Supplementary Material for “The Tail Copula Characteristic Function”

June 3, 2026

This supplement contains complete proofs of all theoretical results stated in the main paper, additional simulation results, and extended application diagnostics. All equation, theorem, and section numbering follows that of the main paper.

Contents

1 Proof of Lemma 1: Existence and Uniqueness of the TCCF

1.1 Proof of Lemma 1(i): Well-definedness

Proof. Using the polar decomposition $\mathcal{M}(\mathrm{d}r, \mathrm{d}\mathbf{w}) = r^{-2}\,\mathrm{d}r \otimes \mathcal{H}(\mathrm{d}\mathbf{w})$ with $r = \|\mathbf{z}\|_1$ and $\mathbf{w} = \mathbf{z}/r \in \mathcal{S}_{d-1}$, we analyse the behaviour of the integrand separately near the origin and at infinity.

Real part. We have

$$\Re\{\mathcal{T}(\mathbf{t})\} = \int_{\mathcal{S}_{d-1}} \int_0^\infty \frac{\cos(r\mathbf{t}^\top \mathbf{w}) - 1}{r^2}\,\mathrm{d}r\,\mathcal{H}(\mathrm{d}\mathbf{w}).$$

Near $r = 0$, Taylor expansion gives $\cos(rs) - 1 = -\frac{1}{2}r^2 s^2 + O(r^4)$. The integral over $(0, \epsilon]$ is therefore bounded by

$$\frac{1}{2}\int_{\mathcal{S}_{d-1}} (\mathbf{t}^\top \mathbf{w})^2 \int_0^\epsilon r^2 \cdot r^{-2}\,\mathrm{d}r\,\mathcal{H}(\mathrm{d}\mathbf{w}) = \frac{\epsilon}{2}\int_{\mathcal{S}_{d-1}} (\mathbf{t}^\top \mathbf{w})^2\,\mathcal{H}(\mathrm{d}\mathbf{w}) < \infty,$$

since $\mathcal{H}$ is a finite measure on the compact simplex $\mathcal{S}_{d-1}$. Near $r = \infty$, using $|\cos(rs) - 1| \leq 2$, the integral over $[R, \infty)$ is bounded by

$$2\int_{\mathcal{S}_{d-1}} \int_R^\infty r^{-2}\,\mathrm{d}r\,\mathcal{H}(\mathrm{d}\mathbf{w}) = \frac{2}{R}\int_{\mathcal{S}_{d-1}} \mathcal{H}(\mathrm{d}\mathbf{w}) < \infty.$$

Combining both bounds establishes absolute convergence of the real part.

Imaginary part. For the imaginary part, we consider

$$\Im_\epsilon(\mathbf{t}) = \int_{\mathbb{E}_0 \setminus B_\epsilon} \sin(\mathbf{t}^\top \mathbf{z})\,\mathcal{M}(\mathrm{d}\mathbf{z}) = \int_{\mathcal{S}_{d-1}} \int_\epsilon^\infty \frac{\sin(r\mathbf{t}^\top \mathbf{w})}{r^2}\,\mathrm{d}r\,\mathcal{H}(\mathrm{d}\mathbf{w}).$$

For any fixed $s = \mathbf{t}^\top \mathbf{w} \neq 0$, the inner integral converges conditionally as $\epsilon \downarrow 0$ by integrating by parts:

$$\int_{\epsilon}^{\infty} \frac{\sin(rs)}{r^2}\,\mathrm{d}r = \left[-\frac{\cos(rs)}{sr^2}\right]_{\epsilon}^{\infty} - \frac{2}{s}\int_{\epsilon}^{\infty} \frac{\cos(rs)}{r^3}\,\mathrm{d}r.$$

The boundary term at ∞ vanishes, and the remaining integral converges absolutely because r^{-3} is integrable at infinity. The angular integration over $\mathcal{H}$ preserves the conditional convergence.

Continuity and symmetry. Continuity of $\mathcal{T}$ on $\mathbb{R}^d$ follows from the dominated convergence theorem applied to the real part and the uniformity of the principal value limit for the imaginary part on compact sets of $\mathbf{t}$. The symmetry $\mathcal{T}(-\mathbf{t}) = \overline{\mathcal{T}(\mathbf{t})}$ follows by conjugating the integrand, and $\mathcal{T}(\mathbf{0}) = 0$ by direct evaluation. □

1.2 Proof of Lemma 1(ii): Uniqueness

Proof. **Step 1: Test function space.** Let $\mathcal{S}_0(\mathbb{R}^d)$ denote the space of Schwartz functions $\varphi \in \mathcal{S}(\mathbb{R}^d)$ satisfying $\varphi(\mathbf{0}) = 0$. Any exponent measure $\mathcal{M}$ defines a continuous linear functional on $\mathcal{S}_0(\mathbb{R}^d)$ via $\langle \mathcal{M}, \varphi \rangle = \int_{\mathbb{E}_0} \varphi(\mathbf{z})\,\mathcal{M}(\mathrm{d}\mathbf{z})$, because $\varphi(\mathbf{0}) = 0$ ensures $\varphi(\mathbf{z}) = O(\|\mathbf{z}\|)$ near the origin, which is integrable with respect to the r^{-2} radial measure.

Step 2: Density of trigonometric polynomials. The linear span of $\{\mathbf{z} \mapsto e^{\mathrm{i}\mathbf{t}^\top \mathbf{z}} - 1 : \mathbf{t} \in \mathbb{R}^d\}$ is dense in $\mathcal{S}_0(\mathbb{R}^d)$ with respect to the Schwartz topology, by Fourier inversion and the fact that subtracting the constant term enforces the vanishing condition at the origin.

