# Scaling limits of discrete copulas are bridged Brownian sheets

Juliana Freire, Nicolau C. Saldanha and Carlos Tomei

January 14, 2016

**Abstract**

For large $n$, take a random $n \times n$ permutation matrix and its associated discrete copula $X_n$. For $a, b = 0, 1, \ldots, n$, let $\mathbf{y}_n(\frac{a}{n}, \frac{b}{n}) = \frac{1}{n}\left(X_{a,b} - \frac{ab}{n}\right)$; define $\mathbf{y}_n : [0,1]^2 \to \mathbb{R}$ by interpolating quadratically on squares of side $\frac{1}{n}$. We prove a Donsker type central limit theorem: $\sqrt{n}\mathbf{y}_n$ approaches a bridged Brownian sheet on the unit square.

*MSC2010*: Primary 60C05, 60F05, 60J65; Secondary 60B20.
*Keywords*: Discrete copulas, Brownian sheets, Donsker theorem.

## 1 Introduction

Mostly following [7], a *discrete copula* is a square matrix $C$ with indices $a, b \in \{0, 1, \ldots, n\}$ and *integer* entries satisfying

(i) $C_{a,0} = C_{0,b} = 0$ for all $a, b$;

(ii) $C_{a,n} = C_{n,a} = a$ for all $a$;

(iii) $C$ is 2-increasing, i.e., if $0 \leq a \leq a' \leq n$ and $0 \leq b \leq b' \leq n$, then

$$C_{a',b'} + C_{a,b} \geq C_{a',b} + C_{a,b'}.$$

Let $S_n$ be symmetric group interpreted as the set of all $n \times n$ permutation matrices; also, denote the set of discrete copulas by $\mathcal{S}_n$. As is well known ([7]), there is a natural bijection $S_n \to \mathcal{S}_n$ with $M \mapsto C$ where $C$ is obtained by (discrete) integration:

$$C_{a,b} = \sum_{a' \leq a;\ b' \leq b} M_{a',b'}.$$

Notice that if $a = 0$ or $b = 0$ the sum is empty and $C_{a,b} = 0$.

A continuous copula is a 1-Lipschitz function $\mathbf{c} : \mathbb{I}^2 \to \mathbb{I}$ which describes the joint distribution of two real-valued random variables $\mathbf{z}_0$ and $\mathbf{z}_1$, assumed
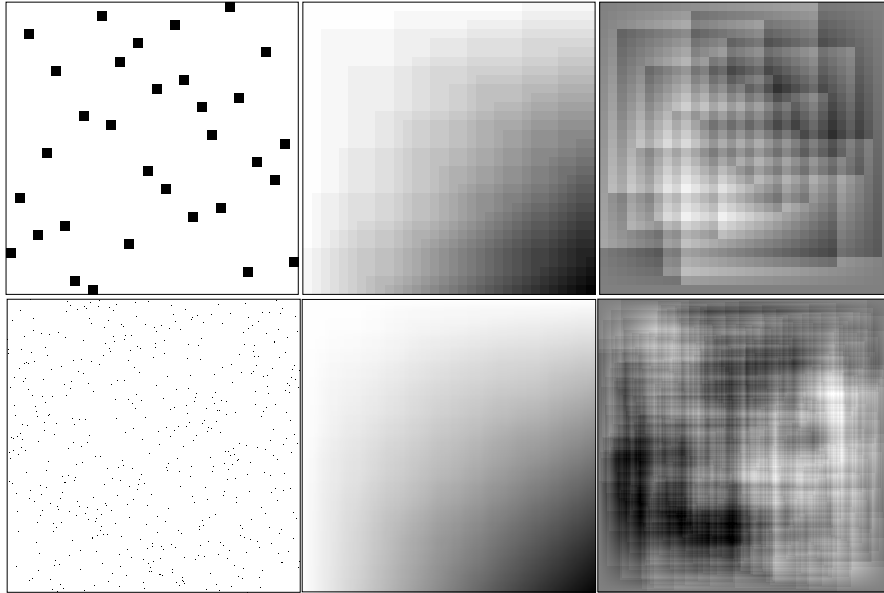
Figure 1: Two permutation matrices, their associated discrete copula $C$ and the (scaled) difference $D = C - C_0$; $n = 32$ and $n = 512$.

here to be both uniform in $\mathbb{I} = [0, 1]$ (for basic information about copulas, see [8]). If $\mathbf{z}_0$ and $\mathbf{z}_1$ are independent then $\mathbf{c}$ is the product copula $\mathbf{c}_0(u, v) = uv$. Being associated with a permutation matrix, a discrete copula describes the joint distribution of two discrete random variables $Z_0$ and $Z_1$, both uniform in $\{1, 2, \ldots, n\}$, when one is a function of the other. Let $C_0$, the *product copula*, be the baricenter of $\mathcal{S}_n$, so that $(C_0)_{a,b} = \frac{ab}{n}$; it is analogous to the (continuous) product copula and indicates that $Z_0$ and $Z_1$ are independent. Clearly, $C_0$ is not a discrete copula: it is an example of Birkhoff copula, a concept to which we shall return in Section 8. We scale each discrete copula $C$ to obtain a continuous copula $\mathbf{c} : \mathbb{I}^2 \to \mathbb{I}$ satisfying

$$\mathbf{c}\left(\frac{a}{n}, \frac{b}{n}\right) = \frac{C_{a,b}}{n},$$

and interpolated on small squares $[\frac{a-1}{n}, \frac{a}{n}] \times [\frac{b-1}{n}, \frac{b}{n}]$ by a polynomial of the form $\mathbf{c}(u, v) = a_{00} + a_{10}u + a_{01}v + a_{11}uv$. The statistics of the differences $D = C - C_0$ and $\mathbf{d} = \mathbf{c} - \mathbf{c}_0$, appropriately scaled, are the main subject of this paper.

In Figure 1, the nonzero entries of a permutation matrix are represented by solid black squares. For the associated discrete copula $C$, each entry is represented by a grey square, with 0 being white and $n$ (the largest value) black. Finally, for the difference $D$, entries with $D_{i,j} \geq \sqrt{n}/2$ are painted black and with $D_{i,j} \leq -\sqrt{n}/2$, white.

2

One of our aims is to obtain a concentration result: a typical discrete copula is near $C_0$. Indeed, in Figure 1 the central squares are almost indistinguishable. We shall consider discrete random variables $X_n$ assuming matrix values in the finite set $\mathcal{S}_n$ with uniform distribution and entries $(X_n)_{a,b} = X_{n,a,b}$. We also consider the differences $Y_n = X_n - C_0$. Just as we extended $C$ or $D$ to $\mathbf{c}$ or $\mathbf{d}$, we define the extensions $\mathbf{x}_n \in C^0(\mathbb{I}^2)$ and $\mathbf{y}_n \in C_0^0(\mathbb{I}^2)$ (i.e., $\mathbf{y}_n|_{\partial \mathbb{I}^2} = 0$) of $X_n$ and $Y_n$.

We shall see that, with high probability, the values assumed by $\mathbf{y}_n$ have order at most $1/\sqrt{n}$. More: $\sqrt{n}\,\mathbf{y}_n$ approaches a bridged Brownian sheet, compatibly with the patterns in the rightmost squares in Figure 1.

Recall that a *Brownian sheet* is a process yielding continuous functions $f_* : \mathbb{I}^2 \to \mathbb{R}$ (see [6]); this generalizes to a two-dimensional domain the more familiar Brownian motion or Wiener process (see [2]). In order to select $f_*$, first generate white noise $\eta$ in the unit square $\mathbb{I}^2$ and then integrate it:

$$f_*(u,v) = \iint_{[0,u] \times [0,v]} \eta\,.$$

In particular, $f_*|_{((\mathbb{I} \times \{0\}) \cup (\{0\} \times \mathbb{I}))} = 0$. We are interested in the *bridged Brownian sheet*, a modified process which obtains functions $f \in C_0^0(\mathbb{I}^2)$. In order to select $f$ from $f_*$, define

$$f(u,v) = f_*(u,v) - u f_*(1,v) - v f_*(u,1) + uv f_*(1,1). \tag{1}$$

It turns out that the values of $f$ on a fixed set of points follow a joint normal distribution (see Equation 8). We are particularly interested in the square grid $I = J = m$, $u_i = \frac{i}{m}$, $v_j = \frac{j}{m}$. Large permutation matrices satisfy similar statistics on coarse $m \times m$ grids: for $\alpha = 19/20$, if $m < n^{(1-\alpha)}$ then the function $\sqrt{n}\,\mathbf{y}_n$ behaves as $f$ in such grids (Lemmas 5.1 and 5.2). Still, the grid is fine enough for errors to be controlled (Lemma 4.2).

We are ready to state our main result. As with Donsker's theorem, this is a kind of central limit theorem, in which a Brownian bridge or bridged sheet is obtained as a limit of discrete processes. We use Dudley's framework as in [3].

**Theorem 1.1.** *There exist probability spaces $\Omega_n$ with the following properties.*

(a) *On $\Omega_n$ there exists a random variable $X_n$ uniformly distributed in $\mathcal{S}_n$.*

(b) *On $\Omega_n$ a bridged Brownian sheet process $f$ in $C_0^0(\mathbb{I}^2)$ is defined.*

(c) *For all $\epsilon > 0$,*
$$\lim_{n \to \infty} \Pr\left[\left\|\sqrt{n}\,\mathbf{y}_n - f\right\|_{C^0} > \epsilon\right] = 0,$$
*where $\mathbf{y}_n$ is defined from $X_n$ as above.*

In Section 2 we study the pointwise behavior of $\mathbf{y}$ for the case of discrete copulas. One of the main results is Lemma 2.5; its Corollary 2.6 shows that,

given $(u, v) \in \mathbb{I}^2$, the distribution of $\sqrt{n}\, \mathbf{y}_n(u, v)$ approaches a normal distribution, in accordance with the results for bridged Brownian sheets mentioned above. In Section 3 we state and prove Theorem 3.1: if $m \in \mathbb{N}$ and points $(u_1, v_1), \ldots, (u_m, v_m) \in \mathbb{I}^2$ are fixed and $n$ goes to infinity, then the joint distribution of $\sqrt{n}\, \mathbf{y}_n(u_i, v_i)$ converges to the corresponding distribution for bridged Brownian sheets. Section 4 is centered on Lemma 4.2, a Hölder type estimate for $|D_{a_2, b_2} - D_{a_1, b_1}|$ provided $|a_2 - a_1|$ and $|b_2 - b_1|$ are small; this goes together with the fact that Brownian sheets are Hölder continuous for any exponent smaller than $\frac{1}{2}$ (see Lemma 4.1). In Section 5 we complete the proof of Theorem 1.1: the key difference from Section 3 is that now the size of the grid goes to infinity as $m \approx n^{\frac{1}{20}}$. Section 6 contains, as an application of Theorem 1.1, a concentration result which makes no reference to bridged Brownian sheets. In Section 7 we briefly discuss a formulation of our result in terms of lozenge tilings of a regular hexagon. Finally, in Section 8, we present a conjecture analogous to Theorem 1.1 for Birkhoff copulas, which are real valued matrices with a definition similar to that of discrete copulas.

## 2 Pointwise convergence

For a positive integer $n$, let $X_n$ be a random discrete copula with uniform distribution, that is, for each copula $C \in \mathcal{S}_n$ (or, equivalently, for each permutation matrix $M \in S_n$) we have $\Pr[X_n = C] = 1/n!$. Given $a, b, c \in \mathbb{N}$, write

$$P(n, a, b, c) = \Pr[X_{n,a,b} = c],$$

the probability that the upper left $a \times b$ submatrix of $M \in S_n$ contains precisely $c$ entries equal to 1. We introduce some heavy handed notation preparing the way to grids, to be considered in later sections.

$$
\left(
\begin{array}{ccccc|ccc}
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\end{array}
\right)
\qquad
\left(
\begin{array}{ccccc|ccc}
0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 \\
1 & 1 & 1 & 2 & 2 & 3 & 3 & 3 \\
1 & 1 & 2 & 3 & 3 & 4 & 4 & 4 \\
\hline
1 & 1 & 2 & 3 & 4 & 5 & 5 & 5 \\
1 & 2 & 3 & 4 & 5 & 6 & 6 & 6 \\
1 & 2 & 3 & 4 & 5 & 6 & 6 & 7 \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\end{array}
\right)
$$

Figure 2: A permutation matrix and its discrete copula: $C_{4,5} = 3$.

