Conditional independence among max-stable laws

Ioannis Papastathopoulos[*],
Kirstin Strokorb[†]

September 16, 2015

Abstract

Let X be a max-stable random vector with positive continuous density. It is proved that the conditional independence of any collection of disjoint subvectors of X given the remaining components implies their joint independence. We conclude that a broad class of tractable max-stable models cannot exhibit an interesting Markov structure.

Keywords: Conditional independence, exponent measure, Markov structure, max-stable random vector, Möbius inversion
2010 MSC: Primary 60G70;
Secondary 62H05

1 Introduction

As pointed out by Dawid (1979) *independence* and *conditional independence* are key concepts in the theory of probability and statistical inference. A collection of (not necessarily real-valued) random variables $Y_1, \dots, Y_k$ on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ are called conditionally independent given the random variable Z (on the same probability space) if

$$\mathbb{P}(Y_1 \in A_1, \dots, Y_k \in A_k \mid Z) = \prod_{i=1}^{k} \mathbb{P}(Y_i \in A_i \mid Z) \qquad \mathbb{P}\text{-a.s.},$$

for any measurable sets $A_1, \dots, A_k$ from the respective state spaces. The conditioning is meant with respect to the σ-algebra generated by Z. A particularly important example for the conditional independence to be an omnipresent attribute are the *Gaussian Markov random fields* that have evolved as a useful tool in spatial statistics (Lauritzen 1996, Rue & Held 2005). Here, the zeroes of the *precision matrix* (the inverse of the covariance matrix) of a Gaussian random vector represent precisely the conditional independence of

[*]School of Mathematics and Maxwell Institute, University of Edinburgh, Edinburgh, EH9 3FD Email: i.papastathopoulos@ed.ac.uk

[†]Institute of Mathematics, University of Mannheim, D-68131 Mannheim, Email: strokorb@math.uni-mannheim.de

the respective components conditioned on the remaing components of the random vector. Hence, sparse precision matrices are desirable for statistical inference.

In the analysis of the extreme values of a distribution (rather than fluctuations around mean values) *max-stable* models have been frequently considered. We refer to Blanchet & Davison (2011), Buishand et al. (2008), Engelke et al. (2014), Naveau et al. (2009) for some spatial applications among many others. Their popularity originates from the fact that max-stable distributions arise precisely as possible limits of location-scale normalizations of i.i.d. random elements. A random vector X is called max-stable if it satisfies the distributional equality $a_n X + b_n \overset{\mathcal{D}}{=} \max(X^{(1)}, \dots, X^{(n)})$ for independent copies $X^{(1)}, \dots, X^{(n)}$ of X for some appropriate normalizing sequences $a_n > 0$ and $b_n \in \mathbb{R}$, where all operations are meant componentwise. If the components X_i of X are *standard Fréchet* distributed, i.e. $\mathbb{P}(X_i \le x) = \exp(-1/x)$ for $x \in (0,\infty)$, we have $a_n = n$ and $b_n = 0$ and the random vector X will be called *simple max-stable*.

Let I be a non-empty finite set. It is well-known (cf. e.g. Resnick (2008)) that the distribution functions G of simple max-stable random vectors $X = (X_i)_{i\in I}$ are in a one-to-one correspondence with Radon measures H on some reference sphere $S_+ = \{\omega \in [0,\infty)^I : \|\omega\| = 1\}$ that satisfy the moment conditions $\int \omega_i H(\mathrm{d}\omega) = 1$, $i \in I$. The correspondence between G and H is given by the relation

$$G(x) = \mathbb{P}(X_i \le x_i,\, i \in I) = \exp\left(-\int_{S_+} \max_{i\in I} \frac{\omega_i}{x_i}\; H(\mathrm{d}\omega)\right), \qquad x \in (0,\infty)^I.$$

Here, $\|\cdot\|$ can be any norm on $\mathbb{R}^I$ and H is often called *angular* or *spectral measure*.

