

The exact Spearman rho–footrule region via optimal transport with applications to finite rankings, mixability, and Chatterjee’s rank correlation

Jonathan Ansari*

Marcus Rockel†

August 21, 2026

Abstract

We solve the open problem of determining the maximal value of Spearman’s rho when Spearman’s footrule is prescribed, thereby completing the exact attainable region of these two quantities. To prove this result, we reformulate the underlying copula optimization problem as an optimal transport problem with a linear moment constraint and construct the unique optimal coupling through a matching feasible dual potential and its contact set. Equivalently, this coupling minimizes the variance of $|U - V|$ among all couplings of $U, V \sim \mathcal{U}(0, 1)$ with prescribed mean $\mathbb{E}|U - V|$. Our main result admits several applications: First, in the context of finite rankings, we obtain an improved Cauchy–Schwarz inequality between Spearman’s footrule distance and the associated quadratic rank difference. Second, in the framework of generalized mixability, we characterize the attainable constant values of $|U' + V'|$, for $U', V' \sim \mathcal{U}(-\frac{1}{2}, \frac{1}{2})$, and determine the minimal quadratic deviation for a given mean. Third, we derive explicit bounds relating Chatterjee’s rank correlation $\xi(X, Y)$, which can detect complex functional dependence of Y on X , to the copula correlation ratio—a rank-based fraction of explained variance—by exploiting their conditional i.i.d. representations in terms of Spearman’s footrule and Spearman’s rho.

Keywords: Chatterjee’s rank correlation; copula correlation ratio; exact attainable region; finite ranking; generalized mixability; Kantorovich duality; optimal transport; shuffle of min; Spearman’s footrule distance; Spearman’s rho

1 Introduction

Measures of association quantify the dependence structure between two random variables X and Y . Classical rank correlations such as Spearman’s rho and Kendall’s tau attain values in $[-1, 1]$ and describe the degree of positive or negative dependence. In particular, they are designed to capture the strength of *monotone* association; see e.g. [11, Section 2.4] for an overview. A different notion of dependence has received considerable attention in recent years. Dependence measures such as Chatterjee’s rank correlation [6] quantify the degree of *functional dependence* of Y on X . They take values in $[0, 1]$, where the value 0 characterizes independence and the value 1 characterizes perfect functional dependence, meaning that Y is almost surely a measurable function of X . Importantly, this function need not be monotone, but can be arbitrarily complex. Such measures therefore detect forms of dependence that may be invisible to classical rank correlations; see [7] for a review.

Generally, measures of association capture distinct aspects of dependence and therefore provide complementary information about the underlying dependence structure. A natural question is to what extent their values constrain one another and how they can be interpreted in terms of each other. Given two measures of association κ_1 and κ_2 , one may consider their *exact attainable region*

$$\Omega_{\kappa_1, \kappa_2} := \{(\kappa_1(C), \kappa_2(C)) : C \in \mathcal{C}\},$$

where $\mathcal{C}$ denotes the class of bivariate copulas. We generally assume that X and Y have a continuous distribution function, so that by Sklar’s theorem (see Eq. (48)) their dependence structure is fully

*Department of Mathematics, Paris Lodron Universität Salzburg, Hellbrunner Straße 34, 5020 Salzburg, Austria.

†Department of Quantitative Finance, Albert-Ludwigs-Universität Freiburg, Rempartstraße 16, 79098 Freiburg, Germany.

captured by a unique copula $C = C_{X,Y}$. The region $\Omega_{\kappa_1, \kappa_2}$ describes all pairs of values that can occur simultaneously and thus provides a sharp quantitative comparison of the two measures. Following the solution of the classical Kendall–Spearman problem by Schreyer et al. [28], exact regions have been determined for several pairs of measures of association; see e.g. [20, 17, 18, 2].

In this paper, we solve the open problem of maximizing Spearman’s rho when the value of Spearman’s footrule is prescribed. While Spearman’s rho is a well-established rank correlation, Spearman’s footrule is less familiar, but no less relevant; see [15] and the applications below. Both admit a representation in terms of the underlying copula $C = C_{X,Y}$ via

$$\varrho(C) := 12 \int_{[0,1]^2} C(u,v) \, d\lambda_2(u,v) - 3, \quad \phi(C) := 6 \int_0^1 C(t,t) \, dt - 2; \quad (1)$$

see [11, 23] for general background. Further, they are affine in C and continuous with respect to uniform convergence, so the region $\Omega_{\varrho, \phi} = \{(\varrho(C), \phi(C)) : C \in \mathcal{C}\}$ is convex and compact. While ϱ attains values from -1 to 1 , ϕ ranges from $-\frac{1}{2}$ to 1 . Kokol Bukovšek and Stopar [19] determined the ϱ -minimal boundary $\underline{\varrho}(x) := \frac{2\sqrt{3}}{9}(1+2x)^{3/2} - 1$ of $\Omega_{\varrho, \phi}$ exactly and established the upper bound $u(x) := 1 - \frac{2}{3}(1-x)^2$, so that

$$\underline{\varrho}(\phi(C)) \leq \varrho(C) \leq u(\phi(C)) \quad \text{for all } C \in \mathcal{C}. \quad (2)$$

The lower bound $\underline{\varrho}(x)$ is attained for all $x \in [-\frac{1}{2}, 1]$ by the Bertino copulas B_x in (50). In contrast, the upper bound $u(x)$ is attained at the countably many points $x_N := 1 - \frac{3}{2N}$, $N \geq 1$, by equidistant even shuffles of min [21], as well as at $x = 1$, and it is not globally sharp. Tschimpke et al. [32] gave alternative proofs of both bounds and constructed the attainable curve in (14), which, however, is below the optimal, yet unknown curve

$$\bar{\varrho}(x) := \max\{\varrho(C) \mid C \in \mathcal{C}, \phi(C) = x\}, \quad x \in [-\frac{1}{2}, 1]. \quad (3)$$

In our main result, Theorem 1.1, we solve the optimization problem (3) and derive a closed-form expression for $\bar{\varrho}$. Further, for every $x \in [-\frac{1}{2}, 1]$, we explicitly construct a copula C_x such that $\phi(C_x) = x$ and $\varrho(C_x) = \bar{\varrho}(x)$ and prove that it is unique. Key is the moment representation

$$\varrho(C) = 1 - 6\mathbb{E}|U - V|^2 \quad \text{and} \quad \phi(C) = 1 - 3\mathbb{E}|U - V| \quad \text{for } (U, V) \sim C; \quad (4)$$

see Lemma 3.2. This allows us to solve the copula maximization problem (3) with tools from optimal transport theory: we minimize $\mathbb{E}|U - V|^2$ over all couplings of $U, V \sim \mathcal{U}(0, 1)$ under the linear constraint $\mathbb{E}|U - V| = m$; see the paragraph after Theorem 1.1 for a sketch of the proof that we carry out in Sections 4 and 5.

The moment representation of ϱ and ϕ in (4) and the closed-form expression of $\bar{\varrho}$ in (9) admit several interesting consequences and applications. First, we obtain an improved Cauchy–Schwarz inequality for finite rankings. For a permutation π of $\{1, \dots, n\}$, let

$$D_\pi := \sum_{i=1}^n |i - \pi(i)| \quad \text{and} \quad S_\pi := \sum_{i=1}^n (i - \pi(i))^2$$

denote the Spearman footrule distance and the corresponding quadratic rank differences, respectively; see e.g. [8, 10]. Then, by Theorem 2.1, the term $V_{\min}(m) \geq 0$ in (18) sharpens the elementary Cauchy–Schwarz inequality $D_\pi^2/n \leq S_\pi$ in [10, p. 268] to

$$\frac{D_\pi^2}{n} + n^3 V_{\min}(m) \leq S_\pi \quad \text{for } m = \frac{D_\pi}{n^2}.$$

We refer to Section 2.1 for details.

A second application arises in the context of mixability. The standard notion of mixability asks whether a sum of identically distributed random variables can be constant; see e.g. [35]. Generalized mixability in [5] extends this concept and studies supermodular functions of random vectors. We consider the supermodular function $\psi(u, v) = |u + v|$ and obtain that

$$\psi(U', V') = m \text{ a.s. for } U', V' \sim \mathcal{U}(-\frac{1}{2}, \frac{1}{2}) \iff m \in \mathfrak{C} := \{0\} \cup \left\{ \frac{1}{2N} : N \in \mathbb{N} \right\}. \quad (5)$$

More generally, we determine, for a prescribed value $m \in [0, \frac{1}{2}]$, the minimal quadratic deviation of $\psi(U', V')$ to m over all couplings of $U', V' \sim \mathcal{U}(-\frac{1}{2}, \frac{1}{2})$; see Theorem 2.4. Interestingly, there are countably many centers m for the $\mathcal{U}(-\frac{1}{2}, \frac{1}{2})$ -distribution with respect to the mixing function ψ , whereas standard mixability for an integrable distribution admits at most one center.

A third application of Theorem 1.1 arises for Chatterjee's rank correlation ξ in (34), recently introduced in [6]. As mentioned above, it ranges from 0 to 1 and can detect arbitrary functional relationships of Y on X . However, there are still open questions, for example, on the interpretation of its values. A natural question is how much variance of Y can be explained by X when the value of $\xi(X, Y)$ is known. To give a first answer to this question, we consider the copula correlation ratio $\eta(X, Y)$ in (39), a rank-transformed fraction of explained variance [29]. Interestingly, ξ and η admit representations in terms of Spearman's footrule and rho via

$$\xi(C) = \phi(C * C) \quad \text{and} \quad \eta(C) = \varrho(C * C); \quad (6)$$

see (35) and (40), where $C * C$ denotes the Markov product in (37). Applying the exact (ϱ, ϕ) -region in Corollary 1.2 and Cauchy–Schwarz inequality, we obtain the bounds

$$\max\{0, \underline{\varrho}(\xi(X, Y))\} \leq \eta(X, Y) \leq \min\{\bar{\varrho}(\xi(X, Y)), 2\xi(X, Y)\}, \quad (7)$$

see Theorem 2.6. These bounds are quite tight due to Proposition 2.10 where we constructed an inner enclosure of the (ξ, η) -region. For an illustration of the ξ - η -bounds, we refer to Figure 2.

1.1 Main result

For $N \in \mathbb{N}$, recall $x_N = 1 - \frac{3}{2N}$ and define

$$I_N := [x_N, x_{N+1}) = \left[1 - \frac{3}{2N}, 1 - \frac{3}{2N+2}\right). \quad (8)$$

These intervals cover $[-\frac{1}{2}, 1)$ and are pairwise disjoint. The following theorem, visualized in Figure 1, is our main result.

Theorem 1.1 (Closed-form expression for $\bar{\varrho}$). *Let $x \in I_N$, put $m := \frac{1-x}{3} \in [\frac{1}{2N+2}, \frac{1}{2N}]$, let $\ell \in \{\frac{1}{2N}, \frac{1}{2N+2}\}$ be an endpoint nearest to m (if the two endpoints are equally near, either choice is allowed; the two values of (9) coincide), and set $\Delta := |m - \ell|$. Then we have*

$$\bar{\varrho}(x) = 1 - 6\ell(2m - \ell) - \frac{4\Delta^{3/2}}{\sqrt{N(N+1)}} = u(x) - \left(\frac{4\Delta^{3/2}}{\sqrt{N(N+1)}} - 6\Delta^2\right). \quad (9)$$

In particular, the symmetric copula C_x in (75) satisfies $\varrho(C_x) = \bar{\varrho}(x)$ and $\phi(C_x) = x$. Moreover, it is the unique copula with these two properties. At the remaining endpoint $x = 1$, we have $\bar{\varrho}(1) = 1$, uniquely attained by the upper Fréchet copula $C_1(u, v) = \min\{u, v\}$.

To prove Theorem 1.1, we solve the copula maximization problem (3) by translating it into an optimal transport problem and then applying standard tools from OT theory. To briefly explain our ideas, let us recall the moment representation ϱ and ϕ in (4). Hence, the ϱ -maximal boundary of $\Omega_{\varrho, \phi}$ corresponds to minimizing $\mathbb{E}|U - V|^2$ under all random vectors (U, V) with $U, V \sim \mathcal{U}(0, 1)$ at a prescribed value $\mathbb{E}|U - V|$. For $m \in [0, \frac{1}{2}]$, this is the OT problem with a linear constraint

$$P(m) := \min_{\pi \in \Pi} \left\{ \int_{[0,1]^2} (b-a)^2 d\pi(a, b) : \int_{[0,1]^2} |b-a| d\pi(a, b) = m \right\}, \quad (10)$$

where Π denotes the set of couplings of two $\mathcal{U}(0, 1)$ distributions, i.e., the set of distributions on $[0, 1]^2$ with uniform marginals. Recall that the distribution function associated with a coupling $\pi \in \Pi$ is a (bivariate) copula. Conversely, the probability measure induced by a bivariate copula is an element of Π . Hence, for $m = \frac{1-x}{3}$, the OT problem (10) is a reformulation of the copula problem (3). To solve the *primal problem* (10), we consider its dual formulation

$$\sup_{\varphi \oplus \psi \leq c_\theta} \int_0^1 \varphi(a) da + \int_0^1 \psi(b) db, \quad (11)$$

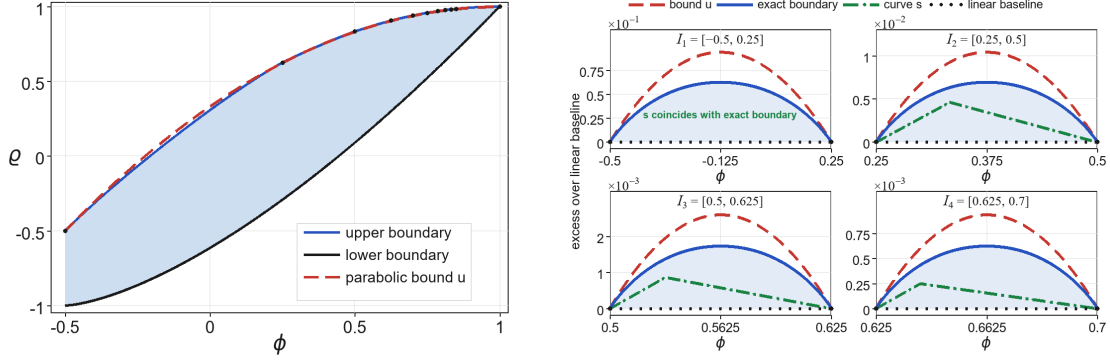


Figure 1: Left: the exact region $\Omega_{\varrho, \phi}$ in (13), bounded above by $\bar{\varrho}$ from (9) and below by the sharp boundary of Kokol Bukovšek and Stopar [19]. The dashed curve is u from (2), and the black dots mark its contact points with $\bar{\varrho}$. Right: on the first four intervals $I_1, \dots, I_4$ defined in (8), the vertical excesses of u , $\bar{\varrho}$, and the attainable curve s of Tschimpke et al. [32] are plotted relative to the piecewise-linear baseline obtained by joining consecutive black contact points in the left panel. The vertical scale is adapted to each interval.

where the optimization is over all continuous functions $\varphi, \psi: [0, 1] \rightarrow \mathbb{R}$ and constants $\theta \in \mathbb{R}$ such that

$$\varphi \oplus \psi(a, b) := \varphi(a) + \psi(b) \leq (b - a)^2 - \theta(|b - a| - m) =: c_\theta(a, b)$$

for all $a, b \in [0, 1]$. Kantorovich duality with linear constraints yields equality of the primal and dual values; see [37, Theorem 2.1].

Since ϱ and ϕ are symmetric, the transport cost c_θ is symmetric. Since, additionally, the optimization in (10) is over distributions having identical marginals, the dual problem (11) can be rewritten as

$$D(m) := \sup_{\substack{f \in C([0, 1]), \theta \in \mathbb{R} \\ f \oplus f \leq c_\theta}} 2 \int_0^1 f(u) du; \quad (12)$$

see Proposition 3.6. Here $C([0, 1])$ denotes the space of continuous real-valued functions on $[0, 1]$. To determine the optimal value $P(m)$, we provide a feasible coupling π_x and a feasible potential (f_x, θ) such that π_x is concentrated on the contact set of (f_x, θ) . In this case, (f_x, θ) maximizes the dual problem and π_x minimizes the primal problem, so that $P(m) = D(m)$; see Corollary 3.7. Then the copula associated with π_x solves the maximization problem (3).

As a consequence of Theorem 1.1, we can determine the exact region of Spearman's rho and Spearman's footrule as follows.

Corollary 1.2 (The exact region $\Omega_{\varrho, \phi}$). *For $\bar{\varrho}$ in (9) and $\underline{\varrho}(x) = \frac{2\sqrt{3}}{9}(1 + 2x)^{3/2} - 1$, we have*

$$\Omega_{\varrho, \phi} = \{(\varrho, x) : x \in [-\tfrac{1}{2}, 1], \underline{\varrho}(x) \leq \varrho \leq \bar{\varrho}(x)\}. \quad (13)$$

Beyond completing the exact (ϱ, ϕ) -region, Theorem 1.1 admits a direct probabilistic interpretation in terms of the first two moments of the absolute rank difference. Before developing this interpretation, we briefly relate Theorem 1.1 to the previously known bounds.

Remark 1.3. *As motivated in the introduction, measures of association provide information about the underlying dependence structure. In particular, the extremal values $\varrho(C) = \pm 1$ determine the copula C as the upper/lower Fréchet copula, respectively. The uniqueness statement in Theorem 1.1 shows that an analogous phenomenon occurs along the upper boundary of the (ϱ, ϕ) -region: if $\phi(C) = x$ and $\varrho(C) = \bar{\varrho}(x)$, then we have $C = C_x$. Thus, every point on the upper boundary uniquely determines the corresponding copula C_x . Conversely, if $(\varrho(C), \phi(C))$ is separated from the boundary of $\Omega_{\varrho, \phi}$, then C cannot be close to either C_x or B_x , where $x = \phi(C)$, with respect to any metric of weak convergence.*

Remark 1.4 (Relation to earlier results). *As mentioned above, Kokol Bukovšek and Stopar [19] determined the sharp lower boundary $\underline{\varrho}$ exactly and proved the universal upper bound u that coincides, due to (9), with $\bar{\varrho}$ at the points $\{x_N\}_{N \geq 1}$ and at 1. They also constructed an attainable curve r , which coincides on the interval $[-\frac{1}{2}, -\frac{1}{8}]$ with $\bar{\varrho}$ and the function s below. Tschimpke et al. [32] improved the attainable curve r by the function*

$$s(x) = \begin{cases} 2x + \frac{1}{2} - \frac{\sqrt{3}}{9}(1+2x)^{3/2}, & x \in [-\frac{1}{2}, -\frac{1}{8}), \\ x + \frac{3}{8} - \frac{\sqrt{6}}{36}(1-4x)^{3/2}, & x \in [-\frac{1}{8}, \frac{1}{4}), \\ u(x_N) + \frac{w_N - u(x_N)}{z_N - x_N}(x - x_N), & x \in [x_N, z_N), \quad N \geq 2, \\ w_N + \frac{u(x_{N+1}) - w_N}{x_{N+1} - z_N}(x - z_N), & x \in [z_N, x_{N+1}), \quad N \geq 2, \\ 1, & x = 1. \end{cases} \quad (14)$$

where $z_N := (2N^2 + N - 4)/(2(N+1)^2)$ and $w_N := (2N^5 + 6N^4 + 3N^3 - 7N^2 - 3N + 1)/(2N^2(N+1)^3)$. They conjecture that s is not optimal at any $x \in (-\frac{1}{8}, 1) \setminus \{x_N : N \geq 2\}$; see [32, Paragraph after Thm. 5.3]. Theorem 1.1 gives the precise resolution:

$$\bar{\varrho}(x) = s(x) \quad \text{for } x \in [-\frac{1}{2}, \frac{1}{4}] \cup \{x_N : N \geq 2\} \cup \{1\}, \quad \bar{\varrho}(x) > s(x) \quad \text{for } x \in \bigcup_{N \geq 2} (x_N, x_{N+1}).$$

Consequently, the function s is also optimal on $(-\frac{1}{8}, \frac{1}{4})$.

1.2 Discussion of Theorem 1.1

For a discussion of the $\bar{\varrho}$ -formula in (9), let us first recall that the universal upper bound u is a direct consequence of the Cauchy–Schwarz inequality. Therefore, define the absolute difference

$$Z := |U - V|, \quad \text{for } (U, V) \sim C. \quad (15)$$

Then, using the representation of ϱ and ϕ in (4) and applying the Cauchy–Schwarz inequality, gives

$$\frac{1 - \varrho(C)}{6} = \mathbb{E}[Z^2] \geq (\mathbb{E}[Z])^2 = \left(\frac{1 - \phi(C)}{3}\right)^2. \quad (16)$$

Solving the above inequality for $\varrho(C)$ yields $\varrho(C) \leq 1 - \frac{2}{3}(1 - \phi(C))^2$. This implies, setting $x = \phi(C)$, the expression

$$u(x) = 1 - \frac{2}{3}(1 - x)^2. \quad (17)$$

Remark 1.5. *The elementary Cauchy–Schwarz argument above, applied directly to the population versions of ϱ and ϕ , provides a considerably shorter proof of the upper bound u established in [19, Theorem 11] and [32, Theorem 3.4].*

The precise sense in which Theorem 1.1 improves Cauchy–Schwarz is worth isolating. For an arbitrary square-integrable nonnegative random variable $\tilde{Z}$ with mean $m = \mathbb{E}\tilde{Z}$, Cauchy–Schwarz inequality (equivalently, nonnegativity of the variance) gives only

$$\mathbb{E}\tilde{Z}^2 \geq m^2.$$

This inequality cannot be improved from the value of m and the boundedness condition $0 \leq \tilde{Z} \leq 1$ alone, since the constant random variable $\tilde{Z} = m$ attains equality. Our variable Z from (15) additionally satisfies $U, V \sim \mathcal{U}(0, 1)$. Thus Z is not an arbitrary bounded random variable: it is the absolute difference induced by a coupling with two prescribed uniform marginals. No independence, symmetry, or particular form of dependence between U and V is assumed. It is precisely this uniform-marginal, or bistochastic, constraint that yields the improvement. The following result gives the sharp strengthening of Cauchy–Schwarz.

Proposition 1.6 (Improvement of Cauchy–Schwarz inequality under uniform marginals).

For $U, V \sim \mathcal{U}(0, 1)$, set $m := \mathbb{E}|U - V|$. Then the following hold true:

- (i) For $m > 0$, choose $N \geq 1$ such that $m \in \left(\frac{1}{2N+2}, \frac{1}{2N}\right]$. Let $\ell \in \{\frac{1}{2N+2}, \frac{1}{2N}\}$ be an endpoint nearest to m , put $\Delta_m := |m - \ell|$, and define

$$V_{\min}(m) := \frac{2\Delta_m^{3/2}}{3\sqrt{N(N+1)}} - \Delta_m^2, \quad \text{and} \quad V_{\min}(0) := 0. \quad (18)$$

Then

$$\mathbb{E}(U - V)^2 \geq m^2 + V_{\min}(m). \quad (19)$$

- (ii) The bound in (19) is sharp for every m . For $m > 0$, it is attained by the copula C_x from Definition 4.1 with $x = 1 - 3m$, and for $m = 0$ by the comonotonicity copula $M(u, v) = \min\{u, v\}$.

