A Online Supplement for “General-Purpose Ranking & Selection for Computer Simulation”

Soonhui Lee¹ and Barry L. Nelson²

¹shlee@unist.ac.kr, ²nelsonb@northwestern.edu

¹School of Business Administration
Ulsan National Institute of Science and Technology
Ulsan, REPUBLIC OF KOREA

²Department of Industrial Engineering and Management Sciences
Northwestern University
Evanston, IL 60208-3119, USA

This appendix provides the mathematical proofs of the main theorems and the experiment results of the algorithms in the paper.

A.1 Auxiliary Results

The following theorem in Bergström (1945) provides a Berry-Esseen type bound in $\mathbb{R}^k$.

Theorem A.1. (Bergström (1945)) Let $F(\mathbf{x})$ be a distribution function in $\mathbb{R}^k$ with the mean values $\underline{\mu}$, finite third moments and the translated second order moments σ_{ij} , such that $Q = \sum_{i,j=1,\dots,k} \sigma_{ij} t_i t_j$ is a positive definite form with the determinant Δ . Further let $\Phi_{\Sigma}(\mathbf{x})$ be a normal distribution function which has the mean values 0 and the same translated second order moments as $F(\mathbf{x})$. Then we have for all $\mathbf{x}$,

$$\left| \Pr \left\{ n^{1/2}(\bar{\mathbf{X}}_n - \underline{\mu}) \leq \mathbf{x} \right\} - \Phi_{\Sigma}(\mathbf{x}) \right| \leq C(k) \rho n^{-1/2} \max_{i,j=1,\dots,k} |[\Sigma^{-1}]_{(i,j)}|^{3/2}$$

where $C(k)$ is a constant only depending on k , $\rho = \sum_{i=1}^k \rho^{(i)}$ with $\rho^{(i)} = \int_{\mathbb{R}} |x_i - E(X_i)|^3 dF_i(x)$, $\Sigma = E(\mathbf{X} - E\mathbf{X})(\mathbf{X} - E\mathbf{X})^{\top}$.

This type of the uniform bound also holds on the probability of any measurable convex set in $\mathbb{R}^k$, as shown in Gotze (1991), Sazonov (1968), and Bergström (1969). Although they present different bounds, they imply the following which we state as Theorem A.2.

Theorem A.2. For all $x \in \mathbb{R}^+$,

$$\left| \Pr \left\{ n^{1/2}|\bar{\mathbf{X}}_n - \underline{\mu}| \leq x \cdot \mathbf{1}_k \right\} - \Gamma_{\Sigma}(x) \right| \leq C_b(k) \rho_F f_1(n) \quad (\text{A.1})$$

where $C_b(k)$ is a constant only depending on k , ρ_F is a function of several moments of the distribution F , and $f_1(n)$ is a function depending only on n , which tends to zero as $n \rightarrow \infty$.

Note that ρ_F in Gotze (1991), Sazonov (1968), and Bergström (1969) is finite when the third moment of F is finite.

Theorem A.3. Pólya’s Theorem (Chow and Teicher (1988)). If X_n converges in distribution to X , and F_X is everywhere continuous, then

$$\sup_x |F_{X_n}(x) - F_X(x)| \rightarrow 0.$$

Definition A.1. *Uniform continuity in probability of $\{Y_n\}$ (Anscombe (1952)). Given any small positive ε and η , there is a large v and small positive c such that, for any $n > v$,*

$$\Pr\{|Y_{n'} - Y_n| < \varepsilon w_n \text{ for all } n' \text{ such that } |n - n'| < cn\} > 1 - \eta \quad (\text{A.2})$$

where $\{w_n\}$ is a sequence of positive numbers.

Remark. We use Definition A.1 with $w_n = n^{-1/2}$ throughout this paper.

Lemma A.1. *Suppose that $\hat{\theta}_{1n}, \hat{\theta}_{2n}, \dots, \hat{\theta}_{kn}$ are k estimators that are individually uniformly continuous in probability. Then $Y_n = \max_i \hat{\theta}_{in}$ and $Y'_n = \max_i |\hat{\theta}_{in}|$ are also uniformly continuous in probability.*

Proof If $\hat{\theta}_{1n}, \hat{\theta}_{2n}, \dots, \hat{\theta}_{kn}$ are k estimators that are individually uniformly continuous in probability, then given ε and η , we can find v_i and small positive c_i such that

$$\Pr\{|\hat{\theta}_{in'} - \hat{\theta}_{in}| < \varepsilon w_n \text{ for all } n' \text{ such that } |n - n'| < c_i n\} > 1 - \eta/(k+1)$$

for any $n > v_i, i = 1, \dots, k$. Let $v = \max_i v_i$ and $c = \min_i c_i$. Then we have

$$\begin{aligned} & \Pr\{|Y_{n'} - Y_n| < \varepsilon w_n \text{ for all } n' \text{ such that } |n - n'| < cn\} \\ & \geq \Pr\{|\hat{\theta}_{in'} - \hat{\theta}_{in}| < \varepsilon w_n \text{ for all } n' \text{ such that } |n - n'| < c_i n, \quad i = 1, 2, \dots, k\} \\ & \geq 1 - k\eta/(k+1) > 1 - \eta. \end{aligned}$$

