

Bridging between short-range and long-range dependence with mixed spatio-temporal Ornstein-Uhlenbeck processes

MICHELE NGUYEN AND ALMUT E. D. VERAART

Department of Mathematics, Imperial College London

Abstract

While short-range dependence is widely assumed in the literature for its simplicity, long-range dependence is a feature that has been observed in data from finance, hydrology, geophysics and economics. In this paper, we extend a Lévy-driven spatio-temporal Ornstein-Uhlenbeck process by randomly varying its rate parameter to model both short-range and long-range dependence. This particular set-up allows for non-separable spatio-temporal correlations which are desirable for real applications, as well as flexible spatial covariances which arise from the shapes of influence regions. Theoretical properties such as spatio-temporal stationarity and second-order moments are established. An isotropic g -class is also used to illustrate how the memory of the process is related to the probability distribution of the rate parameter. We develop a simulation algorithm for the compound Poisson case which can be used to approximate other Lévy bases. The generalised method of moments is used for inference and simulation experiments are conducted with a view towards asymptotic properties.

Keywords: Long range dependence, Ornstein-Uhlenbeck process, spatio-temporal, compound Poisson, generalised method of moments.

Mathematics Subject Classification: 60G10, 60G55, 60G60, 62F10, 62F12

1 Introduction

Lévy-driven Ornstein-Uhlenbeck (OU) processes are popular modelling tools in finance due to their mean-reverting properties and ability to exhibit non-Gaussianity. To encompass the long memory that has been observed in time series of financial volatility, an extension towards randomly-varying rate parameters was introduced in Barndorff-Nielsen (2001). This results in a superposition of OU processes, i.e. a supOU process, which can be seen as another materialisation of the idea in data traffic modelling and hydrology that long memory arises from a hierarchy or aggregation of processes (Doukhan et al. 2002). The supOU process itself has been studied extensively in its univariate and multivariate contexts as well as in its extremal properties (Barndorff-Nielsen & Stelzer 2011, Fasen & Klüppelberg 2007).

OU processes have also been used in the spatio-temporal setting. While Traulsen et al. (2004) studied a product of a one-dimensional spatial OU process with a temporal OU process, Brix & Diggle (2001) considered a multivariate OU process whose vector components correspond to different spatial locations. In the latter, the authors were motivated by environmental epidemiology and used the OU process as the stochastic intensity process of a log-Gaussian Cox process. To model turbulence, Barndorff-Nielsen & Schmiegel (2003) defined a class of spatio-temporal OU (STOU) processes as stochastic integrals with Lévy noise. This can be seen as a direct spatio-temporal extension of the Lévy-driven OU processes used in finance. We call a random field $\{Y_t(\mathbf{x})\}$ in space-time $\mathcal{X} \times \mathcal{T} = \mathbb{R}^d \times \mathbb{R}$ for some $d \in \mathbb{N}$ a *STOU process* if:

$$Y_t(\mathbf{x}) = \int_{A_t(\mathbf{x})} \exp(-\lambda(t-s)) L(d\xi, ds),$$

where $\lambda > 0$ and L is a homogeneous Lévy basis with finite second moments. The integration set or ambit set, $A_t(\mathbf{x}) \subset \mathcal{X} \times \mathcal{T}$, can be interpreted as a causality cone in physics and satisfies the following conditions:

$$\begin{cases} A_t(\mathbf{x}) = A_0(\mathbf{0}) + (\mathbf{x}, t), & \text{(Translation invariant)} \\ A_s(\mathbf{x}) \subset A_t(\mathbf{x}), \forall s < t, \\ A_t(\mathbf{x}) \cap (\mathcal{X} \times (t, \infty)) = \emptyset. & \text{(Non-anticipative)} \end{cases} \quad (1)$$

Further studies have shown that this class of processes exhibits exponential temporal correlation just like the temporal OU process and boasts flexible spatial correlation structures which are determined by the shape of the ambit set (Nguyen & Veraart 2017). In particular, in one spatial dimension, it has been shown that the well-known exponential and Cauchy spatial correlations can be produced. Non-separable spatio-temporal covariances, which are desirable in practice, can also be obtained.

In this paper, we extend the STOU processes by mixing the rate parameter λ . This will enable us to bridge between short-range and long-range dependence structures in space-time. A *mixed spatio-temporal OU (MSTOU) process* is defined by:

$$Y_t(\mathbf{x}) = \int_0^\infty \int_{A_t(\mathbf{x})} \exp(-\lambda(t-s)) L(d\mathbf{\xi}, ds, d\lambda). \quad (2)$$

Now, L is a Lévy basis over the product space of space-time and the parameter space of λ . In addition, it is no longer homogeneous since we typically associate the parameter space with a probability distribution. Depending on the parameters of this distribution, the process has either short-range or long-range dependence. This extension of STOU processes will be useful for applications where long memory has been observed, for example, in hydrology, geophysics and economics (Frías et al. 2008, Doukhan et al. 2002). One may also apply it in a second-order setting, for example, as the intensity process of a logarithmic Gaussian Cox process (Brix & Diggle 2001).

Outline In the next section, we introduce the background required to understand the construction of (2). In Section 3, we derive the key theoretical properties of the MSTOU process. This includes spatio-temporal stationarity and second-order moments. Particular focus is given to the isotropic g -class and we show that long memory can be obtained for specific parameter ranges of the distribution of λ . By way of an example, we contrast the MSTOU process to another way of defining superpositions of STOU processes which is related to the well-known continuous autoregressive (CAR) process. Unlike the MSTOU process, this alternative definition does not model temporal long memory. In Section 4, we look at the case where L is compound Poisson and simulate from the MSTOU process. Unlike the discrete convolution algorithms for the STOU processes in Nguyen & Veraart (2017), we no longer have kernel discretisation error and only have ambit set approximation error that stems from the kernel truncation. The simulation method can also be used to give second-order approximations for other Lévy bases. In Section 5, we apply the generalised method of moments (GMM) to an MSTOU process. The estimators are shown to be consistent and simulation experiments are conducted to illustrate the finite sample behaviour as well as to provide a view towards asymptotic normality. Finally, we conclude and discuss further directions for research in Section 6.

2 Preliminaries

To understand the definition of an MSTOU process in (2), we rely on the $\mathcal{L}_0$ integration theory in Rajput & Rosinski (1989). Let $S = \mathbb{R}^d \times \mathbb{R} \times (0, \infty)$, the product space of space-time and the parameter space of λ . Further denote the Borel σ -algebra of S by $\mathcal{S} = \mathcal{B}(S)$ and let $\mathcal{B}_b(S)$ contain all its Lebesgue-bounded sets. Then, a Lévy basis is defined as follows (Barndorff-Nielsen et al. 2015):

Definition 1 (Lévy basis).

L is a Lévy basis on $(S, \mathcal{S})$ if it is an independently scattered and infinitely divisible random measure. This means that:

1. $L = \{L(E) : E \in \mathcal{B}_b(S)\}$ is a set of $\mathbb{R}$ -valued random variables such that for a sequence of disjoint elements of $\mathcal{B}_b(S)$, $\{E_i : i \in \mathbb{N}\}$:
 - $L(\bigcup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} L(E_j)$ almost surely when $\bigcup_{j=1}^{\infty} E_j \in \mathcal{B}_b(S)$;
 - and for $i \neq j$, $L(E_i)$ and $L(E_j)$ are independent.
2. Let $B_1, \dots, B_m \in \mathcal{B}_b(S)$ for finite $m \in \mathbb{N}$. The random vector $\mathbf{L} = (L(B_1), \dots, L(B_m))$ is infinitely divisible, i.e. for any $n \in \mathbb{N}$, there exists a law μ_n such that the law of $\mathbf{L}$ can be expressed as $\mu = \mu_n^{*n}$, the n -fold convolution of μ_n with itself.

The distributional properties of a Lévy basis can be summarised through its *characteristic quadruplet (CQ)*:

Assumption 1. The CQ of the Lévy basis in (2) is $(a, b, \nu(dz), d\xi ds \pi(d\lambda))$ where $a \in \mathbb{R}$, $b \geq 0$, ν is a Lévy measure and $\int_0^\infty \pi(d\lambda) = 1$.

We define the *Lévy seed* of L to be the random variable L' with the following Lévy-Khintchine (L-K) representation for its cumulant generating function (CGF):

$$C\{\theta; L'\} = \log \left(e^{i\theta L'} \right) = i\theta a - \frac{1}{2}\theta^2 b + \int_{\mathbb{R}} (e^{i\theta z} - 1 - i\theta z \mathbf{1}_{|z| \leq 1}) \nu(dz),$$

where the logarithm used is the distinguished logarithm (see page 33 of Sato (1999)). The CGF of $L(E)$ for $E \in \mathcal{B}_b(S)$ is related to the CGF of L' as follows:

$$C\{\theta; L(E)\} = \int_E C\{\theta; L'\} d\xi ds \pi(d\lambda).$$

The Lévy basis of a MSTOU process is homogeneous, i.e. has stationary distributions, over space-time, but is inhomogeneous over the parameter space of λ in a manner determined by π . In particular, π can be interpreted as a probability measure for the parameter λ and allows for different values of λ in a possibly continuous way over space-time. Typically, we also assume that π has a density with respect to the Lebesgue measure, $f(\lambda)$.

Now that we have defined the required Lévy basis, we use the classical Rajput & Rosinski (1989) construction where the stochastic integral is defined as the probabilistic limit of a simple function. The following result from Theorem 2.7 of Rajput & Rosinski (1989) will be useful later when we show that the MSTOU processes which we construct are well-defined:

Theorem 1. Let L be a Lévy basis on $(S, \mathcal{S})$ whose CQ is $(a, b, \nu(dz), d\xi ds f(\lambda) d\lambda)$. The measurable function $g : (S, \mathcal{S}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is L -integrable if and only if:

$$\int_S |U(g(\xi, s, \lambda))| f(\lambda) d\xi ds d\lambda < \infty, \quad \int_S b |g(\xi, s, \lambda)|^2 f(\lambda) d\xi ds d\lambda < \infty, \quad \text{and} \quad \int_S V_0(g(\xi, s, \lambda)) f(\lambda) d\xi ds d\lambda < \infty,$$

where $U(u) = ua + \int_{\mathbb{R}} (\rho(zu) - u\rho(z)) \nu(dz)$, $\rho(z) = z \mathbf{1}_{|z| \leq 1}$, and $V_0(u) = \int_{\mathbb{R}} \min(1, |zu|^2) \nu(dz)$.

Remark 1. Here, we have used a slightly different truncation function from the one used in Rajput & Rosinski (1989), i.e. $\rho(z) = z$ if $|z| \leq 1$ and $\frac{z}{|z|}$ if $|z| > 1$.

In the following, we make an $\mathcal{L}_2$ assumption because we will conduct a second order estimation for our MSTOU process later. Note that a weaker condition which involves the logarithmic moment is sufficient for the process to be well-defined.

Assumption 2. The Lévy basis in (2) has finite second moments.

With this assumption, the integrability conditions simplify:

Corollary 1. *Under Assumptions 1 and 2, the MSTOU process is well-defined if:*

$$\int_0^\infty \int_{A_t(\mathbf{x})} \exp(-\lambda(t-s)) f(\lambda) d\xi ds d\lambda < \infty. \quad (3)$$

Proof. This is an extension of the proof of existence for canonical STOU processes on pages 3-4 of the supplementary material of Nguyen & Veraart (2017). $\square$

3 Properties

In the following, assume that Corollary 1 holds. In this section, we investigate the theoretical properties of MSTOU processes.

3.1 Finite-dimensional distribution and stationarity

The distribution of an MSTOU process is determined by its ambit set $A_t(\mathbf{x})$ and the CQ of its Lévy basis. A summary of this can be obtained through its generalised cumulant functional (Barndorff-Nielsen et al. 2015, Nguyen & Veraart 2017):

Definition 2 (Generalised cumulant functional).

For a random field in space-time, $Y = \{Y_t(\mathbf{x})\}_{\mathbf{x} \in \mathbb{R}^d, t \in \mathbb{R}}$, let v denote any non-random measure for which:

$$v(Y) = \int_{\mathbb{R}^d \times \mathbb{R}} Y_t(\mathbf{x}) v(d\mathbf{x}, dt),$$

exists almost surely. The *generalised cumulant functional* (GCF) of Y with respect to v is defined as: $C\{\theta; v(Y)\} = \log \mathbb{E} [\exp(i\theta v(Y))]$.

Theorem 2. *Let Y be an MSTOU process defined by (2) and $A = A_0(\mathbf{0})$. Suppose that for all $\xi \in \mathbb{R}^d$, $s \in \mathbb{R}$ and $\lambda \in (0, \infty)$,*

$$h_A(\xi, s, \lambda) = \int_{\mathbb{R}^d \times \mathbb{R}} \mathbf{1}_A(\xi - \mathbf{x}, s - t) \exp(-\lambda(t-s)) v(d\mathbf{x}, dt) < \infty,$$

and $h_A(\xi, s, \lambda)$ is integrable with respect to the Lévy basis L . Then, the GCF of Y with respect to v can be written as:

$$\begin{aligned} C\{\theta; v(Y)\} &= i\theta a \int_S h_A(\xi, s, \lambda) f(\lambda) d\xi ds d\lambda - \frac{1}{2} \theta^2 b \int_S h_A^2(\xi, s, \lambda) f(\lambda) d\xi ds d\lambda \\ &\quad + \int_S \int_{\mathbb{R}} (\exp(i\theta h_A(\xi, s, \lambda) z) - 1 - i\theta h_A(\xi, s, \lambda) z \mathbf{1}_{|z| \leq 1}) \nu(dz) f(\lambda) d\xi ds d\lambda, \end{aligned} \quad (4)$$

where $(a, b, \nu(dz), d\xi ds f(\lambda) d\lambda)$ is the CQ of L .

Proof. This follows the proofs for Proposition 1 and 5 in Barndorff-Nielsen et al. (2015) with h_A being defined differently to account for space-time, the parameter space of λ and the definition of an MSTOU process. Based on our assumptions and Fubini's theorem:

$$v(Y) = \int_{\mathbb{R}^d \times \mathbb{R}} Y_t(\mathbf{x}) v(d\mathbf{x}, dt) = \int_S \int_{\mathbb{R}^d \times \mathbb{R}} \mathbf{1}_A(\xi - \mathbf{x}, s - t) \exp(-\lambda(t-s)) v(d\mathbf{x}, dt) L(d\xi, ds, d\lambda).$$

Using Proposition 2.6 of Rajput & Rosinski (1989), we obtain the expression for the CGF of $v(Y)$. $\square$

For the marginal and joint distributions of MSTOU processes, we use $v(d\mathbf{x}, dt) = \theta_1 \delta_{t_1}(dt) \delta_{\mathbf{x}_1}(d\mathbf{x}) + \dots + \theta_n \delta_{t_n}(dt) \delta_{\mathbf{x}_n}(d\mathbf{x})$, where $\{(\mathbf{x}_j, t_j) : j = 1, \dots, n\}$ is a set of different spatio-temporal locations and $\theta_j \in \mathbb{R}$ for $j = 1, \dots, n$. With this

specification, $C\{1; v(Y)\}$ is the joint cumulant generating function (JCGF) of $Y_{t_1}(\mathbf{x}_1), \dots, Y_{t_n}(\mathbf{x}_n)$.

