Extremes
https://doi.org/10.1007/s10687-025-00517-4

Tail dependence coefficients of moving average processes driven by exponential-tailed Lévy noise

Zhongwei Zhang[1] · David Bolin[2] · Sebastian Engelke[1] · Raphaël Huser[2]

Received: 5 April 2024 / Revised: 10 February 2025 / Accepted: 24 July 2025

Abstract

Moving average processes driven by exponential-tailed Lévy noise are important extensions of their Gaussian counterparts in order to capture deviations from Gaussianity, more flexible dependence structures, and sample paths with jumps. This paper is concerned with the open problem of determining their extremal dependence structure. We leverage the fact that such processes admit approximations on grids or triangulations that are used in practice for efficient simulations and inference. These approximations can be expressed as special cases of a class of linear transformations of independent, exponential-tailed random variables, that bridge asymptotic dependence and independence in a tractable way. This new fundamental result allows us to show that the integral approximation of general moving average processes with exponential-tailed Lévy noise is asymptotically independent when the mesh is fine enough. Under mild assumptions on the kernel function we also derive the limiting residual tail dependence function. For the popular exponential-tailed Ornstein–Uhlenbeck process we prove that it is asymptotically independent, but with a different residual tail dependence function than its Gaussian counterpart.

Keywords Exponential tail · Extremal dependence · Moving average process · Non-Gaussian OU process · Type G Matérn SPDE field

AMS 2000 Subject Classifications 60G70 · 60G51 · 62H20

Extended author information available on the last page of the article

Published online: 10 September 2025

1 Introduction

1.1 General context

Moving average processes, also referred to as process convolutions, are popular and natural constructions for non-stationary and non-Gaussian processes that are widely applied in spatial statistics (Higdon 2002; Cressie and Pavlicová 2002; Rodrigues and Diggle 2010; Ver Hoef and Peterson 2010). They are defined as

$$u(\boldsymbol{s}) = \int_{\mathcal{D}} G(\|\boldsymbol{s} - \boldsymbol{t}\|)\mathcal{M}(\mathrm{d}\boldsymbol{t}), \tag{1}$$

where $\mathcal{D}$ is a Borel subset of $\mathbb{R}^d$ that possibly depends on $\boldsymbol{s}$, $G : \mathbb{R}_+ \to \mathbb{R}_+$ is a measurable $\mathcal{M}$-integrable function, and $\mathcal{M}$ is an infinitely divisible and independently scattered random measure. A classical example in the temporal domain $d = 1$ is the non-Gaussian Ornstein–Uhlenbeck (OU) process developed in Barndorff-Nielsen and Shephard (2001) to model stochastic volatility in financial econometrics; see also an application of this model to longitudinal data in Asar et al. (2020).

Another related approach for constructing non-Gaussian processes is via a stochastic partial differential equation (SPDE) (Barndorff-Nielsen and Shephard 2001; Bolin 2014; Bolin and Wallin 2020). The stationary solution (if it exists) to such an SPDE admits an integral representation of the form (1). In the spatial domain $\mathbb{R}^2$, an important instance of these constructions are the type G Matérn SPDE random fields (Bolin 2014; Bolin and Wallin 2020), a non-Gaussian extension of the popular Gaussian Matérn SPDE random fields (Lindgren et al. 2011).

A key advantage of the SPDE-based process formulation is that one can approximate its solution by using the finite element method to obtain sparsity in the resulting precision matrix (or inverse of the dispersion matrix in the non-Gaussian case), thus achieving computationally efficient simulation and inference (Lindgren et al. 2011; Bolin 2014; Bolin and Wallin 2020). This finite element approximation has the form of a linear transformation $u_n(\boldsymbol{s}) = \sum_{i=1}^{n} a_i(\boldsymbol{s})Y_i$, where the coefficients $a_i(\boldsymbol{s}) \geq 0$ are determined by the constructed basis functions on the triangulated mesh with *n* mesh nodes, and Y_i are independent Gaussian or non-Gaussian random variables, depending on the random measure $\mathcal{M}$. Non-Gaussian SPDE models are widely used in applications whenever the assumption of Gaussianity is too stringent. As an illustration, the left panel of Fig. 1 shows the triangulation constructed in Bolin and Wallin (2020) for Argo float data, a data set that consists of measurements of seawater temperature and salinity in the global ocean and that has motivated the use of non-Gaussian processes due to its non-Gaussian features such as heavier tails than Gaussian distributions (Kuusela and Stein 2018).

1.2 Extremal dependence and main goal of the paper

The dependence structure of SPDE- and integral-based process constructions has mostly been studied in terms of their Pearson's correlation. This linear dependence measure does, however, not fully describe the process with finite-dimensional non-

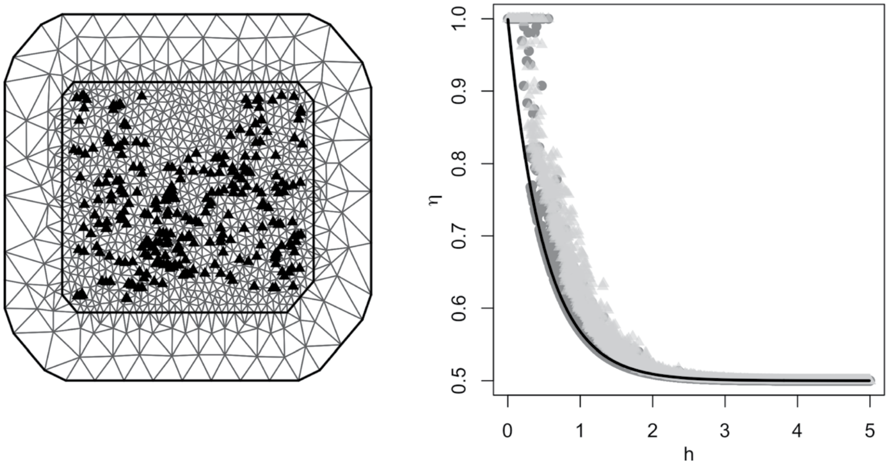


Fig. 1 Observation locations and triangulation mesh used for Argo float data in Bolin and Wallin (2020) (left panel), and its induced residual tail dependence coefficients η (right panel) for all pairs of locations for the finite element (light-grey triangle points) and integral approximations (dark-grey round points) of a type G Matérn SPDE model, and the limiting function (black line)

Gaussian distributions. In fact, the correlation function of these processes is the same regardless of whether Gaussian or non-Gaussian noise is used. Extremal (or tail) dependence describes the strength of dependence in the joint upper or lower tail of a multivariate distribution. It is crucial for risk assessment as it quantifies whether the largest realizations at different locations occur simultaneously. This paper is concerned with the open problem of determining the extremal dependence structure of these popular non-Gaussian processes and their linear approximations that are used in practice.

Let (X_1, X_2) be a random vector with marginal distribution functions F_{X_1} and F_{X_2}, respectively. A commonly used measure of extremal dependence is the (upper) tail dependence coefficient (Coles et al. 1999)

$$\chi = \lim_{q \uparrow 1} \chi(q) = \lim_{q \uparrow 1} \Pr(F_{X_1}(X_1) > q \mid F_{X_2}(X_2) > q), \tag{2}$$

which satisfies $\chi \in [0, 1]$ provided that this limit exists. Since the extremal dependence in the lower tail can be obtained by negating the random vector, here we only focus on the upper tail. We call X_1 and X_2 asymptotically dependent if $\chi > 0$, and asymptotically independent otherwise. In the latter case, the residual tail dependence coefficient η (Ledford and Tawn 1996) measures the second-order extremal dependence behavior and is defined as

$$\eta = \lim_{q \uparrow 1} \frac{\log(1-q)}{\log \Pr(F_{X_1}(X_1) > q, F_{X_2}(X_2) > q)}, \tag{3}$$

provided that this limit exists. We note that $\chi > 0$ implies that $\eta = 1$ and $\eta < 1$ implies that $\chi = 0$; moreover, for $\chi = 0$, the coefficient $1/\eta$ describes the rate of convergence of $\chi(q) \to 0$ as $q \to 1$.

The extremal dependence structure of continuously indexed processes of the form (1) is challenging to derive due to the lack of analytical expressions of the induced multivariate distribution or density functions. We therefore first consider a discretely indexed model of linear transformations

$$\begin{cases} X_1 = a_{11}Y_1 + \cdots + a_{1n}Y_n, \\ X_2 = a_{21}Y_1 + \cdots + a_{2n}Y_n, \end{cases} \tag{4}$$

where $Y_i, i = 1, \ldots, n$, are independent and have exponential tails with the same index, and $a_{ji} \geq 0, j = 1, 2, i = 1, \ldots, n$. This model is motivated by the finite element approximation of the Matérn SPDE fields and integral approximation of general moving average processes. For example, if we approximate (1) by $u_n(\boldsymbol{s}) = \sum_{i=1}^{n} G(\|\boldsymbol{s} - \boldsymbol{t}_i\|)\mathcal{M}(D_i)$ for a partition $\{D_i\}$ of $\mathcal{D}$ and $\boldsymbol{t}_i \in D_i$, and let $X_j = u_n(\boldsymbol{s}_j)$, then $a_{ji} = G(\|\boldsymbol{s}_j - \boldsymbol{t}_i\|)$ and $Y_i = \mathcal{M}(D_i)$. The exponential tail of the noise variables appears naturally in the commonly used non-Gaussian processes. Under mild assumptions we show that (X_1, X_2) is asymptotically dependent if $\arg\max_{i\in\{1,\ldots,n\}} a_{1i} = \arg\max_{i\in\{1,\ldots,n\}} a_{2i}$, and asymptotically independent if $\arg\max_{i\in\{1,\ldots,n\}} a_{1i} \cap \arg\max_{i\in\{1,\ldots,n\}} a_{2i} = \emptyset$. This almost complete characterization of the extremal dependence structure of the model (4) is of independent interest given the popular application of the model in practice. Furthermore, this result also provides a tractable way to bridge asymptotic dependence and independence by simply varying the coefficients in (4). It thus gives a partial answer to the second open problem raised in Nolde and Zhou (2021) who ask how to build flexible parametric models that bridge both extremal dependence regimes.

With this fundamental building block, we then study the extremal dependence structure of general moving average processes with exponential-tailed noise. We show that under mild assumptions on the kernel function and the noise, the integral approximation of such processes is asymptotically independent when the mesh is fine enough, and we derive the limiting residual tail dependence function. Moreover, we prove that the exponential-tailed non-Gaussian OU process (Barndorff-Nielsen and Shephard 2001) is asymptotically independent, but with a different residual tail dependence function than its Gaussian counterpart. As for the Argo float measurements data application, we consider the normal inverse Gaussian (NIG) Matérn SPDE fields, which are a specific class of type G Matérn SPDE fields. The right panel of Fig. 1 shows the residual tail dependence coefficients of the finite element approximation, the integral approximation and its theoretical limiting function, as the triangulation mesh size goes to zero. We conduct additional empirical studies in Section 4.3 and in the Supplementary Material (Section 4) to illustrate our results.

1.3 Related work

There is a vast literature focusing on the extremes of moving average processes, where the interest lies in the tail behavior of the supremum of moving average processes over a compact index set, and the convergence result of the supremums over a sequence of appropriately increasing index sets after certain normalization. For instance, Rootzén (1978) and Leadbetter et al. (1983) studied extremes of discrete- and continuous-time moving average processes with stable noise; Rootzén (1986) studied discrete-time moving average processes with noise having a tail which decays approximately as a polynomial times $\exp(-z^p), p > 0$ as $z \to \infty$; Davis and Resnick (1988) studied extremes of discrete-time moving average processes with convolution-equivalent-tailed noise; an extension of Rootzén (1978) to mixed moving average processes with regularly varying noise was investigated in Fasen (2005, 2006) who studied extremes of continuous-time moving average processes with subexponential noise; Fasen (2009) studied extremes of continuous-time mixed moving average processes with convolution-equivalent-tailed noise; extensions of Fasen (2005, 2006) and Fasen (2009) to spatial random fields are studied in Rønn-Nielsen and Stehr (2022); Stehr and Rønn-Nielsen (2022) and Stehr and Rønn-Nielsen (2021), respectively.

Related literature additionally includes Fasen et al. (2010), Janssen and Drees (2016), Janssen and Drees (2018), Opitz (2018) and Krupskii and Huser (2022), where the focus is on the extremal dependence structures. Specifically, Fasen et al. (2010) studied the first-order extremal dependence structure of discrete- and continuous-time moving average processes with heavy-tailed noise; Janssen and Drees (2016) and Janssen and Drees (2018) studied the extremal behavior of products of powers of regularly varying i.i.d. random variables, which is equivalent to the extremal behavior of linear transformations of exponential-tailed i.i.d. random variables, but the joint probability they investigated is of the form $\Pr(X_1 > s_1 x, X_2 > s_2 x), s_1, s_2 > 0$, differing from $\Pr(F_{X_1}(X_1) > q \mid F_{X_2}(X_2) > q)$ when the marginal distributions of X_1 and X_2 are not the same; Opitz (2018) studied the extremal dependence structure of moving average processes when the kernel function is a special indicator function, and Krupskii and Huser (2022) studied the extremal dependence structure of moving average processes with a special Cauchy noise. By contrast, our interest in this work is the extremal dependence structure of moving average processes when the noise has exponential tails. This relatively light tail allows us to have richer extremal dependence structure in the discrete model, i.e., not only the stronger form of asymptotic dependence, but also the weaker form of asymptotic independence. In the continuous model, the second-order extremal dependence measure, the residual tail dependence function $\eta(\|\boldsymbol{s}_1 - \boldsymbol{s}_2\|)$, turns out to play a more important role than the first-order measure. To the best of our knowledge, the second-order extremal dependence structure induced by general moving average processes has not been studied before.

1.4 Outline

The paper is structured as follows. Section 2 presents the necessary background knowledge on functions with exponential and regularly varying tails. The extremal dependence structure of the discretely indexed model, i.e., linear transformations of the form (4), is discussed in Section 3, followed by that of the moving average processes in Section 4. Section 5 concludes with a discussion. All proofs are given in Section 6 and the Supplementary Material.

