

Online Appendix for “Gradient-Based Myopic Allocation Policy: An Efficient Sampling Procedure in a Low-Confidence Scenario”

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

I. THEORETICAL SUPPLEMENTS

Approximation of IPCS

The predictive distribution of x_{i,n_i+1} is given by (3) in the main body of the note, so

$$\begin{aligned} \int_{x_{i,n_i+1} < r_i^{(n)}} x_{i,n_i+1} P(dx_{i,n_i+1} | \mathcal{E}_n) &= \frac{1}{\sqrt{2\pi v_i^{(n)}}} \int_{-\infty}^{r_i^{(n)}} x \exp\left(-\frac{(x - \mu_i^{(n)})^2}{2v_i^{(n)}}\right) dx \\ &= \mu_i^{(n)} - \frac{1}{\sqrt{2\pi v_i^{(n)}}} \int_{r_i^{(n)}}^{\infty} x \exp\left(-\frac{(x - \mu_i^{(n)})^2}{2v_i^{(n)}}\right) dx = -v_i^{(n)} \phi(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)}) + \mu_i^{(n)} \Phi(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)}), \end{aligned}$$

where $\phi(\cdot | \mu, \sigma^2)$ and $\Phi(\cdot | \mu, \sigma^2)$ are the respective density and distribution functions of the normal distribution with mean μ and variance σ^2 , and

$$\begin{aligned} \int_{x_{i,n_i+1} \geq r_i^{(n)}} x_{i,n_i+1} P(dx_{i,n_i+1} | \mathcal{E}_n) &= \frac{1}{\sqrt{2\pi v_i^{(n)}}} \int_{r_i^{(n)}}^{\infty} x \exp\left(-\frac{(x - \mu_i^{(n)})^2}{2v_i^{(n)}}\right) dx \\ &= \mu_i^{(n)} \bar{\Phi}(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)}) + v_i^{(n)} \phi(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)}), \end{aligned}$$

where $\bar{\Phi}(\cdot) = 1 - \Phi(\cdot)$. Thus, an approximation of (4) in the main body of the note, using a first-order Taylor series

expansion is given by

$$\begin{aligned}
APCS_0(a_{n+1} = i; \mathcal{E}_n) &\doteq \bar{\Phi}(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)}) (P(\mu_i > \mu_j, j \neq i | \mathcal{E}_n) \\
&+ \frac{\partial P(\mu_i > \mu_j, j \neq i | \mathcal{E}_n)}{d\mu_i^{(n)}} \frac{\partial \mu_i^{(n+1)}(x_{i,n_i+1}, n_i + \delta)}{\partial \delta} \Big|_{x_{i,n_i+1}, \delta=0} + \frac{\partial P(\mu_i > \mu_j, j \neq i | \mathcal{E}_n)}{d\vartheta_i^{(n)}} \frac{\partial \vartheta_i^{(n+1)}(n_i + \delta)}{\partial \delta} \Big|_{\delta=0}) \\
&+ \Phi(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)}) (P(\mu_{i_n} > \mu_j, j \neq i_n | \mathcal{E}_n) + \frac{\partial P(\mu_{i_n} > \mu_j, j \neq i_n | \mathcal{E}_n)}{d\mu_i^{(n)}} \frac{\partial \mu_i^{(n+1)}(x_{i,n_i+1}, n_i + \delta)}{\partial \delta} \Big|_{x_{i,n_i+1}, \delta=0} \\
&+ \frac{\partial P(\mu_{i_n} > \mu_j, j \neq i_n | \mathcal{E}_n)}{d\vartheta_i^{(n)}} \frac{\partial \vartheta_i^{(n+1)}(n_i + \delta)}{\partial \delta} \Big|_{\delta=0}) + (\mu_i^{(n)} \bar{\Phi}(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)}) + v_i^{(n)} \phi(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)})) \\
&\times \frac{\partial P(\mu_i > \mu_j, j \neq i | \mathcal{E}_n)}{d\mu_i^{(n)}} \frac{\partial \mu_i^{(n+1)}(x_{i,n_i+1}, n_i + \delta)}{\partial x_{i,n_i+1}} \Big|_{x_{i,n_i+1}, \delta=0} + (\mu_i^{(n)} \Phi(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)}) - v_i^{(n)} \phi(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)})) \\
&\times \frac{\partial P(\mu_{i_n} > \mu_j, j \neq i_n | \mathcal{E}_n)}{d\mu_i^{(n)}} \frac{\partial \mu_i^{(n+1)}(x_{i,n_i+1}, n_i + \delta)}{\partial x_{i,n_i+1}} \Big|_{x_{i,n_i+1}, \delta=0}.
\end{aligned}$$

