Maximum entropy distribution of order statistics with given marginals

CRISTINA BUTUCEA[1], JEAN-FRANÇOIS DELMAS[2,*],
ANNE DUTFOY[3] and RICHARD FISCHER[1,2,3,**]

[1]*LAMA (UPE-MLV), UPE, Marne La Vallée, France. E-mail: cristina.butucea@univ-mlv.fr*
[2]*CERMICS, École des Ponts, UPE, Champs-sur-Marne, France.*
E-mail: *delmas@cermics.enpc.fr; **fischerr@cermics.enpc.fr
[3]*EDF Research & Development, Industrial Risk Management Department, Palaiseau, France.*
E-mail: anne.dutfoy@edf.fr

We consider distributions of ordered random vectors with given one-dimensional marginal distributions. We give an elementary necessary and sufficient condition for the existence of such a distribution with finite entropy. In this case, we give explicitly the density of the unique distribution which achieves the maximal entropy and compute the value of its entropy. This density is the unique one which has a product form on its support and the given one-dimensional marginals. The proof relies on the study of copulas with given one-dimensional marginal distributions for its order statistics.

Keywords: copula; entropy; maximum entropy; order statistics

1. Introduction

Order statistics, an almost surely non-decreasing sequence of random variables, have received a lot of attention due to the diversity of possible applications. If $X = (X_1, \ldots, X_d)$ is a d-dimensional random vector, then its order statistics $X^{\mathrm{OS}} = (X_{(1)}, \ldots, X_{(d)})$ corresponds to the permutation of the components of X in the non-decreasing order, so that $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(d)}$. The components of the underlying random vector X are usually, but not necessarily, independent and identically distributed (i.i.d.). Special attention has been given to extreme values $X_{(1)}$ and $X_{(d)}$, the range $X_{(d)} - X_{(1)}$, or the median value. Direct application of the distribution of the kth largest order statistic occurs in various fields, such as climatology, extreme events, reliability, insurance, financial mathematics. We refer to the monographs of David and Nagaraja [8] and Arnold, Balakrishnan and Nagaraja [1] for a general overview on the subject of order statistics. We are interested in the dependence structure of order statistics, which has received great attention. In the i.i.d. case, Bickel [3] showed that any two order statistics are positively correlated. The copula of the joint distribution of $X_{(1)}$ and $X_{(d)}$ is derived in Schmitz [19] with exact formulas for Kendall's τ and Spearman's ρ. In Avérous, Genest and Kochar [2], it is shown that the dependence of the jth order statistic on the ith order statistic decreases as the distance between i and j increases according to the bivariate monotone regression dependence ordering. The copula connecting the limit distribution of the two largest order statistics, called bi-extremal copula, is given by de Melo Mendes and Sanfins [9] with some additional properties. Exact expressions for Pearson's correlation coefficient, Kendall's τ and Spearman's ρ for any

two order statistics are obtained in Navarro and Balakrishnan [16]. For the non i.i.d. case, Kim and David [14] show that some pairs of order statistics can be negatively correlated, if the underlying random vector is sufficiently negatively dependent. Positive dependence measures for two order statistics are considered in Boland *et al.* [4] when the underlying random variables are independent but arbitrarily distributed or when they are identically distributed but not independent. A generalization of these results for multivariate dependence properties is given by Hu and Chen [11]. See also Dubhashi and Häggström [10] for conditional distribution of order statistics.

Here, we focus on the cumulative distribution function (c.d.f.) of order statistics without referring to an underlying distribution. That is, we consider random vectors $X = (X_1, \ldots, X_d) \in \mathbb{R}^d$ such that a.s. $X_1 \leq \cdots \leq X_d$ and we suppose that the one-dimensional marginal distributions $\mathbf{F} = (\mathbf{F}_i, 1 \leq i \leq d)$ are given, where $\mathbf{F}_i$ is the c.d.f. of X_i. A necessary and sufficient condition for the existence of a joint distribution of order statistics with one-dimensional marginals $\mathbf{F}$ is that they are stochastically ordered, that is $\mathbf{F}_{i-1}(x) \geq \mathbf{F}_i(x)$ for all $2 \leq i \leq d, x \in \mathbb{R}$. With the marginals fixed, the joint distribution of the order statistics can be characterized by the connecting copula of the random vector, which contains all information on the dependence structure of the order statistics. Copulas of order statistics derived from an underlying i.i.d. sample were considered in [2] in order to calculate measures of concordance between any two pairs of order statistics. For order statistics derived from a general parent distribution, Navarro and Spizzichino [17] shows that the copula of the order statistics depends on the marginals and the copula of the parent distribution through an exchangeable copula and the average of the marginals. Construction of some copula of order statistics with given marginals were given in Lebrun and Dutfoy [15].

Our aim is to find the c.d.f. of order statistics of dimension d with fixed marginals which maximizes the relative entropy H_h defined by (13). In an information-theoretic interpretation, the maximum entropy distribution is the least informative among order statistics with given marginals. This problem appears in models where the one-dimensional marginals are well known (either from different experimentation or from physical models) but the dependence structure is unknown, see Butucea *et al.* [7]. In [6], the same authors gave, when it exists, the maximum entropy distribution of $(X_1, \ldots, X_d)$ such that X_i is uniformly distributed on [0, 1] for $1 \leq i \leq d$ and the distribution of $X_{(d)} = \max_{1 \leq i \leq d} X_i$ is given, see Remark 4.8.

For a d-dimensional random variable $X = (X_1, \ldots, X_d)$ with c.d.f. F and copula C_F, the relative entropy of F can be decomposed into the sum of the relative entropy (with respect to a one-dimensional probability density h) of its one-dimensional marginals plus the entropy of C_F, see Lemma 2.1. In our case, since the marginals $\mathbf{F} = (\mathbf{F}_i, 1 \leq i \leq d)$ are fixed, maximizing the entropy of the joint distribution F of an order statistics is equivalent to maximizing the entropy of its copula C_F under constraints, (see Section 2.4). Therefore, we shall find the maximum entropy copula for order statistics with fixed marginal distributions.

The main result of this paper is given by Theorem 5.4. It states that there exists a unique maximum entropy c.d.f. $F_{\mathbf{F}}$ given by (52) if and only if:

$$\sum_{i=1}^{d} H_h(\mathbf{F}_i) - \sum_{i=2}^{d} \int_{\mathbb{R}} \mathbf{F}_i(dt) \big| \log\big(\mathbf{F}_{i-1}(t) - \mathbf{F}_i(t)\big) \big| > -\infty.$$

In this case, $F_{\mathbf{F}}$ is absolutely continuous with density $f_{\mathbf{F}}$ defined as, for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$:

$$f_{\mathbf{F}}(x) = \mathbf{f}_1(x_1) \prod_{i=2}^{d} \frac{\mathbf{f}_i(x_i)}{\mathbf{F}_{i-1}(x_i) - \mathbf{F}_i(x_i)} \exp\left(-\int_{x_{i-1}}^{x_i} \frac{\mathbf{f}_i(s)}{\mathbf{F}_{i-1}(s) - \mathbf{F}_i(s)} \, ds\right) \mathbf{1}_{L^{\mathbf{F}}}(x),$$

where $\mathbf{f}_i$ is the density function of $\mathbf{F}_i$ and $L^{\mathbf{F}} \subset \mathbb{R}^d$ is the set of ordered vectors $(x_1, \ldots, x_d)$, that is $x_1 \leq \cdots \leq x_d$, such that $\mathbf{F}_{i-1}(t) > \mathbf{F}_i(t)$ for all $t \in (x_{i-1}, x_i)$ and $2 \leq i \leq d$. See Example 5.8 for an illustrative example.

The rest of the paper is organized as follows. In Section 2, we introduce the basic notations and give the definition of the objects used in later parts. Section 3 describes the connection between copulas of order statistics with fixed marginals, and symmetric copulas with fixed multidiagonals. The multidiagonal, given by Definition 3.6, is the generalization of the diagonal section for copulas, which received great attention in copula literature. We show that there exists a one-to-one map between these two sets of copulas, see Corollary 3.14. This bijection has good properties with respect to the entropy as explained in Proposition 3.22. In Section 4, we determine the maximum entropy copula with fixed multidiagonal, see Theorem 4.7. Since we obtain a symmetric copula as a result, this is also the maximum entropy symmetric copula with fixed multidiagonal. In Section 5, we use the one-to-one map between the two sets of copulas established in Section 3 to give the maximum entropy copula of order statistics with fixed marginals. We finally obtain the density of the maximum entropy distribution for order statistics with fixed marginals by composing the maximum entropy copula with the marginals, see Theorem 5.4. Section 6 contains the detailed proofs of Theorem 4.7 and other results from Section 4. Section 7 collects the main notations of the paper to facilitate reading.

2. Notations and definitions

2.1. Notations in $\mathbb{R}^d$ and generalized inverse

For a Borel set $A \subset \mathbb{R}^d$, we write $|A|$ for its Lebesgue measure. For $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $y = (y_1, \ldots, y_d) \in \mathbb{R}^d$, we write $x \leq y$ if $x_i \leq y_i$ for all $1 \leq i \leq d$. We define $\min x = \min\{x_i, 1 \leq i \leq d\}$ and $\max x = \max\{x_i, 1 \leq i \leq d\}$ for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$. If J is a real-valued function defined on $\mathbb{R}$, we set $J(x) = (J(x_1), \ldots, J(x_d))$. We shall consider the following subsets of $\mathbb{R}^d$:

$$S = \{(x_1, \ldots, x_d) \in \mathbb{R}^d, x_1 \leq \cdots \leq x_d\} \quad \text{and} \quad \triangle = S \cap I^d,$$

with $I = [0, 1]$. In what follows, usually x, y will belong to $\mathbb{R}^d$, and s, t to $\mathbb{R}$ or I. For a set $A \subset \mathbb{R}$, we note by $A^c = \mathbb{R} \setminus A$ its complementary set.

If J is a bounded non-decreasing càd-làg function defined on $\mathbb{R}$. Its generalized inverse J^{-1} is given by $J^{-1}(t) = \inf\{s \in \mathbb{R}; J(s) \geq t\}$, for $t \in \mathbb{R}$, with the convention that $\inf \varnothing = +\infty$ and $\inf \mathbb{R} = -\infty$. We have for $s, t \in \mathbb{R}$:

$$J(t) \geq s \Leftrightarrow t \geq J^{-1}(s), \qquad J^{-1} \circ J(t) \leq t \quad \text{and} \quad J \circ J^{-1} \circ J(t) = J(t). \tag{1}$$

We define the set of points where J is increasing on their left:

$$I_g(J) = \big\{ t \in \mathbb{R}; u < t \Leftrightarrow J(u) < J(t) \big\}. \tag{2}$$

We have:

$$\mathbf{1}_{(I_g(J))^c} \, dJ = 0 \qquad \text{a.e.,} \tag{3}$$

$$J^{-1}(\mathbb{R}) \subset I_g(J) \cup \{\pm\infty\} \tag{4}$$

and for $s \in \mathbb{R}$, $t \in I_g(J)$:

$$J(t) \le s \Leftrightarrow t \le J^{-1}(s) \quad \text{and} \quad J^{-1} \circ J(t) = t. \tag{5}$$

Notice that if J is continuous in addition, then we have for $t \in J(\mathbb{R})$:

$$J \circ J^{-1}(t) = t. \tag{6}$$

2.2. C.d.f. and copula

Let $X = (X_1, \ldots, X_d)$ be a random vector on $\mathbb{R}^d$. Its cumulative distribution function (c.d.f.), denoted by F is defined by: $F(x) = \mathbb{P}(X \le x)$, $x \in \mathbb{R}^d$. The corresponding one-dimensional marginal c.d.f.s are $(F_i, 1 \le i \le d)$ with $F_i(t) = \mathbb{P}(X_i \le t)$, $t \in \mathbb{R}$. The c.d.f. F is called a copula if X_i is uniform on $I = [0, 1]$ for all $1 \le i \le d$. (Notice a copula is characterized by its values on I^d only.)

We define $\mathcal{L}_d$ as the set of c.d.f.s on $\mathbb{R}^d$ whose one-dimensional marginal c.d.f.s are continuous, and $\mathcal{C} \subset \mathcal{L}_d$ as the subset of copulas. We set $\mathcal{L}_d^0$ (resp. $\mathcal{C}^0$) the subset of absolutely continuous c.d.f. (resp. copulas) on $\mathbb{R}^d$.

Let us define for a c.d.f. F with one-dimensional marginals $(F_i, 1 \le i \le d)$ the function C_F defined on I^d:

$$C_F(y) = F\big(F_1^{-1}(y_1), \ldots, F_d^{-1}(y_d)\big), \qquad y = (y_1, \ldots, y_d) \in I^d. \tag{7}$$

If $F \in \mathcal{L}_d$, then C_F defined by (7) is a copula thanks to (6). According to Sklar's theorem, F is then completely characterized by its one-dimensional marginal c.d.f.s $(F_i, 1 \le i \le d)$ and the associated copula C_F which contains all information on the dependence:

$$F(x) = C_F\big(F_1(x_1), \ldots, F_d(x_d)\big), \qquad x = (x_1, \ldots, x_d) \in \mathbb{R}^d. \tag{8}$$

Equivalently, if $X = (X_1, \ldots, X_d)$ has c.d.f. F, then C_F is the c.d.f. of the random vector:

$$\big(F_1(X_1), \ldots, F_d(X_d)\big). \tag{9}$$

2.3. Order statistics

For a d-dimensional c.d.f. F, we write $\mathbb{P}_F$ for the distribution of a random vector $X = (X_1, \ldots, X_d)$ with c.d.f. F. A d-dimensional c.d.f. F is a c.d.f. of order statistics (and we shall say that X is a vector of order statistics) if $\mathbb{P}_F(X_1 \leq X_2 \leq \cdots \leq X_d) = 1$. Let us denote by $\mathcal{L}_d^{OS} \subset \mathcal{L}_d$ the set of all c.d.f.s of order statistics with continuous one-dimensional marginal c.d.f.s. The d-tuples $(F_i, 1 \leq i \leq d)$ of marginal c.d.f.s then verify $F_{i-1} \geq F_i$ for all $2 \leq i \leq d$. Let $\mathcal{F}_d$ be the set of d-tuples of continuous one-dimensional c.d.f.s compatible with the marginal c.d.f.s of order statistics:

$$\mathcal{F}_d = \big\{ \mathbf{F} = (\mathbf{F}_i, 1 \leq i \leq d) \in (\mathcal{L}_1)^d; \mathbf{F}_{i-1} \geq \mathbf{F}_i, \forall 2 \leq i \leq d \big\}. \tag{10}$$

For a given $\mathbf{F} = (\mathbf{F}_i, 1 \leq i \leq d)$ in $\mathcal{F}_d$, we define the set of c.d.f.s F of order statistics with marginal c.d.f.s $\mathbf{F}$:

$$\mathcal{L}_d^{OS}(\mathbf{F}) = \big\{ F \in \mathcal{L}_d^{OS}; F_i = \mathbf{F}_i, 1 \leq i \leq d \big\}. \tag{11}$$

If $\mathbf{F} \in \mathcal{F}_d$, then we have $\mathcal{L}_d^{OS}(\mathbf{F}) \neq \varnothing$, since the c.d.f. of $(\mathbf{F}_1^{-1}(U), \ldots, \mathbf{F}_d^{-1}(U))$, U uniformly distributed on I, belongs to $\mathcal{L}_d^{OS}(\mathbf{F})$. We define $\mathcal{C}^{OS}(\mathbf{F})$ the set of copulas of order statistics with marginals $\mathbf{F}$:

$$\mathcal{C}^{OS}(\mathbf{F}) = \big\{ C_F \in \mathcal{C}; F \in \mathcal{L}_d^{OS}(\mathbf{F}) \big\}. \tag{12}$$

According to Sklar's theorem, the map $F \mapsto C_F$ is a bijection between $\mathcal{L}_d^{OS}(\mathbf{F})$ and $\mathcal{C}^{OS}(\mathbf{F})$ if $\mathbf{F} \in \mathcal{F}_d$.

2.4. Entropy

Let h be a reference probability density function on $\mathbb{R}$. We define $h^{\otimes d}(x) = \prod_{i=1}^d h(x_i)$ for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$. The relative Shannon-entropy for a c.d.f. $F \in \mathcal{L}_d$ is given by:

$$H_h(F) = \begin{cases} -\infty, & \text{if } F \in \mathcal{L}_d \setminus \mathcal{L}_d^0, \\ -\displaystyle\int_{\mathbb{R}^d} f \log\big(f/h^{\otimes d}\big), & \text{if } F \in \mathcal{L}_d^0, \end{cases} \tag{13}$$

with f the density of F. Notice that $H_h(F) \in [-\infty, 0]$ is well defined. We will use the notation $H_h(X) = H_h(F)$ if X is a random vector with c.d.f. F and $H_h(f) = H_h(F)$ if F has density f. We shall simply write $H(F)$ (resp. $H(X)$ and $H(f)$) instead of $H_h(F)$ (resp. $H_h(X)$ and $H_h(f)$) when $h = \mathbf{1}_{[0,1]}$. Note that $H(F)$ can be finite only if F is the c.d.f. of a probability distribution on $[0, 1]^d$.

According to the next lemma, the relative entropy of any $F \in \mathcal{L}_d^{1c}$ can be decomposed into the relative entropy of the one-dimensional marginals c.d.f. $(F_i, 1 \leq i \leq d)$ and the entropy of the associated copula C_F.

Lemma 2.1. *Let $F \in \mathcal{L}_d^{1c}$. We have*:

$$H_h(F) = H(C_F) + \sum_{i=1}^{d} H_h(F_i). \tag{14}$$

Proof. It is left to the reader to check that F has a density, say f, if and only if F_i has a density, say f_i, for $1 \le i \le d$ and C_F has a density, say c_F. Furthermore, in this case, we have:

$$f(x) = c_F\big(F_1(x_1), \ldots, F_d(x_d)\big) \prod_{i=1}^{d} f_i(x_i) \quad \text{a.e. for } x = (x_1, \ldots, x_d) \in \mathbb{R}^d,$$

as well as, with the convention $0/0 = 0$,

$$c_F(u) = \frac{f(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d))}{\prod_{i=1}^{d} f_i(F_i^{-1}(u_i))} \quad \text{a.e. for } u = (u_1, \ldots, u_d) \in I^d.$$

On the one hand, if F does not have a density then we have $H_h(F) = -\infty$. Since F does not have a density, then one of the F_i or C_F does not have a density either, and then $H(C_F) + \sum_{i=1}^{d} H_h(F_i) = -\infty$. Thus (14) holds.

On the other hand, let us assume that F has a density, say f. Elementary computations give with $x = (x_1, \ldots, x_d)$ and $1 \le i \le d$:

$$H_h(F_i) = -\int_{\mathbb{R}} f_i(x_i) \log\big((f_i/h)(x_i)\big) dx_i = -\int_{\mathbb{R}^d} f(x) \log\big((f_i/h)(x_i)\big) dx.$$

We also have with $u = (u_1, \ldots, u_d)$ and $x = (x_1, \ldots, x_d)$:

$$H(C_F) = -\int_{[0,1]^d} c_F \log(c_F) = -\int \frac{f(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d))}{\prod_{i=1}^{d} f_i(F_i^{-1}(u_i))} \log\big(c_F(u)\big) du$$

$$= -\int f(x) \log\big(c_F(F_1(x_1), \ldots, F_d(x_d))\big) dx,$$

where, for the last equality, we used the change of variable $F_i(x_i) = u_i$ (for $x_i \in I_g(F_i)$) so that $F_i^{-1}(u_i) = F_i^{-1} \circ F_i(x_i) = x_i$ holds $f_i(x_i) dx_i$-a.e and that $f(x) dx = 0$ on $(\otimes_{i=1}^{d} I_g(F_i))^c$. Then use that $f(x) = c_F(F_1(x_1), \ldots, F_d(x_d)) \prod_{i=1}^{d} f_i(x_i)$ a.e. for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ to deduce that $H(C_F) + \sum_{i=1}^{d} H_h(F_i) = -\int f \log(f/h^{\otimes d}) = H_h(F)$. □

Remark 2.2. Notice that if F_i has density f_i for $1 \le i \le d$, then one can choose the reference probability density $h(t) = \frac{1}{d} \sum_{i=1}^{d} f_i(t)$ so that $H_h(F_i) \ge -\log(d)$. In this case, $H_h(F)$ is finite if and only if $H(C_F)$ is finite, and we have:

$$H(C_F) = H_h(F) - \sum_{i=1}^{d} H_h(F_i) = -\int f(x) \log\left(\frac{f(x)}{\prod_{i=1}^{d} f_i(x_i)}\right) dx.$$

Thus, $H(C_F)$ is the relative entropy of the c.d.f. F with respect to the probability distribution with c.d.f. $\otimes_{i=1}^{d} F_i$ of independent real valued random variables with the same one-dimensional marginal as the one with c.d.f. F. This emphasizes the fact that $H_h(F) - \sum_{i=1}^{d} H_h(F_i)$, when it is well defined, does not depend on h.

