Generalised Hyperbolic State-space Models for Inference in Dynamic Systems

Yaman Kındap
Department of Engineering
University of Cambridge
Cambridge, UK
yk392@cam.ac.uk

Simon Godsill
Department of Engineering
University of Cambridge
Cambridge, UK
sjg30@cam.ac.uk

Abstract

In this work we study linear vector stochastic differential equation (SDE) models driven by the generalised hyperbolic (GH) Lévy process for inference in continuous-time non-Gaussian filtering problems. The GH family of stochastic processes offers a flexible framework for modelling of non-Gaussian, heavy-tailed characteristics and includes the normal inverse-Gaussian, variance-gamma and Student-t processes as special cases. We present continuous-time simulation methods for the solution of vector SDE models driven by GH processes and novel inference methodologies using a variant of sequential Markov chain Monte Carlo (MCMC). As an example a particular formulation of Langevin dynamics is studied within this framework. The model is applied to both a synthetically generated data set and a real-world financial series to demonstrate its capabilities.

1 Introduction

In the fields of statistical signal processing and machine learning, the problem of modelling dynamically evolving phenomena can be found in various application domains including financial markets [1], speech recognition [2], health monitoring [3] and more recently in generative models for images [4]. Stochastic differential equation (SDE) models of real-world dynamic phenomena offer a powerful tool for capturing the inherent uncertainty in the estimated behaviour of such systems and provide a probabilistic framework that allows for the quantification and propagation of uncertainty.

Typically the driving (or forcing) function, which characterises the random fluctuations from the estimated deterministic evolution, is chosen as a Brownian motion. However, the Gaussian assumption implicit in the use of Brownian motion is often not justified in real-world applications, as many systems exhibit non-Gaussian features such as asymmetric or heavy-tailed distributions. Even with Brownian-driven paths, discretisation may require complex integrals and introduce further complexity. Instead, inference algorithms that work directly in continuous-time offer an effective alternative to discrete-time algorithms for large, irregularly sampled sequential data sets by reducing sampling costs associated with grid based strategies ([5], Section 4.6, [6]). Recent work on SDE models extends their definition to incorporate non-Gaussian Lévy processes, allowing for a wider range of behaviours including the α-stable and Poisson processes ([7; 8; 9]). This extension enhances the flexibility and applicability of SDE models, enabling them to capture more complex dynamics and uncertainties in real-world systems.

In this work, we study linear SDE models driven by normal variance-mean mixture (NVM) processes, specifically the generalised hyperbolic (GH) Lévy process which specifies a broad class of stochastic processes for varying levels of non-Gaussian, heavy-tailed and asymmetric characteristics, and include the normal inverse-Gaussian, variance-gamma and Student-t processes as special subclasses ([10; 11]). A point process simulation algorithm for the GH process has been recently presented in

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

[12]. Building on this algorithm, we present a simulation method for SDEs driven by GH processes and a novel continuous-time filtering algorithm which enables inference in such systems.

Specifically, the linear vector SDE model is defined as

$$d\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t)dt + \mathbf{L}dW(t) \tag{1}$$

where $\mathbf{x}(t)$ is the D-dimensional state vector at time t, $\mathbf{A}$ and $\mathbf{L}$ are system matrices, and $W(t)$ is a background-driving Lévy process. The observations are assumed to be related to the state vector through

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \varepsilon(t) \tag{2}$$

where $\mathbf{H}$ is the observation matrix and $\varepsilon(t)$ is assumed to be a continuous-time white Gaussian noise with covariance σ_ε^2.

The system of equations of the form presented in Eqs. (1) and (2) define a Lévy state-space model (SSM) [8]. Various well-known models can be treated as special cases of this general structure by choosing the particular forms of $\mathbf{A}$, $\mathbf{L}$ and $\mathbf{H}$. Some example models are the standard linear tracking models [8], continuous-time autoregressive (CAR), the CAR moving-average (CARMA) ([6; 13]) and the Ornstein-Uhlenbeck (OU) processes [14].

As a particular case, we study a Lévy SSM formulation of Langevin dynamics defined as:

$$d\dot{x}(t) = \theta\dot{x}(t)dt + dW(t)$$

where $\theta < 0$ such that the evolution of the process reverts to zero in proportion to the magnitude of $\dot{x}(t)$. Langevin dynamics find application directly in physics and biological sciences as a model of natural phenomena. Furthermore, it can be applied to optimisation and Markov chain Monte Carlo algorithms for the design of gradient-based learning methods [7; 15]. The linear SDE formulation of Langevin dynamics can be expressed as:

$$\begin{bmatrix} dx(t) \\ d\dot{x}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & \theta \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} dt + \begin{bmatrix} 0 \\ 1 \end{bmatrix} dW(t) \tag{3}$$

where by definition $dx(t) = \dot{x}(t)dt$. Hence, a Lévy SSM allows both the state evolution $\dot{x}(t)$ and the integrated state $x(t)$ to be inferred within the same framework.