2 Preliminaries on exponential-tailed functions

A function $h : \mathbb{R} \to \mathbb{R}_+$ with $\lim_{x\to\infty} h(x) = 0$ is said to have an exponential tail with index $\beta \geq 0$, denoted as $h \in \mathscr{L}_\beta$, if

$$\lim_{x\to\infty} h(x-t)/h(x) = \exp(t\beta), \quad t \in \mathbb{R}.$$

A univariate distribution function F is said to have an exponential right tail if its survival function $\bar{F} = 1 - F$ satisfies $\bar{F} \in \mathscr{L}_\beta$ for some $\beta \geq 0$, and F has an exponential left tail if $h(x) = F(-x) \in \mathscr{L}_\beta$. This is different from the conventional notation $F \in \mathscr{L}_\beta$, where the set $\mathscr{L}_\beta$ is restricted to the family of all exponential-right-tailed distribution functions. Our definition of $\mathscr{L}_\beta$ for general, nonnegative functions allows us to also cover exponential-tailed density functions. If $\bar{F} \in \mathscr{L}_\beta$ is such that we further have $\lim_{x\to\infty} \overline{F * F}(x)/\bar{F}(x) = M < \infty$, where $*$ denotes the convolution operator, then F is called convolution tail equivalent, denoted by $\bar{F} \in \mathscr{S}_\beta$. Clearly, $\mathscr{S}_\beta \subset \mathscr{L}_\beta$.

A function $g : \mathbb{R}_+ \to \mathbb{R}_+$ is called regularly varying at ∞ with index $\rho \in \mathbb{R}$ if for any $t > 0$, $\lim_{x\to\infty} g(tx)/g(x) = t^\rho$, in which case we write $g \in \mathrm{RV}_\rho$. If $\rho = 0$, g is called slowly varying. Exponential-tailed and regularly varying functions are intimately related to each other, since for $\beta \geq 0$, $h \in \mathscr{L}_\beta$ if and only if $h \circ \log \in \mathrm{RV}_{-\beta}$. Using this relationship, we obtain a similar representation of exponential-tailed distribution functions to Karamata's representation of regularly varying functions (see Corollary 2.1 in Resnick (2007)), namely for $\bar{F} \in \mathscr{L}_\beta$, we have

$$\bar{F}(x) = a(x) \exp\Big\{ -\int_0^x \beta(v)\mathrm{d}v \Big\}, \quad x \in \mathbb{R}, \tag{5}$$

where $a(x) \to a \in (0,\infty)$ and $\beta(x) \to \beta$ as $x \to \infty$. This will be repeatedly used in our proofs.

Two prominent examples of exponential-tailed distributions are the generalized inverse Gaussian (GIG) distribution and the generalized hyperbolic (GH) distribution (Barndorff-Nielsen 1977; Prause 1999; McNeil et al. 2005). A random variable R is said to have a GIG distribution, denoted as $R \sim \mathrm{GIG}(\lambda, \tau, \psi)$, if its Lebesgue density is

$$f_{\mathrm{GIG}}(x) = \Big(\frac{\psi}{\tau}\Big)^{\lambda/2} \frac{x^{\lambda-1}}{2K_\lambda(\sqrt{\tau\psi})} \exp\Big\{ -\frac{1}{2}\Big(\frac{\tau}{x} + \psi x\Big)\Big\}, \quad x > 0,$$

where K_λ is the modified Bessel function of the second kind with index λ. If a random variable Z has the stochastic representation as a normal mean-variance mixture, i.e.,

$$Z = \mu + \gamma R + \sqrt{R}W, \quad \mu, \gamma \in \mathbb{R},$$

where W is a standard normal random variable and $R \sim \mathrm{GIG}(\lambda, \tau, \psi)$, then Z is said to have the GH distribution, denoted as $Z \sim \mathrm{GH}(\lambda, \tau, \psi, \mu, \gamma)$, and its Lebesgue density is given in the Supplementary Material (Section 5). The admissible parameter values in both the GIG and GH distributions are $\lambda < 0, \tau > 0, \psi \geq 0$, or $\lambda = 0, \tau > 0, \psi > 0$, or $\lambda > 0, \tau \geq 0, \psi > 0$. The special cases $\psi = 0$ and $\tau = 0$ should be understood as limiting cases. Another limiting case, when R is degenerate and thus Z follows a Gaussian distribution, is not considered in this paper as Z does not have an exponential tail in this case.

It is easy to see that the GIG density function (thus, also its distribution function) has exponential tails. The GH density and distribution functions also have exponential tails and the details are given in the Supplementary Material (Section 4). Furthermore, the GIG and GH distributions are convolution tail equivalent if and only if $\lambda < 0$; see Embrechts and Goldie (1982) for more details. The GIG and GH distributions are infinitely divisible, and their associated Lévy processes are widely used in finance; see Eberlein (2001) and McNeil et al. (2005) for an overview. The main examples in Barndorff-Nielsen and Shephard (2001) are the GIG Lévy processes, while the non-Gaussian noise considered in Bolin (2014), Wallin and Bolin (2015), and Bolin and Wallin (2020) is based on certain subclasses of the GH distribution.

3 Linear transformations

3.1 General framework and outline

In this section, we focus on the extremal dependence structure of the random vector $\boldsymbol{X} = (X_1, X_2)$ defined in (4) with $a_{ji} \geq 0$, $\max a_{1i} > 0$, $\max a_{2i} > 0$, and $a_{1i} + a_{2i} > 0$, $j = 1, 2, i = 1, \ldots, n$. We work under the following assumptions throughout the paper unless otherwise stated:

A.1 $Y_i, i = 1, \ldots, n$, are mutually independent with identical distribution function F_Y. If the support of Y_i is $\mathbb{R}$, then F_Y has exponential left and right tails with the same index, i.e., $h(y) = F_Y(-y) \in \mathscr{L}_\beta$ and $\bar{F}_Y \in \mathscr{L}_\beta$ for some $\beta > 0$. If the support of Y_i is $\mathbb{R}_+$, then F_Y has an exponential right tail, i.e., $\bar{F}_Y \in \mathscr{L}_\beta$ for some $\beta > 0$;

A.2 F_Y is absolutely continuous with density f_Y.

The assumption of identical distribution functions in Assumption A.1 can be relaxed to different exponential-tailed distribution functions with the same index β, but for the sake of simplicity, we assume here that they have a common distribution F_Y. The existence of a density in Assumption A.2 is common in practical modeling of multivariate extremes and it allows convenient inference of the gauge function in the geometric approach; see Proposition 2.2 and its discussions in Nolde and Wadsworth (2022).

Although the dependence structure in model (4) is defined via simple linear transformations, surprisingly, nontrivial extremal dependence structures can be obtained. Specifically, it turns out that the extremal dependence structure of

$\boldsymbol{X} = (X_1, X_2)$ mainly depends on the largest coefficients among $a_{1i}, i = 1, \ldots, n$, and $a_{2i}, i = 1, \ldots, n$. We will show that the components of the vector $\boldsymbol{X}$ are asymptotically dependent if there is equality of the maximizing sets

$$\arg \max_{i \in \{1,\ldots,n\}} a_{1i} = \arg \max_{i \in \{1,\ldots,n\}} a_{2i}. \tag{6}$$

We discuss this case in Section 3.2. On the other hand, there is asymptotic independence if these sets have an empty intersection

$$\arg \max_{i \in \{1,\ldots,n\}} a_{1i} \cap \arg \max_{i \in \{1,\ldots,n\}} a_{2i} = \emptyset, \tag{7}$$

which is discussed in Section 3.3. The delicate boundary case where the two maximizing sets $\arg\max_{i \in \{1,\ldots,n\}} a_{1i}$ and $\arg\max_{i \in \{1,\ldots,n\}} a_{2i}$ are not equal but have a non-empty intersection, is discussed in Section 3.4.

It is interesting to note that this behaviour is in sharp contrast to the case of heavy-tailed linear (e.g., Gnecco et al. 2021) and max-linear (e.g. Wang and Stoev 2011) models, where the random variables Y_i have common survival function $\bar{F}_Y \in \mathrm{RV}_{-\beta}$. In this more classical case, only asymptotic dependence (or complete independence) may arise and all coefficients a_{ji} contribute to the corresponding tail dependence coefficient χ.

3.2 Asymptotic dependence

Here we show that $\boldsymbol{X} = (X_1, X_2)$ defined via (4) is asymptotically dependent when (6) holds, and we give the explicit expression of its tail dependence coefficient. All proofs are deferred to Section 6 and Section 1 in the Supplementary Material.

Proposition 3.1 *Let $Y_1, \ldots, Y_n$ be independent and identically distribution with a common distribution function F_Y such that $\bar{F}_Y \in \mathscr{L}_\beta$ for some $\beta > 0$, and let $\boldsymbol{X} = (X_1, X_2)$ be constructed as in (4). If the set equality in (6) holds, then X_1 and X_2 are asymptotically dependent and the tail dependence coefficient of $\boldsymbol{X}$ can be expressed as*

$$\chi = \mathrm{E}\left[\min\left\{\frac{\exp(\beta Z_1)}{M_{Z_1}(\beta)}, \frac{\exp(\beta Z_2)}{M_{Z_2}(\beta)}\right\}\right],$$

where $Z_1 = \sum_{i \notin I_{max}} a_{1i} Y_i / a_{1max}$, $Z_2 = \sum_{i \notin I_{max}} a_{2i} Y_i / a_{2max}$, and M_{Z_j} is the moment generating function of $Z_j, j = 1, 2$. Here the set $I_{\max}$ is the common set of maximizers in (6) and $a_{jmax} = max_{i=1,\ldots,n} a_{ji}$, $j = 1, 2$.

In this asymptotically dependent case where (6) holds, there is a link between our discrete model (4) and the random scale constructions considered in Huser and Wadsworth (2019) and Engelke et al. (2019). As shown in the proof of Proposition 3.1, we can assume without loss of generality that $a_{1\max} = a_{2\max} = 1$. We can then rewrite model (4) as $X_1 = Z_c + Z_1$ and $X_2 = Z_c + Z_2$, where $Z_c = \sum_{i \in I_{\max}} Y_i$, and Z_1 and Z_2 are the same as in Proposition 3.1. Further notice that Z_c has an exponential tail if and only if e^{Z_c} has a regularly varying tail. Hence, if one is interested in the extremal dependence structure of the random vector (e^{X_1}, e^{X_2}), which is the same as that of (X_1, X_2) since extremal dependence is invariant to monotonically increasing marginal transformations, then this problem falls into the setting of random scale constructions. More precisely, we have $(e^{X_1}, e^{X_2}) = e^{Z_c}(e^{Z_1}, e^{Z_2})$ with e^{Z_c} being the shared random component. An application of Proposition 1 in Engelke et al. (2019) yields the asymptotic dependence of (e^{X_1}, e^{X_2}) and thus also of (X_1, X_2), and gives the same tail dependence coefficient as in Proposition 3.1. This alternative proof transforms the sum of exponential-tailed random variables into a product of regularly varying random variables and exploits the properties of regularly varying functions to derive the extremal dependence structure. In comparison, our proof in Proposition 3.1 treats linear transformations of a vector of independent exponential-tailed random variables in a direct manner and reveals many asymptotic properties of exponential-tailed distribution functions.

To further investigate the expression of the tail dependence coefficient χ in Proposition 3.1, we now consider the specific example where the Y_i are GH distributed. Suppose $Y_1, Y_2 \sim \mathrm{GH}(\lambda, \tau, \psi, \mu, \gamma)$ with $\psi > 0, \gamma = 0$, then Supplementary Material (Section 4) shows that their densities and survival functions have exponential tails with the same index, i.e., $f_{Y_i}, \bar{F}_{Y_i} \in \mathscr{L}_{\sqrt{\psi}}$. Proposition 3.1 yields the following result.

Example 1 Let Y_1, Y_2 be independent and identically distributed with a common distribution $\mathrm{GH}(\lambda, \tau, \psi, \mu, \gamma)$, $\psi > 0, \gamma = 0$. Let $X_1 = Y_1 + a_{12}Y_2$, $X_2 = Y_1 + a_{22}Y_2$ with $0 \le a_{12}, a_{22} < 1$. Then (X_1, X_2) is asymptotically dependent with tail dependence coefficient

$$\chi = \frac{\int_{(a_{22}-a_{12})y \le c} \exp(a_{22}\sqrt{\psi}y) F_{Y_1}(\mathrm{d}y)}{M_{Y_1}(a_{22}\sqrt{\psi})} + \frac{\int_{(a_{22}-a_{12})y > c} \exp(a_{12}\sqrt{\psi}y) F_{Y_1}(\mathrm{d}y)}{M_{Y_1}(a_{12}\sqrt{\psi})},$$

where $c = \{\log M_{Y_1}(a_{22}\sqrt{\psi}) - \log M_{Y_1}(a_{12}\sqrt{\psi})\}/\sqrt{\psi}$, and M_{Y_1} is the moment generating function of Y_1.

Clearly, the tail dependence coefficient χ in Example 1 is larger than zero. Also note that when $a_{12} = a_{22}$, $X_1 = X_2$ and thus $\chi = 1$. In fact, one can show that when $a_{12} \ne a_{22}$, χ is always strictly less than 1; see Proposition 4.1 in the Supplementary Material (Section 4). To further investigate the properties of the tail dependence coefficient χ expressed in Example 1, we now assume $0 \le a_{12} < a_{22} < 1$ and examine its limit as $a_{22} \uparrow 1$. This is interesting because in the following subsection we see that when $a_{12} < 1 < a_{22}$, which implies $\arg\max_{i \in \{1,2\}} a_{1i} = 1$ and $\arg\max_{i \in \{1,2\}} a_{2i} = 2$, then X_1 and X_2 are asymptotically independent and necessarily $\chi = 0$. Hence, the investigation of $\lim_{a_{22} \uparrow 1} \chi$ answers the question of whether χ smoothly transitions from asymptotic dependence to independence. It turns out that this is true only in some cases.

