Electronic Companion of "Ensemble Copula Coupling for Multivariate Input Modeling and Uncertainty Quantification"

This electronic companion presents additional background and technical materials, as well as the proofs for all the theorems in the main body of the paper. For simplicity, for $a, b > 0$, we write $a \lesssim b$ if there exists $k > 0$ such that $a \le kb$ throughout this electronic companion.

Appendix EC.1: Bounded Variation and Integration by Parts

EC.1.1. Notation and Definitions In this section, we review multivariate bounded variations. Our main references are Owen [6], Aistleitner and Dick [1], and Radulović et al. [7]. The following notation is borrowed from Radulović et al. [7].

- For any vector $\mathbf{x} \in \mathbb{R}^d$, x_j is its j-th component. We define a hypercube $[\mathbf{x}, \mathbf{y}] := \prod_{i=1}^d [x_i, y_i]$.
- We write an index set $\{1, 2, \cdots, d\}$ as $1:d$ in short.
- For $I \subset 1:d$, $|I|$ means the cardinality of I, and $-I = 1:d - I$.
- If a collection of set $\{I_k\}_{1\le k\le K}$ is a partition of $1:d$, we write it as $I_1 + \cdots + I_K = 1:d$.
- For $I \subset 1:d$, $\mathbf{x}_I$ is a projection of $\mathbf{x}$ on the coordinates in I. That is, $\mathbf{x}_I = (x_i)_{i\in I}$.
- Let $I + J = 1:d$. Then, $\mathbf{x}_I : \mathbf{y}_J$ is a vector whose coordinates are equal to x_i for $i \in I$ and y_i for $i \in J$. For $I_1 + I_2 + I_3 = 1:d$, the vector $\mathbf{x}_{I_1} : \mathbf{y}_{I_2} : \mathbf{z}_{I_3}$ is similarly defined.
- In the next section, we often use $f(\mathbf{x}_{I_1} : \mathbf{a}_{I_2} : \mathbf{b}_{I_3})$ for a function defined on $[\mathbf{a}, \mathbf{b}]$. We abbreviate $f(\mathbf{x}_{I_1} : \mathbf{a}_{I_2} : \mathbf{b}_{I_3})$ as $f_{\mathcal{I}}(\mathbf{x}_{I_1})$ where $\mathcal{I} = \{I_1, I_2, I_3\}$ is a partition of $1:d$.

Definition EC.1.

1. A *ladder* $\mathcal{Y}$ on the interval $[a, b]$ is a set of finitely many elements in $[a, b)$. For each y in $\mathcal{Y}$, its *successor* is defined by $y^+ := \min\{(y, b) \cap \mathcal{Y}\}$. If the set is empty, we set $y^+ = b$. In a higher dimension, a ladder $\mathcal{Y}$ on $[\mathbf{a}, \mathbf{b}] \subset \mathbb{R}^d$ is of the form $\prod_{i=1}^d \mathcal{Y}_i = \{(x_1, \cdots, x_d) | x_i \in \mathcal{Y}_i\}$ where $\mathcal{Y}_i$ is a ladder on $[a_i, b_i]$.
2. The *d-fold alternating sum of f* or *generalized volume of the hypercube based on measure* Δ_f over $[\mathbf{a}, \mathbf{b}]$ is given by $\Delta_f((\mathbf{a}, \mathbf{b}]) = \sum_{I\subset 1:d} (-1)^{|I|} f(\mathbf{a}_I : \mathbf{b}_{-I})$.

Some additional definitions follow.

- *Variation of a function over a ladder* $\mathcal{Y}$ is $V_{\mathcal{Y}}(f) = \sum_{\mathbf{y}\in\mathcal{Y}} |\Delta_f((\mathbf{y}, \mathbf{y}^+])|$;
- *Total variation in the sense of Vitali* is $V(f) := \sup_{\mathcal{Y}} V_{\mathcal{Y}}(f)$;
- *Hardy-Krause variation* is $V_{\text{HK}}(f) = \sum_{I\subset 1:d, I\neq\emptyset} V(f(\cdot : \mathbf{b}_{-I}))$;
- *Hardy-Krause variation anchored at* $\mathbf{a}$ is $V_{\text{HK}\mathbf{a}}(f) = \sum_{I\subset 1:d, I\neq\emptyset} V(f(\cdot : \mathbf{a}_{-I}))$.

Functions with bounded Hardy-Krause variation are simply called BVHK. We note that V_{HK} is the sum of total variations of the function f projected by all possible combinations of coordinates. A useful relationship between V_{HK} and $V_{\mathrm{HK}\mathbf{a}}$ is from Owen [6]:

$$V_{\mathrm{HK}}(f) \le \left(2^d - 1\right) V_{\mathrm{HK}\mathbf{a}}(f) \le \left(2^d - 1\right)^2 V_{\mathrm{HK}}(f).$$

The same reference also explains that if $f(\cdot;\mathbf{b}_{-I})$ has continuous partial derivatives $\frac{\partial^d}{\partial x_1 \cdots \partial x_d} f(\cdot;\mathbf{b}_{-I})$ for all $I \subset 1:d$, then f is BVHK. Another important function class which is BVHK is simple functions of the form $\sum_i c_i \mathbf{1}_{[\mathbf{a}_i,\mathbf{b}_i]}$.

One interesting fact related to BVHK is that it is possible to decompose a BVHK function into the sum of two completely monotone functions, similarly to that a function of bounded variation in one dimension is equal to the sum of two monotone functions.

THEOREM EC.1 **(Aistleitner and Dick [1]).** *Let f be a right-continuous and BVHK function on $[\mathbf{a},\mathbf{b}]$. Then, there is a unique signed measure df on $[\mathbf{a},\mathbf{b}]$ such that $f(\mathbf{x}) = df([\mathbf{a},\mathbf{x}])$ for all $\mathbf{x} \in [\mathbf{a},\mathbf{b}]$. Furthermore, there are two completely monotone functions $f^{\pm}$ such that $f(\mathbf{x}) = f(\mathbf{a}) + f^{+}(\mathbf{x}) - f^{-}(\mathbf{x})$. Equivalently, we have $df = df^{+} - df^{-}$ where $df^{\pm}$ are positive measures corresponding to $f^{\pm}$, respectively.*

EC.1.2. Multivariate Integration by Parts

THEOREM EC.2 **(Radulović et al. [7]).** *For $\mathbf{a},\mathbf{b} \in \mathbb{R}^d$ and functions $f,g : [\mathbf{a},\mathbf{b}] \to \mathbb{R}$, we assume f and g are right-continuous and BVHK. Then, the following formula holds:*

$$\begin{aligned}\int_{(\mathbf{a},\mathbf{b}]} f(\mathbf{x})dg(\mathbf{x}) &= \sum\nolimits_{I_1+I_2+I_3=1:d} (-1)^{|I_1|+|I_2|} \int_{(\mathbf{a}_{I_1},\mathbf{b}_{I_1}]} g(\mathbf{x}_{I_1}- : \mathbf{a}_{I_2} : \mathbf{b}_{I_3}) df(\mathbf{x}_{I_1} : \mathbf{a}_{I_2} : \mathbf{b}_{I_3}) \\ &= \sum\nolimits_{I_1+I_2+I_3=1:d} (-1)^{|I_1|+|I_2|} \int_{(\mathbf{a}_{I_1},\mathbf{b}_{I_1}]} g_{\mathcal{I}}(\mathbf{x}_{I_1}-) df_{\mathcal{I}}(\mathbf{x}_{I_1})\end{aligned}$$

where $g_{\mathcal{I}}(\mathbf{x}_{I_1}-) = \lim_{\mathbf{y}_{I_1} \uparrow \mathbf{x}_{I_1}} g_{\mathcal{I}}(\mathbf{y}_{I_1})$ and $g_{\mathcal{I}}(\mathbf{y}_{I_1}) = g(\mathbf{y}_{I_1} : \mathbf{a}_{I_2} : \mathbf{b}_{I_3})$ for each partition $\mathcal{I}$.

Motivated by Theorem EC.2, we define two functionals Γ_h and $|\Gamma|_h$ by

$$\begin{aligned}\Gamma_h(p) &= \sum\nolimits_{\mathcal{I}} (-1)^{|I_1|+|I_2|} \int_{(\mathbf{a}_{I_1},\mathbf{b}_{I_1}]} p_{\mathcal{I}}(\mathbf{x}_{I_1}) dh_{\mathcal{I}}(\mathbf{x}_{I_1}), \\ |\Gamma|_h(p) &= \sum\nolimits_{\mathcal{I}} \int_{(\mathbf{a}_{I_1},\mathbf{b}_{I_1}]} p_{\mathcal{I}}(\mathbf{x}_{I_1}) |dh|_{\mathcal{I}}(\mathbf{x}_{I_1}).\end{aligned}$$

Here $|dh| = dh^{+} + dh^{-}$ is a total variation measure of dh mentioned in Theorem EC.1. Using these definitions, Theorem EC.2 can be re-written as $\int_{(\mathbf{a},\mathbf{b}]} f(\mathbf{x})dg(\mathbf{x}) = \Gamma_f(g(\cdot-))$. One can easily see that $|\Gamma_h(p)| \le |\Gamma|_h(|p|)$. The next lemma proves some additional results that are useful in our analysis in later sections.

LEMMA EC.1. *Suppose f and g are right-continuous and BVHK on $[\mathbf{a},\mathbf{b}]$. Then, we have*

1. $\left|\Gamma_f(g)\right| \lesssim \sup_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]}\left|g(\mathbf{x})\right|$,
2. $\left|\Gamma\right|_f^q(|g|) \lesssim \left|\Gamma\right|_f(|g|^q)$ *for* $q \geq 1$.

Proof. From the definitions above, we derive that

$$\left|\Gamma_f(g)\right| \leq |\Gamma|_f(|g|) = \sum\nolimits_{\mathcal{I}} \int_{(\mathbf{a}_{I_1},\mathbf{b}_{I_1}]} |g|_{\mathcal{I}}(\mathbf{x}_{I_1})|df|_{\mathcal{I}}(\mathbf{x}_{I_1}) \leq \sup\nolimits_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]}|g(\mathbf{x})| \times A$$

where $A = \sum_{\mathcal{I}} \int_{(\mathbf{a}_{I_1},\mathbf{b}_{I_1}]} |df|_{\mathcal{I}}(\mathbf{x}_{I_1})$ and $\mathcal{I} = \{I_1, I_2, I_3\}$ with $I_1 + I_2 + I_3 = 1:d$. For the first claim, it is enough to show that A is finite. In order to see this, we utilize Theorem EC.1, and Eq.(10), Eq.(47) of Radulović et al. [7] to see

$$\begin{aligned}
A &= \sum\nolimits_{\mathcal{I}, I_1\neq\emptyset} \int_{(\mathbf{a}_{I_1},\mathbf{b}_{I_1}]} |df|_{\mathcal{I}}(\mathbf{x}_{I_1}) \\
&= \sum\nolimits_{\mathcal{I}, I_1\neq\emptyset} \int_{(\mathbf{a}_{I_1},\mathbf{b}_{I_1}]} df^+_{\mathcal{I}}(\mathbf{x}_{I_1}) + df^-_{\mathcal{I}}(\mathbf{x}_{I_1}) \\
&= \sum\nolimits_{\mathcal{I}, I_1\neq\emptyset} \sum_{I_0\subset I_1} (-1)^{|I_0|}\left(f^+_{\mathcal{I}}(\mathbf{a}_{I_0}:\mathbf{b}_{I_1-I_0}) + f^-_{\mathcal{I}}(\mathbf{a}_{I_0}:\mathbf{b}_{I_1-I_0})\right) \\
&= \sum\nolimits_{\mathcal{I}, I_1\neq\emptyset} \Big(V(f^+(\cdot:\mathbf{a}_{I_2}:\mathbf{b}_{I_3})) + V(f^-(\cdot:\mathbf{a}_{I_2}:\mathbf{b}_{I_3}))\Big).
\end{aligned}$$

Here, Eq.(10) applies to the third inequality and Eq.(47) to the last one. One can then re-write the last summation as follows.

$$\begin{aligned}
&= \sum\nolimits_{I_3\subset 1:d, I_3\neq 1:d}\left[\sum\nolimits_{I_1+I_2=-I_3, I_1\neq\emptyset}\Big(V(f^+(\cdot:\mathbf{a}_{I_2}:\mathbf{b}_{I_3})) + V(f^-(\cdot:\mathbf{a}_{I_2}:\mathbf{b}_{I_3}))\Big)\right] \\
&= \sum\nolimits_{I_3\subset 1:d, I_3\neq 1:d}\Big(V_{\mathrm{HK}\mathbf{a}}(f^+(\cdot:\mathbf{b}_{I_3})) + V_{\mathrm{HK}\mathbf{a}}(f^-(\cdot:\mathbf{b}_{I_3}))\Big) \\
&= \sum\nolimits_{I_3\subset 1:d, I_3\neq 1:d} V_{\mathrm{HK}\mathbf{a}}(f(\cdot:\mathbf{b}_{I_3}))
\end{aligned}$$

where the last equality is obtained by Theorem 2 of Aistleitner and Dick [1]. Thanks to the relationship (EC.1.1) between V_{HK} and $V_{\mathrm{HK}\mathbf{a}}$,

$$V_{\mathrm{HK}\mathbf{a}}(f(\cdot:\mathbf{b}_{I_3})) \leq \left(2^{d-|I_3|}-1\right)V_{\mathrm{HK}}(f(\cdot:\mathbf{b}_{I_3})) \leq \left(2^{d-|I_3|}-1\right)V_{\mathrm{HK}}(f).$$

Then, simple calculations show that $A \leq 2^{2d}V_{\mathrm{HK}}(f) < \infty$.

The second claim can be readily verified by combining, first, $|\sum a_i|^q \lesssim \sum|a_i|^q$ and, second, Jensen's inequality for integration: $\int |g|^q d\mu \geq (\int |g| d\mu)^q$ for a measure μ, $q \geq 1$. □

Appendix EC.2: Detailed Review of Empirical Process Theory An important part of the theory is about identifying function classes over which $\mathbb{G}_n$ converges uniformly. There are two relevant concepts:

- $\mathcal{F}$ is *Glivenko-Cantelli* if $\|\mathbb{P}_n - \mathbb{P}\|_{\mathcal{F}} \to 0$ almost surely;
- $\mathcal{F}$ is *Donsker* if $\mathbb{G}_n \Longrightarrow \mathbb{G}$ in $\ell^\infty(\mathcal{F})$ for some Gaussian process $\mathbb{G}$.

The most well-known example is the class of indicator functions $\mathcal{F}_{\text{ind}} = \{f_{\mathbf{x}}(\cdot) = \mathbf{1}\{\cdot \le \mathbf{x}\} | \mathbf{x} \in \mathbb{R}^d\}$. Particularly in one dimension, the next two theorems are classical where one can easily check

$$\|\mathbb{P}_n - \mathbb{P}\|_{\mathcal{F}_{\text{ind}}} = \|F_n - F\|_\infty, \quad \mathbb{G}_n f_{\mathbf{x}} = \sqrt{n}(F_n - F)(\mathbf{x})$$

where F_n is the empirical CDF of n observations and F is the true CDF with the sup norm $\|\cdot\|_\infty$.

Theorem EC.3 (**Glivenko-Cantelli**). *If F_n is an empirical distribution with n observations, then $\left\|F_n - F\right\|_\infty \to 0$ almost surely.*

Theorem EC.4 (**Donsker**). *If F_n is an empirical distribution with n observations, then $\sqrt{n}(F_n - F)$ converges weakly to a Gaussian process.*

In the following, we present a functional delta method (Theorem 20.8 of van der Vaart [11]).

Definition EC.2. Let $\mathbb{D}$ and $\mathbb{E}$ be normed linear spaces. A map $\phi : \mathbb{D}_\phi \subset \mathbb{D} \to \mathbb{E}$ defined on a subset of $\mathbb{D}$ is *Hadamard differentiable at F tangentially to* $\mathbb{D}_0 \subset \mathbb{D}$ if there exists $\phi'_F : \mathbb{D} \to \mathbb{E}$ such that

$$\left\|t^{-1}(\phi(F + th_t) - \phi(F)) - \phi'_F(h)\right\|_{\mathbb{E}} \to 0$$

as $t \to 0$ for every h_t and h such that $F + th_t \in \mathbb{D}_\phi$ and $h_t \to h \in \mathbb{D}_0$

Theorem EC.5 (**van der Vaart [11]**). *Let $\mathbb{D}$ and $\mathbb{E}$ be normed linear spaces. Suppose that F_n's are random elements with values in $\mathbb{D}_\phi \subset \mathbb{D}$, F is nonrandom in $\mathbb{D}_\phi$, and T is a random element with values in $\mathbb{D}_0 \subset \mathbb{D}$, and that they satisfy $r_n(F_n - F) \Rightarrow T$ for some sequence of numbers $r_n \to \infty$. If $\phi : \mathbb{D}_\phi \to \mathbb{E}$ is Hadamard differentiable at F tangentially to $\mathbb{D}_0$, then $r_n(\phi(F_n) - \phi(F)) \Longrightarrow \phi'_F(T)$.*

Appendix EC.3: Technical Proofs

EC.3.1. Validity of Dirichlet Bootstrap This section justifies the resampling ratio for the Dirichlet bootstrap in Example 5. Define $L_{2,1}$ norm of a random variable ξ as $||\xi||_{2,1} := \int_0^\infty \sqrt{\mathrm{P}(|\xi| \ge t)}dt$. In order to guarantee a weak convergence of bootstrap empirical process, a resampling plan $\mathbf{W} = (W_{n1}, W_{n2}, \ldots, W_{nn})$ is enough to satisfy Eq. (3.6.8) in van der Vaart and Wellner [12] given as

$$\begin{aligned}
&\sup\nolimits_n ||W_{n1} - 1||_{2,1} < \infty, \\
&n^{-1/2}\mathrm{E}\left[\max\nolimits_{1\le i\le n} |W_{ni} - 1|\right] \to 0, \\
&n^{-1}\sum\nolimits_{i=1}^{n} (W_{ni} - 1)^2 \to \gamma^{-1} (> 0) \text{ in probability,}
\end{aligned}$$

where γ corresponds to a control parameter of a limit process in Theorem 3. The following statement shows that a hybrid of Dirichlet and subsampling without replacement in Example 5 satisfies (EC.3.1) with $\gamma = a\gamma'/(1+a(1-\gamma'))$.

Proposition EC.1. *Example 5 satisfies* (EC.3.1) *with* $\gamma = a\gamma'/(1+a(1-\gamma'))$.

Proof. When $\gamma' = 1$, Example 3.6.9 in van der Vaart and Wellner [12] states that a ratio of random variables $\frac{Y_i}{\bar{Y}}$ satisfies (EC.3.1) with $\mathrm{E}[Y]^2/\mathrm{Var}[Y] = \gamma$ if $||Y||_{2,1} < \infty$. When $Y \sim \Gamma(a, 1)$, the boundedness comes from the inequality $||\xi||_{2,1} \le \frac{1}{2}||\xi||_4$ given in Exercise 2.9.1 of van der Vaart and Wellner [12] and we further have $\gamma = a$. So, it is enough to consider the case when $\gamma' < 1$. Note that the ratio vector $\left(\frac{Y_1}{\bar{Y}}, \ldots, \frac{Y_{[\gamma' n]}}{\bar{Y}}\right)$ satisfies (EC.3.1), and we will exploit this result. For the first condition, we observe that

$$\begin{aligned}
\sup_n ||W_{n1} - 1||_{2,1}^2 &\le \max\left\{1, \left\|\frac{n}{[\gamma' n]} \sup\nolimits_{1\le i\le[\gamma' n]} \frac{Y_i}{\bar{Y}} - 1\right\|_{2,1}^2\right\} \\
&\le \max\left\{1, \left\|\frac{n}{[\gamma' n]} \left(\sup\nolimits_{1\le i\le[\gamma' n]} \frac{Y_i}{\bar{Y}} - 1\right)\right\|_{2,1}^2 + \left(\frac{n}{[\gamma' n]} - 1\right)^2\right\} < \infty,
\end{aligned}$$

where the second inequality exploits Exercise 2.9.2 in van der Vaart and Wellner [12]. For the second condition, we can see that

$$\max\nolimits_{1\le i\le n} |W_{n1} - 1| \le \max\left\{1, \frac{n}{[\gamma' n]} - 1 + \frac{n}{[\gamma' n]} \max\nolimits_{1\le i\le[\gamma' n]} \left|\frac{Y_i}{\bar{Y}} - 1\right|\right\}$$

and its expectation is bounded from the results of $\left(\frac{Y_1}{\bar{Y}}, \ldots, \frac{Y_{[\gamma' n]}}{\bar{Y}}\right)$. Now, only the third condition remains. To see this, we derive that

$$
\begin{aligned}
&\frac{1}{n}\sum\nolimits_{i=1}^{n}(W_{ni}-1)^2 \\
&= \frac{1}{n}\sum\nolimits_{i=1}^{[\gamma' n]}\left(\frac{n}{[\gamma' n]}\frac{Y_i}{\bar{Y}}-1\right)^2+\frac{n-[\gamma' n]}{n} \\
&= \frac{1}{n}\sum\nolimits_{i=1}^{[\gamma' n]}\left(\frac{n}{[\gamma' n]}\frac{Y_i}{\bar{Y}}-\frac{n}{[\gamma' n]}\right)^2+\frac{2}{n}\left(\frac{n}{[\gamma' n]}-1\right)\underbrace{\sum\nolimits_{i=1}^{[\gamma' n]}\left(\frac{n}{[\gamma' n]}\frac{Y_i}{\bar{Y}}-\frac{n}{[\gamma' n]}\right)}_{=0} \\
&\quad+\frac{[\gamma' n]}{n}\left(\frac{n}{[\gamma' n]}-1\right)^2+\frac{n-[\gamma' n]}{n} \\
&\to \frac{1}{\gamma' a}+\frac{(1-\gamma')^2}{\gamma'}+1-\gamma' = \frac{1+a(1-\gamma')^2+a\gamma'(1-\gamma')}{a\gamma'} = \frac{1+a(1-\gamma')}{a\gamma'},
\end{aligned}
$$

which completes the proof. $\square$

EC.3.2. Consistency of Input Model: Theorem 1

Proof of Theorem 1. In this proof, we only use the ranking structure between n data points $\mathbf{Y}_1, \cdots, \mathbf{Y}_n \in \mathcal{Y}$. The same structure is shared by the transformed sample $\mathbf{U}_1, \ldots, \mathbf{U}_n$ with $\mathbf{U}_i = (F_1(Y_{i1}), \ldots, F_d(Y_{id}))$. One can readily check

$$
C_n^{\mathbf{Y}} = F_n^{\mathbf{Y}}\left(F_{n1}^{\mathbf{Y},-1}, \ldots, F_{nd}^{\mathbf{Y},-1}\right) = U_n\left(U_{n1}^{-1}, \ldots, U_{nd}^{-1}\right),
$$

where U_n is the empirical distribution of $\mathbf{U}_i$'s and U_{nj} is the empirical CDF of the j-th marginal.

Triangle inequality tells us that

$$
\begin{aligned}
||\bar{C}_n^{\mathbf{Y}}(\hat{F}_1, \ldots, \hat{F}_d) - C(F_1, \ldots, F_d)||_\infty &\le ||\bar{C}_n^{\mathbf{Y}} - C||_\infty + ||C(\hat{F}_1, \ldots, \hat{F}_d) - C(F_1, \ldots, F_d)||_\infty \\
&\le ||\bar{C}_n^{\mathbf{Y}} - C||_\infty + \sum\nolimits_{i=1}^{d} ||\hat{F}_i - F_i||_\infty.
\end{aligned}
$$

The last inequality is derived from the fact $|C(\mathbf{u}) - C(\mathbf{v})| \le ||\mathbf{u}-\mathbf{v}||_1$ where $||\cdot||_1$ is L^1-norm. The second term converges to zero in probability by assumption. Now, it suffices to show that

$||\bar{C}_n^{\mathbf{Y}} - C||_\infty$ converges to 0 almost surely. We apply triangle inequality to $||\bar{C}_n^{\mathbf{Y}} - C||_\infty$ again so that

$$\begin{aligned}
||\bar{C}_n^{\mathbf{Y}} - C||_\infty &\le ||C_n^{\mathbf{Y}} - \bar{C}_n^{\mathbf{Y}}||_\infty + ||C_n^{\mathbf{Y}} - C||_\infty \\
&= ||C_n^{\mathbf{Y}} - \bar{C}_n^{\mathbf{Y}}||_\infty + ||U_n(U_{n1}^{-1}, \ldots, U_{nd}^{-1}) - C||_\infty \\
&\le ||C_n^{\mathbf{Y}} - \bar{C}_n^{\mathbf{Y}}||_\infty + ||U_n(U_{n1}^{-1}, \ldots, U_{nd}^{-1}) - C(U_{n1}^{-1}, \ldots, U_{nd}^{-1})||_\infty \\
&\quad + ||C(U_{n1}^{-1}, \ldots, U_{nd}^{-1}) - C||_\infty \\
&= ||C_n^{\mathbf{Y}} - \bar{C}_n^{\mathbf{Y}}||_\infty + ||U_n - C||_\infty + ||C(U_{n1}^{-1}, \ldots, U_{nd}^{-1}) - C||_\infty \\
&\le ||C_n^{\mathbf{Y}} - \bar{C}_n^{\mathbf{Y}}||_\infty + ||U_n - C||_\infty + \sum_{j=1}^d ||U_{nj}^{-1} - u_j||_\infty.
\end{aligned}$$

The first term converges to zero from Lemma EC.3. The second term also converges to zero almost surely by Theorem EC.3. Lastly, the third term vanishes as well, thanks to the law of iterated logarithm for empirical quantiles (see Chapter 13 in Shorack and Wellner [9]). □

EC.3.3. Weak Convergence of Input Model: Theorems 2 and 4 The proofs for weak convergence are based on the functional delta method. For this purpose, we make following definitions.

- $\mathbb{A} = \ell^\infty(\mathbb{R}^d)$: space of bounded functions on $\mathbb{R}^d$
- $\mathbb{A}_1 = \ell^\infty([0,1]^d)$: space of bounded functions on $[0,1]^d$
- $\mathbb{B} = \ell^\infty(\mathbb{R})^d$: space of d bounded functions on $\mathbb{R}$
- $\mathbb{A}_\phi = \{H \in \mathbb{A} | H \text{ is a CDF and } H_j(-\infty) = 0 \text{ for each } j\}$
- $\mathbb{B}_\phi = \{(G_1, \ldots, G_d) \in \mathbb{B} | \text{ each } G_j \text{ is a CDF on } \mathbb{R}\}$
- $\mathbb{A}_0 = \{\alpha \in \mathbb{A} | \alpha \text{ continuous and } \alpha(\infty \cdot \mathbf{1}) = 1, \alpha(\mathbf{x}) = 0 \text{ if some components of } \mathbf{x} \text{ are } -\infty\}$
- $\mathbb{A}_{10} = \{\alpha \in \mathbb{A}_1 | \alpha \text{ continuous and } \alpha(\mathbf{1}) = 1, \alpha(\mathbf{x}) = 0 \text{ if some components of } \mathbf{x} \text{ are } 0\}$
- $\mathbb{B}_0 = \{(\beta_1, \ldots, \beta_d) \in \mathbb{B} | \text{ each } \beta_j \text{ continuous and } \lim_{x \to \pm\infty} \beta_j(x) = 0\}$

Lastly, we set $\mathbb{D} = \mathbb{A} \times \mathbb{B}$, $\mathbb{D}_\phi = \mathbb{A}_\phi \times \mathbb{B}_\phi$, and $\mathbb{D}_0 = \mathbb{A}_0 \times \mathbb{B}_0$. The lemma below is an extended version of Theorem 2.4 in Bücher and Volgushev [2].

LEMMA EC.2. *Consider continuous $(C, F_1, \ldots, F_d) \in \mathbb{D}_\phi$. Suppose that, for each $j = 1, \ldots, d$, the continuous first-order partial derivative $\partial_j C = \partial C / \partial u_j$ exists on $\{\mathbf{u} \in [0,1]^d | u_j \in (0,1)\}$ and $\partial_j C$ is defined as zero on $\{\mathbf{u} \in [0,1]^d | u_j = 0 \text{ or } u_j = 1\}$. The mapping $\phi : \mathbb{D}_\phi \to \ell^\infty(\mathbb{R}^d)$ defined by*

$$\phi(H, G_1, \ldots, G_d) = H(H^{-1}) \circ (G_1, \ldots, G_d), \quad H^{-1} = (H_1^{-1}, \ldots, H_d^{-1})$$

is Hadamard-differentiable at $(F, F_1, \ldots, F_d)$ tangentially to $\mathbb{D}_0$ with

$$\phi'_{(F, F_1, \ldots, F_d)}(\alpha, \beta_1, \ldots, \beta_d) = \alpha - \sum_{j=1}^d \partial_j C(F_1, \ldots, F_d)\alpha_j + \sum_{j=1}^d \partial_j C(F_1, \ldots, F_d)\beta_j$$

for $(\alpha, \beta_1, \ldots, \beta_d) \in \mathbb{D}_0$ *where* $\alpha_j(x_j) = \alpha(\infty, \ldots, \infty, x_j, \infty, \ldots, \infty)$. *Here* H_i*'s are marginal CDFs of* H.

Proof. As the first step, we express $\phi = \phi_2 \circ \phi_1$ where $\mathbb{D}_\phi \xrightarrow{\phi_1} \mathbb{A}_1 \times \mathbb{B}_\phi \xrightarrow{\phi_2} \ell^\infty(\mathbb{R}^d)$ such that

$$\phi_1 : (H, G_1, \cdots, G_d) \mapsto (H(H^{-1}), G_1, \cdots, G_d),$$
$$\phi_2 : (f, G_1, \cdots, G_d) \mapsto f \circ (G_1, \cdots, G_d).$$

Thanks to Theorem 2.4 in Bücher and Volgushev [2], we have that ϕ_1 is Hadamard-differentiable at $(F, F_1, \ldots, F_d)$ tangentially to $\mathbb{D}_0$ with

$$\phi_1{}'_{(F,F_1,\ldots,F_d)}(\alpha, g_1, \ldots, g_d) = \left(\alpha(F_1^{-1}, \ldots, F_d^{-1}) - \sum\nolimits_{j=1}^d \partial_j C \alpha_j(F_j^{-1}), g_1, \ldots, g_d\right)$$

for $\alpha \in \mathbb{A}_0$ and $(g_1, \ldots, g_d) \in \mathbb{B}_0$. We note that α is continuous on $\mathbb{R}^d$ and that each α_j is continuous on $\mathbb{R}$ with $\alpha_j(-\infty) = \alpha_j(\infty) = 0$. Since $\partial_j C$ is continuous whenever $u_j = F_j(x_j) \in (0,1)$ and bounded by 1 for any $\mathbf{u}$, $\alpha(F_1^{-1}, \ldots, F_d^{-1}) - \sum_j \partial_j C \alpha_j(F_j^{-1})$ is continuous and thus uniformly continuous on $[0,1]^d$.

Next, we claim that ϕ_2 is Hadamard-differentiable at $\phi_1(F, F_1, \ldots, F_d) = (C, F_1, \ldots, F_d)$ tangentially to $\mathbb{A}_{10} \times \mathbb{B}_0$ with

$$\phi_2{}'_{(C,F_1,\ldots,F_d)}(f, \beta_1, \ldots, \beta_d) = f(F_1, \ldots, F_d) + \sum\nolimits_{j=1}^d \partial_j C(F_1, \ldots, F_d)\beta_j$$

for $f \in \mathbb{A}_1$ and $(\beta_1, \ldots, \beta_d) \in \mathbb{B}_0$. To see this, let us consider $f_t \in \mathbb{A}_1$ and $(\beta_{t1}, \ldots, \beta_{td}) \in \mathbb{B}$ such that

- $(F_1 + t\beta_{t1}, \ldots, F_d + t\beta_{td}) \in \mathbb{B}_\phi$ for each t,
- $(f_t, \beta_{t1}, \ldots, \beta_{td}) \to (f, \beta_1, \ldots, \beta_d) \in \mathbb{A}_{10} \times \mathbb{B}_0$ as t decreases to zero.

As done in Bücher and Volgushev [2], we look at

$$\frac{1}{t}\left\{\phi_2\Big(C + tf_t, F_1 + t\beta_{t1}, \ldots, F_d + t\beta_{td}\Big) - \phi_2\Big(C, F_1, \ldots, F_d\Big)\right\}$$
$$= f_t \circ \Big(F_1 + t\beta_{t1}, \ldots, F_d + t\beta_{td}\Big) + \frac{1}{t}\left\{C \circ (F_1 + t\beta_{t1}, \ldots, F_d + t\beta_{td}) - C \circ (F_1, \ldots, F_d)\right\}.$$

The uniform convergence of $f_t \to f$, $F_j + t\beta_{tj} \to F_j$ together with uniform continuity of f on $[0,1]^d$ imply that the first term converges to $f \circ (F_1, \ldots, F_d)$ uniformly. For the second term, we apply the Taylor series expansion to obtain

$$\sum\nolimits_{j=1}^d \partial_j C(F_1, \ldots, F_d)\beta_j + \text{remainder}.$$

We can follow the logic of the same reference to show the remainder vanishes uniformly, but we skip the details as the proof is very similar.

Lastly, by the chain rule of Lemma 3.9.3 in van der Vaart and Wellner [12], we can compute the differential of ϕ as

$$\begin{aligned}\phi'_{(C,F_1,\dots,F_d)}(\alpha,\beta_1,\dots,\beta_d) &= \phi2'_{\phi1(F,F_1,\dots,F_d)}\left(\phi1'_{(C,F_1,\dots,F_d)}(\alpha,\beta_1,\dots,\beta_d)\right)\\ &= \phi2'_{(C,F_1,\dots,F_d)}\left(\alpha(F_1^{-1},\dots,F_d^{-1}) - \sum\nolimits_j \partial_j C\alpha_j(F_j^{-1}),\beta_1,\dots,\beta_d\right)\\ &= \alpha - \sum\nolimits_j \partial_j C(F_1,\dots,F_d)\alpha_j + \sum\nolimits_j \partial_j C(F_1,\dots,F_d)\beta_j,\end{aligned}$$

which completes the proof. □

Recall $F_n^{\mathbf{Y}}$ is the empirical distribution of $\{\mathbf{Y}_i\}$, which are i.i.d. observations from distribution F. Given Assumption 1, we see that the rank-order structures of both samples are the same with probability 1 so that the output $\phi(F_n^{\mathbf{Y}},\hat{F}_1,\dots,\hat{F}_d)$ is equivalent to $C_n^{\mathbf{Y}}(\hat{F}_1,\dots,\hat{F}_d)$, where $C_n^{\mathbf{Y}}$ is the empirical copula of $\{\mathbf{Y}_i\}$.

In later sections, we utilize the right continuity of $\bar{C}_n^{\mathbf{Y}}$ to apply integration by parts, Theorem EC.2, whereas we use $C_n^{\mathbf{Y}}$ in the functional delta method, Lemma EC.2. In order to enjoy both, we present the next lemma, which shows there is little difference between the two copulas.

Lemma EC.3. *Let C_n be the empirical copula of an i.i.d. sample $\{\mathbf{X}_i\}$ of size n from a continuous distribution and $\bar{C}_n(\mathbf{u}) = F_n(F_{n1}^{-1}(u_1),\dots,F_{nd}^{-1}(u_d))$. Then, we have $\sup_{\mathbf{u}\in[0,1]^d}\sqrt{n}|C_n(\mathbf{u}) - \bar{C}_n(\mathbf{u})| = o(1)$. Furthermore, suppose that a resampling plan $\mathbf{W}$ with parameter γ satisfies $\max_{i=1,\dots,n} W_{ni}/\sqrt{n} = o_P(1)$. For the resampling versions $C_{n,\mathbf{W}}$ and $\bar{C}_{n,\mathbf{W}}$, we have $\sup_{\mathbf{u}\in[0,1]^d}\sqrt{\gamma n}|C_{n,\mathbf{W}}(\mathbf{u}) - \bar{C}_{n,\mathbf{W}}(\mathbf{u})| = o_P(1)$.*

Proof. We first take the difference between C_n and $\bar{C}_n$:

$$\begin{aligned}&\sup\nolimits_{\mathbf{u}\in[0,1]^d}\left|\sqrt{n}(C_n(\mathbf{u}) - \bar{C}_n(\mathbf{u}))\right|\\ &\le n^{-1/2}\sup\nolimits_{\mathbf{u}\in[0,1]^d}\left|\sum\nolimits_{i=1}^n\left[\mathbf{1}\left\{X_{ij}\le F_{nj}^{-1}(u_j)\ \forall j\right\} - \mathbf{1}\left\{\frac{n}{n+1}F_{nj}(X_{ij})\le u_j\ \forall j\right\}\right]\right|\\ &\le n^{-1/2}\sup\nolimits_{\mathbf{u}\in[0,1]^d}\sum\nolimits_{i=1}^n\mathbf{1}\left\{\forall j,\ F_{nj}(X_{ij})\le\frac{n+1}{n}u_j \text{ and } \exists j \text{ s.t. } F_{nj}(X_{ij})>u_j\right\}\\ &+n^{-1/2}\sup\nolimits_{\mathbf{u}\in[0,1]^d}\left|\sum\nolimits_{i=1}^n\left[\mathbf{1}\left\{X_{ij}\le F_{nj}^{-1}(u_j)\ \forall j\right\} - \mathbf{1}\left\{F_{nj}(X_{ij})\le u_j\ \forall j\right\}\right]\right|\\ &=: \mathrm{I}+\mathrm{II}.\end{aligned}$$

For I in (EC.3.3), we can easily get

$$\begin{aligned}
\mathrm{I} &\leq n^{-1/2} \sup_{\mathbf{u}\in[0,1]^d} \sum_{i=1}^{n} \sum_{j=1}^{d} \mathbf{1}\left\{u_j < F_{nj}(X_{ij}) \leq \frac{n+1}{n} u_j\right\} \\
&\leq n^{-1/2} \sum_{j=1}^{d} \sup_{u_j\in[0,1]} \sum_{i=1}^{n} \mathbf{1}\left\{u_j < F_{nj}(X_{ij}) \leq \frac{n+1}{n} u_j\right\} \leq \frac{d}{\sqrt{n}}.
\end{aligned}$$

The last inequality follows from the observation that each sup term is bounded by 1 because the length of the interval $(u_j, (n+1)/nu_j]$ is less than $1/n$.

For II in (EC.3.3), we remind the reader of the general fact about quantile functions that $F_{nj}^{-1}(u_j) > X_{ij}$ if and only if $u_j > F_{nj}(X_{ij})$. By applying this, we split II and obtain an upper bound by

$$\begin{aligned}
\mathrm{II} &= n^{-1/2} \sup_{\mathbf{u}\in[0,1]^d} \left| \sum_{i=1}^{n} \left[\mathbf{1}\left\{X_{ij} \leq F_{nj}^{-1}(u_j)\ \forall j\right\} - \mathbf{1}\left\{F_{nj}(X_{ij}) \leq u_j\ \forall j\right\} \right] \right| \\
&= n^{-1/2} \sup_{\mathbf{u}\in[0,1]^d} \left| \sum_{i=1}^{n} \left[\mathbf{1}\left\{\forall j, X_{ij} \leq F_{nj}^{-1}(u_j) \text{ and } \exists j \text{ s.t. } X_{ij} = F_{nj}^{-1}(u_j)\right\} \right.\right. \\
&\quad \left.\left. + \mathbf{1}\left\{X_{ij} < F_{nj}^{-1}(u_j)\ \forall j\right\} - \mathbf{1}\left\{F_{nj}(X_{ij}) \leq u_j\ \forall j\right\} \right] \right| \\
&\leq n^{-1/2} \sup_{\mathbf{u}\in[0,1]^d} \sum_{i=1}^{n} \sum_{j=1}^{d} \mathbf{1}\left\{X_{ij} = F_{nj}^{-1}(u_j)\right\} \\
&\quad + n^{-1/2} \sup_{\mathbf{u}\in[0,1]^d} \sum_{i=1}^{n} \left| \mathbf{1}\left\{F_{nj}(X_{ij}) < u_j\ \forall j\right\} - \mathbf{1}\left\{F_{nj}(X_{ij}) \leq u_j\ \forall j\right\} \right| \\
&\leq n^{-1/2} \sum_{j=1}^{d} \sup_{u_j\in[0,1]} \sum_{i=1}^{n} \mathbf{1}\left\{X_{ij} = F_{nj}^{-1}(u_j)\right\} \\
&\quad + n^{-1/2} \sum_{j=1}^{d} \sup_{u_j\in[0,1]} \sum_{i=1}^{n} \mathbf{1}\left\{F_{nj}(X_{ij}) = u_j\right\} \\
&\leq \frac{2d}{\sqrt{n}}.
\end{aligned}$$

These two observations yield the desired upper bound $3d/\sqrt{n}$.

For the resampling case, we can proceed similarly:

$$\begin{aligned}
&\sup_{\mathbf{u}\in[0,1]^d} \left| \sqrt{\gamma n}(C_{n,\mathbf{W}}(\mathbf{u}) - \bar{C}_{n,\mathbf{W}}(\mathbf{u})) \right| \\
&\leq \frac{\sqrt{\gamma}}{\sqrt{n}} \sum_{j=1}^{d} \sup_{u_j\in[0,1]} \sum_{i=1}^{n} W_{ni} \mathbf{1}\left\{u_j < F_{j,\mathbf{W}}(X_{ij}) \leq \frac{n+1}{n} u_j\right\} \\
&+ \frac{\sqrt{\gamma}}{\sqrt{n}} \sum_{j=1}^{d} \sup_{u_j\in[0,1]} \sum_{i=1}^{n} W_{ni} \mathbf{1}\left\{X_{ij} = F_{j,\mathbf{W}}^{-1}(u_j)\right\} + \frac{\sqrt{\gamma}}{\sqrt{n}} \sum_{j=1}^{d} \sup_{u_j\in[0,1]} \sum_{i=1}^{n} W_{ni} \mathbf{1}\left\{F_{j,\mathbf{W}}(X_{ij}) = u_j\right\}.
\end{aligned}$$

In the first term, consider $\mathcal{A} = \{i | u_j < F_{j,\mathbf{W}}(X_{ij}) \leq \frac{n+1}{n} u_j\}$ for a fixed u_j. For convenience, we rename X_{ij} for $i \in \mathcal{A}$ by $\{X_1^*, \ldots, X_a^*\}$ in increasing order with $a = |\mathcal{A}|$. At the same time, we rename W_{ni} by $\{W_1^*, \ldots, W_a^*\}$. As illustrated in Figure EC.1, we see $\sum_{i=2}^{a} W_i^* \leq 1$. Since $W_1^* \leq \max_i W_{ni}$, we

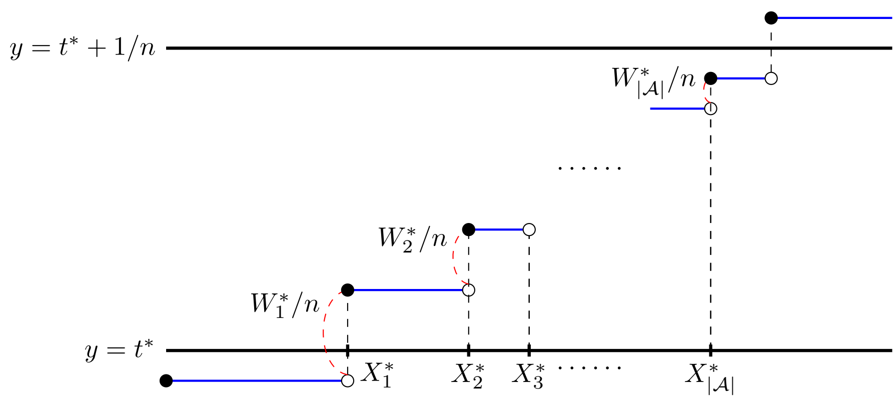

FIGURE EC.1. Figure for proof of Lemma EC.3

get an upper bound $\sqrt{\gamma/nd}(\max_i W_{ni}+1)$ for the first term. On the other hand, it is easy to see that each of the second and third terms is bounded by $\sqrt{\gamma/nd}\max_i W_{ni}$. Finally, we obtain

$$\sup_{\mathbf{u}\in[0,1]^d}\left|\sqrt{\gamma n}(C_{n,\mathbf{W}}(\mathbf{u})-\bar{C}_{n,\mathbf{W}}(\mathbf{u}))\right|\le\sqrt{\gamma/n}\,(3\max_i W_{ni}+1)\,d=o_P(1).\quad\square$$

Proof of Theorems 2 and 4. Recall our notation U_n for the empirical distribution of the $\mathbf{U}_i$'s and its resampling version $U_{n,\mathbf{W}_C}$ under the resampling plan $\mathbf{W}_C$. Thanks to the assumptions, the next relations hold almost surely:

$$C_n^{\mathbf{Y}}(\hat{F}_1,\cdots,\hat{F}_d)=\phi\Big(F_n^{\mathbf{Y}},\hat{F}_1,\cdots,\hat{F}_d\Big),$$
$$C^{\mathbf{Y}}_{n,\mathbf{W}_C}(F_{1,\mathbf{W}},\cdots,F_{d,\mathbf{W}})=\phi\Big(F^{\mathbf{Y}}_{n,\mathbf{W}_C},F_{1,\mathbf{W}_1},\cdots,F_{d,\mathbf{W}_C}\Big)$$

where ϕ is defined as in Lemma EC.2. By the same lemma, ϕ is Hadamard differentiable at $(C,F_1,\ldots,F_d)$. To apply the functional delta method, we should check the joint convergence of empirical processes $T_n,T_{n,\mathbf{W}}\in\mathbb{D}_0$ defined as $T_n=\sqrt{n}(F_n^{\mathbf{Y}}-F,\hat{F}_1-F_1,\ldots,\hat{F}_d-F_d)$ and $T_{n,\mathbf{W}}=\sqrt{n}(F^{\mathbf{Y}}_{n,\mathbf{W}}-F_n^{\mathbf{Y}},\hat{F}_{1,\mathbf{W}}-\hat{F}_1,\ldots,\hat{F}_{d,\mathbf{W}}-\hat{F}_d)$. From (2), we have

$$\hat{F}_j-F_j=\frac{n}{n+n_j}(F^{\mathbf{Y}}_{nj}-F_j)+\frac{n_j}{n+n_j}(F_{n_jj}-F_j).$$

This implies that T_n is a linear transformation of $\widetilde{T}_n$ defined by

$$\widetilde{T}_n:=(\underbrace{\sqrt{n}(F_n^{\mathbf{Y}}-F),\sqrt{n}(F^{\mathbf{Y}}_{n1}-F_1),\ldots,\sqrt{n}(F^{\mathbf{Y}}_{nd}-F_d)}_{:=\widetilde{T}_n^{\mathbf{Y}}},\sqrt{n_1}(F_{n_11}-F_1),\ldots,\sqrt{n_d}(F_{n_dd}-F_d)).$$

Thus, it suffices to show the joint convergence of $\widetilde{T}_n$. Thanks to Donsker's Theorem, we have $\sqrt{n}(F_n^{\mathbf{Y}}-F)\Longrightarrow\mathcal{GP}$, where $\mathcal{GP}$ is the Gaussian process introduced in Assumption 1. From this, the weak limits of $\widetilde{T}_n^{\mathbf{Y}}$ and $\{\sqrt{n_j}(F_{n_jj}-F_j)\}$ are given as

$$\widetilde{T}_n^{\mathbf{Y}}\Longrightarrow(\mathcal{GP},\mathcal{GP}_1,\ldots,\mathcal{GP}_d)\text{ in }\ell^\infty(\mathbb{R}^d\times\mathbb{R}\times\ldots\times\mathbb{R})\ \text{ and }\ \sqrt{n_j}(F_{n_jj}-F_j)\Longrightarrow\mathbb{F}_j^{X_j}\text{ in }\ell^\infty(\mathbb{R}),$$

where $\mathcal{GP}_j = \mathcal{GP}(\infty, \dots, x_j, \dots \infty)$ is a marginal Gaussian process of $\mathcal{GP}$. Since $\widetilde{T}_n^{\mathbf{Y}}$ and $\{\sqrt{n_j}(F_{n_j j} - F_j)\}_{1\le j\le d}$ are mutually independent, the marginal convergence implies joint convergence. Similar arguments hold for the joint convergence of $T_{n,\mathbf{W}}$. As a result, we finally have the joint convergence of T_n and $T_{n,\mathbf{W}}$ as required.

Now we are ready to apply the functional delta method (Theorem EC.5). For the approximating sequence $(F_n^{\mathbf{Y}}, \hat{F}_1, \dots, \hat{F}_d)$ with

- $\sqrt{n}(F_n^{\mathbf{Y}} - F) \Longrightarrow \mathcal{G}P$ in $\ell^\infty(\mathbb{R}^d)$ for the Gaussian process $\mathcal{G}P$;
- $\sqrt{n+n_j}(\hat{F}_j - F_j) \Longrightarrow \mathbb{F}_j$ in $\ell^\infty(\mathbb{R})$ for the Gaussian process $\mathbb{F}_j$ from (6),

the functional delta method implies

$$\begin{aligned}\sqrt{n}\Big(\phi(F_n^{\mathbf{Y}}, \hat{F}_1, \dots, \hat{F}_d) - \phi(F, F_1, \dots, F_d)\Big) &\Longrightarrow \phi'_{(F,F_1,\dots,F_d)}(\mathcal{G}P, w_1\mathbb{F}_1, \dots, w_d\mathbb{F}_d)\\ &= \mathbb{C}(F_1, \dots, F_d) + \sum_{j=1}^{d} w_j \partial C_j(F_1, \dots, F_d)\mathbb{F}_j\end{aligned}$$

from which the limiting process in the theorem statement can be derived from the fact that the limiting empirical copula process $\mathbb{C}$ is written as $\mathbb{C}(F_1, \dots, F_d) = \mathcal{GP} - \sum_{j=1}^{d} \partial_j C(F_1, \dots, F_d)\mathcal{GP}_j$ [8]. Lemma EC.3 allows us to replace $C_n^{\mathbf{Y}}$ with $\bar{C}_n^{\mathbf{Y}}$ and this proves Theorem 2.

The resampling case is identical except for some obvious changes. Hence, we omit the details. □

Lemma EC.4. *For fixed positive integer T, consider the mapping $\psi : \mathbb{A} \to \ell^\infty(\mathbb{R}^{dT})$ defined by*

$$\psi(F)(\mathbf{x}_1, \dots, \mathbf{x}_T) = \prod_{t=1}^{T} F(\mathbf{x}_t) \ \text{ for } F \in \mathbb{A}.$$

Then, $\psi(F)$ is Hadamard differentiable at F tagentially to $\mathbb{A}_0$ with

$$\psi'_F(\tau)(\mathbf{x}_1, \dots, \mathbf{x}_T) = \sum_{t=1}^{T} \tau(\mathbf{x}_t) \prod_{t_1 \neq t} F(\mathbf{x}_{t_1}) \ \text{ for } \tau \in \mathbb{A}_0.$$

Proof. For $\tau_\epsilon \in \mathbb{A}_0$ such that $F + \epsilon\tau_\epsilon \in \mathbb{A}$ and $||\tau_\epsilon - \tau||_\infty \to 0$ as $\epsilon \to 0$, one can easily confirm that

$$\epsilon^{-1}\{\psi(F + \epsilon\tau_\epsilon) - \psi(F)\}(\mathbf{x}_1, \dots, \mathbf{x}_T) = \sum_{t=1}^{T} \tau_\epsilon(\mathbf{x}_t) \prod_{t_1 \neq t} F(\mathbf{x}_{t_1}) + \text{error term},$$

where the error term is given by

$$\text{error term} = \sum_{\ell=2}^{T} \sum_{\substack{S_\tau \cup S_F = \{1,2,\dots,T\},\\ S_\tau \cap S_F = \emptyset, |S_\tau| = \ell}} \epsilon^{\ell-1} \prod_{t_1 \in S_\tau} \tau_\epsilon(\mathbf{x}_{t_1}) \prod_{t_2 \in S_F} F(\mathbf{x}_{t_2}).$$

Since $F(\mathbf{x}) \le 1$, the following upper bound of the error term can be derived:

$$|\text{error term}| \le \sum_{\ell=2}^{T} \sum_{\substack{S_\tau \cup S_F = \{1,2,\ldots,T\}, \\ S_\tau \cap S_F = \emptyset, |S_\tau| = \ell}} \epsilon^{\ell-1} ||\tau_\epsilon||_\infty^\ell,$$

which decays to zero as $\epsilon \to 0$ independent of $(\mathbf{x}_1, \ldots, \mathbf{x}_T)$, i.e., error term $= o(1)$. Hence, we have

$$\begin{aligned}
&\left(\epsilon^{-1}\{\psi(F+\epsilon\tau_\epsilon) - \psi(F)\}(\mathbf{x}_1, \ldots, \mathbf{x}_T) - \psi'_F(\tau)\right)(\mathbf{x}_1, \ldots, \mathbf{x}_T) \\
&= \sum_{t=1}^{T} (\tau_\epsilon(\mathbf{x}_t) - \tau(\mathbf{x}_t)) \prod\nolimits_{t_1 \ne t} F(\mathbf{x}_t) + o(1).
\end{aligned}$$

As $\prod_{t_1 \ne t} F(\mathbf{x}_t) \le 1$ holds for all t, the upper bound of the difference is given by

$$\left\| \epsilon^{-1}\{\psi(F+\epsilon\tau_\epsilon) - \psi(F)\} - \psi'_F(\tau) \right\|_\infty \le T||\tau_\epsilon - \tau||_\infty + o(1).$$

Taking $\epsilon \to 0$ completes the proof. □

EC.3.4. Weak Convergence Results: Theorems 3 and 5. As briefly mentioned in Section EC.1, the multivariate integration by parts plays an important role in understanding the behavior of output measurements in terms of input variability. Some technical approaches are adopted from Radulović et al. [7] and Fermanian et al. [4].

Proof of Theorem 3. Let $[\mathbf{a}, \mathbf{b}]^T$ be a hypercube that contains the support of $h(\mathbf{x}_1, \ldots, \mathbf{x}_T)$. Abusing notation, we denote the synthetic and true distribution of $\mathbf{X}(t)$ by $\hat{F}_t$ and F_t, respectively. We use the subscript j to denote the j-th marginal. By Assumption 3, we are able to apply Theorem EC.2. Then,

$$\begin{aligned}
&\sqrt{n}\left(\int_{[\mathbf{a},\mathbf{b}]^T} h(\mathbf{x}_1, \cdots, \mathbf{x}_T) \prod\nolimits_{t=1}^{T} d\hat{F}_t(\mathbf{x}_t) - \int_{[\mathbf{a},\mathbf{b}]^T} h(\mathbf{x}_1, \cdots, \mathbf{x}_T) \prod\nolimits_{t=1}^{T} dF_t(\mathbf{x}_t)\right) \\
&= \sqrt{n}\Gamma_h\left(\prod\nolimits_{t=1}^{T} \hat{F}_t(\cdot-) - \prod\nolimits_{t=1}^{T} F_t\right) \\
&= \sqrt{n}\Gamma_h\left(\prod\nolimits_{t=1}^{T} \hat{F}_t(\cdot-) - \prod\nolimits_{t=1}^{T} \hat{F}_t\right) + \sqrt{n}\Gamma_h\left(\prod\nolimits_{t=1}^{T} \hat{F}_t - \prod\nolimits_{t=1}^{T} F_t\right) =: \mathrm{I} + \mathrm{II}.
\end{aligned}$$

For II in (EC.3.4), note that $\Gamma_h : \ell^\infty([\mathbf{a}, \mathbf{b}]^T) \to \mathbb{R}$ is Lipschitz by Lemma EC.1, and it is linear by definition. By the functional delta method in Lemmas EC.2 and EC.4 and Theorem 2, $\sqrt{n}\left[\prod_{t=1}^{T} \hat{F}_t - \prod_{t=1}^{T} F_t\right]$ converges weakly to a Gaussian process $\psi'_F(\phi'_{(F,F_1,\ldots,F_d)}(\mathcal{GP}, w_1\mathbb{F}_1, \ldots, w_d\mathbb{F}_d))$ in $\ell^\infty([\mathbf{a}, \mathbf{b}]^T)$. Then, the continuous mapping theorem yields the weak convergence of II, and the linearity of Γ_h implies the limit is Gaussian.

Now, it suffices to show that I converges to 0 in probability. By Lemma EC.1, we can give an upper bound of I by $|\mathrm{I}| \lesssim \sqrt{n}\sup_{\mathbf{x}_t\in[\mathbf{a},\mathbf{b}],1\le t\le T}\left|\prod_{t=1}^T \hat{F}_t(\mathbf{x}_t-) - \prod_{t=1}^T \hat{F}_t(\mathbf{x}_t)\right| \le T\sqrt{n}\sup_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]}\left|\hat{F}(\mathbf{x}-) - \hat{F}(\mathbf{x})\right|$, where, in the last inequality, we utilize the next inequality

$$\left|\prod_{t=1}^T a_t - \prod_{t=1}^T b_t\right| \le \max_{1\le t\le T}\{|a_t|,|b_t|\}^{T-1}\sum_{t=1}^T |a_t - b_t|$$

for any two sequences $\{a_t\}$ and $\{b_t\}$, which is easily verifiable by induction. For the supremum term, similar arguments as in Lemma EC.3 tell us that

$$\begin{aligned}
&\sup_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]}\left|\hat{F}(\mathbf{x}) - \hat{F}(\mathbf{x}-)\right| \\
&= \sup_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]}\left|\bar{C}_n^{\mathbf{Y}}\left(\hat{F}_1(x_1),\ldots,\hat{F}_d(x_d)\right) - \bar{C}_n^{\mathbf{Y}}\left(\hat{F}_1(x_1-),\ldots,\hat{F}_d(x_d-)\right)\right| \\
&= \sup_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]}\left|\frac{1}{n}\sum_{i=1}^n\left[\mathbf{1}\left\{\frac{n}{n+1}F_{nj}^{\mathbf{Y}}(Y_{ij}) \le \hat{F}_j(x_j)\ \forall j\right\} - \mathbf{1}\left\{\frac{n}{n+1}F_{nj}^{\mathbf{Y}}(Y_{ij}) \le \hat{F}_j(x_j-)\ \forall j\right\}\right]\right| \\
&\le \frac{1}{n}\sum_{j=1}^d \sup_{x_j\in[a_j,b_j]}\sum_{i=1}^n \mathbf{1}\left\{\frac{n+1}{n}\hat{F}_j(x_j-) < F_{nj}^{\mathbf{Y}}(Y_{ij}) \le \frac{n+1}{n}\hat{F}_j(x_j)\right\}.
\end{aligned}$$

By (2), the jump size of $\hat{F}_j(x_j)$ is 0 or $(n+n_j)^{-1}$ depending on whether we use a continuous parametric model or an empirical CDF. Thus, the length of the inverval $\left(\frac{n+1}{n}\hat{F}_j(x_j-), \frac{n+1}{n}\hat{F}_j(x_j)\right]$ is bounded by $(n+1)/n(n+n_j)$. Since the jump size of $F_{nj}^{\mathbf{Y}}$ is n^{-1}, the number of $F_{nj}^{\mathbf{Y}}(Y_{ij})$ contained in the target interval is at most $(n+1)/(n+n_j)$. Consequently, we have

$$\sqrt{n}\sup_{\mathbf{x}}\left|\hat{F}(\mathbf{x}) - \hat{F}(\mathbf{x}-)\right| \le \sum_{j=1}^d \frac{2\sqrt{n}}{n+n_j} \to 0$$

thanks to Assumption 2. Therefore, I in (EC.3.4) is $o(1)$, concluding the proof. □

Proof of Theorem 5. As done in the proof of Theorem 3, we denote the synthetic distribution of $\mathbf{X}(t)$ by $\hat{F}_t$. For the resampling distribution $F_{\mathbf{W}}$, we use $F_{n,\mathbf{W},t}$. Then the integration by parts formula yields

$$\eta(F_{\mathbf{W}}) - \eta(\hat{F}) = \Gamma_h\left(\prod_{t=1}^T F_{\mathbf{W},t}(\cdot-) - \prod_{t=1}^T \hat{F}_t(\cdot-)\right).$$

Without the left limit sign, the proof goes exactly the same as in the proof of Theorem 3 by invoking the functional delta method and the continuous mapping theorem.

Given that we already showed $\sqrt{n}\sup_{\mathbf{x}}\left|\hat{F}(\mathbf{x}-) - \hat{F}(\mathbf{x})\right| = o(1)$, it suffices to show

$$\sqrt{\gamma n}\sup_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]}\left|F_{\mathbf{W}}(\mathbf{x}-) - F_{\mathbf{W}}(\mathbf{x})\right| = o_P(1).$$

Since this is trivial if we use continuous parametric models for marginals, we focus on empirical marginals. Observe that

$$\begin{aligned}&\sup_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]}\left|F_{\mathbf{W}}(\mathbf{x}-)-F_{\mathbf{W}}(\mathbf{x})\right|\\&\leq\frac{1}{n}\sum_{j=1}^{d}\sup_{x_j}\sum_{i=1}^{n}W_{i,C}\mathbf{1}\left\{\frac{n+1}{n}F_{j,\mathbf{W}}(x_j-)<F^{\mathbf{Y}}_{n,\mathbf{W}_C}(Y_{ij})\leq\frac{n+1}{n}F_{j,\mathbf{W}}(x_j)\right\}.\end{aligned}$$

Now let $\mathcal{A}=\left\{i|\frac{n+1}{n}F_{j,\mathbf{W}}(x_j-)<F^{\mathbf{Y}}_{n,\mathbf{W}_C}(Y_{ij})\leq\frac{n+1}{n}F_{j,\mathbf{W}}(x_j)\right\}$ for a fixed x_j. The maximal jump size of $F_{j,\mathbf{W}}$ is $\left(\max_{1\leq i\leq n_j}W_{ij}+\max_{1\leq i\leq n}W_{i,C}\right)/(n+n_j)$, thus the maximal interval size becomes $(n+1)/n(n_j+n)\times\left(\max_{1\leq i\leq n_j}W_{ij}+\max_{1\leq i\leq n}W_{i,C}\right)$. Arguments similar to Lemma EC.3 imply that the supremum term in (EC.3.4) is bounded by

$$\frac{n+1}{n+n_j}\left(\max_{1\leq i\leq n_j}W_{ij}+\max_{1\leq i\leq n}W_{i,C}\right)+\max_{1\leq i\leq n}W_{i,C}.$$

This bound together with Assumptions 2 and 4 shows $\sqrt{\gamma n}\sup|F_{\mathbf{W}}(\cdot-)-F_{\mathbf{W}}(\cdot)|=o_P(1)$. □

EC.3.5. Discussion on the Extension of Assumption 3 Our theoretical analysis relies heavily on multivariate integration by parts. For random stopping time T, the performance measure h can be expressed as $h(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_T)=\sum_{t=1}^{\infty}h_t(\mathbf{x}_1,\ldots,\mathbf{x}_t)\mathbf{1}\{T=t\}$. To apply integration by parts requires that for any fixed $T=t$, $h_t(\mathbf{x}_1,\ldots,\mathbf{x}_t)$ is of BVHK. If $\mathrm{P}(T=\infty)>0$, the integration by parts approach becomes inapplicable, because finite-dimensionality is required. Even when $\mathrm{P}(T<\infty)=1$, the conditions ensuring that h_t is BVHK for all t are, to the best of our knowledge, unknown and warrant further investigation.

This issue, however, is easily addressed when T is bounded almost surely, say bounded by some constant M. In such cases, h can be treated as a function mapping $\mathbb{R}^{dM}$ to $\mathbb{R}$, specifically $h(\mathbf{x}_1,\ldots,\mathbf{x}_M)$. Leveraging the fact that the space of BVHK functions is dense in the L^2 space, our ECC framework can be applied by considering the target performance measure as $h(\mathbf{x}_1,\ldots,\mathbf{x}_M)$.

If $\mathrm{P}(T<\infty)=1$, we may apply the ECC framework at the expense of approximation errors; replace h with some approximating BVHK function and truncate the time horizon with a large M. More specifically, if we can find g and M such that

$$\begin{aligned}\mathrm{E}_{\mathbf{X}(t)\sim F}[h(\mathbf{X}(t),\ldots,\mathbf{X}(T))]&\approx\mathrm{E}_{\mathbf{X}(t)\sim F}[h(\mathbf{X}(t),\ldots,\mathbf{X}(T))1\{T\leq M\}]\\&\approx\mathrm{E}_{\mathbf{X}(t)\sim F}[g(\mathbf{X}(1),\ldots,\mathbf{X}(M))],\end{aligned}$$

where g is of BVHK, we can approximate the target performance measure with an arbitrary precision by applying the ECC framework with truncated time horizon, $\min\{T,M\}$.

EC.3.6. L^2 Convergence Results: Theorem 6 Before we consider L^2 convergence, we present a result about the maximum oscillation modulus of empirical process. First, we review a multivariate extension of oscillation modulus provided by Einmahl [3]. Next, we introduce a useful inequality by Segers [8], which we are going to apply to quantify the convergence speed of $M_{nd}(a_n)$ for some $a_n \to 0$.

DEFINITION EC.3. Let $\{\mathbf{U}_1, \cdots, \mathbf{U}_n\}$ be an i.i.d. sample from copula C. The *multivariate oscillation modulus* is defined by

$$M_{nd}(a) = \sup\left\{|\alpha_n(\mathbf{u}) - \alpha_n(\mathbf{v})| : \mathbf{u}, \mathbf{v} \in [0,1]^d \text{ with } \max_{j\in 1:d} |u_j - v_j| \le a\right\}$$

where α_n is the empirical process based on the sample.

PROPOSITION EC.2 **(Segers [8]).** *Let C be a d-variate copula. There exist constants K_1 and K_2 which only depend on d such that* $\mathrm{P}(M_{nd}(a) \ge \lambda) \le \frac{K_1}{a} \exp\left\{-\frac{K_2\lambda^2}{a}\psi\left(\frac{\lambda}{\sqrt{n}a}\right)\right\}$ *for all $a \in (0, 1/2]$ and all $\lambda \ge 0$ where $\psi(x) := 2x^{-2}\{(1+x)\log(1+x) - x\}$.*

Proof of Theorem 6. We adopt the same notation in the proof of Theorem 5. Then, we have

$$\begin{aligned}
\left|\sqrt{n}\left(\eta(F_{\mathbf{W}}) - \eta(\hat{F})\right)\right| &= \left|\sqrt{n}\left(\Gamma_h\Big(\prod_{t=1}^{T} F_{\mathbf{W},t} - \prod_{t=1}^{T} \hat{F}_t\Big)\right) + \varepsilon\right| \\
&\lesssim |\Gamma|_h\left(\sqrt{n}\Big|\prod_{t=1}^{T} F_{\mathbf{W},t} - \prod_{t=1}^{T} \hat{F}_t\Big|\right) + |\varepsilon| \\
&\lesssim T|\Gamma|_h\Big(\sqrt{n}\sum_{t=1}^{T} |F_{\mathbf{W},t} - \hat{F}_t|\Big) + |\varepsilon|,
\end{aligned}$$

where the remainder ε is bounded by the sum of constant multiples of

$$\sqrt{n}\sup_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]}\left|\hat{F}(\mathbf{x}-) - \hat{F}(\mathbf{x})\right|, \quad \sqrt{n}\sup_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]}\left|F_{\mathbf{W}}(\mathbf{x}-) - F_{\mathbf{W}}(\mathbf{x})\right|.$$

Here, the second inequality is from (EC.3.4). The bounds of these two quantities in the proofs of Theorems 3 and 5 can be readily seen to converge to zero in L^p sense thanks to Assumptions 2 and 4. In other words, $\mathrm{E}_{\mathbf{W}}[\varepsilon^p] = o(1)$.

Now, it is enough to check the uniform integrability of $|\Gamma|_h\left(\sqrt{n}|\sum_{t=1}^{T}|F_{\mathbf{W},t} - \hat{F}_t|\right)$. By the triangle inequality, we have

$$\begin{aligned}
\sum_{t=1}^{T}\left|\sqrt{n}\left(F_{\mathbf{W},t} - \hat{F}_t\right)\right| &= \sqrt{n}\sum_{t=1}^{T}\left|\bar{C}^{\mathbf{Y}}_{n,\mathbf{W}_C,t}(F_{1,\mathbf{W},t}, \cdots, F_{d,\mathbf{W},t}) - \bar{C}^{\mathbf{Y}}_{n,t}(\hat{F}_{1,t}, \cdots, \hat{F}_{d,t})\right| \\
&\le \sqrt{n}\sum_{t=1}^{T}\left|\bar{C}^{\mathbf{Y}}_{n,\mathbf{W}_C,t}(F_{1,\mathbf{W},t}, \cdots, F_{d,\mathbf{W},t}) - \bar{C}^{\mathbf{Y}}_{n,t}(F_{1,\mathbf{W},t}, \cdots, F_{d,\mathbf{W},t})\right| \\
&\quad + \sqrt{n}\sum_{t=1}^{T}\left|\bar{C}^{\mathbf{Y}}_{n,t}(F_{1,\mathbf{W},t}, \cdots, F_{d,\mathbf{W},t}) - \bar{C}^{\mathbf{Y}}_{n,t}(\hat{F}_{1,t}, \cdots, \hat{F}_{d,t})\right| \\
&=: \mathrm{I} + \mathrm{II}.
\end{aligned}$$

First, let us consider the uniform integrability with respect to I in (EC.3.6). Consecutive application of the second and the first statements of Lemma EC.1 yields

$$\begin{aligned}
\mathrm{E}_{\mathbf{W}}\left[|\Gamma|_h(\mathrm{I})^{p+\delta}\right] &\lesssim \mathrm{E}_{\mathbf{W}}\left[|\Gamma|_h(I^{p+\delta})\right] \\
&= |\Gamma|_h\left(\mathrm{E}_{\mathbf{W}_1,\cdots,\mathbf{W}_d}\left[\mathrm{E}_{\mathbf{W}_C}\left[I^{p+\delta}\right]\right]\right) \\
&\lesssim \sup\nolimits_{\mathbf{u}\in[0,1]^d}\mathrm{E}_{\mathbf{W}_C}\left[n^{\frac{p+\delta}{2}}\left|\bar{C}^{\mathbf{Y}}_{n,\mathbf{W}_C}(\mathbf{u})-\bar{C}^{\mathbf{Y}}_n(\mathbf{u})\right|^{p+\delta}\right] = O_P(1).
\end{aligned}$$

This is nothing but the crystal ball condition in Assumption 6 with p.

Second, we turn our attention to II in (EC.3.6). For this, recall that the empirical copula process based on $\{\mathbf{Y}_i\}$ is $\mathbb{C}^{\mathbf{Y}}_n = \sqrt{n}(C^{\mathbf{Y}}_n - C)$ where $C^{\mathbf{Y}}_n$ is the empirical copula of $\mathbf{Y}_i$'s and C is the true copula. The RCLL version of this is written as $\bar{\mathbb{C}}^{\mathbf{Y}}_n = \sqrt{n}(\bar{C}^{\mathbf{Y}}_n - C)$. We define yet another process $\widetilde{\mathbb{C}}^{\mathbf{Y}}_n$ by $\widetilde{\mathbb{C}}^{\mathbf{Y}}_n := \alpha_n - \sum_{j=1}^d \partial_j C\alpha_{nj}$ where α_n and α_{nj}'s are the empirical processes of $\mathbf{U}_i$'s and their marginal datasets. Then, we use the triangle inequality to see

$$\begin{aligned}
\mathrm{II} \le &\sum\nolimits_{t=1}^T\left|(\bar{\mathbb{C}}^{\mathbf{Y}}_{n,t}-\widetilde{\mathbb{C}}^{\mathbf{Y}}_{n,t})(F_{1,\mathbf{W},t},\dots,F_{d,\mathbf{W},t})\right| + \sum\nolimits_{t=1}^T\left|(\bar{\mathbb{C}}^{\mathbf{Y}}_{n,t}-\widetilde{\mathbb{C}}^{\mathbf{Y}}_{n,t})(\hat{F}_{1,t},\dots,\hat{F}_{d,t})\right| \\
&+\sum\nolimits_{t=1}^T\left|\widetilde{\mathbb{C}}^{\mathbf{Y}}_{n,t}(F_{1,\mathbf{W},t},\dots,F_{d,\mathbf{W},t})-\widetilde{\mathbb{C}}^{\mathbf{Y}}_{n,t}(\hat{F}_{1,t},\dots,\hat{F}_{d,t})\right| \\
&+\sum\nolimits_{t=1}^T\sqrt{n}\left|C(F_{1,\mathbf{W},t},\dots,F_{d,\mathbf{W},t})-C(\hat{F}_{1,t},\dots,\hat{F}_{d,t})\right|.
\end{aligned}$$

Let us denote the respective expression in each line by (a), (b), and (c). Then, it is enough to show the stochastic boundedness of $\mathrm{E}_{\mathbf{W}}\left[|\Gamma|_h((\mathrm{a}))^{p+\delta}\right]$ to $\mathrm{E}_{\mathbf{W}}\left[|\Gamma|_h((\mathrm{c}))^{p+\delta}\right]$.

For (a), we use Lemma EC.1 to obtain

$$\mathrm{E}_{\mathbf{W}}\left[|\Gamma|_h((\mathrm{a}))^{p+\delta}\right] \lesssim \mathrm{E}_{\mathbf{W}}\left[\sup_{\mathbf{x}}(\mathrm{a})^{p+\delta}\right] \lesssim \sup_{\mathbf{u}}\left|\bar{\mathbb{C}}^{\mathbf{Y}}_n-\widetilde{\mathbb{C}}^{\mathbf{Y}}_n\right|^{p+\delta} \le \left(\sup_{\mathbf{u}}|\bar{\mathbb{C}}^{\mathbf{Y}}_n-\mathbb{C}^{\mathbf{Y}}_n| + \sup_{\mathbf{u}}|\mathbb{C}^{\mathbf{Y}}_n-\widetilde{\mathbb{C}}^{\mathbf{Y}}_n|\right)^{p+\delta}.$$

Lemma EC.3 above and Proposition 3.1 of Segers [8] imply the right hand side is $o_P(1)$.

For (b), we split the term into the terms involving α_n and α_{nj} separately. Using the same lemma, we get for α_n,

$$\begin{aligned}
&\mathrm{E}_{\mathbf{W}}\left[|\Gamma|_h^{p+\delta}\left(\sum\nolimits_{t=1}^T\left|\alpha_{n,t}(F_{1,\mathbf{W},t},\dots,F_{d,\mathbf{W},t})-\alpha_{n,t}(\hat{F}_{1,t},\dots,\hat{F}_{d,t})\right|\right)\right] \\
&\lesssim \mathrm{E}_{\mathbf{W}}\left[|\Gamma|_h\left(\sum\nolimits_{t=1}^T\left|\alpha_{n,t}(F_{1,\mathbf{W},t},\dots,F_{d,\mathbf{W},t})-\alpha_{n,t}(\hat{F}_{1,t},\dots,\hat{F}_{d,t})\right|^{p+\delta}\right)\right] \\
&= \sum\nolimits_{t=1}^T|\Gamma|_h\left(\mathrm{E}_{\mathbf{W}}\left[\left|\alpha_{n,t}(F_{1,\mathbf{W},t},\dots,F_{d,\mathbf{W},t})(\mathbf{x})-\alpha_{n,t}(\hat{F}_{1,t},\dots,\hat{F}_{d,t})(\mathbf{x})\right|^{p+\delta}\right]\right) \\
&\lesssim \sup\nolimits_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]}\mathrm{E}_{\mathbf{W}}\left[\left|\alpha_n(F_{1,\mathbf{W}},\dots,F_{d,\mathbf{W}})(\mathbf{x})-\alpha_n(\hat{F}_1,\dots,\hat{F}_d)(\mathbf{x})\right|^{p+\delta}\right] \\
&\lesssim M_{nd}(a_n)^{p+\delta}+(2\sqrt{n})^{p+\delta}\sum\nolimits_{j=1}^d\sup\nolimits_{x_j\in[a_j,b_j]}\mathrm{P}_{\mathbf{W}}\left[|F_{j,\mathbf{W}}(x_j)-\hat{F}_j(x_j)|\ge a_n\right]
\end{aligned}$$

for a given positive a_n. The second term in the last line is bounded by $2^{p+\delta}db_n = O_P(1)$ by the concentration inequality condition in the theorem statement. For the first term, fix arbitrary small $\varepsilon > 0$ and large M. The condition on a_n implies that there exists some k and N such that $M/\sqrt{n}a_n < k$ for all $n \geq N$. By Proposition EC.2 and the decreasing property of ψ, we then get

$$\mathrm{P}(M_{nd}(a_n) > M) \leq \frac{K_1}{a_n}\exp\left\{-\frac{K_2M^2}{a_n}\psi\left(\frac{M}{\sqrt{n}a_n}\right)\right\} \leq \frac{K_1}{a_n}\exp\left\{-\frac{K_2M^2}{a_n}\psi(kM)\right\}$$

whenever $n \geq N$. The right-hand side can be smaller than ε for all sufficiently large n values because $a_n = o(1)$. Thus we conclude $M_{nd}(a_n) = O_P(1)$. Similar arguments are applicable to α_{nj}'s.

For (c), we use the inequality $|C(\mathbf{u}) - C(\mathbf{v})| \leq \|\mathbf{u}-\mathbf{v}\|_1$ to proceed as follows:

$$\begin{aligned}
\mathrm{E}_{\mathbf{W}}\left[|\Gamma|_h((\mathrm{c}))^{p+\delta}\right] &= \mathrm{E}_{\mathbf{W}}\left[|\Gamma|_h^{p+\delta}(\sqrt{n}\sum\nolimits_{t=1}^{T}\left|C(F_{1,\mathbf{W},t},\ldots,F_{d,\mathbf{W},t}) - C(\hat{F}_{1,t},\ldots,\hat{F}_{d,t})\right|)\right]\\
&\leq \mathrm{E}_{\mathbf{W}}\left[|\Gamma|_h^{p+\delta}\left(\sqrt{n}\sum\nolimits_{t=1}^{T}\sum\nolimits_{j=1}^{d}\left|F_{j,\mathbf{W},t} - \hat{F}_{j,t}\right|\right)\right]\\
&\lesssim |\Gamma|_h\left(\mathrm{E}_{\mathbf{W}}\left[n^{(p+\delta)/2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{j=1}^{d}\left|F_{j,\mathbf{W},t} - \hat{F}_{j,t}\right|^{p+\delta}\right]\right)\\
&\lesssim \sum\nolimits_{j=1}^{d} n^{(p+\delta)/2}\sup\nolimits_{x_j}\mathrm{E}_{\mathbf{W}}\left[\left|F_{j,\mathbf{W}}(x_j) - \hat{F}_j(x_j)\right|^{p+\delta}\right] = O_P(1).
\end{aligned}$$

by the crystal ball condition on the marginals, which completes the proof. □

EC.3.7. Proof of Theorem 7 In this section, we focus on subsampling with/without replacement. To emphasize this, we change our notation for conditional expectation $\mathrm{E}_{\mathbf{W}}$ and conditional probability $\mathrm{P}_{\mathbf{W}}$ to E_* and P_*, respectively, which is common practice in the statistical literature related to bootstrap analysis. When we do not allow replacement, its asymptotic behavior is difficult to handle because of the dependence between redrawn data. To resolve this issue, we utilize two results by Hoeffding.

1 (Hoeffding's Lemma.) *Let $\{c_1,\cdots,c_N\}$ be a population generated in a vector space $\mathbb{V}$. We redraw from the set to construct $\{U_1,\cdots,U_n\}$ and $\{V_1,\cdots,V_n\}$, samples without and with replacement for $n \leq N$. If $\phi : \mathbb{V} \to \mathbb{R}$ is a convex function, then* $\mathrm{E}\left[\phi\left(\sum_{j=1}^n U_j\right)\right] \leq \mathrm{E}\left[\phi\left(\sum_{j=1}^n V_j\right)\right]$ *holds.*

2 (Hoeffding's Inequality.) *Let X be a random variable such that $a \leq X - \mathrm{E}[X] \leq b$. Then, for $s > 0$, we have* $\mathrm{E}[\exp(s(X - \mathrm{E}[X]))] \leq \exp\left(s^2(b-a)^2/8\right)$.

Before we prove our main theorem, we verify the next key lemma, which is about the convergence speed of the difference between resampling and empirical quantiles.

LEMMA EC.5. *Suppose $\mathscr{U}$ is a set of n i.i.d. replicates from the uniform distribution. We redraw m data points with or without replacement from $\mathscr{U}$. Assume the sample size satisfies $m/n = O(1)$. For fixed $p \in [0,1]$ and deterministic $D_1 > 0$, there exists a constant $D = O_P(1)$, independent of p and t, such that*

$$\mathrm{P}_*\left(|v_m^*(p) - v_n(p)| \geq t/\sqrt{m}\right) \leq D\exp(-2t^2) \;\; for\ all\ t \in [0, D_1\sqrt{\log m}].$$

Here, $v_m^*(p)$ and $v_n(p)$ stand for the resampling and empirical p-quantiles, respectively.

Proof. Define G and G_n as the true and empirical distributions of $\mathscr{U}$, respectively. We also define an auxiliary function $p_{m,t} := G_n(v_n(p) + t/\sqrt{m})$ and $V_{j,t}$ as an indicator function which is equal to 1 if and only if the j-th resample is less than or equal to $v_n(p) + t/\sqrt{m}$. By applying the Chebyshev inequality for each $s > 0$, we get

$$\begin{aligned}
&\mathrm{P}_*\left(|v_m^*(p) - v_n(p)| > t/\sqrt{m}\right)\\
&= \mathrm{P}_*\left(v_m^*(p) > v_n(p) + t/\sqrt{m}\right) + \mathrm{P}_*\left(v_m^*(p) < v_n(p) - t/\sqrt{m}\right)\\
&\leq \mathrm{P}_*\left(\textstyle\sum_{j=1}^m V_{j,t} \leq mp\right) + \mathrm{P}_*\left(\textstyle\sum_{j=1}^m V_{j,-t} \geq mp\right)\\
&= \mathrm{P}_*\left(\textstyle\sum_{j=1}^m (V_{j,t} - p_{m,t}) \leq m(p - p_{m,t})\right) + \mathrm{P}_*\left(\textstyle\sum_{j=1}^m (V_{j,-t} - p_{m,-t}) \geq m(p - p_{m,-t})\right)\\
&= \mathrm{P}_*\left(\textstyle\sum_{j=1}^m (-V_{j,t} + p_{m,t}) \geq m(p_{m,t} - p)\right) + \mathrm{P}_*\left(\textstyle\sum_{j=1}^m (V_{j,-t} - p_{m,-t}) \geq m(p - p_{m,-t})\right)\\
&\leq \exp\left(-sm(p_{m,t} - p)\right)\mathrm{E}_*\left[\exp\left(s\textstyle\sum_{j=1}^m (-V_{j,t} + p_{m,t})\right)\right]\\
&+ \exp(-sm(p - p_{m,-t}))\mathrm{E}_*\left[\exp\left(s\textstyle\sum_{j=1}^m (V_{j,-t} - p_{m,-t})\right)\right]\\
&=: \mathrm{I} + \mathrm{II}.
\end{aligned}$$

Let us first consider the case where we permit replacement. By independence and Hoeffding's inequality,

$$\begin{aligned}
\mathrm{I} &= \exp\left(-sm(p_{m,t} - p)\right)\textstyle\prod_{j=1}^m \mathrm{E}_*\left[\exp\left(s(-V_{j,t} + p_{m,t})\right)\right]\\
&\leq \exp\Big(-sm(p_{m,t} - p)\Big)\exp\left(ms^2/8\right).
\end{aligned}$$

The right-hand side is minimized when $s = 4(p_{m,t} - p)$, so that $\mathrm{I} \leq \exp\Big(-2m(p_{m,t} - p)^2\Big)$. Even if replacement is not allowed, we can apply Hoeffding's Lemma with a convex function e^{-sx}. For this, we define c_i as the indicator function that is equal to 1 if and only if the i-th sample point is less than or equal to $v_n(p) + t/\sqrt{m}$. Then, we can still conclude $\mathrm{I} \leq \exp\Big(-2m(p_{m,t} - p)^2\Big)$. By applying the same arguments to II, we obtain the bound

$$\mathrm{P}_*\big(|v_m^*(p) - v_n(p)| > t/\sqrt{m}\big) \le \exp\big(-2m(p_{m,t}-p)^2\big) + \exp\big(-2m(p_{m,-t}-p)^2\big).$$

Now, let us analyze $p_{m,u} - p$ for $|u| \le D_1\sqrt{\log m}$. We divide the difference into three parts:

$$\begin{aligned}
p_{m,u} - p &= G_n(v_n(p) + u/\sqrt{m}) - G(p) \\
&= \Big\{G_n(v_n(p)+u/\sqrt{m}) - G(v_n(p)+u/\sqrt{m}) - G_n(v_n(p)) + G(v_n(p))\Big\} \\
&\quad + \Big\{G(v_n(p)+u/\sqrt{m}) - G(v_n(p))\Big\} + \Big\{G_n(v_n(p)) - G(p)\Big\} \\
&=: \text{(a)} + \text{(b)} + \text{(c)}.
\end{aligned}$$

The second term (b) is equal to $u/\sqrt{m}$ and the third term is bounded by $1/n$ uniformly in p. And (a) is bounded by the oscillation modulus of empirical process $M_{n1}(D_1\sqrt{\log m}/\sqrt{m})/\sqrt{n}$, where M_{n1} is the 1-dimensional maximum oscillation modulus related to the empirical process of $\mathcal{U}$. Let $\epsilon = \sqrt{m}((\text{a})+(\text{c}))$. Then,

$$m(p_{m,u}-p)^2 = u^2 + \epsilon^2 + 2u\epsilon \ge u^2 + \epsilon^2 - 2D_1\sqrt{\log m}|\epsilon|$$

From this observation, it is enough to show that $\epsilon^2 - 2\sqrt{\log m}|\epsilon|$ is bounded in probability by a constant that is independent of p and u. For this, we see

$$\begin{aligned}
&\Big|\epsilon^2 - 2\sqrt{\log m}|\epsilon|\Big| = \Big|m\left((\text{a})^2 + 2(\text{a})\cdot(\text{c}) + (\text{c})^2\right) - 2\sqrt{\log m}\sqrt{m}\cdot|(\text{a})+(\text{c})|\Big| \\
&\le \left(\frac{m}{n}M_{n1}(D_1\sqrt{\log m}/\sqrt{m})^2 + 2\frac{m}{n\sqrt{n}}M_{n1}(D_1\sqrt{\log m}/\sqrt{m}) + \frac{m}{n^2}\right) \\
&+ 2\sqrt{\frac{m\log m}{n}}M_{n1}(D_1\sqrt{\log m}/\sqrt{m}) + 2\frac{\sqrt{m\log m}}{n}.
\end{aligned}$$

The proof would be complete once we show $\sqrt{\log m}M_{n1}(D_1\sqrt{\log m}/\sqrt{m}) = O_P(1)$. By plugging in $a = D_1\sqrt{\log m/m}$ and $\lambda = M/\sqrt{\log m}$ in Proposition EC.2, we get

$$\begin{aligned}
&\mathrm{P}\left(\sqrt{\log m}M_{n1}\left(D_1\sqrt{\frac{\log m}{m}}\right) \ge M\right) \\
&\le \frac{K_1}{D_1}\sqrt{\frac{m}{\log m}}\exp\left\{-\frac{K_2M^2}{D_1}\frac{\sqrt{m}}{(\log m)^{3/2}}\psi\left(\frac{M/D_1}{\log m}\sqrt{\frac{m}{n}}\right)\right\}.
\end{aligned}$$

Since $m/n = O(1)$ by assumption and ψ is decreasing and continuous with $\psi(0) = 1$, $\psi(M/D_1\sqrt{m}/\log m\sqrt{n}) > 1/2$ for all sufficiently large n so that the right hand side is bounded above by

$$\frac{K_1}{D_1}\exp\left\{-K_3\frac{\sqrt{m}}{(\log m)^{3/2}} + \frac{1}{2}\log m - \frac{1}{2}\log\log m\right\}$$

with $K_3 = K_2\frac{M^2}{2D_1}$. It is then easy to check that the exponent diverges to $-\infty$ as $n, m \to \infty$. $\square$

Before the main proof, it is necessary to check whether bootstrap and subsampling satisfy Assumption 4. The next lemma shows that actually a stronger result holds for the two resampling plans.

LEMMA EC.6. *Suppose* $\mathbf{W}_n = (W_{n1}, \cdots, W_{nn})$ *be a resampling plan that redraws* $m = [\gamma n]$ *sample points out of n with or without replacement. For each* $\delta \geq 0$, $\mathrm{E}_{\mathbf{W}}\left[\left|\max W_{nj}/\sqrt{n}\right|^{2+\delta}\right] = o(1)$.

Proof. For resampling without replacement, Example 2 implies that the maximum of W_{ni} is n/m. Therefore, the assertion of the theorem is trivial.

For resampling with replacement, we note that $W_{n1} \stackrel{d}{=} n/m \times \sum_{i=1}^m B_i$, where the B_i's are i.i.d. Bernoulli random variables with success probability $1/n$. By the Chebyshev inequality,

$$\begin{aligned}
\mathrm{P}(W_{n1} > t) &\leq e^{-mt/n}\mathrm{E}\left[e^{\sum_{i-1}^m B_i}\right] = e^{-mt/n}\prod_{i=1}^m \mathrm{E}\left[e^{B_i}\right] = e^{-mt/n}\left(1 + (e-1)\frac{1}{n}\right)^m \\
&\leq e^{-mt/n}\exp\left(\frac{m}{n}(e-1)\right) \\
&\leq e^{-mt/n}\exp\left(\gamma(e-1)\right).
\end{aligned}$$

Thus, $\lim_{t\to\infty} t^{3+\delta}\mathrm{P}(W_{n1} > t) = 0$ for any $\delta \geq 0$. In particular, for each fixed $\delta > 0$, there exists t_0 and N such that

$$t^{3+\delta}\mathrm{P}(W_{n1} > t) \leq \exp\Big((3+\delta)\log t - \gamma t + t/n + \gamma(e-1)\Big) < \frac{1}{2+\delta}, \quad \forall t \geq t_0, \forall n \geq N.$$

Now, for $n > \max\{t_0^{2+\delta}, N\}$, we observe that

$$\begin{aligned}
&\mathrm{E}\left[(\max_{1\leq i\leq n} W_{ni})^{2+\delta}\right] \\
&= (2+\delta)\int_0^\infty t^{1+\delta}\mathrm{P}(\max W_{ni} > t)dt \\
&= (2+\delta)\left\{\int_0^{n^{\frac{1}{2+\delta}}} t^{1+\delta}\mathrm{P}(\max W_{ni} > t)dt + \int_{n^{\frac{1}{2+\delta}}}^\infty t^{1+\delta}\mathrm{P}(\max W_{ni} > t)dt\right\} \\
&\leq (2+\delta)\left\{\int_0^{n^{\frac{1}{2+\delta}}} t^{1+\delta}dt + n\int_{n^{\frac{1}{2+\delta}}}^\infty t^{1+\delta}\mathrm{P}(W_{n1} > t)dt\right\} \\
&\leq n + n^{\frac{1+\delta}{2+\delta}}.
\end{aligned}$$

This implies that $n^{-\frac{2+\delta}{2}}\mathrm{E}\left[(\max_{1\leq i\leq n} W_{ni})^{2+\delta}\right] \leq n^{-\frac{\delta}{2}} + n^{-\frac{\delta}{2}-\frac{1}{2+\delta}} \to 0$, completing the proof. □

Proof of Theorem 7. Thanks to Lemma EC.6, one can see that Assumption 4 holds for all $p \in \{2, 4\}$. It suffices to show Assumption 6. We split the proof into three parts to check Assumption 6.

1. Crystal ball condition for copula

As in the proof of Theorem 6, we can consider n samples $\mathscr{U} := \{\mathbf{U}_1, \ldots, \mathbf{U}_n\}$ from d-variate copula C. We redraw $m = [\gamma n]$ replicates $\left\{\mathbf{U}_1^*, \cdots, \mathbf{U}_m^*\right\}$ from $\mathscr{U}$ with or without replacement. Denote empirical distributions induced by the given data and the redrawn data by G_n and G_m^*, respectively. Also, let $v_n(\mathbf{u}) = (G_{n1}^{-1}(u_1), \cdots, G_{nd}^{-1}(u_d))$ be the empirical quantile vector, and let v_m^* be the v_n's resampling counterpart. Furthermore, denote $C_{n,\gamma}^*$ as the resampling empirical copula. Then, we get

$$C_n(\mathbf{u}) = G_n(v_n(\mathbf{u})), \quad C_{n,\gamma}^* = G_m^*(v_m^*(\mathbf{u})).$$

By Lemma EC.3, it is enough to verify $\mathrm{E}_*\left[\left|\sqrt{\gamma n}\big(C_{n,\gamma}^* - C_n\big)\right|^{p+2}\right] = O_P(1)$ for $p \in \{2, 4\}$. To see this, we divide it into three parts:

$$\begin{aligned}
\mathrm{E}_*\left[\left|\sqrt{\gamma n}\left(C_{n,\gamma}^* - C_n\right)\right|^{p+2}\right] \lesssim (\gamma n)^{(p+2)/2}\mathrm{E}_*\left[\left|G_m^*(v_m^*) - G_n(v_m^*)\right|^{p+2}\right] \qquad \text{(EC.1)}\\
+ (\gamma n)^{(p+2)/2}\mathrm{E}_*\left[\left|(G_n - C)(v_m^*) - (G_n - C)(v_n)\right|^{p+2}\right] \qquad \text{(EC.2)}\\
+ (\gamma n)^{(p+2)/2}\mathrm{E}_*\left[\left|C(v_m^*) - C(v_n)\right|^{p+2}\right]. \qquad \text{(EC.3)}
\end{aligned}$$

Recall that our notation $a \lesssim b$ for $a, b > 0$ if there exists a constant $k > 0$ such that $a \leq kb$.

Now, it suffices to show that (EC.1)-(EC.3) are $O_{\mathrm{P}}(1)$ uniformly in $\mathbf{u}$, i.e., $\sup_{\mathbf{u}\in[0,1]^d}$ (EC.1), $\sup_{\mathbf{u}\in[0,1]^d}$ (EC.2), and $\sup_{\mathbf{u}\in[0,1]^d}$ (EC.3) are all $O_{\mathrm{P}}(1)$, whether we allow replacement or not. First, let us give an upper bound of (EC.1). If we allow replacement, we apply the Dvoretzky-Kiefer-Wolfowitz inequality to deduce

$$\begin{aligned}
\sup_{\mathbf{u}\in[0,1]^d}\text{(EC.1)} &\leq (\gamma n)^{(p+2)/2}\mathrm{E}_*\left[\sup_{\mathbf{u}\in[0,1]^d}\left|G_m^*(\mathbf{u}) - G_n(\mathbf{u})\right|^{p+2}\right]\\
&= (\gamma n)^{(p+2)/2}(p+2)\int_0^\infty t^{p+1}\mathrm{P}_*\Big(\sup_{\mathbf{u}\in[0,1]^d}\left|G_m^*(\mathbf{u}) - G_n(\mathbf{u})\right| > t\Big)dt\\
&\leq (\gamma n)^{(p+2)/2}(p+2)\int_0^\infty 2t^{p+1}\exp\left(-2mt^2\right)dt\\
&\lesssim (\gamma n - 1)^{(p+2)/2}(p+2)\int_0^\infty 2t^{p+1}\exp\left(-2(\gamma n - 1)t^2\right)dt < \infty.
\end{aligned}$$

We conclude that $\sup_{\mathbf{u}\in[0,1]^d}$ (EC.1) is bounded from above by some constant, which is also $O_{\mathrm{P}}(1)$.

If we do not allow replacement, then we can apply Hoeffding's Lemma to $\mathrm{E}_*\left[\sup_{\mathbf{u}}\left|G_m^* - G_n\right|^{p+2}\right]$ so that the expectation is less than or equal to the same expectation with replacement. This leads to the boundedness of $\sup_{\mathbf{u}}$ (EC.1) without replacement as well. Specifically, it is enough to define

$c_i \in \ell^\infty(\mathcal{F}_{\mathrm{ind}})$ as $c_i(\mathbf{1}\{\cdot \le \mathbf{u}\}) = \mathbf{1}\{U_i \le \mathbf{u}\} - G_n(\mathbf{u})$ and the convex functional $\phi : \ell^\infty(\mathcal{F}_{\mathrm{ind}}) \to \mathbb{R}$ such that $\phi(\cdot) = \|\cdot\|_{\mathcal{F}_{\mathrm{ind}}}^{p+2}$.

For fixed $\mathbf{u}$, Lemma EC.5 with $D_1 = \sqrt{p+2}/2$ tells us that there exists $D = O_P(1)$ with

$$\mathrm{P}_*\left(|v_{mj}^*(u_j) - v_{nj}(u_j)| > \sqrt{(p+2)\log m/(4m)}\right) \le Dm^{-(p+2)/2}$$

for all $j = 1, \ldots, d$. Now, we turn to (EC.2). By truncation, we proceed as

$$\begin{aligned}
&\text{(EC.2)}\\
&= (\gamma n)^{(p+2)/2}\mathrm{E}_*\left[\left|(G_n - C)(v_m^*) - (G_n - C)(v_n)\right|^{p+2}\mathbf{1}\left\{\exists j\, |v_{mj}^* - v_{nj}| > \sqrt{(p+2)\frac{\log m}{4m}}\right\}\right]\\
&+ (\gamma n)^{(p+2)/2}\mathrm{E}_*\left[\left|(G_n - C)(v_m^*) - (G_n - C)(v_n)\right|^{p+2}\mathbf{1}\left\{\forall j\, |v_{mj}^* - v_{nj}| \le \sqrt{(p+2)\frac{\log m}{4m}}\right\}\right]\\
&\le 4^{(p+2)}(\gamma n)^{(p+2)/2}\sum\nolimits_{j=1}^d \mathrm{P}_*\left(|v_{mj}^* - v_{nj}| > \sqrt{(p+2)\log m/(4m)}\right)\\
&\quad + \gamma^{(p+2)/2}M_{nd}\left(\sqrt{(p+2)\log m/(4m)}\right)^{p+2}\\
&\le 4^{(p+2)}\gamma^{(p+2)/2}d\frac{Dn^{(p+2)/2}}{m^{(p+2)/2}} + \gamma^{(p+2)/2}M_{nd}\left(\sqrt{\frac{(p+2)\log\gamma n}{4(\gamma n - 1)}}\right)^{p+2}\\
&\le 4^{p+2}dD\left(\frac{m+1}{m}\right)^{(p+2)/2} + \gamma^{(p+2)/2}M_{nd}\left(\sqrt{\frac{(p+2)\log\gamma n}{4(\gamma n - 1)}}\right)^{p+2}.
\end{aligned}$$

It is easy to compute that the last term is $O_P(1)$ and a uniform upper bound in $\mathbf{u}$, thanks to Proposition EC.2 and the definition of M_{n1}.

As for (EC.3), we eliminate copula C by using the inequality $|C(\mathbf{u}) - C(\mathbf{v})| \le \|\mathbf{u} - \mathbf{v}\|_1$.

$$\sup\nolimits_{\mathbf{u}\in[0,1]^d} \text{(EC.3)} \lesssim \sum\nolimits_{i=1}^d n^{(p+2)/2}\sup\nolimits_{u_j\in[0,1]} \mathrm{E}_*\left[\left|v_{mj}^*(u_j) - v_{nj}(u_j)\right|^{p+2}\right].$$

On the other hand, we see that

$$\begin{aligned}
m^{(p+2)/2}\mathrm{E}_*\left[\left|v_{mj}^*(u_j) - v_{nj}(u_j)\right|^{p+2}\right] &= (p+2)\int_0^\infty t^{p+1}\mathrm{P}_*\left(\sqrt{m}\left|v_{mj}^*(u_j) - v_{nj}(u_j)\right| > t\right)dt\\
&= (p+2)\int_0^{2\sqrt{m}} t^{p+1}\mathrm{P}_*\left(\sqrt{m}\left|v_{mj}^*(u_j) - v_{nj}(u_j)\right| > t\right)dt
\end{aligned}$$

because $\left|v_{mj}^{*}-v_{nj}\right| \le 2$. In the last integral, we apply Lemma EC.5 with $D_1 = \sqrt{p+2}/2$, so that

$$
\begin{aligned}
&\int_0^{2\sqrt{m}} t^{p+1}\mathrm{P}_*\left(\sqrt{m}\left|v_{mj}^{*}(u_j)-v_{nj}(u_j)\right|>t\right)dt\\
&=\int_0^{\sqrt{\frac{(p+2)\log m}{4}}} t^{p+1}\mathrm{P}_*\left(\sqrt{m}\left|v_{mj}^{*}(u_j)-v_{nj}(u_j)\right|>t\right)dt\\
&\quad+\int_{\sqrt{(p+2)\log m/4}}^{2\sqrt{m}} t^{p+1}\mathrm{P}_*\left(\sqrt{m}\left|v_{mj}^{*}(u_j)-v_{nj}(u_j)\right|>t\right)dt\\
&\le D\int_0^{\sqrt{\frac{(p+2)\log m}{4}}} t^{p+1}e^{-2t^2}dt+\int_0^{2\sqrt{m}} t^{p+1}\mathrm{P}_*\left(\sqrt{m}\left|v_{mj}^{*}(u_j)-v_{nj}(u_j)\right|>\sqrt{\frac{(p+2)\log m}{4}}\right)dt\\
&\le D\int_0^{\infty} t^{p+1}e^{-2t^2}dt+D\int_0^{2\sqrt{m}}\frac{t^{p+1}}{m^{(p+2)/2}}dt\\
&=D\left(\int_0^{\infty} t^{p+1}e^{-2t^2}dt+\frac{2^{p+2}}{p+2}\right)<\infty,
\end{aligned}
$$

where the last inequality comes from the fact that $\int_0^\infty t^p e^{-2t^2}dt<\infty$ for all $p>0$.

2. Crystal ball condition for marginal

We show the crystal ball condition for marginal: for some $-\infty<a_j\le b_j<\infty$,

$$
\sup_{x_j\in[a_j,b_j]} n^{(p+2)/2}\mathrm{E}_*\left[\left(F_j^*(x_j)-\hat{F}_j(x_j)\right)^{p+2}\right]=O_P(1).
$$

Note that the empirical and resampled CDF, F_j^* and $\hat{F}_j$, are divided into two terms contributed by $\mathbf{Y}$ and $\mathbf{X}_j$. That is, we have

$$
F_j^*=\frac{n}{n+n_j}F_{nj}^{\mathbf{Y},*}+\frac{n_j}{n+n_j}F_{n_jj}^{*}\ \text{ and }\ \hat{F}_j=\frac{n}{n+n_j}F_{nj}^{\mathbf{Y}}+\frac{n_j}{n+n_j}F_{n_jj},
$$

where $F_{nj}^{\mathbf{Y},*}$ and $F_{n_jj}^{*}$ resampled distributions drawn by $\mathcal{Y}$ and $\mathcal{X}_j$ each. Combining this with the Minkowski inequality, (EC.3.7) can be bounded from above as

$$
\begin{aligned}
&\sup_{x_j\in[a_j,b_j]} n^{(p+2)/2}\mathrm{E}_*\left[\left(F_j^*(x_j)-\hat{F}_j(x_j)\right)^{p+2}\right]\\
&\lesssim \sup_{x_j\in[a_j,b_j]} n^{(p+2)/2}\mathrm{E}_*\left[\left(F_{n_jj}^*(x_j)-F_{n_jj}(x_j)\right)^{p+2}\right]\\
&\quad+\sup_{x_j\in[a_j,b_j]} n^{(p+2)/2}\mathrm{E}_*\left[\left(F_{nj}^{\mathbf{Y},*}(x_j)-F_{nj}^{\mathbf{Y}}(x_j)\right)^{p+2}\right].
\end{aligned}
$$

Hence, it is enough to verify that each term in (EC.3.7) is $O_P(1)$. Since both terms exploit the empirical CDF, it suffices to show that the marginal term is $O_P(1)$, and the joint data counterpart can proceed similarly.

Since the case of resampling without replacement can be handled by Hoeffding's Lemma, we focus on the resampling with replacement. Recall $F^*_{n_j j}$ is the empirical CDF based on m_j redrawn data points from the j-th marginal dataset $\{X_{1j}, \ldots, X_{n_j j}\}$. We distinguish the resampled data point by X^*_{ij} for $i = 1, 2, \ldots, m_j$.

When $p = 2$, combining the definition of F^*_j with straightforward expansion yields

$$
\begin{aligned}
& n^2 \mathrm{E}_* \left[\left(F^*_{n_j j}(x_j) - F_{n_j j}(x_j) \right)^4 \right] \\
&= \frac{n^2}{m_j^4} \mathrm{E}_* \left[\sum_{i=1}^{m_j} \left(\mathbf{1}\left\{ X^*_{ij} \le x_j \right\} - F_{n_j j}(x_j) \right)^4 \right] \\
&\quad + 6 \mathrm{E}_* \left[\sum_{k \neq l} \left(\mathbf{1}\left\{ X^*_{kj} \le x_j \right\} - F_{n_j j}(x_j) \right)^2 \left(\mathbf{1}\left\{ X^*_{lj} \le x_j \right\} - F_{n_j j}(x_j) \right)^2 \right] \\
&\le \frac{n^2}{m_j^4} \left\{ 16 m_j + 96 m_j (m_j - 1) \right\} = O_P(1).
\end{aligned}
$$

When $p = 4$, since $m_j F^*_{n_j, j}(x_j)$ follows a binomial distribution with number of trials m_j and success probability $F_{n_j, j}(x_j) = p_j$, one can obtain

$$
\begin{aligned}
n^3 \mathrm{E}_* \left[\left(F^*_{n_j j}(x_j) - F_{n_j j}(x_j) \right)^6 \right] &= \frac{n^3}{m_j^6} m_j p_j (1 - p_j) \times \\
& \Big[1 - 30 p_j (1 - p_j)(1 - 4 p_j (1 - p_j)) + 5 m_j p_j (1 - p_j) [5 - 26 p_j (1 - p_j)] \\
& \quad + 15 m_j^2 p_j^2 (1 - p_j)^2 \Big].
\end{aligned}
$$

From the above result, one can easily confirm $\sup_{x_j \in [a_j, b_j]} n^3 \mathrm{E}_* \left[\left(F^*_{n_j j}(x_j) - F_{n_j j}(x_j) \right)^6 \right] = O_P(1)$, as it converges to a constant as $n \to \infty$. To sum up, we obtain $\sup_{x_j \in [a_j, b_j]} n^{(p+2)/2} \mathrm{E}_* \left[\left(F^*_{n_j j}(x_j) - F_{n_j j}(x_j) \right)^{p+2} \right] < \infty$ for $p \in \{2, 4\}$, as desired.

3. Concentration inequality

Combining (EC.3.7) with triangle inequality, we can see that

$$
\begin{aligned}
& \mathrm{P}_* \left(\left| F^*_j(x_j) - \hat{F}_j(x_j) \right| \ge a_n \right) \\
& \le \mathrm{P}_* \left(\left| F^{\mathbf{Y},*}_{nj}(x_j) - F^{\mathbf{Y}}_{nj}(x_j) \right| \ge a_n \right) + \mathrm{P}_* \left(\left| F^*_{n_j j}(x_j) - F_{n_j j}(x_j) \right| \ge a_n \right).
\end{aligned}
$$

Hence, it is enough to prove the marginal case again.

By Assumption 2, it is straightforward to check $(\sqrt{n} a_n)^{-1} = o(1)$ and $a_n = o(1)$. We aim to show that there exists a constant $b_n = O_P(1)$ such that

$$
\mathrm{P}_* \left(\left| F^*_{n_j j}(x_j) - F_{n_j j}(x_j) \right| \ge a_n \right) \le b_n n^{-(p+2)/2}
$$

for all $x_j \in [a_j, b_j]$. By Chebyshev's inequality, we get

$$\begin{aligned}\mathrm{P}_*\left(F^*_{n_j j}(x_j) - F_{n_j j}(x_j) \geq a_n\right) &\leq \mathrm{P}_*\left(F^*_{n_j j}(x_j) - F_{n_j j}(x_j) \geq \widetilde{a}_j\right)\\ &\leq \exp\left\{-s\widetilde{a}_j\right\} \mathrm{E}_*\left[\exp\left\{s(F^*_j(x_j) - \hat{F}_j(x_j))\right\}\right]\end{aligned}$$

where $\widetilde{a}_j = \sqrt{(p+2)\log m_j/(4m_j)}$. We shall prove that the right-hand side is bounded above by the desired quantity $b_n n^{-(p+2)/2}$ for the resampling with replacement. The other case of no replacement can be taken care of by applying Hoeffding's Lemma again to the expectation part.

When we allow replacement, the independence between redrawn sample points and Hoeffding's inequality imply that

$$\begin{aligned}\mathrm{E}_*\left[\exp\left\{s(F^*_{n_j j}(x_j) - F_{n_j j}(x_j))\right\}\right] &= \mathrm{E}_*\left[\exp\left\{\frac{s}{m_j}\sum_{i=1}^{m_j}\left(\mathbf{1}\left\{X^*_{ij} \leq x_j\right\} - F_{n_j j}(x_j)\right)\right\}\right]\\ &= \prod\nolimits_{i=1}^{m_j} \mathrm{E}_*\left[\exp\left\{\frac{s}{m_j}\left(\mathbf{1}\left\{X^*_{ij} \leq x_j\right\} - F_{n_j j}(x_j)\right)\right\}\right]\\ &\leq \prod\nolimits_{i=1}^{m_j} \exp\left\{\frac{s^2}{8m_j^2}\right\} = \exp\left\{\frac{s^2}{8m_j}\right\}.\end{aligned}$$

Consequently, we obtain an upper bound for the target probability with $s = 4\widetilde{a}_j m_j$ as follows:

$$\mathrm{P}_*\left(F^*_{n_j j}(x_j) - F_{n_j j}(x_j) \geq a_n\right) \leq \exp\left\{-s\widetilde{a}_j + s^2/8m_j\right\} \leq \exp\left\{-2\widetilde{a}_j^2 m_j\right\} = m_j^{-(p+2)/2}.$$

Based on the assumption on data sizes, we can find some b_n such that $m_j^{-(p+2)/2} \leq b_n n^{-(p+2)/2}$.

Since the case $\mathrm{P}_*\left(F^*_j(x_j) - \hat{F}_j(x_j) \leq -a_n\right)$ can be similarly handled, we omit the details. □

EC.3.8. Proof of Corollaries 1 and 2

Proof of Corollary 1. We have already shown that $\gamma n\sigma^2_{n,\mathbf{W}}/\sigma^2 \to 1$ as $n \to \infty$ holds in probability, i.e., $\gamma n\sigma^2_{n,\mathbf{W}} = \sigma^2(1 + o_P(1))$. Hence, it suffices to find technical conditions on B and R that can guarantee $\hat{\sigma}^2_{n,\mathbf{W}}/\sigma^2_{n,\mathbf{W}} \to 1$ as $n \to \infty$ in probability, as functions of n. Recall that $\hat{\sigma}^2_{n,\mathbf{W}}$ is an unbiased estimator of $\sigma^2_{n,\mathbf{W}}$. We first claim the following assertion.

Claim. If we have $\mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}] = o_P(n^{-2})$, then $\hat{\sigma}^2_{n,\mathbf{W}}/\sigma^2_{n,\mathbf{W}} \to 1$ holds in probability. Here, E_{**} and Var_{**} stand for expectation and variance for the simulation and bootstrapping error given input datasets, respectively.

Proof of Claim. Combining $n\sigma^2_{n,\mathbf{W}} = O_P(1)$ with Assumption 3, we have $\mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}]/\sigma^4_{n,\mathbf{W}} = o_P(1)$. For any $k > 0$, we have

$$
\begin{aligned}
&\mathrm{P}\left(|\hat{\sigma}^2_{n,\mathbf{W}}/\sigma^2_{n,\mathbf{W}} - 1| > \epsilon\right) \\
&= \mathrm{P}\left(|\hat{\sigma}^2_{n,\mathbf{W}}/\sigma^2_{n,\mathbf{W}} - 1| > \epsilon, \mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}]/\sigma^4_{n,\mathbf{W}} > k\right) \\
&\quad + \mathrm{P}\left(|\hat{\sigma}^2_{n,\mathbf{W}}/\sigma^2_{n,\mathbf{W}} - 1| > \epsilon, \mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}]/\sigma^4_{n,\mathbf{W}} \le k\right) \\
&\le \mathrm{P}\left(\mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}]/\sigma^4_{n,\mathbf{W}} > k\right) + \mathrm{E}\left[\mathrm{P}_{**}(|\hat{\sigma}^2_{n,\mathbf{W}}/\sigma^2_{n,\mathbf{W}} - 1| > \epsilon)\mathbf{1}\{\mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}]/\sigma^4_{n,\mathbf{W}} \le k\}\right] \\
&\le \mathrm{P}\left(\mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}]/\sigma^4_{n,\mathbf{W}} > k\right) + \mathrm{E}\left[\frac{\mathrm{E}_{**}\left[(\hat{\sigma}^2_{n,\mathbf{W}} - \sigma^2_{n,\mathbf{W}})^2\right]}{\epsilon^2\sigma^4_{n,\mathbf{W}}}\mathbf{1}\{\mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}]/\sigma^4_{n,\mathbf{W}} \le k\}\right] \\
&\le \mathrm{P}\left(\mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}]/\sigma^4_{n,\mathbf{W}} > k\right) + \frac{k}{\epsilon^2}.
\end{aligned}
$$

Taking $n \to \infty$ yields $\limsup_{n\to\infty} \mathrm{P}\left(|\hat{\sigma}^2_{n,\mathbf{W}}/\sigma^2_{n,\mathbf{W}} - 1| > \epsilon\right) \le \frac{k}{\epsilon^2}$ because $\mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}]/\sigma^4_{n,\mathbf{W}} = o_P(1)$. As k is arbitrary, we finally have $\lim_{n\to\infty} \mathrm{P}\left(|\hat{\sigma}^2_{n,\mathbf{W}}/\sigma^2_{n,\mathbf{W}} - 1| > \epsilon\right) = 0$, which completes the proof. □

Analogous to Lam and Qian [5], let us decompose the simulation output as follows:

$$
\begin{aligned}
\delta_{\mathsf{sim}} &= h(\mathbf{Y}_r(1), \dots, \mathbf{Y}_r(T)) - \mathrm{E}_{\mathbf{Y}(t)\sim F_{\mathbf{W}}}[h(\mathbf{Y}(1), \dots, \mathbf{Y}(T))], \\
\delta_{\mathsf{input}} &= \mathrm{E}_{\mathbf{Y}(t)\sim F_{\mathbf{W}}}[h(\mathbf{Y}(1), \dots, \mathbf{Y}(T))] - \mathrm{E}_{\mathbf{Y}(t)\sim \hat{F}}[h(\mathbf{Y}(1), \dots, \mathbf{Y}(T))] \\
&= \eta(F_{\mathbf{W}}) - \eta(\hat{F}),
\end{aligned}
$$

where $\{Y_r(t)\}_{1\le t\le T} \sim F_{\mathbf{W}}$ are a sample of i.i.d. simulation outputs drawn from $F_{\mathbf{W}}$. Here δ_{sim} and δ_{input} represent the error from the simulation experiment and bootstrapping of the input model, respectively. Then, from Equation (10) of Sun et al. [10], we have

$$
\begin{aligned}
&\mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}] \\
&= \frac{1}{B}(\mathrm{E}_{\mathbf{W}}[\delta^4_{\mathsf{input}}] - \mathrm{E}_{\mathbf{W}}[\delta^2_{\mathsf{input}}]^2) + \frac{2}{B(B-1)}\mathrm{E}_{\mathbf{W}}[\delta^2_{\mathsf{input}}]^2 \\
&\quad + \frac{2}{B^2R^2(B-1)}\mathrm{E}_{**}[\delta^2_{\mathsf{sim}}]^2 + \frac{2}{B^2R^3}\mathrm{E}_{**}[\delta^4_{\mathsf{sim}}] \\
&\quad + \frac{2(B+1)}{B^2R(B-1)}\mathrm{E}_{\mathbf{W}}[\delta^2_{\mathsf{input}}]\mathrm{E}_{**}[\delta^2_{\mathsf{sim}}] + \frac{2(BR^2+R^2-4R+3)}{B^2R^3(R-1)}\mathrm{E}_{\mathbf{W}}[\mathrm{E}_{\mathsf{sim}}[\delta^2_{\mathsf{sim}}|F_{\mathbf{W}}]^2] \\
&\quad + \frac{4B+2}{B^2R}\mathrm{E}_{**}[\delta^2_{\mathsf{input}}\delta^2_{\mathsf{sim}}] + \frac{4}{B^2R^2}\mathrm{E}_{**}[\delta_{\mathsf{input}}\delta^3_{\mathsf{sim}}].
\end{aligned}
$$

Here $\mathrm{E}_{\mathsf{sim}}$ represents an expectation conditional on input datasets and $\mathbf{W}$. From the proof of Theorem 6, we have $\gamma n\mathrm{E}_{\mathbf{W}}[\delta^2_{\mathsf{input}}] = \sigma^2(1 + o_P(1))$. Under Assumptions 4 and 6 with $p = 4$, one can confirm L^4 convergence, i.e., $(\gamma n)^2\mathrm{E}_{\mathbf{W}}[\delta^4_{\mathsf{input}}] = O_P(1)$. Additionally, since Assumption 3 implies that h is a bounded function, $\mathrm{E}_{**}[\delta^2_{\mathsf{sim}}] = O(1)$, $\mathrm{E}_{**}[\delta^4_{\mathsf{sim}}] = O(1)$, and $\mathrm{E}_{\mathbf{W}}[\mathrm{E}_{\mathsf{sim}}[\delta^2_{\mathsf{sim}}|F_{\mathbf{W}}]^2] = O(1)$ hold. Consequently, ignoring the constant and considering the leading factors, we finally get

$$\mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}] = O_P\left(\frac{1}{Bn^2} + \frac{1}{BR^2} + \frac{1}{BRn}\right),$$

which is identical to Lam and Qian [5]. From this result, it can be easily seen that $B = \omega(1)$ and $R = \omega(n)$ imply $n^2\mathrm{Var}_{**}[\hat{\sigma}^2_{n,\mathbf{W}}] = o_P(1)$. Combining this with the previous claim proves the statement. □

Proof of Corollary 2. To show this statement, it suffices to show the following central limit theorem of $\hat{\eta}(\hat{F})$, i.e.,

$$\frac{\hat{\eta}(\hat{F}) - \eta(F)}{\sqrt{\gamma\hat{\sigma}^2_{n,\mathbf{W}} + \hat{\sigma}^2_{\mathrm{sim}}/M}} \Longrightarrow N(0,1).$$

From Corollary 1, we have $\gamma n\sigma^2_{n,\mathbf{W}} = \sigma^2(1+o_P(1))$. Combining this with Theorem 3 and Slutsky's Theorem yields $\frac{\eta(\hat{F})-\eta(F)}{\sqrt{\gamma\sigma^2_{n,\mathbf{W}}}} \Rightarrow N(0,1)$ as $n \to \infty$. Further, observe that $\mathrm{Var}[\sqrt{n}(\hat{\eta}(\hat{F}) - \eta(\hat{F}))] = \frac{n}{M}\mathrm{E}[\mathrm{Var}[\eta_r(\hat{F})|\hat{F}]] = O(n/M)$ as the performance measure h is assumed to be defined on the compact support in Assumption 3. Thus, if $M = \omega(n)$, we obtain $\sqrt{n}(\hat{\eta}(\hat{F}) - \eta(\hat{F})) = o_P(1)$.

Putting all the results together, we can get

$$\begin{aligned}
\frac{\hat{\eta}(\hat{F}) - \eta(F)}{\sqrt{\gamma\hat{\sigma}^2_{n,\mathbf{W}} + \hat{\sigma}^2_{\mathrm{sim}}/M}} &= \frac{\sqrt{n}(\hat{\eta}(\hat{F}) - \eta(\hat{F}))}{\sqrt{\gamma n\hat{\sigma}^2_{n,\mathbf{W}} + n\hat{\sigma}^2_{\mathrm{sim}}/M}} + \frac{\sqrt{\gamma n\sigma^2_{n,\mathbf{W}}}}{\sqrt{\gamma n\hat{\sigma}^2_{n,\mathbf{W}} + n\hat{\sigma}^2_{\mathrm{sim}}/M}}\frac{(\eta(\hat{F}) - \eta(F))}{\sqrt{\gamma\sigma^2_{n,\mathbf{W}}}} \\
&= \frac{o_P(1)}{\sqrt{\sigma^2(1+o_P(1)) + o(1)}} + \frac{\sqrt{\sigma^2(1+o_P(1))}}{\sqrt{\sigma^2(1+o_P(1)) + o(1)}}\frac{(\eta(\hat{F}) - \eta(F))}{\sqrt{\gamma\sigma^2_{n,\mathbf{W}}}} \\
&\Longrightarrow N(0,1),
\end{aligned}$$

where the asymptotic notation is derived from $\gamma n\sigma^2_{n,\mathbf{W}} = \sigma^2(1+o_P(1))$ and $n\hat{\sigma}^2_{\mathsf{sim}}/M = O(n/M) = o(1)$. This completes the proof. □

EC.3.9. Proofs of Theorems 8 and Corollary 3 In this subsection, we study the weak and moment convergence results that are required to justify the variance decomposition results introduced in Section 6.2. Recall that we define two functions, $\widetilde{F}^{\mathbf{Y}}_j$ and $\widetilde{F}^{X_j}_j$, as

$$\widetilde{F}^{\mathbf{Y}}_j = \frac{n}{n+n_j}F^{\mathbf{Y}}_{nj} + \frac{n_j}{n+n_j}F_j \quad \text{and} \quad \widetilde{F}^{X_j}_j = \frac{n}{n+n_j}F_j + \frac{n_j}{n+n_j}F_{n_j j}.$$

for any $1 \le j \le d$. In the following, we will investigate the weak convergence of two empirical processes induced by the following two distribution functions defined as

$$\widetilde{F}^{\mathbf{Y}} := \bar{C}_n^{\mathbf{Y}}(\widetilde{F}_1^{\mathbf{Y}}, \ldots, \widetilde{F}_d^{\mathbf{Y}}) \text{ and } \widetilde{F}^{X_j} := C(F_1, \ldots, \widetilde{F}_j^{X_j}, \ldots, F_d).$$

Furthermore, as counterparts of Theorems 5 and 6, we will provide the convergence of $\eta(\widetilde{F}^{\mathbf{Y}})$ and $\eta(\widetilde{F}^{X_j})$, as well as the convergence of their respective Monte Carlo estimators.

Proof of Theorem 8. For each $j \in \{1, 2, \ldots, d\}$, (6) and Donsker's Theorem yield

$$\sqrt{n}\left[(F_n^{\mathbf{Y}}, \widetilde{F}_1^{\mathbf{Y}}, \widetilde{F}_2^{\mathbf{Y}}, \ldots, \widetilde{F}_d^{\mathbf{Y}}) - (F, F_1, F_2, \ldots, F_d)\right] \Longrightarrow (\mathcal{GP}, w_1^2\mathcal{GP}_1, \ldots, w_d^2\mathcal{GP}_d)$$
$$\sqrt{n_j}\left[(F, F_1, F_2, \ldots, \widetilde{F}_j^{X_j}, \ldots, F_d) - (F, F_1, F_2, \ldots, F_d)\right] \Longrightarrow (0, 0, \ldots, (1 - w_j^2)\mathbb{F}_j^{X_j}, \ldots, 0)$$

holds, where $\mathcal{GP}$ and $\mathbb{F}_j^{X_j}$ are Gaussian processes defined in Assumptions 5 and (6), respectively, and recall that $\mathcal{GP}_j$ is defined as the j-th marginal Gaussian process satisfying $\mathcal{GP}_j(x_j) = \mathcal{GP}(\infty, \ldots, x_j, \ldots, \infty)$ or simply $\mathbb{F}_j^{\mathbf{Y}}$ introduced in Section 3. Furthermore, a straightforward computation yields $\widetilde{F}^{\mathbf{Y}} = \phi(F_n^{\mathbf{Y}}, \widetilde{F}_1^{\mathbf{Y}}, \widetilde{F}_{n2}^{\mathbf{Y}}, \ldots, \widetilde{F}_{nd}^{\mathbf{Y}}) + o(1)$ and $\widetilde{F}^{X_j} = \phi(F, F_1, \ldots, \widetilde{F}_j^{X_j}, \ldots, F_d)$, where ϕ is the mapping defined in Lemma EC.2 and the asymptotic relation holds from Lemma EC.3. Applying the delta method in Lemma EC.2, we have

$$\begin{aligned}
\sqrt{n}\left[\widetilde{F}^{\mathbf{Y}} - F\right] \Longrightarrow & \phi'_{(F, F_1, \ldots, F_d)}(\mathcal{GP}, w_1^2\mathcal{GP}_1, \ldots, w_d^2\mathcal{GP}_d), \\
& = \mathcal{GP} - \sum_{j=1}^{d} \partial_i C(F_1, \ldots, F_d)\mathcal{GP}_j + \sum_{j=1}^{d} \partial_j C \times w_j^2 \mathcal{GP}_j \\
& = \mathbb{C} + \sum_{j=1}^{d} w_j^2 \partial_j C(F_1, \ldots, F_d)\mathcal{GP}_j \\
\sqrt{n_j}\left[\widetilde{F}^{X_j} - F\right] \Longrightarrow & (1 - w_j^2)\partial_j C(F_1, \ldots, F_d)\mathbb{F}_j^{X_j}.
\end{aligned}$$

Here $\mathcal{GP} - \sum_{j=1}^{d} \partial_j C \mathcal{GP}_j = \mathbb{C}(F_1, \ldots, F_d)$ as in the proof of Theorem 3. Proceeding similarly to the proof of Theorem EC.5, one can easily confirm

$$\sqrt{n} \sup_{\mathbf{x} \in \operatorname{supp} h} \left| \prod_{t=1}^{T} \widetilde{F}_t^{\mathbf{Y}}(\mathbf{x}_t) - \prod_{t=1}^{T} \widetilde{F}_t^{\mathbf{Y}}(\mathbf{x}_t-) \right| \to 0,$$
$$\sqrt{n_j} \sup_{\mathbf{x} \in \operatorname{supp} h} \left| \prod_{t=1}^{T} \widetilde{F}_t^{X_j}(\mathbf{x}_t) - \prod_{t=1}^{T} \widetilde{F}_t^{X_j}(\mathbf{x}_t-) \right| \to 0.$$

Putting all these results together gives us the following two central limit theorems:

$$\sqrt{n}\left[\eta(\widetilde{F}^{\mathbf{Y}}) - \eta(F)\right] \Longrightarrow \Gamma_h\left(\psi'_F\left(\mathbb{C} + \sum_{j=1}^{d} w_j^2 \partial_j C(F_1, \ldots, F_d)\mathcal{GP}_j\right)\right),$$
$$\sqrt{n_j}\left[\eta(\widetilde{F}^{X_j}) - \eta(F)\right] \Longrightarrow \Gamma_h\left(\psi'_F\left((1 - w_j^2)\partial_j C(F_1, \ldots, F_d)\mathbb{F}_j^{X_j}\right)\right),$$

where ψ'_F is the Hadamard derivative of ψ at F introduced in Lemma EC.4. Since Γ_h and ψ'_F are continuous, the continuous mapping theorem yields that both limits in (EC.3.9) are centered Gaussians, i.e., $\sqrt{n}\left[\eta(\widetilde{F}^{\mathbf{Y}})-\eta(F)\right] \Rightarrow N(0,\sigma^2_{\mathbf{Y}})$ and $\sqrt{n_j}\left[\eta(\widetilde{F}^{X_j})-\eta(F)\right] \Rightarrow N(0,\sigma_j^2)$ for some $\sigma_{\mathbf{Y}}$ and σ_j.

Recall that the weak limit of $\sqrt{n}(\eta(\hat{F}) - \eta(F))$ is given as $\psi'_F(\phi'_{(F,F_1,\ldots,F_d)}(\mathcal{GP}, w_1\mathbb{F}_1,\ldots,w_d\mathbb{F}_d)) = \psi'_F(\mathbb{C}(F_1,\ldots,F_d) + \sum_{j=1}^d w_j\partial_j C\mathbb{F}_j)$. Combining this with (6), one can decompose the limit as

$$\begin{aligned}
&\psi'_F(\mathbb{C}(F_1,\ldots,F_d)+\sum\nolimits_{j=1}^d w_j\partial_j C\mathbb{F}_j),\\
&=\psi'_F(\mathbb{C}(F_1,\ldots,F_d)+\sum\nolimits_{j=1}^d w_j^2\partial_j C(F_1,\ldots,F_d)\mathcal{GP}_j\\
&\quad+\sum\nolimits_{j=1}^d w_j\sqrt{1-w_j^2}\partial_j C(F_1,\ldots,F_d)\mathbb{F}_j^{X_j}).
\end{aligned}$$

Taking a linear functional Γ_h on both sides, combining (EC.3.9) with (EC.3.9) results in

$$\begin{aligned}
N(0,\sigma^2) &\stackrel{d}{=} \Gamma_h\left(\psi'_F(\mathbb{C}(F_1,\ldots,F_d)+\sum\nolimits_{j=1}^d w_j\partial_j C\mathbb{F}_j\right),\\
&\stackrel{d}{=} N(0,\sigma^2_{\mathbf{Y}})+\sum\nolimits_{j=1}^d \frac{w_j}{\sqrt{1-w_j^2}}N(0,\sigma_j^2),
\end{aligned}$$

where all Gaussian distributions are mutually independent. The independence of Gaussian distributions implies $\sigma^2=\sigma^2_{\mathbf{Y}}+\sum_{j=1}^d\frac{w_j^2}{1-w_j^2}\sigma_j^2$. Since $\lim_{n\to\infty}\frac{n}{n_j}=\frac{w_j^2}{1-w_j^2}$ and $\sigma_I^2=\frac{\sigma^2}{n}(1+o_P(1))$, we get $\sigma_I^2=\left(\frac{\sigma^2_{\mathbf{Y}}}{n}+\sum_{j=1}^d\frac{\sigma_j^2}{n_j}\right)(1+o_P(1))$, as desired. □

Proof of Corollary 3. Let us define the resampling counterparts of $\widetilde{F}^{\mathbf{Y}}$ and $\widetilde{F}^{X_j}$ as follows:

$$\widetilde{F}^{\mathbf{Y}}_{\mathbf{W}}=\bar{C}^{\mathbf{Y}}_{n,\mathbf{W}_C}(\widetilde{F}^{\mathbf{Y}}_{1,\mathbf{W}_C},\ldots,\widetilde{F}^{\mathbf{Y}}_{d,\mathbf{W}_C}) \text{ and } \widetilde{F}^{X_j}_{\mathbf{W}_j}=\bar{C}^{\mathbf{Y}}_n(\hat{F}_1,\ldots,\widetilde{F}^{X_j}_{j,\mathbf{W}_j},\ldots,\hat{F}_d),$$

where for each $j\in\{1,2,\ldots,d\}$, $\widetilde{F}^{\mathbf{Y}}_j$ and $\widetilde{F}^{X_j}_{\mathbf{W}_j}$ are defined as

$$\widetilde{F}^{\mathbf{Y}}_{j,\mathbf{W}_C}=\frac{n}{n+n_j}F^{\mathbf{Y}}_{j,\mathbf{W}_C}+\frac{n_j}{n+n_j}F_{n_jj} \text{ and } \widetilde{F}^{X_j}_{\mathbf{W}_j}=\frac{n}{n+n_j}F^{\mathbf{Y}}_{nj}+\frac{n_j}{n+n_j}F^{X_j}_{j,\mathbf{W}_j}.$$

Applying the same approach with Corollary 1, it suffices to show the uniform integrability of $\{\sqrt{rn}(\widetilde{F}^{\mathbf{Y}}_{\mathbf{W}}-\hat{F})\}^4$ and $\{\sqrt{rn}(\widetilde{F}^{X_j}_{\mathbf{W}_j}-\hat{F})\}^4$. Proceeding similarly with the proof of Theorem 6, $\widetilde{F}^{\mathbf{Y}}$

and $\widetilde{F}^{X_j}$ counterparts of Assumption 6 with $p=4$ are sufficient conditions to guarantee the said uniform integrability. More specifically, it is sufficient to show

$$
\begin{aligned}
&\sup_{\mathbf{u}\in[0,1]^d} n^{(4+\delta)/2}\mathrm{E}_{\mathbf{W}_C}\left[\left(\bar{C}^{\mathbf{Y}}_{n,\mathbf{W}_C}(\mathbf{u})-\bar{C}^{\mathbf{Y}}_n(\mathbf{u})\right)^{4+\delta}\right]=O_P(1),\\
&\sup_{\mathbf{x}\in\mathrm{supp}\,h}\max_j\; n^{(4+\delta)/2}\mathrm{E}_{\mathbf{W}}\left[\left(\widetilde{F}^{\mathbf{Y}}_{j,\mathbf{W}_C}(x_j)-\hat{F}_j(x_j)\right)^{4+\delta}\right]=O_P(1),\\
&\sup_{\mathbf{x}\in\mathrm{supp}\,h}\max_j\; n^{(4+\delta)/2}\mathrm{E}_{\mathbf{W}}\left[\left(\widetilde{F}^{X_j}_{\mathbf{W}_j}(x_j)-\hat{F}_j(x_j)\right)^{4+\delta}\right]=O_P(1),\\
&\sup_{\mathbf{x}\in\mathrm{supp}\,h}\max_j\; \mathrm{P}_{\mathbf{W}}\left[\left|\widetilde{F}^{\mathbf{Y}}_{j,\mathbf{W}_C}(x_j)-\hat{F}_j(x_j)\right|\geq a_n\right]\leq b_n n^{-\frac{4+\delta}{2}}\\
&\qquad\text{when } a_n=o(1),\frac{1}{\sqrt{n}a_n}=O(1), b_n=O_P(1),\\
&\sup_{\mathbf{x}\in\mathrm{supp}\,h}\max_j\; \mathrm{P}_{\mathbf{W}}\left[\left|\widetilde{F}^{X_j}_{j,\mathbf{W}_j}(x_j)-\hat{F}_j(x_j)\right|\geq a_n\right]\leq b_n n^{-\frac{4+\delta}{2}}\\
&\qquad\text{when } a_n=o(1),\frac{1}{\sqrt{n}a_n}=O(1), b_n=O_P(1).
\end{aligned}
$$

As we made the same assumptions with Theorem 7, the first line is straightforward thanks to Theorem 7. To see the second and third lines, we observe

$$
\widetilde{F}^{\mathbf{Y}}_{j,\mathbf{W}_C}-\hat{F}_j=\frac{n}{n+n_j}(F^{\mathbf{Y}}_{j,\mathbf{W}_C}-F^{\mathbf{Y}}_{nj}) \;\text{ and }\; \widetilde{F}^{X_j}_{j,\mathbf{W}_j}-\hat{F}_j=\frac{n_j}{n+n_j}(F^{X_j}_{j,\mathbf{W}_j}-F_{n_j j}).
$$

Recall that the constant multiples of these terms appeared in (EC.3.7), and we have shown that each of them is $O_P(1)$ with $p=4$. From this, the second and third lines follow.

Similarly, the fourth and fifth lines have been investigated in (EC.3.7), and each of them satisfies the concentration inequality conditions with the same a_n and b_n. Therefore, by taking the same a_n and b_n with Theorem 7, we can show that the fourth and fifth lines of (EC.3.9) hold, as desired. □

EC.3.10. Proof of Theorem 9 (i) In the proof of Theorem 6, we have shown that

$$
\sup_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]}\mathrm{E}_{\mathbf{W}}\left[\left(\sqrt{n}\left|F_{\mathbf{W}}(\mathbf{x})-\hat{F}(\mathbf{x})\right|\right)^{2+\delta}\right]=O_P(1)
$$

for some $\delta>0$. Combining this with Jensen's inequality, we obtain, for each $\mathbf{x}\in[\mathbf{a},\mathbf{b}]$,

$$
\left|\sqrt{n}\mathrm{E}_{\mathbf{W}}[F_{\mathbf{W}}(\mathbf{x})-\hat{F}(\mathbf{x})]\right|^{2+\delta}\leq\mathrm{E}_{\mathbf{W}}\left[\left|\sqrt{n}\left(F_{\mathbf{W}}(\mathbf{x})-\hat{F}(\mathbf{x})\right)\right|^{2+\delta}\right]=O_P(1),
$$

which implies that $\sqrt{n}\{F_{\mathbf{W}}(\mathbf{x})-\hat{F}(\mathbf{x})\}$ is also uniformly integrable. Hence $\sqrt{n}[F_{\mathbf{W}}(\mathbf{x})-\hat{F}(\mathbf{x})]$ converges to a centered Gaussian distribution in L^1 sense for each $\mathbf{x}\in[\mathbf{a},\mathbf{b}]$, and thus $\lim_{n\to\infty}\sqrt{n}\mathrm{E}_{\mathbf{W}}[F_{\mathbf{W}}(\mathbf{x})-\hat{F}(\mathbf{x})]=0$ in probability.

(ii) The second assertion can also be checked from the similar arguments with Theorem 6. Observe that

$$\begin{aligned}
&\left|\sqrt{n}\left[\eta(\mathrm{E}_{\mathbf{W}}[F_{\mathbf{W}}]) - \eta(\hat{F})\right]\right| \\
&\lesssim \left|\Gamma_h\left(\sqrt{n}\sum\nolimits_{t=1}^{T} \mathrm{E}_{\mathbf{W}}[F_{\mathbf{W},t}] - \hat{F}_t\right)\right| + o_P(1) \\
&= |\Gamma|_h\left(\left|\sqrt{n}\left[\sum\nolimits_{t=1}^{T} \mathrm{E}_{\mathbf{W}}[F_{\mathbf{W},t}] - \hat{F}_t\right]\right|\right) + o_P(1) \\
&\lesssim \sup\nolimits_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]} \mathrm{E}_{\mathbf{W}}\left[\left(\sqrt{n}\left|F_{\mathbf{W}}(\mathbf{x}) - \hat{F}(\mathbf{x})\right|\right)^{2+\delta}\right]^{\frac{1}{2+\delta}} + o_P(1) = O_P(1).
\end{aligned}$$

Therefore, we can apply the dominated convergence. That is, we have

$$\begin{aligned}
\lim_{n\to\infty}\left|\sqrt{n}\left[\eta(\mathrm{E}_{\mathbf{W}}[F_{\mathbf{W}}]) - \eta(\hat{F})\right]\right| &= \lim_{n\to\infty}\left|\Gamma_h\left(\sqrt{n}\left[\prod\nolimits_{t=1}^{T}\mathrm{E}_{\mathbf{W}}[F_{\mathbf{W},t}] - \prod\nolimits_{t=1}^{T}\hat{F}_t\right]\right)\right| \\
&= \left|\Gamma_h\left(\lim_{n\to\infty}\sqrt{n}\left[\prod\nolimits_{t=1}^{T}\mathrm{E}_{\mathbf{W}}[F_{\mathbf{W},t}] - \prod\nolimits_{t=1}^{T}\hat{F}_t\right]\right)\right| = 0,
\end{aligned}$$

where the last equality comes from assertion (i). This completes the proof. □

EC.3.11. Comparison between ECDF and ECC in Terms of Statistical Distances This subsection presents a theoretical result for comparing the ECDF and ECC approaches discussed in Section 7.1. Proposition EC.3 below provides limiting distributions of two statistical distances–Cramér von-Mises (CvM) and Anderson-Darling (AD)–for ECDF and ECC, and then compares their respective expectations.

Observe that both CvM and AD can be written as

$$\mathsf{d}(G,F) = \int (G(\mathbf{x}) - F(\mathbf{x}))^2 \rho(\mathbf{x}) dF(\mathbf{x}),$$

where $\rho(\mathbf{x}) = 1$ for CvM and $\rho(\mathbf{x}) = (F(\mathbf{x})(1 - F(\mathbf{x}))^{-1}$ for AD. Proposition EC.3 below provides the limiting distribution of $n\mathsf{d}(G,F)$ when $G \in \{F_n, \hat{F}\}$, along with their respective expectations.

PROPOSITION EC.3. *If distance $\mathsf{d}(G,F)$ is either CvM or AD, then*

$$\begin{aligned}
&n\mathsf{d}(F_n, F) \Longrightarrow \mathcal{X}^{ecdf} := \int_{\mathbb{R}^d} \mathcal{G}\mathcal{P}(\mathbf{x})^2 \rho(\mathbf{x}) dF(\mathbf{x}), \\
&n\mathsf{d}(\hat{F}, F) \\
&\Longrightarrow \mathcal{X}^{ecc} := \int_{\mathbb{R}^d} (\mathbb{C}(F_1,\ldots,F_d)(\mathbf{x}) + \sum_{j=1}^{d} w_j \partial_j C(F_1,\ldots,F_d)(\mathbf{x})\mathbb{F}_j(x_j))^2 \rho(\mathbf{x}) dF(\mathbf{x}).
\end{aligned}$$

The gap between the expectations of the two limiting distributions can be quantified as

$$\mathrm{E}[\mathcal{X}^{ecc}] - \mathrm{E}[\mathcal{X}^{ecdf}] = -\sum_{j=1}^{d}(1 - w_j^2) a_j - 2\sum_{i<j}(1 - w_i^2)(1 - w_j^2) a_{ij},$$

where $\{a_j\}_{1\le j\le d}$ *and* $\{a_{ij}\}_{1\le i,j\le d}$ *are defined as*

$$a_j = \int_{[0,1]^d} \partial_j C(\mathbf{u})(1-u_j)(2C(\mathbf{u}) - u_j\partial_j C(\mathbf{u}))\tilde{\rho}(\mathbf{u})dC(\mathbf{u}),$$
$$a_{ij} = \int_{[0,1]^d} \partial_j C(\mathbf{u})\partial_i C(\mathbf{u})(u_iu_j - C_{i,j}(u_i,u_j))\tilde{\rho}(\mathbf{u})dC(\mathbf{u}),$$

where $\tilde{\rho}(\mathbf{u})$ *and* $C_{i,j}(u_i,u_j)$ *are defined by* $\tilde{\rho}(\mathbf{u}) = \rho(F_1^{-1}(u_1),\dots,F_d^{-1}(u_d))$ *and* $C_{i,j}(u_i,u_j) = C(1,\dots,u_i,\dots,u_j,\dots,1)$*, respectively.*

Proof of Proposition EC.3. From Donsker's Theorem and Theorem 2, we have $\sqrt{n}(F_n - F) \Longrightarrow \mathcal{GP}$ defined in Assumption 5 and $\sqrt{n}(\hat{F} - F) \Longrightarrow \mathbb{C}(F_1,\dots,F_d) + \sum_{j=1}^d w_j\partial_j C(F_1,\dots,F_d)(\mathbf{x})\mathbb{F}_j$. Applying the continuous mapping theorem proves (EC.3).

Now, we turn our attention to show (EC.3). For notational simplicity, we let $\partial_i C_F := \partial_i C(F_1,\dots,F_d)$. By Fubini's Theorem, we can interchange the expectation and integration in $\mathrm{E}[\mathcal{X}^{\mathrm{ecc}}]$ and $\mathrm{E}[\mathcal{X}^{\mathrm{ecdf}}]$, i.e.,

$$\mathrm{E}[\mathcal{X}^{\mathrm{ecdf}}] = \int_{\mathbb{R}^d} \mathrm{E}\left[\mathcal{GP}(\mathbf{x})^2\right]\rho(\mathbf{x})dF(\mathbf{x}),$$

$$\mathrm{E}[\mathcal{X}^{\mathrm{ecc}}] = \int_{\mathbb{R}^d} \mathrm{E}\left[(\mathbb{C}(F_1,\dots,F_d)(\mathbf{x}) + \sum_{j=1}^d w_j\partial_j C_F(\mathbf{x})\mathbb{F}_j(x_j))^2\right]\rho(\mathbf{x})dF(\mathbf{x}).$$

The inner expectation of (EC.3.11) can be calculated as $\mathrm{E}[\mathcal{GP}(\mathbf{x})^2] = F(\mathbf{x})(1 - F(\mathbf{x}))$. Recall $\mathbb{C}(F_1,\dots,F_d)(\mathbf{x}) + \sum_{j=1}^d w_j\partial_j C_F(\mathbf{x})\mathbb{F}_j(x_j) = \mathcal{GP}(\mathbf{x}) + \sum_{j=1}^d (w_j^2 - 1)\partial_j C_F(\mathbf{x})\mathcal{GP}(x_j) + \sum_{j=1}^d w_j\sqrt{1-w_j^2}\partial_j C_F(\mathbf{x})\mathbb{F}_j^{X_j}(x_j)$. Since $\mathcal{GP}(\mathbf{x}) + \sum_{j=1}^d (w_j^2 - 1)\partial_j C_F(\mathbf{x})\mathcal{GP}(x_j), \mathbb{F}_1^{X_1}(x_1),\dots,\mathbb{F}_d^{X_d}(x_d)$ are mutually independent, the inner expectation of $\mathrm{E}[\mathcal{X}^{\mathrm{ecc}}]$ in (EC.3.11) can be written as

$$\begin{aligned}
&\mathrm{E}\left[\left(\mathbb{C}(F_1,\dots,F_d)(\mathbf{x}) + \sum_{j=1}^d w_j\partial_j C_F(\mathbf{x})\mathbb{F}_j(x_j)\right)^2\right] \\
&= \mathrm{E}\left[\left(\mathcal{GP}(\mathbf{x}) + \sum_{j=1}^d (w_j^2 - 1)\partial_j C_F(\mathbf{x})\mathcal{GP}(x_j)\right)^2\right] + \sum_{j=1}^d (w_j^2-1)^2\partial_j C_F(\mathbf{x})^2\mathrm{E}\left[\left(\mathbb{F}_j^{X_j}(x_j)\right)^2\right].
\end{aligned}$$

After some algebra, (EC.3.11) can be reduced to

$$
\begin{aligned}
\text{(EC.3.11)} = \mathrm{E}\left[\mathcal{GP}(\mathbf{x})^2\right] \\
&+ \sum_{j=1}^{d}(w_j^2-1)^2\partial_j C_F(\mathbf{x})^2 \mathrm{E}\left[\mathcal{GP}(x_j)^2\right] + \sum_{j=1}^{d} w_j^2(1-w_j^2)\partial_j C_F(\mathbf{x})^2 \mathrm{E}\left[\mathbb{F}_j^{X_j}(x_j)^2\right] \\
&+ 2\sum_{j=1}^{d}(w_j^2-1)\partial_j C_F(\mathbf{x})\mathrm{E}[\mathcal{GP}(\mathbf{x})\mathcal{GP}_j(x_j)] \\
&+ 2\sum_{1\le i<j\le d}(w_i^2-1)(w_j^2-1)\partial_i C_F(\mathbf{x})\partial_j C_F(\mathbf{x})\mathrm{E}[\mathcal{GP}_i(x_i)\mathcal{GP}_j(x_j)].
\end{aligned}
$$

A straightforward computation based on the covariance structures of two Gaussian processes $\mathcal{GP}$ and $\mathbb{F}_j^{X_j}$ simplifies (EC.3.11) as

$$
\begin{aligned}
&\mathrm{E}\left[\left(\mathbb{C}(F_1,\ldots,F_d)(\mathbf{x}) + \sum_{j=1}^{d} w_j\partial_j C_F(\mathbf{x})\mathbb{F}_j(x_j)\right)^2\right] \\
&= \mathrm{E}\left[\mathcal{GP}(\mathbf{x})^2\right] + \sum_{j=1}^{d}(w_j^2-1)^2\partial_j C_F(\mathbf{x})^2 F_j(x_j)(1-F_j(x_j)) \\
&+ \sum_{j=1}^{d} w_j^2(1-w_j^2)\partial_j C_F(\mathbf{x})^2(1-F_j(x_j))F_j(x_j) + 2\sum_{j=1}^{d}(w_j^2-1)\partial_j C_F(\mathbf{x})F(\mathbf{x})(1-F_j(x_j)) \\
&+ 2\sum_{1\le i<j\le d}(w_i^2-1)(w_j^2-1)\partial_i C_F(\mathbf{x})\partial_j C_F(\mathbf{x})\left(F_{i,j}(x_i,x_j) - F_i(x_i)F_j(x_j)\right) \\
&= \mathrm{E}[\mathcal{GP}(\mathbf{x})^2] + \sum_{j=1}^{d}(1-w_j^2)\partial_j C_F(\mathbf{x})^2 F_j(x_j)(1-F_j(x_j)) \\
&+ 2\sum_{j=1}^{d}(w_j^2-1)\partial_j C_F(\mathbf{x})F(\mathbf{x})(1-F_j(x_j)) \\
&+ 2\sum_{1\le i<j\le d}\partial_i C_F(\mathbf{x})\partial_j C_F(\mathbf{x})(w_i^2-1)(w_j^2-1)(F_{i,j}(x_i,x_j) - F_i(x_i)F_j(x_j)),
\end{aligned}
$$

where $F_{i,j} = F(\infty, \ldots, x_i, \ldots, x_j, \ldots, \infty)$. Combining (EC.3.11) with (EC.3.11) yields

$$\begin{aligned}
&\mathrm{E}[\mathcal{X}^{\mathrm{ecc}}] - \mathrm{E}[\mathcal{X}^{\mathrm{ecdf}}] \\
&= -\sum_{j=1}^{d}(1-w_j^2)\int \partial_j C_F(\mathbf{x})(2F(\mathbf{x}) - F_j(x_j)\partial_j C_F(\mathbf{x}))(1-F_j(x_j))\rho(\mathbf{x})dF(\mathbf{x}) \\
&\quad -2\sum_{1\le i<j\le d}(w_i^2-1)(w_j^2-1)\int \partial_i C_F(\mathbf{x})\partial_j C_F(\mathbf{x})(F_i(x_i)F_j(x_j) - F_{i,j}(x_i,x_j))\rho(\mathbf{x})dF(\mathbf{x}) \\
&= -\sum_{j=1}^{d}(1-w_j^2)a_j - 2\sum_{1\le i<j\le d}(1-w_i^2)(1-w_j^2)a_{ij}.
\end{aligned}$$

The last line follows from the change of variables, letting $u_j = F_j(x_j)$ for all $1 \le j \le d$. □

Proposition EC.3 implies that the differences in statistical distance from the true input model F to the ECC $\hat{F}$ and ECDF F_n can be expressed as a function of the copula C and the asymptotic data size ratio w_j. In particular, a_j and a_{ij} in (EC.3) represent the first-order and second-order effects of additional input data, respectively. Corollary EC.1 below presents some properties of the limiting distribution for general d.

Corollary EC.1. *Under the same assumptions in Proposition EC.3, the following two statements hold:*

(i) If $a_j > 0$, incorporating additional j-th marginal dataset via the ECC is beneficial locally around $w_j = 1$.

(ii) If $C(\mathbf{u}) = \prod_{i=1}^{d} u_i$ (independent input), then we have $C(\mathbf{u}) = u_j\partial_j C(\mathbf{u})$ and $C_{i,j}(u_i, u_j) = u_i u_j$, i.e., $a_j > 0$ and $a_{ij} = 0$ hold for all $1 \le i, j \le d$ for both CvM and AD.

Proof of Corollary EC.1. (i) Define $f(w_1, \ldots, w_d) = \mathrm{E}[\mathcal{X}^{\mathrm{ecc}}] - \mathrm{E}[\mathcal{X}^{\mathrm{ecdf}}]$. Then, $\nabla_{\mathbf{w}} f(w_1, \ldots, w_d)|_{(w_1,\ldots,w_d)=(1,\ldots,1)} > 0$ holds if $a_j > 0$ for all $1 \le j \le d$, which implies that incorporating marginal datasets via ECC is beneficial from input modeling perspective. To show (ii), if $C(\mathbf{u}) = \prod_{i=1}^{d} u_i$ (independent input), then we have $C(\mathbf{u}) = u_j\partial_j C(\mathbf{u})$ and $C_{i,j}(u_i, u_j) = u_i u_j$, i.e., $a_j > 0$ and $a_{ij} = 0$ holds for all $1 \le i, j \le d$ for both CvM and AD, which proves statement (ii). □

Statement (i) of Proposition EC.3 provides a criterion for determining whether to incorporate the additional j-th marginal dataset: specifically, $a_j > 0$ suggests that ECC should be used to exploit the j-th marginal dataset, whereas $a_j \le 0$ indicates that the additional marginal dataset is not beneficial. As a special case, Statement (ii) indicates that when the underlying distribution is independent, integrating additional marginal datasets via ECC by fusing the marginal datasets with the joint data is always advantageous, regardless of d.

As an illustrative example, we focus on the case when $d = 2$ discussed in Section 7.1. Specifically, for both Clayton and Gaussian copulas, we plot a_1 and a_{12} according to the dependence parameter, say ρ. Given ρ, each a_1 and a_{12} are estimated via Monte Carlo simulation by obtaining 10^6 samples from copula C with parameter ρ. Figures EC.2 and EC.3 summarize the experiments when the underlying copulas are Clayton and Gaussian copulas, respectively. Since both Clayton and Gaussian copulas are symmetric, we have $a_1 = a_2$ so that only a_1 is presented.

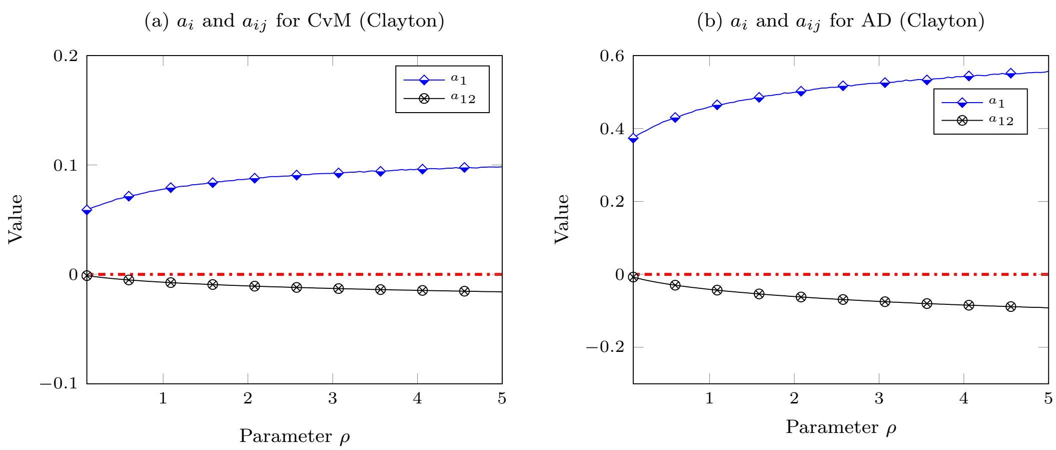

FIGURE EC.2. (Clayton) a_1 and a_{12} according to parameter ρ.

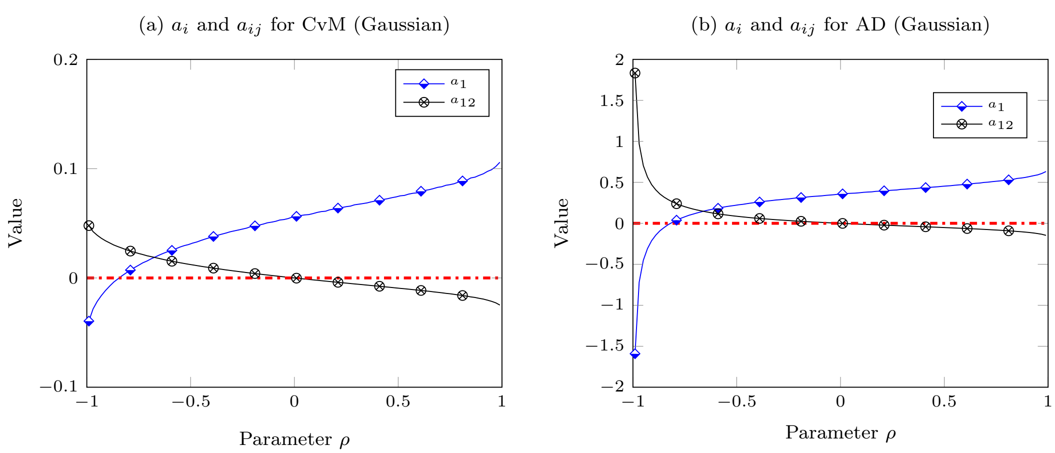

FIGURE EC.3. (Gaussian) a_1 and a_{12} according to parameter ρ.

When $d = 2$, (EC.3) can be reduced to $\mathrm{E}[\mathcal{X}^{\text{ecc}}] - \mathrm{E}[\mathcal{X}^{\text{ecdf}}] = -(1 - w_1^2)a_1 - (1 - w_2^2)a_2 - 2(1 - w_1^2)(1 - w_2^2)a_{12}$. From Figure EC.2, we can confirm $|a_1| > |a_{12}|$ for both CvM and AD, which in turn implies $\mathrm{E}[\mathcal{X}^{\text{ecc}}] - \mathrm{E}[\mathcal{X}^{\text{ecdf}}] < 0$ for any $w_1, w_2 \in [0, 1]$. That is, if the underlying copula is Clayton, the ECC always outperforms ECDF in terms of CvM and AD distances. On the other hand, from Figure EC.3, a similar conclusion can be made for the nonnegatively correlated case, i.e., $\rho \geq 0$. However, under the extremely negatively correlated scenario, $a_1 < 0$ holds, and there exists a combination of w_1 and w_2 where ECC underperforms ECDF in terms of both distances. However, the ECC-based approach is superior from an input modeling perspective for a wider range of ρ.

Appendix EC.4: Additional Numerical Results This section presents additional numerical results for the experiments conducted in Sections 4 and 7.

EC.4.1. Coverage Tests in the Presence of Model Misspecification (Section 4) Figure EC.4 reports the results of the same experiment in Section 4, except that the true dependence structure is not Gaussian but a mixture copula $C = 0.2C_1 + 0.8C_2$, where C_1 is a t-copula with $(\nu, \Sigma) = (10, \Sigma)$ and C_2 is a Clayton copula with $\rho = 3$, with Σ the same as in Section 4. The results in Figure EC.4 are similar to Figure 4.

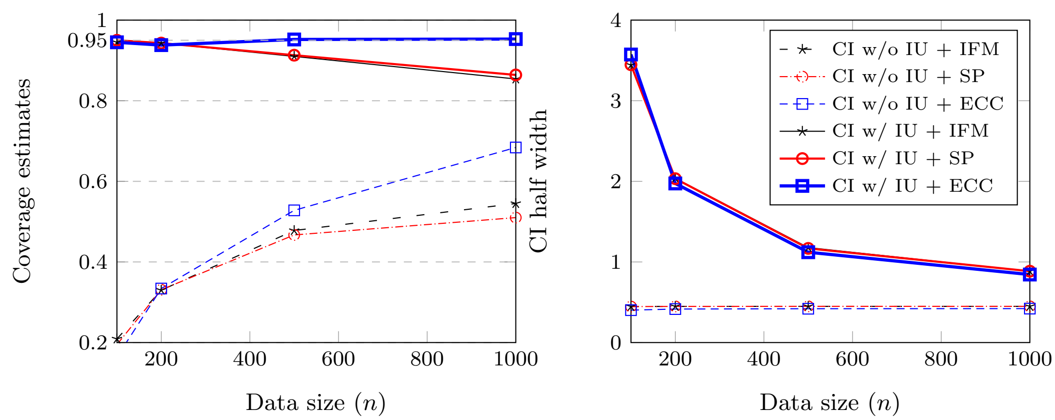

FIGURE EC.4. Coverage test and CI half-width for 95% CI using (11) and (12) under model misspecification

EC.4.2. Clayton Copula Case in Section 7.1 The extra empirical results omitted in Section 7.1 for different (B, R, M) are given in this subsection.

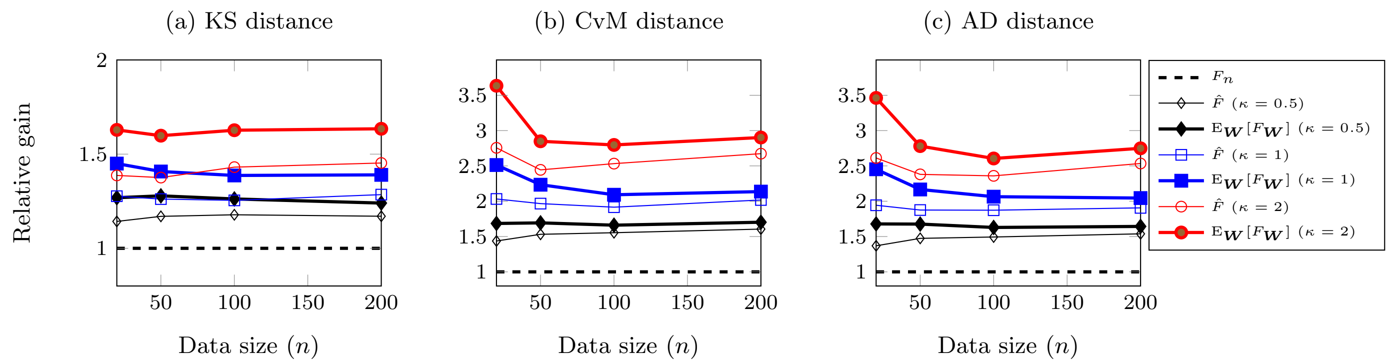

FIGURE EC.5. Relative gain when the underlying copula is Clayton with $\rho = 1$: $\mathrm{d}^{\mathrm{KS}}(G,F) := \max_{\mathbf{x}\in\mathbb{R}^d} |G(\mathbf{x}) - F(\mathbf{x})|$; $\mathrm{d}^{\mathrm{AD}}(G,F) := \int (G(\mathbf{x}) - F(\mathbf{x}))^2/[F(\mathbf{x})(1-F(\mathbf{x}))]\, dF(\mathbf{x})$

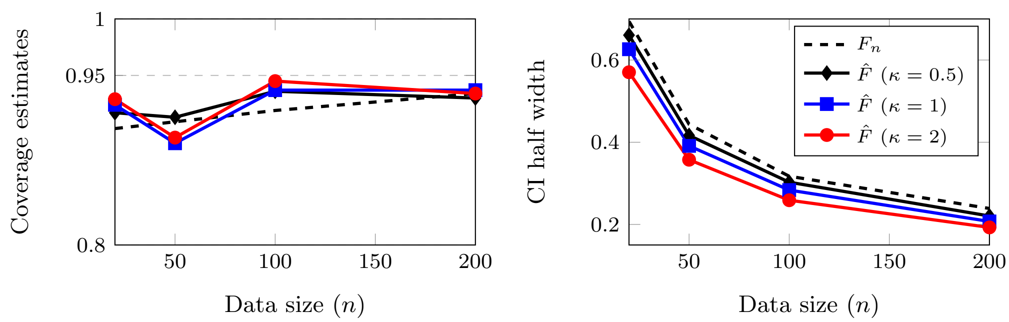

FIGURE EC.6. (Clayton) Comparison of ECDF and ECC when $(B,R,M) = (200,5,500)$

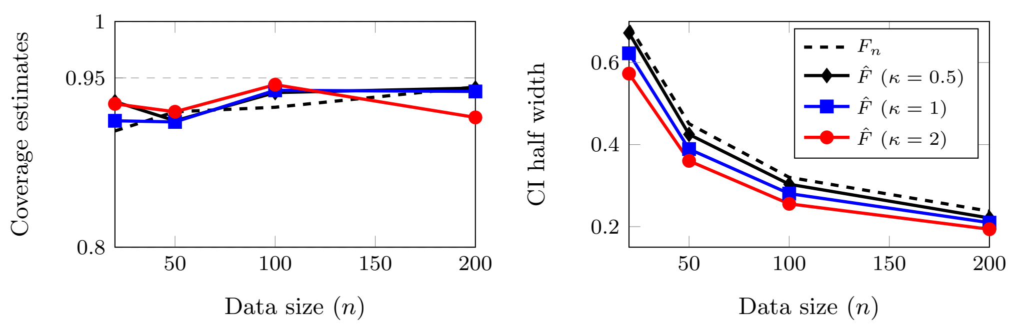

FIGURE EC.7. (Clayton) Comparison of ECDF and ECC when $(B,R,M) = (50,20,500)$

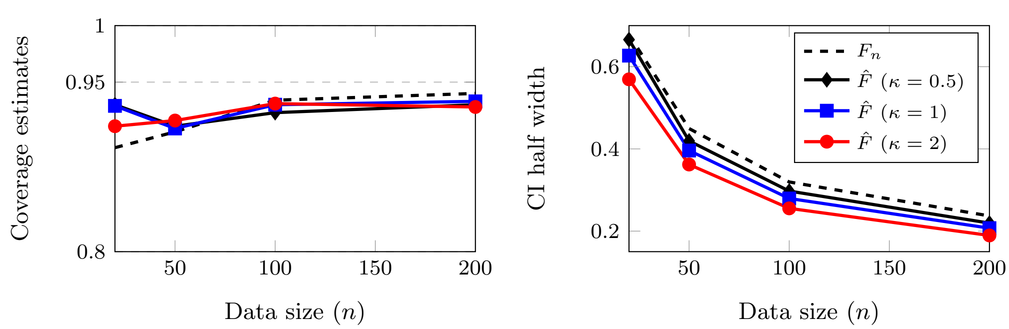

FIGURE EC.8. (Clayton) Comparison of ECDF and ECC when $(B,R,M) = (20,50,500)$

EC.4.3. Gaussian Copula Case in Section 7.1 The empirical results for the case where the underlying copula is Gaussian, which are omitted in Section 7.1, are presented in this section. We have tested different combinations of (B, R, M) and κ parallel to the Clayton copula case.

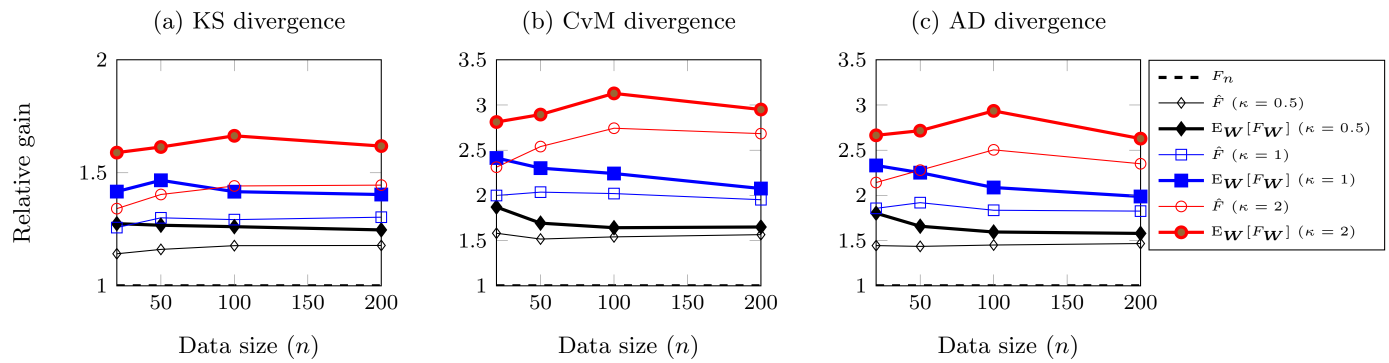

FIGURE EC.9. Relative bias when the underlying copula is a Gaussian copula with $\rho = 0.5$: $\mathsf{d}^{\mathrm{KS}}(G, F) := \max_{\mathbf{x}\in\mathbb{R}^d} |G(\mathbf{x}) - F(\mathbf{x})|$; $\mathsf{d}^{\mathrm{AD}}(G, F) := \int (G(\mathbf{x}) - F(\mathbf{x}))^2 / [F(\mathbf{x})(1 - F(\mathbf{x}))]\, dF(\mathbf{x})$

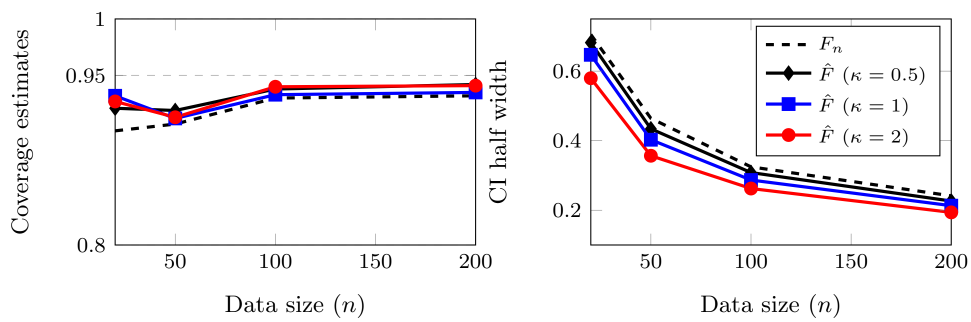

FIGURE EC.10. (Gaussian) Comparison of ECDF and ECC when $(B, R, M) = (100, 10, 500)$

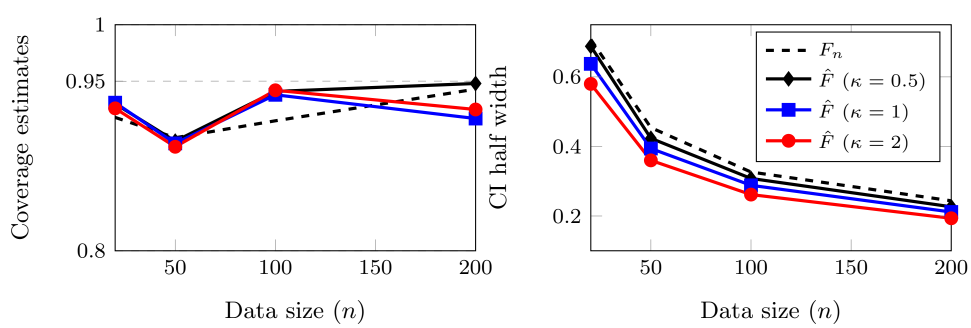

FIGURE EC.11. (Gaussian) Comparison of ECDF and ECC when $(B, R, M) = (200, 5, 500)$

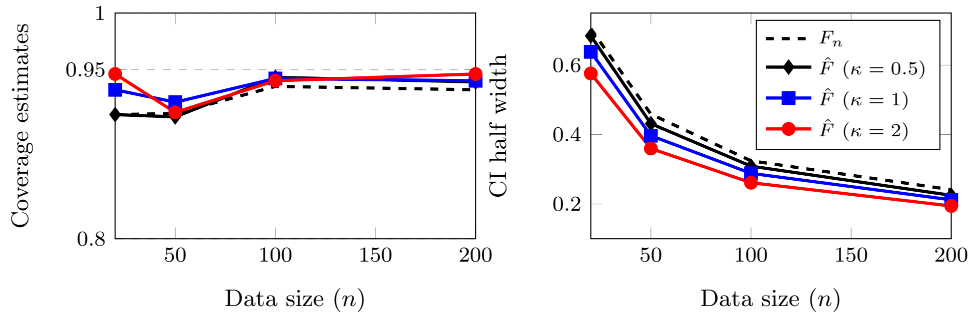

FIGURE EC.12. (Gaussian) Comparison of ECDF and ECC when $(B,R,M) = (50,20,500)$

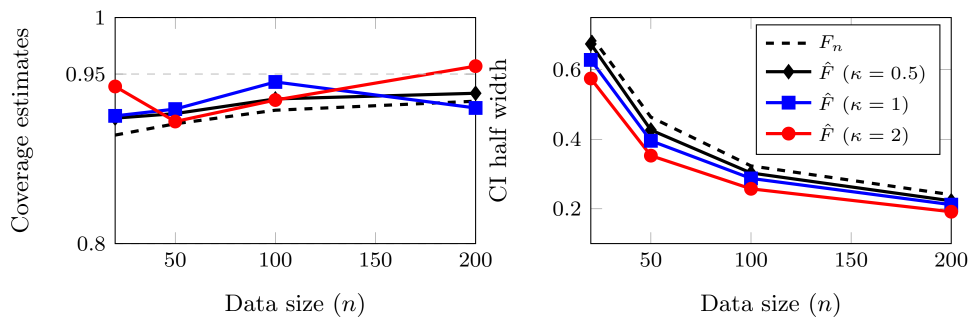

FIGURE EC.13. (Gaussian) Comparison of ECDF and ECC when $(B,R,M) = (20,50,500)$

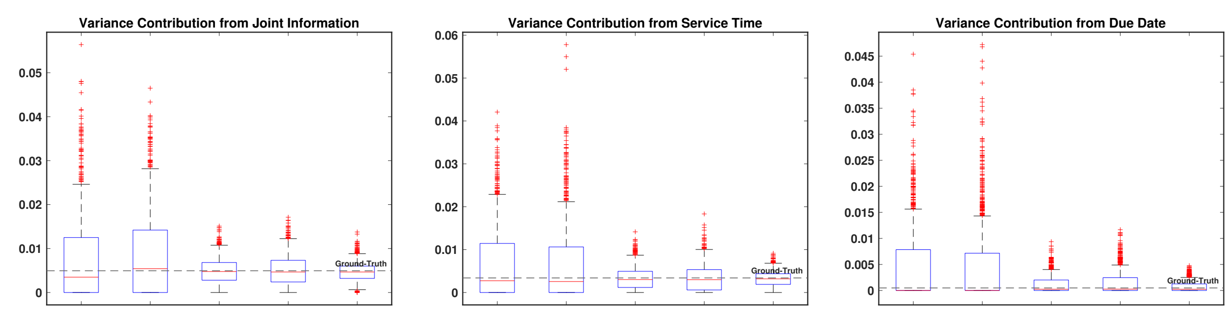

FIGURE EC.14. Boxplots of estimated contributions when $(n, n_S, n_T) = (200, 400, 400)$ and C is Clayton (Plan 1: Classical bootstrap, 2: Bayesian bootstrap, 3: Subsampling with replacement ($\gamma' = 0.2$), 4: Subsampling without replacement ($\gamma' = 0.2$), and 5: Dirichlet subsampling ($(\gamma', a) = (0.2, 1)$))

EC.4.4. Comparison of Resampling Plans Under Different Scenarios. Figures EC.15–EC.21 display the numerical results for different combinations of (B,R) and copulas when $\kappa = 2$.

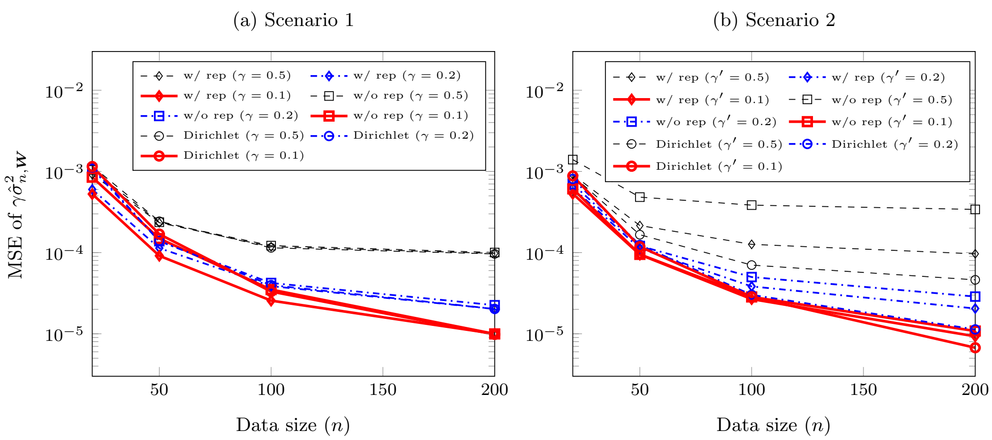

FIGURE EC.15. (Clayton, $(B,R) = (200,5)$) MSE of $\gamma\hat{\sigma}^2_{n,\mathbf{W}}$ for different $\mathbf{W}$ and scenarios.

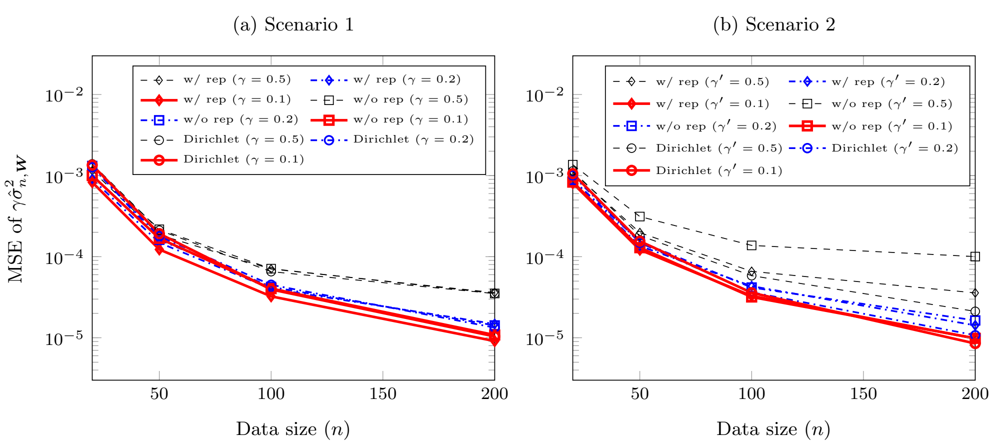

FIGURE EC.16. (Clayton, $(B,R) = (50,20)$) MSE of $\gamma\hat{\sigma}^2_{n,\mathbf{W}}$ for different $\mathbf{W}$ and scenarios.

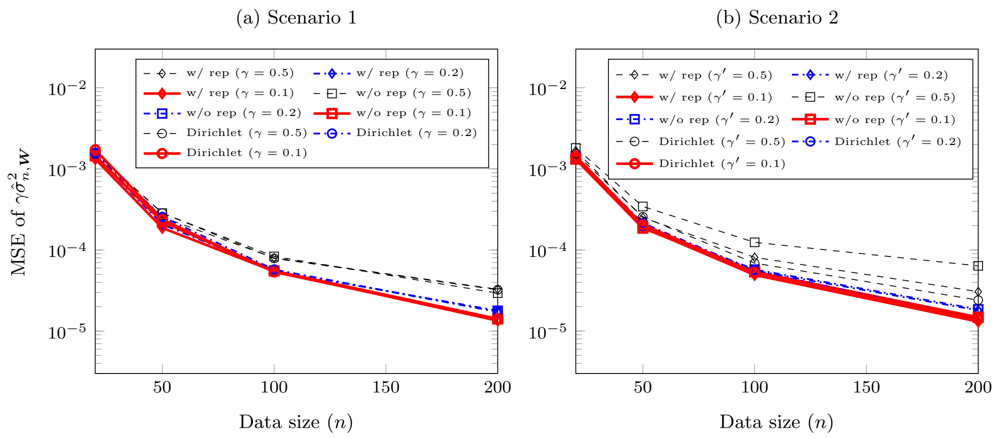

FIGURE EC.17. (Clayton, $(B, R) = (20, 50)$) MSE of $\gamma\hat{\sigma}^2_{n,\mathbf{W}}$ for different $\mathbf{W}$ and scenarios.

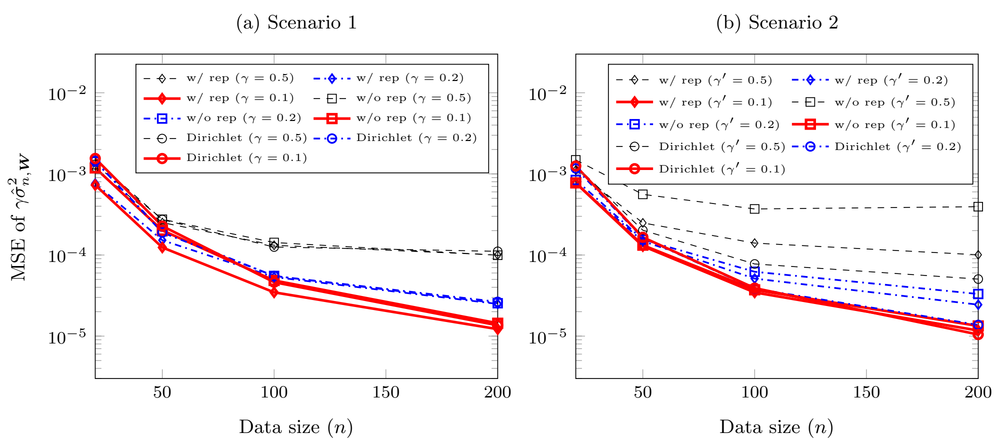

FIGURE EC.18. (Gaussian, $(B, R) = (200, 5)$) MSE of $\gamma\hat{\sigma}^2_{n,\mathbf{W}}$ for different $\mathbf{W}$ and scenarios.

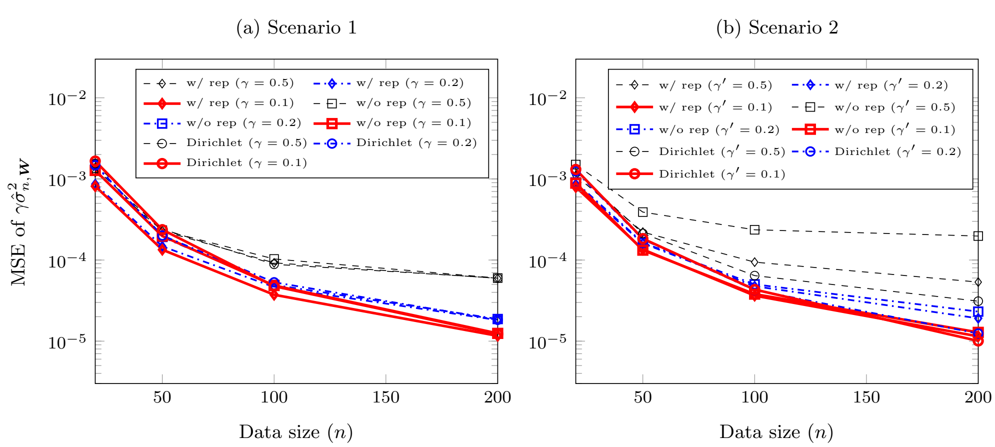

FIGURE EC.19. (Gaussian, $(B, R) = (100, 10)$) MSE of $\gamma\hat{\sigma}^2_{n,\mathbf{W}}$ for different $\mathbf{W}$ and scenarios.

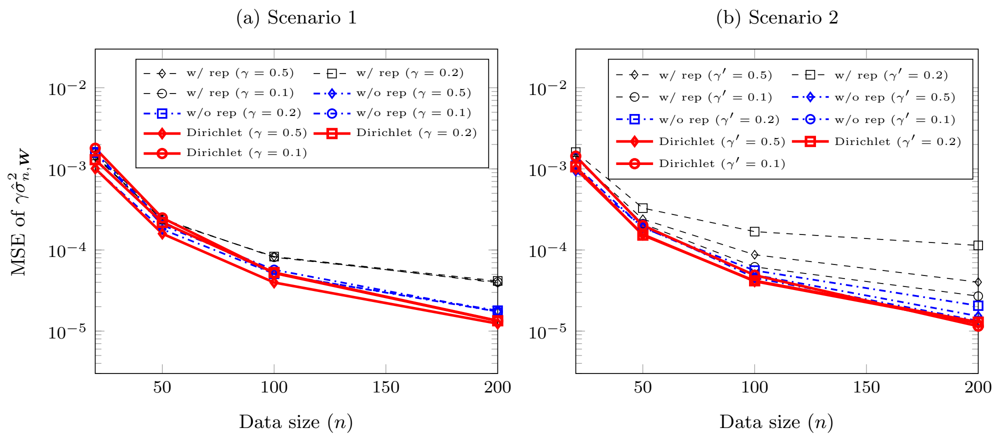

Figure EC.20. (Gaussian, $(B,R) = (50,20)$) MSE of $\gamma\hat{\sigma}^2_{n,\mathbf{W}}$ for different $\mathbf{W}$ and scenarios.

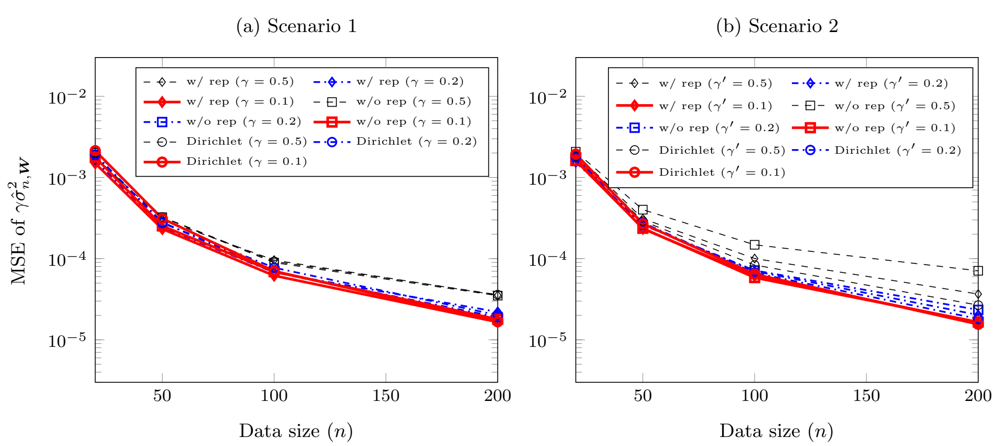

Figure EC.21. (Gaussian, $(B,R) = (20,50)$) MSE of $\gamma\hat{\sigma}^2_{n,\mathbf{W}}$ for different $\mathbf{W}$ and scenarios.

EC.4.5. Additional Experiments for Section 7.2 This section presents the numerical results for in-the-money and out-of-the money cases. For the in-the-money case summarized in Figure EC.22 and Table EC.1, we reach the same conclusion as for the at-the-money case, whereas different behaviors are observed for the out-of-the-money case.

As shown in Figure EC.23, the MoM method outperforms other methods in terms of input variance and MSE for the entire range of n up to 1,000. This is because the out-of-the-money case depends on the tail region. In such scenarios, even if the model is misspecified, the parametric model tends to outperform the nonparametric approach for relatively small n due to the scarcity of data in the tail region. However, since the squared bias of the MoM method does not appear to decrease as n increases, its MSE will also not diminish for large n, so that the ECC-bagged model should eventually outperform the MoM method for sufficiently large n.

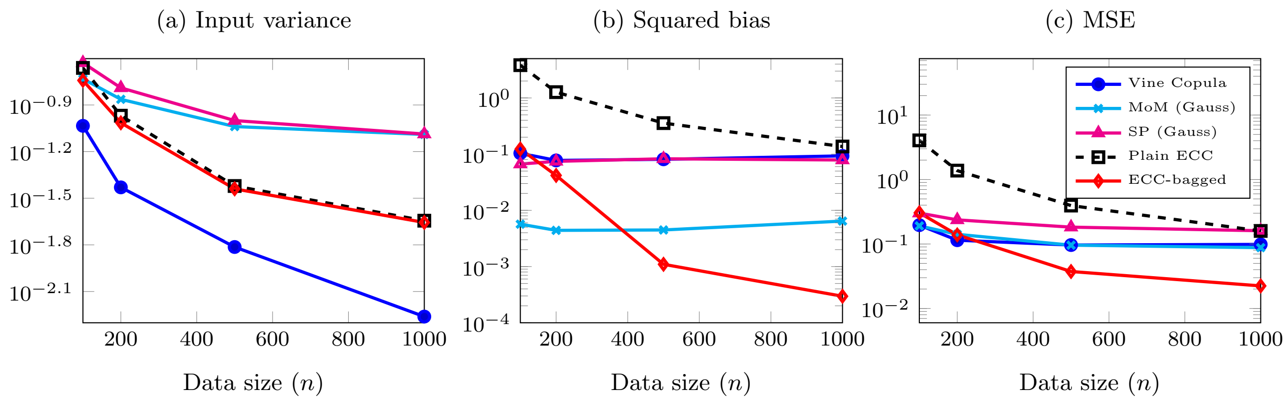

FIGURE EC.22. Comparison of $\hat{F}, \mathrm{E}_{\mathbf{W}}[F_{\mathbf{W}}], F_n^{\text{vine}}$ when $K = 150$

TABLE EC.1. Relative bias of $\eta(\mathrm{E}_{\mathbf{W}}[F_{\mathbf{W}}])$ when $K = 150$

Data size	$\gamma = 1$			$\gamma = 0.2$		
	Sub w/ rep	Sub w/o rep	Bayesian	Sub w/ rep	Sub w/o rep	Dirichlet
$n = 100$	0.54%	0.74%	0.37%	12.77%	13.02%	10.58%
$n = 200$	0.29%	0.39%	0.21%	7.21%	7.31%	6.22%
$n = 500$	0.0076%	0.11%	0.036%	3.28%	3.37%	2.89%
$n = 1000$	0.0015%	0.036%	0.002%	1.79%	1.86%	1.59%

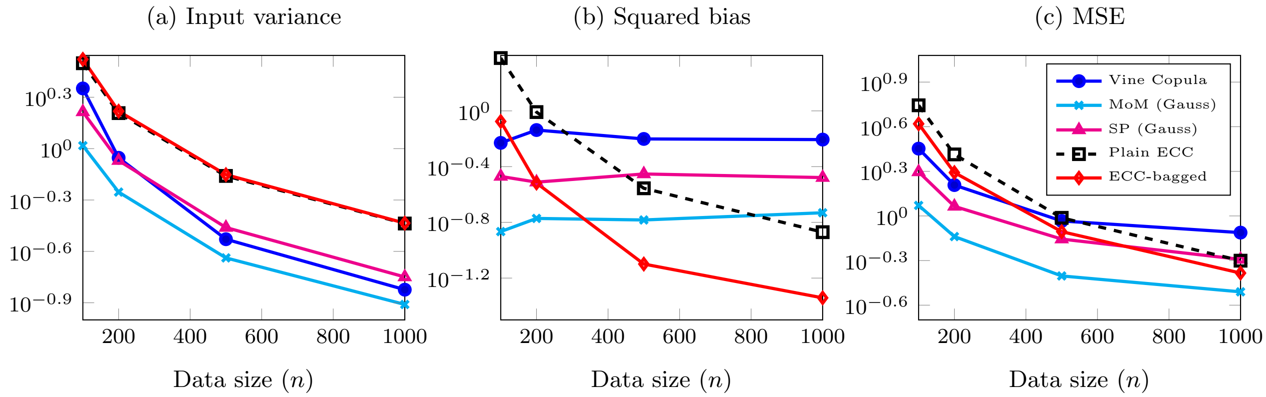

FIGURE EC.23. Comparison of $\hat{F}$ (Plain ECC), $\mathrm{E}_{\mathbf{W}}[F_{\mathbf{W}}]$ (ECC-bagged), and F_n^{vine} ($K = 300, d = 10$)

TABLE EC.2. Relative Bias of $\eta(\mathrm{E}_{\mathbf{W}}[F_{\mathbf{W}}])$ for different resampling plans ($K = 300$)

Data size	$\gamma = 1$			$\gamma = 0.2$		
	Sub w/ rep	Sub w/o rep	Bayesian	Sub w/ rep	Sub w/o rep	Dirichlet
$n = 100$	9.39%	8.16%	10.31%	34.46%	36.7%	22.5%
$n = 200$	5.63%	4.97%	6.18%	22.08%	23.0%	15.5%
$n = 500$	2.42%	2.11%	2.68%	12.14%	12.8%	9.11%
$n = 1000$	2.07%	1.88%	2.21%	6.75%	7.24%	5.07%

References

[1] Aistleitner C, Dick J (2015) Functions of bounded variation, signed measures, and a general Koksma-Hlawka inequality. *Acta Arithmetica* 167(2):143–171.

[2] Bücher A, Volgushev S (2013) Empirical and sequential empirical copula processes under serial dependence. *Journal of Multivariate Analysis* 119:61–70.

[3] Einmahl JHJ (1986) *Multivariate Empirical Processes*. Ph.D. thesis, Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica, Amsterdam.

[4] Fermanian JD, Radulović D, Wegkamp M (2004) Weak convergence of empirical copula processes. *Bernoulli* 10(5):847–860.

[5] Lam H, Qian H (2022) Subsampling to enhance efficiency in input uncertainty quantification. *Operations Research* 70(3):1891–1913.

[6] Owen AB (2005) Multidimensional variation for quasi-Monte Carlo. Fan J, Li G, eds. *Contemporary Multivariate Analysis and Design of Experiments* (World Scientific, Singapore), 49–74.

[7] Radulović D, Wegkamp M, Zhao Y (2017) Weak convergence of empirical copula processes indexed by functions. *Bernoulli* 23(4B):3346–3384.

[8] Segers J (2012) Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli* 18(3):764–782.

[9] Shorack GR, Wellner JA (2009) *Empirical Processes with Applications to Statistics* (SIAM).

[10] Sun Y, Apley DW, Staum J (2011) Efficient nested simulation for estimating the variance of a conditional expectation. *Operations Research* 59(4):998–1007.

[11] van der Vaart AW (2000) *Asymptotic Statistics* (Cambridge, UK: Cambridge University Press).

[12] van der Vaart A, Wellner J (1996) *Weak Convergence and Empirical Processes: With Applications to Statistics* (Springer, New York).