Example 1. Let $\pi(\lambda) = \sum_{k=1}^p q_k \delta_{\lambda_k}(\lambda)$ for $\lambda_k > 0$ with $\lambda_k \neq \lambda_{k'}$ for $k \neq k'$, $q_k > 0$ and $p \in \mathbb{N}$ such that $\sum_{k=1}^p q_k = 1$. This corresponds to a discrete probability measure for λ . By substituting the form of $f(\lambda)d\lambda = \pi(d\lambda)$ in (4), we find that the JCGF of the resulting MSTOU process is equal to:

$$\begin{aligned} C\{1; v(Y)\} = & \sum_{i=1}^p \left(i a q_k \int_{\mathbb{R}^d \times \mathbb{R}} h_A(\boldsymbol{\xi}, s, \lambda_k) d\boldsymbol{\xi} ds - \frac{1}{2} b q_k \int_{\mathbb{R}^d \times \mathbb{R}} h_A^2(\boldsymbol{\xi}, s, \lambda_k) d\boldsymbol{\xi} ds \right. \\ & \left. + \int_{\mathbb{R}^d \times \mathbb{R}} \int_{\mathbb{R}} (\exp(i h_A(\boldsymbol{\xi}, s, \lambda_k) z) - 1 - i h_A(\boldsymbol{\xi}, s, \lambda_k) z \mathbf{1}_{|z| \leq 1}) q_k \nu(dz) d\boldsymbol{\xi} ds \right). \end{aligned}$$

From this expression, we find that the MSTOU process is equal in distribution to the superposition of p independent STOU processes:

$$\sum_{k=1}^p \int_{A_t(\mathbf{x})} \exp(-\lambda_k(t-s)) L^{(k)}(d\boldsymbol{\xi}, ds),$$

where $(L^k)_{k=1, \dots, p}$ are independent homogeneous Lévy bases with characteristic triplets $(q_k a, q_k b, q_k \nu(dz, \cdot))$. We note that the STOU processes have the same ambit set but different rate parameters and possibly different characteristic triplets of their Lévy bases.

Definition 3 (Spatio-temporal stationarity).

Let $x_1, \dots, x_n \in \mathbb{R}^d$ and $t_1, \dots, t_n \in \mathbb{R}$ for $n \in \mathbb{N}$. The spatio-temporal random field $Y_t(\mathbf{x})$ is *stationary in space-time* if the joint distribution of $Y_{t_1}(\mathbf{x}_1), \dots, Y_{t_n}(\mathbf{x}_n)$ is the same as that of $Y_{t_1+\epsilon}(\mathbf{x}_1 + \mathbf{u}), \dots, Y_{t_n+\epsilon}(\mathbf{x}_n + \mathbf{u})$ for $\mathbf{u} \in \mathbb{R}^d$ and $\epsilon \in \mathbb{R}$.

Theorem 3. Let $Y_t(\mathbf{x})$ be an MSTOU process. Then $Y_t(\mathbf{x})$ is stationary in space-time.

Proof. This is analogous to the proof of Theorem 3 in Nguyen & Veraart (2017) with $h_A(\boldsymbol{\xi}, s, \lambda)$ replacing $h_A(\boldsymbol{\xi}, s)$. $\square$

Since Y is stationary, its expectation is the same across space-time locations and the covariance between the process at two locations can be written as a function of their distances apart in space and time:

Corollary 2. Let Y be an MSTOU process defined by (2) and $(a, b, \nu(dz), d\boldsymbol{\xi} ds f(\lambda) d\lambda)$ be the CQ of its Lévy basis L . Then, the mean and spatio-temporal covariance of Y are given by:

$$\begin{aligned} \mathbb{E}[Y_t(\mathbf{x})] &= \left[a + \int_{\mathbb{R}} z \nu(dz) \right] \int_0^\infty \int_{A_t(\mathbf{x})} \exp(-\lambda(t-s)) d\boldsymbol{\xi} ds f(\lambda) d\lambda \\ &= \mathbb{E}[L'] \int_0^\infty \int_{A_t(\mathbf{x})} \exp(-\lambda(t-s)) d\boldsymbol{\xi} ds f(\lambda) d\lambda, \\ \text{and } \text{Cov}(Y_t(\mathbf{x}), Y_{t+d_t}(\mathbf{x} + d_{\mathbf{x}})) &= \left[b + \int_{\mathbb{R}} z^2 \nu(dz) \right] \int_0^\infty \int_{A_t(\mathbf{x}) \cap A_{t+d_t}(\mathbf{x} + d_{\mathbf{x}})} \exp(-2\lambda(t-s) - \lambda d_t) d\boldsymbol{\xi} ds f(\lambda) d\lambda \\ &= \text{Var}(L') \int_0^\infty \int_{A_t(\mathbf{x}) \cap A_{t+d_t}(\mathbf{x} + d_{\mathbf{x}})} \exp(-2\lambda(t-s) - \lambda d_t) d\boldsymbol{\xi} ds f(\lambda) d\lambda, \end{aligned} \quad (5)$$

where $d_t \in \mathbb{R}$ and $d_{\mathbf{x}} \in \mathbb{R}^d$ denote displacements in time and space while L' denotes the Lévy seed.

Proof. For the bivariate CGF, we use the result in Theorem 2 with $v(d\mathbf{x}, dt) = \theta_1 \delta_{t_1}(dt) \delta_{\mathbf{x}_1}(d\mathbf{x}) + \theta_2 \delta_{t_2}(dt) \delta_{\mathbf{x}_2}(d\mathbf{x})$, where $(\mathbf{x}_1, t_1)$ and $(\mathbf{x}_2, t_2)$ denote arbitrary locations in space-time. We also set $\theta = 1$. With these specifications, we find that:

$$h_A(\boldsymbol{\xi}, s, \lambda) = \int_{\mathbb{R}^d \times \mathbb{R}} \mathbf{1}_A(\boldsymbol{\xi} - \mathbf{x}, s - t) \exp(-\lambda(t-s)) v(d\mathbf{x}, dt) = \sum_{i=1}^2 \theta_i \mathbf{1}_A(\boldsymbol{\xi} - \mathbf{x}_i, s - t_i) \exp(-\lambda(t_i - s)).$$

Since we can obtain the covariance structure by differentiating the bivariate CGF with respect to θ_1 and θ_2 , and setting

$\theta_1 = \theta_2 = 0$, we are interested in the cross terms, i.e. the terms in $\theta_1 \theta_2$. The first term in (4) does not contain any cross terms:

$$i\theta a \int_S h_A(\xi, s, \lambda) f(\lambda) d\xi ds d\lambda = ia \sum_{i=1}^2 \theta_i \int_0^\infty \int_{A_{t_i}(\mathbf{x}_i)} \exp(-\lambda(t_i - s)) d\xi ds f(\lambda) d\lambda.$$

A cross term appears in the second term of (4):

$$\begin{aligned} -\frac{1}{2}\theta^2 b \int_S h_A^2(\xi, s, \lambda) f(\lambda) d\xi ds d\lambda &= -\frac{1}{2}b \left[\sum_{i=1}^2 \theta_i^2 \int_0^\infty \int_{A_{t_i}(\mathbf{x}_i)} \exp(-2\lambda(t_i - s)) d\xi ds f(\lambda) d\lambda \right. \\ &\quad \left. + 2\theta_1 \theta_2 \int_0^\infty \int_{A_{t_1}(\mathbf{x}_1) \cap A_{t_2}(\mathbf{x}_2)} \exp(-\lambda(t_1 + t_2 - 2s)) d\xi ds f(\lambda) d\lambda \right], \end{aligned}$$

Differentiating the cross term respect to θ_1 and θ_2 , and setting $\theta_1 = \theta_2 = 0$, we have:

$$-b \int_0^\infty \int_{A_{t_1}(\mathbf{x}_1) \cap A_{t_2}(\mathbf{x}_2)} \exp(-\lambda(t_1 + t_2 - 2s)) d\xi ds f(\lambda) d\lambda. \quad (6)$$

By splitting the integration regions into $A_{t_1}(\mathbf{x}_1) \setminus A_{t_2}(\mathbf{x}_2)$, $A_{t_2}(\mathbf{x}_2) \setminus A_{t_1}(\mathbf{x}_1)$ and $A_{t_1}(\mathbf{x}_1) \cap A_{t_2}(\mathbf{x}_2)$, we can express the last term in (4) as:

$$\begin{aligned} &\int_S \int_{\mathbb{R}} (\exp(i\theta h_A(\xi, s, \lambda)z) - 1 - i\theta h_A(\xi, s, \lambda)z \mathbf{1}_{|z| \leq 1}) \nu(dz) f(\lambda) d\xi ds d\lambda \\ &= \int_0^\infty \int_{\mathbb{R}} \int_{A_{t_1}(\mathbf{x}_1) \setminus A_{t_2}(\mathbf{x}_2)} (\exp(i\theta_1 \exp(-\lambda(t_1 - s))z) - 1 - i\theta_1 \exp(-\lambda(t_1 - s))z \mathbf{1}_{|z| \leq 1}) d\xi ds \nu(dz) f(\lambda) d\lambda \\ &+ \int_0^\infty \int_{\mathbb{R}} \int_{A_{t_2}(\mathbf{x}_2) \setminus A_{t_1}(\mathbf{x}_1)} (\exp(i\theta_2 \exp(-\lambda(t_2 - s))z) - 1 - i\theta_2 \exp(-\lambda(t_2 - s))z \mathbf{1}_{|z| \leq 1}) d\xi ds \nu(dz) f(\lambda) d\lambda \\ &+ \int_0^\infty \int_{\mathbb{R}} \int_{A_{t_1}(\mathbf{x}_1) \cap A_{t_2}(\mathbf{x}_2)} (\exp(i[\theta_1 \exp(-\lambda(t_1 - s)) + \theta_2 \exp(-\lambda(t_2 - s))]z) - 1 \\ &- i[\theta_1 \exp(-\lambda(t_1 - s)) + \theta_2 \exp(-\lambda(t_2 - s))]z \mathbf{1}_{|z| \leq 1}) d\xi ds \nu(dz) f(\lambda) d\lambda. \end{aligned} \quad (7)$$

When we differentiate with respect to θ_1 and θ_2 , set $\theta_1 = \theta_2 = 0$, the first two terms of (7) equal to zero. The same procedure on the last term gives:

$$- \int_0^\infty \int_{\mathbb{R}} \int_{A_{t_1}(\mathbf{x}_1) \cap A_{t_2}(\mathbf{x}_2)} z^2 \exp(-\lambda(t_1 + t_2 - 2s)) d\xi ds \nu(dz) f(\lambda) d\lambda.$$

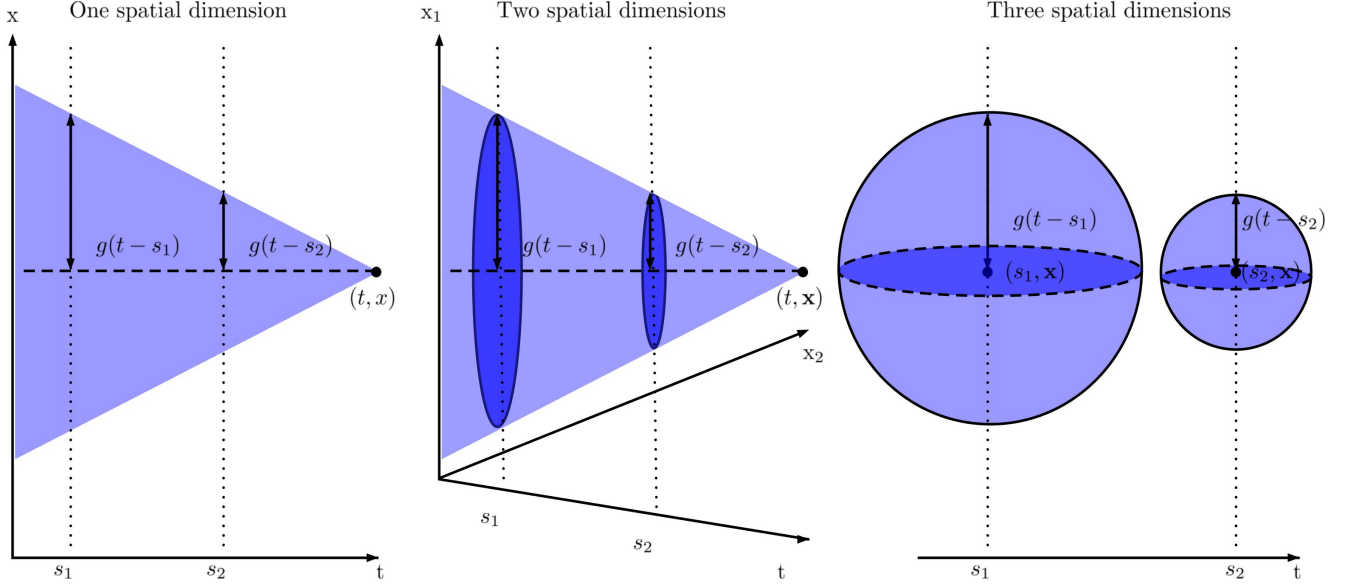
The required expression of the covariance function is obtained by adding this to (6) and multiplying the result by -1 . To obtain the expression for the mean of Y , we differentiate each of the terms in (4) by either θ_1 or θ_2 and set $\theta_1 = \theta_2 = 0$. The mean is then given by multiplying the result by $-i$. $\square$

Remark 2. From (5), we find that the correlation of Y :

$$\text{Corr}(Y_t(\mathbf{x}), Y_{t+d_t}(\mathbf{x} + d_{\mathbf{x}})) = \frac{\int_0^\infty \int_{A_t(\mathbf{x}) \cap A_{t+d_t}(\mathbf{x} + d_{\mathbf{x}})} \exp(-2\lambda(t - s) - \lambda d_t) d\xi ds f(\lambda) d\lambda}{\int_0^\infty \int_{A_t(\mathbf{x})} \exp(-2\lambda(t - s)) d\xi ds f(\lambda) d\lambda}.$$

This means that it depends both on the shape of the integration set $A_t(\mathbf{x})$ and $f(\lambda)$. The additional dependence on the ambit set means that it is harder to establish long memory based on regularly varying characteristics as done in Fasen & Klüppelberg (2007) and Stelzer et al. (2015). Nevertheless, we show in Section 3.3 that long memory can be established in our isotropic g -class of MSTOU processes.