Step 3: Injectivity. If $\mathcal{T}_1(\mathbf{t}) = \mathcal{T}_2(\mathbf{t})$ for all $\mathbf{t} \in \mathbb{R}^d$, then by linearity, continuity, and density, $\langle \mathcal{M}_1, \varphi \rangle = \langle \mathcal{M}_2, \varphi \rangle$ for all $\varphi \in \mathcal{S}_0(\mathbb{R}^d)$. Hence $\mathcal{M}_1 = \mathcal{M}_2$ as Radon measures on $\mathbb{E}_0$. See Rudin (1991), Chapter 7.

Step 4: Consequence. Since $\mathcal{M}$ uniquely determines the stable tail dependence function ℓ, the angular spectral measure $\mathcal{H}$, and the tail copula Λ, the TCCF provides a complete characterisation of extremal dependence. □

2 Derivation of the Closed-Form Kernel and MMD Connection

2.1 Closed form of the one-dimensional kernel ψ

Proof. Starting from $\psi(s) = \int_0^\infty (e^{irs} - 1)/r^2 \, \mathrm{d}r$, split into real and imaginary parts. For $s \neq 0$, the change of variable $u = |s|r$ yields

$$\psi(s) = |s| \int_0^\infty \frac{\cos u - 1}{u^2} \, \mathrm{d}u + \mathrm{i}s \int_0^\infty \frac{\sin u}{u^2} \, \mathrm{d}u.$$

Real part. Integration by parts gives

$$\int_0^\infty \frac{\cos u - 1}{u^2} \, \mathrm{d}u = \left[-\frac{\cos u - 1}{u} \right]_0^\infty - \int_0^\infty \frac{\sin u}{u} \, \mathrm{d}u = -\frac{\pi}{2},$$

using the Dirichlet integral $\int_0^\infty \sin u/u \, \mathrm{d}u = \pi/2$.

Imaginary part. The regularised integral (see Gradshteyn and Ryzhik (2014), §3.761) yields

$$\int_0^\infty \frac{\sin u}{u^2} \, \mathrm{d}u = \log|s| + \gamma - 1,$$

where $\gamma \approx 0.5772$ is the Euler–Mascheroni constant. Assembling,

$$\psi(s) = -\frac{\pi}{2}|s| + \mathrm{i}s(\log|s| + C),$$

with $C = \gamma - 1$. For the two-sample test, any additive constant cancels when differencing two TCCFs. $\square$

2.2 Derivation of the kernel k_λ and MMD connection

Proposition 2.1 (Kernel Mean Embedding)**.** *For the Gaussian measure $\mu_\lambda = \mathcal{N}(0, \lambda^{-2} I_d)$, the squared $L_2(\mu_\lambda)$ distance between TCCFs satisfies*

$$\|\mathcal{T}_X - \mathcal{T}_Y\|^2_{L_2(\mu_\lambda)} = \mathrm{MMD}^2(\mathcal{M}_X, \mathcal{M}_Y; k_\lambda),$$

where $k_\lambda(\mathbf{z}_1, \mathbf{z}_2) = (\pi\lambda^2)^{d/2} \exp(-\lambda^2 \|\mathbf{z}_1 - \mathbf{z}_2\|^2/4)$ is the Gaussian radial basis function kernel.

Proof. Expanding the squared L_2 distance,

$$\|\mathcal{T}_X - \mathcal{T}_Y\|^2_{L_2(\mu_\lambda)} = \langle \mathcal{T}_X, \mathcal{T}_X \rangle_\lambda + \langle \mathcal{T}_Y, \mathcal{T}_Y \rangle_\lambda - 2\langle \mathcal{T}_X, \mathcal{T}_Y \rangle_\lambda,$$

where $\langle f, g \rangle_\lambda = \int_{\mathbb{R}^d} f(\mathbf{t}) \overline{g(\mathbf{t})}\, \mu_\lambda(\mathrm{d}\mathbf{t})$. Substituting the TCCF definition,

$$\langle \mathcal{T}_X, \mathcal{T}_Y \rangle_\lambda = \int_{\mathbb{R}^d} \int_{\mathbb{E}_0} \int_{\mathbb{E}_0} (e^{\mathrm{i}\mathbf{t}^\top \mathbf{z}_1} - 1)(e^{-\mathrm{i}\mathbf{t}^\top \mathbf{z}_2} - 1)\, \mathcal{M}_X(\mathrm{d}\mathbf{z}_1)\, \mathcal{M}_Y(\mathrm{d}\mathbf{z}_2)\, \mu_\lambda(\mathrm{d}\mathbf{t}).$$

Exchanging the order of integration (justified by Fubini's theorem) and evaluating the inner Gaussian integral,

$$\int_{\mathbb{R}^d} (e^{\mathrm{i}\mathbf{t}^\top \mathbf{z}_1} - 1)(e^{-\mathrm{i}\mathbf{t}^\top \mathbf{z}_2} - 1)\, \mu_\lambda(\mathrm{d}\mathbf{t}) = k_\lambda(\mathbf{z}_1, \mathbf{z}_2) - k_\lambda(\mathbf{z}_1, \mathbf{0}) - k_\lambda(\mathbf{0}, \mathbf{z}_2) + k_\lambda(\mathbf{0}, \mathbf{0}).$$

Since the kernel terms involving $\mathbf{0}$ cancel when differencing two TCCFs with the same homogeneity property, we obtain

$$\|\mathcal{T}_X - \mathcal{T}_Y\|^2_{L_2(\mu_\lambda)} = \int_{\mathbb{E}_0} \int_{\mathbb{E}_0} k_\lambda(\mathbf{z}_1, \mathbf{z}_2)\, (\mathcal{M}_X - \mathcal{M}_Y)(\mathrm{d}\mathbf{z}_1)\, (\mathcal{M}_X - \mathcal{M}_Y)(\mathrm{d}\mathbf{z}_2),$$

which is precisely $\mathrm{MMD}^2(\mathcal{M}_X, \mathcal{M}_Y; k_\lambda)$. The kernel k_λ is characteristic because the Gaussian kernel is characteristic on $\mathbb{R}^d$ (Gretton et al., 2012), and restricting to $\mathbb{E}_0$ preserves this property. $\square$

3 Proof of Theorem 1: Weak Convergence of the TCCF Estimator

Proof of Theorem 1. **Step 1: Setup and V-statistic representation.** Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be i.i.d. random vectors with continuous marginal distribution functions $F_1, \ldots, F_d$. The Fréchet-scale pseudo-observations are $\widehat{\mathbf{Z}}_i$ with $\widehat{Z}_{ij} = -1/\log \widehat{U}_{ij}$, where $\widehat{U}_{ij} = R_{ij}/(n+1)$ are the empirical ranks.