Previous works on linear vector SDE models use various particle filtering (PF) strategies for the inference of states in Eq. (1) ([16; 8]). Here, we develop a novel framework using a variant of Sequential Markov chain Monte Carlo (SMCMC) ([17; 18]) for the inference of states in a general Lévy SSM and demonstrate its applicability using Langevin dynamics driven by GH processes.

The paper is organised as follows. In Section 2 we review fundamental results on the simulation of vector SDEs driven by GH processes. In Section 3 we present our novel inference algorithm for the GH SSM and demonstrate its applicability in Section 4 with both a synthetically generated data set and a British Pound - Japanese Yen foreign exchange rate data set.

2 Simulation of GH SSMs

In this section, we review methods of simulation for vector SDEs driven by Lévy processes. Particularly for the GH process, a conditionally Gaussian formulation of the simulation algorithm is introduced which will form the basis of our inference algorithm discussed in the following sections.

The solution to vector SDEs, as shown in Eq. (1), can be expressed as [19]

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-u)}\mathbf{L}dW(u) \tag{4}$$

where $e^{\mathbf{A}t}$ is the matrix exponential of $\mathbf{A}t$ and $\mathbf{x}(0)$ is an initial condition. The second term in the right hand side is a stochastic integral with respect to a Lévy process dW. When studying the solution for specific characterisations of $\mathbf{A}$, $\mathbf{L}$ and $dW(t)$, it is usually convenient to define the stochastic term as:

$$\mathbf{I}(\mathbf{f}_t) = \int_0^T \mathbf{f}_t(u)dW(u)$$

where $\mathbf{f}_t(u) = e^{\mathbf{A}(t-u)}\mathbf{L}\mathbb{1}_{\{u \leq t\}}$ and $T > t$ is some arbitrary boundary point. The integral $\mathbf{I}(\mathbf{f}_t)$ can only be analytically solved when dW is Gaussian. Hence in general $\mathbf{I}(\mathbf{f}_t)$ has to be approximated, e.g. using direct Euler-type discretisations as studied in [6]. In this work, we adopt a point process representation of $\mathbf{I}(\mathbf{f}_t)$ enabled by the series representation of Lévy processes in terms of its jumps times and sizes $\{(V_i, W_i)\}$ as in [12]. When $dW(t)$ is a normal variance-mean (NVM) process, such as the GH process, the integral admits the shot-noise series representation $\mathbf{I}(\mathbf{f}_t) = \sum_{i=1}^{\infty} \mathbf{f}_t(V_i)W_i$ where the jump sizes W_i can be further decomposed as

$$W_i = \mu_W Z_i + \sigma_W \sqrt{Z_i} U_i \tag{5}$$

where $\mu_W \in \mathbb{R}$ is a skewness parameter, $\sigma_W \in (0, \infty)$ is a scale parameter, $V_i \sim \mathrm{Unif}(0, T)$ are jump times, Z_i are the jumps of a subordinator process and $U_i \sim \mathcal{N}(0, 1)$ [20] such that

$$\sum_{i=1}^{n} \mathbf{f}_t(V_i)W_i \overset{n\to\infty}{\to} \int_0^T \mathbf{f}_t(u)dW(u)$$

which converges uniformly on $(0, T]$ to the original integral as $n \to \infty$. Thus the continuous-time process $\mathbf{I}(\mathbf{f}_t)$ can be represented as a countably infinite discrete sum without additional discretisation schemes as in the Gaussian case. The proof of this convergence can be found in Theorem 4 of [20].

To generate random sample paths of the state vector $\mathbf{x}(t)$, the solution shown in Eq. (4) can be incrementally simulated on each interval $(s, t]$ as

$$\mathbf{x}(t) = e^{\mathbf{A}(t-s)}\mathbf{x}(s) + \sum_{i:V_i \in (s,t]} e^{\mathbf{A}(t-V_i)}\mathbf{L}W_i \tag{6}$$

given an initial state vector $\mathbf{x}(s)$ and jumps $\{(V_i, W_i) : V_i \in (s, t]\}$.

Since there are an infinite number of jumps in any interval for the GH process, the summation in Eq. (6) must be truncated to exclude small jumps of $dW(t)$ as discussed in [12]. To improve the approximation, the residual approximation schemes for the series representation of GH processes can be incorporated into the SDE simulation framework. The effect of residual small jumps, defined as $R_t^c = \mathbf{I}(\mathbf{f}_t) - \mathbf{I}^c(\mathbf{f}_t)$ where c is the truncation level, can be included as a standard stochastic integral driven by a Brownian motion. The moments of the residual process for the GH case are derived in [12] and the validity of this residual approximation is studied in [20].

