

# On the exact region determined by Kendall's tau and Spearman's rho

Manuela Schreyer

Roland Paulin<sup>†</sup>

Wolfgang Trutschnig<sup>‡</sup>

*University of Salzburg, Salzburg, Austria.*

**Summary.** Using properties of shuffles of copulas and tools from combinatorics we solve the open question about the exact region  $\Omega$  determined by all possible values of Kendall's  $\tau$  and Spearman's  $\rho$ . In particular, we prove that the well-known inequality established by Durbin and Stuart in 1951 is not sharp outside a countable set, give a simple analytic characterization of  $\Omega$  in terms of a continuous, strictly increasing piecewise concave function, and show that  $\Omega$  is compact and simply connected, but not convex. The results also show that for each  $(x, y) \in \Omega$  there are mutually completely dependent random variables  $X, Y$  whose  $\tau$  and  $\rho$  values coincide with  $x$  and  $y$  respectively.

*Keywords:* Concordance, Copula, Kendall tau, Shuffle, Spearman rho

## 1. Introduction

Kendall's  $\tau$  and Spearman's  $\rho$  are, without doubt, the two most famous nonparametric measures of concordance. Given random variables  $X, Y$  with continuous distribution functions  $F$  and  $G$  respectively, Spearman's  $\rho$  is defined as the Pearson correlation coefficient of the  $\mathcal{U}(0, 1)$ -distributed random variables  $U := F \circ X$  and  $V := G \circ Y$  whereas Kendall's  $\tau$  is given by the probability of concordance minus the probability of discordance, i.e.

$$\begin{aligned}\rho(X, Y) &= 12(\mathbb{E}(UV) - \frac{1}{4}) \\ \tau(X, Y) &= \mathbb{P}((X_1 - X_2)(Y_1 - Y_2) > 0) - \mathbb{P}((X_1 - X_2)(Y_1 - Y_2) < 0),\end{aligned}$$

where  $(X_1, Y_1)$  and  $(X_2, Y_2)$  are independent and have the same distribution as  $(X, Y)$ . Since both measures are scale invariant they only depend on the underlying (uniquely

<sup>†</sup>Supported by the Austrian Science Fund (FWF): P24574

<sup>‡</sup>Corresponding author, Hellbrunnerstr. 34, A-5020 Salzburg, wolfgang.trutschnig@sbg.ac.at

determined) copula  $A$  of  $(X, Y)$ . It is well known and straightforward to verify (Nelsen, 2006) that, given the copula  $A$  of  $(X, Y)$ , Kendall's  $\tau$  and Spearman's  $\rho$  can be expressed as

$$\tau(X, Y) = 4 \int_{[0,1]^2} A(x, y) d\mu_A(x, y) - 1 =: \tau(A) \quad (1)$$

$$\rho(X, Y) = 12 \int_{[0,1]^2} xy d\mu_A(x, y) - 3 =: \rho(A), \quad (2)$$

where  $\mu_A$  denotes the doubly stochastic measure corresponding to  $A$ . Considering that  $\tau$  and  $\rho$  quantify different aspects of the underlying dependence structure (Fredricks and Nelsen, 2007) a very natural question is how much they can differ, i.e. if  $\tau(X, Y)$  is known which values may  $\rho(X, Y)$  assume and vice versa. The first of the following two well-known universal inequalities between  $\tau$  and  $\rho$  goes back to Daniels (1950), the second one to Durbin and Stuart (1951) (for alternative proofs see Kruskal (1958); Genest and Nešlehová (2009); Nelsen (2006)):

$$|3\tau - 2\rho| \leq 1 \quad (3)$$

$$\frac{(1 + \tau)^2}{2} - 1 \leq \rho \leq 1 - \frac{(1 - \tau)^2}{2} \quad (4)$$

The inequalities together yield the set  $\Omega_0$  (see Figure 1) which we will refer to as *classical  $\tau$ - $\rho$  region* in the sequel. Daniels' inequality is known to be sharp (Nelsen, 2006) whereas the first part of the inequality by Durbin and Stuart is only known to be sharp at the points  $\mathbf{p}_n = (-1 + \frac{2}{n}, -1 + \frac{2}{n^2})$  with  $n \geq 2$  (which, using symmetry, is to say that the second part is sharp at the points  $-\mathbf{p}_n$ ). Although both inequalities are known since the 1950s and the interrelation between  $\tau$  and  $\rho$  has received much attention also in recent years, in particular concerning the so-called Hutchinson-Lai conjecture (Fredricks and Nelsen, 2007; Hürlimann, 2003; Balakrishnan and Lai, 2009), to the best of the authors' knowledge the *exact  $\tau$ - $\rho$  region*  $\Omega$ , defined by ( $\mathcal{C}$  denoting the family of all two-dimensional copulas)

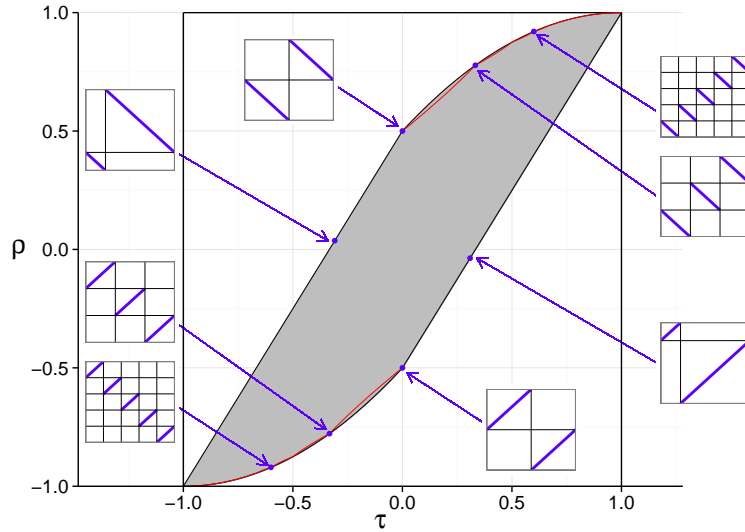
$$\begin{aligned} \Omega &= \{(\tau(X, Y), \rho(X, Y)) : X, Y \text{ continuous random variables}\} \\ &= \{(\tau(A), \rho(A)) : A \in \mathcal{C}\}, \end{aligned} \quad (5)$$

is still unknown.

In this paper we solve the sixty year old question and give a full characterization of  $\Omega$ . We derive a piecewise concave, strictly increasing, continuous function  $\Phi : [-1, 1] \rightarrow [-1, 1]$  and (see Theorem 3.6 and Theorem 5.1) prove that

$$\Omega = \{(x, y) \in [-1, 1]^2 : \Phi(x) \leq y \leq -\Phi(-x)\}. \quad (6)$$

Figure 1 depicts  $\Omega_0$  and the function  $\Phi$  (lower red line), the explicit form of  $\Phi$  is given in eq. (18) and eq. (19). As a byproduct we get that the inequality by Durbin and Stuart is not sharp outside the aforementioned points  $\mathbf{p}_n$  and  $-\mathbf{p}_n$ , that  $\Omega$  is compact and simply connected, but not convex. Moreover, we prove the surprising fact that for each point  $(x, y) \in \Omega$  there exist mutually completely dependent random variables  $X, Y$  for which  $(\tau(X, Y), \rho(X, Y)) = (x, y)$  holds.



**Fig. 1.** The classical  $\tau$ - $\rho$ -region  $\Omega_0$  and some copulas (distributing mass uniformly on the blue segments) for which the inequality by Durbin and Stuart is sharp. The red line depicts the true boundary of  $\Omega$ .

The rest of the paper is organized as follows: Section 2 gathers some notations and preliminaries. In Section 3 we reduce the problem of determining  $\Omega$  to a problem about so-called shuffles of copulas, prove some properties of shuffles and derive the function  $\Phi$ . The main result saying that  $\Omega$  is contained in the right-hand-side of eq. (6) is given in Section 4, tedious calculations needed for the proofs being collected in the Appendix. Section 5 serves to prove equality in eq. (6) and to collect some interesting consequences of this result. Finally, in Section 6 we gather some new conjectures on the exact  $\tau$ - $\rho$  region for well-known subclasses of copulas.

## 2. Notation and Preliminaries

As already mentioned before,  $\mathcal{C}$  will denote the family of all two-dimensional *copulas*, see Durante and Sempi (2015); Embrechts et al. (2003); Nelsen (2006).  $M$  and  $W$  will denote the upper and the lower Fréchet-Hoeffding bound respectively. Given  $A \in \mathcal{C}$  the *transpose*

$A^t \in \mathcal{C}$  of  $A$  is defined by  $A^t(x, y) := A(y, x)$  for all  $x, y \in [0, 1]$ .  $d_\infty$  will denote the uniform distance on  $\mathcal{C}$ ; it is well known that  $(\mathcal{C}, d_\infty)$  is a compact metric space and that  $d_\infty$  is a metrization of weak convergence in  $\mathcal{C}$ . For every  $A \in \mathcal{C}$  the corresponding *doubly stochastic measure* will be denoted by  $\mu_A$ , i.e. we have  $\mu_A([0, x] \times [0, y]) := A(x, y)$  for all  $x, y \in [0, 1]$ .  $\mathcal{P}_\mathcal{C}$  denotes the class of all these doubly stochastic measures.  $\mathcal{B}([0, 1])$  and  $\mathcal{B}([0, 1]^2)$  will denote the Borel  $\sigma$ -fields in  $[0, 1]$  and  $[0, 1]^2$ ,  $\lambda$  and  $\lambda_2$  the Lebesgue measure on  $\mathcal{B}([0, 1])$  and  $\mathcal{B}([0, 1]^2)$  respectively. Instead of  $\lambda$ -a.e. we will simply write a.e. since no confusion will arise.  $\mathcal{T}$  will denote the class of all  $\lambda$ -preserving transformations  $h : [0, 1] \rightarrow [0, 1]$ , i.e. transformations for which the push-forward  $\lambda^h$  of  $\lambda$  via  $h$  coincides with  $\lambda$ ,  $\mathcal{T}_b$  the subclass of all bijective  $h \in \mathcal{T}$ .

For every copula  $A \in \mathcal{C}$  there exists a *Markov kernel* (regular conditional distribution)  $K_A : [0, 1] \times \mathcal{B}([0, 1]) \rightarrow [0, 1]$  fulfilling  $(G_x := \{y \in [0, 1] : (x, y) \in G\})$  denoting the  $x$ -section of  $G \in \mathcal{B}([0, 1]^2)$  for every  $x \in [0, 1]$ )

$$\int_{[0,1]} K_A(x, G_x) d\lambda(x) = \mu_A(G), \quad (7)$$

for every  $G \in \mathcal{B}([0, 1]^2)$ , so, in particular

$$\int_{[0,1]} K_A(x, F) d\lambda(x) = \lambda(F) \quad (8)$$

for every  $F \in \mathcal{B}([0, 1])$ , see Trutschnig (2011). We will refer to  $K_A$  simply as Markov kernel of  $A$ . On the other hand, every Markov kernel  $K : [0, 1] \times \mathcal{B}([0, 1]) \rightarrow [0, 1]$  fulfilling (8) induces a unique element  $\mu \in \mathcal{P}_\mathcal{C}([0, 1]^2)$  via (7). For more details and properties of disintegration we refer to (Kallenberg, 2002; Klenke, 2013).

A copula  $A \in \mathcal{C}$  will be called *completely dependent* if and only if there exists  $h \in \mathcal{T}$  such that  $K(x, E) := \mathbf{1}_E(h(x))$  is a Markov kernel of  $A$  (see Trutschnig (2011) for equivalent definitions and main properties). For every  $h \in \mathcal{T}$  the induced completely dependent copula will be denoted by  $A_h$ . Note that  $h_1 = h_2$  a.e. implies  $A_{h_1} = A_{h_2}$  and that eq. (7) implies  $A_h(x, y) = \lambda([0, x] \cap h^{-1}([0, y]))$  for all  $x, y \in [0, 1]$ . In the sequel  $\mathcal{C}_d$  will denote the family of all completely dependent copulas.  $A_h \in \mathcal{C}_d$  will be called *mutually completely dependent* if we even have  $h \in \mathcal{T}_b$ . Note that in case of  $h \in \mathcal{T}_b$  we have  $A_{h^{-1}} = (A_h)^t$ . Complete dependence is the opposite of independence since it describes the (not necessarily mutual) situation of full predictability/maximum dependence.