By using the chain rule, we have the following expression:

$$\begin{aligned}
APCS_0(a_{n+1} = i; \mathcal{E}_n) &= \bar{\Phi}(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)}) \left[P(\mu_i > \mu_j, j \neq i | \mathcal{E}_n) - \sum_{j \neq i} \frac{\partial P(z_1^{(i)} > 0, \dots, z_{k-1}^{(i)} > 0)}{\partial \pi_j^{(i)}} \frac{\partial \pi_j^{(i)}}{\partial \mu_i^{(n)}} \frac{\mu_i^{(n)} \vartheta_i^{(n)}}{\sigma_i^2} \right. \\
&- \left. \left(\sum_{j \neq i} \frac{\partial P(z_1^{(i)} > 0, \dots, z_{k-1}^{(i)} > 0)}{\partial \pi_j^{(i)}} \frac{\partial \pi_j^{(i)}}{\partial \vartheta_i^{(n)}} + \sum_{j \neq j', j, j' \neq i} \frac{\partial P(z_1^{(i)} > 0, \dots, z_{k-1}^{(i)} > 0)}{\partial \rho_{j,j'}^{(i)}} \frac{\partial \rho_{j,j'}^{(i)}}{\partial \vartheta_i^{(n)}} \right) \frac{(\vartheta_i^{(n)})^2}{\sigma_i^2} \right] \\
&+ \Phi(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)}) \left[P(\mu_{i_n} > \mu_j, j \neq i_n | \mathcal{E}_n) - \sum_{j \neq i_n} \frac{\partial P(z_1^{(i_n)} > 0, \dots, z_{k-1}^{(i_n)} > 0)}{\partial \pi_j^{(i_n)}} \frac{\partial \pi_j^{(i_n)}}{\partial \mu_i^{(n)}} \frac{\mu_i^{(n)} \vartheta_i^{(n)}}{\sigma_i^2} \right. \\
&- \left. \left(\sum_{j \neq i_n} \frac{\partial P(z_1^{(i_n)} > 0, \dots, z_{k-1}^{(i_n)} > 0)}{\partial \pi_j^{(i_n)}} \frac{\partial \pi_j^{(i_n)}}{\partial \vartheta_i^{(n)}} + \sum_{j \neq j', j, j' \neq i_n} \frac{\partial P(z_1^{(i_n)} > 0, \dots, z_{k-1}^{(i_n)} > 0)}{\partial \rho_{j,j'}^{(i_n)}} \frac{\partial \rho_{j,j'}^{(i_n)}}{\partial \vartheta_i^{(n)}} \right) \frac{(\vartheta_i^{(n)})^2}{\sigma_i^2} \right] \\
&+ (\mu_i^{(n)} \bar{\Phi}(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)}) + v_i^{(n)} \phi(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)})) \left(\sum_{j \neq i} \frac{\partial P(z_1^{(i)} > 0, \dots, z_{k-1}^{(i)} > 0)}{\partial \pi_j^{(i)}} \frac{\partial \pi_j^{(i)}}{\partial \mu_i^{(n)}} \frac{\vartheta_i^{(n)}}{\sigma_i^2} \right) \\
&+ (\mu_i^{(n)} \Phi(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)}) - v_i^{(n)} \phi(r_i^{(n)} | \mu_i^{(n)}, v_i^{(n)})) \left(\sum_{j \neq i_n} \frac{\partial P(z_1^{(i_n)} > 0, \dots, z_{k-1}^{(i_n)} > 0)}{\partial \pi_j^{(i_n)}} \frac{\partial \pi_j^{(i_n)}}{\partial \mu_i^{(n)}} \frac{\vartheta_i^{(n)}}{\sigma_i^2} \right).
\end{aligned}$$