A quadruple $\mathbf{b} = (n, a, b, c) \in \mathbb{N}^4$ is a *blocking* if it satisfies the following condition:

$$n > 0, \quad 0 \leq a, b \leq n, \quad \max(0, n - a - b) \leq c \leq \min(a, b). \tag{2}$$

A blocking $\mathbf{b} = (n, a, b, c)$ describes a partitioning of the $n \times n$ square into four rectangles of sizes $\dot a_i \times \dot b_j$, $1 \leq i, j \leq 2$, where

$$\dot a_1 = a_1 = a, \quad \dot a_2 = n - a, \quad \dot b_1 = b_1 = b, \quad \dot b_2 = n - b,$$

together with the number of nonzero entries $\dot c_{i,j}$, where

$$\dot c_{1,1} = c_{1,1} = c, \quad \dot c_{1,2} = a - c, \quad \dot c_{2,1} = b - c, \quad \dot c_{2,2} = n - a - b + c,$$

to fit the $(i, j)$-rectangle for any matrix $M \in S_n$ with $C_{i,j} = c$. In Figure 2,

$$\dot c_{1,1} = 3, \quad \dot c_{1,2} = 1, \quad \dot c_{2,1} = 2, \quad \dot c_{2,2} = 2$$

are indeed the number of nonzero entries in each box. Notice that Condition (2) is equivalent to $\dot c_{i,j} \geq 0$ (for all $i, j$).

An expression for this probability is a simple combinatorial problem.

**Lemma 2.1.** *Let* $\mathbf{b} = (n, a, b, c)$ *be a blocking; then*

$$E[X_{n,a,b}] = (C_0)_{a,b} = \frac{ab}{n};$$

$$P(\mathbf{b}) = P(n, a, b, c) = \frac{a!b!(n-b)!(n-a)!}{n!c!(a-c)!(b-c)!(n-b-a+c)!} = \frac{\dot a_1! \dot a_2! \dot b_1! \dot b_2!}{n! \dot c_{1,1}! \dot c_{1,2}! \dot c_{2,1}! \dot c_{2,2}!}.$$

*Proof.* For the expected value, each one of the first $a$ rows contributes with a 1 with probability $b/n$, obtaining the desired result.

There are $a$ columns in the matrix $M \in S_n$ with a 1 in the first $a$ rows. The total number of ways to choose those columns is $\binom{n}{a}$. In order to have $X_{n,a,b} = c$, $c$ columns must be chosen from the first $b$ columns, and $a - c$ columns from the remaining $n - b$ columns. Thus

$$P(\mathbf{b}) = \frac{\binom{b}{c}\binom{n-b}{a-c}}{\binom{n}{a}},$$

and the result follows by expanding the binomial coefficients. $\qquad\square$

Let $\mathrm{xl}\, t = t \log t$; we use Stirling's approximation formula with tail $\mathrm{st}(k)$:

$$\log(k!) = \mathrm{xl}\, k - k + \frac{1}{2}\log k + \frac{1}{2}\log(2\pi) + \mathrm{st}\, k, \quad 0 < \mathrm{st}\, k < \frac{1}{12k}. \qquad (3)$$

We define the rescalings on $\mathbb{I}$

$$\hat a_i = \frac{\dot a_i}{n}, \quad \hat b_j = \frac{\dot b_j}{n}, \quad \hat c_{i,j} = \frac{\dot c_{i,j}}{n},$$

the auxiliary real numbers

$$\mathcal{C} = \hat a_1 \hat a_2 \hat b_1 \hat b_2 \in [0, 1/16], \quad \mathcal{D} = \hat c_{1,1} \hat c_{1,2} \hat c_{2,1} \hat c_{2,2},$$

and the factorial-free approximation

$$\tilde{P}(\mathbf{b}) = \frac{\sqrt{\mathcal{C}}}{\sqrt{2\pi\mathcal{D}n}} \frac{\hat{a}_1^{\hat{a}_1 n} \hat{a}_2^{\hat{a}_2 n} \hat{b}_1^{\hat{b}_1 n} \hat{b}_2^{\hat{b}_2 n}}{\hat{c}_{1,1}^{\hat{c}_{1,1} n} \hat{c}_{1,2}^{\hat{c}_{1,2} n} \hat{c}_{2,1}^{\hat{c}_{2,1} n} \hat{c}_{2,2}^{\hat{c}_{2,2} n}}$$

$$= \frac{\sqrt{\mathcal{C}}}{\sqrt{2\pi\mathcal{D}n}} \exp\left( n \left( \sum_{1 \le i \le 2} \mathrm{xl}\,\hat{a}_i + \sum_{1 \le j \le 2} \mathrm{xl}\,\hat{b}_j - \sum_{1 \le i,j \le 2} \mathrm{xl}\,\hat{c}_{i,j} \right) \right).$$

The *sparsity* of a blocking $\mathbf{b} = (n, a, b, c)$ is $\lambda(\mathbf{b}) = \min_{i,j} \dot{c}_{i,j}$.

**Lemma 2.2.**

$$\lim_{\lambda(\mathbf{b}) \to \infty} \frac{P(\mathbf{b})}{\tilde{P}(\mathbf{b})} = 1.$$

That is, given $\epsilon > 0$ there exists $\lambda_\epsilon$ such that, if $\lambda(\mathbf{b}) > \lambda_\epsilon$ then

$$\left| \frac{P(\mathbf{b})}{\tilde{P}(\mathbf{b})} - 1 \right| < \epsilon.$$

*Proof.* Use Stirling's formula on Lemma 2.1. Notice that $\lambda(\mathbf{b})$ is the smallest integer whose factorial is taken in the formula of Lemma 2.1. □

We will use another formula for $\tilde{P}$, more amenable to Taylor approximations. From Lemma 2.1, the expected value of $(X_{n,a,b})/n$, the random variable corresponding to $\hat{c}_{1,1}$, equals $\hat{a}\hat{b}$. Inspired by this, we introduce new variables $\hat{l} = \sqrt{n}\,(\hat{c} - \hat{a}\hat{b}) = (c - \frac{ab}{n})/\sqrt{n}$ and, similarly, $\hat{l}_{i,j}$ such that $\hat{c}_{i,j} = \hat{a}_i \hat{b}_j + \hat{l}_{i,j}/\sqrt{n}$. Notice that

$$\sum_i \hat{a}_i = \sum_j \hat{b}_j = 1, \quad \sum_j \hat{c}_{i,j} = \hat{a}_i, \quad \sum_i \hat{c}_{i,j} = \hat{b}_j \tag{4}$$

and therefore

$$\hat{l}_{i,j} = (-1)^{i+j}\hat{l}, \quad \sum_j \hat{l}_{i,j} = \sum_i \hat{l}_{i,j} = 0. \tag{5}$$

We also introduce the corresponding random variables:

$$\hat{h} = \frac{1}{\sqrt{n}} \left( X_{a,b} - \frac{ab}{n} \right);$$

the random variables $\hat{h}_{i,j}$ are defined similarly; thus,

$$P(\mathbf{b}) = \Pr[X_{a,b} = c] = \Pr[\hat{h} = \hat{l}].$$

**Lemma 2.3.**

$$\log(\tilde{P}(\mathbf{b})) = -\frac{1}{2} \log\left( \frac{2\pi\mathcal{D}n}{\mathcal{C}} \right) - n \sum_{i,j} \hat{a}_i \hat{b}_j \, \mathrm{xl}\left( 1 + \frac{\hat{l}_{i,j}}{\sqrt{n}\,\hat{a}_i \hat{b}_j} \right).$$

*Proof.* Write

$$\log(\tilde{P}(\mathbf{b})) = -\frac{1}{2}\log\left(\frac{2\pi\mathcal{D}n}{\mathcal{C}}\right) - nA$$

where

$$A = \sum_{i,j}\mathrm{xl}\,\hat{c}_{i,j} - \sum_{i}\mathrm{xl}\,\hat{a}_i - \sum_{j}\mathrm{xl}\,\hat{b}_j.$$

We have

$$\mathrm{xl}\,\hat{c}_{i,j} = \left(\hat{a}_i\hat{b}_j + \frac{\hat{l}_{i,j}}{\sqrt{n}}\right)\left(\log\hat{a}_i + \log\hat{b}_j + \log\left(1 + \frac{\hat{l}_{i,j}}{\sqrt{n}\,\hat{a}_i\hat{b}_j}\right)\right)$$

$$= \hat{b}_j\,\mathrm{xl}\,\hat{a}_i + \hat{a}_i\,\mathrm{xl}\,\hat{b}_j + \frac{\hat{l}_{i,j}}{\sqrt{n}}\log\hat{a}_i + \frac{\hat{l}_{i,j}}{\sqrt{n}}\log\hat{b}_j$$

$$+ \hat{a}_i\hat{b}_j\,\mathrm{xl}\left(1 + \frac{\hat{l}_{i,j}}{\sqrt{n}\,\hat{a}_i\hat{b}_j}\right).$$

From equations 4 and 5 we have

$$A = \sum_{i,j}\hat{a}_i\hat{b}_j\,\mathrm{xl}\left(1 + \frac{\hat{l}_{i,j}}{\sqrt{n}\,\hat{a}_i\hat{b}_j}\right).$$

$\square$

We expect that appropriate scalings asymptotically yield normal distributions. Consider the residuals $Y_n = X_n - C_0$ and their extension $\mathbf{y} \in C_0^0(\mathbb{I}^2)$. Notice that the values at two points in the the same small square satisfy

$$|\mathbf{y}(u_0, v_0) - \mathbf{y}(u_1, v_1)| < \frac{4}{n}.$$

We shall prove in Corollary 2.6 that, for fixed $(u, v) \in (0, 1)^2$ and $n \to \infty$, $\sqrt{n}\,\mathbf{y}(u, v)$ has normal distribution centered at 0 with variance $\mathcal{C} = u(1 - u)v(1 - v)$. Consistently with Theorem 1.1, this is precisely the distribution of $f(u, v)$ if $f$ is a bridged Brownian sheet.

As a warm-up, we compute an approximation for $P(\mathbf{b}) = \Pr[X_{n,a,b} = c]$ near the mean.

**Lemma 2.4.** *Let $X_n$ be a random discrete copula with uniform distribution in $\mathcal{S}_n$. Given constants $0 < u, v < 1$, let $\mathbf{b}_n = (n, a, b, c)$ for $a = \lfloor un \rfloor$, $b = \lfloor vn \rfloor$ and $c = \lfloor uvn \rfloor$, with the other variables defined as above. Then*

$$\lim_{n \to \infty}\sqrt{2\pi\mathcal{C}n}\,P(\mathbf{b}_n) = 1.$$

*Proof.* Clearly, the sparsity $\lambda(\mathbf{b}_n)$ goes to $\infty$ when $n \to \infty$; in particular, $\mathbf{b}_n$ satisfies Condition (2). In the notation of Lemma 2.3, we claim that

$$\lim_{n \to \infty}\frac{\mathcal{D}}{\mathcal{C}^2} = 1, \quad \lim_{n \to \infty}nA = 0,$$

7

which, given Lemma 2.2, obtains the desired result (we occasionally omit the dependency in $n$). We have $\hat{a} \to u$ and $\hat{b} \to v$ so that $\mathcal{C} \to u(1-u)v(1-v)$ and $\mathcal{D} \to (u(1-u)v(1-v))^2$, from which the first limit follows. Standard manipulation of integer parts yields

$$-\frac{2}{\sqrt{n}} < \hat{l}_{i,j} < \frac{2}{\sqrt{n}}.$$

From the Taylor approximation

$$\mathrm{xl}(1+t) = (1+t)\log(1+t) = t + \frac{t^2}{2} + O(t^3), \tag{6}$$

we may write

$$nA = n \sum_{i,j} \hat{a}_i \hat{b}_j \, \mathrm{xl}\left(1 + \frac{\hat{l}_{i,j}}{\sqrt{n}\,\hat{a}_i\hat{b}_j}\right) = \sqrt{n} \sum_{i,j} \hat{l}_{i,j} + \frac{1}{2} \sum_{i,j} \frac{\hat{l}_{i,j}^2}{\hat{a}_i\hat{b}_j} + O\left(\frac{1}{n^2}\right).$$

Equations 4 and 5 imply that the first summation in the expansion equals zero. The second summation tends to zero. $\qquad\square$

We denote the Gaussian centered at 0 with variance $C$ by

$$\mathrm{ga}_C\, t = \frac{1}{\sqrt{C}}\, \mathrm{ga}\left(\frac{t}{\sqrt{C}}\right) = \frac{1}{\sqrt{C}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2C}\right),$$

with $\mathrm{ga}\, t = \mathrm{ga}_1\, t$; in particular, $\int_{\mathbb{R}} \mathrm{ga}_C = 1$.