Figure 1: Ambit sets in one to three spatial dimensions for a g -class MTOU process with $g(|t-s|) = c|t-s|$ for some $c > 0$.



3.2 Mixing properties

Spatio-temporal stationarity and mixing properties are useful properties to establish the consistency of moment-based estimators such as the GMM estimators which we construct in Section 5. The following definition is adapted from Passeggeri & Veraart (2017):

Definition 4 (Mixing).

Let Y be a stationary spatio-temporal random field and let $(\mathbf{v}_n = (\mathbf{v}_n^{(S)}, v_n^{(T)}))_{n \in \mathbb{N}}$ be a sequence of spatio-temporal lags such that $\lim_{n \rightarrow \infty} \|\mathbf{v}_n\|_\infty = \infty$ where $\|\cdot\|_\infty$ refers to the supremum norm. We define the shift transformation $\theta_{\mathbf{v}}(B)$ such that $\theta_{\mathbf{v}}(B) = \{\omega' \in \Omega : Y_0(\mathbf{0})(\omega') = Y_{v(T)}(\mathbf{v}^{(S)})(\omega) \text{ for } \omega \in B\}$ for any $B \in \sigma_Y$, the σ -algebra generated by Y . We call Y *mixing* if, for all $A, B \in \sigma_Y$:

$$\lim_{n \rightarrow \infty} P(A \cap \theta_{\mathbf{v}_n}(B)) = P(A)P(B).$$

Next, we will show that a large class of MSTOU processes is mixing.

Definition 5 (g -class processes).

Let $t \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^d$ for $d \in \mathbb{N}$. The g -class of MSTOU processes is the set of MSTOU processes where the ambit sets are given by:

$$A_t(\mathbf{x}) = \{(\boldsymbol{\xi}, s) : s \leq t, |\mathbf{x} - \boldsymbol{\xi}| \leq g(|t-s|)\},$$

for some non-negative and non-decreasing function $g : [0, \infty) \rightarrow \mathbb{R}$.

Figure 1 shows the ambit sets for $g(|t-s|) = c|t-s|$ for some $c > 0$ when we have $d = 1, 2$ and 3 . Due to the exponential kernel in the MSTOU integral, the phenomena observed at a spatial location $\mathbf{x} \in \mathbb{R}^d$ is generally more affected by recent events at nearby locations and less affected by older events at locations further away. However, due to the random rate parameters, different events have different levels of influence. By looking at the temporal cross-sections of the ambit sets (i.e. the spatial ranges for fixed $s \in \mathbb{R}$) when s increases towards t , we see that the news from or effects of surrounding locations travel towards the point of interest. Here, the parameter c is related to the speed of this travel. For $d = 1$, it determines the length of the spatial line of influence from past epochs; for $d = 2$, it determines the radius of the circle of influence; and for $d = 3$, it determines the radius of the sphere of influence. Similar interpretations hold for more general g functions since they are non-decreasing. Specifically, we can modulate the behaviour of the travel by setting g to be, for

example, a quadratic function.

Corollary 3. *Under Assumptions 1 and 2, a g -class MSTOU process is well-defined if:*

$$\int_0^\infty \int_{-\infty}^t g^d(|t-s|) \exp(-\lambda(t-s)) f(\lambda) d\lambda ds < \infty. \quad (8)$$

Proof. The conditions follow from (3) since the temporal cross-section of $A_t(\mathbf{x})$ corresponds to the d -dimensional sphere with centre $(\mathbf{x}, s)$ for $s \leq t$ and radius $g(|t-s|)$. $\square$

Example 2. Consider the case with $g(|t-s|) = c|t-s|$ for some $c > 0$. Then, (8) holds when $\int_0^\infty \frac{1}{\lambda^{d+1}} f(\lambda) d\lambda < \infty$. This is fulfilled for example when $f(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$, the Gamma density with shape and rate parameters, $\alpha > d+1$ and $\beta > 0$ since:

$$\int_0^\infty \frac{1}{\lambda^{d+1}} f(\lambda) d\lambda = \frac{\beta^{d+1}}{(\alpha-1) \dots (\alpha-(d+1))} < \infty.$$

Using the one-dimensional case in Theorem 3.6. of Passeggeri & Veraart (2017), we obtain the following result:

Theorem 4. *Let $Y_t(\mathbf{x})$ be a g -class MSTOU process. Then, Y is mixing.*

Proof. A one-dimensional adaptation of Theorem 3.6 of Passeggeri & Veraart (2017) states that the mixed moving average $X_t := \int_{\mathcal{G}} \int_{\mathbb{R}^d} f(\boldsymbol{\zeta}, t-s) L(d\boldsymbol{\zeta}, ds)$ is mixing. Here, $t \in \mathbb{R}^l$ for some $l \in \mathbb{N}$, $\boldsymbol{\zeta} \in \mathcal{G}$ is a varying parameter and $f : \mathcal{G} \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a measurable function. Using $f = f_{\mathbf{x}}$ for fixed $\mathbf{x} \in \mathbb{R}^d$, a g -class MSTOU process can be written as:

$$Y_t(\mathbf{x}) = \int_0^\infty \int_{\mathbb{R}^d \times \mathbb{R}} \mathbf{1}_{A_t(\mathbf{x})}(\boldsymbol{\zeta}, s) e^{-\lambda(t-s)} L(d\boldsymbol{\zeta}, ds, d\lambda) = \int_{\mathcal{G}} \int_{\mathbb{R}} f_{\mathbf{x}}(\boldsymbol{\zeta}, t-s) L(d\boldsymbol{\zeta}, ds),$$

where $\mathcal{G} = \mathbb{R}^d \times (0, \infty)$, $\boldsymbol{\zeta} = (\zeta_1, \zeta_2) = (\boldsymbol{\xi}, \lambda)$ and:

$$f_{\mathbf{x}}(\boldsymbol{\zeta}, t-s) = \mathbf{1}_{t-s>0} \mathbf{1}_{|\mathbf{x}-\boldsymbol{\zeta}_1| \leq g(t-s)} \exp(-\zeta_2(t-s)).$$

So, $\{Y_t(\mathbf{x})\}_{t \in \mathbb{R}}$ is mixing.

We can also show that Y is spatially mixing by using $f = f_t$ for fixed $t \in \mathbb{R}$ to obtain:

$$Y_t(\mathbf{x}) = \int_0^\infty \int_{\mathbb{R}^d \times \mathbb{R}} \mathbf{1}_{A_t(\mathbf{x})}(\boldsymbol{\zeta}, s) e^{-\lambda(t-s)} L(d\boldsymbol{\zeta}, ds, d\lambda) = \int_{\mathcal{G}} \int_{\mathbb{R}^d} f_t(\boldsymbol{\zeta}, \mathbf{x} - \boldsymbol{\xi}) L(d\boldsymbol{\zeta}, d\boldsymbol{\xi}),$$

is mixing. Now, $\mathcal{G} = \mathbb{R} \times (0, \infty)$, $\boldsymbol{\zeta} = (\zeta_1, \zeta_2) = (s, \lambda)$ and:

$$f_t(\boldsymbol{\zeta}, \mathbf{x} - \boldsymbol{\xi}) = \mathbf{1}_{t-\zeta_1>0} \mathbf{1}_{|\mathbf{x}-\boldsymbol{\xi}| \leq g(t-\zeta_1)} \exp(-\zeta_2(t-\zeta_1)).$$

Next, we prove that spatial and temporal mixing individually imply spatio-temporal mixing. Following the notation in Definition 4, let $(\mathbf{v}_n = (v_n^{(S)}, v_n^{(T)}))_{n \in \mathbb{N}}$ be a sequence of spatio-temporal lags such that $\lim_{n \rightarrow \infty} \|\mathbf{v}_n\|_\infty = \infty$ where $\|\cdot\|_\infty$ refers to the supremum norm. Then, by the definition of the supremum norm, we either have that $\lim_{n \rightarrow \infty} \|v_n^{(S)}\|_\infty = \infty$, $\lim_{n \rightarrow \infty} \|v_n^{(T)}\|_\infty = \infty$ or the case where both limits apply.

First, suppose that $\lim_{n \rightarrow \infty} \|v_n^{(S)}\|_\infty = \infty$. Since $\{Y_t(\mathbf{x})\}_{\mathbf{x} \in \mathbb{R}^d}$ is spatially mixing for fixed $t \in \mathbb{R}$, we have that:

$$\lim_{n \rightarrow \infty} P(A \cap \theta_{\tilde{\mathbf{v}}_n}(B)) = P(A)P(B),$$

where $\tilde{\mathbf{v}} = (v_n^{(S)}, 0)$, $\theta_{\tilde{\mathbf{v}}}(B) = \{\omega' \in \Omega : Y_t(\mathbf{0})(\omega') = Y_t(\mathbf{v}^{(S)})(\omega) \text{ for } \omega \in B\}$ for any A, B contained in the σ -algebra generated by the spatial process. As this holds for every $t \in \mathbb{R}$, we have that:

$$\lim_{n \rightarrow \infty} P(A \cap \theta_{\mathbf{v}_n}(B)) = P(A)P(B),$$

where $A, B \in \sigma_Y$ and $\mathbf{v} = (v_n^{(S)}, v_n^{(T)})$ such that $\lim_{n \rightarrow \infty} \|v_n^{(S)}\|_\infty = \infty$ and $\{v_n^{(T)}\}$ is an arbitrary sequence of temporal lags. The shift transformation $\theta_{\mathbf{v}}(B)$ is given by $\{\omega' \in \Omega : Y_0(\mathbf{0})(\omega') = Y_{v_n^{(T)}}(v_n^{(S)})(\omega) \text{ for } \omega \in B\}$ for any $A, B \in \sigma_Y$.

Now, suppose that $\lim_{n \rightarrow \infty} \|v_n^{(T)}\|_\infty = \infty$. Using similar arguments, we can show that:

$$\lim_{n \rightarrow \infty} P(A \cap \theta_{\mathbf{v}_n}(B)) = P(A)P(B),$$

where $\mathbf{v} = (v_n^{(S)}, v_n^{(T)})$ such that $\lim_{n \rightarrow \infty} \|v_n^{(T)}\|_\infty = \infty$ and $\{v_n^{(S)}\}$ is an arbitrary sequence of spatial lags.

The limit result required for spatio-temporal mixing holds for sequences of spatio-temporal lags such that $\lim_{n \rightarrow \infty} \|v_n^{(S)}\|_\infty = \infty$ or $\lim_{n \rightarrow \infty} \|v_n^{(T)}\|_\infty = \infty$. In addition, the case where $\lim_{n \rightarrow \infty} \|v_n^{(S)}\|_\infty = \infty$ and $\lim_{n \rightarrow \infty} \|v_n^{(T)}\|_\infty = \infty$ is covered in the former two cases. Thus, we conclude that Y is spatio-temporally mixing. $\square$

3.3 Isotropy and long memory in the g -class

Since the ambit sets in the g -class are radially symmetric in space, we find that they exhibit spatial isotropy. We prove this key property and explore the long-range dependence structures that the g -class MSTOU processes can generate.

Definition 6 (Isotropy).

Let $t \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^d$. A spatio-temporal process $Y_t(\mathbf{x})$ is called *isotropic* if its spatial covariance:

$$\text{Cov}(Y_t(\mathbf{x}), Y_t(\mathbf{x} + d_{\mathbf{x}})) = C(|d_{\mathbf{x}}|),$$

for some positive definite function C .

Theorem 5. *Let Y be a g -class MSTOU process. Then, Y is isotropic in space.*

Proof. Fix $t \in \mathbb{R}$. From (5), the spatial covariance of Y is given by:

$$\text{Cov}(Y_t(\mathbf{x}), Y_t(\mathbf{x} + d_{\mathbf{x}})) = \text{Var}(L') \int_0^\infty \int_{A_t(\mathbf{x}) \cap A_t(\mathbf{x} + d_{\mathbf{x}})} \exp(-2\lambda(t-s)) d\xi ds f(\lambda) d\lambda,$$

where $d_{\mathbf{x}} \in \mathbb{R}^d$ denotes the spatial displacement vector while L' denotes the Lévy seed of Y .

Suppose first that $d = 1$, i.e. we have one dimensional space. Without loss of generality, let $d_x \geq 0$, then:

$$\begin{aligned} \text{Cov}(Y_t(x), Y_t(x + d_x)) &= \text{Var}(L') \int_0^\infty \int_{-\infty}^{t-g^{-1}(|d_x|/2)} \int_{x+d_x-g(|t-s|)}^{x+g(|t-s|)} \exp(-2\lambda(t-s)) d\xi ds f(\lambda) d\lambda \\ &= \text{Var}(L') \int_0^\infty \int_{-\infty}^{t-g^{-1}(|d_x|/2)} (2g(|t-s|) - |d_x|) \exp(-2\lambda(t-s)) ds f(\lambda) d\lambda \quad (9) \\ &= \text{Var}(L') \int_0^\infty \int_{g^{-1}(|d_x|/2)}^\infty (2g(w) - |d_x|) \exp(-2\lambda w) dw f(\lambda) d\lambda, \end{aligned}$$

where $w = t - s$. Note that $t - g^{-1}(|d_x|/2)$ denotes the largest temporal coordinate of $A_t(\mathbf{x}) \cap A_t(\mathbf{x} + d_{\mathbf{x}})$ since $A_t(\mathbf{x})$ is radially symmetric and translation invariant. Since the spatial covariance of Y is a function of the spatial distance $|d_x|$, Y is isotropic in space.

For general $d \in \mathbb{N}$, replace $(2g(|t-s|) - d_x)$ in (9) with the volume of the intersection of two d -spheres with the same radius $g(|t-s|)$ and centres at $\mathbf{x}$ and $\mathbf{x} + d_{\mathbf{x}} \in \mathbb{R}^d$. This can be written as the volume of two identical spherical caps (Li 2011):

$$\frac{\pi^{(d-1)/2}}{\Gamma(\frac{d-1}{2} + 1)} g^d(|t-s|) B\left(1 - \left(\frac{|d_{\mathbf{x}}|}{2g(|t-s|)}\right)^2; \frac{d+1}{2}, \frac{1}{2}\right),$$

where B denotes the incomplete beta function. Since this quantity is a function of $|d_{\mathbf{x}}|$, Y is isotropic in space for general $d \in \mathbb{N}$. $\square$

Remark 3. In the proof of Theorem 5, we computed the spatial covariance of a g -class MSTOU process. It was seen that this takes a simpler form when $d = 3$ as compared to $d = 2$. This is because we replace $(2g(t - s) - |d_x|)$ in (9) with $\pi(4g(t - s) + |d_x|)(2g(t - s) - |d_x|)^2/12$ when $d = 3$ which is of a simpler functional form than the $2g^2(|t - s|) \cos^{-1}(|d_x|/2g(t - s)) - (|d_x|\sqrt{4g^2(|t - s|) - |d_x|^2})/2$ used for $d = 2$. Thus, for mathematical simplicity, it may be useful to embed data in two spatial dimensions in a modelling scenario with three spatial dimensions. However, checks should be made to ensure that the assumptions on the additional dimension are reasonable for the context.