Tackling the problem of determining the region  $\Omega$ , our main tool will be special members of the class  $\mathcal{C}_d$  usually referred to as *shuffles of the minimum copula  $M$* . Following (Nelsen, 2006) we will call  $h \in \mathcal{T}_b$  a *shuffle* (and  $A_h \in \mathcal{C}_d$  a shuffle of  $M$ ) if there exist  $0 = s_0 < s_1 < \dots < s_{n-1} < s_n = 1$  and  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \in \{-1, 1\}^n$  such that we have  $h'(x) = \varepsilon_i$

for every  $x \in (s_{i-1}, s_i)$ . In case of  $\varepsilon_i = 1$  for every  $i \in \{1, \dots, n\}$  we will call  $h$  *straight shuffle*.  $\mathcal{S}$  will denote the family of all shuffles,  $\mathcal{S}^+$  the family of all straight shuffles. It is well known (Mikusinski et al., 1992; Nelsen, 2006) that  $\mathcal{C}_{\mathcal{S}^+}$ , defined by

$$\mathcal{C}_{\mathcal{S}^+} = \{A_h : h \in \mathcal{S}^+\} \quad (9)$$

is dense in  $(\mathcal{C}, d_\infty)$ . For more general definitions of shuffles we refer to (Durante and Sempi, 2015). Obviously every shuffle  $h \in \mathcal{S}$  can be expressed in terms of vectors  $\mathbf{u} \in \Delta_n, \varepsilon \in \{-1, 1\}^n$  and a permutation  $\pi \in \sigma_n$ , where  $\Delta_n$  denotes the unit simplex  $\Delta_n = \{\mathbf{x} \in [0, 1]^n : \sum_{i=1}^n x_i = 1\}$  and  $\sigma_n$  denotes all bijections on  $\{1, \dots, n\}$ . In fact, choosing suitable  $\mathbf{u} \in \Delta_n, \varepsilon \in \{-1, 1\}^n, \pi \in \sigma_n$ , setting (empty sums are zero by definition)

$$s_k := \sum_{i=1}^k u_i, \quad t_k := \sum_{i=1}^k u_{\pi^{-1}(i)} \quad (10)$$

for every  $k \in \{0, \dots, n\}$ , we have  $s_k - s_{k-1} = u_k = t_{\pi(k)} - t_{\pi(k)-1}$  and on  $(s_{k-1}, s_k)$  the shuffle  $h$  is given by

$$h(x) = h_{\pi, \mathbf{u}, \varepsilon}(x) := \begin{cases} t_{\pi(k)-1} + x - s_{k-1} & \text{if } \varepsilon_k = 1, \\ t_{\pi(k)} - (x - s_{k-1}) & \text{if } \varepsilon_k = -1. \end{cases} \quad (11)$$

In the sequel we will work directly with the function  $h_{\pi, \mathbf{u}, \varepsilon}$ , implicitly defined in eq. (11) since all possible extensions of  $h_{\pi, \mathbf{u}, \varepsilon}$  from  $\bigcup_{i=1}^k (s_{k-1}, s_k)$  to  $[0, 1]$  yield the same copula, which we will denote by  $A_{h_{\pi, \mathbf{u}, \varepsilon}}$ . In case of  $\varepsilon_i = 1$  for every  $i \in \{1, \dots, n\}$  we will simply write  $h_{\pi, \mathbf{u}}$  in the sequel. Note that the chosen representation is not unique, i.e. for given  $\mathbf{u} \in \Delta_n, \varepsilon \in \{-1, 1\}^n, \pi \in \sigma_n$  there always exist  $\mathbf{u}' \in \Delta_m, \varepsilon' \in \{-1, 1\}^m, \pi' \in \sigma_m$  with  $m \neq n$  such that  $h_{\pi, \mathbf{u}, \varepsilon} = h_{\pi', \mathbf{u}', \varepsilon'}$  a.e., implying  $A_{h_{\pi, \mathbf{u}, \varepsilon}} = A_{h_{\pi', \mathbf{u}', \varepsilon'}}$ . So, for instance, the shuffle  $h_{\pi, \mathbf{u}, \varepsilon}$  with  $\pi = (4, 2, 1, 3), \mathbf{u} = (\frac{1}{8}, \frac{3}{8}, \frac{1}{4}, \frac{1}{4})$  and  $\varepsilon = (1, -1, 1, 1)$  and the shuffle  $h_{\pi', \mathbf{u}', \varepsilon'}$  with  $\pi' = (5, 3, 1, 2, 4), \mathbf{u}' = (\frac{1}{8}, \frac{3}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{4})$  and  $\varepsilon' = (1, -1, 1, 1, 1)$  coincide a.e. and induce the same copula.

REMARK 2.1. It might seem more natural to work directly with minimal representations (minimal dimension  $n$ ) and to exclude the case of  $u_k = 0$  for some  $k$  (implying  $(s_{k-1}, s_k) = \emptyset$ ) in the first place – since we will, however, use various compactness arguments in the sequel the chosen representation is more convenient.

### 3. Basic properties of $\Omega$ and some results on shuffles

In this section we will first show that for determining  $\Omega$  it is sufficient to consider straight shuffles, give explicit formulas for  $(\tau(A_h), \rho(A_h))$  for arbitrary  $h \in \mathcal{S}^+$ , and derive a strictly

increasing function  $\Phi : [-1, 1] \rightarrow [-1, 1]$  which, after some change of coordinates, will finally be shown to determine  $\Omega$  in the subsequent section.

We start with some observations about  $\Omega$ . Considering that the mapping  $f : \mathcal{C} \rightarrow [-1, 1]^2$ , defined by  $f(A) = (\tau(A), \rho(A))$ , is continuous w.r.t.  $d_\infty$  (Scarsini, 1984), the compactness of  $(\mathcal{C}, d_\infty)$  implies the compactness of  $\Omega$ . As a consequence, using eq. (5) and the fact that  $\mathcal{C}_{\mathcal{S}^+}$  is dense we immediately get ( $\bar{U}$  denoting the closure of a set  $U$ )

$$\Omega = \overline{\{(\tau(A_h), \rho(A_h)) : h \in \mathcal{S}^+\}}. \quad (12)$$

Based on this, our method of proof will be to construct a compact set  $\Omega_\Phi$  (fully determined by the function  $\Phi$ ) fulfilling  $(\tau(A_h), \rho(A_h)) \in \Omega_\Phi$  for every  $h \in \mathcal{S}^+$  since we then automatically get  $\Omega \subseteq \Omega_\Phi$ .

Being concordance measures,  $\tau$  and  $\rho$  fulfil the axioms mentioned in Scarsini (1984), so  $\Omega$  is also symmetric w.r.t.  $(0, 0)$ . Analogously, it is straightforward to verify that  $\tau(A^t) = \tau(A)$  as well as  $\rho(A^t) = \rho(A)$  holds for every  $A \in \mathcal{C}$ , implying

$$\tau(A_{h^{-1}}) = \tau(A_h), \quad \rho(A_{h^{-1}}) = \rho(A_h) \quad (13)$$

for every  $h \in \mathcal{T}_b$ .

For every  $h \in \mathcal{T}$  define the quantities  $\text{inv}(h)$  and  $\text{invsum}(h)$  (notation loosely based on Sack and Úlfarsson (2011)) by

$$\text{inv}(h) = \int_{[0,1]^2} \mathbf{1}_{[0,x]}(y) \mathbf{1}_{(h(x),1]}(h(y)) d\lambda_2(x, y) \quad (14)$$

$$\text{invsum}(h) = \int_{[0,1]^2} \mathbf{1}_{[0,x]}(y) \mathbf{1}_{(h(x),1]}(h(y))(x - y) d\lambda_2(x, y). \quad (15)$$

LEMMA 3.1. *For every  $h \in \mathcal{T}_b$  the following relations hold:*

$$\tau(A_h) = 4 \int_{[0,1]} A_h(x, h(x)) d\lambda(x) - 1 = 1 - 4 \text{inv}(h)$$

$$\rho(A_h) = 12 \int_{[0,1]} xh(x) d\lambda(x) - 3 = 1 - 12 \text{invsum}(h)$$

Moreover, for every  $h \in \mathcal{T}_b$  we have  $(\text{inv}(h), \text{invsum}(h)) \in [0, \frac{1}{2}] \times [0, \frac{1}{6}]$ .

PROOF. Using disintegration we immediately get

$$\tau(A_h) = 4 \int_{[0,1]} \int_{[0,1]} A_h(x, y) K_{A_h}(x, dy) d\lambda(x) - 1 = 4 \int_{[0,1]} A_h(x, h(x)) d\lambda(x) - 1$$

as well as

$$\text{inv}(h) = \int_{[0,1]} \int_{[0,x]} \left(1 - \mathbf{1}_{[0,h(x)]}(h(y))\right) d\lambda(y) d\lambda(x)$$

$$= \int_{[0,1]} (x - A_h(x, h(x))) d\lambda(x) = \frac{1 - \tau(A_h)}{4}$$

which proves the first identity. The first part of the second one is an immediate consequence of disintegration. To prove the remaining equality use  $\int_{[0,x]} \mathbf{1}_{(h(x),1]}(h(y)) d\lambda(y) = x - A_h(x, h(x))$  and  $\int_{(y,1]} \mathbf{1}_{[0,h(y))}(h(x)) d\lambda(x) = h(y) - A_h(y, h(y))$  to get

$$\begin{aligned} \text{invsum}(h) &= \int_{[0,1]} x(x - A_h(x, h(x)) - (h(x) - A_h(x, h(x)))) d\lambda(x) \\ &= \frac{1}{3} - \int_{[0,1]} xh(x) d\lambda(x). \end{aligned}$$

The fact that  $(\text{inv}(h), \text{invsum}(h)) \in [0, \frac{1}{2}] \times [0, \frac{1}{6}]$  is a direct consequence of  $\Omega \subseteq [-1, 1]^2$ .

As next step we derive explicit formulas for  $\text{inv}(h)$  and  $\text{invsum}(h)$  for the case of  $h$  being a straight shuffle based on which we will afterwards derive the afore-mentioned function  $\Phi$  determining the region  $\Omega$ . To simplify notation define

$$\begin{aligned} I_\pi &= \{\{i, j\} : 1 \leq i < j \leq n, \pi(i) > \pi(j)\} \\ Q_\pi &= \{\{i, j, k\} : 1 \leq i < j < k \leq n, \pi(i) > \pi(j) > \pi(k) \text{ or} \\ &\quad \pi(j) > \pi(k) > \pi(i) \text{ or } \pi(k) > \pi(i) > \pi(j)\}, \end{aligned} \quad (16)$$

as well as

$$a_\pi(\mathbf{u}) = \text{inv}(h_{\pi, \mathbf{u}}), \quad b_\pi(\mathbf{u}) = \text{inv}(h_{\pi, \mathbf{u}}) - 2 \text{invsum}(h_{\pi, \mathbf{u}}) \quad (17)$$

for every  $\pi \in \sigma_n$  and  $\mathbf{u} \in \Delta_n$ . The following lemma (the proof of which is given in the Appendix) holds.

LEMMA 3.2. *For every  $(\pi, \mathbf{u}) \in \sigma_n \times \Delta_n$  the following identities hold:*

$$\begin{aligned} \text{inv}(h_{\pi, \mathbf{u}}) &= a_\pi(\mathbf{u}) = \sum_{i < j, \{i, j\} \in I_\pi} u_i u_j \\ \text{invsum}(h_{\pi, \mathbf{u}}) &= \sum_{i < j, \{i, j\} \in I_\pi} \left( \frac{1}{2} u_i^2 u_j + \frac{1}{2} u_i u_j^2 + \sum_{k: i < k < j} u_i u_j u_k \right) \\ b_\pi(\mathbf{u}) &= \sum_{i < j < k, \{i, j, k\} \in Q_\pi} u_i u_j u_k \end{aligned}$$

REMARK 3.3. Notice that in (Genest and Nešlehová, 2007, Propositions 4 and 5) a slightly different notation ( $h_{i,j}$ ,  $\alpha_{i,j}$  and  $\beta_{i,j}$  instead of  $u_p$ ,  $s_{p-1}$ ,  $t_{\pi(p)-1}$ ) is used to derive analogous formulas. Lemma 3.2 and Theorem 4.1 are the main reason for our choice of notation in this paper: the expressions for  $a_\pi(\mathbf{u})$  and  $b_\pi(\mathbf{u})$ , which are key in the proof of the main result, are simplest possible.

