J. Math. Sci. Univ. Tokyo
(), 1–1.

A novel statistical approach for two-sample testing based on the overlap coefficient

By Atsushi Komaba, Hisashi Johno and Kazunori Nakamoto

Abstract. Here we propose a new nonparametric framework for two-sample testing, named as the OVL-q ($q = 1, 2, \ldots$). This can be regarded as a natural extension of the Smirnov test, which is equivalent to the OVL-1. We specifically focus on the OVL-2, implement its fast algorithm, and show its superiority over other statistical tests in some experiments.

1. Introduction

The overlap coefficient (OVL) is a measure of the similarity between two probability distributions, defined as the common area under their density functions. Previously, we have developed a nonparametric method to estimate the OVL [6].

In any two-sample test for equality of (continuously differentiable) distribution functions, the null hypothesis is equivalent to the OVL being one. To date, however, the OVL has not been the main subject of such hypothesis testing.

The objective of this study is to construct a new nonparametric two-sample test for distribution equality based on the OVL estimation, which will be referred to as the OVL-q ($q = 1, 2, \ldots$). Furthermore, we aim to implement algorithms for the OVL-q and experimentally compare the statistical power of the OVL-1 (which is equivalent to the Smirnov test) and OVL-2, for example, with that of other statistical tests.

In this paper, we start with preliminaries and basic results in Section 2. The algorithms for the OVL-q are described in Section 3. Experimental results are shown in Section 4, and the conclusion follows in Section 5. The proofs of Theorems 2.6 and 3.9 are given in Sections 6 and 7, respectively.

2020 *Mathematics Subject Classification.* Primary 62G10; Secondary 62-04.
Key words: nonparametric statistics, overlap coefficient, two-sample testing.

A system to perform the OVL-1 and OVL-2 is available at `https://fiveseven-lambda.github.io/ovl-test/` along with its source code.

Notation. Throughout this paper, we denote by $\mathbb{Z}$, $\mathbb{N}$, $\mathbb{N}_+$, $\mathbb{Q}$, and $\mathbb{R}$ the sets of integers, nonnegative integers, positive integers, rational numbers, and real numbers, respectively. If $-\infty \le a \le b \le \infty$ and if there is no confusion, we write $[a, b] = \{x : a \le x \le b\}$, $[a, b) = \{x : a \le x < b\}$, $(a, b] = \{x : a < x \le b\}$, and $(a, b) = \{x : a < x < b\}$ as (extended) real intervals. For $q \in \mathbb{N}_+$, we define $\mathbb{R}^q_{\le} = \{(v_1, \ldots, v_q) \in \mathbb{R}^q : v_1 \le \cdots \le v_q\}$. For a set A, $\#A$ denotes the cardinality of A.

2. Analytical framework

2.1. Estimation of the OVL

Definition 2.1. On a probability space $(\Omega, \mathfrak{A}, P)$, let $X_1, \ldots, X_m$ be real random variables with a continuous distribution function F_0, $Y_1, \ldots, Y_n$ be those with F_1, and $X_1, \ldots, X_m, Y_1, \ldots, Y_n$ be mutually independent. The empirical distribution functions corresponding to $\{X_1, \ldots, X_m\}$ and $\{Y_1, \ldots, Y_n\}$ are given by

$$
\begin{aligned}
F_{0,m}(x) &= \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{(-\infty,x]}(X_i) \qquad (x \in \mathbb{R}),\\
F_{1,n}(x) &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{(-\infty,x]}(Y_i) \qquad (x \in \mathbb{R}),
\end{aligned} \tag{1}
$$

respectively, where $\mathbb{1}$ denotes the indicator function. Put $F_0(\infty) = F_1(\infty) = F_{0,m}(\infty) = F_{1,n}(\infty) = 1$ and $F_0(-\infty) = F_1(-\infty) = F_{0,m}(-\infty) = F_{1,n}(-\infty) = 0$.

Definition 2.2. For a real function g on a set A and $x, y \in A$, we write $g|_x^y = g(y) - g(x)$. For $\boldsymbol{v} = (v_1, \ldots, v_q) \in \mathbb{R}^q_{\le}$, define

$$
r(\boldsymbol{v}) = \sum_{i=0}^{q} \min\left\{F_0|_{v_i}^{v_{i+1}}, F_1|_{v_i}^{v_{i+1}}\right\}, \tag{2}
$$

$$
r_{m,n}(\boldsymbol{v}) = \sum_{i=0}^{q} \min\left\{F_{0,m}|_{v_i}^{v_{i+1}}, F_{1,n}|_{v_i}^{v_{i+1}}\right\}, \tag{3}
$$

where $v_0 = -\infty$ and $v_{q+1} = \infty$. Note that $0 \le r(\boldsymbol{v}) \le 1$ and $0 \le r_{m,n}(\boldsymbol{v}) \le 1$ for all $\boldsymbol{v} \in \mathbb{R}^q_\le$. We also define

$$\rho_{q,m,n} = \min_{\boldsymbol{v} \in \mathbb{R}^q_\le} r_{m,n}(\boldsymbol{v}) \in [0,1], \tag{4}$$

which exists because $r_{m,n}$ takes at most finitely many values.

Remark 2.3. Note that $\rho_{q,m,n}$ is measurable on Ω, because $r_{m,n}(\boldsymbol{v})$ is obviously measurable for each $\boldsymbol{v} \in \mathbb{R}^q_\le$ and $\mathbb{R}^q_\le$ in (4) can be replaced by its countable subset $\mathbb{R}^q_\le \cap \mathbb{Q}^q$ (since $F_{0,m}$ and $F_{1,n}$ are right continuous).

Definition 2.4. Suppose ξ is a random variable on $(\Omega, \mathfrak{A}, P)$ taking values in a separable metric space (E, d); $\{\xi_i : i \in \mathbb{N}_+\}$ and $\{\xi'_{i,j} : i, j \in \mathbb{N}_+\}$ are two sequences of random variables on $(\Omega, \mathfrak{A}, P)$ into E. Then we say that $\{\xi_i\}$ and $\{\xi'_{i,j}\}$ *converge almost surely* to ξ if

$$P\left(\left\{\omega \in \Omega : \lim_{i\to\infty} \xi_i(\omega) = \xi(\omega)\right\}\right) = 1,$$
$$P\left(\left\{\omega \in \Omega : \lim_{i,j\to\infty} \xi'_{i,j}(\omega) = \xi(\omega)\right\}\right) = 1,$$

respectively.

Remark 2.5. If F_0 and F_1 are differentiable on $\mathbb{R}$ with continuous derivatives f_0 and f_1, respectively, then the OVL between the two distributions is given by

$$\rho = \int_{-\infty}^{\infty} \min\{f_0(x), f_1(x)\}\, \mathrm{d}x. \tag{5}$$

We call $x \in \mathbb{R}$ a *coincidence point* between f_0 and f_1 if $f_0(x) = f_1(x)$; $x \in \mathbb{R}$ a *crossover point* between f_0 and f_1 if there exists a neighborhood V of x such that for any $a, b \in V$, $(a - x)(b - x) > 0$ if and only if $[f_0(a) - f_1(a)][f_0(b) - f_1(b)] > 0$. The set of crossover points and that of coincidence points are denoted by $C(f_0, f_1)$ and $C'(f_0, f_1)$, respectively. Note that $C(f_0, f_1) \subset C'(f_0, f_1)$.

Theorem 2.6. *Suppose f_0 and f_1 are as in Remark 2.5, $\#C'(f_0, f_1) < \infty$, and $\#C(f_0, f_1) = N < \infty$. Then $\rho_{N,m,n}$ converges almost surely to ρ as $m, n \to \infty$.*

See Section 6 for the proof of Theorem 2.6.

Hereafter, F_0 and F_1 are only assumed to be continuous, unless otherwise noted.

2.2. The OVL-q test

For $q \in \mathbb{N}_+$, we define the OVL-q test statistic as $\rho_{q,m,n}$. Under the null hypothesis $H_0 : F_0 = F_1$, the p-value of $\rho_{q,m,n}$ is given by $p_{q,m,n}(\rho_{q,m,n})$ where

$$p_{q,m,n}(x) = P(\{\omega \in \Omega : \rho_{q,m,n}(\omega) \leq x\}) \qquad (x \in \mathbb{R}), \tag{6}$$

and the lower limit of a $100(1-\alpha)\%$ confidence interval $(0 < \alpha < 1)$ of $\rho_{q,m,n}$ is

$$l_{q,m,n}(\alpha) = \sup\{x \in \mathbb{R} : p_{q,m,n}(x) < \alpha\}. \tag{7}$$

2.3. The Smirnov test

(See [2] for reference.) The Smirnov (or the two-sample Kolmogorov-Smirnov) test statistic is defined as

$$D_{m,n} = \max_{x \in \mathbb{R}} |F_{0,m}(x) - F_{1,n}(x)|.$$

Proposition 2.7. (See [4, Section 3.2] for reference.) The relation $\rho_{1,m,n} = 1 - D_{m,n}$ holds.

Proof. We have

$$\begin{aligned}
\rho_{1,m,n} &= \min_{v \in \mathbb{R}} r_{m,n}(v) \\
&= \min_{v \in \mathbb{R}} \left(\min\left\{F_{0,m}|_{-\infty}^{v}, F_{1,n}|_{-\infty}^{v}\right\} + \min\left\{F_{0,m}|_{v}^{\infty}, F_{1,n}|_{v}^{\infty}\right\}\right) \\
&= \min_{v \in \mathbb{R}} (\min\{F_{0,m}(v), F_{1,n}(v)\} + \min\{1 - F_{0,m}(v), 1 - F_{1,n}(v)\}) \\
&= \min_{v \in \mathbb{R}} (\min\{F_{0,m}(v), F_{1,n}(v)\} + 1 - \max\{F_{0,m}(v), F_{1,n}(v)\}) \\
&= \min_{v \in \mathbb{R}} (1 - |F_{0,m}(v) - F_{1,n}(v)|) \\
&= 1 - \max_{v \in \mathbb{R}} |F_{0,m}(v) - F_{1,n}(v)| \\
&= 1 - D_{m,n}
\end{aligned}$$

by definition. □

Let

$$\widetilde{p}_{m,n}(x) = P(\{\omega \in \Omega : D_{m,n}(\omega) \geq x\}) \qquad (x \in \mathbb{R}).$$

The p-value of $D_{m,n}$ under $H_0 : F_0 = F_1$ is given by $\widetilde{p}_{m,n}(D_{m,n})$. Since $D_{m,n} = 1 - \rho_{1,m,n}$ by Proposition 2.7, we have

$$\widetilde{p}_{m,n}(x) = P(\{\omega \in \Omega : \rho_{1,m,n}(\omega) \leq 1 - x\}) = p_{1,m,n}(1 - x) \qquad (x \in \mathbb{R}).$$

Hence $\widetilde{p}_{m,n}(D_{m,n})$ is equivalent to the p-value of $\rho_{1,m,n}$ under H_0 because

$$\widetilde{p}_{m,n}(D_{m,n}) = p_{1,m,n}(1 - D_{m,n}) = p_{1,m,n}(\rho_{1,m,n}).$$

Therefore, the OVL-1 is equivalent to the Smirnov test.

3. Algorithms for the OVL-q

3.1. Basic principles

Definition 3.1. For $k \in \mathbb{N}_+$, let $\Gamma_k = \{0,1\}^k$ and define $N_1(\boldsymbol{\gamma}) = \sum_{i=1}^{k} \gamma_i$ and $N_0(\boldsymbol{\gamma}) = k - N_1(\boldsymbol{\gamma})$ for $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_k) \in \Gamma_k$. Let $\Gamma_0 = \{e\}$ where e is the empty sequence, and define $N_0(e) = N_1(e) = 0$. Define $\boldsymbol{\gamma}_{i:j} = (\gamma_{i+1}, \ldots, \gamma_j)$ for $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_k) \in \Gamma_k$ ($k \geq 1$) and $i, j \in \{0, \ldots, k\}$ ($i < j$), and $\boldsymbol{\gamma}_{i:i} = e$ for $\boldsymbol{\gamma} \in \Gamma_k$ ($k \geq 0$) and $i \in \{0, \ldots, k\}$. Let $\Gamma_{k,l} = \{\boldsymbol{\gamma} \in \Gamma_{k+l} : N_0(\boldsymbol{\gamma}) = k, N_1(\boldsymbol{\gamma}) = l\}$ for $k, l \in \mathbb{N}$. For $\boldsymbol{\gamma} \in \Gamma_{m,n}$ and $q \in \mathbb{N}_+$, define

$$\widehat{\rho}_q(\boldsymbol{\gamma}) = \min_{0 \leq j_1 \leq \cdots \leq j_q \leq m+n} \widehat{r}_{\boldsymbol{\gamma}}(j_1, \ldots, j_q), \tag{8}$$

where

$$\widehat{r}_{\boldsymbol{\gamma}}(j_1, \ldots, j_q) = \sum_{i=0}^{q} \min\left\{ \widehat{F}_{0,\boldsymbol{\gamma}}\Big|_{j_i}^{j_{i+1}}, \widehat{F}_{1,\boldsymbol{\gamma}}\Big|_{j_i}^{j_{i+1}} \right\}, \tag{9}$$

$$\widehat{F}_{0,\boldsymbol{\gamma}}(i) = \frac{N_0(\boldsymbol{\gamma}_{0:i})}{m}, \quad \widehat{F}_{1,\boldsymbol{\gamma}}(i) = \frac{N_1(\boldsymbol{\gamma}_{0:i})}{n}, \tag{10}$$

$j_0 = 0$, and $j_{q+1} = m + n$. Note that $0 \leq \widehat{r}_{\boldsymbol{\gamma}}(j_1, \ldots, j_q) \leq 1$, and hence

$$0 \leq \widehat{\rho}_q(\boldsymbol{\gamma}) \leq 1. \tag{11}$$

Let $\widehat{\Omega}$ be the set of all $\omega \in \Omega$ such that $X_1(\omega), \ldots, X_m(\omega), Y_1(\omega), \ldots, Y_n(\omega)$ are all distinct. Since F_0 and F_1 are continuous, we can see that

$$P(\widehat{\Omega}) = 1. \tag{12}$$

Hence we can put $\{Z_1, \ldots, Z_{m+n}\} = \{X_1, \ldots, X_m, Y_1, \ldots, Y_n\}$ with $Z_1 < \cdots < Z_{m+n}$ almost surely. We also put $Z_0 = Z_1 - 1$. Now define $\widehat{\boldsymbol{\gamma}} = (\widehat{\gamma}_1, \ldots, \widehat{\gamma}_{m+n}) \in \Gamma_{m,n}$ on $\widehat{\Omega}$ by

$$\widehat{\gamma}_j = \begin{cases} 0 & \text{if} \quad Z_j \in \{X_1, \ldots, X_m\}, \\ 1 & \text{if} \quad Z_j \in \{Y_1, \ldots, Y_n\}. \end{cases}$$

Remark 3.2. By (1) and (10), we have $\widehat{F}_{0,\widehat{\boldsymbol{\gamma}}}(i) = F_{0,m}(Z_i)$ and $\widehat{F}_{1,\widehat{\boldsymbol{\gamma}}}(i) = F_{1,n}(Z_i)$ for all $i \in \{0, \ldots, m+n\}$.