Proposition 3.2 *Let χ be as in Example 1 with $0 \leq a_{12} < a_{22} < 1$. Then*

$$\lim_{a_{22}\uparrow 1} \chi = \begin{cases} \frac{\int_{-\infty}^{c^*} \exp(\sqrt{\psi} y) F_{Y_1}(\mathrm{d}y)}{M_{Y_1}(\sqrt{\psi})} + \frac{\int_{c^*}^{\infty} \exp(a_{12}\sqrt{\psi} y) F_{Y_1}(\mathrm{d}y)}{M_{Y_1}(a_{12}\sqrt{\psi})}, & if \lambda < 0, \\ 0, & if \lambda \geq 0, \end{cases}$$

where $c^* = \lim_{a_{22}\uparrow 1} c/(a_{22} - a_{12}) = \frac{\log\{M_{Y_1}(\sqrt{\psi})\} - \log\{M_{Y_1}(a_{12}\sqrt{\psi}\}}{(1-a_{12})\sqrt{\psi}} < \infty.$

The result in Proposition 3.2 indicates that when $\lambda < 0$, which implies that the distributions of Y_1, Y_2 are convolution tail equivalent (Pakes 2004), there is a discontinuity in χ when a_{22} tends to 1 from below. To illustrate how fast the χ coefficient in Example 1 tends to its limit, we set $a_{12} = 0.3$, and plot χ against a_{22} for various values of λ, τ, ψ in Fig. 2. The results show that λ is the most important parameter that regulates the decay rate of χ as a_{22} increases, and the larger λ the faster χ decays. This indicates that the normal inverse Gaussian distribution, which is a subclass of the GH distribution when λ is fixed to -0.5, might not be a good option to use for modeling extreme events in the presence of both asymptotic dependence and independence, since there is no smooth transition between the two. On the other hand, the variance gamma distribution, another subclass of the GH distribution with $\lambda > 0$, might be a better candidate to consider as there is a smooth transition.

3.3 Asymptotic independence

In this section we show that $\boldsymbol{X}$ defined in (4) is asymptotically independent when (7) holds. We further give the explicit expression of the residual tail dependence coefficient η in certain cases; recall its definition (3) in the introduction.

The idea of this proof is to first expand the bivariate random vector $\boldsymbol{X}$ to n dimensions, then use the geometric approach (Nolde 2014; Nolde and Wadsworth 2022) to cast the computation of the residual tail dependence coefficient η as a convex optimization problem, and finally solve this optimization problem.

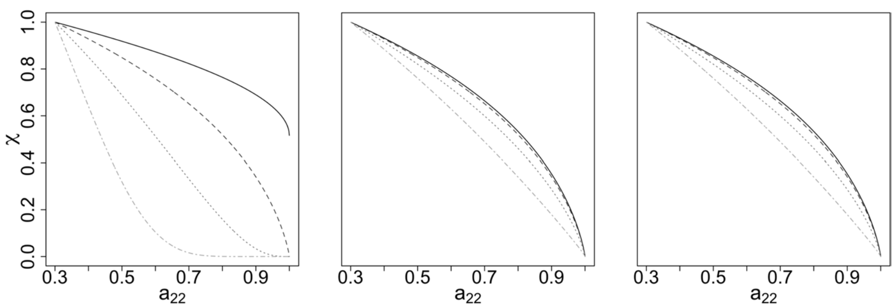


Fig. 2 Values of χ as a function of a_{22} for fixed $a_{12} = 0.3$ and parameters $\tau = 1, \psi = 1, \lambda = -0.5, 1, 5, 30$ (left), $\lambda = 1, \psi = 1, \tau = 0.5, 1, 5, 30$ (center), and $\lambda = 1, \tau = 1, \psi = 0.5, 1, 5, 30$ (right) in Example 1, where lines in each plot are ordered from highest to lowest for increasing values of λ, τ and ψ, respectively

Let $Y_1, \ldots, Y_n$ be a sequence of random variables satisfying Assumptions A.1 and A.2. Let $\boldsymbol{X} = (X_1, X_2)$ be constructed as in (4). Denote

$$c = \min_{\substack{i,j=1,\ldots,n \\ i \neq j}} \left\{ \frac{|\tilde{a}_{2i} - \tilde{a}_{1i}| + |\tilde{a}_{2j} - \tilde{a}_{1j}|}{|\tilde{a}_{2i}\tilde{a}_{1j} - \tilde{a}_{1i}\tilde{a}_{2j}|}, 1/\min(\tilde{a}_{1i}, \tilde{a}_{2i}), 1/\min(\tilde{a}_{1j}, \tilde{a}_{2j}) \right\},$$

where $\tilde{a}_{ji} = a_{ji}/(\max_{r\in\{1,\ldots,n\}} a_{jr}), j = 1, 2, i = 1, \ldots, n$, and the first term in the minimum is set to $+\infty$ whenever $\tilde{a}_{2i}\tilde{a}_{1j} - \tilde{a}_{1i}\tilde{a}_{2j} = 0$. If (6) does not hold, we can permute Y_i such that $\tilde{a}_{11} = \tilde{a}_{22} = 1$, $\tilde{a}_{12}\tilde{a}_{21} < 1$. Then denote

$$\tilde{A} = \begin{pmatrix} 1 & \tilde{a}_{12} & \tilde{a}_{13} & \cdots & \tilde{a}_{1n} \\ \tilde{a}_{21} & 1 & \tilde{a}_{23} & \cdots & \tilde{a}_{2n} \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \in \mathbb{R}^{n\times n},$$

where the off-diagonal entries in $\tilde{A}$ are zero if left unspecified, and denote the row vectors of $\tilde{A}^{-1}$ by $\tilde{\alpha}_i, i = 1, \ldots, n$. We have the following result.

Proposition 3.3 *Let η be the residual tail dependence coefficient of $\boldsymbol{X}$. If the support of Y_i is $\mathbb{R}$, then*

$$\eta = \left(\min_{\boldsymbol{x}\in\mathbb{R}^n: x_1, x_2 \geq 1} \sum_{i=1}^{n} |\tilde{\alpha}_i^\top \boldsymbol{x}| \right)^{-1} = c^{-1}.$$

If the support of Y_i is $\mathbb{R}_+$, then

$$\eta = \left(\min_{\boldsymbol{x}\in\mathbb{R}^n: x_1, x_2 \geq 1, \tilde{A}^{-1}\boldsymbol{x} \geq \mathbf{0}} \sum_{i=1}^{n} \tilde{\alpha}_i^\top \boldsymbol{x} \right)^{-1} \leq c^{-1}.$$

Moreover, if (7) holds then $\eta \leq c^{-1} < 1$, and necessarily X_1 and X_2 are asymptotically independent.

We remark that when (7) does not hold, more assumptions are needed to determine the extremal dependence regime of $\boldsymbol{X}$, i.e., whether the tail dependence coefficient $\chi > 0$ or $\chi = 0$; we refer to Section 3.2 for the asymptotic dependence case and Section 3.4 for the boundary case.

We now consider the specific case that is similar to Example 1 and where a direct application of Proposition 3.3 yields a simple form of η.

Example 2 Let Y_1, Y_2 be independent and identically distributed with a common distribution $\mathrm{GH}(\lambda, \tau, \psi, \mu, \gamma)$, $\psi > 0$, $\gamma = 0$. Let $X_1 = Y_1 + a_{12}Y_2$, $X_2 = a_{21}Y_1 + Y_2$ with $0 \leq a_{12}, a_{21} \leq 1$. Then, the residual tail dependence coefficient of $\boldsymbol{X} = (X_1, X_2)$ is

$$\eta = \frac{1 - a_{12}a_{21}}{2 - a_{12} - a_{21}}.$$

Furthermore, if a_{12}, a_{21} are both strictly less than 1, then $\boldsymbol{X}$ is asymptotically independent.

An interesting observation in Proposition 3.3 and Example 2 is that the expression of η only depends on the coefficients a_{ji}. In other words, this means that as long as we do not change the coefficients a_{ji} in the model (4) and $Y_1, \ldots, Y_n$ satisfy conditions A.1 and A.2, then we always obtain the same residual tail dependence coefficient η, regardless of index β or the exact distribution of Y_i. Another observation of Example 2 is that if we fix a_{12} (or a_{21}), then η is monotonically increasing with respect to a_{21} (or a_{12}). Furthermore, the range of η is [1/2, 1], where 1/2 is achieved when $a_{12} = a_{21} = 0$, and $\eta = 1$ when $a_{12} = 1$ or $a_{21} = 1$.

We now consider another example that clearly distinguishes the extremal dependence measures from dependence measures for the bulk of the distribution.

Example 3 Let $Y_1, \ldots, Y_n \in \mathbb{R}_+$ be independent and identically distributed with a common distribution function F_Y with $\bar{F}_Y \in \mathscr{L}_\beta, \beta > 0$. Let $X_1 = Y_1$, $X_2 = a_{21}Y_1 + Y_2 + a_{23}Y_3 + \cdots + a_{2n}Y_n$ with $0 \leq a_{21}, a_{23}, \ldots, a_{2n} < 1$. Then by Proposition 3.3, the residual tail dependence coefficient of $\boldsymbol{X} = (X_1, X_2)$ is

$$\begin{aligned}
\eta &= \left(\min_{\boldsymbol{x} \in \mathbb{R}^n : x_1, x_2 \geq 1, \tilde{A}^{-1}\boldsymbol{x} \geq \boldsymbol{0}} \sum_{i=1}^{n} \tilde{\alpha}_i^\top \boldsymbol{x} \right)^{-1} \\
&= \left\{ \min_{\boldsymbol{x} \in \mathbb{R}^n : x_1, x_2 \geq 1, x_2 - \sum_{i=1,3,\ldots,n} a_{2i}x_i \geq 0, x_3, \ldots, x_n \geq 0} x_2 + \sum_{i=1,3,\ldots,n} (1 - a_{2i})x_i \right\}^{-1} \\
&= 1/(2 - a_{21}).
\end{aligned}$$

We observe that the expression of η does not depend on *n*, i.e., the number of independent terms in the construction of X_2. Loosely speaking, when more independent terms $a_{2i}Y_i$ with $0 < a_{2i} < 1$ are added to X_2, these added terms do not contribute much to the extremal dependence between X_1 and X_2 and the residual tail dependence coefficient remains the same. This is in clear contrast with the classical Pearson's correlation dependence measure

$$\mathrm{Corr}(X_1, X_2) = a_{21} \left(a_{21}^2 + 1 + \sum_{i=3}^{n} a_{2i}^2 \right)^{-1/2},$$

which decreases monotonically as *n* increases.

3.4 Boundary case and further remarks

In Sections 3.2 and 3.3 we have shown that the discrete model $\boldsymbol{X}$ in (4) is asymptotically dependent or independent when (6) or (7) hold, respectively. The intuition for this phenomenon is that the extremal dependence of (X_1, X_2) is determined by whether the main contribution for the tails of X_1 and X_2 comes from the same components Y_i. If this is the case, then the stronger form of extremal dependence, namely asymptotic dependence, is achieved; otherwise, asymptotic independence is obtained. In the sequel, let $I_j = \arg\max_{i\in\{1,\dots,n\}} a_{ji}$, $j = 1, 2$.

Clearly, there is a boundary case where neither (6) nor (7) hold, that is, we have $I_1 \neq I_2$ and

$$I_{\max} := I_1 \cap I_2 \neq \emptyset.$$

To avoid complications, we here only consider $I_1 \not\subset I_2$ and $I_2 \not\subset I_1$. With the notation of Proposition 3.1 and the paragraph thereafter, without losing generality, we assume that $a_{1\max} = a_{2\max} = 1$. Recall that we can write $(X_1, X_2) = Z_c + (Z_1, Z_2)$, or $(e^{X_1}, e^{X_2}) = e^{Z_c}(e^{Z_1}, e^{Z_2})$. Theorem 3 in Embrechts and Goldie (1980) implies that the survival functions of e^{Z_c}, e^{Z_1}, and e^{Z_2} are all regularly varying with the same index $-\beta$. Hence, Proposition 6 in Engelke et al. (2019) can be applied to determine the extremal dependence structure of (X_1, X_2). More precisely, in this case, the residual tail dependence coefficient of (X_1, X_2) is $\eta = 1$, but more assumptions are needed to determine whether $\chi > 0$ or $\chi = 0$. We leave the other boundary case $I_{\max} \neq \emptyset$, $I_1 \subset I_2$, or *vice versa*, for future research.

Here we have studied the extremal dependence structure of linear transformations, or sums of exponential-tailed random vectors. The results are directly applicable to products of regularly varying random vectors. Indeed, assume that $\bar{Y}_1, \dots, \bar{Y}_n$ are independent copies of a positive random variable Y with absolutely continuous distribution function $F_{\bar{Y}}$, and $\bar{F}_{\bar{Y}} \in \mathrm{RV}_{-\beta}$ with $\beta > 0$. Let $a_{ji} \geq 0$ for $j = 1, 2, i = 1, \dots, n$. Then the product model

$$\begin{cases} \bar{X}_1 = \bar{Y}_1^{a_{11}} \cdots \bar{Y}_n^{a_{1n}}, \\ \bar{X}_2 = \bar{Y}_1^{a_{21}} \cdots \bar{Y}_n^{a_{2n}}, \end{cases}$$

has the same extremal dependence structure as the sum model (4) with $Y_i = \log(\bar{Y}_i)$, and Propositions 3.1 and 3.3 give the extremal dependence coefficients of $(\bar{X}_1, \bar{X}_2)$.

We here restrict our attention in (4) to the case with non-negative coefficients a_{ji}, because both the integral approximation of moving average processes and the finite element approximation of the type G SPDE Matérn fields have non-negative coefficients; see Section 4 for more details. Consequently, the constant c^{-1} in Proposition 3.3 can be shown to have the lower bound 1/2, which implies that if Y_i has support $\mathbb{R}$, then only positive extremal association can be achieved. However, if the interest lies in capturing negative extremal association, i.e., $\eta < 1/2$, then one can achieve this by considering negative coefficients a_{ji} and our proof of Proposition 3.3 can be modified accordingly.

4 Moving average processes

4.1 Main results

We now focus on the extremal dependence structure of moving average processes (1). Note that here $u(\boldsymbol{s})$ refers to the usual definition of stochastic integrals, where in the first step one defines integrals of simple functions, and then a non-random integrand function G is called $\mathcal{M}$-integrable if there exists a sequence of simple functions that converges pointwise to G, such that the limit in probability of the resulting integrals of simple functions exists, and this limit is defined to be the stochastic integral with respect to $\mathcal{M}$ (Rajput and Rosinski 1989). A useful characterization of $\mathcal{M}$-integrable functions is given in (Rajput and Rosinski 1989, Theorem 2.7). Throughout the paper we assume that $u(\boldsymbol{s})$ is well defined, i.e., G is $\mathcal{M}$-integrable.