The g -class MSTOU processes give rise to spatially isotropic covariances which can be helpful simplifications for real world situations. The following examples show that we can construct non-separable covariances using the g -class. This is desirable for many real applications and means that the spatio-temporal covariances cannot be expressed as products of spatial and temporal covariances (Cressie & Wikle 2011). In what follows, we also obtain explicit expressions which are useful for inference.

Example 3. Consider the scenario in Example 2 with $d = 1$. From Corollary 2, we have that:

$$\begin{aligned}
\mathbb{E}[Y_t(x)] &= \mathbb{E}[L'] \int_0^\infty \int_{A_t(x)} \exp(-\lambda(t - s)) d\xi ds f(\lambda) d\lambda, \\
&= \mathbb{E}[L'] \int_0^\infty \frac{2c}{\lambda^2} f(\lambda) d\lambda, \\
&= \frac{2c\beta^2 \mathbb{E}[L']}{(\alpha - 2)(\alpha - 1)} \int_0^\infty \frac{\beta^{\alpha-2}}{\Gamma(\alpha - 2)} \lambda^{\alpha-2-1} e^{-\beta\lambda} d\lambda \\
&= \frac{2c\beta^2 \mathbb{E}[L']}{(\alpha - 2)(\alpha - 1)}, \\
\text{and } \text{Cov}(Y_t(x), Y_{t+d_t}(x + d_x)) &= \text{Var}(L') \int_0^\infty \int_{A_t(x) \cap A_{t+d_t}(x+d_x)} \exp(-2\lambda(t - s) - \lambda d_t) d\xi ds f(\lambda) d\lambda, \\
&= \text{Var}(L') \int_0^\infty \frac{c}{2\lambda^2} \exp(-\lambda \max(|d_t|, |d_x|/c)) f(\lambda) d\lambda \tag{10} \\
&= \text{Var}(L') \frac{c\beta^\alpha}{2(\beta + B)^{\alpha-2}(\alpha - 2)(\alpha - 1)} \int_0^\infty \frac{(\beta + B)^{\alpha-2}}{\Gamma(\alpha - 2)} \lambda^{\alpha-2-1} e^{-(\beta+B)\lambda} d\lambda \\
&= \text{Var}(L') \frac{c\beta^\alpha}{2(\beta + B)^{\alpha-2}(\alpha - 2)(\alpha - 1)} \\
&= \frac{c\beta^\alpha \text{Var}(L')}{2(\beta + \max(|d_t|, |d_x|/c))^{\alpha-2}(\alpha - 2)(\alpha - 1)},
\end{aligned}$$

where $B = \max(|d_t|, |d_x|/c)$ and (10) holds from the results in Example 3 of Nguyen & Veraart (2017). Since the temporal and spatial distances interact in the expression of the covariance, the latter is non-separable.

Example 4. We consider a specific Lévy basis for the case in Example 3. Let L be a spatio-temporal extension of the *compound Poisson Lévy basis* defined in Fasen & Klüppelberg (2007):

$$L(E) = \sum_{k=-\infty}^{\infty} J_k \mathbf{1}_{\{(\Gamma_k, \lambda_k) \in E\}} \text{ for } E \in \mathcal{B}_b(S),$$

where $\{J_k\}_{k \in \mathbb{N}}$ is a sequence of independent, identically distributed (i.i.d.) random variables with distribution function G , $\{\Gamma_k = (\Gamma_k^{(1)}, \Gamma_k^{(2)})\}_{k \in \mathbb{N}}$ denote the spatio-temporal jump locations of a Poisson process $N = (N_t(\mathbf{x}))_{(\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}}$ with intensity μ , and $\{\lambda_k\}_{k \in \mathbb{N}}$ is an i.i.d. sequence with probability density function f . These three components are independent of each other.

From this definition, we see that the Lévy seed is a compound Poisson random variable, i.e. $L' = \sum_{k=1}^{N_1(1)} J_k$. Its mean and variance are:

$$\mathbb{E}[L'] = \mu \mathbb{E}[J_k] \text{ and } \text{Var}(L') = \mu \left(\text{Var}(J_k) + (\mathbb{E}[J_k])^2 \right).$$

Suppose that $J_k \sim \text{Gamma}(\alpha_Z, \beta_Z)$ for $k \in \mathbb{N}$ then:

$$\mathbb{E}[Y_t(x)] = \frac{2c\beta^2\mu\alpha_Z}{(\alpha-2)(\alpha-1)\beta_Z},$$

$$\text{and } \text{Cov}(Y_t(x), Y_{t+d_t}(x+d_x)) = \frac{c\beta^\alpha\mu\alpha_Z(\alpha_Z+1)}{2(\beta+\max(|d_t|, |d_x|/c))^{\alpha-2}(\alpha-2)(\alpha-1)\beta_Z^2}.$$

To compute the spatio-temporal covariance when $g(|t-s|) = c|t-s|$ for $c > 0$ and $d > 1$, we need to consider $A_t(\mathbf{x}) \cap A_t(\mathbf{x}+d_x)$ in two cases: $|d_x| > cd_t$ and $|d_x| \leq cd_t$ for $d_t \geq 0$. For the former case, the intersection begins at time $t^* = t + (d_t - |d_x|/c)/2$ and the temporal cross-section is equal to the volume of the intersection of two d -spheres with centres $\mathbf{x}$ and $\mathbf{x}+d_x$, and radii $g(|t-s|)$ and $g(|t+d_t-s|)$ respectively; for the latter case, the intersection begins at t and the temporal cross-section is equal to the volume of the d -sphere with centre $\mathbf{x}$ and radius $g(|t-s|)$.

It is more complicated to work out the spatio-temporal covariances for a general g function because the forms of $A_t(\mathbf{x}) \cap A_t(\mathbf{x}+d_x)$ would depend on the curvature of the ambit set. Instead of computing spatio-temporal covariances, we now focus on the spatial and the temporal covariances separately in order to establish short-range or long-range dependence.

Definition 7 (Temporal and spatial short/long-range dependence).

The process $\{Y_t(\mathbf{x}) : t \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^d\}$ is said to have *temporal short-range dependence* if:

$$\int_0^\infty \text{Cov}(Y_t(\mathbf{x}), Y_{t+\tau}(\mathbf{x}))d\tau < \infty,$$

and *temporal long-range dependence* if the integral is infinite.

Similarly, an isotropic process has *spatial short-range dependence* if:

$$\int_0^\infty C(r)dr < \infty,$$

where $\text{Cov}(Y_t(\mathbf{x}), Y_t(\mathbf{x}+d_x)) = C(|d_x|)$ and $r = |d_x|$. It is said to have *spatial long-range dependence* if the integral is infinite.

Example 5. Consider the model used in Example 3. Set $r = d_x = 0$ and $\tau = d_t$. Then:

$$\begin{aligned} \int_0^\infty \text{Cov}(Y_t(x), Y_{t+\tau}(x))d\tau &= \frac{c\beta^\alpha \text{Var}(L')}{2(\alpha-2)(\alpha-1)} \int_0^\infty (\beta+\tau)^{-(\alpha-2)}d\tau \\ &= \frac{c\beta^\alpha \text{Var}(L')}{2(\alpha-2)(\alpha-1)} \left[\frac{(\beta+\tau)^{-(\alpha-3)}}{3-\alpha} \right]_0^\infty \\ &= \frac{c\beta^3 \text{Var}(L')}{2(\alpha-2)(\alpha-1)(\alpha-3)}, \end{aligned}$$

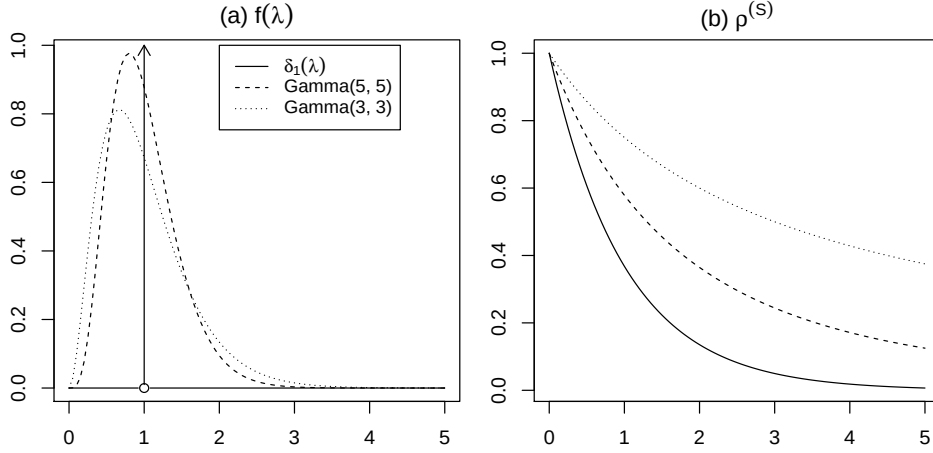
for $\alpha > 3$. For $2 < \alpha \leq 3$, this integral is infinite and the process has temporal long-range dependence. These parameter bounds also apply to spatial long-range dependence since if $r = d_x$ and $\tau = d_t = 0$, we have:

$$\begin{aligned} \int_0^\infty C(r)dr &= \frac{c\beta^\alpha \text{Var}(L')}{2(\alpha-2)(\alpha-1)} \int_0^\infty (\beta+r/c)^{-(\alpha-2)}dr \\ &= \frac{c\beta^\alpha \text{Var}(L')}{2(\alpha-2)(\alpha-1)} \left[\frac{c(\beta+r/c)^{-(\alpha-3)}}{3-\alpha} \right]_0^\infty \\ &= \frac{c^2\beta^3 \text{Var}(L')}{2(\alpha-2)(\alpha-1)(\alpha-3)}, \end{aligned}$$

for $\alpha > 3$. But the integral diverges for $2 < \alpha \leq 3$.

Figure 2 shows three choices of $f(\lambda)$ and the spatial correlation structures of the corresponding MSTOU processes. Here, we have the results for the Dirac delta measure at 1 in bold curves, that for the Gamma(5, 5) density in dashed curves and that for the Gamma(3, 3) density in dotted curves. For all the cases, we have set $c = 1$ so that the temporal correlation

Figure 2: (a) Three choices of $f(\lambda)$ and (b) the spatial correlation structures ($\rho^{(s)}$) of the corresponding MSTOU processes. Since we have set $c = 1$, these share the same forms as the temporal correlations.



function is the same as the spatial one. While the first case leads to exponential correlation, the second and third lead to short-range and long-range dependence respectively.

We can establish long-range dependence for a similar process in three spatial dimensions:

Example 6. Consider three dimensional space ($d = 3$) and the case with $g(|t - s|) = c|t - s|$ for $c > 0$. Let $f(\lambda)$ be the $\text{Gamma}(\alpha, \beta)$ density with $\beta > 0$ and $\alpha > 4$. From the proof of Theorem 5 and Remark 3, we have that the spatial covariance of our process is:

$$\begin{aligned}
 \text{Cov}(Y_t(\mathbf{x}), Y_t(\mathbf{x} + |d_{\mathbf{x}}|)) &= \frac{\pi \text{Var}(L')}{12} \int_0^\infty \int_{|d_{\mathbf{x}}|/2c}^\infty (4cw + |d_{\mathbf{x}}|)(2cw - |d_{\mathbf{x}}|)^2 \exp(-2\lambda w) dw f(\lambda) d\lambda \\
 &= \frac{c^2 \pi \text{Var}(L')}{4} \int_0^\infty \frac{(\lambda |d_{\mathbf{x}}| + 2c)e^{-\lambda |d_{\mathbf{x}}|/c}}{\lambda^4} f(\lambda) d\lambda \\
 &= \frac{\beta^\alpha c^{\alpha-1} \pi \text{Var}(L')}{4(\alpha-4)(\alpha-3)(\alpha-2)(\alpha-1)} (\beta c + |d_{\mathbf{x}}|)^{3-\alpha} (2\beta c + (\alpha-2)|d_{\mathbf{x}}|) \\
 &= \frac{\beta^4 c^3 \pi \text{Var}(L')}{2(\alpha-4)(\alpha-3)(\alpha-2)(\alpha-1)} \left(\frac{\beta c + |d_{\mathbf{x}}|}{\beta c} \right)^{3-\alpha} \left(\frac{2\beta c + (\alpha-2)|d_{\mathbf{x}}|}{2\beta c} \right). \quad (11)
 \end{aligned}$$

Without loss of generality, let $d_t \geq 0$. To compute the temporal covariance, we set $d_{\mathbf{x}} = \mathbf{0}$. Then, $A_t(\mathbf{x}) \cap A_{t+d_t} = A_t(\mathbf{x})$. In this case, the temporal cross-section of $A_t(\mathbf{x})$ corresponds to a sphere with radius $g(|t - s|)$. So:

$$\begin{aligned}
 \text{Cov}(Y_t(\mathbf{x}), Y_{t+d_t}(\mathbf{x})) &= \frac{4\pi \text{Var}(L')}{3} \int_0^\infty \int_0^\infty (cw)^3 \exp(-2\lambda w - \lambda d_t) dw f(\lambda) d\lambda \\
 &= \frac{c^3 \pi \text{Var}(L')}{2} \int_0^\infty \frac{1}{\lambda^4} \exp(-\lambda d_t) f(\lambda) d\lambda \\
 &= \frac{\beta^4 c^3 \pi \text{Var}(L')}{2(\alpha-4)(\alpha-3)(\alpha-2)(\alpha-1)} \left(\frac{\beta}{\beta + d_t} \right)^{\alpha-4}. \quad (12)
 \end{aligned}$$

Using (12), we have that:

$$\begin{aligned}
 \int_0^\infty \text{Cov}(Y_t(\mathbf{x}), Y_{t+\tau}(\mathbf{x})) d\tau &= \int_0^\infty \frac{\beta^4 c^3 \pi \text{Var}(L')}{2(\alpha-4)(\alpha-3)(\alpha-2)(\alpha-1)} \left(\frac{\beta}{\beta + \tau} \right)^{\alpha-4} d\tau \\
 &= \frac{\beta^4 c^3 \pi \text{Var}(L')}{2(\alpha-4)(\alpha-3)(\alpha-2)(\alpha-1)} \left[\frac{(\beta + \tau)^{5-\alpha}}{5-\alpha} \right]_0^\infty \\
 &= \frac{\beta^5 c^3 \pi \text{Var}(L')}{2(\alpha-5)(\alpha-4)(\alpha-3)(\alpha-2)(\alpha-1)},
 \end{aligned}$$

for $\alpha > 5$. For $4 < \alpha \leq 5$, the integral diverges and the process exhibits temporal long-range dependence. Similarly,

using (11), we have:

$$\begin{aligned} & \int_0^\infty C(r) dr \\ &= \int_0^\infty \frac{\beta^4 c^3 \pi \text{Var}(L')}{2(\alpha-4)(\alpha-3)(\alpha-2)(\alpha-1)} \left(\frac{\beta c + r}{\beta c} \right)^{3-\alpha} \left(\frac{2\beta c + (\alpha-2)r}{2\beta c} \right) dr \\ &= \frac{3\beta^5 c^4 \pi \text{Var}(L')}{4(\alpha-5)(\alpha-4)(\alpha-3)(\alpha-2)(\alpha-1)}, \end{aligned}$$

for $\alpha > 5$ and the integral diverges for $4 < \alpha \leq 5$. Under the latter conditions, we have spatial long-range dependence.