Remark 3.3. Under the null hypothesis $H_0 : F_0 = F_1$, we have $\widehat{\boldsymbol{\gamma}}(\widehat{\Omega}) = \Gamma_{m,n}$ and

$$P(\{\omega \in \widehat{\Omega} : \widehat{\boldsymbol{\gamma}}(\omega) = \boldsymbol{\gamma}\}) = (\#\Gamma_{m,n})^{-1} = \binom{m+n}{m}^{-1}$$

for all $\boldsymbol{\gamma} \in \Gamma_{m,n}$.

Proposition 3.4. For $q \in \mathbb{N}_+$, $\widehat{\rho}_q(\widehat{\boldsymbol{\gamma}}) = \rho_{q,m,n} \in [0,1]$.

Proof. By (4) and (8), we have

$$\begin{aligned} \widehat{\rho}_q(\widehat{\boldsymbol{\gamma}}) &= \min_{0 \le j_1 \le \cdots \le j_q \le m+n} \widehat{r}_{\widehat{\boldsymbol{\gamma}}}(j_1, \ldots, j_q) \\ &= \min_{0 \le j_1 \le \cdots \le j_q \le m+n} r_{m,n}(Z_{j_1}, \ldots, Z_{j_q}) \\ &= \min_{(v_1, \ldots, v_q) \in \mathbb{R}^q_{\le}} r_{m,n}(v_1, \ldots, v_q) \\ &= \rho_{q,m,n} \in [0,1], \end{aligned}$$

noting that

$$\begin{aligned} \widehat{r}_{\widehat{\boldsymbol{\gamma}}}(j_1, \ldots, j_q) &= \sum_{i=0}^{q} \min\Big\{ \widehat{F}_{0,\widehat{\boldsymbol{\gamma}}}\big|_{j_i}^{j_{i+1}}, \widehat{F}_{1,\widehat{\boldsymbol{\gamma}}}\big|_{j_i}^{j_{i+1}} \Big\} \\ &= \sum_{i=0}^{q} \min\Big\{ F_{0,m}\big|_{Z_{j_i}}^{Z_{j_{i+1}}}, F_{1,n}\big|_{Z_{j_i}}^{Z_{j_{i+1}}} \Big\} \\ &= r_{m,n}(Z_{j_1}, \ldots, Z_{j_q}) \end{aligned}$$

by (3) and (9) and Remark 3.2, where $j_0 = 0$ and $j_{q+1} = m+n$. □

Theorem 3.5. *Under the null hypothesis $H_0 : F_0 = F_1$, we have*

$$p_{q,m,n}(x) = \frac{\#\{\gamma \in \Gamma_{m,n} : \widehat{\rho}_q(\gamma) \le x\}}{\#\Gamma_{m,n}} \qquad (x \in \mathbb{R})$$

for $q \in \mathbb{N}_+$.

Proof. This is obvious from (6) and (12), Remark 3.3, and Proposition 3.4. □

With this theorem, we can naively perform the OVL-q (see Section 2.2). Let us call this algorithm the *naive OVL-q*. If $q = 2$ and $m = n$, a faster algorithm can be applied, as described in the next subsection. An optimized algorithm for the OVL-1 (equivalent to the Smirnov test; see Section 2.3) has been previously proposed by [8].

3.2. A faster algorithm to calculate $p_{2,n,n}$

Throughout this subsection, we assume that $m = n$ and $H_0 : F_0 = F_1$ hold.

Proposition 3.6. For any $\gamma \in \Gamma_{n,n}$ and $q \in \mathbb{N}_+$, there exists $k \in \{0, \dots, n\}$ such that $\widehat{\rho}_q(\gamma) = k/n$.

Proof. It follows from (8), (9) and (11) that

$$\widehat{\rho}_q(\gamma) = \min_{0 \le j_1 \le \cdots \le j_q \le m+n} \sum_{i=0}^{q} \min\left\{ \widehat{F}_{0,\gamma}|_{j_i}^{j_{i+1}}, \widehat{F}_{1,\gamma}|_{j_i}^{j_{i+1}} \right\} \in [0, 1]$$

where $j_0 = 0$ and $j_{q+1} = m + n$. Noting that

$$\min\left\{ \widehat{F}_{0,\gamma}|_{j_i}^{j_{i+1}}, \widehat{F}_{1,\gamma}|_{j_i}^{j_{i+1}} \right\} \in \left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots \right\}$$

by (10), we obtain the claim. □

Remark 3.7. We can see from Proposition 3.6 that the distribution function $p_{q,n,n}$ in Theorem 3.5 is uniquely determined by the values $p_{q,n,n}(k/n)$ for $k = 0, \dots, n$.

Definition 3.8. Define a sequence $\{Q_i(x)\}$ of polynomials in x inductively by

$$\begin{aligned} Q_1(x) &= Q_0(x) = 1, \\ Q_{i+2}(x) &= Q_{i+1}(x) - xQ_i(x) \qquad (i \in \mathbb{N}). \end{aligned}$$

We denote by $Q_i'(x)$ the derivative of $Q_i(x)$. Note that $Q_0(x), Q_1(x), \ldots$ can be regarded as formal power series. For a formal power series $Q(x)$, we denote by $[x^k]Q(x)$ the coefficient of x^k in $Q(x)$, and by $1/Q(x)$ the multiplicative inverse of $Q(x)$ (if it exists).

We can find $\{Q_i(x)\}$ in [10] as a variation of the Fibonacci polynomials. For each $i \in \mathbb{N}$, we can easily see that $[x^0]Q_i(x) = 1$, and hence $1/Q_i(x)$ exists.

Theorem 3.9. *For $k = 0, \ldots, n$, we have*

$$\#\left\{\gamma \in \Gamma_{n,n} : \widehat{\rho}_2(\gamma) \geq 1 - \frac{k}{n}\right\} = [x^n]\left(\frac{Q_{k+1}'(x)}{Q_k(x)} - \frac{Q_{k+2}'(x)}{Q_{k+1}(x)}\right). \tag{13}$$

See Section 7 for the proof of Theorem 3.9.

Remark 3.10. For $k = 0, \ldots, n$, Theorem 3.5 and Proposition 3.6 imply

$$\begin{aligned} p_{2,n,n}\left(\frac{k}{n}\right) &= 1 - \frac{\#\left\{\gamma \in \Gamma_{n,n} : \widehat{\rho}_2(\gamma) > \frac{k}{n}\right\}}{\#\Gamma_{n,n}} \\ &= 1 - \frac{\#\left\{\gamma \in \Gamma_{n,n} : \widehat{\rho}_2(\gamma) \geq \frac{k+1}{n}\right\}}{\#\Gamma_{n,n}}, \end{aligned}$$

where

$$\#\left\{\gamma \in \Gamma_{n,n} : \widehat{\rho}_2(\gamma) \geq \frac{k+1}{n}\right\} = [x^n]\left(\frac{Q_{n-k}'(x)}{Q_{n-k-1}(x)} - \frac{Q_{n-k+1}'(x)}{Q_{n-k}(x)}\right)$$

if $k \leq n-1$, by Theorem 3.9. It is obvious that $p_{2,n,n}(n/n) = 1$.

Remarks 3.7 and 3.10 imply that we can calculate $p_{2,n,n}$ with the use of $\{Q_i(x)\}$. Let us call this algorithm the *fast OVL*-2. In Section 4.1, we will numerically compare the computation times of the naive and fast OVL-2.

Table 1.

n	Mean computation time [ms] naive OVL-2	fast OVL-2
10	9	0.026
12	135	0.026
14	2153	0.028
16	34361	0.023
500	–	12
1000	–	49
5000	–	1865
10000	–	8027

4. Numerical experiments

4.1. Computation times of the naive and fast OVL-2

We performed the following benchmark test on a personal computer with min 2200 MHz – max 5083 MHz CPU (AMD Ryzen 9 5950X 16-Core Processor), 62.8 GiB RAM, and Linux 5.16.14 (Arch Linux). For each $n \in \{10, 12, 14, 16\}$, we compared the mean computation times of the naive and fast OVL-2 (averaged over 10 computations for the naive; 100000 computations for the fast) to calculate $p_{2,n,n}(1/2)$. We further measured the mean computation time of the fast OVL-2 (averaged over 10 computations) to calculate $p_{2,n,n}(1/2)$ with $n \in \{500, 1000, 5000, 10000\}$. The source code used here was written in Rust (2021 edition, rustc 1.58.1), and is published at `https://github.com/fiveseven-lambda/fast-OVL-benchmark/`.

Table 1 shows the result of the benchmark test. As can be seen, the fast OVL-2 was much faster than the naive OVL-2 (e.g., more than one million times faster to compute $p_{2,16,16}(1/2)$). The calculation of $p_{2,n,n}(1/2)$ with $n \in \{500, 1000, 5000, 10000\}$ was computationally difficult for the naive OVL-2 but easy for the fast OVL-2 (e.g., the fast OVL-2 could compute $p_{2,10000,10000}(1/2)$ in around eight seconds).

4.2. The statistical power of the OVL-2 test

In this experiment, we focused on the case $m = n$ and simulated $X_1, \ldots,$ $X_n, Y_1, \ldots, Y_n$ in Definition 2.1, with f_0 and f_1 in Remark 2.5 being specific functions (described in the next paragraph). The random samples were subjected to the OVL-1, OVL-2, and other statistical tests (i.e., the Welch t [12], two-tailed F [11, Section 6.12], Mann-Whitney U [7], and two-sample Cramér-von Mises test [1]) to verify the null hypothesis $H_0 : F_0 = F_1$ with 95% confidence interval. This trial (from the generation of $2n$ random samples) was repeated 20000 times independently for each $n \in \{2^2, 2^3, \ldots, 2^{12}\}$, and the statistical power (or equivalently the rejection ratio) of each test was calculated. The source code used here was written in Rust (2021 edition, rustc 1.58.1) and Python (version 3.9), and is published at `https://github.com/fiveseven-lambda/OVL-q-test-comparison`.

As probability density functions, we used

$$\mathrm{Normal}_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad (x \in \mathbb{R})$$

with $\mu \in \mathbb{R}$ and $\sigma > 0$ for normal distributions;

$$\mathrm{Trapezoidal}(x) = \begin{cases} (x+2)/2 & \text{if} \quad -2 \le x \le -\sqrt{2}, \\ (2-\sqrt{2})/2 & \text{if} \quad -\sqrt{2} < x \le \sqrt{2}, \\ (-x+2)/2 & \text{if} \quad \sqrt{2} < x \le 2, \\ 0 & \text{if} \quad x < -2 \text{ or } 2 < x \end{cases}$$

for a trapezoidal distribution;

$$\mathrm{Triangular}(x) = \begin{cases} (x+\sqrt{6})/6 & \text{if} \quad -\sqrt{6} \le x \le 0, \\ (-x+\sqrt{6})/6 & \text{if} \quad 0 < x \le \sqrt{6}, \\ 0 & \text{if} \quad x < -\sqrt{6} \text{ or } \sqrt{6} < x \end{cases}$$

for a triangular distribution;

$$\mathrm{Mixed}(x) = \frac{1}{2}\left(\mathrm{Normal}_{-0.8,0.6} + \mathrm{Normal}_{0.8,0.6}\right)(x) \qquad (x \in \mathbb{R})$$

for a mixed normal distribution. As a control function, we fixed $f_0 = \mathrm{Normal}_{0,1}$.

Figures 1 to 5 show the experimental results:

- In the case $f_1 = \mathrm{Normal}_{0,1.1}$ where f_0 and f_1 were the densities of two normal distributions with identical means and different variances, the

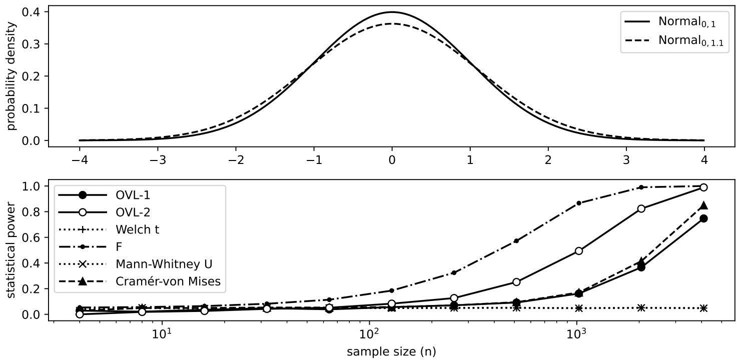

Fig. 1. The random variables $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ were realized with $f_0 = \text{Normal}_{0,1}$ and $f_1 = \text{Normal}_{0,1.1}$, and subjected to the statistical tests (the OVL-1, OVL-2, Welch t, F, Mann-Whitney U, and Cramér-von Mises test) to verify the null hypothesis $H_0 : F_0 = F_1$ with 95% confidence interval. This trial was repeated 20000 times independently for each $n \in \{2^2, 2^3, \ldots, 2^{12}\}$, and the statistical power of each test was evaluated. Note that $\text{Normal}_{0,1}$ has mean 0 and variance 1, while $\text{Normal}_{0,1.1}$ has mean 0 and variance 1.21.