In general, neither does independence imply conditional independence nor does conditional independence imply independence of the subvectors of a random vector. Consider the following two simple examples which illustrate this fact in the case of Gaussian random vectors (Example 1) and max-stable random vectors (Example 2). For notational convenience, we write $X \perp\!\!\!\perp Y$ if X and Y are independent and $X \perp\!\!\!\perp Y \mid Z$ if X and Y are conditionally independent given Z and likewise use the instructive notation $\perp\!\!\!\perp_{i=1}^k X_i$ and $\perp\!\!\!\perp_{i=1}^k X_i \mid Z$ if more than two random elements are involved.

Example 1. Let X_1, X_2, X_3 be three independent standard normal random variables and, moreover, $X_4 = X_1 + X_2$ and $X_5 = X_1 + X_2 + X_3$. Then all subvectors of $(X_i)_{i=1}^5$ are Gaussian and

$$X_1 \perp\!\!\!\perp X_2, \qquad \text{but not} \qquad X_1 \perp\!\!\!\perp X_2 \mid X_5, \tag{1}$$

$$\text{whereas} \qquad X_1 \perp\!\!\!\perp X_5 \mid X_4, \qquad \text{but not} \qquad X_1 \perp\!\!\!\perp X_5. \tag{2}$$

Example 2. Let X_1, X_2, X_3 be three independent standard Fréchet random variables and, moreover, $X_4 = \max(X_1, X_2)$ and $X_5 = \max(X_1, X_2, X_3)$. Then all subvectors of $(X_i)_{i=1,\dots,5}$ are max-stable and both relations (1) and (2) hold true also in this setting.

However, if the distribution of a max-stable random vector has a positive continuous density, then conditional independence of any two subvectors conditioned on the remaining components implies already their independence. To be precise, when we say that a random vector has a *positive continuous density*, we mean that the joint distribution of its components has a positive continuous density. The following theorem is the main result of the present article. If $X = (X_i)_{i\in I}$ is a random vector, we write X_A for the subvector $(X_i)_{i\in A}$ if $A \subset I$. The same convention applies to non-random vectors $x = (x_i)_{i\in I}$.

Theorem 1. *Let $X = (X_i)_{i\in I}$ be a simple max-stable random vector with positive continuous density. Then, for any disjoint non-empty subsets A and B of I, the conditional independence $X_A \perp\!\!\!\perp X_B \mid X_{I\setminus(A\cup B)}$ implies the independence $X_A \perp\!\!\!\perp X_B$.*

A proof of this theorem is given in Section 3. Beforehand, some comments are in order.

(a) First, the requirement of a positive continuous density for X is much less restrictive than requiring the spectral measure H of X to admit such a density, cf. Beirlant et al. (2004) pp. 262-264 and references therein. For instance, fully independent variables $X = (X_i)_{i\in I}$ have a discrete spectral measure, while their density exists and is positive and continuous. A more subtle example is, for instance, the asymmetric logistic model (Tawn 1990), which admits a continuous positive density and whose spectral measure carries mass on all faces of S_+, cf. also Example 3.

(b) Secondly, both random vectors $(X_i)_{i=1,2,5}$ and $(X_i)_{i=1,4,5}$ that were considered in the Gaussian case in Example 1 have a positive continuous density on $\mathbb{R}^d$. Hence, there exists no version for Theorem 1 for the Gaussian case.

(c) Note that the implication of Theorem 1 is the independence of X_A and X_B, not the independence of all three subvectors $X_A, X_B, X_{I\setminus(A\cup B)}$.

(d) By means of the same argument that shows that pairwise independence of the components of a max-stable random vector implies already their joint independence, we may deduce a version of Theorem 1, in which more than two subvectors are considered.