- (iii) The correction term satisfies $V_{\min}(m) \geq 0$ and

$$V_{\min}(m) = 0 \iff m \in \mathfrak{C} := \{0\} \cup \left\{\frac{1}{2N} : N \in \mathbb{N}\right\}. \quad (20)$$

As a consequence of the above result, the ordinary Cauchy–Schwarz inequality

$$\mathbb{E}(U - V)^2 \geq (\mathbb{E}|U - V|)^2 \quad (21)$$

is strict whenever $m = \mathbb{E}|U - V| \notin \mathfrak{C}$. Equality in (21) is equivalent to the constant-displacement condition

$$|U - V| = m \quad \text{almost surely.} \quad (22)$$

For an illustration of the nontrivial equality cases in (21), we refer to Figure 7, where $m = \frac{1}{4}$ and $m = \frac{1}{6}$ occur as endpoint cases of the copula family constructed in Section 4. In terms of the concordance coefficients ϱ and ϕ , the values in Proposition 1.6 are precisely the points at which the ordinary Cauchy–Schwarz bound u touches the sharp boundary $\bar{\varrho}$. We return to the constant-displacement condition (22) from the perspective of generalized mixability in Section 2.2.

Combining the upper and lower boundaries of the exact (ϱ, ϕ) -region, the following results determines the possible variances of $|U - V|$ when its mean is prescribed but the dependence structure of $U, V \sim \mathcal{U}(0, 1)$ is unspecified. Recall that C_x in (75) is ϱ -maximal under the constraint $\phi(C_x) = x$, whereas the Bertino copula B_x in (50) is ϱ -minimal under the same constraint.

Corollary 1.7 (Exact mean–variance region for $|U - V|$). For $U, V \sim \mathcal{U}(0, 1)$ and $m = \mathbb{E}|U - V|$, let $V_{\min}(m)$ be given by (18) and define

$$V_{\max}(m) := \frac{1 - (1 - 2m)^{3/2}}{3} - m^2. \quad (23)$$

Then

$$V_{\min}(m) \leq \text{Var}(|U - V|) \leq V_{\max}(m). \quad (24)$$

The lower bound is attained for $(U, V) \sim C_x$, $x = 1 - 3m$, and the upper bound is attained for $(U, V) \sim B_x$.

Remark 1.8. (a) Every value between the two bounds in (24) is attainable. Indeed, both C_x and B_x satisfy $\phi(C_x) = \phi(B_x) = x = 1 - 3m$, and hence have the same mean absolute difference m . Since both ϕ and ϱ are affine in the copula, convex mixtures of C_x and B_x attain every intermediate variance. For $m = 0$, both bounds are zero, that is $V_{\min}(0) = V_{\max}(0) = 0$.

- (b) Corollary 1.7 provides an equivalent mean–variance interpretation of Theorem 1.1: for every prescribed mean $m = \mathbb{E}|U - V|$, it determines exactly the smallest and largest possible variance of the absolute difference $|U - V|$ among all couplings of two uniform random variables. In particular, the lower boundary $V_{\min}(m)$ is precisely the sharp Cauchy–Schwarz correction from Proposition 1.6.

(c) If the mean m is not prescribed, optimizing the two bounds in (24) over $m \in [0, \frac{1}{2}]$ yields the universal variance bounds

$$0 \leq \text{Var}(|U - V|) \leq \frac{5(3 - \sqrt{5})}{24}. \quad (25)$$

The lower bound is attained whenever $|U - V|$ vanishes almost surely, i.e., $U = V$ almost surely. The upper bound is attained for $(U, V) \sim B_{x_0}$, where $x_0 := \frac{7-3\sqrt{5}}{4}$.

1.3 Organization of the paper

The rest of the paper is organized as follows. Section 2 develops consequences and applications of Theorem 1.1 to finite rankings, generalized mixability, and the relation between Chatterjee's rank correlation and the copula correlation ratio. Section 3 collects notation and reviews the optimal transport concepts used throughout the paper, in particular Kantorovich duality with linear constraints. Section 4 constructs the candidate copula C_x in (75), verifies the moment constraint of the associated coupling, and calculates its rank correlations in Proposition 4.6. Section 5 introduces the dual potential in Definition 5.1, proves equality on the support of the candidate coupling in Lemma 5.4, establishes global dual feasibility in Lemma 5.5, and proves optimality and uniqueness in Proposition 5.8. Finally, the proofs of the results stated in Sections 1 and 2 are collected in Sections 6 and 7, respectively.

2 Consequences and applications

The closed-form expression of $\bar{\varrho}$ in Theorem 1.1 leads to several consequences beyond the comparison of Spearman's rho and Spearman's footrule. We first translate the exact mean–variance region in (24) into inequalities relating absolute and quadratic errors of finite rankings. We then interpret the equality cases of the sharp Cauchy–Schwarz inequality under uniform marginals in terms of generalized mixability and quantify the minimal dispersion when exact mixability is impossible. Finally, we apply the exact (ϱ, ϕ) -region to the comparison of Chatterjee's rank correlation with the copula correlation ratio, two recently studied measures of directed dependence.

2.1 Inequalities for finite rankings

We first translate the variance bounds in Corollary 1.7 into inequalities for two familiar distances between finite rankings. Let $\mathfrak{S}_n$ denote the set of permutations on $\{1, \dots, n\}$ and let $\pi \in \mathfrak{S}_n$. We compare the ranking π with the identity ranking. Two popular unnormalized distances are

$$D_\pi := \sum_{i=1}^n |i - \pi(i)| \quad \text{and} \quad S_\pi := \sum_{i=1}^n (i - \pi(i))^2, \quad (26)$$

which measure absolute and quadratic rank error, respectively; see, e.g., [8, 10]. In particular, $|i - \pi(i)|$ is the absolute *rank displacement* of item i .

To apply the copula bounds, we embed the finite permutation into a coupling with continuous uniform marginals. Let I be uniformly distributed on $\{1, \dots, n\}$, let $W \sim \mathcal{U}(0, 1)$ be independent of I , and define

$$U_\pi := \frac{I - 1 + W}{n}, \quad V_\pi := \frac{\pi(I) - 1 + W}{n}. \quad (27)$$

Then $U_\pi, V_\pi \sim \mathcal{U}(0, 1)$. Indeed, conditionally on $I = i$, the variable U_π is uniform on the i -th interval of the regular partition of $[0, 1]$, whereas V_π is uniform on the $\pi(i)$ -th interval. Since π is a permutation, averaging over I yields uniform marginals. Moreover,

$$U_\pi - V_\pi = \frac{I - \pi(I)}{n},$$

so the coupling preserves the normalized rank displacement exactly. The corresponding copula is a shuffle of M .

Consequently, the mean absolute and mean squared normalized rank displacements are

$$m_\pi := \mathbb{E}|U_\pi - V_\pi| = \frac{D_\pi}{n^2}, \quad q_\pi := \mathbb{E}(U_\pi - V_\pi)^2 = \frac{S_\pi}{n^3}. \quad (28)$$

Here one factor $1/n$ comes from averaging over the n items, while normalizing the rank displacement by n contributes one further factor to m_π and two further factors to q_π . The next result is a discretized version of Corollary 1.7.

Theorem 2.1 (Finite-permutation inequalities). *Every $\pi \in \mathfrak{S}_n$ satisfies*

$$m_\pi^2 + V_{\min}(m_\pi) \leq q_\pi \leq m_\pi^2 + V_{\max}(m_\pi) = \frac{1 - (1 - 2m_\pi)^{3/2}}{3}, \quad (29)$$

where $V_{\min}$ and $V_{\max}$ are given in (18) and (23), respectively.

Remark 2.2. *The lower bound in (29) improves the Cauchy–Schwarz inequality $S_\pi \geq D_\pi^2/n$ to*

$$S_\pi \geq \frac{D_\pi^2}{n} + n^3 V_{\min}(m_\pi). \quad (30)$$

The correction is strictly positive whenever $m_\pi \notin \mathfrak{C}$. For $m_\pi = 1/(2N)$, equality in the ordinary Cauchy–Schwarz inequality is attainable by a permutation of size n if and only if $n/(2N)$ is an integer. Equivalently, there exists a permutation $\pi \in \mathfrak{S}_n$ satisfying

$$|i - \pi(i)| = \frac{n}{2N}, \quad i = 1, \dots, n.$$

Consequently, the bounds in (29) are sharp for the continuum problem of copulas and asymptotically sharp for finite rankings, but need not be optimal for each fixed n .

2.2 Generalized mixability and minimal dispersion

Classical complete mixability asks whether random variables with prescribed marginal distributions can be coupled such that their sum is constant almost surely; see, e.g., [35]. For integrable marginal distributions, the corresponding center is necessarily unique, since

$$X_1 + \dots + X_n = c \quad \text{almost surely} \quad (31)$$

implies $c = \sum_{i=1}^n \mathbb{E}X_i$. Bignozzi and Puccetti [5] introduced the more general notions of Ψ -complete and Ψ -joint mixability, replacing the sum in (31) by a measurable aggregation function $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}$. In particular, a distribution F is Ψ -completely mixable with index n and center c if there exist $X_1, \dots, X_n \sim F$ such that

$$\Psi(X_1, \dots, X_n) = c \quad \text{almost surely.}$$

Unlike for the classical sum, the center of a nonlinear aggregation function need not be determined by the marginal means and, in particular, several centers may be possible.

A particularly important class considered in [5] is given by supermodular aggregation functions. Recall that a function $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}$ is called *supermodular* if

$$\Psi(x) + \Psi(y) \leq \Psi(x \wedge y) + \Psi(x \vee y), \quad x, y \in \mathbb{R}^n,$$

where $x \wedge y$ and $x \vee y$ denote the componentwise minimum and maximum, respectively. For twice continuously differentiable functions, supermodularity is equivalent to $\partial_{ij}\Psi \geq 0$ for all $i \neq j$.

Our main result, Theorem 1.1, admits a natural interpretation in this framework. To make the connection to supermodular aggregation functions explicit, let $U' := U - \frac{1}{2}$ and $V' := \frac{1}{2} - V$. Then U', V' are uniform on $(-1/2, 1/2)$ and

$$|U - V| = |U' + V'|. \quad (32)$$

Consequently, for the supermodular function $\Psi(u, v) := |u + v|$, the constant-distance condition $|U - V| = m$ almost surely is equivalent to $\Psi(U', V') = m$ almost surely. Thus the constant-distance problem can be viewed as a Ψ -complete mixability problem for the centered uniform distribution.

The equality cases of the Cauchy–Schwarz bound that are attainable under the uniform-marginal constraint admit the following reformulation in terms of generalized complete mixability. Recall the set $\mathfrak{C}$ in (20).

Lemma 2.3 (Ψ -complete mixability of the centered uniform distribution). *Let $\Psi(u, v) = |u + v|$. There exist $U', V' \sim \mathcal{U}(-\frac{1}{2}, \frac{1}{2})$ such that $\Psi(U', V') = m$ almost surely if and only if $m \in \mathfrak{C}$.*

Hence, the distribution $\mathcal{U}(-\frac{1}{2}, \frac{1}{2})$ is Ψ -completely mixable with index 2 precisely for the countable set of centers $\mathfrak{C}$.

Indeed, generalized mixability asks whether the aggregate $\Psi(U', V')$ can be made constant. If exact constancy is impossible, it is natural to ask how close one can get to it while prescribing its mean. We therefore define the minimal variance

$$\mathbf{V}(m) := \min \left\{ \text{Var}(\Psi(U', V')) : U', V' \sim \mathcal{U}\left(-\frac{1}{2}, \frac{1}{2}\right), \mathbb{E}\Psi(U', V') = m \right\}, \quad m \in \left[0, \frac{1}{2}\right]. \quad (33)$$

Exact Ψ -complete mixability with center m is equivalent to $\mathbf{V}(m) = 0$. If $\mathbf{V}(m) > 0$, then exact mixability with center m is impossible, and $\mathbf{V}(m)$ quantifies the smallest quadratic dispersion that an aggregate with prescribed mean m can attain. In this sense, $\mathbf{V}(m)$ provides a quantitative relaxation of exact Ψ -complete mixability.

As a consequence of Proposition 1.6, we determine the minimal dispersion exactly as follows.

Theorem 2.4 (Minimal dispersion and generalized mixability). *For $\Psi(u, v) = |u + v|$, we have*

$$\mathbf{V}(m) = V_{\min}(m),$$

where $V_{\min}(m)$ is given by (18). In particular, $\mathbf{V}(m) = 0$ if and only if $m \in \mathfrak{C}$.

Remark 2.5. (a) *Supermodularity provides a natural link between generalized mixability and dependence optimization. Among all couplings with prescribed marginals, for expectations of supermodular functions, the comonotonic coupling is maximal, while the countermonotonic coupling is minimal; see, e.g., [22, Theorems 3.9.8 and 3.9.15]. Extremal expectations of supermodular functions can also be approximated numerically by rearrangement methods [25]. In our setting, $\Psi(u, v) = |u + v|$ is supermodular, and the countermonotonic coupling $V' = -U'$ yields the center 0. Unlike for classical complete mixability, however, centers of generalized Ψ -mixability need not be unique; see, e.g., [5, Example 11] for an example with two distinct centers. Lemma 2.3 shows an even richer phenomenon: for the centered uniform distribution, the set of centers is the countably infinite set $\mathfrak{C}$.*

(b) *Theorem 2.4 refines the binary question of exact Ψ -complete mixability into a quantitative one. The function $m \mapsto \mathbf{V}(m)$ gives the minimal variance of the aggregate $\Psi(U', V')$ under the constraint $\mathbb{E}\Psi(U', V') = m$. Its zeros are precisely the exact mixability centers $m \in \mathfrak{C}$, whereas $\mathbf{V}(m) > 0$ otherwise. Thus Proposition 1.6 determines how closely exact Ψ -complete mixability can be approximated at every prescribed mean aggregate.*

2.3 Chatterjee's rank correlation and the copula correlation ratio

As an application of Theorem 1.1 and the exact (ϱ, ϕ) -region in Corollary 1.2, we now determine bounds for two measures of *directed* dependence recently studied in the statistics literature. To be precise, let us first recall that classical measures of association such as Pearson correlation, Kendall's tau or Spearman's rho and footrule quantify the degree of positive or negative (linear) dependence between two random variables X and Y . However, they fail to detect non-linear and non-monotone relationships, respectively. For example, all these quantities vanish for X standard normal and $Y = X^2$.

Motivated by the seminal papers [9, 6, 4], in the last decade many works have focused on dependence measures κ satisfying the following axioms:

- (I) $\kappa(X, Y) \in [0, 1]$,
- (II) $\kappa(X, Y) = 0$ if and only if X and Y are independent,
- (III) $\kappa(X, Y) = 1$ if and only if Y perfectly depends on X , i.e., there exists a measurable function f such that $Y = f(X)$ almost surely.

Note that the functional relation in (III) is not assumed to be increasing or decreasing. Hence, in contrast to the classical measures of association, suitable dependence measures κ can detect complicated and complex functional dependencies.

The certainly most prominent such dependence measure is Chatterjee's rank correlation ξ [6] whose population version is given by

$$\xi(X, Y) = \frac{\int \text{Var}(P(Y \geq y | X)) dP^Y(y)}{\int \text{Var}(1_{\{Y \geq y\}}) dP^Y(y)}, \quad (34)$$

see [7] for a recent survey and [1, 3, 14, 16, 30, 36] for several extensions and related constructions. Small/Large variability of the conditional survival probability in the numerator of (34) indicates low/strong dependence of Y on X . The extreme cases of independence and perfect dependence are obtained, where $P(Y \geq y | X)$ is constant or coincides with the indicator function in the denominator for all y . To better understand the behavior of ξ and to interpret its values, it is important to compare it with related measures of association. For instance, attainable sets and inequalities with respect to Spearman's rho, Spearman's footrule, and Kendall's tau have been studied in [2, 26, 27].

In the rest of this section, we focus on comparing ξ with the (rank-transformed) fraction of explained variance. This comparison is particularly natural, since ξ measures the strength of functional dependence of Y on X , while the fraction of explained variance quantifies how much of the variance in Y can be accounted for by X . To explain the details, let us assume for simplicity that X and Y have a continuous distribution function. Then ξ depends only on the copula C of (X, Y) , and it reduces to the *Dette-Siburg-Stoimenov measure*

$$\xi(C) = 6 \int_0^1 \int_0^1 (\partial_1 C(u, v))^2 du dv - 2; \quad (35)$$

see [9]. Here, $\partial_1 C$ denotes the partial derivative of C with respect to the first component, which exists outside a Lebesgue null set [23]. Interestingly, the functional ξ in (35) admits a representation via Spearman's footrule through

$$\xi(C) = \phi(C * C); \quad (36)$$

see [13]. Here, $D * E$ denotes (a version of) the Markov product of two bivariate copulas $D, E \in \mathcal{C}$ defined by

$$D * E(u, v) := \int_0^1 \partial_1 D(t, u) \partial_1 E(t, v) dt, \quad (u, v) \in [0, 1]^2. \quad (37)$$

It is well known that the Markov product is a copula that models conditional independence [11, Chapter 5]. In particular, if Y' is a conditionally independent copy of Y given X , then we have

$$(U, U') \sim C * C, \quad \text{where } U := F_Y(Y) \text{ and } U' := F_Y(Y'). \quad (38)$$

Recall that $\phi(C) = 6 \int_0^1 C(t, t) dt - 2$. Hence, by (36), Chatterjee's rank correlation evaluates the Markov product only on its diagonal. While Chatterjee's rank correlation naturally complements Pearson's correlation and Spearman's and Kendall's rank correlation, several questions concerning the interpretation of its population value are open. Since ξ measures the strength of functional dependence of Y on X , a natural question is how much of the variance of the (rank-transformed) response Y can be explained by X when the value $\xi(Y, X)$ is given?

We therefore consider the copula correlation ratio [29, 31] defined by

$$\eta(C) := \eta(X, Y) := \frac{\text{Var}(\mathbb{E}[F_Y(Y) | X])}{\text{Var}(F_Y(Y))}. \quad (39)$$

It quantifies the strength of regression dependence and coincides with the fraction of explained variance or the first-order Sobol' index of the rank-transformed response $U = F_Y(Y)$ with respect to X ; see [14]. The copula correlation ratio satisfies Axioms (I) and (III), but does not characterize independence. Instead it satisfies

$$(II') \quad \eta(X, Y) = 0 \text{ if and only if } \text{Var}(\mathbb{E}[F_Y(Y) | X]) = 0;$$

see [3, Theorem 2.2]. The latter is, in particular, fulfilled if X and Y are independent (but not vice versa).

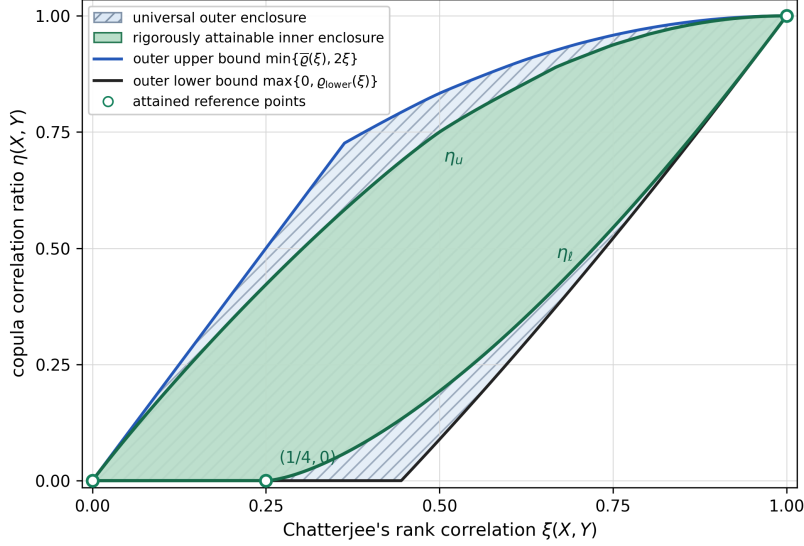


Figure 2: Outer and inner enclosures for $\Omega_{\xi, \eta} = \{(\xi(C), \eta(C)), C \in \mathcal{C}\}$. The hatched blue region is the universal outer enclosure from Theorem 2.6. The solid green region is the rigorously attainable inner enclosure from Proposition 2.10; every point in it is realized by a conditional-i.i.d. model. The two unshaded strips between the inner and outer boundaries remain unresolved.

While Chatterjee's rank correlation can be expressed as Spearman's footrule of the Markov product (see (36)), the copula correlation ratio admits a representation through Spearman's rho of the Markov product via

$$\eta(C) = \frac{\text{Var}(\mathbb{E}[U | X])}{\text{Var}(U)} = \frac{\mathbb{E}[UU'] - \mathbb{E}U\mathbb{E}U'}{\sqrt{\text{Var}(U)}\sqrt{\text{Var}(U')}} = \varrho(U, U') = \varrho(C * C), \quad (40)$$

where U and U' are given as in (38).

As a consequence of the (ϱ, ϕ) -bounds in Theorem 1.1, we determine the following bounds for the copula correlation ratio via Chatterjee's rank correlation. Vice versa, this gives bounds for ξ in terms of η ; see Figure 2 for an illustration.

Theorem 2.6 (ξ - η -bounds). *Assume that (X, Y) has a continuous distribution function. Then we have*

$$\max\{0, \underline{\varrho}(\xi(X, Y))\} \leq \eta(X, Y) \leq \min\{\bar{\varrho}(\xi(X, Y)), 2\xi(X, Y)\}, \quad (41)$$

where $\underline{\varrho}$ and $\bar{\varrho}$ are the bounds for the exact (ϱ, ϕ) -region in Corollary 1.2.

Remark 2.7. (a) The two quantities ξ and η capture different aspects of directed dependence. Chatterjee's ξ measures changes in the entire conditional distribution of the rank $U = F_Y(Y)$, whereas η only measures changes in its conditional mean. Consequently, deviations of $P(U \leq t | X)$ from t may cancel in the latter quantity as we show in Example 2.8.

(b) Theorem 2.6 provides an upper bound for η only. Indeed, it applies Theorem 1.1 to the Markov product $C * C$, but not every copula is of this form. Every such Markov product is symmetric and admits the conditional-i.i.d. representation in (38); in particular, this class is not dense in $\mathcal{C}$, because the symmetric copulas form a closed proper subclass of $\mathcal{C}$. Sharpness of the (ϱ, ϕ) -bounds over all copulas therefore does not imply sharpness under the conditional-i.i.d. restriction; see also Example 2.9.