For $\mathbf{F} = (\mathbf{F}_i, 1 \leq i \leq d) \in \mathcal{F}_d$, we define $\mathbb{J}(\mathbf{F})$ taking values in $[0, +\infty]$ by:

$$\mathbb{J}(\mathbf{F}) = \sum_{i=2}^{d} \int_{\mathbb{R}} \mathbf{F}_i(dt) \big| \log\big(\mathbf{F}_{i-1}(t) - \mathbf{F}_i(t)\big)\big|. \tag{15}$$

Our aim is to find the c.d.f. $F^* \in \mathcal{L}_d^{OS}(\mathbf{F})$ which maximizes the entropy H_h. We shall see that this is possible if and only if $\mathbb{J}(\mathbf{F})$ is finite. From an information theory point of view, this is the distribution which is the least informative among distributions of order statistics with given one-dimensional marginal c.d.f.s $\mathbf{F}$. Since the vector of marginal distribution functions $\mathbf{F}$ is fixed, thanks to (14), we notice that $H_h(F)$ is maximal on $\mathcal{L}_d^{OS}(\mathbf{F})$ if and only if $H(C_F)$ is maximal on $\mathcal{C}^{OS}(\mathbf{F})$. Therefore, we focus on finding the copula $C^* \in \mathcal{C}^{OS}(\mathbf{F})$ which maximizes the entropy H. We will give the solution of this problem in Section 5 under some additional hypotheses on $\mathbf{F}$.

3. Symmetric copulas with given order statistics

In this section, we introduce an operator on the set $\mathcal{C}^{OS}(\mathbf{F})$ of copulas of order statistics with fixed marginal c.d.f.s $\mathbf{F}$. This operator assigns to a copula $C \in \mathcal{C}^{OS}(\mathbf{F})$ the copula of the exchangeable random vector associated to the order statistics with marginal c.d.f.s $\mathbf{F}$ and copula C. We show that this operator is a bijection between $\mathcal{C}^{OS}(\mathbf{F})$ and a set of symmetric copulas which can be characterized by their multidiagonal, which is a generalization of the well-known diagonal section of copulas. This bijection has good properties with respect to the entropy H, giving us a problem equivalent to maximizing H on $\mathcal{C}^{OS}(\mathbf{F})$. We shall solve this problem in Section 4.

3.1. Symmetric copulas

For $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ we define $x^{OS} = (x_{(1)}, \ldots, x_{(d)})$ the ordered vector (increasing order) of x, where $x_{(1)} \leq \cdots \leq x_{(d)}$ and $\sum_{i=1}^{d} \hat{\delta}_{x_i} = \sum_{i=1}^{d} \hat{\delta}_{x_{(i)}}$, with $\hat{\delta}_t$ the Dirac mass at $t \in \mathbb{R}$.

Let $\mathcal{S}_d$ be the set of permutations on $\{1, \ldots, d\}$. For $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $\pi \in \mathcal{S}_d$, we set $x_\pi = (x_{\pi(1)}, \ldots, x_{\pi(d)})$. A function h defined on $\mathbb{R}^d$ is symmetric if $h(x_\pi) = h(x)$ for all $\pi \in \mathcal{S}_d$. A random vector X taking values in $\mathbb{R}^d$ is exchangeable if X_π is distributed as X for all $\pi \in \mathcal{S}_d$. In particular a random vector X taking values in $\mathbb{R}^d$ is exchangeable if and only if its c.d.f. is symmetric. Let $\mathcal{L}_d^{sym} \subset \mathcal{L}_d$ (resp. $\mathcal{C}^{sym} \subset \mathcal{L}_d$) denote the set of symmetric c.d.f. (resp. copulas) on $\mathbb{R}^d$.

Let $F \in \mathcal{L}_d$ and define its symmetrization $F^{sym} \in \mathcal{L}_d^{sym}$ by:

$$F^{sym}(x) = \frac{1}{d!} \sum_{\pi \in \mathcal{S}_d} F(x_\pi), \qquad x \in \mathbb{R}^d. \tag{16}$$

In particular, if X is a random vector taking values in $\mathbb{R}^d$ with c.d.f. F and Π is a random variable independent of X, uniformly distributed on $\mathcal{S}_d$, then X_Π is exchangeable with c.d.f. F^{sym}.

We define the following operator on the set of copulas of order statistics.

Definition 3.1. *Let* $\mathbf{F} \in (\mathcal{L}_1)^d$. *For* $C \in \mathcal{C}$ *we define* $S_{\mathbf{F}}(C)$ *as the copula of the exchangeable random variable* X_Π, *where X is a random vector on $\mathbb{R}^d$ with one-dimensional marginal c.d.f.s $\mathbf{F}$ and copula C and Π is an independent random variable uniform on $\mathcal{S}_d$.*

The application $S_{\mathbf{F}}$ is well-defined on $\mathcal{C}$ and takes values in $\mathcal{C}^{\mathrm{sym}}$. In the above definition, with $\mathbf{F} = (\mathbf{F}_i, 1 \le i \le d)$, the one-dimensional marginal c.d.f.s of X_Π are equal to:

$$G = \frac{1}{d} \sum_{i=1}^{d} \mathbf{F}_i. \tag{17}$$

Since the one-dimensional marginal c.d.f.s $\mathbf{F}_i$ are continuous, we get that G is continuous and thus the c.d.f. of X_Π belongs to $\mathcal{L}_d$. In particular, thanks to Sklar's theorem, the copula of X_Π is indeed uniquely defined.

Combining (8), (7) and (16), we can give an explicit formula for $S_{\mathbf{F}}(C)$:

$$S_{\mathbf{F}}(C)(u) = \frac{1}{d!} \sum_{\pi \in \mathcal{S}_d} C\big(\mathbf{F}_1\big(G^{-1}(u_{\pi(1)})\big), \ldots, \mathbf{F}_d\big(G^{-1}(u_{\pi(d)})\big)\big), \qquad u \in I^d. \tag{18}$$

Remark 3.2. The copula $S_{\mathbf{F}}(C)$ is not equal in general to the exchangeable copula C^{sym} defined similarly to (16) by $C^{\mathrm{sym}} = (1/d!) \sum_{\pi \in \mathcal{S}_d} C(x_\pi)$. However, this is the case if the one-dimensional marginal c.d.f.s $\mathbf{F}_i$ are all equal, in which case $\mathbf{F}_i = G$ for all $1 \le i \le d$.

If X is a random vector on $\mathbb{R}^d$, let $X^{\mathrm{OS}} = (X_{(1)}, \ldots, X_{(d)})$ be the order statistics of X. The proof of the next lemma is elementary.

Lemma 3.3. *Let X be a random vector on $\mathbb{R}^d$ with c.d.f. F and Π a random variable independent of X, uniformly distributed on $\mathcal{S}_d$. We have*:

- *If $F \in \mathcal{L}_d^{\mathrm{OS}}$, then a.s. $(X_\Pi)^{\mathrm{OS}} = X$.*
- *If $F \in \mathcal{L}_d^{\mathrm{sym}}$, then $(X^{\mathrm{OS}})_\Pi$ has the same distribution as X.*

For $\mathbf{F} \in \mathcal{F}_d$, we define the set of copulas $\mathcal{C}^{\mathrm{sym}}(\mathbf{F}) \subset \mathcal{C}^{\mathrm{sym}}$ as the image of $\mathcal{C}^{\mathrm{OS}}(\mathbf{F})$ by the symmetrizing operator $S_{\mathbf{F}}$:

$$\mathcal{C}^{\mathrm{sym}}(\mathbf{F}) = S_{\mathbf{F}}\big(\mathcal{C}^{\mathrm{OS}}(\mathbf{F})\big). \tag{19}$$

The following lemma is one of the main result of this section.

Lemma 3.4. *Let $\mathbf{F} \in \mathcal{F}_d$. The symmetrizing operator $S_{\mathbf{F}}$ is a bijection from $\mathcal{C}^{\mathrm{OS}}(\mathbf{F})$ onto $\mathcal{C}^{\mathrm{sym}}(\mathbf{F})$.*

Proof. Let $C_1, C_2 \in \mathcal{C}^{OS}(\mathbf{F})$ with $S_\mathbf{F}(C_1) = S_\mathbf{F}(C_2)$. Let X and Y be random vectors with one-dimensional marginal c.d.f.s $\mathbf{F}$ and copula C_1, C_2, respectively. Since $C_1, C_2 \in \mathcal{C}^{OS}(\mathbf{F})$, we get that X and Y are order statistics. Notice X_Π and Y_Π have the same one-dimensional marginals according to (17) and same copula given by $S_\mathbf{F}(C_1) = S_\mathbf{F}(C_2)$. Therefore, X_Π and Y_Π have the same distribution. Thus, their corresponding order statistics $(X_\Pi)^{OS}$ and $(Y_\Pi)^{OS}$ have the same distribution. By Lemma 3.3, we get that X and Y have the same distribution as well, which implies $C_1 = C_2$. $\qquad\square$

Remark 3.5. We have in general $\mathcal{C}^{sym}(\mathbf{F}) \neq \mathcal{C}^{OS}(\mathbf{F}) \cap \mathcal{C}^{sym}$. One exception being when the marginal c.d.f.s $\mathbf{F}_i$ are all equal. In this case, both sides reduce to one copula which is the Fréchet-Hoeffding upper bound copula: $C^+(u) = \min u, u \in I^d$.

3.2. Multidiagonals and characterization of $\mathcal{C}^{sym}(\mathbf{F})$

Let $C \in \mathcal{C}$ be a copula and U a random vector with c.d.f. C. The map $t \mapsto C(t, \dots, t)$ for $t \in I$, which is called the diagonal section of C, is the c.d.f. of $\max U$. We shall consider a generalization of the diagonal section of C in the next definition.

Definition 3.6. *Let $C \in \mathcal{C}$ be a copula on $\mathbb{R}^d$ and U a random vector with c.d.f. C. The multidiagonal of the copula C, $\delta_C = (\delta_{(i)}, 1 \leq i \leq d)$, is the d-tuple of the one-dimensional marginal c.d.f.s of $U^{OS} = (U_{(1)}, \dots, U_{(d)})$ the order statistics of U: for $1 \leq i \leq d$*

$$\delta_{(i)}(t) = \mathbb{P}(U_{(i)} \leq t), \qquad t \in I.$$

We denote by $\mathcal{D} = \{\delta_C; C \in \mathcal{C}\}$ the set of multidiagonals. Notice that $\mathcal{D} \subset \mathcal{F}_d$, see Remark 3.7. For $\delta \in \mathcal{D}$ a multidiagonal, we define $\mathcal{C}_\delta = \{C; \delta_C = \delta\}$ the set of copulas with multidiagonal δ.

A characterization of the set $\mathcal{D}$ is given by Theorem 1 of [13]: a vector of functions $\delta = (\delta_{(1)}, \dots, \delta_{(d)})$ belongs to $\mathcal{D}$ if and only if $\delta_{(i)}$ is a one-dimensional c.d.f. and the following conditions hold:

$$\delta_{(i)} \geq \delta_{(i+1)}, \qquad 1 \leq i \leq d-1, \tag{20}$$

$$\sum_{i=1}^{d} \delta_{(i)}(s) = ds, \qquad 0 \leq s \leq 1. \tag{21}$$

Remark 3.7. The condition (21) implies that $\delta_{(i)} \in \mathcal{L}_1, 1 \leq i \leq d$, moreover they are d-Lipschitz. Also, it is enough to know $d-1$ functions from $\delta_{(i)}, 1 \leq i \leq d$, the remaining one is implicitly defined by (21). Condition (20) along with the continuity of $\delta_{(i)}$ implies that any multidiagonal δ_C is compatible with the continuous marginal distributions of an order statistics, therefore $\mathcal{D} \subset \mathcal{F}_d$.

Remark 3.8. Since $\delta_{(i)}, 1 \leq i \leq d$ are non-decreasing and d-Lipschitz, we have for almost every $t \in I: 0 \leq (\delta_{(i)})'(t) \leq d$ and thus $|(\delta_{(i)})'(t) \log((\delta_{(i)})'(t))| \leq d \log(d)$ for $d \geq 2$. We deduce that for $d \geq 2$:

$$\left| H(\delta_{(i)}) \right| \leq d \log(d). \tag{22}$$

Remark 3.9. Let $C \in \mathcal{C}^{\text{sym}}$ be a symmetric copula on $\mathbb{R}^d$ and U a random vector with c.d.f. C. We check that the multidiagonal $\delta_C = (\delta_{(i)}, 1 \leq i \leq d)$ can be expressed in terms of the diagonal sections $(C_{\{i\}}, 1 \leq i \leq d)$ where for $1 \leq i \leq d$:

$$C_{\{i\}}(t) = \mathbb{P}\Big(\max_{1 \leq k \leq i} U_k \leq t \Big) = C(\underbrace{t, \ldots, t}_{i \text{ terms}}, \underbrace{1, \ldots, 1}_{d-i \text{ terms}}), \qquad t \in I.$$

According to 2.8 of [13], we have for $1 \leq i \leq d$:

$$\delta_{(i)}(t) = \sum_{j=i}^{d} (-1)^{j-i} \binom{j-1}{i-1} \binom{d}{j} C_{\{j\}}(t), \qquad t \in I.$$

Conversely, we can express the functions $(C_{\{i\}}, 1 \leq i \leq d)$ with δ_C. For $1 \leq i \leq d$ and $t \in I$, we have:

$$
\begin{aligned}
C_{\{i\}}(t) &= \mathbb{P}\Big(\max_{1 \leq k \leq i} U_k \leq t \Big) \\
&= \sum_{j=i}^{d} \mathbb{P}\Big(U_{(j)} \leq t \,\big|\, \max_{1 \leq k \leq i} U_k = U_{(j)} \Big) \mathbb{P}\Big(\max_{1 \leq k \leq i} U_k = U_{(j)} \Big) \\
&= \sum_{j=i}^{d} \frac{\binom{j-1}{i-1}}{\binom{d}{i}} \delta_{(j)}(t),
\end{aligned}
$$

where we used the definition of $\delta_{(i)}$ and the exchangeability of U for the third equality.

The next technical lemma will be used in forthcoming proofs. Recall that J^{-1} denotes the generalized inverse of a non-decreasing function J, see Section 2.1 for its definition and properties, in particular, $J^{-1} \circ J(t) \leq t$ for $t \in \mathbb{R}$. Recall also that for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, we write $G(x) = (G(x_1), \ldots, G(x_d))$.

Lemma 3.10. *Let $X = (X_1, \ldots, X_d)$ be a random vector on $\mathbb{R}^d$ with one-dimensional marginals c.d.f. $(F_i, 1 \leq i \leq d)$. Set $G = \sum_{i=1}^{d} F_i / d$. We have for $1 \leq i \leq d$:*

$$\mathbb{P}\big(X_i \leq G^{-1} \circ G(t) \big) = \mathbb{P}(X_i \leq t), \qquad t \in \mathbb{R}, \ \text{that is } F_i \circ G^{-1} \circ G = F_i. \tag{23}$$

We also have for $x \in \mathbb{R}^d$:

$$\mathbb{P}\big(G(X) \leq x \big) = \mathbb{P}\big(X \leq G^{-1}(x) \big). \tag{24}$$

Proof. Since G is the average of the non-decreasing functions F_i, if $G(s) = G(s')$ for some $s, s' \in \mathbb{R}$, then we have $F_i(s) = F_i(s')$ for every $1 \leq i \leq d$. Thanks to (1), we have $G \circ G^{-1} \circ G(t) = G(t)$ and thus $F_i \circ G^{-1} \circ G(t) = F_i(t)$. This gives (23).

Recall definition (2) for $I_g(J)$ the set of points where the function J is increasing on their left. Since G is the average of the non-decreasing functions F_i, we deduce that $I_g(G) =$

$\bigcup_{1 \leq i \leq d} I_g(F_i)$. Notice that a.s. X_i belongs to $I_g(F_i)$. Thanks to (5), we get that a.s. $\{G(X) \leq x\} = \{X \leq G^{-1}(x)\}$. This gives (24). $\qquad\square$

We will also require the following lemma.

Lemma 3.11. *Let $X = (X_1, \ldots, X_d)$ be a random vector on $\mathbb{R}^d$ with one-dimensional marginals c.d.f. $(F_i, 1 \leq i \leq d)$. Set $G = \sum_{i=1}^d F_i/d$. We have for $1 \leq i \leq d$:*

$$\left(F_i \circ G^{-1}\right)^{-1} = G \circ F_i^{-1}. \tag{25}$$

Proof. Recall Definition (2) for $I_g(J)$ the set of points where the function J is increasing on their left. Let $1 \leq i \leq d$. Thanks to (4), we have $F_i^{-1}(\mathbb{R}) \subset I_g(F_i) \cup \{\pm\infty\}$. Since G is the average of the non-decreasing functions F_i, we deduce that $I_g(G) = \bigcup_{1 \leq i \leq d} I_g(F_i)$. Thus we get:

$$F_i^{-1}(\mathbb{R}) \subset I_g(G) \cup \{\pm\infty\}, \tag{26}$$

for all $1 \leq i \leq d$. The function $F_i \circ G^{-1}$ is also bounded, non-decreasing and càd-làg therefore we have for $t, s, \in \mathbb{R}$:

$$t \geq \left(F_i \circ G^{-1}\right)^{-1}(s) \quad \Longleftrightarrow \quad F_i \circ G^{-1}(t) \geq s \quad \Longleftrightarrow \quad G^{-1}(t) \geq F_i^{-1}(s)$$

$$\Longleftrightarrow \quad t \geq G \circ F_i^{-1}(s),$$

where we used the equivalence of (1) for the first and second equivalence, (26) and the equivalence of (5) for the last. This gives that $(F_i \circ G^{-1})^{-1} = G \circ F_i^{-1}$. $\qquad\square$

In the following lemma, we show that for $\mathbf{F} \in \mathcal{F}_d$, all copulas in $\mathcal{C}^{\text{sym}}(\mathbf{F})$ share the same multidiagonal denoted by $\delta^{\mathbf{F}}$.

Lemma 3.12. *Let $\mathbf{F} = (\mathbf{F}_i, 1 \leq i \leq d) \in \mathcal{F}_d$. Let $C \in \mathcal{C}^{\text{OS}}(\mathbf{F})$ and U be a random vector with c.d.f. $S_{\mathbf{F}}(C)$. Let $\delta^{\mathbf{F}} = (\delta_{(i)}, 1 \leq i \leq d)$ be the multidiagonal of $S_{\mathbf{F}}(C)$, that is the one-dimensional marginal c.d.f.s of U^{OS}, the order statistics of U. We have that $\delta^{\mathbf{F}}$ does not depend on C and for $1 \leq i \leq d$:*

$$\delta_{(i)} = \mathbf{F}_i \circ G^{-1} \quad \text{and} \quad \delta_{(i)}^{-1} = G \circ \mathbf{F}_i^{-1}, \tag{27}$$

with G given by (17). Furthermore, C is the unique copula of U^{OS}.