The first inequality is easily verified using the definition of Y_n and the second inequality holds by the Bonferroni inequality. Therefore, Y_n is uniformly continuous in probability. Now let v_i, v, c_i , and c be defined as above. Then we have

$$\begin{aligned} & \Pr\{|Y'_{n'} - Y'_n| < \varepsilon w_n \text{ for all } n' \text{ such that } |n - n'| < cn\} \\ & \geq \Pr\{||\hat{\theta}_{in'}| - |\hat{\theta}_{in}|| < \varepsilon w_n \text{ for all } n' \text{ such that } |n - n'| < c_i n, \quad i = 1, 2, \dots, k\} \\ & \geq \Pr\{|\hat{\theta}_{in'} - \hat{\theta}_{in}| < \varepsilon w_n \text{ for all } n' \text{ such that } |n - n'| < c_i n, \quad i = 1, 2, \dots, k\} \\ & \geq 1 - k\eta/(k+1) > 1 - \eta. \end{aligned}$$

□

Remark. Anscombe (1952) shows that $\hat{\theta}_{in}$ is individually uniformly continuous in probability when $\hat{\theta}_{in}$ is a sample quantile or average.

Proposition A.1. *Under assumptions stated in Theorem 2, given $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$, as $n \rightarrow \infty$,*

$$\sup_{\mathbf{y}} |\Pr^*\{\mathbf{G}_n(q) \leq \mathbf{y}\} - \Phi_{\Sigma}(\mathbf{y})| \rightarrow 0$$

where

$$\mathbf{G}_n(q) = \left(\sqrt{n}(\hat{F}_{1n}^{*-1}(q) - \hat{F}_{1n}^{-1}(q)), \dots, \sqrt{n}(\hat{F}_{kn}^{*-1}(q) - \hat{F}_{kn}^{-1}(q)) \right)^\top$$

and Σ is defined as in Theorem 2.

Proof We first consider the univariate case and extend it to the multivariate case. WLOG, suppose that

$$\frac{r-1}{n} < q \leq \frac{r}{n}$$

which implies that $r = \lceil nq \rceil$. To simplify notation, let $F_n = \widehat{F}_n$ and $F_n^* = \widehat{F}_n^*$. For the univariate case, given $X_1, X_2, \dots, X_n (\in \mathbb{R})$, we want to show

$$\Pr^* \left\{ \sqrt{n}(F_n^{*-1}(q) - F_n^{-1}(q)) \leq y \right\} \rightarrow \Phi_\Sigma(y) \quad (\text{A.3})$$

where $\Sigma = q(1-q)/(f(F^{-1}(q)))^2$. Note we have

$$\Pr^* \left\{ \sqrt{n}(F_n^{*-1}(q) - F_n^{-1}(q)) \leq y \right\} = \Pr^* \left\{ F_n^{*-1}(q) \leq \frac{y}{\sqrt{n}} + F_n^{-1}(q) \right\}$$

and let

$$x_n = \frac{y}{\sqrt{n}} + F_n^{-1}(q).$$

On the other hand,

$$\begin{aligned} \Pr^* \left\{ F_n^{*-1}(q) \leq x \right\} &= \Pr^* \left\{ \sum_{i=1}^n \mathbf{I}\{X_i^* \leq x\} \geq r \right\} \\ &= \Pr^* \left\{ \sqrt{n}(F_n^*(x) - F_n(x)) \geq \sqrt{n} \left(\frac{r}{n} - F_n(x) \right) \right\}. \end{aligned} \quad (\text{A.4})$$

It can be shown from Theorem 2.1 in Bickel and Freedman (1981) that $\sqrt{n}(F_n^*(x) - F_n(x))$ converges in distribution to $N(0, \Sigma')$ as $n \rightarrow \infty$ where $\Sigma' = F(x)(1 - F(x))$. After plugging $x = x_n$ into (A.4), we have

$$\Pr^* \left\{ F_n^{*-1}(q) \leq x_n \right\} = \Pr^* \left\{ \sqrt{n}(F_n^*(x_n) - F_n(x_n)) \geq \sqrt{n} \left(\frac{r}{n} - F_n(x_n) \right) \right\}$$

where

$$\sqrt{n} \left(\frac{r}{n} - F_n(x_n) \right) = \frac{\sqrt{n} \left(F_n(F_n^{-1}(q)) - F_n\left(\frac{y}{\sqrt{n}} + F_n^{-1}(q)\right) \right)}{y}$$

noting that $F_n(F_n^{-1}(q)) = r/n$. In Reiss (1989), it has been shown that

$$\frac{F_n(F_n^{-1}(q)) - F_n\left(\frac{y}{\sqrt{n}} + F_n^{-1}(q)\right)}{\frac{y}{\sqrt{n}}} \rightarrow -f(F^{-1}(q))$$

a.s. for any $y \neq 0 (\in \mathbb{R})$. As $n \rightarrow \infty$, $F(x_n)(1 - F(x_n))$ converges to $q(1 - q)$ a.s. This completes the proof for the univariate case. Note that (A.3) has been presented by Bickel and Freedman (1981).

We now extend this result to the multivariate case. Let

$$\mathbf{Y}_j = \left(\mathbf{I}_{\{X_{1j} \leq x_1\}}, \dots, \mathbf{I}_{\{X_{kj} \leq x_k\}} \right)^\top \quad (\text{A.5})$$

and

$$\mathbf{Y}_j^* = \left(\mathbf{I}_{\{X_{1j}^* \leq x_1\}}, \dots, \mathbf{I}_{\{X_{kj}^* \leq x_k\}} \right)^\top$$

defined for $\mathbf{x} \in \mathbb{R}^k$ such that $\text{Var}(\mathbf{Y}_j)$ is positive definite, which implies that $x_i, i = 1, 2, \dots, k$ are not all too small or large. By defining $Z_j = \underline{\lambda}^\top \mathbf{Y}_j$ for non-zero $\underline{\lambda} \in \mathbb{R}^k$, we have $\underline{\lambda}^\top \sqrt{n}(\bar{\mathbf{Y}}^* - \bar{\mathbf{Y}}) = \sqrt{n}(\bar{Z}^* - \bar{Z})$ where we can apply Theorem 2.1 in Bickel and Freedman (1981). Therefore, $\sqrt{n}(\bar{Z}^* - \bar{Z})$ converges in distribution to $N(0, \underline{\lambda}^\top \boldsymbol{\tau}_x \underline{\lambda})$ where $\tau_x(i, j) = F_{ij}(x_i, x_j) - F_i(x_i)F_j(x_j)$. As this holds for each $\underline{\lambda} \in \mathbb{R}^k$, $\sqrt{n}(\bar{\mathbf{Y}}^* - \bar{\mathbf{Y}})$ converges in distribution to $N_k(0, \boldsymbol{\tau}_x)$ by the Cramér-Wold theorem (Billingsley (1995)) as $n \rightarrow \infty$. Let

$$\mathbf{H}_n(\mathbf{x}) = (\sqrt{n}(F_{1n}^*(x_1) - F_{1n}(x_1)), \dots, \sqrt{n}(F_{kn}^*(x_k) - F_{kn}(x_k)))^\top,$$

$$\mathbf{t}_n(\mathbf{x}) = \sqrt{n} \left(\frac{r}{n} - F_{1n}(x_1), \dots, \frac{r}{n} - F_{kn}(x_k) \right)^\top,$$

$\mathbf{y} = (y_1, \dots, y_k) \in \mathbb{R}^k$ and $\mathbf{x} = (x_1, \dots, x_k) \in \mathbb{R}^k$. Then by the result shown above, we know that $\mathbf{H}_n(\mathbf{x})$ converges in distribution to $N_k(0, \boldsymbol{\tau}_x)$ as $n \rightarrow \infty$. From this definition, note that we have

$$\Pr^*\{\mathbf{G}_n(q) \leq \mathbf{y}\} = \Pr^*\{\mathbf{H}_n(\mathbf{x}_n) \geq \mathbf{t}(\mathbf{x}_n)\}$$

where $\mathbf{x}_n = (y_1/\sqrt{n} + F_{1n}^{-1}(q), \dots, y_k/\sqrt{n} + F_{kn}^{-1}(q))^\top$. Let $\tilde{\mathbf{x}} = (F_1^{-1}(q), \dots, F_k^{-1}(q))^\top$ and $\mathbf{Y}_j$ in (A.5) defined on $\tilde{\mathbf{x}}$. Then its covariance matrix is $\boldsymbol{\tau}_{\tilde{\mathbf{x}}}$ where $\tau_{\tilde{\mathbf{x}}}(i, j) = F_{ij}(F_i^{-1}(q), F_j^{-1}(q)) - F_i(F_i^{-1}(q))F_j(F_j^{-1}(q)) = F_{ij}(\theta_i, \theta_j) - q^2 = \sigma_{ij}$ as defined in Theorem 2. Therefore, as $n \rightarrow \infty$, $\mathbf{H}_n(\mathbf{x}_n)$ converges in distribution to $N_k(0, \boldsymbol{\tau}_{\tilde{\mathbf{x}}})$ as $\boldsymbol{\tau}_{\mathbf{x}_n}$ converges to $\boldsymbol{\tau}_{\tilde{\mathbf{x}}}$ a.s. And as $n \rightarrow \infty$, $\mathbf{t}(\mathbf{x}_n)$ converges to $(-f_1(F_1^{-1}(q))y_1, \dots, -f_k(F_k^{-1}(q))y_k)^\top$ a.s. since marginal convergence implies joint convergence in this case. Now it follows that as $n \rightarrow \infty$,