Example 2.8 ($\xi(X, Y) = \frac{1}{4}$ while $\eta(X, Y) = 0$). For $U \sim \mathcal{U}(0, 1)$, set $Y = U$, and take $X = |2U - 1|$. Then $X \sim \mathcal{U}(0, 1)$ and, conditionally on $X = r$, the variable U is equally likely to be $(1 - r)/2$ or $(1 + r)/2$. Hence, $\mathbb{E}[U | X] = 1/2$, so that $\eta(X, Y) = 0$. On the other hand, let U' be a conditionally independent copy of U given X . Then U and U' differ by r with probability $1/2$. This gives

$$\xi(X, Y) = 1 - 3 \mathbb{E}|U - U'| = 1 - 3 \int_0^1 \underbrace{\mathbb{E}[|U - U'| | X = r]}_{=r/2} dr = 1 - \frac{3}{4} = \frac{1}{4}.$$

Thus a positive value of ξ need not imply a positive copula correlation ratio.

Example 2.9 (Non-sharpness of the upper ξ - η bound). *The upper bound in (41) is not sharp. Indeed, recall the Cauchy-Schwarz estimate*

$$\eta(X, Y) \leq 2\xi(X, Y).$$

With $U = F_Y(Y)$, set $g_X(t) := P(U \leq t \mid X) - t$. Equality in the above Cauchy-Schwarz inequality requires g_X to be constant almost everywhere on $[0, 1]$, for almost every X . Since $t \mapsto t + g_X(t)$ is a distribution function, this is possible only for $g_X = 0$, and hence only at $(\xi, \eta) = (0, 0)$.

Further, for $x = \frac{1}{4}$ and $N = 2$, Theorem 1.1 gives $\bar{\varrho}(1/4) = u(1/4) = 5/8$. Thus (41) yields the upper bound

$$\eta \leq \min \left\{ \bar{\varrho} \left(\frac{1}{4} \right), 2 \cdot \frac{1}{4} \right\} = \frac{1}{2}$$

at $\xi = 1/4$, but equality cannot be attained, as discussed before. Moreover, the attainable set of conditional-i.i.d. pairs (ξ, η) can be shown to be compact, so the maximal value of η at $\xi = 1/4$ is strictly smaller than $1/2$. Hence, the upper bound in (41) is not sharp. By contrast, Example 2.8 shows that the lower value $\eta = 0$ is attainable at $\xi = 1/4$.

To complement the outer bounds in (41), we now determine a large set of attainable (ξ, η) -pairs. Following the above setting, we restrict attention to random vectors (X, Y) with continuous marginal distribution functions. We denote the corresponding attainable (ξ, η) -region by

$$\Omega_{\xi, \eta} := \{ (\xi(X, Y), \eta(X, Y)) : F_X \text{ and } F_Y \text{ are continuous} \}. \quad (42)$$

Recall that $\xi(X, Y) = \phi(C * C)$ and $\eta(X, Y) = \varrho(C * C)$, where C is the copula of (X, Y) . Thus every point in $\Omega_{\xi, \eta}$ is generated by a copula of the form $C * C$, or equivalently by a conditional-i.i.d. pair (U, U') as in (38). As discussed above, Theorem 2.6 applies the exact (ϱ, ϕ) -region to the larger class of all copulas and provides an outer enclosure of $\Omega_{\xi, \eta}$ that is not sharp.

To the best of our knowledge, a sharp description of $\Omega_{\xi, \eta}$ is not known. We therefore complement the outer enclosure by constructing an explicit inner region. The natural object for this purpose is the random conditional distribution

$$\mu_X := \text{law}(U \mid X), \quad U = F_Y(Y).$$

Since $U \sim \mathcal{U}(0, 1)$, the conditional distributions satisfy the barycenter condition $\mathbb{E}\mu_X = \lambda$, where λ denotes Lebesgue measure on $[0, 1]$. This means that, for every Borel set $A \subseteq [0, 1]$,

$$\mathbb{E}[\mu_X(A)] = P(U \in A) = \lambda(A).$$

If U' is drawn conditionally independently from the same distribution μ_X , then

$$\xi(X, Y) = 1 - 3\mathbb{E}|U - U'|, \quad \eta(X, Y) = 12 \text{Var} \left(\int_0^1 u \mu_X(du) \right). \quad (43)$$

Hence, ξ reflects the average within-distribution spread of the conditional laws μ_X , whereas η depends only on the variation of their conditional means. This distinction suggests two complementary constructions.

For the lower inner curve η_ℓ in Figure 2, we exploit the fact that η only depends on the variability of the conditional mean. To keep η small, we therefore keep $\mathbb{E}[U \mid X] = 1/2$ whenever possible by grouping ranks into symmetric pairs $\{u, 1 - u\}$. We then reveal U exactly on an increasingly large central interval, while retaining these symmetric two-point conditional distributions outside. This interpolates between the model in Example 2.8 and complete dependence and yields the lower inner curve

$$\eta_\ell(x) := \begin{cases} 0, & 0 \leq x \leq \frac{1}{4}, \\ \left(\frac{4x - 1}{3} \right)^{3/2}, & \frac{1}{4} \leq x \leq 1. \end{cases} \quad (44)$$

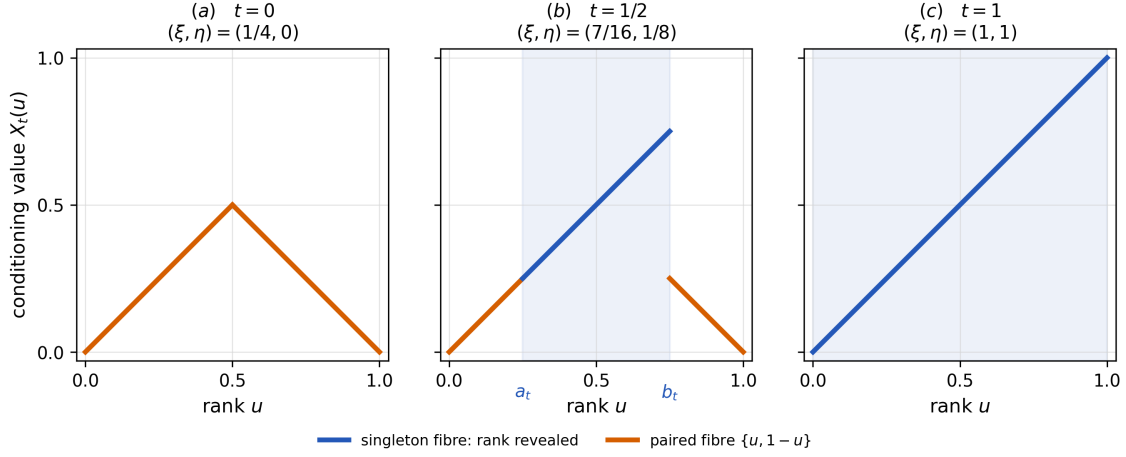


Figure 3: Construction of the lower inner curve η_ℓ in (44). The central interval $[a_t, b_t]$ is revealed exactly, so its conditional fibers are singletons (blue). Outside this interval, the symmetric ranks u and $1-u$ form a common fiber (orange). As t increases from 0 to 1, the revealed interval expands and the construction moves from $(\xi, \eta) = (1/4, 0)$ to complete dependence at $(1, 1)$.

We refer to Figure 3 for a visualization of the above construction. Further details are provided in the proof of Proposition 2.10.

For the upper inner curve η_u in Figure 2, we proceed in the opposite direction. Here the aim is to make the conditional means vary strongly, and hence to make η large, while keeping a controlled amount of variability within the conditional distribution of U given X . For the construction of η_u in (45), we first split the uniform distribution into two conditional laws with different means whose equally weighted mixture recovers the required uniform marginal distribution. We then place affine copies of this binary model into n equal subintervals of $[0, 1]$. The interval index determines the location of U up to an interval of length $1/n$, while the parameter a controls the strength of the binary perturbation, and hence the within-cell variability of U . As n increases, this increasingly fine localization drives both ξ and η towards 1, whereas a determines the position of the resulting (ξ, η) -pair along each branch. To be precise, denote for a set $\mathcal{S} \subseteq [0, 1]^2$ by $\text{conv}(\mathcal{S})$ its convex hull. For $n \in \mathbb{N}$ and $a \in [0, \frac{1}{2}]$, put

$$x_n(a) := 1 - \frac{1 - 2a^2(3 - 4a)}{n}, \quad y_n(a) := 1 - \frac{1 - 12a^2(1 - a)^2}{n^2}.$$

The construction described above generates the set

$$\mathcal{S}_u := \left\{ (x_n(a), y_n(a)) : n \in \mathbb{N}, a \in [0, \tfrac{1}{2}] \right\} \cup \{(1, 1)\},$$

whose upper convexified boundary is the function $\eta_u: [0, 1] \rightarrow [0, 1]$ defined by

$$\eta_u(x) := \max \{y : (x, y) \in \text{conv}(\mathcal{S}_u)\}, \quad 0 \leq x \leq 1. \quad (45)$$

For details of the construction, we refer to the proof of the following proposition, which establishes the corresponding inner enclosure.

Proposition 2.10 (Constructive inner enclosure). *The set $\mathcal{S}_u$ is compact, its projection onto the first coordinate is $[0, 1]$, and η_u is a well-defined concave function. Moreover,*

$$\{(x, y) : 0 \leq x \leq 1, \eta_\ell(x) \leq y \leq \eta_u(x)\} \subseteq \Omega_{\xi, \eta}. \quad (46)$$

Remark 2.11. (a) The functions η_ℓ and η_u provide, respectively, constructive lower and upper inner bounds for the attainable region $\Omega_{\xi, \eta}$. Together with the outer bounds in (41), they yield explicit inner and outer enclosures of the exact (ξ, η) -region; see Figure 2. Numerical evaluation shows that the constructive inner enclosure occupies approximately 81.4% of the area of the outer enclosure. Determining the exact region $\Omega_{\xi, \eta}$, and in particular whether either η_ℓ or η_u is sharp, remains open.

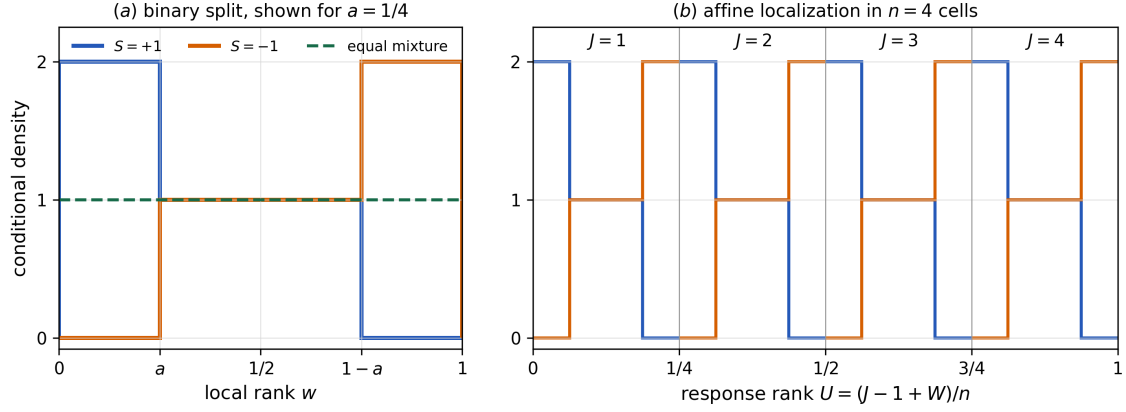


Figure 4: Construction underlying the upper inner curve η_u in (45). Panel (a) shows, for $a = 1/4$, the two mirror-image conditional densities associated with the binary mixing variable $S = +1$ and $S = -1$; their equally weighted mixture is the uniform density; see Eq. (121). Panel (b) shows affine copies of this binary model in $n = 4$ equal cells. The variable J selects the cell and S selects the conditional law within it.

(b) The lower construction illustrates how ξ can increase while η remains small. On the interval $0 \leq x \leq 1/4$, the lower bound satisfies $\eta_\ell(x) = 0$; these values are attained by mixing independence with the model in Example 2.8. For $x \in [1/4, 1]$, progressively revealing U on a central interval increases both measures and yields the curve

$$\eta_\ell(x) = \left(\frac{4x - 1}{3} \right)^{3/2},$$

connecting $(1/4, 0)$ with complete dependence at $(1, 1)$.

(c) The two parameters in the upper construction have distinct roles. The integer n controls the coarse localization of U : the larger n , the smaller the interval of length $1/n$ in which U is known to lie, thereby driving both ξ and η towards 1. By contrast, the parameter a controls the conditional structure within each interval through the strength of the binary perturbation. Thus n governs the global localization of U , whereas a determines the remaining conditional variability within each local cell.

3 Copula and optimal-transport preliminaries

We first collect the copula and measure-theoretic notation used throughout the proofs. Let $\mathbb{N} := \{1, 2, \dots\}$, and let λ and $\lambda_2 := \lambda \otimes \lambda$ denote Lebesgue measure on $[0, 1]$ and $[0, 1]^2$, respectively. We write $\mathcal{U}(0, 1)$ for the uniform distribution on $[0, 1]$. For a Borel set A , the restriction of a measure μ to A is denoted by $\mu|_A$, and $\mathcal{M}(A)$ denotes the set of finite Borel measures on A . If $T: A \rightarrow B$ is Borel measurable and $\mu \in \mathcal{M}(A)$, its pushforward under T is the measure $T_\# \mu$ defined by

$$(T_\# \mu)(E) := \mu(T^{-1}(E)), \quad E \subseteq B \text{ Borel}.$$

We write $\text{proj}_1(u, v) := u$ and $\text{proj}_2(u, v) := v$, and denote the marginals of $\mu \in \mathcal{M}([0, 1]^2)$ by $\mu_i := (\text{proj}_i)_\# \mu$, $i = 1, 2$.

A *bivariate copula* is a bivariate distribution function C on $[0, 1]^2$ with uniform margins, that is,

$$C(u, 1) = u, \quad C(1, v) = v, \quad u, v \in [0, 1]. \quad (47)$$

We denote the class of bivariate copulas by $\mathcal{C}$. The concept of copula is motivated by Sklar's theorem which states that, for every bivariate distribution function H with marginal distribution functions F and G , there exists $C \in \mathcal{C}$ such that

$$H(x, y) = C(F(x), G(y)), \quad x, y \in \mathbb{R}. \quad (48)$$

Further, the copula C is unique on $\text{Ran}(F) \times \text{Ran}(G)$, and hence unique whenever F and G are continuous. Conversely, for any $C \in \mathcal{C}$ and any univariate distribution functions F, G , the function H in (48) is a bivariate distribution function with marginals F and G ; see [23, Thm. 2.3.3]. In particular, if X and Y have continuous distribution functions, their copula coincides with the joint distribution function of the rank-transformed vector $(F_X(X), F_Y(Y))$.

For $C \in \mathcal{C}$, let μ_C denote the Borel probability measure determined by

$$\mu_C([0, u] \times [0, v]) = C(u, v), \quad u, v \in [0, 1].$$

By (47), both marginal distributions of μ_C are given by the Lebesgue measure $\lambda|_{[0,1]} = \mathcal{U}(0, 1)$; such measures are also called *doubly stochastic*. Thus $C \mapsto \mu_C$ is a bijection from $\mathcal{C}$ onto the set

$$\Pi := \{\pi \in \mathcal{M}([0, 1]^2) : \pi_1 = \pi_2 = \lambda\}$$

of couplings of two uniform distributions.

The upper Fréchet copula $M(u, v) := \min\{u, v\}$ models comonotonicity, i.e., perfect positive dependence. A copula C is called a *shuffle of min* if there is a finite interval partition of $[0, 1]$ and a measure-preserving bijection $T: [0, 1] \rightarrow [0, 1]$ that is affine with slope $+1$ or -1 on the interior of each partition interval such that

$$\mu_C = (\text{id}, T)_\# \lambda.$$

Thus the mass of a shuffle of min is concentrated on finitely many line segments of slope $+1$ or -1 ; see [21].

Example 3.1 (Bertino copulas). *For a copula C , the function $\delta_C(t) := C(t, t)$ is called its diagonal. Given a copula diagonal δ , the associated Bertino copula is*

$$B_\delta(u, v) := u \wedge v - \min_{u \wedge v \leq t \leq u \vee v} (t - \delta(t)); \quad (49)$$

see [12, Sec. 5]. The Bertino copula family $(B_x)_{x \in [-1/2, 1]}$ that describes the lower bound $\underline{\varrho}$ due to [19] is obtained from the diagonals

$$\delta_a(t) := \begin{cases} 0, & 0 \leq t \leq a, \\ t - a, & a \leq t \leq 1 - a, \\ 2t - 1, & 1 - a \leq t \leq 1, \end{cases} \quad a \in [0, \tfrac{1}{2}],$$

by setting

$$B_x := B_{\delta_{a_x}} \quad \text{for} \quad a_x := \frac{1}{2} \left(1 - \sqrt{\frac{1+2x}{3}} \right). \quad (50)$$

This is the three-strip shuffle of min from [19, Example 6], which satisfies $\phi(B_x) = x$ and $\varrho(B_x) = \underline{\varrho}(x)$. Figure 5 illustrates how the support of B_x changes along the lower boundary.

For $C \in \mathcal{C}$, let $C^\top(u, v) := C(v, u)$ be its transpose, and call C *symmetric* if $C = C^\top$. With $S(u, v) := (v, u)$, this is equivalent to $S_\# \mu_C = \mu_C$; accordingly, we denote the set of symmetric couplings by

$$\Pi_{\text{sym}} := \{\pi \in \Pi : S_\# \pi = \pi\}.$$

Since ϕ and ϱ in (1) are affine in C and invariant under transposition, symmetrization $C \mapsto \frac{1}{2}(C + C^\top)$ preserves both quantities. Hence all extremal problems below admit a symmetric optimizer. Throughout, $(U, V) \sim C$ means that (U, V) has distribution μ_C .

Lemma 3.2 (Moment representation). *For every $C \in \mathcal{C}$ the identities (4) hold.*

Proof. By Fubini's theorem,

$$\int_{[0,1]^2} C(u, v) d\lambda_2(u, v) = \int_{[0,1]^2} P(U \leq u, V \leq v) d\lambda_2(u, v) = \mathbb{E}[(1 - U)(1 - V)] = \mathbb{E}[UV],$$

where the last equality uses the uniform margins. Moreover, for $D := U - V$, we obtain $2\mathbb{E}[UV] = \mathbb{E}U^2 + \mathbb{E}V^2 - \mathbb{E}D^2 = \frac{2}{3} - \mathbb{E}D^2$, and hence $\varrho = 12\mathbb{E}[UV] - 3 = 1 - 6\mathbb{E}D^2$. Similarly $\int_0^1 C(t, t) dt = \mathbb{E}[1 - \max(U, V)]$ and $\max(U, V) = \frac{U+V+|D|}{2}$, so $\phi = 6(1 - \frac{1}{2} - \frac{1}{2}\mathbb{E}|D|) - 2 = 1 - 3\mathbb{E}|D|$. $\square$

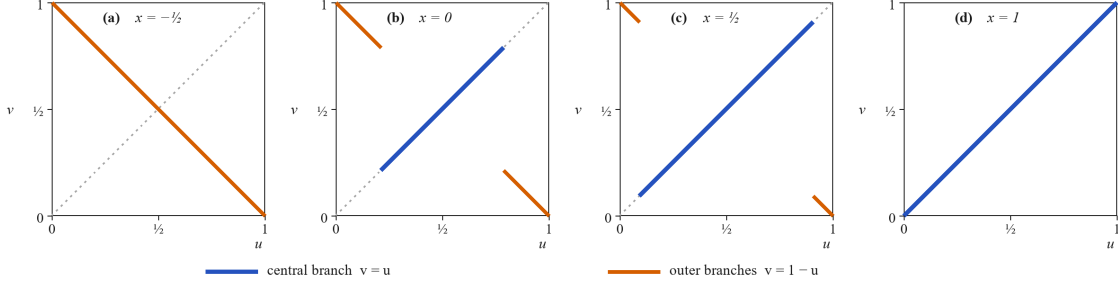


Figure 5: Supports of four lower-bound Bertino copulas $B_x = B_{\delta_{a_x}}$ from (50). The solid blue segment is the central branch $v = u$, the solid orange segments are the outer branches $v = 1 - u$, and the dotted gray diagonal is shown for reference. As x increases, a_x decreases from $\frac{1}{2}$ to 0, deforming the countermonotone copula $W(u, v) = \max\{u + v - 1, 0\}$ into the Min copula M . Since $(\varrho(B_x), \phi(B_x)) = (\varrho(x), x)$, the family traces the lower boundary of $\Omega_{\varrho, \phi}$.

Every $\pi \in \Pi$ satisfies $\pi_1 + \pi_2 = 2\lambda$. We next characterize Π_{sym} through measures on the triangle

$$T := \{(x, y) \in [0, 1]^2 : x \leq y\}.$$

Define

$$\Pi_T := \{\nu \in \mathcal{M}(T) \mid \nu_1 + \nu_2 = \lambda\}. \quad (51)$$

Every $\nu \in \Pi_T$ has total mass $\nu(T) = \frac{1}{2}$, obtained by evaluating $\nu_1 + \nu_2 = \lambda$ on $[0, 1]$.

Lemma 3.3 (Representation of symmetric couplings). *The mapping $\Phi : \Pi_T \rightarrow \Pi_{\text{sym}}$, $\Phi(\nu) := \nu + S_{\#}\nu$, is a bijection.*

Proof. Let $\nu \in \Pi_T$ and set $\pi := \Phi(\nu) = \nu + S_{\#}\nu$. Since $S \circ S = \text{id}$, we obtain

$$S_{\#}\pi = S_{\#}\nu + S_{\#}(S_{\#}\nu) = S_{\#}\nu + \nu = \pi,$$

so that π is symmetric. Moreover, $\pi_1 = \nu_1 + (S_{\#}\nu)_1 = \nu_1 + \nu_2 = \lambda$. Similarly, $\pi_2 = \lambda$. Hence $\pi \in \Pi_{\text{sym}}$, and thus Φ is well defined.

To prove surjectivity, let $\pi \in \Pi_{\text{sym}}$, and denote by $\text{Diag} := \{(x, x) : x \in [0, 1]\}$ the diagonal. Define a measure ν on T by

$$\nu := \pi|_{\{x < y\}} + \frac{1}{2} \pi|_{\text{Diag}}.$$

Since π is symmetric, we have $\pi|_{\{x < y\}} = S_{\#}(\pi|_{\{x > y\}})$. Moreover, S acts as the identity on Diag . Therefore, $\pi = \nu + S_{\#}\nu$. Consequently,

$$\lambda = \pi_1 = \nu_1 + (S_{\#}\nu)_1 = \nu_1 + \nu_2,$$

which shows that $\nu \in \Pi_T$. Thus Φ is surjective.