To reduce the computational complexity of the high-dimensional integral, we use the lower bound of the integral offered by the Slepian inequality (Chick and Inoue, 2001, Branke et al., 2007) as a surrogate in the computational

procedure:

$$\begin{aligned}
& \int_{\{\hat{z}_l^{(b)} > 0, l=1, \dots, k-2\}} \phi_{k-2} \left(\hat{z}_1^{(b)}, \dots, \hat{z}_{k-2}^{(b)} \mid \tilde{\Pi}^{(b)}[j], \tilde{\Sigma}^{(b)}[j] \right) d\hat{z}^{(b)} \\
& \geq \prod_{l=1}^{k-2} \bar{\Phi} \left(0 \mid \tilde{\Pi}_l^{(b)}[j], \tilde{\Sigma}_{l,l}^{(b)}[j] \right), \\
& \int_{\{\hat{z}_l^{(b)} > 0, l=1, \dots, k-3\}} \phi_{k-3} \left(\hat{z}_1^{(b)}, \dots, \hat{z}_{k-3}^{(b)} \mid \hat{\Pi}^{(b)}[j, j'], \hat{\Sigma}^{(b)}[j, j'] \right) d\hat{z}^{(b)} \\
& \geq \prod_{l=1}^{k-3} \bar{\Phi} \left(0 \mid \hat{\Pi}_l^{(b)}[j, j'], \hat{\Sigma}_{l,l}^{(b)}[j, j'] \right),
\end{aligned}$$

where $\Phi(\cdot \mid \tilde{\Pi}_l^{(b)}[j], \tilde{\Sigma}_{l,l}^{(b)}[j])$, $\bar{\Phi}(\cdot \mid \hat{\Pi}_l^{(b)}[j, j'], \hat{\Sigma}_{l,l}^{(b)}[j, j'])$ are the marginal distribution functions of $\hat{z}_l^{(b)}$ and $\hat{z}_l^{(b)}$.

Similarly,

$$\begin{aligned}
& P(\mu_b > \mu_l, l \neq b \mid \mathcal{E}_n) \\
& = P(z_1^{(b)} > 0, \dots, z_{k-1}^{(b)} > 0) \geq \prod_{j=1}^{k-1} \bar{\Phi}(0 \mid \Pi_j^{(b)}, \Sigma_{j,j}^{(b)}),
\end{aligned}$$

where $\Phi(\cdot \mid \Pi_j^{(b)}, \Sigma_{j,j}^{(b)})$ is the marginal distribution function of $z_j^{(b)}$.

With the results in Theorem 1 in the main body of the note and Slepian inequality approximation, we have the following approximation for (4):

