Online Appendix for "Ranking and Selection as Stochastic Control"

Yijie Peng, Edwin K. P. Chong, Chun-Hung Chen and Michael C. Fu

April 8, 2017

Optimal A&S Policy for Three Bernoulli Distributed Alternatives

We show how to optimally allocate three replications among three designs following independent Bernoulli distributions. We assume the unknown parameters θ_i, $i = 1, \ldots, k$, follow the conjugate priors introduced in Section III.B in the main body of the paper. The hyper-parameters for the priors are set as $\alpha_i^{(0)} = \beta_i^{(0)} = 1/2$, $i = 1, 2, 3$. Because of the symmetry in the prior information for different designs, the calculation for many states can be saved.

Figure 1 shows some of the possible states after allocating three replications. We only need to consider three possible sampling allocation combination: $T_1 = 3$, $T_2 = 0$, $T_3 = 0$; $T_1 = 2$, $T_2 = 1$, $T_3 = 0$; $T_1 = 1$, $T_2 = 1$, $T_3 = 1$, since other sampling allocation combinations lead to equivalent states because of the symmetric prior information. For each combination, arbitrarily choosing three nodes with each one from the roots of different trees constitutes a state. For example, for $T_1 = 3$, $T_2 = 0$, $T_3 = 0$, there are four possible states: $(3, 0, 0, 3, 0, 0)$, $(2, 0, 0, 3, 0, 0)$, $(1, 0, 0, 3, 0, 0)$. $(0, 0, 0, 3, 0, 0)$. The optimal selection policy for EOC can be calculated analytically, and the optimal selection policy for PCS is estimated by simulation. By conjugacy, the posterior distributions of θ_i, $i = 1, \ldots, k$, are beta distributions. Expected value functions $V_T(\mathcal{E}_T; i) = \mathbb{E}\left[V(\theta; i)|\mathcal{E}_T\right]$, $i = 1, \ldots, k$, which do not have analytical form, are estimated by running 10^7 macro-experiments, which leads to a precision of up to 10^{-3}. We do not differentiate selection and allocation decisions leading to rewards that are statistically non-differentiable, based on the precision level of numerical experiments. The numerical results are provided as follows:

$$\mathcal{D}_P^*(3, 0, 0, 3, 0, 0) = \mathcal{D}_E^*(3, 0, 0, 3, 0, 0) = 1,$$
$$V_P(3, 0, 0, 3, 0, 0) = 0.670, \quad V_E(3, 0, 0, 3, 0, 0) = -0.049,$$

$$\mathcal{D}_P^*(2, 0, 0, 3, 0, 0) = \mathcal{D}_E^*(2, 0, 0, 3, 0, 0) = 1,$$
$$V_P(2, 0, 0, 3, 0, 0) = 0.372, \quad V_E(2, 0, 0, 3, 0, 0) = -0.190,$$

$$\mathcal{D}_P^*(1, 0, 0, 3, 0, 0) = \mathcal{D}_E^*(1, 0, 0, 3, 0, 0) = 2,$$
$$V_P(1, 0, 0, 3, 0, 0) = 0.404, \quad V_E(1, 0, 0, 3, 0, 0) = -0.248,$$