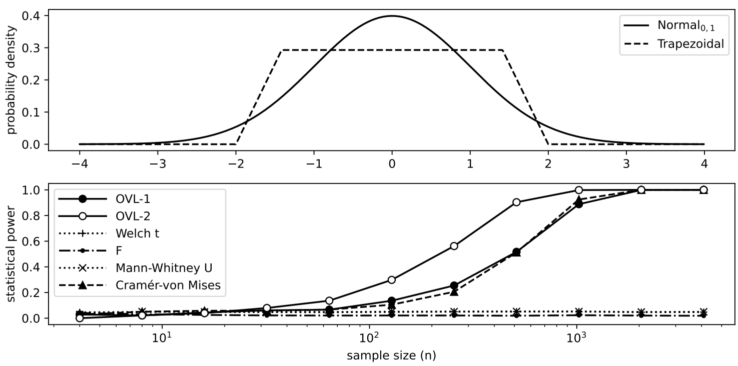

Fig. 2. The random variables $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ were realized with $f_0 = \text{Normal}_{0,1}$ and $f_1 = \text{Trapezoidal}$, and subjected to the statistical tests (the OVL-1, OVL-2, Welch t, F, Mann-Whitney U, and Cramér-von Mises test) to verify the null hypothesis $H_0 : F_0 = F_1$ with 95% confidence interval. This trial was repeated 20000 times independently for each $n \in \{2^2, 2^3, \ldots, 2^{12}\}$, and the statistical power of each test was evaluated. Note that $\text{Normal}_{0,1}$ and Trapezoidal have mean 0 and variance 1.

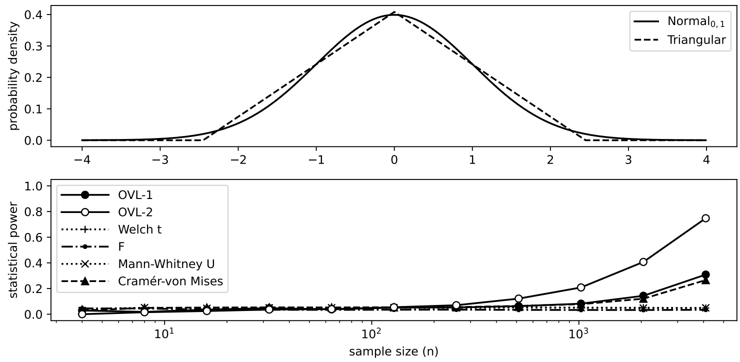

Fig. 3. The random variables $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ were realized with $f_0 = \mathrm{Normal}_{0,1}$ and $f_1 = \mathrm{Triangular}$, and subjected to the statistical tests (the OVL-1, OVL-2, Welch t, F, Mann-Whitney U, and Cramér-von Mises test) to verify the null hypothesis $H_0 : F_0 = F_1$ with 95% confidence interval. This trial was repeated 20000 times independently for each $n \in \{2^2, 2^3, \ldots, 2^{12}\}$, and the statistical power of each test was evaluated. Note that $\mathrm{Normal}_{0,1}$ and Triangular have mean 0 and variance 1.

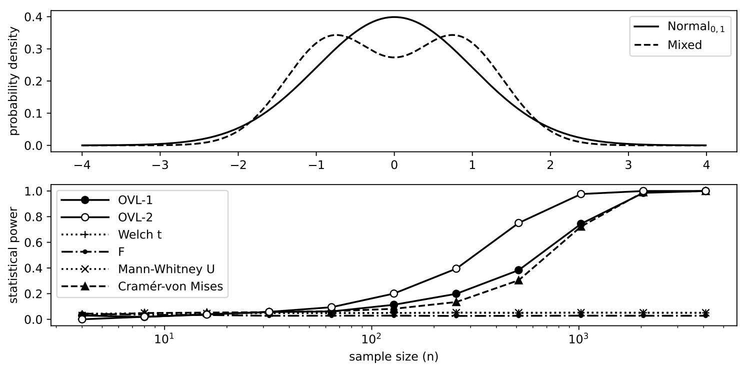

Fig. 4. The random variables $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ were realized with $f_0 = \mathrm{Normal}_{0,1}$ and $f_1 = \mathrm{Mixed}$, and subjected to the statistical tests (the OVL-1, OVL-2, Welch t, F, Mann-Whitney U, and Cramér-von Mises test) to verify the null hypothesis $H_0 : F_0 = F_1$ with 95% confidence interval. This trial was repeated 20000 times independently for each $n \in \{2^2, 2^3, \ldots, 2^{12}\}$, and the statistical power of each test was evaluated. Note that $\mathrm{Normal}_{0,1}$ and Mixed have mean 0 and variance 1.

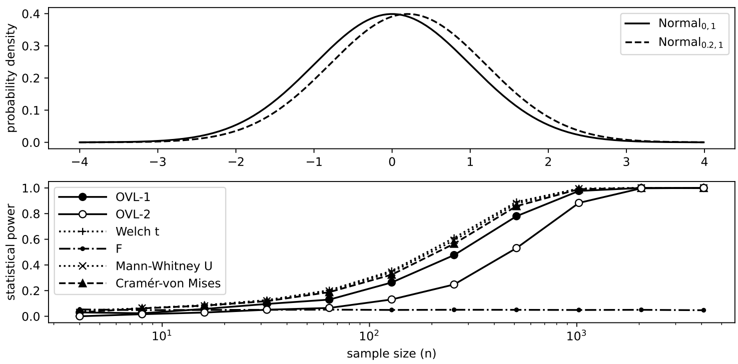

Fig. 5. The random variables $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ were realized with $f_0 = \text{Normal}_{0,1}$ and $f_1 = \text{Normal}_{0.2,1}$, and subjected to the statistical tests (the OVL-1, OVL-2, Welch t, F, Mann-Whitney U, and Cramér-von Mises test) to verify the null hypothesis $H_0 : F_0 = F_1$ with 95% confidence interval. This trial was repeated 20000 times independently for each $n \in \{2^2, 2^3, \ldots, 2^{12}\}$, and the statistical power of each test was evaluated. Note that $\text{Normal}_{0,1}$ has mean 0 and variance 1, while $\text{Normal}_{0.2,1}$ has mean 0.2 and variance 1.

power of the F test was the highest, followed by the OVL-2, Cramér-von Mises, OVL-1, and then Welch t or Mann-Whitney U test (Figure 1).

- In the case $f_1 \in \{\text{Trapezoidal}, \text{Triangular}, \text{Mixed}\}$ where f_0 and f_1 were the densities of two different distributions with identical means and variances, the power of the OVL-2 test was the highest, followed by the OVL-1 or Cramér-von Mises, Welch t or Mann-Whitney U, and then F test (Figures 2 to 4).

- In the case $f_1 = \text{Normal}_{0.2,1}$ where f_0 and f_1 were the densities of two normal distributions with different means and identical variances, the power of the Welch t test was the highest, followed by the Mann-Whitney U, Cramér-von Mises, OVL-1, OVL-2, and then F test (Figure 5).

5. Conclusion

Based on the OVL estimation, we have devised a novel statistical framework for two-sample testing: the OVL-q ($q \in \mathbb{N}_+$), which can be regarded as a natural extension of the Smirnov test (since the OVL-1 is equivalent to the Smirnov test). We have explained and implemented the algorithms for the OVL-q (in particular, the fast OVL-2 algorithm). Furthermore, we have demonstrated the superiority of the OVL-2 over conventional statistical tests in some experiments, although the reason for this is not clear at present.

One limitation is that we are currently unable to rapidly perform the OVL-2 if $m \neq n$ or the OVL-q if $q \geq 3$. To overcome this, we should explore the possibility of expanding fast and exact algorithms for the OVL-q, or should investigate the asymptotic distribution of $\rho_{q,m,n}$ (as $m, n \to \infty$) to approximate the OVL-q in future works. The treatment of ties (which may occur in $\Omega \setminus \widehat{\Omega}$ if F_0 or F_1 is practically discontinuous) is also an important topic of research. In addition, it is meaningful to further evaluate the statistical power of the OVL-q both in simulations and in real observations.

6. Proof for Theorem 2.6

Definition 6.1. In the setting of Definition 2.4, we say that $\{\xi_i\}$ *converges completely* to ξ if

$$\sum_{i=1}^{\infty} P\left(\{\omega \in \Omega : d\left(\xi_i(\omega), \xi(\omega)\right) > \epsilon\}\right) < \infty$$

for any $\epsilon > 0$.

Remark 6.2. (See [5] for reference.) It is well known that $\{\xi_i\}$ converges almost surely to ξ if and only if

$$\lim_{l \to \infty} P\left(\bigcup_{i=l}^{\infty} \{\omega \in \Omega : d\left(\xi_i(\omega), \xi(\omega)\right) > \epsilon\}\right) = 0$$

for any $\epsilon > 0$. Note that if $\{\xi_i\}$ converges completely to ξ, then $\{\xi_i\}$ converges almost surely to ξ.

Theorem 6.3. (The Glivenko-Cantelli theorem. See [9, Theorem A, Section 2.1.4] for the proof.) *As $m \to \infty$ and $n \to \infty$,*

$$\sup_{x\in\mathbb{R}} |F_{0,m}(x) - F_0(x)|, \qquad \sup_{x\in\mathbb{R}} |F_{1,n}(x) - F_1(x)|$$

converge completely to 0*, respectively.*

Lemma 6.4. (See [6, Lemma A.12] for the proof.) If $x, y, z, w \in \mathbb{R}$, then

(a) $|\max\{x, y\} - \max\{z, w\}| \leq |x - z| + |y - w|$,

(b) $|\min\{x, y\} - \min\{z, w\}| \leq |x - z| + |y - w|$.

In accordance with Theorem 2.6, let F_0 and F_1 be differentiable on $\mathbb{R}$ with continuous derivatives f_0 and f_1, respectively, $\#C'(f_0, f_1) < \infty$, $C(f_0, f_1) = \{c_1, \ldots, c_N\}$ with $c_1 < \cdots < c_N$, $\boldsymbol{c} = (c_1, \ldots, c_N)$, $c_0 = -\infty$, and $c_{N+1} = \infty$.

Remark 6.5. It follows from (5) and Definition 2.2 that $r(\boldsymbol{c}) = \rho$.

Definition 6.6. For $q \in \mathbb{N}_+$, define

$$\begin{aligned} \mathcal{V}_q &= \operatorname*{arg\,min}_{\boldsymbol{v}\in\mathbb{R}^q_{\leq}} r(\boldsymbol{v}), \\ \mathcal{V}_{q,m,n} &= \operatorname*{arg\,min}_{\boldsymbol{v}\in\mathbb{R}^q_{\leq}} r_{m,n}(\boldsymbol{v}), \\ \mathcal{C}_q &= \{(c_{i_1}, \ldots, c_{i_q}) : 1 \leq i_1 < \cdots < i_q \leq N\}. \end{aligned}$$

Remark 6.7. It follows from (4), Theorems 6.9 and 6.13, and Corollary 6.12 that $\mathcal{V}_q \neq \emptyset$ and $\mathcal{V}_{q,m,n} \neq \emptyset$ for all $q \in \mathbb{N}_+$. It is obvious that $\mathcal{C}_N = \{\boldsymbol{c}\}$ and $\mathcal{C}_q = \emptyset$ if $q > N$.

Lemma 6.8. Suppose $q \in \mathbb{N}_+$, $\boldsymbol{v} = (v_1, \ldots, v_q) \in \mathbb{R}^q_{\leq}$, $v_0 = -\infty$, and $v_{q+1} = \infty$. If $v_i < c_s < v_{i+1}$ for some $i \in \{0, \ldots, q\}$ and $s \in \{1, \ldots, N\}$, then $r(\boldsymbol{v}) > \rho$.

Proof. Since $\#C'(f_0, f_1) < \infty$, there is an open interval $U \subset (v_i, v_{i+1})$ with $U \cap C'(f_0, f_1) = \{c_s\}$, so that $[f_0(a) - f_1(a)][f_0(b) - f_1(b)] < 0$ for all $a, b \in U$

with $a < c_s < b$. Now fix such a and b. Without loss of generality, we assume that $f_0(a) < f_1(a)$ and $f_0(b) > f_1(b)$. If $F_0|_{v_i}^{v_{i+1}} \leq F_1|_{v_i}^{v_{i+1}}$, then

$$\begin{aligned}
&r(\boldsymbol{v}) - \rho \\
&= \sum_{j=0}^{q} \left(\min\left\{ \int_{v_j}^{v_{j+1}} f_0(x)\,\mathrm{d}x, \int_{v_j}^{v_{j+1}} f_1(x)\,\mathrm{d}x \right\} \right. \\
&\qquad\qquad\qquad\qquad \left. - \int_{v_j}^{v_{j+1}} \min\{f_0(x), f_1(x)\}\,\mathrm{d}x \right) \\
&\geq \min\left\{ \int_{v_i}^{v_{i+1}} f_0(x)\,\mathrm{d}x, \int_{v_i}^{v_{i+1}} f_1(x)\,\mathrm{d}x \right\} - \int_{v_i}^{v_{i+1}} \min\{f_0(x), f_1(x)\}\,\mathrm{d}x \\
&= \int_{v_i}^{v_{i+1}} \left(f_0(x) - \min\{f_0(x), f_1(x)\}\right)\,\mathrm{d}x \\
&\geq \int_{c_s}^{b} \left(f_0(x) - f_1(x)\right)\,\mathrm{d}x \\
&> 0.
\end{aligned}$$

We can similarly prove that $r(\boldsymbol{v}) - \rho > 0$ if $F_0|_{v_i}^{v_{i+1}} > F_1|_{v_i}^{v_{i+1}}$. $\square$

Theorem 6.9. *The minimum of r on $\mathbb{R}^N_{\leq}$ is ρ, which is uniquely attained at $\boldsymbol{c}$.*

Proof. If $\boldsymbol{v} = (v_1, \ldots, v_N) \in \mathbb{R}^N_{\leq}$ with $\boldsymbol{v} \neq \boldsymbol{c}$, then $c_s \notin \{v_1, \ldots, v_N\}$ for some s, hence $v_i < c_s < v_{i+1}$ for some i as in the assumption of Lemma 6.8, so that $r(\boldsymbol{v}) > \rho = r(\boldsymbol{c})$ (note Remark 6.5). $\square$