Corollary 2. *Let $X = (X_i)_{i\in I}$ be a simple max-stable random vector with positive continuous density. Then, for any disjoint non-empty subsets $A_1, \ldots, A_k$ of I, the conditional independence $\perp\!\!\!\perp_{i=1}^k X_{A_i} \mid X_{I\setminus\bigcup_{i=1}^k A_i}$ implies the independence $\perp\!\!\!\perp_{i=1}^k X_{A_i}$.*

(e) The non-degenerate univariate max-stable laws are classified up to location and scale by the one parameter family of extreme value distributions indexed by $\gamma \in \mathbb{R}$

$$F_\gamma(x) = \exp(-(1+\gamma x)^{-1/\gamma}), \qquad x \in \begin{cases} (-1/\gamma, \infty) & \gamma > 0, \\ \mathbb{R} & \gamma = 0, \\ (-\infty, -1/\gamma) & \gamma < 0. \end{cases}$$

Any other (not necessarily simple) max-stable random vector is obtained through a transformation of the marginals that is differentiable and strictly monotone on the respective sub-domain on $\mathbb{R}^d$ (cf. e.g. Resnick (2008) Prop. 5.10). Hence, the above results remain valid for the general class of max-stable random vectors.

(f) Dombry & Eyi-Minko (2014) show that, up to time reversal, only max-autoregressive processes of order one can appear as discrete time stationary max-stable processes that satisfy the first order Markov property. This result indicates already that the conditional independence assumption is to some extent unnatural in presence of the max-stability property.

Example 3. Various classes of tractable max-stable distributions admit a positive continuous density, such that Theorem 1 and Corollary 2 apply. Popular models that are commonly used for statistical inference include the asymmetric logistic model (Tawn 1990), the asymmetric Dirichlet model (Coles & Tawn 1991), the pairwise beta model

(Cooley et al. 2010) and its generalizations involving continuous spectral densities (Ballani & Schlather 2011) in the multivariate case. Moreover, most marginal distributions of spatial models such as the Gaussian max-stable model (Genton et al. 2011, Smith 1990) or the Brown-Resnick model (Hüsler & Reiss 1989, Kabluchko et al. 2009) possess a positive continuous density if the parameters are non-degenerate. Hence, if any of the components of the previously mentioned extreme value models exhibit conditional independence given any of the remaining components, they must be independent.

In the remaining article we subsume auxiliary arguments in Section 2 and give all proofs in Section 3.

2 Preparatory results on max-stable random vectors

Throughout this section let G be the distribution function of a simple max-stable random vector $X = (X_i)_{i\in I}$ that has a positive continuous density. We denote its exponent function by

$$V(x) := -\log G(x) = \int_{S_+} \max_{i\in I}\left(\frac{\omega_i}{x_i}\right) H(\mathrm{d}\omega), \qquad x \in (0,\infty)^I.$$

Lower order marginals G^A that refer to a subset A of I are obtained as $\min_{i\in A^c}(x_i) \to \infty$, where $A^c = I \setminus A$. We write $x_{A^c} \to \infty$ for $\min_{i\in A^c}(x_i) \to \infty$, and with this notation

$$G^A(x_A) := \lim_{x_{A^c}\to\infty} G(x) \quad \text{and} \quad V^A(x_A) := -\log G^A(x_A).$$

Since G is absolutely continuous, the partial derivatives

$$G^A_B(x_A) := \frac{\partial^{|B|}}{\partial x_B} G^A(x_A) \quad \text{and} \quad V^A_B(x_A) := \frac{\partial^{|B|}}{\partial x_B} V^A(x_A)$$

exist and are continuous for $B \subset A$, and the latter V^A_B are homogeneous of order $-(|B|+1)$ (Coles & Tawn 1991). An elementary computation shows that

$$G^A_B(x_A) = W^A_B(x_A) \exp\left(-V^A(x_A)\right),$$

where

$$W^N_M(x_M) = \sum_{\pi\in\Pi(M)} (-1)^{|\pi|} \prod_{J\in\pi} V^N_J(x_N),$$

and $\Pi(M)$ stands for the set of partitions of M for $M \subset N \subset I$.