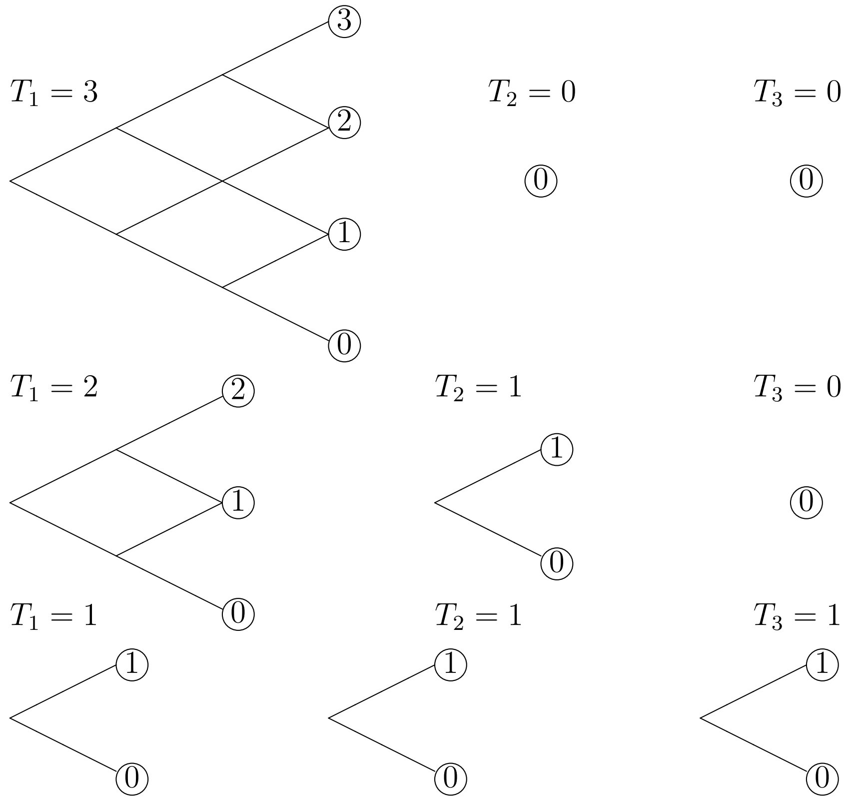

Figure 1: State space for $T = 3$.

$$\mathcal{D}_P^*(0,0,0,3,0,0) = \mathcal{D}_E^*(0,0,0,3,0,0) = 2,$$
$$V_P(0,0,0,3,0,0) = 0.471, \quad V_E(0,0,0,3,0,0) = -0.211,$$

$$\mathcal{D}_P^*(2,1,0,2,1,0) = \mathcal{D}_E^*(2,1,0,2,1,0) = 1,$$
$$V_P(2,1,0,2,1,0) = 0.495, \quad V_E(2,1,0,2,1,0) = -0.097,$$

$$\mathcal{D}_P^*(1,1,0,2,1,0) = \mathcal{D}_E^*(1,1,0,2,1,0) = 2,$$
$$V_P(1,1,0,2,1,0) = 0.595, \quad V_E(1,1,0,2,1,0) = -0.107,$$

$$\mathcal{D}_P^*(0,1,0,2,1,0) = \mathcal{D}_E^*(0,1,0,2,1,0) = 2,$$
$$V_P(0,1,0,2,1,0) = 0.685, \quad V_E(0,1,0,2,1,0) = -0.081,$$

$$\mathcal{D}_P^*(2,0,0,2,1,0) = \mathcal{D}_E^*(2,0,0,2,1,0) = 1,$$
$$V_P(2,0,0,2,1,0) = 0.744, \quad V_E(2,0,0,2,1,0) = -0.048,$$

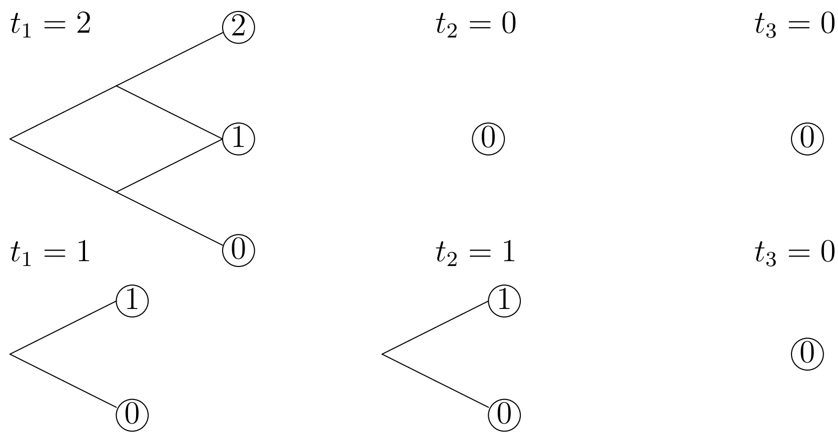

Figure 2: State space for $t = 2$.