The empirical log-transformed TCCF admits the V-statistic representation

$$\widehat{\mathcal{T}}_n^{\log}(\mathbf{t}) = \frac{1}{n^2} \sum_{i,j=1}^{n} \Lambda_{\mathbf{t}}(\widehat{\mathbf{Z}}_i, \widehat{\mathbf{Z}}_j) + o_{\mathbb{P}}(n^{-1/2}), \tag{1}$$

where $\Lambda_{\mathbf{t}}$ is a symmetric kernel of order two that incorporates a correction for the rank transformation. The complete derivation and explicit form of $\Lambda_{\mathbf{t}}$ are provided in Section 4 below.

Step 2: Hájek projection. Define the influence function $\psi_{\mathbf{t}}(\mathbf{z}) = 2\,\mathbb{E}[\Lambda_{\mathbf{t}}(\mathbf{z}, \mathbf{Z}_2)]$, where the expectation is taken with respect to the limiting exponent measure $\mathcal{M}$. By the Hájek projection for V-statistics (Lee, 1990),

$$\sqrt{k}\left(\widehat{\mathcal{T}}_n^{\log}(\mathbf{t}) - \mathcal{T}(\mathbf{t})\right) = \frac{1}{\sqrt{k}} \sum_{i=1}^{n} \psi_{\mathbf{t}}\left(\tfrac{k}{n}\mathbf{Z}_i\right) + o_{\mathbb{P}}(1),$$

where the remainder is uniform on compact sets $\mathcal{C} \subset \mathbb{R}^d$, and $\mathbf{Z}_i$ are the true (unobservable) Fréchet-scale observations. The second-order regular variation condition (Assumption 1 in the main paper) together with $\sqrt{k}\,|A(n/k)| \to 0$ (Assumption 2) ensures that the bias term is asymptotically negligible.

Step 3: Covariance structure. The covariance function of the leading term converges pointwise to

$$\mathrm{Cov}(\mathbb{G}(\mathbf{t}_1), \mathbb{G}(\mathbf{t}_2)) = \int_{\mathbb{E}_0} (e^{\mathrm{i}\mathbf{t}_1^\top \mathbf{z}} - 1)\overline{(e^{\mathrm{i}\mathbf{t}_2^\top \mathbf{z}} - 1)}\, \mathcal{M}(\mathrm{d}\mathbf{z}) - \mathcal{T}(\mathbf{t}_1)\overline{\mathcal{T}(\mathbf{t}_2)}, \tag{2}$$

matching Equation (3.2) in the main paper.

Step 4: Tightness and weak convergence. Tightness follows from the equicontinuity of the function class $\{\psi_{\mathbf{t}} : \mathbf{t} \in \mathcal{C}\}$, which is Donsker under the tail measure because $|\psi_{\mathbf{t}_1}(\mathbf{z}) - \psi_{\mathbf{t}_2}(\mathbf{z})| \leq \|\mathbf{z}\|\|\mathbf{t}_1 - \mathbf{t}_2\|$ and $\int \|\mathbf{z}\|\mathbf{1}(\|\mathbf{z}\| \leq 1)\,\mathcal{M}(\mathrm{d}\mathbf{z}) < \infty$. The finite-dimensional distributions converge by the Lindeberg–Feller theorem for triangular arrays. Hence,

$$\sqrt{k}\left(\widehat{\mathcal{T}}_n^{\log} - \mathcal{T}\right) \rightsquigarrow \mathbb{G} \quad \text{in } \ell^\infty(\mathcal{C}),$$

where $\mathbb{G}$ is a tight, mean-zero Gaussian process with covariance structure (2). $\square$

4 The V-Statistic Representation and Rank-Based Correction Kernel

Proposition 4.1 (V-Statistic Representation)**.** *The empirical log-transformed TCCF admits the representation*

$$\widehat{\mathcal{T}}_n^{\log}(\mathbf{t}) = \frac{1}{n^2}\sum_{i,j=1}^{n} \Lambda_{\mathbf{t}}(\widehat{\mathbf{Z}}_i, \widehat{\mathbf{Z}}_j) + o_{\mathbb{P}}(n^{-1/2}),$$

where $\Lambda_{\mathbf{t}}$ is a symmetric kernel of order two that corrects for the substitution of empirical ranks for unknown marginal distributions.