We first consider the more commonly used case where the domain of integration $\mathcal{D}$ is fixed and does not depend on $\boldsymbol{s}$. To consider a framework that is applicable to general moving average processes which are not necessarily SPDEs, we assume that:

B.1 the function $G(h)$ is non-negative, continuous and strictly decreasing as $h \to \infty$;
B.2 for any bounded Borel set $B \subset \mathcal{D}$, the random variable $\mathcal{M}(B)$ has an absolutely continuous distribution function $F_{\mathcal{M}(B)}$. Furthermore, if the support of $\mathcal{M}(B)$ is $\mathbb{R}_+$, then $F_{\mathcal{M}(B)}$ has an exponential tail, i.e., $\bar{F}_{\mathcal{M}(B)} \in \mathscr{L}_\beta$, $\beta > 0$; if its support is $\mathbb{R}$, then $F_{\mathcal{M}(B)}$ has exponential right and left tails with the same index $\beta > 0$.

We note that for the important class of type G Matérn SPDE random fields on $\mathbb{R}^d$, the function G is non-negative, absolutely continuous, and monotonically decreasing (see Proposition S2 in Section 4 of the Supplementary Material). Assumption B.2 implies that $\mathcal{M}(B)$ has an exponential-tailed distribution with the same index for every bounded Borel set B. This seemingly restrictive assumption is, in fact, a natural result of the convolution-closure property of exponential tails (Embrechts and Goldie 1980, Theorem 3), and it is satisfied for instance for the symmetric NIG Matérn SPDE fields and the symmetric variance Gamma Matérn SPDE fields considered in Bolin (2014), Wallin and Bolin (2015) and Bolin and Wallin (2020).

We now assume that $d = 2$, which is the most important case for spatial applications; the cases $d = 1$ and $d \geq 3$ can be treated analogously. Let $\{D_n : D_n \subset \mathcal{D}\}$ be an increasing sequence of bounded Borel sets in $\mathbb{R}^2$, and $D_n^1, D_n^2, \dots, D_n^J \subset D_n$ be a partition of D_n obtained by triangulating D_n with m_n mesh nodes defined as $M_n = \{\boldsymbol{a}_n^i, i = 1, \dots, m_n\}$. Then we obtain an approximation of $u(\boldsymbol{s})$ as

$$u_n(\boldsymbol{s}) = \int_{\mathcal{D}} \sum_{i=1}^{J} G(\|\boldsymbol{s} - \boldsymbol{d}_i\|) 1_{D_n^i}(\boldsymbol{t}) \mathcal{M}(\mathrm{d}\boldsymbol{t}) = \sum_{i=1}^{J} G(\|\boldsymbol{s} - \boldsymbol{d}_i\|) \mathcal{M}(D_n^i), \quad (8)$$

where $\boldsymbol{d}_i \in D_n^i$, and $1_{D_n^i}$ is the indicator function. We say that the sequence of points M_n is dense in $\mathcal{D}$ if for any point $\boldsymbol{s} \in \mathcal{D}$, there is a sequence $\{\bar{\boldsymbol{a}}_n\}, \bar{\boldsymbol{a}}_n \in M_n$ such that $\lim_{n\to\infty} \|\bar{\boldsymbol{a}}_n - \boldsymbol{s}\| = 0$. Assume that M_n is dense. Then clearly, for any fixed $\boldsymbol{s} \in \mathcal{D}$, $G_n^J(\|\boldsymbol{s} - \boldsymbol{t}\|) = \sum_{i=1}^{J} G(\|\boldsymbol{s} - \boldsymbol{d}_i\|) 1_{D_n^i}(\boldsymbol{t})$ converges to the function

$G(\|\boldsymbol{s}-\boldsymbol{t}\|)$ pointwise as $m_n \to \infty$. Hence, by the definition of stochastic integrals, we know that $u_n(\boldsymbol{s})$ converges to $u(\boldsymbol{s})$ in probability. It follows that $(u_n(\boldsymbol{s}_1), u_n(\boldsymbol{s}_2))$, $\boldsymbol{s}_1 \neq \boldsymbol{s}_2$, converges in probability to $(u(\boldsymbol{s}_1), u(\boldsymbol{s}_2))$.

Furthermore, by Assumption B.2, $\bar{F}_{\mathcal{M}(D_n^i)} \in \mathscr{L}_\beta$. Note also that $\mathcal{M}(D_n^i)$, $i = 1, \ldots, J$, are independent since $\{D_n^i, i = 1, \ldots, J\}$ forms a partition of D_n. Hence, to understand the extremal dependence structure of $(u_n(\boldsymbol{s}_1), u_n(\boldsymbol{s}_2))$, it is sufficient to focus on the coefficients $G(\|\boldsymbol{s}_1 - \boldsymbol{d}_i\|)$ and $G(\|\boldsymbol{s}_2 - \boldsymbol{d}_i\|)$. Since $\boldsymbol{s}_1 \neq \boldsymbol{s}_2$, we know that if the mesh is fine enough, $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ will fall into different triangles $D_n^{i_1}$ and $D_n^{i_2}$. Consequently, we have

$$\arg\max_i G(\|\boldsymbol{s}_1 - \boldsymbol{d}_i\|) = i_1 \neq \arg\max_i G(\|\boldsymbol{s}_2 - \boldsymbol{d}_i\|) = i_2.$$

Proposition 3.3 then yields the asymptotic independence of $(u_n(\boldsymbol{s}_1), u_n(\boldsymbol{s}_2))$ and allows the computation of its residual tail dependence coefficient. We now focus on the limit of the residual tail dependence coefficient of the approximation model $(u_n(\boldsymbol{s}_1), u_n(\boldsymbol{s}_2))$ as $m_n \to \infty$.

Theorem 4.1 *Assume that for the moving average process (1), the support of $\mathcal{M}(B)$ is $\mathbb{R}$ for any bounded Borel set B. Let $u_n(\boldsymbol{s})$ be its integral approximation and let $\eta_n(h)$ be the residual tail dependence coefficient of $(u_n(\boldsymbol{s}_1), u_n(\boldsymbol{s}_2))$ for $h = \|\boldsymbol{s}_1 - \boldsymbol{s}_2\|$. If Assumptions B.1 and B.2 are satisfied and the function G is convex, then if $m_n \to \infty$ and the sequence of mesh nodes M_n is dense in $\mathcal{D}$, the limit of η_n as $n \to \infty$ is*

$$\eta(h) = \frac{1}{2} + \frac{G(h)}{2G(0)}.$$

The idea of our proof is to use Proposition 3.3 to obtain an explicit formula for $\eta_n(h)$ as a minimum (or maximum) over a finite number of terms, and with this, to show that for convex G, the optimum can be achieved as the mesh size tends to zero. We remark that when $G(0) = \infty$, $\eta(h)$ reduces to a constant 1/2. When the function G is non-convex, we conjecture that the limiting residual tail dependence function is

$$\eta(h) = \max\left\{\frac{1}{2} + \frac{G(h)}{2G(0)}, \frac{G(h/2)}{G(0)}\right\}.$$

Our numerical experiments in Section 4.3 seem to support our conjecture, but a rigorous proof would have to use a different technique than the proof of Theorem 4.1, which relies heavily on the convexity of G.

We now consider another important case where the integration domain $\mathcal{D}$ depends on s, namely when $d = 1$ and

$$u(s) = \int_{-\infty}^{s} G(s-t)\mathcal{M}(\mathrm{d}t), \quad s \in (-\infty, T]. \tag{9}$$

This is an interesting case because it is used in practice (Barndorff-Nielsen and Shephard 2001; Ver Hoef and Peterson 2010), and this one-sided integral turns out to yield different residual tail dependence functions.

Let $-n = t_0 < \cdots < t_{n_1} \leq s_1 < \cdots \leq t_{n_2} \leq s_2 < \cdots < t_{m_n} = T$ be an arbitrary partition of $[-n, T]$. Then the approximation (8) becomes

$$u_n(s_1) = \sum_{i=0}^{n_1-1} G(s_1 - t_i)\mathcal{M}([t_i, t_{i+1})), \quad u_n(s_2) = \sum_{i=0}^{n_2-1} G(s_2 - t_i)\mathcal{M}([t_i, t_{i+1})). \quad (10)$$

Clearly when $m_n \to \infty$ such that the sequence of partition points $M_n = \{t_i, i = 0, \ldots, m_n\}$ is dense in $(-\infty, T]$, $(u_n(s_1), u_n(s_2))$ converges in probability to $(u(s_1), u(s_2))$. Note that when the mesh is coarse and there is no partition point in (s_1, s_2), i.e., $t_{n_1} = t_{n_2}$, then the largest coefficients in the expressions of $u_n(s_1)$ and $u_n(s_2)$ are $G(s_1 - t_{n_1-1})$ and $G(s_2 - t_{n_1-1})$ respectively, which correspond to the same variable $\mathcal{M}([t_{n_1-1}, t_{n_1}))$. Hence, Proposition 3.1 gives the asymptotic dependence of $(u_n(t_1), u_n(t_2))$. Otherwise, when there is at least one partition point in (s_1, s_2), $(u_n(t_1), u_n(t_2))$ is asymptotically independent, and in the following, we derive its limiting residual tail dependence function as the mesh size tends to zero.

Theorem 4.2 *Let $u_n(s)$ be the integral approximation (10) of the one-side integral (9), and let $\eta_n(h)$ be the residual tail dependence coefficient of $(u_n(s_1), u_n(s_2))$ for $h = |s_1 - s_2|$. If Assumptions B.1 and B.2 are satisfied, then if $m_n \to \infty$ and the sequence of mesh nodes M_n is dense in $\mathcal{D}$, the limit of η_n as $n \to \infty$ is*

$$\eta(h) = \frac{1}{2 - G(h)/G(0)}.$$

We remark that in Theorem 4.1 and 4.2 we have investigated the limit of the residual tail dependence coefficient of the discrete approximation model as the mesh size tends to zero. Since the models used in practice are the discrete approximation models rather than the limiting continuously indexed process, one might argue that the results for the discretization are more important than the corresponding result for the limiting model. Nevertheless, we tried to establish the link between the extremal dependence structure of the discretization and that of the continuous process, and we found that the extremal dependence structure is not necessarily preserved in the limit due to the unexchangeability of the limits of the approximation and the extreme threshold; see our counterexample Example 4 in Section 4.2. For the non-Gaussian OU process, we are able to link the limit of η_n of the discretization and η of the continuous process thanks to the one-sided integral representation of the OU process, and we now provide more details in the following section.

4.2 Non-Gaussian Ornstein-Uhlenbeck (OU) Process

As the first application of the general results of the previous section, here we study the extremal dependence structure of non-Gaussian OU processes $u(t)$ (Barndorff-Nielsen and Shephard 2001), defined as the stationary solution to the stochastic differential equation (SDE)

$$\mathrm{d}u(t) = -au(t)\mathrm{d}t + \mathrm{d}z(at), \quad a > 0, t \in \mathbb{R}, \tag{11}$$

where $z = \{z(t) : t \in \mathbb{R}\}$ is a Lévy process satisfying $\mathrm{E}[\log\{1 + |z(1)|\}] < \infty$ to guarantee the existence of such a stationary solution. Although the background driving Lévy process z can be chosen arbitrarily, examples considered in Barndorff-Nielsen and Shephard (2001) all have exponential tails for the density of the Lévy measure of $z(1)$.

Non-Gaussian OU processes are generalizations of classical OU processes by replacing the Brownian motion $z(t)$ in the SDE (11) by general Lévy processes. The existence of such processes is established based on the notion of self-decomposability and the stochastic integral representation of self-decomposable random variables (Jurek and Vervaat 1983). More precisely, let V be a self-decomposable random variable (namely for every $\alpha \in (0, 1)$, we have the decomposition $V = \alpha V + \tilde{V}_\alpha$, where $\tilde{V}_\alpha$ is a random variable independent of V), then there exists a Lévy process $z(t)$ and a stationary stochastic process $u(s)$ such that (11) holds for all $a > 0$ and $u(s)$ has the same distribution as V for all $s \geq 0$. Conversely, if z is a Lévy process and $u(t)$ is a stationary stochastic process such that $u(s)$ satisfying (11) for all $a > 0$, then the marginal distribution of $u(s)$ is self-decomposable. We refer to Jurek and Vervaat (1983) and Barndorff-Nielsen and Shephard (2001) for more details.

Importantly, the solution to the SDE (11) can be represented as

$$u(s) = \int_{-\infty}^{s} e^{-a(s-t)}\mathrm{d}z(at) = e^{-as}u(0) + \int_{0}^{s} e^{-a(s-t)}\mathrm{d}z(at), \quad s \geq 0, \tag{12}$$

where $u(0)$ is independent of $\int_0^s e^{-a(s-t)}\mathrm{d}z(at)$. Using this representation, we can show that if the stationary solution $u(s)$ has an absolutely continuous and exponential-tailed distribution function, then the process $u(s)$ is asymptotically independent.

Theorem 4.3 *Let u(s) be a non-Gaussian OU process defined as a stationary solution to (11) and V has the stationary self-decomposable distribution. Suppose F_V, the distribution function of V, is absolutely continuous. If F_V has an exponential tail with index $\beta > 0$ when V has support $\mathbb{R}_+$, or F_V has exponential right and left tails with the same index $\beta > 0$ when V has support $\mathbb{R}$, then $(u(s_1), u(s_2))$ is asymptotically independent for $s_1 \neq s_2$. The corresponding residual tail dependence coefficient is*

$$\eta = \frac{1}{2 - e^{-a|s_2 - s_1|}}.$$

Note that when the Lévy process z in (11) is a Brownian motion, we obtain the classical Gaussian OU process with the correlation between $u(s_1)$ and $u(s_2)$ as $e^{-a|s_2-s_1|}$ and residual tail dependence coefficient as $(1+e^{-a|s_2-s_1|})/2$. Theorem 4.3 shows that when the marginal distribution of the non-Gaussian OU process $u(s)$ has an exponential tail, its induced extremal dependence structure is indeed different from the classical Gaussian OU process, although its correlation function remains the same. More precisely, using a first-order Taylor expansion, one can observe that as $|s_2-s_1|$ tends to zero, the residual tail dependence coefficient η of the Gaussian OU process increases to 1 at a linear rate $a/2$, whilst η of the non-Gaussian OU process increases to 1 at a linear rate a.