$$\mathcal{D}_P^*(1,0,0,2,1,0) = \mathcal{D}_E^*(1,0,0,2,1,0) = 1,$$
$$V_P(1,0,0,2,1,0) = 0.418, \quad V_E(1,0,0,2,1,0) = -0.206,$$

$$\mathcal{D}_P^*(0,0,0,2,1,0) = \mathcal{D}_E^*(0,0,0,2,1,0) = 3,$$
$$V_P(0,0,0,2,1,0) = 0.629, \quad V_E(0,0,0,2,1,0) = -0.103,$$

$$\mathcal{D}_P^*(1,1,1,1,1,1) = \mathcal{D}_E^*(1,1,1,1,1,1) \in \{1,2,3\},$$
$$V_P(1,1,1,1,1,1) = 1/3, \quad V_E(1,1,1,1,1,1) = -0.184,$$

$$\mathcal{D}_P^*(1,1,0,1,1,1) = \mathcal{D}_E^*(1,1,0,1,1,1) \in \{1,2\},$$
$$V_P(1,1,0,1,1,1) = 0.485, \quad V_E(1,1,0,1,1,1) = -0.139,$$

$$\mathcal{D}_P^*(1,0,0,1,1,1) = \mathcal{D}_E^*(1,0,0,1,1,1) = 1,$$
$$V_P(1,0,0,1,1,1) = 0.840, \quad V_E(1,0,0,1,1,1) = -0.036,$$

$$\mathcal{D}_P^*(0,0,0,1,1,1) = \mathcal{D}_E^*(0,0,0,1,1,1) \in \{1,2,3\},$$
$$V_P(1,1,0,1,1,1) = 1/3, \quad V_E(1,1,0,1,1,1) = -0.222\ .$$

For $T_1 = 1$, $T_2 = 1$, $T_3 = 1$, the decision and value function evaluated at other states can be obtained by a symmetry argument. The optimal selection policy is consistent with the selection policy of choosing the design with the largest posterior (sample) mean.

Figure 2 shows some of the possible states after allocating two replications. Other states are equivalent to the states in the figure by a symmetry argument. The allocation decision and value function for each state can be calculated by backward induction as follows:

$$\begin{aligned}
V(2,0,0,2,0,0;1) &= q(1|2,2)\ V(3,0,0,3,0,0) + q(0|2,2)\ V(2,0,0,3,0,0),\\
V_P(2,0,0,2,0,0;1) &= \frac{5}{6}\times 0.670 + \frac{1}{6}\times 0.372 = 0.620,\\
V_E(2,0,0,2,0,0;1) &= \frac{5}{6}\times(-0.049) + \frac{1}{6}\times(-0.190) = -0.072,
\end{aligned}$$

$$\begin{aligned}
V(2,0,0,2,0,0;2) &= q(1|0,0)\ V(2,1,0,2,1,0) + q(0|0,0)\ V(2,0,0,2,1,0),\\
V_P(2,0,0,2,0,0;2) &= \frac{1}{2}\times 0.496 + \frac{1}{2}\times 0.744 = 0.620,\\
V_E(2,0,0,2,0,0;2) &= \frac{1}{2}\times(-0.097) + \frac{1}{2}\times(-0.048) = -0.073,
\end{aligned}$$

where i in the subscript of the predictive probability mass function of $X_{i,t}$ can be dropped due to symmetry, so

$$\begin{aligned}
&a_P^*(2,0,0,2,0,0) = a_E^*(2,0,0,2,0,0) \in \{1,2,3\},\\
&V_P(2,0,0,2,0,0) = 0.620, \quad V_E(2,0,0,2,0,0) = -0.072;
\end{aligned}$$

$$\begin{aligned}
V(1,0,0,2,0,0;1) &= q(1|1,2)\ V(2,0,0,3,0,0) + q(0|1,2)\ V(1,0,0,3,0,0),\\
V_P(1,0,0,2,0,0;1) &= \frac{1}{2}\times 0.372 + \frac{1}{2}\times 0.404 = 0.388,\\
V_E(1,0,0,2,0,0;1) &= \frac{1}{2}\times(-0.190) + \frac{1}{2}\times(-0.248) = -0.219,
\end{aligned}$$

$$\begin{aligned}
V(1,0,0,2,0,0;2) &= q(1|0,0)\ V(1,1,0,2,1,0) + q(0|0,0)\ V(1,0,0,2,1,0),\\
V_P(1,0,0,2,0,0;2) &= \frac{1}{2}\times 0.595 + \frac{1}{2}\times 0.418 = 0.506,\\
V_E(1,0,0,2,0,0;2) &= \frac{1}{2}\times(-0.107) + \frac{1}{2}\times(-0.206) = -0.156,
\end{aligned}$$

so

$$\begin{aligned}
&a_P^*(1,0,0,2,0,0) = a_E^*(1,0,0,2,0,0) \in \{2,3\},\\
&V_P(1,0,0,2,0,0) = 0.506, \quad V_E(1,0,0,2,0,0) = -0.156;
\end{aligned}$$