3.4 Relation to the spatio-temporal $\text{CAR}_\wedge$ process

In Example 1, we saw that when $f(\lambda)$ is concentrated at p distinct values, MSTOU processes are equal in law to a sum of p independent STOU processes. Here, we consider so-called spatio-temporal $\text{CAR}_\wedge(p)$ processes and show that they too can be represented as superpositions of p STOU processes. However, these STOU processes share the same underlying Lévy basis and are correlated.

Definition 8 (Spatio-temporal $\text{CAR}_\wedge(p)$ process).

We call a random field in space-time $(\mathbb{R}^d \times \mathbb{R})$ a *spatio-temporal $\text{CAR}_\wedge(p)$ process* if:

$$Y_t(\mathbf{x}) = \mathbf{b}^T \mathbf{X}_t(\mathbf{x}),$$

where $\mathbf{b} = (1, 0, \dots, 0)^T \in \mathbb{R}^p$ and:

$$\begin{aligned} \mathbf{X}_t(\mathbf{x}) &= \int_{A_t(\mathbf{x})} \exp(A(t-s)) \mathbf{e}_p L(d\boldsymbol{\xi}, ds), \\ \text{with } A &= \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_p & -a_{p-1} & -a_{p-2} & \dots & -a_1 \end{pmatrix}, \end{aligned}$$

$a_1, \dots, a_{p-1} \geq 0$ and $a_p > 0$. Similar to our ambit set for MSTOU processes, $A_t(\mathbf{x})$ satisfies the conditions in (1). Here, L is a homogeneous Lévy basis and $\mathbf{e}_p$ is the p^{th} Euclidean basis vector.

Theorem 6. Suppose that the matrix A has negative and distinct eigenvalues, $\eta_1, \dots, \eta_p$, then the spatio-temporal $\text{CAR}_\wedge(p)$ process defined in Definition 8 can be written as:

$$Y_t(\mathbf{x}) = \sum_{i=1}^p \frac{1}{\prod_{\substack{1 \leq m \leq p \\ m \neq i}} (\eta_i - \eta_m)} \int_{A_t(\mathbf{x})} \exp(\eta_i(t-s)) L(d\boldsymbol{\xi}, ds).$$

Proof. From Remark 2 of Brockwell et al. (2011), the eigenvectors of A are $\mathbf{v}_i = (1, \eta_i, \eta_i^2, \dots, \eta_i^{p-1})^T$ for $i = 1, \dots, p$. With $V = (v_1 \dots v_p)$, we can write:

$$\exp(A(t-s)) \mathbf{e}_p = V \begin{pmatrix} \exp(\eta_1(t-s)) & 0 & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & 0 & \exp(\eta_p(t-s)) \end{pmatrix} V^{-1} \mathbf{e}_p. \quad (13)$$

Since V is the transpose of a Vandermonde matrix, the term “ $V^{-1} \mathbf{e}_p$ ” which corresponds to the last column of V^{-1} now corresponds to the last row of the Vandermonde matrix inverse. We obtain the required result by using the formulae for

these matrix entries in Exercise 40 in Section 1.2.3 of Knuth (1997) and substituting the corresponding expression for (13) in the definition of $Y_t(\mathbf{x})$. $\square$

The definition of a spatio-temporal $\text{CAR}_\wedge(p)$ process is a spatio-temporal extension of the usual $\text{CAR}(p)$ process where an ambit set is incorporated to allow for non-separable spatio-temporal dependence. When L has finite second moments, integrability conditions similar to (3) are required for the process to be well-defined. Just as the $\text{CAR}(p)$ process does not have temporal long-range dependence (see Remark 8 in Fasen & Klüppelberg (2007)), the construction in Theorem 6 does not result in temporal long-range dependence since if the process is well-defined:

$$\int_0^\infty \text{Cov}(Y_t(\mathbf{x}), Y_{t+\tau}(\mathbf{x})) d\tau = \sum_{i=1}^p \frac{\text{Var}(L')}{(-\eta_i) \prod_{\substack{1 \leq m \leq p \\ m \neq i}} (\eta_i - \eta_m)} \int_{A_t(\mathbf{x})} \exp(2\eta_i(t-s)) d\xi ds < \infty.$$

4 Simulation and compound Poisson MSTOU processes

In this section, we develop a simulation algorithm for MSTOU processes which involves the compound Poisson Lévy basis mentioned in Example 4. This is a generalisation and combination of the simulation algorithms in Brockwell & Matsuda (2017) and Fasen & Klüppelberg (2007) for compound Poisson continuous autoregressive moving average (CARMA) random fields on $\mathbb{R}^d$ and positive shot noise processes on $\mathbb{R}$ respectively. As such, the processes that we simulate can be seen as spatio-temporal shot-noise processes.

We call $Y_t(\mathbf{x})$ a *spatio-temporal shot noise process* if:

$$Y_t(\mathbf{x}) = \int_0^\infty \int_{A_t(\mathbf{x})} e^{-\lambda(t-s)} L(d\xi, ds, d\lambda) = \sum_{k=-\infty}^\infty e^{-\lambda_k(t-\Gamma_k^{(2)})} J_k \mathbf{1}_{A_t(\mathbf{x})}(\Gamma_k), \quad (14)$$

where $\{\lambda_k\}$, $\{J_k\}$ and $\{\Gamma_k\}$ are as defined in Example 4. Similar to the approach in Brockwell & Matsuda (2017) for CARMA random fields, we simulate our MSTOU process over a bounded space-time region D . This means that we approximate (14) by:

$$Z_t(\mathbf{x}) = \sum_{k=1}^M e^{-\lambda_k(t-\Gamma_k^{(2)})} J_k \mathbf{1}_{A_t(\mathbf{x})}(\Gamma_k). \quad (15)$$

Here, M denotes the number of jumps in D and $M \sim \text{Poisson}(\mu \text{Leb}(D))$. The M jump locations are uniformly distributed about D .

Let $\{(\mathbf{x}_i, t_j) : i = 1, \dots, n \text{ and } j = 1, \dots, m\}$ denote a spatio-temporal grid in D . Algorithm 1 shows how we can simulate our MSTOU process in the case of one-dimensional space based on this approximation when f is the Gamma(α, β) density and $Z_k \sim \Gamma(\alpha_Z, \beta_Z)$. To extend this to d -dimensional space, we need to use arrays to store the process values instead of a matrix and extend the ‘for’ loop operations.

To reduce kernel truncation error, we should pad the boundaries of D and implement our algorithm on an extended domain. Since our ambit set $A_t(\mathbf{x})$ does not include times after t , we only need to pad the temporal domain from the past. The extent of the padding and its effectiveness depends on the smallest generated value of λ_k : the smaller the minimum value of λ_k , the wider our extended domain needs to be. A good indicator to monitor would be $\exp(-\lambda_{\min} T_{\text{pad}})$ where $\lambda_{\min} > 0$ and $T_{\text{pad}} > 0$ denote the smallest λ_k and the time padding respectively.

We note that unlike the discrete convolution algorithms for STOU processes in Nguyen & Veraart (2017), we do not have kernel discretisation error since we only evaluate the exponential kernel at jump locations. All our simulation error is due to the kernel truncation imposed by the extent of the padding. There is also no ambit set approximation error except that related to the kernel truncation. This means that we are free to choose a grid size for our simulation domain based on our needs. If we want to simulate processes with longer memory than that corresponding to exponential correlations and estimate from our results, we require data over large areas. Thus, we might want to choose large grid sizes in order to cover a large area in a reasonable computational time. On the other hand, if we are not interested in estimating our simulated data, we can choose finer simulation grids over smaller domains.

Algorithm 1 Simulating a space-time positive shot noise process over a bounded domain with one-dimensional space.

```

1:  $M \leftarrow \text{rpois}(1, \mu \text{Leb}(D))$  ▷ Generate the number of jumps from a Poisson distribution.
2:  $\Gamma \leftarrow \text{runif}(M, D)$  ▷ Generate  $M$  spatio-temporal jump locations from a Uniform distribution over  $D$ .
3:  $\Lambda \leftarrow \text{rgamma}(M, \alpha, \beta)$  ▷ Generate  $M$  rate parameters from a  $\text{Gamma}(\alpha, \beta)$  distribution.
4:  $Z \leftarrow \text{rgamma}(M, \alpha_Z, \beta_Z)$  ▷ Generate  $M$  jump values from a  $\text{Gamma}(\alpha_Z, \beta_Z)$  distribution.
5:  $Y \leftarrow \text{matrix}(0, n, m)$  ▷ Create a storage matrix for our simulated data.
6: for  $i = 1, \dots, n$  do
7:   for  $j = 1, \dots, m$  do
8:      $y \leftarrow 0$  ▷ Create a variable for the  $i - j$ th entry of  $Y$ .
9:     for  $k = 1, \dots, M$  do
10:      if  $\Gamma_k \in A_{t_j}(\mathbf{x}_j)$  then
11:         $y \leftarrow y + e^{-\lambda_k(t_j - \Gamma_k^{(2)})} Z_k$  ▷ Add contribution of  $M$ th jump to the process value if it lies in  $A_t(\mathbf{x})$ .
12:      end if
13:    end for
14:     $Y[i, j] \leftarrow y$  ▷ Store the final process value for the  $i - j$ th location.
15:  end for
16: end for

```

Figure 3: Simulating a canonical spatio-temporal shot noise process with $c = 1$: (a) jumps in extended domain $[-40, 140] \times [-40, 100]$ (in red: jumps in the simulation domain); (b) heat plot over $[0, 100] \times [0, 100]$ where the values are generated with a grid spacing of 0.5 units; (c) perspective plot of the same realisation. The parameter values for the Gamma jump and rate parameter distributions are $\alpha = \alpha_z = 3$ and $\beta = \beta_z = 1$, while the rate parameter for the underlying Poisson process is $\mu = 0.2$.

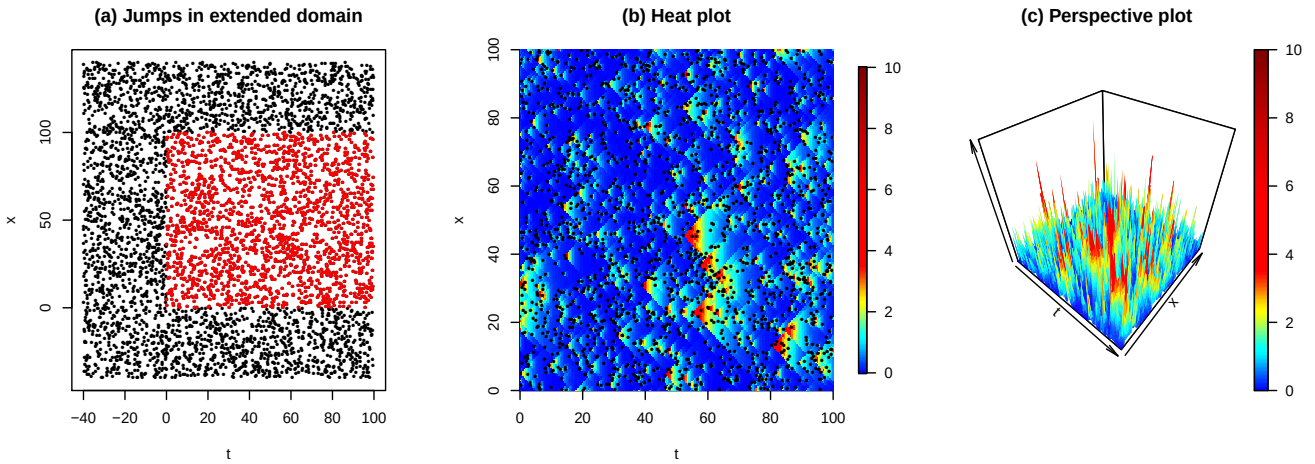
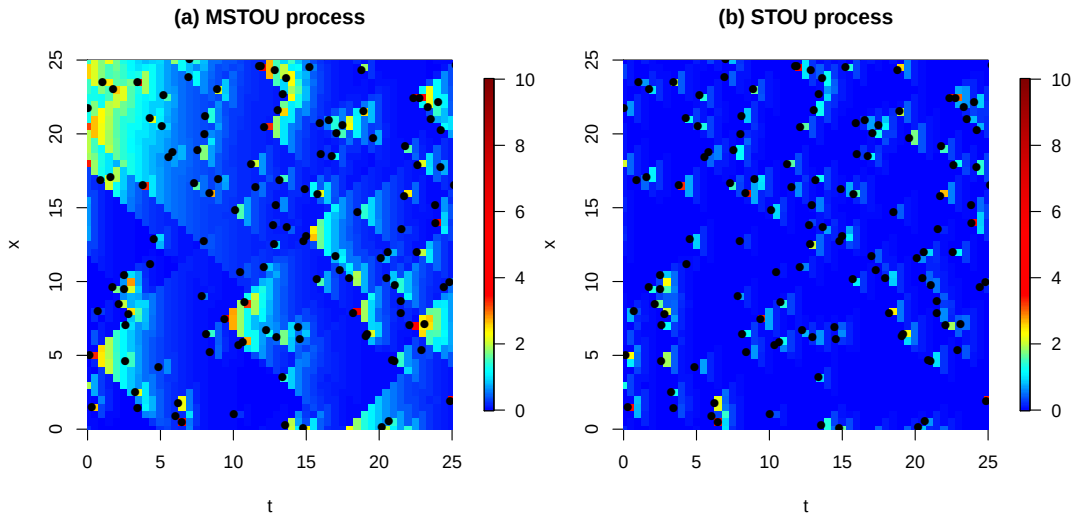


Figure 4: Heat plots over $[0, 25] \times [0, 25]$ of: (a) the MSTOU process and (b) the corresponding STOU process with rate parameter $\int_0^\infty \lambda f(\lambda) d\lambda = \alpha/\beta = 3$. The black dots denote the positions of the jumps in space-time.



