

Article

Not peer-reviewed version

Fitting Copulas with Maximal Entropy

[Milan Bubák](#) and [Mirko Navara](#) *

Posted Date: 14 January 2025

doi: 10.20944/preprints202501.1040.v1

Keywords: maximum entropy estimator; copula; density; convex optimization



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Fitting Copulas with Maximal Entropy

Milan Bubák^{1,†} and Mirko Navara^{1,†,*} 

Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague

* Correspondence: navara@fel.cvut.cz

† These authors contributed equally to this work.

Abstract: We deal with two-dimensional copulas from the perspective of their differential entropy. We formulate a problem of finding a copula with maximum differential entropy when some copula values are given. As expected, the solution is a copula with a piecewise constant density (a checkerboard copula). This allows us to simplify the optimization of the continuous objective function, the differential entropy, to an optimization of finitely many density values. We present several ideas to simplify this problem. It has a feasible numerical solution. We present several instances that admit also closed-form solutions.

Keywords: maximum entropy estimator; copula; density; convex optimization

1. Formulation of the task

1.1. Motivation

Copulas are a successful tool for the description of *any* dependence of continuous random variables. Based on empirical data, we may try to find a copula best describing the underlying distribution.

The standard approach to fitting the *parametric copula classes* is widely used across various copula applications. The hidden assumption behind this is the knowledge of the dependence model for which we try to estimate the parameters. However, there are many applications where this approach is not desirable because of the lack of knowledge, insufficient sample size, or a high dimension of the task. Then we want to find a model based on *general copulas*. When there is no other information than the sample, the empirical or discrete copulas and their continuous extensions can be used. Because of the incomplete information of copula values in between the known values, the maximum entropy principle can also be applied.

We study the problem of finding a joint probability distribution from partial information, when we are given

- (I1) the marginals,
- (I2) the values of the joint cumulative distribution function (*joint cdf*) at finitely many points.

This refers to the situation when we have (or assume to have) complete information about the marginals, but we do not know their dependence. The sample is discrete and it can possibly be discretised to a coarser scale. Thus we know the desired values of the copula only on a *2D rectangular grid* of points resulting from the discretization. This task has a standard solution (a *checkerboard copula*, see below). However, some of these constraints may be missing for various reasons. In particular, we may have too few elements of the sample (possibly none) in some 2D intervals, so we do not consider it a reliable support for modelling the whole copula. We investigate this case.

1.2. Criteria of Optimality

Suppose that we have two continuous random variables, X, Y , with known cumulative distribution functions, F_X, F_Y . They need not be independent and their dependence can be specified, e.g., by their joint cdf, $F_{X,Y}$. According to Sklar's theorem (see [1,2]), it can be determined also by the

respective copula, $C_{X,Y}$. This is the joint cumulative distribution function of transformed random variables $F_X(X)$, $F_Y(Y)$, which have the uniform distribution on the interval $]0, 1[$,¹

$$C_{X,Y}(x, y) = F_{F_X(X), F_Y(Y)}(x, y) = F_{X,Y}(s, t),$$

where s, t are such that $F_X(s) = x$, $F_Y(t) = y$. The copula $C_{X,Y}$ gives “pure” information of the dependence of X, Y , independently of the marginal distributions. It allows the expression of an arbitrary dependence between two continuous random variables.²

Definition 1. If there is a function $c_{X,Y}$ such that the copula can be expressed as its integral,

$$C_{X,Y}(\bar{x}, \bar{y}) = \int_0^{\bar{x}} \int_0^{\bar{y}} c_{X,Y}(x, y) xy,$$

it is called a density of the copula $C_{X,Y}$. The differential entropy of the copula $C_{X,Y}$ is the (Shannon) entropy [3] of its density, given by

$$\text{Ent}(c_{X,Y}) = \int_0^1 \int_0^1 \eta(c_{X,Y}(x, y)) xy, \quad (1)$$

where

$$\eta(w) := \begin{cases} -w \ln w & \text{if } w > 0, \\ 0 & \text{if } w = 0. \end{cases}$$

Remark 1. The differential entropy of a copula does not exist if the copula is not absolutely continuous, i.e., it does not have a density. Our proposed solution avoids this problem.

1.3. Related Work and History of the Problem

One of the earliest pioneers in studying multivariate distributions was Claude E. Shannon, who introduced his groundbreaking work on information theory in [3]. Later, in 1959, Abe Sklar introduced copulas in his seminal work, which laid the foundation for the study of dependency structures in multivariate distributions [2].

The principle of maximum entropy in the choice of a distribution was well defended by Jaynes [4]. Several approaches to this aim have been proposed.

The study of copulas in relation to entropy began to gain attention around 2010, with foundational work by Pougaza et al. [5,6]. They give a nice overview and motivation of the maximum entropy principle.³ They maximize the differential entropy of the original *joint cdf*, while we propose to maximize the differential entropy of the corresponding *copula*.

The early contributions were primarily theoretical. Later on, Ma and Sun [7] introduced the use of the maximum entropy principle to estimate mutual information via copulas. Subsequently, Singh and Zhang [8] expanded the application of copulas connected to entropy for multivariate stochastic modeling in water engineering.

Piantadosi et al [9] (resp. [10]) found bivariate (resp. multivariate) copulas with maximal differential entropy under the knowledge of

- (I1) the marginals,
- (I3) the grade correlation coefficients.

Their solutions are checkerboard copulas.

Recently Lin et al. [11] applied the principle of maximum entropy and found checkerboard copulas as solutions. Their approach differs in the aspect that they start from distributions whose

¹ The bounds, 0 and 1, are omitted for simplification of formulas.

² More generally, an n -copula describes the dependence of n random variables. Here we deal only with 2-copulas.

³ The paper [5] deals equally with several definitions of entropy, here we consider only the original Shannon entropy, as the best-motivated notion.

marginals are not continuous; the discontinuities determine the checkerboard structure. The difficulties of extension of the copula approach to distributions with discontinuous marginals are presented by Genest et al. [12].

While the literature on this topic remains relatively limited, this leaves considerable room for further theoretical advancements and potential new applications.

None of the preceding approaches restricts solutions by prescribed values at some points (I2). Sometimes the values of the joint cdf (and hence also of the copula) at some points are known. Such constraints can originate, e.g., from a discretization of a continuous scale. We want to choose a distribution fitting to these constraints.

Let us choose the distribution whose copula has maximal differential entropy (among those fitting to the constraints). The principle of maximal entropy was successfully used in many other tasks. It expresses our intention “not to include more information than that contained in the constraints.” Choosing this option, we try not to introduce any other dependence than that which follows from the given restrictions. This seems to be a natural criterion for the choice of the model fitting to constraints.

The result can be equivalently described by the joint cdf or by the copula. The differential entropy of the joint cdf suffers some drawbacks, e.g., it is not invariant to the change of scale; the entropies of random variables X and cX , $c \neq 1$, are different. In contrast to it, the copula remains the same even if the random variables are transformed by arbitrary increasing functions f, g ,

$$C_{X,Y}(x,y) = C_{f(X),g(Y)}(x,y).$$

Thus the differential entropy of a copula is a well-defined notion which can be a criterion for the choice of the model.⁴

In this paper, we formulate the task of finding a copula with maximal differential entropy fitting the constraints. We convert it to a finite-dimensional optimization problem. As one of the main contributions, we prove that the result is independent of some inputs. This further simplifies the task. It leads to a system of higher-order polynomial equations which allows for a numerical solution using convex optimization. We show that in the simplest cases, an analytical solution is also feasible.