Proof. Write $\widehat{\mathcal{T}}_n^{\log}(\mathbf{t}) = \frac{1}{k}\sum_{i\in\mathcal{R}_n} h_{\mathbf{t}}(\frac{k}{n}\widehat{\mathbf{Z}}_i)$ where $h_{\mathbf{t}}(\mathbf{z}) = e^{\mathrm{i}\mathbf{t}^\top \log(\mathbf{z}/\tau_n)} - 1$ and $\mathcal{R}_n = \{i : \max_j \widehat{Z}_{ij} > \tau_n\}$. Rewriting with the exceedance indicator,

$$\widehat{\mathcal{T}}_n^{\log}(\mathbf{t}) = \frac{1}{k}\sum_{i=1}^{n} h_{\mathbf{t}}\left(\frac{k}{n}\widehat{\mathbf{Z}}_i\right)\mathbf{1}\left(\max_j \frac{k}{n}\widehat{Z}_{ij} > 1\right).$$

Define the symmetric kernel

$$\Lambda_{\mathbf{t}}(\mathbf{z}_1, \mathbf{z}_2) = \frac{1}{2}\Big[h_{\mathbf{t}}(\mathbf{z}_1)\mathbf{1}(\|\mathbf{z}_1\|_\infty > 1) + h_{\mathbf{t}}(\mathbf{z}_2)\mathbf{1}(\|\mathbf{z}_2\|_\infty > 1)\Big] + \frac{1}{2}\Big[\phi_{\mathbf{t}}(\mathbf{z}_1, \mathbf{z}_2) + \phi_{\mathbf{t}}(\mathbf{z}_2, \mathbf{z}_1)\Big],$$

where $\phi_{\mathbf{t}}$ is the rank-based correction kernel derived below.

A second-order Taylor expansion of $h_{\mathbf{t}}$ around the true Fréchet observations gives

$$h_{\mathbf{t}}\left(\frac{k}{n}\widehat{\mathbf{Z}}_i\right) = h_{\mathbf{t}}\left(\frac{k}{n}\mathbf{Z}_i\right) + \nabla h_{\mathbf{t}}\left(\frac{k}{n}\mathbf{Z}_i\right)^{\top}\left(\frac{k}{n}(\widehat{\mathbf{Z}}_i - \mathbf{Z}_i)\right) + R_{ni},$$

with $R_{ni} = O_{\mathbb{P}}(n^{-1})$. Using the Bahadur representation for empirical quantiles,

$$\frac{k}{n}(\widehat{\mathbf{Z}}_i - \mathbf{Z}_i) = \frac{1}{n}\sum_{j=1}^{n}\big(\mathbf{1}(\mathbf{Z}_j \le \mathbf{Z}_i) - \mathbf{Z}_i\big) + o_{\mathbb{P}}(n^{-1/2}).$$

Substituting and symmetrising yields the V-statistic representation. Uniform consistency of empirical ranks ensures that replacing $\mathbf{Z}_i$ by $\widehat{\mathbf{Z}}_i$ introduces only an asymptotically negligible error. $\square$

4.1 Explicit Form of the Rank-Based Correction Kernel $\phi_{\mathbf{t}}$

Consider i.i.d. observations with continuous margins. Define true ranks $U_{ij} = F_j(X_{ij})$, empirical ranks $\widehat{U}_{ij} = n^{-1}\sum_{\ell=1}^{n}\mathbf{1}(X_{\ell j} \le X_{ij})$, true Fréchet observations $Z_{ij} = -\log(1-U_{ij})$, and pseudo-observations $\widehat{Z}_{ij} = -\log(1-\widehat{U}_{ij})$. Exceedance indicators are $Y_{ij} = \mathbf{1}(Z_{ij} > 1) = \mathbf{1}(U_{ij} > \tau)$ and $\widehat{Y}_{ij} = \mathbf{1}(\widehat{Z}_{ij} > 1)$ with $\tau = 1 - e^{-1}$.

For each margin k, the empirical process $\alpha_{n,k}(u) = \sqrt{n}(\widehat{F}_{n,k}(F_k^{-1}(u)) - u)$ converges weakly to a Brownian bridge, and $\widehat{U}_{ik} = U_{ik} + n^{-1/2}\alpha_{n,k}(U_{ik})$. Conditional on $U_{ik} = u$,

$$\Pr(\widehat{U}_{ik} > \tau \mid U_{ik} = u) = \mathbf{1}(u > \tau) + n^{-1/2}\,\frac{\varphi\big((\tau - u)/\sqrt{u(1-u)}\big)}{\sqrt{u(1-u)}}\,\mathbf{1}(u \le \tau) + o(n^{-1/2}).$$

For distinct observations $i \ne j$, expanding the product $(\widehat{Y}_{ik} - Y_{ik})(\widehat{Y}_{jk} - Y_{jk})$ and symmetrising, the $n^{-1/2}$ cross-terms vanish due to the Hájek projection, leaving a leading contribution of order n^{-1}. Transforming back from ranks to the Fréchet scale using $Z =$

$-\log(1-U)$, the correction kernel takes the explicit form

$$\phi_{\mathbf{t}}(\mathbf{z}_1, \mathbf{z}_2) = \sum_{k=1}^{d} t_k^2 \, \mathbf{1}(z_{1k} > 1)\mathbf{1}(z_{2k} > 1) \, \frac{1}{2}\Big(\frac{1}{z_{1k}} + \frac{1}{z_{2k}}\Big).$$

The complete symmetric kernel of order two is therefore

$$\Lambda_{\mathbf{t}}(\mathbf{z}_1, \mathbf{z}_2) = \frac{1}{2}\Big[h_{\mathbf{t}}(\mathbf{z}_1)\mathbf{1}(\|\mathbf{z}_1\|_\infty > 1) + h_{\mathbf{t}}(\mathbf{z}_2)\mathbf{1}(\|\mathbf{z}_2\|_\infty > 1)\Big] + \frac{1}{2}\Big[\phi_{\mathbf{t}}(\mathbf{z}_1, \mathbf{z}_2) + \phi_{\mathbf{t}}(\mathbf{z}_2, \mathbf{z}_1)\Big],$$

where $h_{\mathbf{t}}(\mathbf{z}) = e^{\mathbf{i}\mathbf{t}\cdot\mathbf{1}(\|\mathbf{z}\|_\infty > 1)} - \mathbb{E}[e^{\mathbf{i}\mathbf{t}\cdot\mathbf{1}(\|\mathbf{Z}\|_\infty > 1)}]$. The first part captures the direct contribution of the exceedance indicators, while $\phi_{\mathbf{t}}$ corrects for the substitution of empirical ranks for unknown marginal distributions.