Theorem 6.10. *If $q \in \{1, \ldots, N-1\}$ and $\boldsymbol{v} \in \mathbb{R}^q_{\leq}$, then $r(\boldsymbol{v}) > r(\boldsymbol{c})$.*

Proof. Since $q < N$, $c_s \notin \{v_1, \ldots, v_q\}$ for some s. The proof is similar as that of Theorem 6.9. $\square$

Theorem 6.11. *If $q \in \{1, \ldots, N-1\}$, then for any $\boldsymbol{v} = (v_1, \ldots, v_q) \in \mathbb{R}^q_{\leq}$, there exists $\boldsymbol{w} = (c_{j_1}, \ldots, c_{j_q})$ with $1 \leq j_1 \leq \cdots \leq j_q \leq N$ such that $r(\boldsymbol{w}) \leq r(\boldsymbol{v})$.*

Proof. Let $\boldsymbol{v} = (v_1, \ldots, v_q) \in \mathbb{R}^q_{\leq}$, $v_0 = -\infty$, $v_{q+1} = \infty$, and

$$\eta(\boldsymbol{v}) = \#\{i \in \{1, \ldots, q\} : v_i \notin C(f_0, f_1)\}.$$

The statement obviously holds when $\eta(\boldsymbol{v}) = 0$. Hence suppose $\eta(\boldsymbol{v}) > 0$. Then we can choose $i \in \{1, \ldots, q\}$ and $s \in \{1, \ldots, N\}$ satisfying $c_{s-1} < v_i < c_s \le v_{i+1}$ or $v_{i-1} \le c_s < v_i < c_{s+1}$. We will only prove the case $c_{s-1} < v_i < c_s \le v_{i+1}$, as the other is similar. Without loss of generality, we may assume that $f_0 \le f_1$ on (c_{s-1}, c_s), so that $F_0|_{c_{s-1}}^{v_i} < F_1|_{c_{s-1}}^{v_i}$ and $F_0|_{v_i}^{c_s} < F_1|_{v_i}^{c_s}$, since $\#C'(f_0, f_1) < \infty$. In the following, we consider the cases (I) $F_0|_{v_{i-1}}^{v_i} \le F_1|_{v_{i-1}}^{v_i}$ and (II) $F_0|_{v_{i-1}}^{v_i} > F_1|_{v_{i-1}}^{v_i}$.

(I) Suppose $F_0|_{v_{i-1}}^{v_i} \le F_1|_{v_{i-1}}^{v_i}$. Then

$$\begin{aligned}
&F_0|_{v_{i-1}}^{c_s} < F_1|_{v_{i-1}}^{c_s}, \\
&F_j|_{v_{i-1}}^{c_s} = F_j|_{v_{i-1}}^{v_i} + F_j|_{v_i}^{c_s} \qquad (j = 0, 1), \\
&F_j|_{c_s}^{v_{i+1}} = F_j|_{v_i}^{v_{i+1}} - F_j|_{v_i}^{c_s} \qquad (j = 0, 1),
\end{aligned}$$

hence

$$\begin{aligned}
\min_j F_j|_{v_{i-1}}^{c_s} + \min_j F_j|_{c_s}^{v_{i+1}} &= F_0|_{v_{i-1}}^{c_s} + \min_j (F_j|_{v_i}^{v_{i+1}} - F_j|_{v_i}^{c_s}) \\
&= F_0|_{v_{i-1}}^{v_i} + F_0|_{v_i}^{c_s} + \min_j (F_j|_{v_i}^{v_{i+1}} - F_j|_{v_i}^{c_s}) \\
&\le F_0|_{v_{i-1}}^{v_i} + F_0|_{v_i}^{c_s} + \min_j F_j|_{v_i}^{v_{i+1}} - F_0|_{v_i}^{c_s} \\
&= F_0|_{v_{i-1}}^{v_i} + \min_j F_j|_{v_i}^{v_{i+1}} \\
&= \min_j F_j|_{v_{i-1}}^{v_i} + \min_j F_j|_{v_i}^{v_{i+1}},
\end{aligned}$$

and setting $\boldsymbol{v}' = (v_1, \ldots, v_{i-1}, c_s, v_{i+1}, \ldots, v_q) \in \mathbb{R}^q_{\le}$ results in $\eta(\boldsymbol{v}') < \eta(\boldsymbol{v})$ and $r(\boldsymbol{v}') \le r(\boldsymbol{v})$.

(II) Suppose $F_0|_{v_{i-1}}^{v_i} > F_1|_{v_{i-1}}^{v_i}$. Since $f_0 \le f_1$ on (c_{s-1}, c_s), we can see that $v_{i-1} < c_{s-1} < v_i$ and $F_0|_{v_{i-1}}^{c_{s-1}} > F_1|_{v_{i-1}}^{c_{s-1}}$. (II-i) First consider the case $F_0|_{v_i}^{v_{i+1}} \le F_1|_{v_i}^{v_{i+1}}$. Then $F_0|_{c_{s-1}}^{v_{i+1}} < F_1|_{c_{s-1}}^{v_{i+1}}$, hence

$$\begin{aligned}
\min_j F_j|_{v_{i-1}}^{c_{s-1}} + \min_j F_j|_{c_{s-1}}^{v_{i+1}} &= F_1|_{v_{i-1}}^{c_{s-1}} + F_0|_{c_{s-1}}^{v_{i+1}} \\
&= F_1|_{v_{i-1}}^{c_{s-1}} + F_0|_{c_{s-1}}^{v_i} + F_0|_{v_i}^{v_{i+1}} \\
&< F_1|_{v_{i-1}}^{c_{s-1}} + F_1|_{c_{s-1}}^{v_i} + F_0|_{v_i}^{v_{i+1}} \\
&= F_1|_{v_{i-1}}^{v_i} + F_0|_{v_i}^{v_{i+1}} \\
&= \min_j F_j|_{v_{i-1}}^{v_i} + \min_j F_j|_{v_i}^{v_{i+1}},
\end{aligned}$$

and setting $\boldsymbol{v}' = (v_1, \dots, v_{i-1}, c_{s-1}, v_{i+1}, \dots, v_q) \in \mathbb{R}^q_\leq$ results in $\eta(\boldsymbol{v}') < \eta(\boldsymbol{v})$ and $r(\boldsymbol{v}') < r(\boldsymbol{v})$. (II-ii) Next consider the case $F_0|_{v_i}^{v_{i+1}} > F_1|_{v_i}^{v_{i+1}}$. (II-ii-a) If there is $x \in (c_{s-1}, v_i)$ such that $F_0|_x^{v_{i+1}} \leq F_1|_x^{v_{i+1}}$, then $F_0|_{v_{i-1}}^x > F_1|_{v_{i-1}}^x$, hence the case (II-i) applies to $\boldsymbol{v}'' = (v_1, \dots, v_{i-1}, x, v_{i+1}, \dots, v_q) \in \mathbb{R}^q_\leq$, where $\eta(\boldsymbol{v}'') = \eta(\boldsymbol{v})$ and

$$\begin{aligned}
r(\boldsymbol{v}'') - r(\boldsymbol{v}) &= \min_j F_j|_{v_{i-1}}^x + \min_j F_j|_x^{v_{i+1}} - \min_j F_j|_{v_{i-1}}^{v_i} - \min_j F_j|_{v_i}^{v_{i+1}} \\
&= F_1|_{v_{i-1}}^x + F_0|_x^{v_{i+1}} - F_1|_{v_{i-1}}^{v_i} - F_1|_{v_i}^{v_{i+1}} \\
&\leq F_1|_{v_{i-1}}^x + F_1|_x^{v_{i+1}} - F_1|_{v_{i-1}}^{v_i} - F_1|_{v_i}^{v_{i+1}} \\
&= 0.
\end{aligned}$$

(II-ii-b) If $F_0|_x^{v_{i+1}} > F_1|_x^{v_{i+1}}$ for any $x \in (c_{s-1}, v_i)$, then $F_0|_{c_{s-1}}^{v_{i+1}} \geq F_1|_{c_{s-1}}^{v_{i+1}}$, and setting $\boldsymbol{v}' = (v_1, \dots, v_{i-1}, c_{s-1}, v_{i+1}, \dots, v_q) \in \mathbb{R}^q_\leq$ results in $\eta(\boldsymbol{v}') < \eta(\boldsymbol{v})$ and

$$\begin{aligned}
r(\boldsymbol{v}') - r(\boldsymbol{v}) &= \min_j F_j|_{v_{i-1}}^{c_{s-1}} + \min_j F_j|_{c_{s-1}}^{v_{i+1}} - \min_j F_j|_{v_{i-1}}^{v_i} - \min_j F_j|_{v_i}^{v_{i+1}} \\
&= F_1|_{v_{i-1}}^{c_{s-1}} + F_1|_{c_{s-1}}^{v_{i+1}} - F_1|_{v_{i-1}}^{v_i} - F_1|_{v_i}^{v_{i+1}} \\
&= 0.
\end{aligned}$$

Taken together, for any $\boldsymbol{v} \in \mathbb{R}^q_\leq$ with $\eta(\boldsymbol{v}) > 0$, there exists $\boldsymbol{v}' \in \mathbb{R}^q_\leq$ such that $\eta(\boldsymbol{v}') < \eta(\boldsymbol{v})$ and $r(\boldsymbol{v}') \leq r(\boldsymbol{v})$. The statement follows by induction. □

Corollary 6.12. If $q \in \{1, \dots, N-1\}$, then there exists $\boldsymbol{c}' \in \mathcal{C}_q$ such that $r(\boldsymbol{c}') = \inf\{r(\boldsymbol{v}) : \boldsymbol{v} \in \mathbb{R}^q_\leq\}$. Furthermore, $r(\boldsymbol{c}') > r(\boldsymbol{c})$.

Proof. Since there are only finitely many choices for $\boldsymbol{w} \in \mathbb{R}^q_\leq$ in Theorem 6.11, we can choose $\boldsymbol{w}' = (c_{i_1}, \dots, c_{i_q}) \in \arg\min_{\boldsymbol{w}} r(\boldsymbol{w})$, where $\boldsymbol{w}$ ranges over the choices. Then $r(\boldsymbol{w}') \leq r(\boldsymbol{v})$ for all $\boldsymbol{v} \in \mathbb{R}^q_\leq$. Suppose $\boldsymbol{w}' \notin \mathcal{C}_q$ and put $A = \{c_{i_1}, \dots, c_{i_q}\}$. Then $\#A < q$, and there exists $A' = \{c_{j_1}, \dots, c_{j_q}\}$ such that $A \subset A'$ and $1 \leq j_1 < \cdots < j_q \leq N$. Putting $\boldsymbol{c}' = (c_{j_1}, \dots, c_{j_q})$, we have $\boldsymbol{c}' \in \mathcal{C}_q$ and $r(\boldsymbol{c}') \leq r(\boldsymbol{w}')$ by definition. Hence $r(\boldsymbol{c}') = r(\boldsymbol{w}') = \min\{r(\boldsymbol{v}) : \boldsymbol{v} \in \mathbb{R}^q_\leq\}$. Furthermore, $r(\boldsymbol{c}') > r(\boldsymbol{c})$ by Theorem 6.10. □

Theorem 6.13. *For $q = N+1, N+2, \dots$, the minimum of r on $\mathbb{R}^q_\leq$ is ρ.*

Proof. Since $\{c_1, \dots, c_N\} \subset \{v_1, \dots, v_q\}$ implies $r(\boldsymbol{c}) = r(v_1, \dots, v_q)$, the claim follows by Remark 6.5 and Lemma 6.8. □

Remark 6.14. For some $q \in \{1, \dots, N-1\}$, $\boldsymbol{v} \in \mathcal{V}_q$ does not necessarily imply $\boldsymbol{v} \in \mathcal{C}_q$. (Note that $\mathcal{V}_N = \mathcal{C}_N = \{\boldsymbol{c}\}$ by Theorem 6.9 and $\mathcal{C}_{N+1} = \mathcal{C}_{N+2} = \cdots = \emptyset$.) Here we give an example for the case where $N = 3$, $q = 2$, and $\mathcal{V}_2 \not\subset \mathcal{C}_2$. Assume that f_0 and f_1 are defined by

$$f_0(x) = \begin{cases} \frac{1-\cos x}{4\pi} & (0 \le x \le 4\pi), \\ 0 & (\text{otherwise}), \end{cases} \qquad f_1(x) = \begin{cases} \frac{1+\cos x}{4\pi} & (\pi \le x \le 5\pi), \\ 0 & (\text{otherwise}). \end{cases}$$

Then $(3\pi/2, 11)$ is in $\mathcal{V}_2$ but not in $\mathcal{C}_2 = \{(3\pi/2, 5\pi/2), (3\pi/2, 7\pi/2), (5\pi/2, 7\pi/2)\}$. (Note that $7\pi/2 = 10.995\ldots < 11$.)

The measurability of $\sup_{\boldsymbol{v}\in\mathbb{R}^q_{\le}} |r_{m,n}(\boldsymbol{v}) - r(\boldsymbol{v})|$ on Ω can be proved by the same argument as in Remark 2.3.