The following lemma gives an estimate for $P(\mathbf{b})$ for large $n$, when $a$ and $b$ are not too close to the boundary values 0 and $n$ (controlled by a parameter $\alpha$) and when $c$ is relatively near its average value $\frac{ab}{n}$ (controlled by $\eta$). The lemma is rather technical and requires additional notation.

A *grid* is a triple $\mathbf{g} = (n, a, b)$; we identify a blocking with a pair $\mathbf{b} = (\mathbf{g}, c)$. Given a grid $\mathbf{g}$, we relate a blocking $\mathbf{b} = (\mathbf{g}, c)$ with a *normalized blocking* $\hat{\mathbf{b}} = (\mathbf{g}, \hat{l})$. Notice that $\hat{l}$ and $\hat{h}$ must belong to the lattice

$$L_{\mathbf{g}} = \frac{1}{\sqrt{n}}\left(\mathbb{Z} - \frac{ab}{n}\right) \subset \mathbb{R},$$

with spacing $1/\sqrt{n}$. We now introduce the first instance of parameters which shall be restricted to smaller regions along the text. Let $\alpha \in (\frac{7}{8}, 1)$; a grid $\mathbf{g}$ is *$\alpha$-regular* if

$$n^\alpha < a, b < n - n^\alpha.$$

Consider the open triangle

$$\Delta_1 = \{(\alpha, \eta) \in \mathbb{R}^2 \mid 0 < 6\eta < 8\alpha - 7 < 1\}; \tag{7}$$

for $(\alpha, \eta) \in \Delta_1$, a normalized blocking $\hat{\mathbf{b}} = (\mathbf{g}, \hat{l})$ is *$(\alpha, \eta)$-standard*, or $\hat{\mathbf{b}} \in \Xi_{\alpha,\eta}$, if $\mathbf{g}$ is $\alpha$-regular and

$$\hat{l} \in L_{\mathbf{g}}, \quad |\hat{l}| < n^\eta.$$

8

**Lemma 2.5.** *For fixed* $(\alpha, \eta) \in \Delta_1$,

$$\lim_{n \to \infty; \, \hat{\mathbf{b}} \in \Xi_{\alpha,\eta}} \frac{\sqrt{n} \, P(\hat{\mathbf{b}})}{\mathrm{ga}_{\mathcal{C}} \, \hat{l}} = 1.$$

We must clarify the meaning of the above limit. In general, let $\Xi$ be a set with a function $n : \Xi \to \mathbb{N}$. Let $f : \Xi \to \mathbb{R}$ be a function and $A \in \mathbb{R}$. Then

$$\lim_{n \to \infty; \, \hat{\mathbf{b}} \in \Xi} f(\hat{\mathbf{b}}) = A$$

means that given $\epsilon > 0$ there exists $n_0$ such that, if $\hat{\mathbf{b}} \in \Xi$ and $n(\hat{\mathbf{b}}) > n_0$ then

$$\left| f(\hat{\mathbf{b}}) - A \right| < \epsilon;$$

in the above lemma, the function $n$ is the obvious one. The set $\Xi_{\alpha,\eta}$ is implicit (and hopefully clear from context) whenever we use notations $\sim$ and little oh. Thus, for instance, the lemma may be stated as

$$P(\hat{\mathbf{b}}) \sim \frac{1}{\sqrt{n}} \, \mathrm{ga}_{\mathcal{C}} \, \hat{l} \quad \text{or} \quad P(\hat{\mathbf{b}}) = (1 + o(1)) \frac{1}{\sqrt{n}} \, \mathrm{ga}_{\mathcal{C}} \, \hat{l}.$$

*Proof.* The hypotheses imply $\eta < 2\alpha - \frac{3}{2}$ and therefore

$$|\hat{l}_{i,j}| < n^\eta < n^{(2\alpha - \frac{3}{2})} < \sqrt{n} \, \hat{a}_i \hat{b}_j$$

so that $\hat{c}_{i,j} = \hat{a}_i \hat{b}_j (1 + o(1))$. In particular, $\lambda(\hat{\mathbf{b}}) \to \infty$ when $n \to \infty$ and therefore, by Lemma 2.2, $P(\hat{\mathbf{b}}) \sim \tilde{P}(\hat{\mathbf{b}})$. We also have $\mathcal{D} \sim \mathcal{C}^2$. In the notation of Lemma 2.3, write

$$\log(P(\hat{\mathbf{b}})) = -\frac{1}{2} \log(2\pi \mathcal{C} n) - nA + o(1),$$

$$A = \sum_{i,j} \hat{a}_i \hat{b}_j \, \mathrm{xl} \left( 1 + \frac{\hat{l}_{i,j}}{\sqrt{n} \hat{a}_i \hat{b}_j} \right), \quad \left| \frac{\hat{l}_{i,j}}{\sqrt{n} \hat{a}_i \hat{b}_j} \right| < n^{\frac{3}{2} + \eta - 2\alpha} = o(1)$$

and we may use the Taylor approximation in Equation 6 (with a universal constant in the $O$) to obtain

$$nA = \sum_{i,j} \sqrt{n} \, \hat{l}_{i,j} + \sum_{i,j} \frac{\hat{l}_{i,j}^2}{\hat{a}_i \hat{b}_j} + O \left( \sum_{i,j} n^{-\frac{1}{2}} \frac{|\hat{l}_{i,j}|^3}{\hat{a}_i^2 \hat{b}_j^2} \right).$$

The first summand equals zero from equation 5. The second equals $\hat{l}^2 / \mathcal{C}$. The third summand goes to zero since each of the four terms is bounded by

$$n^{-\frac{1}{2}} \cdot n^{3\eta} \cdot n^{4(1-\alpha)} = n^{\frac{7}{2} + 3\eta - 4\alpha} = o(1).$$

Thus

$$\log(P(\hat{\mathbf{b}})) = -\frac{1}{2} \log(2\pi \mathcal{C} n) - \frac{\hat{l}^2}{2\mathcal{C}} + o(1),$$

completing the proof. $\qquad\square$

**Corollary 2.6.** *Let* $(u,v) \in (0,1)^2$ *be fixed. For real numbers* $t_- < t_+$, *let*

$$p_n = \Pr\left[t_- < \sqrt{n}\, \mathbf{y}_n(u,v) < t_+\right].$$

*Then*

$$\lim_{n\to\infty} p_n = \int_{t_-}^{t_+} \mathrm{ga}_{\mathcal{C}}(t)dt.$$

*Proof.* Take fixed $(\alpha, \eta) \in \Delta_1$. Given $n$, set $a = \lfloor un \rfloor$, $b = \lfloor vn \rfloor$. Notice that $|y_n(u,v) - y_n(\frac{a}{n}, \frac{b}{n})| < 2$. Let $\mathbf{g}_n = (n, a, b)$. For sufficiently large $n$, the conditions $\hat{l} \in L_{\mathbf{g}_n}$, $t_- < \hat{l} < t_+$ imply that the normalized blocking $\hat{\mathbf{b}}_n = (\mathbf{g}_n, \hat{l})$ is $(\alpha, \eta)$-standard. From Lemma 2.5, since the spacing of $L_{\mathbf{g}_n}$ is $1/\sqrt{n}$, we have

$$\lim_{n\to\infty} p_n = \lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{\substack{t_- < \hat{l} < t_+ \\ \hat{l} \in L_{\mathbf{g}_n}}} \mathrm{ga}_{\mathcal{C}}(\hat{l}) = \int_{t_-}^{t_+} \mathrm{ga}_{\mathcal{C}}(t)dt.$$

$\square$

The following corollary gives an example of an application of Lemma 2.5 for a sequence of blockings $\mathbf{b}_n$. The $\sim$ notation below is therefore used in the more traditional sense of a sequence limit.

**Corollary 2.7.** *For fixed* $(u,v) \in (0,1)^2$, *let* $a = \lfloor nu \rfloor$, $b = \lfloor nv \rfloor$, $\mathcal{C} = u(1-u)v(1-v)$. *Let* $0 \le r < 1/6$ *and* $g(n)$ *be a function such that* $g(n) = o(n^t)$ *for any* $t > 0$. *Then we have*

$$\Pr\left[X_{n,a,b} = \left\lfloor nuv + g(n)n^{r+1/2} \right\rfloor\right] \sim \frac{1}{\sqrt{n}}\, \mathrm{ga}_{\mathcal{C}}\left(g(n)n^r\right).$$

*Proof.* Take $(\alpha, \eta) \in \Delta_1$ with $r < \eta$. For sufficiently large $n$, the blocking

$$\mathbf{b}_n = \left(n, a, b, \left\lfloor nuv + g(n)n^{r+1/2} \right\rfloor\right)$$

is $(\alpha, \eta)$-standard. The result now follows directly from Lemma 2.5. $\square$

Lemma 2.5 above gives a good estimate for $P(\hat{\mathbf{b}})$ when $c$ is relatively near $\frac{ab}{n}$, or, equivalently, when $\hat{l}$ is small. In other words, the probability of $|Y_{n,a,b}|$ (or $|\hat{l}|$) being large is small, but we shall need something more precise. The following lemma addresses these concerns. Notice also that Lemma 2.8 below is not a corollary of Lemma 2.5.

**Lemma 2.8.** *Let* $(\alpha, \eta) \in \Delta_1$ *be fixed. Then there exists* $n_0$ *such that, if* $n > n_0$, *for any* $\alpha$-regular grid $\mathbf{g} = (n, a, b)$ *and* $H \in (1, 2n^\eta)$, *then*

$$\Pr\left[|Y_{n,a,b}| > H\sqrt{\mathcal{C}n}\right] < \frac{1}{H}\exp\left(-\frac{H^2}{2}\right).$$

10

*Proof.* We prove that

$$\Pr\left[Y_{n,a,b} > H\sqrt{\mathcal{C}n}\right] < \frac{1}{2H}\,\exp\left(-\frac{H^2}{2}\right).$$

Let

$$p_l = P(\mathbf{g},c); \quad c_0 = \left\lfloor \frac{ab}{n} + H\sqrt{\mathcal{C}n}\right\rfloor;$$

so that

$$\tilde{p} = \Pr\left[X_{n,a,b} > \frac{ab}{n} + H\sqrt{\mathcal{C}n}\right] = p_{c_0+1} + p_{c_0+2} + \cdots + p_{c_0+k} + \cdots.$$

The blocking $(\mathbf{g}, c_0 + 1)$ is $(\alpha, \eta)$-standard and therefore, from Lemma 2.5,

$$p_{c_0} < (1+o(1))\,\frac{1}{\sqrt{2\pi\mathcal{C}n}}\,\exp\left(-\frac{H^2}{2}\right).$$

Let $\lambda_l = p_{l+1}/p_l$. From Lemma 2.1,

$$\lambda_l = \frac{(a-c)(b-c)}{(l+1)(n-a-b+c+1)},$$

a decreasing function of $c$. Thus, $p_{c_0+k} \leq \lambda_{c_0}^k p_{c_0}$ and therefore

$$\tilde{p} < (1+o(1))\,\frac{1}{1-\lambda_{c_0}}\,\frac{1}{\sqrt{2\pi n\mathcal{C}}}\,\exp\left(-\frac{H^2}{2}\right).$$

We now proceed to estimate $\lambda_{c_0}$. Set $\hat{c}_0 = c_0/n \approx \hat{a}\hat{b} + H\sqrt{\mathcal{C}}n^{-1/2}$.

$$
\begin{aligned}
\lambda_{c_0} &= \frac{(a-c_0)(b-c_0)}{(c_0+1)(n-a-b+c_0+1)}\\
&\approx \frac{(\hat{a}-\hat{c}_0)(\hat{b}-\hat{c}_0)}{\hat{c}_0(1-\hat{a}-\hat{b}+\hat{c}_0)}\\
&\approx \frac{(\hat{a}-\hat{a}\hat{b}-H\sqrt{\mathcal{C}}n^{-1/2})(\hat{b}-\hat{a}\hat{b}-H\sqrt{\mathcal{C}}n^{-1/2})}{(\hat{a}\hat{b}+H\sqrt{\mathcal{C}}n^{-1/2})(1-\hat{a}-\hat{b}+\hat{a}\hat{b}+H\sqrt{\mathcal{C}}n^{-1/2})}.
\end{aligned}
$$

The approximations here are due to taking integer parts and can safely be disconsidered. Expanding, we have

$$
\begin{aligned}
\lambda_{c_0} &= \frac{\mathcal{C}+(2\hat{a}\hat{b}-\hat{a}-\hat{b})H\sqrt{\mathcal{C}}n^{-1/2}+H^2\mathcal{C}n^{-1}}{\mathcal{C}+(1+2\hat{a}\hat{b}-\hat{a}-\hat{b})H\sqrt{\mathcal{C}}n^{-1/2}+H^2\mathcal{C}n^{-1}}\\
&= 1-\left(\frac{H}{\sqrt{\mathcal{C}n}}(1+O(n^{-\epsilon}))\right)
\end{aligned}
$$

where $\epsilon > 0$. We therefore have

$$\frac{1}{1-\lambda_{c_0}} = (1+O(n^{-\epsilon}))\,\frac{\sqrt{\mathcal{C}n}}{H},$$

11

which yields the desired estimate. A similar argument proves that

$$\Pr\left[Y_{n,a,b} < -H\sqrt{\mathcal{C}n}\right] < \frac{1}{2H}\,\exp\left(-\frac{H^2}{2}\right),$$

completing the proof. □

# 3  Grids and joint distributions

In this section we generalize some of the results of Section 2 to *grids*.

