

Functional central limit theorems for single-stage samplings designs

Supplementary Material

Hélène Boistard*, Hendrik Paul Lopuhaä† and Anne Ruiz-Gazen*

*Toulouse School of Economics**, *Delft University of Technology†*

APPENDIX A: PROOFS FOR RESULTS IN THE MAIN TEXT

Proof of Lemma 9.2. We will use the Cramér-Wold device. Note that any linear combination

$$(A.1) \quad a_1 \sqrt{n} \{ \mathbb{F}_N^{\text{HT}}(t_1) - \mathbb{F}_N(t_1) \} + \dots + a_k \sqrt{n} \{ \mathbb{F}_N^{\text{HT}}(t_k) - \mathbb{F}_N(t_k) \}$$

can be written as

$$(A.2) \quad \sqrt{n} \left\{ \frac{1}{N} \sum_{i=1}^N \frac{\xi_i}{\pi_i} V_{ik} - \frac{1}{N} \sum_{i=1}^N V_{ik} \right\},$$

where

$$(A.3) \quad V_{ik} = a_1 \mathbb{1}_{\{Y_i \leq t_1\}} + \dots + a_k \mathbb{1}_{\{Y_i \leq t_k\}} = \mathbf{a}_k^t \mathbf{Y}_{ik}$$

with $\mathbf{Y}_{ik}^t = (\mathbb{1}_{\{Y_i \leq t_1\}}, \dots, \mathbb{1}_{\{Y_i \leq t_k\}})$ and $\mathbf{a}_k^t = (a_1, \dots, a_k)$. For the corresponding design-based variance, we have

$$(A.4) \quad \begin{aligned} nS_N^2 &= \frac{n}{N^2} \sum_{i=1}^N \sum_{j=1}^N \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} V_{ik} V_{jk} \\ &= \mathbf{a}_k^t \left(\frac{n}{N^2} \sum_{i=1}^N \sum_{j=1}^N \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \mathbf{Y}_{ik} \mathbf{Y}_{jk}^t \right) \mathbf{a}_k \rightarrow \mathbf{a}_k^t \boldsymbol{\Sigma}_k^{\text{HT}} \mathbf{a}_k, \end{aligned}$$

ω -almost surely, according to (HT2), where $\boldsymbol{\Sigma}_k^{\text{HT}}$ can be obtained from (3.4). Together with (HT1), it follows that (A.1) converges in distribution to a mean zero normal random variable with variance $\mathbf{a}_k^t \boldsymbol{\Sigma}_k^{\text{HT}} \mathbf{a}_k$. We conclude that (A.1) converges in distribution to $a_1 N_1 + \dots + a_k N_k$, where (N_1, \dots, N_k) has a k -variate mean zero normal distribution with covariance matrix $\boldsymbol{\Sigma}_k^{\text{HT}}$. According to the Cramér-Wold device this proves the lemma. \square

Proof of Proposition 3.1. The proof is similar to that of Theorem 3.1. First consider the case of uniform Y_i 's with $F(t) = t$. We only have to verify the weak convergence of the finite dimensional projections of the process $\mathbb{X}_N = \sqrt{n}(\mathbb{F}_N^{\text{HT}} - \mathbb{F}_N)$. Consider (A.1) represented as in (A.2). From (HT1) and Lemma B.1(ii) in [5] we conclude that (A.1) converges in distribution to a mean zero normal random variable with variance

$$\begin{aligned}\sigma_{\text{HT}}^2 &= \mu_{\pi 1} \mathbb{E}_m [V_{1k}^2] + \mu_{\pi 2} (\mathbb{E}_m [V_{1k}])^2 \\ &= \mu_{\pi 1} \mathbf{a}_k^t \mathbb{E}_m [\mathbf{Y}_{1k} \mathbf{Y}_{1k}^t] \mathbf{a}_k + \mu_{\pi 2} \mathbf{a}_k^t (\mathbb{E}_m \mathbf{Y}_{1k}) (\mathbb{E}_m \mathbf{Y}_{1k})^t \mathbf{a}_k = \mathbf{a}_k^t \boldsymbol{\Sigma}_k \mathbf{a}_k,\end{aligned}$$

where $\boldsymbol{\Sigma}_k$ is the $k \times k$ -matrix with (q, r) -element equal to $\mu_{\pi 1} (t_q \wedge t_r) + \mu_{\pi 2} t_q t_r$. We conclude that (A.1) converges in distribution to $a_1 N_1 + \dots + a_k N_k$, where (N_1, \dots, N_k) has a k -variate mean zero normal distribution with covariance matrix $\boldsymbol{\Sigma}_k$. As in the proof of Lemma 9.2, by means of the Cramér-Wold device this establishes the limit distribution of $(\mathbb{X}_N(t_1), \dots, \mathbb{X}_N(t_k))$, which is the same that of the vector $(\mathbb{G}^{\text{HT}}(t_1), \dots, \mathbb{G}^{\text{HT}}(t_k))$, where \mathbb{G}^{HT} is a mean zero Gaussian process with covariance function $\mathbb{E}_{d,m} \mathbb{G}^{\text{HT}}(s) \mathbb{G}^{\text{HT}}(t) = \mu_{\pi 1} (s \wedge t) + \mu_{\pi 2} st$. From here on, the proof is completely the same as that of Theorem 3.1. \square

Proof of Lemma 9.3. We decompose as follows

$$\begin{aligned}\frac{1}{S_N} \left(\frac{1}{N} \sum_{i=1}^N \frac{\xi_i V_i}{\pi_i} - \mu_V \right) &= \frac{1}{S_N} \left(\frac{1}{N} \sum_{i=1}^N \frac{\xi_i V_i}{\pi_i} - \frac{1}{N} \sum_{i=1}^N V_i \right) \\ &\quad + \frac{1}{\sqrt{n} S_N} \times \frac{\sqrt{n}}{\sqrt{N}} \times \sqrt{N} \left(\frac{1}{N} \sum_{i=1}^N V_i - \mu_V \right).\end{aligned}$$

According to (HT3), the central limit theorem, Slutsky's theorem, and the fact that $n S_N^2 \rightarrow \sigma_{\text{HT}}^2 > 0$ in probability,

$$(A.5) \quad \frac{1}{\sqrt{n} S_N} \times \frac{\sqrt{n}}{\sqrt{N}} \times \sqrt{N} \left(\frac{1}{N} \sum_{i=1}^N V_i - \mu_V \right) \rightarrow N(0, \lambda \sigma_V^2 / \sigma_{\text{HT}}^2),$$

in distribution under \mathbb{P}_m , whereas, thanks to (HT1),

$$(A.6) \quad \frac{1}{S_N} \left(\frac{1}{N} \sum_{i=1}^N \frac{\xi_i V_i}{\pi_i} - \frac{1}{N} \sum_{i=1}^N V_i \right) \rightarrow N(0, 1), \quad \omega - \text{a.s.},$$

in distribution under \mathbb{P}_d . Since the latter limit distribution does not depend on ω , we can apply Theorem 5.1(iii) from [12]. It follows that

$$\frac{1}{S_N} \left(\frac{1}{N} \sum_{i=1}^N \frac{\xi_i V_i}{\pi_i} - \mu_V \right) \rightarrow N(0, 1 + \lambda \sigma_V^2 / \sigma_{\text{HT}}^2),$$

in distribution under $\mathbb{P}_{d,m}$. Together with $nS_N^2 \rightarrow \sigma_{HT}^2$ in probability, this implies that the random variable in (9.9) converges to a mean zero normal random variable with variance $\sigma_{HT}^2 + \lambda\sigma_V^2$. \square

Proof of Lemma 9.4. We will use the Cramér-Wold device. To this end, we determine the limit distribution of $a_1\mathbb{X}_N^F(t_1) + \dots + a_k\mathbb{X}_N^F(t_k)$, for $a_1, \dots, a_k \in \mathbb{R}$ fixed and $\mathbf{a}_k^t = (a_1, \dots, a_k) \neq (0, \dots, 0)$. As in the proof of Lemma 9.2, we consider

$$(A.7) \quad a_1\mathbb{X}_N^F(t_1) + \dots + a_k\mathbb{X}_N^F(t_k) = \sqrt{n} \left(\frac{1}{N} \sum_{i=1}^N \frac{\xi_i}{\pi_i} V_{ik} - \mu_k \right),$$

where V_{ik} is defined in (A.3). We want to apply Lemma 9.3. As in (A.4),

$$(A.8) \quad nS_N^2 \rightarrow \mathbf{a}_k^t \Sigma_k^{\text{HT}} \mathbf{a}_k, \quad \omega - \text{a.s.},$$

where $\mathbf{a}_k^t \Sigma_k^{\text{HT}} \mathbf{a}_k > 0$, thanks to (HT4). This means that, according to Lemma 9.3, the right hand side of (A.7) converges in distribution under $\mathbb{P}_{d,m}$ to a mean zero normal random variable with variance

$$\mathbf{a}_k^t \Sigma_k^{\text{HT}} \mathbf{a}_k + \lambda \left\{ \mathbb{E}_m[V_{1k}^2] - (\mathbb{E}_m[V_{1k}])^2 \right\} = \mathbf{a}_k^t \Sigma_{\text{HT}}^F \mathbf{a}_k,$$

where

$$(A.9) \quad \Sigma_{\text{HT}}^F = \Sigma_k^{\text{HT}} + \lambda \Sigma_F.$$

We conclude that (A.7) converges in distribution to $a_1 N_1 + \dots + a_k N_k$, where (N_1, \dots, N_k) has a mean zero k -variate normal distribution with covariance matrix Σ_{HT}^F . By the Cramér-Wold device, this proves the lemma. \square

Proof of Lemma 9.5. Without loss of generality we may assume $0 < F(t_1) < \dots < F(t_k) < 1$, since we can permute the rows and columns of \mathbf{M} without changing the determinant. For the entries of \mathbf{M} we can distinguish three situations:

1. if $1 \leq j < i \leq k$, then $M_{ij} = aF(t_j) - bF(t_i)F(t_j)$
2. if $1 \leq i = j \leq k$, then $M_{ij} = aF(t_i) - bF(t_i)^2$
3. if $1 \leq i < j \leq k$, then $M_{ij} = aF(t_i) - bF(t_i)F(t_j)$.

Now, for $2 \leq i \leq k$, multiply the i -th row by $F(t_1)/F(t_i)$. This changes the determinant with a factor $F(t_1)^{k-1}/F(t_2) \cdots F(t_k) > 0$, and as a result, all entries in column j , at positions $1 \leq i \leq j \leq k$, are the same: $aF(t_1) - bF(t_1)F(t_j)$. Hence, if we subtract row-2 from row-1, then row-3 from row-2, \dots , and then row- k from row- $(k-1)$, we get a new matrix \mathbf{M}' with a

right-upper triangle consisting of zero's and a main diagonal with elements $M'_{ii} = aF(t_1) - aF(t_1)F(t_i)/F(t_{i+1})$, if $1 \leq i \leq k-1$, and $M'_{kk} = aF(t_1) - bF(t_1)F(t_k)$. It follows that

$$\begin{aligned} \det(\mathbf{M}) &= \frac{F(t_2) \cdots F(t_k)}{F(t_1)^{k-1}} \det(\mathbf{M}') \\ &= a^{k-1} F(t_1) (F(t_2) - F(t_1)) \cdots (F(t_k) - F(t_{k-1})) (a - bF(t_k)) > 0, \end{aligned}$$

since $a > 0$, $0 < F(t_1) < \cdots < F(t_k) < 1$, and $a - bF(t_k) > a - b \geq 0$. \square

Proof of Lemma 9.6. The proof is similar to that of Lemma 9.4. We determine the limit distribution of (A.7). Note that without loss of generality we can assume that $0 \leq F(t_1) \leq \cdots \leq F(t_k) \leq 1$. In contrast with the proof of Lemma 9.4, we now have to distinguish between several cases.

We first consider the situation where all $F(t_i)$'s are distinct and such that $0 < F(t_i) < 1$. From (HT1) and Lemma B.1(ii) we conclude that

$$nS_N^2 \rightarrow \sigma_{\text{HT}}^2 = \mu_{\pi 1} \mathbb{E}_m[V_{1k}^2] + \mu_{\pi 2} (\mathbb{E}_m[V_{1k}])^2 = \mathbf{a}_k^t \boldsymbol{\Sigma}_k \mathbf{a}_k,$$

where

$$(A.10) \quad \boldsymbol{\Sigma}_k = \left(\mu_{\pi 1} F(t_q \wedge t_r) + \mu_{\pi 2} F(t_q) F(t_r) \right)_{q,r=1}^k.$$

First note that

$$\mu_{\pi 1} + \mu_{\pi 2} = \lim_{N \rightarrow \infty} \frac{n}{N^2} \sum_{i=1}^N \sum_{j=1}^N \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} = \lim_{N \rightarrow \infty} \frac{n}{N^2} \text{Var} \left(\sum_{i=1}^N \frac{\xi_i}{\pi_i} \right) \geq 0.$$

Therefore, together with condition (i) we can apply Lemma 9.5 with $a = \mu_{\pi 1}$ and $b = -\mu_{\pi 2}$. It follows that $\boldsymbol{\Sigma}_k$ is positive definite, so that $\sigma_{\text{HT}}^2 > 0$. This means that, according to Lemma 9.3, the right hand side of (A.7) converges in distribution under $\mathbb{P}_{d,m}$ to a mean zero normal random variable with variance $(\mu_{\pi 1} + \lambda) \mathbb{E}_m[V_{1k}^2] + (\mu_{\pi 2} - \lambda) (\mathbb{E}_m[V_{1k}])^2 = \mathbf{a}_k^t \boldsymbol{\Sigma}_{\text{HT}}^F \mathbf{a}_k$, where

$$(A.11) \quad \boldsymbol{\Sigma}_{\text{HT}}^F = \left((\mu_{\pi 1} + \lambda) F(t_q \wedge t_r) + (\mu_{\pi 2} - \lambda) F(t_q) F(t_r) \right)_{q,r=1}^k.$$

We conclude that (A.7) converges in distribution to $a_1 N_1 + \cdots + a_k N_k$, where (N_1, \dots, N_k) has a mean zero k -variate normal distribution with covariance matrix $\boldsymbol{\Sigma}_{\text{HT}}^F$. By means of the Cramér-Wold device, this proves the lemma for the case that $0 < F(t_1) < \cdots < F(t_k) < 1$.

The case that the $F(t_i)$'s are not all distinct, but still satisfy $0 < F(t_i) < 1$, can be reduced to the case where all $F(t_i)$'s are distinct. This can be

seen as follows. For simplicity, suppose $F(t_1) = \dots = F(t_m) = F(t_0)$, with $0 < F(t_0) < F(t_{m+1}) < \dots < F(t_k) < 1$. Then we can write (A.7) as

$$(A.12) \quad a_0 \mathbb{X}_N^F(t_0) + a_{m+1} \mathbb{X}_N^F(t_{m+1}) + \dots + a_k \mathbb{X}_N^F(t_k),$$

where $a_0 = a_1 + \dots + a_m$. As before, with (HT4) and Lemma 9.5, it follows from Lemma 9.3 that (A.12) converges in distribution to a mean zero normal random variable with variance $\mathbf{a}_0^t \boldsymbol{\Sigma}_0^F \mathbf{a}_0$, where $\mathbf{a}_0 = (a_0, a_{m+1}, \dots, a_k)^t$ and

$$\boldsymbol{\Sigma}_0^F = \gamma_{\pi 1} \mathbb{E}_m[\mathbf{Y}_0 \mathbf{Y}_0^t] + (\gamma_{\pi 2} - \lambda) (\mathbb{E}_m[\mathbf{Y}_0]) (\mathbb{E}_m[\mathbf{Y}_0])^t,$$

with $\mathbf{Y}_0 = (\mathbb{1}_{\{Y_i \leq t_0\}}, \mathbb{1}_{\{Y_i \leq t_{m+1}\}}, \dots, \mathbb{1}_{\{Y_i \leq t_k\}})^t$. However, note that

$$\begin{aligned} \mathbf{a}_0^t \mathbf{Y}_0 &= (a_1 + \dots + a_m) \mathbb{1}_{\{Y_i \leq t_0\}} + a_{m+1} \mathbb{1}_{\{Y_i \leq t_{m+1}\}} + \dots + a_k \mathbb{1}_{\{Y_i \leq t_k\}}, \\ &= a_1 \mathbb{1}_{\{Y_i \leq t_1\}} + \dots + a_k \mathbb{1}_{\{Y_i \leq t_k\}} = \mathbf{a}_k^t \mathbf{Y}_{1k}, \end{aligned}$$

where $\mathbf{a}_k = (a_1, \dots, a_k)^t$ and $\mathbf{Y}_{1k} = (\mathbb{1}_{\{Y_i \leq t_1\}}, \dots, \mathbb{1}_{\{Y_i \leq t_k\}})^t$, as before. This means that $\mathbf{a}_0^t \boldsymbol{\Sigma}_0^F \mathbf{a}_0 = \mathbf{a}_k^t \boldsymbol{\Sigma}_{\text{HT}}^F \mathbf{a}_k$, with $\boldsymbol{\Sigma}_{\text{HT}}^F$ from (A.9). It follows that (A.7) converges in distribution to $a_1 N_1 + \dots + a_k N_k$, where (N_1, \dots, N_k) has a mean zero k -variate normal distribution with covariance matrix $\boldsymbol{\Sigma}_{\text{HT}}^F$. By means of the Cramér-Wold device, this proves the lemma for the case $F(t_1) = \dots = F(t_m) = F(t_0) < F(t_{m+1}) < \dots < F(t_k) < 1$. The argument is the same for other cases with multiple $F(t_i) \in (0, 1)$ being equal to each other.

Next, consider the case $F(t_1) = 0$. In this case, $\mathbb{1}_{\{Y_i \leq t_1\}} = 0$ with probability one. This means that the summation on the left hand side of (A.7) reduces to $a_2 \mathbb{X}_N^F(t_2) + \dots + a_k \mathbb{X}_N^F(t_k)$ and

$$(A.13) \quad \boldsymbol{\Sigma}_{\text{HT}} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & \boldsymbol{\Sigma}_{\text{HT},k-1} & & \\ 0 & & & \end{pmatrix},$$

where $\boldsymbol{\Sigma}_{\text{HT},k-1}$ is the matrix in (A.10) based on $0 < F(t_2) < \dots < F(t_k) < 1$. When $\mathbf{a}_{k-1}^t = (a_2, \dots, a_k) \neq (0, \dots, 0)$, then

$$\sigma_{\text{HT}}^2 = \mathbf{a}_k^t \boldsymbol{\Sigma}_{\text{HT}}^F \mathbf{a}_k = \mathbf{a}_{k-1}^t \boldsymbol{\Sigma}_{\text{HT},k-1} \mathbf{a}_{k-1} > 0,$$

because $\boldsymbol{\Sigma}_{\text{HT},k-1}$ is positive definite, due to (HT4) and Lemma 9.5. This allows application of Lemma 9.3 to (A.7). As in the previous cases, we conclude that (A.7) converges in distribution to $a_1 N_1 + \dots + a_k N_k$, where (N_1, \dots, N_k) has a mean zero k -variate normal distribution with covariance matrix $\boldsymbol{\Sigma}_{\text{HT}}^F$ given by (A.9). When $\mathbf{a}_k^t = (a_1, 0, \dots, 0)$, with $a_1 \neq 0$, then

both (A.7) and $a_1N_1 + \dots + a_kN_k$ are equal to zero. According to the Cramér-Wold device, this proves the lemma for the case $F(t_k) = 0$.

It remains to consider the case $F(t_k) = 1$. In this case, the (k, k) -th element of the matrix Σ_{HT} in (A.10) is equal to $\mu_{\pi_1} + \mu_{\pi_2}$. We distinguish between $\mu_{\pi_1} + \mu_{\pi_2} = 0$ and $\mu_{\pi_1} + \mu_{\pi_2} > 0$. In the latter case, from the proof of Lemma 9.5 we find that Σ_{HT} has determinant

$$\mu_{\pi_1}^{k-1} F(t_1) \prod_{i=2}^k (F(t_i) - F(t_{i-1})) (\mu_{\pi_1} + \mu_{\pi_2}) > 0,$$

using (HT4) and $0 < F(t_1) < \dots < F(t_{k-1}) < F(t_k) = 1$. This allows application of Lemma 9.3 to (A.7). As before, we conclude that (A.7) converges in distribution to $a_1N_1 + \dots + a_kN_k$, where (N_1, \dots, N_k) has a k -variate mean zero normal distribution with covariance matrix Σ_{HT}^F from (A.9). According to the Cramér-Wold device, this proves the lemma for the case $F(t_k) = 1$ and $\mu_{\pi_1} + \mu_{\pi_2} > 0$.

Next, consider the case $F(t_k) = 1$ and $\mu_{\pi_1} + \mu_{\pi_2} = 0$. This means

$$(A.14) \quad \Sigma_{\text{HT}} = \begin{pmatrix} & & & 0 \\ & \Sigma_{\text{HT},k-1} & & \vdots \\ & & & 0 \\ 0 & \dots & 0 & 0 \end{pmatrix},$$

where $\Sigma_{\text{HT},k-1}$ is the matrix in (A.10) corresponding to $0 < F(t_1) < \dots < F(t_{k-1}) < 1$. When $\mathbf{a}_{k-1}^t = (a_1, \dots, a_{k-1}) \neq (0, \dots, 0)$, then

$$\sigma_{\text{HT}}^2 = \mathbf{a}_k^t \Sigma_{\text{HT}} \mathbf{a}_k = \mathbf{a}_{k-1}^t \Sigma_{\text{HT},k-1} \mathbf{a}_{k-1} > 0,$$

because $\Sigma_{\text{HT},k-1}$ is positive definite, due to (HT4) and Lemma 9.5. This allows application of Lemma 9.3 to (A.7). As in the previous cases, we conclude that (A.7) converges in distribution to $a_1N_1 + \dots + a_kN_k$, where (N_1, \dots, N_k) has a k -variate mean zero normal distribution with covariance matrix Σ_{HT}^F given by (A.9). When $\mathbf{a}_k^t = (0, \dots, 0, a_k)$, with $a_k \neq 0$, then $a_1N_1 + \dots + a_kN_k = 0$ and

$$a_1 \mathbb{X}_N^F(t_1) + \dots + a_k \mathbb{X}_N^F(t_k) = a_k \sqrt{n} \left(\frac{1}{N} \sum_{i=1}^N \frac{\xi_i}{\pi_i} - 1 \right).$$

converges to zero in probability. The latter follows from the fact that according to (HT1) and Lemma B.1, we have that

$$(A.15) \quad \sqrt{n} \left(\frac{1}{N} \sum_{i=1}^N \frac{\xi_i}{\pi_i} - 1 \right) \rightarrow N(0, \mu_{\pi_1} + \mu_{\pi_2}),$$

in distribution under $\mathbb{P}_{d,m}$. According to the Cramér-Wold device, this proves the lemma for the case $F(t_k) = 1$ and $\mu_{\pi_1} + \mu_{\pi_2} = 0$. Finally, the argument for the case that $F(t_1) = 0$ and $F(t_k) = 1$ simultaneously, either with or without repeated among the $F(t_i)$'s, is completely similar. This finishes the proof. \square

Proof of Proposition 3.2. The proof is similar to that of Theorem 3.2. Tightness is obtained in the same way and the convergence of finite dimensional projections is provided by Lemma 9.6. The theorem now follows from Theorem 13.5 in [3] for the case that the Y_i 's are uniformly distributed on $[0, 1]$. Next, this is extended to Y_i 's with a general c.d.f. F in the same way as in the proof of Theorem 3.1. \square

Proof of Proposition 4.1. The proof is similar to that of Theorem 4.2. We find that the limit behavior of $\sqrt{n}(\mathbb{F}_N^{\text{HJ}} - \mathbb{F}_N)$ is the same as that of the process \mathbb{Y}_N defined in (4.3). When we first consider the case of uniform Y_i 's with $F(t) = t$, tightness of the process \mathbb{Y}_N follows in the same way as in the proof of Theorem 4.2. It remains to establish weak convergence of the finite dimensional projections (A.18). This can be done in the same way as in the proof of Proposition 3.1, but this time with

$$V_{ik} = a_1 (\mathbb{1}_{\{Y_i \leq t_1\}} - t_1) + \cdots + a_k (\mathbb{1}_{\{Y_i \leq t_k\}} - t_k).$$

From (HT1) and Lemma B.1(i) we conclude that (A.19) converges in distribution to a mean zero normal random variable with variance

$$\sigma_{\text{HT}}^2 = \mu_{\pi_1} \mathbb{E}_m [V_{1k}^2] = \mathbf{a}_k^t \tilde{\Sigma}_k \mathbf{a}_k,$$

where $\tilde{\Sigma}_k$ is the $k \times k$ -matrix with (q, r) -element equal to $\mu_{\pi_1} (t_q \wedge t_r - t_q t_r)$. We conclude that (A.19) converges in distribution to $a_1 N_1 + \cdots + a_k N_k$, where (N_1, \dots, N_k) has a k -variate mean zero normal distribution with covariance matrix $\tilde{\Sigma}_k$. By means of the Cramér-Wold device this establishes the limit distribution of (A.18), which is the same as that of the vector $(\mathbb{G}^{\text{HJ}}(t_1), \dots, \mathbb{G}^{\text{HJ}}(t_k))$, where \mathbb{G}^{HJ} is a mean zero Gaussian process with covariance function

$$\mathbb{E}_{d,m} \mathbb{G}^{\text{HJ}}(s) \mathbb{G}^{\text{HJ}}(t) = \mu_{\pi_1} (s \wedge t - st).$$

From here on, the proof is completely the same as that of Theorem 4.2. \square

Remainder of the proof of Theorem 4.1. It remains to prove weak convergence of the finite dimensional projections

$$(A.16) \quad (\mathbb{G}_N^\pi(t_1), \dots, \mathbb{G}_N^\pi(t_k)).$$

To this end we apply the Cramér-Wold device and consider linear combinations

$$(A.17) \quad a_1 \mathbb{G}_N^\pi(t_1) + \cdots + a_k \mathbb{G}_N^\pi(t_k) = \frac{\sqrt{n}}{N} \sum_{i=1}^N \frac{\xi_i}{\pi_i} V_{ik}.$$

Convergence of (A.17), is obtained completely similar to that of (A.7) in Lemma 9.4, but this time with

$$V_{ik} = a_1 (\mathbb{1}_{\{Y_i \leq t_1\}} - t_1) + \cdots + a_k (\mathbb{1}_{\{Y_i \leq t_k\}} - t_k),$$

and $\mu_k = 0$. Using the fact that (HJ4) allows the use of Lemma 9.3, one can deduce that (A.17) converges in distribution under $\mathbb{P}_{d,m}$ to $a_1 N_1 + \cdots + a_k N_k$, where (N_1, \dots, N_k) has a k -variate normal distribution with covariance matrix $\Sigma^\pi = \Sigma_k^{\text{HJ}} + \lambda \Sigma_F$, where Σ_k^{HJ} and Σ_F are given in (4.5) and Lemma 9.4, respectively. By means of the Cramér-Wold device, this proves that (A.16) converges in distribution under $\mathbb{P}_{d,m}$ to a mean zero k -variate normal random vector with covariance matrix Σ^π . This distribution is the same as that of $(\mathbb{G}^\pi(t_1), \dots, \mathbb{G}^\pi(t_k))$, where \mathbb{G}^π is a mean zero Gaussian process with covariance function

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \mathbb{E}_m \left[n \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} (\mathbb{1}_{\{Y_i \leq s\}} - s) (\mathbb{1}_{\{Y_i \leq t\}} - t) \right] \\ + \lambda (s \wedge t - st), \quad s, t \in \mathbb{R}. \end{aligned}$$

Since \mathbb{G}^π is continuous at 1, the theorem then follows from Theorem 13.5 in [3] for the case of uniform Y_i 's. Extension to Y_i 's with a general c.d.f. F is completely similar to the proof of Theorem 3.1.

Remainder of the proof of Theorem 4.2. It remains to prove weak convergence of the finite dimensional projections

$$(A.18) \quad (\mathbb{Y}_N(t_1), \dots, \mathbb{Y}_N(t_k)).$$

As before, we apply the Cramér-Wold device and consider

$$(A.19) \quad a_1 \mathbb{Y}_N(t_1) + \cdots + a_k \mathbb{Y}_N(t_k) = \sqrt{n} \left\{ \frac{1}{N} \sum_{i=1}^N \frac{\xi_i}{\pi_i} V_{ik} - \frac{1}{N} \sum_{i=1}^N V_{ik} \right\},$$

with

$$V_{ik} = a_1 (\mathbb{1}_{\{Y_i \leq t_1\}} - t_1) + \cdots + a_k (\mathbb{1}_{\{Y_i \leq t_k\}} - t_k).$$

Convergence of (A.19) is obtained completely similar to that of (A.2) in the proof of Lemma 9.2. From (HT1) and (HJ2), it follows that (A.19) converges in distribution under $\mathbb{P}_{d,m}$ to $a_1 N_1 + \dots + a_k N_k$, where (N_1, \dots, N_k) has a k -variate normal distribution with covariance matrix Σ_k^{HJ} given in (4.5). By means of the Cramér-Wold device, this proves that (A.18) converges in distribution under $\mathbb{P}_{d,m}$ to a mean zero k -variate normal random vector with covariance matrix Σ_k^{HJ} . This distribution is the same as that of $(\mathbb{G}^{\text{HJ}}(t_1), \dots, \mathbb{G}^{\text{HJ}}(t_k))$, where \mathbb{G}^{HJ} is a mean zero Gaussian process with covariance function

$$\lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \mathbb{E}_m \left[n \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} (\mathbf{1}_{\{Y_i \leq s\}} - s) (\mathbf{1}_{\{Y_j \leq t\}} - t) \right],$$

for $s, t \in \mathbb{R}$. As before, the theorem now follows from Theorem 13.5 in [3] for the case of uniform Y_i 's, and is then extended to Y_i 's with a general c.d.f. F .

□

Proof of Theorem 4.3. The theorem follows directly from relation (4.7) and Theorem 4.1. □

Proof of Proposition 4.2. From relation (4.7) and Theorem 4.1 we know that the limit behavior of $\sqrt{n}(\mathbb{F}_N^{\text{HJ}} - F)$ is the same as that of \mathbb{G}_N^π . Tightness of \mathbb{G}_N^π has been obtained in the proof of Theorem 4.1. It remains to establish weak convergence of (A.16). This can be done in the same way as in the proof of Lemma 9.6, but this time with

$$V_{ik} = a_1 (\mathbf{1}_{\{Y_i \leq t_1\}} - F(t_1)) + \dots + a_k (\mathbf{1}_{\{Y_i \leq t_k\}} - F(t_k))$$

and $\mu_k = 0$. When $0 < F(t_1) < \dots < F(t_k) < 1$, from (HT1) and Lemma B.1 we find that $nS_N^2 \rightarrow \mu_{\pi 1} \mathbb{E}_m[V_{1k}^2] = \mathbf{a}_k^t \Sigma_k \mathbf{a}_k$, where

$$(A.20) \quad \Sigma_k = \mu_{\pi 1} \left(F(t_q \wedge t_r) - F(t_q)F(t_r) \right)_{q,r=1}^k.$$

From condition (i) of Proposition 3.2 and Lemma 9.5, it follows that Σ_k is positive definite, so that $\mathbf{a}_k^t \Sigma_k \mathbf{a}_k > 0$. Hence, according to Lemma 9.3, the right hand side of (A.17) converges in distribution under $\mathbb{P}_{d,m}$ to a mean zero normal random variable with variance $(\mu_{\pi 1} + \lambda) \mathbb{E}_m[V_{1k}^2] = \mathbf{a}_k^t \Sigma_{\text{HJ}}^F \mathbf{a}_k$, where

$$(A.21) \quad \Sigma_{\text{HJ}}^F = \left((\mu_{\pi 1} + \lambda) F(t_q \wedge t_r) \right)_{q,r=1}^k.$$

We conclude that the right hand side of (A.17) converges in distribution to $a_1 N_1 + \dots + a_k N_k$, where (N_1, \dots, N_k) has a mean zero k -variate normal

distribution with covariance matrix Σ_{HJ}^F . By means of the Cramér-Wold device, this proves weak convergence of $(\mathbb{G}_N^\pi(t_1), \dots, \mathbb{G}_N^\pi(t_k))$ for the case that $0 < F(t_1) < \dots < F(t_k) < 1$. As in the proof of Lemma 9.6, the case where the $F(t_i)$'s are not all distinct, but satisfy $0 < F(t_i) < 1$, the case $F(t_1) = 0$, and the case $F(t_k) = 1$, can be reduced to the previous case. From here on, the proof is completely the same as that of Theorem 4.1. \square

Proof of Proposition 5.1. The proposition only needs to be established for the rejective sampling design, as it can be extended to high entropy designs by means of Theorem 5 in [1]. Since the rejective sampling design can be represented as a Poisson sampling design conditionally on the sample size being equal to n , the proof is along the lines of the arguments used in the proof of Theorem 3.2 in [2]. It applies results from [11] on a central limit theorem for sums of functions of independent random variables ξ_1, \dots, ξ_N , conditional on $\xi_1 + \dots + \xi_N = n$. Details are provided in Supplement B in [5]. \square

Proof of Corollary 5.2. As in the proof of Corollary 5.1, we first prove the results for rejective sampling and then extend them to high entropy designs. Completely similar to the proof of Corollary 5.1, conditions (A2)-(A4) imply (C2)-(C4). Furthermore, condition (ii) of Proposition 3.1 is obtained in the same way as in the proof of Corollary 5.1, with $\mu_{\pi 2} = -\alpha$, from conditions (A2)-(A3) and (A5). This proves parts (i)-(iv). \square

Proof of Corollary 6.1. The mapping $\phi : \mathbb{D}_\phi \subset D(\mathbb{R}) \mapsto \mathbb{R}$ is Hadamard-differentiable at F tangentially to the set \mathbb{D}_0 consisting of functions $h \in D(\mathbb{R})$ that are continuous at $F^{-1}(\alpha)$. According to Theorem 3.2, the sequence $\sqrt{n}(\mathbb{F}_N^{\text{HT}} - F)$ converges weakly to a mean zero Gaussian process \mathbb{G}_F^{HT} with covariance structure

$$(A.22) \quad \mathbb{E}_{d,m} \mathbb{G}_F^{\text{HT}}(s) \mathbb{G}_F^{\text{HT}}(t) = (\mu_{\pi 1} + \lambda)F(s \wedge t) + (\mu_{\pi 2} - \lambda)F(s)F(t),$$

for $s, t \in \mathbb{R}$. It then follows from Theorem 3.9.4 in [13], that the random variable $\sqrt{n}(\phi(\mathbb{F}_N^{\text{HT}}) - \phi(F))$ converges weakly to

$$-\beta \frac{f(\beta F^{-1}(\alpha))}{f(F^{-1}(\alpha))} \mathbb{G}_F^{\text{HT}}(F^{-1}(\alpha)) + \mathbb{G}_F^{\text{HT}}(\beta F^{-1}(\alpha)),$$

which has a normal distribution with mean zero and variance

$$\begin{aligned} \sigma_{\text{HT}, \alpha, \beta}^2 &= \beta^2 \frac{f(\beta F^{-1}(\alpha))^2}{f(F^{-1}(\alpha))^2} \mathbb{E} [\mathbb{G}_F^{\text{HT}}(F^{-1}(\alpha))^2] \\ &\quad + \mathbb{E} [\mathbb{G}_F^{\text{HT}}(\beta F^{-1}(\alpha))^2] \\ &\quad - 2\beta \frac{f(\beta F^{-1}(\alpha))}{f(F^{-1}(\alpha))} \mathbb{E} [\mathbb{G}_F^{\text{HT}}(F^{-1}(\alpha)) \mathbb{G}_F^{\text{HT}}(\beta F^{-1}(\alpha))]. \end{aligned}$$

The precise expression can then be derived from (A.22), which proves part one. For part two, write

$$\sqrt{n}(\phi(\mathbb{F}_N^{\text{HT}}) - \phi(\mathbb{F}_N)) = \sqrt{n}(\phi(\mathbb{F}_N^{\text{HT}}) - \phi(F)) + \frac{\sqrt{n}}{\sqrt{N}}\sqrt{N}(\phi(\mathbb{F}_N) - \phi(F)).$$

The process $\sqrt{N}(\mathbb{F}_N - F)$ converges weakly to a mean zero Gaussian process \mathbb{G}_F . Then, Hadamard-differentiability of ϕ together with Theorem 3.9.4 in [13] yields that the sequence $\sqrt{N}(\phi(\mathbb{F}_N) - \phi(F))$ converges weakly to $\phi'_F(\mathbb{G}_F)$. As $n/N \rightarrow 0$, the theorem follows from part one. \square

Proof of Corollary 6.2. The proof is completely the same as that of Corollary 6.1, with the only difference that the covariance structure of the limiting process $\sqrt{n}(\phi(\mathbb{F}_N^{\text{HJ}}) - \phi(F))$ is now given in Theorem 4.3. \square

APPENDIX B: ADDITIONAL TECHNICALITIES

Comment about (C1) on page 6. Condition A3 in [7] requires that

$$(B.1) \quad \lim_{N, n \rightarrow \infty} \mathbb{E}[\pi_i(1 - \pi_i)] = d > 0,$$

where $0 < d \leq 1/4$. The parabola $x \mapsto x(1 - x) - d$ is strictly positive for

$$0 < \frac{1 - \sqrt{1 - 4d}}{2} < x < \frac{1 + \sqrt{1 - 4d}}{2} < 1.$$

According to condition A4 in [7], it holds that $n/N \rightarrow \lambda > 0$. Suppose that the lower bound in (C1) does not hold, so that $N\pi_i/n$ can be arbitrarily small, say

$$\frac{N\pi_i}{n} < \frac{1 - \sqrt{1 - 4d}}{4\lambda}.$$

In that case

$$\lim_{N \rightarrow \infty} \pi_i = \lim_{N \rightarrow \infty} \frac{n}{N} \cdot \frac{N\pi_i}{n} < \lambda \cdot \frac{1 - \sqrt{1 - 4d}}{4\lambda} = \frac{1 - \sqrt{1 - 4d}}{4},$$

which lies left of the smallest zero of the parabola $x(1 - x) - d$. As a consequence

$$\lim_{N, n \rightarrow \infty} \mathbb{E}[\pi_i(1 - \pi_i)] < d,$$

which is in contradiction with (B.1). \square

LEMMA B.1. *Let S_N^2 be defined by (3.2), where V_1, V_2, \dots is a sequence of i.i.d. random variables on $(\Omega, \mathfrak{F}, \mathbb{P}_m)$ with $\mathbb{E}_m[V_1^4] < \infty$. Suppose that n and π_i, π_{ij} , for $i, j = 1, 2, \dots, N$ are deterministic and let $\mathbb{V}_m(S_N^2)$ denote the variance of S_N^2 . If (C1)-(C2) hold, then $n^2\mathbb{V}_m[S_N^2] = O(1/N)$. Then,*

(i) if $\mathbb{E}_m[V_1] = 0$ and condition (i) in Proposition 3.1 holds,

$$nS_N^2 \rightarrow \sigma_{\text{HT}}^2 = \mu_{\pi_1} \mathbb{E}_m[V_1^2], \quad \text{in } \mathbb{P}_m\text{-probability.}$$

(ii) if $\mathbb{E}_m[V_1] \neq 0$ and conditions (i)-(ii) in Proposition 3.1 hold,

$$nS_N^2 \rightarrow \sigma_{\text{HT}}^2 = \mu_{\pi_1} \mathbb{E}_m[V_1^2] + \mu_{\pi_2} (\mathbb{E}_m[V_1])^2, \quad \text{in } \mathbb{P}_m\text{-probability.}$$

PROOF. For any $\epsilon > 0$, by Markov inequality we have

$$(B.2) \quad \mathbb{P}_m \{ |nS_N^2 - \mathbb{E}_m[nS_N^2]| > \epsilon \} < \frac{n^2 \mathbb{V}_m[S_N^2]}{\epsilon^2},$$

where \mathbb{V}_m denotes the variance of S_N^2 under the super-population model. In order to compute $\mathbb{V}_m[S_N^2]$, we first have

$$(B.3) \quad \begin{aligned} \mathbb{E}_m[S_N^2] &= \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \mathbb{E}_m(V_i V_j) \\ &= \frac{\mathbb{E}_m[V_1^2]}{N^2} \sum_{i=1}^N \frac{1 - \pi_i}{\pi_i} + \frac{(\mathbb{E}_m[V_1])^2}{N^2} \sum_{i \neq j} \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j}. \end{aligned}$$

From this, tedious but straightforward calculus leads to the expression for $(\mathbb{E}_m[S_N^2])^2$ and $\mathbb{E}_m[S_N^4]$. One finds

$$N^4 (\mathbb{E}_m[S_N^2])^2 = a_1 (\mathbb{E}_m[V_1])^4 + a_2 \mathbb{E}_m[V_1^2] (\mathbb{E}_m[V_1])^2 + a_3 (\mathbb{E}_m[V_1^2])^2,$$

where, according to (C1)-(C2):

$$\begin{aligned}
a_1 &= \sum_{(i,j,k,l) \in D_{4,N}} \sum \sum \sum \sum \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_k \pi_l} \\
&\quad + 4 \sum_{(i,j,l) \in D_{3,N}} \sum \sum \sum \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \frac{\pi_{il} - \pi_i \pi_l}{\pi_i \pi_l} + 2 \sum_{(i,j) \in D_{2,N}} \sum \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \right)^2 \\
&= \sum_{(i,j,k,l) \in D_{4,N}} \sum \sum \sum \sum \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_k \pi_l} + O(N^3/n^2) + O(N^2/n^2) \\
a_2 &= 2 \sum_{(i,k,l) \in D_{3,N}} \sum \sum \sum \frac{1 - \pi_i}{\pi_i} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_k \pi_l} + 4 \sum_{(i,k) \in D_{2,N}} \sum \sum \frac{1 - \pi_i}{\pi_i} \frac{\pi_{ik} - \pi_i \pi_k}{\pi_i \pi_k} \\
&= 2 \sum_{(i,k,l) \in D_{3,N}} \sum \sum \sum \frac{1 - \pi_i}{\pi_i} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_k \pi_l} + O(N^3/n^2) \\
a_3 &= \sum_{(i,j) \in D_{2,N}} \sum \frac{1 - \pi_i}{\pi_i} \frac{1 - \pi_j}{\pi_j} + \sum_{i=1}^N \left(\frac{1 - \pi_i}{\pi_i} \right)^2 \\
&= \sum_{(i,j) \in D_{2,N}} \sum \frac{1 - \pi_i}{\pi_i} \frac{1 - \pi_j}{\pi_j} + O(N^3/n^2).
\end{aligned}$$

Furthermore,

$$\begin{aligned}
N^4 \mathbb{E}_m [S_N^4] &= b_1 (\mathbb{E}_m [V_1])^4 + b_2 \mathbb{E}_m [V_1^2] (\mathbb{E}_m [V_1])^2 \\
&\quad + b_3 (\mathbb{E}_m [V_1^2])^2 + b_4 \mathbb{E}_m [V_1] \mathbb{E}_m [V_1^3]
\end{aligned}$$

where

$$\begin{aligned}
b_1 &= \sum_{(i,j,k,l) \in D_{4,N}} \sum \sum \sum \sum \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_k \pi_l} + \sum_{i=1}^N \left(\frac{1 - \pi_i}{\pi_i} \right)^2 \\
&= \sum_{(i,j,k,l) \in D_{4,N}} \sum \sum \sum \sum \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_k \pi_l} + O(N^3/n^2) \\
b_2 &= 2 \sum_{(i,k,l) \in D_{3,N}} \sum \sum \sum \frac{1 - \pi_i}{\pi_i} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_k \pi_l} + 4 \sum_{(i,j,l) \in D_{3,N}} \sum \sum \sum \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \frac{\pi_{il} - \pi_i \pi_l}{\pi_i \pi_l} \\
&= 2 \sum_{(i,k,l) \in D_{3,N}} \sum \sum \sum \frac{1 - \pi_i}{\pi_i} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_k \pi_l} + O(N^3/n^2)
\end{aligned}$$

$$\begin{aligned}
b_3 &= \sum_{(i,k) \in D_{2,N}} \sum_{(i,j) \in D_{2,N}} \frac{1-\pi_i}{\pi_i} \frac{1-\pi_k}{\pi_k} + 2 \sum_{(i,j) \in D_{2,N}} \sum_{(i,j) \in D_{2,N}} \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \right)^2 \\
&= \sum_{(i,k) \in D_{2,N}} \sum_{(i,k) \in D_{2,N}} \frac{1-\pi_i}{\pi_i} \frac{1-\pi_k}{\pi_k} + O(N^2/n^2) \\
b_4 &= 4 \sum_{(i,j) \in D_{2,N}} \sum_{(i,j) \in D_{2,N}} \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \frac{1-\pi_j}{\pi_j} = O(N^3/n^2).
\end{aligned}$$

The variance expression for S_N^2 is deduced easily from the previous computations. From the expression derived in [5], we find that $a_i - b_i = O(N^3/n^2)$, for $i = 1, 2, 3$, and $b_4 = O(N^3/n^2)$, so that

$$(B.4) \quad n^2 \mathbb{V}_m[S_N^2] = n^2 \mathbb{E}_m[S_N^4] - n^2 (\mathbb{E}_m[S_N^2])^2 = O(1/N).$$

From (B.2) we conclude that $nS_N^2 - \mathbb{E}_m[nS_N^2]$ tends to zero in \mathbb{P}_m -probability. As a consequence, statements (i) and (ii) follow from (B.3). \square

Proof of Lemma B.1 under (C2).* We used (C2) to bound remainder terms in the coefficients a_i and b_i , but this can also be achieved with (C2*). For the second term in a_1 we get

$$\begin{aligned}
\left| \sum_{(i,j,l) \in D_{3,N}} \sum_{(i,j,l) \in D_{3,N}} \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \frac{\pi_{il} - \pi_i \pi_l}{\pi_i \pi_l} \right| &\leq \sum_{(i,j) \in D_{2,N}} \sum_{(i,j) \in D_{2,N}} \left| \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \right| \cdot \sum_{l \neq i,j} \left| \frac{\pi_{il} - \pi_i \pi_l}{\pi_i \pi_l} \right| \\
&= N \cdot O\left(\frac{N}{n}\right) \cdot O\left(\frac{N}{n}\right) = O\left(\frac{N^3}{n^2}\right),
\end{aligned}$$

by means of (C2*). For the third term in a_1 , we have

$$\begin{aligned}
\sum_{(i,j) \in D_{2,N}} \sum_{(i,j) \in D_{2,N}} \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \right)^2 &\leq \sum_{(i,j) \in D_{2,N}} \sum_{(i,j) \in D_{2,N}} \frac{|\pi_{ij} - \pi_i \pi_j|}{\pi_i \pi_j} \cdot \frac{\pi_{ij}}{\pi_i \pi_j} \\
&= N \cdot O\left(\frac{N}{n}\right) \cdot O\left(\frac{N}{n}\right) = O\left(\frac{N^3}{n^2}\right),
\end{aligned}$$

by means of (C2*) and (C1) and the fact that $\pi_{ij} \leq \pi_i$. For the second term in a_2 we have

$$\begin{aligned}
\left| \sum_{(i,k) \in D_{2,N}} \sum_{(i,k) \in D_{2,N}} \frac{1-\pi_i}{\pi_i} \frac{\pi_{ik} - \pi_i \pi_k}{\pi_i \pi_k} \right| &\leq \sum_{i=1}^N \left(\frac{1}{\pi_i} - 1 \right) \sum_{k \neq i} \frac{|\pi_{ik} - \pi_i \pi_k|}{\pi_i \pi_k} \\
&= O\left(\frac{N^2}{n}\right) \cdot O\left(\frac{N}{n}\right) = O\left(\frac{N^3}{n^2}\right),
\end{aligned}$$

by means of condition (i) and (C2*). For the remainder terms in b_2, b_3, b_4 we obtain bounds for the same quantities, as the previous three. The rest of the proof of Lemma B.1 remains the same. \square

LEMMA B.2. *If $x_N \rightsquigarrow x$ and $y_N \rightsquigarrow y$ in $D[0, 1]$ with the Skorohod metric, and $x, y \in C[0, 1]$, then the sequence $\{x_N + y_N\}$ is also tight in $D[0, 1]$.*

PROOF. We can use Theorem 13.2 from [3]. The first condition follows easily since

$$\sup_{t \in [0, 1]} |x_N(t) + y_N(t)| \leq \sup_{t \in [0, 1]} |x_N(t)| + \sup_{t \in [0, 1]} |y_N(t)|.$$

Because $x_N \rightsquigarrow x$ and $y_N \rightsquigarrow y$ both sequences $\{x_N\}$ and $\{y_N\}$ are tight, so that they satisfy the first condition of Theorem 13.2 individually. For condition (ii) of Theorem 13.2 in [3], choose $\epsilon > 0$. According to (12.7) in [3], for any $0 < \delta < 1/2$,

$$w'_x(\delta) \leq w_x(2\delta).$$

This means that

$$\begin{aligned} \mathbb{P} \{w'_{x_N + y_N}(\delta) \geq \epsilon\} &\leq \mathbb{P} \{w_{x_N + y_N}(2\delta) \geq \epsilon\} \\ &\leq \mathbb{P} \{w_{x_N}(2\delta) \geq \epsilon/2\} + \mathbb{P} \{w_{y_N}(2\delta) \geq \epsilon/2\}. \end{aligned}$$

Consider the first probability. Since $x_N \rightsquigarrow x$ in $D[0, 1]$ with the Skorohod metric, according to the almost sure representation theorem (see, e.g., Theorem 11.7.2 in [9]), there exist \tilde{x}_n and \tilde{x} , having the same distribution as x_N and x , respectively, such that $\tilde{x}_N \rightarrow \tilde{x}$, with probability one, in the Skorohod metric. Because $\tilde{x} \stackrel{d}{=} x$ and $x \in C[0, 1]$, also $\tilde{x} \in C[0, 1]$. Hence, since \tilde{x} is continuous, it follows that

$$(B.5) \quad \sup_{t \in [0, 1]} |\tilde{x}_N(t) - \tilde{x}(t)| \rightarrow 0, \quad \text{with probability one.}$$

We then find that

$$\begin{aligned} \mathbb{P} \{w_{x_N}(2\delta) \geq \epsilon/2\} &= \mathbb{P} \left\{ \sup_{|s-t| < 2\delta} |x_N(s) - x_N(t)| \geq \epsilon/2 \right\} \\ &= \mathbb{P} \left\{ \sup_{|s-t| < 2\delta} |\tilde{x}_N(s) - \tilde{x}_N(t)| \geq \epsilon/2 \right\} \\ &\leq \mathbb{P} \left\{ \sup_{|s-t| < 2\delta} |\tilde{x}(s) - \tilde{x}(t)| \geq \epsilon/4 \right\} \\ &\quad + \mathbb{P} \left\{ \sup_{s \in [0, 1]} |\tilde{x}_N(s) - \tilde{x}(s)| \geq \epsilon/8 \right\} + \mathbb{P} \left\{ \sup_{t \in [0, 1]} |\tilde{x}_N(t) - \tilde{x}(t)| \geq \epsilon/8 \right\}. \end{aligned}$$

The latter two probabilities tend to zero due to (B.5). For the first probability on the right hand side, note that $C[0, 1]$ is separable and complete. This means that each random element in $C[0, 1]$ is tight. Hence, $\tilde{x} \in C[0, 1]$ is tight, so that according to Theorem 7.3 in [3], there exists a $0 < \delta < 1/2$, such that

$$\mathbb{P} \left\{ \sup_{|s-t| < 2\delta} |x(s) - x(t)| \geq \epsilon/4 \right\} = \mathbb{P} \{w_x(2\delta) \geq \epsilon/4\} \leq \eta.$$

We conclude that $\mathbb{P} \{w_{x_N}(2\delta) \geq \epsilon/2\} \rightarrow 0$, and the same result for y_N can be obtained similarly. This proves the lemma. \square

Proof of Proposition 5.1. It suffices to prove (HT1) for rejective sampling. The proof is along the lines of the proof of Theorem 3.2 in [2] and uses results from [11]. To adapt to the notation used in [11], we will show that

$$(B.6) \quad \frac{1}{S_N} \left(\frac{1}{N} \sum_{i=1}^N \frac{\eta_i V_i}{\pi_i} - \frac{1}{N} \sum_{i=1}^N V_i \right) \rightarrow N(0, 1), \quad \omega - \text{a.s.},$$

in distribution under \mathbb{P}_d , where

$$S_N^2 = \text{Var}_d \left[\frac{1}{N} \sum_{i=1}^N \left(\frac{\eta_i}{\pi_i} - 1 \right) V_i \right] = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} V_i V_j.$$

Here, the η_1, \dots, η_N represent the inclusion indicators corresponding to the rejective sampling design. The rejective sampling design can be represented by a Poisson design conditional on the sample size being equal to n (e.g., see [10]) Let ξ_1, \dots, ξ_N denote the indicators of the corresponding Poisson design. Note that $\mathbb{E}_d[\eta_i] = \pi_i$ and $\mathbb{E}_d[\xi_i] = p_i$, where the p_i 's can be chosen such that $\sum_{i=1}^N p_i = n$, and that $d_N = \sum_{i=1}^N \pi_i(1 - \pi_i) \rightarrow \infty$, as a consequence of (B2).

In order to obtain (B.6), it is more convenient to rewrite the left hand side. To this end, note that by means of Theorem 5.1 in [10] and the fact that $\sum_{i=1}^N \eta_i = \sum_{i=1}^N p_i = n$, we can write

$$(B.7) \quad \frac{1}{NS_N} \sum_{i=1}^N (\eta_i - \pi_i) \frac{V_i}{\pi_i} = (1 + o(1)) \frac{1}{NS_N} \sum_{i=1}^N (\eta_i - p_i) \left(\frac{V_i}{p_i} - \theta_N \right)$$

where

$$(B.8) \quad \begin{aligned} \theta_N &= \frac{1}{B_N^2} \sum_{i=1}^N V_i(1 - p_i), \\ B_N^2 &= \sum_{i=1}^N p_i(1 - p_i) = (1 + o(1))d_N, \end{aligned}$$

according to Theorem 5.1 in [10]. The summation on the right hand side of (B.7) is of the form

$$R_N(\eta) = \sum_{m=1}^N f_{m,N}(\eta_m), \quad \text{where } f_{m,N}(y) = \frac{1}{NS_N} (y - p_m) \left(\frac{V_m}{p_m} - \theta_N \right),$$

which is of the type considered in [11]. Furthermore, note that

$$\begin{aligned} \Lambda_N &= \sum_{m=1}^N \mathbb{E}_d [f_{m,N}(\xi_m)] = 0 \\ \gamma_N &= \frac{1}{B_N^2} \sum_{m=1}^N \text{cov}(f_{m,N}(\xi_m), \xi_m) = 0. \end{aligned}$$

Under suitable conditions on that we specify below

$$g_m(y) = f_{m,N}(y) - \mathbb{E}_d f_{m,N}(\xi_m) - \gamma_N(y - \mathbb{E}_d \xi_m) = f_{m,N}(y),$$

according to Theorem 3.1 in [11],

$$(B.9) \quad \frac{R_N(\eta)}{\sigma_N} \rightarrow N(0, 1)$$

in distribution, where

$$\sigma_N^2 = \sum_{m=1}^N \text{Var}[g_m(\xi_m)] = \frac{1}{N^2 S_N^2} \sum_{m=1}^N \left(\frac{V_m}{p_m} - \theta_N \right)^2 p_m (1 - p_m).$$

From Theorem 5.1 and 6.1 in [10], it follows that

$$(B.10) \quad \sigma_N^2 = (1 + o(1)) \frac{1}{N^2 S_N^2} \sum_{m=1}^N \left(\frac{V_m}{\pi_m} - R \right)^2 \pi_m (1 - \pi_m) = 1 + o(1),$$

where $R = d_N^{-1} \sum_{i=1}^N \pi_i (1 - \pi_i)$. Therefore, (B.6) is equivalent with (B.9) and it remains to check the conditions of Theorem 3.1 in [11].

Define (as mentioned in [2], a factor \sqrt{N} after ϵ is missing in [11])

$$\begin{aligned}\mathcal{L}_{1,N}(\epsilon) &= \frac{1}{B_N^3} \sum_{m=1}^N \mathbb{E}_d |\xi_m - p_m|^3 \mathbb{1} \{|\xi_m - p_m| \leq \epsilon B_N\} \\ \mathcal{L}_{2,N}(\epsilon) &= \frac{1}{B_N^2} \sum_{m=1}^N \mathbb{E}_d |\xi_m - p_m|^2 \mathbb{1} \{|\xi_m - p_m| > \epsilon B_N\} \\ L_{2,N}(\epsilon) &= \frac{1}{\sigma_N^2} \sum_{m=1}^N \mathbb{E}_d g_m(\xi_m)^2 \mathbb{1} \{|g_m(\xi_m)| > \epsilon \sigma_N\} \\ M_N(T) &= \inf_{T \leq \tau \leq \pi} \sum_{m=1}^N (1 - |\mathbb{E}_d \exp(i\tau \xi_m)|^2),\end{aligned}$$

if $T \leq \pi$ else $M_N(T) = \infty$. If for arbitrary $\epsilon > 0$,

- (i) $L_{2,N}(\epsilon) \rightarrow 0$,
- (ii) $\mathcal{L}_{2,N}(\epsilon) \rightarrow 0$,
- (iii) $M_N(\pi(4B_N \mathcal{L}_{1,N}(\epsilon))^{-1}) \rightarrow \infty$
- (iv) $\min(B_N, \sqrt{N}) = o(M_N(\pi(4B_N \mathcal{L}_{1,N}(\epsilon))^{-1}))$

then (B.9) holds, according to Theorem 3.1 in [11].

ad(i). Since $|V_i| \leq K$ and $p_m/\pi_m = 1 + o(1)$, according to Theorem 5.1 in [10], together with (C1) it follows that for N sufficiently large

$$|g_m(\xi_m)| \leq \frac{2K}{nS_N} \left(\frac{N}{nK_1} + \frac{N}{B_N^2} \right) \leq \frac{2K}{nS_N} \left(\frac{1}{K_1} + \frac{n}{B_N^2} \right).$$

Together, with condition (B1), there exists $C > 0$, such that

$$L_{2,N}(\epsilon) \leq \frac{C}{n^2 S_N^2} \frac{1}{\sigma_N^2} \sum_{m=1}^N \frac{\mathbb{E}_d g_m(\xi_m)^2}{\epsilon^2 \sigma_N^2} = \frac{C}{n^2 S_N^2} \frac{1}{\epsilon^2 \sigma_N^2} \rightarrow 0$$

according to (B3) and (B.10). This proves condition (i) in [11].

ad(ii). Since $B_N^2 = d_N \rightarrow \infty$, for N sufficiently large, $\{|\xi_m - p_m| > \epsilon B_N\} \subset \{2 > \epsilon B_N\} = \emptyset$, which means that for N sufficiently large $\mathcal{L}_{2,N}(\epsilon) = 0$. This proves condition (ii) in [11].

ad(iii) First note that (see also [2])

$$|\mathbb{E}_d \exp(i\tau \xi_m)|^2 = 1 + 2p_m(1 - p_m)(\cos \tau - 1)$$

so that for $T \in [0, \pi]$,

$$M_N(T) = 2 \inf_{T \leq \tau \leq \pi} (1 - \cos \tau) \sum_{m=1}^N p_m(1 - p_m) = 2B_N^2(1 - \cos T).$$

Because $B_N^2 = d_N \rightarrow \infty$, for N sufficiently large $\mathbb{1}\{|\xi_m - p_m| \leq \epsilon B_N\} = 1$. This means that for N sufficiently large

$$\mathcal{L}_{1,N}(\epsilon) = \frac{1}{B_N^3} \sum_{m=1}^N \mathbb{E}_d |\xi_m - p_m|^3,$$

where

$$\mathbb{E}_d |\xi_m - p_m|^3 = p_m(1 - p_m) \{1 - 2p_m + 2p_m^2\}.$$

It follows that

$$\frac{1}{2}p_m(1 - p_m) \leq \mathbb{E}_d |\xi_m - p_m|^3 \leq p_m(1 - p_m),$$

so that for N sufficiently large, $2 \leq 4B_N \mathcal{L}_{1,N}(\epsilon) \leq 4$, and therefore

$$\begin{aligned} M_N (\pi(4B_N \mathcal{L}_{1,N}(\epsilon))^{-1}) &= 2B_N^2(1 - \cos(\pi(4B_N \mathcal{L}_{1,N}(\epsilon))^{-1})) \\ &\geq 2B_N^2(1 - \cos(\pi/4)) \rightarrow \infty. \end{aligned}$$

This proves condition (iii) in [11].

ad(iv). From the previous computations it follows that

$$\begin{aligned} \frac{\min(B_N, \sqrt{N})}{M_N (\pi(4B_N \mathcal{L}_{1,N}(\epsilon))^{-1})} &\leq \frac{\min(B_N, \sqrt{N})}{2B_N^2(1 - \cos(\pi/4))} \\ &= \frac{1}{2(1 - \cos(\pi/4))} \min\left(1/B_N, \sqrt{N}/B_N^2\right) \rightarrow 0 \end{aligned}$$

according to (B2) and the fact that $B_N^2 = d_N \rightarrow \infty$. This proves condition (iv) in [11]. \square

Proof of (6.2). Following [8], one can write $\phi = \psi_2 \circ \psi_1$, where

$$\begin{aligned} \psi_1(F) &= (F, \beta F^{-1}(\alpha)) \\ \psi_2(F, x) &= F(x). \end{aligned}$$

The Hadamard-derivative of ϕ can then be obtained from the chain rule, e.g., see Lemma 3.9.3 in [13]. According to Lemma 3.9.20 in [13], for $0 < \alpha < 1$ and $F \in \mathbb{D}_\phi$ that have a positive derivative at $F^{-1}(\alpha)$, the map ψ_1 is Hadamard-differentiable at F tangentially to the set of functions $h \in D(\mathbb{R})$ that are continuous at $F^{-1}(\alpha)$ with derivative

$$\psi'_{1,F}(h) = \left(h, -\beta \frac{h(F^{-1}(\alpha))}{f(F^{-1}(\alpha))} \right).$$

It is fairly straightforward to show that for F that are differentiable at x , the mapping ψ_2 is Hadamard-differentiable at (F, x) tangentially to the set of pairs (h, ϵ) , such that h is continuous at x and $\epsilon \in \mathbb{R}$, with derivative

$$\psi'_{2,(F,x)}(h, \epsilon) = \epsilon f(x) + h(x).$$

Then for $F \in \mathbb{D}_\phi$ that are differentiable at $\beta F^{-1}(\alpha)$, the mapping ψ_2 is Hadamard-differentiable at $\psi_1(F) = (F, \beta F^{-1}(\alpha))$. It follows from the chain rule that $\phi(F) = F(\beta F^{-1}(\alpha)) = \psi_2 \circ \psi_1(F)$ is Hadamard-differentiable at F tangentially to the set \mathbb{D}_0 consisting of functions $h \in D(\mathbb{R})$ that are continuous at $F^{-1}(\alpha)$ with derivative

$$\phi'_F(h) = -\beta \frac{f(\beta F^{-1}(\alpha))}{f(F^{-1}(\alpha))} h(F^{-1}(\alpha)) + h(\beta F^{-1}(\alpha)).$$

□

APPENDIX C: FIXED SIZE SAMPLING DESIGNS WITH DETERMINISTIC INCLUSION PROBABILITIES

Conditions (C2)-(C4) put bounds on maximum correlations. This is somewhat restrictive, and bounds on the average correlation may be more suitable for applications. This can indeed be accomplished to some extent for fixed size sampling designs P , with inclusion probabilities π_i that do not depend on ω .

Suppose there exists a $K > 0$, such that for all $N = 1, 2, \dots$,

$$\begin{aligned} \text{(C2*)} \quad & \text{for all } j = 1, 2, \dots, N: \frac{n}{N} \sum_{i \neq j} \left| \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \right| \leq K, \\ \text{(C3*)} \quad & \frac{n}{N^3} \sum_{(i,j,k) \in D_{3,N}} \left| \frac{\pi_{ijk} - \pi_i \pi_j \pi_k}{\pi_i \pi_j \pi_k} \right| \leq K. \\ \text{(C4*)} \quad & \frac{n^2}{N^4} \left| \sum_{(i,j,k,l) \in D_{4,N}} \frac{\mathbb{E}_d [(\xi_i - \pi_i)(\xi_j - \pi_j)(\xi_k - \pi_k)(\xi_l - \pi_l)]}{\pi_i \pi_j \pi_k \pi_l} \right| \leq K. \end{aligned}$$

The summation in (C2*) has a number of terms of the order N . This means that typically the summands must decrease at rate $1/N$. This is comparable to condition (ii) in Proposition 3.1. Similarly for summands in the summation in (C3*). The summands in (C4*) have to overcome a factor of the order N^2 , which will typically not be the case for general sampling designs. However, according to Lemma 2 in [4], the fourth order correlation can be decomposed in terms of the type

$$(-1)^{4-m} \frac{\pi_{i_1 \dots i_m} - \pi_{i_1} \dots \pi_{i_m}}{\pi_{i_1} \dots \pi_{i_m}}, \quad m = 2, 3, 4.$$

Because these terms can be both negative and positive, they may cancel each other in such a way that (C4*) does hold. This is for instance the case for simple random sampling, e.g., see the discussion in Remarks (iii) and (iv) in [6], or for rejective sampling, see Proposition 1 in [4].

By using Lemma 2 in [4] it follows that conditions (C2*)-(C4*) are implied by (C2)-(C4). The following corollary covers the results from Sections 3 and 4 under the weaker conditions (C2*)-(C4*), for fixed size sampling designs with deterministic inclusion probabilities.

COROLLARY C.1. *Let P be a fixed size sampling design with deterministic inclusion probabilities. Suppose that (C1), (C2*)-(C4*), (HT1), hold, as well as conditions (i) and (ii) from Proposition 3.1. Then*

- (i) *the conclusion of Proposition 3.1 holds;*
- (ii) *if (HT3) is satisfied and $\mu_{\pi_1} > 0$, then the conclusion of Proposition 3.2 holds;*
- (iii) *if (HT3) is satisfied, then the conclusions of Propositions 4.1 and 4.2 hold.*

Proof. We first re-prove Lemma 9.1 under conditions (C2*)-(C4*). Because n is deterministic, it can be taken out of the expectation $\mathbb{E}_{d,m}$. When also π_1, \dots, π_N are deterministic, this means that the expectation \mathbb{E}_d over the ξ_i 's can be separated from the expectation \mathbb{E}_m over the A_i 's and B_j 's in (9.2). It follows that

$$(C.1) \quad \begin{aligned} & \frac{1}{N^4} \sum_{i=1}^N \sum_{k=1}^N \mathbb{E}_{d,m} [n^2 \alpha_i^2 \beta_k^2] \\ &= \frac{n^2}{N^4} \sum_{(i,k) \in D_{2,N}} \frac{\mathbb{E}_d [(\xi_i - \pi_i)^2 (\xi_k - \pi_k)^2]}{\pi_i^2 \pi_k^2} p_1 p_2. \end{aligned}$$

Straightforward computation shows that $\mathbb{E}_d(\xi_i - \pi_i)^2 (\xi_k - \pi_k)^2$ equals

$$(\pi_{ik} - \pi_i \pi_k)(1 - 2\pi_i)(1 - 2\pi_k) + \pi_i \pi_k (1 - \pi_i)(1 - \pi_k).$$

The contribution of the last term is

$$\frac{n^2}{N^4} \sum_{(i,k) \in D_{2,N}} \frac{\pi_i \pi_k (1 - \pi_i)(1 - \pi_k)}{\pi_i^2 \pi_k^2} \leq \left(\frac{n}{N^2} \sum_{i=1}^N \left(\frac{1}{\pi_i} - 1 \right) \right)^2 = O(1),$$

according to condition (i) of Proposition 3.1. With (C1) and (C2*), the contribution of the first term is

$$(C.2) \quad \begin{aligned} & \frac{n^2}{N^4} \sum_{(i,k) \in D_{2,N}} \sum_{(i,k) \in D_{2,N}} \frac{(\pi_{ik} - \pi_i \pi_k)(1 - 2\pi_i)(1 - 2\pi_k)}{\pi_i^2 \pi_k^2} \\ & \leq O\left(\frac{N^2}{n^2}\right) \frac{n^2}{N^4} \cdot N \sum_{i \neq k} \left| \frac{\pi_{ik} - \pi_i \pi_k}{\pi_i \pi_k} \right| = O\left(\frac{1}{n}\right). \end{aligned}$$

This establishes (9.4).

For the second (and third) summation on the right hand side of (9.2), we have

$$\begin{aligned} & \frac{1}{N^4} \left| \sum_{i=1}^N \sum_{j \neq i} \sum_{k=1}^N \mathbb{E}_{d,m} [n^2 \alpha_i \alpha_j \beta_k^2] \right| \\ & \leq \frac{n^2}{N^4} \left| \sum_{(i,j,k) \in D_{3,N}} \sum_{(i,j,k) \in D_{3,N}} \mathbb{E}_d \left[\frac{(\xi_i - \pi_i)(\xi_j - \pi_j)(\xi_k - \pi_k)^2}{\pi_i \pi_j \pi_k^2} \right] \right| p_1 p_2 \end{aligned}$$

We still have that $\mathbb{E}_d(\xi_i - \pi_i)(\xi_j - \pi_j)(\xi_k - \pi_k)^2$ equals

$$(1 - 2\pi_k) \mathbb{E}_d(\xi_i - \pi_i)(\xi_j - \pi_j)(\xi_k - \pi_k) + \pi_k(1 - \pi_k) \mathbb{E}_d(\xi_i - \pi_i)(\xi_j - \pi_j).$$

The contribution of the last term is

$$\begin{aligned} & \left| \frac{n^2}{N^4} \sum_{(i,j,k) \in D_{3,N}} \sum_{(i,j,k) \in D_{3,N}} \sum_{(i,j,k) \in D_{3,N}} \left(\frac{1}{\pi_k} - 1 \right) \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \right| \\ & \leq \left| \frac{n}{N^2} \sum_{(i,j) \in D_{2,N}} \sum_{(i,j) \in D_{2,N}} \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \right| \cdot \frac{n}{N^2} \sum_{k=1}^N \left(\frac{1}{\pi_k} - 1 \right) = O(1), \end{aligned}$$

according to conditions (i)-(ii) of Proposition 3.1. From Lemma 2 in [4], we have that $\mathbb{E}_d \xi_i - \pi_i)(\xi_j - \pi_j)(\xi_k - \pi_k)$ splits into

1. $-(\pi_{ij} - \pi_i \pi_j) \pi_k - (\pi_{ik} - \pi_i \pi_k) \pi_j - (\pi_{jk} - \pi_j \pi_k) \pi_i$.
2. $\pi_{ijk} - \pi_i \pi_j \pi_k$.

According to (C1) and (C2*), the contribution of the terms in the first case is of the order $O(1)$ similarly to (C.2), whereas (C1) and (C3*) yield that the contribution of the second case is also of the order $O(1)$. This establishes (9.5).

Finally,

$$\begin{aligned} & \frac{1}{N^4} \sum_{i=1}^N \sum_{j \neq i} \sum_{k=1}^N \sum_{l \neq k} \mathbb{E}_{d,m} [n^2 \alpha_i \alpha_j \beta_k \beta_l] \\ &= \frac{n^2}{N^4} \sum_{(i,j,k,l) \in D_{4,N}} \frac{\mathbb{E}_d [(\xi_i - \pi_i)(\xi_j - \pi_j)(\xi_k - \pi_k)(\xi_l - \pi_l)]}{\pi_i \pi_j \pi_k \pi_l} p_1^2 p_2^2. \end{aligned}$$

Because $0 \leq p_1, p_2 \leq 1$, together with (C4*), we obtain (9.6). Together with (9.4), (9.5) and decomposition (9.2), this proves Lemma 9.1.

Furthermore, at the cost of some extra technicalities, it can be seen that Lemma B.1 in [5] holds with (C2*) and conditions (i)-(ii) from Proposition 3.1 instead of (C2). Details can be found in Supplement B in [5]. From here on, the proofs of Propositions 3.1, 3.2, 4.1, and 4.2 remain the same. \square

REFERENCES

- [1] BERGER, Y. G. (1998). Rate of convergence to normal distribution for the Horvitz-Thompson estimator. *J. Statist. Plann. Inference* **67** 209–226. [MR1624693](#)
- [2] BERTAIL, P., CHAUTRU, E. and CLÉMENTÇON, S. (2016). Empirical processes in survey sampling. DOI:10.1111/sjos.12243.
- [3] BILLINGSLEY, P. (1999). *Convergence of probability measures*, second ed. *Wiley Series in Probability and Statistics: Probability and Statistics*. John Wiley & Sons, Inc., New York A Wiley-Interscience Publication. [MR1700749 \(2000e:60008\)](#)
- [4] BOISTARD, H., LOPUHAÄ, H. P. and RUIZ-GAZEN, A. (2012). Approximation of rejective sampling inclusion probabilities and application to high order correlations. *Electron. J. Stat.* **6** 1967–1983. [MR3020253](#)
- [5] BOISTARD, H., LOPUHAÄ, H. P. and RUIZ-GAZEN, A. (2016). Supplement to "Functional central limit theorems for single-stage samplings designs". DOI:10.1214/16-AOS1507SUPP.
- [6] BREIDT, F. J. and OPSOMER, J. D. (2000). Local polynomial regression estimators in survey sampling. *Ann. Statist.* **28** 1026–1053. [MR1810918 \(2001m:62012\)](#)
- [7] CONTI, P. L., MARELLA, D. and MECATTI, F. (2015). Recovering sampling distributions of statistics of finite populations via resampling: a predictive approach. *Submitted*.
- [8] DELL, F. and D'HAULTFEUILLE, X. (2008). Measuring the evolution of complex indicators: Theory and application to the poverty rate in France. *Ann. Économ. Statist.* **90** 259–290.
- [9] DUDLEY, R. M. (2002). *Real analysis and probability*. *Cambridge Studies in Advanced Mathematics* **74**. Cambridge University Press, Cambridge Revised reprint of the 1989 original. [MR1932358 \(2003h:60001\)](#)
- [10] HÁJEK, J. (1964). Asymptotic theory of rejective sampling with varying probabilities from a finite population. *Ann. Math. Statist.* **35** 1491–1523. [MR0178555 \(31 #2812\)](#)

- [11] MIRAKHMEDOV, S. M., RAO JAMMALAMADAKA, S. and MOHAMED, I. B. (2014). On Edgeworth expansions in generalized urn models. *J. Theoret. Probab.* **27** 725–753. [MR3245983](#)
- [12] RUBIN-BLEUER, S. and SCHIOPU KRATINA, I. (2005). On the two-phase framework for joint model and design-based inference. *Ann. Statist.* **33** 2789–2810. [MR2253102 \(2007k:62025\)](#)
- [13] VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak convergence and empirical processes. Springer Series in Statistics.* Springer-Verlag, New York With applications to statistics. [MR1385671 \(97g:60035\)](#)