Plot (a) in Figure 3 illustrates the simulated jumps in the extended domain of $[-40, 140] \times [-40, 100]$ for a canonical spatio-temporal shot noise process, i.e. the case in one-dimensional space with $A_t(x) = \{(\xi, s) : |x - \xi| < c|t - s|\}$. Here, we have padded the original simulation domain of $[0, 100] \times [0, 100]$ by 40 units in both spatial directions and in the direction towards the past. The rate parameter of the underlying Poisson process is $\mu = 0.2$ and results in 4995 jumps over the extended domain. In this case, we have $\exp(-\lambda_{\min} T_{pad}) = 0.000725$ so the kernel truncation error should be quite small. The other model parameters are $\alpha = \alpha_z = 3$ and $\beta = \beta_z = 1$. In order to cover $[0, 100] \times [0, 100]$ in a reasonable amount of time, we choose a grid size of $\Delta = 0.5$. Plots (b) and (c) show the heat and perspective plots of the corresponding simulation. While the heat plot shows the two-dimensional image of the random field where the colour of each pixel corresponds to the field value at that location, the perspective plot is a three-dimensional graph where the height of the graph at each X - Y coordinate reflects the field value at that location. It is interesting to see that the linear edges of the ambit set are reflected in the heat plot because they determine which jumps affect the process value at a particular location.

For a better understanding of how an MSTOU process works, we zoom into our heat plot and compare our process to a STOU process with the same parameter settings but with its rate parameter set to $\int_0^\infty \lambda f(\lambda) d\lambda$. From Figure 4, we see that the jumps in both processes typically occur at the Poisson jump locations which are denoted by the black dots. However, while the values decay at the same rate for the STOU process in Plot (b), the values decay at varying rates for each jump in the MSTOU process. The lower rate parameters lead to larger clusters which is consistent with the long memory of the process.

Figure 5 shows the series and autocorrelation (ACF) plots for the simulated process at a fixed spatial location ($x = 100$) and a fixed temporal location ($t = 100$). The black curves in the ACF plots denote the theoretical correlations. We see that our simulation replicates the dependence structure of the process quite well although there is some discrepancies at higher lags which are probably due to simulation error, random variation and the lower amount of data for estimation.

For the g -class, we calculate an upper bound for the mean squared error (MSE) as follows:

Theorem 7. *Let $\{Y_t(\mathbf{x})\}_{\mathbf{x} \in \mathbb{R}^d, t \in \mathbb{R}}$ be a spatio-temporal shot noise process in the g -class and let $Z_t(\mathbf{x})$ be its simulation approximation given by (15). Then:*

$$\mathbb{E} \left[(Y_t(\mathbf{x}) - Z_t(\mathbf{x}))^2 \right] \leq \frac{\pi^{d/2} \left(\text{Var}(L') + \mathbb{E}[L']^2 \right)}{\Gamma\left(\frac{d}{2} + 1\right)} \int_0^\infty \left(\int_{\min(T_{pad}, g^{-1}(X_{pad}))}^\infty g^d(w) e^{-2\lambda w} dw \right) f(\lambda) d\lambda, \quad (16)$$

where $X_{pad} > 0$ is the space padding in the simulation.

Proof. By bounding the MSE by that for the boundaries of our simulation domain, we have:

$$\begin{aligned} \mathbb{E} \left[(Y_t(\mathbf{x}) - Z_t(\mathbf{x}))^2 \right] &\leq \mathbb{E} \left[\left(\int_0^\infty \int_{A_t(\mathbf{x}) \setminus [\mathbf{x} - X_{pad}, \mathbf{x} + X_{pad}] \times [t - T_{pad}, t]} e^{-\lambda(t-s)} L(d\boldsymbol{\xi}, ds, d\lambda) \right)^2 \right] \\ &= \left(\text{Var}(L') + \mathbb{E}[L']^2 \right) \int_0^\infty \int_{A_t(\mathbf{x}) \setminus [\mathbf{x} - X_{pad}, \mathbf{x} + X_{pad}] \times [t - T_{pad}, t]} e^{-2\lambda(t-s)} f(\lambda) d\boldsymbol{\xi} ds d\lambda \\ &\leq \frac{\pi^{d/2} \left(\text{Var}(L') + \mathbb{E}[L']^2 \right)}{\Gamma\left(\frac{d}{2} + 1\right)} \int_0^\infty \left(\int_{\min(T_{pad}, g^{-1}(X_{pad}))}^\infty g^d(w) e^{-2\lambda w} dw \right) f(\lambda) d\lambda, \end{aligned}$$

where $w = t - s$, $[\mathbf{x} - X_{pad}, \mathbf{x} + X_{pad}] = [x_1 - X_{pad}, x_1 + X_{pad}] \times \cdots \times [x_d - X_{pad}, x_d + X_{pad}]$ and we have used the fact that the temporal cross-section of the ambit set is the d -dimensional sphere centred at $\mathbf{x}$ with radius $g(|t - s|)$. $\square$

The MSE upper bound (16) shrinks to zero as the padding extents $T_{pad}, X_{pad} \rightarrow \infty$ as we now show for a particular case:

Figure 5: Series and autocorrelation function (ACF) plots for: (a)-(b) $Y_t(100)$ and (c)-(d) $Y_{100}(x)$. Each space and time lag corresponds to one unit. In the ACF plots, the black curves represent the theoretical ACFs.

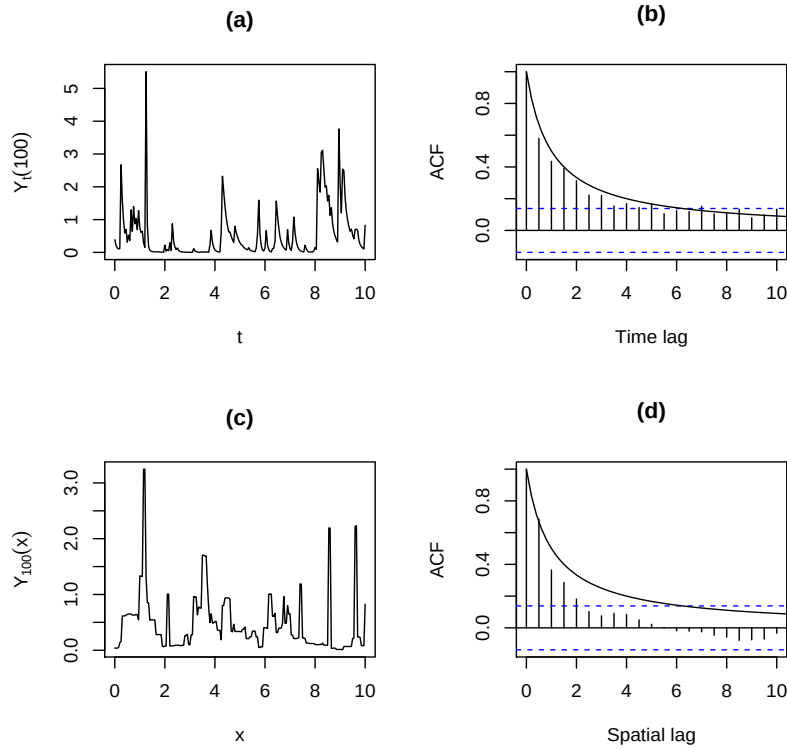
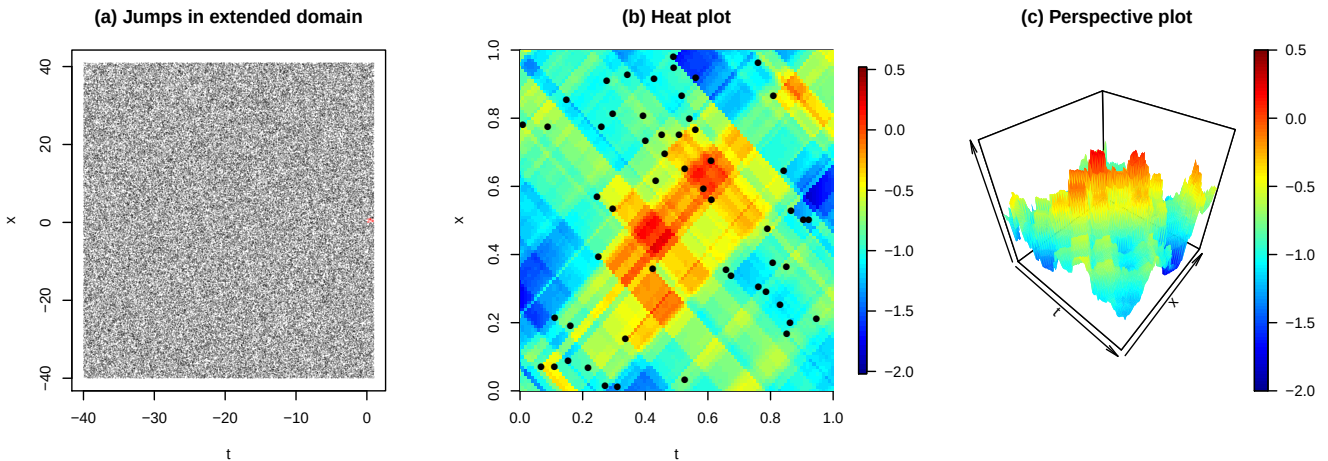


Figure 6: Gaussian Lévy seed approximation: (a) jumps in the extended domain $[-40, 41] \times [-40, 1]$ (in red: jumps in the simulation domain); (b) heat plot over $[0, 1] \times [0, 1]$ for one simulation from a canonical spatio-temporal shot noise process; and (c) the corresponding perspective plot. Here, the grid size is set to 0.01 units, the rate parameter of the underlying Poisson process is $\mu = 40$ and the jumps are normally distributed jumps with mean zero and standard deviation $\sqrt{\frac{\alpha}{\beta\mu}} = \sqrt{3/10}$; These parameter settings mean that the Lévy seed variance is equal to that corresponding to the Gamma distributed jumps in Figure 3, i.e. α/β .



Example 7. Consider the case when $g(|t - s|) = c|t - s|$ for $c > 0$ and $d = 1$. Let $f(\lambda)$ be the Gamma(α, β) density with $\alpha > 2$ and $\beta > 0$. Then, the upper bound on the MSE is given by:

$$\begin{aligned}
 (16) &= c \left(\text{Var}(L') + \mathbb{E}[L']^2 \right) \int_0^\infty \frac{1}{\lambda} \left(\int_{\min(T_{pad}, X_{pad}/c)}^\infty 2\lambda w e^{-2\lambda w} dw \right) f(\lambda) d\lambda \\
 &= \frac{c \left(\text{Var}(L') + \mathbb{E}[L']^2 \right)}{2} \int_0^\infty \frac{1}{\lambda^2} \left(\int_{2\lambda \min(T_{pad}, X_{pad}/c)}^\infty u e^{-u} du \right) f(\lambda) d\lambda \quad \text{where } u = 2\lambda w, \\
 &= \frac{c \left(\text{Var}(L') + \mathbb{E}[L']^2 \right)}{2} \int_0^\infty \frac{1}{\lambda^2} (B\lambda + 1) e^{-B\lambda} f(\lambda) d\lambda \quad \text{where } B = 2 \min(T_{pad}, X_{pad}/c), \\
 &= \frac{c\beta^\alpha \left(\text{Var}(L') + \mathbb{E}[L']^2 \right)}{2(\alpha - 1)} \left[\frac{B}{(\beta + B)^{\alpha-1}} + \frac{1}{(\alpha - 2)(\beta + B)^{\alpha-2}} \right]. \tag{17}
 \end{aligned}$$

From (17), we see that the MSE upper bound converges to zero as T_{pad} and X_{pad} increases. In addition, the rate of convergence increases as α increases. This is in line with the fact that $f(\lambda)$ places more probability weight on larger λ values and large λ values lead to lower kernel truncation error.

As mentioned in Remark 1 of Brockwell & Matsuda (2017) for the simulation of compound Poisson CARMA random fields, we can approximate the first and second moments of other seed distributions by varying the rate of the compound Poisson Lévy seed and the jump distribution. For example, if we allow for positive and negative jumps so that the mean jump size is zero and $\mathbb{E}[Z_k^2] = \sigma^2/\mu$ for μ large, we obtain an approximation of a Gaussian Lévy seed with mean zero and variance σ^2 . More work needs to be done to see what kind of error is incurred by this additional approximation.

In Figure 6, we set $\mu = 40$ and use a Gaussian jump distribution with mean zero and $\sigma^2 = \alpha/\beta = 3$ so that we have the same Lévy seed variance as in the previous case. By virtue of the higher rate parameter, we have more jumps in the same simulation domain (as denoted by the black dots) of smaller size due to the smaller standard deviation. As a result, we get an approximation of the continuous Gaussian Lévy basis.

Remark 4. If we are using the simulation algorithm as an approximation for MSTOU processes driven by Lévy bases other than a compound Poisson one, the approach is related to but not exactly the same as the approximation of infinite activity Lévy processes by a compound Poisson process in Cont & Tankov (2004) since the latter requires a drift term. In the finite variation case (e.g. the IG and Gamma basis), the Lévy-Itô decomposition suggests that we can simulate the process as a sum of a drift term and a compound Poisson process. When the Lévy density $\nu(z)$ exists, the drift term includes the expectation of jumps less than $\epsilon > 0$ which is given by $\int_\epsilon^\infty z\nu(z)dz$ while the compound Poisson process has intensity $U(\epsilon) = \int_\epsilon^\infty \nu(z)dz$ and jump size distribution $p^\epsilon(z) = \frac{\nu(z)\mathbf{1}_{z \geq \epsilon}}{U(\epsilon)}$. As shown in Proposition 6.1 of Cont & Tankov (2004), the error incurred by this approximation can be expressed in terms of ϵ .

Remark 5. The spatio-temporal shot noise process is useful for thinking about data sets for which MSTOU processes are applicable. For example, the jump events could represent flood events in a hydrological setting, news events in house price modelling or outbreaks in an epidemiological context. We note however, that the spatial structure in these applications is very important and more work needs to be done to incorporate covariate information such as the flow direction for flood modelling. Since the MSTOU process can bridge between short-range and long-range dependence, it may be interesting to see if long-range dependence is detected in these different scenarios.

Remark 6. As seen from Algorithm 1, the expected number of iterations required to generate a data set is $\mu \text{Leb}(D) \times n \times m$. This means that the speed of the simulation algorithm depends on the rate parameter of the Poisson process, the extent of the padding, the space-time region we want to cover and the number of simulation grid points. While the first three parameters determine the number of jumps observed and hence the number of additions required for each process value, the latter determines the number of process values to be generated. On a PC with an Intel® Core™i7-3770 CPU Processor @ 3.40GHz, 8GB of RAM and Windows 8.1 64-bit, the data for Figure 3 took about eight minutes to generate while that for Figure 6 took about an hour.