Theorem 6.15. *For $q \in \mathbb{N}_+$, $\sup_{\boldsymbol{v}\in\mathbb{R}^q_{\le}} |r_{m,n}(\boldsymbol{v}) - r(\boldsymbol{v})|$ converges almost surely to 0 as $m, n \to \infty$.*

Proof. For $\boldsymbol{v} \in \mathbb{R}^q_{\le}$, we have

$$\begin{aligned} |r_{m,n}(\boldsymbol{v}) - r(\boldsymbol{v})| &\le \sum_{i=0}^{q} \left| \min\left\{ F_{0,m}|_{v_i}^{v_{i+1}}, F_{1,n}|_{v_i}^{v_{i+1}} \right\} - \min\left\{ F_0|_{v_i}^{v_{i+1}}, F_1|_{v_i}^{v_{i+1}} \right\} \right| \\ &\le \sum_{i=0}^{q} \left(\left| F_{0,m}|_{v_i}^{v_{i+1}} - F_0|_{v_i}^{v_{i+1}} \right| + \left| F_{1,n}|_{v_i}^{v_{i+1}} - F_1|_{v_i}^{v_{i+1}} \right| \right) \end{aligned}$$

by definition and Lemma 6.4. Since

$$\begin{aligned} \left| F_{0,m}|_{v_i}^{v_{i+1}} - F_0|_{v_i}^{v_{i+1}} \right| &\le |F_{0,m}(v_{i+1}) - F_0(v_{i+1})| + |F_{0,m}(v_i) - F_0(v_i)| \\ &\le 2 \sup_{x\in\mathbb{R}} |F_{0,m}(x) - F_0(x)|, \\ \left| F_{1,n}|_{v_i}^{v_{i+1}} - F_1|_{v_i}^{v_{i+1}} \right| &\le |F_{1,n}(v_{i+1}) - F_1(v_{i+1})| + |F_{1,n}(v_i) - F_1(v_i)| \\ &\le 2 \sup_{x\in\mathbb{R}} |F_{1,n}(x) - F_1(x)|, \end{aligned}$$

we obtain

$$\begin{aligned} &\sup_{\boldsymbol{v}\in\mathbb{R}^q_{\le}} |r_{m,n}(\boldsymbol{v}) - r(\boldsymbol{v})| \\ &\le 2(q+1) \left(\sup_{x\in\mathbb{R}} |F_{0,m}(x) - F_0(x)| + \sup_{x\in\mathbb{R}} |F_{1,n}(x) - F_1(x)| \right), \end{aligned}$$

whose right side converges almost surely to 0 as $m, n \to \infty$ by Remark 6.2 and Theorem 6.3. $\square$

Definition 6.16. Let (A, d) be a metric space. We define a discrepancy of $A_1 \subset A$ from $A_2 \subset A$ by

$$D(A_1, A_2) = \sup_{a_1 \in A_1} \left(\inf_{a_2 \in A_2} d(a_1, a_2) \right).$$

If the metric space is $\mathbb{R}^q$ ($q \in \mathbb{N}_+$) with the Euclidean metric, we write D_q in place of D.

Lemma 6.17. Let (A, d) be a metric space. Let g and $g_{i,j}$ ($i, j \in \mathbb{N}_+$) be real functions on A such that $\min\{g(t) : t \in A\}$ and $\min\{g_{i,j}(t) : t \in A\}$ exist. Put $T = \arg\min_{t \in A} g(t)$ and $T_{i,j} = \arg\min_{t \in A} g_{i,j}(t)$. Suppose g is continuous on A, $\sup_{t \in A} |g_{i,j}(t) - g(t)| \to 0$ as $i, j \to \infty$, and there is a compact set $K \subset A$ such that

$$\min\{g(t) : t \in A\} < \inf\{g(t) : t \in A \setminus K\}.$$

Then $D(T_{i,j}, T) \to 0$ as $i, j \to \infty$.

Proof. If $T' = \arg\max_{t \in A}(-g(t))$ and $T'_{i,j} = \arg\max_{t \in A}(-g_{i,j}(t))$, then $T' = T$ and $T'_{i,j} = T_{i,j}$, hence $D(T_{i,j}, T) = D(T'_{i,j}, T') \to 0$ by [6, Lemma A.15] (replace g and g_i with $-g$ and $-g_{i,j}$, respectively). $\square$

Lemma 6.18. There exists a compact set $K \subset \mathbb{R}^N_{\le}$ such that

$$\min\left\{r(\boldsymbol{v}) : \boldsymbol{v} \in \mathbb{R}^N_{\le}\right\} < \inf\left\{r(\boldsymbol{v}) : \boldsymbol{v} \in \mathbb{R}^N_{\le} \setminus K\right\}.$$

Proof. By Theorem 6.9 and Corollary 6.12, there exist

$$M_q = \min\left\{r(\boldsymbol{v}) : \boldsymbol{v} \in \mathbb{R}^q_{\le}\right\} \qquad (q = 1 \dots, N)$$

and $M = \min\{M_1, \dots, M_{N-1}\} > M_N$. Choose $\epsilon > 0$ with $\epsilon < (M - M_N)/3$. We can take $\alpha, \beta \in \mathbb{R}$ with $\alpha < \beta$ such that $F_j(\alpha) < \epsilon$ and $1 - F_j(\beta) < \epsilon$ ($j = 0, 1$), since F_j are nondecreasing functions with $\lim_{x \to -\infty} F_j(x) = 0$ and $\lim_{x \to \infty} F_j(x) = 1$. Let $K = [\alpha, \beta]^N \cap \mathbb{R}^N_{\le}$ and $\boldsymbol{v} = (v_1, \dots, v_N) \in \mathbb{R}^N_{\le} \setminus K$. Then $v_1 < \alpha$ or $v_N > \beta$ holds.

Suppose $v_1 < \alpha$ and put $\boldsymbol{v}' = (v_2, \ldots, v_N)$. Using Lemma 6.4, we obtain

$$\begin{aligned}\left|r(\boldsymbol{v}) - r(\boldsymbol{v}')\right| &= \left|\min_j F_j|_{-\infty}^{v_1} + \min_j F_j|_{v_1}^{v_2} - \min_j F_j|_{-\infty}^{v_2}\right| \\ &\leq \left|\min_j F_j|_{-\infty}^{v_1}\right| + \left|\min_j F_j|_{v_1}^{v_2} - \min_j F_j|_{-\infty}^{v_2}\right| \\ &< \epsilon + \left|F_0|_{v_1}^{v_2} - F_0|_{-\infty}^{v_2}\right| + \left|F_1|_{v_1}^{v_2} - F_1|_{-\infty}^{v_2}\right| \\ &= \epsilon + \left|F_0|_{-\infty}^{v_1}\right| + \left|F_1|_{-\infty}^{v_1}\right| \\ &< 3\epsilon.\end{aligned}$$

Hence $M \leq r(\boldsymbol{v}') \leq |r(\boldsymbol{v}') - r(\boldsymbol{v})| + r(\boldsymbol{v}) < 3\epsilon + r(\boldsymbol{v})$. We can similarly prove that $M < 3\epsilon + r(\boldsymbol{v})$ for the case $v_N > \beta$.

Therefore $M < 3\epsilon + r(\boldsymbol{v})$ for all $\boldsymbol{v} \in \mathbb{R}^N_{\leq} \setminus K$, so that

$$M_N < M - 3\epsilon \leq \inf\{r(\boldsymbol{v}) : \boldsymbol{v} \in \mathbb{R}^N_{\leq} \setminus K\}$$

holds. This is the claim. □

The measurability of $D_N(\mathcal{V}_{N,m,n}, \mathcal{V}_N)$ will be proved at the end of this section.

Theorem 6.19. *As $m, n \to \infty$, $D_N(\mathcal{V}_{N,m,n}, \mathcal{V}_N)$ converges almost surely to 0.*

Proof. In Lemma 6.17, let (A, d) be the subspace $\mathbb{R}^N_{\leq}$ of the Euclidean metric space $\mathbb{R}^N$, $g = r$ (which is continuous on $\mathbb{R}^N_{\leq}$), and $g_{i,j} = r_{i,j}$. Then by Remark 6.7, Theorem 6.15, and Lemma 6.18, the assumptions in Lemma 6.17 are satisfied almost surely, hence $D_N(\mathcal{V}_{N,m,n}, \mathcal{V}_N)$ converges almost surely to 0 as $m, n \to \infty$. □

Corollary 6.20. As $m, n \to \infty$, $\boldsymbol{v}_{m,n} \in \mathcal{V}_{N,m,n}$ converges almost surely to $\boldsymbol{c}$.

Proof. Since $\mathcal{V}_N = \{\boldsymbol{c}\}$ by Theorem 6.9, we have

$$D_N(\mathcal{V}_{N,m,n}, \mathcal{V}_N) = \sup_{\boldsymbol{v} \in \mathcal{V}_{N,m,n}} d(\boldsymbol{v}, \boldsymbol{c}) \geq d(\boldsymbol{v}_{m,n}, \boldsymbol{c}).$$

Hence the claim follows from Theorem 6.19. □

We cannot guarantee that $\boldsymbol{v}_{m,n}$ is necessarily measurable. We mean by "$\boldsymbol{v}_{m,n}$ converges almost surely to $\boldsymbol{c}$" that there exists a measurable set $A \subset \{\omega \in \Omega : \lim_{m,n\to\infty} \boldsymbol{v}_{m,n} = \boldsymbol{c}\}$ with $P(A) = 1$. In fact, we can take $A = \{\omega \in \Omega : \lim_{m,n\to\infty} D_N(\mathcal{V}_{N,m,n}, \mathcal{V}_N) = 0\}$. If $(\Omega, \mathfrak{A}, P)$ is complete, we have $P(\{\omega \in \Omega : \lim_{m,n\to\infty} \boldsymbol{v}_{m,n} = \boldsymbol{c}\}) = 1$.

Theorem 6.21. *As $m, n \to \infty$, $\rho_{N,m,n}$ converges almost surely to ρ.*

Proof. Let $\boldsymbol{v}_{m,n} = (v_1, \ldots, v_N) \in \mathcal{V}_{N,m,n}$, $v_0 = -\infty$, and $v_{N+1} = \infty$. By Definition 2.2, Lemma 6.4, and Theorem 6.9, we have

$$\begin{aligned}|\rho_{N,m,n} - \rho| &= |r_{m,n}(\boldsymbol{v}_{m,n}) - r(\boldsymbol{c})| \\ &\le \sum_{i=0}^{N} \left(\left| F_{0,m}|_{v_i}^{v_{i+1}} - F_0|_{c_i}^{c_{i+1}} \right| + \left| F_{1,n}|_{v_i}^{v_{i+1}} - F_1|_{c_i}^{c_{i+1}} \right| \right),\end{aligned}$$

where

$$\begin{aligned}\left| F_{0,m}|_{v_i}^{v_{i+1}} - F_0|_{c_i}^{c_{i+1}} \right| &= |F_{0,m}(v_{i+1}) - F_{0,m}(v_i) - F_0(c_{i+1}) + F_0(c_i)| \\ &\le |F_{0,m}(v_{i+1}) - F_0(v_{i+1})| + |F_0(v_{i+1}) - F_0(c_{i+1})| \\ &\quad + |F_{0,m}(v_i) - F_0(v_i)| + |F_0(v_i) - F_0(c_i)|\end{aligned}$$

and

$$\begin{aligned}\left| F_{1,n}|_{v_i}^{v_{i+1}} - F_1|_{c_i}^{c_{i+1}} \right| &= |F_{1,n}(v_{i+1}) - F_{1,n}(v_i) - F_1(c_{i+1}) + F_1(c_i)| \\ &\le |F_{1,n}(v_{i+1}) - F_1(v_{i+1})| + |F_1(v_{i+1}) - F_1(c_{i+1})| \\ &\quad + |F_{1,n}(v_i) - F_1(v_i)| + |F_1(v_i) - F_1(c_i)|.\end{aligned}$$

Now recall that we are considering the probability space $(\Omega, \mathfrak{A}, P)$. By Remark 6.2 and Theorem 6.3, there exists $A_1 \in \mathfrak{A}$ with $P(A_1) = 1$ such that for each $\omega \in A_1$,

$$\sup_{x\in\mathbb{R}} |F_{0,m}(x) - F_0(x)| \to 0, \qquad \sup_{x\in\mathbb{R}} |F_{1,n}(x) - F_1(x)| \to 0$$

as $m, n \to \infty$. Since F_0 and F_1 are continuous on $\mathbb{R}$ and $\boldsymbol{v}_{m,n} = (v_1, \ldots, v_N) \to \boldsymbol{c}$ almost surely as $m, n \to \infty$ by Corollary 6.20, there exists $A_2 \in \mathfrak{A}$ with $P(A_2) = 1$ such that for each $\omega \in A_2$,

$$|F_k(v_i) - F_k(c_i)| \to 0 \qquad (k = 0, 1;\ i = 1, \ldots, N)$$

as $m, n \to \infty$. Put $A = A_1 \cap A_2$. Then $P(A) = 1$. For each $\omega \in A$ and for any $\epsilon > 0$, there exist integers $N_1(\omega), N_2(\omega)$ such that $m, n \geq N_1(\omega)$ implies

$$\sup_{x\in\mathbb{R}} |F_{0,m}(x) - F_0(x)| < \epsilon, \qquad \sup_{x\in\mathbb{R}} |F_{1,n}(x) - F_1(x)| < \epsilon$$

and $m, n \geq N_2(\omega)$ implies

$$|F_k(v_i) - F_k(c_i)| < \epsilon \qquad (k = 0, 1;\ i = 1, \dots, N),$$

hence $m, n \geq N(\omega) = \max\{N_1(\omega), N_2(\omega)\}$ implies

$$\left| F_{0,m}|_{v_i}^{v_{i+1}} - F_0|_{c_i}^{c_{i+1}} \right| + \left| F_{1,n}|_{v_i}^{v_{i+1}} - F_1|_{c_i}^{c_{i+1}} \right| < 8\epsilon$$

for $i = 0, \dots, N$, and therefore

$$|\rho_{N,m,n}(\omega) - \rho| < \sum_{i=0}^{N} 8\epsilon = 8(N+1)\epsilon.$$

Since ϵ was arbitrary, $\rho_{N,m,n} \to \rho$ almost surely as $m, n \to \infty$. $\square$

Note that Theorem 6.21 is exactly Theorem 2.6.

Hereafter, we prove that $D_N(\mathcal{V}_{N,m,n}, \mathcal{V}_N)$ is measurable on Ω (related to Theorem 6.19).