Recall that a *bridged Brownian sheet* is a process yielding continuous functions $f \in C_0^0(\mathbb{I}^2)$. We provide a characterization which will be appropriate to our proof. Consider a $I \times J$ *grid* $\mathbf{g}$ of points $(u_i, v_j) \in \mathbb{I}^2$ where

$$0 = u_0 < u_1 < \cdots < u_I = 1, \quad 0 = v_0 < v_1 < \cdots < v_J = 1.$$

In order to generate the values $f(u_i, v_j)$ of a bridged Brownian sheet on the grid, take independent and normally distributed numbers $Z_{i,j}$, $1 \le i \le I$, $1 \le j \le J$, with average 0 and variance 1, to obtain a vector $z \in \mathbb{R}^{IJ}$. Let $V_{\mathbf{g}} \subset \mathbb{R}^{IJ}$ be the subspace of dimension $(I-1)(J-1)$ defined by

$$\sum_i \sqrt{u_i - u_{i-1}}\, z_{i,j} = 0, \quad \sum_j \sqrt{v_j - v_{j-1}}\, z_{i,j} = 0.$$

Project $z \in \mathbb{R}^{IJ}$ orthogonally to $\tilde{z} \in V_{\mathbf{g}}$. Set

$$f(u_i, v_j) = \sum_{i' \le i;\, j' \le j} \sqrt{(u_{i'} - u_{i'-1})(v_{j'} - v_{j'-1})}\, \tilde{z}_{i',j'}. \tag{8}$$

Thus, the values of $f$ on a fixed set of points follow a joint normal distribution. In particular, for fixed $(u,v) \in \mathbb{I}^2$, the random variable $f(u,v)$ has normal distribution with average 0 and variance $u(1-u)v(1-v)$. In Section 5 we focus on square grids $I = J = m$, $u_i = \frac{i}{m}$, $v_j = \frac{j}{m}$.

This section is dedicated to the following result.

**Theorem 3.1.** *Given $(u_1, v_1), \ldots (u_m, v_m) \in (0,1)^2$, the joint distribution of $\sqrt{n}\,\mathbf{y}_n(u_i, v_i)$ converges (when $n \to +\infty$) to the joint distribution of $f(u_i, v_i)$ for a bridged Brownian sheet $f$.*

Notice that the case $m = 1$ of this theorem follows from Section 2.

As before, $n$ is a positive integer and $X$ is a random discrete copula with uniform distribution. A $I \times J$ *grid* is a pair $\mathbf{g} = (a, b)$ of families, $(a_i)_{0 \le i \le I}$ and $(b_j)_{0 \le j \le J}$ of integers with

$$0 = a_0 < \cdots < a_i < \cdots < a_I = n, \quad 0 = b_0 < \cdots < b_j < \cdots < b_J = n.$$

As in Figure 3, the values of $a_i$ and $b_j$ should be interpreted as horizontal and vertical lines dividing a matrix $M$ into *boxes* of dimensions $\dot{a}_i \times \dot{b}_j$ defined by

$$\dot{a}_i = a_i - a_{i-1}, \quad a_{i_0} = \sum_{i \le i_0} \dot{a}_i, \quad \dot{b}_j = b_j - b_{j-1}, \quad b_{j_0} = \sum_{j \le j_0} \dot{b}_j.$$

Given a discrete copula $C$ and a grid $\mathbf{g} = (a, b)$, the number of 1's (of $M$) in box $(i, j)$ is $C_{a_i, b_j} - C_{a_{i-1}, b_j} - C_{a_i, b_{j-1}} + C_{a_{i-1}, b_{j-1}}$; we are interested in the statistics of such numbers.
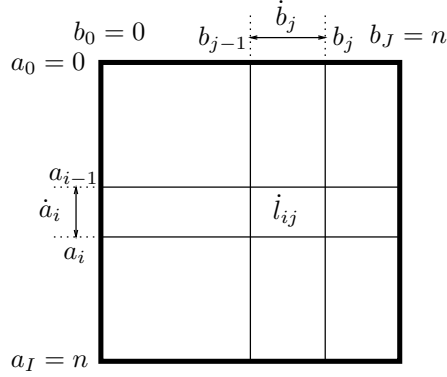


Figure 3: The box $(i, j)$

Let $c = (c_{i,j})$ be a family of integers with

$$i_0 \le i_1, j_0 \le j_1 \Rightarrow c_{i_0, j_0} + c_{i_1, j_1} \ge c_{i_0, j_1} + c_{i_1, j_0},$$

$$c_{i,0} = c_{0,j} = 0, \quad c_{i,J} = a_i, \quad c_{I,j} = b_j.$$

Define
$$\dot{c}_{i,j} = c_{i,j} - c_{i-1,j} - c_{i,j-1} + c_{i-1,j-1}, \quad c_{i_0, j_0} = \sum_{\substack{i \le i_0 \\ j \le j_0}} \dot{c}_{i,j}.$$

For a blocking $\mathbf{b} = (\mathbf{g}, c)$, we are interested in estimating

$$P(\mathbf{b}) = \Pr\left[\forall i, \forall j, X_{a_i, b_j} = c_{i,j}\right];$$

this is the probability that each box $(i, j)$ contains precisely $\dot{c}_{i,j}$ 1's. The situation described in Section 2 corresponds to $I = J = 2$ (the old $a$ is the new $a_1$; the old $c$ is the new $c_{1,1}$). A careful estimate of $P(\mathbf{b})$ will prove Theorem 3.1 for the special case of points in a grid. This special case implies the general one: indeed, it suffices to add points in order to complete a grid.

First a simple remark which shall be used several times: there is nothing special about the upper left corner, and we can count nonzero entries is any rectangular submatrix.

**Lemma 3.2.** *Let $0 < a_1 < a_2 < n$, $0 < b_1 < b_2 < n$ and $c$ be integers. Then*

$$\Pr\left[X_{a_2,b_2} - X_{a_2,b_1} - X_{a_1,b_2} + X_{a_1,b_1} = c\right] = P(n, a_2 - a_1, b_2 - b_1, c).$$

*Proof.* Cyclically permute rows and columns of the corresponding matrix $M$. □

The following lemma generalizes Lemma 2.1.

**Lemma 3.3.**

$$P(\mathbf{b}) = \frac{\left(\prod_{0<i\leq I} \dot{a}_i!\right)\left(\prod_{0<j\leq J} \dot{b}_j!\right)}{n!\left(\prod_{\substack{0<i\leq I \\ 0<j\leq J}} \dot{c}_{i,j}!\right)}.$$

*Proof.* We consider the corresponding permutation matrix by columns: these are the vectors $e_1, \ldots, e_n$, and we have to choose their order. There are $\dot{a}_i$ vectors of *type $i$*: $e_{a_{i-1}+1}, \ldots, e_{a_i}$. The *wide column $j$* is the union of the $\dot{b}_j$ columns $b_{j-1} + 1$ to $b_j$; there are therefore $J$ wide columns. We have to place $\dot{c}_{i,j}$ vectors of type $i$ in the wide column $j$. Thus, for each $i$, we have to choose $\dot{c}_{i,j}$ vectors to go inside each wide column: there are $\left(\prod_i \dot{a}_i!\right)/\left(\prod_{i,j} \dot{c}_{i,j}!\right)$ ways of choosing. Inside each wide column, vectors may be permuted, obtaining the desired count of suitable permutation matrices. □

Rescale to introduce

$$\hat{a}_i = \dot{a}_i/n \in \mathbb{I}, \quad \hat{b}_j = \dot{b}_j/n \in \mathbb{I}, \quad \hat{c}_{i,j} = \dot{c}_{i,j}/n \in \mathbb{I},$$

$$\mathcal{C}_a = \prod_i \hat{a}_i, \quad \mathcal{C}_b = \prod_j \hat{b}_j, \quad \mathcal{D} = \prod_{i,j} \hat{c}_{i,j}$$

and the factorial-free approximation

$$\tilde{P}(\mathbf{g}, c) = \sqrt{\frac{\mathcal{C}_a \mathcal{C}_b}{(2\pi n)^{(I-1)(J-1)} \mathcal{D}}} \frac{\left(\prod_i \hat{a}_i^{(\hat{a}_i n)}\right)\left(\prod_j \hat{b}_j^{(\hat{b}_j n)}\right)}{\prod_{i,j} \hat{c}_{i,j}^{(\hat{c}_{i,j} n)}},$$

or, equivalently, with $\mathrm{xl}(t) = t \log t$,

$$\log(\tilde{P}(\mathbf{g}, c)) = -\frac{(I-1)(J-1)}{2} \log(2\pi n) - \frac{1}{2} \log\left(\frac{\mathcal{C}_a \mathcal{C}_b}{\mathcal{D}}\right)$$

$$+ n\left(\sum_i \mathrm{xl}(\hat{a}_i) + \sum_j \mathrm{xl}(\hat{b}_j) - \sum_{i,j} \mathrm{xl}(\hat{c}_{i,j})\right).$$

The following lemma generalizes Lemma 2.2. The *sparsity* of a blocking $\mathbf{b} = (\mathbf{g}, c)$ is $\lambda(\mathbf{b}) = \min_{i,j} \dot{c}_{i,j}$.

**Lemma 3.4.** *Let $I$ and $J$ be fixed. Then, when $\lambda(\mathbf{b}) \to \infty$, $P(\mathbf{b}) \sim \tilde{P}(\mathbf{b})$.*

14

*Proof.* Use Stirling's formula on Lemma 3.3:

$$\dot{a}_i! \sim (n/e)^{\hat{a}_i n} \hat{a}_i^{\hat{a}_i n} \sqrt{\hat{a}_i} \sqrt{2\pi n}, \quad \dot{b}_j! \sim (n/e)^{\hat{b}_j n} \hat{b}_j^{\hat{b}_j n} \sqrt{\hat{b}_j} \sqrt{2\pi n},$$

$$n! \sim (n/e)^n \sqrt{2\pi n}, \quad \dot{c}_{i,j}! \sim (n/e)^{\hat{c}_{i,j} n} \hat{c}_{i,j}^{\hat{c}_{i,j} n} \sqrt{\hat{c}_{i,j}} \sqrt{2\pi n}.$$

We have $\sum_i \hat{a}_i = \sum_j \hat{b}_j = \sum_{i,j} \hat{c}_{i,j} = 1$ and therefore

$$\frac{\left(\prod_i (n/e)^{\hat{a}_i n}\right)\left(\prod_j (n/e)^{\hat{b}_j n}\right)}{(n/e)^n \left(\prod_{i,j} (n/e)^{\hat{c}_{i,j} n}\right)} = 1.$$

The result follows. $\qquad\square$

At this point in Section 2 we introduce the variable $\hat{l}$ and random variable $\hat{h}$. Here we have reason to introduce slightly different variables $\tilde{l}_{i,j}$ (and corresponding variables $\tilde{h}_{i,j}$) defined by

$$\hat{c}_{i,j} = \hat{a}_i \hat{b}_j + \sqrt{\frac{\hat{a}_i \hat{b}_j}{n}}\, \tilde{l}_{i,j} = \hat{a}_i \hat{b}_j \left(1 + \frac{\tilde{l}_{i,j}}{\sqrt{n \hat{a}_i \hat{b}_j}}\right). \tag{9}$$

We relate a blocking $\mathbf{b} = (\mathbf{g}, c)$ with a *normalized blocking* $\tilde{\mathbf{b}} = (\mathbf{g}, \tilde{l})$. Notice that $\tilde{l}$ belongs to the vector subspace $V_{\mathbf{g}} \subset \mathbb{R}^{IJ}$ of dimension $(I-1)(J-1)$ defined by

$$\sum_i \sqrt{\hat{a}_i}\, \tilde{l}_{i,j} = 0, \quad \sum_j \sqrt{\hat{b}_j}\, \tilde{l}_{i,j} = 0. \tag{10}$$

Furthermore, $\tilde{l}$ belongs to the lattice $L_{\mathbf{g}} \subset V_{\mathbf{g}}$ defined by

$$\tilde{l}_{i,j} \in \frac{1}{\sqrt{n \hat{a}_i \hat{b}_j}} \left(\mathbb{Z} - \frac{\dot{a}_i \dot{b}_j}{n}\right). \tag{11}$$

Thus, the random variable $\tilde{h}$ assumes values in $L_{\mathbf{g}}$.