With obvious notation, we might simply write $\delta^{\mathbf{F}} = \mathbf{F} \circ G^{-1}$, with G given by (17).

Proof of Lemma 3.12. Let X be a random vector of order statistics with marginals $\mathbf{F} \in \mathcal{F}_d$ and copula C. Then $S_{\mathbf{F}}(C)$ is the copula of the exchangeable random vector X_Π, where Π is uniform on $\mathcal{S}_d$ and independent of X. We have already seen in (17) that the one-dimensional marginals of X_Π have the same distribution given by $G \in \mathcal{L}_1$. Thanks to (9), we deduce that the random vector U, with c.d.f. $S_{\mathbf{F}}(C)$, has the same distribution as $G(X_\Pi)$. Since G is non-decreasing, this

implies that the order statistics of U, U^{OS}, has the same distribution as $G((X_\Pi)^{OS})$ that is as $G(X)$, thanks to Lemma 3.3. Then use (24) to get for $x \in \mathbb{R}^d$:

$$\mathbb{P}(U^{OS} \le x) = \mathbb{P}(G(X) \le x) = \mathbb{P}(X \le G^{-1}(x)). \tag{28}$$

This gives the first part of the lemma as the multidiagonal of U is the one-dimensional marginal c.d.f.s of its order statistics. The second equation in (27) is due to Lemma 3.11. The fact that C is the copula of U^{OS} and its uniqueness are due to (28) and the continuity of $\delta_{(i)}$, see Remark 3.7. $\square$

The next proposition shows that the set $\mathcal{C}^{sym}(\mathbf{F})$ is actually the set of symmetric copulas with diagonal section $\delta^{\mathbf{F}}$. This yields the main result of this section given by the subsequent corollary.

Proposition 3.13. *Let* $\mathbf{F} \in \mathcal{F}_d$. *We have* $\mathcal{C}^{sym}(\mathbf{F}) = \mathcal{C}_{\delta^{\mathbf{F}}} \cap \mathcal{C}^{sym}$.

Proof. By Lemma 3.12, we have $\mathcal{C}^{sym}(\mathbf{F}) \subset \mathcal{C}_{\delta^{\mathbf{F}}} \cap \mathcal{C}^{sym}$.

Let $C \in \mathcal{C}_{\delta^{\mathbf{F}}} \cap \mathcal{C}^{sym}$ and U be a random vector with c.d.f. C. Let G be given by (17). Notice that $X = G^{-1}(U)$ is an exchangeable random vector with marginals G and copula C. Thanks to Lemma 3.3, the proof will be complete as soon as we prove that the one-dimensional marginal c.d.f.s of $X^{OS} = (X_{(1)}, \ldots, X_{(d)})$, the order statistics of X, is given by $\mathbf{F}$. Notice $X^{OS} = G^{-1}(U^{OS})$, with U^{OS} the order statistics of U whose one-dimensional marginal c.d.f.s are given by $\delta^{\mathbf{F}}$. We have for $1 \le i \le d$ and $t \in \mathbb{R}$:

$$\mathbb{P}(X_{(i)} \le t) = \mathbb{P}(G^{-1}(U_{(i)}) \le t) = \mathbb{P}(U_{(i)} \le G(t)) = \mathbf{F}_i \circ G^{-1} \circ G(t) = \mathbf{F}_i(t),$$

where we used (1) for the second equality, (27) for the third, and (23) for the last. This finishes the proof. $\square$

Corollary 3.14. *According to Proposition 3.13, (19) and Lemma 3.4, we get that for any* $\mathbf{F} \in \mathcal{F}_d$, *the symmetrizing operator* $S_{\mathbf{F}}$ *is a bijection between* $\mathcal{C}^{OS}(\mathbf{F})$ *and* $\mathcal{C}_{\delta^{\mathbf{F}}} \cap \mathcal{C}^{sym}$.

We end this section by an ancillary result we shall use later.

Lemma 3.15. *Let* $\mathbf{F} \in \mathcal{F}_d$. *We have* $\mathbb{J}(\mathbf{F}) = \mathbb{J}(\delta^{\mathbf{F}})$.

Proof. Let $\mathbf{F} = (\mathbf{F}_i, 1 \le i \le d)$. We get, using (27) and the change of variable $s = G^{-1}(t)$ that:

$$\int_I \delta'_{(i)}(t) \big| \log(\delta_{(i-1)}(t) - \delta_{(i)}(t)) \big| \, dt = \int_{G^{-1}((0,1))} \mathbf{F}_i(ds) \big| \log(\mathbf{F}_{i-1}(s) - \mathbf{F}_i(s)) \big|.$$

Since $d\mathbf{F}_i = 0$ outside $G^{-1}((0,1))$ (as G is increasing as soon as $\mathbf{F}_i$ is increasing), we get that the last integration above is also over $\mathbb{R}$. We deduce that:

$$\mathbb{J}(\delta^{\mathbf{F}}) = \sum_{i=2}^d \int_{\mathbb{R}} \mathbf{F}_i(ds) \big| \log(\mathbf{F}_{i-1}(s) - \mathbf{F}_i(s)) \big| = \mathbb{J}(\mathbf{F}).$$

$\square$

3.3. Density and entropy of copulas in $\mathcal{C}^{\mathrm{sym}}(\mathbf{F})$

We prove in this section that $S_{\mathbf{F}}$ preserves the absolute continuity on $\mathcal{C}^{\mathrm{OS}}(\mathbf{F})$ for $\mathbf{F} \in \mathcal{F}_d$ and the entropy up to a constant. Let us introduce some notation. For marginals $\mathbf{F} \in \mathcal{F}_d$, let

$$\Psi_i^{\mathbf{F}} = \{s \in \mathbb{R}, \mathbf{F}_{i-1}(s) > \mathbf{F}_i(s)\} \qquad \text{for } 2 \le i \le d. \tag{29}$$

The complementary set $(\Psi_i^{\mathbf{F}})^c$ is the collection of the points where $\mathbf{F}_{i-1} = \mathbf{F}_i$. We define $\Sigma^{\mathbf{F}} \subset I$ as:

$$\Sigma^{\mathbf{F}} = \bigcup_{i=2}^{d} \mathbf{F}_i\big((\Psi_i^{\mathbf{F}})^c\big). \tag{30}$$

By Remark 3.7, we have $\mathcal{D} \subset \mathcal{F}_d$, then the definitions (29) and (30) apply for all $\delta \in \mathcal{D}$. In particular, for $\delta = (\delta_{(1)}, \ldots, \delta_{(d)}) \in \mathcal{D}$ the sets Ψ_i^δ, $2 \le i \le d$ are open subsets of I, therefore $(\Psi_i^\delta)^c \cap I$ is a compact subset. This and the continuity of $\delta_{(i)}$ imply that $\delta_{(i)}((\Psi_i^\delta)^c) = \delta_{(i)}((\Psi_i^\delta)^c \cap I)$ is also compact, hence Σ^δ is compact. Notice that $\{0, 1\} \subset \Sigma^\delta$ always holds. We define $\mathcal{C}_\delta^0 = \mathcal{C}_\delta \cap \mathcal{C}^0$ the subset of absolutely continuous copulas with multidiagonal δ and the subset $\mathcal{D}^0 = \{\delta \in \mathcal{D}, \mathcal{C}_\delta^0 \ne \varnothing\}$ of multidiagonals of absolutely continuous copulas. According to Theorem 2 of [13], the multidiagonal δ belongs to $\mathcal{D}^0$ if and only if it belongs to $\mathcal{D}$ and the Lebesgue measure of Σ^δ is zero: $|\Sigma^\delta| = 0$.

Lemma 3.16. *Let $\delta \in \mathcal{D}$. We have $\delta \in \mathcal{D}^0$ if and only if for all $2 \le i \le d$, a.e.:*

$$\delta'_{(i-1)}\mathbf{1}_{(\Psi_i^\delta)^c} = \delta'_{(i)}\mathbf{1}_{(\Psi_i^\delta)^c} = 0. \tag{31}$$

Furthermore, we have that $\mathbb{J}(\delta) < +\infty$ implies $\delta \in \mathcal{D}^0$.

Proof. Let J be a function defined on I, Lipschitz and non-decreasing. Let A be a Borel subset of I. We have:

$$|J(A)| = \int \mathbf{1}_{J(A)}(t)\,dt = \int_0^1 \mathbf{1}_{\{s \in J^{-1} \circ J(A)\}} J'(s)\,ds = \int_0^1 \mathbf{1}_A(s) J'(s)\,ds,$$

where we used (3) and (5) for the last equality. This gives that $|J(A)| = 0$ if and only if a.e. $J'\mathbf{1}_A = 0$. Then use that $\delta \in \mathcal{D}^0$ if and only if $|\delta_{(i)}((\Psi_i^\delta)^c)| = 0$ for all $1 \le i \le d$ and that $\delta_{(i-1)}((\Psi_i^\delta)^c) = \delta_{(i)}((\Psi_i^\delta)^c)$ to conclude that $\delta \in \mathcal{D}^0$ if and only if (31) holds for all $2 \le i \le d$. The last part of the lemma is clear. $\qquad\square$

Definition 3.17. *Let $\mathcal{F}_d^0 \subset \mathcal{F}_d$ be the subset of marginals $\mathbf{F}$ such that there exists an absolutely continuous c.d.f. of order statistics with marginals $\mathbf{F}$, that is $\mathcal{L}_d^{OS}(\mathbf{F}) \cap \mathcal{L}_d^0 \ne \varnothing$.*

In particular, we have $\mathcal{D}^0 \subset \mathcal{F}_d^0$. The next lemma gives a characterization of the set $\mathcal{F}_d^0$.

Lemma 3.18. *Let $\mathbf{F} \in \mathcal{F}_d$. Then $\mathbf{F} \in \mathcal{F}_d^0$ if and only if $\mathbf{F}_i \in \mathcal{L}_1^0$ for $1 \leq i \leq d$ and $|\Sigma^{\mathbf{F}}| = 0$. Furthermore, we have that $\mathbf{F}_i \in \mathcal{L}_1^0$ for $1 \leq i \leq d$ and $\mathbb{J}(\mathbf{F}) < +\infty$ imply $\mathbf{F} \in \mathcal{F}_d^0$.*

Proof. Let $F \in \mathcal{L}_d^{OS}(\mathbf{F})$. We know that $F \in \mathcal{L}_d^0$ if and only if $\mathbf{F}_i \in \mathcal{L}_1^0$ for $1 \leq i \leq d$ and $C_F \in \mathcal{C}^0$, the subset of absolutely continuous copulas (see, for example, [12]). Therefore, $\mathbf{F} \in \mathcal{F}_d^0$ if and only if $\mathbf{F}_i \in \mathcal{L}_1^0$ for $1 \leq i \leq d$ and $\mathcal{C}^{OS}(\mathbf{F}) \cap \mathcal{C}^0 \neq \varnothing$. Recall that $\delta^{\mathbf{F}}$ is defined by (27). We first show that

$$\mathcal{C}^{OS}(\mathbf{F}) \cap \mathcal{C}^0 \neq \varnothing \quad \text{if and only if} \quad \mathcal{C}_{\delta \mathbf{F}}^0 \cap \mathcal{C}^{\text{sym}} \neq \varnothing. \tag{32}$$

Let $C \in \mathcal{C}^{OS}(\mathbf{F}) \cap \mathcal{C}^0$. Then Lemma 3.12 ensures that $S_{\mathbf{F}}(C) \in \mathcal{C}_{\delta \mathbf{F}} \cap \mathcal{C}^{\text{sym}}$. The absolute continuity of $S_{\mathbf{F}}(C)$ is a direct consequence of (18), (27) and Remark 3.7 which ensures that $\delta_{(i)}^{\mathbf{F}}, 1 \leq i \leq d$ are d-Lipschitz, therefore their derivatives exist a.e. on I. This ensures that $\mathcal{C}_{\delta \mathbf{F}}^0 \cap \mathcal{C}^{\text{sym}} \neq \varnothing$.

Conversely, let $C \in \mathcal{C}_{\delta \mathbf{F}}^0 \cap \mathcal{C}^{\text{sym}}$. Let U be a random vector with c.d.f. C. Then its order statistics U^{OS} is also absolutely continuous. Therefore, the copula of U^{OS}, which is $S_{\mathbf{F}}^{-1}(C)$ by Lemma 3.12, is also absolutely continuous. This proves thanks to Proposition 3.13 and Lemma 3.4 that $S_{\mathbf{F}}^{-1}(C) \in \mathcal{C}^{OS}(\mathbf{F}) \cap \mathcal{C}^0$. This gives (32).

Notice that $\mathcal{C}_{\delta \mathbf{F}}^0 \cap \mathcal{C}^{\text{sym}} \neq \varnothing$ is equivalent to $\mathcal{C}_{\delta \mathbf{F}}^0 \neq \varnothing$, since for any $C \in \mathcal{C}_{\delta \mathbf{F}}^0$ we have that C^{sym} defined by (16) belongs to $\mathcal{C}_{\delta \mathbf{F}}^0 \cap \mathcal{C}^{\text{sym}}$. By Theorem 2 of [13], $\mathcal{C}_{\delta \mathbf{F}}^0 \neq \varnothing$ if and only if $\Sigma^{\delta^{\mathbf{F}}}$ has zero Lebesgue measure. The proof is then complete as one can easily verify using (27) that $\Sigma^{\delta^{\mathbf{F}}} = \Sigma^{\mathbf{F}}$ and thanks to Lemma 3.15. $\qquad\square$

From now on, we consider $\mathbf{F} \in \mathcal{F}_d^0$. We give an auxiliary lemma on the support of the copulas in $\mathcal{C}^{OS}(\mathbf{F}) \cap \mathcal{C}^0$.

Lemma 3.19. *Let $\mathbf{F} = (\mathbf{F}_i, 1 \leq i \leq d) \in \mathcal{F}_d^0$ and $C \in \mathcal{C}^{OS}(\mathbf{F}) \cap \mathcal{C}^0$. Then the density of C vanishes a.e. on $I^d \setminus T^{\mathbf{F}}$ with:*

$$T^{\mathbf{F}} = \left\{ u = (u_1, \ldots, u_d) \in I^d; \mathbf{F}_1^{-1}(u_1) \leq \cdots \leq \mathbf{F}_d^{-1}(u_d) \right\}. \tag{33}$$

Proof. Let $X = (X_1, \ldots, X_d)$ be a random vector of order statistics with one-dimensional marginal c.d.f.s $\mathbf{F}$ and copula $C \in \mathcal{C}^0$. Let $U = (U_1, \ldots, U_d)$ be a random vector with c.d.f. C. Then it is distributed as $(\mathbf{F}_1(X_1), \ldots, \mathbf{F}_d(X_d))$, see (9). We get $\mathbb{P}(U \in T^{\mathbf{F}}) = 1$, since X is a vector of order statistics and $X_i \in I_g(\mathbf{F}_i)$ a.s. for $1 \leq i \leq d$. This gives the result. $\qquad\square$

Now we establish the connection between the sets $\mathcal{C}^{OS}(\mathbf{F}) \cap \mathcal{C}^0$ and $\mathcal{C}^{\text{sym}}(\mathbf{F}) \cap \mathcal{C}^0$.

Lemma 3.20. *Let $\mathbf{F} \in \mathcal{F}_d^0$. The symmetrizing operator $S_{\mathbf{F}}$ is a bijection from $\mathcal{C}^{OS}(\mathbf{F}) \cap \mathcal{C}^0$ onto $\mathcal{C}^{\text{sym}}(\mathbf{F}) \cap \mathcal{C}^0$. Moreover, if $C \in \mathcal{C}^{OS}(\mathbf{F}) \cap \mathcal{C}^0$, with density function c, then the density function*

$s_{\mathbf{F}}(C)$ of $S_{\mathbf{F}}(C)$ is given by, for a.e. $u = (u_1, \ldots, u_d) \in I^d$:

$$s_{\mathbf{F}}(C)(u) = \frac{1}{d!} c\big(\delta_{(1)}(u_{(1)}), \ldots, \delta_{(d)}(u_{(d)})\big) \prod_{i=1}^{d} \delta'_{(i)}(u_{(i)}). \tag{34}$$

Let $T^{\mathbf{F}}$ be given by (33). If $C \in \mathcal{C}^{\mathrm{sym}}(\mathbf{F}) \cap \mathcal{C}_0$ with density c, then the density $s_{\mathbf{F}}^{-1}(C)$ of $S_{\mathbf{F}}^{-1}(C)$ is given by, for a.e. $u = (u_1, \ldots, u_d) \in I^d$:

$$s_{\mathbf{F}}^{-1}(C)(u) = d! \frac{c(\delta_{(1)}^{-1}(u_1), \ldots, \delta_{(d)}^{-1}(u_d))}{\prod_{i=1}^{d} \delta'_{(i)} \circ \delta_{(i)}^{-1}(u_i)} \mathbf{1}_{T^{\mathbf{F}}}(u) \mathbf{1}_{\{\prod_{i=1}^{d} \delta'_{(i)} \circ \delta_{(i)}^{-1}(u_i) > 0\}}. \tag{35}$$

Proof. By Proposition 3.13, we deduce that $\mathcal{C}^{\mathrm{sym}}(\mathbf{F}) \cap \mathcal{C}^0 = \mathcal{C}^0_{\delta\mathbf{F}} \cap \mathcal{C}^{\mathrm{sym}}$. Lemma 3.4 and the proof of Lemma 3.18 ensures that $S_{\mathbf{F}}$ is a bijection between $\mathcal{C}^{\mathrm{OS}}(\mathbf{F}) \cap \mathcal{C}^0$ and $\mathcal{C}^{\mathrm{sym}}(\mathbf{F}) \cap \mathcal{C}^0$. The explicit formula (34) can be obtained by taking the mixed derivative of the right hand side of (18). By Lemma 3.19, all the terms in the sum disappear except the one on the right hand side of (34).

To obtain (35), let $C \in \mathcal{C}^{\mathrm{sym}}(\mathbf{F}) \cap \mathcal{C}^0$ with density c, and U be a random vector with c.d.f. C. The order statistics U^{OS} derived from U is also absolutely continuous with cumulative distribution function K, and density function k given by:

$$k(u) = d! c(u) \mathbf{1}_{\triangle}(u), \qquad u \in I^d.$$

By Lemma 3.12, $S_{\mathbf{F}}^{-1}(C)$ is the copula of U^{OS}. From (7), we have for $u = (u_1, \ldots, u_d) \in I^d$:

$$S_{\mathbf{F}}^{-1}(C)(u) = K\big(\delta_{(1)}^{-1}(u_1), \ldots, \delta_{(d)}^{-1}(u_d)\big). \tag{36}$$

According to (5), we deduce that $G^{-1} \circ G \circ \mathbf{F}_i^{-1} = \mathbf{F}_i^{-1}$ on $(0, 1)$. This implies that for $s, t \in (0, 1)$, $1 \le i < j \le d$:

$$
\begin{aligned}
\delta_{(i)}^{-1}(s) \le \delta_{(j)}^{-1}(t) \quad &\Leftrightarrow \quad G \circ \mathbf{F}_i^{-1}(s) \le G \circ \mathbf{F}_j^{-1}(t) \\
&\Rightarrow \quad G^{-1} \circ G \circ \mathbf{F}_i^{-1}(s) \le G^{-1} \circ G \circ \mathbf{F}_j^{-1}(t) \\
&\Leftrightarrow \quad \mathbf{F}_i^{-1}(s) \le \mathbf{F}_j^{-1}(t) \\
&\Rightarrow \quad G \circ \mathbf{F}_i^{-1}(s) \le G \circ \mathbf{F}_j^{-1}(t),
\end{aligned}
$$

where we used (27) for the first equivalence, that G^{-1} is non-decreasing for the first implication and G is non-decreasing for the second. Thus we have, for $s, t \in (0, 1)$, that the two conditions $\delta_{(i)}^{-1}(s) \le \delta_{(j)}^{-1}(t)$ and $\mathbf{F}_i^{-1}(s) \le \mathbf{F}_j^{-1}(t)$ are equivalent. Thus, we deduce that the two sets

$$\big\{(u_1, \ldots, u_d) \in I^d; \delta_{(1)}^{-1}(u_1) \le \cdots \le \delta_{(d)}^{-1}(u_d)\big\}$$

and $T^{\mathbf{F}}$ are equal up to a set of zero Lebesgue measure. Then we deduce (35) from (36). $\qquad\square$

We give a general result on the entropy of an exchangeable random vector and the entropy of its order statistics.