1.4. Notation and Formulation of the Problem

Before the formulation of the task solved in this paper, let us introduce our notation. We restrict attention to 2-copulas, i.e., *binary* functions describing the dependence of *two* random variables. We are trying to find a copula, $C_{X,Y}: [0,1]^2 \rightarrow [0,1]$. As we shall not deal with another copula in the sequel, we shall denote it briefly by C . For the readers not familiar with the notion of a copula, it can be any joint cdf with marginals uniformly distributed on $]0,1[$.⁵ They are characterized by the following necessary and sufficient conditions:

$$C(w,0) = C(0,w) = 0, \quad (2)$$

$$C(w,1) = C(1,w) = w, \quad (3)$$

$$C(x_i, y_j) - C(x_i, y_\ell) - C(x_k, y_j) + C(x_k, y_\ell) \geq 0 \quad \text{for } x_k < x_i, y_\ell < y_j. \quad (4)$$

Inequality (4) means that the probability

$$P(x_k < X < x_i, y_\ell < Y < y_j) = C(x_i, y_j) - C(x_i, y_\ell) - C(x_k, y_j) + C(x_k, y_\ell) \quad (5)$$

is non-negative.

⁴ We acknowledge that Michal Dibaba came up with the idea of using maximum entropy as a criterion for the choice of a copula [13]. The present paper explores this idea in detail; it is based on the bachelor thesis [14].

⁵ Strictly speaking, the joint cdf is defined on the whole plane, but its restriction to the unit square, even to its interior, $]0,1[^2$, determines the copula uniquely.

The constraints are finitely many points in $]0, 1[^2$ where the value of the copula is given. Their first (resp. second) coordinates can be organized in an increasing sequence $(x_1, \dots, x_{I-1})$, resp. $(y_1, \dots, y_{J-1})$. These values determine a rectangular grid. At some of the grid points, (x_i, y_j) , the values $z_{i,j}$ are given,

$$C(x_i, y_j) = z_{i,j}. \quad (6)$$

We denote by G_0 the set of all indices $(i, j) \in \{1, \dots, I-1\} \times \{1, \dots, J-1\}$ for which the values $z_{i,j}$ of the copula are given by (6). To simplify the notation, we define additionally $x_0 = y_0 = 0$, $x_I = y_J = 1$. The corresponding new grid points are at the boundary of the domain of the copula and form the set

$$B_{I,J} = \{(x_i, 0), (x_i, 1) \mid i = 0, \dots, I\} \cup \{(0, y_j), (1, y_j) \mid j = 0, \dots, J\}.$$

The values of the copula at the boundary are given by (2), (3), so we may define $z_{i,j}$ for all $(i, j) \in B_{I,J}$ by

$$z_{i,0} = C(x_i, 0) = 0, \quad z_{0,j} = C(0, y_j) = 0, \quad (7)$$

$$z_{i,I} = C(x_i, 1) = x_i, \quad z_{I,j} = C(1, y_j) = y_j. \quad (8)$$

In total, (6) is required for all (i, j) from the set

$$G = G_0 \cup B_{I,J}$$

and G_0 can be expressed as

$$G_0 = G \setminus B_{I,J} = G \cap]0, 1[^2.$$

In figures, we use a filled disk to denote a grid point at which the value is given (an element of G) and an empty circle to denote a grid point where the value is not restricted. Figures 1 and 2 demonstrate our notation.

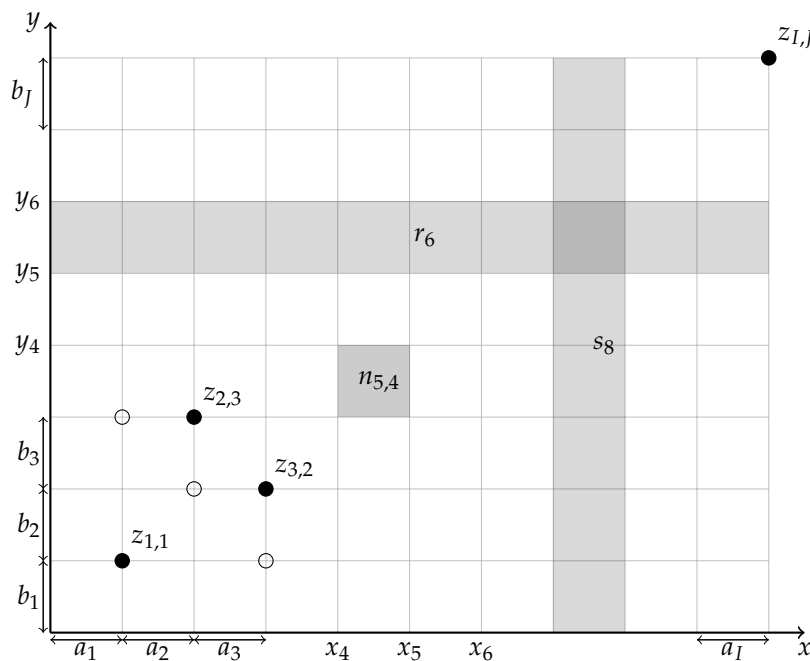


Figure 1. Sample of our notation

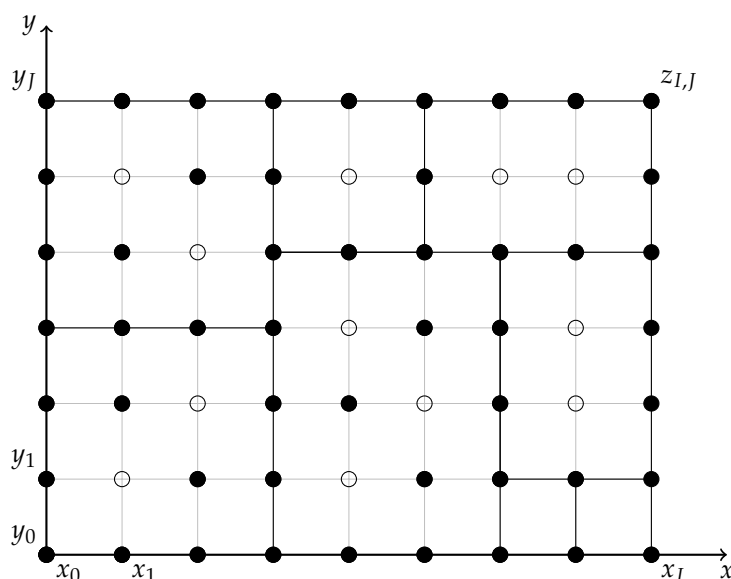


Figure 2. Possible set G of points in which the values of a copula are given and a covering of the grid by disjoint irreducible rectangles (bold lines)

Problem 1. Let $0 = x_0 < x_1 < \dots < x_I = 1$, $0 = y_0 < y_1 < \dots < y_J = 1$, and let there be given values $z_{i,j} \in [0, 1]$ for all (i, j) in some set G such that $B_{I,J} \subseteq G \subseteq \{0, \dots, I\} \times \{0, \dots, J\}$. Find a copula C satisfying (6) for all $(i, j) \in G$. Moreover, among all such copulas, we want to find one with maximum differential entropy.

The necessary and sufficient conditions for a copula, in particular (4), imply necessary conditions for solvability of Problem 1 which later will appear also sufficient.

