A. Statistical Validity Proof for the M_2 Procedure

In order to prove Theorem 1, we need the following two lemmas.

LEMMA A1. *(Jennison et al. 1980) Consider a continuous Brownian Motion process* $\{W(t,\Delta), t \geq 0\}$ *with drift* $\Delta > 0$*, and a continuation region* $H(t)$ *bounded by* $(-h(t), h(t))$*, where* $h(t)$ *is a non-negative function for* $t \geq 0$*. A discrete process is obtained by observing* $W(t,\Delta)$ *over a random, increasing sequence of times* $\{t_i, i = 1, 2, \ldots\}$ *taking values in a given countable set. Let* $T_C = inf\{t : W(t,\Delta) \notin H(t)\}$ *and* $T_D = inf\{t_i : W(t_i,\Delta) \notin H(t_i)\}$*, and assume that* $T_D < \infty$ *almost surely (note that* $T_D \geq T_C$*). The error probabilities are*

$$P\{\varepsilon_C\} \equiv P\{W(T_C,\Delta) \leq -h(T_C)\} = P\{W(T_C,\Delta) < 0\},$$
$$P\{\varepsilon_D\} \equiv P\{W(T_D,\Delta) \leq -h(T_D)\} = P\{W(T_D,\Delta) < 0\}.$$

Consider an outcome $\{(b(t); t \geq 0), \{t_i\}\}$*, where* $b(t)$ *is the path of a Brownian motion. Assume that the conditional distribution of* $\{t_i\}$ *given* $W(t,\Delta) = b(t), t \geq 0$*, is the same as the conditional distribution of* $\{t_i\}$ *given* $W(t,\Delta) = -b(t), t \geq 0$*. Under these conditions,*

$$P\{\varepsilon_D\} \leq P\{\varepsilon_C\}.$$

When each observation is IID normally distributed, the partial sums of the differences behave like Brownian motion process with drift at each integer point. However, Brownian motion with drift is only an approximation for our discrete process. Lemma A1 states that under very general conditions, the probability of incorrect selection does not increase when the Brownian motion process is observed at discrete times compared to the case where the process is observed continuously. Thus, procedures designed for a continuous Brownian motion process with a drift provide an upper bound on the probability of incorrect selection for a discrete process.

LEMMA A2. *(Ross 1996) Suppose that* $W(t,\Delta_i)$ *is a standard Brownian motion process on* $[0,\infty)$ *with drift* Δ_i*. Let* $T_a = \inf\{t : W(t,\Delta_i) = a\}$ *and* $T_b = \inf\{t : W(t,\Delta_i) = b\}$ *be the stopping times when the Brownian motion hits the parallel lines* a *and* b *(*$a < 0 < b$*), respectively. For* $\Delta_1 < 0$*, the following results hold:*

$$P_{\Delta_1}\{T_b < T_a\} = \frac{1 - e^{-2\Delta_1 a}}{e^{-2\Delta_1 b} - e^{-2\Delta_1 a}},$$
$$P_{\Delta_1}\{T_b < \infty\} = \lim_{a \to -\infty} P_{\Delta_1}\{T_b < T_a\} = e^{2\Delta_1 b}.$$

The opposite results hold for $\Delta_2 > 0$:

$$P_{\Delta_2}\{T_a < T_b\} = \frac{1-e^{-2\Delta_2 b}}{e^{-2\Delta_2 a} - e^{-2\Delta_2 b}},$$
$$P_{\Delta_2}\{T_a < \infty\} = \lim_{b\to\infty} P_{\Delta_2}\{T_a < T_b\} = e^{2\Delta_2 a}.$$

Proof of Theorem 1

This proof is similar to the proof of the KN procedure (Kim and Nelson 2001) in its use of the Brownian motion approximation approach. In order to prove that Equations (3)-(5) hold, we need to show that the following holds:

$$\mathrm{P}_{\boldsymbol{\mu}\in\Omega_0}\{\text{Accept } H_1^{\mu}\} \le \alpha_K/2, \text{ where}\Omega\ _0(\delta_\mu) = \{\boldsymbol{\mu} : \mu_i - \mu_j = 0\}; \tag{A.1}$$

$$\mathrm{P}_{\boldsymbol{\mu}\in\Omega_0}\{\text{Accept } H_2^{\mu}\} \le \alpha_K/2, \text{ where}\Omega\ _0(\delta_\mu) = \{\boldsymbol{\mu} : \mu_i - \mu_j = 0\}; \tag{A.2}$$

$$\mathrm{P}_{\boldsymbol{\mu}\in\Omega_1}\{\text{Reject } H_1^{\mu}\} \le \alpha_K, \text{ where}\Omega\ _1(\delta_\mu) = \{\boldsymbol{\mu} : \mu_i - \mu_j \ge \delta_\mu\}; \tag{A.3}$$

$$\mathrm{P}_{\boldsymbol{\mu}\in\Omega_2}\{\text{Reject } H_2^{\mu}\} \le \alpha_K, \text{ where}\Omega\ _2(\delta_\mu) = \{\boldsymbol{\mu} : \mu_i - \mu_j \le -\delta_\mu\}. \tag{A.4}$$

Equations (A.1) and (A.2) imply Equation (3). Equations (A.3)-(A.4) imply Equations (4)-(5), respectively.

Let

$$T = \min\Big\{ r \in \{n_0, n_0+1, \ldots\} : \Big|\sum_{\ell=1}^{r}(X_{i\ell} - X_{j\ell})\Big| \ge \mathrm{OB}_{ij}^{\mu}(r) \text{ or } \Big|\sum_{\ell=1}^{r}(X_{i\ell} - X_{j\ell})\Big| \le \mathrm{IB}_{ij}^{\mu}(r)\Big\}.$$

Note that T is the stage at which the M_2 procedure terminates by leaving the continuation region.

First, we show that Equation (A.1) holds.

$$\begin{aligned}
&\mathrm{P}_{\boldsymbol{\mu}\in\Omega_0}\{\text{Accept } H_1^{\mu}\} = \mathrm{P}_{\boldsymbol{\mu}\in\Omega_0}\Big\{\sum_{\ell=1}^{T}(X_{i\ell} - X_{j\ell}) \ge \mathrm{OB}_{ij}^{\mu}(T)\Big\}\\
&= \mathrm{P}_{\boldsymbol{\mu}\in\Omega_0}\Big\{\sum_{\ell=1}^{T}(X_{i\ell} - X_{j\ell}) \ge Td + \frac{S_{ij}^2(n_0)a(\alpha_K/2)}{4d}\Big\}\\
&= \mathrm{P}_{\boldsymbol{\mu}\in\Omega_0}\Big\{\sum_{\ell=1}^{T}\frac{X_{i\ell} - X_{j\ell}}{\sigma_{ij}} \ge \frac{Td}{\sigma_{ij}} + \frac{S_{ij}^2(n_0)a(\alpha_K/2)}{4d\sigma_{ij}}\Big\}\\
&= \mathrm{E}\Big[\mathrm{P}_{\boldsymbol{\mu}\in\Omega_0}\Big\{\sum_{\ell=1}^{T}\frac{X_{i\ell} - X_{j\ell}}{\sigma_{ij}} \ge \frac{Td}{\sigma_{ij}} + \frac{S_{ij}^2(n_0)a(\alpha_K/2)}{4d\sigma_{ij}}\Big| S_{ij}^2(n_0)\Big\}\Big]\\
&= \mathrm{E}\Big[\mathrm{P}_{\boldsymbol{\mu}\in\Omega_0}\Big\{\sum_{\ell=1}^{T}\frac{X_{i\ell} - X_{j\ell}}{\sigma_{ij}} - \frac{Td}{\sigma_{ij}} \ge \frac{S_{ij}^2(n_0)a(\alpha_K/2)}{4d\sigma_{ij}}\Big| S_{ij}^2(n_0)\Big\}\Big] \qquad \text{(A.5)}
\end{aligned}$$

where $\sigma_{ij}^2 = \text{Var}(X_{i\ell} - X_{j\ell})$.

Let

$$\Delta_1 = \frac{-d}{\sigma_{ij}} \text{ and } b = \frac{S_{ij}^2(n_0)a(\alpha_K/2)}{4d\sigma_{ij}}.$$

Note that when $\boldsymbol{\mu} \in \Omega_0(\delta_\mu)$, $(X_{i\ell} - X_{j\ell})/\sigma_{ij}$ are IID normal with a mean of zero and a variance of one. Hence, $\sum_{\ell=1}^{T}(X_{i\ell} - X_{j\ell})/\sigma_{ij} - Td/\sigma_{ij}$ behaves like a standard Brownian motion, $W(T, \Delta_1)$, with drift$\Delta\ _1 = -d/\sigma_{ij}$ at integer points.

Let $T_b = \inf\{t : W(t, \Delta_1) = b\}$. Following from Equation (A.5),

$$\begin{aligned}
&\text{E}\Big[\text{P}_{\boldsymbol{\mu}\in\Omega_0}\Big\{\sum_{\ell=1}^{T}\frac{X_{i\ell}-X_{j\ell}}{\sigma_{ij}} - \frac{Td}{\sigma_{ij}} \geq \frac{S_{ij}^2(n_0)a(\alpha_K/2)}{4d\sigma_{ij}}\Big|S_{ij}^2(n_0)\Big\}\Big] \\
&\quad \leq \text{E}\Big[\text{P}\{W(T,\Delta_1) \geq b | S_{ij}^2(n_0)\}\Big] && \text{(A.6)} \\
&\quad \leq \text{E}\Big[\text{P}_{\Delta_1}\{T_b < \infty | S_{ij}^2(n_0)\}\Big] && \text{(A.7)} \\
&\quad = \text{E}[\exp(2\Delta_1 b)] && \text{(A.8)} \\
&\quad = \text{E}\Big[\exp\Big(2\frac{-d}{\sigma_{ij}}\frac{S_{ij}^2(n_0)a(\alpha_K/2)}{4d\sigma_{ij}}\Big)\Big] \\
&\quad = \text{E}\Big[\exp\Big(\frac{S_{ij}^2(n_0)(n_0-1)}{\sigma_{ij}^2}\frac{-\Big((2/\alpha_K)^{\frac{2}{n_0-1}}-1\Big)}{2}\Big)\Big] && \text{(A.9)} \\
&\quad = \Big(1 - 2\frac{-\Big((2/\alpha_K)^{\frac{2}{n_0-1}}-1\Big)}{2}\Big)^{-\frac{n_0-1}{2}} \\
&\quad = \Big(1 + (2/\alpha_K)^{\frac{2}{n_0-1}} - 1\Big)^{-\frac{n_0-1}{2}} \\
&\quad = \frac{\alpha_K}{2}.
\end{aligned}$$

The inequality in Equation (A.6) is by Lemma A1. The inequality in Equation (A.7) is because the Brownian motion process may hit IB_{ij}^{μ} before it hits line b. Equation (A.8) is by Lemma A2. In Equation (A.9), $S_{ij}^2(n_0)(n_0-1)/\sigma_{ij}^2$ is chi-squared distributed with $n_0 - 1$ degrees of freedom. The expectation in Equation (A.9) is equivalent to the moment generating function of a chi-squared random variable which is $\text{E}[\exp(\chi_\nu^2 t)] = (1-2t)^{-\nu/2}$ for $t < 1/2$. Hence, we have shown that Equation (A.1) holds. By a similar derivation, we can show that Equation (A.2) holds.

Next, we will show that Equation (A.3) holds.

$$\text{P}_{\boldsymbol{\mu}\in\Omega_1}\{\text{Reject } H_1^{\mu}\}$$

$$\leq \mathrm{P}_{\boldsymbol{\mu}\in\Omega_1}\Big\{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell}) \leq \mathrm{IB}_{ij}^{\mu}(T)\Big\} \tag{A.10}$$

$$= \mathrm{P}_{\boldsymbol{\mu}\in\Omega_1}\Big\{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell}) \leq T(\delta_\mu - d) - \frac{S_{ij}^2(n_0)a(\alpha_K)}{4d}\Big\}$$

$$\leq \mathrm{P}_{\mathrm{LFC}^1}\Big\{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell}) \leq T(\delta_\mu - d) - \frac{S_{ij}^2(n_0)a(\alpha_K)}{4d}\Big\} \tag{A.11}$$

$$= \mathrm{P}_{\mathrm{LFC}^1}\Big\{\sum_{\ell=1}^{T}\frac{X_{i\ell}-X_{j\ell}}{\sigma_{ij}} \leq \frac{T(\delta_\mu - d)}{\sigma_{ij}} - \frac{S_{ij}^2(n_0)a(\alpha_K)}{4d\sigma_{ij}}\Big\}$$

$$= \mathrm{E}\Big[\mathrm{P}_{\mathrm{LFC}^1}\Big\{\sum_{\ell=1}^{T}\frac{X_{i\ell}-X_{j\ell}}{\sigma_{ij}} \leq \frac{T(\delta_\mu - d)}{\sigma_{ij}} - \frac{S_{ij}^2(n_0)a(\alpha_K)}{4d\sigma_{ij}}\Big| S_{ij}^2(n_0)\Big\}\Big]$$

$$= \mathrm{E}\Big[\mathrm{P}_{\mathrm{LFC}^1}\Big\{\sum_{\ell=1}^{T}\frac{X_{i\ell}-X_{j\ell}}{\sigma_{ij}} - \frac{T(\delta_\mu - d)}{\sigma_{ij}} \leq \frac{-S_{ij}^2(n_0)a(\alpha_K)}{4d\sigma_{ij}}\Big| S_{ij}^2(n_0)\Big\}\Big]. \tag{A.12}$$

In Equation (A.10), the partial sum moves to the lower section of the continuation region which includes a possible crossing of IB_{ij}^{μ}, $-\mathrm{IB}_{ij}^{\mu}$ or $-\mathrm{OB}_{ij}^{\mu}$. The inequality is due to the fact that there is not a condition on T as given in the Computation step of the M_2 procedure. The inequality in Equation (A.11) is due to the assumption of least favorable configuration (LFC^1), $\mu_i - \mu_j = \delta_\mu$, in $\Omega_1(\delta_\mu)$.

Let

$$\Delta_2 = \frac{d}{\sigma_{ij}} \text{ and } a = \frac{-S_{ij}^2(n_0)a(\alpha_K)}{4d\sigma_{ij}}.$$

Note that for the LFC^1, $(X_{i\ell} - X_{j\ell})/\sigma_{ij}$ are IID normal with a mean of δ_μ/σ_{ij} and a variance of one. Hence, $\sum_{\ell=1}^{T}(X_{i\ell} - X_{j\ell})/\sigma_{ij} - T(\delta_\mu - d)/\sigma_{ij}$ behaves like a standard Brownian motion, $W(T,\Delta_2)$, with drift$\Delta_{\ 2} = d/\sigma_{ij}$ at integer points.

Let $T_a = \inf\{t : W(t,\Delta_2) = a\}$. Following from Equation (A.12),

$$\mathrm{E}\Big[\mathrm{P}_{\mathrm{LFC}^1}\Big\{\sum_{\ell=1}^{T}\frac{X_{i\ell}-X_{j\ell}}{\sigma_{ij}} - \frac{T(\delta_\mu - d)}{\sigma_{ij}} \leq \frac{-S_{ij}^2(n_0)a(\alpha_K)}{4d\sigma_{ij}}\Big| S_{ij}^2(n_0)\Big\}\Big]$$

$$\leq \mathrm{E}\Big[\mathrm{P}\{W(T,\Delta_2) \leq a | S_{ij}^2(n_0)\}\Big] \tag{A.13}$$

$$\leq \mathrm{E}\Big[\mathrm{P}_{\Delta_2}\{T_a < \infty | S_{ij}^2(n_0)\}\Big] \tag{A.14}$$

$$= \mathrm{E}[\exp(2\Delta_2 a)] \tag{A.15}$$

$$= \mathrm{E}\Big[\exp\Big(2\frac{d}{\sigma_{ij}}\frac{-S_{ij}^2(n_0)a(\alpha_K)}{4d\sigma_{ij}}\Big)\Big]$$

$$= \mathrm{E}\Big[\exp\Big(\frac{S_{ij}^2(n_0)(n_0-1)}{\sigma_{ij}^2}\frac{-\Big((1/\alpha_K)^{\frac{2}{n_0-1}}-1\Big)}{2}\Big)\Big] \tag{A.16}$$

$$= \Big(1-2\frac{-\Big((1/\alpha_K)^{\frac{2}{n_0-1}}-1\Big)}{2}\Big)^{-\frac{n_0-1}{2}}$$

$$= \Big(1+(1/\alpha_K)^{\frac{2}{n_0-1}}-1\Big)^{-\frac{n_0-1}{2}}$$

$$= \alpha_K.$$

The inequality in Equation (A.13) is by Lemma A1. The inequality in Equation (A.14) is because the Brownian motion process may hit OB_{ij}^{μ} before it hits line a. Equation (A.15) is by Lemma A2. In Equation (A.16), $S_{ij}^2(n_0)(n_0-1)/\sigma_{ij}^2$ is chi-squared distributed with n_0-1 degrees of freedom. The expectation in Equation (A.16) is equivalent to the moment generating function of a chi-squared random variable which is $\mathrm{E}[\exp(\chi_\nu^2 t)] = (1-2t)^{-\nu/2}$ for $t<1/2$. Hence, we have shown that Equation (A.3) holds. By a similar derivation, we can show that Equation (A.4) holds.

Proof of Corollary 1

In order to prove that Equation (6) holds, we will show that the following holds

$$\mathrm{P}_{\boldsymbol{\mu}\in\Omega_3}(\text{Accept } H_2^{\mu}) \le \alpha_K/2, \ \text{where}\Omega\ _3(\delta_\mu) = \{\boldsymbol{\mu}: 0<\mu_i-\mu_j<\delta_\mu\}.$$

Then

$$\mathrm{P}_{\boldsymbol{\mu}\in\Omega_3}(\text{Accept } H_2^{\mu})$$

$$\le \mathrm{P}_{\boldsymbol{\mu}\in\Omega_0}(\text{Accept } H_2^{\mu}) \tag{A.17}$$

$$\le \alpha_K/2. \tag{A.18}$$

The inequality in Equation (A.17) holds because the probability of accepting H_2^{μ} increases when $\boldsymbol{\mu}\in\Omega_0(\delta_\mu)$. The inequality in Equation (A.18) holds by Equation (A.2).

Similarly, we can show that Equation (7) holds.

B. M_K Procedure

The algorithm of the M_K procedure is described in detail below.

1. Setup: Place all K systems in the set of initial systems, i.e., $S_{\mathrm{IN}} = \{1,2,\dots,K\}$. Define S_{EQ} as the set of pairs of equivalent systems, and set $S_{\mathrm{EQ}} = \emptyset$. Define S_{DE} as the set of decided systems whose all pairwise tests with the systems in S_{IN} reached to a decision, and set $S_{\mathrm{DE}} = \emptyset$. Specify the confidence level for the M_K procedure, $1-\alpha$, and the mean-IZ parameter, δ_μ ($\delta_\mu > 0$). Set parameter d to $3\delta_\mu/8$. Set α_K to $\alpha/[K(K-1)/2]$. Specify the initial sample size n_0 ($n_0 > 2$).

2. Initialization: Obtain n_0 observations from every system i, $i \in S_{\mathrm{IN}}$. For system pair (i,j), $i,j \in S_{\mathrm{IN}}$, $i<j$, compute the sample variance of differences of observations

$$S_{ij}^2(n_0) = \frac{1}{n_0-1}\sum_{\ell=1}^{n_0}\Big(X_{i\ell} - X_{j\ell} - (\bar{X}_i(n_0) - \bar{X}_j(n_0))\Big)^2, \text{ where } \bar{X}_i(n_0) = \frac{1}{n_0}\sum_{\ell=1}^{n_0} X_{i\ell}.$$

Set $a(\eta) = ((1/\eta)^{2/(n_0-1)} - 1)(n_0-1)$. Set the number of stages, r, to n_0.

3. Computation:

For system pair (i,j), $i,j \in S_{\mathrm{IN}}$, $(i,j) \notin S_{\mathrm{EQ}}$, $i<j$, compute the inner and outer bounds

$$\mathrm{IB}_{ij}^{\mu}(r) = r(\delta_\mu - d) - \frac{S_{ij}^2(n_0)a(\alpha_K)}{4d} \text{ when } r > \frac{S_{ij}^2(n_0)a(\alpha_K)}{4d(\delta_\mu - d)};$$

$$\mathrm{OB}_{ij}^{\mu}(r) = rd + \frac{S_{ij}^2(n_0)a(\alpha_K/2)}{4d}.$$

4. Decision:

For system pair (i,j), $i,j \in S_{\mathrm{IN}}$, $(i,j) \notin S_{\mathrm{EQ}}$, $i<j$:

i) If $-\mathrm{IB}_{ij}^{\mu}(r) \le \sum_{\ell=1}^{r}(X_{i\ell} - X_{j\ell}) \le \mathrm{IB}_{ij}^{\mu}(r)$, then accept $H_0^{\mu}: \mu_i = \mu_j$.

ii) Else if $\sum_{\ell=1}^{r}(X_{i\ell} - X_{j\ell}) \ge \mathrm{OB}_{ij}^{\mu}(r)$, then accept $H_1^{\mu}: \mu_i > \mu_j$.

iii) Else if $\sum_{\ell=1}^{r}(X_{i\ell} - X_{j\ell}) \le -\mathrm{OB}_{ij}^{\mu}(r)$, then accept $H_2^{\mu}: \mu_i < \mu_j$.

5. Consolidation:

For system pair (i,j), $i,j \in S_{\mathrm{IN}}$, $(i,j) \notin S_{\mathrm{EQ}}$, $i<j$:

i) If $\mu_i = \mu_j$ decision is made, then declare systems i and j equivalent and set $S_{\mathrm{EQ}} = S_{\mathrm{EQ}} \cup \{(i,j)\}$.

ii) If $\mu_i > \mu_j$ decision is made, then eliminate system j by setting $S_{\mathrm{IN}} = S_{\mathrm{IN}} \backslash \{j\}$. Delete every system j pair in S_{EQ}.

iii) If $\mu_i < \mu_j$ decision is made, then eliminate system i by setting $S_{\mathrm{IN}} = S_{\mathrm{IN}} \backslash \{i\}$. Delete every system i pair in S_{EQ}.

6. Termination: If $|S_{\mathrm{IN}}| = 1$ or S_{EQ} includes every pair (i,j) where $i,j \in S_{\mathrm{IN}}, i<j$, stop sampling. Return S_{IN} as an M-efficient set.

If the procedure does not terminate, perform the following. For $i \in S_{\mathrm{IN}}$, if every pair (i,j), $j \in S_{\mathrm{IN}}, i \ne j$, is in S_{EQ}, add system i to S_{DE}. Obtain one more observation from every system i, $i \in S_{\mathrm{IN}} \backslash S_{\mathrm{DE}}$, and set $r = r+1$. Go back to the Computation step. $\square$

C. Statistical Validity Proof for the M_2^+ Procedure

In order to prove Theorem 2, we need the following two lemmas.

LEMMA C1. *(Billingsley 1968) If $Y_r \Rightarrow Y$ and $Z_r \to a$ in probability where a is a constant, then $(Y_r, Z_r) \Rightarrow (Y, a)$, where $\Rightarrow$ denotes convergence in distribution.*

LEMMA C2. *(Billingsley 1968) Let $X_{i\ell}, \ell = 1, 2, \ldots,$ be a sequence of IID random variables with finite mean μ_i and finite, positive variance σ_i^2. Define the standardized partial sum for system i as*

$$C_i(t,r) \equiv \frac{\sum_{\ell=1}^{\lfloor rt \rfloor} X_{i\ell} - rt\mu_i}{\sigma_i\sqrt{r}}, \quad 0 \le t \le 1 \text{ and } r \in Z^+,$$

where $\lfloor \cdot \rfloor$ indicates truncation of any fractional part. The probability distribution of $C_i(t,r)$ over $D[0,1]$, the space of functions that are right-continuous and have left-hand limits, converges to that of a standard Brownian motion process, $W(t,0)$, as r increases; i.e., $C_i(\cdot, r) \Rightarrow W(\cdot, 0)$ as $r \to \infty$, where $\Rightarrow$ denotes convergence in distribution.

Lemma C2 is basically the Functional Central Limit Theorem.

Proof of Theorem 2

This proof is similar to the proofs of the KN++ (Kim and Nelson 2006) and WK++ procedures (Wang and Kim 2012) in its use of the Brownian motion approximation approach. In order to prove that Equations (8)-(10) hold, we need to show that the following holds:

$$\limsup_{\delta_\mu \to 0} \sup_{\boldsymbol{\mu} \in \Omega_0(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}(\text{Accept } H_1^\mu) \le \alpha_K/2, \text{where} \Omega_{\ 0}(\delta_\mu) = \{\boldsymbol{\mu} : \mu_i - \mu_j = 0\}; \tag{C.1}$$

$$\limsup_{\delta_\mu \to 0} \sup_{\boldsymbol{\mu} \in \Omega_0(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}(\text{Accept } H_2^\mu) \le \alpha_K/2, \text{where} \Omega_{\ 0}(\delta_\mu) = \{\boldsymbol{\mu} : \mu_i - \mu_j = 0\}; \tag{C.2}$$

$$\limsup_{\delta_\mu \to 0} \sup_{\boldsymbol{\mu} \in \Omega_1(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}(\text{Reject } H_1^\mu) \le \alpha_K, \text{where} \Omega_{\ 1}(\delta_\mu) = \{\boldsymbol{\mu} : \mu_i - \mu_j \ge \delta_\mu\}; \tag{C.3}$$

$$\limsup_{\delta_\mu \to 0} \sup_{\boldsymbol{\mu} \in \Omega_2(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}(\text{Reject } H_2^\mu) \le \alpha_K, \text{where} \Omega_{\ 2}(\delta_\mu) = \{\boldsymbol{\mu} : \mu_i - \mu_j \le -\delta_\mu\}. \tag{C.4}$$

Equations (C.1) and (C.2) imply Equation (8). Equations (C.3)-(C.4) imply Equations (9)-(10), respectively.

Let

$$T \equiv T(\delta_\mu) = \min\Big\{ r \in \{n_0, n_0+1, \ldots\} : \Big| \sum_{\ell=1}^{r} (X_{i\ell} - X_{j\ell}) \Big| \ge \mathrm{OB}_{ij}^{\mu+}(r) \text{ or } \Big| \sum_{\ell=1}^{r} (X_{i\ell} - X_{j\ell}) \Big| \le \mathrm{IB}_{ij}^{\mu+}(r) \Big\}.$$

Thus, $T(\delta_\mu)$ is the stage at which the M_2^+ procedure terminates by leaving the continuation region defined by δ_μ. Let

$$N_{ij}^+(r) = \left\lfloor \frac{S_{ij}^2(r)(g(\alpha_K)+g(\alpha_K/2))}{4d(\delta_\mu - 2d)} \right\rfloor.$$

When $d = 3\delta_\mu/8$,

$$N_{ij}^+(r) = \left\lfloor \frac{8S_{ij}^2(r)(g(\alpha_K)+g(\alpha_K/2))}{3\delta_\mu^2} \right\rfloor.$$

Below we show that $T(\delta_\mu) \to \infty$ as $\delta_\mu \to 0$ with probability 1, which is a fundamental requirement for the convergence of the variance estimator. In the rest of the proof, the focus will be on the LFCs.

When the LFC, $\mu_i - \mu_j = \delta_\mu$, under H_1^μ, is true, the output process $\{X_{i\ell} - X_{j\ell}, \ell = 1, 2, \ldots\}$ can be represented as $\{Z_\ell + \delta_\mu, \ell = 1, 2, \ldots\}$, where $\{Z_\ell, \ell = 1, 2, \ldots\}$ are IID $N(0, \sigma_{ij}^2)$ according to Assumption 1.

First, consider a sample path moving in the direction of $\mathrm{OB}_{ij}^{\mu+}(r)$ for which $T^* \equiv \limsup_{\delta_\mu \to 0} T(\delta_\mu) < \infty$. For this to occur depending on the sample path there must exist $\delta_\mu^* > 0$ such that for all $\delta_\mu \le \delta_\mu^*$,

$$\sum_{\ell=1}^{T^*}(X_{i\ell} - X_{j\ell}) = \sum_{\ell=1}^{T^*}(Z_\ell + \delta_\mu) \ge T^* d + \frac{S_{ij}^2(T^*)g(\alpha_K/2)}{4d}. \tag{C.5}$$

When $d = 3\delta_\mu/8$, the above equation can be rewritten as follows:

$$\sum_{\ell=1}^{T^*} Z_\ell + \frac{5}{8}T^*\delta_\mu \ge \frac{2S_{ij}^2(T^*)g(\alpha_K/2)}{3\delta_\mu}.$$

Note that when T^* is finite, $\sum_{\ell=1}^{T^*} Z_\ell$ is finite with probability 1. As $\delta_\mu \to 0$, $(5/8)T^*\delta_\mu$ converges to zero, but $2S_{ij}^2(T^*)g(\alpha_K/2)/(3\delta_\mu)$ converges to infinity with probability 1. Hence, Equation (C.5) occurs with probability zero.

Second, consider a sample path moving in the direction of $\mathrm{IB}_{ij}^{\mu+}(r)$ for which $T^* \equiv \limsup_{\delta_\mu \to 0} T(\delta_\mu) < \infty$. For this to occur depending on the sample path there must exist $\delta_\mu^* > 0$ such that for all $\delta_\mu \le \delta_\mu^*$,

$$\sum_{\ell=1}^{T^*}(X_{i\ell} - X_{j\ell}) = \sum_{\ell=1}^{T^*}(Z_\ell + \delta_\mu) \le T^*(\delta_\mu - d) - \frac{S_{ij}^2(T^*)g(\alpha_K)}{4d}. \tag{C.6}$$

When $d = 3\delta_\mu/8$, the above equation can be rewritten as follows:

$$\sum_{\ell=1}^{T^*} Z_\ell + \frac{3}{8}T^*\delta_\mu \le -\frac{2S_{ij}^2(T^*)g(\alpha_K)}{3\delta_\mu}.$$

Note that when T^* is finite, $\sum_{\ell=1}^{T^*} Z_\ell$ is finite with probability 1. As $\delta_\mu \to 0$, $(3/8)T^*\delta_\mu$ converges to zero, but $-2S_{ij}^2(T^*)g(\alpha_K)/(3\delta_\mu)$ converges to negative infinity with probability 1. Hence, Equation (C.6) occurs with probability zero.

We can prove a similar contradiction result for the remaining two boundaries of the continuation region, $-\mathrm{OB}_{ij}^{\mu+}(r)$ and $-\mathrm{IB}_{ij}^{\mu+}(r)$. Since the contradiction holds for all four boundaries of the continuation region, these results altogether imply that $T(\delta_\mu)\to\infty$ with probability 1 as $\delta_\mu \to 0$ when the LFC, $\mu_i-\mu_j=\delta_\mu$, under H_1^μ is true.

A similar result can also be proven when the LFC, $\mu_i-\mu_j=-\delta_\mu$, under H_2^μ or the LFC, $\mu_i-\mu_j=0$, under H_0^μ is true. Since we have just shown that $T(\delta_\mu)\to\infty$ with probability 1 as $\delta_\mu\to 0$,

$$S_{ij}^2(T)\to\sigma_{ij}^2 \text{ with probability 1 as } \delta_\mu\to 0 \tag{C.7}$$

when the focus is on the LFCs. Due to Equation (C.7) and Continuous Mapping Theorem (Billingsley 1968), it is also true that $N_{ij}^+(T)\to\infty$ with probability 1 as $\delta_\mu\to 0$.

To establish the asymptotic probability statement in Equation (C.1), we start with $\sup_{\boldsymbol{\mu}\in\Omega_0(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}(\text{Accept } H_1^\mu)$ where $\Omega_0(\delta_\mu)=\{\boldsymbol{\mu}:\mu_i-\mu_j=0\}$.

$$\begin{aligned}
&\sup_{\boldsymbol{\mu}\in\Omega_0(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}(\text{Accept } H_1^\mu) = \sup_{\boldsymbol{\mu}\in\Omega_0(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\Big\{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell}) \ge \mathrm{OB}_{ij}^{\mu+}(T)\Big\}\\
&= \sup_{\boldsymbol{\mu}\in\Omega_0(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\Big\{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell}) \ge Td + \frac{S_{ij}^2(T)g(\alpha_K/2)}{4d}\Big\}\\
&= \sup_{\boldsymbol{\mu}\in\Omega_0(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\Big\{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell}) - (\mu_i-\mu_j)T + (\mu_i-\mu_j)T \ge Td + \frac{S_{ij}^2(T)g(\alpha_K/2)}{4d}\Big\}\\
&= \sup_{\boldsymbol{\mu}\in\Omega_0(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\Big\{\frac{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell}) - (\mu_i-\mu_j)T}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} + \frac{(\mu_i-\mu_j)T}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}\\
&\qquad\qquad \ge \frac{Td}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} + \frac{S_{ij}^2(T)g(\alpha_K/2)}{4d\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}\Big\}\\
&= \mathrm{P}\Big\{\frac{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell}) - (\mu_i-\mu_j)T}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} \ge \frac{Td}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} + \frac{S_{ij}^2(T)g(\alpha_K/2)}{4d\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}\Big\}.
\end{aligned} \tag{C.8}$$

In Equation (C.8), the supremum is attained because $\mu_i-\mu_j$ is replaced with zero which is the LFC, and actually the only configuration, in$\Omega\ _0(\delta_\mu)$. In Equation (C.8), we replace

T with $(N_{ij}^+(T)+1)t$, $0 \le t \le 1$. Then

$$\mathrm{P}\Big\{\frac{\sum_{\ell=1}^{\lfloor (N_{ij}^+(T)+1)t \rfloor}(X_{i\ell}-X_{j\ell})-(\mu_i-\mu_j)(N_{ij}^+(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} \ge \frac{(N_{ij}^+(T)+1)d}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}t+\frac{S_{ij}^2(T)g(\alpha_K/2)}{4d\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}\Big\}$$
$$=\mathrm{P}\{C_{ij}^{\mathrm{OB}}(t,\delta_\mu) \ge B_{ij}^{\mathrm{OB}}(\delta_\mu)t + A_{ij}^{\mathrm{OB}}(\delta_\mu)\},$$

where

$$C_{ij}^{\mathrm{OB}}(t,\delta_\mu)=\frac{\sum_{\ell=1}^{\lfloor (N_{ij}^+(T)+1)t \rfloor}(X_{i\ell}-X_{j\ell})-(\mu_i-\mu_j)(N_{ij}^+(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} \text{ for } 0 \le t \le 1;$$

$$B_{ij}^{\mathrm{OB}}(\delta_\mu)=\frac{(N_{ij}^+(T)+1)d}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}; \text{ and } A_{ij}^{\mathrm{OB}}(\delta_\mu)=\frac{S_{ij}^2(T)g(\alpha_K/2)}{4d\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}.$$

Consider a similar derivation for $\sup_{\boldsymbol{\mu}\in\Omega_0(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\{\text{Accept } H_2^\mu\}$ in Equation (C.2). Next, we continue with $\sup_{\boldsymbol{\mu}\in\Omega_1(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\{\text{Reject } H_1^\mu\}$ in Equation (C.3).

$$\sup_{\boldsymbol{\mu}\in\Omega_1(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\{\text{Reject } H_1^\mu\} \le \sup_{\boldsymbol{\mu}\in\Omega_1(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\Big\{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell}) \le \mathrm{IB}_{ij}^{\mu+}(T)\Big\} \tag{C.9}$$
$$=\sup_{\boldsymbol{\mu}\in\Omega_1(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\Big\{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell}) \le T(\delta_\mu-d)-\frac{S_{ij}^2(T)g(\alpha_K)}{4d}\Big\}$$
$$=\sup_{\boldsymbol{\mu}\in\Omega_1(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\Big\{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell})-(\mu_i-\mu_j)T+(\mu_i-\mu_j)T \le T(\delta_\mu-d)-\frac{S_{ij}^2(T)g(\alpha_K)}{4d}\Big\}$$
$$=\sup_{\boldsymbol{\mu}\in\Omega_1(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\Big\{\frac{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell})-(\mu_i-\mu_j)T}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}+\frac{(\mu_i-\mu_j)T}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} \le \frac{T(\delta_\mu-d)}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}-\frac{S_{ij}^2(T)g(\alpha_K)}{4d\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}\Big\}$$
$$=\mathrm{P}\Big\{\frac{\sum_{\ell=1}^{T}(X_{i\ell}-X_{j\ell})-(\mu_i-\mu_j)T}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}+\frac{\delta_\mu T}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} \le \frac{T(\delta_\mu-d)}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}-\frac{S_{ij}^2(T)g(\alpha_K)}{4d\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}\Big\}. \tag{C.10}$$

In Equation (C.9), the partial sum moves to the lower section of the continuation region which includes a possible crossing of $\mathrm{IB}_{ij}^{\mu+}$, $-\mathrm{IB}_{ij}^{\mu+}$ or $-\mathrm{OB}_{ij}^{\mu+}$. The inequality is due to the fact that there is not a condition on T as given in the Computation step of the M_2^+

procedure. In Equation (C.10), the supremum is attained because $\mu_i - \mu_j$ is replaced with δ_μ which is the LFC in$\Omega_1(\delta_\mu)$.

In Equation (C.10), we replace T with $(N_{ij}^+(T)+1)t$, $0 \le t \le 1$. Then

$$= \mathrm{P}\Big\{ \frac{\sum_{\ell=1}^{\lfloor (N_{ij}^+(T)+1)t \rfloor}(X_{i\ell}-X_{j\ell}) - (\mu_i-\mu_j)(N_{ij}^+(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} + \frac{\delta_\mu(N_{ij}^+(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} \le \frac{(N_{ij}^+(T)+1)(\delta_\mu - d)}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}t - \frac{S_{ij}^2(T)g(\alpha_K)}{4d\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} \Big\}$$

$$= \mathrm{P}\Big\{ C_{ij}^{\mathrm{IB}}(t,\delta_\mu) + \frac{\delta_\mu(N_{ij}^+(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} \le B_{ij}^{\mathrm{IB}}(\delta_\mu)t - A_{ij}^{\mathrm{IB}}(\delta_\mu) \Big\},$$

where

$$C_{ij}^{\mathrm{IB}}(t,\delta_\mu) = \frac{\sum_{\ell=1}^{\lfloor (N_{ij}^+(T)+1)t \rfloor}(X_{i\ell}-X_{j\ell}) - (\mu_i-\mu_j)(N_{ij}^+(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} \text{ for } 0 \le t \le 1;$$

$$B_{ij}^{\mathrm{IB}}(\delta_\mu) = \frac{(N_{ij}^+(T)+1)(\delta_\mu - d)}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}; \text{ and } A_{ij}^{\mathrm{IB}}(\delta_\mu) = \frac{S_{ij}^2(T)g(\alpha_K)}{4d\sigma_{ij}\sqrt{N_{ij}^+(T)+1}}.$$

Consider a similar derivation for $\sup_{\boldsymbol{\mu}\in\Omega_2(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\{\text{Reject } H_2^\mu\}$ in Equation (C.4).

Further define

$$\hat{T}(\delta_\mu) = \min\Big\{ t \in \Big\{\frac{n_0}{N_{ij}^+(T)+1}, \frac{n_0+1}{N_{ij}^+(T)+1}, \ldots, 1\Big\} : |C_{ij}^{\mathrm{OB}}(t,\delta_\mu)| \ge B_{ij}^{\mathrm{OB}}(\delta_\mu)t + A_{ij}^{\mathrm{OB}}(\delta_\mu) \text{ or } \Big| C_{ij}^{\mathrm{IB}}(t,\delta_\mu) + \frac{\delta_\mu(N_{ij}^+(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} \Big| \le B_{ij}^{\mathrm{IB}}(\delta_\mu)t - A_{ij}^{\mathrm{IB}}(\delta_\mu) \Big\}.$$

Clearly, $\hat{T}(\delta_\mu) = T(\delta_\mu)/(N_{ij}^+(T)+1)$. Also, define the stopping time of the corresponding continuous-time process as

$$\tilde{T}(\delta_\mu) = \min\Big\{ t \ge \frac{n_0}{N_{ij}^+(T)+1} : |C_{ij}^{\mathrm{OB}}(t,\delta_\mu)| \ge B_{ij}^{\mathrm{OB}}(\delta_\mu)t + A_{ij}^{\mathrm{OB}}(\delta_\mu) \text{ or } \Big| C_{ij}^{\mathrm{IB}}(t,\delta_\mu) + \frac{\delta_\mu(N_{ij}^+(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} \Big| \le B_{ij}^{\mathrm{IB}}(\delta_\mu)t - A_{ij}^{\mathrm{IB}}(\delta_\mu) \Big\}.$$

Note that for fixed δ_μ, $C_{ij}^{\mathrm{OB}}(\hat{T}(\delta_\mu),\delta_\mu)$ and $C_{ij}^{\mathrm{IB}}(\hat{T}(\delta_\mu),\delta_\mu)$ correspond to the right-hand limit of a point of discontinuity of $C_{ij}^{\mathrm{OB}}(\cdot,\delta_\mu)$ and $C_{ij}^{\mathrm{IB}}(\cdot,\delta_\mu)$, respectively. We can show that

$\hat{T}(\delta_\mu) \to \tilde{T}(\delta_\mu)$ with probability 1 as $\delta_\mu \to 0$ because $1/(N_{ij}^{+}(T)+1) \to 0$ with probability 1 as $\delta_\mu \to 0$. Thus, we can focus on $C_{ij}^{\mathrm{OB}}(\tilde{T}(\delta_\mu), \delta_\mu)$ and $C_{ij}^{\mathrm{IB}}(\tilde{T}(\delta_\mu), \delta_\mu)$ as $\delta_\mu \to 0$.

Since strong consistency implies convergence in probability, Equation (C.7) and Lemma C1 can be applied to $(C_{ij}^{\mathrm{OB}}(t,\delta_\mu), C_{ij}^{\mathrm{IB}}(t,\delta_\mu), S_{ij}^2(T))$. Then Lemma C2 and the random-change-of-time theorem (Billingsley 1968) imply that

$$C_{ij}^{\mathrm{OB}}(t,\delta_\mu) \Rightarrow W(t,0) \text{ as } \delta_\mu \to 0 \text{ and}$$
$$C_{ij}^{\mathrm{IB}}(t,\delta_\mu) + \frac{\delta_\mu (N_{ij}^{+}(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^{+}(T)+1}} \Rightarrow W(t,\Delta) \text{ as } \delta_\mu \to 0,$$

where $\Rightarrow$ denotes convergence in distribution. The computation of the drift Δ is

$$\begin{aligned}
\Delta &= \lim_{\delta_\mu \to 0} \frac{\delta_\mu (N_{ij}^{+}(T)+1)}{\sigma_{ij}\sqrt{N_{ij}^{+}(T)+1}} \\
&= \lim_{\delta_\mu \to 0} \frac{\delta_\mu}{\sigma_{ij}} \sqrt{\frac{8S_{ij}^2(T)(g(\alpha_K)+g(\alpha_K/2))}{3\delta_\mu^2}+1} \\
&= \lim_{\delta_\mu \to 0} \frac{1}{\sigma_{ij}} \sqrt{\frac{8S_{ij}^2(T)(g(\alpha_K)+g(\alpha_K/2))}{3}+\delta_\mu^2} \\
&= \sqrt{(8/3)(g(\alpha_K)+g(\alpha_K/2))}.
\end{aligned}$$

When $d = 3\delta_\mu/8$, the following results are obtained:

$$\begin{aligned}
A_{ij}^{\mathrm{OB}}(\delta_\mu) &= \frac{S_{ij}^2(T)g(\alpha_K/2)}{4d\sigma_{ij}\sqrt{N_{ij}^{+}(T)+1}} = \frac{S_{ij}^2(T)g(\alpha_K/2)}{\sigma_{ij}\sqrt{6S_{ij}^2(T)(g(\alpha_K)+g(\alpha_K/2))+9\delta_\mu^2/4}}, \\
B_{ij}^{\mathrm{OB}}(\delta_\mu) &= \frac{(N_{ij}^{+}(T)+1)d}{\sigma_{ij}\sqrt{N_{ij}^{+}(T)+1}} = \frac{1}{\sigma_{ij}}\sqrt{\frac{3S_{ij}^2(T)(g(\alpha_K)+g(\alpha_K/2))}{8}+\frac{9\delta_\mu^2}{64}}, \\
A_{ij}^{\mathrm{IB}}(\delta_\mu) &= \frac{S_{ij}^2(T)g(\alpha_K)}{4d\sigma_{ij}\sqrt{N_{ij}^{+}(T)+1}} = \frac{S_{ij}^2(T)g(\alpha_K)}{\sigma_{ij}\sqrt{6S_{ij}^2(T)(g(\alpha_K)+g(\alpha_K/2))+9\delta_\mu^2/4}}, \\
B_{ij}^{\mathrm{IB}}(\delta_\mu) &= \frac{(N_{ij}^{+}(T)+1)(\delta_\mu - d)}{\sigma_{ij}\sqrt{N_{ij}^{+}(T)+1}} = \frac{1}{\sigma_{ij}}\sqrt{\frac{25S_{ij}^2(T)(g(\alpha_K)+g(\alpha_K/2))}{24}+\frac{25\delta_\mu^2}{64}}.
\end{aligned}$$

The asymptotic values are as follows:

$$\begin{aligned}
A_{ij}^{\mathrm{OB}} &\equiv \lim_{\delta_\mu \to 0} A_{ij}^{\mathrm{OB}}(\delta_\mu) = g(\alpha_K/2)/\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}, \\
B_{ij}^{\mathrm{OB}} &\equiv \lim_{\delta_\mu \to 0} B_{ij}^{\mathrm{OB}}(\delta_\mu) = \sqrt{6(g(\alpha_K)+g(\alpha_K/2))}/4,
\end{aligned}$$

$$A_{ij}^{\mathrm{IB}} \equiv \lim_{\delta_\mu \to 0} A_{ij}^{\mathrm{IB}}(\delta_\mu) = g(\alpha_K)/\sqrt{6(g(\alpha_K)+g(\alpha_K/2))},$$
$$B_{ij}^{\mathrm{IB}} \equiv \lim_{\delta_\mu \to 0} B_{ij}^{\mathrm{IB}}(\delta_\mu) = 5\sqrt{g(\alpha_K)+g(\alpha_K/2)}/(2\sqrt{6}).$$

Define the mapping $s_{\delta_\mu}: D[0,1] \to R$ such that $s_{\delta_\mu}(C) = C(T_{C,\delta_\mu})$ where

$$T_{C,\delta_\mu} = \inf\{t : |C_{ij}^{\mathrm{OB}}(t,\delta_\mu)| \geq B_{ij}^{\mathrm{OB}}(\delta_\mu)t + A_{ij}^{\mathrm{OB}}(\delta_\mu) \text{ or } \left|C_{ij}^{\mathrm{IB}}(t,\delta_\mu) + \frac{\delta_\mu(N_{ij}^{+}(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^{+}(T)+1}}\right| \leq B_{ij}^{\mathrm{IB}}(\delta_\mu)t - A_{ij}^{\mathrm{IB}}(\delta_\mu)\}$$

for every $C_{ij}^{\mathrm{OB}} \in D[0,1]$, $C_{ij}^{\mathrm{IB}} \in D[0,1]$, and $\delta_\mu > 0$. Define mapping $s(C) = C(T_C)$

$$T_C = \inf\{t : |C_{ij}^{\mathrm{OB}}(t)| \geq B_{ij}^{\mathrm{OB}}t + A_{ij}^{\mathrm{OB}} \text{ or } |C_{ij}^{\mathrm{IB}}(t) + \Delta t| \leq B_{ij}^{\mathrm{IB}}t - A_{ij}^{\mathrm{IB}}\}$$

for every $C_{ij}^{\mathrm{OB}} \in D[0,1]$ and $C_{ij}^{\mathrm{IB}} \in D[0,1]$.

Note that

$$s_{\delta_\mu}(C_{ij}^{\mathrm{OB}}(t,\delta_\mu)) = C_{ij}^{\mathrm{OB}}(\tilde{T},\delta_\mu),$$
$$s(W(\cdot,0)) = W(T_{W(\cdot,0)},0),$$
$$s_{\delta_\mu}\left(C_{ij}^{\mathrm{IB}}(t,\delta_\mu) + \frac{\delta_\mu(N_{ij}^{+}(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^{+}(T)+1}}\right) = C_{ij}^{\mathrm{IB}}(\tilde{T},\delta_\mu) + \frac{\delta_\mu(N_{ij}^{+}(T)+1)\tilde{T}}{\sigma_{ij}\sqrt{N_{ij}^{+}(T)+1}},$$
$$s(W(\cdot,\Delta)) = W(T_{W(\cdot,\Delta)},\Delta).$$

We need to show that

$$s_{\delta_\mu}(C_{ij}^{\mathrm{OB}}(\cdot,\delta_\mu)) \Rightarrow s(W(\cdot,0)) \text{ as } \delta_\mu \to 0 \text{ and} \quad \text{(C.11)}$$
$$s_{\delta_\mu}(G_{ij}(\cdot,\delta_\mu)) \Rightarrow s(W(\cdot,\Delta)) \text{ as } \delta_\mu \to 0, \quad \text{(C.12)}$$

where

$$G_{ij}(t,\delta_\mu) \equiv C_{ij}^{\mathrm{IB}}(t,\delta_\mu) + \frac{\delta_\mu(N_{ij}^{+}(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^{+}(T)+1}} \text{ for } t \in [0,1] \text{ and } \delta_\mu > 0.$$

This follows from Proposition 2 in Kim et al. (2005) which establishes that the extended Continuous Mapping Theorem (Billingsley 1968) applies.

The following completes the asymptotic probability computations in Equations (C.1)-(C.4):

$$\limsup_{\delta_\mu \to 0} \sup_{\boldsymbol{\mu} \in \Omega_0(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\{\text{Accept } H_1^{\mu}\}$$

$$= \limsup_{\delta_\mu \to 0} \mathrm{P}\{C_{ij}^{\mathrm{OB}}(t,\delta_\mu) \geq B_{ij}^{\mathrm{OB}}(\delta_\mu)t + A_{ij}^{\mathrm{OB}}(\delta_\mu)\}$$

$$\leq \mathrm{P}\{W(t,0) \geq B_{ij}^{\mathrm{OB}}t + A_{ij}^{\mathrm{OB}}\} \tag{C.13}$$

$$= \mathrm{P}\Big\{W(t,0) \geq \frac{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}{4}t + \frac{g(\alpha_K/2)}{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}\Big\}$$

$$= \mathrm{P}\Big\{W(t,0) - \frac{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}{4}t \geq \frac{g(\alpha_K/2)}{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}\Big\}. \tag{C.14}$$

The inequality in Equation (C.13) is due to Equation (C.11) and Lemma A1. Let

$$\Delta_1 = -\frac{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}{4} \text{ and } b = \frac{g(\alpha_K/2)}{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}.$$

Let $T_b = \inf\{t : W(t,\Delta_1) = b\}$. Following from Equation (C.14),

$$\mathrm{P}\Big\{W(t,0) - \frac{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}{4}t \geq \frac{g(\alpha_K/2)}{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}\Big\}$$

$$= \mathrm{P}\{W(t,\Delta_1) \geq b\}$$

$$\leq \mathrm{P}_{\Delta_1}\{T_b < \infty\} \tag{C.15}$$

$$= \exp(2\Delta_1 b) \tag{C.16}$$

$$= \exp\Big(2\frac{-\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}{4}\frac{g(\alpha_K/2)}{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}\Big)$$

$$= \exp(-g(\alpha_K/2)/2)$$

$$= \exp(-2\ln(2/\alpha_K)/2)$$

$$= \alpha_K/2.$$

The inequality in Equation (C.15) is because the Brownian motion process may hit IB before it hits line b. Equation (C.16) is due to Lemma A2.

Hence, we have shown that Equation (C.1) holds. By a similar derivation, it can be shown that Equation (C.2) holds. Next, we will show that Equation (C.3) holds.

$$\limsup_{\delta_\mu \to 0} \sup_{\boldsymbol{\mu} \in \Omega_1(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\{\text{Reject } H_1^\mu\}$$

$$\leq \limsup_{\delta_\mu \to 0} \mathrm{P}\Big\{C_{ij}^{\mathrm{IB}}(t,\delta_\mu) + \frac{\delta_\mu(N_{ij}^+(T)+1)t}{\sigma_{ij}\sqrt{N_{ij}^+(T)+1}} \leq B_{ij}^{\mathrm{IB}}(\delta_\mu)t - A_{ij}^{\mathrm{IB}}(\delta_\mu)\Big\}$$

$$\leq \mathrm{P}\{W(t,\Delta) \leq B_{ij}^{\mathrm{IB}}t - A_{ij}^{\mathrm{IB}}\} \tag{C.17}$$

$$= \mathrm{P}\Big\{W(t,\Delta) \leq \frac{5\sqrt{g(\alpha_K)+g(\alpha_K/2)}}{2\sqrt{6}}t - \frac{g(\alpha_K)}{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}\Big\}$$
$$= \mathrm{P}\Big\{W(t,\Delta) - \frac{5\sqrt{g(\alpha_K)+g(\alpha_K/2)}}{2\sqrt{6}}t \leq -\frac{g(\alpha_K)}{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}\Big\}. \quad \text{(C.18)}$$

The inequality in Equation (C.17) is due to Equation (C.12) and Lemma A1. Let

$$\Delta_2 = \Delta - \frac{5\sqrt{g(\alpha_K)+g(\alpha_K/2)}}{2\sqrt{6}} = \frac{3\sqrt{g(\alpha_K)+g(\alpha_K/2)}}{2\sqrt{6}},$$
$$a = -\frac{g(\alpha_K)}{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}.$$

Let $T_a = \inf\{t : W(t,\Delta_2) = a\}$. Following from Equation (C.18),

$$\mathrm{P}\Big\{W(t,\Delta) - \frac{5\sqrt{g(\alpha_K)+g(\alpha_K/2)}}{2\sqrt{6}}t \leq -\frac{g(\alpha_K)}{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}\Big\}$$
$$= \mathrm{P}\Big\{W(t,\Delta_2) \leq a\}$$
$$\leq \mathrm{P}_{\Delta_2}\{T_a < \infty\} \quad \text{(C.19)}$$
$$= \exp(2\Delta_2 a) \quad \text{(C.20)}$$
$$= \exp\Big(2\frac{3\sqrt{g(\alpha_K)+g(\alpha_K/2)}}{2\sqrt{6}}\frac{-g(\alpha_K)}{\sqrt{6(g(\alpha_K)+g(\alpha_K/2))}}\Big)$$
$$= \exp(-g(\alpha_K)/2)$$
$$= \exp(-2\ln(1/\alpha_K)/2)$$
$$= \alpha_K.$$

The inequality in Equation (C.19) is because the Brownian motion process may hit OB before it hits line a. Equation (C.20) is due to Lemma A2.

Hence, we have shown that Equation (C.3) holds. By a similar derivation, it can be shown that Equation (C.4) holds.

Proof of Corollary 2

In order to prove that Equation (11) holds, we will show that

$$\limsup_{\delta_\mu \to 0} \sup_{\boldsymbol{\mu} \in \Omega_3(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\{\text{Accept } H_2^\mu\} \leq \alpha_K/2.$$

Then

$$\limsup_{\delta_\mu \to 0} \sup_{\boldsymbol{\mu} \in \Omega_3(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\{\text{Accept } H_2^\mu\}$$

$$\leq \limsup_{\delta_\mu \to 0} \sup_{\boldsymbol{\mu} \in \Omega_0(\delta_\mu)} \mathrm{P}_{\boldsymbol{\mu}}\{\text{Accept } H_2^\mu\} \tag{C.21}$$

$$\leq \frac{\alpha_K}{2}. \tag{C.22}$$

The inequality in Equation (C.21) holds because the probability of accepting H_2^μ increases when $\boldsymbol{\mu} \in \Omega_0(\delta_\mu)$. Equation (C.22) is due to Equation (C.2).

By a similar derivation, it can be shown that Equation (12) holds.

D. M_K^+ Procedure

The algorithm of the M_K^+ procedure is described in detail below.

1. Setup: Place all K systems in the set of initial systems, i.e., $S_{\mathrm{IN}} = \{1, 2, \ldots, K\}$. Define S_{EQ} as the set of pairs of equivalent systems, and set $S_{\mathrm{EQ}} = \emptyset$. Define S_{DE} as the set of decided systems whose all pairwise tests with the systems in S_{IN} reached to a decision, and set $S_{\mathrm{DE}} = \emptyset$. Specify the confidence level for the M_K^+ procedure, $1-\alpha$, and the mean-IZ parameter, δ_μ ($\delta_\mu > 0$). Set parameter d to $3\delta_\mu/8$. Set α_K to $\alpha/[K(K-1)/2]$. Specify the initial sample size n_0 ($n_0 > 2$).

2. Initialization: Obtain n_0 observations from every system i, i in S_{IN}. Set the number of stages, r, to n_0.

3. Computation: For system pair (i,j), $i,j \in S_{\mathrm{IN}}$, $(i,j) \notin S_{\mathrm{EQ}}$, $i<j$, compute the sample variance of differences of observations

$$S_{ij}^2(r) = \frac{1}{r-1}\sum_{\ell=1}^{r}\Big(X_{i\ell} - X_{j\ell} - (\bar{X}_i(r) - \bar{X}_j(r))\Big)^2, \text{ where } \bar{X}_i(r) = \frac{1}{r}\sum_{\ell=1}^{r} X_{i\ell};$$

and the inner and outer bounds

$$\mathrm{IB}_{ij}^{\mu+}(r) = r(\delta_\mu - d) - \frac{S_{ij}^2(r)g(\alpha_K)}{4d} \text{ when } r > \frac{S_{ij}^2(r)g(\alpha_K)}{4d(\delta_\mu - d)};$$

$$\mathrm{OB}_{ij}^{\mu+}(r) = rd + \frac{S_{ij}^2(r)g(\alpha_K/2)}{4d};$$

where $g(\eta) = 2\ln(1/\eta)$.

4. Decision:

For system pair (i,j), $i,j \in S_{\mathrm{IN}}$, $(i,j) \notin S_{\mathrm{EQ}}$, $i<j$:

i) If $-\mathrm{IB}_{ij}^{\mu+}(r) \leq \sum_{\ell=1}^{r}(X_{i\ell} - X_{j\ell}) \leq \mathrm{IB}_{ij}^{\mu+}(r)$, then accept $H_0^\mu : \mu_i = \mu_j$.

ii) Else if $\sum_{\ell=1}^{r}(X_{i\ell} - X_{j\ell}) \geq \mathrm{OB}_{ij}^{\mu+}(r)$, then accept $H_1^\mu : \mu_i > \mu_j$.

iii) Else if $\sum_{\ell=1}^{r}(X_{i\ell} - X_{j\ell}) \leq -\mathrm{OB}_{ij}^{\mu+}(r)$, then accept $H_2^\mu : \mu_i < \mu_j$.

5. Consolidation:

For system pair (i, j), $i, j \in S_{\text{IN}}$, $(i, j) \notin S_{\text{EQ}}$, $i < j$:

i) If $\mu_i = \mu_j$ decision is made, then declare systems i and j equivalent and set $S_{\text{EQ}} = S_{\text{EQ}} \cup \{(i, j)\}$.

ii) If $\mu_i > \mu_j$ decision is made, then eliminate system j by setting $S_{\text{IN}} = S_{\text{IN}} \backslash \{j\}$. Delete every system j pair in S_{EQ}.

iii) If $\mu_i < \mu_j$ decision is made, then eliminate system i by setting $S_{\text{IN}} = S_{\text{IN}} \backslash \{i\}$. Delete every system i pair in S_{EQ}.

6. Termination: If $|S_{\text{IN}}| = 1$ or S_{EQ} includes every pair (i, j) where $i, j \in S_{\text{IN}}, i < j$, stop sampling. Return S_{IN} as an M-efficient set.

If the procedure does not terminate, perform the following. For $i \in S_{\text{IN}}$, if every pair (i, j), $j \in S_{\text{IN}}, i \neq j$, is in S_{EQ}, add system i to S_{DE}. Obtain one more observation from every system i, $i \in S_{\text{IN}} \backslash S_{\text{DE}}$, and set $r = r + 1$. Go back to the Computation step. □

E. V_K Procedure

The algorithm of the V_K procedure is described in detail below.

1. Setup: Place all K systems in the set of initial systems, i.e., $S_{\text{IN}} = \{1, 2, \ldots, K\}$. Define S_{EQ} as the set of pairs of equivalent systems, and set $S_{\text{EQ}} = \emptyset$. Define S_{DE} as the set of decided systems whose all pairwise tests with the systems in S_{IN} reached to a decision, and set $S_{\text{DE}} = \emptyset$. Specify the confidence level for the V_K procedure, $1 - \alpha$, and the variance-IZ parameter, δ_σ^2 ($\delta_\sigma > 1$). Set parameter λ for the variance hypothesis test to $1 + 0.7(\delta_\sigma - 1)$. Set α_K to $\alpha/[K(K-1)/2]$. Specify the initial sample size n_0 ($n_0 > 2$).

2. Initialization: Obtain n_0 observations from every system i, $i \in S_{\text{IN}}$. Set the number of stages, r, to n_0.

3. Computation:

Compute the inner and outer bounds

$$\text{IB}^\sigma(r) = \frac{\delta_\sigma^2\left(\lambda\alpha_K^{\frac{1}{r-1}} - 1\right)}{\lambda\left(\lambda - \alpha_K^{\frac{1}{r-1}}\right)} \text{ when } r > 1 + \frac{\ln(\alpha_K)}{\ln\left(\frac{\lambda^2+\delta_\sigma^2}{\lambda(\delta_\sigma^2+1)}\right)};$$

$$\text{OB}^\sigma(r) = \frac{\lambda\left(\lambda(2/\alpha_K)^{\frac{1}{r-1}} - 1\right)}{\left(\lambda - (2/\alpha_K)^{\frac{1}{r-1}}\right)} \text{ when } r > 1 + \frac{\ln(2/\alpha_K)}{\ln(\lambda)}.$$

For system i, $i \in S_{\text{IN}}$, if there is a system pair (i,j), $j \in S_{\text{IN}}$, $(i,j) \notin S_{\text{EQ}}$, $i \neq j$, compute the sample variance

$$S_i^2(r) = \frac{1}{r-1}\sum_{\ell=1}^{r}(X_{i\ell} - \bar{X}_i(r))^2, \text{ where } \bar{X}_i(r) = \frac{1}{r}\sum_{\ell=1}^{r} X_{i\ell}.$$

4. Decision:

For system pair (i,j), $i,j \in S_{\text{IN}}$, $(i,j) \notin S_{\text{EQ}}$, $i < j$:

i) If $1/\text{IB}^{\sigma}(r) \le S_i^2(r)/S_j^2(r) \le \text{IB}^{\sigma}(r)$, then accept $H_0^{\sigma}: \sigma_i^2 = \sigma_j^2$.

ii) Else if $S_i^2(r)/S_j^2(r) \ge \text{OB}^{\sigma}(r)$, then accept $H_1^{\sigma}: \sigma_i^2 > \sigma_j^2$.

iii) Else if $S_i^2(r)/S_j^2(r) \le 1/\text{OB}^{\sigma}(r)$, then accept $H_2^{\sigma}: \sigma_i^2 < \sigma_j^2$.

5. Consolidation:

For system pair (i,j), $i,j \in S_{\text{IN}}$, $(i,j) \notin S_{\text{EQ}}$, $i < j$:

i) If $\sigma_i^2 = \sigma_j^2$ decision is made, then declare systems i and j equivalent and set $S_{\text{EQ}} = S_{\text{EQ}} \cup \{(i,j)\}$.

ii) If $\sigma_i^2 > \sigma_j^2$ decision is made, then eliminate system i by setting $S_{\text{IN}} = S_{\text{IN}} \backslash \{i\}$. Delete every system i pair in S_{EQ}.

iii) If $\sigma_i^2 < \sigma_j^2$ decision is made, then eliminate system j by setting $S_{\text{IN}} = S_{\text{IN}} \backslash \{j\}$. Delete every system j pair in S_{EQ}.

6. Termination: If $|S_{\text{IN}}| = 1$ or S_{EQ} includes every pair (i,j) where $i,j \in S_{\text{IN}}, i < j$, stop sampling. Return S_{IN} as a V-efficient set.

If the procedure does not terminate, perform the following. For $i \in S_{\text{IN}}$, if every pair (i,j), $j \in S_{\text{IN}}, i \neq j$, is in S_{EQ}, add system i to S_{DE}. Obtain one more observation from every system i, $i \in S_{\text{IN}} \backslash S_{\text{DE}}$, and set $r = r+1$. Go back to the Computation step. □

F. Relationship of the MV Dominance to Stochastic Dominance

Table F.1 **Relationship of the MV and Stochastic Dominance Criteria Under the Assumption of Normality**

	$\sigma_i < \sigma_j$	$\sigma_i = \sigma_j$	$\sigma_i > \sigma_j$
$\mu_i > \mu_j$	$F_i \underset{\text{MV, SSD}}{>} F_j$	$F_i \underset{\text{MV, SSD, FSD}}{>} F_j$	Nondominant
$\mu_i = \mu_j$	$F_i \underset{\text{MV, SSD}}{>} F_j$	$F_i = F_j$	$F_i \underset{\text{MV, SSD}}{<} F_j$
$\mu_i < \mu_j$	Nondominant	$F_i \underset{\text{MV, SSD, FSD}}{<} F_j$	$F_i \underset{\text{MV, SSD}}{<} F_j$

The MVS procedure selects the best system based on the MV dominance criterion. Since we assume that observations come from normal populations, this best system selection decision is also valid under the first- and/or second-order stochastic dominance criteria. This expanded validation is possible because of the following properties of the normal distribution. For two cumulative normal distributions, F_i and F_j: i) If $\sigma_i = \sigma_j$, the F_i and F_j distributions never intersect; and the system with the larger mean dominates the other system by the first-order stochastic dominance (FSD) criterion. ii) If $\sigma_i \neq \sigma_j$, distributions intersect exactly once; and the system with the larger mean and smaller variance dominates the other system by the second-order stochastic dominance (SSD) criterion (Kroll and Levy 1980). Table F.1 summarizes these relationships. Since FSD implies SSD for all distributions, at the minimum, the MV dominance also implies SSD under the assumption of normality.

G. Experiments for Large Number of Systems

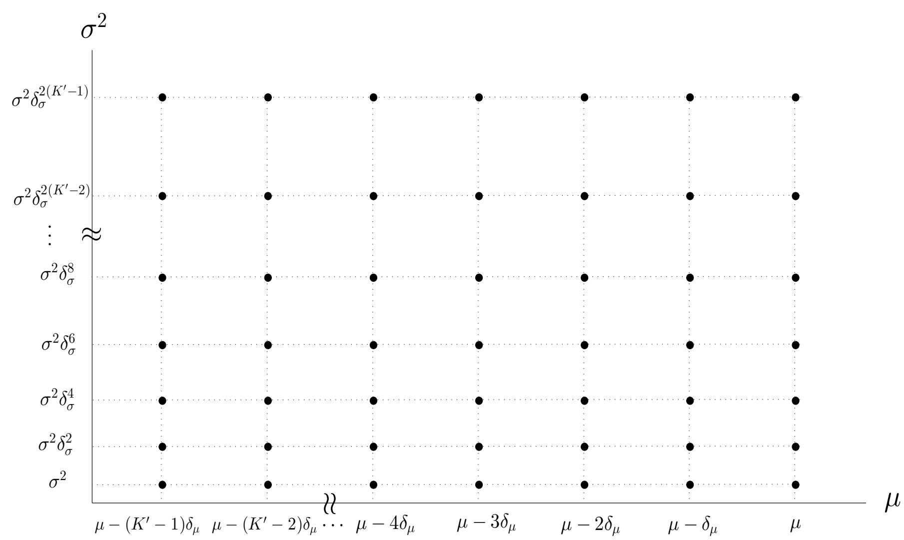

Figure G.1 **Mapping of K Systems in Equation (G.1)**

Experiments 15-19 are designed to test the effectiveness of the two procedures when a large and varied number of systems are tested. In the experiments, the following mean and variance configurations are used:

$$\mu_i = \mu - ((i-1) \bmod K')\delta_\mu \quad \text{and} \quad \sigma_i^2 = \sigma^2 \delta_\sigma^{2\lfloor (i-1)/K' \rfloor}, \tag{G.1}$$

where $i = 1, \ldots, K$ and $K' = \sqrt{K}$. Mapping of K systems is depicted in Figure G.1.

Experiments 15-17 in Table G.1 use the MVS procedure. In Experiment 15, System 1 is the best system. In Experiment 16, System $K+1$ is added to the systems from Experiment

Table G.1 Experiments 15-17 for the MVS Procedure ($\mu = 100, \sigma = 10, \delta_\mu = 3,\ \delta_\sigma = 1.1$)

Exp.	Systems	Correct selection
15	$\mu_i = \mu - ((i-1) \bmod K')\delta_\mu,\ \sigma_i^2 = \sigma^2\delta_\sigma^{2\lfloor (i-1)/K' \rfloor}$	MV-efficient $=\{1\}$
	$i = 1,\ldots,K, K' = \sqrt{K}$	$S_{\text{EQ}} = \emptyset$
		$S_{\text{ND}} = \emptyset$
16	$\mu_i = \mu - ((i-1) \bmod K')\delta_\mu,\ \sigma_i^2 = \sigma^2\delta_\sigma^{2\lfloor (i-1)/K' \rfloor}$	MV-efficient $=\{1, K+1\}$
	$i = 1,\ldots,K, K' = \sqrt{K}$	$S_{\text{EQ}} = \{(1, K+1)\}$
	$\mu_{K+1} = \mu, \sigma_{K+1}^2 = \sigma^2$	$S_{\text{ND}} = \emptyset$
17	$\mu_i = \mu - ((i-1) \bmod K')\delta_\mu,\ \sigma_i^2 = \sigma^2\delta_\sigma^{2\lfloor (i-1)/K' \rfloor}$	MV-efficient $=\{1, K+1\}$
	$i = 1,\ldots,K, K' = \sqrt{K}$	$S_{\text{EQ}} = \emptyset$
	$\mu_{K+1} = \mu + \delta_\mu, \sigma_{K+1}^2 = \sigma^2\delta_\sigma^2$	$S_{\text{ND}} = \{(1, K+1)\}$

Table G.2 Results of Experiments 15-17 for the MVS Procedure ($\mu = 100, \sigma = 10, \delta_\mu = 3,\ \delta_\sigma = 1.1$)

	$K = 100$		$K = 225$		$K = 400$	
Exp.	PCS	SATO	PCS	SATO	PCS	SATO
15	1.000	108,818	1.000	244,563	1.000	522,711
16	1.000	110,985	1.000	246,317	1.000	531,744
17	1.000	107,230	1.000	224,860	1.000	456,837

15 such that Systems 1 and $K+1$ are equivalent. In Experiment 17, System $K+1$ is added to the systems from Experiment 15 such that Systems 1 and $K+1$ are nondominant. The results of Experiments 15-17 for different values of K are shown in Table G.2. All PCS values are above the prespecified confidence level of 0.95.

Table G.3 Experiments 18-19 for the CMVS Procedure ($\mu = 100, \sigma = 10, \delta_\mu = 3, \delta_\sigma = 1.1$)

Exp.	Systems	Correct selection
18	$\mu_i = \mu - ((i-1) \bmod K')\delta_\mu,\ \sigma_i^2 = \sigma^2\delta_\sigma^{2\lfloor (i-1)/K' \rfloor}$	CMV-efficient $=\{0\}$
	$i = 1,\ldots,K, K' = \sqrt{K}$	$S_{\text{EQ}} = \emptyset$
	$\mu_0 = \mu, \sigma_0^2 = \sigma^2$	
19	$\mu_i = \mu - ((i-1) \bmod K')\delta_\mu,\ \sigma_i^2 = \sigma^2\delta_\sigma^{2\lfloor (i-1)/K' \rfloor}$	CMV-efficient $=\{1\}$
	$i = 1,\ldots,K, K' = \sqrt{K}$	$S_{\text{EQ}} = \emptyset$
	$\mu_0 = \mu - \delta_\mu, \sigma_0^2 = \sigma^2\delta_\sigma^2$	

Table G.4 Results of Experiments 18-19 for the CMVS Procedure ($\mu = 100, \sigma = 10, \delta_\mu = 3, \delta_\sigma = 1.1$)

	$K = 100$		$K = 225$		$K = 400$	
Exp.	PCS	SATO	PCS	SATO	PCS	SATO
18	1.000	15,503	1.000	32,245	1.000	55,736
19	0.999	25,491	1.000	44,398	1.000	69,104

Experiments 18-19 in Table G.3 use the CMVS procedure. In Experiment 18, the reference system is set such that all contending systems are risk infeasible. Hence, the reference system is the best system. In Experiment 19, the reference system is set such that K' contending systems with variance σ^2 are risk feasible. When the reference system and the risk-feasible contending systems are compared, System 1 has the largest mean. Hence, System 1 is the best system. The results of Experiments 18-19 for different values of K are shown in Table G.4. All PCS values are above the prespecified confidence level of 0.95.

References

Billingsley, P. 1968. *Convergence of Probability Measures*. John Wiley and Sons, New York.

Jennison, C., I. M. Johnstone, B. W. Turnbull. 1980. Asymptotically optimal procedures for sequential adaptive selection of the best of several normal means. School of OR and IE, Cornell University, Technical Report 463, June 1980.

Kim, S.-H., B. L. Nelson. 2001. A Fully Sequential Procedure for Indifference-Zone Selection in Simulation. *ACM Transactions on Modeling and Computer Simulation*. **11** 251–273.

Kim, S.-H., B. L. Nelson. 2006. On the Asymptotic Validity of Fully Sequential Selection Procedures for Steady-State Simulation. *Operations Research*. **54** 475–488.

Kim, S.-H., B. L. Nelson, and J. R. Wilson. 2005. Some almost-sure convergence properties useful in sequential analysis. *Sequential Analysis*. **24** 411–419.

Kroll, Y., H. Levy. 1980. Stochastic dominance: A review and some new evidence. *Research in Finance*. **2** 163–227.

Ross, S. M. 1996. *Stochastic Processes* (2nd ed.). Wiley.

Wang, H., S.-H. Kim. 2012. Reducing the conservativeness of fully sequential indifference-zone procedures. *IEEE Transactions on Automatic Control*. **58**(6) 1613–1619.