Lemma 3.21. *Let X be a random vector on I^d, X^{OS} the corresponding order statistics and Π an independent uniform random variable on $\mathcal{S}_d$. Then we have*:

$$H\big((X^{\mathrm{OS}})_\Pi\big) = \log(d!) + H\big(X^{\mathrm{OS}}\big).$$

Proof. Let F be the c.d.f. of X^{OS}. If $F \notin \mathcal{L}_d^0$, then the c.d.f. F^{sym} of $(X^{\mathrm{OS}})_\Pi$ given by (16) verifies also $F^{\mathrm{sym}} \notin \mathcal{L}_d^0$, therefore $H((X^{\mathrm{OS}})_\Pi) = H(X^{\mathrm{OS}}) + \log(d!) = -\infty$. If $F \in \mathcal{L}_d^0$ with density function f, then the density function f^{sym} of F^{sym} is given by, for $x \in I^d$:

$$f^{\mathrm{sym}}(x) = \frac{1}{d!} f\big(x^{\mathrm{OS}}\big),$$

where x^{OS} is the ordered vector of x. Therefore, using that $f(x) = 0$ if $x \neq x^{\mathrm{OS}}$, we have:

$$
\begin{aligned}
H\big((X^{\mathrm{OS}})_\Pi\big) &= -\int_{I^d} f^{\mathrm{sym}} \log\big(f^{\mathrm{sym}}\big) \\
&= \log(d!) - \frac{1}{d!} \int_{I^d} f\big(x^{\mathrm{OS}}\big) \log\big(f\big(x^{\mathrm{OS}}\big)\big)\, dx \\
&= \log(d!) - \int_{I^d} f(x) \log\big(f(x)\big)\, dx \\
&= \log(d!) + H\big(X^{\mathrm{OS}}\big). \qquad\qquad \square
\end{aligned}
$$

Now we are ready to give the connection between the entropy of C and $S_{\mathbf{F}}(C)$ for $C \in \mathcal{C}^{\mathrm{OS}}(\mathbf{F})$, which is the main result of this section. Recall the definition of $\delta^{\mathbf{F}} = (\delta_{(i)}^{\mathbf{F}}, 1 \leq i \leq d)$ given in Lemma 3.12 and thanks to Remark 3.8, $H(\delta_{(i)}^{\mathbf{F}})$ is finite for all $1 \leq i \leq d$.

Proposition 3.22. *Let $\mathbf{F} \in \mathcal{F}_d$ and $C \in \mathcal{C}^{\mathrm{OS}}(\mathbf{F})$. Then we have*:

$$H\big(S_{\mathbf{F}}(C)\big) = \log(d!) + H(C) + \sum_{i=1}^{d} H\big(\delta_{(i)}^{\mathbf{F}}\big). \tag{37}$$

Proof. Let U be an exchangeable random vector with c.d.f. $S_{\mathbf{F}}(C)$, and U^{OS} its order statistics. According to Lemma 3.12, U^{OS} has one-dimensional marginal c.d.f.s $\delta^{\mathbf{F}} = (\delta_{(i)}^{\mathbf{F}}, 1 \leq i \leq d)$ and copula C. Therefore, using (14) with $h = \mathbf{1}_{[0,1]}$, we get:

$$H\big(U^{\mathrm{OS}}\big) = H(C) + \sum_{i=1}^{d} H\big(\delta_{(i)}^{\mathbf{F}}\big).$$

On the other hand, since $S_{\mathbf{F}}(C)$ is symmetric, Lemma 3.3 ensures that $(U^{\mathrm{OS}})_{\Pi}$ has the same distribution as U. Therefore Lemma 3.21 gives:

$$H\big(S_{\mathbf{F}}(C)\big) = H(U) = H\big((U^{\mathrm{OS}})_{\Pi}\big) = H\big(U^{\mathrm{OS}}\big) + \log(d!) = H(C) + \sum_{i=1}^{d} H\big(\delta_{(i)}^{\mathbf{F}}\big) + \log(d!). \qquad \square$$

Corollary 3.23. *Since the marginal c.d.f.s $\mathbf{F}$ are fixed, the difference between $H(C)$ and $H(S_{\mathbf{F}}(C))$ is constant for all $C \in \mathcal{C}^{\mathrm{OS}}(\mathbf{F})$. Therefore if the entropy of a copula $C \in \mathcal{C}^{\mathrm{OS}}(\mathbf{F})$ is maximal, then $S_{\mathbf{F}}(C)$ also has maximal entropy on $\mathcal{C}^{\mathrm{sym}}(\mathbf{F}) = \mathcal{C}_{\delta \mathbf{F}} \cap \mathcal{C}^{\mathrm{sym}}$.*

4. Maximum entropy copula with given multidiagonals

This section is a generalization of [6], where the maximum entropy copula with given diagonal section (i.e., given distribution for the maximum of its marginals) is studied.

Recall that multidiagonals of copulas on $\mathbb{R}^d$ are given by Definition 3.6. We recall some further notation: $\mathcal{D}$ denotes the set of multidiagonals; for $\delta \in \mathcal{D}$, $\mathcal{C}_{\delta}$ denotes the subset of copulas with multidiagonal δ; $\mathcal{C}^0$ denotes the subset copulas which are absolutely continuous, and $\mathcal{C}_{\delta}^0 = \mathcal{C}_{\delta} \cap \mathcal{C}^0$. The set $\mathcal{D}^0 \subset \mathcal{D}$ contains all diagonals for which $\mathcal{C}_{\delta}^0 \neq \varnothing$.

We give an explicit formula for C^* such that $H(C^*) = \max_{C \in \mathcal{C}_{\delta}} H(C)$, with H the entropy, see definition (13). Notice that the maximum can be taken over $\mathcal{C}_{\delta}^0$, since the entropy is minus infinity otherwise. When $d = 2$, the problem was solved in [6].

Let $\delta = (\delta_{(i)}, 1 \le i \le d) \in \mathcal{D}$ be a multidiagonal. Since $\delta_{(i)}$, $1 \le i \le d$ are d-Lipschitz, the entropy of $H(\delta_{(i)})$ is well defined and finite, see Remark 3.8 and $\mathbb{J}(\delta)$ given by (15) is also well defined and belongs to $[0, +\infty]$.

The next two lemmas provide sets on which the density of a copula with given multidiagonal is zero. For $\delta \in \mathcal{D}$, let:

$$Z_{\delta} = \big\{ u \in I^d; \text{ there exists } 1 \le i \le d \text{ such that } \delta_{(i)}'(u_{(i)}) = 0 \big\}. \tag{38}$$

Lemma 4.1. *Let $\delta \in \mathcal{D}^0$. Then for all copulas $C \in \mathcal{C}_{\delta}^0$ with density c, we have $c\mathbf{1}_{Z_{\delta}} = 0$ a.e. that is $c(u)\mathbf{1}_{Z_{\delta}}(u) = 0$ for a.e. $u \in I^d$.*

Proof. By definition of $\delta_{(i)}$, we have for all $r \in I$:

$$\int_{I^d} c(u)\mathbf{1}_{\{u_{(i)} \le r\}}\, du = \delta_{(i)}(r) = \int_0^r \delta_{(i)}'(s)\, ds.$$

This implies, by the monotone class theorem, that for all measurable subsets K of I, we have:

$$\int_{I^d} c(u)\mathbf{1}_K(u_{(i)})\, du = \int_K \delta_{(i)}'(s)\, ds.$$

Since $c \ge 0$ a.e., we deduce that a.e. $c(u)\mathbf{1}_{\{\delta_{(i)}'(u_{(i)})=0\}} = 0$ and thus a.e. $c\mathbf{1}_{Z_{\delta}} = 0$. $\qquad \square$

Recall the definition of Ψ_i^δ given by (29) for $2 \le i \le d$. We also define $\Psi_1^\delta = (0, d_1)$ with $d_1 = \inf\{s \in I; \delta_{(1)}(s) = 1\}$ and $\Psi_{d+1}^\delta = (g_{d+1}, 1)$ with $g_{d+1} = \sup\{s \in I; \delta_{(d)}(s) = 0\}$. Since Ψ_i^δ are open subsets of I, there exist at most countably many disjoint intervals $\{(g_i^{(j)}, d_i^{(j)}), j \in J_i\}$ such that

$$\Psi_i^\delta = \bigcup_{j \in J_i} \left(g_i^{(j)}, d_i^{(j)} \right). \tag{39}$$

We denote by $m_i^{(j)} = (g_i^{(j)} + d_i^{(j)})/2$ the midpoint of these intervals for $2 \le i \le d+1$. In particular $m_{d+1} = (1 + g_{d+1})/2$. We also define $m_1 = 0$. For $\delta \in \mathcal{D}$, let:

$$L_\delta = \left\{ u = (u_1, \ldots, u_d) \in I^d; (u_{(i-1)}, u_{(i)}) \subset \Psi_i^\delta \text{ for all } 2 \le i \le d \right\}. \tag{40}$$

We have the following lemma for all absolutely continuous copulas $C \in \mathcal{C}_\delta^0$ with density c.

Lemma 4.2. *Let $\delta \in \mathcal{D}^0$ and $2 \le i \le d$. Then for all copulas $C \in \mathcal{C}_\delta^0$ with density c, we have $c \mathbf{1}_{I \setminus L_\delta} = 0$ a.e., that is for a.e. $u = (u_1, \ldots, u_d) \in I^d$, for all $s \notin \Psi_i^\delta$:*

$$c(u) \mathbf{1}_{\{u_{(i-1)} < s < u_{(i)}\}} = 0. \tag{41}$$

Proof. The complementary set $(\Psi_i^\delta)^c$ is given by:

$$\left(\Psi_i^\delta \right)^c = \overline{\bigcup_{j \in J_i} \left\{ g_i^{(j)}, d_i^{(j)} \right\}}. \tag{42}$$

Let $U = (U_1, \ldots, U_d)$ be a random vector with c.d.f. $C \in \mathcal{C}_\delta^0$. For $2 \le i \le d$ and $s \in \bigcup_{j \in J_i} \{g_i^{(j)}, d_i^{(j)}\}$, that is $\delta_{(i-1)}(s) = \delta_{(i)}(s)$, we have:

$$\mathbb{P}(U_{(i-1)} < s < U_{(i)}) = \mathbb{P}(U_{(i-1)} < s) - \mathbb{P}(U_{(i)} \le s) = \delta_{(i-1)}(s) - \delta_{(i)}(s) = 0.$$

This implies that (41) holds a.e. for all $s \in \bigcup_{j \in J_i} \{g_i^{(j)}, d_i^{(j)}\}$. Since J_i is at most countable, we have for a.e. $u \in I^d$ and for all $s \in \bigcup_{j \in J_i} \{g_i^{(j)}, d_i^{(j)}\}$, that (41) holds. Since for all $u \in I$, $s \notin \Psi_i^\delta$ there exists $s' \in \bigcup_{j \in J_i} \{g_i^{(j)}, d_i^{(j)}\}$ such that

$$\mathbf{1}_{\{u_{(i-1)} < s < u_{(i)}\}} = \mathbf{1}_{\{u_{(i-1)} < s' < u_{(i)}\}},$$

we can conclude that for a.e. $u \in I^d$ and for all $s \notin \Psi_i^\delta$ (41) hold. $\square$

Notice that for all $u = (u_1, \ldots, u_d) \in I^d$:

$$\mathbf{1}_{L_\delta}(u) \le \prod_{i=1}^{d} \mathbf{1}_{\Psi_i^\delta \cap \Psi_{i+1}^\delta}(u_{(i)}). \tag{43}$$

We define the function c_δ on I^d as, for $u = (u_1, \ldots, u_d) \in I^d$:

$$c_\delta(u) = \frac{1}{d!} \mathbf{1}_{L_\delta}(u) \prod_{i=1}^{d} a_i(u_{(i)}), \tag{44}$$

where the function a_i, $1 \le i \le d$, are given by, for $t \in I$:

$$a_i(t) = K_i'(t) e^{K_{i+1}(t) - K_i(t)} \mathbf{1}_{\Psi_i^\delta \cap \Psi_{i+1}^\delta}(t), \tag{45}$$

with for $1 \le i \le d$, $t \in (g_i^{(j)}, d_i^{(j)})$:

$$K_i(t) = \int_{m_i^{(j)}}^{t} \frac{\delta_{(i)}'(s)}{\delta_{(i-1)}(s) - \delta_{(i)}(s)} \, ds \tag{46}$$

and the conventions $\delta_{(0)} = 1$ and $K_{d+1} = 0$. Notice that for $t \in \Psi_1^\delta$:

$$K_1(t) = -\log\big(1 - \delta_{(1)}(t)\big). \tag{47}$$

Remark 4.3. The choice of $m_i^{(j)}$ for the integration lower bound in (46) is arbitrary as any other value in $(g_i^{(j)}, d_i^{(j)})$ would not change the definition of c_δ in (44).

Remark 4.4. For all $1 \le i \le d$, $j \in J_i$, $t \in (m_i^{(j)}, d_i^{(j)})$, we have the following lower bound for $K_i(t)$:

$$K_i(t) \ge \int_{m_i^{(j)}}^{t} \frac{\delta_{(i)}'(s)}{\delta_{(i-1)}(d_i^{(j)}) - \delta_{(i)}(s)} \, ds = \log\left(\frac{\delta_{(i-1)}(d_i^{(j)}) - \delta_{(i)}(m_i^{(j)})}{\delta_{(i-1)}(d_i^{(j)}) - \delta_{(i)}(t)}\right).$$

Since $\delta_{(i)}$ is non-decreasing and $\delta_{(i-1)}(d_i^{(j)}) = \delta_{(i)}(d_i^{(j)})$, we have $\lim_{t \nearrow d_i^{(j)}} K_i(t) = +\infty$.

The following proposition states that c_δ is the density of an absolutely continuous symmetric copula $C_\delta \in \mathcal{C}_\delta^0 \cap \mathcal{C}^{\mathrm{sym}}$. It is more general than the results in [6], where only the diagonal $\delta_{(d)}$ was supposed given.

Proposition 4.5. *Let $\delta \in \mathcal{D}^0$. The function c_δ defined in (44)–(46) is the density of a symmetric copula $C_\delta \in \mathcal{C}_\delta^0 \cap \mathcal{C}^{\mathrm{sym}}$. In addition, we have*:

$$H(C_\delta) = -\mathbb{J}(\delta) + \log(d!) + (d-1) + \sum_{i=1}^{d} H(\delta_{(i)}). \tag{48}$$

The proof of this proposition is given in Section 6.2. The following characterization of C_δ is proved in Section 6.6.

Proposition 4.6. *Let $\delta \in \mathcal{D}^0$. Then C_δ is the only copula in $\mathcal{C}_\delta^0$ whose density is of the form $(1/d!)\mathbf{1}_{L_\delta}(u) \prod_{i=1}^d h_i(u_{(i)})$, where h_i, $1 \le i \le d$ are measurable non-negative functions defined on I.*

The following theorem states that the unique optimal solution of $\max_{C \in \mathcal{C}_\delta} H(C)$, if it exists, is given by C_δ. Its proof is given in Sections 6.7 for case (a) and 6.8 for case (b).

Theorem 4.7. *Let $\delta \in \mathcal{D}$.*

    (a) *If $\mathbb{J}(\delta) = +\infty$, then $\max_{C \in \mathcal{C}_\delta} H(C) = -\infty$.*

    (b) *If $\mathbb{J}(\delta) < +\infty$, then $\delta \in \mathcal{D}^0$, $\max_{C \in \mathcal{C}_\delta} H(C) > -\infty$ and C_δ given in Proposition 4.5 is the unique copula such that $H(C_\delta) = \max_{C \in \mathcal{C}_\delta} H(C)$.*

The copula C_δ will be called the maximum entropy copula with given multidiagonal.

Remark 4.8. In [6], we considered the problem of the maximum entropy copula with given diagonal section, that is when only $\delta_{(d)}$ is fixed. When $d = 2$, the problem considered here coincides with the problem of maximum entropy copula with given diagonal section, see Remark 3.7. For $d > 2$, the constraints of the problem discussed here are more restrictive. With the same techniques it is possible to calculate the maximum entropy copula for which the c.d.f. of the k largest order statistics are given (that is $\delta_{(i)}$ are given for $d - k + 1 \le i \le d$). Reasoning the same way as in the proof of Theorem 4.7, we can deduce that this copula will be of the form $\prod_{i=1}^{d-k} \tilde{b}(u_{(i)}) \prod_{i=d-k+1}^d \tilde{a}_i(u_{(i)})$ on its domain, involving $k + 1$ different functions $\tilde{b}$ and $\tilde{a}_i$, $d - k + 1 \le i \le d$ to compute based on the constraints.

5. Maximum entropy distribution of order statistics with given marginals

We use the results of Section 4 to compute the density of the maximum entropy copula for marginals $\mathbf{F} \in \mathcal{F}_d^0$ with $\mathcal{F}_d^0$ defined in Section 3.3. Recall $\delta^{\mathbf{F}} = (\delta_{(1)}, \ldots, \delta_{(d)}) = \mathbf{F} \circ G^{-1}$ and the definition of $\Sigma^{\delta^{\mathbf{F}}}$ in (30). Recall K_i defined by (46), for $1 \le i \le d$ and $T^{\mathbf{F}}$ defined by (33). We define the function $c_{\mathbf{F}}$ on I^d, for $u = (u_1, \ldots, u_d) \in I^d$:

$$c_{\mathbf{F}}(u) = \prod_{i=2}^d \frac{e^{K_i(\delta_{(i-1)}^{-1}(u_{i-1})) - K_i(\delta_{(i)}^{-1}(u_i))}}{\delta_{(i-1)} \circ \delta_{(i)}^{-1}(u_i) - u_i} \mathbf{1}_{\{u \in T^{\mathbf{F}}; (\delta_{(1)}^{-1}(u_1), \ldots, \delta_{(d)}^{-1}(u_d)) \in L_{\delta^{\mathbf{F}}}\}} \mathbf{1}_{\{\prod_{i=1}^d \delta_{(i)}' \circ \delta_{(i)}^{-1}(u_i) > 0\}}. \quad (49)$$

Recall the function $\mathbb{J}(\delta)$ defined on the set of multidiagonals by (15) and $C_{\delta^{\mathbf{F}}}$ the copula with density given by (44)–(46).

Proposition 5.1. *Let $\mathbf{F} \in \mathcal{F}_d^0$. The function $c_{\mathbf{F}}$ defined by (49) is the density of the copula $C_{\mathbf{F}} = \mathcal{S}_{\mathbf{F}}^{-1}(C_{\delta^{\mathbf{F}}})$ which belongs to $\mathcal{C}^{OS}(\mathbf{F})$. The entropy of $C_{\mathbf{F}}$ is given by:*

$$H(C_{\mathbf{F}}) = d - 1 - \mathbb{J}(\delta^{\mathbf{F}}). \quad (50)$$

Proof. Since $\mathbf{F} \in \mathcal{F}_d^0$, we have that $\delta^{\mathbf{F}} \in \mathcal{D}^0$. According to Proposition 4.5, $c_{\delta \mathbf{F}}$ defined by (44) is the density of a symmetric copula $C_{\delta \mathbf{F}}$ which belongs to $\mathcal{C}^{\mathrm{sym}}(\mathbf{F}) \cap \mathcal{C}^0$, thanks to Proposition 3.13 and Lemma 3.20. According to Lemma 3.20, formula (35) we get that $c_{\mathbf{F}} = s_{\mathbf{F}}^{-1}(C_{\delta \mathbf{F}})$ is therefore the density of a copula $C_{\mathbf{F}}$ which belongs to $\mathcal{C}^{\mathrm{OS}}(\mathbf{F}) \cap \mathcal{C}^0$. Use (35) and (6) to check (49). To conclude, use (37) and (48) to get (50). $\qquad \square$

Analogously to Lemma 4.2, we have the following restriction on the support of all $F \in \mathcal{L}_d^{\mathrm{OS}}(\mathbf{F}) \cap \mathcal{L}_d^0$. Recall the definition of $\Psi_i^{\mathbf{F}}$ in (29). The proof of the next lemma is similar to the proof of Lemma 4.2 and is left to the reader.

