Online Supplement to 'Dynamic Sampling Allocation and Design Selection'

Yijie Peng, Chun-Hung Chen, Michael C. Fu and Jian-Qiang Hu

Appendix A: Proofs in Sections 2, 3 and 4

Proposition A.1. *Assume that there exist $\theta,\ \theta' \in \Theta$ such that $\{\mu_1,\dots,\mu_k\} \subset \theta$, $\mu_{[1]} \geq \cdots \geq \mu_{[k]}$, $\{\mu'_1,\dots,\mu'_k\} \subset \theta'$, $\mu'_{(1)} \geq \cdots \geq \mu'_{(k)}$, and $[1] \neq (1)$. Then there does not exist a policy $(\mathcal{A}^*,\mathcal{D}^*)$ such that $\forall\ \theta \in \Theta$,*

$$Pr(\{\mathcal{D}^*(\mathcal{E}_T^{\mathcal{A}^*}) = [1]\} \mid \theta) \geq Pr(\{\mathcal{D}(\mathcal{E}_T^{\mathcal{A}}) = [1]\} \mid \theta)\ .$$

Proof. Without loss of generality, suppose there exists $\theta \in \Theta$, s.t. $\mu_1 > \mu_i$, $i = 2,\dots,k$, we can define a trivial selection policy for which $\mathcal{D}_{tr}(\cdot) \equiv 1$, which means no matter what information is collected from simulation, design 1 is selected, and $\mathcal{A}$ and $\mathcal{A}'$ are arbitrary allocation policies. Obviously, we have

$$Pr(\{\mathcal{D}_{tr}(\mathcal{E}_T^{\mathcal{A}}) = [1]\} \mid \theta) = 1 \geq Pr(\{\mathcal{D}(\mathcal{E}_T^{\mathcal{A}'}) = [1]\} \mid \theta),$$

which means this trivial policy is the best policy when design 1 is actually the true best, and equality holds only in the case that $\mathcal{D}(\mathcal{E}_T^{\mathcal{A}'}) = \mathcal{D}_{tr}(\mathcal{E}_T^{\mathcal{A}})$, *a.s.*. Therefore, $(\mathcal{A}^*,\mathcal{D}^*)$ cannot dominate $(\mathcal{A},\mathcal{D}_{tr})$ if $\mathcal{D}^*(\mathcal{E}_T^{\mathcal{A}^*}) \neq \mathcal{D}_{tr}(\mathcal{E}_T^{\mathcal{A}})$ (not almost surely equal).

From the assumption for the feasible set Θ, there exists another $\theta \in \Theta$, s.t. $\exists\ i \in \{2,\dots,k\}$, $\mu_1 < \mu_i$. We have

$$Pr(\{\mathcal{D}_{tr}(\mathcal{E}_T^{\mathcal{A}}) = [1]\} \mid \theta) = 0 \leq Pr(\{\mathcal{D}(\mathcal{E}_T^{\mathcal{A}'}) = [1]\} \mid \theta),$$

which means this trivial policy is the worst policy when design 1 is not the true best. Define $\mathcal{A}^e$ to be the allocation policy that equally allocates T replications among k designs, or in other words, $\ulcorner \frac{T}{k} \urcorner$ replications are allocated to each design, where $\ulcorner \cdot \urcorner$ is the operator that truncates to the nearest integer smaller than the argument. Denote $Q^{*\ulcorner \frac{T}{k} \urcorner}$ as the $\ulcorner \frac{T}{k} \urcorner$-fold convolution of distribution Q. Since the mean of the distribution $Q^{*\ulcorner \frac{T}{k} \urcorner}$ is also $(\mu_1,\dots,\mu_k)$ and the mean is the weighted average of the random variable, it's easy to see that

$$Pr(\{\mathcal{D}^m(\mathcal{E}_T^{\mathcal{A}^e}) = [1]\} \mid \theta) = Pr(\bar{m}_{[1]}^{\ulcorner \frac{T}{k} \urcorner} - \bar{m}_{[j]}^{\ulcorner \frac{T}{k} \urcorner} > 0,\ j = 2,\dots,k \mid \theta) > 0,$$

where $\bar{m}_i^n = (\sum_{l=1}^n x_{il})/n$, $i = 1,\dots,k$. Therefore, we have $\mathcal{D}^*(\mathcal{E}_T^{\mathcal{A}^*}) \neq \mathcal{D}_{tr}(\mathcal{E}_T^{\mathcal{A}})$ (not almost surely equal). Combining the analyses above, we can conclude that there does not exist a dominant policy defined above. □

Proof of Proposition 1.

Proof. By definition, we have

$$
\begin{aligned}
&E[Pr(\mu_{\mathcal{D}^*} - \mu_j > 0,\ j \neq \mathcal{D}^* | \mathcal{E}_T^{\mathcal{A}^*})] \\
=&\int \left(\int_{\theta\in\Theta} \mathbf{1}\{\mathcal{D}^*(\mathcal{E}_T^{\mathcal{A}^*}) = [1]\} \cdot P\left(d\theta\ |\mathcal{E}_T^{\mathcal{A}^*}(\hat{\theta})\right)\right) P\left(d\mathcal{E}_T^{\mathcal{A}^*}(\hat{\theta})\right) P(d\hat{\theta}) \\
=&E\left[\mathbf{1}\{\mathcal{D}^*(\mathcal{E}_T^{\mathcal{A}^*}) = [1]\}\right] \geq E\left[\mathbf{1}\{\mathcal{D}^m(\mathcal{E}_T^{\mathcal{A}^e}) = [1]\}\right] \\
=&\int_{\theta\in\Theta} Pr(\bar{m}_{[1]}^{\ulcorner\frac{T}{k}\urcorner} - \bar{m}_{[j]}^{\ulcorner\frac{T}{k}\urcorner},\ j = 2,\ldots,k \mid \theta)\ P(d\theta)\ .
\end{aligned}
$$

Obviously, as $T \to \infty$, we have

$$
Pr(\bar{m}_{[1]}^{\ulcorner\frac{T}{k}\urcorner} - \bar{m}_{[j]}^{\ulcorner\frac{T}{k}\urcorner} > 0,\ j = 2,\ldots,k \mid \theta) \to 1, \quad a.e.,
$$

therefore, $E[\mathbf{1}\{\mathcal{D}^*(\mathcal{E}_T^{\mathcal{A}^*}) = [1]\}] \to 1$, which implies

$$
Pr(\mu_{\mathcal{D}^*} - \mu_j > 0,\ j \neq \mathcal{D}^* | \mathcal{E}_T^{\mathcal{A}^*}) \xrightarrow{p} 1,
$$

where $\xrightarrow{p}$ denote convergence in probability. From (2) in the main body of the paper, we know

$$
\max_{i\in\{1,2,\ldots,k\}} Pr(\mu_i - \mu_j > 0,\ j \neq i | \mathcal{E}_T^{\mathcal{A}^*}) \xrightarrow{p} 1. \tag{b.1}
$$