**Lemma 3.5.**

$$\log(\tilde{P}(\mathbf{g}, \tilde{l})) = \frac{1}{2} \log\left(\frac{\mathcal{C}_a \mathcal{C}_b}{\mathcal{D}}\right) - \frac{(I-1)(J-1)}{2} \log(2\pi n)$$

$$- n \sum_{i,j} \hat{a}_i \hat{b}_j \, \mathrm{xl}\left(1 + \frac{\tilde{l}_{i,j}}{\sqrt{n\,\hat{a}_i \hat{b}_j}}\right).$$

*Proof.* Up to notation, like the proof of Lemma 2.3. $\qquad\square$

We now extend Lemma 2.5 to grids. As in Section 2, let $\alpha \in (\frac{11}{12}, 1)$; a grid $\mathbf{g}$ is $\alpha$-*regular* if

$$\forall (i,j) \ \dot{a}_i, \dot{b}_j > n^\alpha.$$

Consider the open triangle

$$\Delta_2 = \{(\alpha, \eta) \in \mathbb{R}^2 \ | \ 0 < 12\eta < 12\alpha - 11 < 1\} \subset \Delta_1. \tag{12}$$

For $(\alpha, \eta) \in \Delta_2$, a normalized blocking $\tilde{\mathbf{b}} = (\mathbf{g}, \tilde{l})$ is $(\alpha, \eta)$-*standard* (denoted by $\tilde{\mathbf{b}} \in \Xi_{\alpha, \eta}$) if $\mathbf{g}$ is $\alpha$-regular, $\tilde{l} \in L_{\mathbf{g}}$, and

$$\forall (i,j) \ |\tilde{l}_{i,j}| < n^\eta.$$

**Lemma 3.6.** *Fix I, J and $(\alpha, \eta) \in \Delta_2$. If we restrict ourselves to $\Xi_{\alpha, \eta}$ then, as $\lambda(\mathbf{b})$ goes to infinity,*

$$P(\mathbf{g}, \tilde{l}) \sim \frac{1}{\sqrt{(2\pi n)^{(I-1)(J-1)} \mathcal{C}_a^{(J-1)} \mathcal{C}_b^{(I-1)}}} \ \exp\left(-\frac{|\tilde{l}|^2}{2}\right).$$

Here $|\tilde{l}|^2 = \sum_{i,j} \tilde{l}_{i,j}^2$. The notation $\sim$ here means the same as discussed in detail between the statement and proof of Lemma 2.5. Notice that when $\lambda(\mathbf{b})$ goes to infinity, so does $n$. We shall prove in Lemma 3.7 below that, for $\alpha$-regular grids, the condition $|\tilde{h}_{i,j}| < n^\eta$ holds with probability $1 - o(1)$ (when $n$ goes to infinity).

*Proof.* Notice that the conditions imply

$$\frac{\tilde{l}_{i,j}}{\sqrt{n \hat{a}_i \hat{b}_j}} < n^{\eta - \alpha + \frac{1}{2}} = O(n^{-\frac{5}{12}}); \quad \frac{\tilde{l}_{i,j}^3}{\sqrt{n \hat{a}_i \hat{b}_j}} < n^{3\eta - \alpha + \frac{1}{2}} = O(n^{-\frac{1}{4}}).$$

Let

$$\tilde{P}_0(\mathbf{g}, c) = \sqrt{\frac{\mathcal{C}_a \mathcal{C}_b}{(2\pi n)^{(I-1)(J-1)} \mathcal{D}}}, \quad \tilde{P}_1(\mathbf{g}, c) = \frac{\left(\prod_i \hat{a}_i^{(\hat{a}_i n)}\right) \left(\prod_j \hat{b}_j^{(\hat{b}_j n)}\right)}{\prod_{i,j} \hat{c}_{i,j}^{(\hat{c}_{i,j} n)}}.$$

Lemma 3.4 gives the uniform estimate $P(\mathbf{g}, c) \sim \tilde{P}_0(\mathbf{g}, c) \tilde{P}_1(\mathbf{g}, c)$.

We first estimate $\tilde{P}_0$. As before,

$$\hat{c}_{i,j} = \hat{a}_i \hat{b}_j \left(1 + \frac{\tilde{l}_{i,j}}{\sqrt{n \hat{a}_i \hat{b}_j}}\right) = \hat{a}_i \hat{b}_j \left(1 + O(n^{-\frac{5}{12}})\right).$$

We have that $\hat{c}_{i,j} \sim \hat{a}_i \hat{b}_j$ (uniformly) and therefore

$$\mathcal{D} = \prod_{i,j} \hat{c}_{i,j} \sim \prod_{i,j} (\hat{a}_i \hat{b}_j) = \left(\prod_i \hat{a}_i\right)^J \left(\prod_j \hat{b}_j\right)^I = \mathcal{C}_a^J \mathcal{C}_b^I.$$

16

Thus

$$\tilde{P}_0(\mathbf{g}, c) \sim \frac{1}{\sqrt{(2\pi n)^{(I-1)(J-1)} \mathcal{C}_a^{(J-1)} \mathcal{C}_b^{(I-1)}}}.$$

We now consider $\tilde{P}_1$. From Lemma 3.5, using the Taylor approximation for $\mathrm{xl}(1+t)$ (Equation 6),

$$\log\left(\tilde{P}_1(\mathbf{g}, c)\right) = -\sum_{i,j}(n\hat{a}_i\hat{b}_j)\,\mathrm{xl}\left(1 + \frac{\tilde{l}_{i,j}}{\sqrt{n\hat{a}_i\hat{b}_j}}\right)$$

$$= -\sqrt{n}\sum_{i,j}\sqrt{\hat{a}_i\hat{b}_j}\tilde{l}_{i,j} - \sum_{i,j}\frac{\tilde{l}_{i,j}^2}{2} + O(n^{-\frac{1}{4}})$$

or, since $\sum_{i,j}\sqrt{\hat{a}_i\hat{b}_j}\tilde{l}_{i,j} = \sum_i\sqrt{\hat{a}_i}\left(\sum_j\sqrt{\hat{b}_j}\tilde{l}_{i,j}\right) = 0$,

$$\tilde{P}_1(\mathbf{g}, c) \sim \exp\left(-\sum_{i,j}\frac{\tilde{l}_{i,j}^2}{2}\right),$$

completing the proof. $\qquad\square$

**Lemma 3.7.** *Fix $I$, $J$ and $(\alpha, \eta) \in \Delta_2$. Then there exists $n_0$ such that, if $n > n_0$ and if $\mathbf{g}$ is an $\alpha$-regular $I \times J$ grid, then*

$$\Pr\left[\forall(i,j)\,|\tilde{h}_{i,j}| < n^\eta\right] > 1 - \frac{IJ}{n^\eta}\exp\left(-\frac{n^{2\eta}}{2}\right) = 1 - o(1).$$

*Proof.* The idea is to use Lemma 2.8 with $H = n^\eta$. Fix $i$ and $j$. Notice that

$$\tilde{h}_{i,j} = \frac{Y_{n,a_i,b_j}}{\sqrt{n\,\hat{a}_i\,\hat{b}_j}}; \quad \mathcal{C} = \hat{a}_i(1-\hat{a}_i)\hat{b}_j(1-\hat{b}_j).$$

We have

$$\Pr\left[|\tilde{h}_{i,j}| > n^\eta\right] = \Pr\left[|Y_{n,a_i,b_j}| > H\sqrt{n\,\hat{a}_i\,\hat{b}_j}\right]$$

$$\leq \Pr\left[|Y_{n,a_i,b_j}| > H\sqrt{\mathcal{C}\,n}\right] < \frac{1}{H}\exp\left(-\frac{H^2}{2}\right).$$

Adding up for the $IJ$ pairs we complete the proof. $\qquad\square$

*Proof of Theorem 3.1.* Fix $(\alpha, \eta) \in \Delta_2$. By projection onto an appropriate subspace, if Theorem 3.1 holds for the finite family $(u_1, v_1), \ldots, (u_m, v_m)$ it also holds for any subfamily. Assume therefore without loss that the family forms a grid $\mathbf{g}_0$ of points $(u_i, v_j)$, with $i < I$, $j < J$. Write $u_0 = v_0 = 0$, $u_I = v_J = 1$.

17

For a positive integer $n$, set $a_i = \lfloor nu_i \rfloor$ and $b_j = \lfloor nv_j \rfloor$. For large $n$, this defines an $\alpha$-regular grid $\mathbf{g}$ with $\hat{a}_i \approx u_i - u_{i-1}$ and $\hat{b}_j \approx v_j - v_{j-1}$. Thus $V_{\mathbf{g}} \approx V_{\mathbf{g}_0}$. The bridged Brownian sheet defines a normal distribution on $V_{\mathbf{g}_0}$ as in Equation 8. Lemma 3.7 shows that, with probability $1 - o(1)$, a random permutation obtains an $(\alpha, \eta)$-standard normalized blocking $(\mathbf{g}, \tilde{h})$. From Lemma 3.6, such $\tilde{h}$ follow a normal distribution in the lattice $L_{\mathbf{g}} \subset V_{\mathbf{g}}$. Since these lattices get finer when $n$ goes to infinity it follows that the distributions of $\sqrt{n}\,\mathbf{y}_n$ converge to that of $f$, completing the proof. $\qquad\square$

# 4 Hölder estimates

Theorem 1.1 claims that $\mathbf{y}_n$ approaches a bridged Brownian sheet. As is well known, Brownian sheets are (almost surely) Hölder continuous for any exponent smaller than $1/2$ (bridged or not). The following lemma is a formal version of this fact and will be needed in the proof of Theorem 1.1. We use the following technical definition: given $r, \delta, C > 0$, a continuous function $g : \mathbb{I}^2 \to \mathbb{R}$ is $(r, \delta, C)$-Hölder if, for all $(u_0, v_0), (u_1, v_1) \in \mathbb{I}^2$,

$$|u_0 - u_1|, |v_0 - v_1| < r \;\Rightarrow\; |f(u_0, v_0) - f(u_1, v_1)| < C(|u_0 - u_1| + |v_0 - v_1|)^{\frac{1}{2} - \delta}.$$

**Lemma 4.1.** *Let $f : \mathbb{I}^2 \to \mathbb{R}$ be a bridged Brownian sheet. Fix $\delta, C > 0$. Let $p_r$ be the probability that $f$ be $(r, \delta, C)$-Hölder. Then*

$$\lim_{r \searrow 0} p_r = 1.$$

The proof is analogous to the well known case of Brownian motion and is omitted.

Lemma 4.2 below obtains a related estimate for $Y$, to be used in the proof of Theorem 1.1. A discrete copula $C$ is $(\alpha, \epsilon)$-*Hölder* if

$$\forall (a_1, a_2, b_1, b_2), \; \big(|a_2 - a_1|, |b_2 - b_1| < n^\alpha \;\Rightarrow\; |D_{a_2, b_2} - D_{a_1, b_1}| < n^{\frac{\alpha}{2} + \epsilon}\big)$$

(where, as usual, $D_{a,b} = C_{a,b} - \frac{ab}{n}$).

**Lemma 4.2.** *Let $\alpha \in (\frac{7}{8}, 1)$ and $\epsilon > 0$ be fixed. Then, a.a.s., $X$ is $(\alpha, \epsilon)$-Hölder.*

In other words,
$$\lim_{n \to \infty} \Pr[X \text{ is } (\alpha, \epsilon)\text{-Hölder}] = 1.$$

*Proof.* Notice first that it suffices to consider the case $b_1 = b_2$. That is, we prove that a.a.s., for all $a_1, a_2, b$, if $|a_1 - a_2| < n^\alpha$ then $|Y_{a_1, b} - Y_{a_2, b}| < n^{\frac{\alpha}{2} + \epsilon}$. Indeed, symmetry then implies a similar result in the case $a_1 = a_2$. But then, a.a.s.,

$$|Y_{a_2, b_2} - Y_{a_1, b_1}| \le |Y_{a_2, b_2} - Y_{a_2, b_1}| + |Y_{a_2, b_1} - Y_{a_1, b_1}| < 2n^{\frac{\alpha}{2} + \epsilon};$$

changing $\epsilon$ obtains the desired result.