Let us further denote the set of non-empty subsets of I by $\mathcal{C}(I)$. The collection of exponent functions $(V^A)_{A\in\mathcal{C}(I)}$ is in a one-to-one correspondence with its Möbius inversion $(d_A)_{A\in\mathcal{C}(I)}$, i.e., if we set

$$d_A(x) := \sum_{B\in\mathcal{C}(I):A^c\subset B} (-1)^{|B\cap A|+1} V^B(x_B),$$

it follows that V^A can be recovered via

$$V^A(x_A) = \sum_{B \in \mathcal{C}(I): B \cap A \neq \emptyset} d_B(x) \tag{3}$$

(cf. Papastathopoulos & Tawn (2014), Theorem 2 and Schlather & Tawn (2002), Theorem 4 or, more generally, Berge (1971) Chapter 3, Section 2 for the Möbius inversion). Finally, we define

$$\chi_A(x_A) := \lim_{x_{A^c} \to \infty} d_A(x) = \sum_{B \in \mathcal{C}(I): B \subset A} (-1)^{|B|+1} V^B(x_B) = \sum_{B \in \mathcal{C}(I): A \subset B} d_B(x).$$

Then the collection of functions $(\chi_A)_{A \in \mathcal{C}(I)}$ is also in a one-to-one correspondence with $(V^A)_{A \in \mathcal{C}(I)}$ as well as $(d_A)_{A \in \mathcal{C}(I)}$ and the inversions are given by

$$d_A(x) = \sum_{B \in \mathcal{C}(I): A \subset B} (-1)^{|B \setminus A|} \chi_B(x_B),$$
$$V^A(x_A) = \sum_{B \in \mathcal{C}(I): B \subset A} (-1)^{|B|+1} \chi_B(x_B).$$

Further expressions for V^A, d_A and χ_A are collected in Lemma 3. Note that $\chi_A(x_A) \geq d_A(x)$ and thus,

$$d_A = 0 \quad \Leftrightarrow \quad \chi_A = 0. \tag{4}$$

Lemma 3. *The functions V^A and d_A and χ_A (with $A \in \mathcal{C}(I)$) can be expressed in terms of the spectral measure H as follows:*

$$V^A(x_A) = \int_{S_+} \max_{i \in A} \left(\frac{\omega_i}{x_i} \right) H(\mathrm{d}\omega),$$
$$d_A(x) = \int_{S+} \left[\min_{i \in A} \left(\frac{\omega_i}{x_i} \right) - \max_{j \in A^c} \left(\frac{\omega_j}{x_j} \right) \right]_+ H(\mathrm{d}\omega),$$
$$\chi_A(x_A) = \int_{S+} \min_{i \in A} \left(\frac{\omega_i}{x_i} \right) H(\mathrm{d}\omega).$$

Here $z_+ = \max(0, z)$ *and* $\max(\emptyset) = 0$.

It turns out that the following two quantities are closely linked to conditional independence and independence of subvectors of X, respectively. For non-empty disjoint subsets A, B of I and $C = I \setminus (A \cup B)$, we set for $x \in (0, \infty)^I$

$$\begin{aligned} d_{A,B}(x) &:= V^{A \cup C}(x_{A \cup C}) + V^{B \cup C}(x_{B \cup C}) - V(x) - V^C(x_C) \\ &= \sum_{L \in \mathcal{C}(I): L \cap A \neq \emptyset, L \cap B \neq \emptyset, L \cap C = \emptyset} d_L(x), \\ \chi_{A,B}(x_{A \cup B}) &:= \lim_{x_C \to \infty} d_{A,B}(x) = V^A(x_A) + V^B(x_B) - V^{A \cup B}(x_{A \cup B}) \\ &= \sum_{L \in \mathcal{C}(I): L \cap A \neq \emptyset, L \cap B \neq \emptyset} d_L(x), \end{aligned}$$

where the last equalities follow from (3). The following lemma is an analogue of Lemma 3.

