

Strong Approximations for Multivariate Integrated Empirical Processes: Two-Sample Testing, Change-Point Detection, and Inference with Estimated Parameters

Salim Bouzebda

Université de Technologie de Compiègne, LMAC (Laboratory of Applied Mathematics of Compiègne),
Compiègne, France
Email: salim.bouzebda@utc.fr

Abstract. We investigate a broad class of multivariate p -fold integrated empirical processes obtained by iterated lower-orthant Stieltjes integration of the empirical distribution function on $[0, 1]^d$. After deriving exact structural representations for the corresponding integrated distribution and empirical distribution functionals, we show that, in the scalarizable regime (including the univariate and effectively totally ordered/comonotone settings), the p -fold integrated empirical process admits a precise reduction to the ordinary multivariate empirical process, with leading term $(p!)^{-1} F(\mathbf{u})^p \alpha_n(\mathbf{u})$. In the genuinely multivariate partial-order setting, the first-order structure is governed instead by the chain functional and its influence function. This algebraic decomposition permits a systematic transfer of strong approximation results from the classical empirical process to its integrated counterparts. More specifically, on suitable probability spaces, we construct almost sure couplings of the integrated empirical process with appropriately weighted F -Brownian bridges and F -Kiefer processes, and we obtain explicit uniform approximation rates over $[0, 1]^d$. These couplings yield asymptotic approximations for integrated Kolmogorov–Smirnov and Cramér–von Mises type functionals, as well as a functional law of the iterated logarithm. We then exploit the resulting Gaussian approximations in two inferential directions: first, for multivariate two-sample procedures based on modified p -fold integrated empirical contrasts; second, for change-point statistics constructed from sequential integrated empirical discrepancies, for which asymptotically valid Gaussian calibrations are derived. We further analyze plug-in versions of the integrated empirical process in parametric settings and establish asymptotic Gaussian approximations in the presence of estimated nuisance parameters. A simulation study in dimension $d = 2$ complements the theory and illustrates the finite-sample behavior of the proposed procedures under dependence, marginal, and joint perturbation regimes.

Keywords: Multivariate integrated empirical process; strong approximation; Brownian bridge; Kiefer process; law of the iterated logarithm; two-sample testing; change-point analysis; estimated parameters

MSC (2020): 60F17, 62G30, 62G20, 60F15, 62F03, 62F12

1 Introduction

Let $\{\mathbf{X}_i : i \in \mathbb{N}^*\}$ be a sequence of independent and identically distributed (i.i.d.) $\mathbb{R}^d$ -valued random vectors with continuous distribution function

$$F(\mathbf{u}) = \mathbb{P}\{\mathbf{X}_1 \leq \mathbf{u}\}, \quad \mathbf{u} \in [0, 1]^d,$$

where inequalities are understood componentwise: for $\mathbf{x} = (x_1, \dots, x_d)$ and $\mathbf{y} = (y_1, \dots, y_d)$, we write $\mathbf{x} \leq \mathbf{y}$ if and only if $x_j \leq y_j$ for all $j = 1, \dots, d$. Throughout, boldface letters denote vectors,

and we write

$$[\mathbf{0}, \mathbf{u}] := \prod_{j=1}^d [0, u_j].$$

For $n \in \mathbb{N}^*$, the empirical c.d.f. based on $\mathbf{X}_1, \dots, \mathbf{X}_n$ is

$$\mathbb{F}_n(\mathbf{u}) := \frac{1}{n} \# \{1 \leq i \leq n : \mathbf{X}_i \leq \mathbf{u}\} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\mathbf{X}_i \leq \mathbf{u}\}}, \quad \mathbf{u} \in [0, 1]^d,$$

and the associated empirical process is

$$\alpha_n(\mathbf{u}) := \sqrt{n}(\mathbb{F}_n(\mathbf{u}) - F(\mathbf{u})), \quad \mathbf{u} \in [0, 1]^d.$$

The asymptotic theory of α_n indexed by lower orthants is one of the core pillars of empirical process theory, both because of its intrinsic probabilistic structure and because it underlies a vast range of inferential procedures, from goodness-of-fit and two-sample testing to sequential analysis and structural-break detection; see, among many classical sources, [7, 8, 23, 25, 27–29, 40–44].

The present work is devoted to a broad class of *multivariate integrated empirical processes*, obtained by iterating lower-orthant Stieltjes integration with respect to the population and empirical distribution functions. More precisely, for $p \in \mathbb{N}$, define recursively

$$F^{(0)}(\mathbf{u}) := F(\mathbf{u}), \quad \mathbb{F}_n^{(0)}(\mathbf{u}) := \mathbb{F}_n(\mathbf{u}),$$

and

$$F^{(1)}(\mathbf{u}) := \int_{[\mathbf{0}, \mathbf{u}]} F(\mathbf{s}) dF(\mathbf{s}), \quad \mathbb{F}_n^{(1)}(\mathbf{u}) := \int_{[\mathbf{0}, \mathbf{u}]} \mathbb{F}_n(\mathbf{s}) d\mathbb{F}_n(\mathbf{s}),$$

while, for $p \geq 2$,

$$\begin{aligned} F^{(p)}(\mathbf{u}) &:= \int_{[\mathbf{0}, \mathbf{u}]} dF(\mathbf{s}_1) \int_{[\mathbf{0}, \mathbf{s}_1]} dF(\mathbf{s}_2) \cdots \int_{[\mathbf{0}, \mathbf{s}_{p-1}]} F(\mathbf{s}_p) dF(\mathbf{s}_p), \\ \mathbb{F}_n^{(p)}(\mathbf{u}) &:= \int_{[\mathbf{0}, \mathbf{u}]} d\mathbb{F}_n(\mathbf{s}_1) \int_{[\mathbf{0}, \mathbf{s}_1]} d\mathbb{F}_n(\mathbf{s}_2) \cdots \int_{[\mathbf{0}, \mathbf{s}_{p-1}]} \mathbb{F}_n(\mathbf{s}_p) d\mathbb{F}_n(\mathbf{s}_p). \end{aligned}$$

The corresponding integrated empirical process is

$$\alpha_n^{(p)}(\mathbf{u}) := \sqrt{n}(\mathbb{F}_n^{(p)}(\mathbf{u}) - F^{(p)}(\mathbf{u})), \quad \mathbf{u} \in [0, 1]^d. \quad (1.1)$$

These objects are nonlinear lower-orthant transforms of the empirical measure and may be interpreted as iterated order-preserving cumulative functionals of $\mathbb{F}_n$. In particular, they interpolate between the primitive empirical process and progressively more regularized discrepancy processes, while preserving the partial-order geometry of the lower-orthant index class.

A first structural point is decisive. In dimension one, and more generally in situations where the underlying order becomes effectively total, iterated Stieltjes integration collapses to explicit polynomial transforms of the c.d.f. In the genuinely multivariate case, however, the componentwise order is only partial, and the objects $F^{(p)}$ and $\mathbb{F}_n^{(p)}$ are governed not by a scalar algebra in $F(\mathbf{u})$, but by the combinatorics of chains in the lower-orthant order. Thus, any formulation of the theory that identifies $F^{(p)}(\mathbf{u})$ universally with a polynomial in $F(\mathbf{u})$ is, in arbitrary dimension, structurally incorrect. The correct general statement is the following.

Proposition 1.1 (Structural representation and special closed forms). *For each fixed $p \in \mathbb{N}$ and $\mathbf{u} \in [0, 1]^d$, the following identities hold:*

$$F^{(p)}(\mathbf{u}) = \mathbb{P}\{\mathbf{X}_{p+1} \leq \mathbf{X}_p \leq \cdots \leq \mathbf{X}_1 \leq \mathbf{u}\}, \quad (1.2)$$

and, almost surely,

$$\mathbb{F}_n^{(p)}(\mathbf{u}) = \frac{1}{n^{p+1}} \sum_{i_1, \dots, i_{p+1}=1}^n \mathbb{1}_{\{\mathbf{X}_{i_{p+1}} \leq \mathbf{X}_{i_p} \leq \cdots \leq \mathbf{X}_{i_1} \leq \mathbf{u}\}}. \quad (1.3)$$

In particular,

$$0 \leq F^{(p)}(\mathbf{u}) \leq \frac{F(\mathbf{u})^{p+1}}{(p+1)!}.$$

If, in addition, the componentwise order is total on the support of F (for instance in the univariate case $d = 1$, or under a comonotonicity-type assumption), then the closed forms

$$F^{(p)}(\mathbf{u}) = \frac{F(\mathbf{u})^{p+1}}{(p+1)!}, \quad \mathbb{F}_n^{(p)}(\mathbf{u}) = \frac{1}{n^{p+1}} \binom{n\mathbb{F}_n(\mathbf{u}) + p}{p+1} \quad a.s. \quad (1.4)$$

hold for every $\mathbf{u} \in [0, 1]^d$.

The identities (1.2)–(1.3) reveal the correct order-theoretic content of the problem: in full generality, $F^{(p)}(\mathbf{u})$ is a chain probability and $\mathbb{F}_n^{(p)}(\mathbf{u})$ is its empirical V-statistic analogue. Consequently, except in the special regimes covered by (1.4), the integrated functionals cannot be reduced to a universal scalar transform of $F(\mathbf{u})$ alone. This distinction is not a technical subtlety but the precise point at which the multivariate theory departs from the univariate one. It also explains why the asymptotic analysis must be carried out at the process level rather than through pointwise algebraic identities. For multivariate Stieltjes integration over lower orthants we refer to [8]; a more detailed structural discussion is given later in Proposition 7.3.

Integrated empirical processes have a substantial history in the goodness-of-fit literature. In the univariate setting, Henze and Nikitin [32, 33] introduced and analyzed a family of statistics based on integrated empirical processes, including Kolmogorov–Smirnov, Cramér–von Mises, and Watson-type functionals, and showed that the corresponding procedures may possess striking efficiency properties under suitable alternatives. The two-sample counterpart was developed in [34], where integrated two-sample empirical processes were shown to produce competitive and, in certain regimes, superior procedures relative to more classical statistics. A higher-order variant, based on repeated integration, was proposed in [39]. In the multivariate domain, Jing [36, 37] investigated projected integrated empirical processes for equality-of-distribution problems, while [14, 17] developed multivariate two-sample and K -sample methodologies driven by integrated empirical copula processes. Weighted multivariate Cramér–von Mises type statistics under contiguous alternatives were considered in [19]; see also the subsequent developments in [3, 4].

From the probabilistic viewpoint, the central issue is whether the iterated lower-orthant transforms defining $\alpha_n^{(p)}$ preserve the strong approximation architecture of the ordinary empirical process. This is the decisive question for at least three reasons. First, strong approximation by Brownian bridges yields quantitative control of nonlinear functionals and therefore translates directly into asymptotic approximations for integrated Kolmogorov–Smirnov and Cramér–von Mises type statistics. Second, approximation by Kiefer fields transfers almost-sure fluctuation theory, including functional laws of the iterated logarithm, to the integrated setting. Third, once one passes to plug-in regimes with estimated parameters or to sequential schemes relevant for two-sample and change-point analysis, strong approximation becomes the most efficient device for importing Gaussian asymptotics into nonstandard empirical functionals. This paradigm traces back to the

seminal coupling theory of Komlós, Major and Tusnády [38] and to its subsequent extensions in strong approximation theory; see [25, 27, 29, 44].

The first objective of the present paper is therefore to establish strong approximations for the p -fold integrated empirical process. At the Brownian-bridge level, we prove that $\alpha_n^{(p)}$ can be coupled uniformly over $[0, 1]^d$ with an explicit Gaussian transform whose leading term is a multiplicative deformation of the underlying bridge by the deterministic factor $F(\mathbf{u})^p/p!$. At the Kiefer-process level, we derive a natural $\sqrt{n}$ -scale strong approximation from which a functional law of the iterated logarithm follows. These results show that, despite the nonlinear and iterated nature of the lower-orthant transform, the integrated empirical process remains within the reach of the classical strong invariance machinery. In particular, the integrated setting inherits a Gaussian approximation theory sufficiently sharp to support asymptotic analysis of suprema, quadratic functionals, and sequential statistics.

The second objective is inferential. We exploit the preceding strong approximations in two nontrivial statistical directions. The first is the multivariate two-sample problem, where we introduce p -fold integrated discrepancy processes and their power-modified analogues, and derive Gaussian approximations for the associated Kolmogorov–Smirnov and Cramér–von Mises type statistics. The second is the change-point problem, where sequential integrated empirical contrasts furnish natural detectors of structural instability. In the latter case, the approximation by Kiefer-type fields is particularly effective, since it connects the integrated empirical detectors to the classical Gaussian framework underlying asymptotic change-point theory; see [2, 6, 10, 11, 15, 16, 18, 21, 22, 26, 35].

The third objective is to treat the parameter-estimated framework. When the null model belongs to a parametric family and the parameter is unknown, the empirical process is perturbed by plug-in effects whose first-order asymptotics are governed by stochastic linearization and score-type expansions. Starting from Durbin’s classical treatment of empirical processes with estimated parameters [30] and following the strong approximation program of Burke and Csörgő [20], we show that the integrated empirical process with estimated parameters still admits a Gaussian approximation, uniformly over the lower-orthant index set, provided the estimator satisfies the standard asymptotic linear representation and the parametric family enjoys suitable regularity. This extension is indispensable for composite goodness-of-fit hypotheses and for the implementation of fully data-driven integrated procedures.

The contribution of the paper is therefore simultaneously structural, probabilistic, and inferential. Structurally, it formulates the multivariate integrated theory in a manner consistent with the underlying partial order, thereby avoiding an oversimplified scalarization that is valid only in effectively total-order regimes. Probabilistically, it establishes strong Gaussian embeddings for a broad class of iterated lower-orthant empirical functionals. Inferentially, it shows that these embeddings are sharp enough to yield implementable asymptotic theory for multivariate two-sample testing, change-point analysis, and parameter-estimated problems. In this sense, the paper extends the classical strong approximation program from primitive empirical processes to a nonlinear class of multivariate integral transforms whose theory, particularly beyond the copula setting, remains comparatively underdeveloped.

The paper is organized as follows. Section 2 introduces the Gaussian reference processes and develops the strong approximation theory for the p -fold integrated empirical process. Section 3 presents the main statistical applications, including multivariate two-sample testing and change-point detection. Section 4 addresses the parameter-estimated framework and derives the corresponding Gaussian approximations. Section 5 reports a simulation study assessing the finite-sample behavior of the proposed procedures. Section 6 concludes the paper with a discussion and some perspectives for future work. The technical proofs and additional structural developments are deferred to Section 7.

2 Strong approximation

The purpose of this section is to identify the precise Gaussian objects that govern the first-order asymptotics of the p -fold integrated empirical process and to show that, in the scalarizable regime, the latter inherits the full strong approximation machinery of the ordinary multivariate empirical process. At a conceptual level, the problem is subtle. The map $\mathbb{F}_n \mapsto \mathbb{F}_n^{(p)}$ is nonlinear, nonlocal, and defined through iterated lower-orthant Stieltjes integration. A priori, there is no reason for such a transform to preserve the Brownian-bridge/Kiefer structure of the primitive empirical process. The point of the present analysis is that, under the structural reduction available in the effectively totally ordered case, the integrated process is asymptotically nothing but a deterministic multiplicative deformation of the ordinary empirical process. This observation allows one to transport, almost verbatim, the classical strong invariance principles to the integrated setting.

It is important to stress at the outset that the results of this section are proved under the regime in which the iterated integrated functionals admit a scalar polynomial representation in the distribution function itself. This includes, in particular, the univariate case and more generally any setting in which the componentwise order becomes total on the support of the law.

Assumption 1 (Scalarizable integrated regime). Throughout this section, for each fixed $p \in \mathbb{N}^*$, we assume that

$$F^{(p)}(\mathbf{u}) = \frac{F(\mathbf{u})^{p+1}}{(p+1)!}, \quad \mathbb{F}_n^{(p)}(\mathbf{u}) = \frac{\mathbb{F}_n(\mathbf{u})^{p+1}}{(p+1)!} + \sum_{k=1}^p a_k^{(p)} \frac{\mathbb{F}_n(\mathbf{u})^k}{n^{p-k+1}}, \quad (2.1)$$

for every $\mathbf{u} \in [0, 1]^d$, where the coefficients $a_k^{(p)}$ are deterministic constants depending only on p .

Remark 2.1. Assumption 1 is automatic in dimension $d = 1$, and also in comonotonic-type situations where the support is totally ordered under the componentwise order. It is precisely under this hypothesis that the linearization of $\mathbb{F}_n^{(p)}$ around F reduces to the scalar multiplier $\frac{1}{p!}F(\mathbf{u})^p$. In the genuinely multivariate partial-order setting, where $F^{(p)}$ is governed by chain probabilities rather than a universal polynomial in $F(\mathbf{u})$, the first-order term must instead be expressed through the Fréchet derivative of the chain functional. Thus the present section gives the sharp strong approximation theory in the scalarizable regime; the general partial-order case requires an operator-level reformulation.

Remark 2.2. From a functional-analytic viewpoint, the mechanism is a strong-approximation counterpart of the functional delta method. Under (2.1), the map

$$x \mapsto \frac{x^{p+1}}{(p+1)!}$$

has derivative $x \mapsto x^p/p!$, so that the first-order perturbation of $\mathbb{F}_n^{(p)}$ around F is

$$\frac{1}{p!}F(\mathbf{u})^p (\mathbb{F}_n(\mathbf{u}) - F(\mathbf{u})),$$

that is,

$$\frac{1}{p!}F(\mathbf{u})^p \frac{\alpha_n(\mathbf{u})}{\sqrt{n}}.$$

The strong approximation theory below shows that this heuristic linearization is not merely valid in distribution, but can be realized on a common probability space with explicit almost sure rates.

Gaussian reference processes

We begin by fixing the Gaussian processes that will appear as canonical approximants.

The F -Wiener process. Let F be a continuous distribution function on $[0, 1]^d$, and write

$$\mathbf{u} \wedge \mathbf{v} := (\min\{u_1, v_1\}, \dots, \min\{u_d, v_d\}), \quad \mathbf{u}, \mathbf{v} \in [0, 1]^d.$$

A d -variate F -Wiener process, or equivalently an F -Brownian motion indexed by lower orthants, is a centered Gaussian process

$$\mathbb{W}_F = \{\mathbb{W}_F(\mathbf{u}) : \mathbf{u} \in [0, 1]^d\}$$

such that

$$\mathbb{E} \mathbb{W}_F(\mathbf{u}) = 0, \quad \mathbb{E}[\mathbb{W}_F(\mathbf{u})\mathbb{W}_F(\mathbf{v})] = F(\mathbf{u} \wedge \mathbf{v}), \quad \mathbf{u}, \mathbf{v} \in [0, 1]^d,$$

and $\mathbb{W}_F(\mathbf{u}) = 0$ whenever at least one coordinate of $\mathbf{u}$ is equal to 0.

The covariance kernel $F(\mathbf{u} \wedge \mathbf{v})$ is the natural lower-orthant analogue of the kernel $(s, t) \mapsto s \wedge t$ from one-dimensional empirical process theory. In particular, the geometry of the index class is entirely encoded through the partial order $\mathbf{u} \mapsto [\mathbf{0}, \mathbf{u}]$.

The F -Brownian bridge. The F -Brownian bridge is the tied-down Gaussian field

$$\mathbb{B}_F(\mathbf{u}) := \mathbb{W}_F(\mathbf{u}) - F(\mathbf{u})\mathbb{W}_F(1, \dots, 1), \quad \mathbf{u} \in [0, 1]^d.$$

It is centered and has covariance

$$\mathbb{E}[\mathbb{B}_F(\mathbf{u})\mathbb{B}_F(\mathbf{v})] = F(\mathbf{u} \wedge \mathbf{v}) - F(\mathbf{u})F(\mathbf{v}), \quad \mathbf{u}, \mathbf{v} \in [0, 1]^d.$$

This is exactly the covariance structure of the ordinary empirical process α_n , hence $\mathbb{B}_F$ is the intrinsic Gaussian limit attached to the lower-orthant empirical process.

We write $\{\mathbb{B}_{n,F}(\mathbf{u}) : \mathbf{u} \in [0, 1]^d\}$ for a sequence of independent copies of $\mathbb{B}_F$.

The F -Kiefer process. In order to approximate the empirical process simultaneously in its spatial and sample-size arguments, we consider a centered Gaussian field

$$\mathbb{W}_F(t, \mathbf{u}), \quad t \geq 0, \mathbf{u} \in [0, 1]^d,$$

with covariance

$$\mathbb{E}[\mathbb{W}_F(t, \mathbf{u})\mathbb{W}_F(s, \mathbf{v})] = (t \wedge s) F(\mathbf{u} \wedge \mathbf{v}), \quad s, t \geq 0, \mathbf{u}, \mathbf{v} \in [0, 1]^d,$$

and $\mathbb{W}_F(t, \mathbf{u}) = 0$ whenever one of the arguments $t, u_1, \dots, u_d$ is zero. The associated F -Kiefer process is then defined by

$$\mathbf{\Gamma}_F(t, \mathbf{u}) := \mathbb{W}_F(t, \mathbf{u}) - F(\mathbf{u})\mathbb{W}_F(t, 1, \dots, 1), \quad t \geq 0, \mathbf{u} \in [0, 1]^d.$$

Its covariance is

$$\mathbb{E}[\mathbf{\Gamma}_F(t, \mathbf{u})\mathbf{\Gamma}_F(s, \mathbf{v})] = (t \wedge s) (F(\mathbf{u} \wedge \mathbf{v}) - F(\mathbf{u})F(\mathbf{v})).$$

For each fixed $t > 0$, the self-similarity of the Gaussian field yields

$$t^{-1/2}\mathbf{\Gamma}_F(t, \mathbf{u}) \stackrel{\mathcal{L}}{=} \mathbb{B}_F(\mathbf{u}), \quad \mathbf{u} \in [0, 1]^d.$$

Remark 2.3. The distinction between $\mathbb{B}_F$ and $\mathbf{\Gamma}_F$ is not cosmetic. The bridge is the correct approximant for a fixed- n empirical process, whereas the Kiefer field is the correct object for sequential and almost sure fluctuation theory. In particular, laws of the iterated logarithm, weighted change-point limits, and uniform-in- n refinements naturally belong to the Kiefer framework.

Background on these Gaussian fields may be found, for instance, in [1, 12, 13]. As is customary in strong approximation theory, we shall assume without further comment that all processes are realized on a common probability space; see Appendix 2 of [25].

Classical strong couplings for the empirical process

The transfer principle developed below rests on two standard strong approximation results for the ordinary lower-orthant empirical process.

Theorem A ([9]). *On an appropriate probability space, one can construct the empirical processes $\{\alpha_n(\mathbf{u}) : \mathbf{u} \in [0, 1]^d, n \geq 1\}$ together with a sequence of F -Brownian bridges*

$$\{\mathbb{B}_{n,F}(\mathbf{u}) : \mathbf{u} \in [0, 1]^d, n \geq 1\}$$

such that, almost surely as $n \rightarrow \infty$,

$$\sup_{\mathbf{u} \in [0,1]^d} |\alpha_n(\mathbf{u}) - \mathbb{B}_{n,F}(\mathbf{u})| = O\left(n^{-1/(2(2d-1))} \log n\right).$$

Theorem B ([24]). *On an appropriate probability space, one can construct the empirical processes $\{\alpha_n(\mathbf{u}) : \mathbf{u} \in [0, 1]^d, n \geq 1\}$ together with an F -Kiefer process $\{\Gamma_F(n, \mathbf{u}) : n \geq 1, \mathbf{u} \in [0, 1]^d\}$ such that, almost surely as $n \rightarrow \infty$,*

$$\sup_{\mathbf{u} \in [0,1]^d} |\sqrt{n} \alpha_n(\mathbf{u}) - \Gamma_F(n, \mathbf{u})| = O\left(n^{1/2-1/(4d)} (\log n)^{3/2}\right).$$

Remark 2.4. The dimensional deterioration in the rates above is a manifestation of the complexity of the lower-orthant class. From the perspective of the present paper, however, the essential point is not optimality of these exponents, but the fact that the approximation is uniform over $[0, 1]^d$ and strong enough to dominate all nonlinear remainder terms generated by the integrated transform.

First-order reduction of the integrated empirical process

Under Assumption 1, the integrated empirical process

$$\alpha_n^{(p)}(\mathbf{u}) = \sqrt{n}(\mathbb{F}_n^{(p)}(\mathbf{u}) - F^{(p)}(\mathbf{u}))$$

admits an exact algebraic decomposition in terms of α_n and lower-order remainder terms. The leading term is

$$\frac{1}{p!} F(\mathbf{u})^p \alpha_n(\mathbf{u}),$$

which already suggests that the Gaussian approximation of $\alpha_n^{(p)}$ should be obtained by multiplying the Gaussian approximation of α_n by the same deterministic weight.

This is indeed the case.

Theorem 2.5. *Fix $p \in \mathbb{N}^*$. Under Assumption 1, on a suitable probability space one may construct the p -fold integrated empirical process $\{\alpha_n^{(p)} : n \in \mathbb{N}^*\}$ together with a sequence of F -Brownian bridges $\{\mathbb{B}_{n,F} : n \in \mathbb{N}^*\}$ such that*

$$\sup_{\mathbf{u} \in [0,1]^d} \left| \alpha_n^{(p)}(\mathbf{u}) - \frac{1}{p!} F(\mathbf{u})^p \mathbb{B}_{n,F}(\mathbf{u}) \right| = O\left(n^{-1/(2(2d-1))} \log n\right) \quad \text{a.s. as } n \rightarrow \infty.$$

Remark 2.6. Theorem 2.5 shows that the passage from the primitive empirical process to its p -fold integrated counterpart does not alter the strong approximation rate. The nonlinear transform modifies only the deterministic amplitude, through the factor $F(\mathbf{u})^p/p!$, but does not produce a new stochastic universality class. In this sense, the integrated empirical process is asymptotically rigid: its first-order random structure is already completely encoded in the ordinary empirical process.

Remark 2.7. The deterministic weight $F(\mathbf{u})^p$ has a nontrivial statistical meaning. As p increases, the contribution of regions where $F(\mathbf{u})$ is small is progressively attenuated. Thus the integrated process gives increasing emphasis to the interior of the support and decreasing emphasis to the extreme lower tail. This regularizing effect is visible both in supremum-type functionals and in quadratic statistics.

We next state the sequential analogue at the Kiefer-process level.

Theorem 2.8. *Under Assumption 1, on a suitable probability space one may construct the p -fold integrated empirical process $\{\alpha_n^{(p)} : n \in \mathbb{N}^*\}$ together with an F -Kiefer process $\{\Gamma_F(t, \mathbf{u}) : t \geq 0, \mathbf{u} \in [0, 1]^d\}$ such that, almost surely as $n \rightarrow \infty$,*

$$\sup_{\mathbf{u} \in [0, 1]^d} \left| \sqrt{n} \alpha_n^{(p)}(\mathbf{u}) - \frac{1}{p!} F(\mathbf{u})^p \Gamma_F(n, \mathbf{u}) \right| = O\left(n^{1/2-1/(4d)} (\log n)^{3/2}\right).$$

Remark 2.9. Theorem 3.4 is the genuinely dynamic version of Theorem 2.5. Once the integrated empirical process is embedded into a weighted Kiefer field, the entire arsenal of almost sure fluctuation theory becomes available. In particular, one may derive functional laws of the iterated logarithm, weighted sequential limit theorems, and change-point approximations without returning to the nonlinear definition of $\mathbb{F}_n^{(p)}$.

A law of the iterated logarithm

The Kiefer approximation immediately yields a law of the iterated logarithm for the integrated empirical process.

Corollary 2.10. *Let “a.s.” stand for “almost surely”. Under Assumption 1,*

$$\limsup_{n \rightarrow \infty} \left\{ \left(\frac{1}{2 \log \log n} \right)^{1/2} \sup_{\mathbf{u} \in [0, 1]^d} |\alpha_n^{(p)}(\mathbf{u})| \right\} = \sigma \quad a.s.,$$

where

$$\sigma^2 := \sup_{\mathbf{u} \in [0, 1]^d} \text{Var} \left(\frac{1}{p!} F(\mathbf{u})^p \Gamma_F(1, \mathbf{u}) \right).$$

Equivalently,

$$\sigma^2 = \frac{1}{(p!)^2} \sup_{\mathbf{u} \in [0, 1]^d} F(\mathbf{u})^{2p} \left(F(\mathbf{u}) - F(\mathbf{u})^2 \right) = \frac{1}{(p!)^2} \sup_{\mathbf{u} \in [0, 1]^d} F(\mathbf{u})^{2p+1} (1 - F(\mathbf{u})).$$

Remark 2.11. If the image of F over $[0, 1]^d$ is the full interval $[0, 1]$, then the LIL constant is explicit:

$$\sigma^2 = \frac{1}{(p!)^2} \max_{0 \leq x \leq 1} x^{2p+1} (1 - x) = \frac{1}{(p!)^2} \frac{(2p+1)^{2p+1}}{(2p+2)^{2p+2}}.$$

The maximizer occurs at $x = (2p+1)/(2p+2)$, again showing that the dominant oscillations of the integrated process are shifted away from the boundary and toward an interior quantile level.

Application to goodness-of-fit statistics

The strong approximations obtained above have immediate consequences for integrated goodness-of-fit procedures. Consider the null hypothesis

$$\mathcal{H}_0 : F = F_0,$$

where F_0 is a fixed continuous distribution function on $[0, 1]^d$, and define the integrated Kolmogorov–Smirnov and Cramér–von Mises statistics by

$$\mathbf{S}_n^{(p)} := \sup_{\mathbf{u} \in [0, 1]^d} \left| \sqrt{n} (\mathbb{F}_n^{(p)}(\mathbf{u}) - F_0^{(p)}(\mathbf{u})) \right|,$$

and

$$\mathbf{T}_n^{(p)} := n \int_{[0, 1]^d} (\mathbb{F}_n^{(p)}(\mathbf{u}) - F_0^{(p)}(\mathbf{u}))^2 dF_0(\mathbf{u}).$$

Under $\mathcal{H}_0$, Theorem 2.5 yields, almost surely as $n \rightarrow \infty$,

$$\left| \mathbf{S}_n^{(p)} - \frac{1}{p!} \sup_{\mathbf{u} \in [0, 1]^d} |F_0(\mathbf{u})^p \mathbb{B}_{n, F_0}(\mathbf{u})| \right| = O\left(n^{-1/(2(2d-1))} \log n\right), \quad (2.2)$$

$$\left| \mathbf{T}_n^{(p)} - \frac{1}{(p!)^2} \int_{[0, 1]^d} [F_0(\mathbf{u})^p \mathbb{B}_{n, F_0}(\mathbf{u})]^2 dF_0(\mathbf{u}) \right| = O\left(n^{-1/(2(2d-1))} \log n \sqrt{\log \log n}\right). \quad (2.3)$$

Relations (2.2)–(2.3) show that the integrated statistics are asymptotically governed by the same Gaussian bridge as the ordinary empirical statistics, but observed through a deterministic weighting mechanism. Thus integration does not destroy Gaussian calibrability; rather, it changes the functional geometry of the test. In particular, the limiting Kolmogorov statistic is a weighted bridge supremum, while the limiting Cramér–von Mises statistic is a weighted quadratic Gaussian functional.

For inferential purposes, the significance of the preceding theorems is that one does not need to analyze $\mathbb{F}_n^{(p)}$ directly at the nonlinear level. Once the strong approximation is established, asymptotic statements for suprema, quadratic functionals, weighted sequential statistics, and parameter-perturbed versions can all be developed by working with the corresponding weighted Gaussian fields. This reduction is precisely what makes the theory operational in the two-sample, change-point, and estimated-parameter settings studied later in the paper.

Notation. We write $X \stackrel{\mathcal{L}}{=} Y$ for equality in distribution. Unless otherwise specified, all Landau symbols are understood in the almost sure sense.

3 Statistical applications

This section translates the strong approximation results of Section 2 into concrete inferential procedures. The guiding principle is that, once the p -fold integrated empirical process is coupled uniformly with an explicit Gaussian field, the asymptotic analysis of nonlinear functionals may be carried out at the Gaussian level with only asymptotically negligible error. This is precisely the mechanism that allows one to transfer the classical empirical-process methodology to the present integrated setting. Two remarks should be kept in mind throughout. First, in the genuinely multivariate lower-orthant framework, the quantity $F^{(p)}(\mathbf{u})$ is not, in general, a universal scalar polynomial in $F(\mathbf{u})$; see Proposition 1.1 and Proposition 7.3. Consequently, any inferential statement that depends on the deterministic centering must, at the general level, be formulated in terms of $F^{(p)}$ itself. The explicit power expressions in $F(\mathbf{u})$ reappear only under additional structural assumptions, such as effective total ordering on the support. Second, the role of the strong approximation is not merely to identify a weak limit. Its real value is finer: it gives a uniform almost sure reduction of the stochastic problem to an explicit Gaussian surrogate. For supremum-type and quadratic statistics, this is the relevant level of precision.

3.1 The two-sample problem

Let $\mathbf{X}_1, \dots, \mathbf{X}_m$ and $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ be two independent samples of i.i.d. $\mathbb{R}^d$ -valued random vectors with continuous distribution functions F and G , respectively. Denote by $\mathbb{F}_m^{(p)}$ and $\mathbb{G}_n^{(p)}$ the corresponding p -fold integrated empirical distribution functions. We consider the classical homogeneity hypothesis

$$\mathcal{H}'_0 : F = G$$

against general alternatives.

A natural integrated analogue of the ordinary two-sample empirical process is

$$\xi_{m,n}^{(p)}(\mathbf{u}) := \sqrt{\frac{mn}{m+n}} \left(\mathbb{F}_m^{(p)}(\mathbf{u}) - \mathbb{G}_n^{(p)}(\mathbf{u}) \right), \quad \mathbf{u} \in [0, 1]^d.$$

More generally, following the spirit of power-modified discrepancy statistics, we define, for a fixed integer $q \geq 1$,

$$\xi_{m,n}^{(p,q)}(\mathbf{u}) := \sqrt{\frac{mn}{m+n}} \left[\left(\mathbb{F}_m^{(p)}(\mathbf{u}) \right)^q - \left(\mathbb{G}_n^{(p)}(\mathbf{u}) \right)^q \right], \quad \mathbf{u} \in [0, 1]^d.$$

The case $q = 1$ reduces to the unmodified integrated two-sample process.

For later use, set

$$\varphi(m, n) := \max \left\{ m^{-1/(2(2d-1))} \log m, n^{-1/(2(2d-1))} \log n \right\},$$

and

$$\phi(m, n) := \max \left\{ m^{-1/(2(2d-1))} \sqrt{\log \log m}, n^{-1/(2(2d-1))} \sqrt{\log \log n} \right\}.$$

The basic testing functionals are the integrated Kolmogorov–Smirnov and Cramér–von Mises statistics

$$\begin{aligned} \mathbf{S}_{m,n}^{(p,q)} &:= \sup_{\mathbf{u} \in [0,1]^d} |\xi_{m,n}^{(p,q)}(\mathbf{u})|, \\ \mathbf{T}_{m,n}^{(p,q)} &:= \int_{[0,1]^d} \xi_{m,n}^{(p,q)}(\mathbf{u})^2 dF(\mathbf{u}). \end{aligned} \quad (3.1)$$

Under $\mathcal{H}'_0$, the measure F appearing in (3.1) is of course unknown; in practice it may be replaced by the pooled empirical law, or by any consistent estimator, without changing the first-order asymptotics. We retain the population notation here in order to display the asymptotic structure transparently.

The next result is the two-sample counterpart of Theorem 2.5. It is most naturally stated in terms of the deterministic integrated quantity $F^{(p)}$, thereby preserving full compatibility with the genuinely multivariate setting.

Corollary 3.1 (Two-sample Gaussian approximation). *Assume $\mathcal{H}'_0$, and suppose that $m, n \rightarrow \infty$ with*

$$\lambda_{m,n} := \frac{m}{m+n} \longrightarrow \lambda \in (0, 1).$$

On a suitable probability space, one can construct the collection $\{\xi_{m,n}^{(p,q)} : m, n \geq 1\}$ jointly with two independent sequences of F -Brownian bridges

$$\{\mathbb{B}_{m,F}^{(1)}(\mathbf{u}) : \mathbf{u} \in [0, 1]^d, m \geq 1\}, \quad \{\mathbb{B}_{n,F}^{(2)}(\mathbf{u}) : \mathbf{u} \in [0, 1]^d, n \geq 1\},$$

such that, almost surely as $\min(m, n) \rightarrow \infty$,

$$\sup_{\mathbf{u} \in [0,1]^d} \left| \xi_{m,n}^{(p,q)}(\mathbf{u}) - \mathbb{Z}_{m,n}^{(p,q)}(\mathbf{u}) \right| = O(\varphi(m, n)), \quad (3.2)$$

where

$$\mathbb{Z}_{m,n}^{(p,q)}(\mathbf{u}) := \frac{q}{p!} (F^{(p)}(\mathbf{u}))^{q-1} F(\mathbf{u})^p \left(\sqrt{\frac{n}{m+n}} \mathbb{B}_{m,F}^{(1)}(\mathbf{u}) - \sqrt{\frac{m}{m+n}} \mathbb{B}_{n,F}^{(2)}(\mathbf{u}) \right). \quad (3.3)$$

Sketch of proof. Under $\mathcal{H}'_0$, the two samples are independent and have the same law F . Applying Theorem 2.5 separately to $\mathbb{F}_m^{(p)}$ and $\mathbb{G}_n^{(p)}$ yields the approximation

$$\sqrt{m}(\mathbb{F}_m^{(p)} - F^{(p)}) = \frac{1}{p!} F^p \mathbb{B}_{m,F}^{(1)} + r_m, \quad \sqrt{n}(\mathbb{G}_n^{(p)} - F^{(p)}) = \frac{1}{p!} F^p \mathbb{B}_{n,F}^{(2)} + r_n,$$

uniformly in $\mathbf{u}$, with $\sup_{\mathbf{u}} |r_m(\mathbf{u})| = O(m^{-1/(2(2d-1))} \log m)$ and similarly for r_n . A first-order Taylor expansion of $x \mapsto x^q$ about $x = F^{(p)}(\mathbf{u})$, combined with the fact that the quadratic remainder is of smaller order after scaling by $\sqrt{mn/(m+n)}$, yields (3.2). $\square$

Formula (3.3) is the correct general multivariate statement. It is written in terms of $F^{(p)}$, not solely in terms of F , because outside special order regimes the integrated centering is not reducible to a universal scalar polynomial in $F(\mathbf{u})$. This point is conceptually essential in dimension $d \geq 2$.

Remark 3.2 (Specialization to the total-order/comonotone regime). If the componentwise order is total on the support of F , then

$$F^{(p)}(\mathbf{u}) = \frac{F(\mathbf{u})^{p+1}}{(p+1)!},$$

and (3.3) simplifies to

$$\mathbb{Z}_{m,n}^{(p,q)}(\mathbf{u}) = \frac{(p+1)q}{((p+1)!)^q} F(\mathbf{u})^{q(p+1)-1} \left(\sqrt{\frac{n}{m+n}} \mathbb{B}_{m,F}^{(1)}(\mathbf{u}) - \sqrt{\frac{m}{m+n}} \mathbb{B}_{n,F}^{(2)}(\mathbf{u}) \right).$$

This is precisely the scalarized form that one recovers in the univariate case and, more generally, in effectively one-dimensional order structures.

The power q introduces a deterministic sensitivity profile through the factor $(F^{(p)}(\mathbf{u}))^{q-1}$. For $q > 1$, regions where $F^{(p)}(\mathbf{u})$ is small are downweighted, whereas regions carrying substantial lower-orthant mass are accentuated. From a testing perspective, this may improve sensitivity to central or bulk departures, but it may simultaneously reduce responsiveness to sparse tail perturbations. Thus q is not a neutral algebraic decoration: it is a geometric reweighting of the alternative space.

The preceding coupling immediately yields Gaussian approximations for the corresponding supremum and quadratic functionals.

Corollary 3.3 (Integrated two-sample statistics). *Under $\mathcal{H}'_0$, almost surely as $\min(m, n) \rightarrow \infty$,*

$$\left| \mathbf{S}_{m,n}^{(p,q)} - \sup_{\mathbf{u} \in [0,1]^d} \left| \mathbb{Z}_{m,n}^{(p,q)}(\mathbf{u}) \right| \right| = O(\varphi(m, n)),$$

$$\left| \mathbf{T}_{m,n}^{(p,q)} - \int_{[0,1]^d} \mathbb{Z}_{m,n}^{(p,q)}(\mathbf{u})^2 dF(\mathbf{u}) \right| = O(\phi(m, n)).$$

For implementation, the law of $\sup |\mathbb{Z}_{m,n}^{(p,q)}|$ is typically not available in closed form. The strong approximation nevertheless reduces the problem to simulation of an explicit Gaussian process, which is substantially simpler than working directly with the original integrated empirical field. In practice, multiplier or Gaussian-resampling schemes adapted to the lower-orthant structure are natural tools.

The statistic $\mathbf{S}_{m,n}^{(p,q)}$ is sensitive to highly localized discrepancies, whereas $\mathbf{T}_{m,n}^{(p,q)}$ integrates squared departures and therefore emphasizes global, spatially distributed alternatives. The integrated transform indexed by p tends to smooth the raw empirical discrepancy, so that larger values of p may increase robustness to rough local fluctuations while attenuating sharply oscillatory signals. This trade-off should be kept in mind when comparing power across different orders of integration.

3.2 The change-point problem

We now turn to a sequential application of the integrated empirical process, namely the detection of a structural change in distribution. Throughout, $\lfloor x \rfloor$ denotes the integer part of x . Change-point analysis occupies a central place in sequential inference, time-ordered data analysis, econometrics, reliability, and stochastic monitoring, where the fundamental question is whether the probabilistic mechanism generating the observations remains stable over time. Classical references include [18], [26], and [22]; for more recent developments in related directions, see [10], [6], [21], [35], [2], [11], and [15, 16].

In the present multivariate framework, let $\mathbf{X}_1, \dots, \mathbf{X}_n$ be independent $\mathbb{R}^d$ -valued random vectors. The null hypothesis of distributional stability is

$$\mathcal{H}_0'' : \quad \mathbf{X}_1, \dots, \mathbf{X}_n \text{ have common d.f. } F,$$

where F is continuous on $[0, 1]^d$. As is standard in the literature, we first analyze the behavior of the resulting procedures under the benchmark alternative of a *single abrupt change*:

$$\mathcal{H}_1'' : \quad \exists k^* \in \{1, \dots, n-1\} \text{ such that } \mathbf{X}_1, \dots, \mathbf{X}_{k^*} \text{ have d.f. } F, \quad \mathbf{X}_{k^*+1}, \dots, \mathbf{X}_n \text{ have d.f. } G,$$

where F and G are two continuous distribution functions with $F \neq G$. Writing

$$k^* = \lfloor ns^* \rfloor, \quad s^* \in (0, 1),$$

the unknown fraction s^* represents the relative location of the break. The natural sequential detector in our setting is obtained by comparing, for every putative split point $\lfloor ns \rfloor$, the p -fold integrated empirical distributions computed from the first and the last subsamples. More precisely, for each $n \in \mathbb{N}^*$, define the *sequential integrated empirical contrast*

$$\tilde{\alpha}_n^{(p)}(s, \mathbf{u}) := \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{n^{3/2}} \left(\mathbb{F}_{\lfloor ns \rfloor}^{(p)-}(\mathbf{u}) - \mathbb{F}_{n-\lfloor ns \rfloor}^{(p)+}(\mathbf{u}) \right), \quad s \in (0, 1), \mathbf{u} \in [0, 1]^d, \quad (3.4)$$

where $\mathbb{F}_k^{(p)-}$ denotes the p -fold integrated empirical d.f. based on $\mathbf{X}_1, \dots, \mathbf{X}_k$, while $\mathbb{F}_{n-k}^{(p)+}$ denotes the corresponding object based on $\mathbf{X}_{k+1}, \dots, \mathbf{X}_n$. As usual, we adopt the convention

$$\mathbb{F}_0^{(p)-} = \mathbb{F}_0^{(p)+} = 0,$$

so that (3.4) is well defined for all $s \in (0, 1)$ and, in particular, $\tilde{\alpha}_n^{(p)}(s, \mathbf{u}) = 0$ whenever $\lfloor ns \rfloor = 0$. The normalization

$$\lambda_n(s) := \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{n^{3/2}}$$

is the familiar CUSUM scaling: under the null, it balances the variability of the difference between the two subsample processes, whereas under a fixed alternative it produces a deterministic drift of maximal order near the true change-location.

The process $\tilde{\alpha}_n^{(p)}$ should be viewed as the integrated analogue of the classical sequential empirical bridge. The index s plays the role of “time” or putative split location, while $\mathbf{u}$ retains the lower-orthant geometric structure of the empirical process. Thus the detector is genuinely biparametric: it searches simultaneously over the temporal coordinate s and the spatial lower-orthant coordinate $\mathbf{u}$.

Under $\mathcal{H}_0''$, the first and second subsamples are identically distributed, and the strong approximation results of Section 2 suggest that $\tilde{\alpha}_n^{(p)}$ should admit a Kiefer-type Gaussian approximation. To formulate the limiting object, one may construct, on a common probability space, two independent Kiefer processes

$$\{\mathbb{K}_1(t, \mathbf{u}) : t \geq 0, \mathbf{u} \in [0, 1]^d\} \quad \text{and} \quad \{\mathbb{K}_2(t, \mathbf{u}) : t \geq 0, \mathbf{u} \in [0, 1]^d\},$$

corresponding respectively to the two data segments located before and after the candidate split point. The associated Gaussian approximation of $\tilde{\alpha}_n^{(p)}$ is then the sequence

$$\{\mathbb{K}_n^{(p)}(s, \mathbf{u}) : s \in [0, 1], \mathbf{u} \in [0, 1]^d, n \in \mathbb{N}^*\},$$

defined by

$$\mathbb{K}_n^{(p)}(s, \mathbf{u}) = \frac{1}{p!} F(\mathbf{u})^p \overset{\circ}{\mathbb{K}}_n(s, \mathbf{u}), \quad s \in [0, 1], \mathbf{u} \in [0, 1]^d, \quad (3.5)$$

where $\overset{\circ}{\mathbb{K}}_n$ is the stitched Gaussian bridge

$$\overset{\circ}{\mathbb{K}}_n(s, \mathbf{u}) := \begin{cases} \frac{1}{\sqrt{n}} \left[\mathbb{K}_2(\lfloor ns \rfloor, \mathbf{u}) - s(\mathbb{K}_1(\lfloor n/2 \rfloor, \mathbf{u}) + \mathbb{K}_2(\lfloor n/2 \rfloor, \mathbf{u})) \right], & s \in \left[0, \frac{1}{2}\right], \\ \frac{1}{\sqrt{n}} \left[-\mathbb{K}_1(\lfloor n(1-s) \rfloor, \mathbf{u}) + (1-s)(\mathbb{K}_1(\lfloor n/2 \rfloor, \mathbf{u}) + \mathbb{K}_2(\lfloor n/2 \rfloor, \mathbf{u})) \right], & s \in \left[\frac{1}{2}, 1\right]. \end{cases}$$

This representation is the sequential counterpart of the Kiefer approximation from Theorem 2.8, with the deterministic factor $F(\mathbf{u})^p/p!$ inherited from the integrated empirical structure.

The appearance of the factor $F(\mathbf{u})^p/p!$ is not accidental. It reflects the first-order linearization of the p -fold integrated empirical functional around F . In particular, from the viewpoint of local asymptotic theory, the sequential detector inherits the same leading Gaussian modulation as the non-sequential integrated process. This is the precise mechanism through which strong approximation transfers from the primitive empirical process to its change-point version.

The following theorem gives the required uniform approximation.

Theorem 3.4. *On a suitable probability space, it is possible to define $\{\tilde{\alpha}_n^{(p)} : n \in \mathbb{N}^*\}$ jointly with a sequence of Gaussian processes $\{\mathbb{K}_n^{(p)} : n \in \mathbb{N}^*\}$ as above such that, under $\mathcal{H}_0''$, as $n \rightarrow \infty$,*

$$\sup_{s \in (0,1)} \sup_{\mathbf{u} \in [0,1]^d} \left| \tilde{\alpha}_n^{(p)}(s, \mathbf{u}) - \mathbb{K}_n^{(p)}(s, \mathbf{u}) \right| = O\left(n^{-1/(4d)} (\log n)^{3/2}\right).$$

The theorem shows that the sequential integrated empirical detector remains within the classical strong-invariance framework. In particular, sup-functionals and weighted sup-functionals of $\tilde{\alpha}_n^{(p)}$ may be asymptotically calibrated through the corresponding Gaussian limits.

A natural omnibus statistic for testing $\mathcal{H}_0''$ is the integrated Kolmogorov–Smirnov type functional

$$\sigma_n^{(p)} := \sup_{s \in (0,1)} \sup_{\mathbf{u} \in [0,1]^d} \left| \tilde{\alpha}_n^{(p)}(s, \mathbf{u}) \right|. \quad (3.6)$$

Large values of $\sigma_n^{(p)}$ indicate evidence against the null of distributional stability.

The statistic (3.6) is particularly sensitive to alternatives for which the distributional discrepancy accumulates over a substantial portion of the sample, especially when the break occurs away from the endpoints. This is a generic feature of unweighted CUSUM-type procedures: they are most responsive in the central part of the sample, where the variance-stabilizing factor $s(1-s)$ is largest.

The next result is an immediate consequence of Theorem 3.4.

Corollary 3.5. *If $\mathcal{H}_0''$ holds, then, as $n \rightarrow \infty$,*

$$\sigma_n^{(p)} \xrightarrow{\mathcal{L}} \sup_{(s, \mathbf{u}) \in [0, 1]^{d+1}} |\mathbb{K}^{(p)}(s, \mathbf{u})|,$$

where $\mathbb{K}^{(p)} = \{\mathbb{K}^{(p)}(s, \mathbf{u}) : (s, \mathbf{u}) \in [0, 1]^{d+1}\}$ is a centered Gaussian process with covariance function

$$\mathbb{E}[\mathbb{K}^{(p)}(s, \mathbf{u}) \mathbb{K}^{(p)}(s', \mathbf{u}')] = \frac{1}{p!^2} F(\mathbf{u})^p F(\mathbf{u}')^p (F(\mathbf{u} \wedge \mathbf{u}') - F(\mathbf{u})F(\mathbf{u}'))(s \wedge s' - ss').$$

Remark 3.6. The covariance structure in Corollary 3.5 factorizes into two interpretable components: the term

$$F(\mathbf{u} \wedge \mathbf{u}') - F(\mathbf{u})F(\mathbf{u}')$$

is the covariance kernel of the F -Brownian bridge indexed by lower orthants, whereas

$$s \wedge s' - ss'$$

is the covariance kernel of the standard Brownian bridge on $[0, 1]$. Thus the sequential limit is a tensor-type bridge in the variables $(s, \mathbf{u})$, modulated by the integrated factor $F(\mathbf{u})^p/p!$.

Although (3.6) is canonical, it is well known that unweighted supremum statistics may lose sensitivity for changes occurring near the beginning or the end of the sample. This motivates the weighted version

$$\sigma_{n,w}^{(p)} := \sup_{s \in (0, 1)} \sup_{\mathbf{u} \in [0, 1]^d} \frac{|\tilde{\alpha}_n^{(p)}(s, \mathbf{u})|}{w(\lfloor ns \rfloor / n)}, \quad (3.7)$$

where $w : (0, 1) \rightarrow (0, \infty)$ is positive, increasing near 0, decreasing near 1, and satisfies the integrability condition

$$I(w, \varepsilon) := \int_0^1 \exp\left(-\frac{\varepsilon w(s)^2}{s(1-s)}\right) \frac{ds}{s(1-s)} < \infty$$

for some $\varepsilon > 0$. For the role of this condition in weighted empirical-process theory, see Chapter 4 of [25]. A prototypical admissible choice, due to [45], is

$$w(s) = \left(s(1-s) \log \log \frac{1}{s(1-s)} \right)^{1/2}, \quad s \in (0, 1).$$

Under such conditions, and by arguments entirely parallel to those developed in [26], one obtains

$$\sigma_{n,w}^{(p)} \xrightarrow{\mathcal{L}} \sup_{(s, \mathbf{u}) \in [0, 1]^{d+1}} \frac{|\mathbb{K}^{(p)}(s, \mathbf{u})|}{w(s)}.$$

For details in the present integrated framework, we refer to [2].

The weighted statistic (3.7) compensates for the vanishing factor $s(1-s)$ near the boundaries and therefore improves the detectability of early or late changes. In that sense, the weighting device is not merely technical: it changes the geometry of the testing problem by rebalancing the contribution of putative split points across the sample. From a practical viewpoint, the limiting laws in both the weighted and unweighted cases are generally non-pivotal because they depend on the unknown d.f. F . Consequently, implementation typically requires either a suitable resampling scheme, a multiplier approximation, or a transformation to a setting where the limit becomes distribution-free. In copula-based or rank-based versions of the problem, such a reduction is often possible and may lead to more stable finite-sample calibration. The single-change alternative $\mathcal{H}_1''$ serves here as the canonical local model. For multiple change-points, epidemic alternatives, or smooth distributional drift, the same sequential integrated empirical machinery should remain relevant, but the appropriate normalization, localization properties, and argmax analysis become substantially more delicate. These extensions are mathematically natural and, in our view, deserve separate treatment.

4 Strong approximation of the integrated empirical process when parameters are estimated

In this section, we investigate the effect of parameter estimation on the p -fold integrated empirical process. The problem is statistically indispensable, since in virtually all composite goodness-of-fit settings the centering law is not known and must be replaced by a data-driven estimator. Our goal is to show that, under the usual asymptotic linearity and smoothness assumptions, the plug-in integrated empirical process still admits a Gaussian approximation of the same general nature as in the simple-null case.

A structural remark is essential from the outset. In the genuinely multivariate lower-orthant framework, the relevant deterministic object is the integrated distribution functional $F^{(p)}(\theta, \mathbf{u})$, not merely the scalar quantity $F(\theta, \mathbf{u})$. Outside special total-order regimes, $F^{(p)}(\theta, \mathbf{u})$ is not a universal polynomial transform of $F(\theta, \mathbf{u})$, and therefore the first-order plug-in analysis must naturally be formulated at the level of $F^{(p)}$ itself. This distinction is not cosmetic: it is precisely where the general multivariate theory departs from the univariate scalar calculus.

For each $n \in \mathbb{N}^*$, define the p -fold integrated empirical process with estimated parameters by

$$\widehat{\alpha}_n^{(p)}(\mathbf{u}) := \sqrt{n} \left(\mathbb{F}_n^{(p)}(\mathbf{u}) - F^{(p)}(\widehat{\theta}_n, \mathbf{u}) \right), \quad \mathbf{u} \in [0, 1]^d,$$

where $\{\widehat{\theta}_n : n \in \mathbb{N}^*\}$ is a sequence of estimators of the true parameter value $\theta_0 \in \Theta \subset \mathbb{R}^q$, and where

$$\{F(\theta, \mathbf{u}) : \theta \in \Theta, \mathbf{u} \in [0, 1]^d\}$$

is a parametric family of continuous distribution functions. The starting point for the ordinary empirical process with estimated parameters goes back to [30]; strong approximation refinements in this direction were developed by Burke and Csörgő [20]. Our purpose here is to extend that philosophy to the integrated framework.

We now fix some notation.

(4.1) For a vector $V \in \mathbb{R}^q$, $V^\top$ denotes its transpose.

(4.2) On $\mathbb{R}^q$, we use the sup-norm

$$\|(y_1, \dots, y_q)\| := \max_{1 \leq i \leq q} |y_i|.$$

(4.3) If $g(\boldsymbol{\theta}, \mathbf{u})$ is differentiable with respect to $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)$, we write

$$\nabla_{\boldsymbol{\theta}} g(\boldsymbol{\theta}_0, \mathbf{u}) = \left(\frac{\partial}{\partial \theta_1} g(\boldsymbol{\theta}, \mathbf{u}), \dots, \frac{\partial}{\partial \theta_q} g(\boldsymbol{\theta}, \mathbf{u}) \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

for the gradient at $\boldsymbol{\theta}_0$, and $\nabla_{\boldsymbol{\theta}}^2 g(\boldsymbol{\theta}, \mathbf{u})$ for the corresponding Hessian matrix.

(4.4) For a vector-valued function $V = (v_1, \dots, v_q)$, Stieltjes integration is understood componentwise:

$$\int V := \left(\int v_1, \dots, \int v_q \right).$$

We next collect the standing assumptions.

(i) **Asymptotic linearity of the estimator.** For each $n \in \mathbb{N}^*$,

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \ell(\boldsymbol{\theta}_0, \mathbf{X}_i) + \boldsymbol{\varepsilon}_n,$$

where $\ell(\boldsymbol{\theta}_0, \cdot)$ is a measurable $\mathbb{R}^q$ -valued function and $\boldsymbol{\varepsilon}_n \rightarrow 0$ in probability.

(ii) **Centering.**

$$\mathbb{E}[\ell(\boldsymbol{\theta}_0, \mathbf{X}_1)] = \mathbf{0}.$$

(iii) **Finite covariance.** The matrix

$$M(\boldsymbol{\theta}_0) := \mathbb{E}[\ell(\boldsymbol{\theta}_0, \mathbf{X}_1) \ell(\boldsymbol{\theta}_0, \mathbf{X}_1)^\top]$$

is finite and nonnegative definite of size $q \times q$.

(iv) **Smoothness of the integrated functional.** For each $\mathbf{u} \in [0, 1]^d$, the map

$$\boldsymbol{\theta} \mapsto F^{(p)}(\boldsymbol{\theta}, \mathbf{u})$$

is differentiable at $\boldsymbol{\theta}_0$, and the gradient $\nabla_{\boldsymbol{\theta}} F^{(p)}(\boldsymbol{\theta}, \mathbf{u})$ is uniformly continuous in $(\boldsymbol{\theta}, \mathbf{u}) \in \Lambda \times [0, 1]^d$, where Λ is the closure of a neighborhood of $\boldsymbol{\theta}_0$.

(v) **Bounded variation of the influence function.** Each component of $\ell(\boldsymbol{\theta}_0, \mathbf{u})$ has bounded variation in $\mathbf{u}$ on rectangles of $[0, 1]^d$.

Under these assumptions, the ordinary plug-in empirical process admits the usual Gaussian approximation. To formulate the corresponding integrated version, define

$$\mathcal{H}^p(n, \mathbf{u}) := \mathbb{G}_n^{(p)}(\mathbf{u}) - \nabla_{\boldsymbol{\theta}} F^{(p)}(\boldsymbol{\theta}_0, \mathbf{u})^\top \sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0), \quad \mathbf{u} \in [0, 1]^d, \quad (4.1)$$

where $\mathbb{G}_n^{(p)}$ denotes a Gaussian approximation of the centered integrated empirical process

$$\alpha_n^{(p)}(\mathbf{u}) = \sqrt{n} \left(\mathbb{F}_n^{(p)}(\mathbf{u}) - F^{(p)}(\boldsymbol{\theta}_0, \mathbf{u}) \right).$$

In particular, whenever the factorized first-order approximation from Section 2 is valid in the model under consideration, one may take

$$\mathbb{G}_n^{(p)}(\mathbf{u}) = \frac{1}{p!} F(\boldsymbol{\theta}_0, \mathbf{u})^p \mathcal{G}(n, \mathbf{u}),$$

with $\mathcal{G}(n, \mathbf{u})$ the Gaussian surrogate arising in the ordinary parametric empirical-process approximation. Accordingly, define the approximation error

$$\boldsymbol{\varepsilon}_n^{(p)} := \sup_{\mathbf{u} \in [0, 1]^d} \left| \widehat{\alpha}_n^{(p)}(\mathbf{u}) - \mathcal{H}^p(n, \mathbf{u}) \right|.$$

We may now state the parametric plug-in approximation theorem.

Theorem 4.1. *Suppose that assumptions (i), (ii), and (iii) hold. Then*

$$\varepsilon_n^{(p)} \xrightarrow{\mathbb{P}} 0,$$

provided that (iv) and (v) hold and the corresponding ordinary plug-in empirical process admits a valid first-order Gaussian approximation.

Theorem 4.1 says that parameter estimation perturbs the integrated empirical process only through a first-order score-type correction. Thus, despite the nonlinear and order-sensitive nature of the p -fold integration operator, the plug-in effect remains asymptotically linear. This is the decisive fact that makes composite integrated procedures tractable.

Remark 4.2 (Special total-order regime). If the model lies in a total-order or comonotone regime, so that

$$F^{(p)}(\theta, \mathbf{u}) = \frac{F(\theta, \mathbf{u})^{p+1}}{(p+1)!},$$

then

$$\nabla_{\theta} F^{(p)}(\theta_0, \mathbf{u}) = \frac{1}{p!} F(\theta_0, \mathbf{u})^p \nabla_{\theta} F(\theta_0, \mathbf{u}),$$

and the approximation in (4.1) reduces to the familiar scalarized form. This simplification is elegant, but it should be viewed as a structural corollary, not as the generic multivariate statement.

The limiting Gaussian process $\mathcal{H}^P(n, \cdot)$ depends fundamentally on the unknown law $F(\theta_0, \cdot)$ and on the true parameter value θ_0 . Consequently, Theorem 4.1 is not directly implementable for testing the composite hypothesis

$$F \in \{F(\theta, \mathbf{u}) : \mathbf{u} \in [0, 1]^d, \theta \in \Theta\}.$$

A natural way to overcome this difficulty is to replace the unknown parameter by its estimator inside the Gaussian approximation itself. Thus, define the plug-in Gaussian surrogate

$$\widehat{\mathcal{H}}^P(n, \mathbf{u}) := \mathbb{G}_{n, \hat{\theta}_n}^{(p)}(\mathbf{u}) - \nabla_{\theta} F^{(p)}(\hat{\theta}_n, \mathbf{u})^{\top} \sqrt{n}(\hat{\theta}_n - \theta_0), \quad \mathbf{u} \in [0, 1]^d,$$

where $\mathbb{G}_{n, \hat{\theta}_n}^{(p)}$ denotes the same Gaussian construction as in (4.1), but evaluated at the plug-in parameter $\hat{\theta}_n$. Under the regularity conditions ensuring stochastic equicontinuity in the parameter, one obtains

$$\sup_{\mathbf{u} \in [0, 1]^d} \left| \widehat{\mathcal{H}}^P(n, \mathbf{u}) - \mathcal{H}^P(n, \mathbf{u}) \right| \xrightarrow{\mathbb{P}} 0.$$

It follows that

$$\sup_{\mathbf{u} \in [0, 1]^d} \left| \widehat{\alpha}_n^{(p)}(\mathbf{u}) - \widehat{\mathcal{H}}^P(n, \mathbf{u}) \right| \xrightarrow{\mathbb{P}} 0. \quad (4.2)$$

Relation (4.2) is the practically relevant statement: it shows that the plug-in integrated empirical process may be calibrated by a fully data-driven Gaussian approximation. Consequently, Kolmogorov–Smirnov and Cramér–von Mises type statistics built from $\widehat{\alpha}_n^{(p)}$ inherit asymptotically valid composite-null calibrations. From an inferential viewpoint, the message is clear. Under the standard regularity assumptions, one may treat the integrated empirical process with estimated parameters exactly as one treats the ordinary parametric empirical process: the unknown centering distribution modifies the limiting Gaussian law, but does not destroy Gaussian approximability. In practice, this opens the way to parametric bootstrap, multiplier calibration, and Gaussian simulation strategies for composite integrated goodness-of-fit procedures.

5 Simulation study

This section investigates the finite-sample performance of the integrated empirical procedures developed in Sections 2–4. The numerical study serves two complementary purposes. First, it evaluates the practical relevance of the Gaussian approximations derived for the sequential integrated empirical process in the change-point problem. Second, it examines the numerical accuracy of the first-order linearization from Theorem 4.1 in the estimated-parameter setting and compares several feasible calibration strategies for composite testing. All experiments are carried out in dimension $d = 2$. This is the smallest genuinely multivariate framework in which the lower-orthant order is no longer total, so that the simulations retain the essential geometric content of the theory while remaining computationally tractable. We consider integration orders

$$p \in \{0, 1, 2, 3\},$$

where $p = 0$ corresponds to the ordinary empirical-process benchmark. Sample sizes

$$n \in \{100, 200, 500, 1000\}$$

were explored; for concision, the detailed tables below report the representative case $n = 500$, while the overall trends across n are summarized in the discussion. Suprema over $[0, 1]^2$ are approximated on a regular interior grid with 39 points per coordinate, yielding 1521 effective evaluation sites. In preliminary runs, finer grids did not materially alter the reported rankings.

A methodological point deserves explicit mention. Whenever critical values are estimated by Monte Carlo, the calibration sample and the evaluation sample are generated independently. This avoids any artificial reduction in empirical error and allows the reported rejection frequencies to be interpreted as honest finite-sample size and power approximations.

The integration order p is not merely a formal parameter. Although larger values of p smooth the raw empirical fluctuations, the leading Gaussian term is also weighted by $F(\mathbf{u})^p/p!$, which attenuates the effective signal away from the upper corner of the state space. Thus, from a statistical viewpoint, higher-order integration may stabilize some functionals while simultaneously reducing sensitivity to abrupt alternatives.

5.1 Change-point analysis

5.1.1 Experimental design

Let $\mathbf{X}_1, \dots, \mathbf{X}_n$ be a bivariate sample. We test the null hypothesis of distributional homogeneity,

$$\mathcal{H}_0 : \mathbf{X}_1, \dots, \mathbf{X}_n \text{ have common d.f. } F,$$

against the single-break alternative

$$\mathcal{H}_1 : \exists k^* = \lfloor ns^* \rfloor, s^* \in (0, 1), \text{ such that } \mathbf{X}_1, \dots, \mathbf{X}_{k^*} \sim F, \quad \mathbf{X}_{k^*+1}, \dots, \mathbf{X}_n \sim G.$$

The statistics are based on the sequential integrated empirical process $\tilde{\alpha}_n^{(p)}(s, \mathbf{u})$ defined in (3.4). We consider

$$\begin{aligned} \sigma_n^{(p)} &= \sup_{s \in (0, 1)} \sup_{\mathbf{u} \in [0, 1]^2} |\tilde{\alpha}_n^{(p)}(s, \mathbf{u})|, \\ \sigma_{n,w}^{(p)} &= \sup_{s \in (0, 1)} \sup_{\mathbf{u} \in [0, 1]^2} \frac{|\tilde{\alpha}_n^{(p)}(s, \mathbf{u})|}{w(s)}, \quad w(s) = \left(s(1-s) \log \log \frac{1}{s(1-s)} \right)^{1/2}. \end{aligned}$$

The second statistic is the classical endpoint-sensitive weighted version. For each null model and each sample size, critical values are estimated from an independent Monte Carlo calibration sample of size $B_{\text{cal}} = 3000$. Empirical size and power are then evaluated on a second independent sample of size $B_{\text{eval}} = 3000$. At nominal level $\alpha = 0.05$, the Monte Carlo standard error of an empirical rejection frequency is therefore

$$\sqrt{\frac{0.05(0.95)}{3000}} \approx 0.004,$$

so departures from 0.05 of order one percentage point should not be over-interpreted.

The null reference models are:

- (i) **Null_indep**: independent uniform margins;
- (ii) **Null_gauss05**: Gaussian copula with parameter $\rho = 0.5$ and uniform margins;
- (iii) **Null_clayton1**: Clayton copula with parameter $\theta = 1$ and uniform margins.

Under the alternative, we examine:

- (iv) **Dep_shift**: a pure dependence break, from independence to a Gaussian copula with $\rho = 0.7$;
- (v) **Marginal_shift**: a pure marginal break, from uniform margins to Beta(2, 1) margins under independence;
- (vi) **Joint_shift**: a simultaneous break in dependence and margins, from independence with uniform margins to a Gaussian copula with $\rho = 0.6$ and Beta(2, 1) margins.

The break fraction is varied over

$$s^* \in \{0.10, 0.25, 0.50, 0.75, 0.90\}.$$

5.1.2 Null calibration

We begin with the null behavior of the procedures. The goal here is not only to verify approximate level control, but also to check whether the maximization over the putative break fraction produces any systematic localization artifact.

Table 1 reports the empirical size at $n = 500$ under the three null scenarios. The rejection frequencies lie between 0.044 and 0.063, which is fully compatible with nominal level 0.05 at the present Monte Carlo resolution. This remains true across dependence structures, integration orders, and both weighted and unweighted versions. The average estimated break fraction $\hat{k}/n$ stays very close to 1/2 throughout, as expected in the absence of a true structural break.

These findings indicate that the approximation furnished by Theorem 3.4 is already informative at moderate sample sizes, at least from the viewpoint of null calibration.

5.1.3 Power and localization

We next turn to the alternative scenarios. To keep the presentation compact while preserving the localization information, Table 2 reports each configuration in the form

$$\text{power} / \mathbb{E}(|\hat{k}/n - s^*|).$$

The results show that the order of integration has a pronounced effect on finite-sample efficiency. For *pure dependence shifts*, the classical empirical-process statistic $p = 0$ is uniformly the most powerful procedure. This dominance is especially marked for interior breaks: at $s^* = 0.50$, the unweighted power is 0.617 for $p = 0$, compared with 0.279, 0.200, and 0.145 for $p = 1, 2, 3$. The

Table 1: Empirical size of the bivariate change-point statistics at $n = 500$ under the null hypothesis.

Scenario	Version	p	Size	Mean $\widehat{k}/n$
Panel A: Clayton copula ($\theta = 1$)				
	Unweighted	0	0.049	0.496
		1	0.058	0.496
		2	0.052	0.499
		3	0.050	0.497
	Weighted	0	0.053	0.498
		1	0.058	0.497
		2	0.052	0.495
		3	0.052	0.495
Panel B: Gaussian copula ($\rho = 0.5$)				
	Unweighted	0	0.057	0.503
		1	0.044	0.501
		2	0.046	0.503
		3	0.056	0.496
	Weighted	0	0.058	0.498
		1	0.049	0.495
		2	0.049	0.483
		3	0.049	0.483
Panel C: Independence				
	Unweighted	0	0.051	0.498
		1	0.054	0.497
		2	0.063	0.502
		3	0.059	0.497
	Weighted	0	0.048	0.497
		1	0.053	0.494
		2	0.049	0.497
		3	0.049	0.498

Notes. Nominal level $\alpha = 0.05$. Critical values are computed from an independent calibration sample with $B_{\text{cal}} = 3000$ replications, and empirical size is evaluated on an independent sample with $B_{\text{eval}} = 3000$ replications. The Monte Carlo standard error at level 0.05 is approximately 0.004.

same monotone deterioration is visible at $s^* = 0.25$ and $s^* = 0.75$, showing that in this regime higher-order integration primarily attenuates the available signal.

For *marginal shifts*, low-order integration is less harmful, but the dominant procedure remains $p = 0$ in most configurations. The unweighted statistics achieve essentially unit power for central breaks when $p = 0$ and $p = 1$, while the weighted version substantially enhances detection near the boundaries for $p = 0$. However, once $p \geq 2$, the weighted procedures deteriorate sharply and in several cases become scarcely distinguishable from size.

The *joint-shift* setting leads to the same global conclusion. The statistic with $p = 0$ remains uniformly strongest, while $p = 1$ may still be acceptable for interior breaks. Higher-order integration is again clearly unfavorable in finite samples for abrupt structural changes.

The localization behavior is coherent with these power findings. Break estimation is naturally most accurate for interior breaks. Whenever the underlying test retains substantial discriminatory power, weighting improves localization near $s^* = 0.10$ and $s^* = 0.90$, especially for $p = 0$. Once the corresponding power collapses, however, the estimated break location becomes largely noise-driven and should no longer be over-interpreted.

Remark 5.1. From a practical standpoint, the simulations strongly suggest that the cases $p \geq 2$ should not be viewed as routine competitors to the classical sequential empirical process for change-point detection. Their asymptotic validity is clear, but their finite-sample sensitivity under abrupt alternatives is often substantially weaker.

Table 2: Empirical performance of the bivariate change-point statistics at $n = 500$. Each entry is reported as *power / mean absolute break-location error*.

Alternative	Version	p	$s^* = 0.10$	$s^* = 0.25$	$s^* = 0.50$	$s^* = 0.75$	$s^* = 0.90$
Panel A: Dependence shift							
	Unweighted	0	0.073 / 0.353	0.326 / 0.143	0.617 / 0.056	0.321 / 0.143	0.077 / 0.362
		1	0.055 / 0.394	0.124 / 0.214	0.279 / 0.091	0.162 / 0.183	0.061 / 0.372
		2	0.059 / 0.409	0.101 / 0.254	0.200 / 0.119	0.138 / 0.198	0.071 / 0.370
		3	0.056 / 0.440	0.081 / 0.368	0.145 / 0.240	0.117 / 0.298	0.066 / 0.418
	Weighted	0	0.117 / 0.272	0.362 / 0.142	0.550 / 0.096	0.371 / 0.148	0.147 / 0.286
		1	0.048 / 0.425	0.114 / 0.301	0.219 / 0.215	0.186 / 0.247	0.112 / 0.344
		2	0.008 / 0.463	0.011 / 0.484	0.021 / 0.497	0.040 / 0.514	0.054 / 0.519
		3	0.008 / 0.464	0.011 / 0.484	0.020 / 0.497	0.040 / 0.514	0.054 / 0.520
	Panel B: Marginal shift						
	Unweighted	0	0.379 / 0.203	0.999 / 0.042	1.000 / 0.016	1.000 / 0.055	0.430 / 0.221
		1	0.201 / 0.201	0.947 / 0.044	1.000 / 0.027	0.949 / 0.101	0.233 / 0.306
		2	0.170 / 0.194	0.845 / 0.049	0.981 / 0.037	0.795 / 0.137	0.179 / 0.370
		3	0.150 / 0.208	0.710 / 0.058	0.914 / 0.049	0.590 / 0.184	0.125 / 0.506
	Weighted	0	0.860 / 0.044	1.000 / 0.022	1.000 / 0.019	1.000 / 0.027	0.869 / 0.056
		1	0.633 / 0.069	0.962 / 0.034	0.997 / 0.036	0.947 / 0.069	0.327 / 0.223
		2	0.197 / 0.311	0.151 / 0.369	0.133 / 0.458	0.029 / 0.621	0.024 / 0.682
		3	0.195 / 0.312	0.110 / 0.385	0.059 / 0.496	0.023 / 0.624	0.024 / 0.683
	Panel C: Joint shift						
	Unweighted	0	0.215 / 0.215	0.990 / 0.044	1.000 / 0.019	0.991 / 0.060	0.266 / 0.256
		1	0.117 / 0.231	0.829 / 0.053	0.984 / 0.033	0.797 / 0.115	0.142 / 0.333
		2	0.120 / 0.234	0.648 / 0.065	0.905 / 0.046	0.573 / 0.165	0.116 / 0.387
		3	0.111 / 0.239	0.498 / 0.088	0.745 / 0.066	0.380 / 0.246	0.092 / 0.522
	Weighted	0	0.664 / 0.055	0.995 / 0.024	1.000 / 0.023	0.996 / 0.032	0.630 / 0.091
		1	0.432 / 0.098	0.862 / 0.048	0.959 / 0.049	0.789 / 0.090	0.181 / 0.312
		2	0.128 / 0.301	0.097 / 0.385	0.076 / 0.486	0.035 / 0.614	0.025 / 0.675
		3	0.127 / 0.302	0.088 / 0.389	0.055 / 0.497	0.034 / 0.615	0.025 / 0.675

Notes. The reported break-location error is $\mathbb{E}(|\hat{k}/n - s^*|)$. Critical values are inherited from the matched null reference model. The weighted statistic redistributes sensitivity toward the boundaries, but the gain is effectively concentrated in the low-order cases $p = 0$ and, to a lesser extent, $p = 1$.

The conclusions of Table 2 are visually summarized in Figure 1. The top panel confirms satisfactory null calibration, the middle panel illustrates the contrast between interior-oriented unweighted detection and boundary-oriented weighted detection, and the bottom panel shows that localization accuracy mirrors the same pattern.

Bivariate change-point study: classical empirical process ($p = 0$) versus integrated processes ($p = 1, 2, 3$)

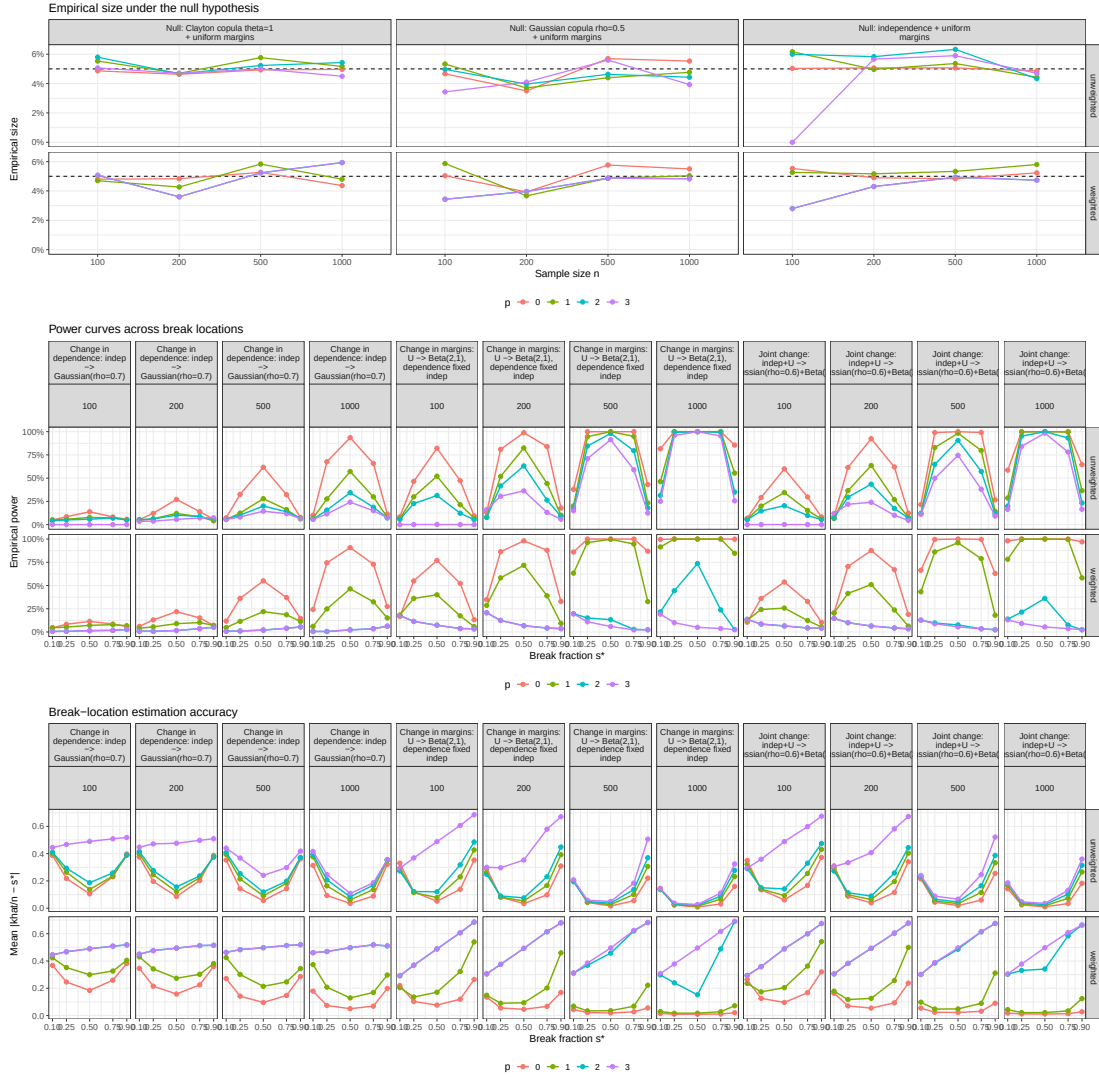


Figure 1: Bivariate change-point study. Top: empirical size under the null hypothesis. Middle: empirical power as a function of the break fraction s^* . Bottom: mean absolute break-location error $\mathbb{E}(|\hat{k}/n - s^*|)$. The weighted procedures improve sensitivity to boundary breaks mainly for low integration orders.

5.2 Estimated-parameter study

5.2.1 Experimental design

We next consider the composite parametric framework. Let

$$\mathbf{X}_1, \dots, \mathbf{X}_n \stackrel{\text{i.i.d.}}{\sim} F(\theta_0, \cdot),$$

where $F(\boldsymbol{\theta}, \cdot)$ belongs to a parametric family of bivariate distributions with continuous marginals. The object of interest is the plug-in integrated empirical process

$$\widehat{\alpha}_n^{(p)}(\mathbf{u}) = \sqrt{n} \left(\mathbb{F}_n^{(p)}(\mathbf{u}) - F^{(p)}(\widehat{\boldsymbol{\theta}}_n, \mathbf{u}) \right), \quad \mathbf{u} \in [0, 1]^2,$$

together with its first-order Gaussian approximation $\mathcal{H}^P(n, \mathbf{u})$ from Theorem 4.1. To quantify the quality of this approximation, we consider the remainder

$$R_n^{(p)} = \sup_{\mathbf{u} \in [0, 1]^2} |\widehat{\alpha}_n^{(p)}(\mathbf{u}) - \mathcal{H}^P(n, \mathbf{u})|.$$

Two copula families are examined:

- **Gaussian copula:**

$$C_\rho(u_1, u_2) = \Phi_\rho(\Phi^{-1}(u_1), \Phi^{-1}(u_2)), \quad \widehat{\rho} = \sin\left(\frac{\pi \widehat{\tau}}{2}\right),$$

where $\widehat{\tau}$ denotes Kendall's sample tau;

- **Clayton copula:**

$$C_\theta(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}, \quad \widehat{\theta} = \frac{2\widehat{\tau}}{1 - \widehat{\tau}} \quad (\widehat{\tau} \geq 0).$$

For composite-null testing, we compare three calibration rules:

- (a) **Oracle:** critical values computed under the true parameter $\boldsymbol{\theta}_0$;
- (b) **Naive plug-in:** critical values obtained by treating $\widehat{\boldsymbol{\theta}}_n$ as fixed and known;
- (c) **Parametric bootstrap:** $B_{\text{boot}} = 199$ bootstrap resamples generated from $F(\widehat{\boldsymbol{\theta}}_n, \cdot)$, with re-estimation performed in each resample.

The distribution of $R_n^{(p)}$ is approximated from $B_{\text{dist}} = 1500$ replications. Empirical size and power are estimated from $B_{\text{eval}} = 300$ replications. At nominal level 0.05, the associated Monte Carlo standard error is about 0.013, so small differences across feasible calibration methods should be interpreted with appropriate caution.

5.2.2 Accuracy of the first-order approximation

We first assess the numerical quality of the approximation from Theorem 4.1. Table 3 reports the empirical distribution of the supremum-norm remainder $R_n^{(p)}$ at $n = 500$.

The message is clear: the first-order Gaussian approximation is already quite accurate at this sample size, and the approximation error decreases markedly as the order of integration increases. For $p = 0$, the mean remainder is of order 10^{-3} to 10^{-2} , depending on the family and the strength of dependence. For $p = 1$, it is smaller still, and for $p \geq 2$ it becomes essentially negligible on the operational scale of the test statistics.

This pattern is fully consistent with the analytic structure of the plug-in expansion. The integrated transform dampens the effect of estimation noise, while stronger dependence slightly magnifies the remainder through the gradient term $\nabla_{\boldsymbol{\theta}} F(\boldsymbol{\theta}_0, \mathbf{u})$.

Table 3: Supremum-norm remainder $R_n^{(p)}$ at $n = 500$ in the estimated-parameter study.

Family	Parameter	p	Mean $R_n^{(p)}$	Median $R_n^{(p)}$	Q95 $R_n^{(p)}$
Panel A: Clayton family					
$\theta_0 = 0.5$		0	0.006	0.003	0.024
		1	0.001	0.001	0.004
		2	0.000	0.000	0.001
		3	0.000	0.000	0.000
$\theta_0 = 1.0$		0	0.007	0.003	0.029
		1	0.002	0.001	0.006
		2	0.000	0.000	0.001
		3	0.000	0.000	0.000
$\theta_0 = 2.0$		0	0.007	0.004	0.026
		1	0.002	0.001	0.008
		2	0.000	0.000	0.002
		3	0.000	0.000	0.000
Panel B: Gaussian family					
$\theta_0 = 0.2$		0	0.002	0.001	0.005
		1	0.001	0.001	0.004
		2	0.000	0.000	0.002
		3	0.000	0.000	0.001
$\theta_0 = 0.5$		0	0.002	0.001	0.007
		1	0.001	0.001	0.005
		2	0.001	0.000	0.002
		3	0.000	0.000	0.001
$\theta_0 = 0.8$		0	0.002	0.001	0.008
		1	0.001	0.001	0.005
		2	0.001	0.000	0.002
		3	0.000	0.000	0.001

Notes. The table reports the empirical distribution of $R_n^{(p)} = \sup_{\mathbf{u} \in [0,1]^2} |\widehat{\alpha}_n^{(p)}(\mathbf{u}) - \mathcal{H}^p(n, \mathbf{u})|$ based on $B_{\text{dist}} = 1500$ replications. The rapid decay of the remainder as p increases is consistent with the smoothing effect of the integrated transform in the plug-in regime.

5.2.3 Null calibration under composite testing

We next compare the calibration rules from the viewpoint of empirical size. Since only 300 evaluation replications are used, the size table should be read as a diagnostic rather than as a high-precision ranking.

Table 4 shows that all three procedures remain on the correct scale, and no gross level distortion is visible. The oracle benchmark behaves as expected. The naive plug-in calibration is sometimes numerically close to the oracle, but this agreement is not systematic. The bootstrap, while still subject to Monte Carlo variability, is the only feasible strategy that is conceptually aligned with the composite-null structure.

Table 4: Empirical size at $n = 500$ in the estimated-parameter study under three calibration rules.

Family	Parameter	p	Oracle	Naive	Bootstrap
Panel A: Clayton family					
	$\theta_0 = 0.5$	0	0.057	0.050	0.047
		1	0.063	0.057	0.050
		2	0.030	0.030	0.037
		3	0.027	0.030	0.033
	$\theta_0 = 1.0$	0	0.037	0.033	0.037
		1	0.033	0.033	0.030
		2	0.027	0.023	0.033
		3	0.033	0.030	0.040
	$\theta_0 = 2.0$	0	0.037	0.030	0.027
		1	0.037	0.037	0.047
		2	0.047	0.047	0.057
		3	0.057	0.057	0.053
Panel B: Gaussian family					
	$\theta_0 = 0.2$	0	0.063	0.053	0.070
		1	0.060	0.070	0.070
		2	0.070	0.073	0.067
		3	0.073	0.067	0.070
	$\theta_0 = 0.5$	0	0.087	0.073	0.073
		1	0.060	0.067	0.060
		2	0.060	0.057	0.063
		3	0.047	0.043	0.053
	$\theta_0 = 0.8$	0	0.047	0.047	0.050
		1	0.060	0.067	0.067
		2	0.043	0.043	0.047
		3	0.037	0.043	0.043

Notes. Empirical size is estimated from $B_{\text{eval}} = 300$ replications. At level 0.05, the Monte Carlo standard error is approximately 0.013. Small differences across calibration methods should therefore not be overstated.

5.2.4 Power under fixed alternatives

Finally, we examine the discriminating power of the corresponding tests. The comparison with the oracle benchmark is particularly informative here, because it separates the intrinsic strength of the statistic from the additional conservativeness induced by parameter estimation and composite-null calibration.

Table 5 shows that the oracle procedure behaves as expected: power is highest for $p = 0$ and decreases as the order of integration increases. This is entirely consistent with the attenuation phenomenon already observed in the change-point setting.

By contrast, the feasible procedures are often much more conservative. In several configurations, the rejection frequencies produced by the naive and bootstrap methods remain close to size even when the oracle power is substantial. Thus, the simulations do *not* suggest that feasible composite-null calibration recovers oracle-level sensitivity at $n = 500$. Rather, they show that process-level approximation accuracy and efficient finite-sample testing are genuinely distinct issues.

Remark 5.2. Theorem 4.1 controls the first-order approximation of the process itself. It does not, by itself, imply that a finite-sample bootstrap implementation will retain strong local power under moderate alternatives. The gap between these two questions is visible in Table 5.

Table 5: Empirical power at $n = 500$ in the estimated-parameter study under three calibration rules.

Family	Alternative (θ_0, θ_1)	p	Oracle	Naive	Bootstrap
Panel A: Clayton family					
	(0.5, 1.0)	0	0.250	0.057	0.063
		1	0.097	0.053	0.060
		2	0.067	0.043	0.050
		3	0.060	0.053	0.060
	(1.0, 1.8)	0	0.263	0.057	0.060
		1	0.110	0.050	0.050
		2	0.093	0.053	0.060
		3	0.080	0.057	0.067
	(1.0, 2.5)	0	0.640	0.047	0.053
		1	0.267	0.073	0.073
		2	0.157	0.063	0.057
		3	0.123	0.067	0.073
Panel B: Gaussian family					
	(0.2, 0.4)	0	0.273	0.073	0.080
		1	0.140	0.067	0.077
		2	0.123	0.067	0.060
		3	0.093	0.067	0.057
	(0.2, 0.6)	0	0.820	0.030	0.033
		1	0.483	0.033	0.037
		2	0.377	0.040	0.040
		3	0.280	0.057	0.057
	(0.5, 0.7)	0	0.363	0.027	0.020
		1	0.217	0.033	0.030
		2	0.157	0.043	0.043
		3	0.157	0.050	0.050

Notes. Power is estimated from $B_{\text{eval}} = 300$ replications. The oracle benchmark reflects the intrinsic discriminatory strength of the statistic, whereas the naive and bootstrap procedures include the additional finite-sample uncertainty generated by parameter estimation and composite-null calibration.

The main features of the estimated-parameter study are summarized in Figure 2. The upper panel confirms the small size of the remainder $R_n^{(p)}$ and its rapid decay with p . The middle panel shows that the three calibration methods are broadly level-correct at the present resolution. The

lower panel illustrates the more delicate finite-sample issue: accurate process approximation does not automatically imply strong testing performance under a composite null.

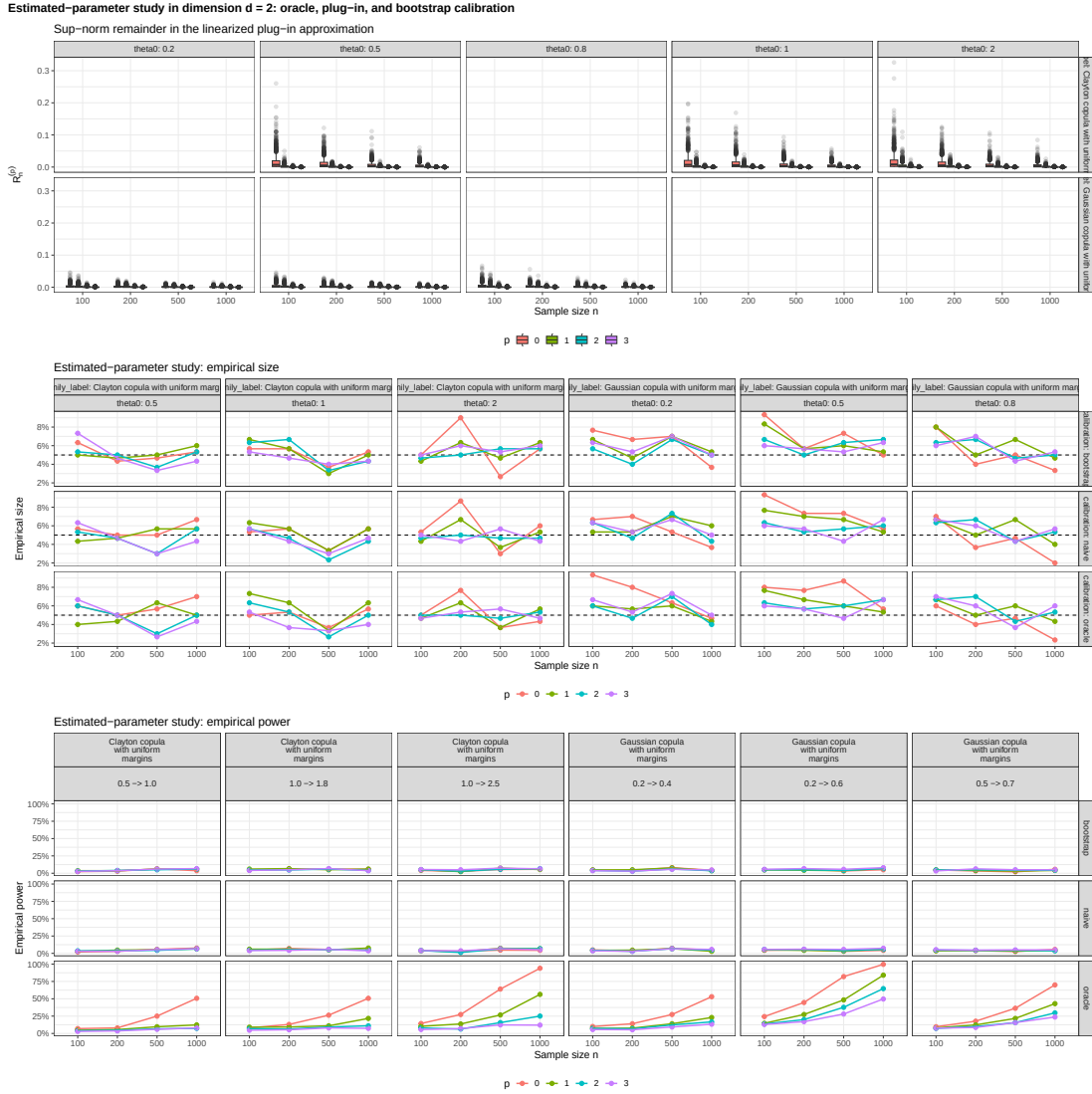


Figure 2: Estimated-parameter study in dimension $d = 2$. Top: distribution of the remainder $R_n^{(p)}$. Middle: empirical size under oracle, naive, and bootstrap calibration. Bottom: empirical power under fixed parametric alternatives. The figure separates process-level approximation accuracy from test-level finite-sample efficiency.

5.3 Discussion and practical implications

The simulations support the theoretical developments, but they also clarify the practical limits of higher-order integration.

1. **Change-point detection.** For abrupt structural breaks, especially when the alternative acts primarily through dependence, the classical empirical process ($p = 0$) remains the most effective procedure. The case $p = 1$ may still be acceptable in some marginal or joint-shift regimes, but the evidence does not favor routine use of $p \geq 2$ in finite samples.

2. **Weighting.** The weighted statistic is genuinely useful when early or late breaks are plausible, but its benefit is concentrated in the low-order procedures. Combined with higher-order integration, weighting may exacerbate signal attenuation.
3. **Plug-in approximation.** Theorem 4.1 is strongly supported numerically. At $n = 500$ in dimension two, the first-order approximation is already accurate, and the remainder becomes negligible for $p \geq 2$.
4. **Composite-null calibration.** Accurate process approximation does not, by itself, guarantee strong finite-sample power. The bootstrap remains the principled feasible calibration device, but the present implementation remains conservative relative to the oracle benchmark.
5. **Choice of integration order.** The simulations consistently indicate that p should be treated as a genuine tuning parameter with real inferential consequences. Higher-order integration is not automatically beneficial; in many practically relevant situations, the low-order procedures remain distinctly preferable.

Overall, the numerical evidence confirms that the integrated empirical framework is mathematically coherent and computationally implementable. At the same time, it makes clear that asymptotic validity and finite-sample efficiency are not synonymous, and that the statistical benefit of integration is strongly regime-dependent.

6 Conclusion and perspectives

The present paper has developed a systematic asymptotic theory for multivariate p -fold integrated empirical processes indexed by lower orthants. The analysis has been carried out simultaneously at three levels: structural, probabilistic, and inferential. At the structural level, the study clarifies the precise nature of the multivariate integrated functionals arising from iterated lower-orthant Stieltjes integration. In particular, it has been shown that, outside effectively totally ordered settings, these quantities cannot in general be reduced to universal scalar polynomial transforms of $F(\mathbf{u})$, but are instead governed by the order-theoretic geometry of multivariate chains. At the probabilistic level, this structural understanding has made it possible to derive strong approximation results which embed the integrated empirical process, on suitable probability spaces, into explicit Gaussian objects with uniform almost sure error control. At the inferential level, these approximations have then been exploited to justify asymptotically valid procedures for multivariate two-sample testing, change-point detection, and settings involving estimated nuisance parameters.

A central message of the paper is that the multivariate integrated empirical process remains asymptotically tractable despite its nonlinear and iterated character. More precisely, the strong approximation results established herein show that, in the scalarizable regime, the integrated process is asymptotically driven by the ordinary multivariate empirical process through a leading Gaussian term modulated by a deterministic factor involving powers of $F(\mathbf{u})$; in the genuinely multivariate partial-order setting, the first-order term is governed by the influence function of the chain functional. This reduction principle is decisive, since it permits the transfer of classical Gaussian approximation techniques from standard empirical process theory to the integrated setting. In particular, the Brownian-bridge approximation provides the appropriate framework for the asymptotic study of integrated Kolmogorov–Smirnov and Cramér–von Mises type functionals, while the Kiefer-process approximation yields a natural sequential Gaussian representation and, in turn, a functional law of the iterated logarithm.

From a methodological standpoint, the results obtained here indicate that integrated empirical procedures should not be regarded merely as formal transforms of the ordinary empirical process, but rather as a genuinely distinct class of stochastic objects whose asymptotic behavior reflects both

the underlying partial order and the smoothing effect induced by repeated integration. This distinction is especially important in the multivariate setting, where the departure from the univariate theory is not purely notational but conceptual. The relevant integrated functionals are intrinsically linked to chain probabilities and V-statistic type representations, and this feature governs both their exact algebraic structure and their asymptotic expansions. In this respect, the present work may be viewed as placing the multivariate theory on a conceptually correct footing, one that is compatible with the geometry of lower-orthant ordering and avoids oversimplified scalar formulations valid only in special regimes.

The statistical consequences of this analysis are equally noteworthy. In the two-sample problem, the Gaussian approximations obtained for integrated discrepancy processes lead to asymptotically justified omnibus procedures whose limiting behavior can be described in explicit functional terms. In the change-point setting, the sequential integrated empirical contrasts yield natural detectors of structural instability, and the corresponding Kiefer-type approximations provide the correct asymptotic calibration for both unweighted and weighted statistics. In the parametric plug-in framework, the results show that the effect of parameter estimation may be absorbed into a first-order linear perturbation, thereby preserving the Gaussian approximation scheme under standard regularity conditions. This last point is of particular importance for composite hypotheses, where purely distribution-free arguments are no longer available and a precise treatment of the estimation effect becomes indispensable.

The numerical study reported in the paper supports the theoretical developments and illustrates the practical relevance of the proposed methodology. At the same time, it also suggests that the role of the integration order p is subtle and should be interpreted with care. The integrated procedures do not uniformly dominate their non-integrated counterparts; rather, they generate a family of statistics whose finite-sample performance depends on the geometry of the alternative, the location of the signal, and the interaction between smoothing and localization. Such behavior is fully consistent with the theoretical structure of the limiting Gaussian terms, whose deterministic weights redistribute sensitivity across the lower-orthant domain.

Several directions for future research emerge naturally from the present work. A first extension would concern observations departing from the i.i.d. framework, including mixing sequences, martingale arrays, random fields, and long-range dependent models. In such contexts, the derivation of strong approximation results for integrated empirical processes would require substantial refinement of the coupling arguments and may reveal new phenomena absent in the independent case. A second direction concerns the deeper analysis of the multivariate chain structure itself. In particular, it would be of significant interest to develop Hoeffding-type projection expansions for the associated V- and U-statistic representations, and to determine more precisely the regimes in which higher-order terms become asymptotically relevant.

Further problems arise on the inferential side. The study of local asymptotic power under contiguous alternatives, the comparison of efficiencies with classical empirical-process-based procedures, and the derivation of bootstrap or multiplier approximations adapted to the integrated setting all appear to be natural and promising continuations. One may also envisage extensions to copula models, censored or truncated data, missing-data mechanisms, and robust goodness-of-fit procedures under contamination. More broadly, it would be worthwhile to investigate whether the framework developed here can be transported to settings where the sample space is no longer Euclidean, for example to functional data, manifold-valued observations, or partially ordered metric spaces.

In conclusion, the theory developed in this paper shows that multivariate integrated empirical processes constitute a mathematically rich and statistically meaningful extension of the classical empirical process. Their study requires a careful combination of order-theoretic analysis, strong approximation techniques, and asymptotic statistical reasoning. The results obtained here provide, in our opinion, a rigorous first-order foundation for this class of processes and demonstrate that

they are sufficiently structured to support substantive statistical applications. At the same time, the range of open problems indicated above suggests that the subject remains far from exhausted and should continue to offer fertile ground for further developments in probability theory and mathematical statistics.

7 Mathematical developments

This section is devoted to the proofs of our results. The previously displayed notations continue to be used in the sequel.

Proposition 7.1. *Assume that the componentwise order is total on the support of F (or, equivalently here, that the scalarizable regime holds). The integrated empirical d.f. $\mathbb{F}_n^{(p)}$ can be expressed by means of $\mathbb{F}_n$ as follows: with probability 1,*

$$\mathbb{F}_n^{(p)}(\mathbf{u}) = \frac{\mathbb{F}_n(\mathbf{u})^{p+1}}{(p+1)!} + \sum_{k=1}^p a_k^{(p)} \frac{\mathbb{F}_n(\mathbf{u})^k}{n^{p-k+1}} \quad \text{for } \mathbf{u} \in [0, 1]^d, n \in \mathbb{N}^*,$$

where the coefficients $a_k^{(p)}$, $1 \leq k \leq p$, are positive integers.

Proposition 7.2. *The p -fold integrated empirical process $\alpha_n^{(p)}$ is related to the empirical process α_n according to, with probability 1,*

$$\alpha_n^{(p)}(\mathbf{u}) = \frac{1}{p!} F(\mathbf{u})^p \alpha_n(\mathbf{u}) + \sum_{k=2}^{p+1} b_k^{(p)} \frac{F(\mathbf{u})^{p+1-k}}{n^{(k-1)/2}} \alpha_n(t)^k + \sum_{k=1}^p a_k^{(p)} \frac{\mathbb{F}_n(\mathbf{u})^k}{n^{p-k+1/2}} \quad (7.1)$$

for $\mathbf{u} \in [0, 1]^d$, $n \in \mathbb{N}^*$, where the coefficients $a_k^{(p)}$, $1 \leq k \leq p$, are those of Proposition 7.1 and the $b_k^{(p)}$, $2 \leq k \leq p+1$, are positive real numbers less than 1.

7.1 Proof of Theorem 2.5

We begin by recalling two elementary but crucial bounds that will be used throughout the proof. Since F is a distribution function on $[0, 1]^d$, we have

$$0 \leq F(\mathbf{u}) \leq 1 \quad \text{and} \quad 0 \leq \mathbb{F}_n(\mathbf{u}) \leq 1, \quad \mathbf{u} \in [0, 1]^d, n \in \mathbb{N}^*,$$

and therefore, for every integer $r \geq 1$,

$$0 \leq F(\mathbf{u})^r \leq 1, \quad 0 \leq \mathbb{F}_n(\mathbf{u})^r \leq 1, \quad \mathbf{u} \in [0, 1]^d. \quad (7.2)$$

By Proposition 7.1, with probability 1, for all $\mathbf{u} \in [0, 1]^d$ and all $n \in \mathbb{N}^*$,

$$\alpha_n^{(p)}(\mathbf{u}) = \frac{1}{p!} F(\mathbf{u})^p \alpha_n(\mathbf{u}) + \sum_{k=2}^{p+1} b_k^{(p)} \frac{F(\mathbf{u})^{p+1-k}}{n^{(k-1)/2}} \alpha_n(\mathbf{u})^k + \sum_{k=1}^p a_k^{(p)} \frac{\mathbb{F}_n(\mathbf{u})^k}{n^{p-k+1/2}}. \quad (7.3)$$

Let

$$G_n(\mathbf{u}) := \frac{1}{p!} F(\mathbf{u})^p \mathbb{B}_{n,F}(\mathbf{u}), \quad \mathbf{u} \in [0, 1]^d.$$

Subtracting $G_n(\mathbf{u})$ from (7.3) yields, almost surely,

$$\alpha_n^{(p)}(\mathbf{u}) - G_n(\mathbf{u}) = \frac{1}{p!} F(\mathbf{u})^p (\alpha_n(\mathbf{u}) - \mathbb{B}_{n,F}(\mathbf{u})) + R_{n,1}(\mathbf{u}) + R_{n,2}(\mathbf{u}), \quad (7.4)$$

where

$$R_{n,1}(\mathbf{u}) := \sum_{k=2}^{p+1} b_k^{(p)} \frac{F(\mathbf{u})^{p+1-k}}{n^{(k-1)/2}} \alpha_n(\mathbf{u})^k, \quad R_{n,2}(\mathbf{u}) := \sum_{k=1}^p a_k^{(p)} \frac{\mathbb{F}_n(\mathbf{u})^k}{n^{p-k+1/2}}.$$

Taking absolute values in (7.4) and then the supremum over $\mathbf{u} \in [0, 1]^d$, we obtain

$$\begin{aligned} \sup_{\mathbf{u} \in [0, 1]^d} |\alpha_n^{(p)}(\mathbf{u}) - G_n(\mathbf{u})| &\leq \frac{1}{p!} \sup_{\mathbf{u} \in [0, 1]^d} |\alpha_n(\mathbf{u}) - \mathbb{B}_{n,F}(\mathbf{u})| \\ &\quad + \sup_{\mathbf{u} \in [0, 1]^d} |R_{n,1}(\mathbf{u})| + \sup_{\mathbf{u} \in [0, 1]^d} |R_{n,2}(\mathbf{u})|. \end{aligned} \quad (7.5)$$

We now bound the remainder terms. Using (7.2) and taking absolute values of the coefficients $a_k^{(p)}$, we have

$$\sup_{\mathbf{u} \in [0, 1]^d} |R_{n,2}(\mathbf{u})| \leq \sum_{k=1}^p |a_k^{(p)}| \frac{1}{n^{p-k+1/2}} \leq \frac{1}{\sqrt{n}} \sum_{k=1}^p |a_k^{(p)}| =: \frac{A_p}{\sqrt{n}}. \quad (7.6)$$

Similarly, since $0 < b_k^{(p)} < 1$ and $0 \leq F(\mathbf{u}) \leq 1$, we obtain

$$\begin{aligned} \sup_{\mathbf{u} \in [0, 1]^d} |R_{n,1}(\mathbf{u})| &\leq \sum_{k=2}^{p+1} \frac{1}{n^{(k-1)/2}} \sup_{\mathbf{u} \in [0, 1]^d} |\alpha_n(\mathbf{u})|^k \\ &\leq \frac{1}{\sqrt{n}} \sum_{k=2}^{p+1} \sup_{\mathbf{u} \in [0, 1]^d} |\alpha_n(\mathbf{u})|^k. \end{aligned} \quad (7.7)$$

The class of lower orthants in $[0, 1]^d$ is a VC class. Hence, by standard law-of-the-iterated-logarithm type bounds for VC-indexed empirical processes,

$$\sup_{\mathbf{u} \in [0, 1]^d} |\alpha_n(\mathbf{u})| = O\left(\sqrt{\log \log n}\right) \quad \text{almost surely.} \quad (7.8)$$

Consequently, for each fixed k ,

$$\sup_{\mathbf{u} \in [0, 1]^d} |\alpha_n(\mathbf{u})|^k = O\left((\log \log n)^{k/2}\right) \quad \text{almost surely,}$$

and therefore (7.7) yields

$$\sup_{\mathbf{u} \in [0, 1]^d} |R_{n,1}(\mathbf{u})| = O\left(\frac{(\log \log n)^{(p+1)/2}}{\sqrt{n}}\right) \quad \text{almost surely.} \quad (7.9)$$

Next, by Theorem A (Borisov's strong approximation), on a suitable probability space,

$$\sup_{\mathbf{u} \in [0, 1]^d} |\alpha_n(\mathbf{u}) - \mathbb{B}_{n,F}(\mathbf{u})| = O\left(n^{-1/(2(2d-1))} \log n\right) \quad \text{almost surely.} \quad (7.10)$$

Combining (7.5), (7.6), (7.9), and (7.10), we obtain

$$\sup_{\mathbf{u} \in [0, 1]^d} |\alpha_n^{(p)}(\mathbf{u}) - G_n(\mathbf{u})| = O\left(n^{-1/(2(2d-1))} \log n\right) + O\left(\frac{(\log \log n)^{(p+1)/2}}{\sqrt{n}}\right) + O\left(\frac{1}{\sqrt{n}}\right)$$

almost surely. Since, for every $d \geq 1$,

$$\frac{(\log \log n)^{(p+1)/2}}{\sqrt{n}} = o\left(n^{-1/(2(2d-1))} \log n\right), \quad n \rightarrow \infty,$$

the Borisov term dominates, and we conclude that

$$\sup_{\mathbf{u} \in [0,1]^d} \left| \alpha_n^{(p)}(\mathbf{u}) - \frac{1}{p!} F(\mathbf{u})^p \mathbb{B}_{n,F}(\mathbf{u}) \right| = O\left(n^{-1/(2(2d-1))} \log n\right) \quad \text{almost surely.}$$

This completes the proof of Theorem 2.5. $\square$

7.2 Proof of Theorem 2.8

Fix $p \in \mathbb{N}^*$. We work on a probability space on which the expansion (7.1) of Proposition 7.1 holds and on which the strong approximation of Theorem B is available. Multiplying (7.1) by $\sqrt{n}$ and rearranging, we obtain, with probability 1, for every $\mathbf{u} \in [0,1]^d$ and $n \in \mathbb{N}^*$,

$$\begin{aligned} \sqrt{n} \alpha_n^{(p)}(\mathbf{u}) &= \frac{1}{p!} F(\mathbf{u})^p \sqrt{n} \alpha_n(\mathbf{u}) + \sum_{k=2}^{p+1} b_k^{(p)} F(\mathbf{u})^{p+1-k} n^{1/2-(k-1)/2} \alpha_n(\mathbf{u})^k \\ &\quad + \sum_{k=1}^p a_k^{(p)} \mathbb{F}_n(\mathbf{u})^k n^{1/2-(p-k+1/2)} \\ &= \frac{1}{p!} F(\mathbf{u})^p \sqrt{n} \alpha_n(\mathbf{u}) + \sum_{k=2}^{p+1} b_k^{(p)} F(\mathbf{u})^{p+1-k} n^{-(k-2)/2} \alpha_n(\mathbf{u})^k \\ &\quad + \sum_{k=1}^p a_k^{(p)} \mathbb{F}_n(\mathbf{u})^k n^{-(p-k)}. \end{aligned} \quad (7.11)$$

Let $\mathbb{K}$ denote the F -Kiefer process in the sense of Theorem B; in particular, we may (and do) take $\mathbb{K} = \mathbb{F}_F$. Subtracting $\frac{1}{p!} F(\mathbf{u})^p \mathbb{K}(n, \mathbf{u})$ from both sides of (7.11), taking absolute values and suprema over $\mathbf{u} \in [0,1]^d$, we obtain

$$\begin{aligned} \sup_{\mathbf{u} \in [0,1]^d} \left| \sqrt{n} \alpha_n^{(p)}(\mathbf{u}) - \frac{1}{p!} F(\mathbf{u})^p \mathbb{K}(n, \mathbf{u}) \right| &\leq \frac{1}{p!} \sup_{\mathbf{u} \in [0,1]^d} \left| \sqrt{n} \alpha_n(\mathbf{u}) - \mathbb{K}(n, \mathbf{u}) \right| \\ &\quad + \sum_{k=2}^{p+1} n^{-(k-2)/2} \sup_{\mathbf{u} \in [0,1]^d} |\alpha_n(\mathbf{u})|^k + \sum_{k=1}^p |a_k^{(p)}| n^{-(p-k)}. \end{aligned} \quad (7.12)$$

Here we used $0 \leq F(\mathbf{u}) \leq 1$, $0 < b_k^{(p)} < 1$ and $0 \leq \mathbb{F}_n(\mathbf{u}) \leq 1$. By Theorem B, almost surely as $n \rightarrow \infty$,

$$\sup_{\mathbf{u} \in [0,1]^d} \left| \sqrt{n} \alpha_n(\mathbf{u}) - \mathbb{K}(n, \mathbf{u}) \right| = O\left(n^{1/2-1/(4d)} (\log n)^{3/2}\right). \quad (7.13)$$

Hence the second term on the right-hand side of (7.12), keeping in mind (7.8), satisfies

$$\begin{aligned} \sum_{k=2}^{p+1} n^{-(k-2)/2} \sup_{\mathbf{u} \in [0,1]^d} |\alpha_n(\mathbf{u})|^k &= \sup_{\mathbf{u} \in [0,1]^d} |\alpha_n(\mathbf{u})|^2 + \sum_{k=3}^{p+1} n^{-(k-2)/2} \sup_{\mathbf{u} \in [0,1]^d} |\alpha_n(\mathbf{u})|^k \\ &= O(\log \log n) + O\left(\sum_{k=3}^{p+1} n^{-(k-2)/2} (\log \log n)^{k/2}\right) \\ &= O(\log \log n), \quad \text{a.s.} \end{aligned}$$

since the sum over $k \geq 3$ is $o(1)$ times a polylogarithmic factor. Finally,

$$\sum_{k=1}^p |a_k^{(p)}| n^{-(p-k)} = O(1).$$

Inserting these bounds into (7.12) and using (7.13), we obtain, almost surely as $n \rightarrow \infty$,

$$\begin{aligned} \sup_{\mathbf{u} \in [0,1]^d} \left| \sqrt{n} \alpha_n^{(p)}(\mathbf{u}) - \frac{1}{p!} F(\mathbf{u})^p \mathbb{K}(n, \mathbf{u}) \right| &= O\left(n^{1/2-1/(4d)} (\log n)^{3/2}\right) + O(\log \log n) + O(1) \\ &= O\left(n^{1/2-1/(4d)} (\log n)^{3/2}\right), \end{aligned}$$

because $\log \log n = o(n^{1/2-1/(4d)} (\log n)^{3/2})$. This completes the proof of Theorem 2.8. $\square$

7.3 Proof of Theorem 3.4

We work under $\mathcal{H}_0''$ and on a probability space on which the identity (7.1) holds simultaneously for the integrated empirical processes built from the two sample segments (the first $\lfloor ns \rfloor$ observations and the last $n - \lfloor ns \rfloor$ observations, for every $s \in (0, 1)$), and on which the strong approximation of Theorem B can be realized for these two segments by means of two independent Kiefer processes $\mathbb{K}_1$ and $\mathbb{K}_2$ as described just before (3.5).

Step 1: Algebraic decomposition of $\tilde{\alpha}_n^{(p)}$. By (1.1) and (3.4), for $s \in (0, 1)$ and $\mathbf{u} \in [0, 1]^d$,

$$\begin{aligned} \tilde{\alpha}_n^{(p)}(s, \mathbf{u}) &= \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{n^{3/2}} \left[\left(\mathbb{F}_{\lfloor ns \rfloor}^{(p)-}(\mathbf{u}) - F^{(p)}(\mathbf{u}) \right) - \left(\mathbb{F}_{n-\lfloor ns \rfloor}^{(p)+}(\mathbf{u}) - F^{(p)}(\mathbf{u}) \right) \right] \\ &= \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{n^{3/2}} \left(\frac{\alpha_{\lfloor ns \rfloor}^{(p)-}(\mathbf{u})}{\sqrt{\lfloor ns \rfloor}} - \frac{\alpha_{n-\lfloor ns \rfloor}^{(p)+}(\mathbf{u})}{\sqrt{n - \lfloor ns \rfloor}} \right). \end{aligned}$$

We now apply Proposition 7.1 to $\alpha_{\lfloor ns \rfloor}^{(p)-}$ and $\alpha_{n-\lfloor ns \rfloor}^{(p)+}$ (with n replaced by $\lfloor ns \rfloor$ and $n - \lfloor ns \rfloor$, respectively). Collecting separately the linear part in $\alpha_{\lfloor ns \rfloor}^{\pm}$ and $\alpha_{n-\lfloor ns \rfloor}^{\pm}$ and the nonlinear remainder terms yields the decomposition (valid a.s. for every $s \in (0, 1)$ and $\mathbf{u} \in [0, 1]^d$)

$$\tilde{\alpha}_n^{(p)}(s, \mathbf{u}) = \text{I}_n(s, \mathbf{u}) - \text{II}_n(s, \mathbf{u}) + \text{III}_n(s, \mathbf{u}) + \text{IV}_n(s, \mathbf{u}), \quad (7.14)$$

where the four terms are defined as follows:

$$\text{I}_n(s, \mathbf{u}) := \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{p! n^{3/2}} \frac{\alpha_{\lfloor ns \rfloor}^{-}(\mathbf{u})}{\sqrt{\lfloor ns \rfloor}} F(\mathbf{u})^p, \quad (7.15)$$

$$\text{II}_n(s, \mathbf{u}) := \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{p! n^{3/2}} \frac{\alpha_{n-\lfloor ns \rfloor}^{+}(\mathbf{u})}{\sqrt{n - \lfloor ns \rfloor}} F(\mathbf{u})^p, \quad (7.16)$$

$$\text{III}_n(s, \mathbf{u}) := \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{n^{3/2}} \sum_{k=2}^{p+1} b_k^{(p)} F(\mathbf{u})^{p+1-k} \left(\frac{\alpha_{\lfloor ns \rfloor}^{-}(\mathbf{u})^k}{\lfloor ns \rfloor^{k/2}} - \frac{\alpha_{n-\lfloor ns \rfloor}^{+}(\mathbf{u})^k}{(n - \lfloor ns \rfloor)^{k/2}} \right),$$

$$\text{IV}_n(s, \mathbf{u}) := \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{n^{3/2}} \sum_{k=1}^p a_k^{(p)} \left(\frac{\mathbb{F}_{\lfloor ns \rfloor}^{-}(\mathbf{u})^k}{\lfloor ns \rfloor^{p-k+1}} - \frac{\mathbb{F}_{n-\lfloor ns \rfloor}^{+}(\mathbf{u})^k}{(n - \lfloor ns \rfloor)^{p-k+1}} \right).$$

Step 2: Uniform reduction to the leading term $\delta_n = \text{I}_n - \text{II}_n$. Set

$$\delta_n(s, \mathbf{u}) := \text{I}_n(s, \mathbf{u}) - \text{II}_n(s, \mathbf{u}).$$

We show that, with probability 1,

$$\sup_{s \in (0,1)} \sup_{\mathbf{u} \in [0,1]^d} |\tilde{\alpha}_n^{(p)}(s, \mathbf{u}) - \delta_n(s, \mathbf{u})| = O\left(\frac{\log \log n}{\sqrt{n}}\right), \quad n \rightarrow \infty. \quad (7.17)$$

By (7.14), it suffices to prove the uniform bounds

$$\sup_{s \in (0,1)} \sup_{\mathbf{u}} |\text{III}_n(s, \mathbf{u})| = O\left(\frac{\log \log n}{\sqrt{n}}\right), \quad \sup_{s \in (0,1)} \sup_{\mathbf{u}} |\text{IV}_n(s, \mathbf{u})| = O\left(\frac{1}{\sqrt{n}}\right).$$

Bound for IV_n. Since $0 \leq \mathbb{F}_m^\pm(\mathbf{u}) \leq 1$ for all m and $\mathbf{u}$, we have

$$|\text{IV}_n(s, \mathbf{u})| \leq \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{n^{3/2}} \sum_{k=1}^p |a_k^{(p)}| \left(\lfloor ns \rfloor^{-(p-k+1)} + (n - \lfloor ns \rfloor)^{-(p-k+1)} \right).$$

As $p - k + 1 \geq 1$, the function $x \mapsto x^{-(p-k+1)}$ is decreasing on $(0, \infty)$, and thus uniformly in $s \in (0, 1)$,

$$\begin{aligned} \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{n^{3/2}} \lfloor ns \rfloor^{-(p-k+1)} &\leq \frac{n - \lfloor ns \rfloor}{n^{3/2}} \leq \frac{1}{\sqrt{n}}, \\ \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{n^{3/2}} (n - \lfloor ns \rfloor)^{-(p-k+1)} &\leq \frac{\lfloor ns \rfloor}{n^{3/2}} \leq \frac{1}{\sqrt{n}}. \end{aligned}$$

Therefore

$$\sup_{s \in (0,1)} \sup_{\mathbf{u} \in [0,1]^d} |\text{IV}_n(s, \mathbf{u})| \leq \frac{1}{\sqrt{n}} \sum_{k=1}^p |a_k^{(p)}| = O\left(\frac{1}{\sqrt{n}}\right). \quad (7.18)$$

Bound for III_n. Using $0 < b_k^{(p)} < 1$ and $0 \leq F(\mathbf{u}) \leq 1$, we obtain

$$\begin{aligned} |\text{III}_n(s, \mathbf{u})| &\leq \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{n^{3/2}} \sum_{k=2}^{p+1} \left(\frac{|\alpha_{\lfloor ns \rfloor}^-(\mathbf{u})|^k}{\lfloor ns \rfloor^{k/2}} + \frac{|\alpha_{n-\lfloor ns \rfloor}^+(\mathbf{u})|^k}{(n - \lfloor ns \rfloor)^{k/2}} \right) \\ &= \sum_{k=2}^{p+1} \left[\frac{n - \lfloor ns \rfloor}{n^{3/2}} \cdot \frac{|\alpha_{\lfloor ns \rfloor}^-(\mathbf{u})|^k}{\lfloor ns \rfloor^{(k-2)/2}} + \frac{\lfloor ns \rfloor}{n^{3/2}} \cdot \frac{|\alpha_{n-\lfloor ns \rfloor}^+(\mathbf{u})|^k}{(n - \lfloor ns \rfloor)^{(k-2)/2}} \right]. \end{aligned}$$

For $k = 2$, the denominators $\lfloor ns \rfloor^{(k-2)/2}$ and $(n - \lfloor ns \rfloor)^{(k-2)/2}$ equal 1, hence

$$\sup_{s \in (0,1)} \sup_{\mathbf{u}} \frac{n - \lfloor ns \rfloor}{n^{3/2}} |\alpha_{\lfloor ns \rfloor}^-(\mathbf{u})|^2 \leq \frac{1}{\sqrt{n}} \sup_{1 \leq m \leq n} \sup_{\mathbf{u} \in [0,1]^d} |\alpha_m(\mathbf{u})|^2,$$

and similarly for the $+$ -term. For $k \geq 3$, since $(k - 2)/2 \geq 1/2$,

$$\sup_{s \in (0,1)} \sup_{\mathbf{u}} \frac{n - \lfloor ns \rfloor}{n^{3/2}} \cdot \frac{|\alpha_{\lfloor ns \rfloor}^-(\mathbf{u})|^k}{\lfloor ns \rfloor^{(k-2)/2}} \leq \frac{1}{n} \sup_{1 \leq m \leq n} \sup_{\mathbf{u} \in [0,1]^d} |\alpha_m(\mathbf{u})|^k,$$

and again similarly for the $+$ -term. Since the class of lower orthants is VC, a uniform law of the iterated logarithm (equivalently, maximal inequalities plus Borel–Cantelli) yields, almost surely,

$$\sup_{1 \leq m \leq n} \sup_{\mathbf{u} \in [0,1]^d} |\alpha_m(\mathbf{u})| = O\left(\sqrt{\log \log n}\right), \quad n \rightarrow \infty.$$

Consequently,

$$\sup_{1 \leq m \leq n} \sup_{\mathbf{u}} |\alpha_m(\mathbf{u})|^2 = O(\log \log n), \quad \sup_{1 \leq m \leq n} \sup_{\mathbf{u}} |\alpha_m(\mathbf{u})|^k = O((\log \log n)^{k/2}) \quad (k \geq 3),$$

almost surely. Putting these estimates together yields

$$\sup_{s \in (0,1)} \sup_{\mathbf{u} \in [0,1]^d} |\text{III}_n(s, \mathbf{u})| = O\left(\frac{\log \log n}{\sqrt{n}}\right), \quad n \rightarrow \infty, \quad (7.19)$$

almost surely. Combining (7.19) and (7.18) proves (7.17).

Step 3: Partial-sum representation of δ_n . Define, for $\mathbf{u} \in [0, 1]^d$,

$$\gamma_n(\mathbf{u}) := \sqrt{n} \alpha_n(\mathbf{u}) = \sum_{i=1}^n (\mathbb{1}_{\{\mathbf{X}_i \leq \mathbf{u}\}} - F(\mathbf{u})),$$

and, for $s \in (0, 1)$,

$$\begin{aligned} \gamma_{\lfloor ns \rfloor}^-(\mathbf{u}) &:= \sqrt{\lfloor ns \rfloor} \alpha_{\lfloor ns \rfloor}^-(\mathbf{u}) = \sum_{i=1}^{\lfloor ns \rfloor} (\mathbb{1}_{\{\mathbf{X}_i \leq \mathbf{u}\}} - F(\mathbf{u})), \\ \gamma_{n-\lfloor ns \rfloor}^+(\mathbf{u}) &:= \sqrt{n - \lfloor ns \rfloor} \alpha_{n-\lfloor ns \rfloor}^+(\mathbf{u}) = \sum_{i=\lfloor ns \rfloor+1}^n (\mathbb{1}_{\{\mathbf{X}_i \leq \mathbf{u}\}} - F(\mathbf{u})). \end{aligned}$$

Then $\gamma_n(\mathbf{u}) = \gamma_{\lfloor ns \rfloor}^-(\mathbf{u}) + \gamma_{n-\lfloor ns \rfloor}^+(\mathbf{u})$, and from (7.15)–(7.16) we rewrite δ_n as

$$\begin{aligned} \delta_n(s, \mathbf{u}) &= \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{p! n^{3/2}} F(\mathbf{u})^p \left(\frac{\gamma_{\lfloor ns \rfloor}^-(\mathbf{u})}{\lfloor ns \rfloor} - \frac{\gamma_{n-\lfloor ns \rfloor}^+(\mathbf{u})}{n - \lfloor ns \rfloor} \right) \\ &= \frac{F(\mathbf{u})^p}{p! \sqrt{n}} \left(\frac{n - \lfloor ns \rfloor}{n} \gamma_{\lfloor ns \rfloor}^-(\mathbf{u}) - \frac{\lfloor ns \rfloor}{n} \gamma_{n-\lfloor ns \rfloor}^+(\mathbf{u}) \right), \end{aligned}$$

which is equivalent to (7.20):

$$\delta_n(s, \mathbf{u}) = \frac{F(\mathbf{u})^p}{p! \sqrt{n}} \left(\gamma_{\lfloor ns \rfloor}^-(\mathbf{u}) - \frac{\lfloor ns \rfloor}{n} \gamma_n(\mathbf{u}) \right) = \frac{F(\mathbf{u})^p}{p! \sqrt{n}} \left(\frac{n - \lfloor ns \rfloor}{n} \gamma_n(\mathbf{u}) - \gamma_{n-\lfloor ns \rfloor}^+(\mathbf{u}) \right). \quad (7.20)$$

Step 4: Strong approximation of δ_n by $\mathbb{K}_n^{(p)}$. By Theorem B applied to the two independent segments, there exist two independent Kiefer processes $\mathbb{K}_1$ and $\mathbb{K}_2$ such that, with probability 1, as $n \rightarrow \infty$,

$$\sup_{\mathbf{u} \in [0,1]^d} \left| \gamma_{\lfloor n/2 \rfloor}^-(\mathbf{u}) - \mathbb{K}_2(\lfloor n/2 \rfloor, \mathbf{u}) \right| = O\left(n^{1/2-1/(4d)} (\log n)^{3/2}\right), \quad (7.21)$$

$$\sup_{\mathbf{u} \in [0,1]^d} \left| \gamma_{n-\lfloor n/2 \rfloor}^+(\mathbf{u}) - \mathbb{K}_1(\lfloor n/2 \rfloor, \mathbf{u}) \right| = O\left(n^{1/2-1/(4d)} (\log n)^{3/2}\right). \quad (7.22)$$

Adding (7.21) and (7.22) yields (7.23):

$$\sup_{\mathbf{u} \in [0,1]^d} \left| \gamma_n(\mathbf{u}) - (\mathbb{K}_1(\lfloor n/2 \rfloor, \mathbf{u}) + \mathbb{K}_2(\lfloor n/2 \rfloor, \mathbf{u})) \right| = O\left(n^{1/2-1/(4d)} (\log n)^{3/2}\right). \quad (7.23)$$

Recall that $\mathbb{K}_n^{(p)}$ is defined in (3.5) by $\mathbb{K}_n^{(p)}(s, \mathbf{u}) = \frac{1}{p!} F(\mathbf{u})^p \overset{\circ}{\mathbb{K}}_n(s, \mathbf{u})$, where $\overset{\circ}{\mathbb{K}}_n$ is the stitched Gaussian field built from $\mathbb{K}_1$ and $\mathbb{K}_2$ (piecewise on $[0, 1/2]$ and $[1/2, 1]$). Using the representation (7.20) together with (7.23) and the uniform bound $0 \leq F(\mathbf{u})^p \leq 1$, we deduce that, with probability 1,

$$\sup_{s \in [0,1]} \sup_{\mathbf{u} \in [0,1]^d} \left| \delta_n(s, \mathbf{u}) - \mathbb{K}_n^{(p)}(s, \mathbf{u}) \right| = O\left(n^{-1/(4d)} (\log n)^{3/2}\right). \quad (7.24)$$

Final step: triangle inequality. Using (7.17) and (7.24), we conclude by the triangle inequality that

$$\begin{aligned} \sup_{s \in (0,1)} \sup_{\mathbf{u} \in [0,1]^d} |\tilde{\alpha}_n^{(p)}(s, \mathbf{u}) - \mathbb{K}_n^{(p)}(s, \mathbf{u})| &\leq \sup_{s \in (0,1)} \sup_{\mathbf{u} \in [0,1]^d} |\tilde{\alpha}_n^{(p)}(s, \mathbf{u}) - \delta_n(s, \mathbf{u})| \\ &\quad + \sup_{s \in (0,1)} \sup_{\mathbf{u} \in [0,1]^d} |\delta_n(s, \mathbf{u}) - \mathbb{K}_n^{(p)}(s, \mathbf{u})| \\ &= O\left(\frac{\log \log n}{\sqrt{n}}\right) + O\left(n^{-1/(4d)} (\log n)^{3/2}\right). \end{aligned}$$

Since $(\log \log n)/\sqrt{n} = o(n^{-1/(4d)} (\log n)^{3/2})$, we obtain

$$\sup_{s \in (0,1)} \sup_{\mathbf{u} \in [0,1]^d} |\tilde{\alpha}_n^{(p)}(s, \mathbf{u}) - \mathbb{K}_n^{(p)}(s, \mathbf{u})| = O\left(n^{-1/(4d)} (\log n)^{3/2}\right),$$

almost surely as $n \rightarrow \infty$. This completes the proof of Theorem 3.4. $\square$

7.4 Proof of Theorem 4.1

We work under assumptions (i)–(v) and on a probability space carrying the Kiefer process $\Gamma_{F\theta_0}$ and the Gaussian objects $\mathbf{W}(\cdot)$, $\mathcal{G}(\cdot, \cdot)$, $\mathcal{H}^p(\cdot, \cdot)$ defined above. Recall

$$\hat{\alpha}_n^{(p)}(\mathbf{u}) = \sqrt{n} \left(\mathbb{F}_n^{(p)}(\mathbf{u}) - F^{(p)}(\hat{\theta}_n, \mathbf{u}) \right), \quad \mathcal{H}^p(n, \mathbf{u}) = \frac{1}{p!} F(\theta_0, \mathbf{u})^p \mathcal{G}(n, \mathbf{u}),$$

and

$$\boldsymbol{\varepsilon}_n^{(p)} = \sup_{\mathbf{u} \in [0,1]^d} \left| \hat{\alpha}_n^{(p)}(\mathbf{u}) - \mathcal{H}^p(n, \mathbf{u}) \right|.$$

Also recall the auxiliary quantity

$$\boldsymbol{\eta}_n := \sup_{\mathbf{u} \in [0,1]^d} |\hat{\alpha}_n(\mathbf{u}) - \mathcal{G}(n, \mathbf{u})|.$$

Using Proposition 7.1 *mutatis mutandis* with $F(\mathbf{u})$ replaced by $F(\hat{\theta}_n, \mathbf{u})$ and $\alpha_n(\mathbf{u})$ replaced by $\hat{\alpha}_n(\mathbf{u})$: for all $\mathbf{u} \in [0, 1]^d$,

$$\hat{\alpha}_n^{(p)}(\mathbf{u}) = \frac{1}{p!} F(\hat{\theta}_n, \mathbf{u})^p \hat{\alpha}_n(\mathbf{u}) + \sum_{k=2}^{p+1} b_k^{(p)} F(\hat{\theta}_n, \mathbf{u})^{p+1-k} \frac{\hat{\alpha}_n(\mathbf{u})^k}{n^{(k-1)/2}} + \sum_{k=1}^p a_k^{(p)} \frac{\mathbb{F}_n(\mathbf{u})^k}{n^{p-k+1/2}}. \quad (7.25)$$

Write

$$\hat{\alpha}_n(\mathbf{u}) = \mathcal{G}(n, \mathbf{u}) + (\hat{\alpha}_n(\mathbf{u}) - \mathcal{G}(n, \mathbf{u})).$$

Substitute this into (7.25) and expand powers using the binomial theorem. Subtracting $\mathcal{H}^p(n, \mathbf{u}) = \frac{1}{p!} F(\theta_0, \mathbf{u})^p \mathcal{G}(n, \mathbf{u})$ yields, almost surely for each $\mathbf{u}$,

$$\begin{aligned} \hat{\alpha}_n^{(p)}(\mathbf{u}) - \mathcal{H}^p(n, \mathbf{u}) &= \frac{1}{p!} F(\hat{\theta}_n, \mathbf{u})^p (\hat{\alpha}_n(\mathbf{u}) - \mathcal{G}(n, \mathbf{u})) \\ &\quad + \frac{1}{p!} \left(F(\hat{\theta}_n, \mathbf{u})^p - F(\theta_0, \mathbf{u})^p \right) \mathcal{G}(n, \mathbf{u}) \\ &\quad + \sum_{k=1}^p a_k^{(p)} \frac{\mathbb{F}_n(\mathbf{u})^k}{n^{p-k+1/2}} \\ &\quad + \sum_{k=2}^{p+1} b_k^{(p)} \frac{F(\hat{\theta}_n, \mathbf{u})^{p+1-k}}{n^{(k-1)/2}} \sum_{i=0}^k \binom{k}{i} \mathcal{G}(n, \mathbf{u})^{k-i} (\hat{\alpha}_n(\mathbf{u}) - \mathcal{G}(n, \mathbf{u}))^i. \end{aligned} \quad (7.26)$$

Let $A_p := \sum_{k=1}^p |a_k^{(p)}|$ and $A_p'' := \max_{0 \leq i \leq k \leq p+1} \binom{k}{i}$. Taking suprema over $\mathbf{u} \in [0, 1]^d$ in (7.26) and using the bounds $0 \leq \mathbb{F}_n(\mathbf{u}) \leq 1$, $0 \leq F(\hat{\theta}_n, \mathbf{u}) \leq 1$, and $0 < b_k^{(p)} < 1$, we obtain almost surely

$$\begin{aligned} \varepsilon_n^{(p)} &\leq \eta_n + \sup_{\mathbf{u} \in [0, 1]^d} |F(\hat{\theta}_n, \mathbf{u})^p - F(\theta_0, \mathbf{u})^p| \sup_{\mathbf{u} \in [0, 1]^d} |\mathcal{G}(n, \mathbf{u})| + \frac{A_p}{\sqrt{n}} \\ &\quad + \frac{A_p''}{\sqrt{n}} \left(\sum_{k=0}^{p+1} \eta_n^k \right) \left(\sum_{k=0}^{p+1} \sup_{\mathbf{u} \in [0, 1]^d} |\mathcal{G}(n, \mathbf{u})|^k \right). \end{aligned} \quad (7.27)$$

Next, since $0 \leq a, b \leq 1$ implies $|a^p - b^p| \leq p|a - b|$, we have

$$\sup_{\mathbf{u}} |F(\hat{\theta}_n, \mathbf{u})^p - F(\theta_0, \mathbf{u})^p| \leq p \sup_{\mathbf{u}} |F(\hat{\theta}_n, \mathbf{u}) - F(\theta_0, \mathbf{u})|.$$

Thus (7.27) implies

$$\varepsilon_n^{(p)} \leq \eta_n + p \Delta_n \sup_{\mathbf{u} \in [0, 1]^d} |\mathcal{G}(n, \mathbf{u})| + \frac{A_p}{\sqrt{n}} + \frac{A_p''}{\sqrt{n}} \left(\sum_{k=0}^{p+1} \eta_n^k \right) \left(\sum_{k=0}^{p+1} \sup_{\mathbf{u} \in [0, 1]^d} |\mathcal{G}(n, \mathbf{u})|^k \right), \quad (7.28)$$

where

$$\Delta_n := \sup_{\mathbf{u} \in [0, 1]^d} |F(\hat{\theta}_n, \mathbf{u}) - F(\theta_0, \mathbf{u})|.$$

Subtle point. All suprema are taken after the purely algebraic manipulations, so no measurability issue occurs: each term is a supremum of measurable functions over a compact set, and the processes admit separable versions (standard for empirical/Kiefer settings).

Recall

$$\mathcal{G}(n, \mathbf{u}) = \frac{1}{\sqrt{n}} \left(\Gamma_{F_{\theta_0}}(n, \mathbf{u}) - \mathbf{W}(n) \nabla_{\theta} F(\theta_0, \mathbf{u})^{\top} \right).$$

Hence, using the triangle inequality,

$$\sup_{\mathbf{u} \in [0, 1]^d} |\mathcal{G}(n, \mathbf{u})| \leq \frac{1}{\sqrt{n}} \sup_{\mathbf{u} \in [0, 1]^d} |\Gamma_{F_{\theta_0}}(n, \mathbf{u})| + \frac{1}{\sqrt{n}} \|\mathbf{W}(n)\| \sup_{\mathbf{u} \in [0, 1]^d} \|\nabla_{\theta} F(\theta_0, \mathbf{u})\|. \quad (7.29)$$

The second factor in the last term is finite by assumption (iv) (uniform continuity on Λ implies boundedness on the compact closure of the relevant domain; in particular $\sup_{\mathbf{u}} \|\nabla_{\theta} F(\theta_0, \mathbf{u})\| < \infty$). Moreover, $\mathbf{W}(\cdot)$ is a finite-dimensional Brownian motion, hence by the one-dimensional law of the iterated logarithm applied componentwise,

$$\|\mathbf{W}(n)\| = O\left(\sqrt{n \log \log n}\right) \quad \text{a.s.}$$

For the Kiefer field term, using that $\Gamma_{F_{\theta_0}}(n, \cdot)$ is a (tied-down) Gaussian process indexed by $[0, 1]^d$ with variance bounded by n , standard Gaussian maximal inequalities (or directly the functional LIL for Kiefer processes; see, e.g., the references used in [25]) yield

$$\sup_{\mathbf{u} \in [0, 1]^d} |\Gamma_{F_{\theta_0}}(n, \mathbf{u})| = O\left(\sqrt{n \log \log n}\right) \quad \text{a.s.}$$

Inserting these bounds into (7.29) gives the almost sure growth control

$$\sup_{\mathbf{u} \in [0, 1]^d} |\mathcal{G}(n, \mathbf{u})| = O\left(\sqrt{\log \log n}\right) \quad \text{a.s.} \quad (7.30)$$

Subtle point. The normalization $n^{-1/2}$ in the definition of $\mathcal{G}$ is essential: without it, the Kiefer field has typical size $\sqrt{n \log \log n}$; dividing by $\sqrt{n}$ yields the natural $\sqrt{\log \log n}$ scale appearing in (7.30).

Let Λ be the closed neighborhood of θ_0 from assumption (iv). Define the event $E_n := \{\widehat{\theta}_n \in \Lambda\}$. By assumption (i) together with $\varepsilon_n \xrightarrow{\mathbb{P}} 0$, we have $\widehat{\theta}_n \xrightarrow{\mathbb{P}} \theta_0$, hence

$$\mathbb{P}(E_n) \longrightarrow 1. \quad (7.31)$$

On the event E_n , for each $\mathbf{u} \in [0, 1]^d$ the mean value theorem in $\mathbb{R}^q$ provides a point $\theta_n^*(\mathbf{u})$ on the segment $[\theta_0, \widehat{\theta}_n]$ (hence still in Λ because Λ is convex after shrinking if necessary, or simply because we may replace Λ by the closed line-segment hull of a neighborhood of θ_0) such that

$$F(\widehat{\theta}_n, \mathbf{u}) - F(\theta_0, \mathbf{u}) = (\widehat{\theta}_n - \theta_0) \nabla_{\theta} F(\theta_n^*(\mathbf{u}), \mathbf{u})^{\top}.$$

Taking suprema over $\mathbf{u}$ and using $\theta_n^*(\mathbf{u}) \in \Lambda$ on E_n , we obtain the uniform bound

$$\Delta_n \mathbb{1}_{E_n} \leq \|\widehat{\theta}_n - \theta_0\| \sup_{\theta \in \Lambda} \sup_{\mathbf{u} \in [0, 1]^d} \|\nabla_{\theta} F(\theta, \mathbf{u})\|. \quad (7.32)$$

The supremum is finite by assumption (iv). Furthermore, by the linear expansion in (i) and the centering/finite second moment conditions (ii)–(iii),

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) = O_{\mathbb{P}}(1), \quad \text{hence} \quad \|\widehat{\theta}_n - \theta_0\| = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right).$$

Combining with (7.32) and (7.31) yields

$$\Delta_n = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right).$$

Therefore, using (7.30),

$$\Delta_n \sup_{\mathbf{u}} |\mathcal{G}(n, \mathbf{u})| = O_{\mathbb{P}}\left(\frac{\sqrt{\log \log n}}{\sqrt{n}}\right) = o_{\mathbb{P}}(1). \quad (7.33)$$

Subtle point. The uniformity in (7.32) is the critical place where assumption (iv) enters: it upgrades pointwise differentiability to a bound that survives the supremum over $\mathbf{u}$.

By Theorem 6.1 of [20] (the strong approximation for the empirical process with estimated parameters), under assumptions (i)–(v) and $\varepsilon_n \xrightarrow{\mathbb{P}} 0$ we have

$$\eta_n \xrightarrow{\mathbb{P}} 0. \quad (7.34)$$

In particular, $\eta_n = O_{\mathbb{P}}(1)$ and $\sum_{k=0}^{p+1} \eta_n^k = O_{\mathbb{P}}(1)$. Now return to (7.28). The first term η_n converges to 0 in probability by (7.34). The drift term is $o_{\mathbb{P}}(1)$ by (7.33). Clearly $A_p/\sqrt{n} = o(1)$. For the last term in (7.28), use (7.30) to get

$$\sum_{k=0}^{p+1} \sup_{\mathbf{u}} |\mathcal{G}(n, \mathbf{u})|^k = O_{\mathbb{P}}\left(\sum_{k=0}^{p+1} (\log \log n)^{k/2}\right) = O_{\mathbb{P}}((\log \log n)^{(p+1)/2}),$$

hence

$$\frac{1}{\sqrt{n}} \left(\sum_{k=0}^{p+1} \eta_n^k \right) \left(\sum_{k=0}^{p+1} \sup_{\mathbf{u}} |\mathcal{G}(n, \mathbf{u})|^k \right) = O_{\mathbb{P}}\left(\frac{(\log \log n)^{(p+1)/2}}{\sqrt{n}}\right) = o_{\mathbb{P}}(1).$$

Therefore every term on the right-hand side of (7.28) converges to 0 in probability, and we conclude that

$$\varepsilon_n^{(p)} \xrightarrow{\mathbb{P}} 0,$$

which completes the proof of Theorem 4.1. $\square$

Proposition 7.3 (Explicit formulas for the p -fold iterated integrals). *Let $\{\mathbf{X}_i\}_{i \geq 1}$ be i.i.d. $\mathbb{R}^d$ -valued random vectors with continuous c.d.f. F supported on $[0, 1]^d$. For $p \in \mathbb{N}$ and $\mathbf{u} \in [0, 1]^d$, define recursively on lower orthants $[\mathbf{0}, \mathbf{u}] := \prod_{j=1}^d [0, u_j]$ by*

$$F^{(0)}(\mathbf{u}) := F(\mathbf{u}), \quad F^{(1)}(\mathbf{u}) := \int_{[\mathbf{0}, \mathbf{u}]} F(\mathbf{s}) dF(\mathbf{s}),$$

and, for $p \geq 2$,

$$F^{(p)}(\mathbf{u}) := \int_{[\mathbf{0}, \mathbf{u}]} dF(\mathbf{s}_1) \int_{[\mathbf{0}, \mathbf{s}_1]} dF(\mathbf{s}_2) \cdots \int_{[\mathbf{0}, \mathbf{s}_{p-1}]} F(\mathbf{s}_p) dF(\mathbf{s}_p).$$

Similarly, for the empirical c.d.f. $\mathbb{F}_n$ based on $\mathbf{X}_1, \dots, \mathbf{X}_n$, define $\mathbb{F}_n^{(p)}$ by replacing F with $\mathbb{F}_n$ in the displays above.

Then the following assertions hold.

(i) **Probabilistic representation.** For every $p \in \mathbb{N}$ and $\mathbf{u} \in [0, 1]^d$,

$$F^{(p)}(\mathbf{u}) = \mathbb{P}\{\mathbf{X}_{p+1} \leq \mathbf{X}_p \leq \cdots \leq \mathbf{X}_1 \leq \mathbf{u}\}. \quad (7.35)$$

(ii) **Closed form under a total order.** Assume that the componentwise order $\leq$ is total on the support of F (for instance, if $d = 1$, or more generally if $\mathbf{X} = g(Z)$ with Z real-valued and $g_1, \dots, g_d$ strictly increasing). Then, for all $p \in \mathbb{N}$ and $\mathbf{u} \in [0, 1]^d$,

$$F^{(p)}(\mathbf{u}) = \frac{F(\mathbf{u})^{p+1}}{(p+1)!}. \quad (7.36)$$

(iii) **Closed form under independent coordinates.** If the coordinates of $\mathbf{X}$ are independent and continuous with marginals $F_1, \dots, F_d$, so that

$$F(\mathbf{u}) = \prod_{j=1}^d F_j(u_j),$$

then, for all $p \in \mathbb{N}$ and $\mathbf{u} \in [0, 1]^d$,

$$F^{(p)}(\mathbf{u}) = \frac{F(\mathbf{u})^{p+1}}{((p+1)!)^d}.$$

(iv) **Empirical closed form in a total-order setting.** Assume the setting of (7.36), and let

$$m(\mathbf{u}) := \#\{1 \leq i \leq n : \mathbf{X}_i \leq \mathbf{u}\} = n \mathbb{F}_n(\mathbf{u}) \quad a.s.$$

Then, for all $p \in \mathbb{N}$,

$$\mathbb{F}_n^{(p)}(\mathbf{u}) = \frac{1}{n^{p+1}} \binom{m(\mathbf{u}) + p}{p+1} = \frac{1}{n^{p+1}} \binom{n \mathbb{F}_n(\mathbf{u}) + p}{p+1} \quad a.s. \quad (7.37)$$

(v) **Empirical representation for general d .** For general $d \geq 1$, without any total-order assumption, one has the exact V-statistic representation

$$\mathbb{F}_n^{(p)}(\mathbf{u}) = \frac{1}{n^{p+1}} \sum_{i_1, \dots, i_{p+1}=1}^n \mathbb{1}\{\mathbf{X}_{i_{p+1}} \leq \cdots \leq \mathbf{X}_{i_1} \leq \mathbf{u}\} \quad a.s. \quad (7.38)$$

If, in addition, one defines the associated U -statistic

$$U_{n,p}(\mathbf{u}) := \frac{1}{(n)_{p+1}} \sum_{\substack{i_1, \dots, i_{p+1}=1 \\ \text{all distinct}}}^n \mathbb{1}\{\mathbf{X}_{i_{p+1}} \leq \dots \leq \mathbf{X}_{i_1} \leq \mathbf{u}\}, \quad (7.39)$$

where $(n)_{p+1} := n(n-1) \cdots (n-p)$, then

$$\mathbb{E} U_{n,p}(\mathbf{u}) = F^{(p)}(\mathbf{u}),$$

and the V -statistic bias satisfies

$$\left| \mathbb{E} F_n^{(p)}(\mathbf{u}) - F^{(p)}(\mathbf{u}) \right| \leq 1 - \frac{(n)_{p+1}}{n^{p+1}} = O\left(\frac{1}{n}\right), \quad n \rightarrow \infty,$$

uniformly in $\mathbf{u} \in [0, 1]^d$.

(vi) **Universal bound and sharpness.** For every $d \geq 1$, $p \in \mathbb{N}$ and $\mathbf{u} \in [0, 1]^d$,

$$0 \leq F^{(p)}(\mathbf{u}) \leq \frac{F(\mathbf{u})^{p+1}}{(p+1)!},$$

and the upper bound is attained under the total-order/comonotone assumption of (7.36).

Proof. We repeatedly use the elementary identity

$$\int_{[0, \mathbf{t}]} g(\mathbf{s}) dF(\mathbf{s}) = \mathbb{E}[g(\mathbf{X}) \mathbb{1}_{\{\mathbf{X} \leq \mathbf{t}\}}], \quad (7.40)$$

valid for every bounded measurable $g : [0, 1]^d \rightarrow \mathbb{R}$, where $\mathbf{X} \sim F$. Its empirical analogue is

$$\int_{[0, \mathbf{t}]} g(\mathbf{s}) dF_n(\mathbf{s}) = \frac{1}{n} \sum_{i=1}^n g(\mathbf{X}_i) \mathbb{1}_{\{\mathbf{X}_i \leq \mathbf{t}\}}. \quad (7.41)$$

Proof of (i). We argue by induction on p . For $p = 0$,

$$F^{(0)}(\mathbf{u}) = F(\mathbf{u}) = \mathbb{P}(\mathbf{X}_1 \leq \mathbf{u}).$$

Assume (7.35) holds for $p-1$. Then, by the recursive definition and (7.40),

$$F^{(p)}(\mathbf{u}) = \mathbb{E}[F^{(p-1)}(\mathbf{X}_1) \mathbb{1}_{\{\mathbf{X}_1 \leq \mathbf{u}\}}].$$

Using the induction hypothesis, with $(\mathbf{X}_2, \dots, \mathbf{X}_{p+1})$ independent of $\mathbf{X}_1$,

$$F^{(p-1)}(\mathbf{X}_1) = \mathbb{P}(\mathbf{X}_{p+1} \leq \dots \leq \mathbf{X}_2 \leq \mathbf{X}_1 \mid \mathbf{X}_1).$$

Hence

$$F^{(p)}(\mathbf{u}) = \mathbb{E}[\mathbb{P}(\mathbf{X}_{p+1} \leq \dots \leq \mathbf{X}_2 \leq \mathbf{X}_1 \mid \mathbf{X}_1) \mathbb{1}_{\{\mathbf{X}_1 \leq \mathbf{u}\}}],$$

and the tower property yields

$$F^{(p)}(\mathbf{u}) = \mathbb{P}(\mathbf{X}_{p+1} \leq \dots \leq \mathbf{X}_2 \leq \mathbf{X}_1 \leq \mathbf{u}).$$

This proves (7.35).

Proof of (ii). Assume the order is total on the support of F . On the event

$$B(\mathbf{u}) := \{\mathbf{X}_1, \dots, \mathbf{X}_{p+1} \leq \mathbf{u}\},$$

continuity of F implies that ties occur with probability zero, so almost surely there is a unique permutation π of $\{1, \dots, p+1\}$ such that

$$\mathbf{X}_{\pi(p+1)} \leq \dots \leq \mathbf{X}_{\pi(1)}.$$

By exchangeability,

$$\mathbb{P}(\mathbf{X}_{\pi(p+1)} \leq \dots \leq \mathbf{X}_{\pi(1)} \leq \mathbf{u}) = \mathbb{P}(\mathbf{X}_{p+1} \leq \dots \leq \mathbf{X}_1 \leq \mathbf{u}).$$

Summing over the $(p+1)!$ possible permutations gives

$$(p+1)! \mathbb{P}(\mathbf{X}_{p+1} \leq \dots \leq \mathbf{X}_1 \leq \mathbf{u}) = \mathbb{P}(B(\mathbf{u})) = F(\mathbf{u})^{p+1},$$

hence (7.36) follows from (7.35).

Proof of (iii). Assume

$$F(\mathbf{u}) = \prod_{j=1}^d F_j(u_j)$$

with continuous marginals and independent coordinates. Set

$$U_{i,j} := F_j(X_{i,j}), \quad 1 \leq i \leq p+1, 1 \leq j \leq d.$$

Then the $U_{i,j}$ are i.i.d. $\mathcal{U}(0, 1)$. The event in (7.35) is equivalent to

$$U_{p+1,j} \leq \dots \leq U_{1,j} \leq F_j(u_j), \quad j = 1, \dots, d.$$

By independence across coordinates,

$$F^{(p)}(\mathbf{u}) = \prod_{j=1}^d \mathbb{P}(U_{p+1,j} \leq \dots \leq U_{1,j} \leq F_j(u_j)).$$

The one-dimensional formula gives

$$\mathbb{P}(U_{p+1,j} \leq \dots \leq U_{1,j} \leq F_j(u_j)) = \frac{F_j(u_j)^{p+1}}{(p+1)!},$$

so

$$F^{(p)}(\mathbf{u}) = \prod_{j=1}^d \frac{F_j(u_j)^{p+1}}{(p+1)!} = \frac{F(\mathbf{u})^{p+1}}{((p+1)!)^d}.$$

Proof of (v). Using (7.41) recursively in the definition of $\mathbb{F}_n^{(p)}$ yields

$$\mathbb{F}_n^{(p)}(\mathbf{u}) = \frac{1}{n^{p+1}} \sum_{i_1, \dots, i_{p+1}=1}^n \mathbb{1}\{\mathbf{X}_{i_{p+1}} \leq \dots \leq \mathbf{X}_{i_1} \leq \mathbf{u}\},$$

which is exactly (7.38). For the U-statistic in (7.39), all indices are distinct, so by independence,

$$\mathbb{E} U_{n,p}(\mathbf{u}) = \mathbb{P}(\mathbf{X}_{p+1} \leq \dots \leq \mathbf{X}_1 \leq \mathbf{u}) = F^{(p)}(\mathbf{u}),$$

by (7.35). To compare the V-statistic and the population quantity, split the sum in (7.38) into tuples with all indices distinct and tuples with at least one collision. Since the kernel is bounded by 1, we obtain

$$\left| \mathbb{E} \mathbb{F}_n^{(p)}(\mathbf{u}) - F^{(p)}(\mathbf{u}) \right| \leq 1 - \frac{(n)_{p+1}}{n^{p+1}},$$

and the right-hand side is $O(n^{-1})$.

Proof of (iv). Assume now the total-order setting of (ii). Let

$$m := m(\mathbf{u}) = n \mathbb{F}_n(\mathbf{u}).$$

Then the event in (7.38) is equivalent to choosing a weakly increasing $(p+1)$ -tuple

$$1 \leq i_{p+1} \leq \cdots \leq i_1 \leq m.$$

Hence

$$\mathbb{F}_n^{(p)}(\mathbf{u}) = \frac{1}{n^{p+1}} \sum_{i_1=1}^m \sum_{i_2=1}^{i_1} \cdots \sum_{i_{p+1}=1}^{i_p} 1.$$

The number of such weakly increasing $(p+1)$ -tuples is $\binom{m+p}{p+1}$, so (7.37) follows.

Proof of (vi). The lower bound is immediate from (7.35). For the upper bound, let Π be a uniform random permutation of $\{1, \dots, p+1\}$, independent of $(\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$. By exchangeability,

$$F^{(p)}(\mathbf{u}) = \mathbb{P}(\mathbf{X}_{\Pi(p+1)} \leq \cdots \leq \mathbf{X}_{\Pi(1)} \leq \mathbf{u}).$$

Conditioning on $(\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$, we get

$$F^{(p)}(\mathbf{u}) = \frac{1}{(p+1)!} \mathbb{E} \left[\sum_{\pi \in \mathfrak{S}_{p+1}} \mathbb{1}\{\mathbf{X}_{\pi(p+1)} \leq \cdots \leq \mathbf{X}_{\pi(1)} \leq \mathbf{u}\} \right].$$

Because ties occur with probability zero, for a fixed realization there is at most one admissible permutation. Therefore

$$\sum_{\pi \in \mathfrak{S}_{p+1}} \mathbb{1}\{\mathbf{X}_{\pi(p+1)} \leq \cdots \leq \mathbf{X}_{\pi(1)} \leq \mathbf{u}\} \leq \mathbb{1}\{\mathbf{X}_1, \dots, \mathbf{X}_{p+1} \leq \mathbf{u}\}$$

almost surely. Taking expectations yields

$$F^{(p)}(\mathbf{u}) \leq \frac{1}{(p+1)!} \mathbb{P}(\mathbf{X}_1, \dots, \mathbf{X}_{p+1} \leq \mathbf{u}) = \frac{F(\mathbf{u})^{p+1}}{(p+1)!}.$$

Sharpness follows immediately from (ii). $\square$

8 Appendix: Multivariate integrated empirical functionals of the first kind

This appendix complements the iterated-integral processes studied in the main text by considering the *first-kind* integrated functionals

$$\int_{[0,t]} F(\mathbf{s})^p dF(\mathbf{s}) \quad \text{and} \quad \int_{[0,t]} (F(\mathbf{t}) - F(\mathbf{s}))^p dF(\mathbf{s}).$$

These are single Stieltjes integrals and should therefore be distinguished from the p -fold iterated integrals $F^{(p)}$ studied in the body of the paper.

Throughout this appendix, $\leq$ denotes the componentwise order on $\mathbb{R}^d$, and

$$[\mathbf{0}, \mathbf{t}] := \prod_{j=1}^d [0, t_j], \quad \mathbf{t} \in [0, 1]^d.$$

We assume that F is a continuous c.d.f. on $[0, 1]^d$, with associated probability measure μ_F , and that $\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2, \dots$ are i.i.d. random vectors with law F . We write

$$\mathbb{F}_n(\mathbf{t}) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\mathbf{X}_i \leq \mathbf{t}\}}, \quad \mathbf{t} \in [0, 1]^d.$$

8.1 Definitions

Definition 8.1 (First-kind integrated functionals). For $p \in \mathbb{N}$ and $\mathbf{t} \in [0, 1]^d$, define

$$\begin{aligned} \tilde{F}^{(p)}(\mathbf{t}) &:= \int_{[\mathbf{0}, \mathbf{t}]} F(\mathbf{s})^p dF(\mathbf{s}), \\ \check{F}^{(p)}(\mathbf{t}) &:= \int_{[\mathbf{0}, \mathbf{t}]} (F(\mathbf{t}) - F(\mathbf{s}))^p dF(\mathbf{s}). \end{aligned}$$

Their empirical analogues are

$$\begin{aligned} \tilde{\mathbb{F}}_n^{(p)}(\mathbf{t}) &:= \int_{[\mathbf{0}, \mathbf{t}]} \mathbb{F}_n(\mathbf{s})^p d\mathbb{F}_n(\mathbf{s}), \\ \check{\mathbb{F}}_n^{(p)}(\mathbf{t}) &:= \int_{[\mathbf{0}, \mathbf{t}]} (\mathbb{F}_n(\mathbf{t}) - \mathbb{F}_n(\mathbf{s}))^p d\mathbb{F}_n(\mathbf{s}). \end{aligned}$$

To avoid confusion with the change-point process $\tilde{\alpha}_n^{(p)}(s, \mathbf{u})$ introduced earlier, we denote the associated empirical processes by

$$\begin{aligned} \tilde{\beta}_n^{(p)}(\mathbf{t}) &:= \sqrt{n}(\tilde{\mathbb{F}}_n^{(p)}(\mathbf{t}) - \tilde{F}^{(p)}(\mathbf{t})), \\ \check{\beta}_n^{(p)}(\mathbf{t}) &:= \sqrt{n}(\check{\mathbb{F}}_n^{(p)}(\mathbf{t}) - \check{F}^{(p)}(\mathbf{t})). \end{aligned}$$

8.2 A basic Stieltjes–expectation identity

Lemma 8.2 (Stieltjes–expectation identities). *Let $g : [0, 1]^d \rightarrow \mathbb{R}$ be bounded and measurable. Then, for every $\mathbf{t} \in [0, 1]^d$,*

$$\begin{aligned} \int_{[\mathbf{0}, \mathbf{t}]} g(\mathbf{s}) dF(\mathbf{s}) &= \mathbb{E}[g(\mathbf{X}) \mathbb{1}_{\{\mathbf{X} \leq \mathbf{t}\}}], \\ \int_{[\mathbf{0}, \mathbf{t}]} g(\mathbf{s}) d\mathbb{F}_n(\mathbf{s}) &= \frac{1}{n} \sum_{i=1}^n g(\mathbf{X}_i) \mathbb{1}_{\{\mathbf{X}_i \leq \mathbf{t}\}}. \end{aligned}$$

Proof. The first identity is simply the definition of integration with respect to the measure induced by F :

$$\int_{[\mathbf{0}, \mathbf{t}]} g(\mathbf{s}) dF(\mathbf{s}) = \int_{[0, 1]^d} g(\mathbf{s}) \mathbb{1}_{[\mathbf{0}, \mathbf{t}]}(\mathbf{s}) d\mu_F(\mathbf{s}) = \mathbb{E}[g(\mathbf{X}) \mathbb{1}_{\{\mathbf{X} \leq \mathbf{t}\}}].$$

The second identity follows similarly from the empirical measure

$$\mu_n := \frac{1}{n} \sum_{i=1}^n \delta_{\mathbf{X}_i}.$$

□

8.3 Canonical representations

Proposition 8.3 (Canonical identities). *For every $p \in \mathbb{N}$ and $\mathbf{t} \in [0, 1]^d$,*

$$\widetilde{F}^{(p)}(\mathbf{t}) = \mathbb{E}[F(\mathbf{X})^p \mathbb{1}_{\{\mathbf{X} \leq \mathbf{t}\}}], \quad (8.1)$$

$$\widetilde{\mathbb{F}}_n^{(p)}(\mathbf{t}) = \frac{1}{n} \sum_{i=1}^n \mathbb{F}_n(\mathbf{X}_i)^p \mathbb{1}_{\{\mathbf{X}_i \leq \mathbf{t}\}}. \quad (8.2)$$

Moreover,

$$\check{F}^{(p)}(\mathbf{t}) = \sum_{k=0}^p (-1)^k \binom{p}{k} F(\mathbf{t})^{p-k} \widetilde{F}^{(k)}(\mathbf{t}), \quad (8.3)$$

$$\check{\mathbb{F}}_n^{(p)}(\mathbf{t}) = \sum_{k=0}^p (-1)^k \binom{p}{k} \mathbb{F}_n(\mathbf{t})^{p-k} \widetilde{\mathbb{F}}_n^{(k)}(\mathbf{t}). \quad (8.4)$$

Proof. The identities (8.1) and (8.2) follow directly from Lemma 8.2 with $g(\mathbf{s}) = F(\mathbf{s})^p$ and $g(\mathbf{s}) = \mathbb{F}_n(\mathbf{s})^p$, respectively.

For (8.3), expand

$$(F(\mathbf{t}) - F(\mathbf{s}))^p = \sum_{k=0}^p (-1)^k \binom{p}{k} F(\mathbf{t})^{p-k} F(\mathbf{s})^k$$

and integrate termwise over $[\mathbf{0}, \mathbf{t}]$ with respect to $dF(\mathbf{s})$. The empirical identity (8.4) is obtained in the same way. $\square$

8.4 Universal bounds

Proposition 8.4 (Elementary bounds). *For every $p \in \mathbb{N}$ and $\mathbf{t} \in [0, 1]^d$,*

$$0 \leq \widetilde{F}^{(p)}(\mathbf{t}) \leq F(\mathbf{t})^{p+1}, \quad 0 \leq \check{F}^{(p)}(\mathbf{t}) \leq F(\mathbf{t})^{p+1}. \quad (8.5)$$

Proof. The lower bounds are immediate. For the upper bounds, note that whenever $\mathbf{s} \leq \mathbf{t}$,

$$0 \leq F(\mathbf{s}) \leq F(\mathbf{t}) \leq 1, \quad 0 \leq F(\mathbf{t}) - F(\mathbf{s}) \leq F(\mathbf{t}).$$

Hence

$$0 \leq F(\mathbf{s})^p \leq F(\mathbf{t})^p, \quad 0 \leq (F(\mathbf{t}) - F(\mathbf{s}))^p \leq F(\mathbf{t})^p.$$

Integrating over $[\mathbf{0}, \mathbf{t}]$ yields

$$\widetilde{F}^{(p)}(\mathbf{t}) \leq F(\mathbf{t})^p \int_{[\mathbf{0}, \mathbf{t}]} dF(\mathbf{s}) = F(\mathbf{t})^{p+1},$$

and similarly for $\check{F}^{(p)}(\mathbf{t})$. $\square$

Remark 8.5. The bound (8.5) is a simple uniform estimate. In the comonotone case below, the exact value is

$$\widetilde{F}^{(p)}(\mathbf{t}) = \check{F}^{(p)}(\mathbf{t}) = \frac{F(\mathbf{t})^{p+1}}{p+1},$$

so the crude upper bound $F(\mathbf{t})^{p+1}$ is generally not attained for $p \geq 1$.

8.5 Closed forms in two benchmark regimes

8.5.1 Comonotone case

Definition 8.6 (Comonotonicity). A random vector $\mathbf{X} \sim F$ is said to be *comonotone* if there exists a real-valued random variable Z and strictly increasing functions $g_1, \dots, g_d$ such that

$$\mathbf{X} = (g_1(Z), \dots, g_d(Z)) \quad \text{a.s.}$$

Proposition 8.7 (Closed forms under comonotonicity). Assume $\mathbf{X}$ is comonotone and F is continuous. Then, for every $p \in \mathbb{N}$ and $\mathbf{t} \in [0, 1]^d$,

$$\tilde{F}^{(p)}(\mathbf{t}) = \check{F}^{(p)}(\mathbf{t}) = \frac{F(\mathbf{t})^{p+1}}{p+1}.$$

Moreover, for every $n \in \mathbb{N}^*$, almost surely,

$$\begin{aligned} \tilde{\mathbb{F}}_n^{(p)}(\mathbf{t}) &= \frac{1}{n^{p+1}} \sum_{k=1}^{m(\mathbf{t})} k^p, \\ \check{\mathbb{F}}_n^{(p)}(\mathbf{t}) &= \frac{1}{n^{p+1}} \sum_{k=0}^{m(\mathbf{t})-1} k^p, \end{aligned}$$

where

$$m(\mathbf{t}) := n \mathbb{F}_n(\mathbf{t}) \in \{0, 1, \dots, n\}.$$

Proof. Write $\mathbf{X} = (g_1(Z), \dots, g_d(Z))$ with Z real-valued and c.d.f. H . Define

$$h(\mathbf{s}) := \min_{1 \leq j \leq d} g_j^{-1}(s_j).$$

Then

$$\{\mathbf{X} \leq \mathbf{s}\} = \{Z \leq h(\mathbf{s})\}, \quad F(\mathbf{s}) = H(h(\mathbf{s})).$$

For fixed $\mathbf{t}$, set $z_{\mathbf{t}} := h(\mathbf{t})$, so that $F(\mathbf{t}) = H(z_{\mathbf{t}})$. Using (8.1),

$$\tilde{F}^{(p)}(\mathbf{t}) = \mathbb{E}[H(Z)^p \mathbb{1}_{\{Z \leq z_{\mathbf{t}}\}}] = \int_{(-\infty, z_{\mathbf{t}}]} H(z)^p dH(z) = \frac{H(z_{\mathbf{t}})^{p+1}}{p+1} = \frac{F(\mathbf{t})^{p+1}}{p+1}.$$

Applying (8.3) and the identity

$$\sum_{k=0}^p (-1)^k \binom{p}{k} \frac{1}{k+1} = \frac{1}{p+1}$$

gives the same formula for $\check{F}^{(p)}(\mathbf{t})$. For the empirical formulas, comonotonicity implies that the sample is totally ordered. Let

$$\mathbf{X}_{(1)} \leq \dots \leq \mathbf{X}_{(n)}$$

be the ordered sample. Continuity implies that ties occur with probability zero, and

$$\mathbb{F}_n(\mathbf{X}_{(k)}) = \frac{k}{n}.$$

If $m(\mathbf{t})$ denotes the number of observations not exceeding $\mathbf{t}$, then

$$\tilde{\mathbb{F}}_n^{(p)}(\mathbf{t}) = \frac{1}{n} \sum_{k=1}^{m(\mathbf{t})} \left(\frac{k}{n}\right)^p = \frac{1}{n^{p+1}} \sum_{k=1}^{m(\mathbf{t})} k^p.$$

Similarly,

$$\check{\mathbb{F}}_n^{(p)}(\mathbf{t}) = \frac{1}{n} \sum_{k=1}^{m(\mathbf{t})} \left(\frac{m(\mathbf{t}) - k}{n}\right)^p = \frac{1}{n^{p+1}} \sum_{k=0}^{m(\mathbf{t})-1} k^p. \quad \square$$

8.5.2 Independent coordinates

Proposition 8.8 (Closed forms under independence). *Assume the coordinates of $\mathbf{X}$ are independent and continuous with marginals $F_1, \dots, F_d$, so that*

$$F(\mathbf{t}) = \prod_{j=1}^d F_j(t_j).$$

Then, for every $p \in \mathbb{N}$ and $\mathbf{t} \in [0, 1]^d$,

$$\tilde{F}^{(p)}(\mathbf{t}) = \frac{F(\mathbf{t})^{p+1}}{(p+1)^d}, \quad (8.6)$$

$$\check{F}^{(p)}(\mathbf{t}) = F(\mathbf{t})^{p+1} \sum_{k=0}^p (-1)^k \binom{p}{k} \frac{1}{(k+1)^d}. \quad (8.7)$$

Proof. Set

$$U_j := F_j(X_j), \quad j = 1, \dots, d.$$

Then the U_j are independent $\mathcal{U}(0, 1)$ random variables, and

$$F(\mathbf{X}) = \prod_{j=1}^d U_j.$$

If we write $v_j := F_j(t_j)$, then

$$\{\mathbf{X} \leq \mathbf{t}\} = \{U_j \leq v_j, 1 \leq j \leq d\}.$$

Therefore, by (8.1),

$$\tilde{F}^{(p)}(\mathbf{t}) = \mathbb{E} \left[\left(\prod_{j=1}^d U_j \right)^p \prod_{j=1}^d \mathbb{1}_{\{U_j \leq v_j\}} \right] = \prod_{j=1}^d \mathbb{E} \left[U_j^p \mathbb{1}_{\{U_j \leq v_j\}} \right] = \prod_{j=1}^d \frac{v_j^{p+1}}{p+1},$$

which gives (8.6). Formula (8.7) then follows from (8.3). $\square$

8.6 Empirical V-statistics, U-statistics, and bias

Proposition 8.9 (V-statistic representation and bias). *For every $p \in \mathbb{N}$ and $\mathbf{t} \in [0, 1]^d$, one has*

$$\tilde{\mathbb{F}}_n^{(p)}(\mathbf{t}) = \frac{1}{n^{p+1}} \sum_{i_0, \dots, i_p=1}^n \mathbb{1}_{\{\mathbf{X}_{i_0} \leq \mathbf{t}\}} \prod_{\ell=1}^p \mathbb{1}_{\{\mathbf{X}_{i_\ell} \leq \mathbf{X}_{i_0}\}}, \quad a.s. \quad (8.8)$$

Thus $\tilde{\mathbb{F}}_n^{(p)}(\mathbf{t})$ is a V-statistic of order $p+1$. Define the associated U-statistic

$$\tilde{U}_{n,p}(\mathbf{t}) := \frac{1}{(n)_{p+1}} \sum_{\substack{i_0, \dots, i_p=1 \\ \text{all distinct}}}^n \mathbb{1}_{\{\mathbf{X}_{i_0} \leq \mathbf{t}\}} \prod_{\ell=1}^p \mathbb{1}_{\{\mathbf{X}_{i_\ell} \leq \mathbf{X}_{i_0}\}},$$

where $(n)_{p+1} := n(n-1) \cdots (n-p)$. Then

$$\mathbb{E} \tilde{U}_{n,p}(\mathbf{t}) = \tilde{F}^{(p)}(\mathbf{t}), \quad (8.9)$$

and

$$\left| \mathbb{E} \tilde{\mathbb{F}}_n^{(p)}(\mathbf{t}) - \tilde{F}^{(p)}(\mathbf{t}) \right| \leq 1 - \frac{(n)_{p+1}}{n^{p+1}} = O\left(\frac{1}{n}\right), \quad n \rightarrow \infty, \quad (8.10)$$

uniformly in $\mathbf{t} \in [0, 1]^d$.

Proof. From (8.2),

$$\widetilde{\mathbb{F}}_n^{(p)}(\mathbf{t}) = \frac{1}{n} \sum_{i_0=1}^n \mathbb{1}_{\{\mathbf{X}_{i_0} \leq \mathbf{t}\}} \left(\frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{\mathbf{X}_j \leq \mathbf{X}_{i_0}\}} \right)^p.$$

Expanding the p th power yields (8.8).

For the U-statistic, all indices are distinct, so by independence,

$$\mathbb{E} \widetilde{U}_{n,p}(\mathbf{t}) = \mathbb{E} \left[\mathbb{1}_{\{\mathbf{X}_0 \leq \mathbf{t}\}} \prod_{\ell=1}^p \mathbb{1}_{\{\mathbf{X}_\ell \leq \mathbf{X}_0\}} \right] = \mathbb{E} [F(\mathbf{X}_0)^p \mathbb{1}_{\{\mathbf{X}_0 \leq \mathbf{t}\}}] = \widetilde{F}^{(p)}(\mathbf{t}),$$

which proves (8.9). Finally, split the V-statistic sum into tuples with all indices distinct and tuples with at least one collision. Since the kernel takes values in $[0, 1]$, the total weight of collision terms is at most $1 - \frac{(n)_{p+1}}{n^{p+1}}$, and (8.10) follows. $\square$

Corollary 8.10 (Asymptotic negligibility of the V-statistic bias). *For every fixed $p \in \mathbb{N}$ and $\mathbf{t} \in [0, 1]^d$,*

$$\sqrt{n} \left| \mathbb{E} \widetilde{\mathbb{F}}_n^{(p)}(\mathbf{t}) - \widetilde{F}^{(p)}(\mathbf{t}) \right| = O\left(\frac{1}{\sqrt{n}}\right) \rightarrow 0.$$

8.7 A safe first-order expansion in the total-order case

The following statement is deliberately restricted to the total-order/comonotone setting. In the general multivariate partial-order setting, no reduction to a univariate power-sum formula is available, so one should not expect the same expansion without additional structure.

Lemma 8.11 (Discrete power-sum approximation). *For every integer $m \geq 0$ and every $p \in \mathbb{N}$,*

$$\left| \sum_{k=1}^m k^p - \frac{m^{p+1}}{p+1} \right| \leq C_p m^p,$$

where $C_p > 0$ depends only on p . Consequently,

$$\left| \frac{1}{n^{p+1}} \sum_{k=1}^m k^p - \frac{1}{p+1} \left(\frac{m}{n}\right)^{p+1} \right| \leq \frac{C_p}{n}.$$

Proof. This is the standard comparison between $\sum_{k=1}^m k^p$ and $\int_0^m x^p dx = m^{p+1}/(p+1)$. $\square$

Proposition 8.12 (First-order expansion in the total-order case). *Assume the total-order/comonotone setting of Proposition 8.7. Fix $p \in \mathbb{N}$ and $\mathbf{t} \in [0, 1]^d$. Then*

$$\widetilde{\beta}_n^{(p)}(\mathbf{t}) = F(\mathbf{t})^p \alpha_n(\mathbf{t}) + O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right), \quad n \rightarrow \infty, \quad (8.11)$$

where

$$\alpha_n(\mathbf{t}) := \sqrt{n} (\mathbb{F}_n(\mathbf{t}) - F(\mathbf{t})).$$

More precisely,

$$\begin{aligned} \widetilde{\beta}_n^{(p)}(\mathbf{t}) &= F(\mathbf{t})^p \alpha_n(\mathbf{t}) + \sum_{k=2}^{p+1} \frac{1}{p+1} \binom{p+1}{k} F(\mathbf{t})^{p+1-k} \frac{\alpha_n(\mathbf{t})^k}{n^{(k-1)/2}} \\ &\quad + O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right), \end{aligned} \quad (8.12)$$

for each fixed $\mathbf{t}$.

Proof. By Proposition 8.7,

$$\widetilde{\mathbb{F}}_n^{(p)}(\mathbf{t}) = \frac{1}{n^{p+1}} \sum_{k=1}^{m(\mathbf{t})} k^p, \quad \widetilde{F}^{(p)}(\mathbf{t}) = \frac{F(\mathbf{t})^{p+1}}{p+1}.$$

Applying Lemma 8.11 with $m = m(\mathbf{t}) = n\mathbb{F}_n(\mathbf{t})$ gives

$$\widetilde{\mathbb{F}}_n^{(p)}(\mathbf{t}) = \frac{\mathbb{F}_n(\mathbf{t})^{p+1}}{p+1} + O_{\mathbb{P}}\left(\frac{1}{n}\right).$$

Therefore

$$\widetilde{\beta}_n^{(p)}(\mathbf{t}) = \sqrt{n} \left(\frac{\mathbb{F}_n(\mathbf{t})^{p+1} - F(\mathbf{t})^{p+1}}{p+1} \right) + O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right).$$

Now write

$$\mathbb{F}_n(\mathbf{t}) = F(\mathbf{t}) + \frac{\alpha_n(\mathbf{t})}{\sqrt{n}},$$

expand $(x + \delta)^{p+1}$ with

$$x = F(\mathbf{t}), \quad \delta = \frac{\alpha_n(\mathbf{t})}{\sqrt{n}},$$

and multiply by $\sqrt{n}/(p+1)$. This gives (8.12). Since, for fixed $\mathbf{t}$, $\alpha_n(\mathbf{t}) = O_{\mathbb{P}}(1)$, the terms with $k \geq 2$ are $O_{\mathbb{P}}(n^{-1/2})$, which proves (8.11). $\square$

Remark 8.13. The expansion in Proposition 8.12 is *not* asserted uniformly over $\mathbf{t} \in [0, 1]^d$, and it is *not* asserted for a general multivariate partial order. Any uniform or genuinely multivariate strong approximation for first-kind functionals would require additional work.

8.8 A general functional framework for the iterated multivariate integrated empirical process

In this subsection we return to the *iterated* multivariate integrated empirical process

$$\alpha_n^{(p)}(\mathbf{u}) := \sqrt{n}(\mathbb{F}_n^{(p)}(\mathbf{u}) - F^{(p)}(\mathbf{u})), \quad \mathbf{u} \in [0, 1]^d,$$

studied in the main text. Thus, unlike the preceding part of the appendix, the present subsection does *not* concern the first-kind functionals $\widetilde{F}^{(p)}$ and $\check{F}^{(p)}$. The purpose here is to formulate a clean measure-theoretic framework for the chain functional underlying $F^{(p)}$, to identify its first-order derivative in the space of signed measures, and to state the corresponding uniform Hájek projection for the empirical process. The central point is that, in the genuinely multivariate partial-order setting, the scalar identity

$$F^{(p)}(\mathbf{u}) = \frac{F(\mathbf{u})^{p+1}}{(p+1)!}$$

fails in general. Accordingly, the first-order asymptotic theory is governed not by a universal scalar multiple of the lower-orthant bridge, but by the influence function of the chain functional.

8.8.1 Measure-theoretic formulation of the chain functional

Let $\mathcal{D} := [0, 1]^d$, equipped with its Borel σ -field $\mathcal{B}(\mathcal{D})$. For a continuous distribution function F on $\mathcal{D}$, let μ_F denote the associated Borel probability measure, so that

$$F(\mathbf{u}) = \mu_F([\mathbf{0}, \mathbf{u}]), \quad \mathbf{u} \in \mathcal{D}.$$

Similarly, if $\mathbf{X}_1, \dots, \mathbf{X}_n$ are i.i.d. with distribution function F , let

$$\mu_n := \frac{1}{n} \sum_{i=1}^n \delta_{\mathbf{X}_i}$$

denote the empirical measure, so that

$$\mathbb{F}_n(\mathbf{u}) = \mu_n([\mathbf{0}, \mathbf{u}]), \quad \mathbf{u} \in \mathcal{D}.$$

Fix $p \in \mathbb{N} \cup \{0\}$, and write

$$m := p + 1.$$

For each $\mathbf{u} \in \mathcal{D}$, define the bounded measurable kernel

$$K_{p,\mathbf{u}}(\mathbf{x}_1, \dots, \mathbf{x}_m) := \mathbb{1}_{\{\mathbf{x}_m \leq \mathbf{x}_{m-1} \leq \dots \leq \mathbf{x}_1 \leq \mathbf{u}\}}, \quad (\mathbf{x}_1, \dots, \mathbf{x}_m) \in \mathcal{D}^m,$$

where all inequalities are understood componentwise. We now define the chain functional on the whole Banach space of finite signed measures.

Definition 8.14 (Chain functional on signed measures). Let $\mathcal{M}(\mathcal{D})$ denote the Banach space of finite signed Borel measures on $\mathcal{D}$, equipped with the total variation norm $\|\cdot\|_{\text{TV}}$. For every $\nu \in \mathcal{M}(\mathcal{D})$, define

$$\Phi_p(\nu)(\mathbf{u}) := \int_{\mathcal{D}^m} K_{p,\mathbf{u}}(\mathbf{x}_1, \dots, \mathbf{x}_m) d\nu(\mathbf{x}_1) \cdots d\nu(\mathbf{x}_m), \quad \mathbf{u} \in \mathcal{D}. \quad (8.13)$$

Since $|K_{p,\mathbf{u}}| \leq 1$, the integral in (8.13) is well defined for every finite signed measure ν , and

$$\|\Phi_p(\nu)\|_\infty \leq \|\nu\|_{\text{TV}}^m.$$

By construction,

$$\Phi_p(\mu_F)(\mathbf{u}) = F^{(p)}(\mathbf{u}), \quad \Phi_p(\mu_n)(\mathbf{u}) = \mathbb{F}_n^{(p)}(\mathbf{u}), \quad \mathbf{u} \in \mathcal{D}.$$

Hence

$$\alpha_n^{(p)}(\mathbf{u}) = \sqrt{n}(\Phi_p(\mu_n)(\mathbf{u}) - \Phi_p(\mu_F)(\mathbf{u})).$$

8.8.2 Fréchet differentiability on the space of signed measures

Set

$$\mathcal{M}_0(\mathcal{D}) := \{H \in \mathcal{M}(\mathcal{D}) : H(\mathcal{D}) = 0\}.$$

The zero-mass condition is natural when one studies local perturbations of a probability measure.

Proposition 8.15 (Fréchet differentiability of the chain functional). Fix $p \geq 0$, let $m = p + 1$, and let $\nu \in \mathcal{M}(\mathcal{D})$. Then the map

$$\Phi_p : \mathcal{M}(\mathcal{D}) \longrightarrow \ell^\infty(\mathcal{D})$$

is Fréchet differentiable at ν , with derivative

$$\dot{\Phi}_{p,\nu}(H)(\mathbf{u}) := \sum_{j=1}^m \int_{\mathcal{D}^m} K_{p,\mathbf{u}}(\mathbf{x}_1, \dots, \mathbf{x}_m) d\nu(\mathbf{x}_1) \cdots d\nu(\mathbf{x}_{j-1}) dH(\mathbf{x}_j) d\nu(\mathbf{x}_{j+1}) \cdots d\nu(\mathbf{x}_m), \quad (8.14)$$

for $H \in \mathcal{M}(\mathcal{D})$ and $\mathbf{u} \in \mathcal{D}$. Moreover,

$$\sup_{\mathbf{u} \in \mathcal{D}} |\Phi_p(\nu + H)(\mathbf{u}) - \Phi_p(\nu)(\mathbf{u}) - \dot{\Phi}_{p,\nu}(H)(\mathbf{u})| \leq \sum_{k=2}^m \binom{m}{k} \|\nu\|_{\text{TV}}^{m-k} \|H\|_{\text{TV}}^k. \quad (8.15)$$

In particular,

$$\frac{\|\Phi_p(\nu + H) - \Phi_p(\nu) - \dot{\Phi}_{p,\nu}(H)\|_{\infty}}{\|H\|_{\text{TV}}} \longrightarrow 0 \quad \text{as } \|H\|_{\text{TV}} \rightarrow 0.$$

Proof. Since Φ_p is a polynomial map of degree m , we expand the m -fold product measure:

$$(\nu + H)^{\otimes m} = \bigotimes_{j=1}^m (\nu + H) = \nu^{\otimes m} + \sum_{j=1}^m \nu^{\otimes(j-1)} \otimes H \otimes \nu^{\otimes(m-j)} + \mathcal{R}(\nu, H),$$

where

$$\mathcal{R}(\nu, H) = \sum_{k=2}^m \sum_{\substack{A \subset \{1, \dots, m\} \\ |A|=k}} \bigotimes_{j=1}^m M_j^{(A)}, \quad M_j^{(A)} = \begin{cases} H, & j \in A, \\ \nu, & j \notin A. \end{cases}$$

Integrating the bounded kernel $K_{p,\mathbf{u}}$ against this decomposition yields

$$\Phi_p(\nu + H)(\mathbf{u}) = \Phi_p(\nu)(\mathbf{u}) + \dot{\Phi}_{p,\nu}(H)(\mathbf{u}) + R_{p,\nu}(H)(\mathbf{u}),$$

where

$$R_{p,\nu}(H)(\mathbf{u}) = \int_{\mathcal{D}^m} K_{p,\mathbf{u}} d\mathcal{R}(\nu, H).$$

Since $|K_{p,\mathbf{u}}| \leq 1$,

$$\sup_{\mathbf{u} \in \mathcal{D}} |R_{p,\nu}(H)(\mathbf{u})| \leq \|\mathcal{R}(\nu, H)\|_{\text{TV}}.$$

Now each tensor product measure appearing in $\mathcal{R}(\nu, H)$ has total variation norm at most $\|\nu\|_{\text{TV}}^{m-k} \|H\|_{\text{TV}}^k$. Summing over all subsets $A \subset \{1, \dots, m\}$ with $|A| = k$ gives (8.15). The final assertion follows immediately. $\square$

Remark 8.16. The extension of Φ_p to all finite signed measures is essential. Indeed, if one defined Φ_p only on probability measures, then for a generic direction $H \in \mathcal{M}_0(\mathcal{D})$ there need not exist any nonzero t such that $\mu_F + tH$ remains a probability measure. Thus the natural linearization must be carried out in the ambient Banach space $\mathcal{M}(\mathcal{D})$, not merely inside the convex set of probabilities.

8.8.3 Influence function

We now specialize Proposition 8.15 to $\nu = \mu_F$. Set

$$\theta_{p,F}(\mathbf{u}) := \Phi_p(\mu_F)(\mathbf{u}) = F^{(p)}(\mathbf{u}), \quad \mathbf{u} \in \mathcal{D}.$$

For $j \in \{1, \dots, m\}$, define

$$h_{j,p,F}(\mathbf{x}, \mathbf{u}) := \int_{\mathcal{D}^{m-1}} K_{p,\mathbf{u}}(\mathbf{y}_1, \dots, \mathbf{y}_{j-1}, \mathbf{x}, \mathbf{y}_j, \dots, \mathbf{y}_{m-1}) d\mu_F(\mathbf{y}_1) \cdots d\mu_F(\mathbf{y}_{m-1}).$$

Thus $h_{j,p,F}(\mathbf{x}, \mathbf{u})$ is the chain probability obtained by fixing the j -th coordinate at $\mathbf{x}$ and integrating the remaining $m-1$ coordinates against μ_F .

Definition 8.17 (Influence function). The influence function of the chain functional at F is defined by

$$\psi_{p,F}(\mathbf{x}, \mathbf{u}) := \dot{\Phi}_{p,\mu_F}(\delta_{\mathbf{x}} - \mu_F)(\mathbf{u}) = \sum_{j=1}^m h_{j,p,F}(\mathbf{x}, \mathbf{u}) - m \theta_{p,F}(\mathbf{u}), \quad \mathbf{x}, \mathbf{u} \in \mathcal{D}. \quad (8.16)$$

Proposition 8.18 (Basic properties of $\psi_{p,F}$). For every $\mathbf{u} \in \mathcal{D}$,

$$\int_{\mathcal{D}} \psi_{p,F}(\mathbf{x}, \mathbf{u}) d\mu_F(\mathbf{x}) = 0. \quad (8.17)$$

Moreover, for every $H \in \mathcal{M}_0(\mathcal{D})$,

$$\dot{\Phi}_{p,\mu_F}(H)(\mathbf{u}) = \int_{\mathcal{D}} \psi_{p,F}(\mathbf{x}, \mathbf{u}) dH(\mathbf{x}), \quad \mathbf{u} \in \mathcal{D}. \quad (8.18)$$

Proof. Fix $\mathbf{u} \in \mathcal{D}$. For each $j \in \{1, \dots, m\}$, Fubini's theorem yields

$$\int_{\mathcal{D}} h_{j,p,F}(\mathbf{x}, \mathbf{u}) d\mu_F(\mathbf{x}) = \int_{\mathcal{D}^m} K_{p,\mathbf{u}}(\mathbf{x}_1, \dots, \mathbf{x}_m) d\mu_F(\mathbf{x}_1) \cdots d\mu_F(\mathbf{x}_m) = \theta_{p,F}(\mathbf{u}).$$

Summing over $j = 1, \dots, m$ proves (8.17). Next, by (8.14),

$$\dot{\Phi}_{p,\mu_F}(H)(\mathbf{u}) = \sum_{j=1}^m \int_{\mathcal{D}} h_{j,p,F}(\mathbf{x}, \mathbf{u}) dH(\mathbf{x}).$$

Since $H(\mathcal{D}) = 0$, we may subtract $m \theta_{p,F}(\mathbf{u})$ inside the integrand without changing the value, and (8.18) follows from (8.16). $\square$

8.8.4 Uniform Hájek projection

We now return to the empirical integrated process. Recall that

$$\mathbb{F}_n^{(p)}(\mathbf{u}) = \Phi_p(\mu_n)(\mathbf{u}), \quad \alpha_n^{(p)}(\mathbf{u}) = \sqrt{n}(\mathbb{F}_n^{(p)}(\mathbf{u}) - F^{(p)}(\mathbf{u})).$$

Equivalently,

$$\mathbb{F}_n^{(p)}(\mathbf{u}) = \frac{1}{n^m} \sum_{i_1, \dots, i_m=1}^n K_{p,\mathbf{u}}(\mathbf{X}_{i_1}, \dots, \mathbf{X}_{i_m}), \quad \mathbf{u} \in \mathcal{D}.$$

This is a V -statistic process indexed by $\mathbf{u} \in \mathcal{D}$. Define the associated U -statistic process

$$U_{n,p}(\mathbf{u}) := \frac{1}{(n)_m} \sum_{\substack{i_1, \dots, i_m=1 \\ \text{all distinct}}}^n K_{p,\mathbf{u}}(\mathbf{X}_{i_1}, \dots, \mathbf{X}_{i_m}), \quad (n)_m := n(n-1) \cdots (n-m+1).$$

Theorem 8.19 (Uniform Hájek projection). Assume that F is continuous. Then, as $n \rightarrow \infty$,

$$\sup_{\mathbf{u} \in \mathcal{D}} \left| \alpha_n^{(p)}(\mathbf{u}) - \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_{p,F}(\mathbf{X}_i, \mathbf{u}) \right| \xrightarrow{\mathbb{P}} 0. \quad (8.19)$$

Proof. We split the proof into three steps.

Step 1: comparison of the V - and U -statistic processes. Since $0 \leq K_{p,\mathbf{u}} \leq 1$,

$$\sup_{\mathbf{u} \in \mathcal{D}} |\mathbb{F}_n^{(p)}(\mathbf{u}) - U_{n,p}(\mathbf{u})| \leq 1 - \frac{(n)_m}{n^m} = O\left(\frac{1}{n}\right).$$

Therefore

$$\sup_{\mathbf{u} \in \mathcal{D}} \sqrt{n} |\mathbb{F}_n^{(p)}(\mathbf{u}) - U_{n,p}(\mathbf{u})| \longrightarrow 0. \quad (8.20)$$

Step 2: first Hoeffding projection. Let

$$\bar{K}_{p,\mathbf{u}}(\mathbf{x}_1, \dots, \mathbf{x}_m) := \frac{1}{m!} \sum_{\sigma \in \mathfrak{S}_m} K_{p,\mathbf{u}}(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(m)})$$

be the symmetrized kernel. Then $U_{n,p}(\mathbf{u})$ is the ordinary U -statistic with symmetric kernel $\bar{K}_{p,\mathbf{u}}$. Its first Hoeffding projection is

$$g_{p,F}(\mathbf{x}, \mathbf{u}) := \int_{\mathcal{D}^{m-1}} \bar{K}_{p,\mathbf{u}}(\mathbf{x}, \mathbf{y}_2, \dots, \mathbf{y}_m) d\mu_F(\mathbf{y}_2) \cdots d\mu_F(\mathbf{y}_m) - \theta_{p,F}(\mathbf{u}).$$

By symmetry,

$$m g_{p,F}(\mathbf{x}, \mathbf{u}) = \sum_{j=1}^m h_{j,p,F}(\mathbf{x}, \mathbf{u}) - m \theta_{p,F}(\mathbf{u}) = \psi_{p,F}(\mathbf{x}, \mathbf{u}).$$

Hence the first-order Hoeffding term is exactly

$$\frac{m}{n} \sum_{i=1}^n g_{p,F}(\mathbf{X}_i, \mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \psi_{p,F}(\mathbf{X}_i, \mathbf{u}). \quad (8.21)$$

Step 3: control of the degenerate remainder. Let

$$\mathcal{K}_p := \{K_{p,\mathbf{u}} : \mathbf{u} \in \mathcal{D}\}.$$

Since p and d are fixed, $\mathcal{K}_p$ is a bounded VC-subgraph class on $\mathcal{D}^m$: each $K_{p,\mathbf{u}}$ is the indicator of a set determined by finitely many coordinatewise inequalities of the form

$$x_{r,\ell} \leq x_{s,\ell}, \quad x_{1,\ell} \leq u_\ell, \quad 1 \leq r, s \leq m, \quad 1 \leq \ell \leq d.$$

Consequently, the symmetrized class

$$\bar{\mathcal{K}}_p := \{\bar{K}_{p,\mathbf{u}} : \mathbf{u} \in \mathcal{D}\}$$

is a bounded VC-type class as well. The Hoeffding decomposition of the U -process $\{U_{n,p}(\mathbf{u}) - \theta_{p,F}(\mathbf{u}) : \mathbf{u} \in \mathcal{D}\}$ therefore takes the form

$$U_{n,p}(\mathbf{u}) - \theta_{p,F}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \psi_{p,F}(\mathbf{X}_i, \mathbf{u}) + \sum_{r=2}^m \binom{m}{r} U_{n,r}^\circ(\mathbf{u}),$$

where, for each $r \geq 2$, $U_{n,r}^\circ$ is the canonical U -process associated with the r -th degenerate Hoeffding projection of $\bar{K}_{p,\mathbf{u}}$. Standard maximal inequalities for canonical U -processes indexed by bounded VC-type classes imply that, for each $r = 2, \dots, m$,

$$\sup_{\mathbf{u} \in \mathcal{D}} |U_{n,r}^\circ(\mathbf{u})| = O_{\mathbb{P}}(n^{-r/2});$$

see, for instance, the maximal inequalities for degenerate U -processes in [5, 31, 46]. Since $r \geq 2$, it follows that

$$\sup_{\mathbf{u} \in \mathcal{D}} \left| U_{n,p}(\mathbf{u}) - \theta_{p,F}(\mathbf{u}) - \frac{1}{n} \sum_{i=1}^n \psi_{p,F}(\mathbf{X}_i, \mathbf{u}) \right| = O_{\mathbb{P}}\left(\frac{1}{n}\right). \quad (8.22)$$

Multiplying by $\sqrt{n}$, and combining (8.20), (8.21), and (8.22), yields (8.19). $\square$

Remark 8.20. Theorem 8.19 is the fundamental first-order statement. It shows that, in the genuinely multivariate setting, the iterated integrated empirical process is asymptotically equivalent to the empirical process indexed by the influence class

$$\Psi_{p,F} := \{\psi_{p,F}(\cdot, \mathbf{u}) : \mathbf{u} \in \mathcal{D}\}.$$

Thus the correct Gaussian limit is determined by $\psi_{p,F}$, not by a universal scalar multiple of the ordinary lower-orthant bridge.

8.8.5 Weak convergence

The preceding projection theorem immediately yields the following consequence.

Corollary 8.21 (Weak convergence under a Donsker condition). *Assume that the class*

$$\Psi_{p,F} := \{\psi_{p,F}(\cdot, \mathbf{u}) : \mathbf{u} \in \mathcal{D}\}$$

is μ_F -Donsker. Then

$$\alpha_n^{(p)} = \sqrt{n}(\mathbb{F}_n^{(p)} - F^{(p)}) \rightsquigarrow \mathbb{G}_{p,F} \quad \text{in } \ell^\infty(\mathcal{D}),$$

where $\mathbb{G}_{p,F}$ is a centered Gaussian process with covariance

$$\text{Cov}(\mathbb{G}_{p,F}(\mathbf{u}), \mathbb{G}_{p,F}(\mathbf{v})) = \int_{\mathcal{D}} \psi_{p,F}(\mathbf{x}, \mathbf{u}) \psi_{p,F}(\mathbf{x}, \mathbf{v}) d\mu_F(\mathbf{x}), \quad \mathbf{u}, \mathbf{v} \in \mathcal{D}. \quad (8.23)$$

Proof. By Theorem 8.19,

$$\alpha_n^{(p)}(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_{p,F}(\mathbf{X}_i, \mathbf{u}) + o_{\mathbb{P}}(1) \quad \text{uniformly in } \mathbf{u} \in \mathcal{D}.$$

If $\Psi_{p,F}$ is μ_F -Donsker, then the empirical process indexed by $\Psi_{p,F}$ converges weakly in $\ell^\infty(\mathcal{D})$ to a centered Gaussian limit $\mathbb{G}_{p,F}$. The covariance formula (8.23) is the usual covariance of the limiting empirical process. $\square$

8.8.6 Specialization to the scalarizable regime

The framework above contains the totally ordered case as a special instance. Assume that the support of μ_F is totally ordered by the componentwise order $\leq$. Then the chain functional reduces to the scalar formula

$$F^{(p)}(\mathbf{u}) = \frac{F(\mathbf{u})^{p+1}}{(p+1)!}, \quad \mathbf{u} \in \mathcal{D}.$$

Accordingly, the influence function collapses to the familiar weighted lower-orthant bridge.

Proposition 8.22 (Influence function in the scalarizable regime). *Assume that the support of μ_F is totally ordered by the componentwise order. Then, for every $\mathbf{x}, \mathbf{u} \in \mathcal{D}$,*

$$\psi_{p,F}(\mathbf{x}, \mathbf{u}) = \frac{F(\mathbf{u})^p}{p!} (\mathbb{1}_{\{\mathbf{x} \leq \mathbf{u}\}} - F(\mathbf{u})). \quad (8.24)$$

Proof. In the totally ordered regime,

$$\theta_{p,F}(\mathbf{u}) = F^{(p)}(\mathbf{u}) = \frac{F(\mathbf{u})^{p+1}}{(p+1)!}.$$

Let $H \in \mathcal{M}_0(\mathcal{D})$. Since

$$F_H(\mathbf{u}) := H([\mathbf{0}, \mathbf{u}]) = \int_{\mathcal{D}} \mathbb{1}_{\{\mathbf{x} \leq \mathbf{u}\}} dH(\mathbf{x}),$$

the Fréchet derivative of the scalar map

$$x \mapsto \frac{x^{p+1}}{(p+1)!}$$

at $x = F(\mathbf{u})$ yields

$$\dot{\Phi}_{p, \mu_F}(H)(\mathbf{u}) = \frac{F(\mathbf{u})^p}{p!} F_H(\mathbf{u}) = \int_{\mathcal{D}} \frac{F(\mathbf{u})^p}{p!} \mathbb{1}_{\{\mathbf{x} \leq \mathbf{u}\}} dH(\mathbf{x}).$$

Because $H(\mathcal{D}) = 0$, we may subtract $\frac{F(\mathbf{u})^{p+1}}{p!}$ inside the integrand without changing the value, and therefore

$$\dot{\Phi}_{p, \mu_F}(H)(\mathbf{u}) = \int_{\mathcal{D}} \frac{F(\mathbf{u})^p}{p!} (\mathbb{1}_{\{\mathbf{x} \leq \mathbf{u}\}} - F(\mathbf{u})) dH(\mathbf{x}).$$

By the uniqueness of the influence representation (8.18), this proves (8.24). $\square$

Remark 8.23. Formula (8.24) is a structural simplification specific to the scalarizable regime. In the genuinely multivariate case, where incomparable support points occur with positive probability, the first-order theory is governed by the full influence function $\psi_{p,F}$, not by the scalar factor $F(\mathbf{u})^p/p!$.

Acknowledgements

We extend our warmest thanks to the Editors-in-Chief of the Journal of Advanced Analysis and Applications for giving us the opportunity and privilege to present this work in this new journal.

Conflict of interest

The author declares no conflict of interest.

References

- [1] R.J. Adler, *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes*, Institute of Mathematical Statistics Lecture Notes–Monograph Series, Vol. 12, Institute of Mathematical Statistics, Hayward, CA, 1990, 160 pp.
- [2] S. Alvarez-Andrade, S. Bouzebda, Some nonparametric tests for change-point detection based on the $\mathbb{P}$ - $\mathbb{P}$ and $\mathbb{Q}$ - $\mathbb{Q}$ plot processes, *Sequential Anal.* **33** (2014), 360–399. DOI: [10.1080/07474946.2014.916930](https://doi.org/10.1080/07474946.2014.916930)
- [3] S. Alvarez-Andrade, S. Bouzebda, A. Lachal, Some asymptotic results for the integrated empirical process with applications to statistical tests, *Comm. Statist. Theory Methods* **46** (2017), 3365–3392. DOI: [10.1080/03610926.2015.1060346](https://doi.org/10.1080/03610926.2015.1060346)
- [4] S. Alvarez-Andrade, S. Bouzebda, A. Lachal, Strong approximations for the p -fold integrated empirical process with applications to statistical tests, *TEST* **27** (2018), 826–849. DOI: [10.1007/s11749-017-0572-0](https://doi.org/10.1007/s11749-017-0572-0)
- [5] M.A. Arcones, E. Giné, Limit theorems for U -processes, *Ann. Probab.* **21** (1993), 1494–1542. DOI: [10.1214/aop/1176989128](https://doi.org/10.1214/aop/1176989128)

- [6] A. Aue, L. Horváth, Structural breaks in time series, *J. Time Series Anal.* **34** (2013), 1–16. DOI: [10.1111/j.1467-9892.2012.00819.x](https://doi.org/10.1111/j.1467-9892.2012.00819.x)
- [7] P.J. Bickel, M.J. Wichura, Convergence criteria for multiparameter stochastic processes and some applications, *Ann. Math. Statist.* **42** (1971), 1656–1670. DOI: [10.1214/aoms/1177693164](https://doi.org/10.1214/aoms/1177693164)
- [8] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York–London–Sydney, 1968, 253 pp.
- [9] I.S. Borisov, Approximation of empirical fields that are constructed from vector observations with dependent coordinates, *Siberian Math. J.* **23** (1982), 615–623. DOI: [10.1007/BF00971279](https://doi.org/10.1007/BF00971279)
- [10] S. Bouzebda, On the strong approximation of bootstrapped empirical copula processes with applications, *Math. Methods Statist.* **21** (2012), 153–188. DOI: [10.3103/S1066530712030015](https://doi.org/10.3103/S1066530712030015)
- [11] S. Bouzebda, Asymptotic properties of pseudo maximum likelihood estimators and test in semi-parametric copula models with multiple change points, *Math. Methods Statist.* **23** (2014), 38–65. DOI: [10.3103/S1066530714010037](https://doi.org/10.3103/S1066530714010037)
- [12] S. Bouzebda, Kac’s representation for empirical copula process from an asymptotic viewpoint, *Statist. Probab. Lett.* **123** (2017), 107–113. DOI: [10.1016/j.spl.2016.12.006](https://doi.org/10.1016/j.spl.2016.12.006)
- [13] S. Bouzebda, On the weak convergence and the uniform-in-bandwidth consistency of the general conditional U -processes based on the copula representation: multivariate setting, *Hacet. J. Math. Stat.* **52** (2023), 1303–1348. DOI: [10.15672/hujms.1134334](https://doi.org/10.15672/hujms.1134334)
- [14] S. Bouzebda, N.-E. El Faouzi, New two-sample tests based on the integrated empirical copula processes, *Statistics* **46** (2012), 313–324. DOI: [10.1080/02331888.2010.512976](https://doi.org/10.1080/02331888.2010.512976)
- [15] S. Bouzebda, A.A. Ferfache, Asymptotic properties of M -estimators based on estimating equations and censored data in semi-parametric models with multiple change points, *J. Math. Anal. Appl.* **497** (2021), Paper No. 124883, 44 pp. DOI: [10.1016/j.jmaa.2020.124883](https://doi.org/10.1016/j.jmaa.2020.124883)
- [16] S. Bouzebda, A.A. Ferfache, Asymptotic properties of semiparametric M -estimators with multiple change points, *Phys. A* **609** (2023), Paper No. 128363, 29 pp. DOI: [10.1016/j.physa.2022.128363](https://doi.org/10.1016/j.physa.2022.128363)
- [17] S. Bouzebda, A. Keziou, T. Zari, K -sample problem using strong approximations of empirical copula processes, *Math. Methods Statist.* **20** (2011), 14–29. DOI: [10.3103/S1066530711010029](https://doi.org/10.3103/S1066530711010029)
- [18] B.E. Brodsky, B.S. Darkhovsky, *Nonparametric Methods in Change-Point Problems*, Mathematics and its Applications, Vol. 243, Kluwer Academic Publishers, Dordrecht, 1993, 209 pp. DOI: [10.1007/978-94-015-8163-9](https://doi.org/10.1007/978-94-015-8163-9)
- [19] S. Bouzebda, T. Zari, Asymptotic behavior of weighted multivariate Cramér–von Mises-type statistics under contiguous alternatives, *Math. Methods Statist.* **22** (2013), 226–252. DOI: [10.3103/S1066530713030046](https://doi.org/10.3103/S1066530713030046)
- [20] M.D. Burke, M. Csörgő, S. Csörgő, P. Révész, Approximations of the empirical process when parameters are estimated, *Ann. Probab.* **7** (1979), 790–810. DOI: [10.1214/aop/1176994939](https://doi.org/10.1214/aop/1176994939)
- [21] J. Chan, L. Horváth, M. Hušková, Darling–Erdős limit results for change-point detection in panel data, *J. Statist. Plann. Inference* **143** (2013), 955–970. DOI: [10.1016/j.jspi.2012.11.004](https://doi.org/10.1016/j.jspi.2012.11.004)
- [22] J. Chen, A.K. Gupta, *Parametric Statistical Change Point Analysis*, Birkhäuser, Boston, MA, 2000, 184 pp. DOI: [10.1007/978-1-4757-3131-6](https://doi.org/10.1007/978-1-4757-3131-6)
- [23] M. Csörgő, A glimpse of the KMT (1975) approximation of empirical processes by Brownian bridges via quantiles, *Acta Sci. Math. (Szeged)* **73** (2007), 349–366.
- [24] M. Csörgő, L. Horváth, A note on strong approximations of multivariate empirical processes, *Stochastic Process. Appl.* **28** (1988), 101–109. DOI: [10.1016/0304-4149\(88\)90068-3](https://doi.org/10.1016/0304-4149(88)90068-3)
- [25] M. Csörgő, L. Horváth, *Weighted Approximations in Probability and Statistics*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Chichester, 1993, 442 pp.
- [26] M. Csörgő, L. Horváth, *Limit Theorems in Change-Point Analysis*, Wiley Series in Probability

- and Statistics, John Wiley & Sons, Chichester, 1997, 414 pp.
- [27] M. Csörgő, P. Révész, *Strong Approximations in Probability and Statistics*, Probability and Mathematical Statistics, Academic Press, New York–London, 1981, 284 pp.
 - [28] S. Csörgő, P. Hall, The Komlós–Major–Tusnády approximations and their applications, *Austral. J. Statist.* **26** (1984), 189–218. DOI: [10.1111/j.1467-842X.1984.tb01233.x](https://doi.org/10.1111/j.1467-842X.1984.tb01233.x)
 - [29] A. DasGupta, *Asymptotic Theory of Statistics and Probability*, Springer Texts in Statistics, Springer, New York, 2008, 722 pp.
 - [30] J. Durbin, Weak convergence of the sample distribution function when parameters are estimated, *Ann. Statist.* **1** (1973), 279–290. DOI: [10.1214/aos/1176342365](https://doi.org/10.1214/aos/1176342365)
 - [31] E. Giné, R. Latała, J. Zinn, Exponential and moment inequalities for U -statistics, in: *High Dimensional Probability II*, Progr. Probab., Vol. 47, Birkhäuser, Boston, MA, 2000, 13–38.
 - [32] N. Henze, Ya.Yu. Nikitin, A new approach to goodness-of-fit testing based on the integrated empirical process, *J. Nonparametr. Statist.* **12** (2000), 391–416. DOI: [10.1080/10485250008832815](https://doi.org/10.1080/10485250008832815)
 - [33] N. Henze, Ya.Yu. Nikitin, Watson-type goodness-of-fit tests based on the integrated empirical process, *Math. Methods Statist.* **11** (2002), 183–202.
 - [34] N. Henze, Ya.Yu. Nikitin, Two-sample tests based on the integrated empirical process, *Comm. Statist. Theory Methods* **32** (2003), 1767–1788. DOI: [10.1081/STA-120022708](https://doi.org/10.1081/STA-120022708)
 - [35] L. Horváth, G. Rice, Extensions of some classical methods in change point analysis, *TEST* **23** (2014), 219–255. DOI: [10.1007/s11749-014-0368-4](https://doi.org/10.1007/s11749-014-0368-4)
 - [36] P. Jing, J. Wang, Testing the equality of multivariate distributions using the bootstrap and integrated empirical processes, *Comm. Statist. Theory Methods* **35** (2006), 661–670. DOI: [10.1080/03610920500498816](https://doi.org/10.1080/03610920500498816)
 - [37] P. Jing, Y. Yang, Testing the equality of multivariate distributions using the integrated empirical processes, *Math. Appl. (Wuhan)* **20** (2007), 614–620.
 - [38] J. Komlós, P. Major, G. Tusnády, An approximation of partial sums of independent RV's and the sample DF. I, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **32** (1975), 111–131. DOI: [10.1007/BF00533093](https://doi.org/10.1007/BF00533093)
 - [39] A. Lachal, Study of some new integrated statistics: computation of Bahadur efficiency, relation with non-standard boundary value problems, *Math. Methods Statist.* **10** (2001), 73–104.
 - [40] D.M. Mason, H.H. Zhou, Quantile coupling inequalities and their applications, *Probab. Surv.* **9** (2012), 439–479. DOI: [10.1214/12-PS198](https://doi.org/10.1214/12-PS198)
 - [41] G. Neuhaus, On weak convergence of stochastic processes with multidimensional time parameter, *Ann. Math. Statist.* **42** (1971), 1285–1295. DOI: [10.1214/aoms/1177693241](https://doi.org/10.1214/aoms/1177693241)
 - [42] F.H. Ruymgaart, Asymptotic normality of nonparametric tests for independence, *Ann. Statist.* **2** (1974), 892–910. DOI: [10.1214/aos/1176342812](https://doi.org/10.1214/aos/1176342812)
 - [43] F.H. Ruymgaart, G.R. Shorack, W.R. van Zwet, Asymptotic normality of nonparametric tests for independence, *Ann. Math. Statist.* **43** (1972), 1122–1135. DOI: [10.1214/aoms/1177692465](https://doi.org/10.1214/aoms/1177692465)
 - [44] G.R. Shorack, J.A. Wellner, *Empirical Processes with Applications to Statistics*, Classics in Applied Mathematics, Vol. 59, SIAM, Philadelphia, PA, 2009, 956 pp. DOI: [10.1137/1.9780898719017](https://doi.org/10.1137/1.9780898719017)
 - [45] B. Szyszkowicz, Asymptotic distributions of weighted pontograms under contiguous alternatives, *Math. Proc. Cambridge Philos. Soc.* **112** (1992), 431–447. DOI: [10.1017/S0305004100071097](https://doi.org/10.1017/S0305004100071097)
 - [46] A.W. van der Vaart, J.A. Wellner, *Weak Convergence and Empirical Processes: With Applications to Statistics*, 2nd ed., Springer Series in Statistics, Springer, Cham, 2023, 679 pp. DOI: [10.1007/978-3-031-29040-4](https://doi.org/10.1007/978-3-031-29040-4)

Submitted: April 2, 2026

Accepted: June 6, 2026

Published: September 28, 2026

© 2026 The Author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. See <http://creativecommons.org/licenses/by/4.0/>.