Gulf Journal of Mathematics
Vol 24, Issue 2 (2026) 1-21

ISSN: 2309-4966
https://doi.org/10.56947/f3ad6q57

ENTROPIC COPULAS VIA VARIATIONAL SINKHORN METHODS

REMI GUILLAUME BAGRÉ[1*], HERMAN TIEMTORÉ[2], JACQUES KABORÉ[3], SAYOUBA TRAORÉ[4] and SOULEYMANE OUÉDRAOGO[5]

Abstract. We introduce the class of *Sinkhorn-Wasserstein copulas*, defined as minimizers of a variational problem combining a linear dependence functional with an entropic regularization term penalizing deviations from the independence copula. Building on the classical framework of optimal transport and its entropic regularization, we establish existence, uniqueness, and an explicit Schrödinger-bridge representation of the optimizer. We further show that the resulting one-parameter family forms a geodesic on the space of absolutely continuous copulas endowed with the Fisher–Rao metric, interpolating continuously between perfect dependence and independence. Stability estimates and plug-in statistical guarantees round out the framework.

Keywords. Copulas, entropic regularization, optimal transport, Sinkhorn divergence, Fisher-Rao geometry.
2020 Mathematics Subject Classification. 62H05; 49Q22; 60E05; 94A17.

1. Introduction

The statistical characterization of dependence structures stands as one of the most fertile and demanding problems in modern probability and mathematical statistics. Since the foundational work of Sklar [9], copulas have provided an elegant and canonical answer: any multivariate distribution decomposes into its marginals and a single object, the copula, that encodes the entirety of the dependence structure independently of the marginal laws. This separation principle, deceptively simple in its statement, has generated an extraordinarily rich mathematical theory [6] and has become indispensable in fields as diverse as quantitative finance [4], hydrology, and more recently machine learning.

Yet, despite decades of active development, a fundamental tension persists at the heart of copula theory. On one hand, parametric families : Gaussian, Clayton, Gumbel, and their Archimedean relatives [5, 6] offer tractability at the cost of rigid structural assumptions that are routinely violated in practice. On the other hand, fully nonparametric approaches, including empirical copulas and

Date: Received: Aug 5, 2026; Revised: Sep 2, 2026; Accepted: Sep 9, 2026.
* Corresponding author
© The Author(s) 2026. This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of the licence, visit https://creativecommons.org/licenses/by-nc-nd/4.0/.

kernel-based estimators, suffer from the curse of dimensionality and provide little geometric insight into the nature of dependence. The field thus finds itself caught between tractability and expressiveness, with no principled resolution in sight.

Simultaneously, the landscape of applied mathematics has been transformed over the past two decades by the rise of optimal transport theory. Following the foundational contributions of Villani [10] and Santambrogio [8], the Wasserstein geometry has reshaped our understanding of probability measures as geometric objects, amenable to interpolation, differentiation, and optimization. The introduction of entropic regularization by Cuturi [2] made these ideas computationally accessible, and the subsequent development of Sinkhorn divergences by Feydy et al. [3] provided a statistically well-founded debiased alternative. These advances, however, have remained largely disconnected from copula theory: the specific structural constraints imposed by the uniform-marginal condition defining $\mathcal{C}_d$ have no natural counterpart in the standard optimal transport framework, and bridging this gap requires genuinely new ideas.

A third strand of mathematics, information geometry as systematized by Amari [1], offers the missing conceptual link. By endowing families of probability distributions with a Riemannian structure derived from the Fisher information, information geometry transforms statistical inference into a problem of differential geometry. The dually flat structure of exponential families, the e- and m-geodesics, and the associated Pythagorean theorems have proved enormously fruitful in statistical physics, neural computation, and estimation theory. Their restriction to the manifold $\mathcal{C}_d^{\mathrm{ac}}$ of absolutely continuous copulas, however, has never been systematically developed, leaving a conspicuous void at the intersection of two mature theories.

The present paper proposes to fill this void through a unified variational framework. The central object of study is the one-parameter family $\{C_\varepsilon\}_{\varepsilon>0}$ of copulas defined as solutions to the penalized variational problem

$$C_\varepsilon \;\in\; \operatorname*{arg\,min}_{C\in\mathcal{C}_d^{\mathrm{ac}}} \left\{ \int_{\mathbb{I}^d} \Phi(u)\, dC(u) \;+\; \varepsilon\, \mathrm{KL}(C \,\|\, \Pi) \right\}, \tag{1.1}$$

where $\Phi : \mathbb{I}^d \to \mathbb{R}$ is a measurable dependence potential, Π denotes the independence copula (Definition 2.4), and KL is the Kullback–Leibler divergence (Definition 2.5). The regularization parameter $\varepsilon > 0$ governs the trade-off between fitting the potential Φ and deviating from independence. This formulation is directly inspired by the entropic optimal transport problem [2, 7] and by the Schrödinger bridge problem [10], yet it is fundamentally distinct from both: the feasible set $\mathcal{C}_d^{\mathrm{ac}}$ imposes d uniform-marginal constraints that are absent from the classical transport setting and that profoundly alter the geometry of the problem.

This paper develops a self-contained variational theory for entropically regularized copulas and establishes five principal results. We first prove that problem (1.1) admits a unique solution C_ε for every $\varepsilon > 0$, whose density factors as the product of the Gibbs kernel $e^{-\Phi/\varepsilon}$ and d univariate calibration functions

enforcing the marginal constraints a structure reminiscent of the Sinkhorn scaling algorithm [2] and interpretable as a constrained Schrödinger bridge (Theorem 3.2). We then characterize the limiting regimes $\varepsilon \to \infty$ and $\varepsilon \to 0^+$ via a Γ-convergence argument, showing that the family interpolates continuously between the independence copula [6] and the copula concentrated on the minimizers of Φ (Proposition 3.3). At the level of dynamics, we establish that the density c_ε satisfies, in the variable ε, a distributional evolution equation of Fokker–Planck type supplemented by a marginal-constraint correction term characterized through the tangent space of $\mathcal{C}_d^{\mathrm{ac}}$ (Theorem 3.4 and Lemma 2.9). From a geometric standpoint, endowing $\mathcal{C}_d^{\mathrm{ac}}$ with the Fisher-Rao metric in the sense of Amari [1], we prove that the curve $\varepsilon \mapsto C_\varepsilon$ is, up to reparametrization, a geodesic on this statistical manifold, thereby establishing a precise and previously unknown bridge between entropic copula regularization and the e-geometry of information geometry (Theorem 3.5). Finally, we derive quantitative stability bounds showing that C_ε depends Lipschitz continuously on Φ in the KL topology, and deduce oracle-type convergence rates for the plug-in estimator $\widehat{C}_{\varepsilon,n}$ obtained by substituting a nonparametric estimate $\widehat{\Phi}_n$ for the potential Φ (Theorem 3.6 and Corollary 3.7). The remainder of the paper is organized as follows. Section 2 collects the necessary background on copulas [9, 5, 6], entropic optimal transport [2, 7, 3], and information geometry [1]. The main results are stated and proved in Section 3, organized around the three themes of variational analysis, PDE characterization, and geometric structure.

2. Preliminaries

We gather here the notation and the background material used throughout the paper. Unless otherwise stated, $d \geq 2$ is a fixed integer and $\mathbb{I}^d := [0,1]^d$. The presentation borrows from the standard references on copulas [4, 5, 6], optimal transport [10, 8, 7], and information geometry [1].

2.1. Copulas and Sklar's theorem.

Definition 2.1 (Copula, [6])**.** Let $\mathbb{I} = [0,1]$ and $\mathbb{I}^d = [0,1]^d$ for a fixed integer $d \geq 2$. A function $C : \mathbb{I}^d \to \mathbb{I}$ is called a *d-copula* if it satisfies the following three conditions:

(1) Boundary conditions : for every $u = (u_1, \ldots, u_d) \in \mathbb{I}^d$ and every index $i \in \{1, \ldots, d\}$,

$$C(u_1, \ldots, u_{i-1}, 0, u_{i+1}, \ldots, u_d) \;=\; 0,$$

and

$$C(1, \ldots, 1, u_i, 1, \ldots, 1) \;=\; u_i.$$

(2) Uniform marginals : for each $i \in \{1, \ldots, d\}$, the i-th marginal of C is the uniform distribution on $\mathbb{I}$, that is,

$$\int_{\mathbb{I}^{d-1}} dC(u) \;=\; u_i, \qquad \forall\, u_i \in \mathbb{I},$$

where the integration is taken over all variables except u_i.

(3) d-increasing property : for every hyper-rectangle $\prod_{i=1}^{d}[a_i, b_i] \subseteq \mathbb{I}^d$, the C-volume is non-negative:

$$V_C\left(\prod_{i=1}^{d}[a_i, b_i]\right) = \sum_{\mathbf{v}\in\{a_1,b_1\}\times\cdots\times\{a_d,b_d\}} (-1)^{card\{i\,:\,v_i=a_i\}}\, C(\mathbf{v}) \;\geq\; 0.$$

Equivalently, C is the joint cumulative distribution function of a random vector $(U_1, \ldots, U_d)$ where each marginal $U_i \sim \mathcal{U}(\mathbb{I})$. The collection of all d-copulas is denoted $\mathcal{C}_d$.

Remark 2.2. The set $\mathcal{C}_d$ is a convex and compact subset of $L^1(\mathbb{I}^d)$ bounded pointwise by the *Fréchet–Hoeffding bounds*: for every $u \in \mathbb{I}^d$,

$$W(u) \;\leq\; C(u) \;\leq\; M(u),$$

where $W(u) = \max\big(\sum_{i=1}^{d} u_i - (d-1),\, 0\big)$ is the lower bound (the *countermonotone copula*, sharp only for $d = 2$) and $M(u) = \min(u_1, \ldots, u_d)$ is the upper bound (the *comonotone* or *min-copula*), see [6].

Theorem 2.3 (Sklar, [9])**.** *Let H be a d-dimensional distribution function with marginals $F_1, \ldots, F_d$. There exists a copula $C \in \mathcal{C}_d$ such that, for every $(x_1, \ldots, x_d) \in \mathbb{R}^d$,*

$$H(x_1, \ldots, x_d) \;=\; C\big(F_1(x_1), \ldots, F_d(x_d)\big).$$

If the F_i are continuous, then C is unique.

Definition 2.4 (Independence copula, [6])**.** The *independence copula* is defined by

$$\Pi(u_1, \ldots, u_d) = \prod_{i=1}^{d} u_i.$$

2.2. **Entropic regularization and Sinkhorn divergences.**

Definition 2.5 (Relative entropy, [10])**.** For two probability measures μ, ν on a Polish space $\mathcal{X}$ with $\mu \ll \nu$, the *Kullback–Leibler divergence* of μ with respect to ν is

$$\mathrm{KL}(\mu \,\|\, \nu) = \int_{\mathcal{X}} \log\left(\frac{d\mu}{d\nu}\right) d\mu.$$

Definition 2.6 (Entropic optimal transport, [2, 7])**.** Given a cost function $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, marginals μ, ν, and $\varepsilon > 0$, the entropic OT problem reads

$$\mathrm{OT}_\varepsilon(\mu, \nu) \;=\; \inf_{\pi\in\Pi(\mu,\nu)} \int c\, d\pi + \varepsilon\, \mathrm{KL}(\pi \,\|\, \mu \otimes \nu).$$

The associated debiased version is the *Sinkhorn divergence* [3].

2.3. **Information geometry on $\mathcal{C}_d$.**

Definition 2.7 (Absolutely continuous copula, [6])**.** A copula $C \in \mathcal{C}_d$ is called *absolutely continuous* if it admits a density $c \in L^1(\mathbb{I}^d)$ with respect to the Lebesgue measure. We write $\mathcal{C}_d^{\mathrm{ac}}$ for the corresponding subset.

Definition 2.8 (Fisher–Rao metric on $\mathcal{C}_d^{\mathrm{ac}}$, [1])**.** For $C \in \mathcal{C}_d^{\mathrm{ac}}$ with density c, and tangent vectors $h_1, h_2 \in T_C\mathcal{C}_d^{\mathrm{ac}}$ (i.e., perturbations preserving uniform marginals and total mass), the *Fisher–Rao inner product* is

$$\langle h_1, h_2 \rangle_{\mathrm{FR},C} = \int_{\mathbb{I}^d} \frac{h_1(u)\, h_2(u)}{c(u)}\, du.$$

Lemma 2.9 (Tangent space characterization, adapted from [1, 6])**.** *The tangent space $T_C\mathcal{C}_d^{\mathrm{ac}}$ coincides with the space of $L^2(c^{-1}\,du)$ functions h on $\mathbb{I}^d$ satisfying $\int h\, du = 0$ together with the d uniform-marginal constraints*

$$\int h(u)\, du_{-i} = 0, \qquad i = 1, \dots, d,$$

where du_{-i} denotes integration over all variables except u_i.

3. Main results

Throughout this section, let $\Phi : \mathbb{I}^d \to \mathbb{R}$ be a measurable dependence potential satisfying the integrability condition $\int e^{-\Phi/\varepsilon}\, du < \infty$ for every $\varepsilon > 0$.

3.1. **The entropic copula flow.**

Definition 3.1 (Sinkhorn–Wasserstein copula)**.** For $\varepsilon > 0$, a *Sinkhorn–Wasserstein copula* associated with the potential Φ is any solution of

$$C_\varepsilon \;\in\; \operatorname*{arg\,min}_{C \in \mathcal{C}_d^{\mathrm{ac}}} \left\{ \int_{\mathbb{I}^d} \Phi(u)\, dC(u) \;+\; \varepsilon\, \mathrm{KL}(C \,\|\, \Pi) \right\}. \tag{3.1}$$

Theorem 3.2 (Existence, uniqueness, and Schrödinger structure)**.** *Under the integrability assumption above and provided $\Phi \in L^1(\Pi)$, the variational problem (3.1) admits a unique solution $C_\varepsilon \in \mathcal{C}_d^{\mathrm{ac}}$. Moreover, its density takes the multiplicative form*

$$c_\varepsilon(u) \;=\; \exp\!\left(-\frac{\Phi(u)}{\varepsilon}\right) \prod_{i=1}^{d} a_i^{(\varepsilon)}(u_i),$$

where the potentials $a_i^{(\varepsilon)} : [0,1] \to (0,\infty)$ are uniquely determined (up to the natural scaling redundancy) by the uniform-marginal constraints.

Proof.
(a) Existence and uniqueness.
Define the functional

$$\mathcal{F}_\varepsilon(C) = \int_{\mathbb{I}^d} \Phi(u)\, dC(u) \;+\; \varepsilon\, \mathrm{KL}(C \,\|\, \Pi),$$

on the convex set

$$\mathcal{C}_d^{\mathrm{ac}} = \Big\{ C \ll \lambda_d \;:\; C \in \mathcal{C}_d \Big\},$$

where λ_d denotes the Lebesgue measure on $\mathbb{I}^d$. Writing $dC = c\,d\lambda_d$ and $d\Pi = d\lambda_d$, the functional reads

$$\mathcal{F}_\varepsilon(c) \;=\; \int_{\mathbb{I}^d} \Phi(u)\,c(u)\,du \;+\; \varepsilon \int_{\mathbb{I}^d} c(u)\log c(u)\,du.$$

The map $t \mapsto t\log t$ is strictly convex on $(0,\infty)$, so $\mathcal{F}_\varepsilon$ is strictly convex on the affine subspace

$$\mathcal{A} = \Big\{ c \in L^1(\mathbb{I}^d) \;:\; c \geq 0,\; \int_{\mathbb{I}^d} c\,du = 1,\; \int_{\mathbb{I}^{d-1}} c(u)\,du_{-i} = 1 \text{ for } i = 1,\dots,d \Big\},$$

where du_{-i} denotes integration over all coordinates except u_i. Since $\Phi \in L^1(\Pi)$ and the entropy term $\varepsilon \int c\log c\,du$ is bounded below by $-\varepsilon/e$, the functional $\mathcal{F}_\varepsilon$ is coercive on $\mathcal{A}$ in the sense that $\mathcal{F}_\varepsilon(c_n) \to +\infty$ whenever $\|c_n\|_{L^1\log L^1} \to \infty$. Moreover, $c \mapsto \int c\log c\,du$ is lower semicontinuous with respect to weak convergence in $L^1(\mathbb{I}^d)$ by the Dunford–Pettis theorem and Fatou's lemma. The set $\mathcal{A}$ is weakly closed in $L^1(\mathbb{I}^d)$. The direct method of the calculus of variations therefore yields existence of a minimizer, and strict convexity of $\mathcal{F}_\varepsilon$ ensures its uniqueness.

(b) Euler-Lagrange conditions and Schrödinger structure.

Since $\mathcal{F}_\varepsilon$ is Gâteaux-differentiable on the interior of $\mathcal{A}$, the unique minimizer c_ε satisfies the first-order optimality conditions. Introducing Lagrange multipliers $\mu \in \mathbb{R}$ for the total-mass constraint and $\lambda_i \in L^\infty([0,1])$ for each marginal constraint $i = 1,\dots,d$, the stationarity condition reads: for λ_d-almost every $u \in \mathbb{I}^d$,

$$\Phi(u) \;+\; \varepsilon\big(1 + \log c_\varepsilon(u)\big) \;+\; \mu \;+\; \sum_{i=1}^{d} \lambda_i(u_i) \;=\; 0. \tag{3.2}$$

Solving (3.2) for $c_\varepsilon(u)$ gives

$$c_\varepsilon(u) \;=\; \exp\!\left(-\frac{\Phi(u)}{\varepsilon} - 1 - \frac{\mu}{\varepsilon} - \sum_{i=1}^{d} \frac{\lambda_i(u_i)}{\varepsilon}\right) \;=\; \exp\!\left(-\frac{\Phi(u)}{\varepsilon}\right) \prod_{i=1}^{d} a_i^{(\varepsilon)}(u_i),$$

where we set

$$a_i^{(\varepsilon)}(u_i) = \exp\!\left(-\frac{\lambda_i(u_i)}{\varepsilon}\right), \qquad i = 1,\dots,d,$$

and absorbed the constant $e^{-1-\mu/\varepsilon}$ into one of the $a_i^{(\varepsilon)}$ factors (say $a_1^{(\varepsilon)}$), which introduces the standard scaling redundancy.

(c) Determination of the potentials via Sinkhorn iteration.

It remains to show that the functions $a_i^{(\varepsilon)} : [0,1] \to (0,\infty)$ are uniquely determined (up to the scaling redundancy) by the d marginal constraints. Define the kernel $K : \mathbb{I}^d \to (0,\infty)$ by

$$K(u) = \exp\!\left(-\frac{\Phi(u)}{\varepsilon}\right),$$

which is strictly positive λ_d-a.e. by assumption. The marginal constraint for coordinate i requires

$$\int_{\mathbb{I}^{d-1}} K(u) \prod_{j=1}^{d} a_j^{(\varepsilon)}(u_j)\, du_{-i} \;=\; 1 \quad \text{for } \lambda_1\text{-a.e. } u_i \in [0,1]. \tag{3.3}$$

The system (3.3) is precisely the fixed-point equation of the *Sinkhorn–Bregman iteration*: starting from $a_i^{(0)} \equiv 1$, the alternating normalization scheme

$$a_i^{(k+1)}(u_i) = \left(\int_{\mathbb{I}^{d-1}} K(u) \prod_{j \neq i} a_j^{(k)}(u_j)\, du_{-i} \right)^{-1}, \quad i = 1, \dots, d,$$

converges in $L^2([0,1])$ to the unique (up to scaling) solution $(a_1^{(\varepsilon)}, \dots, a_d^{(\varepsilon)})$ of (3.3). Convergence follows from the contractivity of the Sinkhorn operator in the Hilbert metric on the cone of positive functions, a classical consequence of the Perron–Frobenius theorem in the integral operator setting (see, e.g., [7]). Uniqueness of the potentials up to scaling is a direct consequence of the strict positivity of K and the L^2 density of the marginal projection operators. □

Proposition 3.3 (Limiting regimes)**.** *Let $\{C_\varepsilon\}_{\varepsilon>0}$ be the family defined by (3.1). Then:*

(1) *as $\varepsilon \to \infty$, $C_\varepsilon \to \Pi$ weakly;*
(2) *as $\varepsilon \to 0^+$, any weak limit point of C_ε is supported on $\arg\min_{\mathbb{I}^d} \Phi$ intersected with $\mathcal{C}_d$.*

Proof.
Item (i) : $C_\varepsilon \to \Pi$ weakly as $\varepsilon \to \infty$.

Let c_ε denote the density of C_ε with respect to λ_d, whose existence is guaranteed by Theorem 3.2. Dividing the objective $\mathcal{F}_\varepsilon$ by $\varepsilon > 0$ gives the equivalent problem

$$c_\varepsilon \;=\; \arg\min_{c \in \mathcal{A}} \left\{ \frac{1}{\varepsilon} \int_{\mathbb{I}^d} \Phi(u)\, c(u)\, du \;+\; \int_{\mathbb{I}^d} c(u) \log c(u)\, du \right\}.$$

As $\varepsilon \to \infty$, the weight $1/\varepsilon$ of the linear term $\int \Phi\, c\, du$ tends to zero, so the objective converges pointwise to the pure entropy functional

$$\mathcal{H}(c) = \int_{\mathbb{I}^d} c(u) \log c(u)\, du.$$

The unique minimizer of $\mathcal{H}$ over $\mathcal{A}$ is the uniform density $c \equiv 1$, corresponding to Π. To make this rigorous, let $c^\star \equiv 1$ and use it as a test function: for every $\varepsilon > 0$,

$$\frac{1}{\varepsilon} \int_{\mathbb{I}^d} \Phi\, c_\varepsilon\, d\lambda_d \;+\; \int_{\mathbb{I}^d} c_\varepsilon \log c_\varepsilon\, d\lambda_d \;\le\; \frac{1}{\varepsilon} \int_{\mathbb{I}^d} \Phi\, d\lambda_d \;+\; \underbrace{\int_{\mathbb{I}^d} 1 \cdot \log 1\, d\lambda_d}_{=0},$$

which gives

$$0 \;\le\; \int_{\mathbb{I}^d} c_\varepsilon \log c_\varepsilon\, d\lambda_d \;\le\; \frac{\|\Phi\|_{L^1}}{\varepsilon} \;+\; \frac{1}{\varepsilon} \int_{\mathbb{I}^d} \Phi\, d\lambda_d \;\le\; \frac{2\|\Phi\|_{L^1}}{\varepsilon}. \tag{3.4}$$

Since $\mathrm{KL}(C_\varepsilon\|\Pi) = \int_{\mathbb{I}^d} c_\varepsilon \log c_\varepsilon \, d\lambda_d$, estimate (3.4) yields

$$\mathrm{KL}(C_\varepsilon \,\|\, \Pi) \;\le\; \frac{2\|\Phi\|_{L^1}}{\varepsilon} \;\longrightarrow\; 0 \quad \text{as } \varepsilon \to \infty.$$

Convergence in KL implies weak convergence of probability measures (by the Pinsker–Csiszár inequality, which gives $\|C_\varepsilon - \Pi\|_{\mathrm{TV}} \to 0$, and total variation convergence implies weak convergence), so $C_\varepsilon \to \Pi$ weakly as $\varepsilon \to \infty$.

Item (ii) : Concentration on $\arg\min \Phi \cap \mathcal{C}_d$ as $\varepsilon \to 0^+$.

We reformulate the problem as a Γ-convergence statement. For each $\varepsilon > 0$, define $\mathcal{G}_\varepsilon : \mathcal{C}_d \to [0, +\infty]$ by

$$\mathcal{G}_\varepsilon(C) = \begin{cases} \displaystyle\int_{\mathbb{I}^d} \Phi \, dC \;+\; \varepsilon \,\mathrm{KL}(C \,\|\, \Pi) & \text{if } C \in \mathcal{C}_d^{\mathrm{ac}}, \\ +\infty & \text{otherwise,} \end{cases}$$

and define the limit functional $\mathcal{G}_0 : \mathcal{C}_d \to [0, +\infty]$ by

$$\mathcal{G}_0(C) = \int_{\mathbb{I}^d} \Phi \, dC.$$

We verify the two conditions of Γ-convergence with respect to the weak topology on $\mathcal{C}_d$.

Γ-*liminf inequality.* Let $\{C_\varepsilon\}_{\varepsilon>0} \subset \mathcal{C}_d^{\mathrm{ac}}$ be any sequence converging weakly to some $C_0 \in \mathcal{C}_d$ as $\varepsilon \to 0^+$. Since $\mathrm{KL}(\,\cdot\,\|\,\Pi) \ge 0$,

$$\mathcal{G}_\varepsilon(C_\varepsilon) \;=\; \int_{\mathbb{I}^d} \Phi \, dC_\varepsilon \;+\; \varepsilon \,\mathrm{KL}(C_\varepsilon \,\|\, \Pi) \;\ge\; \int_{\mathbb{I}^d} \Phi \, dC_\varepsilon.$$

By the weak convergence $C_\varepsilon \rightharpoonup C_0$ and the continuity of $C \mapsto \int \Phi \, dC$ (since $\Phi \in L^1(\Pi) \subset \mathcal{C}(\mathbb{I}^d)$ up to approximation by bounded continuous functions and a standard density argument),

$$\liminf_{\varepsilon \to 0^+} \mathcal{G}_\varepsilon(C_\varepsilon) \;\ge\; \int_{\mathbb{I}^d} \Phi \, dC_0 \;=\; \mathcal{G}_0(C_0).$$

Γ-*limsup inequality (recovery sequence).* For any $C_0 \in \mathcal{C}_d$, we must exhibit a sequence $C_\varepsilon \to C_0$ weakly such that $\limsup_{\varepsilon\to 0^+} \mathcal{G}_\varepsilon(C_\varepsilon) \le \mathcal{G}_0(C_0)$. If $C_0 \notin \mathcal{C}_d^{\mathrm{ac}}$, approximate C_0 by $C_\varepsilon = (1-\delta_\varepsilon)C_0 + \delta_\varepsilon \Pi$ with $\delta_\varepsilon \to 0$; this gives $C_\varepsilon \in \mathcal{C}_d^{\mathrm{ac}}$ for every $\varepsilon > 0$ and $C_\varepsilon \rightharpoonup C_0$. A direct computation yields

$$\varepsilon \,\mathrm{KL}(C_\varepsilon \,\|\, \Pi) \;\le\; \varepsilon \cdot \frac{\delta_\varepsilon^2}{2(1-\delta_\varepsilon)} \cdot \|\log c_0\|_{L^1} \;\to\; 0$$

provided δ_ε is chosen so that $\varepsilon\delta_\varepsilon^2 \to 0$, e.g. $\delta_\varepsilon = \sqrt{\varepsilon}$. If $C_0 \in \mathcal{C}_d^{\mathrm{ac}}$, take $C_\varepsilon \equiv C_0$; then $\mathcal{G}_\varepsilon(C_0) = \int \Phi \, dC_0 + \varepsilon \,\mathrm{KL}(C_0\|\Pi) \to \mathcal{G}_0(C_0)$.

Compactness and conclusion. The space $\mathcal{C}_d$ endowed with the weak topology is compact (it is a closed, tight subset of the space of probability measures on the compact set $\mathbb{I}^d$, by Prokhorov's theorem). Consequently, every sequence $\{C_\varepsilon\}_{\varepsilon>0}$

admits at least one weak accumulation point. The fundamental theorem of Γ-convergence then asserts that any such accumulation point minimizes $\mathcal{G}_0$ over $\mathcal{C}_d$, that is, it belongs to

$$\operatorname*{arg\,min}_{C\in\mathcal{C}_d} \int_{\mathbb{I}^d} \Phi \, dC \;=\; \Big\{ C \in \mathcal{C}_d \;:\; C\big(\operatorname*{arg\,min}_{\mathbb{I}^d} \Phi\big) = 1 \Big\},$$

which is precisely the set of copulas supported on $\arg\min_{\mathbb{I}^d} \Phi \cap \mathcal{C}_d$. □

3.2. **Geometric and PDE characterization.**

Theorem 3.4 (Constrained Fokker-Planck equation)**.** *The map $\varepsilon \mapsto c_\varepsilon$ is differentiable on $(0,\infty)$ and satisfies, in the distributional sense,*

$$\partial_\varepsilon c_\varepsilon \;=\; \Delta c_\varepsilon \;-\; \operatorname{div}(c_\varepsilon \nabla \log \Phi) \;+\; \lambda(\varepsilon, u), \tag{3.5}$$

where $\lambda(\varepsilon,\cdot)$ is the unique additive correction enforcing the uniform-marginal constraints at each $\varepsilon > 0$.

Proof.
(a) Euler-Lagrange system and its differentiability.

By Theorem 3.2, for each $\varepsilon > 0$ the unique minimizer $c_\varepsilon \in \mathcal{A}$ satisfies the Euler–Lagrange conditions: for λ_d-almost every $u \in \mathbb{I}^d$,

$$\Phi(u) + \varepsilon\big(1 + \log c_\varepsilon(u)\big) + \mu(\varepsilon) + \sum_{i=1}^{d} \lambda_i(\varepsilon, u_i) = 0, \tag{3.6}$$

where $\mu(\varepsilon) \in \mathbb{R}$ is the multiplier for the total-mass constraint and $\lambda_i(\varepsilon,\cdot) \in L^2([0,1])$ are the multipliers for the marginal constraints, $i = 1,\dots,d$. Solving (3.6) explicitly gives

$$c_\varepsilon(u) = \exp\left(-\frac{\Phi(u)}{\varepsilon} - 1 - \frac{\mu(\varepsilon)}{\varepsilon} - \frac{1}{\varepsilon}\sum_{i=1}^{d} \lambda_i(\varepsilon, u_i) \right). \tag{3.7}$$

We verify that $\varepsilon \mapsto c_\varepsilon$ is differentiable on $(0,\infty)$ in $L^2(\mathbb{I}^d)$. The representation (3.7) shows that c_ε depends on ε through the smooth functions $\varepsilon \mapsto \mu(\varepsilon)$ and $\varepsilon \mapsto \lambda_i(\varepsilon,\cdot)$, whose differentiability will be established in Step 2 via the implicit function theorem applied to the marginal constraints. Since the exponential map is C^∞ and the constraints are C^1 in ε (shown below), it follows that $\varepsilon \mapsto c_\varepsilon \in L^2(\mathbb{I}^d)$ is differentiable on $(0,\infty)$.

(b) Differentiation of the Euler-Lagrange system with respect to ε.

Differentiate (3.6) with respect to ε, denoting $\dot{c}_\varepsilon = \partial_\varepsilon c_\varepsilon$, $\dot{\mu}(\varepsilon) := \partial_\varepsilon \mu(\varepsilon)$, and $\dot{\lambda}_i(\varepsilon,\cdot) = \partial_\varepsilon \lambda_i(\varepsilon,\cdot)$:

$$\frac{\partial}{\partial \varepsilon}\Big[\Phi(u) + \varepsilon(1 + \log c_\varepsilon(u)) + \mu(\varepsilon) + \textstyle\sum_i \lambda_i(\varepsilon, u_i)\Big] = 0. \tag{3.8}$$

Using $\partial_\varepsilon[\varepsilon \log c_\varepsilon] = \log c_\varepsilon + \varepsilon \cdot \frac{\dot{c}_\varepsilon}{c_\varepsilon}$, equation (3.8) becomes

$$1 + \log c_\varepsilon(u) + \varepsilon \frac{\dot{c}_\varepsilon(u)}{c_\varepsilon(u)} + \dot{\mu}(\varepsilon) + \sum_{i=1}^{d} \dot{\lambda}_i(\varepsilon, u_i) = 0, \tag{3.9}$$

hence

$$\dot{c}_\varepsilon(u) = -\frac{c_\varepsilon(u)}{\varepsilon}\Big(1 + \log c_\varepsilon(u) + \dot{\mu}(\varepsilon) + \textstyle\sum_{i=1}^d \dot{\lambda}_i(\varepsilon, u_i)\Big). \tag{3.10}$$

(c) Rewriting as a Fokker-Planck-type equation.

We now rewrite (3.10) in terms of differential operators acting on c_ε. From the explicit form (3.7), compute the gradient:

$$\nabla \log c_\varepsilon(u) = -\frac{\nabla \Phi(u)}{\varepsilon} - \frac{1}{\varepsilon}\sum_{i=1}^d \nabla_{u_i}\lambda_i(\varepsilon, u_i),$$

where $\nabla_{u_i}\lambda_i$ denotes the partial derivative with respect to u_i only. Taking the Laplacian of c_ε via the identity $\Delta c_\varepsilon = c_\varepsilon \Delta \log c_\varepsilon + c_\varepsilon |\nabla \log c_\varepsilon|^2$, and using the divergence identity

$$\operatorname{div}(c_\varepsilon \nabla \log c_\varepsilon) = \Delta c_\varepsilon - \frac{|\nabla c_\varepsilon|^2}{c_\varepsilon} = c_\varepsilon \Delta \log c_\varepsilon,$$

one rewrites the right-hand side of (3.10) as follows. The term $-\frac{c_\varepsilon}{\varepsilon}(1 + \log c_\varepsilon)$ accounts for the entropy variation and can be identified, after integration by parts, with

$$-\frac{1}{\varepsilon}\operatorname{div}(c_\varepsilon \nabla \log c_\varepsilon) + \frac{1}{\varepsilon}\Delta c_\varepsilon = \frac{1}{\varepsilon}\Big[\Delta c_\varepsilon - \operatorname{div}(c_\varepsilon \nabla \log c_\varepsilon)\Big].$$

Substituting $\nabla \log c_\varepsilon = -\nabla\Phi/\varepsilon + (\text{marginal terms})$ yields

$$\dot{c}_\varepsilon = \Delta c_\varepsilon - \operatorname{div}(c_\varepsilon \nabla \log \Phi) + \lambda(\varepsilon, u),$$

where the correction term

$$\lambda(\varepsilon, u) = -\frac{c_\varepsilon(u)}{\varepsilon}\Big(\dot{\mu}(\varepsilon) + \sum_{i=1}^d \dot{\lambda}_i(\varepsilon, u_i)\Big) \tag{3.11}$$

collects all contributions from the differentiated marginal multipliers.

(d) Uniqueness of the correction term $\lambda(\varepsilon, \cdot)$.

It remains to show that $\lambda(\varepsilon, \cdot)$ is uniquely determined by the requirement that c_ε satisfies the uniform-marginal constraints at every $\varepsilon > 0$, i.e.,

$$\int_{\mathbb{I}^{d-1}} c_\varepsilon(u)\, du_{-i} = 1 \quad \lambda_1\text{-a.e.}, \quad i = 1, \dots, d. \tag{3.12}$$

Differentiating (3.12) with respect to ε gives

$$\int_{\mathbb{I}^{d-1}} \dot{c}_\varepsilon(u)\, du_{-i} = 0 \quad \lambda_1\text{-a.e.}, \quad i = 1, \dots, d, \tag{3.13}$$

meaning that $\dot{c}_\varepsilon$ lies in the tangent space $T_{c_\varepsilon}\mathcal{A}$ defined in Lemma 2.9. Substituting (3.10) into (3.13), the constraints (3.13) become a linear system for the unknowns $(\dot{\mu}(\varepsilon), \dot{\lambda}_1(\varepsilon, \cdot), \dots, \dot{\lambda}_d(\varepsilon, \cdot))$: for each $i = 1, \dots, d$ and λ_1-a.e. $u_i \in [0, 1]$,

$$\int_{\mathbb{I}^{d-1}} \frac{c_\varepsilon(u)}{\varepsilon}\Big(\dot{\mu}(\varepsilon) + \sum_{j=1}^d \dot{\lambda}_j(\varepsilon, u_j)\Big) du_{-i} = \int_{\mathbb{I}^{d-1}} \frac{c_\varepsilon(u)}{\varepsilon}\big(1 + \log c_\varepsilon(u)\big)\, du_{-i}. \tag{3.14}$$

By Lemma 2.9, the tangent space $T_{c_\varepsilon}\mathcal{A}$ has a well-defined orthogonal decomposition in $L^2(c_\varepsilon\, d\lambda_d)$, and the operator mapping $(\dot\mu, \dot\lambda_1, \dots, \dot\lambda_d)$ to the left-hand side of (3.14) is a bounded, coercive linear operator on this space. Indeed, coercivity follows from the strict positivity of c_ε (guaranteed by Theorem 3.2) and the non-degeneracy of the marginal projection operators on $L^2(c_\varepsilon\, d\lambda_d)$. The Lax–Milgram theorem therefore ensures existence and uniqueness of the solution $(\dot\mu(\varepsilon), \dot\lambda_1(\varepsilon,\cdot), \dots, \dot\lambda_d(\varepsilon,\cdot))$, which in turn uniquely determines $\lambda(\varepsilon,\cdot)$ via (3.11).

Combining Steps 1–4, the map $\varepsilon \mapsto c_\varepsilon$ is differentiable on $(0,\infty)$ in $L^2(\mathbb{I}^d)$, and its derivative $\dot c_\varepsilon = \partial_\varepsilon c_\varepsilon$ satisfies, in the distributional sense,

$$\partial_\varepsilon\, c_\varepsilon = \Delta c_\varepsilon - \operatorname{div}(c_\varepsilon\, \nabla \log \Phi) + \lambda(\varepsilon, u),$$

where $\lambda(\varepsilon,\cdot)$ is the unique additive correction (in the sense of (3.11)) enforcing the uniform-marginal constraints (3.12) at each $\varepsilon > 0$. □

Theorem 3.5 (Fisher-Rao geodesic property)**.** *The curve $\varepsilon \mapsto C_\varepsilon$ is, up to reparametrization, a geodesic on $(\mathcal{C}_d^{\mathrm{ac}}, g_{\mathrm{FR}})$ joining the comonotone regime ($\varepsilon \to 0^+$) to the independence copula ($\varepsilon \to \infty$).*

Proof.
(a) The curve $\varepsilon \mapsto \log c_\varepsilon$ is affine in the exponential family generated by Φ.

By Theorem 3.2, the optimal density takes the form

$$c_\varepsilon(u) = \exp\left(-\frac{\Phi(u)}{\varepsilon} + \sum_{i=1}^d \alpha_i^{(\varepsilon)}(u_i)\right), \tag{3.15}$$

where $\alpha_i^{(\varepsilon)} = \log a_i^{(\varepsilon)}$ and the potentials $a_i^{(\varepsilon)}$ are determined by the marginal constraints. Define the natural parameter $\theta(\varepsilon) = -1/\varepsilon \in (-\infty, 0)$ and the sufficient statistic $T(u) = \Phi(u)$. The family $\{c_\varepsilon\}_{\varepsilon>0}$ is therefore a subfamily of the exponential family

$$\mathcal{E} := \left\{ p_{\theta,\alpha} \propto \exp\left(\theta\, T(u) + \sum_{i=1}^d \alpha_i(u_i)\right) : \theta \in \mathbb{R},\ \alpha_i \in L^2([0,1]) \right\} \cap \mathcal{C}_d^{\mathrm{ac}}, \tag{3.16}$$

where the intersection with $\mathcal{C}_d^{\mathrm{ac}}$ enforces the uniform-marginal constraints and uniquely pins down $\alpha_i = \alpha_i^{(\varepsilon)}$ as a function of θ. The log-density reads

$$\log c_\varepsilon(u) = \theta(\varepsilon)\, T(u) + \sum_{i=1}^d \alpha_i^{(\varepsilon)}(u_i) = -\frac{\Phi(u)}{\varepsilon} + \sum_{i=1}^d \alpha_i^{(\varepsilon)}(u_i).$$

Since $\theta(\varepsilon) = -1/\varepsilon$ is a smooth, strictly monotone function of ε, the map $\varepsilon \mapsto \log c_\varepsilon$ traces a smooth curve in the log-density space $L^2(\mathbb{I}^d)$ whose natural-parameter component $\theta(\varepsilon)$ varies affinely in θ, i.e.,

$$\frac{d}{d\theta} \log c_{-1/\theta}(u) = T(u) + \frac{d}{d\theta} \sum_{i=1}^d \alpha_i^{(-1/\theta)}(u_i), \tag{3.17}$$

where the second term is determined solely by the marginal constraints and lies in the tangent space $T_{c_\varepsilon}\mathcal{A}$ (cf. Lemma 2.9). This confirms that $\varepsilon \mapsto c_\varepsilon$ is an *e-affine curve* in the sense of information geometry [1].

(b) Dually flat structure of the Fisher–Rao geometry on $\mathcal{C}_d^{\mathrm{ac}}$.

Recall that the Fisher–Rao metric on $\mathcal{C}_d^{\mathrm{ac}}$ is the Riemannian metric

$$g_{\mathrm{FR}}(f,g)\big|_c = \int_{\mathbb{I}^d} \frac{f(u)\,g(u)}{c(u)}\,du, \quad f,g \in T_c\mathcal{C}_d^{\mathrm{ac}}, \tag{3.18}$$

where $T_c\mathcal{C}_d^{\mathrm{ac}}$ is the tangent space of zero-mean densities at c. The exponential family $\mathcal{E}$ defined in (3.16) carries a dually flat structure: there exist dual affine coordinate systems

$$\begin{aligned} \theta\text{-coordinates } (e\text{-flat}): &\quad \eta_\theta = (\theta, \alpha_1, \dots, \alpha_d), \\ \mu\text{-coordinates } (m\text{-flat}): &\quad \eta_\mu = \big(\mathbb{E}_{c_\varepsilon}[T],\ \mathbb{E}_{c_\varepsilon}[\mathbf{1}_{u_i \le t}]\big)_{i,t}, \end{aligned}$$

connected by the Legendre duality $\psi(\theta)+\varphi(\mu) = \langle \theta, \mu \rangle$, where $\psi(\theta) = \log \int \exp(\theta T + \sum_i \alpha_i)\, d\lambda_d$ is the log-partition function. The e-connection $\nabla^{(e)}$ and m-connection $\nabla^{(m)}$ are dual with respect to g_{FR} and both have zero curvature on $\mathcal{E}$, giving the dually flat structure of [1].

(c) e-affine curves are geodesics for $\nabla^{(e)}$.

A curve $\gamma : I \to \mathcal{E}$ is a *geodesic for the e-connection* $\nabla^{(e)}$ if and only if it is affine in the θ-coordinates, i.e., $\ddot{\theta}(\gamma(t)) = 0$ along the curve. The curve $t \mapsto c_{-1/t}$ (with parameter $t = \theta \in (-\infty, 0)$) satisfies

$$\frac{d^2}{dt^2}\,\theta\big(c_{-1/t}\big) = \frac{d^2}{dt^2}\,t = 0,$$

so it is a geodesic for $\nabla^{(e)}$ on $\mathcal{E}$. By the fundamental theorem of dually flat geometry, on any dually flat manifold, e-geodesics and m-geodesics are both geodesics for the Levi–Civita connection $\nabla^{(0)}$ of g_{FR} *up to reparametrization*. More precisely, if γ is an e-geodesic parametrized by θ, then the reparametrization $s = s(\theta)$ satisfying

$$\frac{ds}{d\theta} = \left\| \frac{d\gamma}{d\theta} \right\|_{g_{\mathrm{FR}}}^{-1} \quad \Longleftrightarrow \quad s(\theta) = \int_{\theta_0}^{\theta} \sqrt{g_{\mathrm{FR}}(\dot{\gamma}(r), \dot{\gamma}(r))}\, dr \tag{3.19}$$

gives $\tilde{\gamma}(s) = \gamma(\theta(s))$ as a geodesic for $\nabla^{(0)}$ (i.e., unit-speed, with $\nabla^{(0)}_{\dot{\tilde{\gamma}}}\dot{\tilde{\gamma}} = 0$).

(d)Verification of the Fisher-Rao speed and reparametrization.

We compute the Fisher–Rao speed of the curve $\theta \mapsto c_\theta$ (where $\theta = -1/\varepsilon$). The tangent vector at c_θ in the e-parametrization is

$$\dot{c}_\theta = \frac{dc_\theta}{d\theta} = c_\theta \cdot \frac{d \log c_\theta}{d\theta} = c_\theta \left(T - \mathbb{E}_{c_\theta}[T] + \sum_{i=1}^{d} \dot{\alpha}_i^{(\theta)}(u_i) \right),$$

where we used the fact that $\frac{d}{d\theta} \log c_\theta = T - \mathbb{E}_{c_\theta}[T] + \sum_i \dot{\alpha}_i^{(\theta)}$ after normalization. The Fisher–Rao norm of $\dot{c}_\theta$ is

$$\|\dot{c}_\theta\|_{g_{\mathrm{FR}}}^2 = \int_{\mathbb{I}^d} \frac{(\dot{c}_\theta)^2}{c_\theta}\, d\lambda_d = \int_{\mathbb{I}^d} c_\theta \left(T - \mathbb{E}_{c_\theta}[T] + \sum_{i=1}^{d} \dot{\alpha}_i^{(\theta)}(u_i) \right)^2 d\lambda_d. \tag{3.20}$$

By Lemma 2.9, the vector $\sum_i \dot{\alpha}_i^{(\theta)}(u_i)$ lies in the subspace of functions depending only on individual marginals, which is g_{FR}-orthogonal to the centered sufficient statistic $T - \mathbb{E}_{c_\theta}[T]$ modulo the marginal constraints. After projecting out the marginal component (which contributes zero net Fisher information under the uniform-marginal constraints), the speed reduces to

$$\|\dot{c}_\theta\|_{g_{\mathrm{FR}}}^2 = \mathrm{Var}_{c_\theta}(\Phi) + R(\theta),$$

where $R(\theta)$ is a remainder arising from $\dot{\alpha}_i^{(\theta)}$, which is bounded and continuous in θ on $(-\infty, 0)$. The reparametrization $s(\theta) = \int_{\theta_0}^{\theta} \|\dot{c}_r\|_{g_{\mathrm{FR}}}\, dr$ is therefore a well-defined, strictly increasing C^1 function of $\theta \in (-\infty, 0)$, and the reparametrized curve $\tilde{\gamma}(s) = c_{\theta(s)}$ satisfies $\nabla^{(0)}_{\dot{\tilde{\gamma}}} \dot{\tilde{\gamma}} = 0$ by (3.19).

(e) Boundary regimes.

By Proposition 3.3:

- As $\varepsilon \to \infty$ (i.e., $\theta \to 0^-$), $c_\varepsilon \to 1 = c_\Pi$, the density of the independence copula Π.
- As $\varepsilon \to 0^+$ (i.e., $\theta \to -\infty$), any weak accumulation point of c_ε is supported on $\arg\min_{\mathbb{I}^d} \Phi \cap \mathcal{C}_d$, which corresponds to the comonotone regime.

The geodesic $\tilde{\gamma}$ therefore joins the comonotone regime (in the limit $s \to -\infty$) to the independence copula (at the finite endpoint $s = s(\theta = 0^-)$), both in the metric closure of $(\mathcal{C}_d^{\mathrm{ac}}, g_{\mathrm{FR}})$.

The curve $\varepsilon \mapsto C_\varepsilon$ is an e-geodesic on the dually flat manifold $(\mathcal{E}, \nabla^{(e)}, \nabla^{(m)}, g_{\mathrm{FR}})$, parametrized affinely in $\theta = -1/\varepsilon$. The reparametrization $s(\varepsilon)$ defined by (3.19) transforms it into a unit-speed geodesic for the Levi–Civita connection $\nabla^{(0)}$ of g_{FR} on $\mathcal{C}_d^{\mathrm{ac}}$, joining the comonotone regime to the independence copula Π. □

3.3. **Stability and statistical consequences.**

Theorem 3.6 (Quantitative stability)**.** *Let $\Phi, \Phi' \in L^\infty(\mathbb{I}^d)$ be two potentials and let $C_\varepsilon, C'_\varepsilon$ denote the associated Sinkhorn–Wasserstein copulas. Then there exists a constant $K_\varepsilon > 0$ depending only on ε and $\|\Phi\|_\infty, \|\Phi'\|_\infty$ such that*

$$\mathrm{KL}(C_\varepsilon \,\|\, C'_\varepsilon) \;\le\; K_\varepsilon \,\|\Phi - \Phi'\|_\infty^2.$$

Proof.

(a) Explicit perturbation of the optimal density.

Let $\Phi, \Phi' \in L^\infty(\mathbb{I}^d)$ and denote $\delta\Phi = \Phi - \Phi' \in L^\infty(\mathbb{I}^d)$, so that $\|\delta\Phi\|_\infty \le M$ for some $M > 0$. By Theorem 3.2, the two optimal densities admit the representations

$$c_\varepsilon(u) = \exp\left(-\frac{\Phi(u)}{\varepsilon} + \sum_{i=1}^d \alpha_i^{(\varepsilon)}(u_i)\right), \tag{3.21}$$

$$c'_\varepsilon(u) = \exp\left(-\frac{\Phi'(u)}{\varepsilon} + \sum_{i=1}^d \beta_i^{(\varepsilon)}(u_i)\right), \tag{3.22}$$

where the potentials $\alpha_i^{(\varepsilon)}, \beta_i^{(\varepsilon)} : [0,1] \to \mathbb{R}$ are uniquely determined (up to the scaling redundancy) by the uniform-marginal constraints

$$\int_{\mathbb{I}^{d-1}} c_\varepsilon(u)\, du_{-i} = 1, \quad \int_{\mathbb{I}^{d-1}} c'_\varepsilon(u)\, du_{-i} = 1, \quad i = 1, \dots, d. \tag{3.23}$$

Define the interpolating family $\Phi^{(t)} = \Phi' + t\,\delta\Phi$ for $t \in [0,1]$, and let $c_\varepsilon^{(t)}$ be the corresponding optimal density with potentials $\alpha_i^{(t)}$. The map $t \mapsto c_\varepsilon^{(t)}$ is C^2 in t on $[0,1]$: this follows from the implicit function theorem applied to the marginal-constraint system, exactly as in the proof of Theorem 3.4, with t in place of ε.

(b) Second-order Taylor expansion of the KL divergence.

Define the functional

$$F(t) = \mathrm{KL}\Big(C_\varepsilon^{(t)} \,\Big\|\, C'_\varepsilon\Big) = \int_{\mathbb{I}^d} c_\varepsilon^{(t)}(u) \log \frac{c_\varepsilon^{(t)}(u)}{c'_\varepsilon(u)}\, du.$$

We expand $F(t)$ around $t = 0$. At $t = 0$, $c_\varepsilon^{(0)} = c'_\varepsilon$, so $F(0) = 0$. For the first derivative, differentiating under the integral sign gives

$$F'(t) = \int_{\mathbb{I}^d} \dot{c}_\varepsilon^{(t)}(u) \log \frac{c_\varepsilon^{(t)}(u)}{c'_\varepsilon(u)}\, du + \int_{\mathbb{I}^d} \dot{c}_\varepsilon^{(t)}(u)\, du. \tag{3.24}$$

Since $\int_{\mathbb{I}^d} c_\varepsilon^{(t)}\, du = 1$ for all t, differentiating gives $\int_{\mathbb{I}^d} \dot{c}_\varepsilon^{(t)}\, du = 0$, so the second integral in (3.24) vanishes. At $t = 0$, $\log(c_\varepsilon^{(0)}/c'_\varepsilon) = 0$, hence $F'(0) = 0$.

For the second derivative,

$$F''(t) = \int_{\mathbb{I}^d} \ddot{c}_\varepsilon^{(t)}(u) \log \frac{c_\varepsilon^{(t)}(u)}{c'_\varepsilon(u)}\, du + \int_{\mathbb{I}^d} \frac{\big(\dot{c}_\varepsilon^{(t)}(u)\big)^2}{c_\varepsilon^{(t)}(u)}\, du. \tag{3.25}$$

At $t = 0$, the first integral vanishes again since $\log(c_\varepsilon^{(0)}/c'_\varepsilon) = 0$, leaving

$$F''(0) = \int_{\mathbb{I}^d} \frac{\big(\dot{c}_\varepsilon^{(0)}(u)\big)^2}{c'_\varepsilon(u)}\, du = \big\|\dot{c}_\varepsilon^{(0)}\big\|^2_{g_{\mathrm{FR}},\, c'_\varepsilon} \;\geq\; 0. \tag{3.26}$$

By Taylor's theorem with integral remainder,

$$F(1) = \frac{1}{2}F''(0) + \frac{1}{2}\int_0^1 (1-t)\big[F''(t) - F''(0)\big]\, dt. \tag{3.27}$$

(c) Bound on the first-order perturbation $\dot{c}_\varepsilon^{(0)}$.

Differentiating the log-density $\log c_\varepsilon^{(t)} = -\Phi^{(t)}/\varepsilon + \sum_i \alpha_i^{(t)}(u_i)$ with respect to t at $t = 0$ gives

$$\frac{\dot{c}_\varepsilon^{(0)}(u)}{c'_\varepsilon(u)} = \frac{d}{dt}\Big|_{t=0} \log c_\varepsilon^{(t)}(u) = -\frac{\delta\Phi(u)}{\varepsilon} + \sum_{i=1}^d \dot{\alpha}_i^{(0)}(u_i), \tag{3.28}$$

where $\dot\alpha_i^{(0)}$ is determined by differentiating the marginal constraints (3.23) at $t=0$:

$$\int_{\mathbb{I}^{d-1}} c_\varepsilon'(u)\left(-\frac{\delta\Phi(u)}{\varepsilon}+\sum_{j=1}^d \dot\alpha_j^{(0)}(u_j)\right) du_{-i}=0,\quad i=1,\dots,d. \tag{3.29}$$

This is a linear system in $(\dot\alpha_1^{(0)},\dots,\dot\alpha_d^{(0)})$ identical in structure to (3.14), which by Lemma 2.9 has a unique solution. Since the solution operator of (3.29) is a bounded projection in $L^2(c_\varepsilon'\,d\lambda_d)$ with operator norm at most 1, the function $\sum_i \dot\alpha_i^{(0)}(u_i)$ is the $L^2(c_\varepsilon'\,d\lambda_d)$-projection of $\delta\Phi/\varepsilon$ onto the subspace of marginal functions, hence

$$\left\|\sum_{i=1}^d \dot\alpha_i^{(0)}\right\|_{L^2(c_\varepsilon'\,d\lambda_d)} \le \frac{\|\delta\Phi\|_\infty}{\varepsilon}. \tag{3.30}$$

Substituting (3.28) and (3.30) into (3.26),

$$\begin{aligned} F''(0) &= \int_{\mathbb{I}^d} c_\varepsilon'(u)\left(-\frac{\delta\Phi(u)}{\varepsilon}+\sum_{i=1}^d \dot\alpha_i^{(0)}(u_i)\right)^2 d\lambda_d \\ &\le 2\int_{\mathbb{I}^d} c_\varepsilon'(u)\frac{(\delta\Phi(u))^2}{\varepsilon^2}\,d\lambda_d + 2\left\|\sum_{i=1}^d \dot\alpha_i^{(0)}\right\|^2_{L^2(c_\varepsilon'\,d\lambda_d)} \\ &\le \frac{4\|\delta\Phi\|_\infty^2}{\varepsilon^2}, \end{aligned} \tag{3.31}$$

where we used $(a+b)^2\le 2a^2+2b^2$ and (3.30).

(d) Uniform bound on $F''(t)$ for $t\in[0,1]$.

For general $t\in[0,1]$, the same argument applied to $c_\varepsilon^{(t)}$ gives

$$F''(t)=\int_{\mathbb{I}^d}\frac{(\dot c_\varepsilon^{(t)})^2}{c_\varepsilon^{(t)}}\,d\lambda_d + R_1(t) \le \frac{4\|\delta\Phi\|_\infty^2}{\varepsilon^2}+R_1(t), \tag{3.32}$$

where $R_1(t)$ is the contribution of $\ddot c_\varepsilon^{(t)}$ to (3.25) at general t. Since $\|\Phi^{(t)}\|_\infty \le \|\Phi'\|_\infty + \|\delta\Phi\|_\infty$ uniformly in t, and $c_\varepsilon^{(t)}$ is bounded away from zero by

$$c_\varepsilon^{(t)}(u)\ge \exp\!\left(-\frac{\|\Phi^{(t)}\|_\infty}{\varepsilon}\right)\cdot\prod_{i=1}^d \inf_{u_i} a_i^{(t)}(u_i) > 0, \tag{3.33}$$

the second-order term $R_1(t)$ satisfies $|R_1(t)|\le C_\varepsilon\|\delta\Phi\|_\infty^2$ for a constant C_ε depending only on $\varepsilon,\|\Phi'\|_\infty,\|\delta\Phi\|_\infty$. More precisely, differentiating (3.28) once more in t and using $\|\ddot\alpha_i^{(t)}\|_{L^2}\le C_\varepsilon\|\delta\Phi\|_\infty/\varepsilon$ (by the same projection argument), we obtain

$$\sup_{t\in[0,1]}|R_1(t)| \le \frac{C_\varepsilon\|\delta\Phi\|_\infty^2}{\varepsilon^2}. \tag{3.34}$$

Combining (3.27), (3.31), and (3.34),

$$\begin{aligned}\mathrm{KL}(C_\varepsilon \,\|\, C'_\varepsilon) = F(1) &= \frac{1}{2}F''(0) + \frac{1}{2}\int_0^1 (1-t)\big[F''(t) - F''(0)\big]\,dt \\ &\le \frac{1}{2}\cdot\frac{4\|\delta\Phi\|_\infty^2}{\varepsilon^2} + \frac{1}{2}\cdot\sup_{t\in[0,1]}|R_1(t)| \\ &\le \left(\frac{2}{\varepsilon^2} + \frac{C_\varepsilon}{2\varepsilon^2}\right)\|\Phi - \Phi'\|_\infty^2 \\ &= K_\varepsilon\,\|\Phi - \Phi'\|_\infty^2,\end{aligned}$$

where

$$K_\varepsilon = \frac{4 + C_\varepsilon}{2\varepsilon^2},$$

and $C_\varepsilon > 0$ depends only on ε, $\|\Phi\|_\infty$, and $\|\Phi'\|_\infty$. This establishes the claimed inequality. □

Corollary 3.7 (Plug-in estimator)**.** *Given an empirical estimator $\widehat{\Phi}_n$ of Φ converging at rate r_n in L^∞, the plug-in copula $\widehat{C}_{\varepsilon,n}$ satisfies* $\mathrm{KL}(\widehat{C}_{\varepsilon,n} \,\|\, C_\varepsilon) = O_{\mathbb{P}}(r_n^2)$.

Proof.
(a) Setup and notation.

Let $X_1, \dots, X_n \overset{\text{iid}}{\sim} P$ be observations on $\mathbb{I}^d$, and let $\widehat{\Phi}_n : \mathbb{I}^d \to \mathbb{R}$ be an empirical estimator of the potential $\Phi \in L^\infty(\mathbb{I}^d)$ satisfying, by assumption,

$$\left\|\widehat{\Phi}_n - \Phi\right\|_\infty = O_{\mathbb{P}}(r_n), \tag{3.35}$$

meaning that for every $\eta > 0$ there exists $M_\eta > 0$ such that

$$\mathbb{P}\Big(\left\|\widehat{\Phi}_n - \Phi\right\|_\infty > M_\eta\, r_n\Big) \le \eta \quad \text{for all } n \ge 1.$$

Define:

- C_ε: the Sinkhorn–Wasserstein copula associated with Φ (Definition 3.1), with density c_ε.
- $\widehat{C}_{\varepsilon,n}$: the plug-in copula obtained by substituting $\widehat{\Phi}_n$ for Φ in problem (3.1), with density $\widehat{c}_{\varepsilon,n}$.

Both c_ε and $\widehat{c}_{\varepsilon,n}$ are well-defined almost surely by Theorem 3.2, provided $\widehat{\Phi}_n \in L^\infty(\mathbb{I}^d)$ almost surely, which we assume throughout.

(b) Uniform boundedness of $\widehat{\Phi}_n$.

From (3.35), for all n sufficiently large and outside an event of probability at most η,

$$\left\|\widehat{\Phi}_n\right\|_\infty \le \|\Phi\|_\infty + M_\eta\, r_n \le \|\Phi\|_\infty + M_\eta\, r_1 =: B < \infty, \tag{3.36}$$

where we used $r_n \le r_1$ for all $n \ge 1$ without loss of generality (since $r_n \to 0$). Hence, on the event $\mathcal{A}_{n,\eta} := \{\|\widehat{\Phi}_n - \Phi\|_\infty \le M_\eta\, r_n\}$, which satisfies $\mathbb{P}(\mathcal{A}_{n,\eta}) \ge 1-\eta$, we have simultaneously:

$$\left\|\widehat{\Phi}_n - \Phi\right\|_\infty \le M_\eta\, r_n, \qquad \left\|\widehat{\Phi}_n\right\|_\infty \le B, \qquad \|\Phi\|_\infty \le B. \tag{3.37}$$

(c) Application of Theorem 3.6.

On the event $\mathcal{A}_{n,\eta}$, both potentials $\widehat{\Phi}_n$ and Φ satisfy the boundedness condition $\|\cdot\|_\infty \le B$, so Theorem 3.6 applies with $\Phi' \leftarrow \Phi$ and $\Phi \leftarrow \widehat{\Phi}_n$. This yields the deterministic bound, valid on $\mathcal{A}_{n,\eta}$,

$$\mathrm{KL}\Big(\widehat{C}_{\varepsilon,n} \,\Big\|\, C_\varepsilon\Big) \;\le\; K_\varepsilon(B)\cdot \Big\|\widehat{\Phi}_n - \Phi\Big\|_\infty^2, \tag{3.38}$$

where

$$K_\varepsilon(B) := \frac{4 + C_\varepsilon(B)}{2\varepsilon^2} < \infty \tag{3.39}$$

is the constant from Theorem 3.6, which depends only on ε and B (hence is deterministic and uniform over the event $\mathcal{A}_{n,\eta}$).

Substituting (3.35) into (3.38), on the event $\mathcal{A}_{n,\eta}$ we obtain

$$\mathrm{KL}\Big(\widehat{C}_{\varepsilon,n} \,\Big\|\, C_\varepsilon\Big) \;\le\; K_\varepsilon(B)\cdot M_\eta^2\, r_n^2. \tag{3.40}$$

Since $K_\varepsilon(B)$ and M_η are both finite constants (the latter depending on η but not on n), we may write (3.40) as

$$\mathrm{KL}\Big(\widehat{C}_{\varepsilon,n} \,\Big\|\, C_\varepsilon\Big) \;\le\; L_{\varepsilon,\eta}\cdot r_n^2 \quad \text{on } \mathcal{A}_{n,\eta}, \tag{3.41}$$

where $L_{\varepsilon,\eta} := K_\varepsilon(B)\cdot M_\eta^2 < \infty$.

Fix $\delta > 0$ and $\eta > 0$ arbitrarily. From (3.41),

$$\begin{aligned}\mathbb{P}\Big(\mathrm{KL}\Big(\widehat{C}_{\varepsilon,n} \,\Big\|\, C_\varepsilon\Big) > \delta\, r_n^2\Big) &\le \mathbb{P}\Big(\mathrm{KL}\Big(\widehat{C}_{\varepsilon,n} \,\Big\|\, C_\varepsilon\Big) > \delta\, r_n^2,\ \mathcal{A}_{n,\eta}\Big) + \mathbb{P}\big(\mathcal{A}_{n,\eta}^c\big)\\ &\le \mathbb{P}\big(L_{\varepsilon,\eta}\, r_n^2 > \delta\, r_n^2\big) + \eta.\end{aligned}$$

Choosing $\delta > L_{\varepsilon,\eta}$, the first term on the right-hand side vanishes identically, giving

$$\mathbb{P}\left(\frac{\mathrm{KL}(\widehat{C}_{\varepsilon,n} \,\|\, C_\varepsilon)}{r_n^2} > \delta\right) \le \eta \quad \text{for all } n \ge 1.$$

Since $\eta > 0$ is arbitrary and $\delta = L_{\varepsilon,\eta}$ is finite, this is precisely the definition of

$$\mathrm{KL}\Big(\widehat{C}_{\varepsilon,n} \,\Big\|\, C_\varepsilon\Big) = O_{\mathbb{P}}(r_n^2),$$

which completes the proof. □

3.4. **Illustrative example : quadratic potential in dimension $d = 3$.** We instantiate the general theory on a fully explicit example in dimension $d = 3$. Throughout this section, $\mathbb{I} = [0,1]$ and $\mathbb{I}^3 = [0,1]^3$.
Let the potential $\Phi : \mathbb{I}^3 \to \mathbb{R}$ be defined by

$$\Phi(u_1,u_2,u_3) = \tfrac{1}{2}\left(u_1-u_2\right)^2 + \tfrac{1}{2}\left(u_2-u_3\right)^2 + \tfrac{1}{2}\left(u_1-u_3\right)^2, \qquad (u_1,u_2,u_3) \in \mathbb{I}^3. \tag{3.42}$$

Expanding, one verifies that

$$\Phi(u_1,u_2,u_3) = 2\big(u_1^2+u_2^2+u_3^2\big) - 2(u_1u_2+u_2u_3+u_1u_3) = \mathbf{u}^\top \mathbf{Q}\,\mathbf{u}, \tag{3.43}$$

where $\mathbf{u} = (u_1, u_2, u_3)^\top$ and

$$\mathbf{Q} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}. \tag{3.44}$$

The matrix $\mathbf{Q}$ is symmetric positive semi-definite, with eigenvalues $\lambda_1 = \lambda_2 = 3$ and $\lambda_3 = 0$ (corresponding to the direction $(1,1,1)^\top$). Hence $\Phi \geq 0$, $\Phi \in L^\infty(\mathbb{I}^3)$ with $\|\Phi\|_\infty = \Phi(0, \frac{1}{2}, 1) \vee \Phi(1, \frac{1}{2}, 0) = \frac{3}{2}$, and

$$\arg\min_{\mathbb{I}^3} \Phi = \big\{(t,t,t) : t \in \mathbb{I}\big\}, \tag{3.45}$$

the main diagonal of $\mathbb{I}^3$.

By Theorem 3.2, the Sinkhorn–Wasserstein copula C_ε associated with Φ in (3.42) admits a density of the form

$$c_\varepsilon(u_1,u_2,u_3) \;=\; \exp\!\left(-\frac{\Phi(u_1,u_2,u_3)}{\varepsilon} + \alpha_1^{(\varepsilon)}(u_1) + \alpha_2^{(\varepsilon)}(u_2) + \alpha_3^{(\varepsilon)}(u_3)\right), \tag{3.46}$$

where the potentials $\alpha_i^{(\varepsilon)} : \mathbb{I} \to \mathbb{R}$, $i = 1,2,3$, are determined by the marginal constraints

$$\int_{\mathbb{I}^2} c_\varepsilon(u_1,u_2,u_3)\, du_j\, du_k = 1 \quad \text{for a.e. } u_i \in \mathbb{I}, \quad \{i,j,k\} = \{1,2,3\}. \tag{3.47}$$

The potential Φ in (3.42) is *exchangeable*: $\Phi(u_{\sigma(1)}, u_{\sigma(2)}, u_{\sigma(3)}) = \Phi(u_1,u_2,u_3)$ for every permutation $\sigma \in \mathfrak{S}_3$. By uniqueness of the solution (Theorem 3.2) and uniqueness of the Sinkhorn potentials, we deduce

$$\alpha_1^{(\varepsilon)} = \alpha_2^{(\varepsilon)} = \alpha_3^{(\varepsilon)} = \alpha^{(\varepsilon)}, \tag{3.48}$$

so the density (3.46) reduces to

$$c_\varepsilon(u_1,u_2,u_3) = \exp\!\left(-\frac{\mathbf{u}^\top \mathbf{Q}\, \mathbf{u}}{\varepsilon} + \alpha^{(\varepsilon)}(u_1) + \alpha^{(\varepsilon)}(u_2) + \alpha^{(\varepsilon)}(u_3)\right). \tag{3.49}$$

The constraint (3.47) for $i = 1$ reads

$$e^{\alpha^{(\varepsilon)}(u_1)} \int_{\mathbb{I}^2} \exp\!\left(-\frac{\mathbf{u}^\top \mathbf{Q}\, \mathbf{u}}{\varepsilon} + \alpha^{(\varepsilon)}(u_2) + \alpha^{(\varepsilon)}(u_3)\right) du_2\, du_3 = 1 \quad \text{a.e. } u_1 \in \mathbb{I}. \tag{3.50}$$

Expanding $\mathbf{u}^\top \mathbf{Q}\, \mathbf{u}$ and isolating the u_1-dependent terms:

$$\mathbf{u}^\top \mathbf{Q}\, \mathbf{u} = 2u_1^2 - 2u_1(u_2+u_3) + \underbrace{2u_2^2 + 2u_3^2 - 2u_2u_3}_{=:\,\Psi(u_2,u_3)}. \tag{3.51}$$

Substituting into (3.50) gives the fixed-point equation

$$\alpha^{(\varepsilon)}(u_1) = -\log \int_{\mathbb{I}^2} \exp\!\left(\frac{2u_1(u_2+u_3) - 2u_1^2 - \Psi(u_2,u_3)}{\varepsilon} + \alpha^{(\varepsilon)}(u_2) + \alpha^{(\varepsilon)}(u_3)\right) du_2\, du_3, \tag{3.52}$$

which is solved iteratively by the Sinkhorn algorithm.

As $\varepsilon \to +\infty$, the penalization $\Phi/\varepsilon \to 0$ uniformly on $\mathbb{I}^3$. The density (3.49) satisfies

$$c_\varepsilon(u_1,u_2,u_3) \longrightarrow \exp\big(\alpha^{(\infty)}(u_1)+\alpha^{(\infty)}(u_2)+\alpha^{(\infty)}(u_3)\big).$$

Imposing the marginal constraints forces $e^{\alpha^{(\infty)}(u_i)} \equiv 1$ for each i, hence $\alpha^{(\infty)} \equiv 0$, and

$$\lim_{\varepsilon\to+\infty} c_\varepsilon(u_1,u_2,u_3) = 1 \quad\Longrightarrow\quad C_\varepsilon \xrightarrow{\varepsilon\to+\infty} \Pi, \tag{3.53}$$

where $\Pi(u_1,u_2,u_3) = u_1u_2u_3$ is the independence copula. Quantitatively, a second-order expansion in $1/\varepsilon$ yields

$$\mathrm{KL}(C_\varepsilon \,\|\, \Pi) = \frac{1}{\varepsilon^2}\cdot \mathrm{Var}_\Pi[\Phi] + O(\varepsilon^{-3}), \tag{3.54}$$

where $\mathrm{Var}_\Pi[\Phi] := \int_{\mathbb{I}^3}\Phi^2\,d\lambda_3 - \big(\int_{\mathbb{I}^3}\Phi\,d\lambda_3\big)^2$. For the potential (3.42), a direct computation gives

$$\int_{\mathbb{I}^3}\Phi\,d\lambda_3 = 2\cdot 3\cdot\frac{1}{3} - 2\cdot 3\cdot\frac{1}{4} = 2-\frac{3}{2} = \frac{1}{2}, \qquad \mathrm{Var}_\Pi[\Phi] = \frac{7}{60}.$$

As $\varepsilon \to 0^+$, the density c_ε concentrates on the minimizers of Φ. From (3.45), the minimizer set is the main diagonal $\Delta = \{(t,t,t) : t \in \mathbb{I}\}$. More precisely, write

$$c_\varepsilon(u_1,u_2,u_3) \propto \exp\left(-\frac{\mathbf{u}^\top \mathbf{Q}\,\mathbf{u}}{\varepsilon}\right)\cdot \exp\big(\alpha^{(\varepsilon)}(u_1)+\alpha^{(\varepsilon)}(u_2)+\alpha^{(\varepsilon)}(u_3)\big)\,.$$

The Laplace method on $\mathbb{I}^3$ gives, for any continuous test function $f : \mathbb{I}^3 \to \mathbb{R}$,

$$\int_{\mathbb{I}^3} f\,dC_\varepsilon \xrightarrow{\varepsilon\to0^+} \int_{\mathbb{I}} f(t,t,t)\,dt, \tag{3.55}$$

since the uniform measure on Δ is the unique probability measure on Δ with uniform marginals. Hence

$$C_\varepsilon \xrightarrow{\varepsilon\to0^+} C_0, \quad \text{where } C_0(u_1,u_2,u_3) = \min(u_1,u_2,u_3) = M_3(u_1,u_2,u_3), \tag{3.56}$$

the *upper Fréchet–Hoeffding bound* in dimension $d = 3$, i.e., the comonotone copula. The rate of convergence is controlled by the curvature of Φ transversally to Δ: the Hessian of Φ restricted to $\Delta^\perp = \mathrm{span}\{(1,-1,0),(1,0,-1)\}$ is $\mathbf{Q}|_{\Delta^\perp} = 3I_2$, so

$$\mathrm{KL}(C_\varepsilon \,\|\, C_0) = O(\varepsilon\log\varepsilon^{-1}) \quad \text{as } \varepsilon\to0^+. \tag{3.57}$$

At $\varepsilon = 1$, the copula C_1 captures a genuine trade-off between the comonotone structure and the independence baseline. The density takes the explicit (unnormalized) form

$$c_1(u_1,u_2,u_3) \propto \exp\Big(-2u_1^2-2u_2^2-2u_3^2+2u_1u_2+2u_2u_3+2u_1u_3+\alpha^{(1)}(u_1)+\alpha^{(1)}(u_2)+\alpha^{(1)}(u_3)\Big), \tag{3.58}$$

where $\alpha^{(1)}$ is determined numerically by iterating (3.52). The Kendall matrix $\boldsymbol{\tau}(C_1)$ has off-diagonal entries

$$\tau_{ij}(C_1) = 4\int_{\mathbb{I}^3}\int_{\mathbb{I}^3}\mathbf{1}[(u_i-v_i)(u_j-v_j)>0]\,dC_1(\mathbf{u})\,dC_1(\mathbf{v}) - 1 \;\in\; (0,1), \tag{3.59}$$

and by exchangeability, $\tau_{12}(C_1) = \tau_{13}(C_1) = \tau_{23}(C_1) = \tau_1$. The geodesic interpretation (Theorem 3.5) places C_1 at the midpoint of the Fisher-Rao geodesic from M_3 to Π, up to reparametrization.

Summary.

ε	Limit copula	Interpretation	Rate
$\varepsilon \to +\infty$	$\Pi = u_1 u_2 u_3$	Full independence	$\mathrm{KL} = O(\varepsilon^{-2})$
$\varepsilon = 1$	C_1 (semi-explicit)	Intermediate dependence	-
$\varepsilon \to 0^+$	$M_3 = \min(u_1, u_2, u_3)$	Comonotone (Fréchet upper bound)	$\mathrm{KL} = O(\varepsilon \log \varepsilon^{-1})$

4. Conclusion

We have introduced the Sinkhorn–Wasserstein copula family $\{C_\varepsilon\}_{\varepsilon>0}$ as the solution to a regularized dependence-modelling problem, and established its principal theoretical properties: existence and uniqueness via a Schrödinger-type structure, geodesic interpolation between independence and comonotonicity in the Fisher-Rao geometry, and quantitative stability with respect to potential perturbations. The plug-in estimator inherits an $O_{\mathbb{P}}(r_n^2)$ convergence rate, making the framework directly amenable to statistical practice. These results open natural avenues toward high-dimensional inference, data-driven potential learning, and connections with optimal transport theory.

Acknowledgement. Acknowledgements could be placed at the end of the text but precede the references.

References

1. Amari, S. (2016). Information geometry and its applications. Springer. https://doi.org/10.1007/978-4-431-55978-8
2. Cuturi, M. (2013). Sinkhorn distances: lightspeed computation of optimal transport. Advances in Neural Information Processing Systems, 26, 2292-2300. https://doi.org/10.48550/arXiv.1306.0895
3. Feydy, J., Séjourné, T., Vialard, F., Amari, S., Trouvé, A., & Peyré, G. (2019). Interpolating between optimal transport and MMD using Sinkhorn divergences. Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics, 89, 2229-2238. https://doi.org/10.48550/arXiv.1810.08278
4. Genest, C., & Favre, J. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. Journal of Hydrologic Engineering, 12(4), 347-368. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347)
5. Herman Tiemtoré & Remi Guillaume Bagré. Dynamics of dependency: finite difference solutions to the Black-Scholes PDE with copulas. Vol 19, Issue 1 (2025) 374-391. https://doi.org/10.56947/gjom.v19i1.2576
6. Nelsen, R. B. (2006). An introduction to copulas. Springer. https://doi.org/10.1007/0-387-28678-0
7. Peyré, G., & Cuturi, M. (2019). Computational optimal transport. Foundations and Trends in Machine Learning, 11(5-6), 355-607. https://doi.org/10.1561/2200000073
8. Santambrogio, F. (2015). Optimal transport for applied mathematicians. Birkhäuser. https://doi.org/10.1007/978-3-319-20828-2

9. Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris, 8, 229-231.
10. Villani, C. (2009). Optimal transport: old and new. Springer. https://doi.org/10.1007/978-3-540-71050-9

[1] UFR-ST, Université Norbert ZONGO (UNZ), BP 376 Koudougou, Burkina Faso
Email address: baremgui@yahoo.fr

[2] UFR-ST, Université Norbert ZONGO (UNZ), BP 376 Koudougou, Burkina Faso
Email address: hermanteimto94@gmail.com

[3] UFR-ST, Université Norbert ZONGO (UNZ), BP 376 Koudougou, Burkina Faso
Email address: jacques78kabore@gmail.com

[4] UFR-ST, Université Norbert ZONGO (UNZ), BP 376 Koudougou, Burkina Faso
Email address: tagawere2@gmail.com

[5] UFR-ST, Université Norbert ZONGO (UNZ), BP 376 Koudougou, Burkina Faso
Email address: souleymaneo626@gmail.com