As pointed out in the Introduction, the first part of inequality (4) is known to be sharp only at the points  $\mathbf{p}_n = (-1 + \frac{2}{n}, -1 + \frac{2}{n^2})$  with  $n \geq 2$ . According to (Nelsen, 2006), or directly using Lemma 3.2, considering  $\boldsymbol{\pi} = (n, n-1, \dots, 2, 1)$  and  $u_1 = u_2 = \dots = u_n = \frac{1}{n}$  we get  $\mathbf{p}_n = (\tau(A_{h_{\boldsymbol{\pi}, \mathbf{u}}}), \rho(A_{h_{\boldsymbol{\pi}, \mathbf{u}}}))$ . Having this, it seems natural to conjecture that all shuffles of the form  $A_{h_{\boldsymbol{\pi}, \mathbf{u}}}$  with

$$\boldsymbol{\pi} = (n, n-1, \dots, 2, 1), \quad u_1 = u_2 = \dots = u_{n-1} = r, u_n = 1 - (n-1)r$$

for some  $n \geq 2$  and  $r \in (\frac{1}{n}, \frac{1}{n-1})$  might also be extremal in the sense that  $(\tau(A_{h_{\boldsymbol{\pi}, \mathbf{u}}}), \rho(A_{h_{\boldsymbol{\pi}, \mathbf{u}}}))$  is a boundary point of  $\Omega$ . The Main content of this paper is the confirmation of this very conjecture. We will assign all shuffles of the just mentioned form the name prototype, calculate  $\tau$  and  $\rho$  explicitly for all prototypes and then, based on these values, derive the function  $\Phi$ .

**DEFINITION 3.4.**  $\boldsymbol{\pi} \in \sigma_n$  will be called decreasing if  $\boldsymbol{\pi} = (n, n-1, \dots, 2, 1)$ . The pair  $(\boldsymbol{\pi}, \mathbf{u}) \in \sigma_n \times \Delta_n$  will be called a prototype if  $\boldsymbol{\pi}$  is decreasing and there exists some  $r \in [\frac{1}{n}, \frac{1}{n-1}]$  such that  $u_1 = u_2 = \dots = u_{n-1} = r$  and  $u_n = 1 - (n-1)r$ . Analogously,  $h \in \mathcal{S}^+$  (and  $A_h \in \mathcal{C}_d$ ) is called a prototype if there exists a prototype  $(\boldsymbol{\pi}, \mathbf{u})$  such that  $h = h_{\boldsymbol{\pi}, \mathbf{u}}$  a.e.

Using the identities from Lemma 3.2 we get the following expressions for prototypes (the proof is given in the Appendix):

**LEMMA 3.5.** Suppose that  $(\boldsymbol{\pi}, \mathbf{u}) \in \sigma_n \times \Delta_n$  is a prototype, then

$$\begin{aligned} \tau(A_{h_{\boldsymbol{\pi}, \mathbf{u}}}) &= 1 - 4(n-1)r + 2r^2n(n-1) \in \left[\frac{2-n}{n}, \frac{2-(n-1)}{n-1}\right] \\ \rho(A_{h_{\boldsymbol{\pi}, \mathbf{u}}}) &= 1 - 2r(n-1)(3 - 3r(n-1) + r^2(n-2)n) \in \left[\frac{2-n^2}{n^2}, \frac{2-(n-1)^2}{(n-1)^2}\right]. \end{aligned}$$

Fix  $n \geq 2$ . Then both functions  $r \mapsto 1 - 4(n-1)r + 2r^2n(n-1)$  and  $r \mapsto 1 - 2r(n-1)(3 - 3r(n-1) + r^2(n-2)n)$  are strictly increasing on  $[\frac{1}{n}, \frac{1}{n-1}]$ . Expressing  $r$  as function of  $\tau$  and substituting the result in the expression for  $\rho$  directly yields

$$\rho(A_{h_{\boldsymbol{\pi}, \mathbf{u}}}) = -1 - \frac{4}{n^2} + \frac{3}{n} + \frac{3\tau(A_{h_{\boldsymbol{\pi}, \mathbf{u}}})}{n} - \frac{n-2}{\sqrt{2n^2\sqrt{n-1}}}(n-2 + n\tau(A_{h_{\boldsymbol{\pi}, \mathbf{u}}}))^{3/2}.$$

Based on this interrelation define  $\Phi_n : [-1 + \frac{2}{n}, 1] \rightarrow [-1, 1]$  by

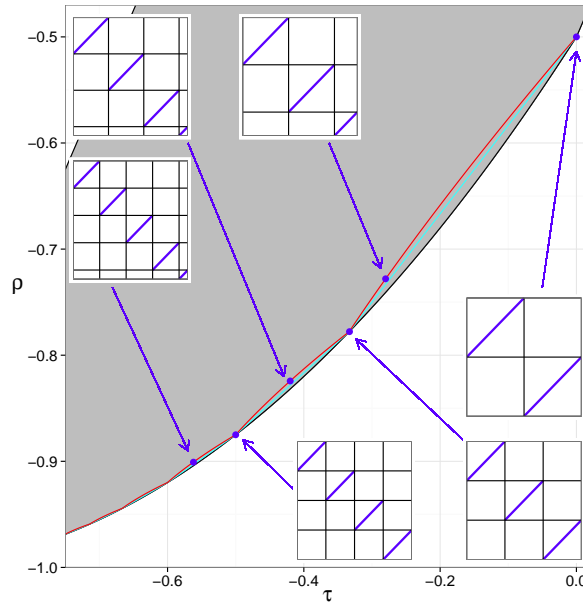
$$\Phi_n(x) = -1 - \frac{4}{n^2} + \frac{3}{n} + \frac{3x}{n} - \frac{n-2}{\sqrt{2n^2\sqrt{n-1}}}(n-2 + nx)^{3/2} \quad (18)$$

and set

$$\Phi(x) = \begin{cases} -1 & \text{if } x = -1, \\ \Phi_n(x) & \text{if } x \in \left[\frac{2-n}{n}, \frac{2-(n-1)}{n-1}\right] \text{ for some } n \geq 2. \end{cases} \quad (19)$$



Since we have  $\Phi_n(\frac{2-n}{n}) = \Phi_{n+1}(\frac{2-n}{n}) = -1 + \frac{2}{n^2}$  for every  $n \geq 1$  this defines a function  $\Phi : [-1, 1] \rightarrow [-1, 1]$ . Notice that  $\Phi_2(x) = -\frac{1}{2} + \frac{3x}{2}$ , i.e. on  $[0, 1]$   $\Phi$  coincides with Daniels' linear bound and for  $x_n = \frac{2-n}{n}$  and  $n \geq 2$  we have  $(x_n, \Phi(x_n)) = \mathbf{p}_n$ , i.e.  $(x_n, \Phi(x_n))$  coincides with the points at which Durbin and Stuart's inequality is known to be sharp. Furthermore, it is straightforward to verify that  $\Phi$  is a strictly increasing homeomorphism on  $[-1, 1]$  which is concave on every interval  $[\frac{2-n}{n}, \frac{2-(n-1)}{n-1}]$  with  $n \geq 2$ . Figure 2 depicts the function  $\Phi$  as well as some prototypes and their corresponding Kendall's  $\tau$  and Spearman's  $\rho$ .



**Fig. 2.** The function  $\Phi$  (red) and some prototypes with their corresponding Kendall's  $\tau$  and Spearman's  $\rho$ . The shaded region depicts the classical  $\tau$ - $\rho$ -region  $\Omega_0$ , straight lines connecting the points  $p_n$  are plotted in green.

Defining the compact set  $\Omega_\Phi$  by

$$\Omega_\Phi = \{(x, y) \in [-1, 1]^2 : \Phi(x) \leq y \leq -\Phi(-x)\}, \quad (20)$$

we can now state the following main result the proof of which is given in the next section.

**THEOREM 3.6.** *The precise  $\tau$ - $\rho$  region  $\Omega$  fulfils  $\Omega \subseteq \Omega_\Phi$ .*

**REMARK 3.7.** The fact that  $\Omega \subseteq \Omega_\Phi$  holds is the principal result of this paper since it improves the classical inequality by Durbin and Stuart mentioned in the Introduction and, more importantly, gives sharp bounds everywhere. In Section 5 we will, however, show that even  $\Omega = \Omega_\Phi$  holds and that for every point  $(x, y) \in \Omega$  there exists a shuffle  $h \in \mathcal{S}$  such that  $(\tau(A_h), \rho(A_h)) = (x, y)$ .

REMARK 3.8. A function similar (but not identical) to  $\Phi$  has appeared in the literature in Shao et al. (2014), where the authors tried to deduce sharp bounds of  $\Omega$  by running simulations (but did not provide any analytic proof). Additionally, it has been brought to our attention during the preparation of this manuscript that Manuel Úbeda-Flores (University of Almería, Spain) already conjectured Theorem 3.6 (with the exact form of  $\Phi$ ) in a unpublished working paper in 2009.

#### 4. Proof of the main theorem

Using the properties of  $\Omega$  mentioned at the beginning of Section 3, Theorem 3.6 is proved if we can show that for every  $h \in \mathcal{S}^+$  we have  $\rho(A_h) \geq \Phi(\tau(A_h))$ . Given Lemma 3.2 it is straightforward to verify that this is equivalent to showing  $\text{invsum}(h) \leq \varphi(\text{inv}(h))$  for every  $h \in \mathcal{S}^+$  where  $\varphi : [0, \frac{1}{2}] \rightarrow [0, \frac{1}{6}]$  is defined by

$$\varphi(x) = \begin{cases} \frac{1}{6} & \text{if } x = \frac{1}{2}, \\ \varphi_n(x) & \text{if } x \in [\frac{1}{2} - \frac{1}{2(n-1)}, \frac{1}{2} - \frac{1}{2n}] \text{ for some } n \geq 2 \end{cases} \quad (21)$$

and  $\varphi_n : [\frac{1}{2} - \frac{1}{2(n-1)}, \frac{1}{2} - \frac{1}{2n}] \rightarrow [0, \frac{1}{6}]$  is given by

$$\varphi_n(x) = \frac{1}{6} + \frac{1}{3n^2} - \frac{1}{2n} + \frac{x}{n} + \frac{n-2}{6n^2\sqrt{n-1}}(n-1-2nx)^{3/2}. \quad (22)$$

Translating this to  $a_\pi(\mathbf{u})$  and  $b_\pi(\mathbf{u})$ , using eq. (17) and defining  $\vartheta : [0, \frac{1}{2}] \rightarrow [0, \frac{1}{6}]$  by  $\vartheta(x) = x - 2\varphi(x)$  we arrive at the following equivalent form of Theorem 3.6:

THEOREM 4.1. *For every  $n \in \mathbb{N}$ ,  $\pi \in \sigma_n$  and  $\mathbf{u} \in \Delta_n$  the following inequality holds:*

$$b_\pi(\mathbf{u}) \geq \vartheta(a_\pi(\mathbf{u})) \quad (23)$$

We are now going to prove this result and start with some first observations and an outline of the structure of the subsequent proof. (i)  $\vartheta$  is continuous and, by calculating the derivative, it is straightforward to see that  $\vartheta$  is non-decreasing. (ii)  $\vartheta(0) = \vartheta(\frac{1}{4}) = 0$  and  $\vartheta(\frac{1}{2}) = \frac{1}{6}$ . (iii) For every prototype  $(\pi, \mathbf{u})$  we have the equality  $b_\pi(\mathbf{u}) = \vartheta(a_\pi(\mathbf{u}))$ . (iv) For given  $n$  and fixed  $\pi \in \sigma_n$  the functions  $\mathbf{v} \mapsto a_\pi(\mathbf{v})$  and  $\mathbf{v} \mapsto b_\pi(\mathbf{v})$  are continuous on  $\Delta_n$ , so there exists some  $\mathbf{u} \in \Delta_n$  minimizing the function  $\mathbf{v} \mapsto b_\pi(\mathbf{v}) - \vartheta(a_\pi(\mathbf{v}))$ . (v) For  $n \leq 2$  the inequality  $b_\pi(\mathbf{u}) \geq \vartheta(a_\pi(\mathbf{u}))$  trivially holds for every  $\pi \in \sigma_n$  and every  $\mathbf{u} \in \Delta_n$ , so **from now on we will only consider the case  $n \geq 3$ .**

The structure of the proof of Theorem 4.1 is as follows:

- (a) Preliminary Step 1: We prove inequality (23) for the case of decreasing  $\pi \in \sigma_n$ .