Notice that for NVM processes, given the set of subordinator jump sizes $\{Z_i\}_{i=1}^{\infty}$, $\mathbf{I}(\mathbf{f}_t)$ is conditionally Gaussian which allows for efficient simulation and inference procedures to be designed. Hence, it is useful to express the integral in Eq. (6) directly as a function of the subordinator jump times and sizes $\{(V_i, Z_i) : V_i \in (s, t]\}$ by substituting Eq. (5) into (6) such that $\mathbf{I}(\mathbf{f}_t)|\{(V_i, Z_i)\} \sim \mathcal{N}(\mathbf{m}, \mathbf{S})$ where

$$\mathbf{m} = \sum_{i:V_i \in (s,t]} \mathbf{f}_t(V_i)\mu_W Z_i \tag{7}$$

$$\mathbf{S} = \sum_{i:V_i \in (s,t]} \mathbf{f}_t(V_i)\mathbf{f}_t(V_i)^T \sigma_W^2 Z_i \tag{8}$$

The conditionally Gaussian formulation of the integral $\mathbf{I}(\mathbf{f}_t)$ leads to a fairly straightforward conditional form for the incremental simulation of the sample paths from Eq. (6) such that

$$p(\mathbf{x}_t|\mathbf{x}_s, \{(V_i, Z_i) : V_i \in (s, t]\}) = \mathcal{N}(e^{\mathbf{A}(t-s)}\mathbf{x}_s + \mathbf{m}, \mathbf{S}) \tag{9}$$

where the notation of the state is simplified to $\mathbf{x}_t := \mathbf{x}(t)$ for brevity. Further details are provided in Appendix A.

3 Inference for GH SSMs

In this section we describe a novel sequential inference algorithm for the states of a Lévy SSM driven by a normal variance-mean process such as the generalised hyperbolic case. We start by briefly reviewing the recursive state estimation problem in the context of Lévy SSMs and subsequently introduce our SMCMC algorithm for this task.

The problem of filtering in a dynamic system can be described as the computation of the posterior distribution of a vector of state variables $\mathbf{x}_t$ conditional on a batch of observations $y_{0:t} := \{y(s_i) :$

$0 \leq s_i \leq t\}$ [21], where here $\{s_i\}$ is a set of discrete observation times for the process. The sequential nature of dynamic systems allow the filtering problem to be solved incrementally for all t. Given an initial filtering estimate at time s, denoted as $p(\mathbf{x}_s|y_{0:s})$, the filtering estimate at time $t > s$ can be obtained using a two step recursive algorithm involving a prediction and a correction step. As discussed later in this section, we will devise a sequential algorithm based on an approximate collapsing of $p(\mathbf{x}_s|y_{0:s})$ onto a single Gaussian.

Since the associated densities of a Lévy SSM involve a set of jump times and magnitudes, define $\{(V_i, Z_i)\}_{(s,t]} := \{(V_i, Z_i) : V_i \in (s, t]\}$ for convenience. Then, the predictive density $p(\mathbf{x}_t|y_{0:s}, \{(V_i, Z_i)\}_{(s,t]})$ can be expressed as $\int f(\mathbf{x}_t|\mathbf{x}_s)p(\mathbf{x}_s|y_{0:s})d\mathbf{x}_s$ where the state transition density $f(\mathbf{x}_t|\mathbf{x}_s)$ can be derived from the forward simulation of the SDE, as specified in Eq. (9). Note that the generic notation for the state transition density $f(\mathbf{x}_t|\mathbf{x}_s)$ omits the dependence on the jumps $\{(V_i, Z_i)\}_{(s,t]}$ for notational convenience.

After an observation at time t is measured, the predictive density can be updated as

$$p(\mathbf{x}_t|y_{0:t}, \{(V_i, Z_i)\}_{(s,t]}) = \frac{p(y_t|\mathbf{x}_t)p(\mathbf{x}_t|y_{0:s}, \{(V_i, Z_i)\}_{(s,t]})}{p(y_t|y_{0:s}, \{(V_i, Z_i)\}_{(s,t]})} \tag{10}$$

Since the jump times and sizes $\{(V_i, Z_i)\}_{(s,t]}$ are latent variables in a Lévy SSM, we want to marginalise over them in practice to obtain the filtering estimate at time t as

$$p(\mathbf{x}_t|y_{0:t}) = \int p(\mathbf{x}_t|y_{0:t}, \{(V_i, Z_i)\}_{(s,t]})p(\{(V_i, Z_i)\}_{(s,t]}|y_{0:t})d\{(V_i, Z_i)\}_{(s,t]} \tag{11}$$

where $\int d\{(V_i, Z_i)\}_{(s,t]}$ is used to indicate a marginalisation over all possible jump sequences in $(s, t]$. The Gaussian noise assumption for the observation density $p(y_t|\mathbf{x}_t)$ and the conditionally Gaussian formulation of the transition density allows both the prediction and update equations to be analytically solved and the resulting densities are still conditionally Gaussian. Hence, given a set of jumps $\{(V_i, Z_i)\}_{(s,t]}$ the filtering problem can be solved by standard Kalman filtering recursions.

Assuming that the previous filtering estimate is approximated as a single Gaussian, $p(\mathbf{x}_s|y_{0:s}) = \mathcal{N}(\mathbf{x}_s; \boldsymbol{\mu}_s, \mathbf{C}_s)$, and given $\{(V_i, Z_i)\}_{(s,t]}$, the prediction equations from time s to t result in a Gaussian density $\mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_{t|s}, \mathbf{C}_{t|s})$ such that ([8; 16]),

$$\boldsymbol{\mu}_{t|s} = \mathbf{F}\boldsymbol{\mu}_s + \mathbf{m} \tag{12}$$

$$\mathbf{C}_{t|s} = \mathbf{F}\mathbf{C}_s\mathbf{F}^T + \mathbf{S} \tag{13}$$

where $\mathbf{F} = e^{\mathbf{A}(t-s)}$ and $\mathbf{m}$, $\mathbf{S}$ are calculated using Eqs. (7) and (8). Then, given the observation matrix $\mathbf{H}$ and noise covariance σ_ε^2 in Eq. (2), the Kalman correction step results in another Gaussian density $\mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_{t|t}, \mathbf{C}_{t|t})$ such that

$$\mathbf{K}_t = \mathbf{C}_{t|s}\mathbf{H}^T(\mathbf{H}\mathbf{C}_{t|s}\mathbf{H}^T + \sigma_\varepsilon^2)^{-1} \tag{14}$$

$$\boldsymbol{\mu}_{t|t} = \boldsymbol{\mu}_{t|s} + \mathbf{K}_t(y_t - \mathbf{H}\boldsymbol{\mu}_{t|s}) \tag{15}$$

$$\mathbf{C}_{t|t} = \mathbf{C}_{t|s} - \mathbf{K}_t\mathbf{H}\mathbf{C}_{t|s} \tag{16}$$

where $\mathbf{K}_t$ is the Kalman gain. Additionally, the marginal-conditional likelihood $p(y_t|y_{0:s}, \{(V_i, Z_i)\}_{(s,t]}, \sigma_\varepsilon)$ can be readily evaluated during the recursions as

$$p(y_t|y_{0:s}, \{(V_i, Z_i)\}_{(s,t]}, \sigma_\varepsilon) = \mathcal{N}(y_t; \mathbf{H}\boldsymbol{\mu}_{t|s}, \mathbf{H}\mathbf{C}_{t|s}\mathbf{H}^T + \sigma_\varepsilon^2) \tag{17}$$

which will be used in later stages of our inference algorithm.

While it is straightforward to solve the filtering problem given a set of jumps $\{(V_i, Z_i)\}_{(s,t]}$ as described above, sampling from the posterior density over the latent jumps in (11) is intractable and hence the estimation of the filtering density also becomes intractable. In previous works, this problem was overcome using particle filtering strategies where each simulated set of jumps $\{(V_i, Z_i)\}_{(s,t]}^{(j)}$ were associated with an importance weight w_j such that (11) can be evaluated as a weighted sum over conditionally Gaussian densities $p(\mathbf{x}_t|y_{0:t}, \{(V_i, Z_i)\}_{(s,t]}^{(j)})$. In this work, we propose a new inference framework based on Monte Carlo sampling strategies where samples from the posterior density over the latent jumps are obtained using an MCMC algorithm at each iteration.

The posterior density over the latent jumps in $(s, t]$ can be expressed through Bayes' theorem as

$$p(\{(V_i, Z_i)\}|y_{0:t}) = \frac{p(y_t|y_{0:s}, \{(V_i, Z_i)\})p(\{(V_i, Z_i)\})}{p(y_t|y_{0:s})} \tag{18}$$

where the likelihood term is evaluated as part of the Kalman filtering recursions and shown in (17). The prior density $p(\{(V_i, Z_i)\}_{(s,t]})$ for the jump sequence is intractable to compute, especially for infinite activity processes such as these. Nevertheless, under the truncated shot-noise model, it would be possible to characterise this prior fully and hence use it to perform MCMC by proposing changes to individual jumps, for example. Here we leave such a possibility for future investigations and rely on forward simulation of the jumps from their prior only in the MCMC.

A simple and effective Metropolis–Hastings (MH) algorithm targeting the posterior density over the latent jumps may be constructed by proposing the jumps from their GIG prior $p(\{(V_i, Z_i)\}_{(s,t]})$ using the algorithms in [12]. In this case MH acceptance probabilities are obtained solely in terms of the marginal-conditional likelihood term $p(y_t|y_{0:s}, \{(V_i, Z_i)\}_{(s,t]})$. For each proposed change $\{(V_i, Z_i)\}^{(\prime)}_{(s,t]} \sim p(\{(V_i, Z_i)\}_{(s,t]})$ to the current state of the jump sequence at iteration j, $\{(V_i, Z_i)\}^{(j)}_{(s,t]}$, the Kalman filtering recursions (12)-(17) are carried out and the new jump sequence $\{(V_i, Z_i)\}^{(\prime)}_{(s,t]}$ is accepted with probability

$$\alpha = \min\left(1, \frac{p(y_t|y_{0:s}, \{(V_i, Z_i)\}^{(\prime)}_{(s,t]})}{p(y_t|y_{0:s}, \{(V_i, Z_i)\}^{(j)}_{(s,t]})}\right) \tag{19}$$

otherwise the move is rejected and the jump sequence remains unchanged. Such an MH algorithm is shown to be effective for producing jump samples from the posterior density in Section 4. The chain is run to convergence, and the chain $\{(V_i, Z_i)\}^{(j)}_{(s,t]}$ and their associated filtering estimates $\mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}^{(j)}_{t|t}, \mathbf{C}^{(j)}_{t|t})$ are combined to make a Gaussian mixture model approximation,

$$p(\mathbf{x}_t|y_{0:t}) \approx \frac{1}{N}\sum_{j=1}^{N} \mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}^{(j)}_{t|t}, \mathbf{C}^{(j)}_{t|t}) \tag{20}$$

where N is total number of iterations of the chain. This approximation is then collapsed onto a single Gaussian, in a spirit similar to [22] for the PF, as $p(\mathbf{x}_t|y_{0:t}) \approx \mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_t, \mathbf{C}_t)$ by moment matching,

$$\boldsymbol{\mu}_t = \frac{1}{N}\sum_{j=1}^{N} \boldsymbol{\mu}^{(j)}_{t|t} \tag{21}$$

$$\mathbf{C}_t = \frac{1}{N}\sum_{j=1}^{N} \left[\mathbf{C}^{(j)}_{t|t} + (\boldsymbol{\mu}^{(j)}_{t|t} - \boldsymbol{\mu}_t)(\boldsymbol{\mu}^{(j)}_{t|t} - \boldsymbol{\mu}_t)^T\right] \tag{22}$$

Hence Eq. (20) forms an approximation of the filtering density in Eq. (11). Furthermore the posterior density over the GIG jumps may also be estimated by storing the underlying samples $\{(V_i, Z_i)\}^{(j)}_{(s,t]}$ from the MCMC algorithm. Note that the conditionally Gaussian formulation of the stochastic integral $\mathbf{I}(\mathbf{f}_t)$ greatly simplifies the design of the filtering algorithm by enabling the use of standard Kalman filtering recursions and also reduces the variance in our estimates compared to algorithms involving direct sampling of the GH process. Furthermore, the Gaussian approximation of the mixture filtering density at $t' > t$ in (20) allows the estimation procedure for $(t, t']$ to be independent of the previous jumps $\{(V_i, Z_i)\}_{(0,t]}$, a simplification compared to more complex SMCMC and PF approaches to this problem. Note that our algorithm, involving the approximation of the filtering density by a single Gaussian, is distinct from the more familiar SMCMC approaches in the literature ([17; 18]), and future work will compare the performance of these algorithms and with regular particle filters [16].

4 Applications

In this section we demonstrate the applicability of the proposed inference algorithm for Lévy SSMs driven by a GH processes. Firstly, we present results of a synthetically generated Langevin dynamics

data set with known and fixed parameters. Here, we display the performance of the proposed sequential MCMC methodology in Section 3 on the estimation of the state vector $\mathbf{x}(t)$. We present further results of our formulation of the Langevin dynamics to a British pound (GBP) to Japanese Yen (JPY) foreign exchange rate data set sampled on 07/01/2013 in Appendix B.