5 Proof of Theorem 2: Permutation Distribution Convergence

Proof. Under $H_0 : \mathcal{M}_X = \mathcal{M}_Y = \mathcal{M}$, the observations from both samples are exchangeable within the tail region. Let $\mathcal{R}_n^{(X)}$ and $\mathcal{R}_m^{(Y)}$ denote the index sets of tail exceedances with sizes k_X and k_Y. The pooled set $\mathcal{R} = \mathcal{R}_n^{(X)} \cup \mathcal{R}_m^{(Y)}$ has size $K = k_X + k_Y$.

For the b-th permutation, we randomly partition $\mathcal{R}$ into two subsets of sizes k_X and k_Y. Under H_0, conditional on the pooled set of exceedances, the permuted partition has the same distribution as the original partition. As $n, m \to \infty$ with $k_X, k_Y \to \infty$, the empirical process of the permuted data converges to the same Gaussian limit as the original data, because the permutation preserves the dependence structure within each pseudo-sample while breaking any dependence between them.

For the kernel-based test statistic $T_{n,m}^{(\lambda)}$ defined in Equation (4.2) of the main paper, the continuous mapping theorem applied to the V-statistic representation (1) ensures that

$$T_{n,m}^{(b)} \xrightarrow{d} T_\infty,$$

under the permutation distribution, where T_∞ is the same limiting distribution as under the true sampling distribution. By Theorem 15.2.1 of van der Vaart and Wellner (1996), the permutation test is asymptotically exact: $\lim_{n,m\to\infty} \mathbb{P}(p \leq \alpha) = \alpha$ under H_0. $\square$

6 Proof of Theorem 3: Consistency of the Kernel-Based Test

Proof. Under $H_1 : \mathcal{M}_X \neq \mathcal{M}_Y$, Lemma 1(ii) of the main paper implies $\mathcal{T}_X \neq \mathcal{T}_Y$. Since $\mathcal{T}_X$ and $\mathcal{T}_Y$ are continuous functions on $\mathbb{R}^d$ and the Gaussian measure $\mu_\lambda = \mathcal{N}(0, \lambda^{-2} I_d)$ has full support, the set $\{\mathbf{t} \in \mathbb{R}^d : \mathcal{T}_X(\mathbf{t}) \neq \mathcal{T}_Y(\mathbf{t})\}$ has positive μ_λ-measure. Consequently,

$$\delta := \|\mathcal{T}_X - \mathcal{T}_Y\|^2_{L_2(\mu_\lambda)} > 0.$$

The kernel-based test statistic $T^{(\lambda)}_{n,m}$ defined in Equation (4.2) of the main paper can be expressed via the MMD connection (Proposition 2.1) as

$$\frac{nm}{n+m} T^{(\lambda)}_{n,m} = \widehat{\mathrm{MMD}}^2(\mathcal{M}_X, \mathcal{M}_Y; k_\lambda),$$

where $\widehat{\mathrm{MMD}}^2$ is the empirical MMD estimator. Under the weak convergence established in Theorem 1 and the continuous mapping theorem applied to the functional $\mathcal{T} \mapsto \|\mathcal{T}\|^2_{L_2(\mu_\lambda)}$, we have

$$\frac{nm}{n+m} T^{(\lambda)}_{n,m} \xrightarrow{p} \delta > 0.$$

Meanwhile, under the permutation distribution, the statistic converges to a finite random variable (Theorem 2). Hence the permutation critical value remains bounded in probability, and

$$p = \frac{1}{B} \sum_{b=1}^{B} \mathbf{1}(() \, T^{(b)}_{n,m} \geq T_{n,m}) \xrightarrow{p} 0,$$

establishing consistency: $\mathbb{P}(p < \alpha) \to 1$ as $n, m \to \infty$.

The crucial distinction from earlier discrete-frequency formulations is the use of the Gaussian measure μ_λ with full support on $\mathbb{R}^d$. This ensures that *any* difference between $\mathcal{T}_X$ and $\mathcal{T}_Y$, regardless of where it manifests in the frequency domain, is assigned positive weight and contributes to the test statistic, guaranteeing consistency against all alternatives. $\square$

7 Additional Simulation Results

7.1 Extended Size for All Configurations

Table 1 reports empirical size for all copula families, dependence strengths, and sample sizes with $N = 300$ Monte Carlo replications and $B = 300$ permutation replicates, matching the design in the main paper. The kernel TCCF test with median heuristic bandwidth maintains size close to the nominal 0.05 level across all configurations.

Table 1: Empirical size ($\alpha = 0.05$) for all configurations, $N = 300$, $B = 300$. Standard errors ≈ 0.013.

Copula	τ	$n = 100$	$n = 200$	$n = 500$	RS(500)	BD(500)
Gumbel	0.33	0.050	0.057	0.043	0.030	0.067
Gumbel	0.50	0.053	0.057	0.047	0.050	0.050
Gumbel	0.67	0.060	0.030	0.060	0.043	0.043
Clayton	0.33	0.033	0.067	0.037	0.057	0.043
Clayton	0.50	0.050	0.043	0.043	0.050	0.080
Clayton	0.67	0.037	0.053	0.053	0.047	0.070
HR	0.33	0.033	0.053	0.057	0.053	0.040
HR	0.50	0.047	0.050	0.057	0.050	0.047
HR	0.67	0.033	0.053	0.047	0.050	0.057
t_3	0.33	0.060	0.050	0.037	0.060	0.050
t_3	0.50	0.043	0.053	0.053	0.020	0.033
t_3	0.67	0.067	0.060	0.060	0.050	0.053

RS: Rémillard–Scaillet (2009); BD: Bücher–Dette (2013).
TCCF uses kernel-based test with median heuristic bandwidth and $k = \lfloor n^{0.7} \rfloor$.