$$\begin{aligned}
V(0,0,0,2,0,0;1) &= q(1|0,2)\ V(1,0,0,3,0,0) + q(0|0,2)\ V(0,0,0,3,0,0),\\
V_P(0,0,0,2,0,0;1) &= \frac{1}{6}\times 0.404 + \frac{5}{6}\times 0.471 = 0.460,\\
V_E(0,0,0,2,0,0;1) &= \frac{1}{6}\times(-0.248) + \frac{5}{6}\times(-0.211) = -0.217,
\end{aligned}$$

$$
\begin{aligned}
&V(0,0,0,2,0,0;2) = q(1|0,0)\ V(0,1,0,2,1,0) + q(0|0,0)\ V(0,0,0,2,1,0),\\
&V_P(0,0,0,2,0,0;2) = \frac{1}{2} \times 0.685 + \frac{1}{2} \times 0.629 = 0.657,\\
&V_E(0,0,0,2,0,0;2) = \frac{1}{2} \times (-0.081) + \frac{1}{2} \times (-0.103) = -0.092,
\end{aligned}
$$

so

$$
\begin{aligned}
&a_P^*(0,0,0,2,0,0) = a_E^*(0,0,0,2,0,0) \in \{2,3\},\\
&V_P(0,0,0,2,0,0) = 0.657, \quad V_E(0,0,0,2,0,0) = -0.092;
\end{aligned}
$$

$$
\begin{aligned}
&V(1,1,0,1,1,0;1) = q(1|1,1)\ V(0,1,0,2,1,0) + q(0|1,1)\ V(0,0,0,2,1,0),\\
&V_P(1,1,0,1,1,0;1) = \frac{3}{4} \times 0.496 + \frac{1}{2} \times 0.595 = 0.521,\\
&V_E(1,1,0,1,1,0;1) = \frac{3}{4} \times (-0.097) + \frac{1}{4} \times (-0.107) = -0.099,
\end{aligned}
$$

$$
\begin{aligned}
&V(1,1,0,1,1,0;3) = q(1|0,0)\ V(1,1,1,1,1,1) + q(0|0,0)\ V(1,1,0,1,1,1),\\
&V_P(1,1,0,1,1,0;3) = \frac{1}{2} \times \frac{1}{3} + \frac{1}{2} \times 0.485 = 0.409,\\
&V_E(1,1,0,1,1,0;3) = \frac{1}{2} \times (-0.184) + \frac{1}{2} \times (-0.139) = -0.162,
\end{aligned}
$$

so

$$
\begin{aligned}
&a_P^*(1,1,0,1,1,0) = a_E^*(1,1,0,1,1,0) \in \{1,2\},\\
&V_P(1,1,0,1,1,0) = 0.521, \quad V_E(1,1,0,1,1,0) = -0.099;
\end{aligned}
$$

$$
\begin{aligned}
&V(1,0,0,1,1,0;1) = q(1|1,1)\ V(2,0,0,2,1,0) + q(0|1,1)\ V(1,0,0,2,1,0),\\
&V_P(1,0,0,1,1,0;1) = \frac{3}{4} \times 0.744 + \frac{1}{4} \times 0.418 = 0.662,\\
&V_E(1,0,0,1,1,0;1) = \frac{3}{4} \times (-0.048) + \frac{1}{2} \times (-0.206) = -0.088,
\end{aligned}
$$

$$
\begin{aligned}
&V(1,0,0,1,1,0;2) = q(1|0,1)\ V(1,1,0,1,2,0) + q(0|0,1)\ V(1,0,0,1,2,0),\\
&V_P(1,0,0,1,1,0;2) = \frac{1}{4} \times 0.595 + \frac{3}{4} \times 0.685 = 0.663,\\
&V_E(1,0,0,1,1,0;2) = \frac{1}{4} \times (-0.107) + \frac{3}{4} \times (-0.081) = -0.088,
\end{aligned}
$$

$$
\begin{aligned}
&V(1,0,0,1,1,0;3) = q(1|0,0)\ V(1,0,1,1,1,1) + q(0|0,0)\ V(1,0,0,1,1,1),\\
&V_P(1,0,0,1,1,0;3) = \frac{1}{2} \times 0.485 + \frac{1}{2} \times 0.840 = 0.663,\\
&V_E(1,0,0,1,1,0;1) = \frac{1}{2} \times (-0.139) + \frac{1}{2} \times (-0.036) = -0.088,
\end{aligned}
$$

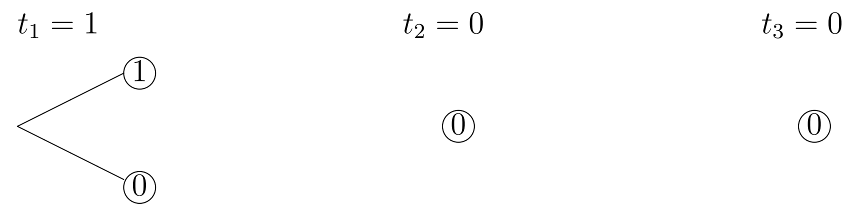

Figure 3: State space for $t = 1$.

so

$$a_P^*(1,0,0,1,1,0) = a_E^*(1,0,0,1,1,0) \in \{1,2,3\},$$
$$V_P(1,0,0,1,1,0) = 0.663, \quad V_E(1,0,0,1,1,0) = -0.088;$$

$$V(0,0,0,1,1,0;1) = q(1|0,1)\ V(1,0,0,2,1,0) + q(0|0,1)\ V(0,0,0,2,1,0),$$
$$V_P(0,0,0,1,1,0;1) = \frac{1}{4} \times 0.418 + \frac{3}{4} \times 0.629 = 0.576,$$
$$V_E(0,0,0,1,1,0;1) = \frac{1}{4} \times (-0.206) + \frac{1}{2} \times (-0.103) = -0.129,$$

$$V(0,0,0,1,1,0;3) = q(1|0,0)\ V(0,0,1,1,1,1) + q(0|0,0)\ V(0,0,0,1,1,1),$$
$$V_P(0,0,0,1,1,0;3) = \frac{1}{2} \times 0.840 + \frac{1}{2} \times \frac{1}{3} = 0.587,$$
$$V_E(0,0,0,1,1,0;1) = \frac{1}{2} \times (-0.139) + \frac{1}{2} \times (-0.036) = -0.129,$$

so

$$a_P^*(0,0,0,1,1,0) = 3, \quad a_E^*(0,0,0,1,1,0) \in \{1,2,3\},$$
$$V_P(0,0,0,1,1,0) = 0.587, \quad V_E(0,0,0,2,0,0) = -0.129\ .$$

Figure 3 shows some of the possible states after allocating one replication. Other states are equivalent to the states in the figure by a symmetry argument. The allocation decision and value function for each state can be calculated by backward induction as follows:

$$V(1,0,0,1,0,0;1) = q(1|1,1)\ V(2,0,0,2,0,0) + q(0|1,1)\ V(1,0,0,2,0,0),$$
$$V_P(1,0,0,1,0,0;1) = \frac{3}{4} \times 0.620 + \frac{1}{4} \times 0.506 = 0.592,$$
$$V_E(1,0,0,1,0,0;1) = \frac{3}{4} \times (-0.072) + \frac{1}{4} \times (-0.156) = -0.093,$$

$$V(1,0,0,1,0,0;1) = q(1|0,0)\ V(1,1,0,1,1,0) + q(0|0,0)\ V(1,0,0,1,1,0),$$
$$V_P(1,0,0,1,0,0;1) = \frac{1}{2} \times 0.521 + \frac{1}{2} \times 0.663 = 0.592,$$
$$V_E(1,0,0,1,0,0;1) = \frac{1}{2} \times (-0.099) + \frac{1}{2} \times (-0.088) = -0.093,$$

so

$$\begin{aligned} &a_P^*(1,0,0,1,0,0) = a_E^*(1,0,0,1,0,0) \in \{1,2,3\}, \\ &V_P(1,0,0,1,0,0) = 0.592, \quad V_E(1,0,0,1,0,0) = -0.093; \end{aligned}$$

$$\begin{aligned} V(0,0,0,1,0,0;1) &= q(1|0,1)\ V(1,0,0,2,0,0) + q(0|0,1)\ V(0,0,0,2,0,0), \\ V_P(0,0,0,1,0,0;1) &= \frac{1}{4} \times 0.506 + \frac{3}{4} \times 0.657 = 0.619, \\ V_E(0,0,0,1,0,0;1) &= \frac{1}{4} \times (-0.156) + \frac{3}{4} \times (-0.092) = -0.108, \end{aligned}$$

$$\begin{aligned} V(0,0,0,1,0,0;2) &= q(1|0,0)\ V(0,1,0,1,1,0) + q(0|0,0)\ V(0,0,0,1,1,0), \\ V_P(0,0,0,1,0,0;2) &= \frac{1}{2} \times 0.663 + \frac{1}{2} \times 0.587 = 0.625, \\ V_E(0,0,0,1,0,0;2) &= \frac{1}{2} \times (-0.088) + \frac{1}{2} \times (-0.129) = -0.108, \end{aligned}$$

so

$$\begin{aligned} &a_P^*(0,0,0,1,0,0) \in \{2,3\}, \quad a_E^*(0,0,0,1,0,0) \in \{1,2,3\}, \\ &V_P(0,0,0,1,0,0) = 0.625, \quad V_E(0,0,0,1,0,0) = -0.108\ . \end{aligned}$$

Therefore, the expected rewards for allocating three replications following the optimal allocation policy are

$$V_P(0,0,0,0,0,0) = 0.609, \quad V_E(0,0,0,0,0,0) = -0.101,$$

contrasted with the expected rewards for allocating three replications following equal allocation:

$$\widetilde{V}_P(0,0,0,0,0,0) = 0.498, \quad \widetilde{V}_E(0,0,0,0,0,0) = -0.145\ .$$

We can see that even for allocating only three replications, the difference between the optimal A&S policy and equal allocation is significant.

Numerical Results in the Presence of Correlation

In this example, we test the performance of AOAP, OCBA, KG, and EA in the setting as Example 1 in the main body of the paper except that we assume the sampling distribution between different alternatives have correlations 0.5. In Figure 4, we can see the performance comparison between different sampling procedure is similar to that in Example 1 of the main body of the paper.

Fitting Weights of Two-Factor Linear VFA in Example 2

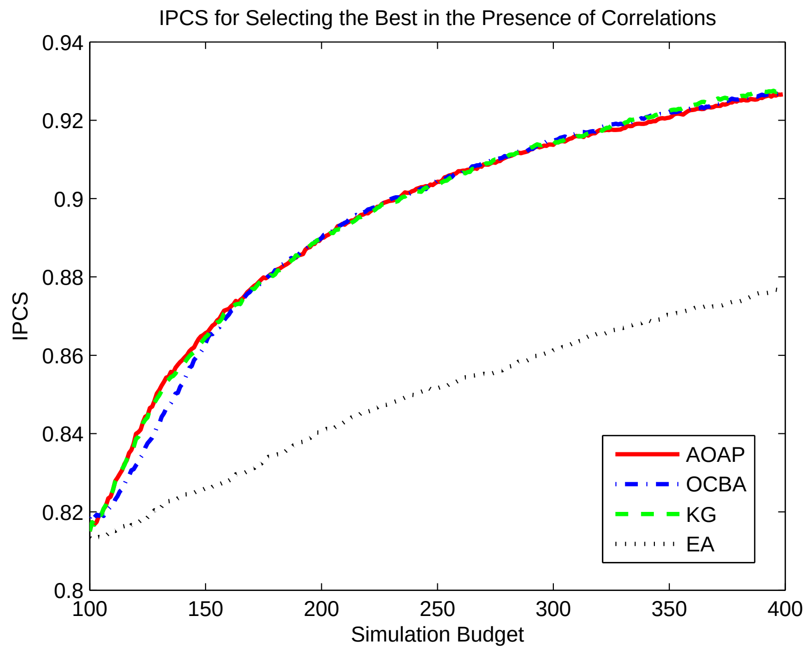

Figure 4: The prior distribution is the normal conjugate prior, with parameters $\mu_i^{(0)} = 0$, and $\sigma_i^{(0)} = 1$, $i = 1, \ldots, 10$. The true variances and correlations are $\sigma_i^2 = 1$, $\rho_{j,j'} = 0.5$, $i = 1, \ldots, 10$, $j, j' = 1, \ldots, 10$, $j \neq j'$. The number of initial replications is $n_0 = 10$ for each alternative. The IPCSs are estimated by 10^5 independent macro replications.

Figure 5 shows the trajectory of SA for fitting the weights of a two-factor Linear VFA in Example 2 of the main body of the paper.