We denote $(T)_i^{\mathcal{A}^*}$ to specify the dependence of the number of replications allocated to each design on the allocation policy. Then we want to prove that as $T \to \infty$, we have $(T)_i^{\mathcal{A}^*} \to \infty$, *a.e.*, $i = 1,\ldots,k$. Otherwise, there exists $i_0 \in \{1,2,\ldots,k\}$ such that for $\mathcal{B} \doteq \bigcap_{T=1}^{\infty}\bigcup_{M=1}^{\infty}\{(T)_{i_0}^{\mathcal{A}^*} \leq M\}$, $Pr(\mathcal{B}) > 0$. Thus

$$
\max_{i\in\{1,2,\ldots,k\}} Pr(\mu_i - \mu_j > 0,\ j \neq i | \mathcal{E}_T^{\mathcal{A}^*}) \leq \max\left(\max_{j\neq i_0} P(\mu_{i_0} > \mu_j | \mathcal{E}_T^{\mathcal{A}^*}), \max_{i\neq i_0} P(\mu_i > \mu_{i_0} | \mathcal{E}_T^{\mathcal{A}^*})\right)\ . \tag{b.2}
$$

As $T \to \infty$, since $(T)_{i_0}^{\mathcal{A}^*}$ doesn't go to infinity in $\mathcal{B}$, by assumptions (ii) and (iii), we have

$$
\begin{aligned}
P(\mu_{i_0} > \mu_j | \mathcal{E}_T^{\mathcal{A}^*}) &= \int P(\mu_{i_0} > \mu_j | \mathcal{E}_T^{\mathcal{A}^*}, \mu_j) P(\mu_j | \mathcal{E}_T^{\mathcal{A}^*}) < 1 - \varepsilon, \quad in\ \mathcal{B}, \\
P(\mu_i > \mu_{i_0} | \mathcal{E}_T^{\mathcal{A}^*}) &= \int P(\mu_i > \mu_{i_0} | \mathcal{E}_T^{\mathcal{A}^*}, \mu_i) P(\mu_i | \mathcal{E}_T^{\mathcal{A}^*}) < 1 - \varepsilon, \quad in\ \mathcal{B}\ .
\end{aligned}
$$

Therefore, we have, as $T \to \infty$,

$$
\begin{aligned}
&\max_{j\neq i_0} P(\mu_{i_0} > \mu_j | \mathcal{E}_T^{\mathcal{A}^*}) < 1 - \varepsilon, \quad in\ \mathcal{B}, \\
&\max_{i\neq i_0} P(\mu_i > \mu_{i_0} | \mathcal{E}_T^{\mathcal{A}^*}) < 1 - \varepsilon, \quad in\ \mathcal{B}\ .
\end{aligned} \tag{b.3}
$$

From (b.2) and (b.3), we have

$$
\max_{i\in\{1,2,\ldots,k\}} Pr(\mu_i - \mu_j > 0,\ j \neq i | \mathcal{E}_T^{\mathcal{A}^*}) < 1 - \varepsilon, \quad in\ \mathcal{B},
$$

which contradicts equation (b.1). Therefore, as $T \to \infty$, we have $(T)_i^{\mathcal{A}^*} \to \infty$, $i = 1, \ldots, k$. By the Law of Large Numbers, we know, as $T \to \infty$, $\mathcal{D}^m(\mathcal{E}_T^{\mathcal{A}^*}) \to [1]$, *a.e.* From Doob's consistency theorem (see Van der Vaart 2000, p.149), as $T \to \infty$, we have

$$(\mu_1, \ldots, \mu_k) \ |\mathcal{E}_T^{\mathcal{A}^*}(\hat{\theta}) \xrightarrow{d} (\hat{\mu}_1, \ldots, \hat{\mu}_k), \quad a.e., \quad i = 1, 2, \ldots, k,$$

where $\xrightarrow{d}$ denote convergence in distribution. Combining the analysis above, we have, as $T \to \infty$,

$$Pr(\mu_{\mathcal{D}^m} - \mu_j > 0, \ j \neq \mathcal{D}^m | \mathcal{E}_T^{\mathcal{A}^*}) \to 1, \quad a.e.,$$

and also notice that

$$\sum_{i=1}^{k} Pr(\mu_i - \mu_j > 0, \ j \neq i | \mathcal{E}_T^{\mathcal{A}^*}) = \int_{\theta \in \Theta} \left(\sum_{i=1}^{k} \mathbf{1}\{\mu_i - \mu_j > 0, \ i \neq j\} \right) P\left(d\theta | \mathcal{E}_T^{\mathcal{A}^*}\right),$$

since for $\mathcal{Z} = \{\theta : \ \mu_i = \mu_j, \ i \neq j | \mathcal{E}_T^{\mathcal{A}^*}\}$, $Pr(\mathcal{Z}) = 0$, and for $\theta \in \mathcal{Z}^c$, there is one and only design that has the largest mean, which also means $\sum_{i=1}^{k} \mathbf{1}\{\mu_i - \mu_j > 0, \ i \neq j\} = 1$. Therefore, we have

$$Pr(\mu_i - \mu_j > 0, \ j \neq i | \mathcal{E}_T^{\mathcal{A}^*}) \to 0, \quad a.e., \quad i \neq \mathcal{D}^m \ .$$

From these results, it is easy to see, as $T \to \infty$,

$$\mathcal{D}^*(\mathcal{E}_T^{\mathcal{A}^*}) - \mathcal{D}^m(\mathcal{E}_T^{\mathcal{A}^*}) \to 0, \quad a.e \ldots$$

□

Proof of Corollary 1.

Proof. To prove the corollary, we only need to check conditions (i)-(iii) hold for the normal-Gamma model. The identifiability condition (i) is obviously satisfied by the normal distribution. Condition (ii) holds because both the prior distribution of the true parameters and sampling of the performance of designs are independent between different alternatives. Also, we know $\mu_i | \mathcal{E}_T$, $i = 1, 2, \ldots, k$, follow t-distributions given by (4) in the main body of the paper. For finite allocation replications to design i, by the definition of t-distribution, it's easy to show that $\forall \ x, \varepsilon < Pr(\mu_i > x | \mathcal{E}_T) < 1 - \varepsilon$. □