Fix $\eta$ such that $(\alpha, \eta) \in \Delta_1$ (as defined in Equation 7). We may assume that $\epsilon < \eta/4$. We first consider $a_1 < a_2$ and $b$ fixed. From Lemma 3.2 we have

$$\Pr\left[|Y_{a_2,b} - Y_{a_1,b}| > n^{\frac{\alpha}{2}+\epsilon}\right] = \Pr\left[|Y_{a_2-a_1,b}| > n^{\frac{\alpha}{2}+\epsilon}\right].$$

Assume first that $n^\alpha < b < n - n^\alpha$. Let $a_3 = a_2 - a_1 + \lfloor n^\alpha \rfloor + 1$ so that $n^\alpha < a_3 < 2n^\alpha$. We apply Lemma 2.8 with $a = a_3$, $b = b$ and $H = n^{\epsilon/2}$: here $\mathcal{C} < 2n^{\alpha-1}$ and $H\sqrt{\mathcal{C}n} < \frac{1}{4}n^{\frac{\alpha}{2}+\epsilon}$. We have

$$\Pr\left[|Y_{a_3,b}| > \frac{1}{4}n^{\frac{\alpha}{2}+\epsilon}\right] < \Pr\left[|Y_{a_3,b}| > H\sqrt{\mathcal{C}n}\right] <$$

$$< n^{-\epsilon/2}\exp\left(-\frac{n^\epsilon}{2}\right) < \exp(-n^{\epsilon/2}).$$

Again from Lemma 3.2,

$$\Pr\left[|Y_{a_3,b} - Y_{a_2-a_1,b}| > \frac{1}{4}n^{\frac{\alpha}{2}+\epsilon}\right] = \Pr\left[|Y_{a_3-(a_2-a_1),b}| > \frac{1}{4}n^{\frac{\alpha}{2}+\epsilon}\right];$$

a computation similar to the previous one gives

$$\Pr\left[|Y_{a_3,b} - Y_{a_2-a_1,b}| > \frac{1}{4}n^{\frac{\alpha}{2}+\epsilon}\right] < \exp(-n^{\epsilon/2})$$

and therefore
$$\Pr\left[|Y_{a_2-a_1,b}| > \frac{1}{2}n^{\frac{\alpha}{2}+\epsilon}\right] < 2\exp(-n^{\epsilon/2}).$$

Consider now $b \le n^\alpha$; let $b_3 = b + \lfloor n^\alpha \rfloor + 1$. From our previous results,

$$\Pr\left[|Y_{a_2-a_1,b_3}| > \frac{1}{2}n^{\frac{\alpha}{2}+\epsilon}\right] < 2\exp(-n^{\epsilon/2});$$

$$\Pr\left[|Y_{a_2-a_1,b_3-b}| > \frac{1}{2}n^{\frac{\alpha}{2}+\epsilon}\right] < 2\exp(-n^{\epsilon/2}).$$

But again from Lemma 3.2,

$$\Pr\left[|Y_{a_2-a_1,b_3-b}| > \frac{1}{2}n^{\frac{\alpha}{2}+\epsilon}\right] = \Pr\left[|Y_{a_2-a_1,b_3} - Y_{a_2-a_1,b}| > \frac{1}{2}n^{\frac{\alpha}{2}+\epsilon}\right],$$

and we have
$$\Pr\left[|Y_{a_2-a_1,b}| > n^{\frac{\alpha}{2}+\epsilon}\right] < 4\exp(-n^{\epsilon/2}).$$

The case $b \ge n - n^\alpha$ is similar.

We now consider all valid triples $(a_1, a_2, b)$: there are fewer that $n^3$ such triples. Thus

$$\Pr\left[\exists (a_1, a_2, b), |a_1 - a_2| < n^\alpha, \ |Y_{a_2-a_1,b}| > n^{\frac{\alpha}{2}+\epsilon}\right] < 4n^3\exp(-n^{\epsilon/2}) \to 0,$$

completing the proof. $\qquad\square$

The following result can be seen as a concentration of measure: almost every $X$ is near the product copula $C_0$. The exponent $1/2$ is sharp; we shall prove a stronger result later.

**Proposition 4.3.** *Let $X$ be a random discrete copula with uniform distribution. Then, a.a.s.,*

$$\forall (a,b), \ |Y_{n,a,b}| < \sqrt{n \log(n)}.$$

*Proof.* Consider $\gamma \in (1/2, 1)$. Consider a subset $M$ of $\{0, 1, \ldots, n\}$ with $m \approx n^{(1-\gamma)}$ elements and gaps between consecutive elements at a bounded distance from $n^\gamma$. With $\alpha = \gamma$, given an arbitrary pair $(a_1, b_1)$ there exists $(a_2, b_2) \in M^2$ with $|a_1 - a_2|, |b_1 - b_2| < n^\alpha$ and therefore, from Lemma 4.2 (a.a.s.),

$$|Y_{a_1,b_1} - Y_{a_2,b_2}| < n^{(1+\gamma)/4}.$$

Set $\alpha = (2 + \gamma)/3$, $\eta = (1 - \gamma)/5$ and $H = 2\sqrt{\log(n)}$. For $(a, b) \in M^2$ we have $\mathcal{C} \le 1/16$ so that $H\sqrt{\mathcal{C}n} < \frac{1}{2}\sqrt{n\log(n)}$. From Lemma 2.8,

$$\Pr\left[|Y_{n,a,b}| > \frac{1}{2}\sqrt{n\log(n)}\right] < \Pr\left[|Y_{n,a,b}| > H\sqrt{\mathcal{C}n}\right] < \frac{1}{2\sqrt{\log(n)}}\, n^{-2}.$$

We now have

$$\Pr\left[\left(\max_{(a,b)\in M^2} |Y_{n,a,b}|\right) > \frac{1}{2}\sqrt{n\log(n)}\right] \le$$

$$\le \sum_{(a,b)\in M^2} \Pr\left[|Y_{n,a,b}| > \frac{1}{2}\sqrt{n\log(n)}\right] \le \frac{1}{2\sqrt{\log(n)}}\, m^2 n^{-2} \to 0.$$

Thus, a.a.s.,

$$\max_{(a,b)} |Y_{n,a,b}| \le \frac{1}{2}\sqrt{n\log(n)} + n^{(1+\gamma)/4} < \sqrt{n\log(n)},$$

as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The weak point in the above proof is handling all points in $M^2$ independently, which we know to be far from true. In the next section we study joint distributions in order to improve this result and our understanding of $X$ and $Y$.

## 5 Large grids

In order to complete the proof of Theorem 1.1, we now consider special grids with $I = J$ growing as a function of $n$. Given $\gamma \in (5/6, 1)$, we construct the $\gamma$-*grid* $\mathbf{g}_\gamma$. Let $m = I = J = \lfloor n^{(1-\gamma)} \rfloor$; for $0 \le i \le n$, let $a_i = b_i = \lfloor \frac{in}{m} + \frac{1}{2} \rfloor$ so that gaps are approximately equal: $\dot{a}_i = a_i - a_{i-1} \in (n^\gamma - 1, n^\gamma + 1)$.

We already established a correspondence between families $(c_{i,j})$ and $(\dot{c}_{i,j})$. Recall that the same information is given by the family $(\tilde{l}_{i,j})$ of real numbers with $\tilde{l} \in L_{\mathbf{g}_\gamma} \subset V_{\mathbf{g}_\gamma}$. Notice that $\sum_i \tilde{l}_{i,j} \approx \sum_j \tilde{l}_{i,j} \approx 0$ and that

$$\dot{c}_{i,j} \approx n^{(2\gamma-1)} + \tilde{l}_{i,j} n^{(\gamma-\frac{1}{2})} = n^{(2\gamma-1)}(1 + \tilde{l}_{i,j} n^{(-\gamma+\frac{1}{2})});$$

the small errors arise from taking integer parts in the definition of $a_i$, $b_j$. As usual, we write $P(\mathbf{g}_\gamma; \tilde{l}) = P(n, a, b, c)$.

In Section 3, we needed Lemma 3.7 in order to prove Theorem 3.1. We follow a similar strategy here, and the analogue of Lemma 3.7 is Lemma 5.1 below. Let $L_{\mathbf{g}_\gamma, \eta} \subset L_{\mathbf{g}_\gamma}$ be the finite set of vectors $\tilde{l}$ such that, for all $(i,j)$, we have $|\tilde{l}_{i,j}| < n^\eta$.

**Lemma 5.1.** *Consider $\gamma \in (5/6, 1)$ and $\eta > 0$ fixed; then*

$$\lim_{n \to \infty} \Pr[\tilde{h} \in L_{\mathbf{g}_\gamma, \eta}] = 1.$$

*Proof.* We will use Lemma 2.8 with $\alpha$ slightly smaller than our $\gamma$. We have $n^\alpha < a_1 = b_1 < n - n^\alpha$ and $|a_1 - n^\gamma| < 1$ and therefore

$$\left| \frac{a_1 b_1}{n} - n^{2\gamma - 1} \right| < 2n^{\gamma - 1} + n^{-1} < 1.$$

We have

$$|\tilde{h}_{1,1}| n^{(\gamma - \frac{1}{2})} = |X_{a_1, b_1} - n^{(2\gamma - 1)}| \leq |Y_{a_1, b_1}| + \left| \frac{a_1 b_1}{n} - n^{(2\gamma - 1)} \right| < |Y_{a_1, b_1}| + 1.$$

Also, $\sqrt{\mathcal{C}n} < n^{(\gamma - \frac{1}{2})}$. Set $H = \frac{1}{2} n^\eta$: from Lemma 2.8,

$$\Pr\left[ |\tilde{h}_{1,1}| \geq n^\eta \right] \leq \Pr\left[ |Y_{a_1, b_1}| > \sqrt{\mathcal{C}n} H \right] < 2n^{-\eta} \exp\left( -\frac{n^{2\eta}}{8} \right).$$

From Lemma 3.2 we know that the other cases are similar, that is,

$$\Pr\left[ |\tilde{h}_{i,j}| \geq n^\eta \right] < 2n^{-\eta} \exp\left( -\frac{n^{2\eta}}{8} \right).$$

for all $(i,j)$ and therefore

$$\Pr\left[ \exists (i,j),\ |\tilde{h}_{i,j}| \geq n^\eta \right] < 2n^{2-\eta} \exp\left( -\frac{n^{2\eta}}{8} \right) \to 0,$$

completing the proof. $\square$

We present a result similar to Lemmas 2.3 and 3.5. For simplicity, from now on we take $\gamma = 19/20$ and $\eta = 1/20$.

21

**Lemma 5.2.** *Set $\gamma = 19/20$ and $\eta = 1/20$. For $n \in \mathbb{N}$, consider the $\gamma$-grid $\mathbf{g}_\gamma$. If $\tilde{l} \in L_{\mathbf{g}_\gamma, \eta}$ then*

$$\log(P(\mathbf{g}_\gamma, \tilde{l})) = C_n - \frac{|\tilde{l}|^2}{2} + o(1),$$

$$C_n = -\frac{(m-1)^2}{2} \log(2\pi) + \left( \frac{m^2 - 1}{2} - m(m-1)\gamma \right) \log n.$$

Notice that the little oh notation is employed in the sense discussed after the statement of Lemma 2.5.