Lemma 5.2. *Let* $\mathbf{F} \in \mathcal{F}_d^0$ *and* $t \in \mathbf{F}_i((\Psi_i^{\mathbf{F}})^c)$ *for some* $2 \leq i \leq d$. *Then we have for all* $F \in \mathcal{L}_d^{\mathrm{OS}}(\mathbf{F}) \cap \mathcal{L}_d^0$ *with density function* f:

$$f(x)\mathbf{1}_{\{x_{i-1} < t < x_i\}} = 0 \qquad \text{for a.e. } x = (x_1, \ldots, x_d) \in S.$$

For $\delta \in \mathcal{D}$, recall the definition of L_δ in (40). Let $L^\delta = L_\delta \cap S$. More generally, for $\mathbf{F} \in \mathcal{F}_d$, we set:

$$L^{\mathbf{F}} = \left\{ x = (x_1, \ldots, x_d) \in S; (x_{i-1}, x_i) \subset \Psi_i^{\mathbf{F}} \text{ for all } 2 \leq i \leq d \right\}. \tag{51}$$

The next lemma establishes the connection between the sets $L_{\delta \mathbf{F}}$ defined by (40) and $L^{\mathbf{F}}$.

Lemma 5.3. *Let* $\mathbf{F} = (\mathbf{F}_1, \ldots, \mathbf{F}_d) \in \mathcal{F}_d^0$ *with density functions* $\mathbf{f}_i$ *for* $1 \leq i \leq d$. *Let* $\delta^{\mathbf{F}} = (\delta_{(1)}, \ldots, \delta_{(d)})$ *given by* (27), $T^{\mathbf{F}}$ *given by* (33) *and* $L_{\delta \mathbf{F}}$ *defined by* (40). *Then for* $\prod_{i=1}^d \mathbf{f}_i(x_i)\,dx_1 \cdots x_d$-*a.e.* $x \in \mathbb{R}^d$ *we have that*

$$\mathbf{1}_{T^{\mathbf{F}}}\big(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_d(x_d)\big)\mathbf{1}_{L_{\delta \mathbf{F}}}\big(\delta_{(1)}^{-1} \circ \mathbf{F}_1(x_1), \ldots, \delta_{(d)}^{-1} \circ \mathbf{F}_d(x_d)\big) = \mathbf{1}_{L^{\mathbf{F}}}(x).$$

Proof. According to (3) and (5), we have $\mathbf{f}_i(t)\,dt$-a.e. that $\mathbf{F}_i^{-1} \circ \mathbf{F}_i(t) = t$. This implies that $\prod_{i=1} \mathbf{f}_i(x_i)\,dx_1 \cdots dx_d$-a.e., $(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_d(x_d))$ belongs to $T^{\mathbf{F}}$ if and only if $x \in S$. Recall the sets $\Psi_i^{\delta^{\mathbf{F}}}$ given by (29). For $\prod_{i=1} \mathbf{f}_i(x_i)\,dx_1 \cdots dx_d$-a.e. $x \in S$, we have:

$$\big(\delta_{(1)}^{-1} \circ \mathbf{F}_1(x_1), \ldots, \delta_{(d)}^{-1} \circ \mathbf{F}_d(x_d)\big) \in L_{\delta \mathbf{F}}$$

$$\iff \quad \big(\delta_{(i-1)}^{-1} \circ \mathbf{F}_{i-1}(x_{i-1}), \delta_{(i)}^{-1} \circ \mathbf{F}_i(x_i)\big) \subset \Psi_i^{\delta^{\mathbf{F}}}, \qquad 2 \leq i \leq d$$

$$\iff \quad \forall t \in \big(\delta_{(i-1)}^{-1} \circ \mathbf{F}_{i-1}(x_{i-1}), \delta_{(i)}^{-1} \circ \mathbf{F}_i(x_i)\big) : \delta_{(i-1)}(t) > \delta_{(i)}(t), \qquad 2 \leq i \leq d$$

$$\iff \quad \forall t \in \big(G \circ \mathbf{F}_{i-1}^{-1} \circ \mathbf{F}_{i-1}(x_{i-1}), G \circ \mathbf{F}_i^{-1} \circ \mathbf{F}_i(x_i)\big) : \mathbf{F}_{i-1} \circ G^{-1}(t) > \mathbf{F}_i \circ G^{-1}(t),$$
$$\qquad 2 \leq i \leq d$$

$$\iff \quad \forall t \in \big(G(x_{i-1}), G(x_i)\big) : \mathbf{F}_{i-1} \circ G^{-1}(t) > \mathbf{F}_i \circ G^{-1}(t), \qquad 2 \leq i \leq d,$$

where the first equivalence comes from the definition of $L_{\delta \mathbf{F}}$, the second from the definition of $\psi_i^{\delta^{\mathbf{F}}}$, the third from (27) and the last from the fact that $\mathbf{f}_i(t)\,dt$-a.e. $\mathbf{F}_i^{-1} \circ \mathbf{F}_i(t) = t$. Consider the

change of variable $s = G^{-1}(t)$. We have by (1):

$$t < G(x_i) \quad \Longleftrightarrow \quad G^{-1}(t) < x_i \quad \Longleftrightarrow \quad s < x_i.$$

Since $x_{i-1} \in I_g(\mathbf{F}_{i-1})\, \mathbf{f}_{i-1}(x_{i-1})\, dx_{i-1}$-a.e., we get $x_{i-1} \in I_g(G)$ and by (5):

$$G(x_{i-1}) < t \quad \Longleftrightarrow \quad x_{i-1} < G^{-1}(t) \quad \Longleftrightarrow \quad x_{i-1} < s.$$

Therefore, we deduce that $\prod_{i=1} \mathbf{f}_i(x_i)\, dx_1 \cdots dx_d$-a.e. $x \in S$:

$$\left(\delta_{(1)}^{-1} \circ \mathbf{F}_1(x_1), \ldots, \delta_{(d)}^{-1} \circ \mathbf{F}_d(x_d) \right) \in L_{\delta \mathbf{F}}$$

$$\Longleftrightarrow \quad \forall s \in (x_{i-1}, x_i) : \mathbf{F}_{-1}(s) > \mathbf{F}_i(s), \qquad 2 \leq i \leq d$$

$$\Longleftrightarrow \quad x \in L^{\mathbf{F}}. \qquad\qquad\qquad \square$$

Using Proposition 3.22, we check that the copula $C_{\mathbf{F}}$ maximizes the entropy over the set $\mathcal{C}^{\mathrm{OS}}(\mathbf{F})$. For $\mathbf{F} \in \mathcal{F}_d^0$, we define the c.d.f. $F_{\mathbf{F}}$ as, for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$:

$$F_{\mathbf{F}}(x) = C_{\mathbf{F}}\big(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_d(x_d) \big). \tag{52}$$

Let $\mathbf{f}_i$ denote the density function of $\mathbf{F}_i$ when it exists. Let us further note for $2 \leq i \leq d, t \in \mathbb{R}$:

$$\ell_i(t) = \frac{\mathbf{f}_i(t)}{\mathbf{F}_{i-1}(t) - \mathbf{F}_i(t)}. \tag{53}$$

When the densities $\mathbf{f}_i$ exist for all $1 \leq i \leq d$, we define the function $f_{\mathbf{F}}$ for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ as:

$$f_{\mathbf{F}}(x) = \mathbf{f}_1(x_1) \prod_{i=2}^{d} \ell_i(x_i) \exp\left(-\int_{x_{i-1}}^{x_i} \ell_i(s)\, ds \right) \mathbf{1}_{L^{\mathbf{F}}}(x), \tag{54}$$

with $L^{\mathbf{F}}$ given by (51). The next theorem asserts that the c.d.f. $F_{\mathbf{F}}$ maximizes the entropy over the set $\mathcal{L}_d^{\mathrm{OS}}(\mathbf{F})$ and that its density is $f_{\mathbf{F}}$. Recall $\mathbb{J}$ defined by (15). Recall that h is an arbitrary probability density on $\mathbb{R}$.

Theorem 5.4. *Let* $\mathbf{F} = (\mathbf{F}_i, 1 \leq i \leq d) \in \mathcal{F}_d$.

(a) *If there exists* $1 \leq i \leq d$ *such that* $H_h(\mathbf{F}_i) = -\infty$, *or if* $\mathbb{J}(\mathbf{F}) = +\infty$, *then we have*

$$\max_{F \in \mathcal{L}_d^{\mathrm{OS}}(\mathbf{F})} H_h(F) = -\infty.$$

(b) *If* $H_h(\mathbf{F}_i) > -\infty$ *for all* $1 \leq i \leq d$, *and* $\mathbb{J}(\mathbf{F}) < +\infty$, *then we have* $\mathbf{F} \in \mathcal{F}_d^0$, $\max_{F \in \mathcal{L}_d^{\mathrm{OS}}(\mathbf{F})} H_h(F) > -\infty$, *and* $F_{\mathbf{F}}$ *defined in* (52) *is the unique c.d.f. in* $\mathcal{L}_d^{\mathrm{OS}}(\mathbf{F})$ *such*

that $H_h(F_{\mathbf{F}}) = \max_{F \in \mathcal{L}_d^{OS}(\mathbf{F})} H_h(F)$. Furthermore, the density function of $F_{\mathbf{F}}$ exists, and is given by $f_{\mathbf{F}}$ defined in (54). *We also have*:

$$H_h(F_{\mathbf{F}}) = d - 1 + \sum_{i=1}^{d} H_h(\mathbf{F}_i) - \mathbb{J}(\mathbf{F}).$$

Proof. The proof of case (a) is postponed to Section 6.7.

We shall assume that $H_h(\mathbf{F}_i) > -\infty$ for all $1 \le i \le d$ and $\mathbb{J}(\delta^{\mathbf{F}}) < +\infty$. This implies that the densities $\mathbf{f}_i$ of $\mathbf{F}_i$ exist for $1 \le i \le d$ and, thanks to Lemma 3.18, that $\mathbf{F} \in \mathcal{F}_d^0$. Let $F_{\mathbf{F}}$ be defined by (52), that is the c.d.f. with copula $C_{\mathbf{F}}$ from Proposition 5.1 and one-dimensional marginal c.d.f.s $\mathbf{F}$. Thanks to Proposition 5.1, we have $F_{\mathbf{F}} \in \mathcal{L}_d^{OS}(\mathbf{F})$.

We deduce from (14), Propositions 3.13 and 3.22, Theorem 4.7 case (b) and Proposition 5.1 that $F_{\mathbf{F}}$ is the only c.d.f. such that $H_h(F_{\mathbf{F}}) = \max_{F \in \mathcal{L}_d^{OS}(\mathbf{F})} H_h(F)$. We deduce from (14), (50) and Lemma 3.15 that:

$$H_h(F_{\mathbf{F}}) = d - 1 + \sum_{i=1}^{d} H_h(\mathbf{F}_i) - \mathbb{J}(\mathbf{F}).$$

Since the copula $C_{\mathbf{F}}$ is absolutely continuous with density $c_{\mathbf{F}}$ given in (49), we deduce from (52) that $F_{\mathbf{F}}$ has density $f_{\mathbf{F}}$ given by, for a.e. $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$:

$$f_{\mathbf{F}}(x) = c_{\mathbf{F}}\big(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_d(x_d)\big) \prod_{i=1}^{d} \mathbf{f}_i(x_i). \tag{55}$$

Recall the expression (49) of $c_{\mathbf{F}}$ as well as K_i defined by (46), for $1 \le i \le d$. Using the change of variable $s = G^{-1}(t)$ and (23), we get (similarly to the proof of Lemma 3.15):

$$K_i \circ \delta_{(i)}^{-1} \circ \mathbf{F}_i(x_i) - K_i \circ \delta_{(i-1)}^{-1} \circ \mathbf{F}_{i-1}(x_{i-1}) = \int_{\mathbf{F}_{i-1}^{-1} \circ \mathbf{F}_{i-1}(x_{i-1})}^{\mathbf{F}_i^{-1} \circ \mathbf{F}_i(x_i)} \ell_i(s) \, ds. \tag{56}$$

Using (23), we also get:

$$\frac{\mathbf{f}_i(x_i)}{\delta_{(i-1)} \circ \delta_{(i)}^{-1} \circ \mathbf{F}_i(x_i) - \mathbf{F}_i(x_i)} = \frac{\mathbf{f}_i(x_i)}{\mathbf{F}_{i-1} \circ \mathbf{F}_i^{-1} \circ \mathbf{F}_i(x_i) - \mathbf{F}_i(x_i)}. \tag{57}$$

According to (5), we have $\mathbf{f}_i(t) \, dt$-a.e. that $\mathbf{F}_i^{-1} \circ \mathbf{F}_i(t) = t$. For $1 \le i \le d$, we have from (5) that $\prod_{i=1}^{d} \mathbf{f}_i(x_i) \, dx_1 \cdots dx_d$-a.e.:

$$\mathbf{1}_{\{\prod_{i=1}^{d} \delta'_{(i)} \circ \delta_{(i)}^{-1}(\mathbf{F}_i(x_i)) > 0\}} = 1. \tag{58}$$

We deduce from (49), (55), (56), (57), (58) and Lemma 5.3 that a.e. for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$:

$$f_{\mathbf{F}}(x) = \mathbf{f}_1(x_1) \prod_{i=2}^{d} \ell_i(x_i) e^{-\int_{x_{i-1}}^{x_i} \ell_i(s) \, ds} \mathbf{1}_{L^{\mathbf{F}}}(x). \qquad \square$$

Remark 5.5. We deduce from the proof of Theorem 5.4 case (b) and Proposition 5.1, that if $\mathbf{F} = (\mathbf{F}_i, 1 \leq i \leq d) \in \mathcal{F}_d^0$, then $f_\mathbf{F}$ defined by (54) is a probability density function on $S \subset \mathbb{R}^d$.

Remark 5.6. The density $f_\mathbf{F}$ has a product form on $L^\mathbf{F}$, that is it can be written as, for a.e. $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$:

$$f_\mathbf{F}(x) = \prod_{i=1}^d p_i(x_i) \mathbf{1}_{L^\mathbf{F}}(x), \tag{59}$$

where the functions $(p_i, 1 \leq i \leq d)$ are measurable and non-negative.

In addition to Remark 5.6, the next corollary asserts that $F_\mathbf{F}$ is the only element of $\mathcal{L}_d^{OS}(\mathbf{F})$, whose density has a product form.

Corollary 5.7. *Let $\mathbf{F} \in \mathcal{F}_d^0$. Let $F \in \mathcal{L}_d^{OS}(\mathbf{F})$ be an absolutely continuous c.d.f. with density f given by, a.e. for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$: $f(x) = \prod_{i=1}^d h_i(x_i) \mathbf{1}_{L^\mathbf{F}}(x)$, with h_i, $1 \leq i \leq d$ some measurable non-negative functions on $\mathbb{R}$. Then we have $F = F_\mathbf{F}$ on $\mathbb{R}^d$.*

Proof. Let $X = (X_1, \ldots, X_d)$ be an order statistic with c.d.f. F, and F^{sym} the c.d.f. of X_Π given by (16), with Π uniform on $\mathcal{S}_d$ and independent of X. Then the c.d.f. F^{sym} is also absolutely continuous, and its density f^{sym} is given by:

$$f^{\mathrm{sym}}(x) = \frac{1}{d!} \prod_{i=1}^d h_i(x_{(i)}) \mathbf{1}_{L^\mathbf{F}}(x^{OS}), \tag{60}$$

where x^{OS} is the ordered vector of x. The one-dimensional marginal c.d.f.s of X_Π are all equal to G given by (17). Let $C \in \mathcal{C}^{OS}(\mathbf{F}) \cap \mathcal{C}^0$ denote the copula of F. Then according to (7), the copula $S_\mathbf{F}(C)$ of X_Π is given by, for a.e. $u = (u_1, \ldots, u_d) \in I^d$:

$$S_\mathbf{F}(C)(u) = F^{\mathrm{sym}}(G^{-1}(u)).$$

Therefore its density $s_\mathbf{F}(C)$ can be expressed as:

$$s_\mathbf{F}(C)(u) = \frac{f^{\mathrm{sym}}(G^{-1}(u))}{\prod_{i=1}^d g \circ G^{-1}(u_i)} \prod_{i=1}^d \mathbf{1}_{\{g \circ G^{-1}(u_i) > 0\}}$$

$$= \frac{1}{d!} \mathbf{1}_{L^\mathbf{F}}(G^{-1}(u^{OS})) \prod_{i=1}^d \frac{h_i \circ G^{-1}(u_{(i)})}{g \circ G^{-1}(u_{(i)})} \mathbf{1}_{\{g \circ G^{-1}(u_{(i)}) > 0\}},$$

where g is the density of G. Notice that for $x = (x_1, \ldots, x_d) \in S$, we have a.e.:

$$\mathbf{1}_{\{\prod_{i=1}^d h_i(x_i) > 0\}} \leq \mathbf{1}_{\{\prod_{i=1}^d f_i(x_i) > 0\}},$$

with f_i the density of X_i. Therefore, by Lemma 5.3 and since G is continuous, we have that $\prod_{i=1}^d h_i \circ G^{-1}(u_{(i)}) \, du_1 \cdots du_d$-a.e.:

$$\mathbf{1}_{L^{\mathrm{F}}}\big(G^{-1}\big(u^{\mathrm{OS}}\big)\big) = \mathbf{1}_{L_{\delta \mathrm{F}}}\big(\delta_{(1)}^{-1} \circ \delta_{(1)}(u_{(1)}), \ldots, \delta_{(d)}^{-1} \circ \delta_{(d)}(u_{(d)})\big).$$

By Lemma 3.12, $S_{\mathbf{F}}(C)$ belongs to $\mathcal{C}_{\delta \mathrm{F}}^0$ and thus $s_{\mathbf{F}}(C) = 0$ a.e. on $Z_{\delta \mathrm{F}}$ defined by (38). Then use (3) and (5) to get that $\delta_{(i)}^{-1} \circ \delta_{(i)}(u_{(i)}) = u_{(i)}$ a.e. on $Z_{\delta \mathrm{F}}^c$. This gives:

$$\mathbf{1}_{L^{\mathrm{F}}}\big(G^{-1}\big(u^{\mathrm{OS}}\big)\big) = \mathbf{1}_{L_{\delta \mathrm{F}}}\big(u^{\mathrm{OS}}\big) = \mathbf{1}_{L_{\delta \mathrm{F}}}(u)$$

that is $s_{\mathbf{F}}(C)$ is of the form $s_{\mathbf{F}}(C)(u) = (1/d!)\mathbf{1}_{L_{\delta \mathrm{F}}}(u) \prod_{i=1}^d \bar{h}_i(u_{(i)})$ for some measurable non-negative functions $(\bar{h}_i, 1 \le i \le d)$. Then, thanks to Proposition 4.6, we get that $S_{\mathbf{F}}(C) = C_{\delta \mathrm{F}}$. Then, use Proposition 5.1 to get that $F = F_{\mathbf{F}}$. $\qquad\square$

Example 5.8. We consider the following example. Let $+\infty > \lambda_1 > \cdots > \lambda_d > 0$ and for $1 \le i \le d$ let $\mathbf{F}_i$ be the c.d.f. of the exponential distribution with mean $1/\lambda_i$ and density $\mathbf{f}_i(t) = \lambda_i e^{-\lambda_i t}\mathbf{1}_{\{t>0\}}$. Notice that $\mathbf{F}_{i-1} > \mathbf{F}_i$ on $(0, +\infty)$, so that $L^{\mathbf{F}} = \{(x_1, \ldots, x_d) \in \mathbb{R}^d; 0 \le x_1 \le \cdots \le x_d\}$. It is easy to check that $\mathbb{J}(\mathbf{F}) < +\infty$ with $\mathbf{F} = (\mathbf{F}_i, 1 \le i \le d)$. Elementary computations yield that the maximum entropy density of the order statistic $(X_1, \ldots, X_d)$, where X_i has distribution $\mathbf{F}_i$, is given by:

$$f_{\mathbf{F}}(x_1, \ldots, x_d) = \mathbf{1}_{L^{\mathbf{F}}}(x)\lambda_1 e^{-\Delta_2 x_1}\big(1 - e^{-\Delta_2 x_1}\big)^{\lambda_2/\Delta_2} \prod_{i=2}^d \lambda_i e^{-\Delta_{i+1} x_i} \frac{(1 - e^{-\Delta_{i+1} x_i})^{\lambda_{i+1}/\Delta_{i+1}}}{(1 - e^{-\Delta_i x_i})^{\lambda_{i-1}/\Delta_i}},$$

where $\Delta_i = \lambda_{i-1} - \lambda_i$ for $1 \le i \le d+1$ and $\lambda_{d+1} = 0$.