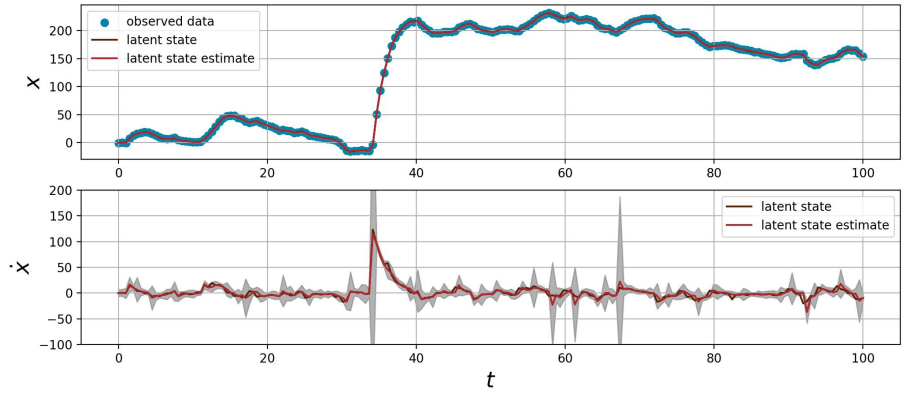

Figure 1: Estimation of states using sequential MCMC for synthetically generated data. With Lévy SSM parameters $\theta = -0.5$, $\mu = 0$, $\sigma = 1$, $\sigma_\varepsilon = 0.1$ and GIG parameters $\lambda = -0.4$, $\gamma = 0.01$, $\delta = 1.$.

For the synthetically generated data set, the parameters of the driving GH process are chosen as $\lambda = -0.4$, $\gamma = 0.01$ and $\delta = 1.0$ which are defined in terms of the underlying GIG subordinator process. Furthermore, $\beta = \mu_W = 0$ which implies that the process is symmetric and the additional scale parameter σ_W in Eq. (5) is chosen to be 1.0. The damping parameter θ for the evolution of the acceleration of a particle moving under Langevin dynamics is set to -0.5 and the observation noise variance σ_ε^2 is set to 0.1. The associated SDE is initialised as $\mathbf{x}(0) = \mathbf{0}$ and the process is simulated between $(0, 100]$ with 200 linearly spaced observations drawn according to Eq. (2) with $\mathbf{H} = [1 \quad 0]$.

The latent states $x(t)$, $\dot{x}(t)$ and the observed data $y(t)$ are visualised in Fig. 1 together with the results of our sequential MCMC filter. Notice that all of the observations are shown in the first plot with blue dots and the second plot showing the velocity of the particle is completely unobserved. Both plots include sequentially computed latent state estimates in addition to 3 standard deviation ranges in grey showing the uncertainty in our estimates. At each time interval 100 samples are generated from the jump proposal and the associated conditional densities are accepted with probability as in (19).

The results suggest that our proposed algorithm is able to accurately infer the true latent states. Specifically, the jump causing a large change in velocity around $t = 35$ is successfully estimated and the model is able to track the state without any lags. The effect of this jump on $x(t)$ can also be seen as a rapid upwards change in position in the first plot. Furthermore, due to the damping effect in Langevin dynamics the velocity $\dot{x}(t)$ gradually reverts to its zero mean levels.

5 Conclusions

In this work, we introduced a Lévy SSM driven by a GH process which provides a flexible representation of continuous-time stochastic linear systems with non-Gaussian properties. We introduced a novel sequential inference algorithm for this model that may be readily generalised to Lévy SSM driven by normal variance-mean processes. A specific formulation of Langevin dynamics is used to demonstrate the applicability of Lévy SSMs and our sequential MCMC algorithm in Section 4. However, it is worth noting that the class of Lévy SSMs is able to represent a wider range of models such as the standard object tracking including the random walk, constant velocity, constant acceleration models. Together with the broad class of processes represented by a GH process, our work enables modelling of a wide range of real-world phenomena with applications in a variety of fields.

References

[1] E. F. Fama, "The behavior of stock-market prices," *The Journal of Business*, vol. 38, no. 1, pp. pp. 34–105, 1965.

[2] D. Yu and L. Deng, *Automatic speech recognition*. Springer, 2016, vol. 1.

[3] J. A. Quinn, C. K. Williams, and N. McIntosh, "Factorial Switching Linear Dynamical Systems Applied to Physiological Condition Monitoring," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 9, pp. 1537–1551, 2009.

[4] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli, "Deep Unsupervised Learning using Nonequilibrium Thermodynamics," in *Proceedings of the 32nd International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 37. PMLR, 07–09 Jul 2015, pp. 2256–2265.

[5] S. Särkkä and L. Svensson, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2023, vol. 17.

[6] S. Godsill and G. Yang, "Bayesian Inference for Continuous-Time ARMA Models Driven by Non-Gaussian Lévy Processes," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 5, 2006, pp. V–V.

