x_i$.

We now proceed to prove the convergence of Alg. 5 using Rosiński (2001) Th. 4.1 (A). Note that the proof is here presented for the first time.

Lemma 1 *The Generalised Shot noise process* $X(t) = \sum_i H(\Gamma_i, e_i)\mathbb{I}(V_i \leq t)$ *generated according to Algorithm 5 converges a.s. to the Poisson point process with Lévy density* $Q(x)$.

Proof Algorithm 5 generates, for each point x_i, an auxiliary marking variable $z_i \sim Q_0(z_i|x_i)$, and an acceptance variable $a_i \in \{0, 1\}$ that is binomial with $\mathbb{P}\{a_i = 1\} = Q(x_i, z_i)/Q_0(x_i, z_i)$. Thus set $e_i = (z_i, a_i) \in (0, \infty) \times \{0, 1\}$ and hence

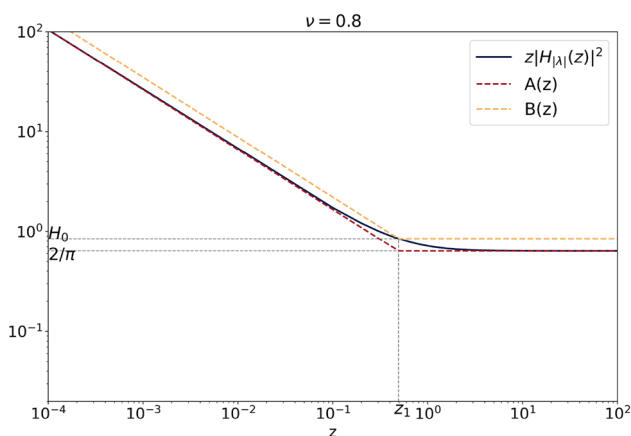$$H(\gamma, (z_i, a_i)) = h_0(\gamma_i)a_i$$

Fig. 1 Plot of Bessel function bounds, $\nu = 0.8$. z_0 set equal to z_1 and $z_1 = \left(\frac{2^{1-2\nu}\pi}{\Gamma(\nu)^2} \right)^{1/(1-2\nu)}$

Fig. 2 Plot of Bessel function bounds, $\nu = 0.3$. z_0 set equal to z_1 and $z_1 = \left(\frac{2^{1-2\nu}\pi}{\Gamma(\nu)^2} \right)^{1/(1-2\nu)}$

with resulting probability kernel

$$
\begin{aligned}
\sigma(\gamma, A_0) &= \int_0^\infty \mathbb{I}(h_0(\gamma) \in A_0)\, (Q(h_0(\gamma), z) \\
&\quad / Q_0(h_0(\gamma), z))\, Q_0(z | h_0(\gamma))\, dz \\
&= \mathbb{I}(h_0(\gamma) \in A_0)\, Q(h_0(\gamma)) / Q_0(h_0(\gamma))
\end{aligned}
$$

for all measurable sets A_0 not containing 0. Hence the resulting Lévy measure is

$$
\begin{aligned}
Q_1(A_0) &= \int_0^\infty \mathbb{I}(h_0(\gamma) \in A_0)\, Q(h_0(\gamma)) / Q_0(h_0(\gamma))\, d\gamma \\
&= \int_{x \in A_0} Q(x)\, dx
\end{aligned}
$$

as required. The remaining conditions in Rosiński (2001) Th. 4.1 (A) are simply that $Q(\cdot)$ is a Lévy density, which is true by construction (it is a subordinator and hence satisfies (3)), and a second technical condition that is always satisfied by subordinators, see Rosiński (2001) Remark 4.1. □

A new set of bounds on $z|H_\nu(z)|^2$ is now given in Theorem 1 below, which will be used in bounding the overall function (10). The bounds are graphically illustrated for the two distinct parameter ranges in Figs. 1 and 2. The proof of the theorem follows a similar scheme to Theorem 2 in Godsill and Kındap (2021) and is hence only briefly stated:

Theorem 1 *Choose a point $z_0 \in (0, \infty)$ and compute $H_0 = z_0 |H_\nu(z_0)|^2$. This will define the corner point on a piecewise lower or upper bound. Choose now any $0 \le z_1 \le \left(\frac{2^{1-2\nu}\pi}{\Gamma(\nu)^2} \right)^{1/(1-2\nu)}$ and define the following functions:*

$$
A(z) = \begin{cases} \frac{2}{\pi}\left(\frac{z_1}{z}\right)^{2\nu-1}, & z < z_1 \\ \frac{2}{\pi}, & z \ge z_1 \end{cases}
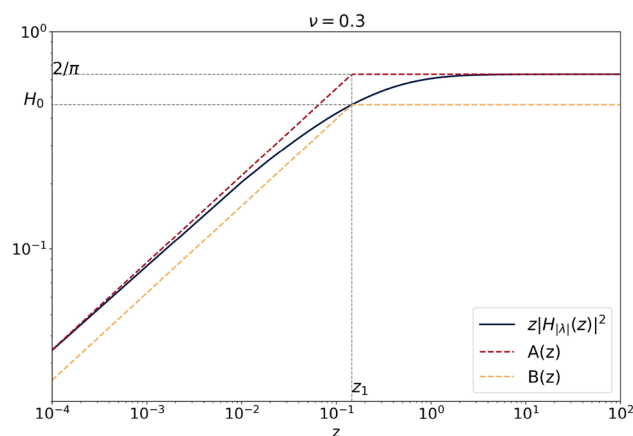$$

and

$$
B(z) = \begin{cases} H_0 \left(\frac{z_0}{z}\right)^{2\nu-1}, & z < z_0 \\ H_0, & z \ge z_0 \end{cases}
$$

Then, for $0 < \nu \le 0.5$,

$$
A(z) \ge z|H_\nu(z)|^2 \ge B(z) \tag{11}
$$

and for $\nu \ge 0.5$,

$$
A(z) \le z|H_\nu(z)|^2 \le B(z) \tag{12}
$$

with all inequalities becoming equalities when $\nu = 0.5$, and both $A(z)$ bounds (left side inequalities) becoming tight at $z = 0$ and $z = \infty$.

Proof The proof is an obvious extension of Theorem 2 in Godsill and Kındap (2021) where we now allow for a range of values $0 \le z_1 \le \left(\frac{2^{1-2\nu}\pi}{\Gamma(\nu)^2} \right)^{1/(1-2\nu)}$. This follows from the fact that any value of z_1 less than $\left(\frac{2^{1-2\nu}\pi}{\Gamma(\nu)^2} \right)^{1/(1-2\nu)}$, the function $A(z)$ is shifted to the left and lies below of those plotted in Fig. 1 and above of those plotted in Fig. 2, hence providing a valid but less tight bounding function. □

Remark 1 Choice of any $z_1 < \left(\frac{2^{1-2\nu}\pi}{\Gamma(\nu)^2} \right)^{1/(1-2\nu)}$ leads to a poorer (less tight) bound on the true function $z|H_\nu(z)|^2$. In particular, as $z_1 \to 0$ we obtain the crude and well known (Watson 1944, Section 13.75) bound $A(z) = \frac{2}{\pi}$, which was employed in a first version of our method for $|\lambda| > 0.5$ (Godsill and Kındap 2021, Theorem 1). This asymptotic bound forms the loosest bound $A(z)$ and all other valid choices of z_1 yield increasingly tight bounds on the required Lévy density $Q_{GIG}(x, z)$, which we employ in our new and improved sampling algorithms for $|\lambda| > 0.5$.

Corollary 1 *For any positive z and fixed $|\lambda|$, the following bounds are obtained by replacing $z|H_\nu(z)|^2$ with $A(z)$ and $B(z)$ in the definition of $Q_{GIG}(x, z)$ (10). For the case $|\lambda| \geq 0.5$ we have:*

$$Q_{GIG}^B(x, z) \leq Q_{GIG}(x, z) \leq Q_{GIG}^A(x, z) \tag{13}$$

and for $0 < |\lambda| \leq 0.5$ we have:

$$Q_{GIG}^A(x, z) \leq Q_{GIG}(x, z) \leq Q_{GIG}^B(x, z) \tag{14}$$

with equality being achieved in both cases for $|\lambda| = 0.5$. Here $Q_{GIG}^A(x, z)$ is defined as:

$$Q_{GIG}^A(x, z) = \begin{cases} \dfrac{e^{-x\gamma^2/2}}{\pi x} \dfrac{z^{2|\lambda|-1} e^{-\frac{z^2 x}{2\delta^2}}}{z_1^{2|\lambda|-1}}, & z < z_1 \\[4mm] & \tag{15a} \\[2mm] \dfrac{e^{-x\gamma^2/2}}{\pi x} e^{-\frac{z^2 x}{2\delta^2}}, & z \geq z_1 \tag{15b} \end{cases}$$

and $Q_{GIG}^B(x, z)$ defined as:

$$Q_{GIG}^B(x, z) = \begin{cases} \dfrac{2e^{-x\gamma^2/2}}{\pi^2 x} \dfrac{z^{2|\lambda|-1} e^{-\frac{z^2 x}{2\delta^2}}}{H_0 z_0^{2|\lambda|-1}}, & z < z_0 \\[4mm] & \tag{16a} \\[2mm] \dfrac{2e^{-x\gamma^2/2}}{\pi^2 x} \dfrac{e^{-\frac{z^2 x}{2\delta^2}}}{H_0}, & z \geq z_0 \tag{16b} \end{cases}$$

Remark 2 Setting $z_0 = z_1$, it can be clearly seen that the ratio $Q_{GIG}^A(x, z)/Q_{GIG}^B(x, z) = \pi H_0/2$, is a constant independent of the value of x and z. This fact, which can be clearly visualised in Figs. 1 and 2 (note the log-scale), is utilised later in our development of a retrospective 'squeezed' sampler, see Sect. 6.

Corollary 2 *The bound in Eq. (15a) can be rewritten in factorised form as*

$$\begin{aligned} Q_{N_1}^A(x, z) &= \frac{e^{-x\gamma^2/2}}{\pi x} \frac{z^{2|\lambda|-1} e^{-\frac{z^2 x}{2\delta^2}}}{z_1^{2|\lambda|-1}} \mathbb{I}_{0<z<z_1} \\ &= \frac{e^{-x\gamma^2/2}}{\pi x^{1+|\lambda|}} \frac{(2\delta^2)^{|\lambda|} \gamma(|\lambda|, z_1^2 x/(2\delta^2))}{2 z_1^{2|\lambda|-1}} \\ &\qquad \frac{\Gamma(|\lambda|) \sqrt{Ga}(z||\lambda|, x/(2\delta^2))}{\gamma(|\lambda|, z_1^2 x/(2\delta^2))} \mathbb{I}_{0<z<z_1} \\ &= Q_{N_1}^A(x) Q_{N_1}^A(z|x) \end{aligned}$$

where

$$Q_{N_1}^A(z|x) = \frac{\Gamma(|\lambda|) \sqrt{Ga}(z||\lambda|, x/(2\delta^2))}{\gamma(|\lambda|, z_1^2 x/(2\delta^2))} \mathbb{I}_{0<z<z_1}$$

is a conditional right-truncated square-root gamma density[2] with its associated normalising constant. The marginal term $Q_{N_1}^A(x)$ is a modified tempered $|\lambda|$-stable process.[3]

Corollary 3 *The bound in Eq. (15b) can be rewritten in a similar way as*

$$\begin{aligned} Q_{N_2}^A(x, z) &= \frac{e^{-x\gamma^2/2}}{\pi x} e^{-\frac{z^2 x}{2\delta^2}} \mathbb{I}_{z \geq z_1} \\ &= \frac{e^{-x\gamma^2/2}}{\pi x^{3/2}} \frac{(2\delta^2)^{0.5} \Gamma(0.5, z_1^2 x/(2\delta^2))}{2} \\ &\qquad \frac{\Gamma(0.5) \sqrt{Ga}(z|0.5, x/(2\delta^2))}{\Gamma(0.5, z_1^2 x/(2\delta^2))} \mathbb{I}_{z \geq z_1} \\ &= Q_{N_2}^A(x) Q_{N_2}^A(z|x) \end{aligned}$$

where

$$Q_{N_2}^A(z|x) = \frac{\Gamma(0.5) \sqrt{Ga}(z|0.5, x/(2\delta^2))}{\Gamma(0.5, z_1^2 x/(2\delta^2))} \mathbb{I}_{z \geq z_1}$$

is a conditional left-truncated square-root gamma density with its associated normalising constant. The marginal term $Q_{N_2}^A(x)$ is a modified tempered 0.5-stable process.

In Corollaries 2 and 3, the point process intensities correspond marginally to a (modified) tempered stable process in x, and conditionally to a truncated $\sqrt{Ga}$ density for z. This feature enables sampling from the dominating bivariate point process $Q_{GIG}^A(x, z)$ by first sampling x and then, conditional on the value of x, sampling a corresponding z value. Here, for the parameter range $|\lambda| \geq 0.5$, we are extending the approach previously derived only for the parameter range $0 < |\lambda| < 0.5$ (Godsill and Kındap 2021), (and here summarised in Sect. 4.3). Notice that allowing $z_1 \to 0$, as discussed in the Remark following Theorem 1, will result in our previous crude bounding function for the parameter range $|\lambda| > 0.5$ (Godsill and Kındap 2021, Corollary 1).

Since the functions $Q_{GIG}^A(x, z)$ and $Q_{GIG}^B(x, z)$ form both upper and lower bounds on the target point process

---

[2] The square-root gamma density $\sqrt{Ga}$ is the density of $X^{0.5}$ when $X \sim Ga(x|\alpha, \beta)$, which has the probability density function $\sqrt{Ga}(x|\alpha, \beta) = \frac{2\beta^\alpha}{\Gamma(\alpha)} x^{2\alpha-1} e^{-\beta x^2}$.

[3] The lower/upper incomplete gamma functions are defined, for $\text{Re}(s) > 0$, as:

$$\gamma(s, x) = \int_0^x t^{s-1} e^{-t} dt, \quad \Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt.$$

$Q_{GIG}(x, z)$, see (13) and (14), we are able to construct effective sampling algorithms for all parameter ranges $|\lambda| > 0$ based upon rejection sampling ideas, see Sects. 4.1 and 4.3. Furthermore, in Sect. 6 we use the corresponding lower bounds to create efficient 'squeezed' versions of these algorithms. In the case of $0 < |\lambda| \leq 0.5$, the new algorithm is a direct development of that presented in Godsill and Kındap (2021), while in the case $|\lambda| > 0.5$ the algorithm now follows the same structure as for the other parameter range, in contrast with the previous approach from Godsill and Kındap (2021) that uses a cruder bound. In all parameter ranges we propose significant improvements over the previous work, including better bounds for simulation of the marginal process, adaptive truncation, simulation of a Gaussian residual term and squeezed rejection sampling.

In the next section we show how to move from simulations of the underlying GIG process towards our ultimate aim here, which is simulation of the GH process.

4 Simulating GH processes

In this section simulation algorithms for the GH Lévy process are presented. Our approach relies on the definition of a GH process as a subordinated Brownian motion where the subordinator is a GIG process (Barndorff-Nielsen and Halgreen 1977; Barndorff-Nielsen 1978). In this approach, jumps sizes $\{x_i\}$ are first generated from the underlying GIG process with intensity Q_{GIG}, see previous sections for full details. Then, jumps for the corresponding GH process are obtained as

$$w_i = \mu + \beta x_i + \sigma \sqrt{x_i} u_i, \quad u_i \overset{iid}{\sim} N(0, 1) \tag{17}$$

for some $\beta \in \mathbb{R}$ and $\sigma > 0$.

The conditional simulation of the GH process is common to all parameter regimes and is presented in Algorithm 6. Given jump times and magnitudes (V_i, W_i) the corresponding value of the GH Lévy process at t is given in Eq. (5).

We now detail the methods for generation of the underlying GIG process.

4.1 The case for $|\lambda| \geq 0.5$

Here a new algorithm is presented for simulation in the parameter range $|\lambda| \geq 0.5$, based on the bound $Q_{GIG}^A(x, z)$ derived in previous sections, which is an improved bound compared with that in Godsill and Kındap (2021) Algorithm 3. The process associated with the Lévy density $Q_{GIG}^A(x, z)$ can be considered as a marked point process split into two independent point processes N_1 and N_2 having factorised (marginal-conditional) intensity functions as given in Corollaries 2 and 3, respectively.

Algorithm 6 Simulation of GH process.

1. Generate a large number of points x_i from the GIG Lévy process with intensity Q_{GIG}, see Algorithms 8 and 10 for $|\lambda| \geq 0.5$, or Algorithms 12 and 14 for $0 < |\lambda| \leq 0.5$,
2. If $\lambda > 0$, draw an additional set of points $\{x'_j\}$ from the gamma process $\frac{\lambda e^{-x\gamma^2/2}}{x}$, $x > 0$, see (9) and take the union $\{x_i\} \leftarrow \{x_i\} \cup \{x'_j\}$,
3. For each point x_i, draw an independent and identically distributed random variate $u_i \sim \mathcal{N}(0, 1)$,
4. The corresponding GH jump sizes w_i are obtained as:

   $$w_i = \mu + \beta x_i + \sigma \sqrt{x_i} u_i$$

5. For each jump size w_i, draw an independent and identically distributed jump time $v_i \sim \mathcal{U}(0, T)$.
6. The process $W(t)$ is then obtained as:

   $$W(t) = \sum_{i=1}^{\infty} w_i \mathcal{I}(v_i \leq t)$$

Both N_1 and N_2 correspond to a marginal modified tempered stable process for x and a conditional truncated $\sqrt{Ga}$ density for z. Each simulated pair (x, z) is accepted with probability equal to the ratio $Q_{GIG}(x, z) / Q_{GIG}^A(x, z)$. As a result of the piecewise form of $Q_{GIG}^A(x, z)$, the accept/reject steps for N_1 and N_2 may be treated independently and the union of points from the two processes forms the final set of GIG points. The thinning probabilities for points drawn from $Q_{N_1}^A$ and $Q_{N_2}^A$ are then:

$$\frac{Q_{GIG}(x, z)}{Q_{N_1}^A(x, z)} = \frac{2}{\pi |H_{|\lambda|}(z)|^2 \left(\frac{z^{2|\lambda|}}{z_1^{2|\lambda|-1}} \right)} \tag{18}$$

$$\frac{Q_{GIG}(x, z)}{Q_{N_2}^A(x, z)} = \frac{2}{\pi z |H_{|\lambda|}(z)|^2} \tag{19}$$

Due to the presence of upper and lower incomplete gamma functions in the marginal point process envelopes $Q_{N_1}^A(x)$ and $Q_{N_2}^A(x)$ defined as

$$Q_{N_1}^A(x) = \frac{e^{-x\gamma^2/2}}{\pi x^{1+|\lambda|}} \frac{(2\delta^2)^{|\lambda|} \gamma(|\lambda|, z_1^2 x/(2\delta^2))}{2 z_1^{2|\lambda|-1}} \tag{20}$$

and

$$Q_{N_2}^A(x) = \frac{e^{-x\gamma^2/2}}{\pi x^{3/2}} \frac{(2\delta^2)^{0.5} \Gamma(0.5, z_1^2 x/(2\delta^2))}{2} \tag{21}$$

direct simulation from these Lévy densities are still intractable and hence dominating processes and associated thinning methods are required.

For $Q_{N_1}^A(x)$ the following bound is used to formulate a tractable dominating process (Neuman 2013a, Theorem 4.1):

$$\frac{a\gamma(a,x)}{x^a} \leq \frac{(1+ae^{-x})}{(1+a)} \tag{22}$$

so that the dominating process can be expressed as:

$$
\begin{aligned}
Q_{N_1}^A(x) &\leq \frac{e^{-x\gamma^2/2}}{\pi x} \frac{z_1(1+|\lambda|e^{-(z_1^2 x)/(2\delta^2)})}{2|\lambda|(1+|\lambda|)} \\
&= \frac{z_1}{2\pi(1+|\lambda|)} \left(\frac{e^{-x\gamma^2/2}}{|\lambda|x} + \frac{e^{-x(\gamma^2/2+z_1^2/(2\delta^2))}}{x} \right) \\
&= Q_{N_1}^{A,d}(x)
\end{aligned}
\tag{23}
$$

Notice that the point process associated with $Q_{N_1}^{A,d}(x)$ may be considered as the union of two independent gamma processes. Points are then independently accepted with probability $Q_{N_1}^A(x)/Q_{N_1}^{A,d}(x)$. The corresponding algorithm is given in Algorithm 7.

Algorithm 7 Sampling from $Q_{N_1}^A(x)$.

---

1. $N = \emptyset$,
2. Generate a gamma process N_{Ga}^1 having parameters $a_1 = \frac{z_1}{2\pi|\lambda|(1+|\lambda|)}$ and $\beta_1 = \gamma^2/2$,
3. Generate a gamma process N_{Ga}^2 having parameters $a_2 = \frac{z_1}{2\pi(1+|\lambda|)}$ and $\beta_2 = \gamma^2/2 + z_1^2/(2\delta^2)$,
4. $N = N_{Ga}^1 \cup N_{Ga}^2$,
5. For each $x \in N$ accept with probability

$$\frac{(2\delta^2)^{|\lambda|}\gamma(|\lambda|,(z_1^2 x)/(2\delta^2))|\lambda|(1+|\lambda|)}{x^{|\lambda|}z_1^{2|\lambda|}(1+|\lambda|\exp\left(-z_1^2 x/(2\delta^2)\right))}$$

otherwise reject and delete x from N.

Having simulated the x values from the marginal point process associated with $Q_{N_1}^A(x)$, the corresponding z values are simulated from a right-truncated square-root gamma density as in Corollary 2 and accept-reject steps are carried out to obtain samples from the N_1 point process. The complete procedure is outlined in Algorithm 8.

For the simulation of $Q_{N_2}^A(x)$ in (21), a bound on the term $\Gamma(0.5, z_1^2 x/(2\delta^2))$ is established by using the well-known equivalence $\Gamma(0.5, x) = \sqrt{\pi}\,\mathrm{erfc}(\sqrt{x})$ where $\mathrm{erfc}(\cdot)$ is the complementary error function. Two valid upper bounds on the gamma function are then obtained directly from Chiani et al. (2003) as

$$\Gamma(0.5, x) \leq \sqrt{\pi}\left[\frac{1}{2}e^{-2x} + \frac{1}{2}e^{-x}\right] \leq \sqrt{\pi}e^{-x} \tag{24}$$

While the first inequality is tighter, using it requires the simulation of two TS processes instead of a single process.

Algorithm 8 Generation of N_1 for $|\lambda| \geq 0.5$

---

1. $N_1 = \emptyset$,
2. Simulate x_i from the marginal point process associated with $Q_{N_1}^A(x)$ as given in Algorithm 7,
3. For each x_i, simulate a z_i from a truncated square-root gamma density

$$\frac{\Gamma(|\lambda|)\sqrt{\mathrm{Ga}}(z||\lambda|, x_i/(2\delta^2))}{\gamma(|\lambda|, z_1^2 x_i/(2\delta^2))}\mathcal{I}_{0<z<z_1}$$

4. With probability

$$\frac{Q_{GIG}(x_i, z_i)}{Q_{N_1}^A(x_i, z_i)} = \frac{2}{\pi|H_{|\lambda|}(z_i)|^2\left(\frac{z_i^{2|\lambda|}}{z_1^{2|\lambda|-1}}\right)}$$

accept x_i, i.e. set $N_1 = N_1 \cup x_i$, otherwise discard x_i.

Hence in the current implementation the right hand bound in Eq. (24) is chosen and the associated dominating point process envelope is then given by

$$
\begin{aligned}
Q_{N_2}^A(x) &\leq \frac{\delta e^{-\left[\frac{z_1^2}{2\delta^2}+\frac{\gamma^2}{2}\right]x}}{\sqrt{2\pi}x^{3/2}} \\
&= Q_{N_2}^{A,d}(x)
\end{aligned}
\tag{25}
$$

which can be simulated as a TS process and for each x_i the probability of acceptance is $\Gamma(0.5, z_1^2 x_i/(2\delta^2))/(\sqrt{\pi}e^{-z_1^2 x_i/(2\delta^2)})$. The corresponding algorithm is given in Algorithm 9.

Algorithm 9 Sampling from $Q_{N_2}^A(x)$.

---

1. Generate a tempered stable process N_{MTS} with parameters $C = \frac{\delta}{\sqrt{2\pi}}$, $\alpha = 0.5$ and $\beta = \frac{z_1^2}{2\delta^2} + \frac{\gamma^2}{2}$ using Algorithm 3.
2. For each point $x \in N_{MTS}$, accept with probability $\Gamma(0.5, z_1^2 x/(2\delta^2))/(\sqrt{\pi}e^{-z_1^2 ps2x/(2\delta^2)})$, otherwise reject and delete x from N_{MTS}.

Using the simulated values x_i, the corresponding z_i values are generated from the conditional left-truncated square-root gamma density $Q_{N_2}^A(z|x)$ and the whole procedure is outlined in Algorithm 10.

Note that the bound shown in Eq. (24) is a significant improvement over the bound based on the complete gamma function as used in Alg. 7 of Godsill and Kındap (2021). The choice of using a single TS process instead of the two TS processes associated with the sharper inequality is due to ease of implementation (fewer independent point processes to generate). However we do note that the right hand bound in Eq. (24) would lead to additional point rejections, and hence in some critical applications the tighter bound may be preferred, and parallel processing of the two generated TS

Algorithm 10 Generation of N_2 for $|\lambda| \geq 0.5$

1. $N_2 = \emptyset$,
2. Simulate x_i from the marginal point process associated with $Q^A_{N_2}(x)$ as shown in Algorithm 9,
3. For each x_i, simulate a z_i from a truncated square-root gamma density

$$\frac{\Gamma(0.5)\sqrt{\mathrm{Ga}}(z|0.5, x_i/(2\delta^2))}{\Gamma(0.5, z_1^2 x_i/(2\delta^2))}\mathcal{I}_{z \geq z_1}$$

4. With probability

$$\frac{Q_{GIG}(x_i, z_i)}{Q^A_{N_2}(x_i, z_i)} = \frac{2}{\pi z_i |H_{|\lambda|}(z_i)|^2}$$

accept x_i, i.e. set $N_2 = N_2 \cup x_i$, otherwise discard x_i.

---

processes might indeed reduce computational burden even further.

Finally, the set of points $N = N_1 \cup N_2$ is a realisation of jump magnitudes corresponding to a GIG process having intensity function $Q_{GIG}(x)$. The associated GH process may be obtained using Algorithm 6.

Remark 1 ote that whenever $\lambda > 0$, the set of points generated from the process with intensity function in Eq. (9) is added (by taking a union operation) to the set of points coming from $Q_{GIG}(x)$ in Algorithms 8 and 10 for $|\lambda| \geq 0.5$, or Algorithms 12 and 14 for $0 < |\lambda| \leq 0.5$, to obtain the full set of jumps $\{x_i\}$ from the required GIG process.

Remark 2 otice that for $\gamma = 0$, the gamma process N^1_{Ga} in Algorithm 7 is not well-defined since its rate parameter is zero and hence N_1 cannot be simulated. In this case, set $z_1 = 0$ and then only samples from N_2 are required. In this case the conditional density for z becomes the complete square-root gamma density instead of a truncated one and the marginal modified tempered stable process for x reduces to a standard TS process that does not require any further thinning. The simulation algorithm then becomes equivalent to Alg. 3 in Godsill and Kındap (2021). For all other values of γ we set $z_1 \leq \left(\frac{2^{1-2\nu}\pi}{\Gamma(\nu)^2}\right)^{1/(1-2\nu)}$ (with $\nu = |\lambda|$), in order to achieve the improved bound according to Theorem 1, with the best bound corresponding to $z_1 = \left(\frac{2^{1-2\nu}\pi}{\Gamma(\nu)^2}\right)^{1/(1-2\nu)}$, the case plotted in Fig. 1. Lower values of z_1 correspond to moving the lower bound in Fig. 1 to the left: clearly still a valid bound but suboptimal.

4.2 Acceptance rates for simulation from $Q^A_{GIG}(x, z)$

We now analyse the acceptance rates for the new procedure. This will enable a quantitative comparison with our previous methods in Godsill and Kındap (2021). The acceptance probabilities for the two point processes N_1 and N_2 are obtained

from (18) and (19) as

$$\rho_1(x, z) = \frac{2}{\pi |H_{|\lambda|}(z)|^2 \left(\frac{z^{2|\lambda|}}{z_1^{2|\lambda|-1}}\right)}$$

$$\rho_2(x, z) = \frac{2}{\pi z |H_{|\lambda|}(z)|^2}.$$

The expected value of the acceptance rates for fixed x may be evaluated w.r.t. the sampling densities for random variable Z, i.e. $Q^A_{N_1}(z|x)$ (20) and $Q^A_{N_2}(z|x)$ (21):

$$\mathbb{E}[\rho_1(x, Z)] = \int_0^{z_1} \frac{2}{\pi z |H_{|\lambda|}(z)|^2 \left(\frac{z}{z_1}\right)^{2|\lambda|-1}}$$
$$\frac{\Gamma(|\lambda|)\sqrt{\mathrm{Ga}}(z||\lambda|, x/(2\delta^2))}{\gamma(|\lambda|, z_1^2 x/(2\delta^2))} dz$$

$$\mathbb{E}[\rho_2(x, Z)] = \int_{z_1}^{\infty} \frac{2}{\pi z |H_{|\lambda|}(z)|^2}$$
$$\frac{\Gamma(0.5)\sqrt{\mathrm{Ga}}(z|0.5, x/(2\delta^2))}{\Gamma(0.5, z_1^2 x/(2\delta^2))} dz$$

However, the presence of the term $z|H_{|\lambda|}(z)|^2$ in both integrals makes them intractable. The expected acceptance rates may then be bounded by using the same functions $A(z)$ and $B(z)$, introduced in Theorem 1, to replace $z|H_{|\lambda|}(z)|^2$. Note that only the lower bound on the acceptance rates are of interest here since upper bounding both expectations using $A(z)$ leads to a trivial upper bound of 1 on the acceptance rates.

The acceptance rates associated with the N_1 and N_2 processes can be lower bounded using Theorem 2 and the proof is provided in the Appendix.

Theorem 2 *Choose a point* $z_0 \in [0, \infty)$, *compute* $H_0 = z_0 |H_{|\lambda|}(z_0)|^2$ *and fix* $z_1 = \left(\frac{2^{1-2\nu}\pi}{\Gamma(\nu)^2}\right)^{1/(1-2\nu)}$ *with* $\nu = |\lambda|$. *For any fixed* x *and* $|\lambda| \geq 0.5$, *the following lower bounds on* $\mathbb{E}[\rho_1(x, Z)]$ *and* $\mathbb{E}[\rho_2(x, Z)]$ *apply:*

$$\mathbb{E}[\rho_1(x, Z)] \geq \begin{cases} \frac{2}{\pi H_0}\left[\left(\frac{z_1}{z_0}\right)^{2|\lambda|-1}\frac{\gamma(|\lambda|, \frac{z_0^2 x}{2\delta^2})}{\gamma(|\lambda|, \frac{z_1^2 x}{2\delta^2})} + \left(\frac{z_1^2 x}{2\delta^2}\right)^{|\lambda|-0.5} \right. \\ \left. \frac{\left(\gamma(0.5, \frac{z_1^2 x}{2\delta^2}) - \gamma(0.5, \frac{z_0^2 x}{2\delta^2})\right)}{\gamma(|\lambda|, \frac{z_1^2 x}{2\delta^2})}\right], z_0 \in [0, z_1) \\ \frac{2}{\pi H_0}\left(\frac{z_1}{z_0}\right)^{2|\lambda|-1} \qquad\qquad , z_0 \in [z_1, \infty) \end{cases}$$
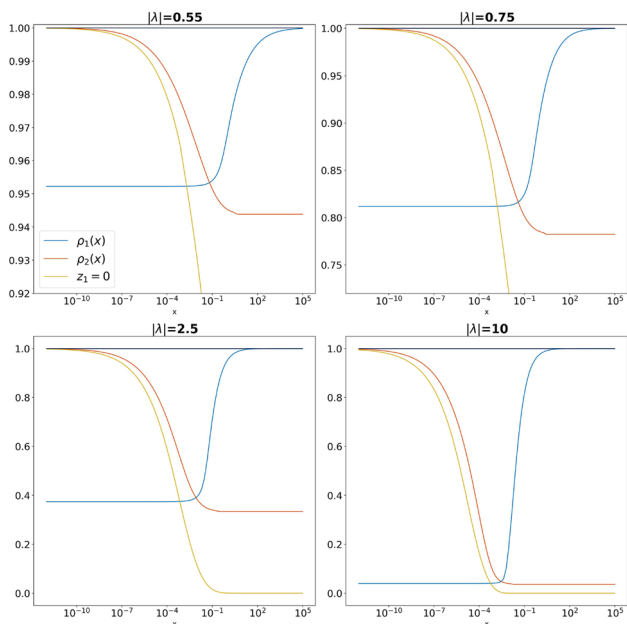
$$(26)$$

Fig. 3 Plot of optimised lower bounds on $\rho_1(x)$ and $\rho_2(x)$, for various $|\lambda| > 0.5$. $\delta = 0.1$ in all cases and bounds do not depend on γ

$$
\mathbb{E}\left[\rho_2(x, Z)\right] \geq
\begin{cases}
\dfrac{2}{\pi H_0} & , z_0 \in [0, z_1] \\[2ex]
\dfrac{2}{\pi H_0}\left[\dfrac{\Gamma(0.5, \frac{z_0^2 x}{2\delta^2})}{\Gamma(0.5, \frac{z_1^2 x}{2\delta^2})} + \left(\dfrac{z_0^2 x}{2\delta^2}\right)^{0.5 - |\lambda|} \right. \\[2ex]
\left. \dfrac{\left(\gamma(|\lambda|, \frac{z_0^2 x}{2\delta^2}) - \gamma(|\lambda|, \frac{z_1^2 x}{2\delta^2})\right)}{\Gamma(0.5, \frac{z_1^2 x}{2\delta^2})}\right], & z_0 \in (z_1, \infty)
\end{cases}
\tag{27}
$$

The proof for Theorem 2 can be found in the Appendix. Note that the corner point $z_0 \in (0, \infty)$ may be chosen arbitrarily, while z_1 is a fixed quantity in order to ensure a correct bounding function. Therefore the lower bound may be optimised for both N_1 and N_2 w.r.t. z_0 for each x value, leading to:

$$
\rho_1(x) := \max_{z_0}\{\mathbb{E}\left[\rho_1(x, Z)\right]\}
$$

and

$$
\rho_2(x) := \max_{z_0}\{\mathbb{E}\left[\rho_2(x, Z)\right]\}
$$

Note that the acceptance rates for our previous algorithm Section 3.1.1 of Godsill and Kındap (2021) can be considered as a limiting case of the new procedure corresponding to $z_1 = 0$, and hence the new acceptance rates are at least as large as the previous rates for each $x > 0$. In Fig. 3, the optimised lower and upper bounds on the acceptance rates for the new procedure are shown. Additionally for comparison, the lower bounds on the previous algorithm ($z_1 = 0$) are

plotted in Fig. 3. This illustrates that the new procedure is a significant improvement in terms of the acceptance rate for N_2 (i.e. $E[\rho_2(x, Z)]$) for small $|\lambda|$ values, and a slight improvement for larger $|\lambda|$ values.

For each fixed x value, the optimised lower bounds $\rho_i(x)$ in Fig. 3 are obtained using an implementation of Sequential Least Squares Programming (SLSQP) in the standard Python library SciPy. The optimiser is applied to Eqs. (26) and (27) to obtain $\rho_1(x)$ and $\rho_2(x)$. Note that the sequence of x values in Fig. 3 are log-linearly spaced. Additionally, for several fixed x and λ values, we plot the lower bounds in Eqs. (26) and (27) as a function of z_0 in Figs. 4 and 5. These show a clearly defined optimum that is not centred on any obvious solution such as z_1, although $z_0 = z_1$ would be a reasonable first guess if optimisation were to be avoided. These show that the optimal point of z_0 lies to the left of z_1 for all N_1 cases in Fig. 4, and to the right of z_1 in N_2 cases, Fig. 5.

Here we provide lower bounds on the expected acceptance probability across the whole range $x \in (0, \infty)$. We can observe from Fig. 3 that the optimised lower bounds $\rho_i(x)$ exhibit monotonicity (not proven) and limiting behaviour as $x \to 0$ and $x \to \infty$ and that these limits can thus be postulated as a uniform lower bound on the average acceptance probability. The limits for each $\rho_i(x)$ can be found from simple asymptotic expansions of the incomplete gamma functions[4] as,

$$
\lim_{x \to 0} \mathbb{E}\left[\rho_1(x, Z)\right] \geq
\begin{cases}
\dfrac{2}{\pi H_0} & , z_0 \in [0, z_1) \\[2ex]
\dfrac{2}{\pi H_0}\left(\dfrac{z_1}{z_0}\right)^{2|\lambda| - 1} & , z_0 \in [z_1, \infty)
\end{cases}
$$

so that it can be seen that the best (highest) lower bound is $\lim_{x \to 0} \mathbb{E}\left[\rho_1(x, Z)\right] \geq 2/(\pi H_0)$, for any $z_0 \leq z_1$, the term z_1/z_0 being less than unity for $z_0 > z_1$ and the power $2|\lambda| - 1$ being greater than 0 in the second case.

Similarly for $\rho_2(x)$, we have:

$$
\lim_{x \to \infty} \mathbb{E}\left[\rho_2(x, Z)\right] \geq
\begin{cases}
\dfrac{2}{\pi H_0} & , z_0 \in [0, z_1) \\[2ex]
\dfrac{2}{\pi H_0}\left(\dfrac{z_1}{z_0}\right)^{2|\lambda| - 1} & , z_0 \in [z_1, \infty)
\end{cases}
$$

and once again the best (highest) lower bound is seen to be $\lim_{x \to 0} \mathbb{E}\left[\rho_2(x, Z)\right] \geq 2/(\pi H_0)$, for any $z_0 \leq z_1$.

In both cases it is then clear that the optimal choice of $z_0 \in [0, z_1]$ is $z_0 = z_1$, since $H_0 = z_0|H_{|\lambda|}(z_0)|^2$ is monotonically decreasing as a function of z_0 (see, informally, Fig. 1, and more formally the monotonicity arguments of Theorem 1 of Godsill and Kındap (2021)). Thus the optimal lower bound on expected acceptance probability is found as, for both N_1 and N_2,

---

[4] $\gamma(s, x) \overset{x \to 0}{\to} x^s/s$ and $\Gamma(s, x) \overset{x \to \infty}{\to} x^{s-1}e^{-x}$.
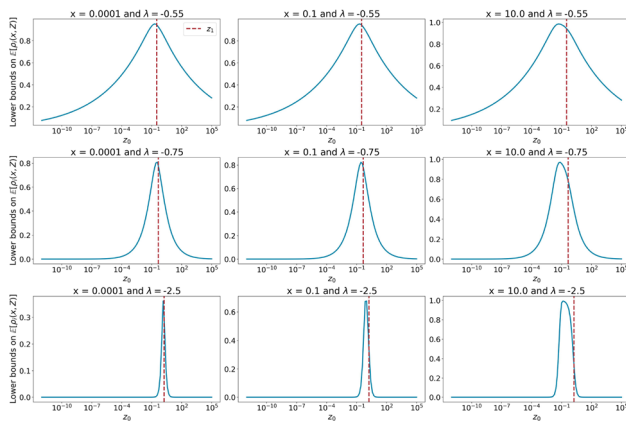
Fig. 4 Plot of the lower bounds for N_1 in Eq. (26) as a function of $z_0 \in [10^{-12}, 10^5]$
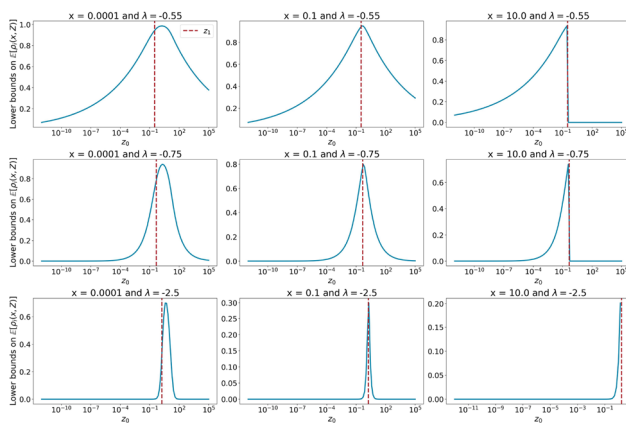

Fig. 5 Plot of the lower bounds for N_2 in Eq. (27) as a function of $z_0 \in [10^{-12}, 10^5]$

$$\frac{2}{\pi z_1 |H(z_1)|^2}$$

and we would postulate that this is a lower bound on the expected acceptance probability for any value of x (observing the apparent monotonic behaviour of the average acceptance probabilities in Fig. 3). A more detailed analysis beyond the current scope would study the expected acceptance rate as a function of the truncation level, see Eq.(20) and Figs. 5 and 6 from Godsill and Kındap (2021) for a possible approach.

4.3 The case of $0 < |\lambda| \leq 0.5$

This parameter range was covered in Godsill and Kındap (2021) and we review the basics here for completeness. We provide several improvements to this approach, including the more efficient sampling of point process N_1 as two gamma processes, improved bounds on the incomplete gamma functions, as well as the adaptive truncation method, residual approximation and the squeezed sampling methods detailed in subsequent sections.

The process associated with the upper bounding Lévy density $Q_{GIG}^B(x, z)$ for this parameter range, see (16a) and (16b), can be considered once again as a marked point process split into two independent point processes N_1 and N_2 having factorised intensity functions as given in Corollary 2 of Godsill and Kındap (2021) as

$$N_1 : \quad \frac{e^{-x\gamma^2/2}}{\pi^2 x^{1+|\lambda|}} \frac{(2\delta^2)^{|\lambda|} \gamma(|\lambda|, z_0^2 x/(2\delta^2))}{H_0 z_0^{2|\lambda|-1}}$$

$$\frac{\Gamma(|\lambda|)\sqrt{Ga}(z||\lambda|, x/(2\delta^2))}{\gamma(|\lambda|, z_0^2 x/(2\delta^2))} \mathbb{I}_{z<z_0}$$

$$= Q_{N_1}^B(x) Q_{N_1}^B(z|x)$$

$$N_2 : \quad \frac{e^{-x\gamma^2/2}}{\pi^2 x^{3/2}} \frac{(2\delta^2)^{0.5} \Gamma(0.5, z_0^2 x/(2\delta^2))}{H_0}$$

$$\frac{\Gamma(0.5)\sqrt{Ga}(z|0.5, x/(2\delta^2))}{\Gamma(0.5, z_0^2 x/(2\delta^2))} \mathbb{I}_{z \geq z_0}$$

$$= Q_{N_2}^B(x) Q_{N_2}^B(z|x)$$

with

$$Q_{N_1}^B(x) = \frac{e^{-x\gamma^2/2}}{\pi^2 x^{1+|\lambda|}} \frac{(2\delta^2)^{|\lambda|} \gamma(|\lambda|, z_0^2 x/(2\delta^2))}{H_0 z_0^{2|\lambda|-1}} \tag{28}$$

and

$$Q_{N_2}^B(x) = \frac{e^{-x\gamma^2/2}}{\pi^2 x^{3/2}} \frac{(2\delta^2)^{0.5} \Gamma(0.5, z_0^2 x/(2\delta^2))}{H_0} \tag{29}$$

Again, N_1 and N_2 correspond to a marginal modified tempered stable process for x and a conditional truncated $\sqrt{Ga}$ density for z. The upper and lower incomplete gamma functions in the marginal point process envelopes $Q_{N_1}^B(x)$ and $Q_{N_2}^B(x)$ require the use of dominating processes and thinning methods similar to Sect. 4.1.

Using the bound in Eq. (22) the density $Q_{N_1}^B(x)$ can be transformed into two independent gamma processes and the methodology is summarised in Algorithm 11. This simulation algorithm improves upon the method shown in Algorithm 4 of Godsill and Kındap (2021) by transforming the problem of simulating a tempered stable process into that of simulating two gamma processes, which are known to converge rapidly in terms of the number of points required in Eq. (5). The convergence of these sums are discussed further in Sect. 5. Having simulated points x_i from the marginal density $Q_{N_1}^B(x)$, z_i values are simulated from a right-truncated square-root gamma density and the accept-reject step for each x_i is performed as shown in Algorithm 12.

For $Q_{N_2}^B(x)$, the incomplete gamma function is upper bounded by the complete gamma function to produce a process that is tractable for simulation and the corresponding

---

Algorithm 11 Sampling from $Q_{N_1}^B(x)$.

1. $N = \emptyset$,
2. Generate a gamma process N_{Ga}^1 having parameters $a_1 = \frac{z_0}{\pi^2 H_0 |\lambda|(1+|\lambda|)}$ and $\beta_1 = \gamma^2/2$,
3. Generate a gamma process N_{Ga}^2 having parameters $a_2 = \frac{z_0}{\pi^2 H_0(1+|\lambda|)}$ and $\beta_2 = \gamma^2/2 + z_0^2/(2\delta^2)$,
4. $N = N_{Ga}^1 \cup N_{Ga}^2$,
5. For each $x \in N$ accept with probability

$$\frac{(2\delta^2)^{|\lambda|} \gamma(|\lambda|, (z_0^2 x)/(2\delta^2))|\lambda|(1+|\lambda|)}{x^{|\lambda|} z_0^{2|\lambda|}(1+|\lambda| \exp\left(-z_0^2 x/(2\delta^2)\right))}$$

otherwise reject and delete x from N.

Algorithm 12 Generation of N_1 for $0 < |\lambda| \leq 0.5$

1. $N_1 = \emptyset$,
2. Simulate x_i from the marginal point process associated with $Q_{N_1}^B(x)$ as shown in Algorithm 11,
3. For each x_i, simulate a z_i from a truncated square-root gamma density

$$\frac{\Gamma(|\lambda|)\sqrt{\mathrm{Ga}}(z||\lambda|, x_i/(2\delta^2))}{\gamma(|\lambda|, z_0^2 x_i/(2\delta^2))} \mathcal{I}_{0 < z < z_0}$$

4. With probability

$$\frac{Q_{GIG}(x, z)}{Q_{N_1}^B(x, z)} = \frac{H_0}{|H_{|\lambda|}(z_i)|^2 \left(\frac{z_i^{2|\lambda|}}{z_0^{2|\lambda|-1}}\right)}$$

accept x_i, i.e. set $N_1 = N_1 \cup x_i$, otherwise discard x_i.

---

algorithm is given in Algorithm 13. For each x_i value simulated from the marginal density, a corresponding z value is sampled from the conditional left-truncated square-root gamma density and the whole methodology is outlined in Algorithm 14.

Note that the bound defined in Eq. (24) could also be used in this parameter regime to obtain a more efficient sampling algorithm; however, the acceptance rates using the simpler bound are found to work well.

Algorithm 13 Sampling from $Q_{N_2}^B(x)$.

1. Generate a tempered stable process N_{MTS} with parameters $C = \frac{(2\delta^2)^{0.5}\Gamma(0.5)}{\pi^2 H_0}$, $\alpha = 0.5$ and $\beta = \gamma^2/2$,
2. For each point $x \in N_{MTS}$, accept with probability $\Gamma(0.5, z_0^2 x/(2\delta^2))/\Gamma(0.5)$, otherwise reject and delete x from N_{MTS}.

Lastly, the set of points $N = N_1 \cup N_2$ is a realisation of jump magnitudes corresponding to a GIG process having intensity function $Q_{GIG}(x)$ and once again the corresponding GH process may be obtained using Algorithm 6.

Algorithm 14 Generation of N_2 for $0 < |\lambda| \leq 0.5$

1. $N_2 = \emptyset$,
2. Simulate x_i from the marginal point process associated with $Q_{N_2}^B(x)$ as shown in Algorithm 13,
3. For each x_i, simulate a z_i from a truncated square-root gamma density

$$\frac{\Gamma(0.5)\sqrt{\mathrm{Ga}}(z|0.5, x_i/(2\delta^2))}{\Gamma(0.5, z_0^2 x_i/(2\delta^2))} \mathcal{I}_{z \geq z_0}$$

4. With probability

$$\frac{Q_{GIG}(x, z)}{Q_{N_2}^B(x, z)} = \frac{H_0}{z_i |H_{|\lambda|}(z_i)|^2}$$

accept x_i, i.e. set $N_2 = N_2 \cup x_i$, otherwise discard x_i.

---

5 Adaptive truncation and Gaussian approximation of residuals

The shot noise methods discussed in previous sections involve infinite series of decreasing random variables and in practice must be truncated after a finite number of terms. In this section we propose novel methods for adaptive determination of the number of terms required in the truncated series. This adaptive truncation can both save substantially on the computational burden of generating very long shot noise series, and also ensure (probabilistically) a specified error tolerance. Furthermore, we provide lower bounds on the mean and variance of the GIG and GH residual sequences which will be used in order to approximate the residual term as Brownian motion once the adaptive truncation procedure is terminated.

The adaptive truncation and residual approximation methods studied in this section are designed to match the moments of a realisation from a GH process at time t to its theoretical moments. For a subordinator Lévy process $X(t)$ with Lévy measure $Q(dx)$ and finite first and second moments, the mean and variance of the subordinator may be expressed as Barndorff-Nielsen and Shephard (2012)

$$\mathbb{E}[X(t)] = \frac{t}{T} \int_0^\infty x Q(dx) \tag{30}$$

and

$$\mathrm{Var}[X(t)] = \frac{t}{T} \int_0^\infty x^2 Q(dx) \tag{31}$$

where the distribution of the process at time T, denoted as $X(T)$, defines the associated random variable. These integrals are intractable for the GIG case. Hence upper and lower bounds on these integrals are studied.

For normal variance-mean processes $W(t)$, such as the GH process, the associated Lévy measure can be expressed

as a function of the subordinator Lévy measure $Q_{GIG}(dx)$ as in (4). Hence the mean and variance of the GH process can be found in terms of the moments of a GIG process as

$$\mathbb{E}[W(t)] = \frac{t}{T}\beta\mathbb{E}[X(T)] \tag{32}$$

and

$$\mathrm{Var}[W(t)] = \frac{t}{T}\left(\beta^2\mathrm{Var}[X(T)] + \sigma^2\mathbb{E}[X(T)]\right) \tag{33}$$

where β and σ are the skewness and scale parameters as defined in (17).

In Sect. 5.1 we study adaptive truncation of infinite series for subordinator processes and provide associated algorithms for the GIG case. In principle the adaptive truncation scheme can also be described in terms of the GH moments using (32) and (33). This transformation to the GH moments is explicitly required for the Gaussian approximation of residual moments of the GH process and lower bounds on these residuals are studied in Sect. 5.2. The Gaussian approximation we use is motivated by proofs of the convergence of normal variance-mean mixture residuals to a Brownian motion in all cases of the GH process except for the normal-gamma process, and these results will be presented in a forthcoming publication.

5.1 Adaptive truncation of shot noise series

The shot noise series for a subordinator $X(t)$ with its jumps truncated at ε may be defined as

$$X^\varepsilon(t) = \sum_{\{i:x_i \geq \varepsilon\}} x_i\mathbb{I}_{V_i \leq t} \tag{34}$$

The difference of $X(t)$ and the truncated series $X^\varepsilon(t)$ characterises the residual error caused by truncation and may be expressed as a random process $R^\varepsilon(t)$ such that

$$
\begin{aligned}
R^\varepsilon(t) &= X(t) - X^\varepsilon(t) \\
&= \sum_{\{i:x_i < \varepsilon\}} x_i\mathbb{I}_{V_i \leq t}
\end{aligned}
\tag{35}
$$

where ε is the value at which the jump magnitude sequence $\{x_i\}$ are stochastically truncated (i.e. the truncated series X^ε has a random number of terms with x_i greater than or equal to ε).

The statistical properties of the residual error $R^\varepsilon(t)$ as a function of ε can be used to study the convergence of the truncated series to the Lévy process $X(t)$. The number of terms used in the approximation of $X(t)$ may be dynamically adjusted depending on the particular realisations of $\{x_i\}$ and the required precision of approximation. Theorem 3

below and its Corollary describes the construction of a probabilistic bound on the residual error caused by truncation of a subordinator process in terms of upper bounds on its residual moments and provide the residual mean and variance for the tempered stable and gamma processes which are used as dominating processes for sampling the GIG process.

Note that the bound in the Theorem below is a point-wise bound at a particular time t, whereas ideally a *pathwise* bound might be desired that applies across all times. Martingale inequalities can be used in principle to achieve this, see Wolpert (2021), although these may not be directly applicable here as we do not in general have an exact characterisation of the residual mean and variance for the GIG process, only upper and lower bounds on these.

Theorem 3 *For the residual error $R^\varepsilon(t)$ associated with truncation of a subordinator process, the following probabilistic bound applies for any $E > \bar{\mu}_\varepsilon$ and truncation level $\varepsilon > 0$:*

$$Pr\left(R^\varepsilon(t) \geq E | X^\varepsilon(t)\right) \leq \frac{\bar{\sigma}_\varepsilon^2}{(E - \bar{\mu}_\varepsilon)^2} \tag{36}$$

where $\bar{\mu}_\varepsilon \geq \mu_\varepsilon$ and $\bar{\sigma}_\varepsilon \geq \sigma_\varepsilon$ are upper bounds on $\mu_\varepsilon = \mathbb{E}[R^\varepsilon(t)]$ and $\sigma_\varepsilon^2 = var(R^\varepsilon(t))$, and E is a threshold that may depend on the random realisation $X^\varepsilon(t)$.

Proof The mean and variance of a subordinator process $X(t)$ are given in (30) and (31). Similarly, the mean and variance of the truncated process residual $R^\varepsilon(T)$ can be found as

$$\mu_\varepsilon = \mathbb{E}[R^\varepsilon(t)] = \frac{t}{T}\int_0^\varepsilon x\,Q(dx) \tag{37}$$

and

$$\sigma_\varepsilon^2 = \mathrm{Var}(R^\varepsilon(t)) = \frac{t}{T}\int_0^\varepsilon x^2\,Q(dx) \tag{38}$$

where both of these integrals are well-defined and finite for any valid subordinator and $0 < \varepsilon < \infty$, by condition (3).

Now, using the expected value and standard deviation of $R^\varepsilon(t)$ we may bound the residual error using concentration inequalities. Specifically, Chebyshev's inequality states that for a random variable $R^\varepsilon(t)$ with finite expected value μ_ε and finite non-zero variance σ_ε^2

$$Pr\left(|R^\varepsilon(t) - \mu_\varepsilon| \geq k\sigma_\varepsilon\right) \leq \frac{1}{k^2}$$

where the probability is conditional on a random realisation of $X^\varepsilon(t)$. We require here only the right tail probability mass corresponding to the event $R^\varepsilon(t) - \mu_\varepsilon \geq k\sigma_\varepsilon$, and this is clearly less than or equal to the probability of the event

$|R^\varepsilon(t) - \mu_\varepsilon| \geq k\sigma_\varepsilon$; hence rearranging we arrive at

$$\Pr\left(R^\varepsilon(t) \geq \mu_\varepsilon + k\sigma_\varepsilon\right) \leq \frac{1}{k^2}$$

Now, if we have instead upper bounds $\mu_\varepsilon \leq \bar{\mu}_\varepsilon$ and $\sigma_\varepsilon \leq \bar{\sigma}_\varepsilon$, it is clear that $\mu_\varepsilon + k\sigma_\varepsilon \leq \bar{\mu}_\varepsilon + k\bar{\sigma}_\varepsilon$ and so

$$\Pr\left(R^\varepsilon(t) \geq \bar{\mu}_\varepsilon + k\bar{\sigma}_\varepsilon\right) \leq \frac{1}{k^2} \tag{39}$$

Finally, a simple rearrangement with $E = \bar{\mu}_\varepsilon + k\bar{\sigma}_\varepsilon$ leads to the theorem as stated. □

Corollary 4 *Our simulation algorithms for the GIG process involve thinning/rejection sampling operations in order to generate point processes N_1 and N_2 from gamma and tempered stable dominating processes, see Algorithms* 7, 9, 11 *and* 13. *By construction, the resulting thinned processes have Lévy density $Q(x)$ strictly less than or equal to that of the dominating process in each case, say $Q^0(x)$, i.e. $Q(x) \leq Q^0(x)$. Hence the means and variances of the truncation error, calculated using* (37) *and* (38), *are strictly less than or equal to those of the corresponding dominating gamma and tempered stable processes, and we may thus take the means and variances of the underlying TS or gamma processes as the upper bounds $\bar{\mu}_\varepsilon$ and $\bar{\sigma}_\varepsilon$ required in Theorem* 3.

For the TS process the expected value and variance of the residual process is found as

$$\mu_{TS}(t) = \frac{t}{T}\int_0^\varepsilon C x^{-\alpha} e^{-\beta x} dx$$
$$= \frac{tC\beta^{\alpha-1}}{T}\gamma(1-\alpha, \beta\varepsilon) \tag{40}$$

and

$$\sigma_{TS}^2(t) = \frac{t}{T}\int_0^\varepsilon C x^{1-\alpha} e^{-\beta x} dx$$
$$= \frac{tC\beta^{\alpha-2}}{T}\gamma(2-\alpha, \beta\varepsilon) \tag{41}$$

In limit as $\beta \to 0$ we obtain the stable subordinator whose moments can be obtained either directly or as limits of the above TS case using $\frac{\gamma(s,x)}{x^s} \to \frac{1}{s}$ as $x \to 0$, giving:

$$\mu_S(t) = \frac{tC\varepsilon^{1-\alpha}}{T(1-\alpha)} \tag{42}$$

and

$$\sigma_S^2(t) = \frac{tC\varepsilon^{2-\alpha}}{T(2-\alpha)} \tag{43}$$

Similarly for the gamma process the expected value and variance of the residual process can be found as

$$\mu_{Ga}(t) = \frac{t}{T}\int_0^\varepsilon C e^{-\beta x} dx$$
$$= \frac{tC}{T\beta}\gamma(1, \beta\varepsilon) \tag{44}$$

and

$$\sigma_{Ga}^2(t) = \frac{t}{T}\int_0^\varepsilon C x e^{-\beta x} dx$$
$$= \frac{tC}{T\beta^2}\gamma(2, \beta\varepsilon) \tag{45}$$

Take, for example, generation of N_2 in Algorithms 9 *and* 10. *The starting point is generation of a TS process with parameters $C = \frac{\delta}{\sqrt{2\pi}}$, $\alpha = 0.5$ and $\beta = \frac{z_1^2}{2\delta^2} + \frac{\gamma^2}{2}$, implemented using Algorithm* 3. *The mean and variance for the truncated residual of this process are obtained from* (40) *and* (41). *Algorithms* 9 *and* 10 *then perform random thinning on the TS points. Hence the resulting process N_2 has truncated residual with mean and variance no larger than those of the corresponding TS process. Thus Theorem* 3 *applies, using the TS mean and variance as the upper bounds $\bar{\mu}_\varepsilon$ and $\bar{\sigma}_\varepsilon^2$. The other cases of N_1 and N_2 simulation follow a similar argument, using the appropriate gamma or TS process to generate upper bounds on the moments required for Theorem* 3.

Finally, a probabilistic upper bound on the GIG residual may be obtained by adding the upper bounds on the means and variances for N_1 and N_2, since the two point processes are independent.

Corollary 5 *Improved residual errors and corresponding bounds on these are available if the mean μ_ϵ is available, as then the improved estimate $\hat{X}^\epsilon(t) = X^\epsilon(t) + \mu_\epsilon$ may be formed as proposed in* Asmussen and Rosiński (2001). *In our GH case however we only have upper and lower bounds $\underline{\mu}_\epsilon \leq \mu_\epsilon \leq \bar{\mu}_\epsilon$ established in Theorems* 3 *and* 4 *(see below). In this case it would seem appropriate to take a conservative line and substitute the lower bound $\underline{\mu}_\epsilon$ in place of μ_ϵ. Then Theorem* 3 *can be modified in step* (39) *as follows,*

$$Pr\left(|R^\varepsilon(t) - \underline{\mu}_\varepsilon| \geq \bar{\mu}_\varepsilon - \underline{\mu}_\varepsilon + k\bar{\sigma}_\varepsilon\right) \leq \frac{1}{k^2} \tag{46}$$

To justify this, use Chebyshev directly to give

$$Pr(|R^\varepsilon(t) - \mu_\varepsilon| \geq k\sigma_\varepsilon) \leq 1/k^2$$

But we have

$$A := \{|R^\varepsilon(t) - \underline{\mu}_\varepsilon| \geq \bar{\mu}_\varepsilon - \underline{\mu}_\varepsilon + k\sigma_\varepsilon\}$$
$$\subseteq \{|R^\varepsilon(t) - \mu_\varepsilon| \geq k\sigma_\varepsilon\} := B$$

and hence $Pr(A) \leq Pr(B)$ from which (46) *follows.*

Rearranging this expression with $E = \bar{\mu}_\varepsilon - \underline{\mu}_\varepsilon + k\bar{\sigma}_\varepsilon$ a new expression is obtained, valid for $E + \underline{\mu}_\varepsilon - \bar{\mu}_\varepsilon > 0$:

$$Pr\left(|R^\varepsilon(t) - \underline{\mu}_\varepsilon| \geq E\right) \leq \frac{\bar{\sigma}_\varepsilon^2}{(E + \underline{\mu}_\varepsilon - \bar{\mu}_\varepsilon)^2}$$

This expression would then be recommended for practical use in adaptive truncation schemes, yielding always smaller probabilities of exceedance than Theorem 3 *for fixed threshold E since $E + \underline{\mu}_\epsilon - \bar{\mu}_\varepsilon \geq E - \bar{\mu}_\varepsilon$.*

Using the probabilistic bounds given in Theorem 3 and its Corollaries, an adaptive truncation scheme can be devised to determine a suitable value for ε for *each* generated realisation of the process. As ε decreases, so we accumulate sequentially the realised value of $X^\varepsilon(t)$ (or its mean-adjusted version from Corollary 5) according to Eq. (34). A tolerance $E = \tau X^\varepsilon(t)$ is chosen, where $0 < \tau \ll 1$, which is designed to truncate the series once the predicted residual has become very small in comparison with the series realised to level ε. Then a probability threshold, $p_T \ll 1$ is chosen for comparison with $Pr(R^\varepsilon(t) \geq E)$, in order to decide when to terminate the simulation. A generic adaptive truncation scheme is outlined in Algorithm 15 for a point process N associated with a subordinator Lévy process $X(t)$ having Lévy density $Q(x)$.

Algorithm 15 Simulation of N with adaptive truncation using tolerance τ and probability threshold p_T for a point process with Lévy density $Q()$ having moment bounds $\bar{\mu}_{\varepsilon_n}$ and $\bar{\sigma}_{\varepsilon_n}$.

---

1. Define a decreasing truncation schedule $\varepsilon_0 = \infty > \varepsilon_1 > \varepsilon_2 > \cdots > \varepsilon_n > \ldots$ and initialise $E_0 = 0$, $N = \emptyset$. Set $n = 1$, $B =$ False.
2. While $B =$ False,
   - If $\frac{\bar{\sigma}_{\varepsilon_n}^2}{(\tau E_n - \bar{\mu}_{\varepsilon_n})^2} > p_T$
     Simulate points $\{x_i\}$ from $Q(x)\mathcal{I}_{x \in (\varepsilon_n, \varepsilon_{n-1}]}$ using Algorithms 1 and 2 (with $a = \varepsilon_n$ and $b = \varepsilon_{n-1}$),
     Set $N = N \cup \{x_i\}$,
     Set $E_n = E_{n-1} + \sum_i x_i$,
   - Else $B =$ True.
   - $n = n + 1$.
3. Return N.

The GH simulation algorithms studied in this work and Godsill and Kındap (2021) are made up of two independent point processes N_1 and N_2. An adaptive truncation algorithm such as Algorithm 15 can be applied separately to each process to obtain the resulting jumps. This misses a trick however, since either series could in principle be truncated even earlier once its residual error is very small relative to the accumulated sum of both N_1 and N_2. There are many

---

Algorithm 16 Simulation of $N = \cup_{k=1}^K N_k$ with Lévy density $Q(x) = \sum_{k=1}^K Q_k(x)$, $x > 0$, having moment bounds $\bar{\mu}_{\varepsilon_n}^k$ and $\bar{\sigma}_{\varepsilon_n}^k$ with adaptive truncation using tolerance τ and probability threshold p_T.

---

1. Define a decreasing truncation schedule $\varepsilon_0 = \infty > \varepsilon_1 > \varepsilon_2 > \ldots > \varepsilon_n > \ldots$ and initialise $E = 0$, $N_k = \emptyset$. Set $n = 1$, $B_k =$ False, $B =$ False.
2. While $B =$ False
   - For $k = 1$ to K,
     If $\frac{\bar{\sigma}_{\varepsilon_n}^{k2}}{(\tau E_n - \bar{\mu}_{\varepsilon_n}^k)^2} > p_T$
       Simulate points $\{x_i\}$ from $Q_k(x)\mathcal{I}_{x \in (\varepsilon_n, \varepsilon_{n-1}]}$ using e.g. Algorithms 1 and 2 (with $a = \varepsilon_n$ and $b = \varepsilon_{n-1}$)
       Set $N = N \cup \{x_i\}$.
       Set $E = E + \sum_i x_i$
     Else $B_k =$ True.
   - $B = \wedge_k B_k$, Logical AND operation
   - $n = n + 1$
3. Return N.

---

workable schemes based around this idea and one possible such approach is presented in Algorithm 16. It is presented in a general form that can apply to the parallel simulation and adaptive truncation of K independent subordinator point processes N_k having Lévy densities $Q_k(x)$ and overall Lévy density $Q(x) = \sum_{k=1}^K Q_k(x)$.

For N_1 in the most general settings of Algorithm 8 and 12, the dominating process is made up of two independent gamma processes N_{Ga}^1 and N_{Ga}^2. In the case of N_2 in both settings, Algorithms 10 and 14, a single dominating tempered stable process is required. Hence an efficient method of simulation is running Algorithm 16 on these 3 independent dominating processes. The residual means and variances of the tempered stable and gamma processes required by Algorithm 15 and Algorithm 16 are shown in Corollary 4. It is worth noting that the convergence of a gamma process is typically significantly faster than a tempered stable process and so the N_1 process tends to terminate much sooner than N_2.

Remark 3 In the edge parameter setting $\lambda < 0$ and $\gamma = 0$, the marginal point process simulation methods shown in Algorithms 7 and 11 are not valid as a result of N_{Ga}^1 becoming undefined for $\gamma = 0$. For this setting, Alg. 3 of Godsill and Kındap (2021) may be used together with the adaptive truncation and residual approximation methods introduced in this section.

For this parameter setting the tempered stable process defined in Alg. 3 of Godsill and Kındap (2021) becomes a stable process since the tempering parameter β is equal to 0. The associated residual moments of a stable process are presented in Corollary 4 which should be used to implement Step 2) of Algorithm 15.

5.2 Gaussian approximation of residual errors

Here we present a Brownian motion approximation method for the residual error $R^\varepsilon(t)$ of a GIG or GH process caused by the truncation of a shot noise series as defined in Eq. (35). Such an approach is well known from previous work, see e.g. Asmussen and Rosiński (2001), but here we propose an intermediate solution in which a Brownian motion is injected whose drift and variance are lower bounds compared with the exact result, which is intractable in general for the GIG and GH processes.

The theoretical mean and variance of the residual error for the GH case can be found as a function of the mean and variance of an associated GIG residual error $R^\varepsilon(t)$ using Eqs. (32) and (33). Hence similar to Sect. 5.1, we provide lower bounds $\underline{\mu_\varepsilon}, \underline{\sigma_\varepsilon^2}$ on the mean and variance of a residual GIG process as a function of the truncation level ε. Together with the upper bounds discussed in Corollary 4, these lower bounds characterise the residual error of truncating the infinite shot noise series for the GIG and GH cases. Using the residual approximation module the series representation of the GIG Lévy process $X(t)$ can be expressed as

$$X(t) \approx \frac{t\underline{\mu_\varepsilon}}{T} + \frac{\underline{\sigma_\varepsilon}}{\sqrt{T}} B(t) + X^\varepsilon(t)$$

where $X^\varepsilon(t)$ is computed in the usual way as $\sum_{\{i:x_i \geq \varepsilon\}} x_i \mathbb{I}_{V_i \leq t}$, $B(t)$ is an independent standard Brownian motion term and $\underline{\mu_\varepsilon}, \underline{\sigma_\varepsilon^2}$ are lower bounds on the mean and variance of the residual error $R^\varepsilon(T)$ given a truncation level ε. Note that this approximation can technically become negative because the Brownian motion term is unconstrained, which is undesirable for the positive-valued process $X(t)$. This effect will however become negligible as $\varepsilon \to 0$ in the GIG case, see convergence results in Asmussen and Rosiński (2001), but in any case we show this result only for completeness and the approximation we implement is for $W(t)$ itself, which is permitted to become negative or positive, see Eq. (49) below.

According to Eqs. (32) and (33), the lower bounds on the moments of the residual error $R^\varepsilon_W(t)$ of the associated GH process $W(t)$ can be obtained as

$$\mathbb{E}[R^\varepsilon_W(t)] \geq \frac{t}{T} \beta \underline{\mu_\varepsilon} \tag{47}$$

$$\mathrm{Var}[R^\varepsilon_W(t)] \geq \frac{t}{T} \left(\beta^2 \underline{\sigma_\varepsilon^2} + \sigma^2 \underline{\mu_\varepsilon} \right) \tag{48}$$

Hence for the GH process the same procedure is adopted to obtain an approximation as

$$W(t) \approx \frac{t}{T} \beta \underline{\mu_\varepsilon} + \frac{\sqrt{\beta^2 \underline{\sigma_\varepsilon^2} + \sigma^2 \underline{\mu_\varepsilon}}}{\sqrt{T}} B(t) + W^\varepsilon(t) \tag{49}$$

The approximation of the residual error is in addition to the adaptive truncation methods described in Sect. 5.1 that provide a specified level of truncation ε for the simulation of sample paths from a GH process. In Theorem 4 below we provide the required lower bounds on the mean and variance of a residual GIG process with Lévy density $Q_{GIG}(x)$ as a function of ε. The derivation of these bounds are presented for a specific T as it is straightforward to scale the moments according to time. The lower bounds are then used to evaluate the mean and variance of the residual error in the GH process simulation and approximate the contribution of the residual small jumps according to Eq. (49). A Brownian motion approximation to the residual is known to hold for many shot noise series, see Asmussen and Rosiński (2001), with the gamma process being a well-known exception that does not converge to a Gaussian as $\varepsilon \to 0$. In our own work we have proven convergence of the shot noise series for the GH process to a Brownian motion in all cases except the normal-gamma, and these results will be presented in a future publication.

Theorem 4 *Given a truncation level ε, a residual sequence $R^\varepsilon(T)$ of GIG jumps with mean $\mu_\varepsilon = E[R^\varepsilon(T)]$ and variance $\sigma_\varepsilon^2 = Var[R^\varepsilon(T)]$ may be lower bounded by:*

$$\underline{\mu_\varepsilon} \leq \mu_\varepsilon, \quad \underline{\sigma_\varepsilon^2} \leq \sigma_\varepsilon^2$$

where the bounds are defined as:

$$\underline{\mu_\varepsilon} = \begin{cases} \dfrac{C_{Ga}^B \gamma(1, \beta_{Ga}^B \varepsilon)}{\beta_{Ga}^B} + \dfrac{C_{TS}^B \gamma(0.5, \beta_{TS}^B \varepsilon)}{\beta_{TS}^{B\ 0.5}}, & |\lambda| \geq 0.5 \\[3mm] \dfrac{C_{Ga}^A \gamma(1, \beta_{Ga}^A \varepsilon)}{\beta_{Ga}^A} + \dfrac{C_{TS}^A \gamma(0.5, \beta_{TS}^A \varepsilon)}{\beta_{TS}^{A\ 0.5}}, & |\lambda| < 0.5 \end{cases}$$

$$\underline{\sigma_\varepsilon^2} = \begin{cases} \dfrac{C_{Ga}^B \gamma(2, \beta_{Ga}^B \varepsilon)}{\beta_{Ga}^{B\ 2}} + \dfrac{C_{TS}^B \gamma(1.5, \beta_{TS}^B \varepsilon)}{\beta_{TS}^{B\ 1.5}}, & |\lambda| \geq 0.5 \\[3mm] \dfrac{C_{Ga}^A \gamma(2, \beta_{Ga}^A \varepsilon)}{\beta_{Ga}^{A\ 2}} + \dfrac{C_{TS}^A \gamma(1.5, \beta_{TS}^A \varepsilon)}{\beta_{TS}^{A\ 1.5}}, & |\lambda| < 0.5 \end{cases}$$

where

$$C_{Ga}^A = \frac{z_1}{2\pi|\lambda|} \quad and \quad \beta_{Ga}^A = \frac{\gamma^2}{2} + \frac{|\lambda|}{(1+|\lambda|)} \frac{z_1^2}{2\delta^2}$$

$$C_{Ga}^B = \frac{z_0}{\pi^2 H_0 |\lambda|} \quad and \quad \beta_{Ga}^B = \frac{\gamma^2}{2} + \frac{|\lambda|}{(1+|\lambda|)} \frac{z_0^2}{2\delta^2}$$

$$C_{TS}^A = \frac{\delta\sqrt{e}\sqrt{\beta_0 - 1}}{\pi\beta_0} \quad and \quad \beta_{TS}^A = \frac{\gamma^2}{2} + \frac{\beta_0 z_1^2}{2\delta^2}$$

$$C_{TS}^B = \frac{2\delta\sqrt{e}\sqrt{\beta_0 - 1}}{\pi^2 H_0 \beta_0} \quad and \quad \beta_{TS}^B = \frac{\gamma^2}{2} + \frac{\beta_0 z_0^2}{2\delta^2}$$

and β_0 is a free parameter such that $\beta_0 > 1$.

The proof for Theorem 4 can be found in the Appendix. Note that the lower bounds presented in Theorem 4 are valid only for the $\lambda < 0$ setting. The additional mean and variance components present for $\lambda > 0$ are associated with a gamma process with Lévy density (9) and residual moments given by Eqs. (44) and (45). These terms are simply added to the lower bounds $\underline{\mu_\varepsilon}$ and $\underline{\sigma_\varepsilon^2}$ in order to obtain the final bounds on the GIG process.

Remark 4 For the special case of $|\lambda| = 0.5$, the GIG density $Q_{GIG}(x)$ has a functional form equivalent to a TS process with intensity function

$$\frac{e^{-x\gamma^2/2}}{x^{3/2}} \frac{\delta \Gamma(1/2)}{\sqrt{2\pi}}$$

Thus both the simulation algorithms in Sect. 4, and the adaptive truncation and residual approximation methods presented in this Section are significantly simplified. It is straightforward to simulate a TS process using Algorithm 3 and the residual moments shown in Eqs. (40) and (41) provide exact expressions for the residual moments of the equivalent GIG process.

The true mean and variance of the residual moments replace the upper and lower bounds required for determining the adaptive truncation level ε and the associated moments of the approximating Brownian term. In this case Eq. (49) exactly matches the first and second moments of the residual process.

6 Squeezed rejection sampling

In this section, we provide a practical extension to the sampling algorithms discussed in Sect. 4 that is designed to increase the efficiency of simulation. The above methods for sampling of N_1 and N_2 in both parameter settings, given in Algorithms 8, 10 and 12, 14, involve a computationally expensive pair of steps in the sampling of a truncated gamma random variate (Step 3) and a pointwise evaluation of the Hankel function (Step 4). Notice however, that Theorem 1 provides both lower *and* upper bounds on the term $z|H_\nu(z)|^2$, which allows us to specify *squeezing* functions on $Q_{GIG}(x, z)$,[5] as given in (15a)–(16b). This allows for a

labour-saving *retrospective sampling* procedure in which, for a fixed fraction of points x_i, we may replace the simulation of a conditional random variable z and rejection sampling based on its value (Steps 3 and 4) with a simple 1-step accept/reject and no requirement to sample z or evaluate $H_{|\lambda|}$. Considering first the case $|\lambda| \geq 0.5$ and the sampling of process N_1, see Algorithm 8, we have generated at Step 2 a single point realisation x_i from the process $Q_{N_1}^A(x)$. Now consider Step 4, which accepts x_i with probability

$$\frac{Q_{GIG}(x_i, z_i)}{Q_{N_1}^A(x_i, z_i)} \geq \frac{Q_{GIG}^B(x_i, z_i)}{Q_{GIG}^A(x_i, z_i)} = \frac{2}{\pi H_0}$$

where we have used the squeezing inequality (13) and where the final equality applies when we set $z_0 = z_1$, see Remark 2. This implies that we may carry out a retrospective sampling step: draw a uniform random variate W_i on $[0, 1]$, test whether it is less than or equal to $\frac{2}{\pi H_0}$, and if so, accept x_i with no Steps 3 and 4 required. If $W_i > \frac{2}{\pi H_0}$, carry out steps 3 and 4, using the same realised W_i to carry out the test in Step 4.[6] The modified version of Algorithm 8 using squeezed rejection sampling is presented in Algorithm 17.

Algorithm 17 Squeezed generation of N_1 for $|\lambda| \geq 0.5$

1. $N_1 = \emptyset$,
2. Simulate x_i from the marginal point process associated with $Q_{N_1}^A(x)$ as given in Algorithm 7,
3. Generate $w_i \sim \mathcal{U}[0, 1]$,
4. If $w_i \leq \frac{2}{\pi H_0}$, accept x_i, i.e. set $N_1 = N_1 \cup x_i$,
   Else,

   (a) For each rejected x_i, simulate a z_i from a truncated square-root gamma density

   $$\frac{\Gamma(|\lambda|)\sqrt{\text{Ga}}(z||\lambda|, x_i/(2\delta^2))}{\gamma(|\lambda|, z_1^2 x_i/(2\delta^2))} \mathcal{I}_{0 < z < z_1}$$

   (b) If

   $$w_i \leq \frac{Q_{GIG}(x_i, z_i)}{Q_{N_1}^A(x_i, z_i)} = \frac{2}{\pi |H_{|\lambda|}(z_i)|^2 \left(\frac{z_i^{2|\lambda|}}{z_1^{2|\lambda|-1}}\right)}$$

   accept x_i, i.e. set $N_1 = N_1 \cup x_i$; otherwise discard x_i.

An exactly similar modification applies for generation of N_2 in Algorithm 10 for $|\lambda| > 0.5$. In the other parameter range $|\lambda| < 0.5$, the bounds are reversed, see (14) and so Step 4 in the squeezed sampler is replaced with 'if $w_i \leq \frac{\pi H_0}{2}$'; otherwise squeezed procedures for N_1 and N_2 in this parameter range are modified exactly as in Algorithm 17.

---

[5] see Devroye (1986) Section 3.6 for application of the Squeeze principle in the setting of standard random variate generation.

[6] The method is termed 'retrospective' because we only need to generate z and test the exact acceptance probability following a much simpler initial accept/reject step.
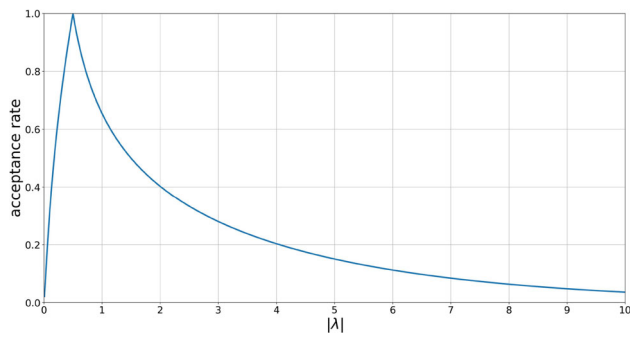
Fig. 6 Fractional computational saving for the retrospective sampling procedure, as a function of $|\lambda|$

We can see that the savings arising from this method could be substantial in cases where $\frac{2}{\pi H_0}$ is close to unity, saving almost all of the heavy computations in the original Steps 3 and 4. This will occur when $|\lambda|$ is close to 0.5, and improvements will lessen as $|\lambda|$ moves away from this value. The fraction of saved computation is shown as a function of $|\lambda|$ are shown in Fig. 6, exhibiting a broad range of λ values for which savings are useful.

7 Simulations

In this section we present results on the accuracy and efficiency of our proposed improvements to the simulation of a generalised hyperbolic process in Sect. 4 including the modifications discussed in Sect. 5. Particularly, we present the results of applying the Gaussian approximation of the residual process studied in Sect. 5.2 to our novel adaptive truncation method and compare these results to the method in Godsill and Kındap (2021). Additionally, we use a QQ (quantile-quantile) plot to compare the marginal distribution of randomly sampled GH processes generated up to $T = 1$ against exact samples from the GH distribution generated using random variable samplers as in Devroye (2014), Statovic (2017)). Note that while these methods are able to generate samples for a specific $t = T$, our method is able to generate the entire path of the process up to $t = T$ and hence information about the dynamics within the interval $(0, T)$ is made available. Furthermore, we show histograms and continuous-time sample paths for multiple parameter settings including special cases as the normal-inverse Gaussian and Student-t processes.

Since there is no known closed-form cumulative distribution function (CDF) for the GH distribution, the accuracy of the simulated sample paths are measured using the two-sample Kolmogorov-Smirnov (KS) test which compares the empirical distribution functions of the sample paths at $T = 1$ and exact samples from the GH distribution. These tests involve 10^6 independent samples for each method and param-

Table 1 The results of two-sample KS tests for adaptive simulation algorithm (using Algorithms 16)

λ	KS statistic	Time/sample	τ
-0.4	0.00301	8.27×10^{-4}	0.01
-0.4	0.00442	7.77×10^{-4}	0.1
-0.8	0.00305	1.34×10^{-3}	0.01
-0.8	0.00578	9.87×10^{-4}	0.1
-2.5	0.00369	3.20×10^{-3}	0.01
-2.5	0.00976	1.27×10^{-3}	0.1
-10	0.00485	9.28×10^{-3}	0.01
-10	0.01038	2.01×10^{-3}	0.1

No residual approximation, i.e. using $W^\varepsilon(t)$ to approximate $W(t)$

Table 2 The results of two-sample KS tests for adaptive simulation (Algorithm 16 plus Corollary 5) with residual approximation algorithm, using Eq. (49)

λ	KS statistic	Time/sample	τ
-0.4	0.00178	7.98×10^{-4}	0.01
-0.4	0.00246	6.57×10^{-4}	0.1
-0.8	0.00269	9.34×10^{-4}	0.01
-0.8	0.00179	7.89×10^{-4}	0.1
-2.5	0.00325	2.22×10^{-3}	0.01
-2.5	0.00487	9.71×10^{-4}	0.1
-10	0.00276	1.09×10^{-2}	0.01
-10	0.00835	1.72×10^{-3}	0.1

Table 3 The results of two-sample KS tests for the simulation algorithm in Godsill and Kındap (2021)

λ	KS statistic	Time/sample
-0.4	0.00292	1.72×10^{-2}
-0.8	0.00435	8.37×10^{-3}
-2.5	0.00494	1.09×10^{-2}
-10	0.00272	2.13×10^{-2}

eter setting and the results are shown in Table 1 for the simulation procedure using only adaptive determination of the number of jumps, in Table 2 for the improvements in convergence by approximating the residual small jumps. The results of the same test using the methods in Godsill and Kındap (2021) are repeated and shown in Table 3 for comparison.

For the proposed adaptive truncation algorithm in Algorithm 16, the probability threshold $p_T = 0.05$ is found to work well and is used for all cases throughout the section. The results in both Tables 1 and 2 show clear trade-offs between the accuracy of the distribution at $T = 1$ and the time required per sample for a given λ value. This can be observed by comparing the KS statistic and time per sam-
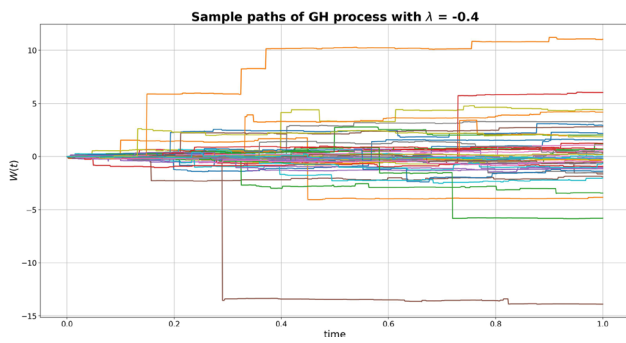
Fig. 7 Pathwise simulations of the GH process for $\lambda = -0.4$, $\delta = 1.0$, $\gamma = 0.1$ and $\beta = 0$
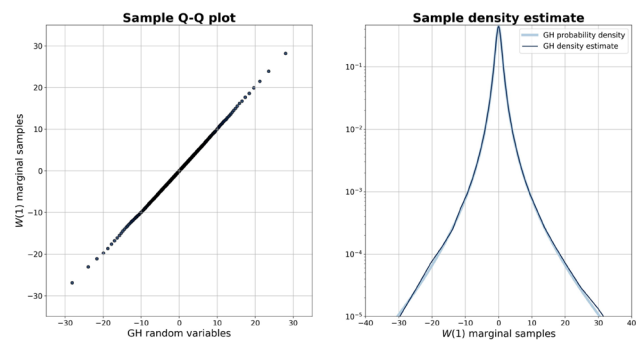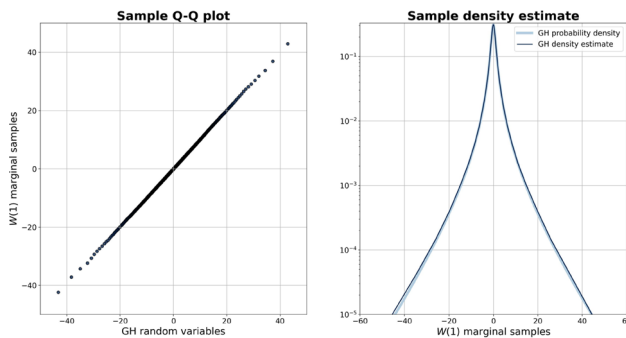

Fig. 8 Simulation comparison between the shot noise generated GH process and GH random variates, $\lambda = -0.4$, $\gamma = 0.1$, $\delta = 1$, $\beta = 0$. The adaptive truncation parameters are $p_T = 0.05$ and $\tau = 0.01$. Left hand panel: QQ plot comparing our shot noise method (y-axis) with random samples of the GH density generated using a random variate generator (x-axis). Right hand panel: Normalised histogram density estimate for our method compared with the true GH density function

Fig. 9 Pathwise simulations of the GH process for $\lambda = -0.8$, $\delta = 1.0$, $\gamma = 0.1$ and $\beta = 0$

Fig. 10 Simulation comparison between the shot noise generated GH process and GH random variates, $\lambda = -0.8$, $\gamma = 0.1$, $\delta = 1$, $\beta = 0$. The adaptive truncation parameters are $p_T = 0.05$ and $\tau = 0.01$. Left hand panel: QQ plot comparing our shot noise method (y-axis) with random samples of the GH density generated using a random variate generator (x-axis). Right hand panel: Normalised histogram density estimate for our method compared with the true GH density function

Comparing the results in Table 1, where no residual approximation is applied i.e. $W(t)$ is approximated as $W^{\varepsilon}(t)$, against the results in Table 2, where the residual part of the process is approximated as in Eq. (49) and the adaptive truncation procedure includes the adjustment in Corollary 5, it can be seen that there is a significant improvement in the KS statistic for all parameter settings. Particularly, the results suggest that the improvement in large tolerance parameter cases where $\tau = 0.1$ are substantial. Furthermore, no significant increase in time per sample can be observed when applying residual approximation methods. Note that to compare our new methods with the method studied in Godsill and Kındap (2021) producing a fixed number M of jumps, the adaptive truncation algorithm in Algorithm 16 is limited to producing a certain maximum number $M = 10^4$ of jump magnitudes per sample path while running the experiments. Comparing the results in Tables 2 and 3 it can be seen that improved convergence results can be obtained by the novel algorithms introduced in this work while also providing significantly reduced time complexity.

To present a more intuitive understanding of the accuracy of our methods, QQ plots of sample paths generated up to $T = 1$ and samples from a random variable generator are shown in addition to histograms with the true probability density function overlaid. The parameter values in the examples are selected to reflect the different characteristic behaviour of the GH process as well as edge cases such as the normal-inverse Gaussian process ($\lambda = -1/2$), and the Student-t process ($\gamma = 0$, $\lambda \leq 0$).

For the case $|\lambda| > 0.5$, Figs. 10, 12, 14 shows the QQ plot and histogram for different parameter settings and Figs. 9, 11, 13 presents sample paths from our new adaptive simulation algorithm with residual approximation. The corresponding processes are simulated using Algorithms 7, 8, 9 and 10. Similarly, Figs. 7, 8 presents an example in the $0 < |\lambda| <$
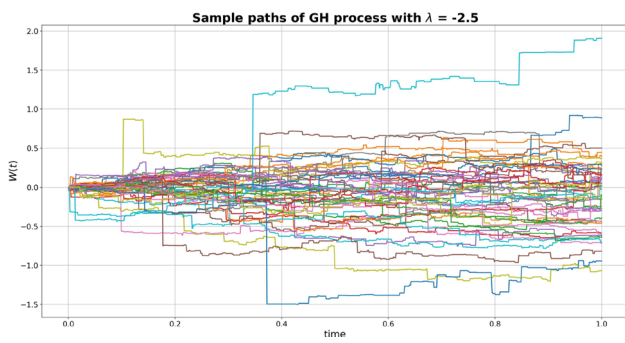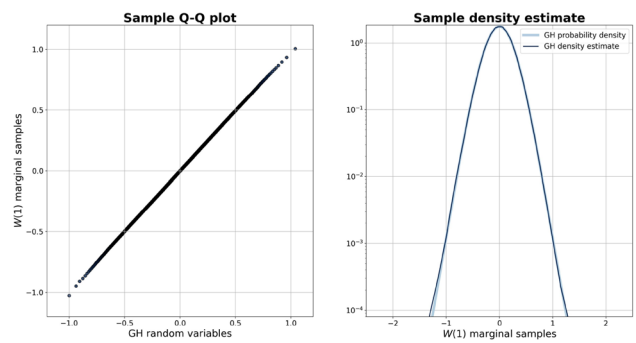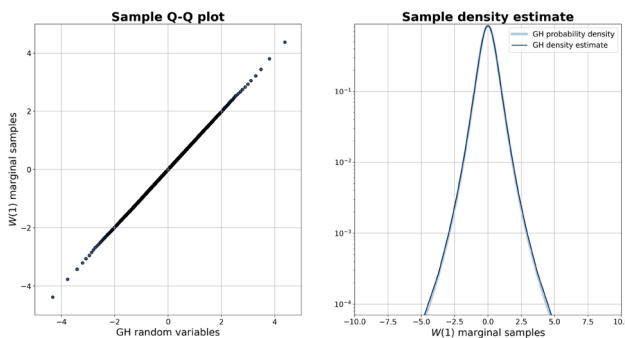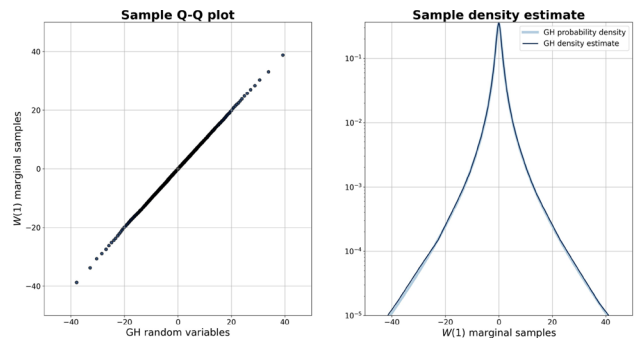
ple for different tolerance parameters τ. A large tolerance results in worse convergence but allows faster simulation, and hence the tolerance parameter may be adjusted depending on the requirements of the application. Furthermore as $|\lambda|$ increases both the time required per sample and the KS statistic increases as a result of the reduced acceptance rates of the simulation algorithm as plotted in Fig. 3.

Fig. 11 Pathwise simulations of the GH process for $\lambda = -2.5, \delta = 1.0,$ $\gamma = 0.1$ and $\beta = 0$

Fig. 12 Simulation comparison between the shot noise generated GH process and GH random variates, $\lambda = -2.5, \gamma = 0.1, \delta = 1, \beta = 0$. The adaptive truncation parameters are $p_T = 0.05$ and $\tau = 0.1$. Left hand panel: QQ plot comparing our shot noise method (y-axis) with random samples of the GH density generated using a random variate generator (x-axis). Right hand panel: Normalised histogram density estimate for our method compared with the true GH density function
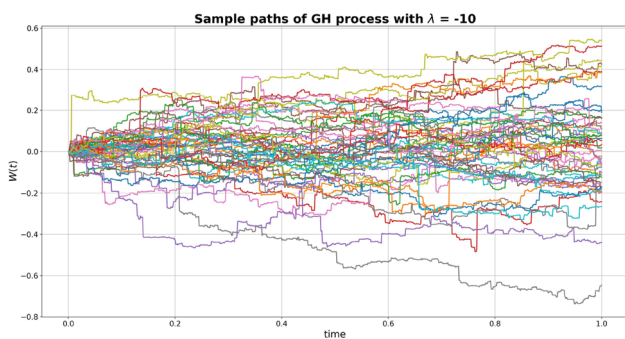
Fig. 13 Pathwise simulations of the GH process for $\lambda = -10, \delta = 1.0,$ $\gamma = 0.1$ and $\beta = 0$

Fig. 14 Simulation comparison between the shot noise generated GH process and GH random variates, $\lambda = -10, \gamma = 0.1, \delta = 1, \beta = 0$. The adaptive truncation parameters are $p_T = 0.05$ and $\tau = 0.1$. Left hand panel: QQ plot comparing our shot noise method (y-axis) with random samples of the GH density generated using a random variate generator (x-axis). Right hand panel: Normalised histogram density estimate for our method compared with the true GH density function

Fig. 15 Simulation comparison between the shot noise generated NIG process and NIG random variates, $\lambda = -0.5, \gamma = 0.1, \delta = 1, \beta = 0$. The adaptive truncation parameters are $p_T = 0.05$ and $\tau = 0.01$. Left hand panel: QQ plot comparing our shot noise method (y-axis) with random samples of the NIG density generated using a random variate generator (x-axis). Right hand panel: Normalised histogram density estimate for our method compared with the true NIG density function

0.5 setting and the corresponding simulation methods are described in Algorithms 11, 12, 13 and 14. The number of samples from each simulation method is $N = 10^6$ and the sample path plots show randomly selected 50 paths.

The normal-inverse Gaussian (NIG) distribution, or the distribution of NIG processes at $T = 1$, forms an exponential family and hence all of its moments have analytical expressions Barndorff-Nielsen (1997b). As a result of its tractable probabilistic properties, the NIG process finds application in modelling turbulence and financial data (Barndorff-Nielsen 1997a; Rydberg 1997). The NIG process is in fact a special case of the GH process where $\lambda = -0.5$. This results in the bounds given in Corollary 15b being exactly equal to the GIG density and thus the acceptance rate of points simulated from Alg. 3 of Godsill and Kındap (2021) is 1.0. The QQ plot, density estimate and sample paths for this parameter setting are shown in Figs. 15 and 16.

Another special case of the GH process is the Student-t process where $\lambda < 0, \gamma = 0$ and $\delta^2 = -2\lambda$. The Student-t distribution is parameterised using a single parameter ν, called the degrees of freedom, which is a positive real number. This parameter is related to the usual parameters of a GH process such that $\lambda = -\nu/2$ and $\delta = \sqrt{\nu}$.

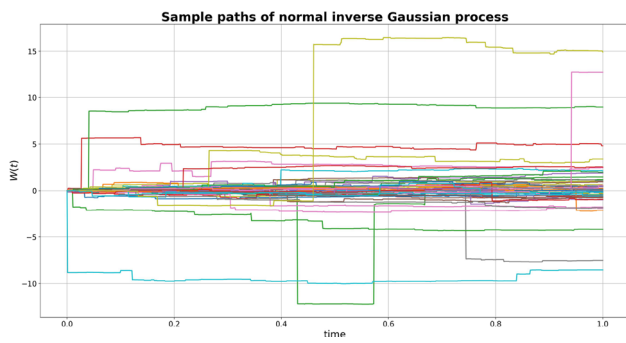As remarked in Sect. 5, it is not possible to simulate a Student-t process using Algorithms 7, 8, 9 and 10 because

Fig. 16 Pathwise simulations of the NIG process for $\lambda = -0.5$, $\gamma = 0.1$, $\delta = 1$, $\beta = 0$
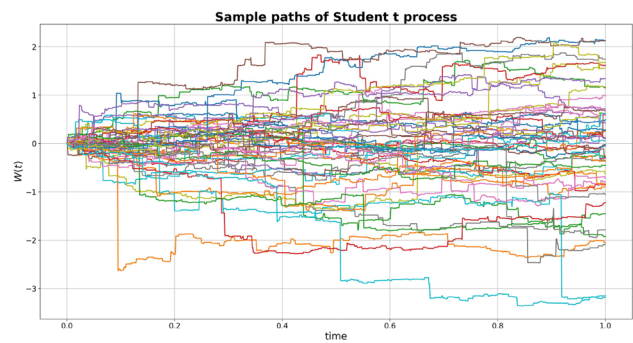

Fig. 18 Pathwise simulations of the Student t process for $\nu = 5$ and $\beta = 0$. ($\lambda = -2.5$, $\delta = \sqrt{5}$, $\gamma = 0$)
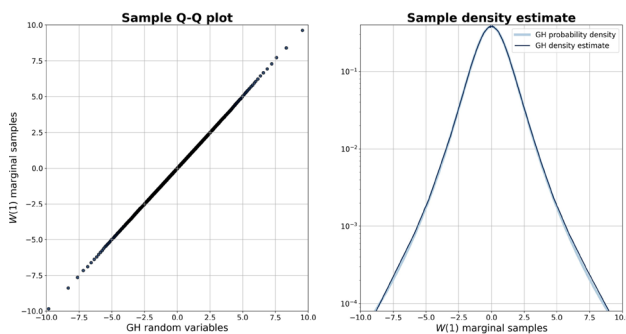

Fig. 17 Simulation comparison between the shot noise generated student-t process and student-t random variates, $\lambda = -2.5$, $\gamma = 0$, $\delta = \sqrt{5}$, $\beta = 0$. The adaptive truncation parameters are $p_T = 0.05$ and $\tau = 0.1$. Left hand panel: QQ plot comparing our shot noise method (y-axis) with random samples of the student-t density generated using a random variate generator (x-axis). Right hand panel: Normalised histogram density estimate for our method compared with the true student-t density function
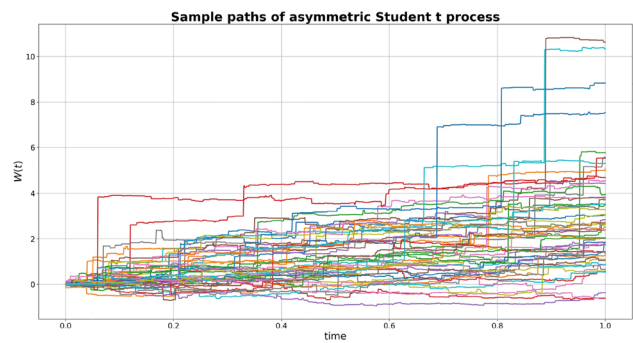
Fig. 19 Pathwise simulations of the asymmetric Student t process for $\nu = 5$ and $\beta = 2$. ($\lambda = -2.5$, $\delta = \sqrt{5}$, $\gamma = 0$)

Fig. 20 Simulation comparison between the shot noise generated asymmetric student-t process and student-t random variates, $\lambda = -2.5$, $\gamma = 0$, $\delta = \sqrt{5}$, $\beta = 2$. The adaptive truncation parameters are $p_T = 0.05$ and $\tau = 0.1$. Left hand panel: QQ plot comparing our shot noise method (y-axis) with random samples of the asymmetric student-t density generated using a random variate generator (x-axis). Right hand panel: Normalised histogram density estimate for our method compared with the true student-t density function

the corresponding gamma processes are not well-defined in this case. Instead, for this parameter setting Algorithm 3 of Godsill and Kındap (2021) is used together with the adaptive truncation and residual approximation methods presented in Sect. 5 to produce the jumps from the subordinator GIG process. The QQ plot, density estimate and sample paths for the resulting samples from the Student t process are shown in Figs. 17 and 18. Note that removing the condition $\delta^2 = -2\lambda$ still results in a well defined Lévy process with its marginal distribution parameterised by Eq. (3.11) in Eberlein and Hammerstein (2004).