In the particular case $\lambda_i = (d - i + 1)\lambda$ for some $\lambda > 0$, we get:

$$f_{\mathbf{F}}(x_1, \ldots, x_d) = \mathbf{1}_{L^{\mathbf{F}}}(x)\, d!\lambda^d e^{-\lambda x_1}\big(1 - e^{-\lambda x_1}\big)^{d-1} \prod_{i=2}^d \frac{e^{-\lambda x_i}}{(1 - e^{-\lambda x_i})^2}.$$

By considering the change of variable $u_i = 1 - e^{-\lambda x_i}$, we get the following result. For $1 \le i \le d$ let $\mathbf{F}_i$ be the c.d.f. of the $\beta(1, d - i + 1)$ distribution with density $\mathbf{f}_i(t) = (d - i + 1)(1 - t)^{d-i}\mathbf{1}_{(0,1)}(t)$. Notice that $\mathbf{F}_{i-1} > \mathbf{F}_i$ on $(0, 1)$. The maximum entropy density of the order statistic $(U_1, \ldots, U_d)$, where U_i has distribution $\mathbf{F}_i$, is given by:

$$f_{\mathbf{F}}(u_1, \ldots, u_d) = \mathbf{1}_{\{0 < u_1 < \cdots < u_d < 1\}}\, d!\, u_1^{d-1} \prod_{i=2}^d \frac{1}{u_i^2}.$$

Elementary computations give $H(F_{\mathbf{F}}) = -\log(d!) + 2d - (d + 1)\sum_{i=1}^d (1/i)$.

6. Proofs

6.1. Preliminary notations for the optimization problem

Recall notations from Sections 2 and 3. In particular, if $u = (u_1, \ldots, u_d) \in I^d$ then $u^{OS} = (u_{(1)}, \ldots, u_{(d)})$ denote the ordered vector of u.

In order to apply the technique established in [5], we introduce the linear functional $\mathcal{A} = (\mathcal{A}_i, 1 \leq i \leq 2d) : L^1(I^d) \to L^1(I)^{2d}$ as, for $f \in L^1(I^d)$ and $r \in I$:

$$\mathcal{A}_i(f)(r) = \int_{I^d} f(u) \mathbf{1}_{\{u_i \leq r\}} \, du \quad \text{and} \quad \mathcal{A}_{d+i}(f)(r) = \int_{I^d} f(u) \mathbf{1}_{\{u_{(i)} \leq r\}} \, du \qquad \text{for } 1 \leq i \leq d.$$

Let $\delta = (\delta_{(i)}, 1 \leq i \leq d) \in \mathcal{D}^0$ be a multidiagonal, see Definition 3.6. We set $b^\delta = (b_i, 1 \leq i \leq 2d)$ given by $b_i = \mathrm{id}_I$ the identity function on I and $b_{d+i} = \delta_{(i)}$, for $1 \leq i \leq d$. If, for $c \in L^1(I^d)$, we have $\mathcal{A}_i(c) = b_i, 1 \leq i \leq d$ and $c \geq 0$ a.e., then we deduce that c is the density of an absolutely continuous copula, say C. If we further have $\mathcal{A}_{d+i}(c) = b_{d+i}$, for $1 \leq i \leq d$, then δ is the multidiagonal of C.

Lemma 6.1. *Let $\delta \in \mathcal{D}^0$ and $b^\delta = (b_i, 1 \leq i \leq 2d)$. If $c \in L^1(I^d)$ is non-negative, symmetric and satisfies $\mathcal{A}_{d+i}(c) = b_{d+i}$ for $1 \leq i \leq d$, then c is the density of a copula with multidiagonal δ.*

Proof. The symmetry and non-negativity of c as well as the condition $\mathcal{A}_{d+1}(c)(1) = \int_{I^d} c = b_{d+1}(1) = 1$ ensures that c is a density function of an exchangeable random vector $V = (V_1, \ldots, V_d)$ on I^d. Recall $V^{OS} = (V_{(1)}, \ldots, V_{(d)})$ denotes the corresponding order statistics. By symmetry, the lemma is proved as soon as we check that $\mathcal{A}_1(c) = b_1$. We have for $r \in I$:

$$\mathcal{A}_1(c)(r) = \mathbb{P}(V_1 \leq r) = \sum_{i=1}^{d} \mathbb{P}(V_{(i)} \leq r | V_1 = V_{(i)}) \mathbb{P}(V_1 = V_{(i)}) = \sum_{i=1}^{d} \delta_{(i)}(r) \frac{1}{d} = r,$$

where we used the exchangeability of V and the definition of $\delta_{(i)}$ for the third equality, and (21) for the last. This gives $\mathcal{A}_1(c) = b_1$. $\square$

6.2. Proof of Proposition 4.5

Let $\delta \in \mathcal{D}^0$. Lemma 3.16 implies that $\delta_{(i)}((\Psi_i^\delta)^c)$ has zero Lebesgue measure for all $2 \leq i \leq d$ with Ψ_i^δ given by (29). By construction, the function c_δ defined by (44) is non-negative, symmetric and well defined a.e. on I^d. Recall the notation $(g_i^{(j)}, d_i^{(j)})$ used in (39) and the definition (45) of the functions a_i. We define the functions B_i on I as, for $1 \leq i \leq d+1$, $t \in (g_i^{(j)}, d_i^{(j)})$ (with the conventions $\Psi_1^\delta = (0, d_1)$, $\Psi_{d+1}^\delta = (g_{d+1}, 1)$):

$$B_{d+1}(t) = 1 \quad \text{and} \quad B_i(t) = \int_t^{d_i^{(j)}} a_i(s) B_{i+1}(s) \, ds \qquad \text{for } 1 \leq i \leq d. \tag{61}$$

For $t \in (\Psi_i^\delta)^c$, we set $B_i(t) = 0$. Recall K_i defined in (46) for $1 \le i \le d+1$ with the convention $K_{d+1} = 0$. We show that B_i can be simply expressed by K_i on Ψ_i^δ.

Lemma 6.2. *Let $1 \le i \le d+1$ and $t \in \Psi_i^\delta$. Then we have:*

$$B_i(t) = \exp\bigl(-K_i(t)\bigr). \tag{62}$$

Proof. For $i = d+1$, the result is trivial. We proceed by induction on i. We suppose that $B_{i+1}(t) = \exp(-K_{i+1}(t))$ holds for some $1 \le i \le d$, and all $t \in \psi_{i+1}^\delta$. We have for $t \in (g_i^{(j)}, d_i^{(j)})$:

$$
\begin{aligned}
B_i(t) &= \int_t^{d_i^{(j)}} a_i(s) B_{i+1}(s) \, ds \\
&= \int_t^{d_i^{(j)}} K_i'(s) e^{K_{i+1}(s) - K_i(s)} \mathbf{1}_{\Psi_i^\delta \cap \Psi_{i+1}^\delta}(s) B_{i+1}(s) \, ds \\
&= \int_t^{d_i^{(j)}} K_i'(s) e^{-K_i(s)} \mathbf{1}_{\Psi_i^\delta \cap \Psi_{i+1}^\delta}(s) \, ds \\
&= \int_t^{d_i^{(j)}} K_i'(s) e^{-K_i(s)} \, ds \\
&= \exp\bigl(-K_i(t)\bigr),
\end{aligned}
$$

where we used the definition of a_i given by (45) for the second equality, the induction hypothesis for the third equality, $(t, d_i^{(j)}) \subset \Psi_i^\delta$ and Lemma 3.16 for the fourth equality, and finally Remark 4.4 for the fifth equality. This ends the induction. $\square$

Similarly, we define the functions E_i on I as, for $0 \le i \le d$ as for $t \in (g_{i+1}^{(j)}, d_{i+1}^{(j)})$:

$$E_0(t) = 1, \quad \text{and} \quad E_i(t) = \int_{g_{i+1}^{(j)}}^t a_i(s) E_{i-1}(s) \, ds \qquad \text{for } 1 \le i \le d. \tag{63}$$

For $t \in (\Psi_{i+1}^\delta)^c$ we set $E_i(t) = 0$. The next lemma gives a simple formula for E_i on Ψ_{i+1}^δ.

Lemma 6.3. *Let $0 \le i \le d$ and $t \in \Psi_{i+1}^\delta$. Then we have:*

$$E_i(t) = \bigl(\delta_{(i)}(t) - \delta_{(i+1)}(t)\bigr) \exp\bigl(K_{i+1}(t)\bigr). \tag{64}$$

Proof. For $i = 0$, the result is clear thanks to the convention $\delta_{(0)} = 1$ and (47). We proceed by induction on i. We suppose that $E_{i-1}(t) = (\delta_{(i-1)}(t) - \delta_{(i)}(t)) \exp(K_i(t))$ holds for some $1 \le i \le d$, and all $t \in \psi_i^\delta$. Let us denote $h_i = \delta_{(i-1)} - \delta_{(i)}$. Before computing $E_i(t)$ for $t \in (g_{i+1}^{(j)}, d_{i+1}^{(j)})$,

we give an alternative expression for $\exp(K_i(s))$ for $s \in (g_{i+1}^{(j)}, t)$:

$$
\begin{aligned}
\mathrm{e}^{K_{i+1}(s)} &= \exp\!\left(-\int_{m_{i+1}^{(j)}}^{t} \frac{h_{i+1}'(u)}{h_{i+1}(u)} + \int_{s}^{t} \frac{h_{i+1}'(u)}{h_{i+1}(u)} + \int_{m_{i+1}^{(j)}}^{s} \frac{\delta_{(i)}'(u)}{h_{i+1}(u)}\, du\right) \\
&= \frac{h_{i+1}(t)}{h_{i+1}(s)} \exp\!\left(-\int_{m_{i+1}^{(j)}}^{t} \frac{h_{i+1}'(u)}{h_{i+1}(u)} + \int_{m_{i+1}^{(j)}}^{s} \frac{\delta_{(i)}'(u)}{h_{i+1}(u)}\, du\right).
\end{aligned}
\tag{65}
$$

Then we have for $t \in (g_{i+1}^{(j)}, d_{i+1}^{(j)})$:

$$
\begin{aligned}
E_i(t) &= \int_{g_{i+1}^{(j)}}^{t} a_i(s) E_{i-1}(s)\, ds \\
&= \int_{g_{i+1}^{(j)}}^{t} K_i'(s) \mathrm{e}^{K_{i+1}(s)-K_i(s)} \mathbf{1}_{\Psi_i^\delta \cap \Psi_{i+1}^\delta}(s) E_{i-1}(s)\, ds \\
&= \int_{g_{i+1}^{(j)}}^{t} K_i'(s) \mathrm{e}^{K_{i+1}(s)} h_i(s) \mathbf{1}_{\Psi_i^\delta \cap \Psi_{i+1}^\delta}(s)\, ds \\
&= \int_{g_{i+1}^{(j)}}^{t} \delta_{(i)}'(s) \mathrm{e}^{K_{i+1}(s)} \mathbf{1}_{\Psi_i^\delta \cap \Psi_{i+1}^\delta}(s)\, ds \\
&= h_{i+1}(t) \exp\!\left(-\int_{m_{i+1}^{(j)}}^{t} \frac{h_{i+1}'(u)}{h_{i+1}(u)}\, du\right) \int_{g_{i+1}^{(j)}}^{t} \left(\frac{\delta_{(i)}'(s)}{h_{i+1}(s)} \exp\!\left(\int_{m_{i+1}^{(j)}}^{s} \frac{\delta_{(i)}'(u)}{h_{i+1}(u)}\, du\right)\right) ds \\
&= h_{i+1}(t) \exp\!\left(-\int_{m_{i+1}^{(j)}}^{t} \frac{h_{i+1}'(u) - \delta_{(i)}'(u)}{h_{i+1}(u)}\, du\right) \\
&= h_{i+1}(t) \exp\!\left(K_{i+1}(t)\right),
\end{aligned}
$$

where we used the definition of a_i given by (45) for the second equality, the induction hypothesis for the third equality, Lemma 3.16 and (65) for the fifth equality, and for the seventh equality we use that, for $t \in (g_{i+1}^{(j)}, m_{i+1}^{(j)})$ (similarly to Remark 4.4):

$$
\int_{m_{i+1}^{(j)}}^{t} \frac{\delta_{(i)}'(s)}{h_{i+1}(s)}\, ds \le \int_{m_{i+1}^{(j)}}^{t} \frac{\delta_{(i)}'(s)}{\delta_{(i)}(s) - \delta_{(i+1)}(g_{i+1}^{(j)})}\, ds = \log\!\left(\frac{\delta_{(i)}(t) - \delta_{(i+1)}(g_{i+1}^{(j)})}{\delta_{(i)}(m_{i+1}^{(j)}) - \delta_{(i+1)}(g_{i+1}^{(j)})}\right),
$$

giving $\lim_{t \searrow g_{i+1}^{(j)}} \int_{m_{i+1}^{(j)}}^{t} \frac{\delta_{(i)}'(s)}{h_{i+1}(s)}\, ds = -\infty$.   □

The following lemma justifies the introduction of the functions B_i, E_i.

Lemma 6.4. *We have with $u_{(0)} = 0$ for $1 \le i \le d$, $t \in \Psi_i^\delta$:*

$$\int_{I^d} c_\delta(u) \mathbf{1}_{\{u_{(i-1)} \le t \le u_{(i)}\}} \, du = B_i(t) E_{i-1}(t). \tag{66}$$

Proof. The definition (61) of B_i for $1 \le i \le d$ gives that for $t \in I$:

$$B_i(t) = \int a_i(r_i) a_{i+1}(r_{i+1}) \cdots a_d(r_d) \mathbf{1}_{\{t \le r_i \le r_{i+1} \le \cdots \le r_d \le 1\}} \mathbf{1}_{\{[t,r_i) \subset \Psi_i^\delta\}} \prod_{j=i}^{d-1} \mathbf{1}_{\{(r_j, r_{j+1}) \subset \Psi_{j+1}^\delta\}} \, dr,$$

with $r = (r_i, r_{i+1}, \ldots, r_d) \in I^{d-i+1}$. Similarly, we have for $1 \le i \le d$, $t \in I$ that $E_{i-1}(t)$ is equal to:

$$\int a_1(q_1) a_2(q_2) \cdots a_{i-1}(q_{i-1}) \mathbf{1}_{\{0 \le q_1 \le q_2 \le \cdots \le q_{i-1} \le t\}} \mathbf{1}_{\{(q_{i-1}, t] \subset \Psi_i^\delta\}} \prod_{j=1}^{i-2} \mathbf{1}_{\{(q_j, q_{j+1}) \subset \Psi_{j+1}^\delta\}} \, dq,$$

with $q = (q_1, q_2, \ldots, q_{i-1}) \in I^{i-1}$. Multiplying $B_i(t)$ with $E_{i-1}(t)$ gives:

$$B_i(t) E_{i-1}(t) = \int_\Delta \prod_{j=1}^d a_j(u_j) \mathbf{1}_{\{u_{i-1} \le t \le u_i\}} \prod_{j=1}^{d-1} \mathbf{1}_{\{(u_j, u_{j+1}) \subset \Psi_{j+1}^\delta\}} \, du$$

$$= \int_\Delta \prod_{j=1}^d a_j(u_j) \mathbf{1}_{\{u_{i-1} \le t \le u_i\}} \mathbf{1}_{L^\delta}(u) \, du$$

$$= d! \int_\Delta c_\delta(u) \mathbf{1}_{\{u_{i-1} \le t \le u_i\}} \, du$$

$$= \int_{I^d} c_\delta(u) \mathbf{1}_{\{u_{(i-1)} \le t \le u_{(i)}\}} \, du,$$

where we used the symmetry of c_δ for the fourth equality. □

Lemma 6.4 with $i = 1$ ensures that $\int_{I^d} c_\delta(u) \, du = \lim_{t \searrow 0} B_1(t) E_0(t) = 1$, that is c_δ a probability density function on I^d. Now we compute $\mathcal{A}_{d+1}(c_\delta)$. We have, for $t \in \Psi_i^\delta$:

$$\mathcal{A}_{d+1}(c_\delta)(t) = \int_{I^d} c_\delta(u) \mathbf{1}_{\{u_{(1)} \le t\}} \, du$$

$$= 1 - \int_{I^d} c_\delta(u) \mathbf{1}_{\{u_{(1)} \ge t\}} \, du = 1 - B_1(t) E_0(t)$$

$$= \delta_{(1)}(t),$$

where we used Lemma 6.4 with $i = 1$ for the third equality, then (62) and (47) for the fourth equality. By continuity this gives $\mathcal{A}_{d+1}(c_\delta) = \delta_{(1)}$ on I. For $2 \le i \le d$, we have by induction

for $t \in \Psi_i^\delta$:

$$\mathcal{A}_{d+i}(c_\delta)(t) = \int_{I^d} c_\delta(u) \mathbf{1}_{\{u_{(i)} \leq t\}} \, du$$

$$= \int_{I^d} c_\delta(u) \mathbf{1}_{\{u_{(i-1)} \leq t\}} \, du - \int_{I^d} c_\delta(u) \mathbf{1}_{\{u_{(i-1)} \leq t \leq u_{(i)}\}} \, du$$

$$= \mathcal{A}_{d+i-1}(c_\delta)(t) - B_i(t) E_{i-1}(t)$$

$$= \delta_{(i-1)}(t) - \left(\delta_{(i-1)}(t) - \delta_{(i)}(t) \right)$$

$$= \delta_{(i)}(t),$$

where we used the induction and Lemma 6.4 for the third equality, as well as (62) and (64) for the fourth. By continuity, we obtain $\mathcal{A}_{d+i}(c_\delta) = \delta_{(i)}$ on I. Then use Lemma 6.1 to get that c_δ is the density of a (symmetric) copula, say C_δ, with multidiagonal δ.

To conclude, we compute the entropy $H(C_\delta) = - \int_{I^d} c_\delta \log(c_\delta)$.

Lemma 6.5. *We have*:

$$H(C_\delta) = \log(d!) + \sum_{i=1}^{d} H(\delta_{(i)}) + (d-1) - \mathbb{J}(\delta).$$

Proof. Recall that for $u \in L_\delta$:

$$\log(c_\delta(u)) = -\log(d!) + \sum_{i=1}^{d} \log(\delta'_{(i)}(u_{(i)})) - \sum_{i=2}^{d} \log(\delta_{(i-1)}(u_{(i)}) - \delta_{(i)}(u_{(i)}))$$

$$- \sum_{i=2}^{d} \left(K_i(u_{(i)}) - K_i(u_{(i-1)}) \right),$$

where we used (47) to express $a_1 = \delta'_{(1)} e^{K_2}$ a.e., so that the sums in the last two terms start at $i = 2$.