Finally, suppose that $\nu, \tilde{\nu} \in \Pi_T$ satisfy $\nu + S_{\#}\nu = \tilde{\nu} + S_{\#}\tilde{\nu}$. Restricting this identity to $\{x < y\}$ gives $\nu|_{\{x < y\}} = \tilde{\nu}|_{\{x < y\}}$, because $S_{\#}\nu$ and $S_{\#}\tilde{\nu}$ are supported on $\{x \geq y\}$. Restricting it to Diag yields $2\nu|_{\text{Diag}} = 2\tilde{\nu}|_{\text{Diag}}$. Hence $\nu = \tilde{\nu}$, and therefore Φ is injective. $\square$

Using the representation of symmetric couplings in Lemma 3.3, the ϱ -maximal value in dependence on $\phi(C) = x$ can be written as an optimal transport problem via the measures Π_T :

Lemma 3.4 (Representation of $\bar{\varrho}$). *Let $x \in [-\frac{1}{2}, 1]$. For the ϱ -maximal value in (3), we have*

$$\bar{\varrho}(x) = 1 - 12 \min \left\{ \int_T (b - a)^2 d\nu(a, b) : \nu \in \Pi_T, \int_T (b - a) d\nu(a, b) = \frac{m}{2} \right\}, \quad (52)$$

where Π_T is defined in (51) and $m := \frac{1-x}{3}$.

Proof. Given a symmetric $C \in \mathcal{C}$, let $\nu := \frac{1}{2}(\min, \max)_{\#} \mu_C$ on T . For integrable $g : [0, 1] \rightarrow \mathbb{R}$,

$$\begin{aligned} \int_T [g(a) + g(b)] d\nu(a, b) &= \frac{1}{2} \int_{[0, 1]^2} [g(u \wedge v) + g(u \vee v)] d\mu_C(u, v) \\ &= \frac{1}{2} \int_{[0, 1]^2} [g(u) + g(v)] d\mu_C(u, v) = \int_0^1 g(t) d\lambda(t), \end{aligned}$$

so $\nu \in \Pi_T$. In particular, taking $g \equiv 1$ gives $2\nu(T) = 1$. Moreover, $\int_T h(b - a) d\nu(a, b) = \frac{1}{2} \mathbb{E}[h(|U - V|)]$ for $(U, V) \sim C$ and for every bounded Borel function $h : [0, \infty) \rightarrow \mathbb{R}$. Conversely, given $\nu \in \Pi_T$, the measure $\mu(A) := \int_T [\mathbf{1}_A(a, b) + \mathbf{1}_A(b, a)] d\nu(a, b)$ is symmetric with uniform margins by the same computation, hence doubly stochastic, and reverses the correspondence. Now apply Lemma 3.2 and the symmetrization remark above; the constraint $\phi = x$ becomes $\mathbb{E}|D| = m$, i.e. $\int_T (b - a) d\nu(a, b) = \frac{m}{2}$. $\square$

Problem (52) is an infinite-dimensional linear program of transportation type: mass on the triangle T with a coupled marginal constraint, cost $(b - a)^2$ and one moment constraint. Sections 4 and 5 solve it. We remark that a second reduction is available and was used for independent numerical validation: ϕ depends on C only through the diagonal $\delta_C(t) = C(t, t)$, symmetrization preserves the diagonal, and among symmetric copulas with diagonal δ the diagonal copula $E_\delta(u, v) = \min\{u, v, \frac{\delta(u) + \delta(v)}{2}\}$ is pointwise maximal, see [24, 33, 12]. Since ϱ is increasing with respect to the pointwise order, the ϱ -maximal boundary equals $\max\{\varrho(E_\delta)\}$ over all diagonals with prescribed integral. The objective is a concave functional of δ .

Turning to the dual problem (12), we introduce the following notion of feasibility. A pair (f, θ) is called *feasible* for the dual problem (12) if it satisfies the inequality constraint

$$f(a) + f(b) + \theta((b - a) - m) \leq (b - a)^2 \quad \text{for all } (a, b) \in T. \quad (53)$$

By symmetry, this is equivalent to the full-square constraint in (12). Similarly, we call a measure $\pi \in \Pi$ *feasible* for the primal problem (10) if it satisfies the fixed-moment constraint (56) below. Since the optimization problem (10) is symmetric, we may equivalently focus on feasible measures $\nu \in \Pi_T$ satisfying

$$\int_T (b - a) d\nu(a, b) = \frac{m}{2}.$$

For the next result, recall that $D(m)$ is the value of the dual formulation where we optimize with respect to one feasible potential.

Lemma 3.5. *For any feasible potential (f, θ) and for any feasible $\nu \in \Pi_T$, we have*

$$\int_0^1 f(u) du \leq \int_T (b - a)^2 d\nu(a, b) \quad (54)$$

and thus $D(m) \leq P(m)$.

Proof. Since $\nu_1 + \nu_2 = \lambda$, we obtain

$$\begin{aligned} \int_0^1 f(u) du &= \int_0^1 f(a) d\nu_1(a) + \int_0^1 f(b) d\nu_2(b) \\ &= \int_T (f(a) + f(b)) d\nu(a, b) \\ &= \int_T [f(a) + f(b) + \theta((b - a) - m)] d\nu(a, b) \\ &\leq \int_T (b - a)^2 d\nu(a, b), \end{aligned}$$

because $\nu(T) = \frac{1}{2}$ and feasibility gives $\int_T ((b - a) - m) d\nu = \frac{m}{2} - m\nu(T) = 0$.

Now let $\pi \in \Pi$ be feasible for the primal problem. Its symmetrization $\pi^{\text{sym}} := \frac{1}{2}(\pi + S_{\#}\pi)$ has the same cost and moment. The associated triangular measure

$$\nu_\pi := \frac{1}{2}(\min, \max)_{\#} \pi^{\text{sym}}$$

is feasible by Lemma 3.4. Applying the inequality just proved yields

$$2 \int_0^1 f(u) du \leq 2 \int_T (b-a)^2 d\nu_\pi(a,b) = \int_{[0,1]^2} (b-a)^2 d\pi(a,b).$$

Taking the supremum over feasible potentials and the minimum over feasible couplings proves $D(m) \leq P(m)$. $\square$

The classical Kantorovich duality theorem identifies the primal and dual values under standard regularity assumptions [34, Theorem 5.10]. For the present setting, the constrained extension in [37, Theorem 2.1] can be stated as follows. For every cost $c \in C([0,1]^2)$ and every linear subspace $W \subset C([0,1]^2)$, we have the strong duality

$$\inf_{\substack{\pi \in \Pi \\ \int w d\pi = 0 \text{ for all } w \in W}} \int_{[0,1]^2} c d\pi = \sup_{\substack{\varphi, \psi \in C([0,1]), \\ \varphi \oplus \psi + w \leq c \text{ for all } w \in W}} \left\{ \int_0^1 \varphi d\lambda + \int_0^1 \psi d\lambda \right\}. \quad (55)$$

The additional linear constraint in our primal problem (10) fixes the first absolute moment of the difference $U - V$:

$$\int_{[0,1]^2} |b-a| d\pi(a,b) = m. \quad (56)$$

The following proposition identifies the left-hand side of (55) with $P(m)$ in (10) and its right-hand side, after symmetrization, with $D(m)$ in (12). The annihilator constraint on the left-hand side reproduces (56) and the two-potential supremum on the right-hand side equals the one-potential value $D(m)$. Therefore, (55) reduces to the following strong duality result.

Proposition 3.6 (Kantorovich duality with linear constraint). *For every $m \in [0, \frac{1}{2}]$, the values of the primal problem (10) and the dual problem (12) coincide, i.e., $P(m) = D(m)$.*

It is important to mention that minimizers of the primal problem exist here by compactness, whereas the supremum in the dual problem need not be attained, even when strong duality holds. If, however, there is some feasible pair (f, θ) and some feasible measure $\nu \in \Pi_T$ such that equality holds in (54), then ν is optimal for the triangular problem (52), and $\nu + S_\# \nu$ is optimal for the primal problem (10). Furthermore, (f, θ) is dual optimal in the sense that the supremum in (12) is attained. This motivates us to consider the *contact set*

$$\Gamma_{f,\theta} := \{(a,b) \in T \mid f(a) + f(b) + \theta((b-a) - m) = (b-a)^2\}, \quad (57)$$

for any feasible (f, θ) . The following result is a direct consequence of Lemma 3.5 and Proposition 3.6. It states, in particular, that if ν is concentrated on the contact set of (f, θ) , then it is optimal.

Corollary 3.7 (Optimality via contact set). *Let $\nu \in \Pi_T$ be feasible and let (f, θ) be feasible. If $\nu(T \setminus \Gamma_{f,\theta}) = 0$, then ν is optimal for (52), $\nu + S_\# \nu$ is optimal for (10), and (f, θ) is optimal for (12).*

To solve the maximization problem (3), we will apply Corollary 3.7. To be precise, in Section 4, we provide a feasible candidate solution $\nu \in \Pi_T$ and determine the associated copula C_x . In Section 5, we provide a feasible pair (f, θ) and then show that ν is concentrated on the contact set of (f, θ) . Consequently, C_x is optimal and, as we will verify, it satisfies $\phi(C_x) = x$ and $\varrho(C_x) = \bar{\varrho}(x)$, so that $\bar{\varrho}$ determines the upper boundary of $\Omega_{\varrho, \phi}$.

4 Construction of optimal copulas

In this section, we construct a family of distributions $\pi_x \in \Pi$, $x \in [-\frac{1}{2}, 1)$, such that the associated copula family $(C_x)_{x \in [-\frac{1}{2}, 1]}$ satisfies $\phi(C_x) = x$. In Proposition 5.8, we will show that these copulas are optimal in the sense that they maximize Spearman's rho, i.e. $\varrho(C_x) = \bar{\varrho}(x)$. The case $x = 1$ is trivial, where we define $C_x(u, v) := \min\{u, v\}$. Hence, we assume $x \in [-\frac{1}{2}, 1)$ in the following.

Recall that, for $(U, V) \sim C$, we have $\phi(C) = 1 - 3\mathbb{E}|U - V|$. Then, for fixed $\phi(C) = x$, the first moment of $|U - V|$ is given by

$$m := m(x) = \frac{1-x}{3}. \quad (58)$$

We aim to minimize the second moment of $U - V$ over all copulas with fixed first absolute moment m . In view of (58), choose $N \in \mathbb{N}$ such that

$$x \in I_N = \left[1 - \frac{3}{2N}, 1 - \frac{3}{2N+2}\right). \quad (59)$$

Note that $I_1 = [-\frac{1}{2}, \frac{1}{4})$, $I_2 = [\frac{1}{4}, \frac{1}{2})$, $I_3 = [\frac{1}{2}, \frac{5}{8})$, ..., and $\bigcup_{N \in \mathbb{N}} I_N = [-\frac{1}{2}, 1)$. Defining

$$R := \frac{1}{2N+2} \quad \text{and} \quad L := \frac{1}{2N}, \quad (60)$$

the interval in (59) can also be written as $I_N = [1 - 3L, 1 - 3R)$. Now, we divide I_N into two intervals of equal length, where the left and right part are denoted by

$$I_N^L := \left[1 - 3L, 1 - \frac{3}{2}(L + R)\right) \quad \text{and} \quad I_N^R := \left[1 - \frac{3}{2}(L + R), 1 - 3R\right), \quad (61)$$

respectively. Define

$$d_N := \frac{1}{L - R} = 2N(N + 1) \quad (62)$$

as the reciprocal of the interval length of $[R, L]$. Then we have $d_N \lambda([R, L]) = d_N(L - R) = 1$. Further, it is $\lambda(I_N) = 3/d_N$. Note that

$$x \in I_N^L \iff m \in \left(\frac{L + R}{2}, L\right] \quad \text{and} \quad x \in I_N^R \iff m \in \left(R, \frac{L + R}{2}\right].$$

To define the candidate couplings $(\pi_x)_x$ via measures (ν_x) in Π_T , we introduce three functions depending on the constraint x as follows:

$$\begin{aligned} v = v(x) &:= \begin{cases} \sqrt{\frac{2(L - m(x))}{d_N}}, & \text{for } x \in I_N^L, \\ \sqrt{\frac{2(m(x) - R)}{d_N}}, & \text{for } x \in I_N^R. \end{cases} \\ p = p(x) &:= 1 - d_N v(x), \\ \theta = \theta(x) &:= \begin{cases} 2L - v(x), & \text{for } x \in I_N^L, \\ 2R + v(x), & \text{for } x \in I_N^R. \end{cases} \end{aligned} \quad (63)$$

The measure ν_x is now defined by

$$\nu_x := \begin{cases} \nu_x^L, & \text{for } x \in I_N^L, \\ \nu_x^R, & \text{for } x \in I_N^R, \end{cases} \quad (64)$$

with

$$\nu_x^L := \sum_{k=0}^{N-1} (A_k^L)_\#(\lambda|_{[0, pL]}) + \sum_{k=0}^{N-1} \alpha_k (B_k^L)_\#(\lambda|_{[0, Nv]}) + \sum_{k=0}^{N-1} \beta_k (C_k^L)_\#(\lambda|_{[0, Nv]}), \quad (65)$$

$$\nu_x^R := \sum_{k=0}^N (A_k^R)_\#(\lambda|_{[0, pR]}) + \sum_{k=0}^{N-1} \alpha_k (B_k^R)_\#(\lambda|_{[0, Nv]}) + \sum_{k=0}^{N-1} \beta_k (C_k^R)_\#(\lambda|_{[0, Nv]}) \quad (66)$$

for weights $\alpha_k := 1 - \frac{k}{N}$ and $\beta_k := \frac{k+1}{N}$. Further, for

$$\ell_L(t) := L - v + \frac{t}{N} \quad \text{and} \quad \ell_R(t) := R + \frac{t}{N}, \quad t \in [0, Nv], \quad (67)$$

the pushforward measures are defined via the mappings

$$A_k^L(s) := (k\theta + Nv + s, k\theta + Nv + s + L), \quad 0 \leq s \leq pL, \quad (68)$$

$$B_k^L(t) := (k\theta + t, k\theta + t + \ell_L(t)), \quad 0 \leq t \leq Nv, \quad (69)$$

$$C_k^L(t) := (k\theta + t + \ell_L(t), (k+1)\theta + t), \quad 0 \leq t \leq Nv, \quad (70)$$

$$A_k^R(s) := (k\theta + s, k\theta + s + R), \quad 0 \leq s \leq pR, \quad (71)$$

$$B_k^R(t) := (k\theta + pR + t, k\theta + pR + t + \ell_R(t)), \quad 0 \leq t \leq Nv, \quad (72)$$

$$C_k^R(t) := (k\theta + pR + t + \ell_R(t), (k+1)\theta + pR + t), \quad 0 \leq t \leq Nv. \quad (73)$$

In Lemma 4.3 below, we show that $\nu_x \in \Pi_T$. Hence, by Lemma 3.3, the following measure is a coupling in Π_{sym} :

$$\pi_x := \nu_x + S_{\#}\nu_x. \quad (74)$$

Since π_x has $\mathcal{U}(0, 1)$ -marginals, we may define the associated copula C_x as follows.

Definition 4.1 (Optimal copula C_x). *For $x \in [-\frac{1}{2}, 1)$, the copula associated with π_x is defined by*

$$C_x(u, v) := \pi_x([0, u] \times [0, v]), \quad (u, v) \in [0, 1]^2. \quad (75)$$

For $x = 1$, we define $C_1(u, v) := \min\{u, v\}$ as the upper Fréchet copula.

Remark 4.2.

- (a) Since π_x is symmetric, also C_x is symmetric. In Proposition 4.6, we will show that $\phi(C_x) = x$ and we will determine a closed-form expression for $\varrho(C_x)$. In Proposition 5.8, we verify that $\varrho(C) \leq \varrho(C_x)$ for every copula C with $\phi(C) = x$; hence, C_x is optimal in the sense that it maximizes Spearman's ϱ .
- (b) Figure 6 displays the support of ν_x for one right-half parameter choice, together with the corresponding dual potential. Figure 7 displays the support of $\pi_x = \nu_x + S_{\#}\nu_x$, which is the distribution underlying C_x . For $x \in \{1 - \frac{3}{2N} : N \in \mathbb{N}\} = \{-\frac{1}{2}, \frac{1}{4}, \frac{1}{2}, \frac{5}{8}, \dots\}$, C_x reduces to an equidistant even shuffle of min copula for which optimality has been shown in [19, Theorem 11 and Example 12].

For proving optimality of C_x , we begin by showing that ν_x defined in (64) is in Π_T .

Lemma 4.3. *For every $x \in [-\frac{1}{2}, 1)$, we have $\nu_x \in \Pi_T$.*

Proof. Recall that $N \in \mathbb{N}$ is chosen such that $x \in I_N$. Further, recall v, p, θ from (63). By the definitions of I_N^L and I_N^R in (61), we have $0 \leq v \leq L - R = \frac{1}{d_N}$. Consequently,

$$0 \leq p = 1 - d_N v \leq 1.$$

Moreover, $\alpha_k = 1 - \frac{k}{N} \geq 0$ and $\beta_k = \frac{k+1}{N} \geq 0$ for $k = 0, \dots, N-1$. Thus, ν_x is a finite nonnegative Borel measure.

It remains to show that ν_x is supported on T and satisfies

$$(\text{proj}_1)_{\#}\nu_x + (\text{proj}_2)_{\#}\nu_x = \lambda. \quad (76)$$

We treat the left and right parts separately.

Case 1 (the right-side part): Let $x \in I_N^R$. Since $d_N R = N$, we obtain $pR = (1 - d_N v)R = R - Nv$, and therefore

$$pR + Nv = R. \quad (77)$$

Furthermore, using $\theta = 2R + v$, we have

$$\begin{aligned} N\theta + R + pR &= N(2R + v) + R + pR \\ &= 2NR + Nv + R + pR \\ &= 2(N+1)R = 1. \end{aligned} \quad (78)$$

The coordinate differences appearing in the definition of ν_x^R are given by

$$\begin{aligned} \text{proj}_2(A_k^R(s)) - \text{proj}_1(A_k^R(s)) &= R, \\ \text{proj}_2(B_k^R(t)) - \text{proj}_1(B_k^R(t)) &= \ell_R(t) = R + \frac{t}{N}, \\ \text{proj}_2(C_k^R(t)) - \text{proj}_1(C_k^R(t)) &= \theta - \ell_R(t) = R + v - \frac{t}{N}. \end{aligned}$$

Hence all these lengths belong to the interval $[R, R + v]$ and are nonnegative.

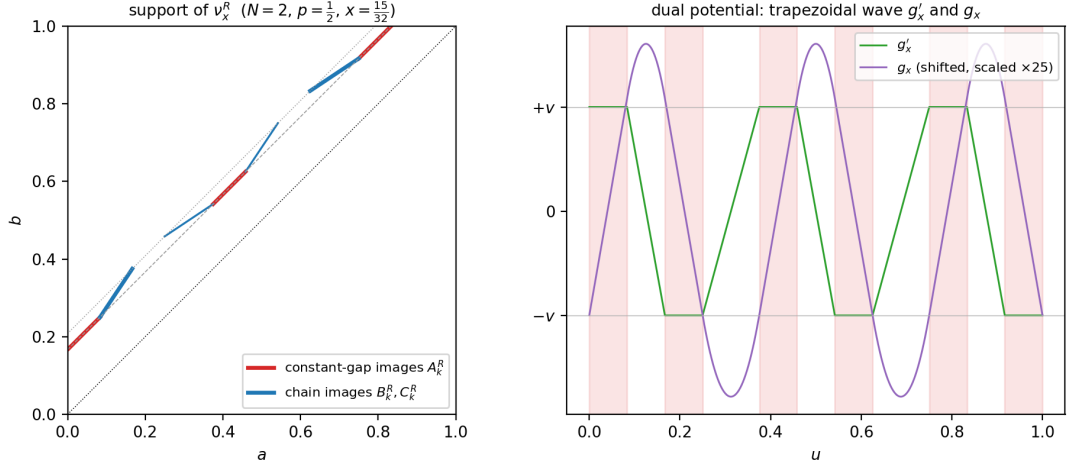


Figure 6: Left: support of the measure $\nu_x = \nu_x^R$ from (64) and (66), for $N = 2$ and $p = \frac{1}{2}$ (equivalently, $x = \frac{15}{32}$), inside $T = \{(a, b) : a \leq b\}$. The red segments are the images of the maps A_k^R and have constant difference $b - a = R$. The blue segments are the images of B_k^R and C_k^R ; their differences $b - a$ fill the interval $[R, R + v]$. Thus, projection by $q(a, b) = b - a$ yields the right-half distance law in Lemma 4.5. Grey lines mark $b - a = R$ (dashed) and $b - a = R + v$ (dotted). Right: the trapezoidal derivative g'_x and the θ -periodic dual potential g_x from Definition 5.1; the shaded intervals are the first- and second-coordinate ranges of the constant-difference segments.

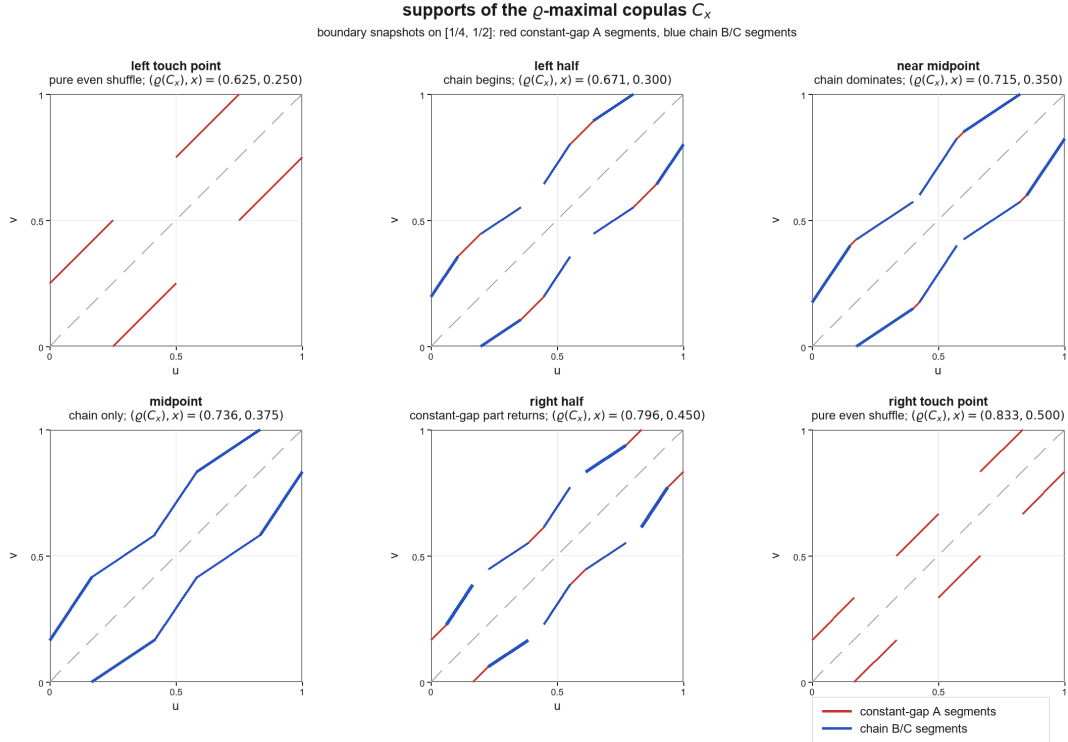


Figure 7: Six supports of $\pi_x = \nu_x + S_{\#}\nu_x$, the joint distributions associated with the ρ -maximal copulas C_x , along the boundary segment $x \in [\frac{1}{4}, \frac{1}{2}]$. Red denotes the images of the maps A_k^L or A_k^R , whereas blue denotes the images of B_k^L, C_k^L or B_k^R, C_k^R , as defined in (68)–(73), together with their reflections under S . On the red segments, $b - a$ is constant. At the two touch points $x = \frac{1}{4}$ and $x = \frac{1}{2}$, the supports are equidistant even shuffles. Toward the midpoint, the parameter v increases and the remaining mass $p = 1 - d_N v$ on the constant-difference segments decreases; see (62) and (63). At the midpoint $x = \frac{3}{8}$, we have $v = L - R = 1/d_N$ and hence $p = 0$, so only the blue segments remain.