5 Two-step iterated GMM estimation

In this section, we apply the two-step iterated GMM to MSTOU processes. As mentioned in Section 3 of Stelzer et al. (2015) for supOU processes, this is a semi-parametric estimation method since we conduct inference based on second order moments. However, if we assume a particular distribution for the Lévy seed L' which is characterised by two parameters, we can estimate these parameters directly.

5.1 The method

For illustrative purposes, we focus on the case in Example 3. We are interested in estimating $\beta = (\alpha, \beta, c, \mathbb{E}[L'], \text{Var}(L'))$. Suppose that we have data on an $N \times N$ space-time grid with origin (x_0, t_0) and grid size $\Delta > 0$, we define the vector:

$$Y_t(x)^{(m)} := (Y_t(x), Y_t(x + \Delta), \dots, Y_t(x + m\Delta), \dots, Y_{t+\Delta}(x), \dots, Y_{t+m\Delta}(x)),$$

for $t \in \{t_0, \dots, t_0 + (N - m)\Delta\}$ and $x \in \{x_0, \dots, x_0 + (N - m)\Delta\}$. Next, we construct the following moment function:

$$f_Y(Y_t(x)^{(m)}, \beta) = \begin{pmatrix} f_{\mathbb{E}}(Y_t(x)^{(m)}, \beta) \\ f_{\text{Var}}(Y_t(x)^{(m)}, \beta) \\ f_{X,1}(Y_t(x)^{(m)}, \beta) \\ \vdots \\ f_{X,m}(Y_t(x)^{(m)}, \beta) \\ f_{T,1}(Y_t(x)^{(m)}, \beta) \\ \vdots \\ f_{T,m}(Y_t(x)^{(m)}, \beta) \end{pmatrix}, \quad (18)$$

$$\text{where } f_{\mathbb{E}}(Y_t(x)^{(m)}, \beta) = Y_t(x) - \frac{2c\beta^2 \mathbb{E}[L']}{(\alpha - 2)(\alpha - 1)},$$

$$f_{\text{Var}}(Y_t(x)^{(m)}, \beta) = Y_t(x)^2 - \frac{c\beta^2 \text{Var}(L')}{2(\alpha - 2)(\alpha - 1)} - \left(\frac{2c\beta^2 \mathbb{E}[L']}{(\alpha - 2)(\alpha - 1)} \right)^2,$$

$$f_{X,h}(Y_t(x)^{(m)}, \beta) = Y_t(x)Y_t(x + h\Delta) - \frac{c\beta^\alpha \text{Var}(L')}{2(\beta + h\Delta/c)^{\alpha-2}(\alpha - 2)(\alpha - 1)} - \left(\frac{2c\beta^2 \mathbb{E}[L']}{(\alpha - 2)(\alpha - 1)} \right)^2,$$

$$f_{T,h}(Y_t(x)^{(m)}, \beta) = Y_t(x)Y_{t+h\Delta}(x) - \frac{c\beta^\alpha \text{Var}(L')}{2(\beta + h\Delta)^{\alpha-2}(\alpha - 2)(\alpha - 1)} - \left(\frac{2c\beta^2 \mathbb{E}[L']}{(\alpha - 2)(\alpha - 1)} \right)^2,$$

where $h = 1, \dots, m$ and $m \geq 2$ is an integer.

The GMM estimator of β is given by:

$$\hat{\beta}_N = \text{argmin}_{\beta} \{g_{N,m}(Y, \beta)\}' W_N \{g_{N,m}(Y, \beta)\},$$

$$\text{where } g_{N,m}(Y, \beta) = \frac{1}{(N - m)^2} \sum_{i=1}^{N-m} \sum_{j=1}^{N-m} f_Y(Y_{t_0+i\Delta}(x_0 + j\Delta)^{(m)}, \beta). \quad (19)$$

In the first step of the GMM procedure, we set $W_N = I$, the $2(1 + m) \times 2(1 + m)$ identity matrix to find $\hat{\beta}_{1,N}$, the first step estimator. In the second step of the GMM procedure, we set W_N to be $\hat{S}_N^{-1}$ where:

$$\hat{S}_N := \frac{1}{(N - m)^2} \sum_{i=1}^{N-m} \sum_{j=1}^{N-m} f_Y(Y_{t_0+i\Delta}(x_0 + j\Delta)^{(m)}, \hat{\beta}_{1,N}) f_Y(Y_{t_0+i\Delta}(x_0 + j\Delta)^{(m)}, \hat{\beta}_{1,N})'.$$

We note that $\hat{S}_N$ is an estimator for $\bar{V}_N = N \text{Var}(g_{N,m}(Y, \beta_0))$. Improvements can be made by considering the autocor-

relation effects.

Theorem 8. *Let $Y_t(x)$ be the MSTOU process defined in Example 3, $m \geq 2$ be a fixed integer and $f_Y(Y_t(x)^{(m)}, \beta)$ be as defined in (18). Then, the true parameter vector β_0 is identifiable, i.e. $\mathbb{E}[f_Y(Y_t(x)^{(m)}, \beta)] = \mathbf{0}$ for all (x, t) if and only if $\beta = \beta_0$.*

Proof. This is similar to the arguments used to establish identifiability of the GMM estimator for the supOU process in Proposition 3.3 of Stelzer et al. (2015). When $m \geq 2$, we can use the temporal correlations at two different time lags to identify α and β uniquely. These can then be used to determine c from the spatial correlation. Lastly, $\mathbb{E}[L']$ and $\text{Var}(L')$ can be found through the mean and variance. $\square$

In practice, we typically impose some parameter bounds in order to find $\hat{\beta}_N$ via some optimisation procedure. Hence, the following assumption seems reasonable:

Assumption 3. The parameter space Θ is compact.

Provided that this space is large enough to include the true parameter vector, β_0 , weak consistency of the GMM estimators can be established from the spatio-temporal mixing properties of our MSTOU process.

Theorem 9. *Let $Y_t(x)$ be the MSTOU process defined in Example 3, $m \geq 2$ be a fixed integer and $f_Y(Y_t(x)^{(m)}, \beta)$ be as defined in (18). Further, assume that Assumption 3 holds. Then, the GMM estimator, $\hat{\beta}_N$, is weakly consistent, i.e. $\lim_{N \rightarrow \infty} P(|\hat{\beta}_N - \beta_0| > 0) = 0$.*

Proof. We follow the steps in the proof of Theorem 1.1 in Mátyás (1999). This involves checking Assumptions 1-6 in Mátyás (1999).

By construction, Assumption 1.1(i) in Mátyás (1999) holds. Assumption 1.1(ii) in Mátyás (1999) is also satisfied since by Theorem 8, the true parameter vector β_0 is identifiable.

Recall the definition of $g_{N,m}(Y, \beta)$ given in (19). Then, Assumption 1.2 in Mátyás (1999) requires that each vector component of $(g_{N,m}(Y, \beta) - \mathbb{E}[f_Y(Y_t(x)^{(m)}, \beta)])$ converges uniformly in probability to zero for all $\beta \in \Theta$, i.e. $\sup_{\beta \in \Theta} |g_{N,m}^{(i)}(Y, \beta) - \mathbb{E}[f_Y^{(i)}(Y_t(x)^{(m)}, \beta)]| \xrightarrow{P} 0$ as $N \rightarrow \infty$ for $i = 1, \dots, 2(m+1)$. Based on the argument on pages 14-15 of Mátyás (1999), we check that the stated Assumptions 1.4-1.6 hold. Assumption 1.4 in Mátyás (1999) corresponds to Assumption 3 above. For Assumption 1.5, we show that each vector component of $(g_{N,m}(Y, \beta) - \mathbb{E}[f_Y(Y_t(x)^{(m)}, \beta)])$ converges in probability to zero pointwise on Θ as $N \rightarrow \infty$.

Consider $\frac{1}{(N-m)} \sum_{i=1}^{N-m} f_{\mathbb{E}}(Y_{t_0+i\Delta}(x_0 + j\Delta)^{(m)}, \beta)$ for some $j \in \{1, \dots, N-m\}$. Since $\{Y_t(x_0 + j\Delta)\}_{t \in \mathbb{R}}$ is mixing, the sampled process is ergodic and the ergodic average converges in probability to $\mathbb{E}[f_{\mathbb{E}}(Y_t(x_0 + j\Delta), \beta)]$ as $N \rightarrow \infty$. In addition, $\frac{1}{(N-m)} \sum_{i=1}^{N-m} \mathbb{E}[f_{\mathbb{E}}(Y_t(x_0 + j\Delta), \beta)] = \mathbb{E}[f_{\mathbb{E}}(Y_t(x), \beta)]$ because $Y_t(x)$ has spatial stationarity. One can show similar convergences for other vector components of $g_{N,m}(Y, \beta)$ by using an additional fact: suppose that $\{Z_t\}_{t \in \mathbb{R}} = g(\{Y_t(x), Y_{t+\Delta}(x), \dots\})$ where $g: \mathbb{R}^\infty \rightarrow \mathbb{R}$ is a measurable function, then $\{Z_t\}_{t \in \mathbb{R}}$ is an ergodic process (see for example, Theorem 36.4 on page 495 of Billingsley (1995)). Thus, Assumption 1.5 in Mátyás (1999) holds.

Next, we check Assumption 1.6 of Mátyás (1999). From the argument on page 17 of Mátyás (1999), it is sufficient to check a stochastic Lipschitz-type assumption for each vector component of $f_Y(Y_t(x)^{(m)}, \beta)$. By construction, the random terms involved cancel out. For example, for the first component:

$$\left| f_{\mathbb{E}}(Y_t(x)^{(m)}, \beta_1) - f_{\mathbb{E}}(Y_t(x)^{(m)}, \beta_2) \right| = \left| \frac{2c_2\beta_2\mathbb{E}[L']_2}{(\alpha_2-2)(\alpha_2-1)} - \frac{2c_1\beta_1\mathbb{E}[L']_1}{(\alpha_1-2)(\alpha_1-1)} \right|,$$

where $\beta_i = (\alpha_i, \beta_i, c_i, \mathbb{E}[L']_i, \text{Var}(L')_i)$ for $i = 1, 2$.

This means that the stochastic Lipschitz-type condition reduces to a Lipschitz continuity condition on the non-random terms in each vector component of $f_Y(Y_t(x)^{(m)}, \beta)$. Since these terms have bounded first partial derivatives, they are Lipschitz continuous and Assumption 1.6 of Mátyás (1999) holds.

Lastly, we need to show that Assumption 1.3 in Mátyás (1999) is fulfilled. In the first step of our GMM procedure, we used $W_N = I$. Since this is a constant and positive definite matrix, Assumption 1.3 is fulfilled and the first step estimator

$\hat{\beta}_{1,N}$ is consistent for β_0 . In the second step of our GMM procedure, we used $W_N = \hat{S}_N^{-1}$. To show that Assumption 1.3 is satisfied, it is sufficient to show that $S_N \rightarrow S$, as $N \rightarrow \infty$ where S is a constant and positive definite matrix. Indeed, since $\{Y_t(x)\}$ is mixing, $\hat{\beta}_{1,N}$ is consistent and each matrix entry of $\hat{S}_N$ involves measurable functions of the process at different, finite lags, $\hat{S}_N \rightarrow \text{Var}(f_Y(Y_t(x)^{(m)}, \hat{\beta}_{1,n}))$, a constant and positive definite matrix.

Since we have checked all the necessary conditions, the consistency of the GMM estimator follows from the same steps in the proof of Theorem 1.1 in Mátyás (1999) when we replace T , θ and A_T by N , β and W respectively. $\square$

Remark 7. To extend the GMM approach to $d > 1$, we can add additional observations corresponding to other spatial directions to $Y_t(\mathbf{x})^{(m)}$ and adapt $f_{X,h}$ and $f_{T,h}$ accordingly.

5.2 Simulation experiments

GMM estimators are known to be asymptotically normal under certain assumptions (Mátyás 1999). For example, one typically requires that a central limit theorem (CLT) holds for $f_Y(Y_t(x)^{(m)}, \beta)$. However, little work has been done in establishing CLTs for general supOU processes, much less MSTOU processes. So far, it has been shown that CLTs hold for supOU processes when $f(\lambda)$ is a discrete probability distribution with finite support but may not hold under infinite support (Grahovac et al. 2016). It is likely that similar results hold for MSTOU processes but proving them is out of the scope of this paper. Instead, to illustrate our method and strengthen our conjectures about the asymptotic properties, we conduct simulation studies.

We set the simulation domain to $D = [0, 100] \times [0, 100]$, the intensity of the underlying Poisson process to $\mu = 0.2$, the rate parameter of the Gamma distribution for λ to $\beta = 1$, the shape parameter of the ambit set to $c = 1$ and use a $N(0, 15)$ jump distribution. The padding extents and grid size are chosen to be $X_{pad} = T_{pad} = 40$ and $\Delta = 0.5$ respectively. 100 data sets are generated for the short-range dependence case with $\alpha = 5$ and the long-range dependence case with $\alpha = 3$. To have a properly overidentified system and avoid high dimensional matrices, we choose $m = 3$ and conduct the two-step GMM estimation as laid out in Section 5.1. The “DEoptim” function of the “DEoptim” R package was used to perform global optimisation over the parameter space $[2, 35] \times [0, 35] \times [0, 5] \times [-2.5, 2.5] \times [0, 15]$. Figure 7 shows box plots of the estimates for the short-range and long-range dependence scenarios in the top and bottom rows respectively. For a closer look at where the majority of the estimates lie, we have omitted one, five and four outliers for $\hat{\alpha}$, $\hat{\beta}$ and $\hat{c}$ in long-range dependence setting. From the plots, we see that the true parameter values (denoted by the red horizontal lines) lie well within the range of the estimates. We also notice that when the data has short-range dependence, α is never estimated to be lower than 3, the boundary value for long-range dependence (denoted by the dotted blue line). This ability to distinguish between the two forms of dependence is desirable in practice. However, we also note that it is one-sided since $\hat{\alpha} > 3$ for many long-range dependent data sets. There is also some skewness and bias in the estimates which one might expect since we are not in the asymptotic regime.

To illustrate the consistency of our estimators, we repeat our estimation on subsets of our simulated data over the reduced space-time region $[25, 75] \times [50, 100]$. The box plots of the results are shown in Figure 8. As before, a few outliers have been removed to enable us to zoom into majority of the estimates. From the plots, we see that the ranges and bias of the estimates are larger than those for the full data sets. In line with the fact that our estimators are consistent, the estimates become closer to the true values as more data is included in inference in both dependence scenarios. It is also interesting to note that even for the reduced data sets, $\hat{\alpha}$ does not drop below $\alpha = 3$ under short-range dependence.

Next, we look at the normal quantile-quantile (QQ) plots for the full data sets in Figure 9. Apart from $\hat{b}$, there are stronger deviations from normality under the long-range dependence than short-range dependence. Just as the asymptotic distributions for partial sums of one-dimensional transformations of Gaussian processes with finite second moments depend on the Hurst parameter (see Theorem 3.1 of Beran (1994)), it seems reasonable to hypothesize that the asymptotic distribution of the sample mean of a MSTOU process will depend on the strength of the dependence.

Figure 7: Box plots of GMM parameter estimates from 100 data sets over $[0, 100] \times [0, 100]$, shown without extreme outliers. The top row corresponds to the case of short-range dependence ($\alpha = 5$) while the bottom row corresponds to long-range dependence ($\alpha = 3$). The red horizontal lines denote the true parameter values and the blue dotted line in Plot (a) denotes $\alpha = 3$, the boundary value for long-range dependence.

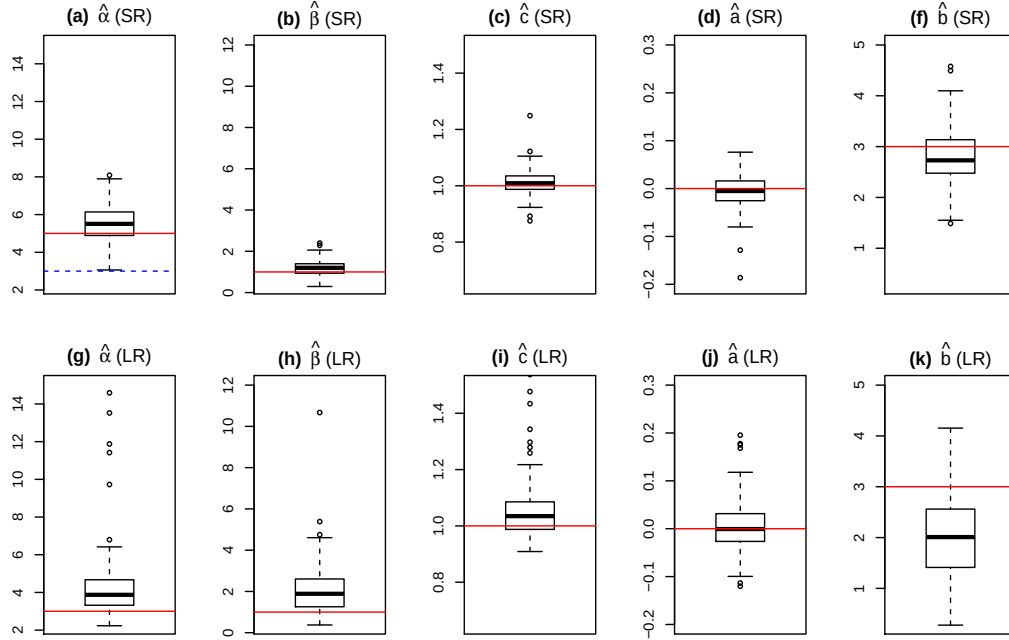


Figure 8: Box plots of GMM parameter estimates from 100 data sets over $[25, 75] \times [50, 100]$, shown without extreme outliers. The top row corresponds to the case of short-range dependence ($\alpha = 5$) while the bottom row corresponds to long-range dependence ($\alpha = 3$). The red horizontal lines denote the true parameter values and the blue dotted line in Plot (a) denotes $\alpha = 3$, the boundary value for long-range dependence.

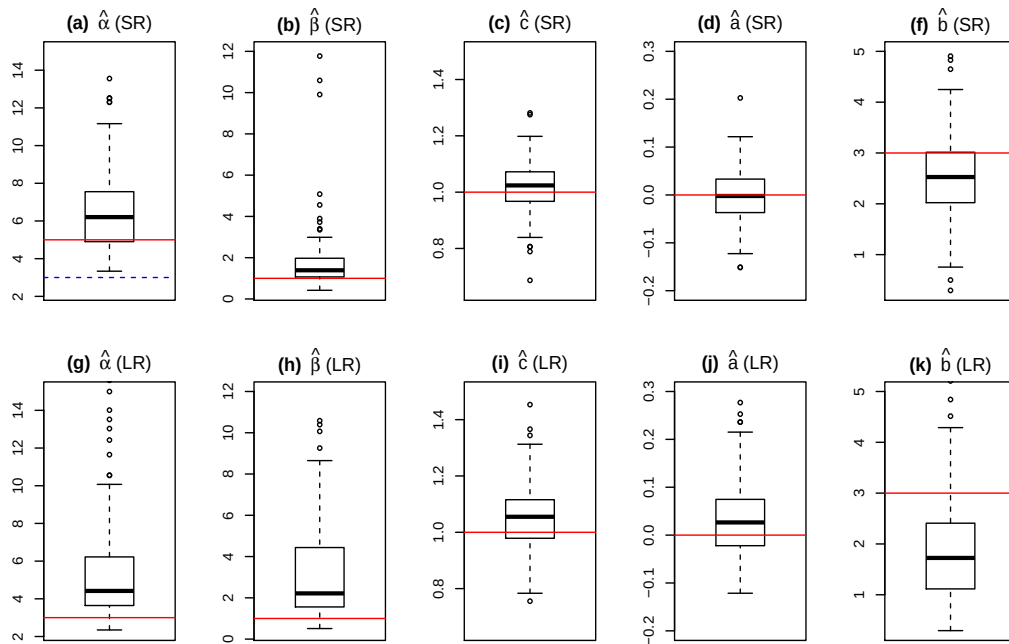
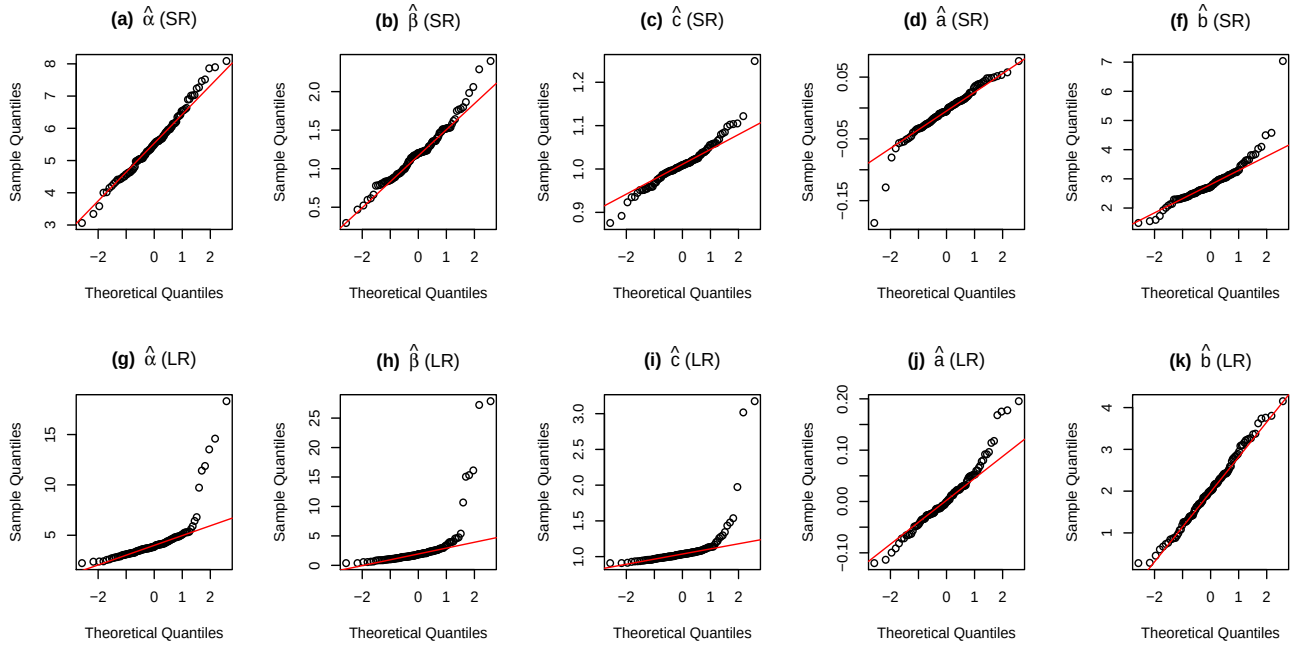


Figure 9: Normal QQ plots of GMM parameter estimates from 100 data sets over $[0, 100] \times [0, 100]$. The top row corresponds to the case of short-range dependence ($\alpha = 5$) while the bottom row corresponds to long-range dependence ($\alpha = 3$).



6 Conclusion and further work

The mixed spatio-temporal Ornstein-Uhlenbeck (MSTOU) process is an extension of the STOU process studied in Barndorff-Nielsen & Schmiegel (2003) and Nguyen & Veraart (2017). While the highlight of this set up is the ability to encompass both short-range and long-range dependence, the MSTOU process also retains the ability to create non-separable spatio-temporal covariances and flexible spatial covariances. This was illustrated for an isotropic class of MSTOU processes, known as the g -class, in Section 3.3.

After developing the theory for MSTOU processes in Sections 2 and 3, we presented a simulation algorithm in Section 4. Unlike the discrete convolution algorithms for STOU process in Nguyen & Veraart (2017), our algorithm does not suffer from the kernel discretisation errors. Instead, the simulation error depends on the kernel truncation: for g -class processes with compound Poisson Lévy bases, an upper bound for the mean squared error was shown to shrink to zero as the simulation padding extents increase to infinity. As already mentioned, it will be useful to better understand the implications of approximating other Lévy bases using our simulation algorithm.

Since we derived the stationarity and second order moments of our processes in Section 3, we applied the two-step iterated generalised method of moments (GMM) to an MSTOU process in Section 5. Promising results were obtained from the simulation experiments. These illustrate the consistency of our estimators and suggest that asymptotic normality depends on the strength of the dependence. More work needs to be done in order to formally establish the latter property. In particular, it would be useful to determine the asymptotic distributions of the sample averages of MSTOU processes.

So far, we have focused mostly on isotropic MSTOU processes. Extending our results to anisotropy via geometric or coordinate-wise means is an interesting direction for further research. While the former assumes isotropy for transformed space-time coordinates, the latter assumes isotropy in individual spatial directions only.

Acknowledgements

We would like to thank the editor, associate editor and two anonymous referees for their constructive feedback. M. Nguyen is grateful to Imperial College for her PhD scholarship which supported this research. A.E.D. Veraart acknowledges financial support by a Marie Curie FP7 Integration Grant (grant agreement number PCIG11-GA-2012-321707) within the 7th European Union Framework Programme.

References

- Barndorff-Nielsen, O. E. (2001), ‘Superposition of Ornstein–Uhlenbeck type processes’, *Theory of Probability & Its Applications* **45**(2), 175–194.
- Barndorff-Nielsen, O. E., Benth, F. E. & Veraart, A. E. D. (2015), ‘Recent advances in ambit stochastics with a view towards tempo-spatial stochastic volatility/intermittency’, *Banach Center Publ.* **104**, 25–60.
- Barndorff-Nielsen, O. E. & Schmiegel, J. (2003), ‘Lévy-based tempo-spatial modelling; with applications to turbulence’, *Uspekhi Mat. Nauk* **159**, 65–91.
- Barndorff-Nielsen, O. E. & Stelzer, R. (2011), ‘Multivariate supOU processes’, *Ann. Appl. Probab.* **21**(1), 140–182.
- Beran, J. (1994), *Statistics for long-memory processes*, Vol. 61, CRC press.
- Billingsley, P. (1995), *Probability and measure*, 3rd edn, John Wiley & Sons.
- Brix, A. & Diggle, P. J. (2001), ‘Spatiotemporal prediction for log-Gaussian Cox processes’, *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63**(4), 823–841.
- Brockwell, P. J., Davis, R. A. & Yang, Y. (2011), ‘Estimation for non-negative Lévy-driven CARMA processes’, *Journal of Business & Economic Statistics* **29**(2), 250–259.
- Brockwell, P. & Matsuda, Y. (2017), ‘Continuous auto-regressive moving average random fields on $\mathbb{R}^n$ ’, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*.
- Cont, R. & Tankov, P. (2004), *Financial modelling with jump processes*, CRC press.
- Cressie, N. & Wikle, C. K. (2011), *Statistics for spatio-temporal data*, John Wiley & Sons.
- Doukhan, P., Oppenheim, G. & Taqqu, M. (2002), *Theory and applications of long-range dependence*, Springer Science & Business Media.
- Fasen, V. & Klüppelberg, C. (2007), Extremes of supOU Processes, in F. E. Benth, G. Di Nunno, T. Lindstrøm, B. Øksendal & T. Zhang, eds, ‘Stochastic Analysis and Applications: The Abel Symposium 2005’, Springer, Berlin, pp. 339–359.
- Frías, M. P., Ruiz-Medina, M. D., Alonso, F. J. & Angulo, J. M. (2008), ‘Spectral-marginal-based estimation of spatiotemporal long-range dependence’, *Communications in Statistics—Theory and Methods* **38**(1), 103–114.
- Grahovac, D., Leonenko, N. N., Sikorskii, A. & Tešnjak, I. (2016), ‘Intermittency of superpositions of Ornstein–Uhlenbeck type processes’, *Journal of Statistical Physics* **165**(2), 390–408.
- Knuth, D. E. (1997), *The Art of Computer Programming: Volume 1: Fundamental Algorithms*, 3rd edn, Addison Wesley Longman Publishing Co., Inc. Redwood City, CA, USA.
- Li, S. (2011), ‘Concise formulas for the area and volume of a hyperspherical cap’, *Asian Journal of Mathematics and Statistics* **4**(1), 66–70.

- Mátyás, L. (1999), *Generalized method of moments estimation*, Vol. 5, Cambridge University Press.
- Nguyen, M. & Veraart, A. E. D. (2017), ‘Spatio-temporal Ornstein–Uhlenbeck processes: theory, simulation and statistical inference’, *Scandinavian Journal of Statistics* **44**(1), 46–80.
- Passeggeri, R. & Veraart, A. E. D. (2017), ‘Mixing properties of multivariate infinitely divisible random fields’. Preprint available at arXiv:1704.02503.
- Rajput, B. S. & Rosinski, J. (1989), ‘Spectral representations of infinitely divisible processes’, *Probability Theory and Related Fields* **82**(3), 451–487.
- Sato, K. (1999), *Lévy processes and infinitely divisible distributions*, Cambridge University Press.
- Stelzer, R., Tosstorff, T. & Wittlinger, M. (2015), ‘Moment based estimation of supOU processes and a related stochastic volatility model’, *Statistics & Risk Modeling* **32**(1), 1–24.
- Traulsen, A., Lippert, K. & Behn, U. (2004), ‘Generation of spatiotemporal correlated noise in 1+ 1 dimensions’, *Phys. Rev. E* **69**(2), 026116–026124.
- Michele Nguyen, Department of Mathematics, Imperial College London, 180 Queen’s Gate, SW7 2AZ London, UK.
Email: michele.nguyen09@imperial.ac.uk