- (b) Preliminary Step 2: We analyze how, for fixed  $\boldsymbol{\pi} \in \sigma_n$ , the quantities  $a_{\boldsymbol{\pi}}(\mathbf{u})$  and  $b_{\boldsymbol{\pi}}(\mathbf{u})$  change if  $\mathbf{u} \in \Delta_n$  changes.
- (c) Induction Step 1: Assuming that the result is true for all  $(\boldsymbol{\pi}, \mathbf{u}) \in \sigma_m \times \Delta_m$  with  $m < n$  we prove inequality (23) for  $(\boldsymbol{\pi}, \mathbf{u}) \in \sigma_n \times \Delta_n$  under the hypothesis that there either exist (i)  $p < q < r$  such that  $\boldsymbol{\pi}(r) > \boldsymbol{\pi}(q) > \boldsymbol{\pi}(p)$  or (ii)  $p < q < r < s$  such that  $\boldsymbol{\pi}(q) > \boldsymbol{\pi}(p) > \boldsymbol{\pi}(s) > \boldsymbol{\pi}(r)$  holds.
- (d) Induction Step 2: Assuming that the result is true for all  $(\boldsymbol{\pi}, \mathbf{u}) \in \sigma_m \times \Delta_m$  with  $m < n$  we prove inequality (23) for  $(\boldsymbol{\pi}, \mathbf{u}) \in \sigma_n \times \Delta_n$  with  $\boldsymbol{\pi}$  not fulfilling the hypothesis in Induction Step I.

*Preliminary Step 1:* Consider  $n \geq 3$  and  $\boldsymbol{\pi} = (n, n-1, \dots, 2, 1)$ . Note that in this situation we have  $e_1(\mathbf{u}) = 1, e_2(\mathbf{u}) = a_{\boldsymbol{\pi}}(\mathbf{u}), e_3(\mathbf{u}) = b_{\boldsymbol{\pi}}(\mathbf{u})$  for every  $\mathbf{u} \in \Delta_n$ , where  $e_i$  denotes the  $i$ -th elementary symmetric polynomial for  $i \in \{1, 2, 3\}$ , i.e.  $e_1(\mathbf{v}) := \sum_i v_i, e_2(\mathbf{v}) := \sum_{i < j} v_i v_j$  and  $e_3(\mathbf{v}) := \sum_{i < j < k} v_i v_j v_k$  for every  $\mathbf{v} \in \mathbb{R}^n$ . Hence  $a_{\boldsymbol{\pi}}(\mathbf{u})$  and  $b_{\boldsymbol{\pi}}(\mathbf{u})$  do not change if we reorder the coordinates of  $\mathbf{u}$ .

LEMMA 4.2. *Suppose that  $n \geq 3, \boldsymbol{\pi} = (n, n-1, \dots, 2, 1), c_2 \in a_{\boldsymbol{\pi}}(\Delta_n)$  and that  $\mathbf{u} \in \Delta_n$  fulfils  $b_{\boldsymbol{\pi}}(\mathbf{u}) = \min\{b_{\boldsymbol{\pi}}(\mathbf{v}) : \mathbf{v} \in \Delta_n \cap (a_{\boldsymbol{\pi}})^{-1}(\{c_2\})\}$  as well as  $u_1 \geq \dots \geq u_n \geq 0$ . Then there exists  $m \in \{1, \dots, n\}$  such that  $u_i = 0$  for every  $i > m$ , and  $u_1 = \dots = u_{m-1} \geq u_m$ .*

PROOF. Note that the continuity of  $b_{\boldsymbol{\pi}}$  and the compactness of  $\Delta_n \cap (a_{\boldsymbol{\pi}})^{-1}(\{c_2\})$  imply the existence of the minimum. We first prove the statement for the case  $n = 3$  and suppose that  $\mathbf{u}$  is a minimizer fulfilling  $u_3 \geq u_2 \geq u_1 \geq 0$ . Define a polynomial  $f : \mathbb{R} \rightarrow \mathbb{R}$  by

$$f(T) = (T - u_1)(T - u_2)(T - u_3) = T^3 - T^2 + c_2 T - e_3(\mathbf{u}),$$

and let  $D_f$  denote the discriminant of  $f$ . It is well known that  $D_f > 0$  if and only if  $f$  has three distinct real zeros and that in case of  $D_f \neq 0$  locally the zeros of  $f$  are smooth (so in particular continuous) functions of the coefficients of  $f$ .

Suppose that  $u_1 > u_2 > u_3 > 0$ . Then  $D_f > 0$ . Let  $f_{\epsilon}(T) = T^3 - T^2 + c_2 T - (e_3(\mathbf{u}) - \epsilon)$ , then for small enough values of  $\epsilon > 0$ , the polynomial  $f_{\epsilon}$  has three distinct, positive real zeros:  $u_{\epsilon,1}, u_{\epsilon,2}, u_{\epsilon,3}$ . Then  $u_{\epsilon,1} + u_{\epsilon,2} + u_{\epsilon,3} = 1$  and  $u_{\epsilon,1}u_{\epsilon,2} + u_{\epsilon,2}u_{\epsilon,3} + u_{\epsilon,3}u_{\epsilon,1} = c_2$ , while  $u_{\epsilon,1}u_{\epsilon,2}u_{\epsilon,3} = e_3(\mathbf{u}) - \epsilon < e_3(\mathbf{u})$ , which is a contradiction. So either  $u_3 = 0$  or  $u_1 = u_2$  or  $u_1 > u_2 = u_3 > 0$ . In the first two cases we are done, so suppose that  $u_1 > u_2 = u_3 > 0$ . Then  $1 = u_1 + 2u_2$  and  $c_2 = 2u_1u_2 + u_2^2$ . Suppose that  $u_1 \geq 4u_2$ . Then  $1 \geq 4c_2$ , so there are unique  $y_1 \geq y_2 \geq 0$  such that  $y_1 + y_2 = 1$  and  $y_1y_2 = c_2$ . Let  $y_3 = 0$ , then considering  $\mathbf{y} = (y_1, y_2, y_3)$  we get  $e_1(\mathbf{y}) = 1, e_2(\mathbf{y}) = c_2$ , and  $e_3(\mathbf{y}) = 0 < e_3(\mathbf{u})$ , which is a contradiction. So  $4u_2 > u_1 > u_2$ . Let  $y_1 = y_2 = \frac{2u_1 + u_2}{3}$  and  $y_3 = \frac{4u_2 - u_1}{3}$ . Then  $y_1, y_2, y_3 \geq 0, e_1(\mathbf{y}) = 1,$

$e_2(\mathbf{y}) = c_2$ , and  $e_3(\mathbf{y}) = \frac{1}{27}(2u_1 + u_2)^2(4u_2 - u_1) = e_3(\mathbf{u}) - \frac{4}{27}(u_1 - u_2)^3 < e_3(\mathbf{u})$ , which is a contradiction. This proves the claim for  $n = 3$ .

Suppose indirectly that the statement is false for some  $n > 3$ . Then there are  $i < j < k$  such that  $u_i > u_j \geq u_k > 0$ . Setting  $\bar{u}_l := \frac{u_l}{u_i + u_j + u_k}$  for every  $l \in \{i, j, k\}$  obviously  $\bar{u}_i + \bar{u}_j + \bar{u}_k = 1$ . Applying the case  $n = 3$  to  $\bar{u}_i, \bar{u}_j, \bar{u}_k$  yields  $\bar{y}_i, \bar{y}_j, \bar{y}_k \in [0, 1]$  such that  $\bar{y}_i + \bar{y}_j + \bar{y}_k = \bar{u}_i + \bar{u}_j + \bar{u}_k$ ,  $\bar{y}_i\bar{y}_j + \bar{y}_j\bar{y}_k + \bar{y}_k\bar{y}_i = \bar{u}_i\bar{u}_j + \bar{u}_j\bar{u}_k + \bar{u}_k\bar{u}_i$  and  $\bar{y}_i\bar{y}_j\bar{y}_k < \bar{u}_i\bar{u}_j\bar{u}_k$ . Setting  $y_l = u_l$  for every  $l \in \{1, \dots, n\} \setminus \{i, j, k\}$  and  $y_l = \bar{y}_l(u_i + u_j + u_k)$  for every  $l \in \{i, j, k\}$  finally yields  $e_1(\mathbf{y}) = e_1(\mathbf{u})$ ,  $e_2(\mathbf{y}) = e_2(\mathbf{u})$  and  $e_3(\mathbf{y}) < e_3(\mathbf{u})$ , which is a contradiction.

**COROLLARY 4.3.** *Suppose that  $n \geq 3$  and that  $\boldsymbol{\pi} = (n, n-1, \dots, 2, 1)$ . Then  $b_{\boldsymbol{\pi}}(\mathbf{u}) \geq \vartheta(a_{\boldsymbol{\pi}}(\mathbf{u}))$  holds for every  $\mathbf{u} \in \Delta_n$ .*

*Preliminary Step 2:* We investigate how, for fixed  $\boldsymbol{\pi} \in \sigma_n$ , the quantities  $a_{\boldsymbol{\pi}}(\mathbf{u})$  and  $b_{\boldsymbol{\pi}}(\mathbf{u})$  change if  $\mathbf{u} \in \Delta_n$  changes. To do so, temporarily extend  $a_{\boldsymbol{\pi}}$  and  $b_{\boldsymbol{\pi}}$  to the full  $\mathbb{R}^n$  using the identities in Lemma 3.2. The following lemmata (whose proof is given in the Appendix) will be crucial in the sequel.

**LEMMA 4.4.** *Suppose that  $n \geq 3$  and that  $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n) \in \mathbb{R}^n$  fulfils  $\sum_i \delta_i = 0$ . Then for every  $t \in \mathbb{R}$  the following identities hold:*

$$a_{\boldsymbol{\pi}}(\mathbf{u} + t\boldsymbol{\delta}) - a_{\boldsymbol{\pi}}(\mathbf{u}) = \alpha_1 t + \alpha_2 t^2 \quad (24)$$

$$b_{\boldsymbol{\pi}}(\mathbf{u} + t\boldsymbol{\delta}) - b_{\boldsymbol{\pi}}(\mathbf{u}) = \beta_1 t + \beta_2 t^2 + \beta_3 t^3 \quad (25)$$

where

$$\begin{aligned} \alpha_1 &= \sum_i a_i \delta_i, & \alpha_2 &= \sum_{i < j, \{i, j\} \in I_{\boldsymbol{\pi}}} \delta_i \delta_j, \\ \beta_1 &= \sum_i b_i \delta_i, & \beta_2 &= \sum_{i < j} c_{i, j} \delta_i \delta_j, & \beta_3 &= \sum_{i < j < k, \{i, j, k\} \in Q_{\boldsymbol{\pi}}} \delta_i \delta_j \delta_k, \end{aligned}$$

and

$$a_i = \sum_{j: \{i, j\} \in I_{\boldsymbol{\pi}}} u_j, \quad b_i = \sum_{j < k, \{i, j, k\} \in Q_{\boldsymbol{\pi}}} u_j u_k, \quad c_{i, j} = c_{j, i} = \sum_{k: \{i, j, k\} \in Q_{\boldsymbol{\pi}}} u_k.$$

**LEMMA 4.5.** *Suppose that  $n \geq 3$  and that  $\boldsymbol{\pi} \in \sigma_n$ . If  $p, q, r \in \{1, \dots, n\}$  are distinct elements such that  $\{p, q, r\} \notin Q_{\boldsymbol{\pi}}$ , then  $c_{p, r} + c_{q, r} \geq c_{p, q} \geq 0$ .*

We now state two conditions for  $\boldsymbol{\pi}$  that imply the existence of a direction  $\boldsymbol{\delta} \in \mathbb{R}^n \setminus \{0\}$  with  $\sum_i \delta_i = 0$  such that  $t \mapsto a_{\boldsymbol{\pi}}(\mathbf{u} + t\boldsymbol{\delta}) - a_{\boldsymbol{\pi}}(\mathbf{u})$  is identical to zero for every  $t$  and  $t \mapsto b_{\boldsymbol{\pi}}(\mathbf{u} + t\boldsymbol{\delta}) - b_{\boldsymbol{\pi}}(\mathbf{u})$  is of degree two and concave.

LEMMA 4.6. *Suppose that  $n \geq 3$ , that  $\pi \in \sigma_n$ , and that one of the following two conditions holds:*

- (i) *There exist  $p, q, r \in \{1, 2, \dots, n\}$  with  $p < q < r$  and  $\pi(r) > \pi(q) > \pi(p)$ .*
- (ii) *There exist  $p, q, r, s \in \{1, 2, \dots, n\}$  with  $p < q < r < s$  and  $\pi(q) > \pi(p) > \pi(s) > \pi(r)$ .*

*Then there exists  $\delta \in \mathbb{R}^n \setminus \{0\}$  such that the coefficients in (24) and (25) fulfil  $\alpha_1 = \alpha_2 = \beta_3 = 0$  and  $\beta_2 \leq 0$ .*

*Induction Step 1:* We prove the induction step for every  $\pi \in \sigma_n$  fulfilling one of the conditions in Lemma 4.6.