Denote the covariance of two r.v.'s X and Y by

$$cov(X, Y) \doteq E[(X - EX)(Y - EY)] \ .$$

Lemma A.1. *Suppose $y_j \sim N(0, \sigma_j^2)$, $j = 1, \ldots, k$, $y_{j'}$ and y_j, $j \neq j'$, are mutually independent. Denote*

$$c_{j,j'}^{(i)} \doteq cov(y_i - y_j, y_i - y_{j'}), \ j, j' = 1, \ldots, i - 1,$$
$$c_{j,j'}^{(i)} \doteq cov(y_i - y_{j+1}, y_i - y_{j'+1}), \ j, j' = i, \ldots, k - 1 \ .$$

and similar definition for $c^{(i')}_{j,j'}$. *If* $\sigma_i^2 > \sigma_{i'}^2$, *then*

$$c^{(i)}_{j,j'} > c^{(i')}_{j,j'}, \quad j, j' = 1, \ldots, k-1.$$

Proof. Without loss of generality, suppose $\sigma_1^2 > \sigma_2^2$ and we only need to discuss the case $i = 1$ and $i' = 2$. Then, we have $(y_1 - y_2, \ldots, y_1 - y_k) \sim N(0, \Sigma_1)$ with

$$\Sigma_1 = \begin{pmatrix} \sigma_1^2 + \sigma_2^2 & \sigma_1^2 & \cdots & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 + \sigma_3^2 & \cdots & \sigma_1^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1^2 & \sigma_1^2 & \cdots & \sigma_1^2 + \sigma_k^2 \end{pmatrix},$$

and $(y_2 - y_1, y_2 - y_3, \ldots, y_1 - y_k) \sim N(0, \Sigma_1)$ with

$$\Sigma_2 = \begin{pmatrix} \sigma_2^2 + \sigma_1^2 & \sigma_2^2 & \cdots & \sigma_2^2 \\ \sigma_2^2 & \sigma_2^2 + \sigma_3^2 & \cdots & \sigma_2^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_2^2 & \sigma_2^2 & \cdots & \sigma_2^2 + \sigma_k^2 \end{pmatrix}.$$

The conclusion can be easily obtained by the assumption of the lemma and the expression of Σ_1 and Σ_2 given above. ∎

Proposition A.2. *Suppose* $y_j \sim N(0, \sigma_j^2)$, $j = 1, \ldots, k$, y_j *and* $y_{j'}$, $j \neq j'$, *are mutually independent. If* $\sigma_1^2 > \sigma_2^2 > \cdots > \sigma_k^2$, *then*

$$Pr(y_1 - y_j > 0, \ j \neq 1) > Pr(y_2 - y_j > 0, \ j \neq 2) > \cdots > Pr(y_k - y_j > 0, \ j \neq k) .$$

The conclusion of the proposition is the direct consequence of Lemma A.1 and Plackett-Slepian Comparison Lemma (see page 44 in Azaïs and Wschebor 2009).

Proof of Proposition 2.

Proof. For case (i), if $k = 2$, we have

$$\mu_1 - \mu_2 | \mathcal{E}_T \sim N(\mu_1^{(T)} - \mu_2^{(T)}, (\sigma_1^{(T)})^2 + (\sigma_2^{(T)})^2) .$$

If $\mu_1^{(T)} > \mu_2^{(T)}$, by the symmetry of the normal distribution, we have

$$Pr(\mu_1 - \mu_2 > 0, | \mathcal{E}_T) > Pr(\mu_2 - \mu_1 > 0, | \mathcal{E}_T),$$

so $\forall \ \omega \in \Omega$,

$$\mathcal{D}^*(\mathcal{E}_T(\omega)) = \arg\max_{i=1,2} \mu_i^{(T)} .$$

For case (ii), if $\omega \in \Delta$, we have $\sigma_1^{(T)} = \ldots = \sigma_k^{(T)}$. For $i = 1, \ldots, k$, the covariance matrix of vector $(\mu_i - \mu_j)_{j \neq i} | \mathcal{E}_T$ can be denoted as Σ. Then, we have

$$(\mu_i - \mu_j)_{j \neq i} | \mathcal{E}_T \sim N(\mu^{(T)}, \Sigma),$$

where $\mu^{(T)} = (\mu_i^{(T)} - \mu_j^{(T)})_{j \neq i}$. $Pr(\mu_{\mathcal{D}} - \mu_j > 0,\ j \neq \mathcal{D}|\mathcal{E}_T)$ is an increasing function of $\mu_{\mathcal{D}}^{(T)} - \mu_j^{(T)}$, $j \neq \mathcal{D}$; therefore, $\forall\ \omega \in \Delta$

$$\mathcal{D}^*(\mathcal{E}_T(\omega)) = \arg \max_{i=1,\ldots,k} \mu_i^{(T)} \ .$$

If $\omega \in \Delta^c$, $(\sigma_1^{(T)}(\omega), \ldots, \sigma_k^{(T)}(\omega))$ is a fixed value, and we can define $\Lambda^\omega \doteq \{\omega' \in \Delta^c :\ \sigma_j^{(T)} = \sigma_j^{(T)}(\omega),\ j = 1, \ldots, k\}$. Suppose $y_j \sim N(0, (\sigma_j^{(T)}(\omega))^2)$, $j = 1, \ldots, k$, and y_j and $y_{j'}$, $j \neq j'$, are mutually independent. Then we have

$$(y_i - y_j)_{j \neq i} \sim N(0, \Sigma_i) \ .$$

Define $P_i^\omega \doteq Pr(y_i - y_j > 0,\ j \neq i)$, and

$$i_1^\omega \doteq \min \arg \max_{i=1,\ldots,k} P_i^\omega,$$
$$i_2^\omega \doteq \min \arg \min_{i=1,\ldots,k} P_i^\omega \ .$$

From Proposition A.2, we know, for $\omega \in \Delta^c$

$$P_{i_1^{\tilde{\omega}}}^{\tilde{\omega}} > P_j^{\tilde{\omega}}, \quad j \neq i_1^{\tilde{\omega}}.$$

For $\tilde{\omega} \in \Upsilon \doteq \{\omega' :\ \mu_1^{(T)}(\omega') = \ldots = \mu_k^{(T)}(\omega')\}$,

$$Pr(\mu_i - \mu_j > 0,\ j \neq i|\mathcal{E}_T(\tilde{\omega})) = P_i^{\tilde{\omega}},\ i = 1, \ldots, k.$$

For $\hat{\omega} \in \Delta^c \bigcap \Upsilon$, we have

$$Pr(\mu_{i_1^{\hat{\omega}}} - \mu_j > 0,\ j \neq i_1^{\hat{\omega}}|\mathcal{E}_T(\hat{\omega})) > Pr(\mu_i - \mu_j > 0,\ j \neq i|\mathcal{E}_T(\hat{\omega})),\ i \neq i_1^{\hat{\omega}}.$$

Because $Pr(\mu_i - \mu_j > 0,\ j \neq i|\mathcal{E}_T)$ is a continuous function of $(\mu_1^{(T)}, \ldots, \mu_k^{(T)})$ and increasing with respect to $\mu_i^{(T)} - \mu_j^{(T)}$, $j \neq i$, $\forall\ \omega \in \Delta^c$ there exists $\varepsilon_\omega > 0$, $\forall\ \hat{\omega} \in \Upsilon_1^\omega \doteq \Delta^c \bigcap \tilde{\Upsilon}_1^\omega$, where

$$\tilde{\Upsilon}_1^\omega \doteq \{\omega' \in \Lambda^\omega :\ \mu_{i_2^\omega}^{(T)}(\omega') > \mu_{i_1^\omega}^{(T)}(\omega') > \mu_{i_2^\omega}^{(T)}(\omega') - \varepsilon_\omega,\ \mu_{i_1^\omega}^{(T)}(\omega') > \mu_j^{(T)}(\omega'),\ j \neq i_1^\omega, i_2^\omega\},$$

such that

$$Pr(\mu_{i_1^{\hat{\omega}}} - \mu_j > 0,\ j \neq i_1^{\hat{\omega}}|\mathcal{E}_T(\hat{\omega})) > Pr(\mu_i - \mu_j > 0,\ j \neq i|\mathcal{E}_T(\hat{\omega})), \quad i \neq i_1^{\hat{\omega}} \ .$$