We first show that the function $u \mapsto c_\delta(u) \log(\delta'_{(i)}(u_{(i)}))$ belongs to $L^1(I^d)$ for all $1 \leq i \leq d$. Since $\mathcal{A}_{d+i}(c_\delta) = \delta_{(i)}$, we deduce that for $1 \leq i \leq d$ and any measurable non-negative function h defined on I:

$$\int_{I^d} c_\delta(u) h(u_{(i)}) \, du = \int_I \delta'_{(i)}(t) h(t) \, dt. \tag{67}$$

In particular, we get:

$$\int_{I^d} c_\delta(u) \left| \log(\delta'_{(i)}(u_{(i)})) \right| du = \int_I \delta'_{(i)}(t) \left| \log(\delta'_{(i)}(t)) \right| dt,$$

which is finite thanks to Remark 3.8. Therefore, the function $u \mapsto c_\delta(u) \log(\delta'_{(i)}(u_{(i)}))$ is indeed in $L^1(I^d)$, and its integral $J_{1,i}$ is given by:

$$J_{1,i} = \int_{I^d} c_\delta(u) \log(\delta'_{(i)}(u_{(i)})) \, du = \int_I \delta'_{(i)}(t) \log(\delta'_{(i)}(t)) \, dt = -H(\delta_{(i)}).$$

We proceed by showing that $u \mapsto c_\delta(u)(K_i(u_{(i)}) - K_i(u_{(i-1)}))$ belongs to $L^1(I^d)$ for $2 \le i \le d$. Since this is a non-negative function, a direct calculation of its integral $J_{2,i}$ gives:

$$J_{2,i} = d! \int_\Delta c_\delta(u)(K_i(u_i) - K_i(u_{i-1})) \, du$$

$$= \int_{I^2} (E_{i-2}a_{i-1})(u_{i-1})(K_i(u_i) - K_i(u_{i-1}))(a_i B_{i+1})(u_i) \mathbf{1}_{\{u_{i-1} \le u_i, (u_{i-1}, u_i) \subset \Psi_i^\delta\}} \, du_{i-1} \, du_i,$$

where we used the symmetry of c_δ for the first equality; the definition of the functions B_i and E_i given by (61) and (63) for the second equality. Using (45), (62), (64) and Lemma 3.16, we have:

$$E_{i-2}a_{i-1} = \delta'_{(i-1)} e^{K_i} \mathbf{1}_{\Psi_{i-1}^\delta \cap \Psi_i^\delta} = \delta'_{(i-1)} e^{K_i} \quad \text{and} \quad a_i B_{i+1} = K_i' e^{-K_i} \mathbf{1}_{\Psi_i^\delta \cap \Psi_{i+1}^\delta} = K_i' e^{-K_i}.$$

Therefore, we have:

$$J_{2,i} = \sum_{j \in J_i} \int_{g_i^{(j)}}^{d_i^{(j)}} \delta'_{(i-1)}(u_{i-1}) \left(\int_{u_{i-1}}^{d_i^{(j)}} K_i'(u_i)(K_i(u_i) - K_i(u_{i-1})) e^{K_i(u_{i-1}) - K_i(u_i)} \, du_i \right) du_{i-1}$$

$$= \sum_{j \in J_i} \int_{g_i^{(j)}}^{d_i^{(j)}} \delta'_{(i-1)}(u_{i-1}) \left(\int_0^{+\infty} s e^{-s} \, ds \right) du_{i-1}$$

$$= \sum_{j \in J_i} \int_{g_i^{(j)}}^{d_i^{(j)}} \delta'_{(i-1)}(u_{i-1}) \, du_{i-1}$$

$$= 1,$$

where we applied the change of variable $s = K_i(u_i) - K_i(u_{i-1})$ and Remark 4.4 for the fourth equality; finally Lemma 3.16 for the sixth equality.

Let us define $J_{3,i}$, $2 \le i \le d$ as:

$$J_{3,i} = -\int_{I^d} c_\delta(u) \log(\delta_{(i-1)}(u_{(i)}) - \delta_{(i)}(u_{(i)})) \, du.$$

Notice that the integrand is non-positive a.e., since for $t \in I$, $2 \le i \le d$, we have $\delta_{(i-1)}(t) - \delta_{(i)}(t) \le 1$. Therefore, we get by (67):

$$J_{3,i} = \int_{I^d} c_\delta(u) |\log(\delta_{(i-1)}(u_{(i)}) - \delta_{(i)}(u_{(i)}))| \, du = \int_I \delta'_{(i)}(t) |\log(\delta_{(i-1)}(t) - \delta_{(i)}(t))| \, dt.$$

Notice that $J_{3,i} \in [0, +\infty]$ and $\sum_{i=2}^{d} J_{3,1} = \mathbb{J}(\delta)$. The results on $J_{1,i}$, $J_{2,i}$ and $J_{3,i}$ imply that we can decompose $H(C_\delta)$ as:

$$H(C_\delta) = \log(d!) - \sum_{i=1}^{d} J_{1,i} + \sum_{i=2}^{d} J_{2,i} - \sum_{i=2}^{d} J_{3,i} = \log(d!) + \sum_{i=1}^{d} H(\delta_{(i)}) + (d-1) - \mathbb{J}(\delta). \qquad \square$$

6.3. The optimization problem

Let $\delta \in \mathcal{D}^0$. Recall notation from Section 6.1. The problem of maximizing H over $\mathcal{C}_\delta^0$ can be written as an optimization problem (P^δ) with infinite dimensional constraints:

$$\text{maximize } H(c) \text{ subject to } \begin{cases} \mathcal{A}(c) = b^\delta, \\ c \geq 0 \text{ a.e. and } c \in L^1(I^d). \end{cases} \qquad (P^\delta)$$

Notice that if $f \in L^1(I^d)$ is non-negative and solves $\mathcal{A}(f) = b^\delta$, then f is the density of a copula. We say that a function f is feasible for (P^δ) if $f \in L^1(I^d)$, $f \geq 0$ a.e., $\mathcal{A}(f) = b^\delta$ and $H(f) > -\infty$. We say that f is an optimal solution of (P^δ) if f is feasible and $H(f) \geq H(g)$ for all g feasible. The next proposition gives conditions which ensure the existence of an optimal solution.

Proposition 6.6. *Let $\delta \in \mathcal{D}^0$. If there exists c feasible for (P^δ), then there exists a unique optimal solution to (P^δ) and it is symmetric.*

Proof. Since $\mathcal{A}(f) = b^\delta$ implies $\mathcal{A}_1(f)(1) = b_1(1)$ that is $\int_{I^d} f(x)\, dx = 1$, we can directly apply Corollary 2.3 of [5] which states that if there exists a feasible c, then there exists a unique optimal solution to (P^δ). Since the constraints of (P^δ) are symmetric, such as the functional H, we deduce that if c^* is the optimal solution, then so is c_π^* defined for $\pi \in \mathcal{S}_d$ and $u \in I^d$ as $c_\pi^*(u) = c^*(u_\pi)$. By uniqueness of the optimal solution, we deduce that $c^* = c_\pi^*$ for all permutations $\pi \in \mathcal{S}_d$; hence c^* is symmetric. $\qquad \square$

Combining Lemmas 4.2 and 4.1 gives the following corollary on the support of any c verifying $\mathcal{A}(c) = b^\delta$.

Corollary 6.7. *Let $\delta \in \mathcal{D}^0$. If $c \in L^1(I^d)$ is non-negative and verifies $\mathcal{A}(c) = b^\delta$, then $c = 0$ a.e. on $Z_\delta \cup L_\delta^c$ with L_δ defined by (40) and $L_\delta^c = I^d \setminus L_\delta$.*

6.4. Reduction of the optimization problem (P^δ)

Let $\delta \in \mathcal{D}^0$. Since the optimal solution of (P^δ) is symmetric, see Proposition 6.6, we can reduce the optimization problem by considering it on the simplex $\triangle$. We define μ to be the Lebesgue measure restricted to $(Z_\delta^c \cap L_\delta) \cap \triangle$: $\mu(du) = \mathbf{1}_{(Z_\delta^c \cap L_\delta) \cap \triangle}(u)\, du$. We define, for $f \in L^1(I^d)$:

$$H^\mu(f) = -\int_{I^d} f(u) \log\big(f(u)\big) \mu(du).$$

From Corollary 6.7, we can deduce that if $c \in L^1(I^d)$ is non-negative symmetric and solves $\mathcal{A}(c) = b^\delta$, then:

$$H(c) = d! H^\mu(c). \tag{68}$$

Let us also define, for $f \in L^1(I^d)$, $1 \leq i \leq d$, $r \in I$:

$$\mathcal{A}_i^\mu(c)(r) = d! \int_{I^d} c(u) \mathbf{1}_{\{u_i \leq r\}} \mu(du).$$

We shall consider the restricted optimization problem (P_μ^δ) given by:

$$\text{maximize } H^\mu(c) \text{ subject to } \begin{cases} \mathcal{A}^\mu(c) = \delta, \\ c \geq 0 \ \mu\text{-a.e. and } c \in L^1(I^d). \end{cases} \tag{P_μ^δ}$$

We have the following equivalence between (P^δ) and (P_μ^δ). Recall that u^{OS} denotes the ordered vector of $u \in \mathbb{R}^d$.

Corollary 6.8. *Let $\delta \in \mathcal{D}^0$. If c is the optimal solution of (P^δ) then it is also an optimal solution to (P_μ^δ). If $\hat{c}$ is an optimal solution of (P_μ^δ), then c, defined by $c(u) = \hat{c}(u^{\mathrm{OS}}) \mathbf{1}_{Z_\delta^c \cap L_\delta}(u)$ is the optimal solution to (P^δ).*

Notice the Corollary implies that (P_μ^δ) has a μ-a.e. unique optimal solution: if c_1 and c_2 are two optimal solutions of (P_μ^δ) then μ-a.e. $c_1 = c_2$. Thanks to Proposition 6.6 and (68), Corollary 6.8 is a direct consequence of the following lemma that establishes the connection between the constraints.

Lemma 6.9. *Let $\delta \in \mathcal{D}^0$. For $c \in L^1(I^d)$ symmetric and non-negative the following two conditions are equivalent:*

1. $\mathcal{A}(c) = b^\delta$.
2. $\mathcal{A}^\mu(c) = \delta$ and $c = 0$ a.e. on $Z_\delta \cup L_\delta^c$.

Proof. Assume that $\mathcal{A}(c) = b^\delta$. We have, by Corollary 6.7, that $c = 0$ a.e. on $Z_\delta \cup L_\delta^c$. This and the symmetry of c gives, for $1 \leq i \leq d$, $r \in I$:

$$\mathcal{A}_i^\mu(c)(r) = d! \int_{I^d} c(u) \mathbf{1}_{\{u_{(i)} \leq r\}} \mathbf{1}_\triangle(u) \, du = \int_{I^d} c(u) \mathbf{1}_{\{u_{(i)} \leq r\}} \, du = \delta_{(i)}(r).$$

On the other hand, let us assume that $\mathcal{A}^\mu(c) = \delta$ and $c = 0$ a.e. on $Z_\delta \cup L_\delta^c$. We have, for $1 \leq i \leq d$, $r \in I$:

$$\mathcal{A}_{d+i}(c)(r) = \int_{I^d} c(u) \mathbf{1}_{\{u_{(i)} \leq r\}} \mathbf{1}_{Z_\delta^c \cap L_\delta}(u) \, du = d! \int_{I^d} c(u) \mathbf{1}_{\{u_i \leq r\}} \mu(du) = \delta_{(i)}(r),$$

where we used $c = 0$ a.e. on $Z_\delta \cup L_\delta^c$ for the first equality, the symmetry of c and the definition of μ for the second, and $\mathcal{A}^\mu(c) = \delta$ for the third. Lemma 6.1 ensures then that $\mathcal{A}_i(c) = b_i$ for $1 \leq i \leq d$. This ends the proof. $\qquad \square$

6.5. Solution for the reduced optimization problem (P_μ^δ)

Let $\delta \in \mathcal{D}^0$. We compute $(\mathcal{A}^\mu)^* : L^\infty(I)^d \to L^\infty(I^d)$ the adjoint of $\mathcal{A}^\mu$. For $\lambda = (\lambda_i, 1 \leq i \leq d) \in L^\infty(I)^d$ and $f \in L^1(I^d)$, we have:

$$\langle (\mathcal{A}^\mu)^*(\lambda), f \rangle = \langle \lambda, \mathcal{A}^\mu(f) \rangle = \sum_{i=1}^d \int_I \lambda_i(r) \int_{I^d} f(u) \mathbf{1}_{\{u_i \leq r\}} \, d\mu(u) \, dr$$

$$= \int_{I^d} f(u) \sum_{i=1}^d \Lambda_i(u_i) \, d\mu(u),$$

where we used the definition of the adjoint operator for the first equality, Fubini's theorem for the second, and the following definition of the functions $(\Lambda_i, 1 \leq i \leq d)$ for the third:

$$\Lambda_i(t) = \int_I \lambda_i(r) \mathbf{1}_{\{r \geq t\}} \, dr, \qquad t \in I.$$

Thus, we have for $\lambda \in L^\infty(I)^d$ and $u = (u_1, \ldots, u_d) \in I^d$:

$$(\mathcal{A}^\mu)^*(\lambda)(u) = \sum_{i=1}^d \Lambda_i(u_i). \tag{69}$$

We will use Theorem 2.9. from [5] on abstract entropy minimization, which we recall here, adapted to the context of (P_μ^δ).

Theorem 6.10 (Borwein, Lewis and Nussbaum). *Suppose there exists $c > 0$ μ-a.e. which is feasible for (P_μ^δ). Then there exists a μ-a.e. unique optimal solution, c^*, of (P_μ^δ). Furthermore, we have $c^* > 0$ μ-a.e. and there exists a sequence $(\lambda^n, n \in \mathbb{N}^*)$ of elements of $L^\infty(I)^d$ such that:*

$$\int_{I^d} c^*(u) \big| (\mathcal{A}^\mu)^*(\lambda^n)(u) - \log(c^*(u)) \big| \mu(du) \xrightarrow[n \to \infty]{} 0. \tag{70}$$

Now we are ready to prove that the optimal solution c^* of (P_μ^δ) is the product of measurable univariate functions.

Lemma 6.11. *Let $\delta \in \mathcal{D}^0$. Suppose that there exists $c > 0$ μ-a.e. which is feasible for (P_μ^δ). Then there exist non-negative, measurable functions $(a_i^*, 1 \leq i \leq d)$ defined on I such that $a_i^*(s) = 0$ if $\delta'_{(i)}(s) = 0$ and the function c^* defined a.e. on I^d by:*

$$c^*(u) = \frac{1}{d!} \mathbf{1}_{L_\delta}(u) \prod_{i=1}^d a_i^*(u_i)$$

is the optimal solution to (P_μ^δ).

Proof. According to Theorem 6.10, there exists a sequence $(\lambda^n, n \in \mathbb{N})$ of elements of $L^\infty(I)^d$ such that the optimal solution, say c^*, satisfies (70). This implies, thanks to (69), that there exist d sequences $(\Lambda_i^n, n \in \mathbb{N}^*, 1 \leq i \leq d)$ of elements of $L^\infty(I)$ such that the following convergence holds in $L^1(I^d, c^*\mu)$:

$$\sum_{i=1}^d \Lambda_i^n(u_i) \xrightarrow[n \to \infty]{\log} \big(c^*(u)\big). \tag{71}$$

We first assume that there exist Λ_i, $1 \leq i \leq d$ measurable functions defined on I such that μ-a.e. on S:

$$\sum_{i=1}^d \Lambda_i(u_i) = \log\big(c^*(u)\big). \tag{72}$$

Set $a_i^* = \sqrt[d]{d!} \exp(\Lambda_i)$ so that μ-a.e. on S:

$$c^*(u) = \frac{1}{d!} \prod_{i=1}^d a_i^*(u_i). \tag{73}$$

Recall $\mu(du) = \mathbf{1}_{(Z_\delta^c \cap L_\delta) \cap \triangle}(u)\, du$. From the definition (38) of Z_δ, we deduce that without loss of generality, we can assume that $a_i^*(u_i) = 0$ if $\delta'_{(i)}(u_i) = 0$. Therefore, we obtain $c^*(u) = (1/d!)\mathbf{1}_{L_\delta}(u) \prod_{i=1}^d a_i^*(u_i)$ for $u \in I^d$.

To complete the proof, we now show that (72) holds for Λ_i, $1 \leq i \leq d$ measurable functions. We introduce the notation $u_{(-i)} = (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_d) \in I^{d-1}$. Let us define the probability measure $P(du) = c^*(u)\mu(du)/\int_{I^d} c^*(y)\mu(dy)$ on I^d. We fix j, $1 \leq j \leq d$. In order to apply Proposition 2 of [18], which ensures the existence of the limiting measurable functions Λ_i, $1 \leq i \leq d$, we first check that P is absolutely continuous with respect to $P_1^j \otimes P_2^j$, where $P_1^j(du_{(-j)}) = \int_{u_j \in I} P(du_{(-j)}\, du_j)$ and $P_2^j(du_j) = \int_{u_{(-j)} \in I^{d-1}} P(du_{(-j)}\, du_j)$ are the marginals of P. Notice that there exists a non-negative density function h such that $P(du) = h(u_{(-j)}, u_j)\, du_{(-j)}\, du_j$. Let $h_1(u_{(-j)}) = \int h(u_{(-j)}, u_j)\, du_j$ and $h_2(u_j) = \int h(u_{(-j)}, u_j)\, du_{(-j)}$ denote the density of the marginals P_1^j and P_2^j. Then the density of the product measure $P_1^j \otimes P_2^j$ is given by $P_1^j \otimes P_2^j(du) = h_1(u_{(-j)})h_2(u_j)\, du_{(-j)}\, du_j$. The support of the density h is noted by $T_0 = \{u \in I^d; h(u) > 0\}$, and the support of the marginals are noted by $T_1 = \{v \in I^{d-1}; h_1(v) > 0\}$ and $T_2 = \{t \in I; h_2(t) > 0\}$. With this notation, we have that a.e. $T_0 \subset T_1 \times T_2$ (that is $T_0 \cap (T_1 \times T_2)^c$ is of zero Lebesgue measure). If $A \subset I^d$ is such that $\int \mathbf{1}_A(u)h_1(u_{(-j)})h_2(u_j)\, du_{(-j)}\, du_j = 0$, then we also have $\int \mathbf{1}_{A \cap (T_1 \times T_2)}(u)h_1(u_{(-j)})h_2(u_j)\, du_{(-j)}\, du_j = 0$. Since $h_1 h_2$ is positive on $T_1 \times T_2$, this implies that $A \cap (T_1 \times T_2)$ has zero Lebesgue measure. Therefore, we have:

$$\int \mathbf{1}_A(u)h(u)\, du = \int \mathbf{1}_{A \cap (T_1 \times T_2)}(u)h(u)\, du + \int \mathbf{1}_{A \setminus (T_1 \times T_2)}(u)h(u)\, du = 0,$$

since $h = 0$ a.e. on $A \setminus (T_1 \times T_2)$. This proves that P is absolutely continuous with respect to $P_1^j \otimes P_2^j$. Then according to Proposition 2 of [18], (71) implies that there exist measurable

functions Φ_j and $\tilde{\Lambda}_j$ defined respectively, on I^{d-1} and I, such that $c^*\mu$-a.e. on $\triangle$:

$$\log\big(c^*(u)\big) = \Phi_j(u_{(-j)}) + \tilde{\Lambda}_j(u_j).$$

As μ-a.e. $c^* > 0$, this equality holds μ-a.e. on S. Since we have such a representation for every $1 \leq j \leq d$, we can easily verify that $\log(c^*(u)) = \sum_{i=1}^{d} \Lambda_i(u_i)$ μ-a.e. with $\tilde{\Lambda}_j = \Lambda_j$ up to an additive constant. $\qquad\square$

6.6. Proof of Proposition 4.6

Let $\delta \in \mathcal{D}^0$. Recall that u^{OS} denotes the ordered vector of $u \in \mathbb{R}^d$. Let c be the density of a symmetric copula in $\mathbb{R}^d$ such that $\mathcal{A}(c) = b^\delta$ and c is of product form, that is, thanks to Corollary 6.8, $c(u) = c^*(u^{\mathrm{OS}})$ with

$$c^*(u) = \frac{1}{d!}\mathbf{1}_{L_\delta}(u)\prod_{i=1}^{d} a_i^*(u_i), \qquad u = (u_1, \ldots, u_d) \in \triangle,$$

where a_i^*, $1 \leq i \leq d$ are measurable non-negative functions defined on I. In this section, we shall prove that c equals c_δ defined by (44); that is, for all $1 \leq i \leq d$, a_i^* is a.e. equal, up to a multiplicative constant, to a_i defined in (45). This will prove Proposition 4.6.