Using (77), the coordinate images are

$$\begin{aligned}
(\text{proj}_1 \circ A_k^R)([0, pR]) &= [k\theta, k\theta + pR], \\
(\text{proj}_2 \circ A_k^R)([0, pR]) &= [k\theta + R, k\theta + R + pR], \\
(\text{proj}_1 \circ B_k^R)([0, Nv]) &= [k\theta + pR, k\theta + R], \\
(\text{proj}_2 \circ B_k^R)([0, Nv]) &= [k\theta + R + pR, (k+1)\theta], \\
(\text{proj}_1 \circ C_k^R)([0, Nv]) &= [k\theta + R + pR, (k+1)\theta], \\
(\text{proj}_2 \circ C_k^R)([0, Nv]) &= [(k+1)\theta + pR, (k+1)\theta + R].
\end{aligned}$$

Together with (78), this shows that all coordinates lie in $[0, 1]$, and thus all mappings take values in T .

To verify (76), set

$$\gamma_x^R := (\text{proj}_1)_\# \nu_x^R + (\text{proj}_2)_\# \nu_x^R.$$

We show that $\gamma_x^R = \lambda$.

We repeatedly use the following elementary change-of-variables fact: if J is an interval and $g(t) = a + ct$, with $c > 0$, then the push-forward of $w \lambda|_J$ under g has density w/c on $g(J)$.

For $k = 0, \dots, N-1$, decompose

$$[k\theta, (k+1)\theta] = J_{k,1} \cup J_{k,2} \cup J_{k,3} \cup J_{k,4},$$

where

$$\begin{aligned}
J_{k,1} &= [k\theta, k\theta + pR], \\
J_{k,2} &= [k\theta + pR, k\theta + R], \\
J_{k,3} &= [k\theta + R, k\theta + R + pR], \\
J_{k,4} &= [k\theta + R + pR, (k+1)\theta].
\end{aligned}$$

On $J_{k,1}$, the first coordinate of A_k^R in (71) has slope 1 and weight 1. Since, on $J_{k,1}$, only the first coordinate of A_k^R contributes, γ_x^R has density 1 on $J_{k,1}$.

On $J_{k,2}$, the first coordinate of B_k^R contributes density α_k . If $k \geq 1$, the second coordinate of C_{k-1}^R contributes density β_{k-1} . Since

$$\alpha_k + \beta_{k-1} = 1 - \frac{k}{N} + \frac{k}{N} = 1,$$

the total density is 1. For $k = 0$, only B_0^R contributes, and $\alpha_0 = 1$. Hence, γ_x^R has density 1 on $J_{k,2}$.

On $J_{k,3}$, only the second coordinate of A_k^R contributes. It has slope 1 and weight 1, and therefore γ_x^R has density 1 on $J_{k,3}$.

Finally, $J_{k,4}$ is covered by the second coordinate of B_k^R and the first coordinate of C_k^R . Both maps have slope

$$1 + \ell'_R(t) = 1 + \frac{1}{N} = \frac{N+1}{N}.$$

Hence their combined density equals

$$\frac{\alpha_k + \beta_k}{(N+1)/N} = \frac{1 - \frac{k}{N} + \frac{k+1}{N}}{(N+1)/N} = 1.$$

It remains to consider the terminal interval $[N\theta, 1]$. By (78),

$$[N\theta, 1] = [N\theta, N\theta + pR] \cup [N\theta + pR, N\theta + R] \cup [N\theta + R, 1].$$

The first and third intervals are covered by the two coordinate projections of A_N^R , each with density 1. The middle interval is covered by the second coordinate of C_{N-1}^R , whose weight is $\beta_{N-1} = 1$. Thus, γ_x^R has density 1 on all of $[0, 1]$, and therefore $(\text{proj}_1)_\# \nu_x^R + (\text{proj}_2)_\# \nu_x^R = \lambda$. follows.

Case 2 (the left-side part): Let $x \in I_N^L$. Since $d_N L = N + 1$, we have $pL = (1 - d_N v)L = L - (N + 1)v$, and hence

$$Nv + pL = L - v. \tag{79}$$

Moreover, using $\theta = 2L - v$, we have

$$N\theta + Nv = N(2L - v) + Nv = 2NL = 1. \quad (80)$$

The coordinate differences of the pairs appearing in ν_x^L are

$$\begin{aligned} \text{proj}_2(A_k^L(s)) - \text{proj}_1(A_k^L(s)) &= L, \\ \text{proj}_2(B_k^L(t)) - \text{proj}_1(B_k^L(t)) &= \ell_L(t) = L - v + \frac{t}{N}, \\ \text{proj}_2(C_k^L(t)) - \text{proj}_1(C_k^L(t)) &= \theta - \ell_L(t) = L - \frac{t}{N}. \end{aligned}$$

Thus all differences belong to the interval $[L - v, L]$ and are non-negative.

By (79), the coordinate images are

$$\begin{aligned} \text{proj}_1 \circ B_k^L([0, Nv]) &= [k\theta, k\theta + Nv], \\ \text{proj}_1 \circ A_k^L([0, pL]) &= [k\theta + Nv, k\theta + L - v], \\ \text{proj}_2 \circ B_k^L([0, Nv]) &= [k\theta + L - v, k\theta + L + Nv], \\ \text{proj}_1 \circ C_k^L([0, Nv]) &= [k\theta + L - v, k\theta + L + Nv], \\ \text{proj}_2 \circ A_k^L([0, pL]) &= [k\theta + L + Nv, (k+1)\theta], \\ \text{proj}_2 \circ C_k^L([0, Nv]) &= [(k+1)\theta, (k+1)\theta + Nv]. \end{aligned}$$

Together with (80), this shows that all coordinates lie in $[0, 1]$, and hence all mappings take values in T .

To verify (76), define

$$\gamma_x^L := (\text{proj}_1)_\# \nu_x^L + (\text{proj}_2)_\# \nu_x^L.$$

For $k = 0, \dots, N-1$, decompose

$$[k\theta, (k+1)\theta] = K_{k,1} \cup K_{k,2} \cup K_{k,3} \cup K_{k,4},$$

where

$$\begin{aligned} K_{k,1} &= [k\theta, k\theta + Nv], \\ K_{k,2} &= [k\theta + Nv, k\theta + L - v], \\ K_{k,3} &= [k\theta + L - v, k\theta + L + Nv], \\ K_{k,4} &= [k\theta + L + Nv, (k+1)\theta]. \end{aligned}$$

On $K_{k,1}$, the first coordinate of B_k^L contributes density α_k . For $k \geq 1$, the second coordinate of C_{k-1}^L contributes density β_{k-1} . Hence the total density is $\alpha_k + \beta_{k-1} = (1 - \frac{k}{N}) + \frac{k}{N} = 1$. For $k = 0$, only B_0^L contributes, and $\alpha_0 = 1$.

On $K_{k,2}$, the first coordinate of A_k^L has slope 1 and weight 1, and therefore contributes density 1. On $K_{k,3}$, the second coordinate of B_k^L and the first coordinate of C_k^L both have slope

$$1 + \ell'_L(t) = \frac{N+1}{N}.$$

Their combined density is

$$\frac{\alpha_k + \beta_k}{(N+1)/N} = 1.$$

On $K_{k,4}$, the second coordinate of A_k^L has slope 1 and weight 1, and hence contributes density 1. Finally, by (80), the remaining terminal interval is

$$[N\theta, 1] = [N\theta, N\theta + Nv].$$

It is covered by the second coordinate of C_{N-1}^L , whose weight is $\beta_{N-1} = 1$. Consequently, γ_x^L has density 1 on $[0, 1]$, and thus $(\text{proj}_1)_\# \nu_x^L + (\text{proj}_2)_\# \nu_x^L = \lambda$. $\square$

Next, we show that ν_x is feasible.

Lemma 4.4 (Feasibility of ν_x). *For every $x \in [-\frac{1}{2}, 1)$, the measure ν_x is feasible for $m = \frac{1-x}{3}$, i.e.,*

$$\int_T (b-a) \, d\nu_x(a,b) = \frac{m}{2}.$$

Proof. Recall the definition of the weights $\alpha_k = 1 - \frac{k}{N}$ and $\beta_k = \frac{k+1}{N}$ occurring in (65) and (66). Then we have

$$\sum_{k=0}^{N-1} \alpha_k = \sum_{k=0}^{N-1} \beta_k = \frac{N+1}{2}.$$

First, let $x \in I_N^R$. The coordinate differences are generated by A_k^R , B_k^R , and C_k^R are R , $\ell_R(t)$, and $\theta - \ell_R(t)$, respectively. Consequently,

$$\begin{aligned} \int_T (b-a) \, d\nu_x^R(a,b) &= \sum_{k=0}^N \int_0^{pR} R \, ds + \sum_{k=0}^{N-1} \alpha_k \int_0^{Nv} \ell_R(t) \, dt + \sum_{k=0}^{N-1} \beta_k \int_0^{Nv} (\theta - \ell_R(t)) \, dt \\ &= (N+1)pR^2 + \frac{N+1}{2} \int_0^{Nv} \theta \, dt = \frac{pR}{2} + \frac{d_N \theta v}{4}, \end{aligned}$$

where we use $(N+1)R = \frac{1}{2}$ and $d_N = 2N(N+1)$ by (60) and (62). Since, by (63), $p = 1 - d_N v$ and $\theta = 2R + v$, we obtain

$$\frac{pR}{2} + \frac{d_N \theta v}{4} = \frac{1}{2} \left(pR + \frac{d_N \theta v}{2} \right) = \frac{1}{2} \left(R - d_N R v + d_N R v + \frac{d_N v^2}{2} \right) = \frac{1}{2} \left(R + \frac{d_N v^2}{2} \right).$$

By definition of $v(x)$ on I_N^R in (63), it follows that $m(x) = R + \frac{d_N v^2}{2}$. Hence,

$$\int_T (b-a) \, d\nu_x^R(a,b) = \frac{m(x)}{2}.$$

Now let $x \in I_N^L$. The differences of the coordinates generated by A_k^L , B_k^L , and C_k^L are L , $\ell_L(t)$, and $\theta - \ell_L(t)$, respectively. Thus,

$$\begin{aligned} \int_T (b-a) \, d\nu_x^L(a,b) &= \sum_{k=0}^{N-1} \int_0^{pL} L \, ds + \sum_{k=0}^{N-1} \alpha_k \int_0^{Nv} \ell_L(t) \, dt + \sum_{k=0}^{N-1} \beta_k \int_0^{Nv} (\theta - \ell_L(t)) \, dt \\ &= NpL^2 + \frac{N+1}{2} \int_0^{Nv} \theta \, dt = \frac{pL}{2} + \frac{d_N \theta v}{4}, \end{aligned}$$

where we used $NL = \frac{1}{2}$. Since $p = 1 - d_N v$ and $\theta = 2L - v$, it follows that

$$\frac{pL}{2} + \frac{d_N \theta v}{4} = \frac{1}{2} \left(pL + \frac{d_N \theta v}{2} \right) = \frac{1}{2} \left(L - d_N L v + d_N L v - \frac{d_N v^2}{2} \right) = \frac{1}{2} \left(L - \frac{d_N v^2}{2} \right).$$

By definition of $v(x)$ on I_N^L , it follows that $m(x) = L - \frac{d_N v^2}{2}$. Therefore,

$$\int_T (b-a) \, d\nu_x^L(a,b) = \frac{m(x)}{2}.$$

This proves the assertion. $\square$

To calculate Spearman's rho and Spearman's footrule for C_x , we determine the distribution of $|U - V|$ for $(U, V) \sim C_x$. We denote by δ_y the Dirac measure in y and write $\lambda|_I$ for the restriction of Lebesgue measure to an interval I .

Lemma 4.5. *Consider the coupling π_x defined in (74). For $(U, V) \sim \pi_x$, we have*

$$\mathcal{L}_{\pi_x}(|U - V|) = \begin{cases} p\delta_R + d_N \lambda|_{[R, R+v]}, & \text{if } x \in I_N^R, \\ p\delta_L + d_N \lambda|_{[L-v, L]}, & \text{if } x \in I_N^L, \end{cases} \quad (81)$$

where R , L , I_N^R , I_N^L , d_N , p , and v are defined in (60)–(63).

Proof. Define the mappings

$$q : T \rightarrow [0, 1], \quad q(a, b) := b - a,$$

and

$$d : [0, 1]^2 \rightarrow [0, 1], \quad d(a, b) := |b - a|.$$

Since ν_x has support in T , we have $d = q$ on the support of ν_x . Moreover, $d \circ S = d$. Hence,

$$d_{\#}\pi_x = d_{\#}\nu_x + d_{\#}(S_{\#}\nu_x) = q_{\#}\nu_x + (d \circ S)_{\#}\nu_x = 2q_{\#}\nu_x. \quad (82)$$

Thus, $2q_{\#}\nu_x$ is the distribution of $|U - V|$. We determine this distribution separately for the two parts of I_N .

Case 1: $x \in I_N^R$. Recall that the coordinate differences of the pairs generated by A_k^R , B_k^R , and C_k^R are R , $\ell_R(t) = R + \frac{t}{N}$, and $\theta - \ell_R(t) = R + v - \frac{t}{N}$, respectively. Therefore, for every bounded Borel function $h : [0, 1] \rightarrow \mathbb{R}$,

$$\begin{aligned} \int_{[0,1]} h(r) \, d(q_{\#}\nu_x^R)(r) &= \sum_{k=0}^N \int_0^{pR} h(R) \, ds \\ &\quad + \sum_{k=0}^{N-1} \alpha_k \int_0^{Nv} h(R + \frac{t}{N}) \, dt \\ &\quad + \sum_{k=0}^{N-1} \beta_k \int_0^{Nv} h(R + v - \frac{t}{N}) \, dt. \end{aligned}$$

Using $(N+1)R = \frac{1}{2}$ and $\sum_{k=0}^{N-1} \alpha_k = \sum_{k=0}^{N-1} \beta_k = \frac{N+1}{2}$, together with the changes of variables $r = R + \frac{t}{N}$ and $r = R + v - \frac{t}{N}$, we obtain

$$\int_{[0,1]} h(r) \, d(q_{\#}\nu_x^R)(r) = \frac{p}{2}h(R) + N(N+1) \int_R^{R+v} h(r) \, dr = \frac{p}{2}h(R) + \frac{d_N}{2} \int_R^{R+v} h(r) \, dr.$$

Consequently,

$$q_{\#}\nu_x^R = \frac{p}{2}\delta_R + \frac{d_N}{2}\lambda|_{[R, R+v]}.$$

By (82), the distribution of $|U - V|$ is therefore

$$\mathcal{L}_{\pi_x}(|U - V|) = d_{\#}\pi_x = p\delta_R + d_N\lambda|_{[R, R+v]}.$$

Case 2: $x \in I_N^L$. Recall that the coordinate differences of the pairs generated by A_k^L , B_k^L , and C_k^L are L , $\ell_L(t) = L - v + \frac{t}{N}$, and $\theta - \ell_L(t) = L - \frac{t}{N}$, respectively. Thus, for every bounded Borel function h ,

$$\begin{aligned} \int_{[0,1]} h(r) \, d(q_{\#}\nu_x^L)(r) &= \sum_{k=0}^{N-1} \int_0^{pL} h(L) \, ds \\ &\quad + \sum_{k=0}^{N-1} \alpha_k \int_0^{Nv} h(L - v + \frac{t}{N}) \, dt \\ &\quad + \sum_{k=0}^{N-1} \beta_k \int_0^{Nv} h(L - \frac{t}{N}) \, dt. \end{aligned}$$

Since $NL = \frac{1}{2}$, the same change-of-variables argument as before gives $q_{\#}\nu_x^L = \frac{p}{2}\delta_L + \frac{d_N}{2}\lambda|_{[L-v, L]}$. Therefore,

$$\mathcal{L}_{\pi_x}(|U - V|) = d_{\#}\pi_x = p\delta_L + d_N\lambda|_{[L-v, L]}.$$

□

Proposition 4.6 (Spearman's ϱ and Spearman's footrule for C_x). *Let $x \in [-\frac{1}{2}, 1)$. For the copula C_x in (75), we have*

$$\begin{aligned} \phi(C_x) &= x \quad \text{and} \\ \varrho(C_x) &= \begin{cases} 1 - 6L(2m - L) - \frac{4(L-m)^{3/2}}{\sqrt{N(N+1)}}, & \text{for } x \in I_N^L, \\ 1 - 6R(2m - R) - \frac{4(m-R)^{3/2}}{\sqrt{N(N+1)}}, & \text{for } x \in I_N^R. \end{cases} \end{aligned}$$

At $x = 1$, one has $\phi(C_1) = \varrho(C_1) = 1$.

Proof. For $x = 1$, Definition 4.1 gives the comonotonicity copula, under which $U = V$ almost surely; in this case, the assertion follows from Lemma 3.2. Let now $x \in [-\frac{1}{2}, 1)$. The joint distribution of $(U, V) \sim C_x$ is by definition of the copula C_x in (75) the coupling $\pi_x = \nu_x + S_{\#}\nu_x$. To determine $\phi(C_x)$ and $\varrho(C_x)$, we need to determine the first and second moment of $|U - V|$.

Case 1: $x \in I_N^R$. For the first moment, we obtain from Lemma 4.5 that

$$\mathbb{E}|U - V| = pR + d_N \int_R^{R+v} r \, dr = (1 - d_N v)R + d_N \left(Rv + \frac{v^2}{2} \right) = R + \frac{d_N v^2}{2} = m,$$

where the last equality follows from the definition of $v(x)$ on I_N^R . Hence,

$$\phi(C_x) = 1 - 3\mathbb{E}|U - V| = 1 - 3m = x.$$

For the second moment, we obtain from Lemma 4.5 that

$$\begin{aligned} \mathbb{E}(U - V)^2 &= pR^2 + d_N \int_R^{R+v} r^2 \, dr \\ &= (1 - d_N v)R^2 + d_N \left(R^2 v + Rv^2 + \frac{v^3}{3} \right) \\ &= R^2 + d_N Rv^2 + \frac{d_N v^3}{3} = R(2m - R) + \frac{d_N v^3}{3}. \end{aligned}$$

Since $v = \sqrt{\frac{2(m-R)}{d_N}}$ and $d_N = 2N(N+1)$ (see (63) and (62)), we have

$$\frac{d_N v^3}{3} = \frac{2(m-R)^{3/2}}{3\sqrt{N(N+1)}}.$$

Using $\varrho(C_x) = 1 - 6\mathbb{E}(U - V)^2$, we conclude that

$$\varrho(C_x) = 1 - 6R(2m - R) - \frac{4(m-R)^{3/2}}{\sqrt{N(N+1)}}.$$

Case 2: $x \in I_N^L$. For the first moment, we obtain from Lemma 4.5 that

$$\mathbb{E}|U - V| = pL + d_N \int_{L-v}^L r \, dr = (1 - d_N v)L + d_N \left(Lv - \frac{v^2}{2} \right) = L - \frac{d_N v^2}{2} = m.$$

Hence,

$$\phi(C_x) = 1 - 3m = x.$$

For the second moment, we again use Lemma 4.5 and obtain

$$\begin{aligned} \mathbb{E}(U - V)^2 &= pL^2 + d_N \int_{L-v}^L r^2 \, dr \\ &= (1 - d_N v)L^2 + d_N \left(L^2 v - Lv^2 + \frac{v^3}{3} \right) \\ &= L^2 - d_N Lv^2 + \frac{d_N v^3}{3} = L(2m - L) + \frac{d_N v^3}{3}. \end{aligned}$$

Since $v = \sqrt{\frac{2(L-m)}{d_N}}$, it follows that

$$\frac{d_N v^3}{3} = \frac{2(L-m)^{3/2}}{3\sqrt{N(N+1)}}.$$

This yields

$$\varrho(C_x) = 1 - 6L(2m - L) - \frac{4(L-m)^{3/2}}{\sqrt{N(N+1)}},$$

which proves the assertion. $\square$

5 Dual potential and optimality

Up to this point, for every $x \in [-\frac{1}{2}, 1)$, we have constructed a copula C_x that satisfies $\phi(C_x) = x$, and we have derived an explicit expression for $\varrho(C_x)$. By definition of $\bar{\varrho}(x)$, we necessarily have

$$\varrho(C_x) \leq \bar{\varrho}(x).$$

To prove equality, it remains to show that every copula C satisfying $\phi(C) = x$ also fulfills

$$\varrho(C) \leq \varrho(C_x).$$

In this section, we establish the upper bound by constructing a dual feasible potential whose contact set contains the support of ν_x . In Corollary 3.7, we established the contact-set optimality criterion for the centered dual potential. We now construct, for each x , an explicit dual potential matching the candidate measure ν_x . It is more convenient to work with the non-centered potential g_x , satisfying

$$g_x(a) + g_x(b) + \theta(b-a) \leq (b-a)^2 \quad \text{for all } (a, b) \in T; \quad (83)$$

recall $T = \{(a, b) \in [0, 1]^2 : a \leq b\}$. The corresponding centered potential is

$$f_x = g_x + \theta m/2.$$

Hence, (f_x, θ) is feasible for the centered dual problem if and only if g_x satisfies (83).

Now, fix $x \in [-\frac{1}{2}, 1)$, choose $N \in \mathbb{N}$ such that $x \in I_N$, and recall the parameters

$$m = m(x), \quad v = v(x), \quad p = p(x), \quad \theta = \theta(x)$$

defined in (58) and (63). The definitions imply $0 \leq v \leq L - R$. If $x \in I_N^R$, then $\theta = 2R + v > v$. If $x \in I_N^L$, then $\theta - v = 2(L - v) \geq 2R > 0$. Thus, in both cases, $\theta > v \geq 0$, and in particular $\theta > 0$. Define the modified cost function

$$k_x(r) := r^2 - \theta r, \quad r \in \mathbb{R}. \quad (84)$$

Then (83) is equivalent to

$$g_x(a) + g_x(b) \leq k_x(b-a) \quad \text{for all } (a, b) \in T. \quad (85)$$

The choice of θ is adapted to the construction of ν_x . Specifically, we have

$$k_x(\theta - r) = k_x(r) \quad \text{for all } r \in \mathbb{R}. \quad (86)$$

We shall construct g_x such that

$$g_x(a) + g_x(b) = k_x(b-a) \quad \nu_x\text{-almost surely}, \quad (87)$$

while (85) holds globally on T . Then, by Corollary 3.7, the two properties (85) and (87) imply that ν_x is optimal. Recall $v = v(x)$, $p = p(x)$, and $\theta = \theta(x)$ in (63).