[7] U. Şimşekli, "Fractional Langevin Monte carlo: Exploring Levy driven stochastic differential equations for Markov chain Monte Carlo," in *Proceedings of the 34th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, D. Precup and Y. W. Teh, Eds., vol. 70. PMLR, 06–11 Aug 2017, pp. 3200–3209. [Online]. Available: https://proceedings.mlr.press/v70/simsekli17a.html

[8] S. Godsill, M. Riabiz, and I. Kontoyiannis, "The Lévy State Space Model," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, 2019, pp. 487–494.

[9] R. Gan, B. I. Ahmad, and S. J. Godsill, "Lévy State-Space Models for Tracking and Intent Prediction of Highly Maneuverable Objects," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 4, pp. 2021–2038, 2021.

[10] O. Barndorff-Nielsen and C. Halgreen, "Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 38, pp. 309–311, 1977.

[11] N. Shephard, O. E. Barndorff-Nielsen, and U. of Aarhus, "Normal Modified Stable Processes," University of Oxford, Department of Economics, Economics Series Working Papers 72, Jul. 2001. [Online]. Available: https://ideas.repec.org/p/oxf/wpaper/72.html

[12] Y. Kındap and S. Godsill, "Point process simulation of generalised hyperbolic Lévy processes," arXiv 2303.10292, 2023.

[13] P. Brockwell, "Lévy-Driven Carma Processes," *Annals of the Institute of Statistical Mathematics*, vol. 53, pp. 113–124, 02 2001.

[14] O. E. Barndorff-Nielsen and N. Shephard, "Non-Gaussian Ornstein-Uhlenbeck-Based Models and Some of Their Uses in Financial Economics," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 63, no. 2, pp. 167–241, 2001. [Online]. Available: http://www.jstor.org/stable/2680596

[15] G. O. Roberts and R. L. Tweedie, "Exponential convergence of langevin distributions and their discrete approximations," *Bernoulli*, pp. 341–363, 1996.

[16] J. Johnson, Y. Kındap, and S. Godsill, "Inference for Variance-Gamma Driven Stochastic Systems," 2023, https://easychair.org/publications/preprint/1JVC.

[17] F. Septier, S. K. Pang, A. Carmi, and S. Godsill, "On MCMC-Based particle methods for Bayesian filtering: Application to multitarget tracking," in *2009 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2009, pp. 360–363.

[18] R. M. Clavería and S. J. Godsill, "Sequential MCMC Methods for Audio Signal Enhancement," in *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 856–860.

[19] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications (Universitext)*, 6th ed. Springer, Jan. 2014.

[20] M. T. Costa, I. Kontoyiannis, and S. Godsill, "Generalised shot noise representations of stochastic systems driven by non-Gaussian Lévy processes," arXiv 2305.05931, 2023.

[21] O. Cappe, S. J. Godsill, and E. Moulines, "An Overview of Existing Methods and Recent Advances in Sequential Monte Carlo," *Proceedings of the IEEE*, vol. 95, no. 5, pp. 899–924, 2007.

[22] J. Kotecha and P. Djuric, "Gaussian particle filtering," *IEEE Transactions on Signal Processing*, vol. 51, no. 10, pp. 2592–2601, 2003.

A Appendix for the Simulation of GH SSMs

For the Langevin model shown in (3), the exact forms for $e^{\mathbf{A}(t-s)}$, $\mathbf{f}_t(V_i)$ and $\mathbf{f}_t(V_i)\mathbf{f}_t(V_i)^T$ are

$$e^{\mathbf{A}(t-s)} = \begin{bmatrix} 1 & \frac{1}{\theta}\left(f_t(s)-1\right) \\ 0 & f_t(s) \end{bmatrix}$$

$$\mathbf{f}_t(V_i) = \begin{bmatrix} \frac{1}{\theta}(f_t(V_i)-1) \\ f_t(V_i) \end{bmatrix}$$

$$\mathbf{f}_t(V_i)\mathbf{f}_t(V_i)^T = \begin{bmatrix} \frac{1}{\theta^2}\left(f_t(V_i)-1\right)^2 & \frac{1}{\theta}f_t(V_i)\left(f_t(V_i)-1\right) \\ \frac{1}{\theta}f_t(V_i)\left(f_t(V_i)-1\right) & f_t(V_i)^2 \end{bmatrix}$$

where $f_t(x) = e^{\theta(t-x)}$ for convenience. These terms are subsequently substituted into (7), (8) and (9) to simulate incremental sample paths from the Langevin model. This procedure is summarised in Alg. 1.

Algorithm 1 Incremental sampling from the GH SSM, $\{\mathbf{x}(t;\mu_W,\sigma_W^2,\lambda,\delta,\gamma,\theta) : s \leq t\}$, given $\mathbf{x}_s$.