$$\begin{aligned}
& APCS(a_{n+1} = i; \mathcal{E}_n) = \bar{\Phi}(r_i^{(n)} \mid \mu_i^{(n)}, v_i^{(n)}) \left[\prod_{j=1}^{k-1} \bar{\Phi}(0 \mid \Pi_j^{(i)}, \Sigma_{j,j}^{(i)}) - \sum_{j \neq i} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\pi_j^{(i)})^2}{2}} \right. \\
& \times \prod_{l=1}^{k-2} \bar{\Phi} \left(0 \mid \tilde{\Pi}_l^{(i)}[j], \tilde{\Sigma}_{l,l}^{(i)}[j] \right) \frac{\partial \pi_j^{(i)}}{\partial \mu_i^{(n)}} \frac{\mu_i^{(n)} \vartheta_i^{(n)}}{\sigma_i^2} - \left(\sum_{j \neq i} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\pi_j^{(i)})^2}{2}} \prod_{l=1}^{k-2} \bar{\Phi} \left(0 \mid \tilde{\Pi}_l^{(i)}[j], \tilde{\Sigma}_{l,l}^{(i)}[j] \right) \frac{\partial \pi_j^{(i)}}{\partial \vartheta_i^{(n)}} \right. \\
& \left. \left. + \sum_{j \neq j', j, j' \neq i} \phi_2(0 \mid \Pi_{j,j'}^{(i)}, \Sigma_{j,j',j,j'}^{(i)}) \times \prod_{l=1}^{k-3} \bar{\Phi} \left(0 \mid \hat{\Pi}_l^{(i)}[j, j'], \hat{\Sigma}_{l,l}^{(i)}[j, j'] \right) \frac{\partial \rho_{j,j'}^{(i)}}{\partial \vartheta_i^{(n)}} \right) \frac{(\vartheta_i^{(n)})^2}{\sigma_i^2} \right] \\
& + \bar{\Phi}(r_i^{(n)} \mid \mu_i^{(n)}, v_i^{(n)}) \left[\prod_{j=1}^{k-1} \bar{\Phi}(0 \mid \Pi_j^{(i_n)}, \Sigma_{j,j}^{(i_n)}) - \sum_{j \neq i_n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\pi_j^{(i_n)})^2}{2}} \prod_{l=1}^{k-2} \bar{\Phi} \left(0 \mid \tilde{\Pi}_l^{(i_n)}[j], \tilde{\Sigma}_{l,l}^{(i_n)}[j] \right) \frac{\partial \pi_j^{(i_n)}}{\partial \mu_i^{(n)}} \frac{\mu_i^{(n)} \vartheta_i^{(n)}}{\sigma_i^2} \right. \\
& - \left(\sum_{j \neq i_n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\pi_j^{(i_n)})^2}{2}} \prod_{l=1}^{k-2} \bar{\Phi} \left(0 \mid \tilde{\Pi}_l^{(i_n)}[j], \tilde{\Sigma}_{l,l}^{(i_n)}[j] \right) \frac{\partial \pi_j^{(i_n)}}{\partial \vartheta_i^{(n)}} + \sum_{j \neq j', j, j' \neq i_n} \phi_2(0 \mid \Pi_{j,j'}^{(i_n)}, \Sigma_{j,j',j,j'}^{(i_n)}) \right. \\
& \left. \left. \times \prod_{l=1}^{k-3} \bar{\Phi} \left(0 \mid \hat{\Pi}_l^{(i_n)}[j, j'], \hat{\Sigma}_{l,l}^{(i_n)}[j, j'] \right) \frac{\partial \rho_{j,j'}^{(i_n)}}{\partial \vartheta_i^{(n)}} \right) \frac{(\vartheta_i^{(n)})^2}{\sigma_i^2} \right] + \left(\mu_i^{(n)} \bar{\Phi}(r_i^{(n)} \mid \mu_i^{(n)}, v_i^{(n)}) + v_i^{(n)} \phi(r_i^{(n)} \mid \mu_i^{(n)}, v_i^{(n)}) \right) \\
& \times \sum_{j \neq i} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\pi_j^{(i)})^2}{2}} \prod_{l=1}^{k-2} \bar{\Phi} \left(0 \mid \tilde{\Pi}_l^{(i)}[j], \tilde{\Sigma}_{l,l}^{(i)}[j] \right) \frac{\partial \pi_j^{(i)}}{\partial \mu_i^{(n)}} \frac{\vartheta_i^{(n)}}{\sigma_i^2} + \left(\mu_i^{(n)} \bar{\Phi}(r_i^{(n)} \mid \mu_i^{(n)}, v_i^{(n)}) - v_i^{(n)} \phi(r_i^{(n)} \mid \mu_i^{(n)}, v_i^{(n)}) \right) \\
& \times \sum_{j \neq i_n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\pi_j^{(i_n)})^2}{2}} \prod_{l=1}^{k-2} \bar{\Phi} \left(0 \mid \tilde{\Pi}_l^{(i_n)}[j], \tilde{\Sigma}_{l,l}^{(i_n)}[j] \right) \frac{\partial \pi_j^{(i_n)}}{\partial \mu_i^{(n)}} \frac{\vartheta_i^{(n)}}{\sigma_i^2},
\end{aligned}$$

which gives the same formula of APCS in the main body of the note by rearranging the terms in above formula.

II. NUMERICAL SUPPLEMENTS

Example 1:

This is the same example in the main body of the note, but we compare G-MAP with two well-known upper confidence interval (UCB) algorithms (see Auer et al., 2002):

$$a_{n+1} = \arg \max_{1, \dots, k} \left\{ \bar{m}_i^{n_i} + \sqrt{\frac{2 \log n}{n_i}} \right\},$$

and

$$a_{n+1} = \arg \max_{1, \dots, k} \left\{ \bar{m}_i^{n_i} + 4 \sqrt{\frac{\bar{v}_i^{n_i} \log(n-1)}{n_i - 1}} \right\},$$

where $\bar{v}_i^{n_i}$ is the sample variance of alternative i estimated by the allocated n_i replications. The first algorithm is denoted as UCB1, and the second algorithm is denoted as UCB2. They both decrease the IPCS in this low-confidence scenario from Figure 1.