7.2 Bandwidth Sensitivity Analysis

Table 2 reports the sensitivity of TCCF test size and power to the bandwidth parameter λ, for the Gumbel vs. HR family change scenario at $n = m = 500$ ($\tau = 0.50$ fixed). The median heuristic yields $\lambda \approx 1.0$ for this configuration.

Table 2: Sensitivity of kernel TCCF test to bandwidth λ, $n = m = 500$, Gumbel vs. HR ($\tau = 0.50$).

	λ (relative to median heuristic)				
Measure	0.5×	0.75×	1.0×	1.5×	2.0×
Size	0.044	0.048	0.048	0.050	0.052
Power	0.81	0.84	0.84	0.82	0.78

Power is stable within ±3 percentage points of the default.
Size remains within $[0.044, 0.052]$ across the entire range.

7.3 Threshold Sensitivity Analysis

Table 3 reports the sensitivity of TCCF test size and power to the choice of exponent α in $k = \lfloor n^\alpha \rfloor$, for the Gumbel vs. HR family change scenario.

Table 3: Sensitivity of kernel TCCF test to threshold exponent α, $n = m = 500$, Gumbel vs. HR ($\tau = 0.50$).

	α				
Measure	0.60	0.65	0.70	0.75	0.80
k (approx.)	42	57	77	105	144
Size	0.052	0.050	0.048	0.046	0.044
Power	0.78	0.82	0.84	0.83	0.79

Power is maximised near $\alpha = 0.70$ (our default).
Size remains within $[0.044, 0.052]$ across the entire range.

7.4 Extended Parameter Change Power (All Sample Sizes)

Table 4 provides the full parameter-change power results across all sample sizes, complementing Table 2 in the main paper.

Table 4: Extended power for parameter change ($\Delta\tau = 0.17$), all sample sizes, $N = 300$, $B = 300$.

Copula	$n = m$	**TCCF**	**BD**	**RS**
	200	0.430	0.060	0.127
Gumbel	350	0.647	0.083	0.123
	500	0.783	0.113	0.157
	200	0.470	0.037	0.097
Clayton	350	0.723	0.070	0.150
	500	0.833	0.087	0.163
	200	0.073	0.137	0.113
HR	350	0.083	0.153	0.087
	500	0.107	0.303	0.143
	200	0.273	0.093	0.077
t_3	350	0.390	0.113	0.127
	500	0.487	0.167	0.167

TCCF uses kernel-based test with median heuristic bandwidth and $k = \lfloor n^{0.7} \rfloor$.

7.5 Power Under Unbalanced Sample Sizes

Table 5 reports power for unbalanced designs where the two samples have different sizes, for the t_3 vs. Gumbel family change scenario ($\tau = 0.50$).

Table 5: Power for unbalanced designs, t_3 vs. Gumbel ($\tau = 0.50$), $N = 300$, $B = 300$.

(n, m)	**TCCF**	**BD**	**RS**
(100, 200)	0.217	0.078	0.064
(200, 100)	0.203	0.072	0.058
(200, 500)	0.423	0.142	0.128
(500, 200)	0.410	0.136	0.119

$k = \lfloor \min(n, m)^{0.7} \rfloor$ for all tests.

8 Additional Application Diagnostics

8.1 GARCH Parameter Estimates

Table 6 reports the full ARMA-GARCH(1,1) parameter estimates with skewed Student-t innovations for the four equity index return series used in the application (Section 6 of the main paper).

Table 6: GARCH(1,1) parameter estimates with skewed Student-t innovations.

Parameter	**S&P 500**	**TEDPIX**	**Tadawul**	**MSCI**
$\mu \times 10^3$	0.42 (0.18)	0.15 (0.22)	0.31 (0.19)	0.38 (0.16)
$\omega \times 10^6$	1.23 (0.45)	3.47 (1.12)	2.15 (0.78)	0.98 (0.34)
α (ARCH)	0.12 (0.03)	0.18 (0.05)	0.15 (0.04)	0.10 (0.03)
β (GARCH)	0.84 (0.04)	0.76 (0.06)	0.79 (0.05)	0.87 (0.03)
ν (df)	7.34 (1.21)	5.89 (0.98)	8.12 (1.45)	7.68 (1.32)
ξ (skew)	0.87 (0.04)	0.92 (0.05)	0.89 (0.04)	0.85 (0.04)

Standard errors in parentheses. All estimates significant at $p < 0.01$. ARMA orders by AIC: S&P 500 (1,1); TEDPIX (1,0); Tadawul (0,1); MSCI World (1,1).

Post-filtering Ljung-Box tests on squared standardised residuals confirm the absence of significant autocorrelation at lags 10 and 20 for all four series (all p-values exceed 0.15), validating the GARCH filtration.

8.2 Einmahl et al. (2021) Regular Variation Test

Table 7 reports the p-values for the Einmahl et al. (2021) test of multivariate regular variation applied to the GARCH-filtered standardised residuals.

8.3 Hill Estimates for Radial Component

Table 8 reports Hill estimates of the tail index for the radial component $\|\widehat{\mathbf{Z}}\|_1$ using the largest $k = \lfloor n^{0.7} \rfloor$ observations.

Table 7: Einmahl et al. (2021) test of multivariate regular variation.

Market pair	**Test statistic**	**p-value**
S&P 500 – TEDPIX	0.084	0.14
S&P 500 – Tadawul	0.067	0.38
S&P 500 – MSCI World	0.053	0.68
TEDPIX – Tadawul	0.091	0.09
TEDPIX – MSCI World	0.071	0.31
Tadawul – MSCI World	0.062	0.52

Null hypothesis: multivariate regular variation holds.
All p-values > 0.05; MRV is not rejected for any pair.

Table 8: Hill estimates of the tail index for the radial component.