1. Generate jump times and sizes $\{(V_i, Z_i) : V_i \in (s,t]\}$ from a GIG process using Algs. 8 and 10 of [12] for $|\lambda| \geq 0.5$, or Algs. 12 and 14 of [12] for $0 < |\lambda| \leq 0.5$,
2. Compute the moments $\mathbf{m}$ and $\mathbf{S}$ of the associated conditional stochastic integral using the jumps $\{(V_i, Z_i) : V_i \in (s,t]\}$ and Eqs. (7), (8),
3. Using $\mathbf{x}_s$, sample a conditionally Gaussian random variable $\mathbf{x}_t \sim \mathcal{N}(e^{\mathbf{A}(t-s)}\mathbf{x}_s + \mathbf{m}, \mathbf{S})$.

B Appendix for Applications

For the foreign exchange data set, the original data was downsampled to include every 500th irregularly arriving order price for a period of 113 minutes. The observed data used for the experiment includes 200 points in total. The results are shown in Fig. 2 where the x-axis is in seconds. The parameters of the driving GH process as well as the Lévy SSM parameters are tuned by hand using grid-based strategies that maximise the average log marginal likelihood of the data.

The inferred velocity of the rates (lower plot) include rapid shifts, most notably around $t = 4200$, which suggest that a Brownian-driven model would be inappropriate for such a heavy-tailed process. The state on the top plot is able to follow the price accurately without losing track. Hence, the ability to sequentially estimate large deviations in state will likely be of significant assistance in momentum-based trading strategies.

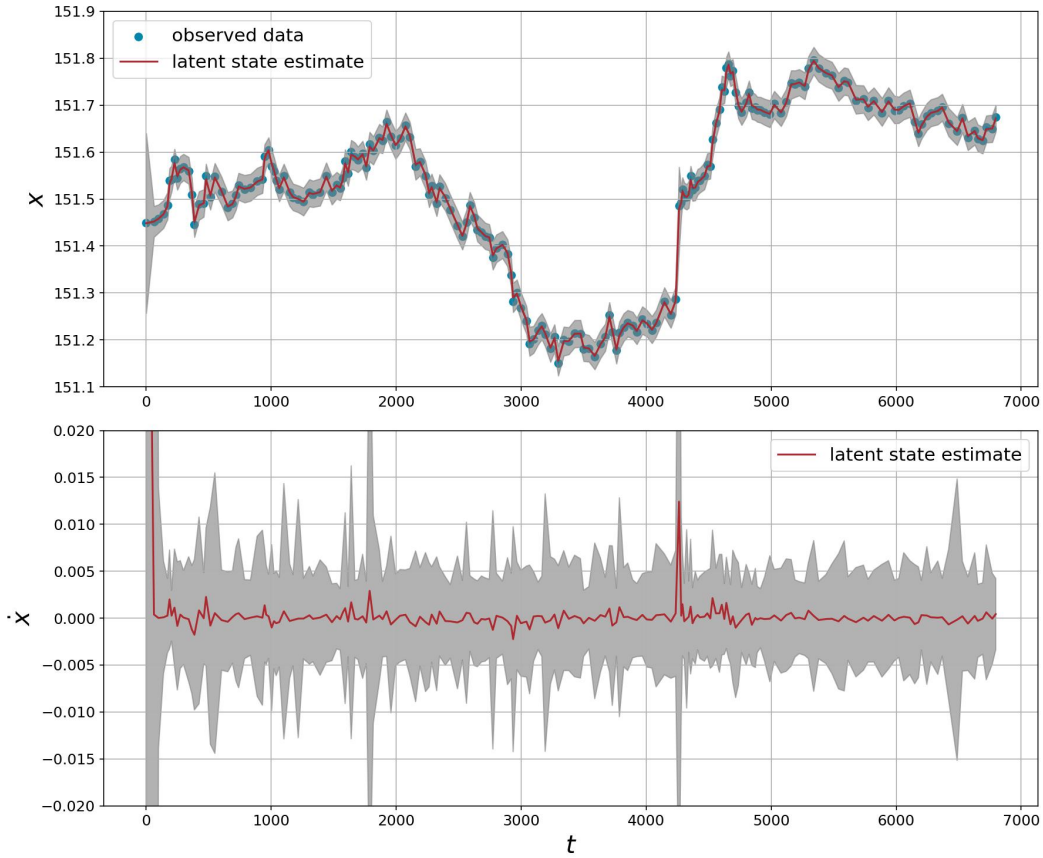

Figure 2: Estimation of states using sequential MCMC for historical intraday foreign exchange rates of British pound (GBP) to Japanese Yen (JPY) sampled on 07/01/2013. With Lévy SSM parameters $\theta = -0.1$, $\mu = 0$, $\sigma = 10^{-4}$, $\sigma_\varepsilon = 10^{-2}$ and GIG parameters $\lambda = -0.6$, $\gamma = 0.001$, $\delta = 1$.