Proposition 1. A necessary condition for the existence of a solution to Problem 1 is the conjunction of (7), (8), and

(N) If $(i, j), (i, \ell), (k, j), (k, \ell) \in G$, $k < i$, $\ell < j$, then $z_{i,j} - z_{i,\ell} - z_{k,j} + z_{k,\ell} \geq 0$.

In the sequel, we assume that the constraints satisfy these conditions.

2. General Results

2.1. Reformulation to a Finite-Dimensional Optimization

As formulated, Problem 1 looks like a task from variational calculus. We shall transform it into a finite-dimensional optimization problem.

It is well-known that among all absolutely continuous distributions on a bounded set, the uniform distribution has the highest differential entropy (if the integral of the density over the whole domain is fixed). We shall derive a consequence of this principle. To simplify its formulation, we denote the lengths of intervals between neighboring grid points:

$$\begin{aligned} a_i &= x_i - x_{i-1}, \\ b_j &= y_j - y_{j-1} \end{aligned}$$

for all $i \in \{1, \dots, I\}$, $j \in \{1, \dots, J\}$ (see Figure 1).

Proposition 2. Let $i \in \{1, \dots, I\}$, $j \in \{1, \dots, J\}$. The restriction of the copula C to the two-dimensional open interval $]x_{i-1}, x_i[\times]y_{j-1}, y_j[$ contributes to the copula differential entropy by

$$\int_{x_{i-1}}^{x_i} \int_{y_{j-1}}^{y_j} \eta(c_{X,Y}(x, y)) dx dy. \quad (9)$$

Among all copulas with a given value

$$\begin{aligned} n_{i,j} &:= C(x_i, y_j) - C(x_i, y_{j-1}) - C(x_{i-1}, y_j) + C(x_{i-1}, y_{j-1}) \\ &= P(x_{i-1} < X < x_i, y_{j-1} < Y < y_j), \end{aligned} \quad (10)$$

the contribution (9) is maximal iff its density is constant,

$$c_{X,Y}(x, y) = \frac{n_{i,j}}{(x_i - x_{i-1})(y_j - y_{j-1})} = \frac{n_{i,j}}{a_i b_j}, \quad (11)$$

almost everywhere on $]x_{i-1}, x_i[\times]y_{j-1}, y_j[$. Without loss of generality, we may choose the copula with density (11) for all $(x, y) \in]x_{i-1}, x_i[\times]y_{j-1}, y_j[$.

Proposition 2 determines the solution to Problem 1 (almost everywhere) when all values

$$C(x_i, y_j), \quad C(x_i, y_{j-1}), \quad C(x_{i-1}, y_j), \quad C(x_{i-1}, y_{j-1})$$

are given, i.e., $(i, j), (i, j-1), (i-1, j), (i-1, j-1) \in G$. Even when this is not the case, there are some optimal values $n_{i,j}, (i, j) \in \{1, \dots, I\} \times \{1, \dots, J\}$, and the copula density on $(x_{i-1}, x_i) \times (y_{j-1}, y_j)$ can be chosen according to (11) because there are no other restrictions on its values. This observation allows us to restrict attention to copulas with piecewise constant densities, so-called *checkerboard copulas*. Lin et al. [11] also found checkerboard copulas by maximization of the differential entropy, although they solved a different task.

It remains to find the optimal values of finitely many densities on intervals (rectangles) $]x_{i-1}, x_i[\times]y_{j-1}, y_j[$ for $(i, j) \in \{1, \dots, I\} \times \{1, \dots, J\}$. Equivalently, we look for the values $n_{i,j}$ defined by (10). From them, the values of the copula at grid points can be computed as sums

$$C(x_i, y_j) = \sum_{u=1}^i \sum_{v=1}^j n_{u,v}. \quad (12)$$

In particular, for $i = I$, resp. $j = J$, we obtain

$$\begin{aligned} \sum_{u=1}^I \sum_{v=1}^j n_{u,v} &= C(x_I, y_j) = C(1, y_j) = y_j, \\ \sum_{u=1}^i \sum_{v=1}^J n_{u,v} &= C(x_i, y_J) = C(x_i, 1) = x_i. \end{aligned}$$

These formulas can be equivalently expressed using row and column sums:

$$\begin{aligned} \sum_{u=1}^I n_{u,j} &= y_j - y_{j-1}, \\ \sum_{v=1}^J n_{i,v} &= x_i - x_{i-1}. \end{aligned}$$

Remark 2. We do not deal with the values of density on the boundaries of these intervals. The boundaries are of measure zero, so they do not influence the results. In fact, the densities at these boundaries are not uniquely defined.

Copulas that do not have densities or entropies (see Remark 1) do not seem to be good candidates for approximation and cannot compete with our proposed solution.

The differential entropy of the checkerboard copula is

$$\begin{aligned}\text{Ent}(c_{X,Y}) &= \int_0^1 \int_0^1 \eta(c_{X,Y}(x,y)) \, dx \, dy \\ &= \sum_{i=1}^I \sum_{j=1}^J a_i b_j \eta\left(\frac{n_{i,j}}{a_i b_j}\right) \\ &= - \sum_{i=1}^I \sum_{j=1}^J n_{i,j} \ln\left(\frac{n_{i,j}}{a_i b_j}\right),\end{aligned}$$

provided that all $n_{i,j} \neq 0$ (otherwise, the corresponding summand of the differential entropy is 0, i.e., minimal). We shall deal with the following reformulation:

Problem 2. Let $0 = x_0 < x_1 < \dots < x_I = 1$, $0 = y_0 < y_1 < \dots < y_J = 1$, and let there be given values $z_{i,j} \in [0, 1]$ for all (i, j) in some set G such that $B_{I,J} \subseteq G \subseteq \{0, \dots, I\} \times \{0, \dots, J\}$. Suppose that values $z_{i,j}$, $(i, j) \in G$, satisfy the necessary conditions from Proposition 1. Find values $n_{i,j} \in [0, 1]$, $(i, j) \in \{1, \dots, I\} \times \{1, \dots, J\}$, satisfying

$$\sum_{u=1}^I n_{u,j} = y_j - y_{j-1}, \quad (13)$$

$$\sum_{v=1}^J n_{i,v} = x_i - x_{i-1}, \quad (14)$$

$$\sum_{u=1}^i \sum_{v=1}^j n_{u,v} = z_{i,j} \text{ for all } (i, j) \in G,^6 \quad (15)$$

and such that the differential entropy

$$\text{Ent}(c_{X,Y}) = - \sum_{i=1}^I \sum_{j=1}^J n_{i,j} \ln\left(\frac{n_{i,j}}{(x_i - x_{i-1})(y_j - y_{j-1})}\right) \quad (16)$$

is maximal.

Remark 3. If $G = \{0, \dots, I\} \times \{0, \dots, J\}$, i.e., if all values $z_{i,j}$, $(i, j) \in \{1, \dots, I\} \times \{1, \dots, J\}$, are given, then (10) and (6) determine all $n_{i,j}$ and the solution.

2.2. Decomposition of the Problem

We concentrate on rectangles with the following property:

- (B) The values of the copula at all grid points at the boundary of the rectangle $[x_k, x_K] \times [y_\ell, y_L]$ are given, i.e.,

$$\{(x_i, y_\ell), (x_i, y_L) \mid i = k, \dots, K\} \cup \{(x_k, y_j), (x_K, y_j) \mid j = \ell, \dots, L\} \subseteq G.$$

Then the optimization of the copula inside such a rectangle is independent of the rest of the domain and can be solved separately. E.g., in Figure 2, rectangles satisfying (B) are drawn by thick lines. The whole domain $[0, 1] \times [0, 1]$ has the values on the entire boundary given, so it satisfies (B).

Let us consider a rectangle satisfying (B). If all its grid points in some row or column (not at the boundary) are in G , then we can split it into two disjoint rectangles satisfying (B). (The choice need not be unique.) If this is not the case (i.e., there is a missing value in each row and each column⁷),

⁶ If $i = 0$ (resp. $j = 0$), then $n(0, j)$ (resp. $n(i, 0)$) are not defined, but the sum from 1 to 0 is considered empty and equal to 0 = $z_{0,j}$ (resp. $0 = z_{i,0}$).

⁷ The boundary rows and columns are not considered.

we call the rectangle *irreducible*. Irreducible rectangles can be equivalently characterized as minimal rectangles satisfying (B). The whole domain can be covered by disjoint irreducible rectangles. This may allow us to decompose the problem into simpler ones, dealing with each irreducible rectangle separately. Figure 2 shows eight irreducible rectangles. Each component of the decomposed task has the following formulation. The active restrictions refer to grid points with indices from the set $G \cap \{k, \dots, K\} \times \{\ell, \dots, L\}$. Notice that the monotonicity of the values $z_{i,j}$ needs to be formulated explicitly, while in Problem 2 it followed from (7).

Problem 3. Let $0 \leq x_k < \dots < x_K \leq 1$, $0 \leq y_\ell < \dots < y_L \leq 1$, and let $G \subseteq \{k, \dots, K\} \times \{\ell, \dots, L\}$ be a set such that

$$\{(x_i, \ell), (x_i, L) \mid i = 0, \dots, I\} \cup \{(k, y_j), (K, y_j) \mid j = 0, \dots, J\} \subseteq G.$$

Suppose that values $z_{i,j} \in [0, 1]$ are given for all $(i, j) \in G$ so that sequences $(z_{i,\ell})_{i=k, \dots, K}$, $(z_{k,j})_{j=\ell, \dots, L}$ are nondecreasing and (N) is satisfied. Find values $n_{i,j} \in [0, 1]$, $(i, j) \in \{k+1, \dots, K\} \times \{\ell+1, \dots, L\}$, satisfying

$$\sum_{u=k+1}^i \sum_{v=\ell+1}^j n_{u,v} = z_{i,j} - z_{i,\ell} - z_{k,j} + z_{k,\ell} \quad (17)$$

for all $(i, j) \in G$ and such that the contribution to the differential entropy

$$- \sum_{i=k+1}^K \sum_{j=\ell+1}^L n_{i,j} \ln \left(\frac{n_{i,j}}{(x_i - x_{i-1})(y_j - y_{j-1})} \right) \quad (18)$$

is maximal.

In the sequel, we analyze special cases of irreducible rectangles. Before that, we simplify indexing by shifting the rectangle to the origin.

2.3. Shifted Indexing

If a rectangle $[x_k, x_K] \times [y_\ell, y_L]$ satisfies (B), the corresponding values $n_{i,j}$, $(i, j) \in \{k+1, \dots, K\} \times \{\ell+1, \dots, L\}$, are bounded only by the row and column sums

$$r_j = \sum_{i=k+1}^K n_{i,j} = z_{K,j} - z_{K,j-1} - z_{k,j} + z_{k,j-1}, \quad j \in \{\ell+1, \dots, L\}, \quad (19)$$

$$s_i = \sum_{j=\ell+1}^L n_{i,j} = z_{i,L} - z_{i-1,L} - z_{i,\ell} + z_{i-1,\ell}, \quad i \in \{k+1, \dots, K\}, \quad (20)$$

which are non-negative values satisfying the equality

$$\sum_{j=\ell}^L r_j = \sum_{i=k}^K s_i = \sum_{j=\ell+1}^L \sum_{i=k+1}^K n_{i,j} = z_{K,L} - z_{k,L} - z_{K,\ell} + z_{k,\ell}. \quad (21)$$

The same task is obtained if we take a rectangle of the same size with the lower left vertex at the origin and impose the same requirements on the row and column sums. We shall apply this simplification, and instead of the original task for a general rectangle, we shall, without loss of generality, solve the case when $k = \ell = 0$ in the above notation (with the new upper bounds still denoted by K, L ; the above general indexing will not be used anymore). The boundary grid points form the set

$$B_{K,L} = \{(x_i, 0), (x_i, L) \mid i = 0, \dots, I\} \cup \{(0, y_j), (K, y_j) \mid j = 0, \dots, J\}.$$

The situation simplifies not only in indexing (starting from 0, resp. 1), but also the values at the bottom and left edge become zero. The values at the top and right bound must be modified to general values $z_{i,L}, z_{K,j}$, which are not determined by x_i, y_j like they were in (8). The modified task is as follows.

Problem 4. Let $0 = x_0 < \dots < x_K \leq 1$, $0 = y_0 < \dots < y_L \leq 1$, and let there be given values $z_{i,j} \in [0, 1]$ for all (i, j) in some set G such that $B_{K,L} \subseteq G \subseteq \{0, \dots, K\} \times \{0, \dots, L\}$.⁸ Suppose that values $z_{i,j}, (i, j) \in G$, satisfy the necessary conditions (N) and (7) for $i \in \{0, \dots, K\}$ and $j \in \{0, \dots, L\}$. Find values $n_{i,j} \in [0, 1], (i, j) \in \{1, \dots, K\} \times \{1, \dots, L\}$, satisfying

$$\sum_{u=1}^i \sum_{v=1}^j n_{u,v} = z_{i,j} \quad \text{for all } (i, j) \in G \cap]0, K] \times]0, L] \quad (22)$$

and such that the contribution to the differential entropy

$$-\sum_{i=1}^K \sum_{j=1}^L n_{i,j} \ln \left(\frac{n_{i,j}}{(x_i - x_{i-1})(y_j - y_{j-1})} \right) \quad (23)$$

is maximal.

As in [11], this is an optimization of a concave function that has a maximum that can be found numerically (applying standard convex optimization with the opposite sign). For numerical solutions of convex tasks, standard references are, e.g., [15–18]. We shall show that analytical solutions exist in some cases and we demonstrate them to support the intuition about the task.

3. Less General Tasks

In this section, we collect results related to Problem 4. Throughout this section, we consider an irreducible rectangle $[0, x_K] \times [0, y_L]$.

3.1. Rectangle with no Given Values Inside

Suppose that the set $G = B_{K,L}$ has an empty intersection with the interior of the rectangle $]0, x_K[\times]0, y_L[$. This is an instance of Problem 4; equations (22) reduce to monotonicity.

Problem 5. Let $0 = x_0 < \dots < x_K \leq 1$, $0 = y_0 < \dots < y_L \leq 1$, and let

$$G = \{(x_i, 0), (x_i, L) \mid i = 0, \dots, I\} \cup \{(0, y_j), (K, y_j) \mid j = 0, \dots, J\}.$$

Suppose that values $z_{i,j} \in [0, 1]$ are given for all $(i, j) \in G$ so that (7) is satisfied for all $i \in \{0, \dots, K\}$ and $j \in \{0, \dots, L\}$ and sequences $(z_{i,L})_{i=0}^K, (z_{K,j})_{j=0}^L$ are nondecreasing. Find values $n_{i,j} \in [0, 1], (i, j) \in \{1, \dots, K\} \times \{1, \dots, L\}$, satisfying

$$\sum_{u=1}^K n_{u,j} = z_{K,j}, \quad (24)$$

$$\sum_{v=1}^L n_{i,v} = z_{i,L}, \quad (25)$$

and such that the contribution to the differential entropy

$$-\sum_{i=1}^K \sum_{j=1}^L n_{i,j} \ln \left(\frac{n_{i,j}}{(x_i - x_{i-1})(y_j - y_{j-1})} \right) \quad (26)$$

⁸ From now on, we denote this set again by G , although this was used for all given values from $\{0, \dots, I\} \times \{0, \dots, J\}$ before.

is maximal.