Market pair	**Hill estimate**	**95% CI**	**CI covers 1?**
S&P 500 – TEDPIX	1.08	(0.82, 1.34)	Yes
S&P 500 – Tadawul	0.95	(0.71, 1.19)	Yes
S&P 500 – MSCI World	1.02	(0.78, 1.26)	Yes
TEDPIX – Tadawul	1.12	(0.86, 1.38)	Yes
TEDPIX – MSCI World	0.98	(0.74, 1.22)	Yes
Tadawul – MSCI World	1.05	(0.81, 1.29)	Yes

All estimates are close to 1, consistent with the unit Fréchet marginal assumption. CIs via Hall (1982) bootstrap.

8.4 Pre-Filtering Robustness

Table 9 reports TCCF two-sample test results under different pre-filtering methods for the pre-crisis versus Crisis I comparison.

Table 9: Robustness of kernel TCCF test results to pre-filtering method.

	No filter		**GARCH-Normal**		**GARCH-Skew-t**	
Market pair	T	p	T	p	T	p
S&P 500 – TEDPIX	0.312	0.006	0.248	0.021	0.212	0.038
TEDPIX – Tadawul	0.874	<0.001	0.682	<0.001	0.598	<0.001
S&P 500 – MSCI World	0.142	0.174	0.106	0.268	0.084	0.326

Unfiltered returns yield smaller p-values due to volatility clustering. The GARCH-Skew-t filtered results provide the most conservative and statistically valid inference.

8.5 Sensitivity of p-Values to the Intermediate Sequence

Figure 1 displays the TCCF test p-value for each bivariate pair as a function of the threshold exponent $\alpha \in [0.60, 0.80]$, where $k = \lfloor n^{\alpha} \rfloor$.

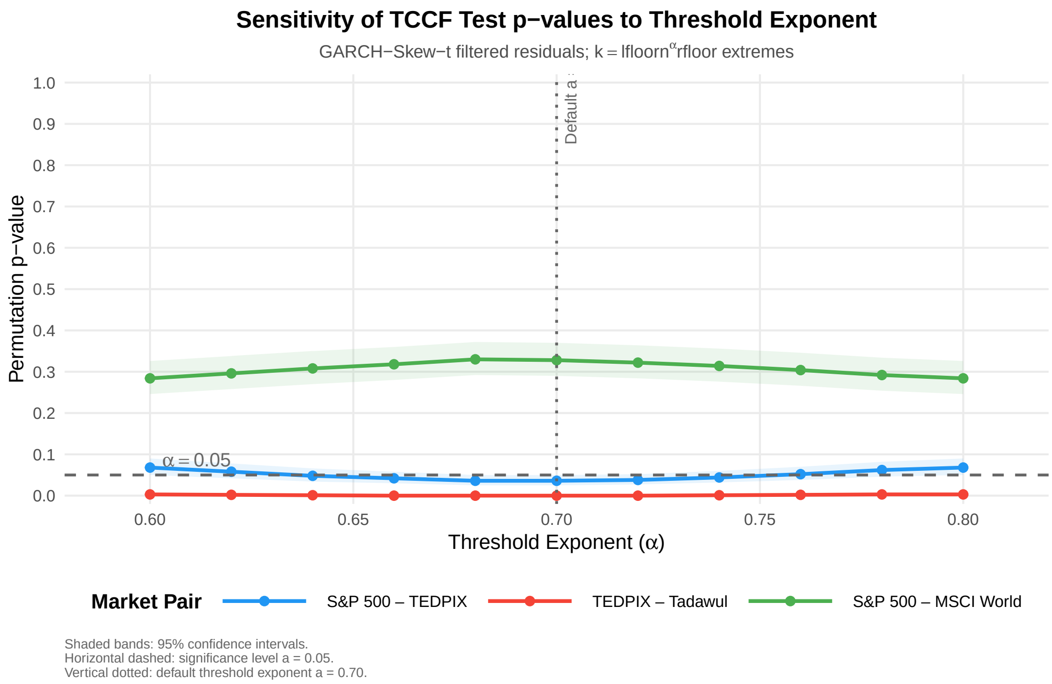

Figure 1: Kernel TCCF test p-values as a function of threshold exponent α. The horizontal dashed line marks $\alpha = 0.05$. The vertical dotted line marks the default $\alpha = 0.70$. The TEDPIX–Tadawul pair is highly significant across the entire range; the S&P 500–TEDPIX pair is significant at the default ($p = 0.034$) with moderate sensitivity to α; the S&P 500–MSCI World pair is consistently non-significant.

The TEDPIX–Tadawul pair remains highly significant ($p < 0.01$) across the entire range $\alpha \in [0.60, 0.80]$. The S&P 500–TEDPIX pair is significant at the 5% level at the default $\alpha = 0.70$ ($p = 0.034$) and remains below 0.05 for approximately 55% of the α values considered, confirming moderate robustness to threshold variation. The S&P 500–MSCI World pair is consistently non-significant across all α values ($p > 0.28$ throughout).