Numerical Results for An AOAP with A Nonlinear VFA

Figure 6 shows the trajectory of G-MCL for a nonlinear VFA with activation function $K(z) = 1 - \exp(-z)$ in the same setting as the first scenario in Example 2. The initial point is chosen as $(w_1^{(0)}, w_2^{(0)}) = (1.8, 1)$, and the rest initialization of the algorithm is set the same as that in Example 2. The final fitting results are $w_1^* \approx 1.77$ and $w_2^* \approx 0.48$.

In Figure 7, we can see the performances of the AOAPs using the linear VFA (AOAP-L) and the nonlinear VFA (AOAP-NL) have comparable performances, and the former has a slight edge over the latter. Similar observations can be seen in Figure 8 that shows the numerical performances of AOAP-L and AOAP-NL in the second scenario of Example 2.

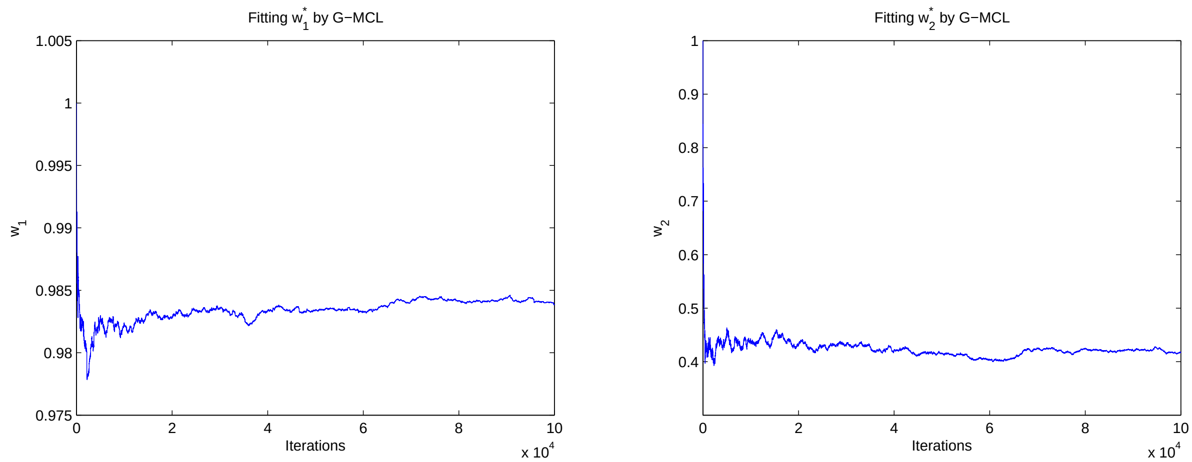

Figure 5: Fitting parameters of a two-factors linear VFA by G-MCL.

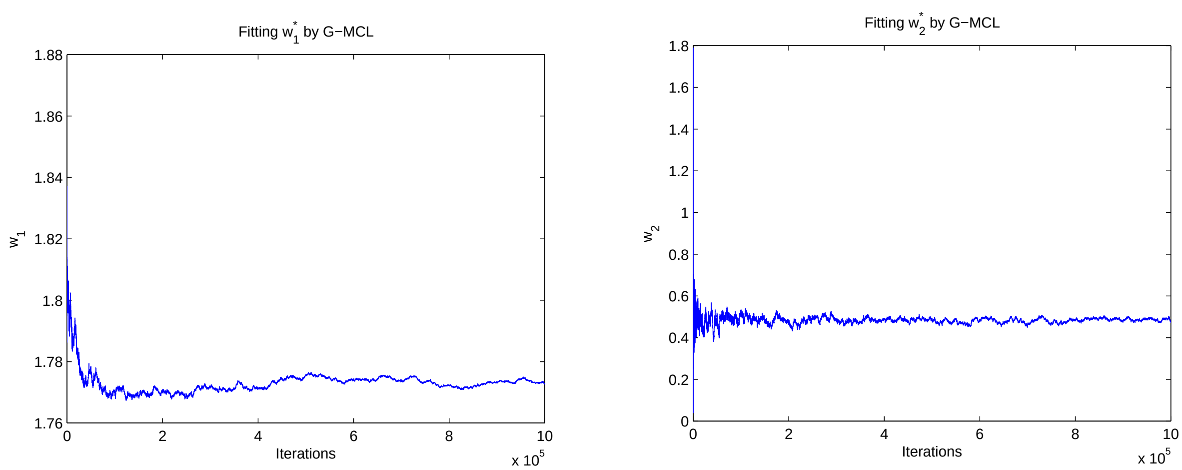

Figure 6: Fitting parameters of a two-factor nonlinear VFA by G-MCL.

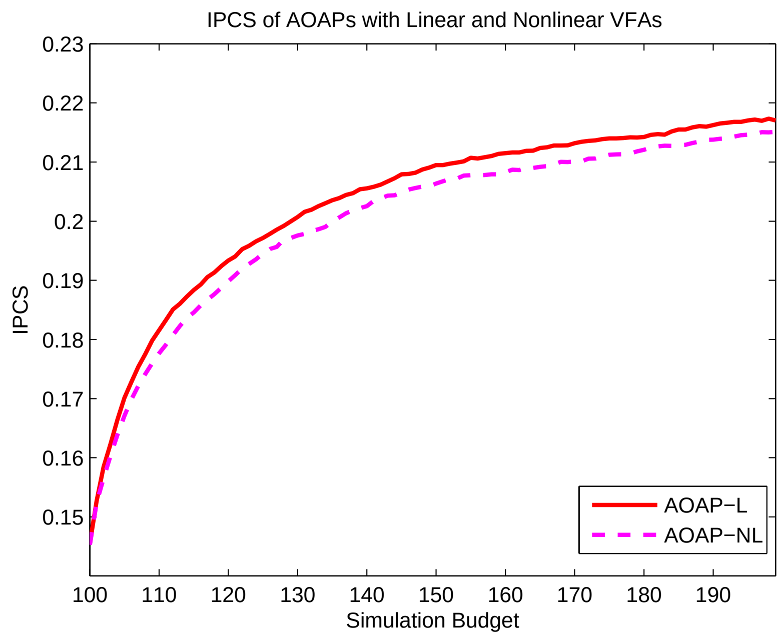

Figure 7: The prior distribution is the normal conjugate prior, with parameters $\mu_i^{(0)} = 0$, $i = 1, \dots, 10$, $\sigma_1^{(0)} = 0.02$, and $\sigma_i^{(0)} = 0.01$, $i = 2, \dots, 10$. The true variances are $\sigma_i = 1$, $i = 1, \dots, 10$. The number of initial replications is $n_0 = 10$ for each alternative. IPCSs are estimated by 10^5 independent macro replications.

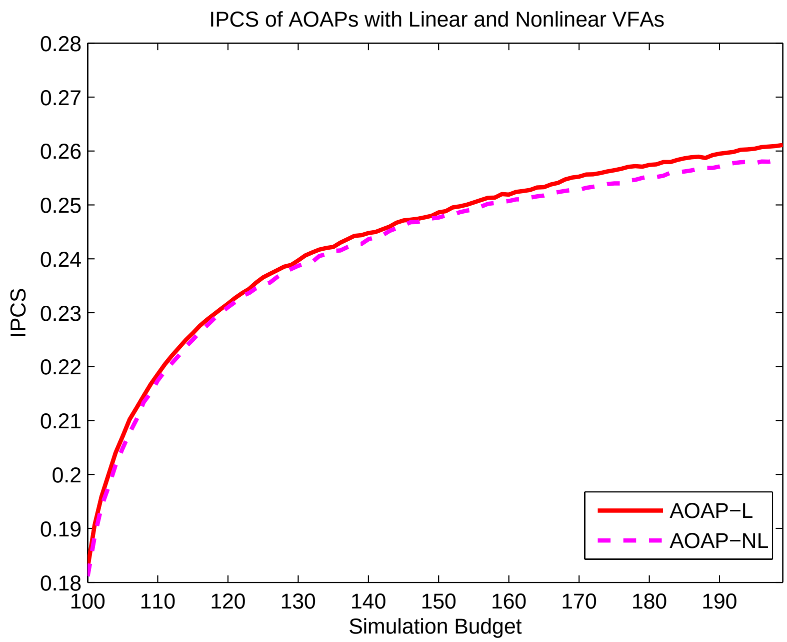

Figure 8: The prior distribution is the normal conjugate prior, with parameters $\mu_i^{(0)} = 0$, $i = 1, \ldots, 10$, $\sigma_1^{(0)} = 0.08$, and $\sigma_i^{(0)} = 0.04$, $i = 2, \ldots, 10$. The true variances are $\sigma_i = 1$, $i = 1, \ldots, 10$. The number of initial replications is $n_0 = 10$ for each alternative. IPCSs are estimated by 10^5 independent macro replications.