The solution need not be a copula with a constant density at $]0, x_K[\times]0, y_L[$ because we must keep the correct row and column sums.

We use Lagrange multipliers $\lambda_1, \dots, \lambda_K, \mu_1, \dots, \mu_L$; the Lagrange function is

$$L(n_{i,j}, \lambda_i, \mu_j) = \sum_{i=1}^K \sum_{j=1}^L -n_{i,j} \ln \frac{n_{i,j}}{a_i b_j} + \sum_{i=1}^K \lambda_i \left(\sum_{j=1}^L n_{i,j} - s_i \right) + \sum_{j=1}^L \mu_j \left(\sum_{i=1}^K n_{i,j} - r_j \right).$$

We put all partial derivatives equal to zero and obtain a system of equations

$$\frac{\partial L}{\partial n_{i,j}} = -1 - \ln \frac{n_{i,j}}{a_i b_j} + \lambda_i + \mu_j = 0,$$

where $i \in \{1, 2, \dots, K\}, j \in \{1, 2, \dots, L\}$. Equivalently,

$$\ln \frac{n_{i,j}}{a_i b_j} = -1 + \lambda_i + \mu_j.$$

Subtracting the last ones, we obtain

$$\begin{aligned} \ln \frac{n_{i,j}}{a_i b_j} - \ln \frac{n_{K,j}}{a_K b_j} &= \lambda_i - \lambda_K, \\ \ln \frac{n_{i,j}}{a_i b_j} - \ln \frac{n_{i,L}}{a_i b_L} &= \mu_j - \mu_L, \end{aligned}$$

and express $n_{i,j}$ using $n_{K,j}$ or $n_{i,L}$ as

$$\begin{aligned} n_{i,j} &= n_{K,j} \frac{a_i}{a_K} \exp(\lambda_i - \lambda_K), \\ n_{i,j} &= n_{i,L} \frac{b_j}{b_L} \exp(\mu_j - \mu_L). \end{aligned}$$

We see that all rows (resp. columns) are multiples of the last one and the matrix composed of all $n_{i,j}$ is a dyad (a product of two vectors),

$$\begin{bmatrix} n_{1,1} & \dots & n_{1,K} \\ \vdots & \ddots & \vdots \\ n_{L,1} & \dots & n_{L,K} \end{bmatrix} = c \begin{bmatrix} r_1 \\ \vdots \\ r_m \end{bmatrix} \begin{bmatrix} s_1 & \dots & s_t \end{bmatrix}, \quad (27)$$

where r_j are the row sums, s_i the column sums, and the constant c can be determined from the known total sum

$$\begin{aligned} c \sum_{j=1}^L \sum_{i=1}^K s_i r_j &= \sum_{j=1}^L r_j = \sum_{i=1}^K s_i, \\ c &= \frac{\sum_{j=1}^L r_j}{\sum_{i=1}^L \sum_{j=1}^K s_j r_i} = \frac{1}{\sum_{i=1}^K s_i}. \end{aligned}$$

This is an expected result that says that the probability density of a result in an elementary rectangle is the product of the probability densities of the marginal distributions.

3.2. Methodology of Solving Particular Cases

Here we collect ideas used to solve particular cases described in the appendices. We consider Problem 4. This leads to a system of linear equations; after the Gauss–Jordan elimination, it could look like

$$M = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & \dots & 0 & 1 & C_1 \\ 0 & 1 & 1 & 0 & 0 & \dots & 0 & 0 & C_2 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 & 1 & C_3 \\ 0 & 0 & 0 & 0 & 1 & \dots & 0 & 1 & C_4 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 & 1 & C_N \end{bmatrix},$$

where $C_1, \dots, C_N$ are some constants computed from the values $z_{i,j}$. Until we consider the maximized differential entropy, some variables $n_{i,j}$ can be chosen and the rest can be computed from them. This describes all possible solutions to the system of linear equations (ignoring their bounds). To maximize the differential entropy, we compute the partial derivatives of the contribution to the differential entropy and put them equal to zero. We obtain

$$N = KL - \text{card } G = KL - K - L + 1 - o$$

equations, where $o = \text{card}(G \cap]0, x_K[\times]0, y_L[)$ is the number of given values inside the rectangle.

Theorem 1. *The optimal solution to Problem 4 is independent of the values $a_1, \dots, a_K, b_1, \dots, b_L$.*

Proof. Recall that

$$n_{i,j} = z_{i,j} - z_{i-1,j} - z_{i,j-1} + z_{i-1,j-1}. \quad (28)$$

This leads to an optimization in those variables $z_{i,j} = C(x_i, y_j)$ where the copula values are not given $((i, j) \notin G)$.

In the criterion, each variable $z_{i,j}$, $(i, j) \in G \setminus B_{K,L}$, occurs in four equations; for $n_{i,j}$ and $n_{i+1,j+1}$ with coefficient 1 and for $n_{i+1,j}$ and $n_{i,j+1}$ with coefficient -1 .

We solve this task by computing the partial derivatives w.r.t. unknown values $z_{i,j}$ and putting them zero,

$$\begin{aligned} & \ln \frac{z_{i,j} - z_{i-1,j} - z_{i,j-1} + z_{i-1,j-1}}{a_i b_j} + 1 - \ln \frac{z_{i,j+1} - z_{i-1,j+1} - z_{i,j} + z_{i-1,j}}{a_i b_{j+1}} - 1 \\ & - \ln \frac{z_{i+1,j} - z_{i,j} - z_{i+1,j-1} + z_{i,j-1}}{a_{i+1} b_j} - 1 + \ln \frac{z_{i+1,j+1} - z_{i+1,j} - z_{i,j+1} + z_{i,j}}{a_{i+1} b_{j+1}} + 1 = 0. \end{aligned} \quad (29)$$

This can be expressed as

$$\begin{aligned} & \frac{z_{i,j} - z_{i-1,j} - z_{i,j-1} + z_{i-1,j-1}}{a_i b_j} \cdot \frac{z_{i+1,j+1} - z_{i+1,j} - z_{i,j+1} + z_{i,j}}{a_{i+1} b_{j+1}} \\ & = \frac{z_{i,j+1} - z_{i-1,j+1} - z_{i,j} + z_{i-1,j}}{a_i b_{j+1}} \cdot \frac{z_{i+1,j} - z_{i,j} - z_{i+1,j-1} + z_{i,j-1}}{a_{i+1} b_j}. \end{aligned}$$

Multiplying this by $a_i b_j a_{i+1} b_{j+1}$, we obtain an equation independent of a_i, b_j, a_{i+1} , and b_{j+1} . This holds for all $z_{i,j}$ and $n_{i,j}$, $(i, j) \in \{1, \dots, K-1\} \times \{1, \dots, L-1\} \setminus G$. Due to (28), also the values $n_{i,j}$ are independent of $a_1, \dots, a_K, b_1, \dots, b_L$. $\square$

Besides the independence of the result on interval lengths $a_1, \dots, a_K, b_1, \dots, b_L$, we proved that each equation in the system describing entropy maximization is of the form

$$\prod_{\kappa_1 \in Z_+} (n_{i,j} + B_{\kappa_1}) = \prod_{\kappa_2 \in Z_-} (-n_{i,j} + B_{\kappa_2}),$$

where Z_+ is the set of indices (i, j) for which $n_{i,j}$ occurs with the positive sign and Z_- is the set of indices for which $n_{i,j}$ occurs with the negative sign. Apparently, $\text{card } Z_+ = \text{card } Z_-$, therefore the units in (29) cancelled.