LEMMA 4.7. *Suppose that  $n \geq 3$  and that  $b_\omega(\mathbf{v}) \geq \vartheta(a_\omega(\mathbf{v}))$  holds for all  $(\omega, \mathbf{v}) \in \sigma_m \times \Delta_m$  with  $m < n$ . If  $\pi \in \sigma_n$  fulfils one of the conditions in Lemma 4.6 then  $b_\pi(\mathbf{u}) \geq \vartheta(a_\pi(\mathbf{u}))$  for every  $\mathbf{u} \in \Delta_n$ .*

PROOF. Suppose that  $\pi \in \sigma_n$  fulfils one of the conditions in Lemma 4.6 and consider  $\mathbf{u} \in \Delta_n$ . If  $u_k = 0$  for some  $k \in \{1, \dots, n\}$  then, defining  $(\pi', \mathbf{v}) \in \sigma_{n-1} \times \Delta_{n-1}$  by  $v_i = u_i$  for  $i < k$  and  $v_i = u_{i+1}$  for  $i \geq k$  as well as

$$\pi'(i) = \begin{cases} \pi(i) & \text{if } i < k \text{ and } \pi(i) < \pi(k) \\ \pi(i) - 1 & \text{if } i < k \text{ and } \pi(i) > \pi(k), \\ \pi(i+1) & \text{if } i \geq k \text{ and } \pi(i+1) < \pi(k) \\ \pi(i+1) - 1 & \text{if } i \geq k \text{ and } \pi(i+1) > \pi(k), \end{cases}$$

we immediately get  $b_\pi(\mathbf{u}) = b_{\pi'}(\mathbf{v}) \geq \vartheta(a_{\pi'}(\mathbf{v})) = \vartheta(a_\pi(\mathbf{u}))$ .

Suppose now that  $\mathbf{u} \in (0, 1)^n$  and, using Lemma 4.6, choose  $\delta \in \mathbb{R}^n \setminus \{0\}$  such that  $\beta_2 \leq 0$  and  $a_\pi(\mathbf{u} + t\delta) = a_\pi(\mathbf{u})$  and  $b_\pi(\mathbf{u} + t\delta) - b_\pi(\mathbf{u}) = \beta_1 t + \beta_2 t^2$  for all  $t \in \mathbb{R}$ . Considering  $\mathbf{u} \in (0, 1)^n$  there are  $t_0 < 0 < t_1$  such that  $\mathbf{u} + t\delta \in [0, 1]^n$  if and only if  $t \in [t_0, t_1]$ . Concavity of  $t \mapsto b_\pi(\mathbf{u} + t\delta)$  implies that  $b_\pi(\mathbf{u} + t_0\delta) \leq b_\pi(\mathbf{u})$  or  $b_\pi(\mathbf{u} + t_1\delta) \leq b_\pi(\mathbf{u})$ . Moreover there are  $i, j$  such that  $(\mathbf{u} + t_0\delta)_i = 0$  and  $(\mathbf{u} + t_1\delta)_j = 0$  by construction, so we can proceed as in the first step of the proof and use induction to get  $b_\pi(\mathbf{u}) \geq \vartheta(a_\pi(\mathbf{u}))$ .

*Induction Step 2:* As a final step we concentrate on permutations  $\pi \in \sigma_n$  not fulfilling any of the two conditions in 4.6 and start with the following definition and the subsequent lemma (whose proof can be found in the Appendix).

DEFINITION 4.8. *A permutation  $\pi \in \sigma_l$  is called almost decreasing if there is at most one  $i \in \{1, \dots, l-1\}$  so that  $\pi(i) < \pi(i+1)$ .*

LEMMA 4.9. *Let  $l \geq 1$  and  $\pi \in \sigma_l$ . Then the following two conditions are equivalent:*

- *There are no  $1 \leq p < q < r \leq l$  so that  $\pi(p) < \pi(q) < \pi(r)$ , and there are no  $1 \leq p < q < r < s \leq l$  so that  $\pi(r) < \pi(s) < \pi(p) < \pi(q)$ .*
- *$\pi$  or  $\pi^{-1}$  is almost decreasing.*

Having this characterization we can now prove the remaining induction step for those  $\pi \in \sigma_n$  fulfilling that  $\pi$  or  $\pi^{-1}$  is almost decreasing. Notice that w.l.o.g. we may assume that  $\pi \in \sigma_n$  is almost decreasing since defining  $\mathbf{v} \in \Delta_n$  by  $v_i = u_{\pi^{-1}(i)}$  for every  $i \in \{1, \dots, n\}$  yields  $a_\pi(\mathbf{u}) = a_{\pi^{-1}}(\mathbf{v})$  as well as  $b_\pi(\mathbf{u}) = b_{\pi^{-1}}(\mathbf{v})$ . Both subsequent lemmata are therefore only stated and proved for almost decreasing  $\pi$ .

LEMMA 4.10. *Suppose that  $n \geq 3$  and that  $b_\omega(\mathbf{v}) \geq \vartheta(a_\omega(\mathbf{v}))$  holds for all  $(\omega, \mathbf{v}) \in \sigma_m \times \Delta_m$  with  $m < n$ . If  $\pi \in \sigma_n$  is almost decreasing with  $\pi(1) = n$  or  $\pi(n) = 1$  then  $b_\pi(\mathbf{u}) \geq \vartheta(a_\pi(\mathbf{u}))$  holds for every  $\mathbf{u} \in \Delta_n$ .*

PROOF. As before we may assume  $\mathbf{u} \in (0, 1)^n$ . Suppose that  $\pi(1) = n$ . Defining  $(\pi', \mathbf{u}') \in \sigma_{n-1} \times \Delta_{n-1}$  by  $\pi'(i) = \pi(i+1)$  and  $u'_i = \frac{u_{i+1}}{1-u_1}$  for every  $i \in \{1, \dots, n-1\}$  and considering

$$a_{\pi'}(\mathbf{u}') = \frac{1}{(1-u_1)^2} \sum_{2 \leq i < j \leq n: \{i,j\} \in I_\pi} u_i u_j$$

yields that  $a_\pi(\mathbf{u}) = (1-u_1)^2 a_{\pi'}(\mathbf{u}') + u_1(1-u_1)$ . Analogously, using

$$b_{\pi'}(\mathbf{u}') = \frac{1}{(1-u_1)^3} \sum_{2 \leq i < j < k \leq n: \{i,j,k\} \in Q_\pi} u_i u_j u_k$$

we get  $b_\pi(\mathbf{u}) = (1-u_1)^3 b_{\pi'}(\mathbf{u}') + u_1(1-u_1)^2 a_{\pi'}(\mathbf{u}')$ . To simplify notation let  $\tilde{\pi}_k$  denote the decreasing permutation in  $\sigma_k$  for every  $k \in \mathbb{N}$ . Choose  $u''_1, \dots, u''_{n-1} \in \Delta_{n-1}$  such that  $a_{\tilde{\pi}_{n-1}}(\mathbf{u}'') = a_{\pi'}(\mathbf{u}')$  and  $b_{\tilde{\pi}_{n-1}}(\mathbf{u}'') = \vartheta(a_{\pi'}(\mathbf{u}'))$ . Define  $\tilde{\mathbf{u}} = (\tilde{u}_1, \dots, \tilde{u}_n)$  by  $\tilde{u}_1 = u_1$  and  $\tilde{u}_i = (1-u_1)u''_{i-1}$  for every  $i \in \{2, \dots, n\}$ . Then  $\sum_{i=1}^n \tilde{u}_i = u_1 + (1-u_1) \sum_{i=1}^{n-1} u''_i = 1$  and we get

$$a_{\tilde{\pi}_n}(\tilde{\mathbf{u}}) = (1-u_1)^2 a_{\tilde{\pi}_{n-1}}(\mathbf{u}'') + u_1(1-u_1) = a_\pi(\mathbf{u})$$

as well as

$$\begin{aligned} b_{\tilde{\pi}_n}(\tilde{\mathbf{u}}) &= \sum_{1 < i < j < k: \{i,j,k\} \in Q_{\tilde{\pi}_n}} \tilde{u}_i \tilde{u}_j \tilde{u}_k + \tilde{u}_1 \sum_{1 < j < k: \{j,k\} \in I_{\tilde{\pi}_n}} \tilde{u}_i \tilde{u}_j \\ &= (1-u_1)^3 b_{\tilde{\pi}_{n-1}}(\mathbf{u}'') + u_1(1-u_1)^2 a_{\tilde{\pi}_{n-1}}(\mathbf{u}''). \end{aligned}$$

Altogether this yields

$$\begin{aligned}
 b_{\pi}(\mathbf{u}) &= (1 - u_1)^3 b_{\pi'}(\mathbf{u}') + u_1(1 - u_1)^2 a_{\pi'}(\mathbf{u}') \\
 &\geq (1 - u_1)^3 \vartheta(a_{\pi'}(\mathbf{u}')) + u_1(1 - u_1)^2 a_{\pi'}(\mathbf{u}') \\
 &= (1 - u_1)^3 b_{\tilde{\pi}_{n-1}}(\mathbf{u}'') + u_1(1 - u_1)^2 a_{\tilde{\pi}_{n-1}}(\mathbf{u}'') = b_{\tilde{\pi}_n}(\tilde{\mathbf{u}}) \\
 &\geq \vartheta(a_{\tilde{\pi}_n}(\tilde{\mathbf{u}})) = \vartheta(a_{\pi}(\mathbf{u})).
 \end{aligned}$$

The proof of the case  $\pi(n) = 1$  is completely analogous.

The following final lemma assures that in case of almost decreasing  $\pi \in \sigma_n$  with  $\pi(1) \neq n$  and  $\pi(n) \neq 1$  we cannot be on the boundary of  $\Omega_{\Phi}$ . Note that in the proof we do not make use of the induction hypothesis.

LEMMA 4.11. *Suppose that  $n \geq 3$  and that  $\pi \in \sigma_n$  is almost decreasing with  $\pi(1) \neq n$  and  $\pi(n) \neq 1$ . Then for every  $\mathbf{u} \in \Delta_n \cap (0, 1)^n$  we have*

$$b_{\pi}(\mathbf{u}) - \vartheta(a_{\pi}(\mathbf{u})) > \min \left\{ b_{\omega}(\mathbf{v}) - \vartheta(a_{\omega}(\boldsymbol{\pi})) : \omega \in \sigma_n, \mathbf{v} \in \Delta_n \right\}$$

PROOF. First note that the existence of the minimum is assured by the fact that  $\sigma_n$  is finite and  $\Delta_n$  is compact. Set  $k := \pi^{-1}(1)$ . Then  $1 = \pi(k) < \pi(k-1) < \dots < \pi(1) < n$  and  $1 < \pi(n) < \dots < \pi(k+2) < \pi(k+1)$ , so  $\pi(k+1) = n$ . Define  $(\pi', \mathbf{u}') \in \sigma_n \times \Delta_n$  as follows:  $\pi'(i) = \pi(i)$  for  $i \notin \{k, k+1\}$ ,  $\pi'(k) = \pi(k+1) = n$  and  $\pi'(k+1) = \pi(k) = 1$ ;  $u'_i = u_i$  for  $i \notin \{k, k+1\}$ ,  $u'_k = u_{k+1}$  and  $u'_{k+1} = u_k$ . Then it is straightforward to verify that  $a_{\pi'}(\mathbf{u}') - a_{\pi}(\mathbf{u}) = u_k u_{k+1}$  and

$$b_{\pi'}(\mathbf{u}') - b_{\pi}(\mathbf{u}) = - \sum_{i \neq k, k+1} u_k u_{k+1} u_i,$$

holds, which, considering that  $n \geq 3$  implies  $a_{\pi'}(\mathbf{u}') > a_{\pi}(\mathbf{u})$  and  $b_{\pi'}(\mathbf{u}') < b_{\pi}(\mathbf{u})$ . Having this we get  $b_{\pi}(\mathbf{u}) - \vartheta(a_{\pi}(\mathbf{u})) > b_{\pi'}(\mathbf{u}') - \vartheta(a_{\pi'}(\mathbf{u}'))$  since  $\vartheta$  is non-decreasing, which completes the proof.

Since Lemma 4.11 implies that in order to prove inequality (23) for every  $\pi \in \sigma_n$  and  $\mathbf{u} \in \Delta_n$  it is not necessary to consider almost decreasing permutations  $\pi$  with  $\pi(1) \neq n$  and  $\pi(n) \neq 1$  the proof of Theorem 4.1 (hence the one of Theorem 3.6) is complete.

## 5. Additional related results

So far we have shown that  $\Omega \subseteq \Omega_{\Phi}$ . We now prove that the two sets are in fact identical.