Definition 5.1 (Optimal dual potential). For $x \in [-\frac{1}{2}, 1)$, define g_x first on one period $[0, \theta]$ and then extend it θ -periodically to $\mathbb{R}$; we use the same symbol g_x for its restriction to $[0, 1]$. If $x \in I_N^R$, put $c_x^R := \frac{k_x(R) - vpR}{2}$ and define

$$g_x(u) := \begin{cases} c_x^R + vu, & 0 \leq u < pR, \\ c_x^R + vpR + v(u - pR) - \frac{(u - pR)^2}{N}, & pR \leq u < R, \\ c_x^R + vpR - v(u - R), & R \leq u < R + pR, \\ c_x^R - v(u - R - pR) + \frac{(u - R - pR)^2}{N + 1}, & R + pR \leq u \leq \theta. \end{cases} \quad (88)$$

If $x \in I_N^L$, put $c_x^L := \frac{k_x(L) + vpL}{2}$ and define

$$g_x(u) := \begin{cases} c_x^L + vu - \frac{u^2}{N}, & 0 \leq u \leq Nv, \\ c_x^L - v(u - Nv), & Nv \leq u \leq L - v, \\ c_x^L - vpL - v(u - L + v) + \frac{(u - L + v)^2}{N + 1}, & L - v \leq u \leq L + Nv, \\ c_x^L - vpL + v(u - L - Nv), & L + Nv \leq u \leq \theta. \end{cases} \quad (89)$$

Remark 5.2. The choice of c_x^R is equivalently characterized by

$$g_x(0) + g_x(R) = k_x(R). \quad (90)$$

and c_x^L is determined by

$$g_x(Nv) + g_x(Nv + L) = k_x(L). \quad (91)$$

Some properties of g_x are given in the following Lemma; see also Figure 6.

Lemma 5.3. The function g_x in Definition 5.1 is θ -periodic and continuously differentiable with derivative

$$g'_x(u) = \begin{cases} v, & 0 \leq u < pR, \\ v - \frac{2}{N}(u - pR), & pR \leq u < R, \\ -v, & R \leq u < R + pR, \\ -v + \frac{2}{N + 1}(u - R - pR), & R + pR \leq u \leq \theta. \end{cases} \quad (92)$$

on $[0, \theta]$ for $x \in I_N^R$, and

$$g'_x(u) = \begin{cases} v - \frac{2}{N}u, & 0 \leq u < Nv, \\ -v, & Nv \leq u < L - v, \\ -v + \frac{2}{N + 1}(u - L + v), & L - v \leq u < L + Nv, \\ v, & L + Nv \leq u \leq \theta. \end{cases} \quad (93)$$

on $[0, \theta]$ for $x \in I_N^L$.

Proof. For $x \in I_N^R$, recall that $pR + Nv = R$ and $\theta = 2R + v$ by (77) and (63). Straightforward calculations show that g'_x is continuous and piecewise linear with the expressions in (92) and (93). Thus, g_x is continuously differentiable. To show that g_x is θ -periodic, consider

$$\begin{aligned} \int_0^\theta g'_x(u) \, du &= \int_0^{pR} v \, du + \int_{pR}^R \left(v - \frac{2}{N}(u - pR) \right) \, du \\ &\quad - \int_R^{R+pR} v \, du + \int_{R+pR}^\theta \left(-v + \frac{2}{N+1}(u - R - pR) \right) \, du. \end{aligned}$$

The second integral vanishes because its integrand decreases linearly from v to $-v$ on an interval of length Nv . Similarly, the fourth integral vanishes because its integrand increases linearly from $-v$ to v on an interval of length $(N+1)v$. The first and third integrals cancel. Hence, $\int_0^\theta g'_x(u) du = 0$.

For $x \in I_N^L$, recall that $Nv + pL = L - v$ and $\theta = 2L - v$ from (79) and (63). Again, straightforward calculations show that g'_x is continuous and piecewise linear so that g_x is continuously differentiable. To show that g_x is θ -periodic, consider

$$\begin{aligned} \int_0^\theta g'_x(u) du &= \int_0^{Nv} \left(v - \frac{2}{N}u \right) du - \int_{Nv}^{L-v} v du \\ &\quad + \int_{L-v}^{L+Nv} \left(-v + \frac{2}{N+1}(u - L + v) \right) du + \int_{L+Nv}^\theta v du. \end{aligned}$$

The first and third integrals vanish because their integrands are linear and have endpoint values $v, -v$ and $-v, v$, respectively. The second and fourth intervals both have length

$$L - v - Nv = \theta - L - Nv = pL.$$

Hence, the second and fourth integrals cancel, and therefore $\int_0^\theta g'_x(u) du = 0$. $\square$

In the following result, we verify that ν_x in (64) is concentrated on the set

$$\{(a, b) \in T : g_x(a) + g_x(b) = k_x(b - a)\}.$$

For $f_x = g_x + \theta m/2$, this is equivalent to concentration on the centered contact set $\Gamma_{f_x, \theta}$.

Lemma 5.4 (Equality on the support of ν_x). *For every $x \in [-\frac{1}{2}, 1)$, we have*

$$g_x(a) + g_x(b) = k_x(b - a) \quad \text{for } \nu_x\text{-almost every } (a, b) \in T. \quad (94)$$

Proof. Recall the definitions of the mappings $A_k^{L/R}$, $B_k^{L/R}$, and $C_k^{L/R}$. By definition, the measure ν_x is concentrated on the images of these mappings. We verify (94) separately on each of the corresponding families of line segments.

By Lemma 5.3, the function g_x is θ -periodic. Moreover, every mapping with index k is obtained from the corresponding mapping with index 0 by translating both coordinates by $k\theta$. For instance,

$$A_k^{L/R}(s) = A_0^{L/R}(s) + (k\theta, k\theta),$$

and analogously for $B_k^{L/R}$ and $C_k^{L/R}$. Since a common translation does not change the coordinate difference and

$$g_x(u + k\theta) = g_x(u) \quad \text{for all } u \in \mathbb{R},$$

it is sufficient to consider the case $k = 0$.

Case 1: Let $x \in I_N^R$. For the A_0^R -segments, define

$$H_A^R(s) := g_x(s) + g_x(s + R) - k_x(R), \quad 0 \leq s \leq pR.$$

Since $s \in [0, pR]$ and $s + R \in [R, R + pR]$, it follows from (92) that $g'_x(s) = v$ and $g'_x(s + R) = -v$. Consequently,

$$(H_A^R)'(s) = g'_x(s) + g'_x(s + R) = 0.$$

Thus, H_A^R is constant. By the normalization condition (90), $H_A^R(0) = g_x(0) + g_x(R) - k_x(R) = 0$. Hence,

$$g_x(s) + g_x(s + R) = k_x(R) \quad (95)$$

for every $s \in [0, pR]$. Since $A_0^R(s) = (s, s + R)$ has coordinate difference R , this proves (94) on the A_0^R -segments.

Next, consider the B_0^R -segments. Set

$$a(t) := pR + t, \quad r(t) := \ell_R(t) = R + \frac{t}{N}, \quad 0 \leq t \leq Nv,$$

and define

$$H_B^R(t) := g_x(a(t)) + g_x(a(t) + r(t)) - k_x(r(t)).$$

Since $\theta = 2R + v$ by (63), we have

$$k'_x(r(t)) = 2r(t) - \theta = -v + \frac{2t}{N}.$$

Moreover, by (92),

$$g'_x(a(t)) = v - \frac{2t}{N} = -k'_x(r(t)).$$

Furthermore, $a(t) + r(t) = R + pR + \frac{N+1}{N}t$, and hence

$$g'_x(a(t) + r(t)) = -v + \frac{2t}{N} = k'_x(r(t)).$$

Using $a'(t) = 1$ and $r'(t) = \frac{1}{N}$, we obtain

$$\begin{aligned} (H_B^R)'(t) &= g'_x(a(t)) + g'_x(a(t) + r(t)) \left(1 + \frac{1}{N}\right) - k'_x(r(t)) \frac{1}{N} \\ &= -k'_x(r(t)) + k'_x(r(t)) \left(1 + \frac{1}{N}\right) - k'_x(r(t)) \frac{1}{N} = 0. \end{aligned}$$

Thus, H_B^R is constant. At $t = 0$, we have $a(0) = pR$ and $r(0) = R$, and therefore

$$B_0^R(0) = (pR, pR + R) = A_0^R(pR).$$

The equality (95) already proved on the A_0^R -segment consequently gives

$$H_B^R(0) = g_x(pR) + g_x(pR + R) - k_x(R) = 0.$$

Hence,

$$g_x(a(t)) + g_x(a(t) + r(t)) = k_x(r(t)) \quad (96)$$

for every $t \in [0, Nv]$. Since $B_0^R(t) = (a(t), a(t) + r(t))$ and the coordinate difference of this pair is $r(t)$, this proves (94) on the B_k^R -segments.

Finally, the pair generated by C_0^R is $C_0^R(t) = (a(t) + r(t), \theta + a(t))$. By the θ -periodicity of g_x ,

$$g_x(\theta + a(t)) = g_x(a(t)).$$

Using the equality (96) already established on the B_0^R -segment, we therefore obtain

$$g_x(a(t) + r(t)) + g_x(\theta + a(t)) = g_x(a(t) + r(t)) + g_x(a(t)) = k_x(r(t)).$$

The coordinate difference of the C_0^R -pair is

$$(\theta + a(t)) - (a(t) + r(t)) = \theta - r(t).$$

Since $k_x(\theta - r) = k_x(r)$, by (86), it follows that

$$g_x(a(t) + r(t)) + g_x(\theta + a(t)) = k_x(\theta - r(t)).$$

This proves (94) on the C_k^R -segments.

Case 2: Let $x \in I_N^L$. For the A_0^L -segments, define

$$H_A^L(s) := g_x(Nv + s) + g_x(Nv + s + L) - k_x(L), \quad 0 \leq s \leq pL.$$

Since $Nv + pL = L - v$ by (79), we have $Nv + s \in [Nv, L - v]$ and $Nv + s + L \in [L + Nv, \theta]$. It follows from (93) that $g'_x(Nv + s) = -v$ and $g'_x(Nv + s + L) = v$. Consequently,

$$(H_A^L)'(s) = g'_x(Nv + s) + g'_x(Nv + s + L) = 0.$$

Thus, H_A^L is constant. By the normalization condition (91),

$$H_A^L(0) = g_x(Nv) + g_x(Nv + L) - k_x(L) = 0.$$

Hence,

$$g_x(Nv + s) + g_x(Nv + s + L) = k_x(L) \quad (97)$$

for every $s \in [0, pL]$. Since $A_0^L(s) = (Nv + s, Nv + s + L)$ has coordinate difference L , this proves (94) on the A_k^L -segments.

Next, consider the B_0^L -segments. Set

$$a(t) := t, \quad r(t) := \ell_L(t) = L - v + \frac{t}{N}, \quad 0 \leq t \leq Nv,$$

and define

$$H_B^L(t) := g_x(a(t)) + g_x(a(t) + r(t)) - k_x(r(t)).$$

Since $\theta = 2L - v$, we have

$$k'_x(r(t)) = 2r(t) - \theta = -v + \frac{2t}{N}.$$

Moreover, by (93),

$$g'_x(a(t)) = v - \frac{2t}{N} = -k'_x(r(t)).$$

Furthermore, $a(t) + r(t) = L - v + \frac{N+1}{N}t$, and hence

$$g'_x(a(t) + r(t)) = -v + \frac{2t}{N} = k'_x(r(t)).$$

Using $a'(t) = 1$ and $r'(t) = \frac{1}{N}$, we obtain

$$\begin{aligned} (H_B^L)'(t) &= g'_x(a(t)) + g'_x(a(t) + r(t)) \left(1 + \frac{1}{N}\right) - k'_x(r(t)) \frac{1}{N} \\ &= -k'_x(r(t)) + k'_x(r(t)) \left(1 + \frac{1}{N}\right) - k'_x(r(t)) \frac{1}{N} = 0. \end{aligned}$$

Thus, H_B^L is constant. At $t = Nv$, we have $a(Nv) = Nv$ and $r(Nv) = L$ and therefore

$$B_0^L(Nv) = (Nv, Nv + L) = A_0^L(0).$$

The equality (97) already proved on the A_0^L -segment consequently gives

$$H_B^L(Nv) = g_x(Nv) + g_x(Nv + L) - k_x(L) = 0.$$

Hence,

$$g_x(a(t)) + g_x(a(t) + r(t)) = k_x(r(t)) \quad (98)$$

for every $t \in [0, Nv]$. Since $B_0^L(t) = (a(t), a(t) + r(t))$ and the coordinate difference of this pair is $r(t)$, this proves (94) on the B_k^L -segments.

Finally, the pair generated by C_0^L is

$$C_0^L(t) = (a(t) + r(t), \theta + a(t)).$$

By the θ -periodicity of g_x ,

$$g_x(\theta + a(t)) = g_x(a(t)).$$

Using the equality in (98) already established on the B_0^L -segment, we therefore obtain

$$g_x(a(t) + r(t)) + g_x(\theta + a(t)) = g_x(a(t) + r(t)) + g_x(a(t)) = k_x(r(t)).$$

The coordinate difference of the C_0^L -pair is

$$(\theta + a(t)) - (a(t) + r(t)) = \theta - r(t).$$

Since $k_x(\theta - r) = k_x(r)$ by (86), it follows that

$$g_x(a(t) + r(t)) + g_x(\theta + a(t)) = k_x(\theta - r(t)),$$

which proves (94) on the C_k^L -segments.

Since ν_x is concentrated on the images of the mappings $A_k^{L/R}$, $B_k^{L/R}$, and $C_k^{L/R}$, the equality (94) holds ν_x -almost surely. $\square$

Next we show dual feasibility.

Lemma 5.5 (Global dual feasibility). *For every $x \in [-\frac{1}{2}, 1)$, the candidate potential g_x satisfies*

$$g_x(a) + g_x(b) \leq k_x(b - a) \quad \text{for all } (a, b) \in T. \quad (99)$$

Consequently, the centered potential

$$f_x(u) := g_x(u) + \frac{\theta m}{2}, \quad u \in [0, 1], \quad (100)$$

satisfies $f_x(a) + f_x(b) + \theta((b - a) - m) \leq (b - a)^2$ for all $(a, b) \in T$, and hence (f_x, θ) is feasible.

Proof. Extend g_x and g'_x θ -periodically to $\mathbb{R}$. For $u \in \mathbb{R}$ and $r \geq 0$, define

$$G_x(u, r) := k_x(r) - g_x(u) - g_x(u + r). \quad (101)$$

It suffices to prove that

$$G_x(u, r) \geq 0 \quad \text{for all } u \in \mathbb{R}, r \geq 0. \quad (102)$$

We first reduce the problem to $r \in [0, \theta]$. Write $r = r_0 + n\theta$ and $r_0 \in [0, \theta)$ for $n \in \mathbb{N}_0$. By the θ -periodicity of g_x (Lemma 5.3), we have $g_x(u + r) = g_x(u + r_0)$. Moreover, by definition of k_x in (84),

$$\begin{aligned} G_x(u, r_0 + n\theta) - G_x(u, r_0) &= k_x(r_0 + n\theta) - k_x(r_0) \\ &= (r_0 + n\theta)^2 - \theta(r_0 + n\theta) - r_0^2 + \theta r_0 = n\theta(2r_0 + (n - 1)\theta) \geq 0. \end{aligned} \quad (103)$$

Consequently, $G_x(u, r) \geq G_x(u, r_0)$ and it is enough to establish (102) for $0 \leq r \leq \theta$.

Fix $u \in \mathbb{R}$. Since g'_x is continuous and piecewise linear, the function $r \mapsto G_x(u, r)$ is continuously differentiable and piecewise twice differentiable. Its first derivative is

$$\partial_r G_x(u, r) = 2r - \theta - g'_x(u + r). \quad (104)$$

The slopes of g'_x belong to the set $\left\{0, -\frac{2}{N}, \frac{2}{N+1}\right\}$. Hence, wherever the second derivative exists,

$$\partial_r^2 G_x(u, r) = 2 - g''_x(u + r) \geq 2 - \frac{2}{N+1} > 0.$$

It follows that $r \mapsto G_x(u, r)$ is strictly convex on $[0, \theta]$. Define

$$r_*(u) := \frac{\theta - g'_x(u)}{2}. \quad (105)$$

Since, as a consequence of the representation of g'_x in (92) and (93), using the definition of θ in (63), one can show that

$$|g'_x(u)| \leq v < \theta \quad (106)$$

which yields $r_*(u) \in (0, \theta)$.

We next prove the identity

$$g'_x(u + r_*(u)) = -g'_x(u), \quad (107)$$

Both g'_x and, by (105), r_* are θ -periodic. Given $u \in \mathbb{R}$, choose $u_0 \in [0, \theta)$ with $u - u_0 \in \theta\mathbb{Z}$. Then $g'_x(u) = g'_x(u_0)$, $r_*(u) = r_*(u_0)$, and

$$g'_x(u + r_*(u)) = g'_x(u_0 + r_*(u_0)).$$

It therefore suffices to prove (107) for $u \in [0, \theta]$, which is done by the following case split.

Case 1: $x \in I_N^R$. Recall from (63) and (77) that

$$\theta = 2R + v \quad \text{and} \quad R - pR = Nv, \quad \text{so} \quad \theta - R - pR = (N + 1)v.$$

If $u \in [0, pR]$, then $g'_x(u) = v$, and hence $r_*(u) = R$. By (92), this gives

$$g'_x(u + r_*(u)) = -v = -g'_x(u).$$

If $u = pR + s$, $0 \leq s \leq Nv$, then by (92) and (105) $g'_x(u) = v - \frac{2s}{N}$ and $r_*(u) = R + \frac{s}{N}$. Thus,

$$u + r_*(u) = R + pR + \frac{N+1}{N}s.$$

This point belongs to $[R + pR, \theta]$, and therefore

$$g'_x(u + r_*(u)) = -v + \frac{2}{N+1} \left(\frac{N+1}{N}s \right) = -v + \frac{2s}{N} = -g'_x(u).$$

If $u = R + s$, $0 \leq s \leq pR$, then $g'_x(u) = -v$, and hence $r_*(u) = R + v$. Since $u + r_*(u) = \theta + s$, periodicity gives

$$g'_x(u + r_*(u)) = g'_x(s) = v = -g'_x(u).$$

Finally, if $u = R + pR + s$, $0 \leq s \leq (N+1)v$. Then $g'_x(u) = -v + \frac{2s}{N+1}$ and $r_*(u) = R + v - \frac{s}{N+1}$. Consequently,

$$u + r_*(u) = \theta + pR + \frac{N}{N+1}s.$$

By periodicity of g'_x ,

$$g'_x(u + r_*(u)) = g'_x \left(pR + \frac{N}{N+1}s \right).$$

Since $0 \leq \frac{N}{N+1}s \leq Nv$, the argument above belongs to $[pR, R]$, and thus

$$g'_x(u + r_*(u)) = v - \frac{2}{N} \left(\frac{N}{N+1}s \right) = v - \frac{2s}{N+1} = -g'_x(u).$$

This proves (107) in the right-hand case.

Case 2: $x \in I_N^L$. Recall that

$$\theta = 2L - v \quad \text{and} \quad L - v - Nv = pL, \quad \text{so} \quad \theta - L - Nv = pL.$$

If $u = s$, $0 \leq s \leq Nv$, then, by (93) and (105), $g'_x(u) = v - \frac{2s}{N}$ and $r_*(u) = L - v + \frac{s}{N}$. Hence,

$$u + r_*(u) = L - v + \frac{N+1}{N}s,$$

which belongs to $[L - v, L + Nv]$. Therefore,

$$g'_x(u + r_*(u)) = -v + \frac{2}{N+1} \left(\frac{N+1}{N}s \right) = -v + \frac{2s}{N} = -g'_x(u).$$

If $u = Nv + s$, $0 \leq s \leq pL$, then $g'_x(u) = -v$, and hence $r_*(u) = L$. Since

$$u + r_*(u) = L + Nv + s \in [L + Nv, \theta],$$

we obtain

$$g'_x(u + r_*(u)) = v = -g'_x(u).$$

If $u = L - v + s$, $0 \leq s \leq (N+1)v$, then $g'_x(u) = -v + \frac{2s}{N+1}$ and $r_*(u) = L - \frac{s}{N+1}$. Thus,

$$u + r_*(u) = \theta + \frac{N}{N+1}s.$$

By periodicity of g'_x ,

$$g'_x(u + r_*(u)) = g'_x\left(\frac{N}{N+1}s\right).$$

Since $0 \leq \frac{N}{N+1}s \leq Nv$, we obtain

$$g'_x(u + r_*(u)) = v - \frac{2}{N} \left(\frac{N}{N+1}s \right) = v - \frac{2s}{N+1} = -g'_x(u).$$

Finally, if $u = L + Nv + s$, $0 \leq s \leq pL$, then $g'_x(u) = v$, and hence $r_*(u) = L - v$. Since

$$u + r_*(u) = \theta + Nv + s,$$

periodicity gives

$$g'_x(u + r_*(u)) = g'_x(Nv + s) = -v = -g'_x(u).$$

Thus, (107) also holds in the left-hand case.

Now, using (104), (105), and (107), we obtain

$$\partial_r G_x(u, r_*(u)) = 2r_*(u) - \theta - g'_x(u + r_*(u)) = -g'_x(u) + g'_x(u) = 0.$$

By strict convexity, $r_*(u)$ is the unique minimizer of $r \mapsto G_x(u, r)$ on $[0, \theta]$.

It remains to determine the minimum value. Define

$$H_x(u) := G_x(u, r_*(u)).$$

Since g'_x is continuous and piecewise linear, the function r_* is Lipschitz continuous. Consequently, H_x is locally Lipschitz and hence absolutely continuous on compact intervals. At every point at which the relevant derivatives exist, the chain rule gives

$$H'_x(u) = k'_x(r_*(u))r'_*(u) - g'_x(u) - g'_x(u + r_*(u))(1 + r'_*(u)).$$

By the definition of k_x and r_* in (84) and (105),

$$k'_x(r_*(u)) = 2r_*(u) - \theta = -g'_x(u).$$

Together with (107), this yields

$$H'_x(u) = -g'_x(u)r'_*(u) - g'_x(u) + g'_x(u)(1 + r'_*(u)) = 0$$

almost everywhere. Hence, H_x is constant.

If $x \in I_N^R$, then $g'_x(0) = v$ and $r_*(0) = \frac{\theta-v}{2} = R$. By the normalization condition (90),

$$H_x(0) = k_x(R) - g_x(0) - g_x(R) = 0.$$

If $x \in I_N^L$, then $g'_x(Nv) = -v$ and $r_*(Nv) = \frac{\theta+v}{2} = L$. By the normalization condition (91),

$$H_x(Nv) = k_x(L) - g_x(Nv) - g_x(Nv + L) = 0.$$

Thus, in both cases, $H_x(u) = 0$ for all $u \in \mathbb{R}$. Since $r_*(u)$ is the minimizer of $r \mapsto G_x(u, r)$, we obtain

$$G_x(u, r) \geq G_x(u, r_*(u)) = 0 \tag{108}$$

for every $u \in \mathbb{R}$ and every $r \in [0, \theta]$. By the initial reduction, the same inequality holds for every $r \geq 0$.

