# D-OPTIMAL DESIGNS FOR ORDINAL RESPONSE EXPERIMENTS

Huiping Dang[1], Jun Yu[2], Fasheng Sun[1]

*KLAS and School of Mathematics and Statistics, Northeast Normal University*[1],

*School of Mathematics and Statistics, Beijing Institute of Technology*[2]

**Supplementary Material**

## S1 Proofs

### S1.1 Proof of Theorem 1

In order to prove Theorem 1, two lemmas are needed. We first give the representations of $\Delta_k(\boldsymbol{x})$ needed to prove the lemma. For convenience, we will denote $g_t(\boldsymbol{x}), \Delta_J(\boldsymbol{x})$ as $g_t, \Delta_J$ and omit $\boldsymbol{x}$ from all notations.

For $\Delta_k$, denote $A = \prod_{t=1}^{k-2}(1-g_t) + \Delta_{k-2}g_{k-2}g_{k-1}$ and $B = \prod_{t=1}^{k-1}(1-g_t) + \prod_{t=1}^{k-3}(1-g_t)g_{k-1} + \left[\Delta_{k-2} - \prod_{t=1}^{k-3}(1-g_t)\right]g_{k-2}g_{k-1}$. When $l > k$, $\prod_{t=l}^{k}(1-g_t) = 1$ and $\prod_{t=l}^{k} g_t = 1$. Then

$$A = \prod_{t=1}^{k-2}(1-g_t) + \Delta_{k-2}g_{k-2}g_{k-1}$$

$$= \sum_{l=1}^{\lceil \frac{k}{2} \rceil} \left( \prod_{t=1}^{k-2l} (1-g_t) \prod_{t=k-2(l-1)}^{k-1} g_t \right)$$

$$= \begin{cases} \prod_{t=1}^{k-2}(1-g_t) + \prod_{t=1}^{k-4}(1-g_t) \prod_{t=k-2}^{k-1} g_t + \cdots + \prod_{t=1}^{k-1} g_t & \text{k is odd,} \\ \prod_{t=1}^{k-2}(1-g_t) + \prod_{t=1}^{k-4}(1-g_t) \prod_{t=k-2}^{k-1} g_t + \cdots + \prod_{t=2}^{k-1} g_t & \text{k is even.} \end{cases}$$

$$B = \prod_{t=1}^{k-1}(1-g_t) + \prod_{t=1}^{k-3}(1-g_t) g_{k-1} + \left[ \Delta_{k-2} - \prod_{t=1}^{k-3}(1-g_t) \right] g_{k-2} g_{k-1}$$

$$= \sum_{l=1}^{\lceil \frac{k+1}{2} \rceil} \left( \prod_{t=1}^{k+1-2l} (1-g_t) \prod_{t=k+1-2(l-1)}^{k-1} g_t \right)$$

$$= \begin{cases} \prod_{t=1}^{k-1}(1-g_t) + \prod_{t=1}^{k-3}(1-g_t) g_{k-1} + \cdots + \prod_{t=2}^{k-1} g_t & \text{k is odd,} \\ \prod_{t=1}^{k-1}(1-g_t) + \prod_{t=1}^{k-3}(1-g_t) g_{k-1} + \cdots + \prod_{t=1}^{k-1} g_t & \text{k is even.} \end{cases}$$

Let $\boldsymbol{\gamma} \in \mathbb{R}^p$ denote the parameter vector, i.e., $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_p)^\top = (\theta_1, \ldots, \theta_{J-1}, \boldsymbol{\beta}^\top)^\top$.

For an exact design

$$\xi_{\text{exact}} = \begin{pmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_m \\ n_1 & \cdots & n_m \end{pmatrix},$$

the corresponding information matrix is given in the following lemma.

**Lemma S1.** *Suppose Assumptions 1 and 2 hold, the Fisher information matrix for Model (2.2) under the exact design $\xi_{exact}$ can be written as*

$$M(\xi_{exact}) = \sum_{i=1}^{m} n_i M(\boldsymbol{x}_i),$$

*where* $M(\boldsymbol{x}_i) = (m_{i_{st}})_{1\le s,t\le p}$ *is a* $p \times p$ *matrix with*

$$m_{i_{st}} = \sum_{j=1}^{J} \frac{1}{\pi_j(\boldsymbol{x}_i)} \frac{\partial \pi_j(\boldsymbol{x}_i)}{\partial \gamma_s} \frac{\partial \pi_j(\boldsymbol{x}_i)}{\partial \gamma_t}.$$

Lemma S1 comes from supplementary material of Ai et al. (2023).