Instead of specifying the self-decomposable distribution function of $u(s)$, one can alternatively define the non-Gaussian OU process by specifying the background driving Lévy process z, which reduces to specifying the distribution of $z(1)$. Although the relationship between the tail of $z(1)$ and that of $u(t)$ is unclear, $u(t)$ and $z(1)$ are closely linked by their Lévy measures U_u and $U_{z(1)}$ through the equation $U_u([x,\infty)) = \int_1^\infty s^{-1} U_{z(1)}([sx,\infty))\mathrm{d}s$ (Barndorff-Nielsen 1998, Theorem 2.2). Using this link, we show in the following that under mild assumptions, a convolution equivalent tail of $z(1)$ implies a convolution equivalent tail for $u(t)$.

Proposition 4.4 *Denote the distribution function of z(1) and u(t) by $F_{z(1)}$ and F_u, respectively. Suppose that the function $U_{z(1)}([x,\infty))$ is continuous in x on $(0,\infty)$. If $\bar{F}_{z(1)} \in \mathscr{S}_\beta$, $\beta \geq 0$, then $\bar{F}_u \in \mathscr{S}_\beta$. Moreover, $\bar{F}_{z(t)} \in \mathscr{S}_\beta$, where $F_{z(t)}$ denotes the distribution function of z(t). Hence, all increments of the Lévy process z(t) have exponential-tailed distribution functions with the same index β.*

One example is given by $U_{z(1)}([x,\infty)) = rx^{-\epsilon}\exp(-\beta x)$, where $\epsilon > 1$ and $r, \beta > 0$; see Example 2.4.1 in Barndorff-Nielsen and Shephard (2001). Another example is the so-called CGMY Lévy process (Carr et al. 2002) with $U_{z(1)}([x,\infty)) = \int_x^\infty rw^{-1-b}\exp(-\beta w)\mathrm{d}w$, where $r, \beta > 0$, and $0 < b < 2$. The restriction $\epsilon > 1$ and $b > 0$ implies that the normalized Lévy measure of $z(1)$ is convolution tail equivalent and hence $\bar{F}_{z(1)} \in \mathscr{S}_\beta$ (Shimura and Watanabe 2005, Theorem B).

Now we assume that $z(1)$ has support $\mathbb{R}_+$ and has a convolution tail equivalent distribution $F_{z(1)}$ with $\bar{F}_{z(1)} \in \mathscr{S}_\beta$, $\beta > 0$, the function $U_{z(1)}([x,\infty))$ is continuous in x on $(0,\infty)$, and the distribution function of $z(t)$ is absolutely continuous for all t. That is, Assumption B.2 is satisfied. Then we are ready to link our results in Theorem 4.2 and Theorem 4.3.

Clearly, for the OU processes, the integrand function in (12) satisfies Assumption B.1. Hence, Theorem 4.2 implies that the limiting residual tail dependence function of the approximation model of the form (10) is $\eta(h) = 1/(2-e^{-ah})$. On the other hand, the convolution equivalent tail of $z(1)$ implies a convolution equivalent tail for $u(s)$. If we further assume that the stationary distribution of $u(s)$ is absolutely continuous, then Theorem 4.3 gives its residual tail dependence function as $\eta(h) = 1/(2-e^{-ah})$, which coincides with the limit of its approximation model. In Fig. 3, we illustrate this convergence of the residual tail dependence coefficient

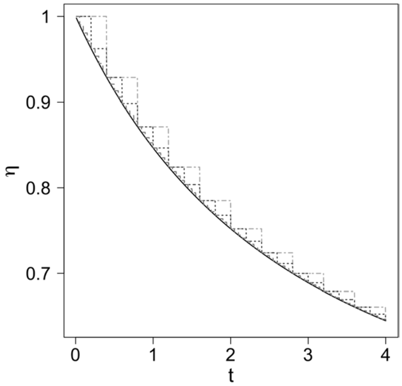


Fig. 3 The residual tail dependence coefficient η of the true non-Gaussian OU process (black line), and its integral approximations for three different, equidistant partitions of [0, *T*] with mesh length $\Delta \in \{0.4, 0.2, 0.05\}$, where the lines are ordered from highest to lowest for decreasing Δ

η of the approximating model to the true non-Gaussian OU process for $a = 0.2$, $s_1 = 0, T = 4$, and three equidistant partitions of the interval [0, *T*] with mesh length $\Delta = 0.4, 0.2, 0.05$.

The preceding analysis seemingly indicates that the extremal dependence structure is preserved in the limit for convergent (in probability) random vectors. It is however important to note that in general, convergence in probability or even almost surely does not (!) necessarily imply convergence of the corresponding tail dependence coefficients, as shown in the following counterexample.

Example 4 Let $X_{1,n} = X/n + \epsilon_1, X_{2,n} = X/n + \epsilon_2$, where $X, \epsilon_1, \epsilon_2$ are independent, X has a regularly varying survival function $\bar{F}_X \in \mathrm{RV}_\rho$, and $\epsilon_i, i = 1, 2$ are standard normal random variables. It is straightforward to see that $(X_{1,n}, X_{2,n})$ converges almost surely to the limiting random vector (ϵ_1, ϵ_2). However, the extremal dependence is clearly not preserved in the limit, since by Proposition 4 in Engelke et al. (2019) we know that for any finite *n*, $(X_{1,n}, X_{2,n})$ is asymptotically dependent with tail dependence coefficient $\chi = 1$, whilst (ϵ_1, ϵ_2) is asymptotically independent with residual tail dependence coefficient $\eta = 1/2$.

4.3 Type G matérn SPDE random fields

The second application of our general results is the popular class of type G Matérn SPDE random fields defined as the stationary solution to the SPDE

$$(\kappa^2 - \Delta)^{\alpha/2} u(\boldsymbol{s}) = \dot{\mathcal{M}}(\boldsymbol{s}), \quad \boldsymbol{s} \in \mathbb{R}^d, \tag{13}$$

where $\alpha = \nu + d/2$, $d = 1, 2, \ldots$ is the dimension, $\nu > 0$ is the smoothness parameter, $\kappa > 0$ is the range parameter, Δ is the Laplacian and $\dot{\mathcal{M}}$ is the so-called type G Lévy noise (Rosinski 1991). The solution $u(\boldsymbol{s})$ can be expressed as process convolutions

$$u(\boldsymbol{s}) = \int_{\mathbb{R}^d} G(\|\boldsymbol{s}-\boldsymbol{t}\|)\mathcal{M}(\mathrm{d}\boldsymbol{t}), \quad \boldsymbol{s} \in \mathbb{R}^d, \tag{14}$$

where $\mathcal{M}$ is the random measure associated with the noise $\dot{\mathcal{M}}$, and G is the Green's function of the differential operator in (13) of the form

$$G(\|\boldsymbol{s}-\boldsymbol{t}\|) = \frac{2^{1-\frac{\alpha-d}{2}}}{(4\pi)^{d/2}\Gamma(\alpha/2)\kappa^{\alpha-d}}(\kappa\|\boldsymbol{s}-\boldsymbol{t}\|)^{\frac{\alpha-d}{2}}K_{\frac{\alpha-d}{2}}(\kappa\|\boldsymbol{s}-\boldsymbol{t}\|), \tag{15}$$

with Γ being the gamma function and K being the modified Bessel function of the second kind.

Notably, for the important class of type G Matérn SPDE random fields, the function G is absolutely continuous and monotonically decreasing (see Proposition S2 in the Supplementary Material). This implies that Assumption B.1 is satisfied. Moreover, Assumption B.2 is satisfied for the symmetric NIG Matérn SPDE fields and variance Gamma Matérn SPDE fields considered in Bolin (2014), Wallin and Bolin (2015), and Bolin and Wallin (2020). Hence, if we consider its integral approximation of the form (8), Theorem 4.1 gives the limiting residual tail dependence function for convex G.

We remark that for $d = 2$, the specific Green's function G in (15) is convex when $\alpha \leq 3$ and $G(0)$ is bounded only when $\alpha > 2$. This implies that when $\alpha = 2$, which is the case commonly used in practice, if the constructed mesh is very fine, then the resulting discrete approximation model would have $\eta \approx 1/2$ (near independence) between all pairs of locations, regardless of the distance between them. From a practical point of view, this indicates that the case $\alpha = 2$ might not be suitable for modeling extremal dependence, whilst $\alpha = 3$ might be more useful. We also remark that the frequently used finite element approximation is in the form of linear transformations as well, and more details are given in the Supplementary Material (Section 3).

We conduct numerical experiments to illustrate our results. We consider the NIG Matérn SPDE model with range parameter $\kappa = 2$, smoothness parameter $\alpha = 2$, and NIG noise location parameter $\mu = 0$, skewness parameter $\gamma = 0$, and shape parameters $\psi = \tau = 1$. We first consider its finite element approximation and examine the convergence of the pre-asymptotic tail dependence coefficient $\chi(q)$ as $q \to 1$. We randomly select 100 sites in the unit square and consider a fine mesh constructed based on a 1600-node lattice with outer extensions; see the left panel of Fig. 4. Then we simulate 10^6 observations at each site, and compute the empirical tail dependence coefficient $\chi(q)$ with respect to different probability levels q. Fig. 4 (right panel) shows $\chi(q)$ for four different pairs of locations with different distances. The plot indicates that the two pairs with longer distances are asymptotically independent and also depicts the decay rate of $\chi(q)$ as $q \to 1$. The other two pairs with shorter distances are also asymptotically independent by Proposition 3.3, but their corresponding $\chi(q)$ decays at a much slower rate and much more simulations are needed to show that its limit is zero.

We now compare the integral approximation with the finite element approximation and examine the effect of the smoothness parameter α on the extremal dependence structure of the approximation models. We choose the same SPDE range parameter,

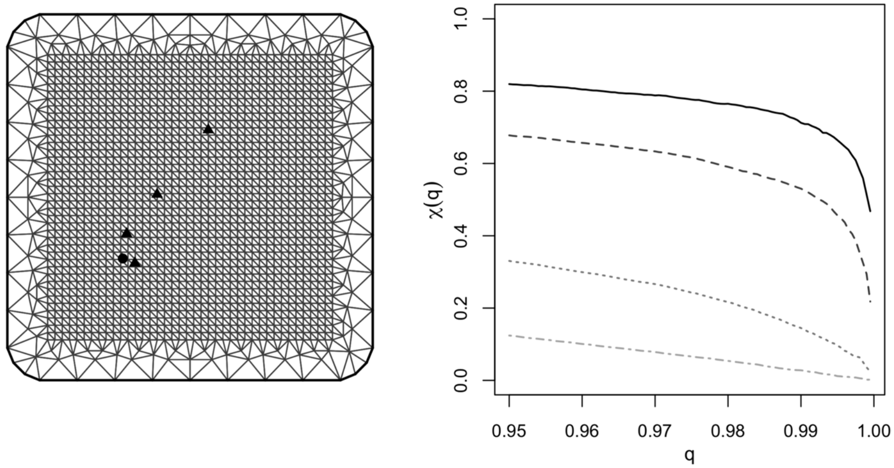


Fig. 4 Left panel: constructed mesh and selected pairs of sites with different distances (triangle points to the round point) in the first numerical experiment in Section 4.3; right panel: empirical tail dependence coefficient $\chi(q)$ for different probability levels q, of the finite element approximation of the NIG Matérn SPDE model. Four pairs of locations with different distances are plotted, with the lines ordered from top to bottom for increasing distances

NIG noise parameters and the fine mesh constructed based on 1600 lattice nodes as in the first numerical experiment, and consider more sites, namely 225 randomly selected sites in the unit square, and smoothness parameter $\alpha = 2, \ldots, 5$. We numerically compute the residual tail dependence coefficient η of all pairs of sites for both approximations using the formula from Proposition 3.3.

Figure 5 depicts η against the distance between all pairs of sites. A first observation is that with such a fine mesh the difference between the integral approximation and the finite element approximation is negligible. When $\alpha = 2$, i.e., the function G in (15) is convex but unbounded at 0, the residual tail dependence function $\eta(h)$ is still fairly far from its limiting function (as the mesh size tends to 0), which is $\eta(h) \equiv 1/2$. On the other hand, when $\alpha = 3$, i.e., G is convex and bounded at 0, then the function $\eta(h)$ of both approximation models has almost converged to its limiting function. While the cases $\alpha = 4, 5$ are not covered by our theory since G is noncon-

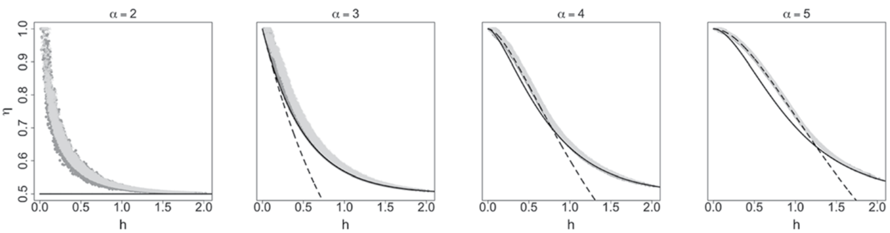


Fig. 5 Coefficients for the finite element approximation (light-grey triangle points) and integral approximation (dark-grey round points) of the type G Matérn SPDE model based on a fine mesh, and the functions (black solid line) and (black dashed line). Panels from left to right correspond to different smoothness parameters

vex, the η values of the approximation models seem to converge to our conjectured limiting function, namely $\eta(h) = \max\{1/2 + G(h)/\{2G(0)\}, G(h/2)/G(0)\}$. This provides numerical evidence for our conjecture.

5 Conclusion

Linear transformations of a random vector with independent components are classical constructions to capture complex correlation dependence structures due to their simplicity and analytical tractability. For instance, the approximation models used in practice in the well-known SPDE approach (Lindgren et al. 2011; Bolin 2014) are of this form. In this paper, we derived the first results on the extremal dependence structure induced by such constructions when the independent components have exponential tails. These general results are leveraged to study the extremal dependence structure of moving average processes driven by exponential-tailed Lévy noise. In particular, the classical exponential-tailed non-Gaussian OU processes are shown to be asymptotically independent, but with a different residual tail dependence function than their Gaussian counterpart. As for the type G Matérn random fields, or more general moving average processes, we have shown that under certain assumptions on the kernel function and the noise process, the integral approximation is asymptotically independent when the mesh is fine enough and the limiting residual tail dependence function is derived.