The only problem is that $n_{i,j}$ may depend also on other variables. Thus the summands B_* contain not only constants C_* , but also other independent variables.

In the appendices, we present explicit solutions to some specific instances of Problem 4. They lead to higher-order algebraic equations.

4. Conclusion

We formulated the task of fitting a continuous copula to finitely many given values in such a way that the entropy of the copula density is maximal. This is motivated by situations when some of the values are prescribed and the rest of the copula should be “as independent as possible”, with the intention not to include any other dependence than that contained in the constraints. We simplified the task with several hints, showed that it has a unique solution (because it is equivalent to a convex optimization problem), and demonstrated that a closed-form solution can be found analytically in some simple cases, although it may require solving higher-order polynomials. We propose this concept as an alternative to current approaches, which also maximize the entropy, but use different restrictions of the admitted copulas.

Author Contributions: Writing – original draft, Milan Bubák and Mirko Navara. All authors will be updated at each stage of manuscript processing, including submission, revision, and revision reminder, via emails from our system or the assigned Assistant Editor.

Funding: This research was supported by the CTU institutional support (Future Fund).

Data Availability Statement: This research is not accompanied by any additional data.

Acknowledgments: The authors are grateful to Michal Dibala who inspired and initiated this research. They thank the anonymous reviewers for their comments which helped to improve the original manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

In Problem 4, we put $K = L = 3$, $G = B_{K,L} \cup \{(1, 1), (2, 2)\}$ (see Figure A1).

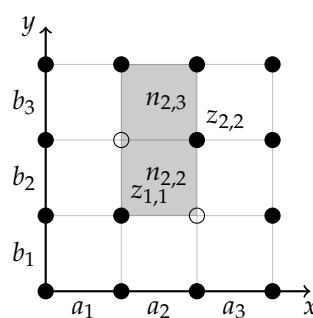


Figure A1. Example 1

The task can be formulated as follows.

$$\begin{aligned} \max_{n_{i,j}} \quad & \sum_{i=1}^3 \sum_{j=1}^3 -n_{i,j} \ln \frac{n_{i,j}}{a_i b_j}, \quad i, j \in \{1; 2; 3\}, \\ \text{under conditions} \quad & n_{i,j} \geq 0, \\ & n_{1,1} = z_{1,1}, \\ & n_{1,1} + n_{2,1} + n_{3,1} = z_{3,1}, \\ & n_{1,1} + n_{2,1} + n_{1,2} + n_{2,2} = z_{2,2}, \\ & n_{1,1} + n_{2,1} + n_{3,1} + n_{1,2} + n_{2,2} + n_{3,2} = z_{3,2}, \\ & n_{1,1} + n_{1,2} + n_{1,3} = z_{1,3}, \\ & n_{1,1} + n_{2,1} + n_{1,2} + n_{2,2} + n_{1,3} + n_{2,3} = z_{2,3}, \\ & n_{1,1} + n_{2,1} + n_{3,1} + n_{1,2} + n_{2,2} + n_{3,2} + n_{1,3} + n_{2,3} + n_{3,3} = z_{3,3}. \end{aligned}$$

We take $n_{2,2}$, $n_{2,3}$ as variables allowing us to express the general solution of the system of linear equations; we introduce new constants C_* as abbreviations.

$$\begin{aligned} n_{1,1} &= z_{1,1} = C_1, \\ n_{2,1} &= -n_{2,2} - n_{2,3} \underbrace{-z_{1,3} + z_{2,3}}_{C_2} = -n_{2,2} - n_{2,3} + C_2, \\ n_{3,1} &= n_{2,2} + n_{2,3} \underbrace{-z_{1,1} + z_{3,1} + z_{1,3} - z_{2,3}}_{C_3} = n_{2,2} + n_{2,3} + C_3, \\ n_{1,2} &= n_{2,3} \underbrace{-z_{1,1} + z_{2,2} + z_{1,3} - z_{2,3}}_{C_4} = n_{2,3} + C_4, \\ n_{3,2} &= -n_{2,2} - n_{2,3} \underbrace{+z_{1,1} - z_{3,1} - z_{2,2} + z_{3,2} - z_{1,3} + z_{2,3}}_{C_5} = -n_{2,2} - n_{2,3} + C_5, \\ n_{1,3} &= -n_{2,3} \underbrace{-z_{2,2} + z_{2,3}}_{C_6} = -n_{2,3} + C_6, \\ n_{3,3} &= \underbrace{z_{2,2} - z_{3,2} - z_{2,3} + z_{3,3}}_{C_7} = C_7. \end{aligned}$$

Substituting this in the optimality criterion and putting its partial derivatives equal to zero, we obtain

$$\begin{aligned} & \ln \frac{C_2 - n_{2,2} - n_{2,3}}{a_2 b_1} + 1 - \ln \frac{C_3 + n_{2,2} + n_{2,3}}{a_3 b_1} - 1 - \ln \frac{n_{2,2}}{a_2 b_2} - 1 \\ & + \ln \frac{C_5 - n_{2,2} - n_{2,3}}{a_3 b_2} + 1 = 0, \\ & \ln \frac{C_2 - n_{2,2} - n_{2,3}}{a_2 b_1} + 1 - \ln \frac{C_3 + n_{2,2} + n_{2,3}}{a_3 b_1} - 1 - \ln \frac{C_4 + n_{2,3}}{a_1 b_2} - 1 \\ & + \ln \frac{C_5 - n_{2,2} - n_{2,3}}{a_3 b_2} + 1 + \ln \frac{C_6 - n_{2,3}}{a_1 b_3} + 1 - \ln \frac{n_{2,3}}{a_2 b_3} - 1 = 0. \end{aligned}$$

These equations can be simplified to

$$\begin{aligned} \ln \frac{(C_2 - n_{2,2} - n_{2,3})(C_5 - n_{2,2} - n_{2,3})}{a_2 b_1 a_3 b_2} &= \ln \frac{(C_3 + n_{2,2} + n_{2,3})n_{2,2}}{a_3 b_1 a_2 b_2}, \\ \ln \frac{(C_2 - n_{2,2} - n_{2,3})(C_5 - n_{2,2} - n_{2,3})(C_6 - n_{2,3})}{a_2 b_1 a_3 b_2 a_1 b_3} &= \ln \frac{(C_3 + n_{2,2} + n_{2,3})(C_4 + n_{2,3})n_{2,3}}{a_3 b_1 a_1 b_2 a_2 b_3}. \end{aligned}$$

The divisors are equal and thus cancel out, as shown in the proof of Theorem 1, and this leads to the expression

$$\begin{aligned} (C_2 - n_{2,2} - n_{2,3})(C_5 - n_{2,2} - n_{2,3}) &= (C_3 + n_{2,2} + n_{2,3})n_{2,2}, \\ (C_2 - n_{2,2} - n_{2,3})(C_5 - n_{2,2} - n_{2,3})(C_6 - n_{2,3}) &= (C_3 + n_{2,2} + n_{2,3})(C_4 + n_{2,3})n_{2,3}. \end{aligned} \quad (A1)$$