**Remark S1.** From Lemma S1, the Fisher information matrix for Model (2.2) under an approximate design

$$\xi = \begin{pmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_m \\ \\ \omega_1 & \cdots & \omega_m \end{pmatrix}$$

can be written as

$$M(\xi) = \sum_{i=1}^{m} \omega_i M(\boldsymbol{x}_i).$$

Let $\partial \boldsymbol{\pi}(\boldsymbol{x})/\partial \boldsymbol{\gamma}^\top$ denote a $J \times p$ matrix, whose $(l, j)$th entry is $\partial \pi_l(\boldsymbol{x})/\partial \gamma_j$, where $\boldsymbol{x} \in \mathcal{X}$ is a design point. We have the following lemma.

**Lemma S2.** *For Model (2.2),*

$$\frac{\partial \boldsymbol{\pi}(\boldsymbol{x})}{\partial \boldsymbol{\gamma}^\top} = U(\boldsymbol{x}) H(\boldsymbol{x}), \tag{S1}$$

*where* $U(\boldsymbol{x})$ *and* $H(\boldsymbol{x})$ *are defined in Section (2.2).*

*Proof of Lemma S2.* Introduce the following two functions needed for the proof.

$$f_{t-c} = \begin{cases} -\frac{1}{1-g_t} & t - c < 0, \\ & \qquad\qquad c \ \text{ is a constant.} \\ \frac{1}{g_t} & t - c \ge 0, \end{cases}$$

$$y_{t-c} = \begin{cases} -\frac{1}{1-g_t} & t-c<0, \\ 0 & t-c=0, \quad c \text{ is a constant.} \\ \frac{1}{g_t} & t-c>0, \end{cases}$$

And denote $I(j < J-2(l-1))$, $I(j \geq J-2(l-1))$ as $I_{j<J-2(l-1)}$, $I_{j\geq J-2(l-1)}$.

The rest is similar.

We first prove that the following equation

$$\frac{\partial \pi_1}{\partial \theta_j} = u_{1j}, \tag{S2}$$

holds for $j = 1, \ldots, J-1$.

From Section 2.2, we have

$$\Delta_J = \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right), \quad \pi_1 = \frac{\prod_{t=1}^{J-1} g_t}{\Delta_J}.$$

Then

$$\begin{aligned} \frac{\partial \pi_1}{\partial \theta_j} &= \left[ \left( \frac{\partial}{\partial \theta_j} \prod_{t=1}^{J-1} g_t \right) \Delta_J - \frac{\partial \Delta_J}{\partial \theta_j} \prod_{t=1}^{J-1} g_t \right] \Delta_J^{-2} \\ &= \left[ \left( \frac{\partial g_j}{\partial \theta_j} f_{j-1} \prod_{t=1}^{J-1} g_t \right) \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) \right. \\ &\quad \left. - \frac{\partial g_j}{\partial \theta_j} \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( y_{j-(J-2l+1)} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) \prod_{t=1}^{J-1} g_t \right] \Delta_J^{-2} \\ &= \left[ g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) f_{j-1} \prod_{t=1}^{J-1} g_t \right. \end{aligned}$$

$$
\begin{aligned}
&-g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( y_{j-(J-2l+1)} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) g_j f_{j-1} \prod_{t=1}^{J-1} g_t \Bigg] \Delta_J^{-2} \\
= &\Bigg[ g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( I_{j<J-2(l-1)} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) f_{j-1} \prod_{t=1}^{J-1} g_t \\
&+ g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( I_{j\geq J-2(l-1)} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) f_{j-1} \prod_{t=1}^{J-1} g_t \\
&- g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( I_{j<J-2(l-1)} y_{j-(J-2l+1)} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) g_j f_{j-1} \prod_{t=1}^{J-1} g_t \\
&-g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( I_{j\geq J-2(l-1)} y_{j-(J-2l+1)} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) g_j f_{j-1} \prod_{t=1}^{J-1} g_t \Bigg] \Delta_J^{-2} \\
= &\Bigg[ g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( I_{j<J-2(l-1)} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) f_{j-1} \prod_{t=1}^{J-1} g_t \\
&-g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( I_{j<J-2(l-1)} y_{j-(J-2l+1)} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) g_j f_{j-1} \prod_{t=1}^{J-1} g_t \Bigg] \Delta_J^{-2} \\
= &\Bigg[ g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( I_{j\leq J-2l} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t + I_{j=J-2l+1} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) \\
&\times f_{j-1} \prod_{t=1}^{J-1} g_t - g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( I_{j\leq J-2l} y_{j-(J-2l+1)} g_j \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) f_{j-1} \prod_{t=1}^{J-1} g_t \Bigg] \Delta_J^{-2} \\
= &\Bigg[ g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \Bigg( -I_{j\leq J-2l} y_{j-(J-2l+1)} (1-g_j) \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \\
&+ I_{j=J-2l+1} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \Bigg) f_{j-1} \prod_{t=1}^{J-1} g_t \\
&-g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( I_{j\leq J-2l} y_{j-(J-2l+1)} g_j \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) f_{j-1} \prod_{t=1}^{J-1} g_t \Bigg] \Delta_J^{-2}
\end{aligned}
$$