In terms of statistical modeling, linear transformations of exponential-tailed random vectors have the potential to bridge asymptotic dependence and independence in a tractable way, and models with such features are in pursuit in the extremes community (Nolde and Zhou 2021). In fact, this desirable property distinguishes them from other marginal distributions. For instance, linear combinations of heavy-tailed random variables can only result in asymptotic dependence or complete independence. On the other hand, linear transformations of Gaussian random vectors exclusively yield asymptotic independence or complete dependence. Therefore an interesting and natural question is whether exponential tails are the only ones that can exhibit both extremal dependence classes under linear transformations. Another interesting future research direction is to apply the non-Gaussian SPDE models for extremes, as this has not been investigated in the literature yet.

6 Proofs

Proof of Proposition 3.3 If (6), then by Proposition 3.1 we know $\chi > 0$, which implies that $\eta = 1$. Otherwise, we can permute Y_i such that $1 \in \arg\max_{i\in\{1,\dots,n\}} a_{1i}$, $2 \in \arg\max_{i\in\{1,\dots,n\}} a_{2i}$,but$2 \notin \arg\max_{i\in\{1,\dots,n\}} a_{1i}or1 \notin \arg\max_{i\in\{1,\dots,n\}} a_{2i}$. Since extremal dependence is a copula property and $(X_1/a_{11}, X_2/a_{22})$ has the same extremal dependence structure as $\boldsymbol{X}$, we can assume $a_{11} = a_{22} = 1$ and $0 \leq a_{ji} \leq 1$, $j = 1, 2$, $i = 1, \dots, n$. Hence, we further have $a_{12}a_{21} < 1$.

The strategy for this proof relies on augmenting the model in the following way

$$\begin{cases} X_1 &= Y_1 + a_{12}Y_2 + \cdots + a_{1n}Y_n, \\ X_2 &= a_{21}Y_1 + Y_2 + a_{23}Y_3 + \cdots + a_{2n}Y_n, \\ X_3 &= Y_3, \\ &\vdots \\ X_n &= Y_n. \end{cases}$$

Since $Y_i, i = 1, \ldots, n$ are independent and have a common distribution function F_Y with $\bar{F}_Y \in \mathscr{L}_\beta$, by Theorem 3 in Embrechts and Goldie (1980), we know that $\bar{F}_{X_i} \in \mathscr{L}_\beta, i = 1, \ldots, n$. Using Lemma 7 in the Supplementary Material, we have

$$1 - F_{X_i}(x) = e^{-h(x)}, \quad h(x) \sim \beta x, \qquad \text{as } x \to \infty.$$

Now for the square matrix

$$A = \begin{pmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & 1 & a_{23} & \cdots & a_{2n} \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \in \mathbb{R}^{n \times n},$$

where unspecified entries are set to zero, it can be shown that its determinant is $|A| = 1 - a_{12}a_{21}$, and its inverse is

$$A^{-1} = \frac{1}{1 - a_{12}a_{21}} \begin{pmatrix} 1 & -a_{12} & a_{12}a_{23} - a_{13} & \cdots & a_{12}a_{2n} - a_{1n} \\ -a_{21} & 1 & a_{21}a_{13} - a_{23} & \cdots & a_{21}a_{1n} - a_{2n} \\ & & 1 - a_{12}a_{21} & & \\ & & & \ddots & \\ & & & & 1 - a_{12}a_{21} \end{pmatrix}.$$

Denote $\boldsymbol{X} = (X_1, X_2)$, $\tilde{\boldsymbol{X}} = (X_1, \ldots, X_n)$ and the probability density function of $\tilde{\boldsymbol{X}}$ as $f_{\tilde{\boldsymbol{X}}}$. If the support of Y_i is $\mathbb{R}$, then we have

$$\lim_{t \to \infty} \frac{-\log f_{\tilde{\boldsymbol{X}}}(t\boldsymbol{x})}{h(t)} = \lim_{t \to \infty} \frac{\beta t(|\boldsymbol{\alpha}_1^\top \boldsymbol{x}| + \cdots + |\boldsymbol{\alpha}_n^\top \boldsymbol{x}|)}{\beta t} = |\boldsymbol{\alpha}_1^\top \boldsymbol{x}| + \cdots + |\boldsymbol{\alpha}_n^\top \boldsymbol{x}| =: \tilde{g}(\boldsymbol{x}),$$

where $\boldsymbol{\alpha}_i, i = 1, \ldots, n$ are the ith row vector of the matrix A^{-1}. By Proposition 2.2 in Nolde and Wadsworth (2022) or Proposition 3.1 in Nolde (2014), a sequence of scaled random samples $N_k = \{\tilde{\boldsymbol{X}}_1/r_k, \ldots, \tilde{\boldsymbol{X}}_k/r_k\}$ from $f_{\tilde{\boldsymbol{X}}}$ converges in probability onto a limit set $\tilde{G} = \{\boldsymbol{x} \in \mathbb{R}^d : \tilde{g}(\boldsymbol{x}) \leq 1\}$. Then using Proposition 2.4 in Nolde and Wadsworth (2022) we know that, for $\boldsymbol{X}$, which is a two-dimensional subvector of $\tilde{\boldsymbol{X}}$, sample clouds from $\boldsymbol{X}$ converge onto the limit set $G = \{x_1, x_2 \in \mathbb{R} : g(x_1, x_2) \leq 1\}$ with gauge function

$$g(x_1, x_2) = \min_{x_3, \ldots, x_n} \tilde{g}(\boldsymbol{x}).$$

Therefore, by Proposition 3.4 in Nolde and Wadsworth (2022), we have

$$\begin{aligned}
\eta^{-1} &= \min_{x_1,x_2\geq 1} g(x_1,x_2) \\
&= \min_{x_1,x_2\geq 1}\min_{x_3,\dots,x_n} \frac{|x_1-a_{12}x_2+(a_{12}a_{23}-a_{13})x_3+\cdots+(a_{12}a_{2n}-a_{1n})x_n|}{1-a_{12}a_{21}}+ \\
&\quad \frac{|x_2-a_{21}x_1+(a_{21}a_{13}-a_{23})x_3+\cdots+(a_{21}a_{1n}-a_{2n})x_n)|}{1-a_{12}a_{21}}+\sum_{i=3}^{n}|x_i|.
\end{aligned}$$

Using Lemma 4 and 5 in the Supplementary Material, we have

$$\begin{aligned}
\eta &= \Bigg\{\min_{x_1,x_2\geq 1}\min_{\substack{i,j=1,\dots,n\\ i\neq j}} \frac{|a_{2i}x_1-a_{1i}x_2|+|a_{2j}x_1-a_{1j}x_2|}{|a_{2i}a_{1j}-a_{1i}a_{2j}|}\Bigg\}^{-1} \\
&= \Bigg\{\min_{\substack{i,j=1,\dots,n\\ i\neq j}}\min_{x_1,x_2\geq 1} \frac{|a_{2i}x_1-a_{1i}x_2|+|a_{2j}x_1-a_{1j}x_2|}{|a_{2i}a_{1j}-a_{1i}a_{2j}|}\Bigg\}^{-1} \\
&= \Bigg[\min_{\substack{i,j=1,\dots,n\\ i\neq j}}\min\Bigg\{\frac{|a_{2i}-a_{1i}|+|a_{2j}-a_{1j}|}{|a_{2i}a_{1j}-a_{1i}a_{2j}|},\frac{1}{\min(a_{1i},a_{2i})},\frac{1}{\min(a_{1j},a_{2j})}\Bigg\}\Bigg]^{-1} \\
&= c^{-1}.
\end{aligned}$$

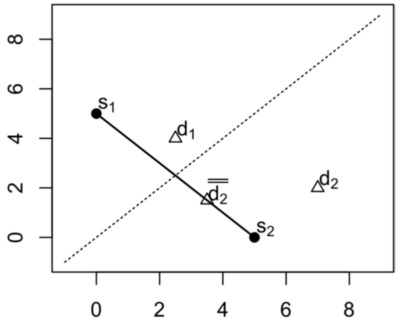
Similarly, when the support of Y_i is $\mathbb{R}_+$, the gauge function of $\tilde{\boldsymbol{X}}$ is $\tilde{g}(\boldsymbol{x})=\sum_{i=1}^{n}\boldsymbol{\alpha}_1^\top\boldsymbol{x}$ for $A^{-1}\boldsymbol{x}\geq\boldsymbol{0}$. Then the gauge function of $\boldsymbol{X}$ becomes $g(x_1,x_2)=\min_{x_3,\dots,x_n\in\mathbb{R}:A^{-1}\boldsymbol{x}\geq\boldsymbol{0}}\sum_{i=1}^{n}\boldsymbol{\alpha}_1^\top\boldsymbol{x}$. Hence, in this case,

$$\begin{aligned}
\eta &= \Big(\min_{\boldsymbol{x}\in\mathbb{R}^n:x_1,x_2\geq 1,A^{-1}\boldsymbol{x}\geq\boldsymbol{0}}\sum_{i=1}^{n}\boldsymbol{\alpha}_1^\top\boldsymbol{x}\Big)^{-1} = \Big(\min_{\boldsymbol{x}\in\mathbb{R}^n:x_1,x_2\geq 1,A^{-1}\boldsymbol{x}\geq\boldsymbol{0}}\sum_{i=1}^{n}|\boldsymbol{\alpha}_1^\top\boldsymbol{x}|\Big)^{-1} \\
&\leq \Big(\min_{\boldsymbol{x}\in\mathbb{R}^n:x_1,x_2\geq 1}\sum_{i=1}^{n}|\boldsymbol{\alpha}_1^\top\boldsymbol{x}|\Big)^{-1}=c^{-1}.
\end{aligned}$$

If (7), then by Lemma 6 in the Supplementary Material, we have $\eta\leq c^{-1}<1$ and thus X_1 and X_2 are asymptotically independent. Notice that although we have assumed $a_{11}=\max_{i\in\{1,\dots,n\}}a_{1i}$ and $a_{22}=\max_{i\in\{1,\dots,n\}}a_{2i}$ at the beginning of this proof, they are only needed to ensure that $a_{11}a_{22}-a_{12}a_{21}\neq 0$, which implies that the matrix *A* is nonsingular, and they do not affect the expression of η. Therefore, the proof is complete.

Proof of Theorem 4.1 By Proposition 3.3, we know that the residual tail dependence coefficient of $(u_n(\boldsymbol{s}_1),u_n(\boldsymbol{s}_2))$ is

Fig. 6 Explaination for the derivation of the limiting η in the type G model



$$\eta_n(\boldsymbol{s}_1, \boldsymbol{s}_2) = \max_{i,j \in \{1,\dots,J\}, i \neq j} \left\{ \frac{|\tilde{a}_{1i}\tilde{a}_{2j} - \tilde{a}_{1j}\tilde{a}_{2i}|}{|\tilde{a}_{1i} - \tilde{a}_{2i}| + |\tilde{a}_{1j} - \tilde{a}_{2j}|}, \tilde{a}_{1i} \wedge \tilde{a}_{2i}, \tilde{a}_{1j} \wedge \tilde{a}_{2j} \right\},$$

where $\tilde{a}_{ri} = \frac{G(\|\boldsymbol{s}_r - \boldsymbol{d}_i\|)}{G(l_r)}, r = 1, 2, i = 1, \dots, J$ and $l_r = \min_{i \in \{1,\dots,J\}} \|\boldsymbol{s}_r - \boldsymbol{d}_i\|$. If $l_1 \neq l_2$, say $l_1 < l_2$, then one can add one or more mesh nodes close to $\boldsymbol{s}_2$ such that $l_1 = l_2'$ holds in the resulting mesh. So without loss of generality, we assume that $l_1 = l_2 = l$.

For fixed $l > 0$, it is clear that $\tilde{a}_{1i} \wedge \tilde{a}_{2i} \leq G(\|\boldsymbol{s}_1 - \boldsymbol{s}_2\|/2)/G(l)$, where the equality is achieved when $\boldsymbol{d}_i$ is at the midpoint of $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$, i.e., $\|\boldsymbol{s}_1 - \boldsymbol{d}_i\| = \|\boldsymbol{s}_2 - \boldsymbol{d}_i\| = \|\boldsymbol{s}_1 - \boldsymbol{s}_2\|/2$. Since our interest is in the case when $l \to 0$, we assume that l is small and $l < \|\boldsymbol{s}_1 - \boldsymbol{s}_2\|/2$. We now focus on the term $\frac{|\tilde{a}_{1i}\tilde{a}_{2j} - \tilde{a}_{1j}\tilde{a}_{2i}|}{|\tilde{a}_{1i} - \tilde{a}_{2i}| + |\tilde{a}_{1j} - \tilde{a}_{2j}|}$. Substitute $\tilde{a}_{ri} = \frac{G(\|\boldsymbol{s}_r - \boldsymbol{d}_i\|)}{G(l)}$, and denote this term by f, i.e.,

$$f(\boldsymbol{d}_1, \boldsymbol{d}_2) = \frac{|G(\|\boldsymbol{d}_1 - \boldsymbol{s}_1\|)G(\|\boldsymbol{d}_2 - \boldsymbol{s}_2\|) - G(\|\boldsymbol{d}_1 - \boldsymbol{s}_2\|)G(\|\boldsymbol{d}_2 - \boldsymbol{s}_1\|)|}{\{|G(\|\boldsymbol{d}_1 - \boldsymbol{s}_1\|) - G(\|\boldsymbol{d}_1 - \boldsymbol{s}_2\|)| + |G(\|\boldsymbol{d}_2 - \boldsymbol{s}_1\|) - G(\|\boldsymbol{d}_2 - \boldsymbol{s}_2\|)|\}G(l)}.$$

Then we need to find the maximum of f for $\boldsymbol{d}_1, \boldsymbol{d}_2 \in \mathbb{R}^2 \setminus (\mathcal{B}(\boldsymbol{s}_1, l) \cup \mathcal{B}(\boldsymbol{s}_2, l))$, where $\mathcal{B}(\boldsymbol{s}_r, l) = \{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x} - \boldsymbol{s}_r\| < l\}$ is the open ball of radius l centered at $\boldsymbol{s}_r$.