We divide the second equation by the first,

$$C_6 - n_{2,3} = \frac{(C_4 + n_{2,3})n_{2,3}}{n_{2,2}},$$

and $n_{2,2}$ can be expressed as

$$n_{2,2} = \frac{(C_4 + n_{2,3})n_{2,3}}{(C_6 - n_{2,3})}. \quad (A2)$$

Substituting to (A1) and expanding, we obtain an equation for $n_{2,3}$,

$$\begin{aligned} & (C_3 - C_4 - C_6)n_{2,3}^3 \\ & + (C_2C_6 + C_2C_4 + C_2C_5 + C_6^2 + C_4C_6 + C_5C_6 - C_3C_6 + C_4C_5 + C_3C_4)n_{2,3}^2 \\ & + (-C_2C_6^2 - C_2C_4C_6 - C_2C_5C_6 - C_5C_6^2 - C_4C_5C_6 - C_3C_4C_6)n_{2,3} \\ & + C_2C_5C_6^2 = 0. \end{aligned}$$

This is an algebraic equation of order three. It can be solved analytically using Cardano formulas. A problem remains that it can have up to three roots; we want one that is real. The cubic coefficient is

$$C_3 - C_4 - C_6 = z_{3,1} - z_{2,3}.$$

If it is nonzero, there is at least one real root. Moreover, we need it in the interval $[0, 1]$. Substituting to (A2), we obtain $n_{2,2}$, which should be also in the interval $[0, 1]$, as well as all other unknowns. However, we know (from the convexity of the task, resp. concavity of the maximized criterion) that a solution with this property exists.

Appendix B

In Problem 4, we put $K = 3$, $L = 4$, $G = B_{K,L} \cup \{(1, 2), (2, 1), (2, 3)\}$ (see Figure A2).

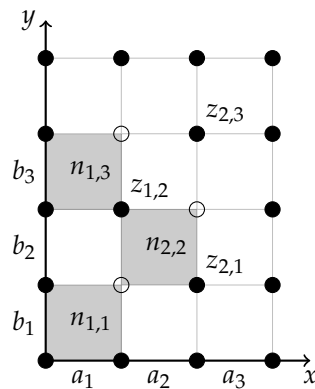


Figure A2. Example 3

The task can be formulated as follows.

$$\begin{aligned} \max_{n_{i,j}} \quad & \sum_{i=1}^3 \sum_{j=1}^4 -n_{i,j} \ln \frac{n_{i,j}}{a_i b_j}, \quad i \in \{1; 2; 3\}, \quad j \in \{1; 2; 3; 4\}, \\ \text{under conditions} \quad & n_{i,j} \geq 0, \\ & n_{2,1} = C_1 - n_{1,1}, \\ & n_{1,2} = C_2 - n_{1,1}, \\ & n_{3,2} = C_3 + n_{1,1} - n_{2,2}, \\ & n_{2,3} = C_4 + n_{1,1} - n_{2,2} - n_{1,3}, \\ & n_{3,3} = C_5 - n_{1,1} + n_{2,2}, \\ & n_{1,4} = C_6 - n_{1,3}, \\ & n_{2,4} = C_7 + n_{1,3}, \end{aligned}$$

where the constants C_* are some sums of $\pm z_{i,j}$ as in the previous example. We take $n_{1,1}$, $n_{2,2}$, $n_{1,3}$ as variables allowing us to express the general solution of the system of linear equations.

We put the partial derivatives equal to zero and eliminate the logarithms. We obtain products of factors in the form of a sum of some $\pm n_{i,j}$ and a constant C_* . We also know that the result is independent of the values a_i , b_j . In particular, we obtain

$$\begin{aligned} \frac{\partial}{\partial n_{1,1}} : \quad & (C_1 - n_{1,1})(C_2 - n_{1,1})(C_5 - n_{1,1} + n_{2,2}) \\ & = n_{1,1}(C_3 + n_{1,1} - n_{2,2})(C_4 + n_{1,1} - n_{2,2} - n_{1,3}), \\ \frac{\partial}{\partial n_{2,2}} : \quad & (C_3 + n_{1,1} - n_{2,2})(C_4 + n_{1,1} - n_{2,2} - n_{1,3}) = (C_5 - n_{1,1} + n_{2,2})n_{2,2}, \\ \frac{\partial}{\partial n_{1,3}} : \quad & (C_4 + n_{1,1} - n_{2,2} - n_{1,3})(C_6 - n_{1,3}) = n_{1,3}(C_7 + n_{1,3}). \end{aligned} \quad (\text{A3})$$

From (A3) we separate the factors containing only the variable $n_{1,1}$,

$$\frac{(C_1 - n_{1,1})(C_2 - n_{1,1})}{n_{1,1}} = \frac{(C_3 - n_{2,2} + n_{1,1})(C_4 + n_{1,1} - n_{2,2} - n_{1,3})}{C_5 - n_{1,1} + n_{2,2}}. \quad (\text{A4})$$

The second equation in (A3) can be written as

$$\frac{(C_3 - n_{2,2} + n_{1,1})(C_4 + n_{1,1} - n_{2,2} - n_{1,3})}{C_5 - n_{1,1} + n_{2,2}} = n_{2,2}. \quad (\text{A5})$$

We substitute (A4) to the left-hand side and express $n_{2,2}$ as

$$n_{2,2} = \frac{(C_1 - n_{1,1})(C_2 - n_{1,1})}{n_{1,1}}. \quad (\text{A6})$$

The third equation in (A3) can be rewritten as

$$\begin{aligned} C_4 + n_{1,1} - n_{2,2} - n_{1,3} &= \frac{n_{1,3}(C_7 + n_{1,3})}{C_6 - n_{1,3}}, \\ C_4 + n_{1,1} - n_{2,2} &= \frac{n_{1,3}(C_7 + n_{1,3}) + n_{1,3}(C_6 - n_{1,3})}{C_6 - n_{1,3}}, \\ C_4 + n_{1,1} - n_{2,2} &= \frac{n_{1,3}(C_7 + C_6)}{C_6 - n_{1,3}}, \\ C_4 + n_{1,1} - n_{2,2} &= -(C_7 + C_6) + \frac{(C_7 + C_6)C_6}{C_6 - n_{1,3}}. \end{aligned} \quad (\text{A7})$$

In the last equation of (A7), we substitute $n_{2,2}$ from (A6) and obtain $n_{1,3}$:

$$\begin{aligned}
 C_4 + n_{1,1} - \frac{(C_1 - n_{1,1})(C_2 - n_{1,1})}{n_{1,1}} &= -(C_7 + C_6) + \frac{(C_7 + C_6)C_6}{C_6 - n_{1,3}}, \\
 \frac{(C_4 + n_{1,1})n_{1,1} - (C_1 - n_{1,1})(C_2 - n_{1,1}) + (C_7 + C_6)n_{1,1}}{n_{1,1}} &= \frac{(C_7 + C_6)C_6}{C_6 - n_{1,3}}, \\
 \frac{n_{1,1}(C_1 + C_2 + C_4 + C_6 + C_7) - C_1C_2}{n_{1,1}} &= \frac{(C_7 + C_6)C_6}{C_6 - n_{1,3}}, \\
 \frac{n_{1,1}}{n_{1,1}(C_1 + C_2 + C_4 + C_6 + C_7) - C_1C_2} &= \frac{C_6 - n_{1,3}}{(C_7 + C_6)C_6}, \\
 \frac{n_{1,1}(C_7 + C_6)C_6}{n_{1,1}(C_1 + C_2 + C_4 + C_6 + C_7) - C_1C_2} &= C_6 - n_{1,3}, \\
 \frac{C_6(n_{1,1}(C_1 + C_2 + C_4 + C_6 + C_7) - C_1C_2) - n_{1,1}(C_7 + C_6)C_6}{n_{1,1}(C_1 + C_2 + C_4 + C_6 + C_7) - C_1C_2} &= n_{1,3}, \\
 \frac{C_6(n_{1,1}(C_1 + C_2 + C_4) - C_1C_2)}{n_{1,1}(C_1 + C_2 + C_4 + C_6 + C_7) - C_1C_2} &= n_{1,3}.
 \end{aligned} \tag{A8}$$