Lemma 4. *The functions $d_{A,B}$ and $\chi_{A,B}$ can be expressed in terms of the spectral measure H as follows*

$$d_{A,B}(x) = \int_{S^+} \left[\min\left(\max_{i \in A} \left(\frac{\omega_i}{x_i}\right), \max_{i \in B} \left(\frac{\omega_i}{x_i}\right) \right) - \max_{j \in C} \left(\frac{\omega_j}{x_j}\right) \right]_+ H(\mathrm{d}\omega),$$

$$\chi_{A,B}(x_{A \cup B}) = \int_{S^+} \min\left(\max_{i \in A} \left(\frac{\omega_i}{x_i}\right), \max_{i \in B} \left(\frac{\omega_i}{x_i}\right) \right) H(\mathrm{d}\omega).$$

Note that $\chi_{A,B}(x_{A \cup B}) \geq d_{A,B}(x)$ implies similarly to (4) that

$$d_{A,B} = 0 \quad \Leftrightarrow \quad \chi_{A,B} = 0. \tag{5}$$

General expressions for the regular conditional distributions for the distribution of a max-stable process conditioned on a finite number of sites that are based on hitting scenarios of Poisson point process representations have been computed in Dombry & Eyi-Minko (2013), Oesting (2015), Oesting & Schlather (2014) under mild regularity assumptions or in Wang & Stoev (2011) for spectrally discrete max-stable random vectors.

Let again A and B be non-empty disjoint subsets of I. Since we assumed a positive continuous density for G (and hence also for its marginals), the numerators and denominators in

$$G(x_A | x_B) := \frac{G_B^{A \cup B}(x_{A \cup B})}{G_B^B(x_B)} = \exp\left(-\left[V^{A \cup B}(x_{A \cup B}) - V^B(x_B)\right]\right) \frac{W_B^{A \cup B}(x_{A \cup B})}{W_B^B(x_B)}$$

are non-zero and continuous for $x \in (0, \infty)^I$ and the expression $G(x_A|x_B)$ constitutes a regular version of the conditional probability $\mathbb{P}(X_A \leq x_A | X_B = x_B)$.

Proposition 5. *The functions $\chi_{A,B}$ and $d_{A,B}$ are connected with the independence and conditional independence of the respective subvectors of X as follows.*

a) $X_A \perp\!\!\!\perp X_B \mid X_{I \setminus (A \cup B)} \quad \Rightarrow \quad d_{A,B} = 0.$

b) $X_A \perp\!\!\!\perp X_B \quad \Leftrightarrow \quad \chi_{A,B} = 0.$

Remark. The assumption that G admits a positive continuous density on $(0, \infty)^I$ is crucial for part a) to hold true. It fails in Example 2.

Moreover, it is a simple consequence of Berman (1961/1962) and de Haan (1978) that the pairwise independence of any disjoint subvectors of the simple max-stable random vector X implies already their joint independence.

Lemma 6. *If $X_{A_1}, \ldots, X_{A_k}$ are pairwise independent subvectors of a simple max-stable random vector X (for necessarily disjoint $A_i \subset I$), then they are jointly independent.*

3 Proofs

Proof of Lemma 3. The first equation is clear from the definition of V^A. The relation for d_A can be obtained as follows.

$$\begin{aligned} d_A(x) &= \sum_{B\in\mathcal{C}(I):A^c\subset B} (-1)^{|B\cap A|+1} V^B(x_B) \\ &= \int_{S_+} \sum_{B\in\mathcal{C}(I):A^c\subset B} (-1)^{|B\cap A|+1} \max_{i\in B}\left(\frac{\omega_i}{x_i}\right) H(\mathrm{d}\omega) \\ &= \int_{S+} \left[\min_{i\in A}\left(\frac{\omega_i}{x_i}\right) - \max_{j\in A^c}\left(\frac{\omega_j}{x_j}\right)\right]_+ H(\mathrm{d}\omega). \end{aligned}$$

In order to obtain the last equality, we denote $a_i = \omega_i/x_i$ and distinguish two cases:
1st case: $A = I$. Then

$$\sum_{B\in\mathcal{C}(I):A^c\subset B} (-1)^{|B\cap A|+1} \max_{i\in B}(a_i) = \sum_{B\in\mathcal{C}(I)} (-1)^{|B|+1} \max_{i\in B}(a_i) = \min_{i\in I}(a_i).$$

2nd case: $A \neq I$. Then set $b := \max_{i\in A^c} a_i$ and $c_i := \max(a_i, b)$, such that

$$\begin{aligned} &\sum_{B\in\mathcal{C}(I):A^c\subset B} (-1)^{|B\cap A|+1} \max_{i\in B}(a_i) = \sum_{B\in\mathcal{C}(I):A^c\subset B, B\neq A^c} (-1)^{|B\cap A|+1} \max_{i\in B\cap A}(c_i) - b \\ &= \sum_{U\subset A:U\neq\emptyset} (-1)^{|U|+1} \max_{i\in U}(c_i) - b = \min_{i\in A}(c_i) - b \\ &= \min_{i\in A}(\max(a_i, b)) - b = \max\left(\min_{i\in A}(a_i), b\right) - b = \left(\min_{i\in A}(a_i) - b\right)_+. \end{aligned}$$

The expression for χ_A follows immediately. □

Proof of Lemma 4. Similar to the proof of Lemma 3, the relation for $d_{A,B}$ follows from

$$\begin{aligned} d_{A,B}(x) &= V^{A\cup C}(x_{A\cup C}) + V^{B\cup C}(x_{B\cup C}) - V(x) - V^C(x_C) \\ &= \int_{S_+} \max_{i\in A\cup C}\left(\frac{\omega_i}{x_i}\right) + \max_{i\in B\cup C}\left(\frac{\omega_i}{x_i}\right) - \max_{i\in I}\left(\frac{\omega_i}{x_i}\right) - \max_{i\in C}\left(\frac{\omega_i}{x_i}\right) H(\mathrm{d}\omega) \\ &= \int_{S+} \left[\min\left(\max_{i\in A}\left(\frac{\omega_i}{x_i}\right), \max_{i\in B}\left(\frac{\omega_i}{x_i}\right)\right) - \max_{j\in C}\left(\frac{\omega_j}{x_j}\right)\right]_+ H(\mathrm{d}\omega), \end{aligned}$$

where the last equality is obtained from

$$\begin{aligned} &\max_{i\in A\cup C}(a_i) + \max_{i\in B\cup C}(a_i) - \max_{i\in I}(a_i) - \max_{i\in C}(a_i) \\ &= \min\left(\max_{i\in A\cup C}(a_i), \max_{i\in B\cup C}(a_i)\right) - \max_{i\in C}(a_i) \\ &= \max\left[\min\left(\max_{i\in A}(a_i), \max_{i\in B}(a_i)\right), \max_{i\in C}(a_i)\right] - \max_{i\in C}(a_i) \\ &= \left[\min\left(\max_{i\in A}(a_i), \max_{i\in B}(a_i)\right) - \max_{i\in C}(a_i)\right]_+ \end{aligned}$$

if we denote $a_i = \omega_i/x_i$. The expression for $\chi_{A,B}$ follows immediately. □

Proof of Proposition 5. a) As before, let $C = I \setminus (A \cup B)$. Since $G(x) = \exp(-V(x))$ has a positive continuous density, we have that the conditional independence $X_A \perp\!\!\!\perp X_B \mid X_C$ for $C = I \setminus (A \cup B)$ implies that for all $x \in (0, \infty)^I$

$$G(x_A | X_C)\, G(x_B | X_C) = G(x_{A\cup B} | X_C) \qquad \mathbb{P}\text{-a.s.}\,.$$

Since X_C has a positive continuous density with respect to the Lebesgue-measure on $(0, \infty)^C$, it follows that

$$G(x_A | x_C)\, G(x_B | x_C) = G(x_{A\cup B} | x_C) \qquad \text{for all } x \in Q,$$

where Q is a dense subset of $(0, \infty)^I$. By the continuity of these expressions in $x \in (0, \infty)^I$, the equality holds for all $x \in (0, \infty)^I$ and is equivalent to

$$\exp\left(d_{A,B}(x)\right) = \frac{W_C^{A\cup C}(x_{A\cup C}) W_C^{B\cup C}(x_{B\cup C})}{W_C^{A\cup B\cup C}(x_{A\cup B\cup C}) W_C^{C}(x_C)}, \qquad x \in (0, \infty)^I. \tag{6}$$

Here, $d_{A,B} \geq 0$ and $d_{A,B}$ is homogeneous of order -1, while the components V_J^N that build the terms W_M^N are homogeneous of order $-(|J| + 1)$. Now, replacing x by $t^{-1}x$ for $t > 0$ in (6), we see that the left-hand side grows exponentially in the variable t as t tends to ∞ if $d_{A,B}(x) > 0$, while the right-hand side exhibits at most polynomial growth. Therefore, $d_{A,B}(x) = 0$ for $x \in (0, \infty)^I$.

b) Both sides are equivalent to $G^A(x_A) G^B(x_B) = G^{A\cup B}(x_{A\cup B})$ for all $x \in (0, \infty)^I$. □

Proof of Theorem 1. The hypothesis follows from Proposition 5 and (5). □

Proof of Lemma 6. It suffices to show that for $x_{A_i} \in (0, \infty)^{A_i}$, $i = 1, \ldots, k$ and $r \in (0, \infty)$

$$\mathbb{P}\left(X_{A_1} \leq x_{A_1}, \ldots, X_{A_k} \leq x_{A_k}\right) = \prod_{i=1}^{k} \mathbb{P}\left(X_{A_i} \leq x_{A_i}\right).$$

Using the notation $r_i = \sum_{j_i \in A_i} x_{j_i}^{-1}$, $u_{j_i} = (r_i x_{j_i})^{-1}$ for $j_i \in A_i$ and $Y_i = \max_{j_i \in A_i} u_{j_i} X_{j_i}$, $i = 1, \ldots, k$, we can rewrite this equality in the form

$$\mathbb{P}\left(Y_1 \leq r_1^{-1}, \ldots, Y_k \leq r_k^{-1}\right) = \prod_{i=1}^{k} \mathbb{P}\left(Y_i \leq r_i^{-1}\right),$$

where the random vector $(Y_1, \ldots, Y_k)$ is simple max-stable (de Haan 1978) and has pairwise independent components due to our assumptions. Hence, by Berman (1961/1962) Theorem 2, the Y_i are jointly independent, which entails the relation above. □

Proof of Corollary 2. $\perp\!\!\!\perp_{i=1}^k X_{A_i} \mid X_{I\setminus \bigcup_{i=1}^k A_i}$ implies $X_{A_{i_1}} \perp\!\!\!\perp X_{A_{i_2}} \mid X_{I\setminus \bigcup_{i=1}^k A_i}$ for $i_1 \neq i_2$ and hence $X_{A_{i_1}} \perp\!\!\!\perp X_{A_{i_2}}$ by Theorem 1. The hypothesis follows if we apply Lemma 6 to the X_{A_i}, $i = 1, \ldots, k$. □

Acknowledgments The authors would like to thank an anonymous referee for helpful comments improving the clarity of the paper. IP acknowledges funding from the SuSTaIn program - Engineering and Physical Sciences Research Council grant EP/D063485/1 - at the School of Mathematics of the University of Bristol. Part of this paper was written during a visit of KS at the University of Bristol funded by the SuSTaIn program and during a visit of IP at the University of Mannheim funded by the German Research Foundation DFG through the RTG 1953. IP and KS thank them for their generous hospitality.

References

Ballani, F. & Schlather, M. (2011), 'A construction principle for multivariate extreme value distributions', *Biometrika* **98**(3), 633–645. 4

Beirlant, J., Goegebeur, Y., Teugels, J. & Segers, J. (2004), *Statistics of Extremes*, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Chichester. 3

Berge, C. (1971), *Principles of combinatorics*, Translated from the French. Mathematics in Science and Engineering, Vol. 72, Academic Press, New York-London. 5

Berman, S. M. (1961/1962), 'Convergence to bivariate limiting extreme value distributions', *Ann. Inst. Statist. Math.* **13**, 217–223. 6, 8

Blanchet, J. & Davison, A. C. (2011), 'Spatial modeling of extreme snow depth', *Ann. Appl. Stat.* **5**(3), 1699–1725. 2

Buishand, T. A., de Haan, L. & Zhou, C. (2008), 'On spatial extremes: with application to a rainfall problem', *Ann. Appl. Stat.* **2**(2), 624–642. 2

Coles, S. G. & Tawn, J. A. (1991), 'Modelling extreme multivariate events', *J. Roy. Statist. Soc. Ser. B* **53**(2), 377–392. 3, 4

Cooley, D., Davis, R. A. & Naveau, P. (2010), 'The pairwise beta distribution: a flexible parametric multivariate model for extremes', *J. Multivariate Anal.* **101**(9), 2103–2117. 4

Dawid, A. P. (1979), 'Conditional independence in statistical theory', *J. Roy. Statist. Soc. Ser. B* **41**(1), 1–31. 1

de Haan, L. (1978), 'A characterization of multidimensional extreme-value distributions', *Sankhyā Ser. A* **40**(1), 85–88. 6, 8

Dombry, C. & Eyi-Minko, F. (2013), 'Regular conditional distributions of continuous max-infinitely divisible random fields', *Electron. J. Probab.* **18**, no. 7, 21. 6

Dombry, C. & Eyi-Minko, F. (2014), 'Stationary max-stable processes with the Markov property', *Stochastic Process. Appl.* **124**(6), 2266–2279. 3

Engelke, S., Malinowski, A., Oesting, M. & Schlather, M. (2014), 'Statistical inference for max-stable processes by conditioning on extreme events', *Adv. in Appl. Probab.* **46**(2), 478–495. 2

Genton, M. G., Ma, Y. & Sang, H. (2011), 'On the likelihood function of Gaussian max-stable processes', *Biometrika* **98**(2), 481–488. 4

Hüsler, J. & Reiss, R.-D. (1989), 'Maxima of normal random vectors: between independence and complete dependence', *Statist. Probab. Lett.* **7**(4), 283–286. 4

Kabluchko, Z., Schlather, M. & de Haan, L. (2009), 'Stationary max-stable fields associated to negative definite functions', *Ann. Probab.* **37**(5), 2042–2065. 4

Lauritzen, S. L. (1996), *Graphical models*, Vol. 17 of *Oxford Statistical Science Series*, The Clarendon Press, Oxford University Press, New York. Oxford Science Publications. 1

Naveau, P., Guillou, A., Cooley, D. & Diebolt, J. (2009), 'Modelling pairwise dependence of maxima in space', *Biometrika* **96**(1), 1–17. 2

Oesting, M. (2015), 'On the distribution of a max-stable process conditional on max-linear functionals', *Statist. Probab. Lett.* **100**, 158–163. 6

Oesting, M. & Schlather, M. (2014), 'Conditional sampling for max-stable processes with a mixed moving maxima representation', *Extremes* **17**(1), 157–192. 6

Papastathopoulos, I. & Tawn, J. A. (2014), 'Dependence properties of multivariate max-stable distributions', *J. Multivariate Anal.* **130**, 134–140. 5

Resnick, S. I. (2008), *Extreme values, regular variation and point processes*, Springer Series in Operations Research and Financial Engineering, Springer, New York. Reprint of the 1987 original. 2, 3

Rue, H. & Held, L. (2005), *Gaussian Markov Random Fields*, Vol. 104 of *Monographs on Statistics and Applied Probability*, Chapman & Hall/CRC, Boca Raton, FL. Theory and applications. 1

Schlather, M. & Tawn, J. (2002), 'Inequalities for the extremal coefficients of multivariate extreme value distributions', *Extremes* **5**(1), 87–102. 5

Smith, R. L. (1990), Max-stable processes and spatial extremes, Technical report, University of North Carolina. 4

Tawn, J. A. (1990), 'Modelling multivariate extreme value distributions', *Biometrika* **77**, 245–53. 3

Wang, Y. & Stoev, S. (2011), 'Conditional sampling for spectrally discrete max-stable random fields', *Adv. in Appl. Probab.* **43**(2), 461–483. 6