Denote the line segment from $\boldsymbol{s}_1$ to $\boldsymbol{s}_2$ by $LS(\boldsymbol{s}_1, \boldsymbol{s}_2)$, and its perpendicular bisector by $PB(\boldsymbol{s}_1, \boldsymbol{s}_2)$. If $\boldsymbol{d}_1 \in PB(\boldsymbol{s}_1, \boldsymbol{s}_2)$ and $\boldsymbol{d}_2 \in PB(\boldsymbol{s}_1, \boldsymbol{s}_2)$, then clearly the function f is not well-defined. By Proposition 3.3 and the discussion before it, we know that in this case $\tilde{a}_{1i}\tilde{a}_{2j} - \tilde{a}_{1j}\tilde{a}_{2i} = 0$ and thus the first term, f, is set to zero, implying that these points can be neglected since $f \leq \tilde{a}_{1i} \wedge \tilde{a}_{2i}$. If only one of $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ is on the line $PB(\boldsymbol{s}_1, \boldsymbol{s}_2)$, then we have $f(\boldsymbol{d}_1, \boldsymbol{d}_2) = G(\|\boldsymbol{s}_1 - \boldsymbol{s}_2\|/2)/G(l)$.

For $\boldsymbol{d}_1, \boldsymbol{d}_2 \notin PB(\boldsymbol{s}_1, \boldsymbol{s}_2)$, without loss of generality we assume that $\boldsymbol{d}_1$ is closer to $\boldsymbol{s}_1$ than to $\boldsymbol{s}_2$, i.e., $\|\boldsymbol{d}_1 - \boldsymbol{s}_1\| < \|\boldsymbol{d}_1 - \boldsymbol{s}_2\|$, as shown in Fig. 6. One can show that if $\boldsymbol{d}_2$ is also closer to $\boldsymbol{s}_1$ than to $\boldsymbol{s}_2$, then its symmetric point $\bar{\boldsymbol{d}}_2$ about the line $PB(\boldsymbol{s}_1, \boldsymbol{s}_2)$ satisfies $f(\boldsymbol{d}_1, \bar{\boldsymbol{d}}_2) \geq f(\boldsymbol{d}_1, \boldsymbol{d}_2)$. This implies that to obtain the maximum of f, it is sufficient to consider points $\boldsymbol{d}_1, \boldsymbol{d}_2$ located on the two sides of the line $PB(\boldsymbol{s}_1, \boldsymbol{s}_2)$. Hence we can assume that $\boldsymbol{d}_1$ is closer to $\boldsymbol{s}_1$ than to $\boldsymbol{s}_2$, and $\boldsymbol{d}_2$ is closer to $\boldsymbol{s}_2$ than to $\boldsymbol{s}_1$. Since G is strictly decreasing on $[0, \infty)$, the function f can then be simplified to

$$f(\boldsymbol{d}_1, \boldsymbol{d}_2) = \frac{G(\|\boldsymbol{d}_1 - \boldsymbol{s}_1\|)G(\|\boldsymbol{d}_2 - \boldsymbol{s}_2\|) - G(\|\boldsymbol{d}_1 - \boldsymbol{s}_2\|)G(\|\boldsymbol{d}_2 - \boldsymbol{s}_1\|)}{\{G(\|\boldsymbol{d}_1 - \boldsymbol{s}_1\|) - G(\|\boldsymbol{d}_1 - \boldsymbol{s}_2\|) + G(\|\boldsymbol{d}_2 - \boldsymbol{s}_1\|) - G(\|\boldsymbol{d}_2 - \boldsymbol{s}_2\|)\}G(l)}.$$

Furthermore, due to the convexity of G, it is straightforward to show that for $\boldsymbol{d} \in \mathbb{R}^2 \setminus \{\mathcal{B}(\boldsymbol{s}_1, l) \cup \mathcal{B}(\boldsymbol{s}_2, l)\}$, the range of $|G(\|\boldsymbol{d} - \boldsymbol{s}_1\|) - G(\|\boldsymbol{d} - \boldsymbol{s}_2\|)|$ is $[0, G(l) - G(\|\boldsymbol{s}_1 - \boldsymbol{s}_2\| - l)]$ with its minimum obtained on the line $PB(\boldsymbol{s}_1, \boldsymbol{s}_2)$ and its maximum at the intersection points of the line passing through $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ and the two circles with radius l centered at $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$. Now if we fix the point $\boldsymbol{d}_1$, then by the intermediate value theorem (Tao 2016, Theorem 9.7.1), for any $\boldsymbol{d}_2$ such that $\|\boldsymbol{d}_2 - \boldsymbol{s}_2\| < \|\boldsymbol{d}_2 - \boldsymbol{s}_1\|$ and $\boldsymbol{d}_2 \notin \mathcal{B}(\boldsymbol{s}_2, l)$, there always exists one point $\bar{\bar{\boldsymbol{d}}}_2$ on the line segment $LS((\boldsymbol{s}_1 + \boldsymbol{s}_2)/2, \boldsymbol{s}_2) \setminus \mathcal{B}(\boldsymbol{s}_2, l)$ such that $G(\|\boldsymbol{d}_2 - \boldsymbol{s}_1\|) - G(\|\boldsymbol{d}_2 - \boldsymbol{s}_2\|) = G(\|\bar{\bar{\boldsymbol{d}}}_2 - \boldsymbol{s}_1\|) - G(\|\bar{\bar{\boldsymbol{d}}}_2 - \boldsymbol{s}_2\|)$, and for $i = 1, 2$, $G(\|\bar{\bar{\boldsymbol{d}}}_2 - \boldsymbol{s}_i\|) \geq G(\|\boldsymbol{d}_2 - \boldsymbol{s}_i\|)$, yielding that $f(\boldsymbol{d}_1, \bar{\bar{\boldsymbol{d}}}_2) \geq f(\boldsymbol{d}_1, \boldsymbol{d}_2)$. Hence, it is sufficient to consider points $\boldsymbol{d}_1 \in LS(\boldsymbol{s}_1, (\boldsymbol{s}_1 + \boldsymbol{s}_2)/2) \setminus \mathcal{B}(\boldsymbol{s}_1, l)$ and $\boldsymbol{d}_2 \in LS(\boldsymbol{s}_2, (\boldsymbol{s}_1 + \boldsymbol{s}_2)/2) \setminus \mathcal{B}(\boldsymbol{s}_2, l)$, and the optimization problem can be rewritten as

$$\sup_{x_1, x_2 \in [l, h)} \bar{f}(x_1, x_2) = \sup_{x_1, x_2 \in [l, h)} \frac{G(x_1)G(x_2) - G(2h - x_1)G(2h - x_2)}{\{G(x_1) - G(2h - x_1) + G(x_2) - G(2h - x_2)\}G(l)}.$$

Since G is absolutely continuous, we have the partial derivative of $\bar{f}$ with respect to x_1 as

$$\frac{\partial \bar{f}}{\partial x_1} = a(x_1, x_2)[G'(x_1)\{G(x_2) - G(2h - x_2)\} - G'(2h - x_1)\{G(x_1) - G(2h - x_2)\}],$$

where $a(x_1, x_2) = \frac{G(x_2) - G(2h - x_2)}{\{G(x_1) - G(2h - x_1) + G(x_2) - G(2h - x_2)\}^2 G(l)} > 0$. For $x \in [l, h)$, due to the convexity of G, we have that $G'(x) \leq G'(2h - x)$. Hence,

$$\frac{\partial \bar{f}}{\partial x_1} \leq a(x_1, x_2)G'(2h - x_1)\{G(x_2) + G(2h - x_2) - G(x_1) - G(2h - x_1)\}.$$

As G is monotonically decreasing, we know that $G'(2h - x_1) \leq 0$. Furthermore, the convexity of G implies that the function $g(x) = G(x) + G(2h - x)$ is monotonically decreasing on $[l, h)$. Therefore, for fixed $x_2 \in [l, h)$, we know that if $x_1 \geq x_2$, then $\frac{\partial \bar{f}}{\partial x_1} \leq 0$. This implies that

$$\bar{f}(x_1, x_2) \leq \bar{f}(x_2, x_2) = \{G(x_2) + G(2h - x_2)\}/\{2G(l)\} \leq \{G(l) + G(2h - l)\}/\{2G(l)\},$$

where the maximum is obtained when $x_1 = x_2 = l$.

Above all, we know that for $\boldsymbol{d}_1, \boldsymbol{d}_2 \in \mathbb{R}^2 \setminus (\mathcal{B}(\boldsymbol{s}_1, l) \cup \mathcal{B}(\boldsymbol{s}_2, l))$,

$$
\begin{aligned}
f(\boldsymbol{d}_1, \boldsymbol{d}_2) &\leq \max\left\{\frac{G(\|\boldsymbol{s}_1-\boldsymbol{s}_2\|/2)}{G(l)}, \frac{G(l)+G(\|\boldsymbol{s}_1-\boldsymbol{s}_2\|-l)}{2G(l)}\right\} \\
&= 1/2 + G(\|\boldsymbol{s}_1-\boldsymbol{s}_2\|-l)/G(l),
\end{aligned}
$$

where the last equality holds due to the convexity of G, and the maximum of f is achieved when $\boldsymbol{d}_1, \boldsymbol{d}_2$ are the intersection points of the line segment $LS(\boldsymbol{s}_1, \boldsymbol{s}_2)$ and the two circles with radius l centered at $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$. Hence, as the mesh becomes finer and the set of mesh nodes M_n becomes denser in $\mathcal{D}$, which means that $l \to 0$, the residual tail dependence function $\eta(\boldsymbol{s}_1, \boldsymbol{s}_2)$ tends to $1/2 + G(\|\boldsymbol{s}_1-\boldsymbol{s}_2\|)/G(0)$. This completes the proof.

***Proof of Theorem* 4.2** We first consider the case when $\mathcal{M}(B)$ has support $\mathbb{R}$ for any bounded Borel set $B \subset \mathbb{R}$. By Proposition 3.3, the residual tail dependence coefficient of the approximation model $(u_n(s_1), u_n(s_2))$ of the form (10) is

$$
\eta = c^{-1} = \max_{i,j=0,\ldots,n_2-1, i\neq j}\left\{\frac{|a_{2i}a_{1j}-a_{1i}a_{2j}|}{|a_{2i}-a_{1i}|+|a_{2j}-a_{1j}|}, \min(a_{1i}, a_{2i}), \min(a_{1j}, a_{2j})\right\},
$$

with $a_{1i} = G(s_1-t_i)/G(s_1-t_{n_1-1})$, $0 \leq i \leq n_1-1$, and for $n_1 \leq i \leq n_2-1$, $a_{1i} = 0$, and $a_{2i} = G(s_2-t_i)/G(s_2-t_{n_2-1})$, $0 \leq i \leq n_2-1$.

Since G is strictly decreasing, the sequences $\{a_{1i}, 0 \leq i \leq n_1-1\}$ and $\{a_{2i}, 0 \leq i \leq n_2-1\}$ are strictly increasing. Consequently, this yields that $\max_{i=0,\ldots,n_2-1}\min(a_{1i}, a_{2i}) = a_{2(n_1-1)}$. For the first term inside the maximum operation of the expression of η, we have

$$
\begin{aligned}
&\frac{|a_{2i}a_{1j}-a_{1i}a_{2j}|}{|a_{2i}-a_{1i}|+|a_{2j}-a_{1j}|} \\
\leq &\frac{\max(a_{1j}, a_{2j})|a_{2i}-a_{1i}|}{|a_{2i}-a_{1i}|+|a_{2j}-a_{1j}|} = \frac{\max(a_{1j}, a_{2j})}{1+|a_{2j}-a_{1j}|/|a_{2i}-a_{1i}|} \\
\leq &\frac{\max(a_{1j}, a_{2j})}{1+|a_{2j}-a_{1j}|} = \frac{1}{1+\{1-\min(a_{1j}, a_{2j})\}/\max(a_{1j}, a_{2j})} \\
\leq &\frac{1}{2-a_{2(n_1-1)}},
\end{aligned}
$$

where the equality is obtained when $i = n_2-1$ and $j = n_1-1$. Thus we have η for $(u_n(s_1), u_n(s_2))^\top$ as

$$
\eta = \max\left(\frac{1}{2-a_{2(n_1-1)}}, a_{2(n_1-1)}\right) = \frac{1}{2-a_{2(n_1-1)}}.
$$

If we construct the matrix A as in Proposition 3.3, then the result $c = (2-a_{2(n_1-1)})$ implies that the problem $\min_{\boldsymbol{x}\in\mathbb{R}^{n_2-1}: x_1\geq 1, x_2\geq 1}\sum_{i=1}^{n_2-1}|\boldsymbol{\alpha_i}^\top\boldsymbol{x}|$ achieves its global minimum at $\boldsymbol{x}^* = (1,1,0,\ldots,0)$. Since $\boldsymbol{x}^*$ satisfies that $A^{-1}\boldsymbol{x} \geq 0$, we know that

$\min_{\boldsymbol{x}\in\mathbb{R}^{n_2-1}:x_1\geq 1,x_2\geq 1,A^{-1}\boldsymbol{x}\geq\mathbf{0}}\sum_{i=1}^{n_2-1}|\boldsymbol{\alpha_i}^\top\boldsymbol{x}|$ also achieves its global minimum at $\boldsymbol{x}^*$. Hence, when $\mathcal{M}(B)$ has support $\mathbb{R}_+$, we also have $\eta=c^{-1}=1/(2-a_{2(n_1-1)})$.

Therefore, as $m_n\to\infty$ and the set of mesh nodes M_n becomes denser in $\mathcal{D}$, the limiting residual tail dependence function is $1/\{2-G(h)/G(0)\}$.