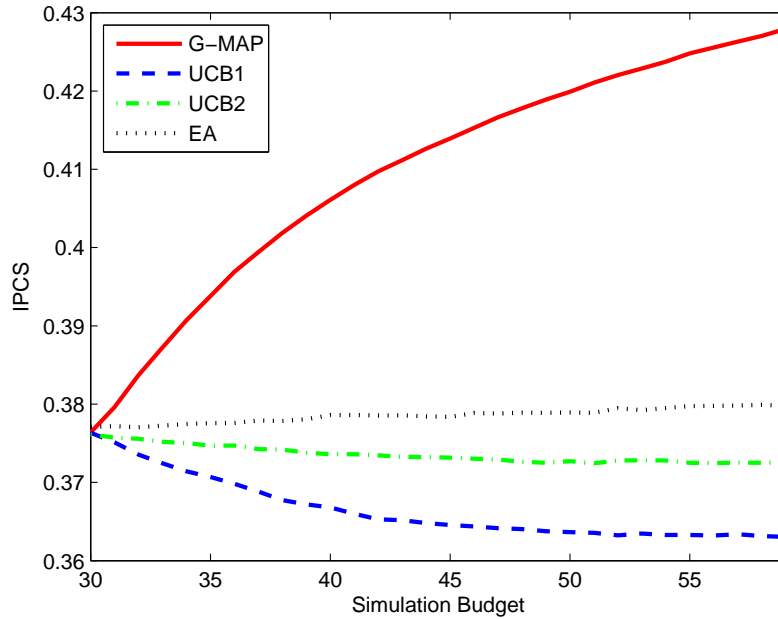


Fig. 1. The prior distribution is the normal conjugate prior, with hyper-parameters $\mu_i^{(0)} = 0$, $i = 1, 2, 3$, $\sigma_1^{(0)} = 0.01$, and $\sigma_2^{(0)} = \sigma_3^{(0)} = 0.001$. The true variances are $\sigma_i = 1$, $i = 1, 2, 3$. The number of initial replications is $n_0 = 10$. The IPCS is estimated by 10^6 independent macro replications (precision $\pm 10^{-3}$).

Example 2:

In this example, there are three alternatives with true means following conjugate priors with hyper-parameters $\mu_i^{(0)} = 0$, $i = 1, 2, 3$, $\sigma_1^{(0)} = \sqrt{0.1}$, and $\sigma_2^{(0)} = \sigma_3^{(0)} = 0.1\sqrt{0.1}$. Similar to the first example, the first alternative is more likely to be the best, but this example is somewhere between a low-confidence and high-confidence scenario.