$$\Pr^*\{\mathbf{G}_n(q) \leq \mathbf{y}\} \rightarrow \Phi_\Sigma(\mathbf{y}).$$

Thus as $n \rightarrow \infty$, we have

$$\sup_{\mathbf{y}} |\Pr^*\{\mathbf{G}_n(q) \leq \mathbf{y}\} - \Phi_\Sigma(\mathbf{y})| \rightarrow 0$$

by Theorem A.3. □

Lemma A.2. Let $F_{\mathbf{X}_k}$ be the cdf of $\mathbf{X}_k$ where $\mathbf{X}_k = (X_1, X_2, \dots, X_k)^\top$ is a random vector in $\mathbb{R}^k$; for $\mathbf{x}_k = (x_1, x_2, \dots, x_k)^\top$ and $x_{k+1} \in \mathbb{R}$, let

$$F_{\mathbf{X}_k}(\mathbf{x}_k) = \Pr\{\mathbf{X}_k \leq \mathbf{x}_k\}$$

and

$$F_{\mathbf{X}_k, X_{k+1}}(\mathbf{x}_{k+1}) = \Pr\{\mathbf{X}_k \leq \mathbf{x}_k, X_{k+1} \leq x_{k+1}\}.$$

Let $\mathbf{L} = (x_{11}, x_{21}, \dots, x_{k1})^\top$ and $\mathbf{U} = (x_{12}, x_{22}, \dots, x_{k2})^\top$. Then for $x_{i1} < x_{i2}$, $i = 1, 2, \dots, k$,

$$\Pr\{\mathbf{L} < \mathbf{X}_k \leq \mathbf{U}\} = \sum_{i_1=1}^2 \cdots \sum_{i_k=1}^2 (-1)^{i_1 + \dots + i_k} F_{\mathbf{X}_k}(x_{1i_1}, x_{2i_2}, \dots, x_{ki_k}). \quad (\text{A.6})$$

Proof For $k = 1$, $\Pr\{x_{11} < X_1 \leq x_{12}\} = F_{X_1}(x_{12}) - F_{X_1}(x_{11})$. If (A.6) holds for $k = t$, we show that it holds for $k = t + 1$ in the following:

$$\Pr\{(x_{11}, \dots, x_{t1})^\top < (X_1, \dots, X_t)^\top \leq (x_{12}, \dots, x_{t2})^\top, x_{t+1,1} < X_{t+1} \leq x_{t+1,2}\}$$

$$\begin{aligned}
&= \sum_{i_1=1}^2 \cdots \sum_{i_t=1}^2 (-1)^{i_1+\cdots+i_t} F_{\mathbf{X}_t, \mathbf{X}_{t+1}}(x_{1i_1}, \dots, x_{ti_t}, x_{t+1,2}) \\
&\quad - \sum_{i_1=1}^2 \cdots \sum_{i_t=1}^2 (-1)^{i_1+\cdots+i_t} F_{\mathbf{X}_t, \mathbf{X}_{t+1}}(x_{1i_1}, \dots, x_{ti_t}, x_{t+1,1}) \\
&= \sum_{i_1=1}^2 \cdots \sum_{i_t=1}^2 \sum_{i_{t+1}=1}^2 (-1)^{i_1+\cdots+i_t+i_{t+1}} F_{\mathbf{X}_t, \mathbf{X}_{t+1}}(x_{1i_1}, \dots, x_{ti_t}, x_{t+1, i_{t+1}}).
\end{aligned}$$

□

Note that when $\mathbf{X}$ is a continuous random vector, $\Pr\{\mathbf{L} < \mathbf{X} \leq \mathbf{U}\} = \Pr\{\mathbf{L} \leq \mathbf{X} \leq \mathbf{U}\}$. If $\mathbf{X}$ is discrete valued, then there exists a vector $\boldsymbol{\varepsilon} > 0$ small enough that $\Pr\{\mathbf{L} \leq \mathbf{X} \leq \mathbf{U}\} = \Pr\{\mathbf{L} - \boldsymbol{\varepsilon} < \mathbf{X} \leq \mathbf{U}\}$. Then we can apply (A.6) with $\mathbf{L}' = \mathbf{L} - \boldsymbol{\varepsilon}$ replacing $\mathbf{L}$. Therefore, for both cases, $\Pr\{\mathbf{L} \leq \mathbf{X} \leq \mathbf{U}\}$ can be written as a linear combination of a finite number of orthant probabilities.

A.2 Proof of Theorem 1

Part (a): It follows from Theorem A.2 for all $x \in \mathbb{R}^+$ that

$$|\Pr^*\{n^{1/2}(\bar{\mathbf{X}}_n^* - \bar{\mathbf{X}}_n) \leq x \cdot \mathbf{1}_k\} - \Gamma_{\Sigma_n}(x)| \leq C_b(k) \rho_{F_n} f_1(n) \quad (\text{A.7})$$

where $C_b(k)$ is a constant only depending on k , $\Sigma_n = \sum_{j=1}^n (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})^\top / n$, and $\rho_{F_n} < \infty$ when $E^*|\bar{\mathbf{X}}_n^* - \bar{\mathbf{X}}_n|^3 < \infty$ with large enough n . Let P_n^* denote the probability used in the bootstrap stopping variable as follows,

$$P_n^* = \Pr^*\{|\bar{\mathbf{X}}_n^* - \bar{\mathbf{X}}_n| \leq \delta \cdot \mathbf{1}_k\}.$$

Applying (A.7) with $x = \delta n^{1/2}$ we obtain

$$P_n^* = \Gamma_{\Sigma_n}(\delta n^{1/2}) + o(1) \quad (\text{A.8})$$

a.s. as $n \rightarrow \infty$. The inequalities $P_{N^*}^* \geq 1 - \alpha$ and $P_{N^*-1}^* < 1 - \alpha$ hold with the bootstrap stopping variable N^* given by

$$N^* = \inf\{n \geq n_0 : P_n^* \geq 1 - \alpha\}. \quad (\text{A.9})$$

Note that as $\delta \rightarrow 0$, we have $N^* \rightarrow \infty$ since $N^* \geq n_0$ and n_0 grows as $\delta \rightarrow 0$. It follows from (A.8), (A.9), and the fact that $N^* \rightarrow \infty$ as $\delta \rightarrow 0$ that

$$\begin{aligned}
P_{N^*}^* &= \Gamma_{\Sigma_{N^*}}(\delta N^{*1/2}) + o(1) \geq 1 - \alpha \\
P_{N^*-1}^* &= \Gamma_{\Sigma_{N^*-1}}(\delta (N^* - 1)^{1/2}) + o(1) < 1 - \alpha
\end{aligned}$$

a.s. as $\delta \rightarrow 0$; equivalently,

$$\delta N^{*1/2} \geq \Gamma_{\Sigma_{N^*}}^{-1}(1 - \alpha + o(1)), \quad (\text{A.10})$$

$$\delta (N^* - 1)^{1/2} < \Gamma_{\Sigma_{N^*-1}}^{-1}(1 - \alpha + o(1)). \quad (\text{A.11})$$

Since $\Sigma_{N^*} \rightarrow \Sigma$ as $\delta \rightarrow 0$, it follows from (A.10), (A.11) and the continuous mapping theorem (Billingsley (1995)) that we have

$$\delta N^{*1/2} \rightarrow a_{1-\alpha} = \Gamma_{\Sigma}^{-1}(1-\alpha) \quad a.s.$$

as $\delta \rightarrow 0$.

Part(b): Under our assumptions about $\mathbf{X}$, as $n \rightarrow \infty$, we have

$$\Pr \left\{ n^{1/2}(\bar{\mathbf{X}}_n - \Theta) \leq x \cdot \mathbf{1}_k \right\} \rightarrow \Phi_{\Sigma}(x \cdot \mathbf{1}_k)$$

from the multivariate central limit theorem. Let $Y_n = \max_i |\bar{X}_{in} - \theta_i|$. Then by continuous mapping theorem, we have

$$n^{1/2}Y_n \rightarrow Z \text{ in distribution} \quad (\text{A.12})$$

where $Z = \max\{|Z_1|, \dots, |Z_k|\}$ and the distribution of $\mathbf{Z} = (Z_1, \dots, Z_k)^{\top}$ is Φ_{Σ} . As $\delta \rightarrow 0$, we are interested in $\Pr\{Y_{N^*} \leq \delta\}$ and show that it can be stated in the form of Theorem 1 of Anscombe (1952) as follows:

$$\begin{aligned} \Pr\{Y_{N^*} \leq \delta\} &= \Pr \left\{ \sqrt{\frac{a_{1-\alpha}^2}{\delta^2}} \cdot Y_{N^*} \leq \sqrt{\frac{a_{1-\alpha}^2}{\delta^2}} \cdot \delta \right\} \\ &= \Pr \left\{ \frac{a_{1-\alpha}}{\delta} \cdot Y_{N^*} \leq a_{1-\alpha} \right\} \rightarrow \Gamma_{\Sigma}(a_{1-\alpha}) \end{aligned}$$

where $\lceil a_{1-\alpha}^2/\delta^2 \rceil = n_r$, $N^* = N_r$, and $w_{n_r} = \delta/a_{1-\alpha}$ in Anscombe's result. Then it follows from Lemma A.1 and (A.12) that

$$\Pr \left\{ |\bar{\mathbf{X}}_{N^*} - \Theta| \leq \delta \cdot \mathbf{1}_k \right\} \rightarrow 1 - \alpha \quad (\text{A.13})$$

as $\delta \downarrow 0$. □

A.3 Proof of Theorem 2

Part (a): Let P_n^* denote the probability used in the bootstrap stopping variable as follows,

$$P_n^* = \Pr^* \{ |\hat{\Theta}(\mathbf{X}_n^*) - \hat{\Theta}(\mathbf{X}_n)| \leq \delta \cdot \mathbf{1}_k \}.$$

where $\hat{\Theta}(\mathbf{X}_n^*)$ and $\hat{\Theta}(\mathbf{X}_n)$ are the sample q th quantiles based on $\mathbf{X}_n^*$ and $\mathbf{X}_n$, respectively. Analogous to the Proof of Theorem 1, it follows from Lemma A.2 and Proposition A.1 that

$$\sup_x |\Pr^*(\mathbf{G}_n(q) \leq x \cdot \mathbf{1}_k) - \Phi_{\Sigma}(x \cdot \mathbf{1}_k)| \rightarrow 0.$$

By letting $x = \delta n^{1/2}$, we obtain

$$P_n^* = \Gamma_{\Sigma}(\delta n^{1/2}) + o(1) \quad (\text{A.14})$$

a.s. as $n \rightarrow \infty$. From here on, the argument follows that of Theorem 1 (a). The inequalities $P_{N^*}^* \geq 1 - \alpha$ and $P_{N^*-1}^* < 1 - \alpha$ hold with the bootstrap stopping variable N^* given by

$$N^* = \inf\{n \geq n_0 : P_n^* \geq 1 - \alpha\}. \quad (\text{A.15})$$

It follows from (A.14), (A.15), and the fact that $N^* \rightarrow \infty$ as $\delta \rightarrow 0$ that

$$\begin{aligned} P_{N^*}^* &= \Gamma_\Sigma(\delta N^{*1/2}) + o(1) \geq 1 - \alpha \\ P_{N^*-1}^* &= \Gamma_\Sigma(\delta(N^* - 1)^{1/2}) + o(1) < 1 - \alpha \end{aligned}$$

a.s. as $\delta \rightarrow 0$; equivalently,

$$\delta N^{*1/2} \geq \Gamma_\Sigma^{-1}(1 - \alpha + o(1)), \quad (\text{A.16})$$

$$\delta(N^* - 1)^{1/2} < \Gamma_\Sigma^{-1}(1 - \alpha + o(1)). \quad (\text{A.17})$$

As $\delta \rightarrow 0$, it follows from (A.16), (A.17) that

$$\delta N^{*1/2} \rightarrow a_{1-\alpha} = \Gamma_\Sigma^{-1}(1 - \alpha) \quad a.s.$$

as $\delta \rightarrow 0$.

Part (b): As $n \rightarrow \infty$, we have

$$\Pr\{n^{1/2}(\widehat{\Theta}(\underline{\mathbf{X}}_n) - \Theta) \leq x \cdot \mathbf{1}_k\} \rightarrow \Phi_\Sigma(x \cdot \mathbf{1}_k),$$

as shown in Babu and Rao (1988). Let $Y_n = \max_i |\widehat{\theta}(X_{in}) - \theta_i|$. Then by the continuous mapping theorem, we have

$$n^{1/2}Y_n \rightarrow Z \text{ in distribution} \quad (\text{A.18})$$

where $Z = \max\{|Z_1|, \dots, |Z_k|\}$ and the distribution function of $\mathbf{Z} = (Z_1, \dots, Z_k)^\top$ is Φ_Σ . Then using the same argument as in the proof of Part (b) of Theorem 1, we have

$$\Pr\{Y_{N^*} \leq \delta\} \rightarrow \Gamma_\Sigma(a_{1-\alpha})$$

as $\delta \downarrow 0$. Then it follows from Lemma A.1 and (A.18) that

$$\Pr\left\{|\widehat{\Theta}(\underline{\mathbf{X}}_{N^*}) - \Theta| \leq \delta \cdot \mathbf{1}_k\right\} \rightarrow 1 - \alpha \quad (\text{A.19})$$

as $\delta \downarrow 0$. □

A.4 Proof of Corollary 2

Let $\mathbf{Y}_j = \mathbf{A}\mathbf{X}_j$ and P_n^* denote the probability used in the bootstrap stopping variable in (10) as follows,

$$P_n^* = \Pr^*\left\{|\mathbf{A}\bar{\mathbf{X}}_n^* - \mathbf{A}\bar{\mathbf{X}}_n| \leq \delta \cdot \mathbf{1}_{\binom{k}{2}}\right\} = \Pr^*\left\{|\bar{\mathbf{Y}}_n^* - \bar{\mathbf{Y}}_n| \leq \delta \cdot \mathbf{1}_{\binom{k}{2}}\right\}.$$

It follows from Theorem A.2 for all $x \in \mathbb{R}^+$ that

$$\left|\Pr^*\left\{n^{1/2}(\bar{\mathbf{Y}}_n^* - \bar{\mathbf{Y}}_n) \leq x \cdot \mathbf{1}_{\binom{k}{2}}\right\} - \Gamma_{\mathbf{A}\Sigma_n\mathbf{A}^\top}(x)\right| \leq C_b(k)\rho_{F_n}f_1(n) \quad (\text{A.20})$$

where $C_b(k)$ is a constant only depending on k , $\Sigma_n = \sum_{j=1}^n (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})^\top / n$, and $\rho_{F_n} < \infty$ when $E^* |\bar{\mathbf{Y}}_n^* - \bar{\mathbf{Y}}_n|^3 < \infty$ with large enough n . Applying (A.20) with $x = \delta n^{1/2}$ we obtain

$$P_n^* = \Gamma_{\mathbf{A}\Sigma_n\mathbf{A}^\top}(\delta n^{1/2}) + o(1)$$

a.s. as $n \rightarrow \infty$. The inequalities $P_{N_A^*}^* \geq 1 - \alpha$ and $P_{N_A^*-1}^* < 1 - \alpha$ hold with the bootstrap stopping variable N_A^* given by

$$N_A^* = \inf\{n \geq n_0 : P_n^* \geq 1 - \alpha\}.$$

Using the same arguments as in the proof of Theorem 1 (a) and (b), we have

$$\delta^2 N_A^* \rightarrow a_{1-\alpha}^2 \quad a.s.$$

where $a_{1-\alpha} = \Gamma_{\mathbf{A}\Sigma\mathbf{A}^\top}^{-1}(1 - \alpha)$ and

$$\Pr\left\{|\mathbf{A}\hat{\Theta}(\underline{\mathbf{X}}_{N_A^*}^*) - \mathbf{A}\Theta| \leq \delta \cdot \mathbf{1}_{\binom{k}{2}}\right\} \rightarrow 1 - \alpha$$

as $\delta \downarrow 0$. □

A.5 Proof of Corollary 3

Proof As shown in Proposition A.1, we have

$$\Pr^*\{n^{1/2}(\hat{\Theta}(\underline{\mathbf{X}}_n^*) - \hat{\Theta}(\underline{\mathbf{X}}_n)) \leq \mathbf{x}\} \rightarrow \Phi_\Sigma(\mathbf{x})$$

for all continuity points $\mathbf{x}$ as $n \rightarrow \infty$. It follows that

$$\Pr^*\{n^{1/2}(\mathbf{A}\hat{\Theta}(\underline{\mathbf{X}}_n^*) - \mathbf{A}\hat{\Theta}(\underline{\mathbf{X}}_n)) \leq \mathbf{x}\} \rightarrow \Phi_{\mathbf{A}\Sigma\mathbf{A}^\top}(\mathbf{x})$$

by the continuous mapping theorem. Therefore, as $n \rightarrow \infty$

$$\sup_{\mathbf{x}} |\Pr^*\{\mathbf{A}\hat{\Theta}(\underline{\mathbf{X}}_n^*) - \mathbf{A}\hat{\Theta}(\underline{\mathbf{X}}_n) \leq \mathbf{x}\} - \Phi_{\mathbf{A}\Sigma\mathbf{A}^\top}(\mathbf{x})| \rightarrow 0$$

by Theorem A.3. We redefine P_n^* as follows,

$$P_n^* = \Pr^*\left\{|\mathbf{A}\hat{\Theta}(\underline{\mathbf{X}}_n^*) - \mathbf{A}\hat{\Theta}(\underline{\mathbf{X}}_n)| \leq \delta \cdot \mathbf{1}_{\binom{k}{2}}\right\}.$$

Using the same arguments as in the proof of Theorem 2 (a) and (b), we have

$$\delta^2 N_A^* \rightarrow a_{1-\alpha}^2 \quad a.s.$$

where $a_{1-\alpha} = \Gamma_{\mathbf{A}\Sigma\mathbf{A}^\top}^{-1}(1 - \alpha)$ and

$$\Pr\left\{|\mathbf{A}\hat{\Theta}(\underline{\mathbf{X}}_{N_A^*}^*) - \mathbf{A}\Theta| \leq \delta \cdot \mathbf{1}_{\binom{k}{2}}\right\} \rightarrow 1 - \alpha$$

as $\delta \downarrow 0$. □

A.6 Results for Selecting the Best Mean or the Best 0.8 quantile without CRN

This section contains the experiment results of the bootstrap R&S procedure without CRN combined with the jump-ahead procedure, which have been referred in Section 6. The corresponding results with CRN are presented in Tables 1–4 of Section 6.

Table A.1: Empirical results for means from 100 macroreplications for half exponential and half normal distributions in the slippage configuration without CRN.

k	n_0	Ave N^*	PCS	Est Coverage	True Coverage	Ave Jumps
2(n)	50	10550	0.99	0.96	0.92	5.18
2(n)	100	10306	0.94	0.96	0.92	5.11
2(e)	50	10971	1.00	0.97	0.99	5.22
2(e)	100	11061	1.00	0.96	0.99	5.27
4(n)	50	20138	1.00	0.96	0.89	7.11
4(n)	100	20501	1.00	0.96	0.94	6.88
4(e)	50	20752	0.97	0.96	0.93	7.31
4(e)	100	20723	0.97	0.96	0.95	6.94
10(n)	50	37383	1.00	0.96	0.98	9.03
10(n)	100	37510	1.00	0.96	0.94	8.62
10(e)	50	37745	0.99	0.96	0.96	8.91
10(e)	100	37717	1.00	0.96	0.95	8.58
20(n)	50	50217	1.00	0.96	0.93	10.2
20(n)	100	50029	1.00	0.96	0.97	9.21
20(e)	100	50257	1.00	0.96	0.97	9.36
20(e)	50	50229	1.00	0.96	0.95	9.99

Table A.2: Empirical results for means from 100 macroreplications for half exponential and half normal distributions in the monotone decreasing means configuration without CRN.

k	n_0	Ave N^*	PCS	Est Coverage	True Coverage	Ave Jumps
4(n)	50	21276	1.00	0.96	0.92	7.54
4(n)	100	21239	1.00	0.96	0.96	6.95
4(e)	50	21974	1.00	0.96	0.95	7.38
4(e)	100	21760	1.00	0.96	0.94	6.76
10(n)	50	43913	1.00	0.96	0.98	0.35
10(n)	100	44056	1.00	0.96	0.97	8.68
10(e)	50	45917	1.00	0.96	0.95	9.38
10(e)	100	45349	1.00	0.96	0.98	8.51
20(n)	50	72426	1.00	0.96	0.96	10.43
20(n)	100	72471	1.00	0.96	0.90	10.03
20(e)	50	74808	1.00	0.96	0.95	10.72
20(e)	100	73802	1.00	0.96	0.98	9.69

Table A.3: Empirical results for means from 100 macroreplications for negative binomial distributions in the slippage configuration without CRN.

k	θ	Average N^*	PCS	Est. Coverage	True Coverage	Ave Jumps
2	1	1729	0.98	0.97	0.97	4.47
2	10	15842	0.99	0.96	0.97	5.24
4	1	2757	0.96	0.96	0.93	5.36
4	10	26980	1.00	0.96	0.97	6.29
10	1	4064	0.99	0.96	0.95	6.09
10	10	40837	0.99	0.96	0.95	7.72
20	1	5125	1.00	0.96	0.93	6.80
20	10	51189	1.00	0.96	0.97	8.98

Table A.4: Empirical results for quantiles ($q = 0.8$) from 100 macroreplications for half exponential and half normal distributions in the slippage configuration without CRN.

k	n_0	Ave N^*	PCS	Est Coverage	True Coverage	Ave Jumps
2(n)	50	22355	0.97	0.97	0.92	5.72
2(n)	100	23156	0.99	0.97	0.95	5.01
2(e)	50	22304	0.97	0.96	0.94	5.82
2(e)	100	24103	0.96	0.96	0.95	5.14
4(n)	50	43351	1.00	0.96	0.95	7.72
4(n)	100	42708	1.00	0.96	0.95	7.31
4(e)	50	44194	0.95	0.96	0.93	7.84
4(e)	100	43824	0.99	0.96	0.99	7.38
10(n)	50	80487	1.00	0.96	0.97	9.56
10(n)	100	79610	1.00	0.96	0.98	9.14
10(e)	50	79611	1.00	0.96	0.97	9.48
10(e)	100	79276	1.00	0.96	0.97	8.84
20(n)	50	107071	1.00	0.96	0.98	10.61
20(n)	100	106991	1.00	0.96	0.97	10.07
20(e)	50	108670	1.00	0.96	0.98	10.93
20(e)	100	107608	1.00	0.96	0.94	9.98