Proof of Theorem 4.3 Without loss of generality we assume that $0\leq s_1<s_2$. The representation (12) implies that $u(s_2)$ can be represented as the convolution of two independent random variables

$$u(s_2)=e^{-a(s_2-s_1)}u(s_1)+\int_{s_1}^{s_2}e^{-a(s_2-s)}\mathrm{d}z(at).$$

Since $u(s)=V$ and $\bar{F}_V\in\mathscr{L}_\beta$, we know that $u(s_1),u(s_2)$ have exponential tails with index β and $e^{-a(s_2-s_1)}u(s_1)$ has an exponential tail with index $e^{a(s_2-s_1)}\beta$. Denote $V_1=e^{-a(s_2-s_1)}u(s_1)$. The self-decomposability of V_1 yields its infinite divisibility, which further implies that the Wiener condition is satisfied, i.e. $M_{V_1}(\beta+it)\neq 0,t\in\mathbb{R}$, where M_{V_1} is the moment generating function of V_1; see Theorem 25.17 of Sato (1999) for more details. Then Lemma 2.5 in Pakes (2004) yields that $\int_{s_1}^{s_2}e^{-a(s_2-s)}\mathrm{d}z(at)$ must have an exponential tail with index β. Therefore, the result in Proposition 3.3 and Example 2 gives the asymptotic independence between $u(s_1)$ and $u(s_2)$, and their residual tail dependence coefficient is $\eta=1/(2-e^{-a(s_2-s_1)})$.

Proof of Proposition 4.4 By Sgibnev (1990) and Shimura and Watanabe (2005), we know that an infinitely divisible distribution is convolution tail equivalent if and only if its normalized Lévy measure is convolution tail equivalent. That is, for any $t\in\mathbb{R}$,

$$\bar{F}_{z(1)}\in\mathscr{S}_\beta\qquad\Leftrightarrow\qquad\frac{\mathbf{1}_{x>1}U_{z(1)}([x,\infty))}{U_{z(1)}([1,\infty))}\in\mathscr{S}_\beta.$$

From Barndorff-Nielsen (1998) we know that if $U_{z(1)}([x,\infty))$ is continuous in x on $(0,\infty))$, then we have the relation $U_u([x,\infty))=\int_x^\infty s^{-1}U_{z(1)}([s,\infty))\mathrm{d}s$. Hence, for $x>1$,

$$U_u([\log x,\infty))=\int_{\log x}^\infty s^{-1}U_{z(1)}([s,\infty))\mathrm{d}s=\int_x^\infty U_{z(1)}([\log t,\infty))/(t\log t)\mathrm{d}t.$$

Since $\frac{\mathbf{1}_{x>1}U_{z(1)}([x,\infty))}{U_{z(1)}([1,\infty))}\in\mathscr{S}_\beta$, we have that $U_{z(1)}([\log t,\infty))\in\mathrm{RV}_{-\beta}$, which further implies that $U_{z(1)}([\log t,\infty))/(t\log t)\in\mathrm{RV}_{-\beta-1}$. By Karamata's theorem (cf. Theorem 2.1 in Resnick (2007)),

$$\lim_{x\to\infty}\frac{U_{z(1)}([\log x,\infty))/\log x}{\int_x^\infty U_{z(1)}([\log t,\infty))/(t\log t)\mathrm{d}t}=\beta.$$

Hence,

$$\lim_{x\to\infty}\frac{U_u([\log x,\infty))}{U_{z(1)}([\log t,\infty))}=\lim_{x\to\infty}\frac{1}{\beta\log x}=1/\beta.$$

Therefore, using Theorem 1 in Cline (1986), we have that $\frac{\mathbf{1}_{x>1}U_u([x,\infty))}{U_u([1,\infty))}\in\mathscr{S}_\beta$ and thus $\bar{F}_u\in\mathscr{S}_\beta$.

Denote the Lévy measure of $z(t)$ as $U_{z(t)}$. Since $F_{z(1)}\in\mathscr{S}_\beta$, we know that $\frac{\mathbf{1}_{x>0}U_{z(1)}}{U_{z(1)}([1,\infty))}\in\mathscr{S}_\beta$. By the definition of Lévy process, we have $U_{z(t)}=tU_{z(1)}$, for any $t>0$. Hence, using Theorem 1 in Cline (1986), we know that $\frac{\mathbf{1}_{x>0}U_{z(t)}([x,\infty))}{U_{z(t)}([1,\infty))}\in\mathscr{S}_\beta$ and thus $\bar{F}_{z(t)}\in\mathscr{S}_\beta$.

7 Supplementary information

The Proofs of Proposition 3.1 and 3.2, further details of the finite element approximation of the SPDE model, additional numerical experiments and simple theoretical results are given in the Supplementary Material.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10687-025-00517-4.

Acknowledgements We are grateful for Jonathan Tawn's comments on an early version of the manuscript. We also thank Ioannis Papastathopoulos for pointing out one part in our proof of Proposition 3.3 was not clearly written. We would also like to thank the Associate Editor and three anonymous reviewers for their constructive comments. All the comments have led to an improvement of our manuscript. This work started when Zhongwei Zhang was a PhD student at KAUST, and continued during his postdoc at the University of Geneva. Zhongwei Zhang thanks KAUST and the University of Geneva for providing the environment to work on this project.

Author Contributions All authors contributed to the conception and proof of this work. All authors contributed to the writing of the manuscript. All authors reviewed the manuscript.

Funding Zhongwei Zhang and Raphaël Huser were partly supported by funding from the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No. OSR-CRG2020-4394.
Sebastian Engelke was supported by Swiss National Science Foundation (Grant 186858).

Data Availability No datasets were generated or analysed during the current study.

Code availability The code is available on GitHub (https://github.com/Zhongwei-Zhang/Tail_dependence_moving_average_processes).

Declarations

Competing interests The authors declare no competing interests.

Ethics approval and consent to participate All authors approve and consent to participate.

Consent for publication All authors consent for publication.

Materials availability Not applicable.

References

Asar, Ö., Bolin, D., Diggle, P.J., Wallin, J.: Linear mixed effects models for non-Gaussian continuous repeated measurement data. J. R. Stat. Soc.: Ser. C: Appl. Stat. **69**(5), 1015–1065 (2020)
Barndorff-Nielsen, O.E.: Exponentially decreasing distributions for the logarithm of particle size. Proc. A **353**, 401–419 (1977)
Barndorff-Nielsen, O.E.: Processes of normal inverse Gaussian type. Finance Stoch. **2**, 41–68 (1998)
Barndorff-Nielsen, O.E., Shephard, N.: Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. J. R. Stat. Soc. Ser. B Stat Methodol. **63**(2), 167–241 (2001)
Bolin, D.: Spatial Matérn fields driven by non-Gaussian noise. Scand. J. Stat. **41**, 557–579 (2014)
Bolin, D., Wallin, J.: Multivariate type G Matérn stochastic partial differential equation random fields. J. R. Stat. Soc. Ser. B Stat Methodol. **82**(1), 215–239 (2020)
Carr, P., Geman, H., Madan, D.B., Yor, M.: The fine structure of asset returns: An empirical investigation. J. Bus. **75**(2), 305–333 (2002)
Cline, D.B.H.: Convolution tails, product tails and domains of attraction. Probab. Theory Relat. Fields **72**, 529–557 (1986)
Coles, S., Heffernan, J., Tawn, J.A.: Dependence measures for extreme value analyses. Extremes **2**(4), 339–365 (1999)
Cressie, N., Pavlicová, M.: Calibrated spatial moving average simulations. Stat. Model. **2**, 267–279 (2002)
Davis, R., Resnick, S.: Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution. Stoch. Process. Appl. **30**, 41–68 (1988)
Eberlein, E.: Application of generalized hyperbolic Lévy motions to finance. In: Barndorff-Nielsen, O.E., Resnick, S.I., Mikosch, T. (eds.) Lévy Processes - Theory and Applications, pp. 319–336. Birkhäuser, Boston, MA (2001)
Embrechts, P., Goldie, C.M.: On closure and factorization properties of subexponential and related distributions. J. Aust. Math. Soc. **29**, 243–256 (1980)
Embrechts, P., Goldie, C.M.: On convolution tails. Stoch. Process. Appl. **13**, 263–278 (1982)
Engelke, S., Opitz, T., Wadsworth, J.L.: Extremal dependence of random scale constructions. Extremes **22**, 623–666 (2019)
Fasen, V.: Extremes of regularly varying Lévy-driven mixed moving average processes. Adv. Appl. Probab. **37**, 993–1014 (2005)
Fasen, V.: Extremes of subexponential Lévy driven moving average processes. Stoch. Process. Appl. **116**, 1066–1087 (2006)
Fasen, V.: Extremes of Lévy driven mixed MA processes with convolution equivalent distributions. Extremes **12**, 265–296 (2009)
Fasen, V., Klüppelberg, C., Schlather, M.: High-level dependence in time series models. Extremes **13**, 1–33 (2010)
Gnecco, N., Meinshausen, N., Peters, J., Engelke, S.: Causal discovery in heavy-tailed models. Ann. Statist. **49**(3), 1755–1778 (2021)
Higdon, D.: Space and space-time modeling using process convolutions. In: Anderson, C.W., Barnett, V., Chatwin, P.C., El-Shaarawi, A.H. (eds.) Quantitative Methods for Current Environmental Issues, pp. 37–56. Springer, London (2002)
Huser, R., Wadsworth, J.L.: Modeling spatial processes with unknown extremal dependence class. J. Amer. Statist. Assoc. **114**(525), 434–444 (2019)
Janssen, A., Drees, H.: A stochastic volatility model with flexible extremal dependence structure. Bernoulli **22**(3), 1448–1490 (2016)
Janssen, A., Drees, H.: Joint exceedances of random products. Ann. Inst. H. Poincaré Probab. Statist. **54**(1), 437–465 (2018)
Jurek, Z.J., Vervaat, W.: An integral representation for selfdecomposable Banach space valued random variables. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **62**, 247–262 (1983)

Krupskii, P., Huser, R.: Modeling spatial tail dependence with Cauchy convolution processes. Electron. J. Stat. **16**, 6135–6174 (2022)
Kuusela, M., Stein, M.L.: Locally stationary spatio-temporal interpolation of Argo profiling float data. Proc. R. Soc. A. **474**, 20180400 (2018)
Leadbetter, M.R., Lindgren, G., Rootzén, H.: Extremes and Related Properties of Random Sequences and Processes. Springer, New York, NY (1983)
Ledford, A.W., Tawn, J.A.: Statistics for near independence in multivariate extreme values. Biometrika **83**(1), 169–187 (1996)
Lindgren, F., Rue, H., Lindström, J.: An explicit link between the Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B Stat Methodol. **73**, 423–498 (2011)
McNeil, A.J., Frey, R., Embrechts, P.: Quantitative Risk Management: Concepts. Techniques and Tools. Princeton University Press, Princeton, NJ (2005)
Nolde, N.: Geometric interpretation of the residual dependence coefficient. J. Multivar. Anal. **123**, 85–95 (2014)
Nolde, N., Wadsworth, J.L.: Linking representations for multivariate extremes via a limit set. Adv. Appl. Probab. **54**(3), 688–717 (2022)
Nolde, N., Zhou, C.: Extreme value analysis for financial risk management. Annu. Rev. Stat. Appl. **8**, 217–240 (2021)
Opitz, T.: Spatial random field models based on Lévy indicator convolutions. arXiv:1710.06826 (2018)
Pakes, A.G.: Convolution equivalence and infinity divisibility. J. Appl. Probab. **41**(2), 407–424 (2004)
Prause, K.: The generalized hyperbolic model: Estimation, financial derivatives, and risk measures. PhD thesis, Albert-Ludwigs-Universität Freiburg (1999)
Rajput, B.S., Rosinski, J.: Spectral representations of infinitely divisible processes. Probab. Theory Relat. Fields **82**, 451–487 (1989)
Resnick, S.I.: Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer, New York, NY (2007)
Rodrigues, A., Diggle, P.J.: A class of convolution-based models for spatio-temporal processes with non-seperable covariance structure. Scand. J. Stat. **37**, 553–567 (2010)
Rønn-Nielsen, A., Stehr, M.: Extremes of Lévy-driven spatial random fields with regularly varying Lévy measure. Stoch. Process. Appl. **150**, 19–49 (2022)
Rootzén, H.: Extremes of moving averages of stable processes. Ann. Prob. **6**(5), 847–869 (1978)
Rootzén, H.: Extreme value theory for moving average processes. Ann. Prob. **14**(2), 612–652 (1986)
Rosinski, J.: On a class of infinitely divisible processes represented as mixtures of Gaussian processes. In: Cambanis, S., Samorodnitsky, G., Taqqu, M.S. (eds.) Stable Processes and Related Topics: A Selection of Papers from the Mathematical Sciences Institute Workshop, January 9–13, 1990, pp. 27–41. Birkhäuser, Boston, MA (1991)
Sato, K.: Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press, Cambridge (1999)
Sgibnev, M.S.: Asymptotics of infinitely divisible distributions on $\mathbb{R}$. Sib. Math. J. **31**, 115–119 (1990)
Shimura, T., Watanabe, T.: Infinite divisibility and generalized subexponentiality. Bernoulli **11**(3), 445–469 (2005)
Stehr, M., Rønn-Nielsen, A.: Extreme value theory for spatial random fields - with application to a Lévy-driven field. Extremes **24**, 753–795 (2021)
Stehr, M., Rønn-Nielsen, A.: Extremes of subexponential Lévy-driven random fields in the Gumbel domain of attraction. Extremes **25**, 79–105 (2022)
Tao, T.: Analysis I, 3rd edn. Springer, Singapore (2016)
Ver Hoef, J.M., Peterson, E.E.: A moving average approach for spatial statistical models for stream networks. J. Amer. Statist. Assoc. **105**(489), 6–18 (2010)
Wallin, J., Bolin, D.: Geostatistical modelling using non-Gaussian Matérn fields. Scand. J. Stat. **42**, 872–890 (2015)
Wang, Y., Stoev, S.A.: Conditional sampling for spectrally discrete max-stable random fields. Adv. Appl. Probab. **43**, 461–483 (2011)

Authors and Affiliations

Zhongwei Zhang[1] · **David Bolin**[2] · **Sebastian Engelke**[1] · **Raphaël Huser**[2]

✉ Zhongwei Zhang
zhongwei.zhangss@gmail.com

David Bolin
david.bolin@kaust.edu.sa

Sebastian Engelke
sebastian.engelke@unige.ch

Raphaël Huser
raphael.huser@kaust.edu.sa

[1] Research Institute for Statistics and Information Science, University of Geneva, Geneva 1205, Switzerland

[2] Statistics Program, CEMSE Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia