E-Companion to "The (Surprising) Sample Optimality of Greedy Procedures for Large-Scale Ranking and Selection"

by Zaile Li, Weiwei Fan, and L. Jeff Hong

EC.1. Technical Supplements for Section 4

EC.1.1. Proof of Lemma 1

We prove the lemma by contradiction. For any particular sample path $\{\bar{Z}(n;\omega), n=1,2,\ldots\}$ of the running-average process, we let $M(\omega)=\arg\min_{n\geq 1}\bar{Z}(n;\omega)$ denote the location of the minima. Suppose that the lemma is violated. Then, there must exist a collection of sample paths $\Omega=\{\omega: \bar{Z}(n;\omega), n=1,2,\ldots\}$ such that

$$M(\omega)=\infty \text{ for any } \omega\in\Omega, \text{ and } \Pr\{\omega\in\Omega\}>0. \tag{EC.1.1}$$

Further, we let $M_s(\omega)=\arg\min_{1\leq n\leq s}\bar{Z}(n;\omega)$. If $M_s(\omega)$ has tied values, we set it as the smallest one. It is clear that, for any $\omega\in\Omega$, $\{M_s(\omega), s=1,2,\ldots\}$ is a monotonically non-decreasing sequence. In other words, for any $\omega\in\Omega$,

$$M_s(\omega)\uparrow M(\omega)=\infty, \text{ as } s\uparrow\infty, \text{ and } \bar{Z}(M_1(\omega);\omega)\geq \bar{Z}(M_2(\omega);\omega)\geq\cdots\geq \bar{Z}(M_s(\omega);\omega)\geq\ldots.$$

By the strong law of large number, $\bar{Z}(M_s(\omega);\omega)\downarrow 0$ as $s\uparrow\infty$. Therefore, we have that $\bar{Z}(M_s(\omega);\omega)=\min_{1\leq n\leq s}\bar{Z}(n;\omega)\geq \lim_{s\uparrow\infty}\bar{Z}(M_s(\omega);\omega)=0$ for any $s\geq 1$, or equivalently, $\min_{n\geq 1}\bar{Z}(n;\omega)\geq 0$. Then, in the consideration of Lemma 2, we can derive

$$\Pr\{\omega\in\Omega\}\leq \Pr\left\{\min_{n\geq 1}\bar{Z}(n;\omega)\geq 0\right\}=\Pr\{N(0)=\infty\}=\exp\left(-\sum_{n=1}^{\infty}n^{-1}\mathrm{P}(\bar{X}(n)\leq 0)\right)=0.$$

Clearly, this contradicts with Equation (EC.1.1). The proof is completed. □

EC.1.2. Proof of Lemma 2

In this proof, we only prove that $C(x)=\exp\left(\sum_{n=1}^{\infty}\frac{1}{n}\Phi\left(-\sqrt{n}x\right)\right)$ is continuous on $x\in(0,\infty)$. The other results are direct consequences of Corollaries 8.39 and 8.44 in Siegmund (1985). Let

$$f_m(x)=\sum_{n=1}^{m}\frac{1}{n}\Phi\left(-\sqrt{n}x\right) \text{ and } f(x)=\lim_{m\to\infty}f_m(x)=\sum_{n=1}^{\infty}\frac{1}{n}\Phi\left(-\sqrt{n}x\right).$$

Given any constant $\delta>0$, it is clear that

$$|f_m(x)-f(x)|=\sum_{n=m+1}^{\infty}\frac{1}{n}\Phi\left(-\sqrt{n}x\right)\leq \sum_{n=m+1}^{\infty}\frac{1}{n}\Phi\left(-\sqrt{n}\theta\right),\forall x\in[\theta,\infty)$$

Therefore, $f_m(x)$ uniformly converges to $f(x)$ on $x\in[\theta,\infty)$. Further, by Theorem 7.12 of Rudin (1976), $C(x)$ is continuous on $[\theta,\infty)$. Notice that the continuity holds for any arbitrarily small $\theta>0$, then the conclusion is drawn. □

EC.1.3. Proof of Theorem 1

As stated in Equation (3.2), the PCS can be formulated as follows,

$$\mathrm{PCS} \geq \Pr\left\{ ck \geq \arg\min_{n\geq 1} \bar{X}_1(n) + \sum_{i=2}^{k} N_i\left(\min_{n\geq 1} \bar{X}_1(n)\right)\right\}.$$

Then, to complete the proof, it suffices to build several useful properties on the sum of the boundary-crossing times of non-best alternatives, $\sum_{i=2}^{k} N_i\left(\min_{n\geq 1} \bar{X}_1(n)\right)$.

To begin with, we first set up some necessary notations. For any given $c > C(\gamma/\overline{\sigma})$, we can arbitrarily select a constant ϵ with $0 < 2\epsilon < c - C(\gamma/\overline{\sigma})$. Since $C(\cdot)$ is a continuous and monotonically decreasing function by Lemma 2, we can always choose a unique positive constant $\gamma_\epsilon^- \in (0, \gamma)$ such that

$$C\left(\frac{\gamma-\gamma_\epsilon^-}{\overline{\sigma}}\right) = c - 2\epsilon. \tag{EC.1.2}$$

Suppose that the condition $A_\epsilon = \{\min_{n\geq 1} \bar{X}_1(n) > \mu_1 - \gamma_\epsilon^-\}$ holds. Because $\mu_1 - \mu_i \geq \gamma$ and $\sigma_i^2 \leq \overline{\sigma}^2$, we have from Equation (4.4) that

$$\sum_{i=2}^{k} N_i\left(\min_{n\geq 1} \bar{X}_1(n)\right) \leq \sum_{i=2}^{k} N_i\left(\mu_1 - \gamma_\epsilon^-\right) \leq \sum_{i=2}^{k} \tilde{N}_i\left(\frac{\gamma-\gamma_\epsilon^-}{\overline{\sigma}}\right).$$

Notice that $\tilde{N}_i\left(\frac{\gamma-\gamma_\epsilon^-}{\overline{\sigma}}\right), i = 2, 3, \ldots, k$, are independent and identically distributed, we further have, as $k \to \infty$,

$$\frac{1}{k}\sum_{i=2}^{k} N_i\left(\min_{n\geq 1} \bar{X}_1(n)\right) \leq \frac{1}{k}\sum_{i=2}^{k} \tilde{N}_i\left(\frac{\gamma-\gamma_\epsilon^-}{\overline{\sigma}}\right) \xrightarrow{\text{a.s.}} \mathrm{E}\left[\tilde{N}_i\left(\frac{\gamma-\gamma_\epsilon^-}{\overline{\sigma}}\right)\right] = C\left(\frac{\gamma-\gamma_\epsilon^-}{\overline{\sigma}}\right) = c - 2\epsilon, \tag{EC.1.3}$$

according to the strong law of large numbers. The last equality above holds by the choice of γ_ϵ^- given in Equation (EC.1.2). We let

$$\Omega_\epsilon = \left\{\frac{1}{k}\sum_{i=2}^{k} N_i\left(\min_{n\geq 1} \bar{X}_1(n)\right) - (c - 2\epsilon) \leq \epsilon\right\}.$$

Then by the definition of the strong law of almost sure convergence and Equation (EC.1.3),

$$\begin{aligned}
\lim_{k\to\infty} \Pr\{\Omega_\epsilon | A_\epsilon\} &\geq \lim_{k\to\infty} \Pr\left\{\frac{1}{k}\sum_{i=2}^{k} \tilde{N}_i\left(\frac{\gamma-\gamma_\epsilon^-}{\overline{\sigma}}\right) - (c-2\epsilon) \leq \epsilon \middle| A_\epsilon\right\} \\
&= \lim_{k\to\infty} \Pr\left\{\frac{1}{k}\sum_{i=2}^{k} \tilde{N}_i\left(\frac{\gamma-\gamma_\epsilon^-}{\overline{\sigma}}\right) - (c-2\epsilon) \leq \epsilon\right\} \\
&= 1.
\end{aligned} \tag{EC.1.4}$$

Now we are ready to study the PCS formulated in Equation (3.2). To be specific,

$$
\begin{aligned}
\text{PCS} &\geq \Pr\left\{ ck \geq \arg\min_{n\geq 1} \bar{X}_1(n) + \sum_{i=2}^{k} N_i\left(\min_{n\geq 1}\bar{X}_1(n)\right)\right\} \\
&\geq \Pr\left\{\left\{ ck \geq \arg\min_{n\geq 1} \bar{X}_1(n) + \sum_{i=2}^{k} N_i\left(\min_{n\geq 1}\bar{X}_1(n)\right)\right\}\cap A_\epsilon \cap \Omega_\epsilon\right\} \\
&\geq \Pr\left\{\left\{ ck \geq \arg\min_{n\geq 1} \bar{X}_1(n) + (c-2\epsilon)k + \epsilon k\right\}\cap A_\epsilon \cap \Omega_\epsilon\right\} \quad (\text{by definition of } \Omega_\epsilon) \\
&= \Pr\left\{\left\{ \epsilon k \geq \arg\min_{n\geq 1} \bar{X}_1(n)\right\}\cap A_\epsilon \cap \Omega_\epsilon\right\} \\
&\geq \Pr\{A_\epsilon \cap \Omega_\epsilon\} - \Pr\left\{\arg\min_{n\geq 1}\bar{X}_1(n) > \epsilon k\right\} \\
&= \Pr\{\Omega_\epsilon | A_\epsilon\}\Pr\{A_\epsilon\} - \Pr\left\{\arg\min_{n\geq 1}\bar{X}_1(n) > \epsilon k\right\},
\end{aligned}
$$

where the last inequality arises due to the fact that $\Pr\{A\cap B\} = \Pr\{B\} - \Pr\{A^c \cap B\} \geq \Pr\{B\} - \Pr\{A^c\}$. From Equation (EC.1.4), letting $k\to\infty$ in above gives rise to

$$
\begin{aligned}
\liminf_{k\to\infty} \text{PCS} &\geq \limsup_{k\to\infty} \Pr\{\Omega_\epsilon|A_\epsilon\}\Pr\{A_\epsilon\} - \Pr\left\{\arg\min_{n\geq 1}\bar{X}_1(n) = \infty\right\} \\
&= \Pr\left\{\min_{n\geq 1}\bar{X}_1(n) > \mu_1 - \gamma_\epsilon^-\right\} - \Pr\left\{\arg\min_{n\geq 1}\bar{X}_1(n) = \infty\right\}.
\end{aligned} \tag{EC.1.5}
$$

Notice that the above statement holds for any ϵ sufficiently close to zero. Besides, it is clear that $C(\cdot)$ is continuous, and therefore, $\lim_{\epsilon\to 0}\gamma_\epsilon^- = \gamma_0$ where γ_0 is a positive constant given by $C\left(\frac{\gamma-\gamma_0}{\bar{\sigma}}\right) = c$. Then, by Lemma 1, letting $\epsilon \to 0$ yields

$$
\begin{aligned}
\liminf_{k\to\infty} \text{PCS} &\geq \Pr\left\{\min_{n\geq 1}\bar{X}_1(n) > \mu_1 - \gamma_0\right\} - \Pr\left\{\arg\min_{n\geq 1}\bar{X}_1(n) = \infty\right\} \\
&= \Pr\left\{\min_{n\geq 1}\bar{X}_1(n) > \mu_1 - \gamma_0\right\} \\
&= \left[C\left(\frac{\gamma_0}{\sigma_1}\right)\right]^{-1},
\end{aligned}
$$

This concludes the proof of the theorem. □

EC.1.4. Proof of Theorem 2

As stated in Section 4.3, it sufficient to show the following statement

$$
\liminf_{k\to\infty} \text{PICS} \geq \Pr\left\{\min_{n\geq 1}\bar{X}_1(n) < \mu_1 - \gamma_0\right\} = \Pr\{\exists\, n\geq 1 \text{ s.t. } \bar{X}_1(n) < \mu_1 - \gamma_0\}. \tag{EC.1.6}
$$

By Lemma 2, $C(\cdot)$ is a strictly decreasing and continuous function in $(0,\infty)$. Then, for any constant ϵ with $0 < 2\epsilon < c - C(\gamma/\sigma)$, we can always choose $0 < \gamma_\epsilon^-, \gamma_\epsilon^+ < \gamma$ such that

$$
C\left(\frac{\gamma-\gamma_\epsilon^-}{\sigma}\right) = c - 2\epsilon \text{ and } C\left(\frac{\gamma-\gamma_\epsilon^+}{\sigma}\right) = c + 2\epsilon. \tag{EC.1.7}
$$

Further, it is easy to check that $\gamma_\epsilon^- < \gamma_0 < \gamma_\epsilon^+$ and $\gamma_\epsilon^-, \gamma_\epsilon^+ \to \gamma_0$ as $\epsilon \to 0$.

Let $m_\epsilon = \inf\{n \geq 1 : \bar{X}_1(n) < \mu_1 - \gamma_\epsilon^+\}$. Suppose for now the condition $B_\epsilon = \{m_\epsilon < \infty, \min_{1\leq n\leq m_\epsilon - 1} \bar{X}_1(n) \geq \mu_1 - \gamma_\epsilon^-\}$ is forced. When $m_\epsilon = 1$, we set $\bar{X}_1(m_\epsilon - 1) = \bar{X}_1(0) = \infty$. From the boundary-crossing perspective, the greedy procedure would produce an incorrect selection whenever the sample mean of alternative 1 falls below $\mu_1 - \gamma_\epsilon^+$ but the remaining sampling budget is not enough to let alternative 1 become the sample best again. In other words, we may write

$$\text{PICS} \geq \Pr\left\{\sum_{i=2}^k N_i\left(\min_{1\leq n\leq m_\epsilon-1} \bar{X}_1(n)\right) + m_\epsilon \leq B < \sum_{i=2}^k N_i\left(\min_{1\leq n\leq m_\epsilon} \bar{X}_1(n)\right) + m_\epsilon\right\}. \quad \text{(EC.1.8)}$$

Under the SC-CV and condition B_ϵ, the sum of boundary-crossing times satisfies that, as $k \to \infty$,

$$\begin{aligned}
&\frac{1}{k-1}\sum_{i=2}^k N_i\left(\min_{1\leq n\leq m_\epsilon-1} \bar{X}_1(n)\right)\\
&\leq \frac{1}{k-1}\sum_{i=2}^k N_i\left(\mu_1 - \gamma_\epsilon^-\right) = \frac{1}{k-1}\sum_{i=2}^k \tilde{N}_i\left(\frac{\gamma - \gamma_\epsilon^-}{\sigma}\right) \to C\left(\frac{\gamma - \gamma_\epsilon^-}{\sigma}\right) = c - 2\epsilon, \ \text{a.s.}
\end{aligned}$$

by the strong law of large numbers and Equation (EC.1.7). Likewise, as $k \to \infty$,

$$\begin{aligned}
&\frac{1}{k-1}\sum_{i=2}^k N_i\left(\min_{1\leq n\leq m_\epsilon} \bar{X}_1(n)\right)\\
&\geq \frac{1}{k-1}\sum_{i=2}^k N_i\left(\mu_1 - \gamma_\epsilon^+\right) = \frac{1}{k-1}\sum_{i=2}^k \tilde{N}_i\left(\frac{\gamma - \gamma_\epsilon^+}{\sigma}\right) \to C\left(\frac{\gamma - \gamma_\epsilon^+}{\sigma}\right) = c + 2\epsilon, \ \text{a.s.}
\end{aligned}$$

Therefore, given any $\epsilon > 0$, the collections of sample paths denoted by

$$\begin{aligned}
\Omega_\epsilon^- &= \left\{\frac{1}{k-1}\sum_{i=2}^k N_i\left(\min_{1\leq n\leq m_\epsilon-1} \bar{X}_1(n)\right) - (c - 2\epsilon) \leq \epsilon\right\}\\
\Omega_\epsilon^+ &= \left\{\frac{1}{k-1}\sum_{i=2}^k N_i\left(\min_{1\leq n\leq m_\epsilon} \bar{X}_1(n)\right) - (c + 2\epsilon) \geq -\epsilon\right\}
\end{aligned} \quad \text{(EC.1.9)}$$

satisfy that

$$\lim_{k\to\infty} \Pr\{\Omega_\epsilon^+ | B_\epsilon\} = \lim_{k\to\infty} \Pr\{\Omega_\epsilon^- | B_\epsilon\} = 1. \quad \text{(EC.1.10)}$$

Now we are ready to study the PICS in Equation (EC.1.8). Plugging Equation (EC.1.8) and Equation (EC.1.9) into Equation (EC.1.8) gives

$$\begin{aligned}
\text{PICS} &\geq \Pr\left\{\left\{\sum_{i=2}^k N_i\left(\min_{1\leq n\leq m_\epsilon-1} \bar{X}_1(n)\right) + m_\epsilon \leq B < \sum_{i=2}^k N_i\left(\min_{1\leq n\leq m_\epsilon} \bar{X}_1(n)\right) + m_\epsilon\right\} \bigcap \Omega_\epsilon^+ \bigcap \Omega_\epsilon^- \bigcap B_\epsilon\right\}\\
&\geq \Pr\left\{\{(k-1)(c-\epsilon) + m_\epsilon \leq ck < (k-1)(c+\epsilon) + m_\epsilon\} \cap \Omega_\epsilon^+ \cap \Omega_\epsilon^- \cap B_\epsilon\right\}\\
&= \Pr\left\{\{c - (k-1)\epsilon < m_\epsilon \leq c + (k-1)\epsilon\} \cap \Omega_\epsilon^+ \cap \Omega_\epsilon^- \cap B_\epsilon\right\}\\
&\geq \Pr\left\{\{c - (k-1)\epsilon < m_\epsilon \leq c + (k-1)\epsilon\} \cap B_\epsilon\right\} - \Pr\left\{(\Omega_\epsilon^+)^c \cap B_\epsilon\right\} - \Pr\left\{(\Omega_\epsilon^-)^c \cap B_\epsilon\right\}\\
&= \Pr\left\{\{c - (k-1)\epsilon < m_\epsilon \leq c + (k-1)\epsilon\} \cap B_\epsilon\right\} - \Pr\left\{(\Omega_\epsilon^+)^c | B_\epsilon\right\}\Pr\{B_\epsilon\} - \Pr\left\{(\Omega_\epsilon^-)^c | B_\epsilon\right\}\Pr\{B_\epsilon\},
\end{aligned}$$

where the last inequality holds due to the fact that

$$\begin{aligned}\Pr\{A\cap B\cap C\cap D\} &= \Pr\{A\cap D\} - \Pr\{A\cap (B\cap C)^c\cap D\}\\ &\geq \Pr\{A\cap D\} - \Pr\{(B\cap C)^c\cap D\}\\ &= \Pr\{A\cap D\} - \Pr\{(B^c\cap D)\cup(C^c\cap D)\}\\ &\geq \Pr\{A\cap D\} - \Pr\{B^c\cap D\} - \Pr\{C^c\cap D\}.\end{aligned}$$

Taking $k\to\infty$ on both sides, we can derive

$$\liminf_{k\to\infty}\ \text{PICS} \geq \Pr\left\{m_\epsilon<\infty, \min_{1\leq n\leq m_\epsilon-1}\bar{X}_1(n)\geq \mu_1-\gamma_\epsilon^-\right\} \tag{EC.1.11}$$

from Equation (EC.1.10). The inequality above holds for any sufficiently small $\epsilon>0$. Recall that $\gamma_\epsilon^-\to\gamma_0$ as $\epsilon\to 0$, and accordingly, $m_\epsilon\to m=\inf\{n\geq 1: \bar{X}_1(n)<\mu_1-\gamma_0\}$ almost surely. Letting $\epsilon\to 0$ on the right hand side of the inequality above, we readily obtain that

$$\begin{aligned}\liminf_{k\to\infty}\ \text{PICS} &\geq \limsup_{\epsilon\to 0}\Pr\left\{m_\epsilon<\infty, \min_{1\leq n\leq m_\epsilon-1}\bar{X}_1(n)\geq \mu_1-\gamma_\epsilon^-\right\}\\ &= \Pr\left\{m<\infty, \min_{1\leq n\leq m-1}\bar{X}_1(n)\geq\mu_1-\gamma_0\right\} = \Pr\{\exists\, n\geq 1 \text{ s.t. } \bar{X}_1(n)<\mu_1-\gamma_0\}.\end{aligned}$$

The proof is completed. □

EC.1.5. Proof of Proposition 1

To analyze the greedy procedure's PGS from the boundary-crossing perspective, we regard the minimum of the running maximum of the good alternatives, i.e., $\min_{r\geq 1}\bar{Y}_\delta(r)$, as the boundary. Then, we prove the conclusion based on the PGS statement in Equation (4.7), i.e.,

$$\text{PGS}\geq \Pr\left\{\sum_{i\in\mathcal{N}} N_i\left(\min_{r\geq 1}\bar{Y}_\delta(r)\right)+\arg\min_{r\geq 1}\bar{Y}_\delta(r)\leq B\right\}.$$

Notice that Equation (4.7) exhibits a similar form to the PCS statement in Equation (3.2). Following almost the same analysis of Theorem 1, we obtain an analogous result of Equation (EC.1.5) as

$$\liminf_{k\to\infty}\text{PGS}\geq \limsup_{k\to\infty}\Pr\left\{\min_{r\geq 1}\bar{Y}_\delta(r)>\mu_1-\delta_\epsilon^-\right\}-\liminf_{k\to\infty}\Pr\left\{\arg\min_{r\geq 1}\bar{Y}_\delta(r)=\infty\right\} \tag{EC.1.12}$$

where $0<\epsilon<(c-C(\delta/\overline{\sigma}))/2$ and δ_ϵ^- is chosen to satisfy $C\left(\frac{\delta-\delta_\epsilon^-}{\overline{\sigma}}\right)=c-2\epsilon$. By the continuity of $C(\cdot)$ stated in Lemma 2, $\delta_\epsilon^-\to\delta_0$, as $\epsilon\to 0$. Letting $\epsilon\to 0$ in Equation (EC.1.12) yields

$$\liminf_{k\to\infty}\text{PGS}\geq \limsup_{k\to\infty}\Pr\left\{\min_{r\geq 1}\bar{Y}_\delta(r)>\mu_1-\delta_0\right\}-\liminf_{k\to\infty}\Pr\left\{\arg\min_{r\geq 1}\bar{Y}_\delta(r)=\infty\right\}. \tag{EC.1.13}$$

Notice that Equation (4.8) provides a reasonable upper bound for $\arg\min_{r\geq 1}\bar{Y}_\delta(r)$. Specifically,

$$\arg\min_{r\geq 1}\bar{Y}_\delta(r) \leq \sum_{i\in\mathcal{G}}\arg\min_{n\geq 1}\bar{X}_i(n),$$

Then, we can conclude that $\Pr\left\{\arg\min_{r\geq 1}\bar{Y}_\delta(r)=\infty\right\} \leq \sum_{i\in\mathcal{G}}\Pr\left\{\arg\min_{n\geq 1}\bar{X}_i(n)=\infty\right\}$, which amounts to zero by Lemma 1 and the assumption that $\limsup_{k\to\infty}|\mathcal{G}|<\infty$. Meanwhile, because $\min_{r\geq 1}\bar{Y}_\delta(r)=\max_{i\in\mathcal{G}}\min_{n\geq 1}\bar{X}_i(n)$, Equation (EC.1.13) can be further written as

$$\liminf_{k\to\infty}\text{PGS} \geq \limsup_{k\to\infty}\Pr\left\{\min_{r\geq 1}\bar{Y}_\delta(r)>\mu_1-\delta_0\right\} = \limsup_{k\to\infty}\Pr\left\{\exists\, i\in\mathcal{G}:\min_{n\geq 1}\bar{X}_i(n)>\mu_1-\delta_0\right\}$$

It is clear that the right-hand-side probability is at least $\Pr\left\{\min_{n\geq 1}\bar{X}_1(n)>\mu_1-\delta_0\right\}$. The proof is completed. □

EC.2. Technical Supplements for Section 5

EC.2.1. Proof of Lemma 3

Lemma 3 is a generalization of Lemma 1. Similar to the proof of Lemma 1, we prove Lemma 3 by contradiction. For any $n_0\geq 1$, we let $M(\omega;n_0)=\arg\min_{n\geq n_0}\bar{Z}(n;\omega)$ denote the location of the minima of the running-average process $\{\bar{Z}(n;\omega), n=n_0,n_0+1,\ldots\}$. Suppose that the lemma is violated. Then, there must exist a collection of sample paths $\Omega(n_0)=\{\omega:\bar{Z}(n;\omega), n=n_0,n_0+1,\ldots\}$ such that

$$M(\omega;n_0)=\infty \text{ for any } \omega\in\Omega(n_0), \text{ and } \Pr\{\omega\in\Omega(n_0)\}>0. \tag{EC.2.1}$$

By repeating similar arguments in Section EC.1.1 for the case $n_0=1$, we can derive

$$\Pr\{\omega\in\Omega(n_0)\} \leq \Pr\left\{\min_{n\geq n_0}\bar{Z}(n;\omega)\geq 0\right\} = \Pr\{N(0;n_0)=\infty\}.$$

where $N(0;n_0)$ is defined in Section 5.2 and we have that $\Pr\{N(0;n_0)=\infty\}=0$ by Lemma 4. Clearly, this contradicts with Equation (EC.2.1). The proof is completed. □

EC.2.2. Proof of Theorem 3

Before moving to show Theorem 3, we first prepare the following Lemma 7.

Lemma 7 *Given any positive integer n_0, let $N(x;n_0)=\inf\{n\geq n_0:\bar{Z}(n)<x\}$. Then, $C(x;n_0)=\mathrm{E}[N(x;n_0)]$ is a strictly decreasing and continuous function on $x\in(0,\infty)$.*

Proof: Firstly, we show that $C(x;n_0)$ is strictly decreasing in x for any fixed n_0. Set x_1 and x_2 with $0<x_1<x_2<\infty$. Taking a sample-path viewpoint, it is straightforward to see that $N(x_1;n_0)\geq N(x_2;n_0)$. Moreover, we find that the strict inequality holds with a non-zero probability, i.e.,

$$\begin{aligned}\Pr\{N(x_1;n_0)>N(x_2;n_0)\} &\geq \Pr\{N(x_1;n_0)>n_0, N(x_2;n_0)=n_0\}\\ &= \Pr\left\{x_1\leq\bar{Z}(n_0)<x_2\right\} = \Phi\left(\sqrt{n_0}x_2\right)-\Phi\left(\sqrt{n_0}x_1\right)>0.\end{aligned}$$

As a consequence, we can conclude that $C(x_1;n_0) > C(x_2;n_0)$.

Secondly, we prove the continuity of $C(x;n_0)$ with respect to x. For any $x_0 \geq \epsilon > 0$, we denote a sequence of numbers $\{x_m, m = 1,2,\ldots\} \subset [\epsilon, \infty)$ such that $\lim_{m\to\infty} x_m = x_0$. Recalling that $N(x;n_0)$ is decreasing in x, we can derive

$$|C(x_m;n_0)| = \mathrm{E}[N(x_m;n_0)] \leq \mathrm{E}[N(\epsilon;n_0)].$$

From Lemma 4, $\mathrm{E}[N(\epsilon;n_0)] = C(\epsilon;n_0) < \infty$. Further, by the Lebesgue's dominated convergence theorem,

$$\lim_{m\to\infty} C(x_m;n_0) = \lim_{m\to\infty} \mathrm{E}[N(x_m;n_0)] = \mathrm{E}\left[\lim_{m\to\infty} N(x_m;n_0)\right] = \mathrm{E}\left[N(x_0;n_0)\right] = C(x_0;n_0).$$

In other words, $C(x;n_0)$ is continuous in $x \in [\epsilon,\infty)$. Because the continuity holds for any $\epsilon > 0$, we conclude that $C(x;n_0)$ is continuous in $x \in (0,\infty)$. The proof is completed. □

Proof of Theorem 3: The proof is conducted based on the proofs of Theorems 1 and 2. To analyze the explore-first greedy procedure's PCS from the boundary-crossing perspective, we regard the minimum of the running average of the best alternative after the exploration phase, i.e., $\min_{n\geq n_0} \bar{X}_1(n)$, as the boundary. Indeed, following almost the same analysis, we can conclude the following result which is analogous to Equation (EC.1.5)

$$\liminf_{k\to\infty} \text{ PCS } \geq \Pr\left\{\min_{n\geq n_0} \bar{X}_1(n) > \mu_1 - \gamma_\epsilon^-\right\} - \Pr\left\{\arg\min_{n\geq n_0} \bar{X}_1(n) = \infty\right\}, \tag{EC.2.2}$$

where

$$C\left(\frac{\gamma - \gamma_\epsilon^-}{\bar{\sigma}}; n_0\right) = c - 2\epsilon. \tag{EC.2.3}$$

By the continuity of $C(\cdot;n_0)$ stated in Lemma 7, it is easy to check that $\gamma_\epsilon^- \to \gamma_0$ as $\epsilon \to 0$. Letting $\epsilon \to$ in Equation (EC.2.2), we can obtain

$$\begin{aligned}\liminf_{k\to\infty} \text{ PCS} &\geq \Pr\left\{\min_{n\geq n_0} \bar{X}_1(n) > \mu_1 - \gamma_0\right\} - \Pr\left\{\arg\min_{n\geq n_0} \bar{X}_1(n) = \infty\right\} \\ &= \Pr\left\{\min_{n\geq n_0} \bar{X}_1(n) > \mu_1 - \gamma_0\right\}.\end{aligned} \tag{EC.2.4}$$

The second inequality arises from Lemma 3.

In addition, we define $m_\epsilon(n_0) = \inf\{n \geq n_0 : \bar{X}_1(n) < \mu_1 - \gamma_\epsilon^-\}$ with γ_ϵ^- defined in Equation (EC.2.3). Under the SC-CV, we derive the analogous result of Equation (EC.1.11), namely,

$$\liminf_{k\to\infty} \text{ PICS} \geq \Pr\left\{m_\epsilon(n_0) < \infty, \min_{1\leq n\leq m_\epsilon(n_0)-1} \bar{X}_1(n) \geq \mu_1 - \gamma_\epsilon^-\right\}.$$

Recalling that $\gamma_\epsilon^- \to \gamma_0$ as $\epsilon \to 0$, we can have $m_\epsilon(n_0) \to m(n_0) = \inf\{n \geq n_0 : \bar{X}_1(n) < \mu_1 - \gamma_0\}$ almost surely. Therefore, by letting $\epsilon \to 0$, the above statement is further written as

$$\liminf_{k\to\infty} \text{ PICS} \geq \Pr\left\{m(n_0) < \infty, \min_{1\leq n\leq m(n_0)-1} \bar{X}_1(n) \geq \mu_1 - \gamma_0\right\} = \Pr\left\{m(n_0) < \infty\right\}. \tag{EC.2.5}$$

Considering Equation (EC.2.4) and Equation (EC.2.5), the conclusion is drawn. □

EC.2.3. The EFG Procedure's Sample Optimality Regarding PGS

Proposition 2 *Suppose that $\sigma_i^2 \le \bar{\sigma}^2 < \infty$ for all $i = 1, \ldots, k$ and $\limsup_{k\to\infty} |\mathcal{G}| < \infty$ for a given $\delta > 0$. If the total sampling budget B satisfies $B/k = n_0 + n_g$ and $n_g > C\left(\frac{\delta}{\bar{\sigma}}; n_0\right) - n_0$, the PGS of the explore-first greedy procedure satisfies*

$$\liminf_{k\to\infty} \text{PGS} \ge \limsup_{k\to\infty} \Pr\left\{\exists\ i \in \mathcal{G} : \min_{n\ge n_0} \bar{X}_i(n) > \mu_1 - \delta_0\right\} \ge \Pr\left\{\min_{n\ge n_0} \bar{X}_1(n) > \mu_1 - \delta_0\right\},$$

where δ_0 is a positive constant such that $\delta_0 \in (0, \delta)$ and $C\left(\frac{\delta-\delta_0}{\bar{\sigma}}; n_0\right) = n_0 + n_g$.

Proof. For the explore-first greedy procedure, we again let $\bar{Y}_\delta(r) = \max_{i\in\mathcal{G}} \bar{X}_i(n_i)$, where $n_i \ge n_0$ denotes the sample size of alternative i after the exploration phase and $r = \sum_{i\in\mathcal{G}} n_i$. Note that $r \ge |\mathcal{G}|n_0$. To analyze the explore-first greedy procedure's PGS from the boundary-crossing perspective, we regard the minimum of the running maximum of the good alternatives after the exploration phase, i.e., $\min_{r\ge|\mathcal{G}|n_0} \bar{Y}_\delta(r)$, as the boundary. Moreover, for each alternative $i \in \mathcal{N}$, let $N_i(x; n_0) = \inf\{n \ge n_0 : \bar{X}_i(n) < x\}$ denote the boundary-crossing time after n_0 observations *w.r.t.* a boundary x. Then, analogous to Equation (4.7), we have

$$\text{PGS} \ge \Pr\left\{\sum_{i\in\mathcal{N}} N_i\left(\min_{r\ge|\mathcal{G}|n_0} \bar{Y}_\delta(r); n_0\right) + \arg\min_{r\ge|\mathcal{G}|n_0} \bar{Y}_\delta(r) \le B\right\}. \tag{EC.2.6}$$

We prove the conclusion based on the PGS statement in Equation (EC.2.6). Notice that Equation (EC.2.6) exhibits a similar form to the PCS statement in Equation (5.2). Following the same logic of obtaining Equation (EC.1.12) for proving Proposition 1, we can obtain an analogous result of Equation (EC.2.2) as follows,

$$\liminf_{k\to\infty} \text{PGS} \ge \limsup_{k\to\infty} \Pr\left\{\min_{r\ge|\mathcal{G}|n_0} \bar{Y}_\delta(r) > \mu_1 - \delta_\epsilon^-\right\} - \liminf_{k\to\infty} \Pr\left\{\arg\min_{r\ge|\mathcal{G}|n_0} \bar{Y}_\delta(r) = \infty\right\}, \tag{EC.2.7}$$

where $0 < \epsilon < (c - C(\delta/\bar{\sigma}; n_0))/2$ and δ_ϵ^- is chosen to satisfy $C\left(\frac{\delta-\delta_\epsilon^-}{\bar{\sigma}}; n_0\right) = c - 2\epsilon$. By the continuity of $C(\cdot; n_0)$ stated in Lemma 7, it is easy to check that $\delta_\epsilon^- \to \delta_0$ as $\epsilon \to 0$. Letting $\epsilon \to 0$ in Equation (EC.2.7) yields

$$\liminf_{k\to\infty} \text{PGS} \ge \limsup_{k\to\infty} \Pr\left\{\min_{r\ge|\mathcal{G}|n_0} \bar{Y}_\delta(r) > \mu_1 - \delta_0\right\} - \liminf_{k\to\infty} \Pr\left\{\arg\min_{r\ge|\mathcal{G}|n_0} \bar{Y}_\delta(r) = \infty\right\}, \tag{EC.2.8}$$

Notice that one reasonable upper bound for $\arg\min_{r\ge|\mathcal{G}|n_0} \bar{Y}_\delta(r)$ is that

$$\arg\min_{r\ge|\mathcal{G}|n_0} \bar{Y}_\delta(r) \le \sum_{i\in\mathcal{G}} \arg\min_{n\ge n_0} \bar{X}_i(n),$$

i.e., the process $\left\{\bar{Y}_\delta(r), r \ge |\mathcal{G}|n_0\right\}$ reaches its minimum before all $\left\{\bar{X}_i(n), n \ge n_0\right\}$ processes reach their minimums for all $i \in \mathcal{G}$. Then, we can conclude that $\Pr\left\{\arg\min_{r\ge|\mathcal{G}|n_0} \bar{Y}_\delta(r) = \infty\right\} \le$

$\sum_{i\in\mathcal{G}} \Pr\left\{\arg\min_{n\geq n_0} \bar{X}_i(n)=\infty\right\}$, which amounts to zero by Lemma 3 and the assumption that $\limsup_{k\to\infty}|\mathcal{G}|<\infty$. Meanwhile, because $\min_{r\geq|\mathcal{G}|n_0}\bar{Y}_\delta(r)=\max_{i\in\mathcal{G}}\min_{n\geq n_0}\bar{X}_i(n)$, Equation (EC.2.8) can be further written as

$$\liminf_{k\to\infty}\mathrm{PGS}\geq\limsup_{k\to\infty}\Pr\left\{\min_{r\geq|\mathcal{G}|n_0}\bar{Y}_\delta(r)>\mu_1-\delta_0\right\}=\limsup_{k\to\infty}\Pr\left\{\exists i\in\mathcal{G}:\min_{n\geq n_0}\bar{X}_i(n)>\mu_1-\delta_0\right\}.$$

It is clear that the right-hand-side probability is at least $\Pr\left\{\min_{n\geq n_0}\bar{X}_1(n)>\mu_1-\delta_0\right\}$. The proof is completed. □

EC.2.4. Proof of Lemma 6

For any fixed n_0, the function $C(x;n_0)$ for any $x>0$ can be given by

$$\begin{aligned}C(x;n_0)-n_0&=\mathrm{E}\left[\inf\{n\geq n_0:\bar{Z}(n)<x\}\right]-n_0\\&=\mathrm{E}\left[\inf\{n\geq 0:\bar{Z}(n+n_0)<x\}\right]:=\mathrm{E}[M(x;n_0)]\end{aligned}$$

where we denote $M(x;n_0)=\inf\{n\geq 0:\bar{Z}(n+n_0)<x\}$ above for simplicity of notation. By this definition, we must have the following fact holds: when $M(n_0)\geq n\geq 1$ is true, we must have that the event $\bar{Z}(n_0+n-1)\geq x$ holds. This indicates that

$$\mathrm{E}[M(n_0;x)]=\sum_{n=1}^{\infty}\Pr\{M(n_0;x)\geq n\}\leq\sum_{n=1}^{\infty}\Pr\{\bar{Z}(n+n_0-1)\geq x\}=\sum_{n=n_0}^{\infty}\Pr\{\bar{Z}(n)\geq x\}=\sum_{n=n_0}^{\infty}\bar{\Phi}(\sqrt{n}x),$$

where $\bar{\Phi}(\cdot)$ denotes the complementary cumulative distribution function of a standard normal distribution. Because $\bar{\Phi}(x)\leq\exp(-x^2/2)$ for $x>0$, we further have

$$\mathrm{E}[M(n_0;x)]\leq\sum_{n=n_0}^{\infty}\exp(-nx^2/2)=\frac{\exp\left(-n_0x^2/2\right)}{1-\exp(-x^2/2)}.$$

Set $x=(\gamma-\gamma_0)/\overline{\sigma}$, and we can have

$$n_g=C\left(\frac{\gamma-\gamma_0}{\overline{\sigma}}\right)-n_0=\mathrm{E}\left[M\left(\frac{\gamma-\gamma_0}{\overline{\sigma}};n_0\right)\right]\leq\frac{\exp\left(-n_0\frac{(\gamma-\gamma_0)^2}{2\overline{\sigma}^2}\right)}{1-\exp\left(-\frac{(\gamma-\gamma_0)^2}{2\overline{\sigma}^2}\right)}.\tag{EC.2.9}$$

Letting $\kappa=\frac{(\gamma-\gamma_0)^2}{2\overline{\sigma}^2}$ and $\beta=(1-e^{-\kappa})^{-1}$ leads to the conclusion. The proof is completed. □

EC.2.5. The Enhanced Explore-First Greedy (EFG$^+$) Procedure and Proof of Theorem 5

Procedure 3 **Enhanced Explore-First Greedy (EFG$^+$) Procedure**

Require: k alternatives $X_1,\dots,X_k$, the total sampling budget $B=(n_{sd}+n_0+n_g)k$, and the total number of groups G satisfying that $2^G-1\le k$ and $2\le G\le n_0$.

For each alternative $i=1,\dots,k$, take n_{sd} independent observations $X_{i1},\dots,X_{in_{sd}}$ and set $\bar{X}_i^{sd}=\frac{1}{n_{sd}}\sum_{j=1}^{n_{sd}}X_{ij}$.

According to $\bar{X}_i^{sd}$, sort the alternatives in descending order as $\{(1),(2),\dots,(k)\}$.

Let I^r denote the group of alternatives for $r=1,\dots,G$. Set $\Delta=2^G-1$. Then, let $I^1=\{(1),(2),\dots,(\lfloor k/\Delta\rfloor)\}$, $I^r=\{(\lfloor k2^{r-2}/\Delta\rfloor+1),\dots,(\lfloor k2^{r-1}/\Delta\rfloor)\}$ for $r=2,\dots,G-1$, and let $I^G=\{(\lfloor k2^{G-2}/\Delta\rfloor+1),\dots,(k)\}$.

for $r\in\{1,2,3,\dots,G\}$ **do**

  Set the sample size allocated to each alternative $i\in I^r$ as $n^r=\left\lfloor\frac{n_0(2^G-1)}{G2^{r-1}}\right\rfloor$.

  For each alternative $i\in I^r$, take n^r independent observations $X_{i1},\dots,X_{in^r}$, set $\bar{X}_i(n^r)=\frac{1}{n^r}\sum_{j=1}^{n^r}X_{ij}$ and let $n_i=n^r$.

end for

while $\sum_{i=1}^k n_i<(n_0+n_g)k$ **do**

  Let $s=\arg\max_{i=1,\dots,k}\bar{X}_i(n_i)$ and take one observation x_s from alternative s;

  Update $\bar{X}_s(n_s+1)=\frac{1}{n_s+1}\left[n_s\bar{X}_s(n_s)+x_s\right]$ and let $n_s=n_s+1$;

end while

Select $\arg\max_{i\in\{1,\dots,k\}}\bar{X}_i(n_i)$ as the best.

In the following, we first show that the total amount of observations allocated in the exploration phase, $\sum_{r=1}^G n^r|I^r|$, does not exceed the given exploration budget n_0k. Specifically, we have,

$$\sum_{r=1}^{G}n^r|I^r|=\sum_{r=1}^{G-1}n^r|I^r|+n^G\times\left(k-\sum_{r=1}^{G-1}|I^r|\right)=\sum_{r=1}^{G-1}(n^r-n^G)|I^r|+n^Gk.$$

As $n^r\ge n_G$ and $|I^r|$ is chosen as $\left\lfloor\frac{k2^{r-1}}{2^G-1}\right\rfloor$ for $r=1,\dots,G-1$, we can further derive

$$\begin{aligned}\sum_{r=1}^{G}n^r|I^r|&\le\sum_{r=1}^{G-1}(n^r-n^G)\frac{k2^{r-1}}{2^G-1}+n^Gk=\sum_{r=1}^{G-1}n^r\frac{k2^{r-1}}{2^G-1}-n^Gk+n^Gk\\&=\sum_{r=1}^{G}\left\lfloor\frac{n_0(2^G-1)}{G2^{r-1}}\right\rfloor\frac{k2^{r-1}}{2^G-1}\le\sum_{r=1}^{G}\frac{n_0(2^G-1)}{G2^{r-1}}\frac{k2^{r-1}}{2^G-1}=n_0k.\end{aligned}\qquad\text{(EC.2.10)}$$

Now, we study the PCS of the EFG$^+$ procedure. For each alternative i, we denote its allocated sample size in the exploration phase by e_i. For simplicity, we let $\mathbf{e}$ represent the vector $[e_1,\dots,e_k]$. Applying the boundary-crossing perspective, we can formulate the PCS of the EFG$^+$ procedure as

$$\text{PCS} \ge \mathrm{E}_{\mathbf{e}}\left[\Pr\left\{(n_0+n_g)k \ge \arg\min_{n\ge e_1}\bar{X}_1(n)+\sum_{i=2}^{k}N_i\left(\min_{n\ge e_1}\bar{X}_1(n);e_i\right)\,\Big|\,\mathbf{e}\right\}\right]. \quad \text{(EC.2.11)}$$

The above PCS lower bound is harder to analyze compared to that for the Greedy and EFG procedures. It is because the e_i are mutually correlated random variables as they are calculated based on all alternatives' ranking information in the seeding phase. Thus, we consider eliminating the impact of the $\mathbf{e}$ on the PCS lower bound inside the expectation and thus simplify the analysis.

Notice the fact that $n^G \le e_i \le n^1$, and thus $\arg\min_{n\ge n^1}\bar{X}_1(n) \ge \arg\min_{n\ge e_1}\bar{X}_1(n) \ge \min_{n\ge n^G}\bar{X}_1(n)$ from the sample-path viewpoint. Then, for the PCS lower bound in (EC.2.11), we further have

$$\begin{aligned}\text{PCS} &\ge \mathrm{E}_{\mathbf{e}}\left[\Pr\left\{(n_0+n_g)k \ge \arg\min_{n\ge n^1}\bar{X}_1(n)+\sum_{i=2}^{k}N_i\left(\min_{n\ge n^G}\bar{X}_1(n);e_i\right)\,\Big|\,\mathbf{e}\right\}\right]\\ &\ge \Pr\left\{(n_0+n_g)k \ge \arg\min_{n\ge n^1}\bar{X}_1(n)+\sum_{i=2}^{k}N_i\left(\min_{n\ge n^G}\bar{X}_1(n);n^1\right)\right\}. \end{aligned}\quad \text{(EC.2.12)}$$

The second inequality holds because $N_i(x;n)$ is non-decreasing in n for any fixed x. We note that the PCS lower bound in Equation (EC.2.12) shares almost the same form as the PCS lower bound of the EFG procedure in Equation (5.2). Thus, we can prove Theorem 5 by similar arguments of showing the sample optimality of the EFG procedure stated in EC.2.2.

Specifically, facilitated by the condition that the boundary $\min_{n\ge n^G}\bar{X}_1(n)$ is always above the level $\mu_1-\gamma_\epsilon^-$ where γ_ϵ^- is the constant satisfying

$$C\left(\frac{\gamma-\gamma_\epsilon^-}{\bar{\sigma}};n^1\right)=n_0+n_g-2\epsilon, \quad \text{(EC.2.13)}$$

we can get the following result which is analogous to Equation (EC.2.2),

$$\liminf_{k\to\infty}\text{PCS} \ge \Pr\left\{\min_{n\ge n^G}\bar{X}_1(n)>\mu_1-\gamma_\epsilon^-\right\}-\Pr\left\{\arg\min_{n\ge n^1}\bar{X}_1(n)=\infty\right\}. \quad \text{(EC.2.14)}$$

Notice that $\gamma_\epsilon^-\to\gamma_0$ as $\epsilon\to 0$, where γ_0 is the unique constant satisfying $C\left(\frac{\gamma-\gamma_0}{\bar{\sigma}};n^1\right)=n_0+n_g$. Then, by letting $\epsilon\to 0$ in Equation (EC.2.14) and applying Lemma 3, we can further derive

$$\begin{aligned}\liminf_{k\to\infty}\text{PCS} &\ge \Pr\left\{\min_{n\ge n^G}\bar{X}_1(n)>\mu_1-\gamma_0\right\}-\Pr\left\{\arg\min_{n\ge n^1}\bar{X}_1(n)=\infty\right\}\\ &= \Pr\left\{\min_{n\ge n^G}\bar{X}_1(n)>\mu_1-\gamma_0\right\}.\end{aligned}\quad \text{(EC.2.15)}$$

The conclusion of interest is now drawn. □

EC.3. Additional Numerical Experiments & Results

EC.3.1. Additional Numerical Results for the Configurations EM-IV, EM-DV

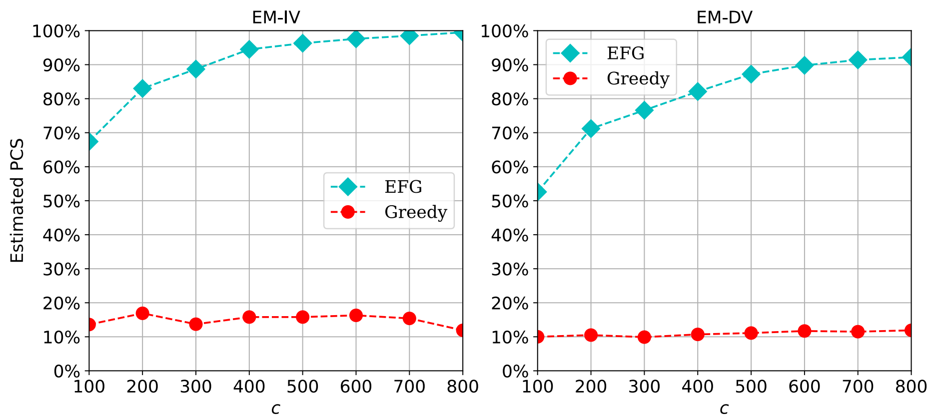

Figure EC.1 Estimated PCS of the EFG procedure and the greedy procedure for different values of c**.** $k = 8192$**.**

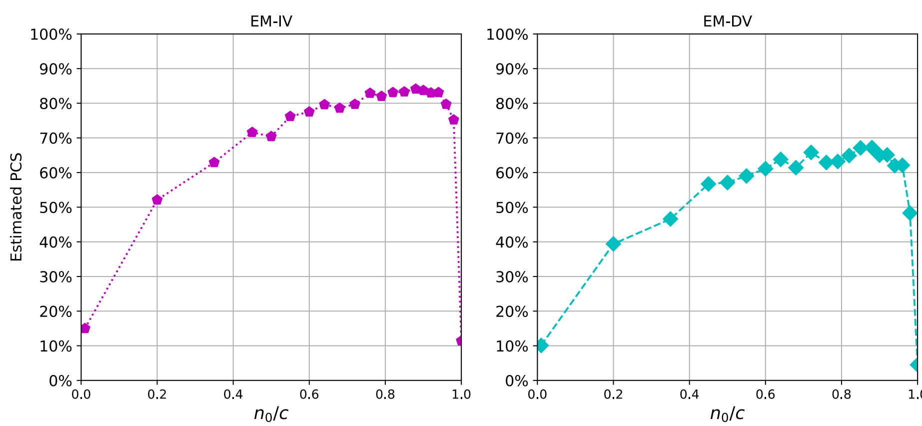

Figure EC.2 Estimated PCS of the EFG procedure for different values of $p = n_0/c$**.** $c = 200$ **and** $k = 8192$**.**

EC.3.2. The Slippage Configuration is Least Favorable for the Greedy Procedures

Theorem 2, together with Theorem 3, show that for the greedy procedures, when $k \to \infty$, the slippage configuration is least favorable among all configurations of means satisfying Assumption 1. We now provide additional numerical results to validate this result and show that the slippage configuration is also least favorable for a finite k.

We consider the following problem configuration

$$\mu_1 = 0.1, \mu_i = -\frac{\lambda(k-2)}{k}, \forall i = 2, \ldots, k; \sigma_i^2 = 1, \forall i = 1, \ldots, k,$$

where $\lambda > 0$ can be a constant or a function of k. In the configuration, the gap between the best and second-best means, i.e., $\mu_2 - \mu_1$, remains 0.1, regardless of the choice of λ. Besides, the inferior means of the non-best alternatives are distributed in the set $[-\lambda, 0]$. In this experiment, we consider the following five choices of $\lambda \in \{0, 0.2, 2, \sqrt{k}, 0.1k\}$. When $\lambda = 0$, the mean configuration is the slippage configuration. As λ is increased, the inferior means become more dispersed. When $\lambda = 0.1k$, the inferior means are progressively worse as k grows.

For each choice of λ, we test the greedy procedure and the EFG procedure with the same experiment setting used in Section 6.1. We then plot the estimated PCS of the two procedures against different k for each λ in Figure EC.3. From Figure EC.3, we can find that for both procedures, the PCS under $\lambda = 0$ is the lowest for different values of k. This shows that the slippage configuration is least favorable for the greedy procedures. Furthermore, as the means of the alternatives become more dispersed, the PCS of the procedures become higher.

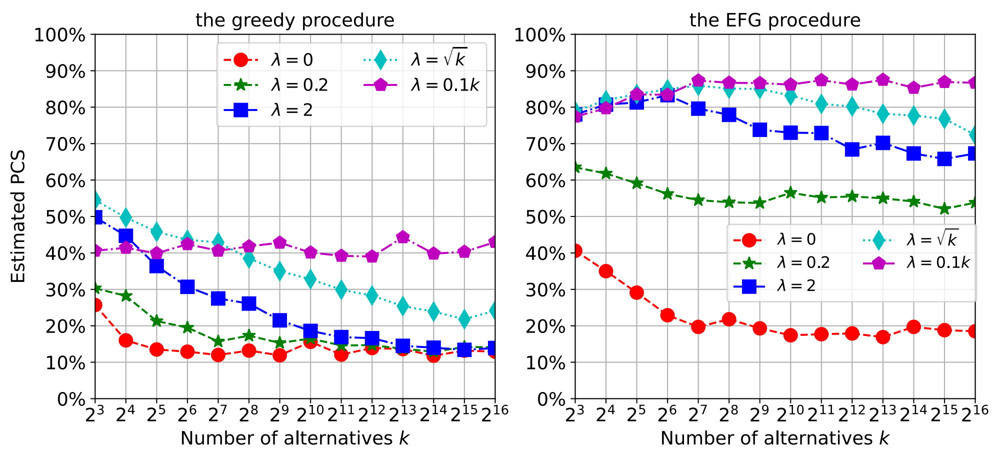

Figure EC.3 Estimated PCS of the greedy procedure and the EFG procedure under different mean configurations.

EC.3.3. Budget Allocation between Exploration and Exploitation

In Section 6.2.2, using the problems with $k = 8192$ alternatives, we show that the PCS of the EFG procedure may be insensitive to the value of p, i.e., the proportion of the exploration budget to the total sampling budget, over a wide near-optimal range of p. Now we study how the optimal p and the near-optimal range of p vary for different k.

In this experiment, we again use the configurations SC-CV, EM-CV, EM-IV and EM-DV, and set the number of alternatives as $k = 2^l$ with l ranging from 3 to 14. For each k, we let $p = 0.02a$ where a is an integer and let a increase from 0 to 50. For every combination of k and p, we use a total sampling budget $B = 200k$ and estimate the PCS based on 2000 independent macro replications. Recall that the PCS curve against different values of p has an inverted U-shape. For each k, we can find an optimal value of p that maximizes the PCS and then estimate a near-optimal range of p by setting the lower (upper) bound of the range as the first (last) p that maintains a PCS within 10% to the maximal PCS. For each problem configuration, we plot the optimal p and the near-optimal range of p against different k in Figure EC.4.

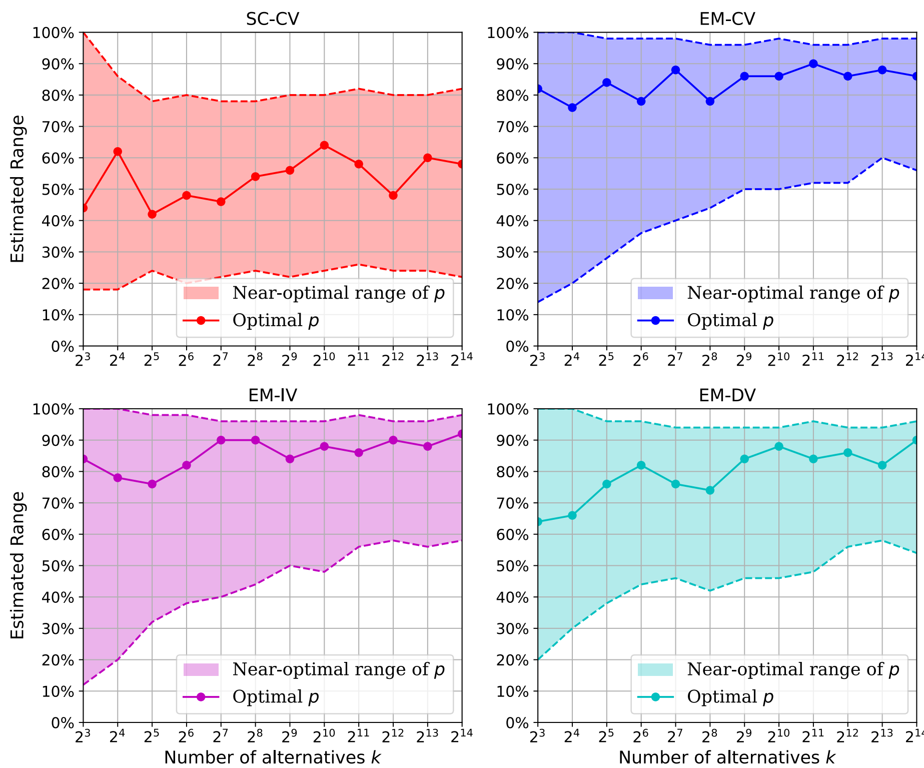

Figure EC.4 **Estimated optimal p and the near-optimal range under different configurations.**

From Figure EC.4, we can observe that the PCS of the EFG procedure is not sensitive to the value of p over a wide range for both small-scale and large-scale problems. For all four configurations, as k grows, the near-optimal range of p shrinks slowly. However, after k is large enough, e.g., $k \geq 2^{13} = 8192$, the range will remain almost the same. Notice that under SC-CV, the optimal p and the lower and upper bounds of the near-optimal range are typically much lower than those under other configurations. This is because SC-CV is a much harder configuration under which the EFG procedure needs more greedy budget to maintain the sample optimality.

EC.3.4. Comparison between Sample-Optimal Fixed-Budget R&S Procedures

EC.3.4.1. Procedures in Comparison. The necessary introduction and the implementation settings of the compared procedures are summarized as follows.

- **EFG procedure and EFG$^+$ procedure**. For the EFG procedure, when solving the practical problem, we allocate 90% of the total sampling budget to the exploration phase. Then, for the EFG$^+$ procedure, we allocate 20% of the total sampling budget to the seeding phase and another

10% to the greedy phase. Moreover, we let $G = 11$, which is selected to let each group have at least one alternative for the problem instances tested in the experiments.

• **FBKT procedure and FBKT-Seeding procedure**. We have briefly introduced the FBKT procedure in Section 1. Readers can go there for a review. The FBKT-Seeding procedure improves the FBKT procedure with a seeding approach, and it is expected to perform better than the FBKT procedure when all alternatives' means are scattered over a wide range. Readers can refer to Hong et al. (2022) for more details. When implementing the procedures, we follow the default settings extracted from the paper.

• **Sequential halving (SH) procedure and the modified SH procedure**. The original SH procedure is developed by Karnin et al. (2013), and the modified procedure is introduced in Appendix B.6 of Zhao et al. (2023). They both proceed in $L = \lceil \log_2 k \rceil$ rounds. In each round l, $l = 1, \ldots, L$, the (modified) SH procedure allocates a sampling budget $\left(T_\ell = \frac{B}{L}\right)$ $T_\ell = \left\lfloor \frac{B}{81k} \cdot \left(\frac{16}{9}\right)^{\ell-1} \cdot \ell \right\rfloor$ equally to the survived alternatives and eliminates the lower half in terms of the sample means. Then, the procedures select the unique alternative that survives all the L rounds as the best.

EC.3.4.2. The Throughput Maximization Problem. The throughput maximization involves a three-stage flow line with a single service station per stage and each service station has *i.i.d.* exponential service time. For stage i, $i = 1, 2, 3$, the service rate is denoted as x_i. In stage 1, there are an infinite number of jobs waiting to be processed in front of the service station. Each job needs to pass through the three stages in sequence and be processed by all three service stations in turn. For stage i, $i = 2, 3$, the buffer size (or storage space) b_i is finite, and it includes the position in the service station. The problem adopts a service protocol called *production blocking*. It means that for stage i, $i = 1, 2$, if the downstream stage is full, the service station is blocked, then it cannot transfer a finished job to the next stage and serve a new job until the buffer of the downstream stage becomes available.

The goal of the problem is to maximize the expected throughput of the three-stage flow line by allocating a total service rate $S_1 \in \mathbb{Z}_+^5$ among the three service stations and a total buffer size $S_2 \in \mathbb{Z}_+^5$ among stage 2 and stage 3. Formally, we write the optimization problem as

$$\begin{aligned}
\max_{x} \quad & \mathrm{E}\left[f\left(x;\xi\right)\right] \\
s.t. \quad & x_1 + x_2 + x_3 = S_1 \\
& b_2 + b_3 = S_2 \\
& x = (x_1, x_2, x_3, b_2, b_3) \in \mathbb{Z}_+^5
\end{aligned}$$

where f represents the flow line's throughput and ξ represents the randomness. A nice feature of the problem is that for each alternative (i.e., feasible solution), we can evaluate its true mean throughput by solving the balancing equations of the underlying Markov chains (see p. 189 of Buzacott and Shanthikumar (1993)). Then, we can easily find the best alternative or identify good alternatives given the IZ parameter δ.

EC.3.4.3. Implementation Settings. For each alternative, every time we generate an observation of the throughput, we independently simulate the flow line until 1050 jobs are finished and estimate the throughput based on the last 50 jobs. In every experiment, we run 500 independent macro replications to estimate the PCS or the PGS, and round the results to two decimal places.

EC.3.4.4. A Comparison Using Configurations with Bounded Means. We now compare the sample-optimal fixed-budget R&S procedures on the configurations EM-IV and EM-DV used previously in Section 6. A common feature of these configurations is that the means of all alternatives are bounded between -1 and 0.1 regardless of how large k is. For each configuration, we set $k = 2^l$ with l ranging from 11 to 16 and set $B = 100k$ for each k. The estimated PCS based on 1000 independent macro replications of the six procedures against different k under each configuration is plotted in Figure EC.5.

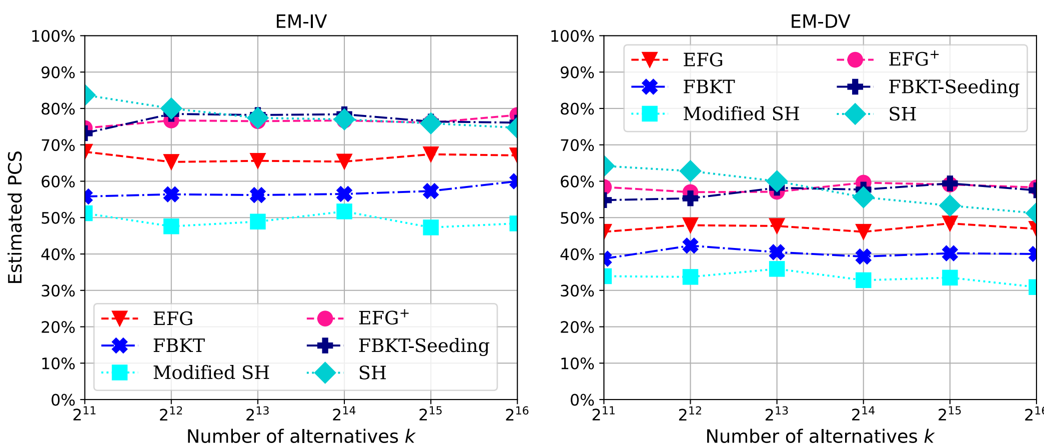

Figure EC.5 A comparison between the EFG procedures, the FBKT procedures, and the SH procedures.

We have the following findings from Figure EC.5. First, the EFG$^+$ procedure performs significantly better than the EFG procedure for the configurations with bounded means. Under both configurations, the gap between the PCS of the two procedures remains around 10% as k grows. Second, when k is large enough, the EFG$^+$ procedure achieves the same level of PCS as the FBKT-Seeding procedure and the SH procedure under EM-IV and a higher PCS than the SH procedure under EM-DV. This result indicates that, although the greedy procedures achieve the sample optimality in a completely different and surprisingly simple manner, they have the potential to perform better than the median elimination procedures.

EC.3.5. Budget Allocation Among the Alternatives

In this subsection, we study how the EFG procedure allocates the sampling budget to all the alternatives. Since the exploration phase is simply an EA phase, we focus on the budget allocation in the greedy phase. Specifically, we study the budget allocation using the following three different problem configurations with a unit variance:

- the slippage configuration where $\mu_1 = 0.1, \mu_i = 0, \forall i = 2, \ldots, k$;
- the configuration with bounded means where $\mu_1 = 0.1, \mu_i = -\frac{2(k-2)}{k}, \forall i = 2, \ldots, k$;
- the configuration with progressively worse means where $\mu_1 = 0.1, \mu_i = -0.1(k-2), \forall i = 2, \ldots, k$.

We test the EFG procedure for each configuration with the same implementation setting used in Section 6.1 and analyze the budget allocation based on 1000 independent macro replications. For a start, we plot the PCS of the EFG procedure under each configuration against different k in Figure EC.6. The PCS curves show that the EFG procedure is sample optimal for all three configurations. Besides, the PCS of the EFG procedure becomes higher as the means become more scattered from the slippage configuration to the configuration with progressively worse means. Next, we plot the average proportion of the non-best alternatives that obtain observations in the greedy phase and the average minimal mean of these alternatives against different k in Figure EC.7. We then plot the average sample size obtained in the greedy phase of the non-best alternatives ever selected in the greedy phase against different k in Figure EC.8, which also displays how the proportion of the greedy budget allocated to the best alternative changes as k increases.

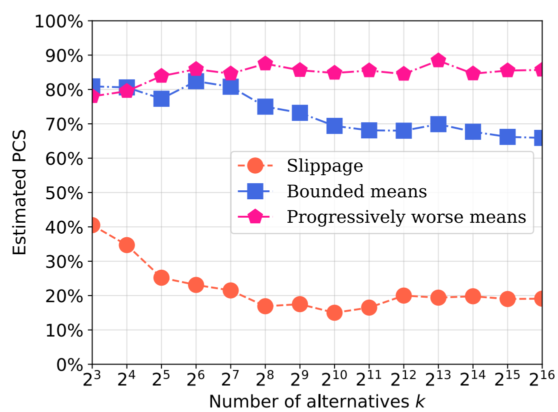

Figure EC.6 Estimated PGS of the EFG procedure under different configurations.

We summarize the findings from Figure EC.7 as follows. First, in the greedy phase of the EFG procedure, only a small proportion of the non-best alternatives obtain observations. Even under the slippage configuration, the proportion is only about 24%. Furthermore, as the means of the alternatives become more dispersed, the aforementioned proportion becomes smaller. Second, the

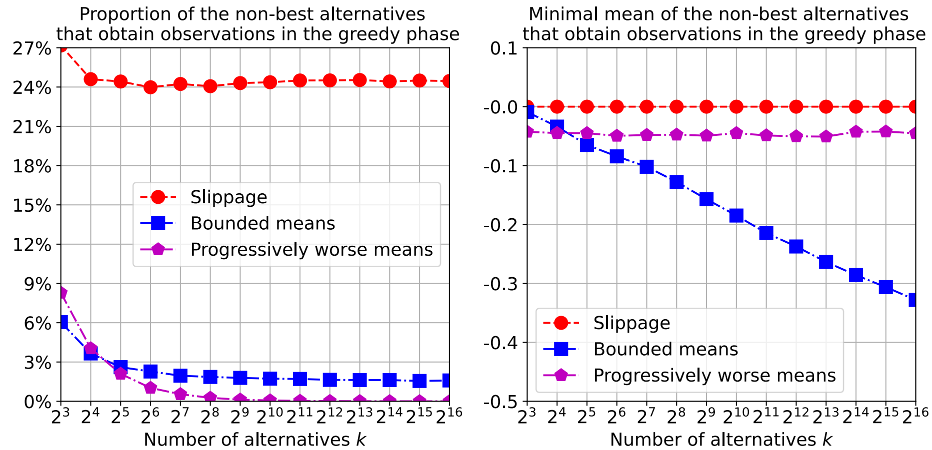

Figure EC.7 **Characterizations of the budget allocation in the greedy phase for the EFG procedure: Part I.**

non-best alternatives that obtain observations in the greedy phase typically have a mean performance very close to the best mean (0.1 in this experiment). For instance, under the configuration with progressively worse means, only a few top alternatives whose means are within 0.2 to the best mean on average obtain observations in the greedy phase.

The above results indicate that the purely exploitative budget allocation of the EFG procedure in the greedy phase may be very efficient. Intuitively speaking, for the non-best alternatives, even if the sample mean is upper-biased after the exploration phase, only the ones close to the best alternative may have a larger sample mean than the best alternative and need more observations to reduce the estimation bias. As demonstrated before, in the greedy phase, the EFG procedure will focus on those alternatives that are worthy of further sampling. Notice that the intuition here may also explain the PCS result shown in Figure EC.6. As the means become more scattered, there will be fewer alternatives that are strong competitors to the best, making it easier for the procedure to reduce the bias causing a possible wrong selection and select the true best.

We obtain the following findings from Figure EC.8. First, for each non-best alternative, once it is selected at any stage in the greedy phase, it may be selected many more times in the subsequent stages. This may be because the non-best alternatives that obtain observations in the greedy phase typically bear a very competitive (e.g., highly upper-biased) sample mean after the exploration phase; then, it may require a considerable number of additional observations to debias the sample mean for each of them to finish the boundary-crossing process. Furthermore, we notice that for the configuration with progressively worse means, the average sample size received in the greedy phase of the non-best alternatives ever selected in the greedy phase grows linearly in k. This may be because even when k is large, as discussed above, only a few top alternatives obtain observations in the greedy phase; then, they share all the greedy budget $n_g k$, which also grows linearly in k. Second, under all three configurations, the best alternative obtains a fixed proportion of the

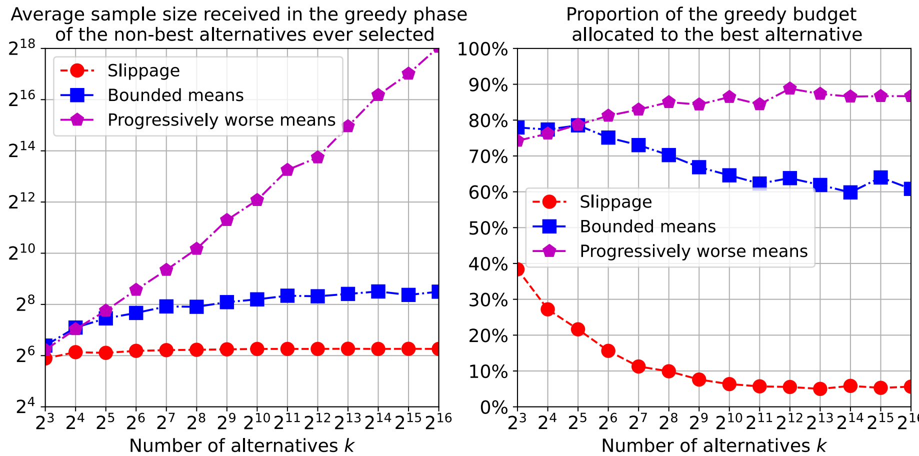

Figure EC.8 Characterizations of the budget allocation in the greedy phase for the EFG procedure: Part II.

sampling budget in the greedy phase, and the proportion will converge as k increases. Notice that under both non-slippage configurations, the proportion exceeds 50% for all values of k, indicating that we may allocate too much sampling budget to the greedy phase.

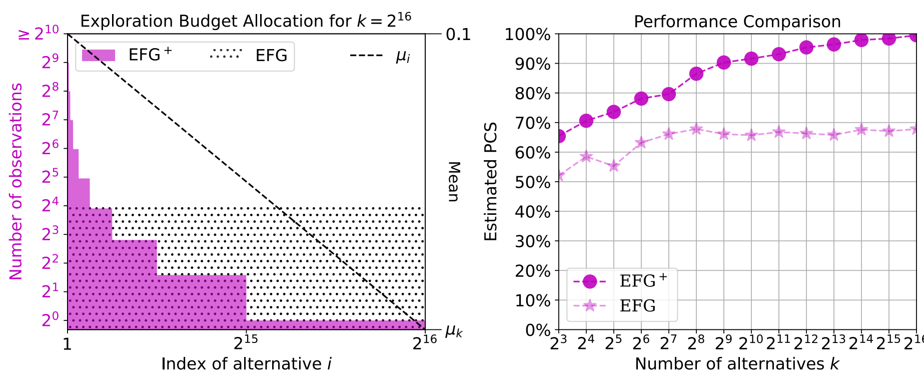

Figure EC.9 A comparison between the EFG procedure and the EFG$^+$ procedure under the configuration with progressively worse means.

To end this subsection, we note that for the configuration with progressively worse means, the EFG procedure may be inefficient in the exploration phase. Under the configuration, a large portion of the alternatives may be clearly inferior; for each of them, only a few observations may be sufficient to reveal its inferiority. However, the EFG procedure allocates the exploration budget equally to all alternatives, which is extravagant since the majority of the total sampling budget is used for exploration. This inefficiency further motivates the use of a seeding phase in the EFG$^+$ procedure which is introduced in Section 5.6. To illustrate the effect of the seeding phase, we compare the exploration budget allocation of the EFG procedure and the EFG$^+$ procedure (using an additional 10% of the sampling budget B in the seeding phase) for $k = 2^{16}$ and $B = 20k$ in Figure EC.9, which also displays the procedures' PCS for different k. From Figure EC.9, we can see that compared to the EFG procedure, the EFG$^+$ procedure allocates much less budget to the inferior alternatives (indexed with large i), validating the seeding phase's usefulness for enhancing the exploration. Consequently, the EFG$^+$ procedure obtains a much higher PCS than the EFG procedure.

EC.3.6. On Assumption 1 and the PGS of the EFG Procedure

Recall that our sample optimality results on the PCS rely on Assumption 1 that requires the gap between the best and second-best means to remain above a positive constant no matter how large k is. Besides, the results on the PGS rely on the assumption that the number of good alternatives remains bounded as k grows to infinity. We now consider problem configurations violating the structural assumptions to study the EFG procedure's performance. In the experiments, unless otherwise specified, we estimate the PCS and PGS for the tested procedure based on 1000 independent macro replications and set the IZ parameter δ as 0.05. Regarding notation, for different k, we use $\gamma(k)$ to denote the mean gap between the best and second-best alternatives and $g(k)$ to denote the total number of good alternatives.

EC.3.6.1. Problem Configurations with Randomly Generated Means. We first consider the following two problem configurations: (1) the configuration with means *i.i.d.* generated from a Normal distribution and a common variance (Normal-CV) under which $\mu_i \sim \text{Normal}(0,1)$, $\sigma_i^2 = 9$, $\forall i = 1,\dots,k$; (2) the configuration with means *i.i.d.* generated from a Beta distribution and a common variance (Beta-CV) under which $\mu_i \sim \text{Beta}(1.5,2)$, $\sigma_i^2 = 9$, $\forall i = 1,\dots,k$. Notice that for the two configurations, $\gamma(k)$ is the difference between the two order statistics $\mu_{(k)}$ and $\mu_{(k-1)}$ of the randomly generated means. For the two configurations, we plot the $\gamma(k)$ and the $g(k)$ for different δ estimated based on one particular sample path against k in Figure EC.10 and Figure EC.11, respectively. Histograms of means for $k = 2^{16}$ are also visualized there.

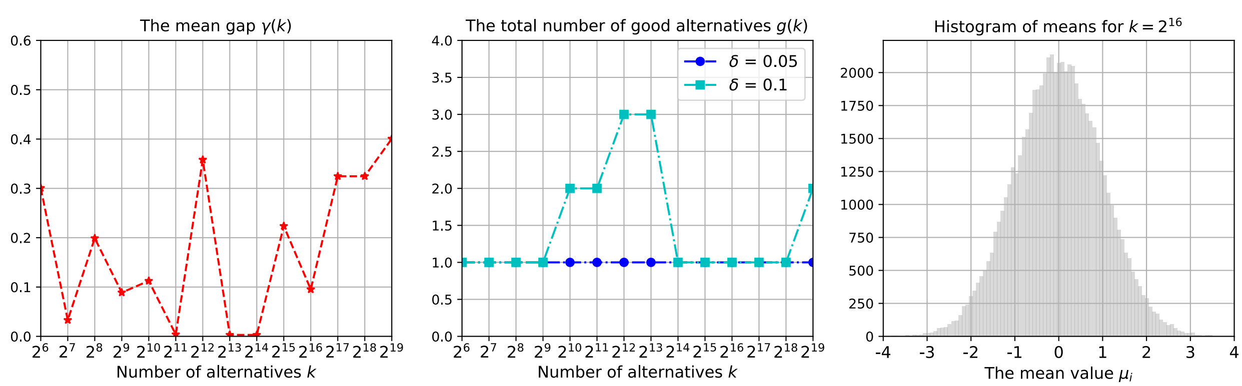

Figure EC.10 Characterizations of the configuration Normal-CV.

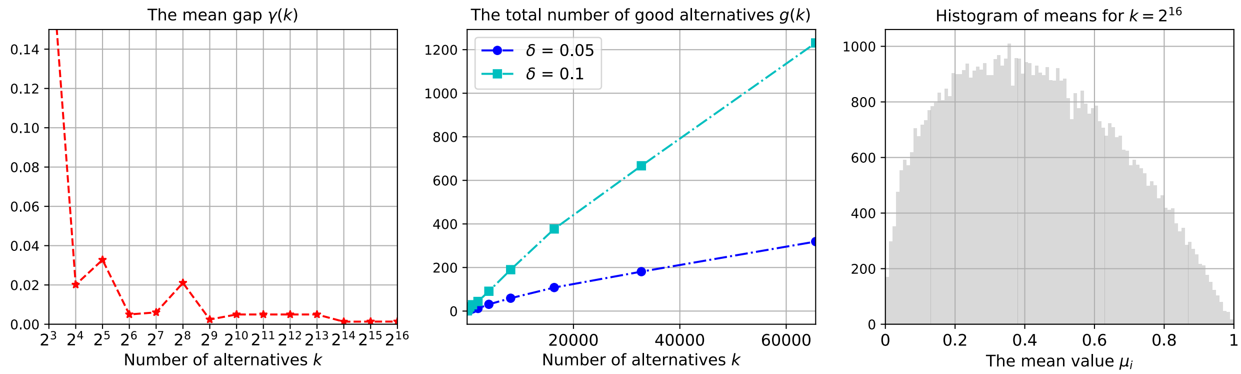

Figure EC.11 Characterizations of the configuration Beta-CV.

From Figure EC.10, we can find that under Normal-CV, $\gamma(k)$ fluctuates up and down as k increases, and it does not remain above some positive constant as required by Assumption 1. However, it still exhibits a key structure of Assumption 1: it does not shrink to zero as k increases. This may be because, for the normal distribution, the distribution of the difference of two order statistics may only depend on the order lag, which is known to hold for the exponential distribution (Balakrishnan and Cohen 2014). Additionally, under Normal-CV, $g(k)$ shows no apparent growth as k increases. Interestingly, the configuration Beta-CV exhibits completely different features. From Figure EC.11, we can see that as k increases, the mean gap $\gamma(k)$ shrinks to zero quickly, and $g(k)$ grows linearly in k.

We test the EFG procedure and compare it with the EA procedure for the two configurations. In the experiment, we let $k = 2^l$ with l ranging from 3 to 16, and for each k, we let $B = 100k$ and allocate $10k$ to the greedy phase for the EFG procedure. Then, for the two configurations, we plot the estimated PCS and PGS of the two procedures against different k in Figure EC.12 and Figure EC.13, respectively. We have the following findings from Figure EC.12. First, under Normal-CV, which violates Assumption 1, the EFG procedure may also be sample optimal for the PCS (PGS) as the estimated PCS (PGS) remains above 60% (70%) for all values of k. Second, as k grows, the PCS of the EA procedure does not plummet to zero, as in Figure 3. This may be because, under Normal-CV, the number of alternatives generated that are close to the best does not increase as k grows, which may also lead to the non-diminishing PGS of the EA procedure. However, despite this, the EFG procedure still significantly outperforms the EA procedure for both the PCS and PGS, showing the substantial performance gains of adding a greedy phase.

We obtain three findings from Figure EC.13. First, under Beta-CV, both the PCS of the EFG and EA procedures diminish to zero as k grows. While the EFG procedure's PCS remains higher than that of the EA procedure, this shows that when $\gamma(k) \to 0$, the EFG procedure may fail to maintain a non-zero PCS given a total sampling budget B that is $O(k)$. Furthermore, when $\gamma(k) \to 0$, the

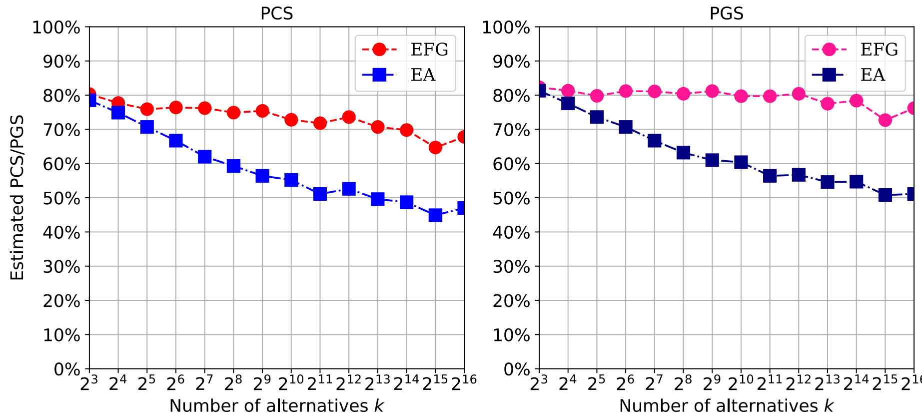

Figure EC.12 A comparison between the EA procedure and the EFG procedure under Normal-CV.

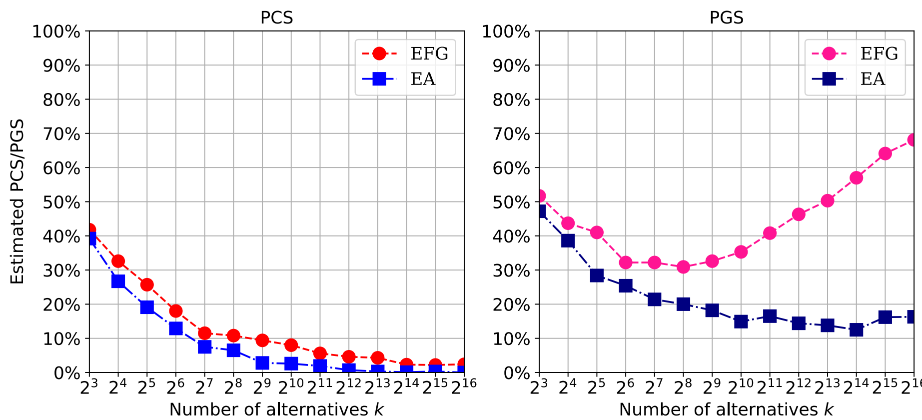

Figure EC.13 A comparison between the EA procedure and the EFG procedure under Beta-CV.

attainable lower bound of the growth order of the total sampling budget required for a positive PCS may no longer be $O(k)$. In such scenarios, we conjecture that the EFG procedure may also achieve the sample optimality, i.e., attain the new minimal order of the total sampling budget. However, we will not go further on this issue because the objective of selecting a good alternative may be more meaningful when $\gamma(k) \to 0$.

Second, under Beta-CV, the PGS of the EA procedure does not decrease to zero, which is not surprising. When $g(k)$ grows at the order $O(k)$ as under Beta-CV, obtaining a non-zero PGS is trivial. One can even randomly select an alternative, and the resulting PGS $g(k)/k$ is positive no matter how large k is. Therefore, when $g(k)$ grows at the order of $O(k)$, a persistent non-zero PGS may not indicate the sample optimality. Last, the EFG procedure's PGS exhibits an interesting monotonic increase after $k \geq 2^8$, and it becomes much larger than the EA procedure's PGS when k is large. Again, this demonstrates the greedy phase's power in improving the sample efficiency.

We note here that in the above experiment, $g(k)$ either remains bounded or grows linearly in k. However, it may only grow at a sub-linear order such that $g(k) \to \infty$ but $g(k)/k \to 0$ as $k \to \infty$. The growth order of $g(k)$ may have a significant impact on the PGS. Therefore, we conduct additional experiments to compare the PGS of the EFG procedure under different growth orders of $g(k)$. See EC.3.6.2 for the details and the results.

EC.3.6.2. Problem Configurations with Explicit Orders of $g(k) \to \infty$. To characterize the impact of the growth order of $g(k)$ on the PGS of the EFG procedure, we consider the following configuration with a common unit variance and

$$\mu_1 = \delta,\ \mu_2 = \delta - \frac{\delta}{k},\ \mu_i \sim \mathcal{U}(0, \mu_2),\ \forall i = 3, \dots, g(k);\ \mu_j \sim \mathcal{U}(-1, 0),\ \forall j = g(k)+1, \dots, k \quad \text{(EC.3.1)}$$

where $\mathcal{U}(a,b)$ is the uniform distribution with the support $[a,b]$ and $g(k)$ is a function of k. In this configuration, $\gamma(k)$ converges to zero quickly as k increases and our focus is only on the PGS. In our experiment, we let $\delta = 0.05$, and consider the following two choices of $g(k)$: (1) $g(k) = \lceil 0.05k \rceil$; (2) $g(k) = \lceil 0.5\sqrt{k} \rceil$. When k is very small, the total number of good alternatives is set as $\max\{g(k), 2\}$. In this experiment, for each k, we let $B = 100k$ and allocate $10k$ to the greedy phase for the EFG procedure. See Figure EC.14 for the comparison between the PGS under $g(k) = \lceil 0.05k \rceil$ and the PGS under $g(k) = \lceil 0.5\sqrt{k} \rceil$ that are estimated based on 2000 independent macro replications.

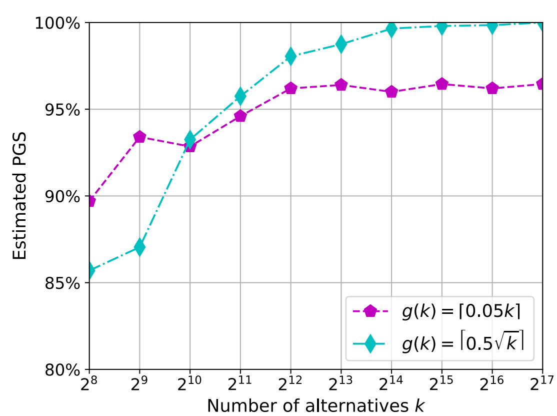

Figure EC.14 Estimated PGS of the EFG procedure under different $g(k)$.

From Figure EC.14, we can see that the PGS under $g(k) = \lceil 0.05k \rceil$ becomes lower than that under $g(k) = \lceil 0.5\sqrt{k} \rceil$ when $k \geq 2^{10}$. The former grows to 1 as k increases, while the latter remains around 97%. This may be because, when $g(k) = \lceil 0.5\sqrt{k} \rceil$, the good alternatives only represent a negligible proportion of the alternatives (notice that $\frac{g(k)}{k} \to 0$ as $k \to \infty$), and they together may only use a small proportion of the total sampling budget when k is large, e.g., $1k$. If so, the left budget is sufficient to let all inferior alternatives finish their boundary-crossing processes and ensure a correct selection as $k \to \infty$. However, when $g(k) = \lceil 0.05k \rceil$, this is less likely to happen as the good alternatives may wastefully occupy too much sampling budget in the selection process.

The phenomenon discussed above indicates that under the asymptotic regime of $k \to \infty$, if the number of good alternatives is allowed to be unbounded, the sample optimality regarding the PGS should be defined with more delicacy. For instance, when $g(k)$ grows at a sub-linear order as $k \to \infty$, for an R&S procedure, being sample optimal may require its PGS to converge to 1 as $k \to \infty$ within a total sampling budget $B = ck$ for some positive constant c. Then, characterizing the EFG procedure's performance in terms of redefined sample optimality becomes an interesting and important problem. However, as argued in Section 4.5, we need a better upper bound of the sampling budget allocated to the good alternatives in the boundary-crossing process. New technical treatment required for resolving this issue may lie beyond the scope of this paper, and we leave it as a future work.

EC.4. Parallelization of the EFG$^+$ Procedure

In this section, we first introduce the parallel version of the EFG$^+$ procedure, namely the EFG^{++} procedure. We then test its performance in solving the large-scale throughput maximization problem and its parallel efficiency when implemented in a master-worker parallel computing environment. Lastly, we show how to reduce the communication overhead and achieve asynchronization for the EFG^{++} procedure to improve the parallel efficiency.

EC.4.1. Description of the EFG^{++} procedure

A detailed description of the EFG^{++} procedure is in Procedure 4. We design the procedure following a master-worker paradigm of parallel computing. In such a paradigm, the unique master processor (e.g., CPU) controls the main computing logic and manages a bunch of parallel worker processors to execute specific parallelizable tasks, e.g., simulating the alternatives. We implement the procedure in the Python language, and the program codes are available at https://github.com/largescaleRS/greedy-procedures.

As a parallel version of the EFG$^+$ procedure, the EFG^{++} procedure also has three phases: the seeding phase, exploration phase, and greedy phase. In Procedure 4, the three phases are described in lines 1-5, lines 6-12, and lines 13-20, respectively. For the seeding and exploration phases, the key to efficient parallelization is load balancing of simulation tasks among the available worker processors, i.e., allocating different processors (almost) the same number of observations to simulate. Thus, in both phases, the master processor needs to find a load-balancing simulation task assignment scheme and then assign the tasks to the worker processors accordingly. To facilitate this, we also design a simple yet efficient sequential-filling algorithm for solving the task assignment schemes and introduce it in Section EC.4.4. In Procedure 4, the algorithm is used in line 1 for the seeding phase and in line 8 for the exploration phase. In each of the two phases, after simulation tasks are assigned, the master processor waits for all the worker processors to finish and return the

Procedure 4 Parallel Explore-First Greedy (EFG^{++}) Procedure

Require: k alternatives $X_1, \ldots, X_k$, q parallel worker processors, the total sampling budget $B = (n_{sd} + n_0 + n_g)k$, the total number of groups G, and the in-processor mini-batch size z.

Find a task assignment scheme $\{n_{i,j}\}_{i=1,\ldots,k;j=1,\ldots,q}$ by Algorithm 1 satisfying

$$\sum_{i=1}^{k} n_{i,j} \le \lceil n_{sd}k/q \rceil, \quad \sum_{j=1}^{q} n_{i,j} = n_{sd}, \quad n_{i,j} \in \mathbb{N}_+, \quad \forall\, i = 1, \ldots, k,\ j = 1, \ldots, q. \tag{EC.4.1}$$

for **processor** $j \in \{1, 2, 3, \ldots, q\}$ **do**

  For each alternative $i = 1, \ldots, k$, if $n_{i,j} > 0$, take $n_{i,j}$ independent observations $X_{i1}, \ldots, X_{in_{i,j}}$ and set $Y^{sd}_{i,j} = \sum_{l=1}^{n_{i,j}} X_{il}$; otherwise, let $Y^{sd}_{i,j} = 0$. Return $\{Y^{sd}_{i,j}\}_{i=1,\ldots,k}$.

end for

Block until all $\{Y^{sd}_{i,j}\}_{i=1,\ldots,k}, j = 1, \ldots, q$ are received. For each alternative $i = 1, \ldots, k$, let $\bar{X}^{sd}_i = \frac{\sum_{j=1}^{q} Y^{sd}_{i,j}}{n_{sd}}$.

According to $\bar{X}^{sd}_i$, sort the alternatives in descending order as $\{(1), (2), \ldots, (k)\}$.

Let I^r denote the group of alternatives for $r = 1, \ldots, G$. Set $\Delta = 2^G - 1$. Then, let $I^1 = \{(1), (2), \ldots, (\lfloor k/\Delta \rfloor)\}$, $I^r = \{(\lfloor k2^{r-2}/\Delta \rfloor + 1), \ldots, (\lfloor k2^{r-1}/\Delta \rfloor)\}$ for $r = 2, \ldots, G-1$, and let $I^G = \{(\lfloor k2^{G-2}/\Delta \rfloor + 1), \ldots, (k)\}$.

Find a task assignment scheme $\{n_{i,j}\}_{i=1,\ldots,k;j=1,\ldots,q}$ by Algorithm 1 satisfying

$$\sum_{i=1}^{k} n_{i,j} \le \lceil n_0 k/q \rceil, \quad \sum_{j=1}^{q} n_{i,j} = \sum_{r=1}^{G} n^r \mathbb{I}_{i \in I^r}, \quad n_{i,j} \in \mathbb{N}_+, \quad \forall\, i = 1, \ldots, k, j = 1, \ldots, q. \tag{EC.4.2}$$

for processor $j \in \{1, 2, 3, \ldots, q\}$ **do**

  For each alternative $i = 1, \ldots, k$, if $n_{i,j} > 0$, take $n_{i,j}$ independent observations $X_{i1}, \ldots, X_{in_{i,j}}$ and set $Y_{i,j} = \sum_{l=1}^{n_{i,j}} X_{il}$; otherwise, let $Y_{i,j} = 0$. Return $\{Y_{i,j}\}_{i=1,\ldots,k}$.

end for

Block until all $\{Y_{i,j}\}_{i=1,\ldots,k}, j = 1, \ldots, q$ are received. For each alternative $i = 1, \ldots, k$, let $n_i = \sum_{r=1}^{G} n^r \mathbb{I}_{i \in I^r}$ and $\bar{X}_i(n_i) = \frac{\sum_{j=1}^{q} Y_{i,j}}{n_i}$.

while $\sum_{i=1}^{k} n_i + qz \le (n_0 + n_g)k$ **do**

  Let $s = \arg\max_{i=1,\ldots,k} \bar{X}_i(n_i)$;

  for processor $j \in \{1, 2, 3, \ldots, q\}$ **do**

    For alternative s, take z independent observations $X_{s1}, \ldots, X_{sz}$ and return $Y_j = \sum_{l=1}^{z} X_{sl}$;

  end for

  Block until all $Y_j, j = 1, \ldots, q$ are received. Let $\bar{X}_s(n_s) = \frac{n_s \bar{X}_s(n_s) + \sum_{j=1}^{q} Y_j}{n_s + qz}$ and $n_s = n_s + qz$.

end while

Select $\arg\max_{i=1,\ldots,k} \bar{X}_i(n_i)$ as the best.

results. Then, the master processor updates the sample sizes and sample means for all alternatives and enters the next phase.

As introduced in Section 5.7, we consider using a batching approach to parallelize the fully sequential greedy phase for the EFG^{++} procedure. In the batched greedy phase of the procedure, at each stage, the current-best alternative s will be simulated more than once, and the simulation task will be equally assigned to the worker processors. Then, after all the worker processors finish simulating alternative s and return the results, the master processor updates the sample size and sample mean for alternative s and enters the next stage. When the total sampling budget is exhausted, the master processor delivers the alternative with the largest sample mean as the best. For the worker processors, we use z to denote the number of observations to simulate at each stage and name it the *in-processor mini-batch size*. With the presence of q worker processors, the batch size of the greedy phase equals $q \times z$. In our implementation, we regard z as a parameter of the procedure.

In this section's experiments, we are concerned about not only one procedure's performance in terms of the PCS and PGS but also its parallel efficiency when implemented in parallel computing environments. Following Ni et al. (2017) and Hong et al. (2022), we quantify one procedure's parallel efficiency in a master-worker parallel computing environment using the measure *utilization* defined by

$$\text{Utilization} = \frac{\text{Total simulation time}}{\text{Wall-clock time} \times \text{Number of parallel worker processors}},$$

where the total simulation time is the total system time used by all the parallel worker processors for simulating all the observations. The higher the utilization is, the more efficient the procedure is in using the available processors, and the more suitable the procedure is for parallelization. Notice that we can estimate the utilization for each phase of the EFG^{++} procedure.

EC.4.2. Performance and Parallel Efficiency of the EFG^{++} procedure

We test the performance and parallel efficiency of the EFG^{++} procedure on the throughput maximization (TP) problem, which has been introduced in Section EC.3.4.2. In our experiments, we use a total sampling budget $B = 100k$ and allocate 20%, 70%, and 10% of the total sampling budget to the seeding phase, exploration phase, and greedy phase, respectively. Furthermore, unless otherwise specified, we estimate the concerned quantities like the PCS, PGS, wall-clock time, and the utilization of the procedure based on 300 independent macro replications.

EC.4.2.1. Testing the EFG^{++} Procedure on the TP Problem. We first test the EFG^{++}procedure using the problem with $k = 11774$ alternatives. When running the procedure, we try four choices for the number of available worker processors $q \in \{5, 10, 20, 40\}$. For each q, we set

$z = 1$, i.e., let each worker processor simulate the current best only once at each stage of the greedy phase. Then, we report the PCS, PGS, wall-clock time, and utilization of the EFG^{++}procedure for different q in Table EC.1.

# of worker processors q	In-processor mini-batch size z	PCS	PGS ($\delta = 0.01$)	Wall-clock time (s)	Utilization
1	1	0.42	0.94	107.572	83.85%
5	1	0.44	0.96	27.346	65.19%
10	1	0.46	0.96	19.059	46.13%
20	1	0.40	0.95	15.306	28.78%
40	1	0.47	0.92	13.132	16.68%

Table EC.1 Performance and parallel efficiency of the EFG^{++} procedure on the TP problem with 11774 alternatives.

From Table EC.1, we have the following findings. First, there is no apparent decrease in the PCS and PGS as the number of worker processors grows from 1 to 40. This indicates that using a parallelizable batched greedy phase will not influence the performance of the EFG^{++} procedure when the batch size is not large. Second, the utilization of the EFG^{++} procedure is less than 90% even when $q = 1$. This is due to the fact that the time spent to find the current best at each stage of the greedy phase is not negligible in this medium-size problem. Third, when the number of worker processors is small (e.g., 5 or 10), the EFG^{++} procedure can efficiently reduce the wall-clock time. In Table EC.1, when the number of worker processors increases from 1 to 10, the wall-clock time is reduced from 107.572 seconds to 13.132 seconds. However, if the number of processors continues to increase, the procedure can not be further effectively accelerated. As a result, the utilization keeps decreasing to a deficient level as q grows. To understand this phenomenon, we report the wall-clock times of different phases for the EFG^{++} procedure in Table EC.2.

# of worker	In-processor	Wall-clock time (s)		
processors q	mini-batch size z	seeding phase	exploration phase	greedy phase
1	1	16.799	56.840	33.933
5	1	3.442	11.812	12.092
10	1	1.762	5.957	11.341
20	1	0.894	2.997	11.415
40	1	0.492	1.551	11.089

Table EC.2 Each phase's wall-clock time of the EFG^{++} procedure on the TP problem with 11774 alternatives.

From Table EC.2, we can see that both the seeding and exploration phases are parallelized very efficiently. As the number of worker processors used q is doubled for the two phases, the wall-clock time is approximately halved. The problem lies in the greedy phase. When $q \geq 10$, increasing q

shows almost no impact on the wall-clock time of the greedy phase. This may arise due to the frequent communication between the master processors and the worker processors in the greedy phase. In this experiment, we set $z = 1$ for different q. Then, in the greedy phase, to obtain one single simulation observation of the alternatives, the master processor has to communicate with one worker processor twice, once to specify which alternative to simulate and once to receive the observation from the worker processor, no matter how many worker processors there are.

EC.4.2.2. Improving Utilization by Reducing Communication. A simple way to reduce the communication cost of running the batched greedy phase in parallel is to enlarge the value of z. By doing so, the batch size of each stage is increased, and the total number of rounds of communications is reduced. To show the impact of z, we also test the EFG^{++} procedure with $z = 10$ for $q = 10$, 20 and also 40. The estimated PCS, PGS, wall-clock time and utilization for different z and q are summarized in Table EC.3. We also keep track of the greedy phase's wall-clock times and report them in Table EC.3.

# of worker processors q	In-processor mini-batch size z	PCS	PGS ($\delta = 0.01$)	The greedy phase's wall-clock time (s)	Wall-clock time (s)	Utilization
10	1	0.46	0.96	12.884	19.059	46.13%
10	10	0.44	0.96	1.286	8.673	94.77%
20	1	0.40	0.95	11.415	15.306	28.78%
20	10	0.48	0.94	1.315	5.423	78.72%
40	1	0.47	0.92	11.089	13.132	16.68%
40	10	0.46	0.95	1.131	3.174	66.65%

Table EC.3 Examining the impact of in-processor mini-batch size z on the performance and parallel efficiency of the EFG^{++} procedure using the TP problem with 11774 alternatives.

We have the following findings from Table EC.3. First, we find that for each value of q, increasing z does not reduce the PCS and PGS of the EFG^{++} procedure. Recall that the stage batch size of the batched greedy phase m equals $q \times z$. This finding shows that even when the batch size is large, e.g., $m = 400$ in the last row of Table EC.3, using a batched greedy phase for parallelization can deliver the same level of selection accuracy. Second, increasing z can effectively reduce the communication overhead between processors. For each value of q, when we increase z from 1 to 10, the wall-clock time of the greedy phase is significantly reduced. As a result, we reduce the wall-clock time of the procedure and improve the parallel efficiency. These results indicate that with a proper choice of z (e.g., 10), we can parallelize the EFG^{++} procedure efficiently without damaging its selection accuracy in solving large-scale R&S problems.

EC.4.3. Asynchronization of the EFG^{++} procedure

EC.4.3.1. Procedure Design and the Performance. We now introduce how to implement the EFG^{++} procedure in an asynchronized fashion. Notice that for both the seeding phase and exploration phase, synchronization of sampling information only happens once at the end of the phase. Therefore, it suffices to consider asynchronizing the batched greedy phase only. Our design is straightforward. At the beginning of the greedy phase, the master processor still requires every worker processor to take z observations from the first current-best alternative. Immediately after that, whenever the master processor receives the simulation results from any worker processor, it updates the sample size and sample mean for the simulated alternative and asks the idle processor to take z observations from the new current-best alternative. After the total sampling budget is exhausted, the procedure delivers the alternative with the largest sample mean as the best. We name the new parallel procedure with an asynchronized greedy phase the Asyn-EFG^{++} procedure. A formal description of the procedure is in Procedure 5.

# of worker processors q	In-processor mini-batch size z	PCS	PGS ($\delta = 0.01$)	Wall-clock time (s)	Utilization
5	1	0.46	0.95	29.900	59.05%
5	10	0.44	0.92	17.497	95.07%
10	1	0.43	0.95	20.297	41.59%
10	10	0.44	0.96	8.673	94.77%
20	1	0.44	0.94	16.824	25.17%
20	10	0.45	0.95	5.432	76.46%
40	1	0.45	0.98	15.604	13.66%
40	10	0.45	0.97	3.340	61.62%

Table EC.4 Performance and parallel efficiency of the Asyn-EFG^{++} procedure on the TP problem with 11774 alternatives.

Now we test the selection accuracy of the Asyn-EFG^{++} procedure. Here, we use the same problem and implementation settings used in Section EC.4.2. The estimated PCS, PGS and wall-clock times for different q and z are summarized in Table EC.4. From Table EC.4, we can see that all the results regarding the performance and parallel efficiency of the EFG^{++} procedure and the impact of z obtained in Section EC.4.2 also hold for the Asyn-EFG^{++} procedure.

EC.4.3.2. Random Simulation Times. In the above experiments, it only takes 0.07ms on average to take one observation from the alternatives[6]. When simulating the alternatives is so fast, synchronization in the greedy phase may be fine. To show the value of asynchronization, we conduct the following experiment. For each worker processor, every time it generates one simulation

[6] We achieve this speed by utilizing the Python library Cython to compile the simulation program as faster C codes.

Procedure 5 Asynchronized Parallel Explore-First Greedy (Asyn-EFG^{++}) Procedure

Require: k alternatives $X_1, \ldots, X_k$, q parallel worker processors, the total sampling budget $B = (n_{sd} + n_0 + n_g)k$, the total number of groups G, and the in-processor mini-batch size z.

Execute lines 1 - 12 of Procedure 4.
Let $s = \arg\max_{i=1,\ldots,k} \bar{X}_i(n_i)$;
for processor $j \in \{1, 2, 3, \ldots, q\}$ **do**
  For alternative s, take z independent observations $X_{s1}, \ldots, X_{sz}$, return s, j and $r = \sum_{l=1}^{z} X_{sl}$;
end for
while $\sum_{i=1}^{k} n_i + qz \le (n_0 + n_g)k$ **do**
  Block until receive from any processor the alternative id s, the sum r and the processor id j;
  Let $\bar{X}_s(n_s) = \frac{n_s \bar{X}_s(n_s) + r}{n_s + z}$ and $n_s = n_s + z$;
  Let $s = \arg\max_{i=1,\ldots,k} \bar{X}_i(n_i)$;
  Let **processor** j **do**
    For alternative s, take z independent observations $X_{s1}, \ldots, X_{sz}$, return s, j and $r = \sum_{l=1}^{z} X_{sl}$;
end while
for $t \in \{1, 2, 3, \ldots, q\}$ **do**
  Block until receive from any processor the alternative id s, the sum r and the processor id j;
  Let $\bar{X}_s(n_s) = \frac{n_s \bar{X}_s(n_s) + r}{n_s + z}$ and $n_s = n_s + z$;
end for
Select $\arg\max_{i=1,\ldots,k} \bar{X}_i(n_i)$ as the best.

observation, it sleeps for T ms, which is *i.i.d.* generated from the uniform distribution with the support $[0.5, 1.5]$. We use the problem instance with $k = 41624$ alternatives and 40 worker processors in the experiment. For both the EFG^{++} procedure and the Asyn-EFG^{++} procedure, we estimate the utilization for each of the three phases and summarize them in Table EC.5.

In-processor mini-batch size z	Procedure	Utilization		
		seeding phase	exploration phase	greedy phase
1	EFG^{++}	99.20%	99.60%	46.71%
	Asyn-EFG^{++}	99.18%	99.57%	40.48%
10	EFG^{++}	99.18%	99.57%	80.45%
	Asyn-EFG^{++}	99.17%	99.56%	99.48%

Table EC.5 Each phase's parallel efficiency of the EFG^{++} procedure and the Asyn-EFG^{++} procedure on the TP problem with 41624 alternatives.

We obtain the following findings from Table EC.5. First, the seeding and exploration phases can be parallelized very efficiently for both procedures. The utilization in the two phases are all above 99%. Second, for the Asyn-EFG^{++} procedure, increasing z can also help increase the utilization in the greedy phase due to the reduction of communication overhead. Third, the impact of asynchronization on the parallel efficiency of the EFG^{++} procedure in the greedy phase varies for different values of z. When $z = 1$, the Asyn-EFG^{++} procedure's utilization in the greedy phase can be lower than that of the EFG^{++} procedure. This may be because, in the greedy phase, the Asyn-EFG^{++} procedure has to update the current best alternative every time a worker processor returns the simulation results, thereby incurring a larger overhead of finding the *argmin* than the EFG^{++} procedure. However, when z is increased to 10, asynchronizing the greedy phase improves the utilization significantly, showing that the profit of asynchronization surpasses the excess overhead of frequent updating the current best alternative. From the results, we conclude that with a proper choice of z, asynchronization can improve the EFG^{++} procedure's robustness to random and unequal simulation times.

EC.4.4. Load-balancing Simulation Task Assignment

The sequential-filling algorithm for load-balancing simulation task assignment among the available worker processors is shown in Algorithm 1. Given $\sum_{i=1}^{k} n_i$ observations to simulate in total, each of the q processors should simulate no more than $\lceil \sum_{i=1}^{k} n_i/q \rceil$ observations. Informally, we regard $\lceil \sum_{i=1}^{k} n_i/q \rceil$ as the maximum volume of each processor. In the procedure, we use $l_i, i = 1, \ldots, k$ to denote the unassigned number of observations for alternative i, and use $v_j, j = 1, \ldots, q$ to denote the unfulfilled "volume" of processor j. The procedure proceeds by filling the simulation task into different processors sequentially and uses $n_{i,j}$ to keep track of how many observations of alternative i are assigned to processor j.

This algorithm is used in the seeding and exploration phases of the EFG^{++} procedure and the Asyn-EFG^{++} procedure. It works very efficiently in our experiments. As displayed in Table EC.5, the utilization of the two procedures in the seeding and exploration phases are always above 99%.

Algorithm 1 Sequential-Filling Algorithm

Require: The number of alternatives k, the number of processors q, and the sample size of each alternative n_i, $i = 1, \ldots, k$.

Let $l_i = n_i$, $v_j = \lceil \sum_{i=1}^{k} n_i / q \rceil$, and $n_{i,j} = 0$ for $i = 1, \ldots, k$ and $j = 1, \ldots, q$. Set counter $j = 0$.
for $i \in \{1, 2, 3, \ldots, q\}$ **do**
  while True do
    if $l_i \leq v_j$ **then**
      Let $n_{i,j} = n_{i,j} + l_j$ and $v_j = v_j - l_i$;
      break
    else
      Let $n_{i,j} = n_{i,j} + v_j$, $l_i = l_i - v_j$, and update the counter $j = j + 1$.
    end if
  end while
end for
Return $\{n_{i,j}\}_{i=1,\ldots,k; j=1,\ldots,q}$.

References

Balakrishnan N, Cohen AC (2014) *Order statistics & inference: estimation methods* (Elsevier).

Buzacott J, Shanthikumar J (1993) *Stochastic Models of Manufacturing Systems* (Prentice Hall).

Hong LJ, Jiang G, Zhong Y (2022) Solving large-scale fixed-budget ranking and selection problems. *INFORMS Journal on Computing* 34(6):2930–2949.

Karnin Z, Koren T, Somekh O (2013) Almost optimal exploration in multi-armed bandits. *Proceedings of the 30th International Conference on Machine Learning*, volume 28, 1238–1246 (PMLR).

Ni EC, Ciocan DF, Henderson SG, Hunter SR (2017) Efficient ranking and selection in parallel computing environments. *Operations Research* 65(3):821–836.

Rudin W (1976) *Principles of mathematical analysis*, volume 3 (McGraw-hill New York).

Siegmund D (1985) *Sequential analysis: tests and confidence intervals* (Springer Science & Business Media).

Zhao Y, Stephens C, Szepesvari C, Jun KS (2023) Revisiting simple regret: fast rates for returning a good arm. *Proceedings of the 40th International Conference on Machine Learning*, volume 202, 42110–42158 (PMLR).