Finally, let $(a, b) \in T$. Setting $u = a$ and $r = b - a$, we obtain

$$0 \leq G_x(a, b - a) = k_x(b - a) - g_x(a) - g_x(b),$$

which proves (99).

Since $k_x(b - a) = (b - a)^2 - \theta(b - a)$, we have

$$g_x(a) + g_x(b) + \theta(b - a) \leq (b - a)^2.$$

Since $f_x(u) = g_x(u) + \frac{\theta m}{2}$ by (100), this is equivalent to

$$f_x(a) + f_x(b) + \theta((b - a) - m) \leq (b - a)^2.$$

Hence, (f_x, θ) is feasible. $\square$

For proving uniqueness of the optimal copula C_x in Proposition 5.8, we use the following elementary lemma. It states that a finite signed measure cannot be invariant, up to sign, under a map that shifts every point by a strictly positive amount, unless the measure is identically zero.

Lemma 5.6. *Let $A \subseteq [0, 1]$ be Borel and let $\tau : A \rightarrow [0, 1]$ be Borel. Suppose that, for some $q > 0$, $\tau(a) \geq a + q$ for every $a \in A$. If a finite signed Borel measure μ on A satisfies*

$$\mu = \varepsilon \tau_{\#} \mu \quad \text{as measures on } [0, 1]$$

for some $\varepsilon \in \{-1, 1\}$, where μ is extended by zero outside A , then $\mu = 0$.

Proof. Extend μ by zero on A^c to a finite signed Borel measure on $[0, 1]$. The identity $\mu = \varepsilon \tau_{\#} \mu$ means that, for every Borel set $B \subseteq [0, 1]$,

$$\mu(B) = \varepsilon \mu(\tau^{-1}(B)). \quad (109)$$

Choose $N \in \mathbb{N}$ minimal such that $Nq > 1$, and define

$$\begin{aligned} E_j &:= [jq, (j+1)q), \quad j = 0, \dots, N-2, \\ E_{N-1} &:= [(N-1)q, 1]. \end{aligned}$$

Then $(E_j)_{j=0}^{N-1}$ is a partition of $[0, 1]$.

We show inductively that the restriction of μ to each E_j vanishes. For $j = 0$, let $B \subseteq E_0$ be Borel. If $a \in A$, then

$$\tau(a) \geq a + q \geq q,$$

and hence $\tau(a) \notin E_0$. Thus $\tau^{-1}(B) = \emptyset$, and (109) yields $\mu(B) = 0$. Consequently, $\mu|_{E_0} = 0$.

Now let $j \in \{1, \dots, N-1\}$ and suppose that $\mu|_{E_0 \cup \dots \cup E_{j-1}} = 0$. For a Borel set $B \subseteq E_j$ and $a \in \tau^{-1}(B)$, we have

$$\tau(a) \in E_j \quad \text{and} \quad a \leq \tau(a) - q.$$

If $j \leq N-2$, then $\tau(a) < (j+1)q$, and therefore $a < jq$.

For $j = N-1$, the minimality of N gives $Nq > 1$, and hence

$$a \leq 1 - q < (N-1)q.$$

Thus, in either case,

$$\tau^{-1}(B) \subseteq E_0 \cup \dots \cup E_{j-1}.$$

By the induction hypothesis, $\mu(\tau^{-1}(B)) = 0$, and (109) implies $\mu(B) = 0$. Hence $\mu|_{E_j} = 0$.

We conclude, that by induction, μ vanishes on every E_j , and since these sets partition $[0, 1]$, we must have $\mu = 0$. $\square$

We make use of the following characterization of the contact set for the centered dual potentials constructed in Lemma 5.5. Recall $\Gamma_{f_x, \theta} = \{(a, b) \in T : f_x(a) + f_x(b) + \theta((b-a) - m) = (b-a)^2\}$.

Lemma 5.7 (Characterization of the contact set of (f_x, θ)). *Let $x < 1$, and let (f_x, θ) be the feasible dual potential in (100). Define*

$$r_x(a) := \frac{\theta - g'_x(a)}{2}, \quad \tau_x(a) := a + r_x(a), \quad \text{and} \quad A_x := \{a \in [0, 1] : \tau_x(a) \leq 1\}. \quad (110)$$

Then the contact set of the dual potential (f_x, θ) is given by

$$\Gamma_{f_x, \theta} = \{(a, \tau_x(a)) : a \in A_x\}.$$

In particular, for every $a \in [0, 1]$, there exists at most one $b \geq a$ such that $(a, b) \in \Gamma_{f_x, \theta}$.

Proof. Recall (100) that $f_x = g_x + \frac{\theta m}{2}$. Hence, writing $r = b - a$, the slack in the constraint of the dual problem (12) is

$$(b-a)^2 - f_x(a) - f_x(b) - \theta((b-a) - m) = k_x(r) - g_x(a) - g_x(a+r) = G_x(a, r);$$

where the first equality follows from definition of f_x and k_x in (100) and (84). The last equality is by definition of G_x in (101). Consequently,

$$(a, b) \in \Gamma_{f_x, \theta} \iff G_x(a, b - a) = 0. \quad (111)$$

Notice that r_x in (110) coincides with the function r_* introduced in (105). By the strict convexity argument following (104), for every fixed a , the map $r \mapsto G_x(a, r)$ is strictly convex on $[0, \theta]$. Moreover, (104), (105), and (107) imply that $r_x(a) \in (0, \theta)$ is its unique minimizer, while (108) gives

$$G_x(a, r) = 0 \iff r = r_x(a), \quad r \in [0, \theta].$$

It remains to exclude further zeros for $r > \theta$. Write

$$r = r_0 + n\theta, \quad r_0 \in [0, \theta), \quad n \in \mathbb{N}_0.$$

By (103),

$$G_x(a, r) - G_x(a, r_0) = n\theta(2r_0 + (n-1)\theta).$$

For $n \geq 1$, the right-hand side of the above equation is strictly positive except when $(n, r_0) = (1, 0)$. In this exceptional case,

$$G_x(a, \theta) = G_x(a, 0) > 0,$$

because the unique zero of $G_x(a, \cdot)$ on $[0, \theta]$ is the interior point $r_x(a)$. Therefore,

$$G_x(a, r) = 0 \iff r = r_x(a)$$

throughout the admissible range $0 \leq r \leq 1 - a$.

Combining this with (111), equality in the dual constraint holds precisely when

$$b = a + r_x(a) = \tau_x(a).$$

The condition $b \leq 1$ is equivalent to $a \in A_x$, and hence $\Gamma_{f_x, \theta} = \{(a, \tau_x(a)) : a \in A_x\}$. $\square$

Proposition 5.8 (Optimality and uniqueness of the copulas C_x). *For every $x \in [-\frac{1}{2}, 1]$, the copula C_x in (75) uniquely maximizes Spearman's ϱ among all copulas C satisfying $\phi(C) = x$. In particular,*

$$\varrho(C_x) = \bar{\varrho}(x).$$

Proof. We first show optimality of C_x for $x \in [-\frac{1}{2}, 1)$. Therefore, recall $m = m(x) = \frac{1-x}{3}$ and $\theta = \theta(x)$ due to (58) and (63). By Lemma 4.3 and Lemma 4.4, the measure ν_x defined in (64) satisfies

$$\nu_x \in \Pi_T \quad \text{and} \quad \int_T (b - a) \, d\nu_x(a, b) = \frac{m}{2}.$$

Consider the centered potential $f_x(u) = g_x(u) + \frac{\theta m}{2}$ for $u \in [0, 1]$. Then, by Lemma 5.5, the pair (f_x, θ) is feasible, i.e.,

$$f_x(a) + f_x(b) + \theta((b - a) - m) \leq (b - a)^2 \quad \text{for all } (a, b) \in T.$$

Moreover, Lemma 5.4 yields

$$g_x(a) + g_x(b) = k_x(b - a) = (b - a)^2 - \theta(b - a)$$

for ν_x -almost every $(a, b) \in T$. Consequently,

$$f_x(a) + f_x(b) + \theta((b - a) - m) = g_x(a) + g_x(b) + \theta(b - a) = (b - a)^2$$

for ν_x -almost every $(a, b) \in T$. Thus, ν_x is concentrated on the contact set of the feasible pair (f_x, θ) .

Corollary 3.7 therefore implies that ν_x is optimal for the triangular problem, i.e.,

$$\min_{\substack{\nu \in \Pi_T \\ \int_T (b-a) \, d\nu = m/2}} \int_T (b - a)^2 \, d\nu = \int_T (b - a)^2 \, d\nu_x.$$

By the representation of $\bar{\varrho}$ in (52), it follows that

$$\bar{\varrho}(x) = 1 - 12 \int_T (b-a)^2 \, d\nu_x(a, b).$$

On the other hand, since $\pi_x = \nu_x + S_{\#}\nu_x$, we have

$$\varrho(C_x) = 1 - 6 \int_{[0,1]^2} (b-a)^2 \, d\pi_x(a, b) = 1 - 12 \int_T (b-a)^2 \, d\nu_x(a, b).$$

Therefore, $\varrho(C_x) = \bar{\varrho}(x)$.

We next prove uniqueness of the optimizer C_x for $x \in [-\frac{1}{2}, 1)$. The argument proceeds in two steps. First, we prove uniqueness of the optimizer of the triangular problem. By Lemma 5.7, the contact set of the dual pair (f_x, θ) is the graph of the map τ_x . Since this dual pair is optimal, complementary slackness forces every triangular optimizer to be concentrated on this graph. The difference of two such optimizers is therefore determined by a finite signed measure μ on the first coordinate. The common marginal constraint then implies $\mu = -(\tau_x)_{\#}\mu$. Since τ_x shifts every point by a strictly positive amount, Lemma 5.6 yields $\mu = 0$.

In the second step, we exclude nonsymmetric maximizers of the original copula problem. Any maximizer has the same symmetrization as the candidate optimizer. Hence, its difference from the candidate is an antisymmetric signed measure concentrated on the contact graph and its transpose. Restricting this measure to the upper triangle yields an invariance relation under τ_x , and a second application of Lemma 5.6 completes the proof.

For the first step, recall from (110) that $\tau_x(a) = a + \frac{\theta - g'_x(a)}{2}$. Since $|g'_x| \leq v < \theta$ by (106), we have

$$\tau_x(a) \geq a + q_x \quad \text{for } q_x := \frac{\theta - v}{2} > 0. \quad (112)$$

We first note that the particular dual pair (f_x, θ) is optimal. Indeed, (f_x, θ) is dual feasible by Lemma 5.5, while ν_x is primal feasible. Moreover, Lemma 5.4 shows that equality in the dual constraint holds ν_x -almost surely. Consequently, the primal value attained by ν_x coincides with the dual value attained by (f_x, θ) . Thus, ν_x is primal optimal and (f_x, θ) is dual optimal.

Let now $\tilde{\nu}$ be any optimizer of the triangular problem in (52). Since (f_x, θ) is dual optimal, equality of the primal and dual values yields

$$0 = \int_T \left[(b-a)^2 - f_x(a) - f_x(b) - \theta((b-a) - m) \right] \tilde{\nu}(da, db).$$

The integrand is nonnegative by dual feasibility. Hence it must vanish $\tilde{\nu}$ -almost surely, so $\tilde{\nu}$ is concentrated on the contact set of (f_x, θ) . By Lemma 5.7,

$$\Gamma_{f_x, \theta} = \{(a, \tau_x(a)) : a \in A_x\}.$$

The same is true for ν_x . Now, set

$$\zeta := \tilde{\nu} - \nu_x, \quad \mu := (\text{proj}_1)_{\#}\zeta.$$

Since both $\tilde{\nu}$ and ν_x are concentrated on $\Gamma_{f_x, \theta}$, so is the signed measure ζ . Moreover, the projection onto the first coordinate is a bijection from $\Gamma_{f_x, \theta} = \{(a, \tau_x(a)) : a \in A_x\}$ onto A_x , with inverse $a \mapsto (a, \tau_x(a))$. Consequently,

$$\zeta = (\text{id}, \tau_x)_{\#}\mu, \quad (\text{proj}_2)_{\#}\zeta = (\tau_x)_{\#}\mu. \quad (113)$$

Since $\tilde{\nu}$ and ν_x are in Π_T and thus satisfy the same coupled marginal constraint in (51), their difference satisfies

$$(\text{proj}_1)_{\#}\zeta + (\text{proj}_2)_{\#}\zeta = 0. \quad (114)$$

Therefore, equations (113) and (114) yield

$$\mu = -(\tau_x)_{\#}\mu.$$

By (112), the map τ_x satisfies the assumptions of Lemma 5.6. Applying the lemma with $\varepsilon = -1$ gives $\mu = 0$, and hence $\zeta = 0$. Consequently, $\tilde{\nu} = \nu_x$. Thus the optimizer of the triangular problem is unique.

In the second step, we exclude nonsymmetric maximizers of the original copula problem. Let $\tilde{\pi}$ be the coupling associated with any maximizing copula at the prescribed value x , and define its symmetrization by

$$\tilde{\pi}^{\text{sym}} := \frac{1}{2}(\tilde{\pi} + S_{\#}\tilde{\pi}).$$

Since both the objective and the constraint are symmetric, $\tilde{\pi}^{\text{sym}}$ is again optimal. Its triangular representative is therefore ν_x by the uniqueness proved above. Hence, by Lemma 3.3, $\tilde{\pi}^{\text{sym}} = \pi_x$.

Now, set

$$\delta := \tilde{\pi} - \pi_x.$$

Since π_x is symmetric, the preceding identity implies $S_{\#}\delta = -\delta$. Moreover, both marginals of δ vanish, and

$$\tilde{\pi} + S_{\#}\tilde{\pi} = 2\pi_x. \quad (115)$$

Consider the symmetrized contact set $K := \Gamma_{f_x, \theta} \cup S_{\#}\Gamma_{f_x, \theta}$. Since the triangular representative ν_x of π_x is concentrated on $\Gamma_{f_x, \theta}$, the symmetric coupling π_x is concentrated on K . Hence $\pi_x(K^c) = 0$. Evaluating (115) on K^c gives

$$\tilde{\pi}(K^c) + S_{\#}\tilde{\pi}(K^c) = 0.$$

Both terms are nonnegative, and hence $\tilde{\pi}(K^c) = 0$. Thus also δ is concentrated on K .

By (112), $\Gamma_{f_x, \theta}$ lies strictly above the diagonal, whereas $S(\Gamma_{f_x, \theta})$ lies strictly below it. In particular, K does not intersect the diagonal. Define

$$\eta := \delta|_{\{a < b\}}.$$

Since $S_{\#}\delta = -\delta$ and δ has no mass on the diagonal, we obtain

$$\delta = \eta - S_{\#}\eta.$$

Since the first marginal of δ vanishes, it follows that $(\text{proj}_1)_{\#}\eta = (\text{proj}_2)_{\#}\eta$. By the concentration of δ on K established above and by (112), the restriction $\eta = \delta|_{\{a < b\}}$ is concentrated on $\Gamma_{f_x, \theta}$, which, by Lemma 5.7, is the graph of τ_x . Setting

$$\mu := (\text{proj}_1)_{\#}\eta,$$

we therefore have

$$\eta = (\text{id}, \tau_x)_{\#}\mu \quad \text{and} \quad (\tau_x)_{\#}\mu = (\text{proj}_2)_{\#}\eta = (\text{proj}_1)_{\#}\eta = \mu.$$

Using (112), we may apply Lemma 5.6 once more, now with $\varepsilon = 1$. This gives $\mu = 0$, and hence $\eta = 0$. Therefore $\delta = 0$, so $\tilde{\pi} = \pi_x$. Thus the maximizing coupling, and hence the maximizing copula, is uniquely given by C_x .

Finally, let $x = 1$. Then $\phi(C) = 1$ implies $\mathbb{E}|U - V| = 0$, and hence $U = V$ almost surely. Thus, the only possible copula is the comonotonic copula $C_1(u, v) = \min\{u, v\}$ for which $\varrho(C_1) = 1 = \bar{\varrho}(1)$. This completes the proof. $\square$

6 Proofs of Section 1

Proof of Theorem 1.1. Fix $x \in [-\frac{1}{2}, 1)$, and let $N \in \mathbb{N}$ be such that $x \in I_N = \left[1 - \frac{3}{2N}, 1 - \frac{3}{2N+2}\right)$. Recall $m = \frac{1-x}{3}$, $L = \frac{1}{2N}$, and $R = \frac{1}{2N+2}$ as well as $I_N^L = \left[1 - 3L, 1 - \frac{3}{2}(L + R)\right)$ and $I_N^R = \left[1 - \frac{3}{2}(L + R), 1 - 3R\right)$.

If $x \in I_N^L$, then

$$m \in \left(\frac{L + R}{2}, L\right].$$

Hence, in this case, it is $\ell = L$ and $\Delta = L - m$. By Proposition 4.6, we then have

$$\varrho(C_x) = 1 - 6L(2m - L) - \frac{4(L - m)^{3/2}}{\sqrt{N(N+1)}} = 1 - 6\ell(2m - \ell) - \frac{4\Delta^{3/2}}{\sqrt{N(N+1)}}.$$

If $x \in I_N^R$, then

$$m \in \left(R, \frac{L+R}{2}\right].$$

In this case, it is $\ell = R$ and $\Delta = m - R$. Again by Proposition 4.6, it follows that

$$\varrho(C_x) = 1 - 6R(2m - R) - \frac{4(m - R)^{3/2}}{\sqrt{N(N+1)}} = 1 - 6\ell(2m - \ell) - \frac{4\Delta^{3/2}}{\sqrt{N(N+1)}}.$$

At the common boundary point of I_N^L and I_N^R , we have $m = \frac{L+R}{2}$. In this case, both L and R are equally close to m , and $L(2m - L) = LR = R(2m - R)$, while $L - m = m - R = \frac{L-R}{2}$. Hence, the two expressions coincide.

By Proposition 5.8, the copula C_x uniquely maximizes Spearman's rho among all copulas with Spearman's footrule equal to x . Consequently,

$$\bar{\varrho}(x) = \varrho(C_x) = 1 - 6\ell(2m - \ell) - \frac{4\Delta^{3/2}}{\sqrt{N(N+1)}}.$$

In particular, the maximum is uniquely attained by the symmetric copula C_x .

It remains to verify the second representation in (9). Recall that

$$u(x) = 1 - \frac{2}{3}(1 - x)^2 = 1 - 6m^2.$$

Since $m = \ell \pm \Delta$, we have in either case $m^2 = \ell(2m - \ell) + \Delta^2$. Therefore,

$$\begin{aligned} \bar{\varrho}(x) &= 1 - 6\ell(2m - \ell) - \frac{4\Delta^{3/2}}{\sqrt{N(N+1)}} \\ &= 1 - 6m^2 + 6\Delta^2 - \frac{4\Delta^{3/2}}{\sqrt{N(N+1)}} \\ &= u(x) - \left(\frac{4\Delta^{3/2}}{\sqrt{N(N+1)}} - 6\Delta^2 \right). \end{aligned}$$

This proves (9). The assertion at $x = 1$ follows from Proposition 5.8. $\square$

Proof of Corollary 1.2. The sharp lower bound for Spearman's rho in terms of Spearman's footrule gives

$$\varrho(C) \geq \underline{\varrho}(x) = \frac{2\sqrt{3}}{9}(1 + 2x)^{3/2} - 1$$

for every copula C satisfying $\phi(C) = x$, and this lower bound is attained for every $x \in [-\frac{1}{2}, 1]$; see [19, Theorem 8]. Together with the upper bound established in Theorem 1.1, every copula C satisfies

$$\underline{\varrho}(\phi(C)) \leq \varrho(C) \leq \bar{\varrho}(\phi(C)).$$

It follows that

$$\Omega_{\varrho, \phi} \subseteq \{(\varrho, x) : x \in [-\frac{1}{2}, 1], \underline{\varrho}(x) \leq \varrho \leq \bar{\varrho}(x)\}.$$

To prove the reverse inclusion, fix $x \in [-\frac{1}{2}, 1]$. Let $C_x^+ := C_x$ be the copula in (75) attaining the upper boundary, and let B_x be the Bertino copula in (50) attaining the lower boundary. Thus

$$\phi(C_x^+) = \phi(B_x) = x, \quad \varrho(C_x^+) = \bar{\varrho}(x), \quad \text{and} \quad \varrho(B_x) = \underline{\varrho}(x).$$

For $t \in [0, 1]$, define

$$C_{x,t} := tC_x^+ + (1 - t)B_x.$$

Since the class of copulas is convex, $C_{x,t}$ is again a copula. Moreover, both ϕ and ϱ are affine in the copula. Hence,

$$\phi(C_{x,t}) = t\phi(C_x^+) + (1-t)\phi(B_x) = x, \quad \text{and} \quad \varrho(C_{x,t}) = t\bar{\varrho}(x) + (1-t)\underline{\varrho}(x).$$

As t ranges over $[0, 1]$, the latter expression ranges over the entire interval $[\underline{\varrho}(x), \bar{\varrho}(x)]$. Thus, every point between the lower and upper boundary is attained.

Finally, for $x = 1$, the identity

$$\phi(C) = 1 \implies \mathbb{E}|U - V| = 0$$

implies $U = V$ almost surely. Hence, the only possible copula is M , and $\varrho(M) = \phi(M) = 1$. Therefore,

$$\Omega_{\varrho, \phi} = \left\{ (\varrho, x) : x \in [-\tfrac{1}{2}, 1], \frac{2\sqrt{3}}{9}(1+2x)^{3/2} - 1 \leq \varrho \leq \bar{\varrho}(x) \right\},$$

as claimed. $\square$

Proof of Proposition 1.6. For $m = 0$, we have $V_{\min}(0) = 0$, and (19) reduces to $\mathbb{E}(U - V)^2 \geq 0$. The bound is attained by the comonotone coupling $U = V$.

Now let $m > 0$ and put $x = 1 - 3m$. By (4), maximizing $\varrho(C)$ under the constraint $\phi(C) = x$ is equivalent to minimizing $\mathbb{E}(U - V)^2$ over all couplings $U, V \sim \mathcal{U}(0, 1)$ satisfying $\mathbb{E}|U - V| = m$. Hence, Theorem 1.1 yields

$$\min_{\substack{U, V \sim \mathcal{U}(0, 1) \\ \mathbb{E}|U - V| = m}} \mathbb{E}(U - V)^2 = \ell(2m - \ell) + \frac{2\Delta_m^{3/2}}{3\sqrt{N(N+1)}}.$$

Since $\ell(2m - \ell) = m^2 - \Delta_m^2$, we obtain

$$\min_{\substack{U, V \sim \mathcal{U}(0, 1) \\ \mathbb{E}|U - V| = m}} \mathbb{E}(U - V)^2 = m^2 + \frac{2\Delta_m^{3/2}}{3\sqrt{N(N+1)}} - \Delta_m^2 = m^2 + V_{\min}(m),$$

which proves (19). By Theorem 1.1, the minimum is attained by C_x , proving sharpness.