For $\hat{\omega} \in \Xi_1 \doteq \bigcup_{\omega \in \Lambda_1} \Upsilon_1^\omega$, we have

$$\mathcal{D}^*(\mathcal{E}_T(\hat{\omega})) = i_1^{\hat{\omega}}, \quad \arg \max_{i=1,\ldots,k} \mu_i^{(T)} = i_2^{\hat{\omega}},$$

and $i_1^{\hat{\omega}} \neq i_2^{\hat{\omega}}$; therefore,

$$\mathcal{D}^*(\mathcal{E}_T(\hat{\omega})) \neq \arg \max_{i=1,\ldots,k} \mu_i^{(T)} \ .$$

Denote $\mathcal{F}(X)$ as the σ-algebra generated by random variable X. We have

$$\mathcal{F}(\sigma_1^{(T)}, \ldots, \sigma_k^{(T)}) = \mathcal{F}((T)_1, \ldots, (T)_k),$$

and

$$\mathcal{F}(\mu_1^{(T)}, \ldots, \mu_k^{(T)}, \sigma_1^{(T)}, \ldots, \sigma_k^{(T)}) = \mathcal{F}(x_{1,1}, \ldots, x_{1,(T)_1}, \ldots, x_{k,1}, \ldots, x_{k,(T)_k}, (T)_1, \ldots, (T)_k) \ .$$

Then, $\mathcal{F}(\sigma_1^{(T)}, \ldots, \sigma_k^{(T)}) \subset \mathcal{F}(\mu_1^{(T)}, \ldots, \mu_k^{(T)}, \sigma_1^{(T)}, \ldots, \sigma_k^{(T)})$, and it is easy to show, $\forall\ \omega \in \Delta^c$,

$$\mathcal{P}(\Xi | \sigma_1^{(T)}, \ldots, \sigma_k^{(T)})(\omega) > 0,$$

where $\Xi \doteq \Xi_1 \bigcup \Xi_2$. Therefore, if $\mathcal{P}(\Delta^c) > 0$, we have

$$\mathcal{P}(\Xi) = \int_{w \in \Delta^c} \mathcal{P}(\Xi | \sigma_1^{(T)}, \ldots, \sigma_k^{(T)}) \mathcal{P}(d\omega) > 0 \ .$$

□

<u>**Proof of Proposition 3**</u>.

Proof. Under the assumptions,

$$(\bar{m}_j^{(T)} - \bar{m}_i^{(T)}) | \theta \sim N\left(\mu_j - \mu_i, \sigma_{ij}^2 / T\right),$$

where

$$\sigma_{ij}^2 \doteq \frac{\sigma_i^2}{w_i} + \frac{\sigma_j^2}{w_j} \ .$$

In addition, we can easily obtain

$$\begin{aligned} Pr\left(\bar{m}_i^{(T)} \le \bar{m}_j^{(T)} \mid \theta\right) =& E\left[\mathbf{1}\left\{T\left(\bar{m}_j^{(T)} - \bar{m}_i^{(T)}\right) > 0\right\} \mid \theta\right] \le E\left[\exp\left(T\left(\bar{m}_j^{(T)} - \bar{m}_j^{(T)}\right)\right) \mid \theta\right] \\ =& \exp\left(-T(\mu_i - \mu_j) + \sigma_{ij}^2\right), \end{aligned}$$

and the following inequality:

$$\frac{1}{T} \mid \log\left(Pr\left(\bar{m}_i^{(T)} \le \bar{m}_j^{(T)} \mid \theta\right)\right) \mid \le |\mu_i| + |\mu_j| + \frac{\sigma_i^2}{(s)_i} + \frac{\sigma_j^2}{(s)_j} \ .$$

Therefore, we have

$$\begin{aligned} &\frac{1}{T} \mid \log E\left[\mathbf{1}\{\mathcal{D}^m(\mathcal{E}_T) \ne [1]\} \mid \theta, \mathcal{E}_s\right] \mid \le \frac{1}{T} \sum_{j \ne i} \mid \log Pr\left(\bar{m}_i^{(T)} \le \bar{m}_j^{(T)} \mid \theta\right) \mid \\ &\le \sum_{j \ne i} \left(|\mu_i| + |\mu_j| + \frac{\sigma_i^2}{(s)_i} + \frac{\sigma_j^2}{(s)_j}\right) \ . \end{aligned}$$

Under the condition $\alpha_i^{(s)} > 1$, $i = 1, \ldots, k$, from the fact that μ_i follows t-distribution (4) and the expectation of σ_i^2 given by (5) in the main body of the paper, we know $E\left[|\mu_i| \mid \mathcal{E}_s\right] < \infty$, and $E\left[\sigma_i^2 \mid \mathcal{E}_s\right] < \infty$, $i = 1, \ldots, k$.

In other words, $\frac{1}{T} \mid \log E\left[\mathbf{1}\{\mathcal{D}^m(\mathcal{E}_T) \ne [1]\} \mid \theta, \mathcal{E}_s\right] \mid$ is dominated by an integrable function. Applying the Dominated Convergence Theorem, the limit and expectation can be exchanged and the conclusion in the proposition holds. □

Proof of Proposition 4.

Proof. By the assumption of the proposition, from Glynn and Juneja (2004), we have

$$G_{ij}(w_i, w_j; \theta) = \frac{(\mu_i - \mu_j)^2}{2(\sigma_i^2/w_i + \sigma_j^2/w_j)}, \quad j \neq i,$$

which is continuous with respect to μ_i and σ_i^2, $i = 1, \ldots, k$. Obviously, $\max(y_1, y_2)$ is a continuous function of (y_1, y_2), and by induction, $\max(y_1, \ldots, y_{k-1})$ is a continuous function of $(y_1, \ldots, y_{k-1})$. Therefore, $\min_{j \neq i} G_{ij}(w_i, w_j; \theta)$ is continuous with respect to μ_i and σ_i^2, $i = 1, \ldots, k$. Denote the true value of parameter as $\theta = (\mu_1, \ldots, \mu_k, \sigma_1^2, \ldots, \sigma_k^2)$. By the strong law of large numbers

$$\lim_{s\to\infty} \theta^{(s)} = \lim_{s\to\infty} (\mu_1^{(s)}, \ldots, \mu_k^{(s)}, \upsilon_1^{(s)}, \ldots, \upsilon_k^{(s)}) = \theta, \quad a.s.$$

Then from Continuous Mapping Theorem, we have

$$\lim_{s\to\infty} ULDR\left(W, \theta^{(s)}\right) = ULDR\left(W, \theta\right), \quad a.s.,$$

where $ULDR\left(W, \theta\right) \doteq \min_{j\neq[1]} G_{[1]j}(w_{[1]}, w_j; \theta)$ and $[1] \doteq \arg\max_{i=1,\ldots,k} \mu_i$. Define

$$\Psi(\theta) \doteq \sum_{i=1,\ldots,k} \prod_{j\neq i} \mathbf{1}\{\mu_i > \mu_j\} \left(\min_{j\neq i} G_{ij}(w_i, w_j; \theta) \right),$$

and we have

$$\begin{aligned} E[|E[\Psi(\theta)]|\mathcal{E}_s] &\leq E[E[|\Psi(\theta)||\mathcal{E}_s]] = E[|\Psi(\theta)|] \\ &\leq E\left[\sum_{i,j=1,\ldots,k} |G_{ij}(w_i, w_j; \theta)|\right] \leq C \sum_{i,j=1,\ldots,k} E\left[\mu_i^4 + \mu_j^4\right] E\left[\frac{1}{\sigma_i^2 \sigma_j^2}\right] < \infty, \end{aligned}$$

where C is a positive constant. The first inequality is obtained by Jansen's inequality, the equality uses iterated conditional expectation, the second one applies Cauchy inequality, and the third inequality holds because of the assumption of the Normal-Gamma prior. By Doob's Martingale Convergence Theorem (Doob, 1953), we have

$$\lim_{s\to\infty} ULDR\left(W, \mathcal{E}_s\right) = \lim_{s\to\infty} E[\Psi(\theta)|\mathcal{E}_s] = E[\Psi(\theta)|\mathcal{E}_\infty] = ULDR\left(W, \theta\right), \ a.s,$$

where $\mathcal{E}_\infty$ is the minimal σ-algebra generated by $\{\mathcal{E}_s\}_{s\in\mathbb{Z}^+}$. It leads to the conclusion of the theorem. □

Proposition A.3. *For a fixed $W = (w_1, \ldots, w_k) \in \mathcal{S}$, with the "most starving" sequential rule defined in Section 4.2. of the main body of the paper, we have*

$$\lim_{s\to\infty} \tilde{w}_i^{(s)} \to w_i,$$

where $\tilde{w}_i^{(s)} = (s)_i/s$, $i = 1, 2, \ldots, k$.

Proof. Suppose $\mathcal{N} \subset \{1, 2, \ldots, k\}$ is the set of indices of designs that stop receiving allocation of replications at some point. Obviously, for $j \in \mathcal{N}$, we have $\lim_{s\to\infty} \tilde{w}_j^{(s)} = 0$. We also know there exist $i \in \{1, 2, \ldots, k\}$ such that

$$D_i^{(s)} \doteq (1 + \frac{1}{s-1}) \cdot w_i - \tilde{w}_i^{(s-1)} > 0. \tag{b.21}$$

Otherwise, we would have

$$1 = \sum_{i=1}^{k} \left(s \cdot w_i - (s-1) \cdot \tilde{w}_i^{(s-1)} \right) \leq 0,$$

which is a contradiction. Therefore, in the sth step, it is only possible to allocate a replication to the design i such that $D_i^{(s)} > 0$.

For $i \in \mathcal{N}^c \doteq \{1, 2, \ldots, k\} \backslash \mathcal{N}$, define $\tau_i^{(s)} \doteq \max\{t \in \{1, 2, \ldots, s\} : \ (t+1) \cdot \tilde{w}_i^{(t+1)} - t \cdot \tilde{w}_i^{(t)} = 1\}$. By the definition of $\mathcal{N}^c$, we have $\tau_i^{(s)} \to \infty$ $(s \to \infty)$. For $s \in \{t \in \mathbb{Z}^+ : \ t \cdot \tilde{w}_i^{(t)} - (t-1) \cdot \tilde{w}_i^{(t-1)} = 0\}$, we have

$$\tilde{w}_i^{(s)} = \frac{s-1}{s} \tilde{w}_i^{(s-1)} \leq \tilde{w}_i^{(s-1)} \ . \tag{b.22}$$

For $s \in \{t \in \mathbb{Z}^+ : \ t \cdot \tilde{w}_i^{(t)} - (t-1) \cdot \tilde{w}_i^{(t-1)} = 1\}$, we have

$$\tilde{w}_i^{(s)} = \frac{s-1}{s} \tilde{w}_i^{(s-1)} + \frac{1}{s} \ . \tag{b.23}$$

Then, from (b.21), (b.22) and (b.23), we have

$$\tilde{w}_i^{(s)} \leq \tilde{w}_i^{(s-1)} \leq \cdots \leq \tilde{w}_i^{\tau_i^{(s)}+1} \leq \frac{\tau_i^{(s)}}{\tau_i^{(s)}+1} \tilde{w}_i^{(\tau_i^{(s)})} + \frac{1}{\tau_i^{(s)}+1} \leq w_i + \frac{1}{\tau_i^{(s)}+1} \ .$$

For $j \in \mathcal{N}$, $\exists\ S > 0$, for $s > S$, we have $\tilde{w}_j^{(s)} \leq w_j$. Then, $\forall\ i \in \{1, 2, \ldots, k\}$, from (b.21) and (b.22), we have

$$w_i = \tilde{w}_i^{(s)} + \sum_{j \neq i} (\tilde{w}_j^{(s)} - w_j) \leq \tilde{w}_i^{(s)} + \sum_{j \in \mathcal{N}^c} \frac{1}{\tau_j^{(s)}+1} \ .$$

Combining the two inequalities above, we have

$$w_i - \sum_{j \in \mathcal{N}^c} \frac{1}{\tau_j^{(s)}+1} \leq \tilde{w}_i^{(s)} \leq w_i + \frac{1}{\tau_i^{(s)}+1} \ .$$

Therefore, we know $\mathcal{N} = \emptyset$, and

$$\lim_{s\to\infty} \tilde{w}_i^{(s)} \to w_i, \quad i = 1, 2, \ldots, k.$$

$\square$

Proposition A.4. *For a fixed $W = (w_1, \ldots, w_k) \in \mathcal{S}$, under the sequential rule defined in Section 4.2. of the main body of the paper, we have*

(i) $\lim_{s\to\infty} \tilde{w}_i^{(s)} \to w_i$, *a.e.,* $\quad i = 1, \ldots, k$;

(ii) $E[\tilde{w}_i^{(s)}] = w_i, \quad i = 1, \ldots, k$;

(iii) $\forall\ \varepsilon > 0$, *we have*

$$\lim_{s\to\infty} \frac{1}{s} \log Pr(||\tilde{W}^{(s)} - W|| \geq \varepsilon) = -\inf_{\tilde{W}\in\mathcal{O}_\varepsilon} h(\tilde{W}),$$

where $h(\tilde{W}) \doteq \sum_{i=1}^{k} \tilde{w}_i \log(\tilde{w}_i/w_i)$ *and* $\mathcal{O}_\varepsilon \doteq \{\tilde{W} \in \mathcal{S} :\ ||\tilde{W} - W|| \geq \varepsilon\}$.

In Proposition A.4, conclusion (i) states that the randomized sequential rule is eligible, which can be proved by the law of large numbers. Conclusion (ii) states that, following the randomized sequential rule, $\tilde{w}_i^{(s)}$ is an unbiased estimator of w_i, and conclusion (iii) is the large deviations rate for the multinomial distribution, which provides the convergence rate of $\lim_{s\to\infty} \tilde{w}_i^{(s)} = w_i$ in probability. $h(\tilde{W})$ is called the rate function, and it is the Kullback-Leibler information number for the multinomial distribution. The proof for conclusion (iii) can found in Varadhan (2008).

Appendix B: More Numerical Experiments

Example B.1.

This example is similar to the Example 2 in the main body of the paper but has more competing designs (100 designs). The prior distribution is the Normal-Gamma conjugate prior with parameters given as: $\mu_i^{(0)} = 0.1 \times i$, $\kappa_i^{(0)} = 10^3$, $\alpha_i^{(0)} = 3$, and $\beta_i^{(0)} = 10^3 \times i$, $i = 1, 2, \ldots, 100$. From (4) and (5) in the main body of the paper, we know the expectations of the true means $E[\mu_i] = 0.1 \times i$, the variances of the true mean $Var(\mu_1) = i/2$, and the expectation of true variance $E[\sigma_i^2] = 500 \times i$, $i = 1, 2, \ldots, 100$.

From Table 1 and Figure 1, we can see all the allocation policies with APPROX selection policy outperform those with the MEAN selection policy. Moreover, GBBA and KG which include prior information in the allocation procedure are superior to the rest and EA has the worst performance over all the selection policies; KG is better than GBBA at the beginning and is surpassed by the latter when simulation budget exceeds 300.

Example B.2.

The difference between this example and Example 2 in the main body of the paper is that for the latter the design with the largest prior mean has the largest prior variance while the design with the smallest prior mean has the largest prior variance in this example. The prior distribution is the normal-Gamma conjugate prior with parameters given as: $\mu_i^{(0)} = -0.3 \times i$,

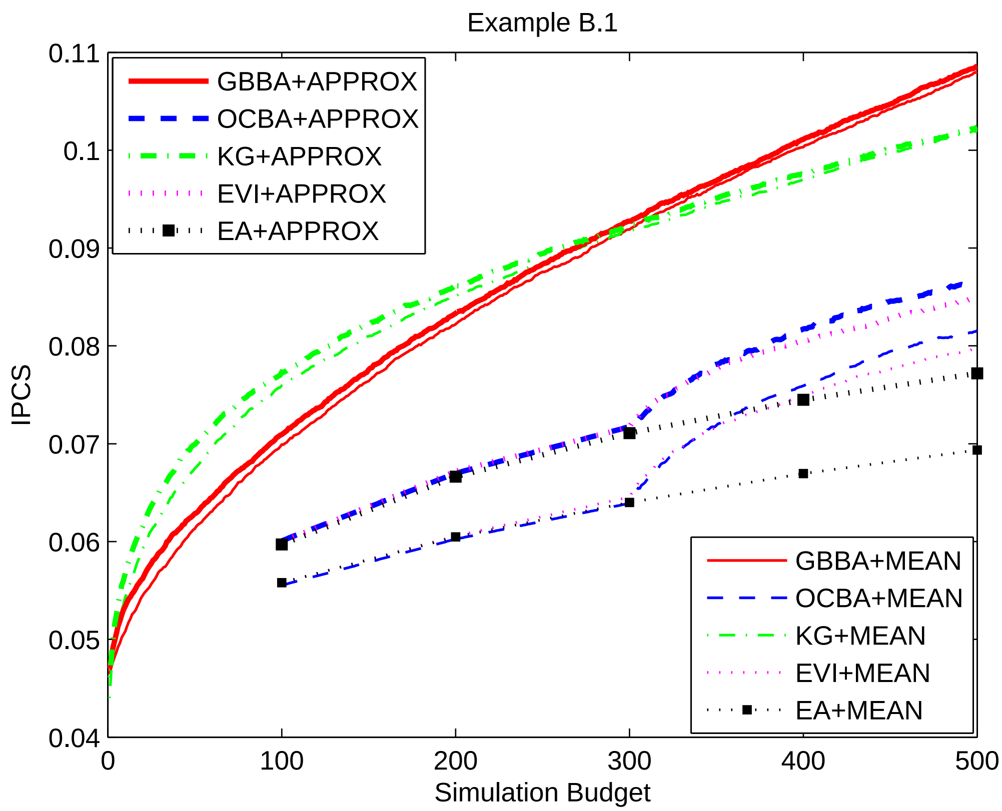

Figure 1: The prior distribution is the normal-Gamma conjugate prior, with parameters $\mu_i^{(0)} = 0.1 \times i$, $\kappa_i^{(0)} = 10^3$, $\alpha_i^{(0)} = 3$, and $\beta_i^{(0)} = 10^3 \times i$, $i = 1, 2, \ldots, 100$. The IPCSs are estimated by 10^6 independent macro replications (precision $\pm$ 0.001 or better).

$T = 100$	GBBA	OCBA	KG	EVI	EA
APPROX	0.071	0.060	0.077	0.060	0.060
MEAN	0.070	0.056	0.076	0.056	0.056
$T = 300$	GBBA	OCBA	KG	EVI	EA
APPROX	0.093	0.072	0.092	0.072	0.071
MEAN	0.092	0.064	0.092	0.065	0.064
$T = 500$	GBBA	OCBA	KG	EVI	EA
APPROX	0.109	0.087	0.102	0.085	0.077
MEAN	0.108	0.082	0.102	0.080	0.069

Table 1: The prior distribution is the normal-Gamma conjugate prior, with parameters $\mu_i^{(0)} = 0.1 \times i$, $\kappa_i^{(0)} = 10^3$, $\alpha_i^{(0)} = 3$, and $\beta_i^{(0)} = 10^3 \times i$, $i = 1, 2, \ldots, 100$. The IPCSs are estimated by 10^6 independent macro replications (precision $\pm$ 0.001 or better).

$\kappa_i^{(0)} = 10^3$, $\alpha_i^{(0)} = 3$, and $\beta_i^{(0)} = 10^3 \times i$, $i = 1, 2, \ldots, 10$. From (4) and (5) in the main body of the paper, we know the expectations of the true means $E[\mu_i] = -0.3 \times i$, the variances of the true mean $Var(\mu_i) = i/2$, and the expectation of true variance $E[\sigma_i^2] = 500 \times i$, $i = 1, 2, \ldots, 10$.