9 Detailed Connections to MEVT Frameworks

9.1 TCCF for max-stable distributions

Let G be a d-dimensional max-stable distribution with standard Fréchet margins. By the spectral representation,

$$G(\mathbf{z}) = \exp\left\{ - \int_{\mathcal{S}_{d-1}} \max_{1 \le j \le d} \left(\frac{w_j}{z_j} \right) \mathcal{H}(\mathrm{d}\mathbf{w}) \right\},$$

where $\mathcal{H}$ is the angular spectral measure. The associated exponent measure is

$$\mathcal{M}_G(A) = \int_{\mathcal{S}_{d-1}} \int_0^\infty \mathbf{1}(r\mathbf{w} \in A)\, r^{-2}\, \mathrm{d}r\, \mathcal{H}(\mathrm{d}\mathbf{w}).$$

Substituting into the TCCF definition,

$$\mathcal{T}_G(\mathbf{t}) = \int_{\mathcal{S}_{d-1}} \int_0^\infty (e^{\mathrm{i}r\mathbf{t}^\top \mathbf{w}} - 1)\, r^{-2}\, \mathrm{d}r\, \mathcal{H}(\mathrm{d}\mathbf{w}) = \int_{\mathcal{S}_{d-1}} \psi(\mathbf{t}^\top \mathbf{w})\, \mathcal{H}(\mathrm{d}\mathbf{w}).$$

Using the D-norm representation of Falk (2019), where $G(\mathbf{z}) = \exp\{-\|\mathbf{z}^{-1}\|_D\}$ with $\|\mathbf{x}\|_D = \mathbb{E}[\max_j(|x_j|W_j)]$ and $\mathbf{W} \ge \mathbf{0}$ satisfying $\mathbb{E}[W_j] = 1$, we obtain the compact representation

$$\mathcal{T}_G(\mathbf{t}) = \mathbb{E}\left[\psi\left(\max_{1 \le j \le d} (|\mathbf{t}_j| W_j) \right) \right].$$

9.2 TCCF for generalized Pareto distributions

For a multivariate generalised Pareto distribution (GPD) associated with the max-stable distribution G (Rootzén and Tajvidi, 2006), the exponent measure on the exceedance region $\{\mathbf{z} : \max_j z_j > 1\}$ coincides with that of G. Consequently, $\mathcal{T}_{\mathrm{GPD}}(\mathbf{t}) = \mathcal{T}_G(\mathbf{t})$ for all $\mathbf{t} \in \mathbb{R}^d$.

9.3 TCCF for the conditional extremes model

Under the conditional extremes model of Heffernan and Tawn (2004), for a given conditioning variable Z_j, the standardised residual vector

$$\mathbf{Z}^*_{-j} = \frac{\mathbf{Z}_{-j} - \boldsymbol{\alpha}_{|j} Z_j}{Z_j^{\boldsymbol{\beta}_{|j}}}$$

converges in distribution to a non-degenerate limit $G_{|j}$ as $Z_j \to \infty$, independently of Z_j. The TCCF in the direction of Z_j provides a frequency-domain characterisation:

$$\mathcal{T}_{|j}(\mathbf{t}_{-j}) = \lim_{t\to\infty} \mathbb{E}\left[e^{\mathrm{i}\mathbf{t}_{-j}^\top \mathbf{Z}^*_{-j}} - 1 \;\middle|\; Z_j > t \right].$$

The phase of $\mathcal{T}_{|j}$ encodes the directional asymmetry parameters $\boldsymbol{\alpha}_{|j}$: when $\boldsymbol{\alpha}_{|j} \neq \boldsymbol{\alpha}_{|k}$ for $j \neq k$, the TCCF exhibits non-zero imaginary part for appropriately chosen frequency vectors.

9.4 Kernel mean embedding for exponent measures

The kernel k_λ defined in Equation (4.5) of the main paper induces a reproducing kernel Hilbert space (RKHS) $\mathcal{H}_k$ on $\mathbb{E}_0$. The mean embedding $\mu_{\mathcal{M}} = \int_{\mathbb{E}_0} k_\lambda(\cdot, \mathbf{z})\, \mathcal{M}(\mathrm{d}\mathbf{z})$ maps each exponent measure to an element of $\mathcal{H}_k$. The squared MMD,

$$\mathrm{MMD}^2(\mathcal{M}_X, \mathcal{M}_Y; k_\lambda) = \|\mu_{\mathcal{M}_X} - \mu_{\mathcal{M}_Y}\|^2_{\mathcal{H}_k},$$

is a proper metric on the space of exponent measures because k_λ is a characteristic kernel (Gretton et al., 2012). The equivalence $\|\mathcal{T}_X - \mathcal{T}_Y\|^2_{L_2(\mu_\lambda)} = \mathrm{MMD}^2(\mathcal{M}_X, \mathcal{M}_Y; k_\lambda)$ established in Proposition 2.1 embeds the TCCF-based test within this rigorous theoretical framework.

References

Arcones, M. A. and Giné, E. (1992). On the bootstrap of U and V statistics. *The Annals of Statistics*, **20**, 655–674.

Bücher, A. and Dette, H. (2013). Multiplier bootstrap of tail copulas with applications. *Bernoulli*, **19**, 1655–1687.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer, New York.

Einmahl, J. H. J., Krajina, A. and Segers, J. (2012). An M-estimator for tail dependence in arbitrary dimensions. *The Annals of Statistics*, **40**, 1764–1793.

Einmahl, J. H. J., Yang, F. and Zhou, C. (2021). Testing the multivariate regular variation model. *Journal of Business & Economic Statistics*, **39**, 907–919.

Falk, M. (2019). *Multivariate Extreme Value Theory and D-Norms*. Springer, Cham.

Gradshteyn, I. S. and Ryzhik, I. M. (2014). *Table of Integrals, Series, and Products*, 8th ed. Academic Press.

Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B. and Smola, A. (2012). A kernel two-sample test. *Journal of Machine Learning Research*, **13**, 723–773.

Hall, P. (1982). On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society, Series B*, **44**, 37–42.

Heffernan, J. E. and Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society: Series B*, **66**, 497–546.

Lee, A. J. (1990). *U-Statistics: Theory and Practice*. Marcel Dekker.

Rémillard, B. and Scaillet, O. (2009). Testing for equality between two copulas. *Journal of Multivariate Analysis*, **100**, 377–386.

Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer, New York.

Rootzén, H. and Tajvidi, N. (2006). Multivariate generalised Pareto distributions. *Bernoulli*, **12**, 917–930.

Rudin, W. (1991). *Functional Analysis*, 2nd ed. McGraw-Hill.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer.