Online Supplement for “Bayesian Closed-loop Robust Process Design Considering Model Uncertainty and Data Quality” by “Linhan Ouyang, Jianxiong Chen, Yizhong Ma, Chanseok Park, Jionghua (Judy) Jin”

1. Computational cost analysis in Section 5.1

The computational times are presented in Table 1.

Table 1. Summary statistics of the average time (unit: second) by different online design approaches

$m_0 = 50$			$m_0 = 100$		
	L_d	L_p		L_d	L_p
Min	0.405	0.466	Min	0.449	0.409
Median	0.423	0.506	Median	0.553	0.471
Mean	0.421	0.502	Mean	0.545	0.468
Max	0.442	0.665	Max	0.858	0.648

An interesting phenomenon can be found that the computational cost of the proposed online approach L_p is smaller than that of the existing online approach L_d under $m_0 = 100$. It is because the model parameters are updated at each run and the computation is mainly used for estimating model parameters. There is no doubt that the computational cost of the online approach L_d is smaller than that of the online approach L_p when most of the online data are used in the updating process (e.g., $m_0 = 50$), since examining the data quality for every incoming response data also needs extra time. The p value for the case $m_0 = 50$ is 0.189, which indicate that there is no significant difference between the two online approaches. Furthermore, the p value for the case $m_0 = 100$ is 0.000, which means that the proposed online approach L_p at $m_0 = 100$ does not necessarily increase the computational cost since its computational cost is smaller than that of the online approach L_d.

2. Detailed proofs for some problems

Derivation of Eq. (7): the meaning of Eq. (7) is posterior distribution of the model parameters. By adding and subtracting $\mathbf{X}_t\hat{\boldsymbol{\beta}}_t$ in each parenthesis of Eq. (6), it yields

$$\begin{aligned}
&[(\mathbf{y}_t - \mathbf{X}_t\hat{\boldsymbol{\beta}}_t)^T(\mathbf{y}_t - \mathbf{X}_t\hat{\boldsymbol{\beta}}_t) + (\boldsymbol{\beta}_t - \hat{\boldsymbol{\beta}}_t)^T \mathbf{X}_t^T\mathbf{X}_t(\boldsymbol{\beta}_t - \hat{\boldsymbol{\beta}}_t)]^{-n_0/2} \\
&= [(n_0 - r - 1)s^2 + (\boldsymbol{\beta}_t - \hat{\boldsymbol{\beta}}_t)^T \mathbf{X}_t^T\mathbf{X}_t(\boldsymbol{\beta}_t - \hat{\boldsymbol{\beta}}_t)]^{-n_0/2} \\
&= \left[1 + \frac{(\boldsymbol{\beta}_t - \hat{\boldsymbol{\beta}}_t)^T \mathbf{X}_t^T\mathbf{X}_t(\boldsymbol{\beta}_t - \hat{\boldsymbol{\beta}}_t)}{(n_0 - r - 1)s^2}\right]^{-n_0/2} \left[(n_0 - r - 1)s^2\right]^{n_0/2} \\
&\propto \left[1 + \frac{(\boldsymbol{\beta}_t - \hat{\boldsymbol{\beta}}_t)^T \mathbf{X}_t^T\mathbf{X}_t(\boldsymbol{\beta}_t - \hat{\boldsymbol{\beta}}_t)}{(n_0 - r - 1)s^2}\right]^{-n_0/2} \left[s^2\left|\mathbf{X}_t^T\mathbf{X}_t\right|^{-1}\right]^{-1/2}
\end{aligned} \tag{A.1}$$

where the estimated variance s^2 and the estimated model parameters $\hat{\boldsymbol{\beta}}_t$ can be calculated by $(\mathbf{y}_t - \mathbf{X}_t\hat{\boldsymbol{\beta}}_t)^T(\mathbf{y}_t - \mathbf{X}_t\hat{\boldsymbol{\beta}}_t)/(n_0 - r - 1)$ and $\hat{\boldsymbol{\beta}}_t = (\mathbf{X}_t^T\mathbf{X}_t)^{-1}\mathbf{X}_t^T\mathbf{y}_t$, respectively. It can be seen from Eq. (7) that $\left[1 + \frac{(\boldsymbol{\beta}_t - \hat{\boldsymbol{\beta}}_t)^T \mathbf{X}_t^T\mathbf{X}_t(\boldsymbol{\beta}_t - \hat{\boldsymbol{\beta}}_t)}{(n_0 - r - 1)s^2}\right]^{-[n_0-(r+1)+(r+1)]/2} \left[s^2\left|\mathbf{X}_t^T\mathbf{X}_t\right|^{-1}\right]^{-1/2}$ has the form of a multivariate student t distribution with degrees of freedom $n_0 - r - 1$. Then, Eq. (7) can be rewritten as $\boldsymbol{\beta}_t \big| \mathbf{y}_t, \mathbf{X}_t \sim t_{n_0-r-1}\left[\hat{\boldsymbol{\beta}}_t,\, s^2(\mathbf{X}_t^T\mathbf{X}_t)^{-1}\right]$.

***Derivation of* Theorem 1**: We can calculate the following expression

$$\begin{aligned}
&(\mathbf{B} + \mathbf{C}\mathbf{D}^T)[\mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{C}(1 + \mathbf{D}^T\mathbf{B}^{-1}\mathbf{C})^{-1}\mathbf{D}^T\mathbf{B}^{-1}] \\
&= \mathbf{E} - (\mathbf{C} + \mathbf{C}\mathbf{D}^T\mathbf{B}^{-1}\mathbf{C})(1 + \mathbf{D}^T\mathbf{B}^{-1}\mathbf{C})^{-1}\mathbf{D}^T\mathbf{B}^{-1} + \mathbf{C}\mathbf{D}^T\mathbf{B}^{-1} \\
&= \mathbf{E}
\end{aligned}$$

where $\mathbf{E}$ is the identity matrix. Based on the invertible matrix theorem, the inverse of the matrix $(\mathbf{B} + \mathbf{C}\mathbf{D}^T)$ can be given as $\mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{C}(1 + \mathbf{D}^T\mathbf{B}^{-1}\mathbf{C})^{-1}\mathbf{D}^T\mathbf{B}^{-1}$.

***Derivation of* Lemma 1**: Since $\mathbf{P}_{t+1} = (\mathbf{X}_{t+1}^T\mathbf{X}_{t+1})^{-1}$, the updated model parameters can be calculated by $\hat{\boldsymbol{\beta}}_{t+1|t} = (\mathbf{X}_{t+1}^T\mathbf{X}_{t+1})^{-1}\mathbf{X}_{t+1}^T\mathbf{y}_{t+1}$ based on ordinary least square method. Let $\mathbf{X}_{t+1} = \begin{bmatrix}\mathbf{X}_t^T & \mathbf{x}_t^*\end{bmatrix}^T$,

$\mathbf{y}_{t+1} = \begin{bmatrix} \mathbf{y}_t^T & y_{t+1} \end{bmatrix}^T$, $\mathbf{x}_t^*$ is the optimal input settings at step t, and y_{t+1} is the new measurement of the response at step (t+1). The updated model parameters can be written as $\hat{\boldsymbol{\beta}}_{t+1|t} = [\mathbf{X}_t^T \mathbf{X}_t + \mathbf{x}_t^* \mathbf{x}_t^{*T}]^{-1} [\mathbf{X}_t^T \mathbf{y}_t + \mathbf{x}_t^* y_{t+1}]$.

Let $\mathbf{P}_t = (\mathbf{X}_t^T \mathbf{X}_t)^{-1}$, $\mathbf{P}_{t+1}$ can be obtained as $\mathbf{P}_t - \mathbf{P}_t \mathbf{x}_t^* k^{-1} \mathbf{x}_t^{*T} \mathbf{P}_t$ based on Theorem 1. The calculation process for $\hat{\boldsymbol{\beta}}_{t+1|t}$ can be rewritten as

$$
\begin{aligned}
\hat{\boldsymbol{\beta}}_{t+1|t} &= \mathbf{P}_{t+1} [\mathbf{X}_t^T \mathbf{y}_t + \mathbf{x}_t^* y_{t+1}] \\
&= (\mathbf{P}_t - \mathbf{P}_t \mathbf{x}_t^* k^{-1} \mathbf{x}_t^{*T} \mathbf{P}_t) [\mathbf{X}_t^T \mathbf{y}_t + \mathbf{x}_t^* y_{t+1}] \\
&= \hat{\boldsymbol{\beta}}_t + \mathbf{P}_t \mathbf{x}_t^* y_{t+1} - \mathbf{P}_t \mathbf{x}_t^* k^{-1} \mathbf{x}_t^{*T} \hat{\boldsymbol{\beta}}_t - \mathbf{P}_t \mathbf{x}_t^* k^{-1} (k-1) y_{t+1} \\
&= \hat{\boldsymbol{\beta}}_t + \mathbf{P}_t \mathbf{x}_t^* k^{-1} [y_{t+1} - \mathbf{x}_t^{*T} \hat{\boldsymbol{\beta}}_t]
\end{aligned}
$$