It is common to apply the Student-t distribution to financial data sets as an alternative to the Gaussian distribution to account for the heavier tails observed in asset returns. There has been numerous studies that suggest that the asymmetric Student-t distribution, $\beta \neq 0$, provides a better fit to financial data sets compared to its symmetric counterpart (Zhu and Galbraith 2010; Alberg et al. 2008; Aas and Haff 2006). To the best of our knowledge, we present the sample paths of an asymmetric Student-t process for the first time in Fig. 19 together with the resulting marginal QQ plot and density esti-
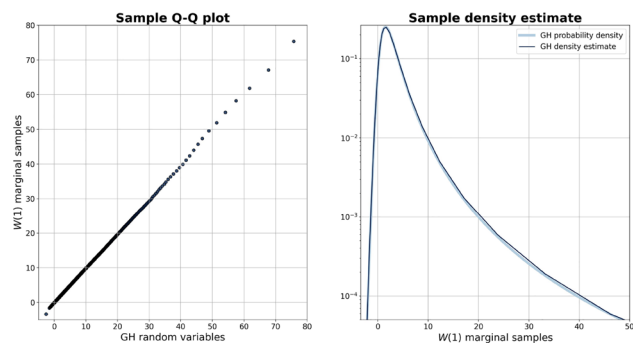
mate in Fig. 20. The marginal density of this limiting case of asymmetric GH processes are given in Eq. (3.9) of Eberlein and Hammerstein (2004).

8 Conclusions

The point process representation of a generalised hyperbolic process and the generalised shot noise methods developed in this work provide the the first complete methodology for simulation of generalised hyperbolic (GH) Lévy processes, giving a unified framework for a very broad range of heavy-tailed and semi-heavy-tailed non-Gaussian processes. The continuous time formulation, simulating directly in continuous time path space, can be employed for accurate uncertainty propagation, path visualisation, modelling and inference, especially for irregularly sampled time series datasets.

The presented methods are based on the subordination of a Brownian motion by the generalised inverse Gaussian (GIG) process, and we have here provided novel improvements in GIG process simulation compared with our previous work (Godsill and Kındap 2021). We have proved also that these series representations are almost surely convergent, verifying the conditions presented in Rosiński (2001). In addition to these improvements we present a novel scheme for adaptive truncation of the random shot noise representation based on probabilistic exceedance bounds, relying on new upper and lower bound expressions for the moments of truncated residual of the shot noise process; these truncations methods are shown to reduce computational burden dramatically without noticeable compromising of accuracy. Further computational savings are made through the use of squeezed rejection sampling, again based on our lower and upper moment bounds. The new GH process simulators developed in this work are used as a fundamental building block for modelling of stochastic differential equations (SDEs) driven by GH processes in Kındap and Godsill (2023), in a spirit similar to Godsill et al. (2019), where the conditionally Gaussian form of our models is of great benefit in inference for states and parameters for GH process-driven SDEs, finding application in spatial tracking, finance and vibration data modelling, to list only a few possibilities. These models and their applications will be further studied in future publications.

Declarations

A: Proof of Theorem 2

Proof For the case $|\lambda| \geq 0.5$, $B(z)$ forms an upper bound on $z|H_{|\lambda|}(z)|^2$ as stated in Eq. (12). The resulting integral obtained by replacing $z|H_{|\lambda|}(z)|^2$ with $B(z)$ is therefore a lower bound for both $\mathbb{E}[\rho_1(x, Z)]$ and $\mathbb{E}[\rho_2(x, Z)]$. As the function $B(z)$ has a piecewise form dependent on the choice of z_0 and the integrals are defined in terms of z_1, the lower bounds form a piecewise function dependent on their relative values.

Assuming $z_0 \leq z_1$, the interval $(0, z_1)$ is partitioned into two non-overlapping intervals $(0, z_0)$ and (z_0, z_1). The constituent functions of $B(z)$ is used to replace $z|H_{|\lambda|}(z)|^2$ in both partitions of $\mathbb{E}[\rho_1(x, Z)]$ as follows

$$
\mathbb{E}[\rho_1(x, Z)]
$$

$$
\geq \int_0^{z_0} \frac{2\,\Gamma(|\lambda|)\sqrt{\text{Ga}}(z||\lambda|, \frac{x}{2\delta^2})}{\pi\,H_0 \left(\frac{z_0}{z}\right)^{2|\lambda|-1} \left(\frac{z}{z_1}\right)^{2|\lambda|-1} \gamma(|\lambda|, \frac{z_1^2 x}{2\delta^2})}\,dz
$$

$$
+ \int_{z_0}^{z_1} \frac{2\,\Gamma(|\lambda|)\sqrt{\text{Ga}}(z||\lambda|, \frac{x}{2\delta^2})}{\pi\,H_0 \left(\frac{z}{z_1}\right)^{2|\lambda|-1} \gamma(|\lambda|, \frac{z_1^2 x}{2\delta^2})}\,dz
$$

$$
= \frac{2^2 \left(\frac{x}{2\delta^2}\right)^{|\lambda|}}{\pi\,H_0} \frac{\left(\frac{z_1}{z_0}\right)^{2|\lambda|-1}}{\gamma(|\lambda|, \frac{z_1^2 x}{2\delta^2})} \int_0^{z_0} z^{2|\lambda|-1} e^{-\frac{z^2 x}{2\delta^2}}\,dz
$$

$$
+ \frac{2^2 \left(\frac{x}{2\delta^2}\right)^{|\lambda|}}{\pi\,H_0} \frac{z_1^{2|\lambda|-1}}{\gamma(|\lambda|, \frac{z_1^2 x}{2\delta^2})} \int_{z_0}^{z_1} e^{-\frac{z^2 x}{2\delta^2}}\,dz
$$

$$
= \frac{2}{\pi\,H_0}\Bigg(\left(\frac{z_1}{z_0}\right)^{2|\lambda|-1} \frac{\gamma(|\lambda|, \frac{z_0^2 x}{2\delta^2})}{\gamma(|\lambda|, \frac{z_1^2 x}{2\delta^2})}
$$

$$
+ \left(\frac{z_1^2 x}{2\delta^2}\right)^{|\lambda|-0.5} \frac{\left[\gamma(0.5, \frac{z_1^2 x}{2\delta^2}) - \gamma(0.5, \frac{z_0^2 x}{2\delta^2})\right]}{\gamma(|\lambda|, \frac{z_1^2 x}{2\delta^2})} \Bigg)
$$

For $\mathbb{E}[\rho_2(x, Z)]$, the same assumption that $z_0 \leq z_1$ leads to a simple lower bound as

$$
\mathbb{E}[\rho_2(x, Z)] \geq \frac{2}{\pi\,H_0} \int_{z_1}^{\infty} \frac{\Gamma(0.5)\sqrt{\text{Ga}}(z|0.5, \frac{x}{2\delta^2})}{\Gamma(0.5, \frac{z_1^2 x}{2\delta^2})}\,dz
$$

$$= \frac{2^2}{\pi H_0} \frac{\left(\frac{x}{2\delta^2}\right)^{0.5}}{\Gamma(0.5, \frac{z_1^2 x}{2\delta^2})} \int_{z_1}^{\infty} e^{-\frac{z^2 x}{2\delta^2}} dz$$

$$= \frac{2}{\pi H_0}$$

For the case of $z_0 \geq z_1$, replacing $z|H_{|\lambda|}(z)|^2$ with $B(z)$ again leads to a simple lower bound on $\mathbb{E}[\rho_1(x, Z)]$ as

$$\mathbb{E}[\rho_1(x, Z)]$$

$$\geq \int_0^{z_1} \frac{2\,\Gamma(|\lambda|)\sqrt{\mathrm{Ga}}(z||\lambda|, \frac{x}{2\delta^2})}{\pi H_0 \left(\frac{z_0}{z}\right)^{2|\lambda|-1} \left(\frac{z}{z_1}\right)^{2|\lambda|-1} \gamma(|\lambda|, \frac{z_1^2 x}{2\delta^2})} dz$$

$$= \frac{2^2 \left(\frac{x}{2\delta^2}\right)^{|\lambda|} z_1^{2|\lambda|-1}}{\pi H_0 \gamma(|\lambda|, \frac{z_1^2 x}{2\delta^2}) z_0^{2|\lambda|-1}} \int_0^{z_1} z^{2|\lambda|-1} e^{-\frac{z^2 x}{2\delta^2}} dz$$

$$= \frac{2}{\pi H_0} \left(\frac{z_1}{z_0}\right)^{2|\lambda|-1}$$

Lastly, the lower bound on $\mathbb{E}[\rho_2(x, Z)]$ for $z_0 \geq z_1$ is obtained by partitioning (z_1, ∞) into two non-overlapping intervals (z_1, z_0), (z_0, ∞), and the constituent functions of $B(z)$ is used to replace $z|H_{|\lambda|}(z)|^2$ as

$$\mathbb{E}[\rho_2(x, Z)] \geq \int_{z_1}^{z_0} \frac{2\,\Gamma(0.5)\sqrt{\mathrm{Ga}}(z|0.5, \frac{x}{2\delta^2})}{\pi H_0 \left(\frac{z_0}{z}\right)^{2|\lambda|-1} \Gamma(0.5, \frac{z_1^2 x}{2\delta^2})} dz$$

$$+ \int_{z_0}^{\infty} \frac{2}{\pi H_0} \frac{\Gamma(0.5)\sqrt{\mathrm{Ga}}(z|0.5, \frac{x}{2\delta^2})}{\Gamma(0.5, \frac{z_1^2 x}{2\delta^2})} dz$$

$$= \frac{2^2 \left(\frac{x}{2\delta^2}\right)^{0.5}}{\pi H_0 z_0^{2|\lambda|-1} \Gamma(0.5, \frac{z_1^2 x}{2\delta^2})} \int_{z_1}^{z_0} z^{2|\lambda|-1} e^{-\frac{z^2 x}{2\delta^2}} dz$$

$$+ \frac{2^2 \left(\frac{x}{2\delta^2}\right)^{0.5}}{\pi H_0 \Gamma(0.5, \frac{z_1^2 x}{2\delta^2})} \int_{z_0}^{\infty} e^{-\frac{z^2 x}{2\delta^2}} dz$$

$$= \frac{2}{\pi H_0} \left(\frac{\Gamma(0.5, \frac{z_0^2 x}{2\delta^2})}{\Gamma(0.5, \frac{z_1^2 x}{2\delta^2})} \right.$$

$$\left. + \left(\frac{z_0^2 x}{2\delta^2}\right)^{0.5-|\lambda|} \frac{\left[\gamma(|\lambda|, \frac{z_0^2 x}{2\delta^2}) - \gamma(|\lambda|, \frac{z_1^2 x}{2\delta^2}) \right]}{\Gamma(0.5, \frac{z_1^2 x}{2\delta^2})} \right)$$

$\square$

B: Proof of Theorem 4

Proof In principle the residual moments of a Lévy process can be obtained directly from Eqs. (37) and (38) as before.

However, once again these cannot be evaluated exactly for most parameter settings. Instead we provide lower bounds for the integrand term $Q_{GIG}(x)$ which are obtained from the marginal point processes $Q_{N_1/N_2}^{A/B}(x)$ defined in Eqs. (20), (21), (28) and (29).

The starting point is the bivariate Lévy density $Q_{GIG}(x, z)$, and the lower bounds on this function given in Corollary 1. For the case $|\lambda| \geq 0.5$, the marginal $Q_{GIG}(x)$ is bounded as

$$Q_{N_1}^B(x) + Q_{N_2}^B(x) \leq Q_{GIG}(x)$$

Hence for $n \in \{1, 2\}$, bounds on the residual moments may be found by studying each term in

$$\int_0^{\varepsilon} x^n Q_{N_1}^B(x)dx + \int_0^{\varepsilon} x^n Q_{N_2}^B(x)dx$$

$$\leq \int_0^{\varepsilon} x^n Q_{GIG}(x)dx$$

The first integral over $Q_{N_1}^B(x)$ is still intractable and hence has to be further lower bounded. A lower bound on the incomplete gamma function is defined in Neuman (2013b), Th. 4.1)[7] as

$$\frac{(z_0^2 x/(2\delta^2))^{|\lambda|}}{|\lambda|} \exp\left(\frac{-|\lambda| z_0^2 x/(2\delta^2)}{|\lambda| + 1} \right) \leq \gamma(|\lambda|, z_0^2 x/(2\delta^2))$$

Therefore a lower bounding density is obtained as

$$\frac{z_0}{\pi^2 H_0 |\lambda|} x^{-1} \exp\left(-\left[\frac{\gamma^2}{2} + \frac{|\lambda|}{(1+|\lambda|)} \frac{z_0^2}{2\delta^2} \right] x \right) \leq Q_{N_1}^B(x)$$

which is the density of a gamma process. The associated residual moments have already been given in Theorem 3 and shown in Eqs. (44) and (45). The associated mean and variance of the residual of $Q_{N_1}^B$ for a truncation level ε are then given directly as

$$\mu_{N_1}^B(\varepsilon) = \frac{C_{Ga}^B}{\beta_{Ga}^B} \gamma\left(1, \beta_{Ga}^B \varepsilon\right)$$

$$\sigma_{N_1}^{B\,2}(\varepsilon) = \frac{C_{Ga}^B}{\beta_{Ga}^{B\,2}} \gamma\left(2, \beta_{Ga}^B \varepsilon\right)$$

where

$$C_{Ga}^B = \frac{z_0}{\pi^2 H_0 |\lambda|} \quad \text{and} \quad \beta_{Ga}^B = \frac{\gamma^2}{2} + \frac{|\lambda|}{(1+|\lambda|)} \frac{z_0^2}{2\delta^2}$$

---

[7] Note that an alternative bound is given by $(1 - e^{-x})/x \leq a/x^a \gamma(a, x)$ for $0 \leq a \leq 1$.

The density $Q_{N_2}^B$ is a modified tempered stable characterised by the upper incomplete gamma function, which can be lower bounded as (Chang et al. 2011, Th. 2)

$$\sqrt{2e}\frac{\sqrt{\beta_0 - 1}}{\beta_0} \exp\left(-\beta_0 \frac{z_0^2 x}{2\delta^2}\right) \leq \Gamma(0.5, \frac{z_0^2 x}{2\delta^2})$$

where $\beta_0 > 1$ and $x \geq 0$. The free parameter β_0 can be adaptively optimised across the whole domain. The bound can be derived using the well-known equivalence $\Gamma(0.5, x) = \sqrt{\pi}\,\mathrm{erfc}(\sqrt{x})$.

Hence a lower bound on $Q_{N_2}^B$ is obtained as

$$\frac{2\delta\sqrt{e}\sqrt{\beta_0 - 1}}{\pi^2 H_0 \beta_0} x^{-3/2} \exp\left(-\left[\frac{\gamma^2}{2} + \frac{\beta_0 z_0^2}{2\delta^2}\right]x\right) \leq Q_{N_2}^B(x)$$

which is the density of a tempered stable process and the associated residual moments are previously derived in the proof of Theorem 3 and shown in Eqs. (40) and (41). The lower bound on the residual mean and variance of $Q_{N_2}^B$ for a truncation level ε are then obtained directly as

$$\mu_{N_2}^B(\varepsilon) = C_{TS}^B \beta_{TS}^{B}{}^{-0.5} \gamma\left(0.5, \beta_{TS}^B \varepsilon\right)$$

$$\sigma_{N_2}^B{}^2(\varepsilon) = C_{TS}^B \beta_{TS}^{B}{}^{-1.5} \gamma\left(1.5, \beta_{TS}^B \varepsilon\right)$$

where

$$C_{TS}^B = \frac{2\delta\sqrt{e}\sqrt{\beta_0 - 1}}{\pi^2 H_0 \beta_0} \quad \text{and} \quad \beta_{TS}^B = \frac{\gamma^2}{2} + \frac{\beta_0 z_0^2}{2\delta^2}$$

Finally the lower bound on the residual mean and variance of the GIG density can be expressed in terms of the residual means and variances of the lower bounding gamma and TS processes as $\underline{\mu_\varepsilon} = \mu_{N_1}^B(\varepsilon) + \mu_{N_2}^B(\varepsilon)$ and $\underline{\sigma_\varepsilon^2} = \sigma_{N_1}^B{}^2(\varepsilon) + \sigma_{N_2}^B{}^2(\varepsilon)$.