Definition 6.22. Let $(Z_1, \gamma_1), \dots, (Z_{m+n}, \gamma_{m+n})$ be the rearrangement of $(X_1, 0), \dots, (X_m, 0), (Y_1, 1), \dots, (Y_n, 1)$ with $Z_1 \leq \dots \leq Z_{m+n}$, $\mathcal{T}$ be the set of tuples $(t_1, \dots, t_l)$ of positive integers with $t_1 + \dots + t_l = m + n$, $\mathbb{R}^{m+n}_{(t_1,\dots,t_l)}$ the set of real $m+n$-tuples $(v_1, \dots, v_{m+n})$ with $v_1 = \dots = v_{t_1} < v_{t_1+1} = \dots = v_{t_1+t_2} < \dots < v_{t_1+\dots+t_{l-1}+1} = \dots = v_{m+n}$, and $\mathcal{S}_{(t_1,\dots,t_l)} = (\{0,1\}^{t_1}/\sim) \times \dots \times (\{0,1\}^{t_l}/\sim)$, where $\{0,1\}^t/\sim$ denotes the t-th symmetric product of $\{0,1\}$. For $\boldsymbol{t} \in \mathcal{T}$ and $\boldsymbol{s} \in \mathcal{S}_{\boldsymbol{t}}$, let $\Omega_{\boldsymbol{t}} = \{\omega \in \Omega : (Z_1, \dots, Z_{m+n}) \in \mathbb{R}^{m+n}_{\boldsymbol{t}}\}$, $\Omega_{\boldsymbol{t},\boldsymbol{s}} = \{\omega \in \Omega_{\boldsymbol{t}} : (\gamma_1, \dots, \gamma_{m+n}) \text{ corresponds to } \boldsymbol{s}\}$. Put $I_0 = (-\infty, Z_1)$ and $I_i = [Z_i, Z_{i+1})$ for $i = 1, \dots, m+n$ where $Z_{m+n+1} = \infty$. Denote by $\mathcal{J}$ the set of N-tuples $(j_1, \dots, j_N)$ of integers with $0 \leq j_1 \leq \dots \leq j_N \leq m+n$. For $(j_1 \dots, j_N) \in \mathcal{J}$, define $I_{(j_1,\dots,j_N)} = (I_{j_1} \times \dots \times I_{j_N}) \cap \mathbb{R}^N_{\leq}$.

Remark 6.23. Since $\mathbb{R}^{m+n}_{\boldsymbol{t}}$ are measurable and pairwise disjoint for $\boldsymbol{t} \in \mathcal{T}$, so are $\Omega_{\boldsymbol{t}}$. In addition, $\mathbb{R}^{m+n}_{\leq} = \bigcup_{\boldsymbol{t}\in\mathcal{T}} \mathbb{R}^{m+n}_{\boldsymbol{t}}$ implies that $\Omega = \bigcup_{\boldsymbol{t}\in\mathcal{T}} \Omega_{\boldsymbol{t}}$, where $\Omega_{\boldsymbol{t}}$ equals the disjoint union of $\Omega_{\boldsymbol{t},\boldsymbol{s}} \in \mathfrak{A}$ over $\boldsymbol{s} \in \mathcal{S}_{\boldsymbol{t}}$. Besides, for any

$\boldsymbol{t} \in \mathcal{T}$, there exists a nonempty set $\mathcal{J}_{\boldsymbol{t}} \subset \mathcal{J}$ such that, on the event $\Omega_{\boldsymbol{t}}$, $\mathbb{R}^N_{\leq}$ equals the disjoint union of $I_{\boldsymbol{j}}$ $(\neq \emptyset)$ over $\boldsymbol{j} \in \mathcal{J}_{\boldsymbol{t}}$.

For $\boldsymbol{t} \in \mathcal{T}$ and $\boldsymbol{s} \in \mathcal{S}_{\boldsymbol{t}}$, consider the event $\Omega_{\boldsymbol{t},\boldsymbol{s}}$ in the case $\Omega_{\boldsymbol{t},\boldsymbol{s}} \neq \emptyset$. Then $r_{m,n}$ is constant on $I_{\boldsymbol{j}}$ for any $\boldsymbol{j} \in \mathcal{J}_{\boldsymbol{t}}$, since it depends only on the rank statistics of $X_1, \ldots, X_m, Y_1, \ldots, Y_n$. Furthermore, there exists a nonempty set $\mathcal{J}_{\boldsymbol{t},\boldsymbol{s}} \subset \mathcal{J}_{\boldsymbol{t}}$ such that $\mathcal{V}_{N,m,n} = \arg\min_{\boldsymbol{v} \in \mathbb{R}^N_{\leq}} r_{m,n}(\boldsymbol{v})$ equals the disjoint union of $I_{\boldsymbol{j}}$ $(\neq \emptyset)$ over $\boldsymbol{j} \in \mathcal{J}_{\boldsymbol{t},\boldsymbol{s}}$.

Theorem 6.24. *$D_N(\mathcal{V}_{N,m,n}, \mathcal{V}_N)$ is measurable on Ω.*

Proof. Let $\boldsymbol{t} \in \mathcal{T}$ and $\boldsymbol{s} \in \mathcal{S}_{\boldsymbol{t}}$. Since Ω equals the disjoint union of $\Omega_{\boldsymbol{t},\boldsymbol{s}} \in \mathfrak{A}$ over $\boldsymbol{t} \in \mathcal{T}$ and $\boldsymbol{s} \in \mathcal{S}_{\boldsymbol{t}}$ by Remark 6.23, it suffices to prove the measurability of $D_N(\mathcal{V}_{N,m,n}, \mathcal{V}_N)$ on each $\Omega_{\boldsymbol{t},\boldsymbol{s}}$. In the rest of the proof, we restrict $D_N(\mathcal{V}_{N,m,n}, \mathcal{V}_N)$ to $\Omega_{\boldsymbol{t},\boldsymbol{s}}$, which gives $D_N(\mathcal{V}_{N,m,n}, \mathcal{V}_N) = \max_{\boldsymbol{j} \in \mathcal{J}_{\boldsymbol{t},\boldsymbol{s}}} D_N(I_{\boldsymbol{j}}, \{\boldsymbol{c}\})$.

If there exists $\boldsymbol{j} = (j_1, \ldots, j_N) \in \mathcal{J}_{\boldsymbol{t},\boldsymbol{s}}$ such that $j_1 = 0$ or $j_N = m+n$, then $I_{\boldsymbol{j}} = ((-\infty, Z_1) \times I_{j_2} \times \cdots \times I_{j_N}) \cap \mathbb{R}^N_{\leq}$ or $I_{\boldsymbol{j}} = (I_{j_1} \times \cdots \times I_{j_{N-1}} \times [Z_{m+n}, \infty)) \cap \mathbb{R}^N_{\leq}$, so that $D_N(\mathcal{V}_{N,m,n}, \mathcal{V}_N) = \infty$ (the measurability is obvious).

Next, let $\boldsymbol{j} = (j_1, \ldots, j_N) \in \mathcal{J}_{\boldsymbol{t},\boldsymbol{s}}$ with $1 \leq j_1 \leq \cdots \leq j_N \leq m+n-1$. Then the closure $\overline{I_{\boldsymbol{j}}}$ of $I_{\boldsymbol{j}}$ equals $([Z_{j_1}, Z_{j_1+1}] \times \cdots \times [Z_{j_N}, Z_{j_N+1}]) \cap \mathbb{R}^N_{\leq}$. Since $I_{\boldsymbol{j}} \neq \emptyset$, we have $Z_{j_k} < Z_{j_k+1}$ for all $k = 1, \ldots, N$. Put $V_{\boldsymbol{j}} = \{(v_1, \ldots, v_N) \in \overline{I_{\boldsymbol{j}}} : v_i \in \{Z_{j_i}, Z_{j_i+1}\} \text{ for } i = 1, \ldots, N\}$. We can see that $\overline{I_{\boldsymbol{j}}}$ is the convex hull of the finite vertex set $V_{\boldsymbol{j}}$, so that $\sup_{\boldsymbol{v} \in \overline{I_{\boldsymbol{j}}}} d(\boldsymbol{v}, \boldsymbol{c}) = \max_{\boldsymbol{v} \in V_{\boldsymbol{j}}} d(\boldsymbol{v}, \boldsymbol{c})$ since $\overline{I_{\boldsymbol{j}}} \supset V_{\boldsymbol{j}}$ by definition and the closed ball with center $\boldsymbol{c}$ and radius $\max_{\boldsymbol{v} \in V_{\boldsymbol{j}}} d(\boldsymbol{v}, \boldsymbol{c})$ contains $\overline{I_{\boldsymbol{j}}}$. Noting that $\sup_{\boldsymbol{v} \in I_{\boldsymbol{j}}} d(\boldsymbol{v}, \boldsymbol{c}) = \sup_{\boldsymbol{v} \in \overline{I_{\boldsymbol{j}}}} d(\boldsymbol{v}, \boldsymbol{c})$, we obtain $D_N(I_{\boldsymbol{j}}, \{\boldsymbol{c}\}) = \max_{\boldsymbol{v} \in V_{\boldsymbol{j}}} d(\boldsymbol{v}, \boldsymbol{c})$. Since $d(\boldsymbol{v}, \boldsymbol{c})$ for each $\boldsymbol{v} \in V_{\boldsymbol{j}}$ is obviously measurable on $\Omega_{\boldsymbol{t},\boldsymbol{s}}$, $D_N(I_{\boldsymbol{j}}, \{\boldsymbol{c}\})$ and also $D_N(\mathcal{V}_{N,m,n}, \mathcal{V}_N)$ are measurable on $\Omega_{\boldsymbol{t},\boldsymbol{s}}$. $\square$

7. Proof for Theorem 3.9

Here we assume the same setting as in Section 3.2. We denote by $R[x]$ and $R[[x]]$ the rings of polynomials and formal power series in x over a ring R, respectively.

Definition 7.1. For $\boldsymbol{\gamma} \in \Gamma_{k,k}$ $(k \in \mathbb{N})$, define

$$\begin{aligned} \delta_{\boldsymbol{\gamma}}(i) &= N_0(\boldsymbol{\gamma}_{0:i}) - N_1(\boldsymbol{\gamma}_{0:i}) && (i = 0, 1, \ldots, 2k), \\ d_{\boldsymbol{\gamma}}(i,j) &= |\delta_{\boldsymbol{\gamma}}(i)| + |\delta_{\boldsymbol{\gamma}}(i) - \delta_{\boldsymbol{\gamma}}(j)| + |\delta_{\boldsymbol{\gamma}}(j)| && (i, j = 0, 1, \ldots, 2k). \end{aligned}$$

Note that $\delta_{\boldsymbol{\gamma}}(0) = \delta_{\boldsymbol{\gamma}}(2k) = 0$,

$$\delta_{\boldsymbol{\gamma}}(i) = k\left(\widehat{F}_{0,\boldsymbol{\gamma}}(i) - \widehat{F}_{1,\boldsymbol{\gamma}}(i)\right) \qquad (k > 0), \tag{14}$$

and

$$d_{\boldsymbol{\gamma}}(i,j) = d_{\boldsymbol{\gamma}}(j,i) \tag{15}$$

by definition.

Lemma 7.2. For all $\boldsymbol{\gamma} \in \Gamma_{n,n}$,

$$\widehat{\rho}_2(\boldsymbol{\gamma}) = 1 - \frac{1}{2n} \max_{0 \le j_1, j_2 \le 2n} d_{\boldsymbol{\gamma}}(j_1, j_2).$$

Proof. Let us put

$$\widehat{s}_{\boldsymbol{\gamma}}(j_1, j_2) = \sum_{i=0}^{2} \max\left\{ \widehat{F}_{0,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}}, \widehat{F}_{1,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}} \right\} \qquad (0 \le j_1 \le j_2 \le 2n),$$

where $j_0 = 0$, $j_3 = 2n$. Then

$$\begin{aligned} &\widehat{s}_{\boldsymbol{\gamma}}(j_1, j_2) + \widehat{r}_{\boldsymbol{\gamma}}(j_1, j_2) \\ &= \sum_{i=0}^{2} \left(\max\left\{ \widehat{F}_{0,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}}, \widehat{F}_{1,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}} \right\} + \min\left\{ \widehat{F}_{0,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}}, \widehat{F}_{1,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}} \right\} \right) \\ &= \sum_{i=0}^{2} \left(\widehat{F}_{0,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}} + \widehat{F}_{1,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}} \right) \\ &= 2. \end{aligned}$$

On the other hand, by (14),

$$
\begin{aligned}
&\widehat{s}_{\boldsymbol{\gamma}}(j_1, j_2) - \widehat{r}_{\boldsymbol{\gamma}}(j_1, j_2) \\
&= \sum_{i=0}^{2} \left(\max\left\{ \widehat{F}_{0,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}}, \widehat{F}_{1,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}} \right\} - \min\left\{ \widehat{F}_{0,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}}, \widehat{F}_{1,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}} \right\} \right) \\
&= \sum_{i=0}^{2} \left| \widehat{F}_{0,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}} - \widehat{F}_{1,\boldsymbol{\gamma}}|_{j_i}^{j_{i+1}} \right| \\
&= \frac{1}{n} \sum_{i=0}^{2} |\delta_{\boldsymbol{\gamma}}(j_{i+1}) - \delta_{\boldsymbol{\gamma}}(j_i)| \\
&= \frac{d_{\boldsymbol{\gamma}}(j_1, j_2)}{n}.
\end{aligned}
$$

Hence we have

$$
\widehat{r}_{\boldsymbol{\gamma}}(j_1, j_2) = 1 - \frac{d_{\boldsymbol{\gamma}}(j_1, j_2)}{2n},
$$

so that

$$
\begin{aligned}
\widehat{\rho}_2(\boldsymbol{\gamma}) &= \min_{0 \le j_1 \le j_2 \le 2n} \widehat{r}_{\boldsymbol{\gamma}}(j_1, j_2) \\
&= 1 - \frac{1}{2n} \max_{0 \le j_1 \le j_2 \le 2n} d_{\boldsymbol{\gamma}}(j_1, j_2) \\
&= 1 - \frac{1}{2n} \max_{0 \le j_1, j_2 \le 2n} d_{\boldsymbol{\gamma}}(j_1, j_2)
\end{aligned}
$$

by (15). $\square$

Definition 7.3. For $\boldsymbol{\gamma} \in \Gamma_{k,k}$ $(k \in \mathbb{N})$, define

$$
\overline{\delta}_{\boldsymbol{\gamma}} = \max_{0 \le i \le 2k} \delta_{\boldsymbol{\gamma}}(i), \qquad \underline{\delta}_{\boldsymbol{\gamma}} = \min_{0 \le i \le 2k} \delta_{\boldsymbol{\gamma}}(i).
$$

Note that $\underline{\delta}_{\boldsymbol{\gamma}} \le 0 \le \overline{\delta}_{\boldsymbol{\gamma}}$ since $\delta_{\boldsymbol{\gamma}}(0) = 0$.