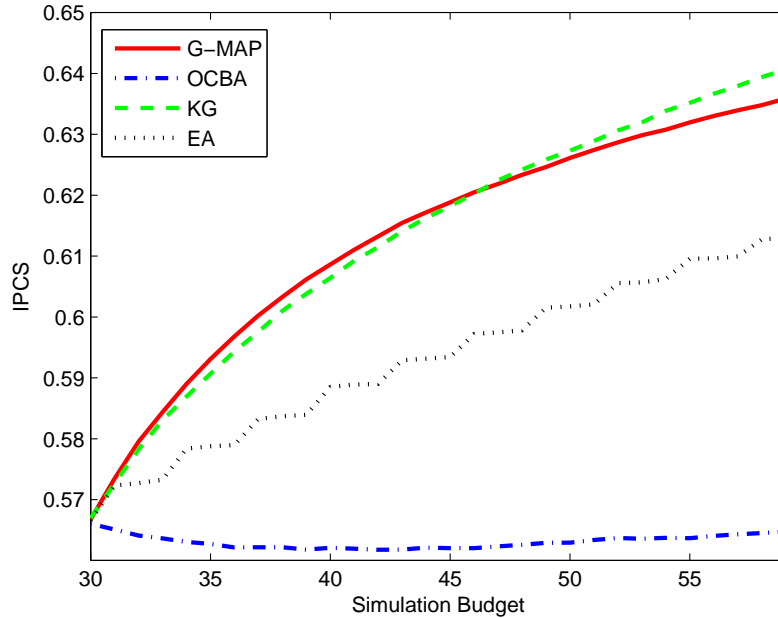


Fig. 2. The prior distribution is the normal conjugate prior, with hyper-parameters $\mu_i^{(0)} = 0$, $i = 1, 2, 3$, $\sigma_1^{(0)} = \sqrt{0.1}$, and $\sigma_2^{(0)} = \sigma_3^{(0)} = 0.1\sqrt{0.1}$. The true variances are $\sigma_i = 1$, $i = 1, 2, 3$. The number of initial replications is $n_0 = 10$. The IPCS is estimated by 10^6 independent macro replications (precision $\pm 10^{-3}$).

From Figure 2, we can see that the IPCS of OCBA again decreases as the simulation budget grows, but not as dramatically as in the first example. The IPCS of EA follows a zigzag path, but the trend is upward. G-MAP and KG are significantly better than EA. G-MAP has a slight edge over KG at the beginning but falls behind KG after the simulation budget reaches around 45, which might indicate that the approximation in G-MAP overstates the influence of induced correlations relative to the means and variances in the high-confidence scenario.

Example 3:

This example is analogous to the example in the main body of the note but with more alternatives. There are ten alternatives with true means following conjugate priors with hyper-parameters $\mu_i^{(0)} = 0$, $i = 1, \dots, 10$, $\sigma_1^{(0)} = 0.01$, and $\sigma_i^{(0)} = 0.001$, $i = 2, \dots, 10$. In Figure 3. we can see the numerical results are similar to the example in the main body of the note.

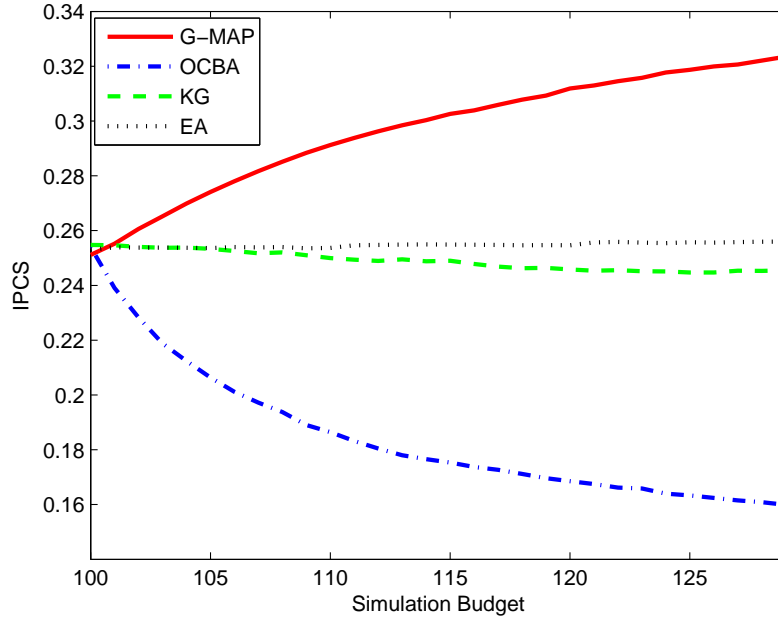


Fig. 3. The prior distribution is the normal conjugate prior, with hyper-parameters $\mu_i^{(0)} = 0$, $i = 1, \dots, 10$, $\sigma_1^{(0)} = 0.01$, and $\sigma_i^{(0)} = 0.001$, $i = 2, \dots, 10$. The true variances are $\sigma_i = 1$, $i = 1, \dots, 10$. The number of initial replications is $n_0 = 10$. The IPCS is estimated by 10^5 independent macro replications (precision $\pm 10^{-2}/\sqrt{10}$).

REFERENCES

- Auer, P., Cesa-Bianchi, N., and Fischer, P. (2002). Finite-time analysis of the multiarmed bandit problem. *Mach. Learn.*, 47(2-3):235–256.
- Branke, J., Chick, S. E., and Schmidt, C. (2007). Selecting a selection procedure. *Manag. Sci.*, 53(12):1916–1932.
- Chick, S. E. and Inoue, K. (2001). New two-stage and sequential procedures for selecting the best simulated system. *Oper. Res.*, 49(5):732–743.