Recall the definitions of $g_i^{(j)}, m_i^{(j)}, d_i^{(j)}$ from Section 4, for $1 \leq i \leq d + 1$. We deduce from (43) that:

$$c^*(u) = \frac{1}{d!}\mathbf{1}_{L_\delta}(u)\prod_{i=1}^{d} a_i^*(u_i)\mathbf{1}_{\Psi_i^\delta \cap \Psi_{i+1}^\delta}(u_i), \qquad u = (u_1, \ldots, u_d) \in \triangle.$$

We deduce also from Lemma 6.9 that $\mathcal{A}^\mu(c^*) = \delta$. We introduce the following family of functions:

$$B_{d+1}^*(t) = E_0^*(t) = 1,$$

and for $1 \leq i \leq d$, $t \in (g_i^{(j)}, d_i^{(j)})$ and $t' \in (g_{i+1}^{(j)}, d_{i+1}^{(j)})$:

$$B_i^*(t) = \int_t^{d_i^{(j)}} a_i^*(s)B_{i+1}^*(s)\,ds, \qquad E_i^*(t') = \int_{g_{i+1}^{(j)}}^{t'} a_i^*(s)E_{i-1}^*(s)\,ds.$$

Recall the functions B_i, for $1 \leq i \leq d + 1$, and E_i, for $0 \leq i \leq d$ defined by (61) and (63). We will prove by (downward) induction on $i \in \{1, \ldots, d + 1\}$ that:

$$B_i^*(t) = B_i^*\big(m_i^{(j)}\big)B_i(t), \qquad t \in \big(g_i^{(j)}, d_i^{(j)}\big). \tag{74}$$

For $i = d + 1$, it trivially holds. Let us assume that (74) holds for $i + 1$, $d \geq i \geq 1$. Recall the convention $K_{d+1} = 0$, $\delta_{(d+1)} = 0$ and $\delta_{(0)} = 1$. Arguing as in the proof of Lemma 6.4, we deduce

from $\mathcal{A}_i^\mu(c^*) = \delta_{(i)}$ that for $r \in \Psi_i^\delta$:

$$\delta_{(i)}(r) = d! \int_{I^d} c^*(u)\mathbf{1}_{\{u_{(i)} \leq r\}} \mu(du)$$

$$= d! \int_{I^d} c^*(u)\mathbf{1}_{\{u_{(i+1)} \leq r\}} \mu(du) + \int_\triangle \mathbf{1}_{L_\delta}(u) \prod_{j=1}^{d} \left(a_j^*(u_j)\mathbf{1}_{\Psi_j^\delta \cap \Psi_{j+1}^\delta}(u_j)\right)\mathbf{1}_{\{u_i \leq r \leq u_{i+1}\}} \, du$$

$$= \delta_{(i+1)}(r) + B_{i+1}^*(r)E_i^*(r).$$

This gives on Ψ_i^δ:

$$\delta_{(i)} - \delta_{(i+1)} = B_{i+1}^* E_i^*. \tag{75}$$

Notice (75) holds for $i = d$ thanks to the conventions. We get on Ψ_i^δ:

$$\delta_{(i)}' - \delta_{(i+1)}' = -K_{i+1}' B_{i+1}^* E_i^* + B_{i+1}^* a_i^* E_{i-1}^* = -\delta_{(i+1)}' + B_{i+1}^* a_i^* E_{i-1}^*,$$

where we took the derivative in (75), twice the induction hypothesis for B_{i+1}^* and (62) for the first equality; then (46) and (75) for the second. We deduce that on Ψ_i^δ:

$$\delta_{(i)}' = B_{i+1}^* a_i^* E_{i-1}^*. \tag{76}$$

On Ψ_i^δ, we can divide (76) by (75) and get, thanks to (46):

$$\frac{a_i^* B_{i+1}^*}{B_i^*} = \frac{\delta_{(i)}'}{\delta_{(i-1)} - \delta_{(i)}} = K_i'.$$

Notice that $(B_i^*)' = -a_i^* B_{i+1}^*$. So using the representation (62) of B_i, we get that (74) holds for i. Thus (74) holds for $1 \leq i \leq d + 1$. Then use (74) as well as $(B_i^*)' = -a_i^* B_{i+1}^*$ and $B_i' = -a_i B_{i+1}$ to get that for $t \in (g_i^{(j)}, d_i^{(j)}) \cap (g_{i+1}^{(k)}, d_{i+1}^{(k)})$:

$$a_i^*(t) = \frac{B_i^*(m_i^{(j)})}{B_{i+1}^*(m_{i+1}^{(k)})} a_i(t).$$

Therefore if $u = (u_1, \ldots, u_d) \in L^\delta$, we have:

$$\prod_{i=1}^{d} a_i^*(u_{(i)}) = \frac{B_1^*(m_1)}{B_{d+1}^*(m_{d+1})} \prod_{i=1}^{d} a_i(u_{(i)}),$$

since when $u \in L_\delta$, $u_{(i-1)}$ and $u_{(i)}$ belong to the same interval $(g_i^{(j)}, d_i^{(j)})$ for $2 \leq i \leq d$. This ensures that c_δ and c^* are densities of probability function which differ by a multiplicative constant, therefore they are equal. This ends the proof of Proposition 4.6.

6.7. Proof of case (a) for Theorems 4.7 and 5.4

We first consider the case $d = 2$. Let $\delta \in \mathcal{D}^0$ with $\mathbb{J}(\delta) = +\infty$. Recall $\mathbb{J}(\delta)$ is defined by (15). We have:

$$
\begin{aligned}
\mathbb{J}(\delta) &= -\int_I \delta'_{(2)}(t) \log\big(2\big(t - \delta_{(2)}(t)\big)\big)\, dt \\
&= -\log(2) - \int_I \log\big(t - \delta_{(2)}(t)\big)\, dt + \int_I \big(1 - \delta'_{(2)}(t)\big) \log\big(t - \delta_{(2)}(t)\big)\, dt \\
&= -\log(2) - \int_I \log\big(t - \delta_{(2)}(t)\big)\, dt + \big[\big(t - \delta_{(2)}(t)\big) \log\big(t - \delta_{(2)}(t)\big) - \big(t - \delta_{(2)}(t)\big)\big]_0^1 \\
&= -\log(2) - \int_I \log\big(t - \delta_{(2)}(t)\big)\, dt,
\end{aligned}
$$

where we used $\delta_{(1)} + \delta_{(2)} = 2t$ for the first equality, $\delta_{(2)}(1) = 1$ and $\delta_{(2)}(0) = 0$ for the second and last. In particular, we obtain that $\mathbb{J}(\delta)$ is equal to $-\log(2) + \mathcal{J}(\delta_{(2)})$, with $\mathcal{J}$ as also defined by (1) in [6]. Therefore, we deduce case (a) of Theorem 4.7 (for $d = 2$) from case (a) of Theorem 2.4 in [6]. Then, we get from (14) and Theorem 4.7 case (a) that $H_h(F) = -\infty$ for all $F \in \mathcal{L}_2^{OS}(\mathbf{F})$. This proves case (a) for Theorem 5.4 (for $d = 2$).

We then consider the case $d \geq 2$. Let $\delta \in \mathcal{D}^0$ with $\mathbb{J}(\delta) = +\infty$. This implies that there exists $2 \leq i \leq d$ such that $\int_I \delta'_{(i)}(t)|\log(\delta_{(i-1)}(t) - \delta_{(i)}(t))|\, dt = +\infty$. Set $\mathbf{F} = (\delta_{(i-1)}, \delta_{(i)})$ and notice that $\mathbf{F}$ belongs to $\mathcal{F}_2$ as $\delta_{(i)}$ is d-Lipschitz. Since $\int_I \delta'_{(i)}(t)|\log(\delta_{(i-1)}(t) - \delta_{(i)}(t))|\, dt = +\infty$, we deduce from the first part of this Section that $\max_{F \in \mathcal{L}_2^{OS}(\mathbf{F})} H_h(F) = -\infty$.

Consider a copula C belonging to $\mathcal{C}_\delta \cap \mathcal{C}^{\mathrm{sym}}$ and U a random vector on I^d with c.d.f. C. According to Lemma 3.21 and Lemma 3.3, as C is symmetric, we have:

$$
H\big(U^{OS}\big) = H(U) - \log(d!) = H(C) - \log(d!).
$$

It is easy to check that if $X = (X_1, \ldots, X_d)$ is a random vector on I^d and $2 \leq i \leq d$, then we have $H((X_{i-1}, X_i)) \geq H(X)$. This implies that, for $V = (U_{i-1}^{OS}, U_i^{OS})$,

$$
H(V) \geq H(C) - \log(d!).
$$

Since the c.d.f. of U_ℓ^{OS} is $\delta_{(\ell)}$ as $C \in \mathcal{C}_\delta$, we deduce the c.d.f. of V belongs to $\mathcal{L}_2^{OS}(\mathbf{F})$, and thus $H(V) = -\infty$. This implies that $H(C) = -\infty$. Thanks to Proposition 6.6 which states that the entropy is maximal on symmetric copulas, we deduce that:

$$
\max_{C \in \mathcal{C}_\delta} H(C) = \max_{C \in \mathcal{C}_\delta \cap \mathcal{C}^{\mathrm{sym}}} H(C) = -\infty.
$$

This proves cases (a) for Theorem 4.7. Then, we get from (14) that $H_h(F) = -\infty$ for all $F \in \mathcal{L}_d^{OS}(\mathbf{F})$. This proves case (a) for Theorem 5.4.

6.8. Proof of Theorem 4.7, case (b)

Let $\delta \in \mathcal{D}$ with $\mathbb{J}(\delta) < +\infty$. Thanks to Lemma 3.16, $\mathbb{J}(\delta) < +\infty$ implies that $\delta \in \mathcal{D}^0$. By construction, c_δ introduced in Proposition 4.5 verifies μ-a.e. $c_\delta > 0$. The density c_δ is a feasible solution to the problem (P_μ^δ). Theorem 6.10 ensures the existence of a unique optimal solution c^*. Furthermore, by Lemma 6.11, we have that there exist non-negative, measurable functions a_i^*, $1 \le i \le d$, such that $c^*(u) = (1/d!)\mathbf{1}_{L_\delta}(u) \prod_{i=1}^d a_i^*(u_i)$ μ-a.e. By Corollary 6.8, the optimal solution c of (P^δ) is given by, for $u = (u_1, \ldots, u_d)$:

$$c(u) = c^*(u^{\mathrm{OS}})\mathbf{1}_{Z_\delta^c \cap L_\delta}(u) = \frac{1}{d!}\mathbf{1}_{L_\delta}(u) \prod_{i=1}^d a_i^*(u_{(i)})\mathbf{1}_{\{\delta'_{(i)}(u_{(i)}) \neq 0\}}.$$

Since c is of product form, Proposition 4.6 yields that $c = c_\delta$ a.e., therefore C_δ is the unique copula achieving $H(C_\delta) = \max_{C \in \mathcal{C}_\delta} H(C)$.

7. Overview of the notations

- $\mathcal{F}_d$: set of continuous one-dimensional marginal c.d.f.s $\mathbf{F} = (\mathbf{F}_1, \ldots, \mathbf{F}_d)$ of d-dimensional order statistics, see (10).
- $\mathcal{F}_d^0$: set of continuous one-dimensional marginal c.d.f.s $\mathbf{F} = (\mathbf{F}_1, \ldots, \mathbf{F}_d)$ of d-dimensional abs. cont. order statistics, see Definition 3.17.
- $H_h(F)$: the relative entropy (with respect to the reference probability density h) of the random variable corresponding to the c.d.f. F, see (13).
- $H(F) = H_h(F)$ when $h = \mathbf{1}_{[0,1]}$ and F is the c.d.f. of a random variable taking values in $[0, 1]^d$.
- $\mathbb{J}(\mathbf{F})$: the quantity appearing in the expression of the entropy of the maximum entropy distribution of order statistics with marginal c.d.f.s $\mathbf{F} \in \mathcal{F}_d$, see (15).
- $\mathcal{L}_d$: set of c.d.f.s on $\mathbb{R}^d$ with continuous one-dimensional marginal c.d.f.s.
- $\mathcal{L}_d^0$: set of absolutely continuous c.d.f.s on $\mathbb{R}^d$.
- $\mathcal{L}_d^{\mathrm{OS}}$: set of c.d.f.s of d-dimensional order statistics with continuous one dimensional marginal c.d.f.s.
- $\mathcal{L}_d^{\mathrm{OS}}(\mathbf{F})$: set of c.d.f.s of d-dimensional order statistics with marginal c.d.f.s $\mathbf{F}$, see (11).
- $\mathcal{L}_d^{\mathrm{sym}}$: set of symmetric c.d.f.s on $\mathbb{R}^d$ with continuous one-dimensional marginal c.d.f.s.
- F^{sym}: the symmetrization of the c.d.f. F, see (16).
- $S_\mathbf{F}$: symmetrizing operator on copulas, associated to the marginal c.d.f.s $\mathbf{F}$, see Definition 3.1.
- $\mathcal{C}$: set of all copulas.
- $\mathcal{C}^0$: set of absolutely continuous copulas.
- $\mathcal{C}^{\mathrm{OS}}(\mathbf{F})$: set of copulas of order statistics with marginal c.d.f.s $\mathbf{F}$, see (12).
- $\mathcal{C}^{\mathrm{sym}}$: set of symmetric (permutation invariant) copulas.
- $\mathcal{C}^{\mathrm{sym}}(\mathbf{F})$: image of the set $\mathcal{C}^{\mathrm{OS}}(\mathbf{F})$ by the operator $S_\mathbf{F}$, see (19). It is the set of symmetric copulas with multidiagonal $\delta^\mathbf{F}$.
- $\mathcal{C}_\delta$: set of copulas with multidiagonal δ, see Section 3.2.

- $\mathcal{C}_\delta^0$: set of abs. cont. copulas with multidiagonal δ, see Section 3.2.
- $\mathcal{D}$: set of multidiagonals of copulas, see Section 3.2.
- $\mathcal{D}^0$: set of multidiagonals of abs. cont. copulas, see Section 3.2.
- $\Psi_i^{\mathbf{F}}$: set of points $t \in \mathbb{R}$ for which the marginal c.d.f.s $\mathbf{F} = (\mathbf{F}_1, \ldots, \mathbf{F}_d)$ verify $\mathbf{F}_{i-1}(t) > \mathbf{F_i}(t)$, see (29).
- $T^{\mathbf{F}}$: set of points $u = (u_1, \ldots, u_d) \in I^d$ for which $\mathbf{F}_1^{-1}(u_1) \leq \cdots \leq \mathbf{F}_d^{-1}(u_d)$, see (33). The density of all copulas in $\mathcal{C}^{\mathrm{OS}}(\mathbf{F})$ vanishes outside $T^{\mathbf{F}}$.
- $L^{\mathbf{F}}$: set of ordered vectors $x \in \mathbb{R}^d$ such that the marginal c.d.f.s $\mathbf{F} = (\mathbf{F}_1, \ldots, \mathbf{F}_d)$ verify $\mathbf{F}_{i-1}(t) > \mathbf{F}_i(t)$ for all $t \in (x_{i-1}, x_i)$, $2 \leq i \leq d$, see (51). The density of any abs. cont. c.d.f. in $\mathcal{L}_d^{\mathrm{OS}}(\mathbf{F})$ vanishes outside $L^{\mathbf{F}}$.
- L_δ: set of points $u = (u_1, \ldots, u_d) \in I^d$ for which all points $t \in (u_{(i-1)}, u_{(i)})$ verify $\delta_{(i-1)}(t) > \delta_{(i)}(t)$ for all $2 \leq i \leq d$, see (40). The density of any copula in $\mathcal{C}_\delta^0$ vanishes outside L_δ.
- Z_δ: set of points $u = (u_1, \ldots, u_d) \in I^d$ such that $\delta'_{(i)}(u_{(i)}) = 0$ for some $1 \leq i \leq d$, see (38). The density of any copula in $\mathcal{C}_\delta^0$ vanishes on Z_δ.

Acknowledgements

References

[1] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1992). *A First Course in Order Statistics*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics* **54**. New York: Wiley. MR1178934

[2] Avérous, J., Genest, C. and Kochar, S.C. (2005). On the dependence structure of order statistics. *J. Multivariate Anal.* **94** 159–171. MR2161215

[3] Bickel, P.J. (1967). Some contributions to the theory of order statistics. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability* (*Berkeley, Calf.*, 1965/66), *Vol. i*: *Statistics* 575–591. Berkeley, CA: Univ. California Press. MR0216701

[4] Boland, P.J., Hollander, M., Joag-Dev, K. and Kochar, S. (1996). Bivariate dependence properties of order statistics. *J. Multivariate Anal.* **56** 75–89. MR1380182

[5] Borwein, J.M., Lewis, A.S. and Nussbaum, R.D. (1994). Entropy minimization, *DAD* problems, and doubly stochastic kernels. *J. Funct. Anal.* **123** 264–307. MR1283029

[6] Butucea, C., Delmas, J.-F., Dutfoy, A. and Fischer, R. (2015). Maximum entropy copula with given diagonal section. *J. Multivariate Anal.* **137** 61–81. MR3332799

[7] Butucea, C., Delmas, J.-F., Dutfoy, A. and Fischer, R. (2015). Nonparametric estimation of distributions of order statistics with application to nuclear engineering. In *Safety and Reliability of Complex Engineered Systems*: *ESREL* 2015 (L. Podofillini, B. Sudret, B. Stojadinovic, E. Zio and W. Kröger, eds.) CRC Press. URL.

[8] David, H.A. and Nagaraja, H.N. (1970). *Order Statistics*. New York: Wiley.

[9] de Melo Mendes, B.V.d.M. and Sanfins, M.A. (2007). The limiting copula of the two largest order statistics of independent and identically distributed samples. *Braz. J. Probab. Stat.* **21** 85–101. MR2396715

[10] Dubhashi, D. and Häggström, O. (2008). A note on conditioning and stochastic domination for order statistics. *J. Appl. Probab.* **45** 575–579. MR2426853

[11] Hu, T. and Chen, H. (2008). Dependence properties of order statistics. *J. Statist. Plann. Inference* **138** 2214–2222. MR2406434

[12] Jaworski, P. (2009). On copulas and their diagonals. *Inform. Sci.* **179** 2863–2871. MR2547755

[13] Jaworski, P. and Rychlik, T. (2008). On distributions of order statistics for absolutely continuous copulas with applications to reliability. *Kybernetika* (*Prague*) **44** 757–776. MR2488903

[14] Kim, S.H. and David, H.A. (1990). On the dependence structure of order statistics and concomitants of order statistics. *J. Statist. Plann. Inference* **24** 363–368. MR1046972

[15] Lebrun, R. and Dutfoy, A. (2014). Copulas for order statistics with prescribed margins. *J. Multivariate Anal.* **128** 120–133. MR3199832

[16] Navarro, J. and Balakrishnan, N. (2010). Study of some measures of dependence between order statistics and systems. *J. Multivariate Anal.* **101** 52–67. MR2557618

[17] Navarro, J. and Spizzichino, F. (2010). On the relationships between copulas of order statistics and marginal distributions. *Statist. Probab. Lett.* **80** 473–479. MR2593588

[18] Rüschendorf, L. and Thomsen, W. (1993). Note on the Schrödinger equation and I-projections. *Statist. Probab. Lett.* **17** 369–375. MR1237783

[19] Schmitz, V. (2004). Revealing the dependence structure between $X_{(1)}$ and $X_{(n)}$. *J. Statist. Plann. Inference* **123** 41–47. MR2058120