$T = 50$	GBBA	OCBA	KG	EVI	EA
APPROX	0.213	0.207	0.201	0.206	0.207
MEAN	0.205	0.201	0.201	0.200	0.201
$T = 250$	GBBA	OCBA	KG	EVI	EA
APPROX	0.241	0.232	0.218	0.227	0.221
MEAN	0.232	0.220	0.213	0.212	0.207
$T = 450$	GBBA	OCBA	KG	EVI	EA
APPROX	0.262	0.253	0.234	0.244	0.233
MEAN	0.254	0.240	0.227	0.227	0.216

Table 2: The prior distribution is the normal-Gamma conjugate prior, with parameters $\mu_i^{(0)} = -0.3 \times i$, $\kappa_i^{(0)} = 10^3$, $\alpha_i^{(0)} = 3$, and $\beta_i^{(0)} = 10^3 \times i$, $i = 1, 2, \ldots, 10$. The IPCSs are estimated by 10^6 independent macro replications (precision $\pm$ 0.001 or better).

From Table 2 and Figure 2, we can see all the allocation policies with APPROX selection policy outperform those with the MEAN selection policy, GBBA has the best performances, and EA has the worst among five allocation policies.

Example B.3.

This example has the same setting as Example 2 in the main body of the paper. We compare the approximate optimal selection policy choosing from the top-three designs with that considering all k designs as the candidates to be selected, and the most-starving rule

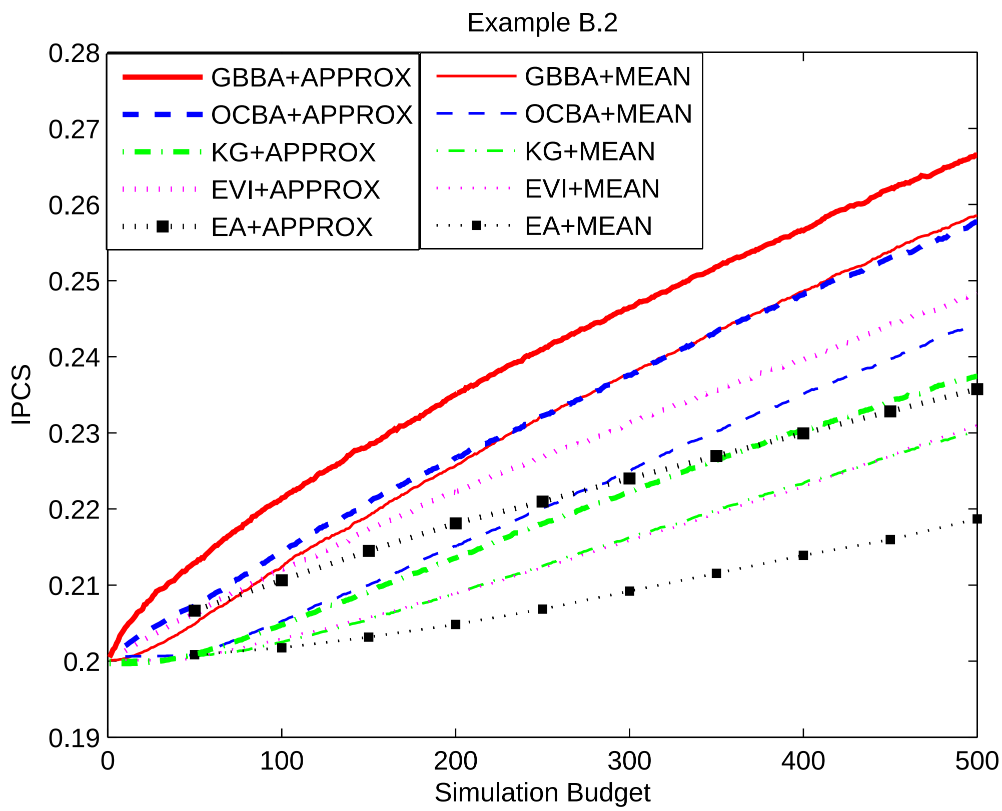

Figure 2: The prior distribution is the normal-Gamma conjugate prior, with parameters $\mu_i^{(0)} = -0.3 \times i$, $\kappa_i^{(0)} = 10^3$, $\alpha_i^{(0)} = 3$, and $\beta_i^{(0)} = 10^3 \times i$, $i = 1, 2, \dots, 10$. The IPCSs are estimated by 10^6 independent macro replications (precision $\pm$ 0.001 or better).

with the randomized rule, introduced in Section 3.3 and Section 4.2, respectively in the main body of the paper.

$T = 50$	GBBA	OCBA	KG	EVI	EA
APPROX-2	0.238	0.231	0.237	0.231	0.231
APPROX-1	0.233	0.226	0.233	0.226	0.227
MEAN	0.221	0.213	0.226	0.213	0.212
$T = 250$	GBBA	OCBA	KG	EVI	EA
APPROX-2	0.275	0.260	0.269	0.261	0.254
APPROX-1	0.270	0.252	0.268	0.253	0.250
MEAN	0.262	0.240	0.264	0.240	0.234
$T = 450$	GBBA	OCBA	KG	EVI	EA
APPROX-2	0.300	0.279	0.291	0.279	0.267
APPROX-1	0.296	0.273	0.290	0.272	0.264
MEAN	0.290	0.262	0.288	0.260	0.249

Table 3: The prior distribution is the normal-Gamma conjugate prior, with parameters $\mu_i^{(0)} = 0.1 \times i$, $\kappa_i^{(0)} = 10^3$, $\alpha_i^{(0)} = 3$, and $\beta_i^{(0)} = 10^3 \times i$, $i = 1, 2, \ldots, 10$. The IPCSs are estimated by 10^6 independent macro replications (precision $\pm$ 0.001 or better).

$T = 50$	R-GBBA	S-GBBA	R-OCBA	S-OCBA
APPROX	0.233	0.233	0.227	0.226
MEAN	0.221	0.222	0.212	0.213
$T = 250$	R-GBBA	S-GBBA	R-OCBA	S-OCBA
APPROX	0.270	0.268	0.255	0.252
MEAN	0.263	0.261	0.242	0.240
$T = 450$	R-GBBA	S-GBBA	R-OCBA	S-OCBA
APPROX	0.296	0.293	0.276	0.273
MEAN	0.290	0.288	0.266	0.262

Table 4: The prior distribution is the normal-Gamma conjugate prior, with parameters $\mu_i^{(0)} = 0.1 \times i$, $\kappa_i^{(0)} = 10^3$, $\alpha_i^{(0)} = 3$, and $\beta_i^{(0)} = 10^3 \times i$, $i = 1, 2, \ldots, 10$. The IPCSs are estimated by 10^6 independent macro replications (precision $\pm$ 0.001 or better).

In Figure 3, APPROX-1 stands for the approximate selection policy which chooses the top-three designs with largest posterior mean and is introduced in the second paragraph of Section 3.3 (at the bottom of page 13) in the main body of the paper, and APPROX-2 denotes the scheme which considers all k designs as the candidates to be selected and is introduced in Section 3.3 (at the top of page 14). From Table 3 and Figure 2, we can see APPROX-2 is better than APPROX-1 and both of the approximate selection policies are better than MEAN which chooses the design with the largest posterior mean.

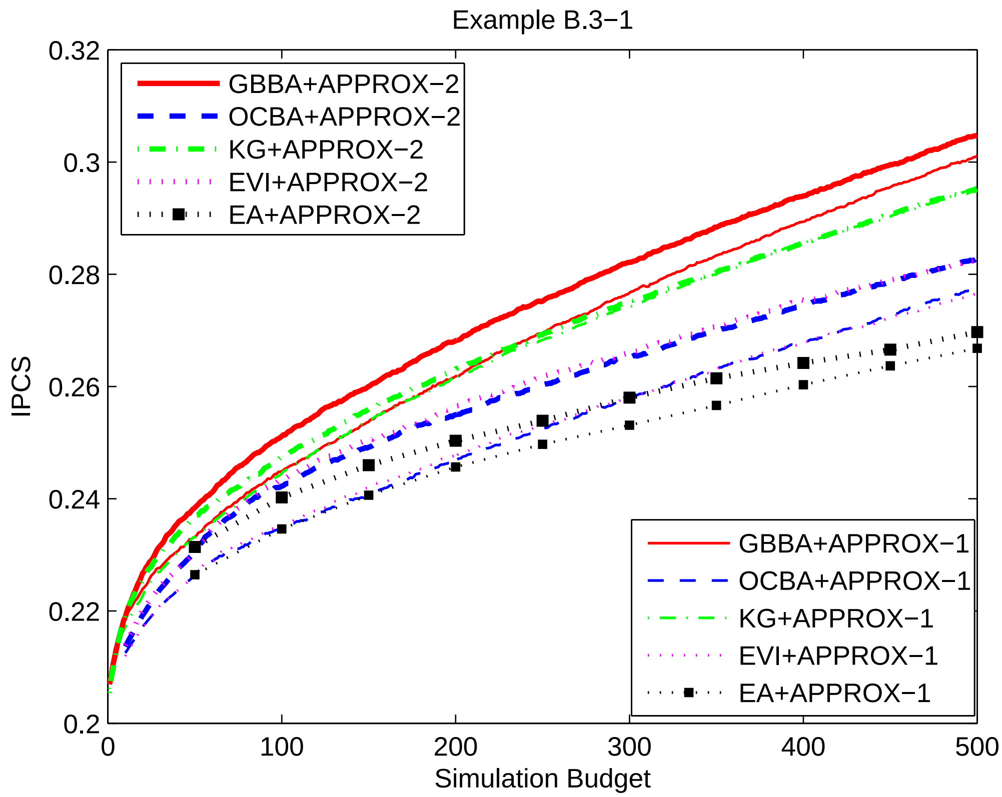

Figure 3: The prior distribution is the normal-Gamma conjugate prior, with parameters $\mu_i^{(0)} = 0.1 \times i$, $\kappa_i^{(0)} = 10^3$, $\alpha_i^{(0)} = 3$, and $\beta_i^{(0)} = 10^3 \times i$, $i = 1, 2, \ldots, 10$. The IPCSs are estimated by 10^6 independent macro replications (precision $\pm$ 0.001 or better).

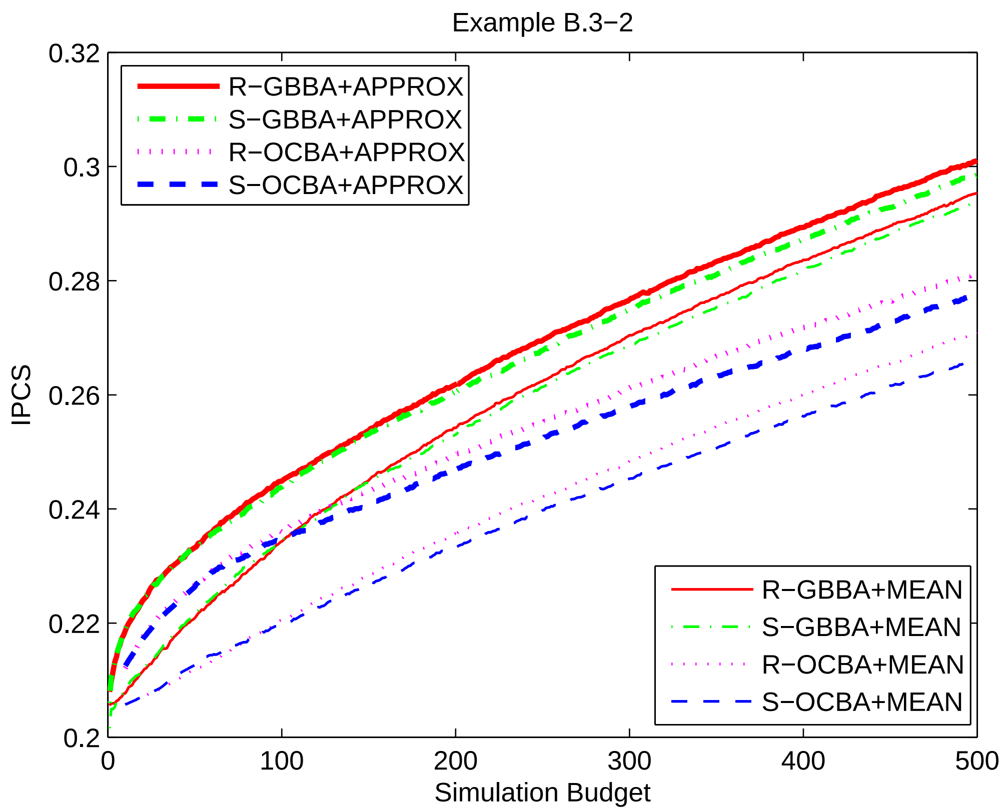

Figure 4: The prior distribution is the normal-Gamma conjugate prior, with parameters $\mu_i^{(0)} = 0.1 \times i$, $\kappa_i^{(0)} = 10^3$, $\alpha_i^{(0)} = 3$, and $\beta_i^{(0)} = 10^3 \times i$, $i = 1, 2, \ldots, 10$. The IPCSs are estimated by 10^6 independent macro replications (precision $\pm$ 0.001 or better).

In Figure 4, R-GBBA and R-OCBA stand for allocation policies which use randomized sequential rule combined with the asymptotic sampling statistic introduced in equation (13) of the main body of the paper and the OCBA formula, respectively; S-GBBA and S-OCBA stand for allocation policies which use the "most starving" rule combined with the asymptotic sampling statistic and OCBA formula, respectively. From Table 4 and Figure 4, R-GBBA and R-OCBA are slightly better than S-GBBA and S-OCBA, respectively, over all selection policies.

References

Azaïs, J.-M. and Wschebor, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. John Wiley and Sons.

Doob, J. L. (1953). *Stochastic Processes*, volume 101. New York Wiley.

Glynn, P. W. and Juneja, S. (2004). A large deviations perspective on ordinal optimization. In *Proceedings of the 2004 Winter Simulation Conference*, pages 577–585. IEEE.

Van der Vaart, A. W. (2000). *Asymptotic Statistics*. Cambridge University Press.

Varadhan, S. (2008). Large deviations. *The Annals of Probability*, 36(2):397–419.