THEOREM 5.1. *The precise  $\tau$ - $\rho$  region  $\Omega$  coincides with  $\Omega_{\Phi}$ .  $\Omega$  is not convex.*

PROOF. The construction of  $\Phi$  implies the existence of a family  $(A_t)_{t \in [0,1]}$  of shuffles of  $M$  fulfilling that the map  $t \mapsto A_t$  is continuous on  $[0, 1]$  (w.r.t.  $d_\infty$ ) and that

$$\gamma(t) := (\tau(A_t), \rho(A_t)) = \begin{cases} (4t - 1, \Phi(4t - 1)) & \text{if } t \in [0, \frac{1}{2}] \\ (3 - 4t, -\Phi(4t - 3)) & \text{if } t \in [\frac{1}{2}, 1]. \end{cases} \quad (26)$$

Obviously the curve  $\gamma : [0, 1] \rightarrow [-1, 1]^2$  is simply closed and rectifiable. For every  $s \in [0, 1]$  consider the similarities  $f_s, g_s : [0, 1]^2 \rightarrow [0, 1]^2$ , given by  $f_s(x, y) = s(x, y)$  and  $g_s(x, y) = (1 - s)(x, y) + (s, s)$  and define the operator  $O_s : \mathcal{C} \rightarrow \mathcal{C}$  implicitly via

$$\mu_{O_s(A)} = s\mu_M^{f_s} + (1 - s)\mu_A^{g_s}.$$

Notice that  $O_s(A)$  is usually referred to as the ordinal sum of  $M, A$  with respect to the partition  $[0, s), (s, 1]$ , see Durante and Sempi (2015). Then we have  $d_\infty(O_s(A), O_s(B)) \leq d_\infty(A, B)$  for all  $A, B \in \mathcal{C}$  and every  $s \in [0, 1]$ , and the mapping  $s \mapsto O_s(A)$  is continuous for every  $A \in \mathcal{C}$ . Consequently, the function  $H : [0, 1]^2 \rightarrow [-1, 1]^2$ , given by

$$H(s, t) = (\tau(O_s(A_t)), \rho(O_s(A_t)))$$

is continuous and fulfils, firstly, that  $H(0, t) = \gamma(t)$  and  $H(1, t) = (1, 1)$  for every  $t \in [0, 1]$  and, secondly, that  $H(s, 0) = H(s, 1)$  for all  $s \in [0, 1]$ . In other words,  $H$  is a homotopy and  $\gamma$  is homotopic to the point  $(1, 1)$  (see Figure 3), implying  $\Omega = \Omega_\Phi$ . Since  $\Phi$  is strictly concave on each interval  $[\frac{2-n}{n}, \frac{2-(n-1)}{n-1}]$  with  $n \geq 3$ ,  $\Omega = \Omega_\Phi$  cannot be convex.

Considering that the operator  $O_s : \mathcal{C} \rightarrow \mathcal{C}$  maps the family of all shuffles of  $M$  into itself for every  $s \in [0, 1]$  the proof of Theorem 5.1 has the following surprising byproduct:

**COROLLARY 5.2.** *For every point  $(x, y) \in \Omega$  there is a shuffle  $h \in \mathcal{S}$  such that we have  $(\tau(A_h), \rho(A_h)) = (x, y)$ .*

Additionally, Theorem 5.1 also implies the following result concerning the possible range of Spearman's  $\rho$  if Kendall's  $\tau$  is known (and vice versa):

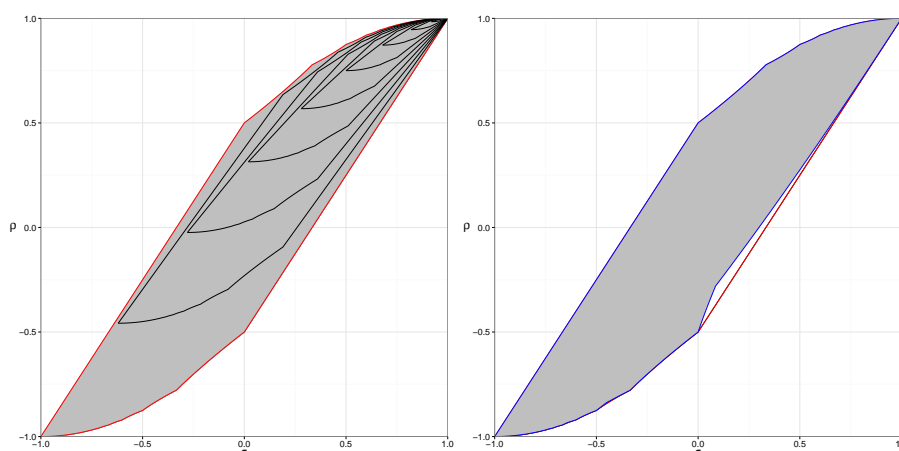
**COROLLARY 5.3.** *Suppose that  $X, Y$  are continuous random variables with  $\tau(X, Y) = \tau_0$ . Then  $\rho(X, Y) \in [\Phi(\tau_0), -\Phi(-\tau_0)]$ .*

**REMARK 5.4.** Due to the simple analytic form of  $\Phi$  it is straightforward to verify that

$$\lambda_2(\Omega) = \frac{4}{5} - \frac{4}{5} \zeta(3) + \frac{2}{15} \pi^2 \approx 1.1543,$$

whereby  $\zeta(3) = \sum_{i=1}^{\infty} \frac{1}{i^3}$ . Considering that  $\lambda_2(\Omega_0) = \frac{7}{6} \approx 1.1667$  this underlines the quality of the classical inequalities.





**Fig. 3.** The curves  $\gamma_s(t) = H(s, t)$  for  $t \in [0, 1]$  and  $s \in \{\frac{1}{10}, \dots, \frac{9}{10}\}$  with  $H$  being the homotopy used in the proof of Theorem 5.1 (left panel); the region corresponding to conjecture **(C1)** for the class of exchangeable copulas (right panel). The red lines depict the boundary of  $\Omega$ .

## 6. Conclusions and future work

Although Kendall's  $\tau$  and Spearman's  $\rho$  are both measures of concordance, they quantify different aspects of the dependence structure (Fredricks and Nelsen, 2007). The results established in this paper do not only answer the open question about how much they can differ in a definitive manner, they also show for which dependence structures the discrepancy is maximal and, more surprisingly, that mutually completely dependent random variables cover all possible joint values of  $\tau$  and  $\rho$ . Although mutual complete dependence might be considered a highly atypical dependence structure, it naturally appears in various practical settings, in particular in optimization problems like for instance in worst-case value-at-risk scenarios (Makarov, 1981).

To the best of the authors' knowledge the combination of combinatorial and continuity arguments used in the proofs is novel. The authors conjecture that this combination may also prove useful in various other problems - in particular concerning open questions in complete mixability, where many important recent results are based on combinatorial arguments and discretizations (Wang, 2015).

Having at hand a full characterization of  $\Omega$  the question naturally arises, how much  $\tau(A)$  and  $\rho(A)$  may differ if it is known that  $A \in \mathcal{C}$  is an element of a given subclass of copulas. The authors are convinced that the problem of determining the exact  $\tau$ - $\rho$  region for subclasses may in some cases be even more difficult than answering the sixty year old question about  $\Omega$  has been, in particular for classes where there are no simple explicit

formulas for  $\tau$  or  $\rho$ . Based on numerous simulations and analytic calculations the authors conjecture that the following inequalities hold for the class of exchangeable copulas  $\mathcal{C}_{ex}$ , the class of extreme-value-copulas  $\mathcal{C}_{ev}$  and the class of Archimedean copulas  $\mathcal{C}_{ar}$ :

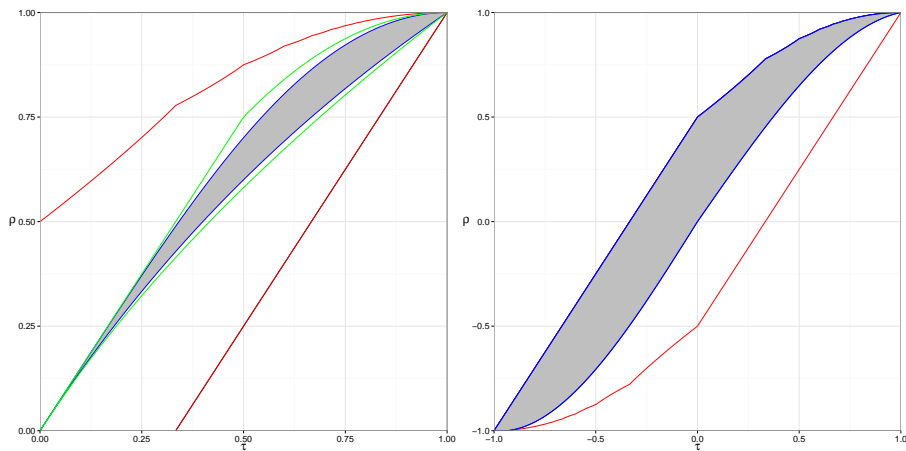
**(C1)** Conjectured (sharp) inequality for  $A \in \mathcal{C}_{ex}$ :  $L_{ex}(\tau) \leq \rho \leq -\Phi(-\tau)$ , where

$$L_{ex}(\tau) = \begin{cases} \Phi(\tau) & \text{if } \tau \in \left[\frac{2-n}{n}, \frac{2-(n-1)}{n-1}\right], n \geq 3 \text{ odd,} \\ \min \left\{ \frac{3\tau}{1+n} - \frac{2+(n-1)n}{(1+n)^2} - \frac{(n-3)(-1+n+(1+n)\tau)^{3/2}}{2(n+1)^2\sqrt{n-1}}, \right. \\ \quad \left. \frac{3\tau}{n} - \frac{(n-2)^2+n}{n^2} - \frac{(n-4)(-2+n+n\tau)^{3/2}}{2n^2\sqrt{n-2}} \right\}, & \text{if } \tau \in \left[\frac{2-n}{n}, \frac{2-(n-1)}{n-1}\right], n \geq 3 \text{ even,} \\ \frac{-1+6\tau-3\tau^{3/2}}{2}, & \text{if } \tau \in \left[0, \frac{3-2\sqrt{2}}{2}\right], \\ -\frac{4}{9} + \tau + \frac{(1+3\tau)^{3/2}}{18}, & \text{if } \tau \in \left[\frac{3-2\sqrt{2}}{2}, 1\right]. \end{cases}$$

**(C2)** Conjectured inequality for  $A \in \mathcal{C}_{ev}$ :  $L_{ev}(\tau) \leq \rho \leq R_{ev}(\tau)$ , where  $L_{ev}(\tau) = \frac{3\tau}{2+\tau}$  and  $R_{ev}(\tau) = \frac{3\tau}{2+\tau^3} - \frac{1}{3}(1-\tau)^2\tau^4$

**(C3)** Conjectured inequality for  $A \in \mathcal{C}_{ar}$ :  $L_{ar}(\tau) \leq \rho < -\Phi(-\tau)$ , where  $L_{ar}(\tau) = \frac{7\tau-2\tau^3}{5}\mathbf{1}_{[0,1]}(\tau) + \frac{31\tau-11\tau^3}{20}\mathbf{1}_{[-1,0)}(\tau)$

Notice that the set determined by the inequalities in (C2) is strictly contained in the region determined by the famous Hutchinson-Lai inequalities (Hürlimann, 2003). Furthermore we remark that we were able to prove that for each  $x \in [-1, 1]$  there exists an exchangeable copula  $A \in \mathcal{C}_{ex}$  with  $(\tau(A), \rho(A)) = (x, L_{ex}(x))$  and a sequence  $(A_n)_{n \in \mathbb{N}}$  of Archimedean copulas such that  $(\tau(A_n), \rho(A_n)) \rightarrow (x, -\Phi(-x))$  for  $n \rightarrow \infty$ . The right-hand side of Figure 3 and Figure 4 depict the corresponding regions.



**Fig. 4.** The region corresponding to conjecture **(C2)** for the class of extreme-value copulas (left panel) and the one corresponding to conjecture **(C3)** for the class of Archimedean copulas (right panel); the green lines depict the Hutchinson-Lai inequalities, the red ones the boundary of  $\Omega$ .