We have both $n_{2,2}$ in (A6) and $n_{1,3}$ in (A8) expressed as functions of $n_{1,1}$. This can be substituted to the first equation in (A3), which yields

$$\begin{aligned}
 L : & (C_1 - n_{1,1})(C_2 - n_{1,1}) \left(C_5 - n_{1,1} + \frac{(C_1 - n_{1,1})(C_2 - n_{1,1})}{n_{1,1}} \right), \\
 R : & n_{1,1} \left(C_3 - \frac{(C_1 - n_{1,1})(C_2 - n_{1,1})}{n_{1,1}} + n_{1,1} \right) \left(C_4 + n_{1,1} - \frac{(C_1 - n_{1,1})(C_2 - n_{1,1})}{n_{1,1}} \right. \\
 & \left. - \frac{C_6(n_{1,1}(C_1 + C_2 + C_4) - C_1C_2)}{n_{1,1}(C_1 + C_2 + C_4 + C_6 + C_7) - C_1C_2} \right),
 \end{aligned} \tag{A9}$$

where L (resp. R) denotes the left-hand (resp. right-hand) side of the equation. We modify this expression to discuss its analytical solvability:

$$\begin{aligned}
 L : & \frac{(n_{1,1}^2 - (C_1 + C_2)n_{1,1} + C_1C_2)((-C_1 - C_2 + C_5)n_{1,1} - C_1C_2)}{n_{1,1}}, \\
 R : & ((C_1 + C_2 + C_3)n_{1,1} - C_1C_2) \\
 & \cdot \frac{(n_{1,1}^2(C_1 + C_2 + C_4)(C_1 + C_2 + C_4 + C_7) - n_{1,1}C_1C_2(2C_1 + 2C_2 + 2C_4 + 2C_6 + C_7) + C_1^2C_2^2)}{n_{1,1}(n_{1,1}(C_1 + C_2 + C_4 + C_6 + C_7) - C_1C_2)}.
 \end{aligned} \tag{A10}$$

Multiplication by a common divisor on both sides gives

$$\begin{aligned}
 L : & (n_{1,1}^2 - (C_1 + C_2)n_{1,1} + C_1C_2)((-C_1 - C_2 + C_5)n_{1,1} - C_1C_2) \\
 & \cdot ((C_1 + C_2 + C_4 + C_6 + C_7)n_{1,1} - C_1C_2), \\
 R : & ((C_1 + C_2 + C_3)n_{1,1} - C_1C_2) \left(n_{1,1}^2(C_1 + C_2 + C_4)(C_1 + C_2 + C_4 + C_7) \right. \\
 & \left. - n_{1,1}C_1C_2(2C_1 + 2C_2 + 2C_4 + 2C_6 + C_7) + C_1^2C_2^2 \right).
 \end{aligned} \tag{A11}$$

After expansion, the left-hand side is a polynomial of order 4 and the right-hand side is a polynomial of order 3. Their difference is a polynomial of order 4, which is solvable analytically.

The remaining variables are obtained by substitution to (A6) and (A8). This results in the set of all solutions; substitution of the feasible solution to the equation for differential entropy allows us to determine the maximum.

References

1. Nelsen, R.B. *An Introduction to Copulas*, 2nd ed.; Vol. 139, *Lecture Notes in Statistics*, Springer, New York, NY, 2006.
2. Sklar, A. *Fonctions de répartition à n dimensions et leurs marges*; Vol. 8, Publ. Inst. Statist. Univ. Paris, 1959; pp. 229–231.
3. Shannon, C.E. A Mathematical Theory of Communication. *The Bell System Technical Journal* **1948**, *27*, 379–423.
4. Jaynes, E.T. Information theory and statistical mechanics. *Physical Review* **1957**, *106*, 620–628.
5. Pougaza, D.B.; Mohammad-Djafari, A. Maximum Entropies Copulas. In Proceedings of the 30th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, France, 2010; pp. 329–336, [https://pubs.aip.org/aip/acp/article-pdf/1305/1/329/11567694/329_1_online.pdf].
<https://doi.org/10.1063/1.3573634>.
6. Pougaza, D.B.; Mohammad-Djafari, A.; cois Bercher, J.F. Link between copula and tomography. *Pattern Recognition Letters* **2010**, *31*, 2258–2264. <https://doi.org/10.1016/j.patrec.2010.05.001.hal-00509705>.
7. Ma, J.; Sun, Z. Mutual information is copula entropy. *Tsinghua Science & Technology* **2011**, *16*, 51–54. [https://doi.org/https://doi.org/10.1016/S1007-0214\(11\)70008-6](https://doi.org/https://doi.org/10.1016/S1007-0214(11)70008-6).
8. Singh, V.P.; Zhang, L. Copula–entropy theory for multivariate stochastic modeling in water engineering. *Geoscience Letters* **2018**, *5*. <https://doi.org/10.1186/s40562-018-0105-z>.
9. Piantadosi, J.; Howlett, P.; Boland, J.W. Matching the grade correlation coefficient using a copula with maximum disorder. *Journal of Industrial and Management Optimization* **2007**, *3*, 305–312.
10. Piantadosi, J.; Howlett, P.; Borwein, J. Copulas with Maximum Entropy. *Optimization Letters* **2012**, *6*, 99–125.
11. Lin, L.; Wang, R.; Zhang, R.; Zhao, C. The checkerboard copula and dependence concepts. *ArXiv e-prints*, [[arXiv:2404.15023](https://arxiv.org/abs/2404.15023)]. Accessed 2024/10/07.
12. Genest, C.; Nešlehová, J.G.; Rémillard, B. Asymptotic behavior of the empirical multilinear copula process under broad conditions. *Journal of Multivariate Analysis* **2017**, *159*, 82–110.
13. Dibala, M.; Navara, M. Discrete Copulas and Maximal Entropy Principle. In Proceedings of the Copulas and Their Applications, Almeria, Spain, 2017; p. 24.
14. Bubák, M. Copulas with Maximal Entropy (in Czech). BSc. Thesis, Czech Technical University in Prague, 2024, [<http://hdl.handle.net/10467/115430>].
15. Bertsekas, D.P.; Nedić, A.; Ozdaglar, A.E. *Convex Analysis and Optimization*; Athena Scientific, 2003.
16. Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press, 2004.
17. Hiriart-Urruty, J.; Lemaréchal, C. *Fundamentals of Convex Analysis*; Grundlehren Text Editions, Springer, 2004.
18. Rockafellar, R.T. *Convex Analysis*; Princeton Mathematical Series, Princeton University Press, 1970.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.