*Proof.* We know from Lemma 3.3 that

$$\log(P(\mathbf{g}_\gamma, \tilde{l})) = \sum_i \log(\dot{a}_i!) + \sum_j \log(\dot{b}_j!) - \log(n!) - \sum_{i,j} \log(\dot{c}_{i,j}!);$$

the idea is again to use Stirling approximation but since the number of terms is unbounded we must proceed more carefully. Expand using Equation 3,

$$\log(k!) = \mathrm{xl}(k) - k + \frac{1}{2} \log(k) + \frac{1}{2} \log(2\pi) + \mathrm{st}(k), \quad 0 < \mathrm{st}(k) < 1/(12n),$$

and consider the five such terms from right to left.

$$|T_5| = \left| \sum_i \mathrm{st}(\dot{a}_i) + \sum_j \mathrm{st}(\dot{b}_j) - \mathrm{st}(n) - \sum_{i,j} \mathrm{st}(\dot{c}_{i,j}) \right| <$$

$$< \frac{1}{12} \left( \sum_i \frac{1}{\dot{a}_i} + \sum_j \frac{1}{\dot{b}_j} + \frac{1}{n} + \sum_{i,j} \frac{1}{\dot{c}_{i,j}} \right) <$$

$$< \frac{1}{6} \left( \frac{n^{(1-\gamma)} + 1}{n^\gamma - 1} + \frac{n^{(1-\gamma)} + 1}{n^\gamma - 1} + \frac{1}{n} + \frac{(n^{(1-\gamma)} + 1)^2}{n^{(2\gamma-1)} - n^{1/20} n^{(\gamma - \frac{1}{2})}} \right) \to 0,$$

so the usual Stirling approximation is safe after all. The next term $T_4$ equals

$$\sum_i \frac{\log(2\pi)}{2} + \sum_j \frac{\log(2\pi)}{2} - \left( \frac{\log(2\pi)}{2} \right) - \sum_{i,j} \frac{\log(2\pi)}{2} = -\frac{(m-1)^2}{2} \log(2\pi).$$

Also, since $|\log(n^{-\gamma} \dot{a}_i)| < 2n^{-\gamma}$ and $|\log(n^{(1-2\gamma)} \dot{c}_{i,j})| < Cn^{(\frac{11}{20} - \gamma)}$,

$$T_3 = \sum_i \frac{\log \dot{a}_i}{2} + \sum_j \frac{\log \dot{b}_j}{2} - \left( \frac{\log n}{2} \right) - \sum_{i,j} \frac{\log \dot{c}_{i,j}}{2} =$$

$$= \sum_i \frac{\gamma \log n}{2} + \sum_j \frac{\gamma \log n}{2} - \left( \frac{\log n}{2} \right) - \sum_{i,j} \frac{(2\gamma - 1) \log n}{2} +$$

$$+ \sum_i \frac{\log(n^{-\gamma} \dot{a}_i)}{2} + \sum_j \frac{\log(n^{-\gamma} \dot{b}_j)}{2} - \sum_{i,j} \frac{\log(1 + \tilde{l}_{i,j} n^{(-\gamma + \frac{1}{2})})}{2} =$$

$$= \left( \frac{m^2 - 1}{2} - m(m-1)\gamma \right) \log n + o(1).$$

We have
$$T_2 = -\sum_i \dot{a}_i - \sum_j \dot{b}_j + n + \sum_{i,j} \dot{c}_{i,j} = 0.$$

Simplifications similar to the proof of Lemma 2.3 yields that $T_1$ equals

$$\sum_i \mathrm{xl}(\dot{a}_i) + \sum_j \mathrm{xl}(\dot{b}_j) - \mathrm{xl}(n) - \sum_{i,j} \mathrm{xl}(\dot{c}_{i,j}) = -\sum_{i,j} \frac{\dot{a}_i \dot{b}_j}{n} \, \mathrm{xl}\left(1 + \sqrt{\frac{n}{\dot{a}_i \dot{b}_j}}\, \tilde{l}_{i,j}\right).$$

In preparation to use Equation 6, $\mathrm{xl}(1+t) = t + t^2/2 + O(t^3)$, notice that

$$m = O(n^{(1-\gamma)}), \quad \left|\sqrt{\frac{n}{\dot{a}_i \dot{b}_j}}\, \tilde{l}_{i,j}\right| = O(n^{\frac{11}{20}-\gamma}).$$

Thus

$$T_1 = -\sum_{i,j} \sqrt{\frac{\dot{a}_i \dot{b}_j}{n}}\, \tilde{l}_{i,j} - \sum_{i,j} \frac{\tilde{l}_{i,j}^2}{2} + O(n^{(\frac{53}{20}-3\gamma)}) = -\sum_{i,j} \frac{\tilde{l}_{i,j}^2}{2} + o(1),$$

completing the proof. $\qquad\square$

Use the lattice $L_{\mathbf{g}_\gamma}$ to define the Voronoi decomposition of the space $V_{\mathbf{g}_\gamma}$:

$$V_{\mathbf{g}_\gamma} = \bigcup_{\tilde{l} \in L_{\mathbf{g}_\gamma}} R_{\tilde{l}};$$

$$v \in R_{\tilde{l}} \iff \left(\forall \tilde{l}' \in L_{\mathbf{g}_\gamma},\ |v - \tilde{l}| \leq |v - \tilde{l}'|\right).$$

Notice that the interiors of the cells $R_{\tilde{l}}$ are pairwise disjoint. We only need a crude estimate of the diameter $\mathrm{diam}(R_{\tilde{l}})$ of a Voronoi cell.

**Lemma 5.3.** *Set $\gamma = 19/20$. We have* $\mathrm{diam}(R_{\tilde{l}}) = O(n^{(\frac{5}{2}-3\gamma)})$.

*Proof.* Given a point $\tilde{z} \in V_{\mathbf{g}_\gamma}$, we obtain a nearby $\tilde{l} \in L_{\mathbf{g}_\gamma}$ (usually not the nearest one!). We shall use the conditions in displays 10 and 11 which define $V_{\mathbf{g}_\gamma}$ and $L_{\mathbf{g}_\gamma}$. Set, for $1 \leq i, j < m$ (that is, for entries in the $(m-1) \times (m-1)$ leading principal minor), $\tilde{l}_{i,j}$ to be the element of the set

$$\frac{1}{\sqrt{n\hat{a}_i\hat{b}_j}}\left(\mathbb{Z} - \frac{\dot{a}_i \dot{b}_j}{n}\right)$$

nearest to $\tilde{z}_{i,j}$. Use all but two of the equations defining $V_{\mathbf{g}_\gamma}$,

$$\sum_j \sqrt{\hat{b}_j}\, \tilde{l}_{i,j} = 0, \quad \sum_i \sqrt{\hat{a}_i}\, \tilde{l}_{i,j} = 0,$$

23

for $1 \le i < m$ and $1 \le j < m$, to obtain $\tilde{l}_{i,m}$ and $\tilde{l}_{m,j}$. Finally, set $\tilde{l}_{m,m}$ to satisfy the two remaining equations,

$$\sum_i \sqrt{\hat{a}_i}\, \tilde{l}_{i,m} = \sum_j \sqrt{\hat{b}_j}\, \tilde{l}_{m,j} = 0,$$

thus defining $\tilde{l} \in L_{\mathbf{g}_\gamma}$. For $1 \le i, j < m$,

$$|\tilde{z}_{i,j} - \tilde{l}_{i,j}| < 1/\sqrt{n\hat{a}_i\hat{b}_j} < 2n^{(\frac{1}{2} - \gamma)}.$$

For $1 \le i < m$, thus, $|\tilde{z}_{i,m} - \tilde{l}_{i,m}| < 2mn^{(\frac{1}{2} - \gamma)} < 4n^{(\frac{3}{2} - 2\gamma)}$. A similar estimate holds for $|\tilde{z}_{m,j} - \tilde{l}_{m,j}|$ (with $1 \le j < m$). Finally, $|\tilde{z}_{m,m} - \tilde{l}_{m,m}| < 8n^{(\frac{5}{2} - 3\gamma)}$ and, computing the euclidean norm of the difference,

$$|\tilde{z} - \tilde{l}| < 16n^{(\frac{5}{2} - 3\gamma)},$$

completing the proof. $\qquad\square$

Endow the vector space $V_{\mathbf{g}_\gamma}$ with a unit normal measure: for an open subset $A \subseteq V_{\mathbf{g}_\gamma}$, set

$$\mu(A) = (2\pi)^{-\frac{(m-1)^2}{2}} \int_A \exp\left(-\frac{|\tilde{w}|^2}{2}\right) d\tilde{w}.$$

**Lemma 5.4.** *Set* $\gamma = 19/20$ *and* $\eta = 1/20$. *For* $n \in \mathbb{N}$, *consider the* $\gamma$-*grid* $\mathbf{g}_\gamma$. *If* $\tilde{l} \in L_{\mathbf{g}_\gamma, \eta}$ *then*

$$P(\mathbf{g}_\gamma, \tilde{l}) \sim \mu(R_{\tilde{l}}).$$

The $\sim$ symbol above is used in the sense discussed after the statement of Lemma 2.5. More explicitly: given $\epsilon > 0$ there exists $N$ such that if $n > N$ and $\tilde{l} \in L_{\mathbf{g}_\gamma, \eta}$ then

$$1 - \epsilon < \frac{P(\mathbf{g}_\gamma, \tilde{l})}{\mu(R_{\tilde{l}})} < 1 + \epsilon.$$

*Proof.* Notice that the condition that $|\tilde{l}_{i,j}| < n^\eta$ for all $(i,j)$ implies that $|\tilde{l}|^2 < m^2 n^{2\eta} < 4n^{(2 - 2\gamma + 2\eta)}$ and therefore $|\tilde{l}| < 2n^{(1 - \gamma + \eta)}$. In this region,

$$\int_{R_{\tilde{l}}} \exp\left(-\frac{|\tilde{w}|^2}{2}\right) d\tilde{w} \sim \mathrm{Vol}(R_{\tilde{l}}) \exp\left(-\frac{|\tilde{l}|^2}{2}\right).$$

Indeed, for $\tilde{w} \in R_{\tilde{l}}$, we have (from Lemma 5.3) $|\tilde{w} - \tilde{l}| < Cn^{(\frac{5}{2} - 3\gamma)}$ and $|\tilde{l}|, |\tilde{w}| < 4n^{(1 - \gamma + \eta)}$. Thus $||\tilde{w}|^2 - |\tilde{l}|^2| < ||\tilde{w}| + |\tilde{l}|| \cdot ||\tilde{w}| - |\tilde{l}|| < 8Cn^{(\frac{7}{2} - 4\gamma + \eta)} \to 0$, proving the claim.

Thus, from Lemma 5.2, assuming $\tilde{l} \in L_{\mathbf{g}_\gamma, \eta}$,

$$P(\mathbf{g}_\gamma, \tilde{l}) \sim \tilde{C}_n\, \mu(R_{\tilde{l}}) = \tilde{C}_n\, (2\pi)^{-\frac{(m-1)^2}{2}} \int_{R_{\tilde{l}}} \exp\left(-\frac{|\tilde{w}|^2}{2}\right) d\tilde{w}$$
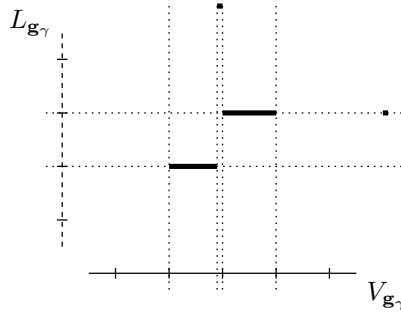
Figure 4: Projecting $\Omega_n$ on $V_{\mathbf{g}_\gamma}$ and $L_{\mathbf{g}_\gamma}$

for some positive constant $\tilde{C}_n$. On the other hand, we know from Lemma 5.1 that

$$\sum_{\tilde{l} \in L_{\mathbf{g}_\gamma, \eta}} \mu(R_{\tilde{l}}) \to 1, \qquad \sum_{\tilde{l} \in L_{\mathbf{g}_\gamma, \eta}} P(\mathbf{g}_\gamma, \tilde{l}) \to 1.$$

Thus, $\tilde{C}_n$ can be taken to be 1, completing the proof. $\qquad\qquad\square$

Notice that we computed the approximate value of $\mathrm{Vol}(R_{\tilde{l}})$:

$$\mathrm{Vol}(R_{\tilde{l}}) \sim n^{\left( \frac{m^2 - 1}{2} - m(m-1)\gamma \right)}.$$

We are ready to present the proof of Theorem 1.1.

*Proof.* We describe the space $\Omega_n$ and define the random variables $X_n$ (which obtains a discrete copula) and $f$ (the bridged Brownian sheet). First, we obtain random variables $\tilde{z} \in V_{\mathbf{g}_\gamma}$ and $\tilde{h} \in L_{\mathbf{g}_\gamma}$ (towards samplings of $f$ and $\sqrt{n}\, \mathbf{y}_n$) with

$$\Pr_{\Omega_n}[\tilde{z} \in A] = \mu(A), \quad A \subset V_{\mathbf{g}_\gamma}; \qquad \Pr_{\Omega_n}[\tilde{h} = \tilde{l}] = P(\mathbf{g}_\gamma, \tilde{l}).$$

These two random variables are far from independent. In a way, we would like to always have $\tilde{z} \in R_{\tilde{h}}$ but that is not quite possible since in principle $\mu(R_{\tilde{l}}) \neq P(\mathbf{g}_\gamma, \tilde{l})$.

Figure 4 shows the beginning of the construction of $\Omega_n$ and these two random variables. Sometimes, as in the figure, we have $\mu(R_{\tilde{l}}) > P(\mathbf{g}_\gamma, \tilde{l})$. For some values of $\tilde{z} \in R_{\tilde{l}}$, we must therefore have $\tilde{h} \neq \tilde{l}$. In the second step, $\mu(R_{\tilde{l}}) < P(\mathbf{g}_\gamma, \tilde{l})$ and, for some values of $\tilde{z}$ in other Voronoi cells, we have $\tilde{h} = \tilde{l}$. We know, however, from Lemmas 5.1 and 5.4 that, following this construction,

$$\lim_{n \to \infty} \Pr_{\Omega_n}[\tilde{z} \in R_{\tilde{h}}] = 1.$$

Let $u_i = v_i = a_i/n$ so that $\mathbf{g}_\gamma$ can be interpreted as a grid in the unit square $[0, 1]^2$. Given $\tilde{h} = \tilde{l} \in L_{\mathbf{g}_\gamma}$, uniformly select $X_n$ from the set of discrete copulas

$C$ for which, as in Equation 9,

$$\hat{c}_{i,j} = \hat{a}_i \hat{b}_j + \sqrt{\frac{\hat{a}_i \hat{b}_j}{n}} \, \tilde{l}_{i,j}$$

and therefore

$$\sqrt{n} \, \mathbf{y}_n(u_i, v_j) = \sum_{i' \le i;\, j' \le j} \sqrt{\hat{a}_{i'} \hat{b}_{j'}} \, \tilde{h}_{i',j'}.$$

As in Equation 8, set the values of $f$ on $\mathbf{g}_\gamma$ to be

$$f(u_i, v_j) = \sum_{i' \le i;\, j' \le j} \sqrt{\hat{a}_{i'} \hat{b}_{j'}} \, \tilde{z}_{i',j'}.$$

From this point on, obtain a bridged Brownian sheet $f$ in the usual way. This completes the construction of $\Omega_n$ and of the random variables $X_n$ and $f$. Items (a) and (b) in the statement of the theorem are trivial; we are left with proving (c).

We may disregard subsets of $\Omega_n$ whose probability tends to zero (as a function of $n$). We may therefore assume that $\tilde{z} \in R_{\tilde{h}}$ and, from Lemma 5.3, that $|\tilde{z} - \tilde{h}| < n^{(\frac{5}{2} - 3\gamma)}$. For a point $(u_i, v_j)$ on the grid $\mathbf{g}_\gamma$ we have

$$\sqrt{n} \, \mathbf{y}_n(u_i, v_j) - f(u_i, v_j) = \sum_{i' \le i;\, j' \le j} \sqrt{\hat{a}_{i'} \hat{b}_{j'}} \, (\tilde{h}_{i',j'} - \tilde{z}_{i',j'})$$

and therefore, from Lemma 5.3,

$$\left| \sqrt{n} \, \mathbf{y}_n(u_i, v_j) - f(u_i, v_j) \right| = \left| \langle M, (\tilde{h} - \tilde{z}) \rangle \right| \le 2|(\tilde{h} - \tilde{z})| \le 2n^{(\frac{5}{2} - 3\gamma)}. \quad (13)$$

Here $M_{i',j'} = \sqrt{\hat{a}_{i'} \hat{b}_{j'}}$ for $i' \le i$, $j' \le j$ and 0 otherwise; clearly, $|M|^2 = \sum (M_{i',j'})^2 \le 4$.

For an arbitrary point $(u, v) \in \mathbb{I}^2$, let $(u_i, v_j)$ be the point in the grid $\mathbf{g}_\gamma$ closest to $(u, v)$. Apply Lemma 4.2 with $\alpha = \gamma = 19/20$ and $\epsilon = 1/100$ to deduce that, a.a.s., $X_n$ is $(\alpha, \epsilon)$-Hölder and therefore, for all $(u, v) \in \mathbb{I}^2$,

$$|\sqrt{n} \, \mathbf{y}_n(u, v) - \sqrt{n} \, \mathbf{y}_n(u_i, v_j)| < n^{(\frac{\gamma}{2} + \epsilon - \frac{1}{2})}. \quad (14)$$

Furthermore, from Lemma 4.1 with $r = 4n^{-\frac{1}{20}}$, $\delta = \frac{1}{5}$ and $C = \frac{1}{4}$, a.a.s. we have that, for all $(u, v) \in \mathbb{I}^2$,

$$|f(u, v) - f(u_i, v_j)| < n^{(\frac{\gamma}{2} + \epsilon - \frac{1}{2})}. \quad (15)$$

Thus, from Equations 13, 14 and 15, we a.a.s. have

$$|\sqrt{n} \mathbf{y}_n - f|_{C^0} < 2n^{(\frac{\gamma}{2} + \epsilon - \frac{1}{2})} + 2n^{(\frac{5}{2} - 3\gamma)} \to 0,$$

completing the proof. $\qquad \square$

26

# 6    A concentration result

In this section we present some applications of the main theorem. We recall some basic statistics for a bridged Brownian sheet $f \in C_0^0(\mathbb{I}^2)$ ([2], [3]). For any fixed $p \in [1, +\infty]$

$$\Pr\left[\|f\|_{L^p} < r\right] = \rho^p(r)$$

defines a strictly increasing continuous function $\rho^p : [0, +\infty) \to \mathbb{I}$ with

$$\lim_{r \to 0} \rho^p(r) = 0; \quad \lim_{r \to +\infty} \rho^p(r) = 1.$$

Theorem 1.1 obtains similar results for discrete copulas. For $C \in \mathcal{S}_n$ and $p \in [1, +\infty)$ define

$$d^p(C) = \left( \frac{1}{n^2} \sum_{a,b} \left| \frac{C_{a,b} - \frac{ab}{n}}{\sqrt{n}} \right|^p \right)^{1/p} ; \qquad d^\infty(C) = \max_{a,b} \left| \frac{C_{a,b} - \frac{ab}{n}}{\sqrt{n}} \right|.$$

**Corollary 6.1.** *Let $p \in [1, +\infty]$ and $r \in (0, 1)$ be fixed. Then*

$$\lim_{n \to \infty} \Pr\left[d^p(X) < r\right] = \rho^p(r).$$

*Proof.* Theorem 1.1 directly yields the case $p = \infty$. For the general case, notice that $\|\sqrt{n}\, \mathbf{y}_n - f\|_{C_0} < \epsilon$ implies $\|\sqrt{n}\, \mathbf{y}_n - f\|_{L^p} < \epsilon$.    □

A weaker statement is the following concentration result, which makes no reference to bridged Brownian sheets.

**Corollary 6.2.** *Let $g : \mathbb{N} \to \mathbb{R}$ be a function with $\lim g(n) = +\infty$. Let $X_n$ be a random discrete copula in $\mathcal{S}_n$. Then*

$$\lim_{n \to \infty} \Pr\left[\|\mathbf{y}_n\|_{C^0} > \frac{g(n)}{\sqrt{n}}\right] = 0.$$

*Proof.* The result follows from Corollary 6.1 and $\lim_{r \to +\infty} \rho^\infty(r) = 1$.    □

# 7    Discrete copulas and lozenge tilings

The graph of a discrete copula can also be seen as a pile of cubes. Figure 5 exemplifies this correspondence for a permutation in $S_8$. The north-east corner of the square on the left is taken to the south vertex of the bottom of the cube. The bullets correspond to the black squares.

This defines a bijection between $\mathcal{S}_n$ and a certain set of tilings of a regular hexagon of side $n - 1$. Clearly, not every gravitationally stable pile of cubes corresponds to a permutation: the grey walls are forced and the three configurations to the right of Figure 5 are forbidden (in the white part). It is not hard to prove that these conditions characterize tilings corresponding to copulas.

Theorem 1.1 implies that if we randomly select a tiling corresponding to a copula the surface of the pile of cubes is essentially an affine transformation of a Brownian sheet. The reader may contrast this result with the celebrated Arctic circle theorem (see [5]): in our scenario there are no "frozen" regions.
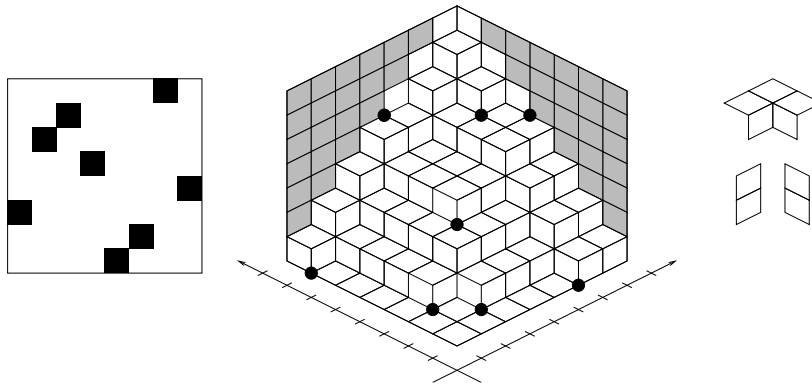
Figure 5: A permutation and its representation as a lozenge tiling of a subset of the hexagon; the configurations to the right are forbidden.

# 8    Birkhoff copulas

A *Birkhoff copula* is a matrix $C$ with *real* entries satisfying conditions (i), (ii) and (iii) in the definition of discrete copula. Discrete copulas are associated with permutation matrices. The more general Birkhoff copulas correspond to doubly stochastic matrices. More precisely, the *Birkhoff polytope* $B_n$ is the convex set of $n \times n$ doubly stochastic matrices, whose vertices are the permutation matrices. The counterpart to $\mathcal{S}_n$ is its convex hull $\mathcal{B}_n$, the polytope of Birkhoff copulas. The natural bijection $B_n \to \mathcal{B}_n$ takes $M$ to $C$, where again $C$ is obtained by integration:

$$C_{a,b} = \sum_{a' \leq a;\ b' \leq b} M_{a',b'} \,.$$

Thus, a Birkhoff copula indicates the joint distribution of two discrete random variables $Z_0$ and $Z_1$, both uniform in $\{1, 2, \ldots, n\}$ but usually not independent.

For the Birkhoff case, numerical experiments (using a Metropolis style algorithm) obtain uniformly distributed random elements of $\mathcal{B}_n$, the set of Birkhoff copulas. We then study their statistics, particularly the analogue of Theorem 3.1. The results ratify the following conjecture, a counterpart of Theorem 1.1.

**Conjecture 8.1.** *There exist probability spaces $\Omega_n$ with the following properties.*

(a)  *On $\Omega_n$ there exists a random variable $X_n$ uniformly distributed in $\mathcal{B}_n$.*

(b)  *On $\Omega_n$ a sample-continuous bridged Brownian sheet process $f$ in $C_0^0(\mathbb{I}^2)$ is defined.*

(c)  *For all $\epsilon > 0$, $\lim_{n \to \infty} \Pr\left[\|n\,\mathbf{y}_n - f\|_{C^0} > \epsilon\right] = 0$.*

More precisely: we select points $(u_1, v_1), \ldots, (u_m, v_m) \in (0,1)^2$ and compare the empirically observed joint distribution of $\mathbf{y}_n(u_i, v_i)$ with the joint normal distribution predicted by the conjecture: they match.

# References

[1] Bradley, R. C., *Basic properties of strong mixing conditions. A survey and some open questions*, Probability Surveys, 2, 107–144, 2005.

[2] Dudley, R. M., *Real analysis and probability*, Cambridge University Press, 2002.

[3] Dudley, R. M., *Uniform Central Limit Theorems*, Cambridge University Press, 1999.

[4] Erdely, A., González-Barrios, J. M. and Nelsen, R. B., *Symmetries of random discrete copulas*, Kybernetika, 44:6, 846–863, 2008.

[5] Jockusch, W., Propp, J. and Shor, P., *Random Domino Tilings and the Arctic Circle Theorem*, arXiv:math/9801068.

[6] Léandre, R., *The geometry of Brownian surfaces*, Probability Surveys, 3, 37–88, 2006.

[7] Mayor, G., Suñer, J. and Torrens, J., *Copula-like operations on finite settings*, IEEE Transactions on Fuzzy Systems, 13:4, 468–477, 2005.

[8] Nelsen, R. B., *An Introduction to Copulas*, Springer, 2006.

Juliana Freire, Nicolau C. Saldanha and Carlos Tomei, PUC-Rio
jufreire@gmail.com
saldanha@puc-rio.br; http://www.nicolausaldanha.com
tomei@mat.puc-rio.br

Departamento de Matemática, PUC-Rio
R. Marquês de S. Vicente 225, Rio de Janeiro, RJ 22453-900, Brazil