Lemma 7.4. For all $\boldsymbol{\gamma} \in \Gamma_{n,n}$,

$$
\max_{0 \le i,j \le 2n} d_{\boldsymbol{\gamma}}(i, j) = 2\big(\overline{\delta}_{\boldsymbol{\gamma}} - \underline{\delta}_{\boldsymbol{\gamma}}\big).
$$

Proof. Denote $\overline{\delta}_{\gamma,i,j} = \max\{\delta_\gamma(i), \delta_\gamma(j)\}$ and $\underline{\delta}_{\gamma,i,j} = \min\{\delta_\gamma(i), \delta_\gamma(j)\}$. Note that $\overline{\delta}_{\gamma,i,j} + \underline{\delta}_{\gamma,i,j} = \delta_\gamma(i) + \delta_\gamma(j)$ and $\overline{\delta}_{\gamma,i,j} - \underline{\delta}_{\gamma,i,j} = |\delta_\gamma(i) - \delta_\gamma(j)|$.

If $\delta_\gamma(i) > 0$ and $\delta_\gamma(j) > 0$, then

$$\begin{aligned} d_\gamma(i,j) &= \delta_\gamma(i) + \delta_\gamma(j) + |\delta_\gamma(i) - \delta_\gamma(j)| \\ &= \overline{\delta}_{\gamma,i,j} + \underline{\delta}_{\gamma,i,j} + \overline{\delta}_{\gamma,i,j} - \underline{\delta}_{\gamma,i,j} \\ &= 2\overline{\delta}_{\gamma,i,j} \\ &\le 2\overline{\delta}_\gamma \\ &\le 2\big(\overline{\delta}_\gamma - \underline{\delta}_\gamma\big). \end{aligned}$$

If $\delta_\gamma(i) < 0$ and $\delta_\gamma(j) < 0$, then

$$\begin{aligned} d_\gamma(i,j) &= -\delta_\gamma(i) - \delta_\gamma(j) + |\delta_\gamma(i) - \delta_\gamma(j)| \\ &= -\big(\overline{\delta}_{\gamma,i,j} + \underline{\delta}_{\gamma,i,j}\big) + \overline{\delta}_{\gamma,i,j} - \underline{\delta}_{\gamma,i,j} \\ &= -2\underline{\delta}_{\gamma,i,j} \\ &\le -2\underline{\delta}_\gamma \\ &\le 2\big(\overline{\delta}_\gamma - \underline{\delta}_\gamma\big). \end{aligned}$$

If $\delta_\gamma(i)\delta_\gamma(j) \le 0$, then

$$\begin{aligned} d_\gamma(i,j) &= 2|\delta_\gamma(i) - \delta_\gamma(j)| \\ &\le 2\big(\overline{\delta}_\gamma - \underline{\delta}_\gamma\big). \end{aligned}$$

Taken together, we have $d_\gamma(i,j) \le 2\big(\overline{\delta}_\gamma - \underline{\delta}_\gamma\big)$ in general. On the other hand, if $\delta_\gamma(i) = \overline{\delta}_\gamma$ and $\delta_\gamma(j) = \underline{\delta}_\gamma$, then

$$\begin{aligned} d_\gamma(i,j) &= \big|\overline{\delta}_\gamma\big| + \big|\overline{\delta}_\gamma - \underline{\delta}_\gamma\big| + \big|\underline{\delta}_\gamma\big| \\ &= \overline{\delta}_\gamma + \big(\overline{\delta}_\gamma - \underline{\delta}_\gamma\big) - \underline{\delta}_\gamma \\ &= 2\big(\overline{\delta}_\gamma - \underline{\delta}_\gamma\big). \end{aligned}$$

This completes the proof. □

Theorem 7.5. *For all* $\gamma \in \Gamma_{n,n}$,

$$\widehat{\rho}_2(\gamma) = 1 - \frac{\overline{\delta}_\gamma - \underline{\delta}_\gamma}{n}.$$

Proof. This follows immediately from Lemmas 7.2 and 7.4. □

The following arguments (from Definition 7.6 to Theorem 7.15) refer to [3, Section I].

Definition 7.6. A *combinatorial class* is a set $\mathcal{A}$ on which a *size function* $|\cdot| : \mathcal{A} \to \mathbb{N}$ is defined so that $\{\alpha \in \mathcal{A} : |\alpha| = k\}$ is finite for all $k \in \mathbb{N}$. Unless confusion arises, we simply say a *class* instead of a combinatorial class.

Any subset $\mathcal{B} \subset \mathcal{A}$ is also a class with its size function defined as in $\mathcal{A}$. The *counting sequence* $\{a_k\}$ of $\mathcal{A}$ is defined by

$$a_k = \#\{\alpha \in \mathcal{A} : |\alpha| = k\} \qquad (k \in \mathbb{N}),$$

and the *ordinary generating function (OGF)* $A(x) \in \mathbb{Z}[[x]]$ of $\mathcal{A}$ is by

$$A(x) = \sum_{k=0}^{\infty} a_k x^k.$$

Definition 7.7. Let $\mathcal{A}$ and $\mathcal{B}$ be two classes. A map $\phi : \mathcal{A} \to \mathcal{B}$ is called a *homomorphism* between $\mathcal{A}$ and $\mathcal{B}$ if $|\alpha| = |\phi(\alpha)|$ for all $\alpha \in \mathcal{A}$. If, in addition, ϕ is bijective, then we call ϕ an *isomorphism*, say that $\mathcal{A}$ and $\mathcal{B}$ are *isomorphic*, or write $\mathcal{A} \cong \mathcal{B}$.

Remark 7.8. Let $\mathcal{A}$ and $\mathcal{B}$ be two classes, $\{a_k\}$ and $\{b_k\}$ their counting sequences, and $A(x)$ and $B(x)$ their OGFs, respectively. We can easily see that the following three statements are equivalent:

1. $\mathcal{A} \cong \mathcal{B}$.

2. $a_k = b_k$ for all $k \in \mathbb{N}$.

3. $A(x) = B(x)$.

Definition 7.9. The *neutral class* $\mathcal{E}$ and the *atomic class* $\mathcal{Z}$ are classes with $\#\mathcal{E} = \#\mathcal{Z} = 1$, $|\varepsilon| = 0$ ($\varepsilon \in \mathcal{E}$), and $|\zeta| = 1$ ($\zeta \in \mathcal{Z}$).

Remark 7.10. The OGFs of $\mathcal{E}$ and $\mathcal{Z}$ are 1 and x in $\mathbb{Z}[[x]]$, respectively.

Definition 7.11. Let $\{A_i\}$ be a set of classes, where i runs over some index set I. If $\mathcal{B} = \{(i, \alpha) : i \in I,\ \alpha \in \mathcal{A}_i\}$ is also a class with its size function defined by $|(i, \alpha)| = |\alpha|$, then we call $\mathcal{B}$ the *combinatorial sum* (or simply the *sum*) of $\{\mathcal{A}_i\}$ and write $\mathcal{B} = \bigsqcup_{i \in I} \mathcal{A}_i$. In particular, if $I = \{1, \ldots, k\}$, then $\mathcal{B}$ is always a class, and we may write $\mathcal{B} = \mathcal{A}_1 + \cdots + \mathcal{A}_k$.

Definition 7.12. The *Cartesian product* (or simply the *product*) $\mathcal{A}_1 \times \cdots \times \mathcal{A}_k$ of k classes $\mathcal{A}_1, \ldots, \mathcal{A}_k$ is the class $\{(\alpha_1, \ldots, \alpha_k) : \alpha_1 \in \mathcal{A}_1, \ldots, \alpha_k \in \mathcal{A}_k\}$ whose size function is defined by $|(\alpha_1, \ldots, \alpha_k)| = |\alpha_1| + \cdots + |\alpha_k|$.

For a class $\mathcal{A}$ and $k \in \mathbb{N}_+$, we may write $\mathcal{A}^k$ instead of $\mathcal{A} \times \cdots \times \mathcal{A}$ (k times). Let $\mathcal{A}^0 = \mathcal{E} = \{\varepsilon\}$. If a class $\mathcal{B} = \bigsqcup_{i \in \mathbb{N}} \mathcal{A}^i$ exists, then we call $\mathcal{B}$ a *sequence class* of $\mathcal{A}$, and write $\mathcal{B} = \textsc{Seq}(\mathcal{A})$.

Remark 7.13. If $A_1(x), \ldots, A_k(x)$ are the OGFs of classes $\mathcal{A}_1, \ldots, \mathcal{A}_k$, respectively, then the OGFs of $\mathcal{A}_1 + \ldots + \mathcal{A}_k$ and $\mathcal{A}_1 \times \cdots \times \mathcal{A}_k$ are $A_1(x) + \cdots + A_k(x)$ and $A_1(x) \cdots A_k(x)$, respectively.

Theorem 7.14. (See [3, Section I.2.1] for reference.) *Let $\{a_i\}$ be the counting sequence of a class $\mathcal{A}$. Then $\textsc{Seq}(\mathcal{A})$ exists if and only if $a_0 = 0$.*

Theorem 7.15. (See [3, Section I.2.2, Theorem I.1] for the proof.) *Let $A(x)$ be the OGF of a class $\mathcal{A}$ and assume that $\textsc{Seq}(\mathcal{A})$ exists. Then the OGF of $\textsc{Seq}(\mathcal{A})$ is $1/(1 - A(x))$.*

Definition 7.16. Define a class $\mathcal{G}$ by

$$\mathcal{G} = \bigcup_{i=0}^{\infty} \Gamma_{i,i},$$
$$|\boldsymbol{\gamma}| = i \qquad (\boldsymbol{\gamma} \in \Gamma_{i,i}).$$

For $k, l \in \mathbb{N}$, let $\mathcal{G}_{k,l} = \{\boldsymbol{\gamma} \in \mathcal{G} : -k \le \underline{\delta}_{\boldsymbol{\gamma}},\ \overline{\delta}_{\boldsymbol{\gamma}} \le l\}$ and $G_{k,l}(x)$ be the OGF of $\mathcal{G}_{k,l}$. For $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_{2i}) \in \Gamma_{i,i}$ ($i \ge 1$), define

$$\lambda^+(\boldsymbol{\gamma}) = (0, \gamma_1, \ldots, \gamma_{2i}, 1) \in \Gamma_{i+1,i+1},$$
$$\lambda^-(\boldsymbol{\gamma}) = (1, \gamma_1, \ldots, \gamma_{2i}, 0) \in \Gamma_{i+1,i+1}.$$

Put $\lambda^+(e) = (0, 1) \in \Gamma_{1,1}$ and $\lambda^-(e) = (1, 0) \in \Gamma_{1,1}$ for $e \in \Gamma_{0,0}$. Note that λ^+ and λ^- are injective on $\mathcal{G}$.

Lemma 7.17. For any $\mathcal{H} \subset \mathcal{G}$, $\mathcal{H} \times \mathcal{Z} \cong \lambda^+(\mathcal{H}) \cong \lambda^-(\mathcal{H})$.

Proof. Since $|(\gamma,\zeta)| = |\gamma| + |\zeta| = |\gamma| + 1 = |\lambda^+(\gamma)|$ for all $(\gamma,\zeta) \in \mathcal{H} \times \mathcal{Z}$, the bijection $\nu^+ : \mathcal{H} \times \mathcal{Z} \to \lambda^+(\mathcal{H})$ defined by $\nu^+(\gamma,\zeta) = \lambda^+(\gamma)$ is a homomorphism, hence $\mathcal{H} \times \mathcal{Z} \cong \lambda^+(\mathcal{H})$. Similarly, the bijection $\nu^- : \mathcal{H} \times \mathcal{Z} \to \lambda^-(\mathcal{H})$ defined by $\nu^-(\gamma,\zeta) = \lambda^-(\gamma)$ is a homomorphism, hence $\mathcal{H} \times \mathcal{Z} \cong \lambda^-(\mathcal{H})$. □

Corollary 7.18. If $H(x)$ is the OGF of $\mathcal{H} \subset \mathcal{G}$, then the OGFs of $\lambda^+(\mathcal{H})$ and $\lambda^-(\mathcal{H})$ are both equal to $xH(x)$.

Proof. By Remark 7.8 and Lemma 7.17, the OGFs of $\lambda^+(\mathcal{H})$ and $\lambda^-(\mathcal{H})$ are equal to that of $\mathcal{H} \times \mathcal{Z}$, which equals $xH(x)$ by Remarks 7.10 and 7.13. □

Lemma 7.19. For all $k,l \in \mathbb{N}$, we have $\mathcal{G}_{k+1,0} \cong \textsc{Seq}(\lambda^-(\mathcal{G}_{k,0}))$, $\mathcal{G}_{0,l+1} \cong \textsc{Seq}(\lambda^+(\mathcal{G}_{0,l}))$, and $\mathcal{G}_{k+1,l+1} \cong \textsc{Seq}(\lambda^-(\mathcal{G}_{k,0}) + \lambda^+(\mathcal{G}_{0,l}))$.

Proof. Define a map $\sigma : \mathcal{G}_{k+1,0} \to \textsc{Seq}(\lambda^-(\mathcal{G}_{k,0}))$ by $\sigma(e) = (0,\varepsilon)$ where $\varepsilon \in \mathcal{E} = (\lambda^-(\mathcal{G}_{k,0}))^0$, and by

$$\sigma(\gamma) = \Big(p, (\gamma_{j_0:j_1}, \dots, \gamma_{j_{p-1}:j_p})\Big) \qquad (\gamma \in \Gamma_{i,i};\ i \geq 1)$$

where $\{j_0,\dots,j_p\} = \{j \in \{0,\dots,2i\} : \delta_\gamma(j) = 0\}$ and $0 = j_0 < \cdots < j_p = 2i$. It follows from definition that σ is bijective and $|\sigma(\gamma)| = |\gamma|$ for all $\gamma \in \mathcal{G}_{k+1,0}$, so that σ is an isomorphism, i.e., $\mathcal{G}_{k+1,0} \cong \textsc{Seq}(\lambda^-(\mathcal{G}_{k,0}))$. We can similarly show that $\mathcal{G}_{0,l+1} \cong \textsc{Seq}(\lambda^+(\mathcal{G}_{0,l}))$ and $\mathcal{G}_{k+1,l+1} \cong \textsc{Seq}(\lambda^-(\mathcal{G}_{k,0}) + \lambda^+(\mathcal{G}_{0,l}))$. □

Lemma 7.20. For all $k,l \in \mathbb{N}_+$,

$$Q_{k+1}(x)Q_{l+1}(x) - x^2Q_{k-1}(x)Q_{l-1}(x) = Q_{k+l+1}(x). \tag{16}$$

Proof. Since $Q_2(x) = Q_1(x) - xQ_0(x) = 1 - x$ by Definition 3.8, we have

$$\begin{aligned} Q_{k+1}(x)Q_2(x) &= Q_{k+1}(x) - xQ_{k+1}(x) \\ &= Q_{k+1}(x) - x(Q_k(x) - xQ_{k-1}(x)) \\ &= Q_{k+1}(x) - xQ_k(x) + x^2Q_{k-1}(x) \\ &= Q_{k+2}(x) + x^2Q_{k-1}(x) \qquad (17) \\ &= Q_{k+2}(x) + x^2Q_{k-1}(x)Q_0(x), \end{aligned}$$

hence (16) holds for $l = 1$. We also have

$$\begin{aligned}
Q_{k+1}(x)Q_3(x) &= Q_{k+1}(x)(Q_2(x) - xQ_1(x)) \\
&= Q_{k+1}(x)Q_2(x) - xQ_{k+1}(x) \\
&= Q_{k+2}(x) + x^2Q_{k-1}(x) - xQ_{k+1}(x) \\
&= Q_{k+3}(x) + x^2Q_{k-1}(x) \\
&= Q_{k+3}(x) + x^2Q_{k-1}(x)Q_1(x)
\end{aligned}$$

by (17), hence (16) holds for $l = 2$.

By Definition 3.8, we have $\boldsymbol{Q}_{j+1}(x) = R\boldsymbol{Q}_j(x)$ for all $j \in \mathbb{N}$, where

$$\boldsymbol{Q}_i(x) = \begin{pmatrix} Q_i(x) \\ Q_{i+1}(x) \end{pmatrix}, \quad R = \begin{pmatrix} 0 & 1 \\ -x & 1 \end{pmatrix}.$$

If (16) holds for $l = i \in \mathbb{N}_+$ and $l = i + 1$, or equivalently

$$Q_{k+1}(x)\boldsymbol{Q}_{i+1}(x) - x^2Q_{k-1}(x)\boldsymbol{Q}_{i-1}(x) = \boldsymbol{Q}_{k+i+1}(x)$$

holds, then

$$\begin{aligned}
&Q_{k+1}(x)\boldsymbol{Q}_{i+2}(x) - x^2Q_{k-1}(x)\boldsymbol{Q}_i(x) \\
&= Q_{k+1}(x)R\boldsymbol{Q}_{i+1}(x) - x^2Q_{k-1}(x)R\boldsymbol{Q}_{i-1}(x) \\
&= R(Q_{k+1}(x)\boldsymbol{Q}_{i+1}(x) - x^2Q_{k-1}(x)\boldsymbol{Q}_{i-1}(x)) \\
&= R\boldsymbol{Q}_{k+i+1}(x) \\
&= \boldsymbol{Q}_{k+i+2}(x),
\end{aligned}$$

hence (16) holds for $l = i + 2$. The claim follows by induction. □

Lemma 7.21. For all $k \in \mathbb{N}$,

$$\sum_{0 \le i < k} Q_i(x)Q_{k-i-1}(x) = -Q'_{k+1}(x).$$

Proof. Define $\mathfrak{Q}(x, t) \in (\mathbb{Z}[x])[[t]]$ as

$$\mathfrak{Q}(x, t) = \sum_{k=0}^{\infty} Q_k(x)t^k.$$

Since

$$\mathfrak{Q}(x,t) = Q_0(x) + Q_1(x)t + \sum_{k=2}^{\infty} Q_k(x)t^k,$$
$$t\mathfrak{Q}(x,t) = Q_0(x)t + \sum_{k=2}^{\infty} Q_{k-1}(x)t^k,$$
$$xt^2\mathfrak{Q}(x,t) = x\sum_{k=2}^{\infty} Q_{k-2}(x)t^k,$$

we have

$$\begin{aligned}(1-t+xt^2)\mathfrak{Q}(x,t) &= Q_0(x) + Q_1(x)t - Q_0(x)t \\ &\qquad + \sum_{k=2}^{\infty}(Q_k(x) - Q_{k-1}(x) + xQ_{k-2}(x))t^k \\ &= 1\end{aligned}$$

by Definition 3.8, hence

$$\mathfrak{Q}(x,t) = \frac{1}{1-t+xt^2}. \tag{18}$$

Therefore

$$\begin{aligned}\sum_{k=0}^{\infty} Q_k'(x)t^k &= \frac{\partial}{\partial x}\mathfrak{Q}(x,t) = -\frac{t^2}{(1-t+xt^2)^2} \\ &= -t^2(\mathfrak{Q}(x,t))^2 = -\sum_{k=0}^{\infty}\left(\sum_{0\le i<k} Q_i(x)Q_{k-i-1}(x)\right)t^{k+1},\end{aligned}$$

which implies the claim. □

Proposition 7.22. For all $k \in \mathbb{N}$,

$$G_{k,0}(x) = G_{0,k}(x) = \frac{Q_k(x)}{Q_{k+1}(x)}. \tag{19}$$

Proof. Since $G_{0,0}(x) = 1 = Q_0(x)/Q_1(x)$, (19) holds for $k = 0$.

Suppose (19) holds for some $k \in \mathbb{N}$. Since $\mathcal{G}_{k+1,0} \cong \textsc{Seq}(\lambda^-(\mathcal{G}_{k,0}))$ by Lemma 7.19, we have

$$G_{k+1,0}(x) = \frac{1}{1 - xG_{k,0}(x)} = \frac{Q_{k+1}(x)}{Q_{k+1}(x) - xQ_k(x)} = \frac{Q_{k+1}(x)}{Q_{k+2}(x)}$$

by Definition 3.8, Theorem 7.15, and Corollary 7.18. Similarly, since $\mathcal{G}_{0,k+1} \cong \textsc{Seq}(\lambda^+(\mathcal{G}_{0,k}))$ by Lemma 7.19, we obtain

$$G_{0,k+1}(x) = \frac{Q_{k+1}(x)}{Q_{k+2}(x)}.$$

Therefore, (19) holds for $k+1$ in place of k, and the proof is complete. □

Proposition 7.23. For all $k, l \in \mathbb{N}$,

$$G_{k,l}(x) = \frac{Q_k(x)Q_l(x)}{Q_{k+l+1}(x)}. \tag{20}$$

Proof. If $k = 0$ or $l = 0$, then (20) holds by Proposition 7.22.

Suppose $k \geq 1$ and $l \geq 1$. Since $\mathcal{G}_{k,l} \cong \textsc{Seq}(\lambda^-(\mathcal{G}_{k-1,0}) + \lambda^+(\mathcal{G}_{0,l-1}))$ by Lemma 7.19, we have

$$G_{k,l}(x) = \frac{1}{1 - xG_{k-1,0}(x) - xG_{0,l-1}(x)}$$

by Remark 7.13, Theorem 7.15, and Corollary 7.18, where

$$G_{k-1,0}(x) = \frac{Q_{k-1}(x)}{Q_k(x)}, \qquad G_{0,l-1}(x) = \frac{Q_{l-1}(x)}{Q_l(x)}$$

by Proposition 7.22. Hence

$$\begin{aligned} G_{k,l}(x) &= \frac{Q_k(x)Q_l(x)}{Q_k(x)Q_l(x) - xQ_{k-1}(x)Q_l(x) - xQ_k(x)Q_{l-1}(x)} \\ &= \frac{Q_k(x)Q_l(x)}{(Q_k(x) - xQ_{k-1}(x))(Q_l(x) - xQ_{l-1}(x)) - x^2Q_{k-1}(x)Q_{l-1}(x)} \\ &= \frac{Q_k(x)Q_l(x)}{Q_{k+1}(x)Q_{l+1}(x) - x^2Q_{k-1}(x)Q_{l-1}(x)} \\ &= \frac{Q_k(x)Q_l(x)}{Q_{k+l+1}(x)} \end{aligned}$$

by Definition 3.8 and Lemma 7.20. □

Proposition 7.24. For $k \in \mathbb{N}$, the OGF of $\widetilde{\mathcal{G}}_k = \{\boldsymbol{\gamma} \in \mathcal{G} : \overline{\delta}_{\boldsymbol{\gamma}} - \underline{\delta}_{\boldsymbol{\gamma}} \le k\}$ is

$$\widetilde{G}_k(x) = \frac{Q'_{k+1}(x)}{Q_k(x)} - \frac{Q'_{k+2}(x)}{Q_{k+1}(x)}.$$

Proof. For $i, j \in \mathbb{N}$, let $\widetilde{\mathcal{G}}_{i,j} = \{\boldsymbol{\gamma} \in \mathcal{G} : -i = \underline{\delta}_{\boldsymbol{\gamma}}, \overline{\delta}_{\boldsymbol{\gamma}} \le j\}$ and $\widetilde{G}_{i,j}$ be the OGF of $\widetilde{\mathcal{G}}_{i,j}$. Since $\widetilde{\mathcal{G}}_{0,j} = \mathcal{G}_{0,j}$, $\mathcal{G}_{i,j} \cong \mathcal{G}_{i-1,j} + \widetilde{\mathcal{G}}_{i,j}$ if $i \ge 1$, and $\widetilde{\mathcal{G}}_k = \bigsqcup_{i=0}^{k} \widetilde{\mathcal{G}}_{i,k-i}$ by definition, we have

$$\begin{aligned}
\widetilde{G}_k(x) &= \sum_{i=0}^{k} \widetilde{G}_{i,k-i}(x) \\
&= G_{0,k}(x) + \sum_{1 \le i < k+1} (G_{i,k-i}(x) - G_{i-1,k-i}(x)) \\
&= \sum_{0 \le i < k+1} G_{i,k-i}(x) - \sum_{0 \le i < k} G_{i,k-i-1}(x) \\
&= \sum_{0 \le i < k+1} \frac{Q_i(x) Q_{k-i}(x)}{Q_{k+1}(x)} - \sum_{0 \le i < k} \frac{Q_i(x) Q_{k-i-1}(x)}{Q_k(x)} \\
&= \frac{Q'_{k+1}(x)}{Q_k(x)} - \frac{Q'_{k+2}(x)}{Q_{k+1}(x)}
\end{aligned}$$

by Remark 7.13, Lemma 7.21, and Proposition 7.23. □

Theorem 7.25. *For $k = 0, \ldots, n$, we have*

$$\#\left\{\boldsymbol{\gamma} \in \Gamma_{n,n} : \widehat{\rho}_2(\boldsymbol{\gamma}) \ge 1 - \frac{k}{n}\right\} = [x^n]\left(\frac{Q'_{k+1}(x)}{Q_k(x)} - \frac{Q'_{k+2}(x)}{Q_{k+1}(x)}\right).$$

Proof. We have

$$\begin{aligned}
\#\left\{\boldsymbol{\gamma} \in \Gamma_{n,n} : \widehat{\rho}_2(\boldsymbol{\gamma}) \ge 1 - \frac{k}{n}\right\} &= \#\{\boldsymbol{\gamma} \in \Gamma_{n,n} : \overline{\delta}_{\boldsymbol{\gamma}} - \underline{\delta}_{\boldsymbol{\gamma}} \le k\} \\
&= \#\{\boldsymbol{\gamma} \in \widetilde{\mathcal{G}}_k : |\boldsymbol{\gamma}| = n\} \\
&= [x^n] \widetilde{G}_k(x) \\
&= [x^n]\left(\frac{Q'_{k+1}(x)}{Q_k(x)} - \frac{Q'_{k+2}(x)}{Q_{k+1}(x)}\right)
\end{aligned}$$

by Theorem 7.5 and Proposition 7.24. □

Note that Theorem 7.25 is exactly Theorem 3.9.

Acknowledgement

This study was partially supported by JSPS KAKENHI Grant Numbers JP21K15762, JP15K04814, and JP20K03509.

References

[1] Anderson, T. W., On the distribution of the two-sample Cramér-von Mises criterion, Ann. Math. Stat. **33** (1962), 1148–1159.

[2] Berger, V. W. and Y. Zhou, Kolmogorov-Smirnov Test: Overview, Wiley StatsRef: Statistics Reference Online (2014).

[3] Flajolet, P. and R. Sedgewick, Analytic Combinatorics, Cambridge University Press (2009).

[4] Hand, D. J., Assessing the performance of classification methods, Int. Stat. Rev. **80** (2012), 400–414.

[5] Hsu, P. L. and H. Robbins, Complete convergence and the law of large numbers, Proc. Nat. Acad. Sci. U.S.A. **33** (1947), 25–31.

[6] Johno, H. and K. Nakamoto, Decision tree-based estimation of the overlap of two probability distributions, J. Math. Sci. Univ. Tokyo **30** (2023), 21–54.

[7] Mann, H. B. and D. R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, Ann. Math. Stat. **18** (1947), 50–60.

[8] Schröer, G. and D. Trenkler, Exact and randomization distributions of Kolmogorov-Smirnov tests two or three samples, Comput. Stat. Data Anal. **20** (1995), 185–202.

[9] Serfling, R. J., Approximation Theorems of Mathematical Statistics, John Wiley & Sons, Inc. (1980).

[10] Sloane, N. J. A. and The OEIS Foundation Inc., entry A115139 in the On-Line Encyclopedia of Integer Sequences, `https://oeis.org/A115139`.

[11] Snedecor, G. W. and W. G. Cochran, Statistical Methods, Iowa State University Press (1989).

[12] Welch, B. L., The generalization of 'Student's' problem when several different population variances are involved, Biometrika **34** (1947) 28–35.

Atsushi Komaba and Hisashi Johno
Department of Radiology
Faculty of Medicine
University of Yamanashi
Shimokato 1110, Chuo, Yamanashi 409-3898, Japan
Email: fiveseven.lambda@gmail.com
johnoh@yamanashi.ac.jp

Kazunori Nakamoto
Center for Medical Education and Sciences
Faculty of Medicine
University of Yamanashi
Shimokato 1110, Chuo, Yamanashi 409-3898, Japan
Email: nakamoto@yamanashi.ac.jp