## 7. Appendix

PROOF (OF LEMMA 3.2). Since it is straightforward to verify the first inequality we start with the proof of the second one. Using  $s_j^2 - s_{j-1}^2 = (s_j - s_{j-1})(s_j + s_{j-1}) = u_j(2 \sum_{k < j} u_k + u_j)$  we get

$$\begin{aligned}
 \text{invsum}(h_{\pi, \mathbf{u}}) &= \sum_{j=1}^n \int_{s_{j-1}}^{s_j} \int_0^1 \mathbf{1}_{[0, y)}(x) \mathbf{1}_{(h_{\pi, \mathbf{u}}(y), 1]}(h_{\pi, \mathbf{u}}(x)) (y - x) d\lambda(x) d\lambda(y) \\
 &= \sum_{j=1}^n \int_{s_{j-1}}^{s_j} \sum_{\substack{i: i < j, \\ \pi(i) > \pi(j)}} \int_{s_{i-1}}^{s_i} (y - x) d\lambda(x) d\lambda(y) \\
 &= \sum_{j=1}^n \sum_{\substack{i: i < j, \\ \pi(i) > \pi(j)}} \int_{s_{j-1}}^{s_j} \left( y u_i - \frac{1}{2} u_i^2 - u_i \sum_{k: k < i} u_k \right) d\lambda(y) \\
 &= \sum_{j=1}^n \sum_{\substack{i: i < j, \\ \pi(i) > \pi(j)}} \left( \frac{1}{2} u_i u_j \left( 2 \sum_{l: l < j} u_l + u_j \right) - \frac{1}{2} u_i^2 u_j - u_i u_j \sum_{k: k < i} u_k \right) \\
 &= \sum_{\substack{i < j: \\ \pi(i) > \pi(j)}} \left( u_i u_j \sum_{l: l < j} u_l + \frac{1}{2} u_i u_j^2 - \frac{1}{2} u_i^2 u_j - u_i u_j \sum_{k: k < i} u_k \right) \\
 &= \sum_{\substack{i < j: \\ \pi(i) > \pi(j)}} \left( \frac{1}{2} u_i u_j^2 + \frac{1}{2} u_i^2 u_j + \sum_{k: i < k < j} u_i u_j u_k \right).
 \end{aligned}$$

The third equality follows from

$$\begin{aligned}
 b_{\pi}(\mathbf{u}) &= \text{inv}(h_{\pi, \mathbf{u}}) - 2 \text{invsum}(h_{\pi, \mathbf{u}}) \\
 &= \sum_{k < i < j, \pi(i) > \pi(j)} u_i u_j u_k + \sum_{i < k < j, \pi(i) > \pi(j)} u_i u_j u_k + \sum_{i < j < k, \pi(i) > \pi(j)} u_i u_j u_k \\
 &\quad + \sum_{i < j, \pi(i) > \pi(j)} u_i^2 u_j + \sum_{i < j, \pi(i) > \pi(j)} u_i u_j^2 \\
 &\quad - \sum_{i < j, \pi(i) > \pi(j)} u_i^2 u_j - \sum_{i < j, \pi(i) > \pi(j)} u_i u_j^2 - 2 \sum_{i < k < j, \pi(i) > \pi(j)} u_i u_j u_k \\
 &= \sum_{i < j < k, \pi(j) > \pi(k)} u_i u_j u_k + \sum_{i < j < k, \pi(i) > \pi(j)} u_i u_j u_k - \sum_{i < j < k, \pi(i) > \pi(k)} u_i u_j u_k \\
 &= \sum_{\substack{i < j < k, \\ \pi(i) > \pi(j) > \pi(k)}} u_i u_j u_k + \sum_{\substack{i < j < k, \\ \pi(j) > \pi(i) > \pi(k)}} u_i u_j u_k + \sum_{\substack{i < j < k, \\ \pi(j) > \pi(k) > \pi(i)}} u_i u_j u_k \\
 &\quad + \sum_{\substack{i < j < k, \\ \pi(k) > \pi(i) > \pi(j)}} u_i u_j u_k + \sum_{\substack{i < j < k, \\ \pi(i) > \pi(k) > \pi(j)}} u_i u_j u_k + \sum_{\substack{i < j < k, \\ \pi(i) > \pi(j) > \pi(k)}} u_i u_j u_k \\
 &\quad - \sum_{\substack{i < j < k, \\ \pi(j) > \pi(i) > \pi(k)}} u_i u_j u_k - \sum_{\substack{i < j < k, \\ \pi(i) > \pi(j) > \pi(k)}} u_i u_j u_k - \sum_{\substack{i < j < k, \\ \pi(i) > \pi(k) > \pi(j)}} u_i u_j u_k \\
 &= \sum_{\substack{i < j < k, \\ \pi(i) > \pi(j) > \pi(k)}} u_i u_j u_k + \sum_{\substack{i < j < k, \\ \pi(j) > \pi(k) > \pi(i)}} u_i u_j u_k + \sum_{\substack{i < j < k, \\ \pi(k) > \pi(i) > \pi(j)}} u_i u_j u_k.
 \end{aligned}$$

PROOF (OF LEMMA 3.5). To simplify calculations let  $e_i(\mathbf{v})$  denote the  $i$ -th elementary symmetric polynomial for  $i \in \{1, 2, 3\}$  and  $\mathbf{v} \in \mathbb{R}^n$ , i.e.  $e_1(\mathbf{v}) = \sum_i v_i$ ,  $e_2(\mathbf{v}) = \sum_{i < j} v_i v_j$  and  $e_3(\mathbf{v}) = \sum_{i < j < k} v_i v_j v_k$ . Using  $\sum_i v_i^2 = e_1(\mathbf{v})^2 - 2e_2(\mathbf{v})$  it follows that

$$\begin{aligned} \text{inv}(h_{\pi, \mathbf{u}}) &= e_2(\mathbf{u}) = \frac{1}{2} \left( e_1(\mathbf{u})^2 - \sum_i u_i^2 \right) = \frac{1}{2} \left( 1 - \sum_i u_i^2 \right) \\ &= \frac{1}{2} \left( 1 - (n-1)r^2 - (1 - (n-1)r)^2 \right) = r(n-1) - \frac{1}{2}r^2n(n-1). \end{aligned}$$

Moreover, considering  $r \in [\frac{1}{n}, \frac{1}{n-1}]$  we get  $\text{inv}(h_{\pi, \mathbf{u}}) \in [\frac{1}{2} - \frac{1}{2(n-1)}, \frac{1}{2} - \frac{1}{2n}]$ , implying  $\tau(A_{h_{\pi, \mathbf{u}}}) \in [-1 + \frac{2}{n}, -1 + \frac{2}{n-1}]$ , which completes the proof of the first assertion. Using  $\sum_i v_i^3 = e_1(\mathbf{v})^3 - 3e_1(\mathbf{v})e_2(\mathbf{v}) + 3e_3(\mathbf{v})$  and Lemma 3.2 moreover it follows that

$$\begin{aligned} \text{invsum}(h_{\pi, \mathbf{u}}) &= \frac{1}{2} \text{inv}(h_{\pi, \mathbf{u}}) - \frac{1}{2} b_{\pi}(h) = \frac{1}{2} e_2(\mathbf{u}) - \frac{1}{2} e_3(\mathbf{u}) \\ &= \frac{1}{2} e_2(\mathbf{u}) - \frac{1}{2} \left( \frac{1}{3} \sum_i u_i^3 - \frac{1}{3} e_1(\mathbf{u})^3 + e_1(\mathbf{u})e_2(\mathbf{u}) \right) \\ &= \frac{1}{2} e_2(\mathbf{u}) - \frac{1}{6} \sum_i u_i^3 + \frac{1}{6} - \frac{1}{2} e_2(\mathbf{u}) = \frac{1}{6} - \frac{1}{6} \sum_i u_i^3 \\ &= \frac{1}{6} - \frac{1}{6} \left( (n-1)r^3 + (1 - (n-1)r)^3 \right). \end{aligned}$$

Again considering  $r \in [\frac{1}{n}, \frac{1}{n-1}]$  we get  $\text{invsum}(h_{\pi, \mathbf{u}}) \in [\frac{1}{6} - \frac{1}{6(n-1)^2}, \frac{1}{6} - \frac{1}{6n^2}]$ , implying  $\rho(A_{h_{\pi, \mathbf{u}}}) \in [-1 + \frac{2}{n^2}, -1 + \frac{2}{(n-1)^2}]$ , which completes the proof.

PROOF (OF LEMMA 4.4). The expression for  $a_{\pi}(\mathbf{u} + t\boldsymbol{\delta}) - a_{\pi}(\mathbf{u})$  is easily verified:

$$\begin{aligned} a_{\pi}(\mathbf{u} + t\boldsymbol{\delta}) - a_{\pi}(\mathbf{u}) &= \sum_{i < j, \{i, j\} \in I_{\pi}} ((u_i + \delta_i t)(u_j + \delta_j t) - u_i u_j) \\ &= t \sum_{i < j, \{i, j\} \in I_{\pi}} \delta_j u_i + \delta_i u_j + t^2 \sum_{i < j, \{i, j\} \in I_{\pi}} \delta_i \delta_j \\ &= t \left( \sum_{j < i, \{i, j\} \in I_{\pi}} \delta_i u_j + \sum_{i < j, \{i, j\} \in I_{\pi}} \delta_i u_j \right) + t^2 \alpha_2 \\ &= t \sum_{i=1}^n \delta_i \sum_{j: \{i, j\} \in I_{\pi}} u_j + t^2 \alpha_2 = t \sum_{i=1}^n \delta_i a_i + t^2 \alpha_2 = \alpha_1 t + \alpha_2 t^2 \end{aligned}$$

To derive the expression for  $b_{\pi}(\mathbf{u} + t\boldsymbol{\delta}) - b_{\pi}(\mathbf{u})$  notice that

$$\begin{aligned} b_{\pi}(\mathbf{u} + t\boldsymbol{\delta}) - b_{\pi}(\mathbf{u}) &= \sum_{i < j < k, \{i, j, k\} \in Q_{\pi}} ((u_i + \delta_i t)(u_j + \delta_j t)(u_k + \delta_k t) - u_i u_j u_k) \\ &= \sum_{i < j < k, \{i, j, k\} \in Q_{\pi}} \delta_i \delta_j \delta_k t^3 + \delta_i \delta_j u_k t^2 + \delta_i \delta_k u_j t^2 \\ &\quad + \delta_j \delta_k u_i t^2 + \delta_k u_i u_j t + \delta_j u_i u_k t + \delta_i u_j u_k t \end{aligned}$$

$$\begin{aligned}
 &= t^3 \sum_{i < j < k, \{i,j,k\} \in Q_\pi} \delta_i \delta_j \delta_k \\
 &\quad + t^2 \sum_{i < j < k, \{i,j,k\} \in Q_\pi} \delta_i \delta_j u_k + \delta_i \delta_k u_j + \delta_j \delta_k u_i \\
 &\quad + t \sum_{i < j < k, \{i,j,k\} \in Q_\pi} \delta_k u_i u_j + \delta_j u_i u_k + \delta_i u_j u_k = \beta_1 t + \beta_2 t^2 + \beta_3 t^3
 \end{aligned}$$

since

$$\begin{aligned}
 &\sum_{\substack{i < j < k: \\ \{i,j,k\} \in Q_\pi}} \delta_i \delta_j u_k + \delta_i \delta_k u_j + \delta_j \delta_k u_i \\
 &= \sum_{\substack{i < j < k: \\ \{i,k,j\} \in Q_\pi}} \delta_i \delta_j u_k + \sum_{\substack{i < k < j: \\ \{i,k,j\} \in Q_\pi}} \delta_i \delta_j u_k + \sum_{\substack{k < i < j: \\ \{k,i,j\} \in Q_\pi}} \delta_i \delta_j u_k \\
 &= \sum_{i < j} \delta_i \delta_j \left( \sum_{\substack{i < j < k: \\ \{i,k,j\} \in Q_\pi}} u_k + \sum_{\substack{i < k < j: \\ \{i,k,j\} \in Q_\pi}} u_k + \sum_{\substack{k < i < j: \\ \{k,i,j\} \in Q_\pi}} u_k \right) \\
 &= \sum_{i < j} \delta_i \delta_j \sum_{k: \{i,k,j\} \in Q_\pi} u_k = \sum_{i < j} \delta_i \delta_j c_{i,j} = \beta_2
 \end{aligned}$$

and

$$\begin{aligned}
 &\sum_{\substack{i < j < k: \\ \{i,j,k\} \in Q_\pi}} \delta_k u_i u_j + \delta_j u_i u_k + \delta_i u_j u_k \\
 &= \sum_{\substack{i < j < k: \\ \{i,j,k\} \in Q_\pi}} \delta_k u_i u_j + \sum_{\substack{i < j < k: \\ \{i,j,k\} \in Q_\pi}} \delta_j u_i u_k + \sum_{\substack{i < j < k: \\ \{i,j,k\} \in Q_\pi}} \delta_i u_j u_k \\
 &= \sum_{\substack{j < k < i: \\ \{i,j,k\} \in Q_\pi}} \delta_i u_j u_k + \sum_{\substack{j < i < k: \\ \{i,j,k\} \in Q_\pi}} \delta_i u_j u_k + \sum_{\substack{i < j < k: \\ \{i,j,k\} \in Q_\pi}} \delta_i u_j u_k \\
 &= \sum_{i=1}^n \delta_i \left( \sum_{\substack{j < k < i: \\ \{i,j,k\} \in Q_\pi}} u_j u_k + \sum_{\substack{j < i < k: \\ \{i,j,k\} \in Q_\pi}} u_j u_k + \sum_{\substack{i < j < k: \\ \{i,j,k\} \in Q_\pi}} u_j u_k \right) \\
 &= \sum_{i=1}^n \delta_i \sum_{j < k} u_j u_k = \sum_{i=1}^n \delta_i b_i = \beta_1.
 \end{aligned}$$