For the case $0 < |\lambda| \leq 0.5$, the marginal $Q_{GIG}(x)$ is bounded as

$$Q_{N_1}^A(x) + Q_{N_2}^A(x) \leq Q_{GIG}(x)$$

Similarly for $n \in \{1, 2\}$, lower bounds on the moments may be found as

$$\int_0^\varepsilon x^n Q_{N_1}^A(x)dx + \int_0^\varepsilon x^n Q_{N_2}^A(x)dx$$
$$\leq \int_0^\varepsilon x^n Q_{GIG}(x)dx$$

The lower bounds on the incomplete gamma functions established for the previous parameter setting can be directly substituted into $Q_{N_1}^A(x)$ and $Q_{N_2}^A(x)$. The resulting densities

are again characterised as a gamma process and a tempered stable process respectively such that

$$\frac{z_1}{2\pi|\lambda|}x^{-1} \exp\left(-\left[\frac{\gamma^2}{2} + \frac{|\lambda|}{(1+|\lambda|)}\frac{z_1^2}{2\delta^2}\right]x\right) \leq Q_{N_1}^A(x)$$

$$\frac{\delta\sqrt{e}\sqrt{\beta_0 - 1}}{\pi\beta_0}x^{-3/2} \exp\left(-\left[\frac{\gamma^2}{2} + \frac{\beta_0 z_1^2}{2\delta^2}\right]x\right) \leq Q_{N_2}^A(x)$$

Hence the associated residual moments of these processes can be directly found again using Eqs. (44), (45) and (40), (41). The corresponding lower bound on the residual mean and variance of the GIG process for the case $0 < |\lambda| \leq 0.5$ and truncation level ε are then found as

$$\underline{\mu_\varepsilon} = \frac{C_{Ga}^A \gamma\left(1, \beta_{Ga}^A \varepsilon\right)}{\beta_{Ga}^A} + \frac{C_{TS}^A \gamma\left(0.5, \beta_{TS}^A \varepsilon\right)}{\beta_{TS}^{A}{}^{0.5}}$$

$$\underline{\sigma_\varepsilon^2} = \frac{C_{Ga}^A \gamma\left(2, \beta_{Ga}^A \varepsilon\right)}{\beta_{Ga}^{A}{}^2} + \frac{C_{TS}^A \gamma\left(1.5, \beta_{TS}^A \varepsilon\right)}{\beta_{TS}^{A}{}^{1.5}}$$

$\square$

References

Aas, K., Haff, I.H.: The generalized hyperbolic skew student's t-distribution. J. Financ. Econom. **4**(2), 275–309 (2006). https://doi.org/10.1093/jjfinec/nbj006

Achim, A., Buxton, B., Tzagkarakis, G., Tsakalides, P.: Compressive sensing for ultrasound RF echoes using α-Stable Distributions. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, pp. 4304–4307 (2010)

Achim, A., Bezerianos, A., Tsakalides, P.: Novel Bayesian multiscale method for speckle removal in medical ultrasound images. IEEE Trans. Med. Imaging **20**(8), 772–783 (2001)

Achim, A., Kuruoğlu, E.E., Zerubia, J.: SAR image filtering based on the heavy-tailed Rayleigh model. IEEE Trans. Image Process. **15**(9), 2686–2693 (2006)

Alberg, D., Shalit, H., Yosef, R.: Estimating stock market volatility using asymmetric Garch models. Appl. Financ. Econ. **18**(15), 1201–1208 (2008). https://doi.org/10.1080/09603100701604225

Amini, A., Unser, M.: Sparsity and infinite divisibility. IEEE Trans. Inf. Theory **60**(4), 2346–2358 (2014)

Asmussen, S., Rosiński, J.: Approximations of small jumps of Lévy processes with a view towards simulation. J. Appl. Probab. **38**(2), 482–493 (2001)

Azzaoui, N., Clavier, L.: Statistical channel model based on α-stable random processes and application to the 60 GHz ultra wide band channel. IEEE Trans. Commun. **58**(5), 1457–1467 (2010)

Barndorff-Nielsen, O.: Hyperbolic distributions and distributions on hyperbolae. Scan. J. Stat. **5**(3), 151–157 (1978)

Barndorff-Nielsen, O.E.: Processes of normal inverse Gaussian type. Finance Stoch. **2**(1), 41–68 (1997a)

Barndorff-Nielsen, O.E.: Normal inverse gaussian distributions and stochastic volatility modelling. Scand. J. Stat. **24**(1), 1–13 (1997b)

Barndorff-Nielsen, O., Halgreen, C.: Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **38**, 309–311 (1977)

Barndorff-Nielsen, O.E., Shephard, N.: Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial

economics. J. R. Stat. Soc. Ser. B (Stat. Methodol.) **63**(2), 167–241 (2001). https://doi.org/10.1111/1467-9868.00282

Barndorff-Nielsen, O.E., Shephard, N.: Basics of Levy processes. Economics Group, Nuffield College, University of Oxford, Economics Papers 2012-W06 (2012). https://ideas.repec.org/p/nuf/econwp/1206.html

Barndorff-Nielsen, O., Mikosch, T., Resnick, S.: Lévy Processes: Theory and Applications. Birkhäuser, Boston (2001)

Barndorff-Nielsen, O.E., Shephard, N.G., et al.: Basics of Lévy processes. Citeseer (2012)

Barth, A., Stein, A.: Approximation and simulation of infinite-dimensional levy processes. Stoch. Partial Differ. Equ. Anal. Comput. **6**, 286–334 (2017)

Bertoin, J.: Lévy Processes, ser. Cambridge Tracts in Mathematics, vol. 121. Cambridge University Press, London (1997)

Borak, S., Misiorek, A., Weron, R.: Models for heavy-tailed asset returns, Statistical Tools for Finance and Insurance, 02 (2011)

Carrillo, R.E., Ramirez, A.B., Arce, G.R., Barner, K.E., Sadler, B.M.: Robust compressive sensing of sparse signals: a review. EURASIP J. Adv. Signal Process. **2016**(1), 108 (2016)

Chang, S.-H., Cosman, P.C., Milstein, L.B.: Chernoff-type bounds for the gaussian error function. IEEE Trans. Commun. **59**(11), 2939–2944 (2011)

Chen, X., Wang, Z.J., McKeown, M.J.: Asymptotic analysis of robust lassos in the presence of noise with large variance. IEEE Trans. Inf. Theory **56**(10), 5131–5149 (2010)

Chiani, M., Dardari, D., Simon, M.: New exponential bounds and approximations for the computation of error probability in fading channels. IEEE Trans. Wirel. Commun. **2**(4), 840–845 (2003)

Cont, R., Tankov, P.: Financial Modelling with Jump Processes. Chapman & Hall/CRC, Boca Raton (2003)

Costa, M.T., Kontoyiannis, I., Godsill, S.: Generalised shot noise representations of stochastic systems driven by non-Gaussian Lévy processes, arXiv:2305.05931 (2023)

de Freitas, M.L., Egan, M., Clavier, L., Goupil, A., Peters, G.W., Azzaoui, N.: Capacity bounds for additive symmetric α-stable noise channels. IEEE Trans. Inf. Theory **63**, 5115–5123 (2017)

Devroye, L.: Non-Uniform Random Variate Generation. Springer, Berlin (1986)

Devroye, L.: Random variate generation for the generalized inverse Gaussian distribution. Stat. Comput. **24**, 239–246 (2014)

Eberlein, E.: Application of generalized hyperbolic Lévy motions to finance. In: Barndorff-Nielsen, O., Resnick, S., Mikosch, T. (eds.) Lévy Process. Birkhäuser, Boston, MA (2001)

Eberlein, E., Hammerstein, E.A.V.: Generalized hyperbolic and inverse Gaussian distributions: limiting cases and approximation of processes. In: Dalang, R.C., Dozzi, M., Russo, F. (eds.) Seminar on Stochastic Analysis, Random Fields and Applications IV, pp. 221–264. Basel, Birkhäuser (2004)

Eberlein, E., Keller, U.: Hyperbolic distributions in finance. Bernoulli **1**(3), 281–299 (1995). https://doi.org/10.3150/bj/1193667819

Fahs, J., Abou-Faycal, I.: On the capacity of additive white alpha-stable noise channels. In: IEEE International Symposium on Information Theory Proceedings. IEEE 2012, pp. 294–298 (2012)

Fama, E.F.: The behavior of stock-market prices. J. Bus. **38**(1), 34–105 (1965)

Ferguson, T.S., Klass, M.J.: A representation of independent increment processes without Gaussian components. Ann. Math. Stat. **43**(5), 1634–1643 (1972). https://doi.org/10.1214/aoms/1177692395

Godsill, S., Riabiz, M., Kontoyiannis, I.: The Lévy state space model. In: 2019 53rd Asilomar Conference on Signals, Systems, and Computers, pp. 487–494 (2019)

Godsill, S., Kındap, Y.: Point process simulation of generalised inverse Gaussian processes and estimation of the Jaeger integral. Stat. Comput. **32**(1), 13 (2021). https://doi.org/10.1007/s11222-021-10072-0

Godsill, S., Rayner, P.: Statistical reconstruction and analysis of autoregressive signals in impulsive noise using the Gibbs sampler. IEEE Trans. Speech Audio Process. **6**(4), 352–372 (1998)

Grabchak, M.: Rejection sampling for tempered Lévy processes. Stat. Comput. **29**, 549–558 (2019)

Hörmann, W., Leydold, J.: Generating generalized inverse Gaussian random variates. Stat. Comput. **24**(4), 547–557 (2013)

Kallenberg, O.: Foundations of Modern Probability, 2nd edn. Springer, Berlin (2002)

Katz, R.W., Brown, B.G.: Extreme events in a changing climate: variability is more important than averages. Clim. Change **21**(3), 289–302 (1992)

Katz, R.W., Parlange, M.B., Naveau, P.: Statistics of extremes in hydrology. Adv. Water Resour. **25**(8–12), 1287–1304 (2002)

Ken-Iti, S.: Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press, London (1999)

Kındap, Y., Godsill, S.: Generalised hyperbolic state-space models for inference in dynamic systems, arXiv arXiv:2309.11422 (2023)

Kingman, J.F.C.: Poisson Processes, ser Oxford Studies in Probability. Clarendon Press, London (1992)

Lemke, T., Godsill, S.: Inference for models with asymmetric α-stable noise processes. In: Unobserved Components and Time Series Econometrics, ch. 9. Oxford University Press (2015)

Lemke, T., Riabiz, M., Godsill, S.J.: Fully Bayesian inference for α-stable distributions using a Poisson series representation. Digit. Signal Process. **47**, 96–115 (2015)

Lewis, P.A.W., Shedler, G.S.: Simulation of nonhomogeneous Poisson processes by thinning. Naval Res. Logist. Q. **26**(3), 403–413 (1979)

Liebeherr, J., Burchard, A., Ciucu, F.: Delay bounds in communication networks with heavy-tailed and self-similar traffic. IEEE Trans. Inf. Theory **58**(2), 1010–1024 (2012)

Lombardi, M.J., Godsill, S.J.: On-line Bayesian estimation of signals in symmetric α-stable noise. IEEE Trans. Signal Process. **54**(2), 775–779 (2006)

Lopes, M.E.: Unknown sparsity in compressed sensing: denoising and inference. IEEE Trans. Inf. Theory **62**(9), 5145–5166 (2016)

Madan, D.B., Seneta, E.: The variance gamma (v.g.) model for share market returns. J. Bus. **63**(4), 511–24 (1990)

Mandelbrot, B.: New methods in statistical economics. J. Polit. Econ. **71**(5), 421–440 (1963)

Neuman, E.: Inequalities and bounds for the incomplete gamma function. Res. Math. **63**(3), 1209–1214 (2013b)

Neuman, E.: Inequalities and bounds for the incomplete gamma function. Res. Math. **63**, 06 (2013a)

Nikias, C.L., Shao, M.: Signal Processing with Alphastable Distributions and Applications, Series Adaptive and Learning Systems for Signal Processing, Communications, and Control. Wiley, New York (1995)

Øksendal, B.: Stochastic Differential Equations: An Introduction with Applications (Universitext), 6th edn. Springer, Berlin (2014)

Qu, Y., Dassios, A., Zhao, H.: Exact simulation of Ornstein–Uhlenbeck tempered stable processes. J. Appl. Probab. **58**, 347–351 (2021)

Riabiz, M., Ardeshiri, T., Godsill, S.: Convergence results for tractable inference in α-stable stochastic processes. In: 2017 22nd International Conference on Digital Signal Processing (DSP), pp. 1–5 (2017)

Riabiz, M., Ardeshiri, T., Kontoyiannis, I., Godsill, S.: Nonasymptotic Gaussian approximation for inference with stable noise. IEEE Trans. Inf. Theory **66**, 4966–4991 (2020)

Rosiński, J.: Series Representations of Lévy Processes from the Perspective of Point Processes, pp. 401–415. Birkhäuser, Boston, MA (2001). https://doi.org/10.1007/978-1-4612-0197-7_18

Rosiński, J.: Series representations of Lévy processes from the perspective of point processes. In: Barndorff-Nielsen, O., Resnick,

S., Mikosch, T. (eds.) Lévy Processes, pp. 401–415. Birkhauser, Boston (2001)

Rydberg, T.H.: The normal inverse Gaussian Lévy process: simulation and approximation. Commun. Stat. Stoch. Models **13**(4), 887–910 (1997). https://doi.org/10.1080/15326349708807456

Sabino, P.: Exact simulation of normal tempered stable processes of ou type with applications. Stat. Comput. **32**, 09 (2022)

Samorodnitsky, G., Taqqu, M.S.: Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance. CRC Press, Boca Raton (1994)

Shah, A., Wilson, A., Ghahramani, Z.: Student-t Processes as Alternatives to Gaussian Processes. In: Kaski, S., Corander, J. (Eds) Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics, ser. Proceedings of Machine Learning Research, vol. 33. Reykjavik, Iceland: PMLR, pp. 877–885 (2014). Available: https://proceedings.mlr.press/v33/shah14.html

Shevlyakov, G., Kim, K.: Robust minimax detection of a weak signal in noise with a bounded variance and density value at the center of symmetry. IEEE Trans. Inf. Theory **52**(3), 1206–1211 (2006)

Simon, T.: Subordination in the wide sense for Lévy processes. Probab. Theory Relat. Fields **115**, 445–477 (1999)

Solin, A., Särkkä, S.: State space methods for efficient inference in student-t process regression. In: Lebanon, G., Vishwanathan, S.V.N. (Eds.) Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics, ser. Proceedings of Machine Learning Research, vol. 38. PMLR, San Diego, pp. 885–893 (2015)

Statovic, gigrnd(p, a, b, samplesize). MATLAB Central File Exchange (2017). https://www.mathworks.com/matlabcentral/fileexchange/53594-gigrnd-p-a-b-samplesize

Tzagkarakis, G.: Bayesian Compressed Sensing Using α-stable Distributions, Ph.D. dissertation, Department of Computer Science, University of Crete, Crete, Greece (2009)

Unser, M., Tafti, P.D.: An Introduction to Sparse Stochastic Processes. Cambridge University Press, London (2014)

Unser, M., Tafti, P., Sun, Q.: A unified formulation of Gaussian versus sparse stochastic processes—2014; Part I: continuous-domain theory. IEEE Trans. Inf. Theory **60**(3), 1945–1962 (2014a)

Unser, M., Tafti, P., Amini, A., Kirshner, H.: A unified formulation of Gaussian versus sparse stochastic processes—2014; Part II: discrete-domain theory. IEEE Trans. Inf. Theory **60**(5), 3036–3051 (2014b)

Veraart, A.E., Winkel, M.: Time change. Encyclopedia of Quantitative Finance **4**, 1812–1816 (2010)

Warren, D.J., Thomas, J.B.: Asymptotically robust detection and estimation for very heavy-tailed noise. IEEE Trans. Inf. Theory **37**(3), 475–481 (1991)

Watson, G.: A Treatise on the Theory of Bessel Functions, 2nd edn. Cambridge University Press, London (1944)

Wolpert, R., Ickstadt, K.: Poisson/gamma random field models for spatial statistics. Biometrika **85**(2), 251–267 (1998a). https://doi.org/10.1093/biomet/85.2.251

Wolpert, R.L., Ickstadt, K.: Simulation of Lévy Random Fields, pp. 227–242. Springer, New York (1998b). https://doi.org/10.1007/978-1-4612-1732-9_12

Wolpert, R.L.: Lévy Random Fields (2021). https://www2.stat.duke.edu/courses/Spring16/sta961/lec/levy.pdf

Zhang, S.: Exact simulation of tempered stable Ornstein–Uhlenbeck processes. J. Stat. Comput. Simul. **81**(11), 1533–1544 (2011). https://doi.org/10.1080/00949655.2010.494247

Zhou, Z., Yu, J.: Estimation of block sparsity in compressive sensing, arXiv preprint arXiv:1701.01055 (2017)

Zhu, D., Galbraith, J.: A generalized asymmetric Student-t distribution with application to financial econometrics. J. Econom. **157**(2), 297–305 (2010)