It remains to characterize the zeros of the correction term. Since ℓ is an endpoint nearest to m , we have $0 \leq \Delta_m \leq \frac{1}{4N(N+1)}$. Consequently,

$$V_{\min}(m) = \Delta_m^{3/2} \left(\frac{2}{3\sqrt{N(N+1)}} - \sqrt{\Delta_m} \right),$$

and

$$\frac{2}{3\sqrt{N(N+1)}} - \sqrt{\Delta_m} \geq \frac{1}{6\sqrt{N(N+1)}} > 0.$$

Thus $V_{\min}(m) = 0$ if and only if $\Delta_m = 0$. By the choice $m \in (\frac{1}{2N+2}, \frac{1}{2N}]$, this is equivalent to $m = 1/(2N)$. Together with the case $m = 0$, this proves $V_{\min}(m) = 0$ if and only if $m \in \mathfrak{C}$. $\square$

Proof of Corollary 1.7. Put $Z := |U - V|$, $m := \mathbb{E}Z$, and $x := 1 - 3m$. By Lemma 3.2,

$$\text{Var}(Z) = \frac{1 - \varrho(C)}{6} - m^2.$$

Substituting the upper boundary $\bar{\varrho}(x)$ from Theorem 1.1 gives $V_{\min}(m)$, while substituting

$$\underline{\varrho}(1 - 3m) = 2(1 - 2m)^{3/2} - 1$$

gives $V_{\max}(m)$. Both bounds are attained by C_x and B_x , respectively.

It remains to maximize $V_{\max}$ on $[0, \frac{1}{2}]$. Its derivative is

$$V'_{\max}(m) = \sqrt{1 - 2m} - 2m,$$

so its unique maximizer is $m_0 = (\sqrt{5} - 1)/4$. A direct substitution gives

$$V_{\max}(m_0) = \frac{5(3 - \sqrt{5})}{24},$$

and $x_0 = 1 - 3m_0 = (7 - 3\sqrt{5})/4$. This proves (25) and the equality statements. $\square$

7 Proofs of Section 2

Proof of Theorem 2.1. We embed the permutation into a copula. Partition $(0, 1]$ into the rank cells $I_i := ((i-1)/n, i/n]$, and define the measure-preserving map

$$Q_\pi(0) := 0, \quad Q_\pi(u) := u + \frac{\pi(i) - i}{n}, \quad u \in I_i.$$

If U is uniform on $[0, 1]$ and $V = Q_\pi(U)$, then V is also uniform, so the copula of (U, V) is a shuffle of M [21]. Conditional on $U \in I_i$, the difference $V - U$ is the constant $(\pi(i) - i)/n$. Therefore

$$\mathbb{E}|U - V| = m_\pi, \quad \mathbb{E}(U - V)^2 = q_\pi.$$

In other words, the normalized finite-ranking costs in (28) are exactly the two copula moments considered above. Since

$$q_\pi = \mathbb{E}|U - V|^2 = m_\pi^2 + \text{Var}(|U - V|),$$

Corollary 1.7 gives (29). $\square$

Proof of Lemma 2.3. Let $U := U' + \frac{1}{2}$ and $V := \frac{1}{2} - V'$. Then $U, V \sim \mathcal{U}(0, 1)$ and, by (32),

$$\Psi(U', V') = |U - V|.$$

Hence, the assertion is equivalent to the existence of a coupling $U, V \sim \mathcal{U}(0, 1)$ such that $|U - V| = m$ almost surely. This is equivalent to equality in the Cauchy–Schwarz inequality (21). By Proposition 1.6, such equality is attainable if and only if $m \in \mathfrak{C}$. $\square$

Proof of Theorem 2.4. For $U', V' \sim \mathcal{U}(-\frac{1}{2}, \frac{1}{2})$, define $U := U' + \frac{1}{2}$ and $V := \frac{1}{2} - V'$. Then $U, V \sim \mathcal{U}(0, 1)$ and, by (32), $\Psi(U', V') = |U - V|$. Hence,

$$\mathbb{V}(m) = \min \{ \text{Var}(|U - V|) : U, V \sim \mathcal{U}(0, 1), \mathbb{E}|U - V| = m \}.$$

Since $\text{Var}(|U - V|) = \mathbb{E}(U - V)^2 - m^2$, Proposition 1.6 implies $\mathbb{V}(m) = V_{\min}(m)$. The final assertion follows from Proposition 1.6(iii). $\square$

Proof of Theorem 2.6. Let C be the copula of (X, Y) . Then Corollary 1.2 implies for $x = \xi(C) = \phi(C * C)$ that

$$\underline{\varrho}(x) \leq \varrho(C * C) \leq \overline{\varrho}(x).$$

By non-negativity of the copula correlation ratio and its representation in (40), we obtain

$$\max\{0, \underline{\varrho}(\xi(X, Y))\} \leq \eta(X, Y) \leq \overline{\varrho}(\xi(X, Y)). \quad (116)$$

For $U = F_Y(Y)$, the statement now follows from

$$\eta(X, Y) = 12\mathbb{E} \left(\int_0^1 (P(U \leq t | X) - t) dt \right)^2 \leq 12\mathbb{E} \int_0^1 (P(U \leq t | X) - t)^2 dt = 2\xi(X, Y),$$

applying Cauchy–Schwarz inequality. $\square$

Proof of Proposition 2.10. We first show that $\Omega_{\xi, \eta}$ is convex. Let (X_1, Y_1) and (X_2, Y_2) , with continuous marginal distribution functions, attain (ξ_1, η_1) and (ξ_2, η_2) , respectively. Set

$$U_j := F_{Y_j}(Y_j), \quad T_j := F_{X_j}(X_j), \quad j = 1, 2.$$

Then $U_j, T_j \sim \mathcal{U}(0, 1)$. Moreover, since F_{X_j} is continuous, conditioning on T_j is equivalent to conditioning on X_j , up to null sets.

Let $B \sim \text{Bernoulli}(\alpha)$, $\alpha \in [0, 1]$, be independent of both models, with $P(B = 1) = \alpha$, and define

$$\tilde{X} := \begin{cases} \alpha T_1, & B = 1, \\ \alpha + (1 - \alpha)T_2, & B = 0, \end{cases} \quad \tilde{U} := \begin{cases} U_1, & B = 1, \\ U_2, & B = 0. \end{cases}$$

Conditionally on $B = 1$, $\tilde{X}$ is uniform on $[0, \alpha]$, whereas conditionally on $B = 0$, it is uniform on $[\alpha, 1]$. Hence $\tilde{X} \sim \mathcal{U}(0, 1)$. Moreover, since $U_1, U_2 \sim \mathcal{U}(0, 1)$, also $\tilde{U} \sim \mathcal{U}(0, 1)$. Thus, setting $\tilde{Y} := \tilde{U}$, both marginal distributions of $(\tilde{X}, \tilde{Y})$ are continuous.

On $\{B = 1\}$, we have $(\tilde{U} \mid \tilde{X}) \stackrel{d}{=} (U_1 \mid X_1)$, where as on $\{B = 0\}$, it is $(\tilde{U} \mid \tilde{X}) \stackrel{d}{=} (U_2 \mid X_2)$. Let $\tilde{U}'$ denote a conditionally independent copy of $\tilde{U}$ given $\tilde{X}$. Then

$$\mathbb{E}|\tilde{U} - \tilde{U}'| = \alpha \mathbb{E}|U_1 - U_1'| + (1 - \alpha) \mathbb{E}|U_2 - U_2'|,$$

and therefore

$$\xi(\tilde{X}, \tilde{Y}) = \alpha \xi_1(X_1, Y_1) + (1 - \alpha) \xi_2(X_2, Y_2). \quad (117)$$

Further, for $M_j := \mathbb{E}[U_j \mid X_j]$, we have $\mathbb{E}M_j = 1/2$ and thus

$$\text{Var}(\mathbb{E}[\tilde{U} \mid \tilde{X}]) = \alpha \text{Var}(M_1) + (1 - \alpha) \text{Var}(M_2).$$

This gives

$$\eta(\tilde{X}, \tilde{Y}) = \alpha \eta_1 + (1 - \alpha) \eta_2. \quad (118)$$

Combining (117) and (118) yields $\alpha(\xi_1, \eta_1) + (1 - \alpha)(\xi_2, \eta_2) \in \Omega_{\xi, \eta}$, which proves that $\Omega_{\xi, \eta}$ is convex.

For the lower inner curve η_ℓ , let $t \in [0, 1]$, let $U \sim \mathcal{U}(0, 1)$, and put $a_t := (1 - t)/2$ and $b_t := (1 + t)/2$. Define

$$X_t := \begin{cases} \min\{U, 1 - U\}, & U \notin [a_t, b_t], \\ U, & U \in [a_t, b_t]. \end{cases}$$

Outside the interval $[a_t, b_t]$, the level sets $\{u : X_t(u) = x\}$ of X_t are the symmetric pairs $\{x, 1 - x\}$, whereas inside the interval they are singletons. The distribution function of X_t is continuous. Consequently, for $x < a_t$, conditionally on $X_t = x$, the rank U is equally likely to be x or $1 - x$, whereas for $x \in [a_t, b_t]$, it is uniquely determined by $U = x$. Direct integration yields

$$\xi(X_t, U) = \frac{1 + 3t^2}{4} \quad \text{and} \quad \eta(X_t, U) = 12 \text{Var}(\mathbb{E}[U \mid X_t]) = t^3.$$

Eliminating t gives

$$\eta = \left(\frac{4\xi - 1}{3} \right)^{3/2}, \quad \frac{1}{4} \leq \xi \leq 1,$$

which is the second branch of η_ℓ . The endpoint $t = 0$ gives $(1/4, 0)$, while independence gives $(0, 0)$. By convexity, the entire segment $\{(x, 0) : 0 \leq x \leq 1/4\}$ is attainable by models with continuous marginals, proving the first branch of η_ℓ .

For the upper inner curve η_u , fix $a \in [0, 1/2]$ and set

$$h_a(s) := \min\{s, a, 1 - s\}, \quad 0 \leq s \leq 1.$$

For $\sigma \in \{-1, 1\}$, define

$$G_\sigma(s) := s + \sigma h_a(s). \quad (119)$$

The functions G_{-1} and G_1 are distribution functions and satisfy

$$\frac{G_{-1}(s) + G_1(s)}{2} = s. \quad (120)$$

Let S be uniformly distributed on $\{-1, 1\}$ and, conditionally on $S = \sigma$, let U have distribution function G_σ . Then, by (120), $U \sim \mathcal{U}(0, 1)$.

To obtain a continuous conditioning variable, let $R \sim \mathcal{U}(0, 1)$ be independent of (S, U) and define

$$X := \begin{cases} R/2, & S = -1, \\ (1 + R)/2, & S = 1. \end{cases}$$

Then $X \sim \mathcal{U}(0, 1)$, and S is measurable with respect to X . Moreover, the additional variable R carries no information about U given S . Hence

$$P(U \leq s \mid X) = P(U \leq s \mid S) = G_S(s). \quad (121)$$

Thus, two conditionally independent draws given X define the same conditional-i.i.d. model as conditioning directly on S . Since

$$\int_0^1 h_a(s)^2 \, ds = a^2 - \frac{4}{3}a^3 \quad \text{and} \quad \int_0^1 h_a(s) \, ds = a(1 - a),$$

substitution into (43) gives

$$\xi_0(a) := 2a^2(3 - 4a) \quad \text{and} \quad \eta_0(a) := 12a^2(1 - a)^2.$$

Now let J be uniformly distributed on $\{1, \dots, n\}$, independently of S , and let W denote the response rank in the preceding binary model, i.e.,

$$\text{law}(W \mid S = \sigma) = G_\sigma, \quad \sigma \in \{-1, 1\}.$$

Define

$$U := \frac{J - 1 + W}{n}.$$

Thus, conditionally on $J = j$, the binary model is mapped affinely from $[0, 1]$ onto the interval $[(j - 1)/n, j/n]$. The conditioning state is therefore the pair (J, S) : the variable J determines the interval containing U , while S determines the conditional distribution of U within that interval.

To encode this discrete state by a continuous conditioning variable, set

$$K := 2(J - 1) + 1_{\{S=1\}},$$

and let $R \sim \mathcal{U}(0, 1)$ be independent of (J, S, U) . Then

$$X := \frac{K + R}{2n}$$

is uniformly distributed on $[0, 1]$ and determines the pair (J, S) . Hence the resulting model has continuous marginals and belongs to $\Omega_{\xi, \eta}$.

Conditionally on (J, S) , let W' be an independent copy of W and put $U' = (J - 1 + W')/n$. Then

$$|U - U'| = \frac{1}{n}|W - W'|,$$

and hence

$$1 - \xi = 3\mathbb{E}|U - U'| = \frac{1 - \xi_0(a)}{n}.$$

Therefore,

$$\xi = 1 - \frac{1 - \xi_0(a)}{n} = x_n(a).$$

For the copula correlation ratio, let $\mu_S := \mathbb{E}[W \mid S]$. Then

$$\mathbb{E}[U \mid J, S] = \frac{J - 1 + \mu_S}{n} = \frac{J - \frac{1}{2}}{n} + \frac{\mu_S - \frac{1}{2}}{n}.$$

Thus the conditional mean consists of the midpoint of the J -th interval and the binary perturbation inherited from the base model. Since J and S are independent,

$$\text{Var}(\mathbb{E}[U \mid J, S]) = \frac{1}{n^2} (\text{Var}(J - 1) + \text{Var}(\mu_S)) = \frac{1}{12n^2} (n^2 - 1 + \eta_0(a)),$$

where we used $\text{Var}(J - 1) = (n^2 - 1)/12$ and $\eta_0(a) = 12 \text{Var}(\mu_S)$. Consequently,

$$\eta = 12 \text{Var}(\mathbb{E}[U \mid J, S]) = 1 - \frac{1 - \eta_0(a)}{n^2} = y_n(a).$$

For every fixed $n \in \mathbb{N}$, the curve $\{(x_n(a), y_n(a)) : a \in [0, \frac{1}{2}]\}$ is compact, and these curves converge uniformly to $(1, 1)$ as $n \rightarrow \infty$. Hence $\mathcal{S}_u$ is compact. Since $\Omega_{\xi, \eta}$ is convex, $\text{conv}(\mathcal{S}_u) \subseteq \Omega_{\xi, \eta}$. Moreover, the projection of $\mathcal{S}_u$ onto its first coordinate is $[0, 1]$. Indeed, for $x \in [0, 1)$, write $q = 1 - x$ and choose $n = \lfloor \frac{1}{q} \rfloor$. Then $1 - nq \in [0, 1/2]$, which belongs to the range of $a \mapsto 2a^2(3 - 4a)$ for $a \in [0, 1/2]$. Thus there exists $a \in [0, 1/2]$ such that $x_n(a) = x$. Since $\text{conv}(\mathcal{S}_u)$ is compact, the maximum defining $\eta_u(x)$ is attained. Its concavity follows from the convexity of $\text{conv}(\mathcal{S}_u)$.

It remains to verify that the lower construction does not exceed the upper one. Fix x and choose (n, a) such that $x_n(a) = x$. Put $r := 1 - \xi_0(a)$, so that $1 - x = r/n$. For $x \geq 1/4$,

$$\eta_\ell(x) = \left(1 - \frac{4r}{3n}\right)^{3/2} \leq 1 - \frac{4r}{3n} \leq 1 - \frac{1 - \eta_0(a)}{n^2} = y_n(a),$$

where the second inequality follows from

$$\frac{4}{3}(1 - \xi_0(a)) - (1 - \eta_0(a)) = \frac{1 + 12a^2 - 40a^3 + 36a^4}{3} \geq 0.$$

For $x \leq 1/4$, the same ordering follows from nonnegativity. Finally, convexity of $\Omega_{\xi, \eta}$ fills the vertical segment between the attained lower and upper points for every fixed x , proving (46). $\square$

Acknowledgement

The first author was funded in whole by the Austrian Science Fund (FWF) [10.55776/PAT1669224] project *Stochastic orders for functional dependence*.

References

- [1] Jonathan Ansari and Sebastian Fuchs. A direct extension of Azadkia & Chatterjee’s rank correlation to multi-response vectors. arXiv preprint arXiv:2212.01621v4, 2025.
- [2] Jonathan Ansari and Marcus Rockel. The exact region and an inequality between Chatterjee’s and Spearman’s rank correlations. *J. Multivariate Anal.*, 214:19, 2026. Id/No 105630.
- [3] Jonathan Ansari, Patrick B. Langthaler, Sebastian Fuchs, and Wolfgang Trutschnig. Quantifying and estimating dependence via sensitivity of conditional distributions. *Bernoulli*, 32(1): 179–204, 2026.
- [4] Mona Azadkia and Sourav Chatterjee. A simple measure of conditional dependence. *Ann. Stat.*, 49(6):3070–3102, 2021.
- [5] Valeria Bignozzi and Giovanni Puccetti. Studying mixability with supermodular aggregating functions. *Stat. Probab. Lett.*, 100:48–55, 2015.
- [6] Sourav Chatterjee. A new coefficient of correlation. *J. Am. Stat. Assoc.*, 116(536):2009–2022, 2021.
- [7] Sourav Chatterjee. A survey of some recent developments in measures of association. In *Probability and stochastic processes. A volume in honour of Rajeeva L. Karandikar*, pages 109–128. Singapore: Springer, 2024.
- [8] Stéphan Cléménçon, Marine Depecker, and Nicolas Vayatis. Ranking forests. *J. Mach. Learn. Res.*, 14:39–73, 2013.
- [9] Holger Dette, Karl F. Siburg, and Pavel A. Stoimenov. A copula-based non-parametric measure of regression dependence. *Scand. J. Stat.*, 40(1):21–41, 2013.

- [10] Persi Diaconis and Ronald L. Graham. Spearman’s footrule as a measure of disarray. *J. R. Stat. Soc., Ser. B*, 39(2):262–268, 1977.
- [11] Fabrizio Durante and Carlo Sempi. *Principles of copula theory*. Boca Raton, FL: CRC Press, 2016.
- [12] Juan Fernández Sánchez and Wolfgang Trutschnig. Some members of the class of (quasi-)copulas with given diagonal from the Markov kernel perspective. *Commun. Stat., Theory Methods*, 45(5):1508–1526, 2016.
- [13] Sebastian Fuchs. Quantifying directed dependence via dimension reduction. *J. Multivariate Anal.*, 201:21, 2024. Id/No 105266.
- [14] Fabrice Gamboa, Pierre Gremaud, Thierry Klein, and Agnès Lagnoux. Global sensitivity analysis: a novel generation of mighty estimators based on rank statistics. *Bernoulli*, 28(4): 2345–2374, 2022.
- [15] Christian Genest, Johanna Nešlehová, and Noomen Ben Ghorbal. Spearman’s footrule and Gini’s gamma: a review with complements. *J. Nonparametr. Stat.*, 22(8):937–954, 2010.
- [16] Zhen Huang, Nabarun Deb, and Bodhisattva Sen. Kernel partial correlation coefficient – a measure of conditional dependence. *J. Mach. Learn. Res.*, 23(216):1–58, 2022.
- [17] Damjana Kokol Bukovšek and Blaž Mojskerc. On the exact region determined by Spearman’s footrule and Gini’s gamma. *J. Comput. Appl. Math.*, 410:114212, 2022.
- [18] Damjana Kokol Bukovšek and Nik Stopar. On the exact regions determined by Kendall’s tau and other concordance measures. *Mediterr. J. Math.*, 20(3):147, 2023.
- [19] Damjana Kokol Bukovšek and Nik Stopar. On the exact region determined by Spearman’s rho and Spearman’s footrule. *J. Comput. Appl. Math.*, 437:115463, 2024.
- [20] Damjana Kokol Bukovšek, Tomaž Košir, Blaž Mojskerc, and Matjaž Omladič. Spearman’s footrule and Gini’s gamma: local bounds for bivariate copulas and the exact region with respect to Blomqvist’s beta. *J. Comput. Appl. Math.*, 390:113385, 2021.
- [21] Piotr Mikusiński, Howard Sherwood, and Michael D. Taylor. Shuffles of Min. *Stochastica*, 13 (1):61–74, 1992.
- [22] Alfred Müller and Dietrich Stoyan. *Comparison methods for stochastic models and risks*. Wiley Ser. Probab. Stat. Chichester: Wiley, 2002.
- [23] Roger B. Nelsen. *An introduction to copulas*. Springer Ser. Stat. New York, NY: Springer, 2nd edition, 2006.
- [24] Roger B. Nelsen, José Juan Quesada-Molina, José Antonio Rodríguez-Lallena, and Manuel Úbeda-Flores. Best-possible bounds on sets of bivariate distribution functions. *J. Multivariate Anal.*, 90(2):348–358, 2004.
- [25] Giovanni Puccetti and Ludger Rüschendorf. Computation of sharp bounds on the expected value of a supermodular function of risks with given marginals. *Commun. Stat., Simul. Comput.*, 44(3):705–718, 2015.
- [26] Marcus Rockel. On the exact region between Chatterjee’s rank correlation and Spearman’s footrule. *J. Comput. Appl. Math.*, 485:19, 2026. Id/No 117466.
- [27] Marcus Rockel. Kendall and Spearman bounds for Chatterjee’s rank correlation under positive dependence. arXiv preprint arXiv:2606.22074, 2026.
- [28] Manuela Schreyer, Roland Paulin, and Wolfgang Trutschnig. On the exact region determined by Kendall’s τ and Spearman’s ρ . *J. R. Stat. Soc., Ser. B, Stat. Methodol.*, 79(2):613–633, 2017.
- [29] Jia-Han Shih and Takeshi Emura. On the copula correlation ratio and its generalization. *J. Multivariate Anal.*, 182:15, 2021. Id/No 104708.

- [30] Christopher Strothmann, Holger Dette, and Karl Friedrich Siburg. Rearranged dependence measures. *Bernoulli*, 30(2):1055–1078, 2024.
- [31] Engin A. Sungur. A note on directional dependence in regression setting. *Commun. Stat., Theory Methods*, 34(9-10):1957–1965, 2005.
- [32] Marco Tschimpke, Manuela Schreyer, and Wolfgang Trutschnig. Revisiting the region determined by Spearman’s ρ and Spearman’s footrule ϕ . *J. Comput. Appl. Math.*, 457:116259, 2025.
- [33] Manuel Úbeda-Flores. On the best-possible upper bound on sets of copulas with given diagonal sections. *Soft Comput.*, 12(10):1019–1025, 2008.
- [34] Cédric Villani. *Optimal transport. Old and new*, volume 338 of *Grundlehren Math. Wiss.* Berlin: Springer, 2009.
- [35] Ruodu Wang. Current open questions in complete mixability. *Probab. Surv.*, 12:13–32, 2015.
- [36] Johannes C. W. Wiesel. Measuring association with Wasserstein distances. *Bernoulli*, 28(4):2816–2832, 2022.
- [37] D. A. Zaev. On the Monge–Kantorovich problem with additional linear constraints. *Math. Notes*, 98(5):725–741, 2015.