PROOF (OF LEMMA 4.5). The inequality  $c_{p,q} \geq 0$  immediately follows from  $u_1, \dots, u_n \geq 0$ . Let

$$\iota(i, j) = \iota(j, i) = \begin{cases} 1 & \text{if } (i < j \text{ and } \pi(i) > \pi(j)) \text{ or } (j < i \text{ and } \pi(j) > \pi(i)), \\ -1 & \text{otherwise} \end{cases}$$

and

$$\gamma_{i,j,k} = \frac{1}{2}(1 + \iota(i, j)\iota(i, k)\iota(j, k)) \in \{0, 1\}$$

for every  $i, j, k \in \{1, \dots, n\}$ . Then we have  $\{i, j, k\} \in Q_\pi$  if and only if  $\iota(i, j)\iota(j, k)\iota(i, k) = 1$  if and only if  $\gamma_{i,j,k} = 1$ . Therefore  $c_{i,j} = \sum_k \gamma_{i,j,k} u_k$  for every  $i, j$ . So

$$c_{p,r} + c_{q,r} - c_{p,q} = \sum_i (\gamma_{p,r,i} + \gamma_{q,r,i} - \gamma_{p,q,i}) u_i$$

and it is enough to prove  $\gamma_{p,r,i} + \gamma_{q,r,i} \geq \gamma_{p,q,i}$  for every  $i$ . If  $\gamma_{p,q,i} = 0$  or  $\gamma_{p,r,i} = 1$  or  $\gamma_{q,r,i} = 1$ , then this is clear. So suppose indirectly that  $\gamma_{p,q,i} = 1$  and  $\gamma_{p,r,i} = \gamma_{q,r,i} = 0$ . Then  $\iota(p, q)\iota(p, i)\iota(q, i) = 1$ ,  $\iota(p, r)\iota(p, i)\iota(r, i) = -1$  and  $\iota(q, r)\iota(q, i)\iota(r, i) = -1$ . Multiplying these together, we get that  $\iota(p, q)\iota(p, r)\iota(q, r) = 1$ , so  $\gamma_{p,q,r} = 1$ . However  $\{p, q, r\} \notin Q_\pi$ , so  $\gamma_{p,q,r} = 0$ , contradiction.

PROOF (OF LEMMA 4.6). (i) Suppose that there are  $p < q < r$  such that  $\pi(r) > \pi(q) > \pi(p)$ . Let  $\delta_i = 0$  for every  $i \neq p, q, r$ . We can fix a nonzero solution  $(\delta_p, \delta_q, \delta_r) \in \mathbb{R}^3 \setminus \{0\}$  to the following system of homogeneous linear equations:

$$\delta_p + \delta_q + \delta_r = 0, \quad a_p \delta_p + a_q \delta_q + a_r \delta_r = 0.$$

Then  $\sum_i \delta_i = 0$ ,  $\alpha_1 = 0$ , moreover  $\{p, q\}, \{p, r\}, \{q, r\} \notin I_\pi$  and  $\{p, q, r\} \notin Q_\pi$ , so  $\alpha_2 = \beta_3 = 0$ . The numbers  $\delta_p \delta_q, \delta_p \delta_r, \delta_q \delta_r$  cannot be all negative, so e.g.,  $\delta_p \delta_q \geq 0$ . Then using Lemma 4.5 we obtain

$$\begin{aligned} \beta_2 &= c_{p,q} \delta_p \delta_q + c_{p,r} \delta_p \delta_r + c_{q,r} \delta_q \delta_r \leq (c_{p,r} + c_{q,r}) \delta_p \delta_q + c_{p,r} \delta_p \delta_r + c_{q,r} \delta_q \delta_r \\ &= c_{p,r} \delta_p (\delta_q + \delta_r) + c_{q,r} \delta_q (\delta_p + \delta_r) = -c_{p,r} \delta_p^2 - c_{q,r} \delta_q^2 \leq 0. \end{aligned}$$

Now suppose that there are  $p < q < r < s$  such that  $\pi(q) > \pi(p) > \pi(s) > \pi(r)$ . Let  $\delta_i = 0$  for  $i \neq p, q, r, s$ . We can fix a nonzero solution  $(\delta_p, \delta_q, \delta_r, \delta_s) \in \mathbb{R}^4 \setminus \{0\}$  to the following system of homogeneous linear equations:

$$\delta_p + \delta_q = 0, \quad \delta_r + \delta_s = 0, \quad a_p \delta_p + a_q \delta_q + a_r \delta_r + a_s \delta_s = 0.$$

Then  $\sum_i \delta_i = 0$ ,  $\alpha_1 = 0$ , and  $\alpha_2 = \delta_p \delta_r + \delta_p \delta_s + \delta_q \delta_r + \delta_q \delta_s = (\delta_p + \delta_q)(\delta_r + \delta_s) = 0$ . Moreover  $\{p, q, r\}, \{p, q, s\}, \{p, r, s\}, \{q, r, s\} \notin Q_\pi$ , so  $\beta_3 = 0$ . We claim that

$$\beta_2 = -c_{p,q} \delta_p^2 + (c_{p,r} + c_{q,s} - c_{p,s} - c_{q,r}) \delta_p \delta_r - c_{r,s} \delta_r^2 \leq 0.$$

Let  $d = c_{p,r} + c_{q,s} - c_{p,s} - c_{q,r}$ . Lemma 4.5 implies that  $c_{p,q} \geq |c_{q,r} - c_{p,r}|$  and  $c_{p,q} \geq |c_{p,s} - c_{q,s}|$ , so  $2c_{p,q} \geq |c_{q,r} - c_{p,r}| + |c_{p,s} - c_{q,s}| \geq |d|$ . Similarly,  $c_{r,s} \geq |c_{p,r} - c_{p,s}|$  and  $c_{r,s} \geq |c_{q,s} - c_{q,r}|$ , so  $2c_{r,s} \geq |d|$  too. Since either  $-d \leq 0$  or  $d \leq 0$ , the equations

$$\begin{aligned} 2\beta_2 &= -2c_{p,q} \delta_p^2 + 2d \delta_p \delta_r - 2c_{r,s} \delta_r^2 = -d(\delta_p - \delta_r)^2 - (2c_{p,q} - d) \delta_p^2 - (2c_{r,s} - d) \delta_r^2 \\ &= d(\delta_p + \delta_r)^2 - (2c_{p,q} + d) \delta_p^2 - (2c_{r,s} + d) \delta_r^2 \end{aligned}$$

imply that  $\beta_2 \leq 0$ .

PROOF (OF LEMMA 4.9). Each of the two conditions is true for  $\pi$  if and only if it is true for  $\pi^{-1}$ . It is easy to see that the second condition implies the first one. Conversely, suppose that  $\pi$  (and hence also  $\pi^{-1}$ ) satisfies the first condition. We prove by induction on  $l$ . The statement is trivial for  $l = 1$ , so let  $l \geq 2$ . If  $\pi(l) = 1$  then we can use the induction hypothesis for  $\pi|_{\{1, \dots, l-1\}} \in \sigma_{l-1}$ . If  $\pi(1) = l$  then we can use the induction hypothesis for  $\pi' \in \sigma_{l-1}$ , where  $\pi'(i) = \pi(i+1) - 1$  for every  $i \in \{1, \dots, l-1\}$ . So we may assume  $\pi(1) \neq l$  and  $\pi(l) \neq 1$ .

Suppose that  $\pi^{-1}(l) > \pi^{-1}(1)$  and set  $k = \pi^{-1}(l)$ . If  $\pi^{-1}(1) < k-1 < k$ , then  $1 = \pi(\pi^{-1}(1)) < \pi(k-1) < \pi(k) = l$ , which contradicts the condition on  $\pi$ . Consider  $\pi^{-1}(1) = k-1$ . If  $i < j < k$ , then we cannot have  $\pi(i) < \pi(j)$ , because then we would have  $\pi(i) < \pi(j) < \pi(k) = l$ , contradicting the condition on  $\pi$ . If  $k-1 < i < j$  then we cannot have  $\pi(i) < \pi(j)$ , because then we would have  $1 = \pi(\pi^{-1}(1)) = \pi(k-1) < \pi(i) < \pi(j)$ , contradicting the condition on  $\pi$ . So  $\pi(i) > \pi(i+1)$  for every  $i \in \{1, \dots, l-1\} \setminus \{k\}$ , hence  $\pi$  is almost decreasing.

Now suppose that  $\pi^{-1}(l) < \pi^{-1}(1)$ . If  $\pi(l) < \pi(1)$ , then the condition on  $\pi$  is false for  $p = 1$ ,  $q = \pi^{-1}(l)$ ,  $r = \pi^{-1}(1)$ ,  $s = l$ . So  $\pi(l) > \pi(1)$  and  $(\pi^{-1})^{-1}(l) > (\pi^{-1})^{-1}(1)$ . Applying the previous paragraph to  $\pi^{-1}$  shows that  $\pi^{-1}$  is almost decreasing.

## References

- Balakrishnan, N. and Lai, C.-D. (2009) *Continuous Bivariate Distributions*. Springer.
- Daniels, H. (1950) Rank correlation and population models. *Journal of the Royal Statistical Society. Series B (Methodological)*, **12**, 171–191.
- Durante, F. and Sempi, C. (2015) *Principles of Copula Theory*. Chapman and Hall/CRC.
- Durbin, J. and Stuart, A. (1951) Inversions and rank correlation coefficients. *Journal of the Royal Statistical Society. Series B (Methodological)*, 303–309.
- Embrechts, P., Lindskog, F. and McNeil, A. (2003) Modelling dependence with copulas and applications to risk management. *Handbook of heavy tailed distributions in finance*, **8**, 329–384.
- Fredricks, G. A. and Nelsen, R. B. (2007) On the relationship between Spearman's rho and Kendall's tau for pairs of continuous random variables. *Journal of Statistical Planning and Inference*, **137**, 2143–2150.

- Genest, C. and Nešlehová, J. (2007) A primer on copulas for count data. *ASTIN Bulletin*, **37**, 475–515.
- (2009) Analytical proofs of classical inequalities between Spearman’s  $\rho$  and Kendall’s  $\tau$ . *Journal of Statistical Planning and Inference*, **139**, 3795–3798.
- Hürlimann, W. (2003) Hutchinson-Lai’s conjecture for bivariate extreme value copulas. *Statistics and Probability Letters*, **61**, 191–198.
- Kallenberg, O. (2002) *Foundations of modern probability*. Springer Science & Business Media.
- Klenke, A. (2013) *Probability theory: a comprehensive course*. Springer Science & Business Media.
- Kruskal, W. H. (1958) Ordinal measures of association. *Journal of the American Statistical Association*, **53**, 814–861.
- Makarov, G. (1981) Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed. *Theory of Probability & Its Applications*, **26**, 803–806.
- Mikusinski, P., Sherwood, H. and Taylor, M. (1992) Shuffles of Min. *Stochastica: revista de matemática pura y aplicada*, **13**, 61–74.
- Nelsen, R. B. (2006) *An introduction to copulas*. Springer Science & Business Media.
- Sack, J. and Úlfarsson, H. (2011) Refined inversion statistics on permutations. *arXiv preprint arXiv:1106.1995*.
- Scarsini, M. (1984) On measures of concordance. *Stochastica*, **8**, 201–218.
- Shao, W., Guo, G., Zhao, G. and Meng, F. (2014) Simulated annealing for the bounds of Kendall’s  $\tau$  and Spearman’s  $\rho$ . *Journal of Statistical Computation and Simulation*, **84**, 2688–2699.
- Trutschnig, W. (2011) On a strong metric on the space of copulas and its induced dependence measure. *Journal of Mathematical Analysis and Applications*, **384**, 690–705.
- Wang, R. (2015) Current open questions in complete mixability. *Probability Surveys*, **12**, 13–32.