$$
\begin{aligned}
&= \left[ g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( -I_{j \le J-2l} y_{j-(J-2l+1)} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right.\right. \\
&\quad \left.\left. + \ I_{j=J-2l+1} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) f_{j-1} \prod_{t=1}^{J-1} g_t \right] \Delta_J^{-2} \\
&= \left[ g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( I_{j<J-2(l-1)} \prod_{t=1,t\neq j}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) \prod_{t=1,t\neq j}^{J-1} g_t \right] \Delta_J^{-2} \\
&= \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( I_{j<J-2(l-1)} \prod_{t=1,t\neq j}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) g_j' \prod_{t=1,t\neq j}^{J-1} g_t \Delta_J^{-2} \\
&= u_{1j},
\end{aligned}
$$

which implies Equation (S2) holds for $j = 1, \ldots, J-1$.

Secondly, we prove that the following equation

$$
\frac{\partial \pi_{j+1}}{\partial \theta_j} = u_{j+1,j}, \tag{S3}
$$

holds for $j = 1, \ldots, J-1$.

From Section 2.2, we have

$$
\pi_{j+1} = \frac{\prod_{t=1}^{j} (1-g_t) \prod_{t=j+1}^{J-1} g_t}{\Delta_J}.
$$

Then

$$
\begin{aligned}
\frac{\partial \pi_{j+1}}{\partial \theta_j} &= \left[ \frac{\partial}{\partial \theta_j} \left( \prod_{t=1}^{j} (1-g_t) \prod_{t=j+1}^{J-1} g_t \right) \Delta_J - \frac{\partial \Delta_J}{\partial \theta_j} \prod_{t=1}^{j} (1-g_t) \prod_{t=j+1}^{J-1} g_t \right] \Delta_J^{-2} \\
&= \left[ -\frac{\partial g_j}{\partial \theta_j} \prod_{t=1}^{j-1} (1-g_t) \prod_{t=j+1}^{J-1} g_t \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) \right.
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{\partial g_j}{\partial \theta_j}\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(y_{j-(J-2l+1)}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1}g_t\right)(1-g_j)\prod_{t=1}^{j-1}(1-g_t)\prod_{t=j+1}^{J-1}g_t\Bigg]\Delta_J^{-2}\\
=&-\Bigg[g_j'\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(I_{j\le J-2l}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1}g_t\right)\prod_{t=1}^{j-1}(1-g_t)\prod_{t=j+1}^{J-1}g_t\\
&+g_j'\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(I_{j>J-2l}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1}g_t\right)\prod_{t=1}^{j-1}(1-g_t)\prod_{t=j+1}^{J-1}g_t\\
&+g_j'\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(y_{j-(J-2l+1)}I_{j\le J-2l}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1}g_t\right)(1-g_j)\prod_{t=1}^{j-1}(1-g_t)\prod_{t=j+1}^{J-1}g_t\\
&+g_j'\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(y_{j-(J-2l+1)}I_{j>J-2l}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1}g_t\right)(1-g_j)\prod_{t=1}^{j-1}(1-g_t)\prod_{t=j+1}^{J-1}g_t\Bigg]\Delta_J^{-2}\\
=&-\Bigg[g_j'\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(I_{j>J-2l}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1}g_t\right)\prod_{t=1}^{j-1}(1-g_t)\prod_{t=j+1}^{J-1}g_t\\
&+g_j'\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(y_{j-(J-2l+1)}I_{j>J-2l}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1}g_t\right)(1-g_j)\prod_{t=1}^{j-1}(1-g_t)\prod_{t=j+1}^{J-1}g_t\Bigg]\Delta_J^{-2}\\
=&-\Bigg[g_j'\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(I_{j=J-2l+1}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1}g_t\right)\prod_{t=1}^{j-1}(1-g_t)\prod_{t=j+1}^{J-1}g_t\\
&+g_j'\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(I_{j\ge J-2(l-1)}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1}g_t\right)\prod_{t=1}^{j-1}(1-g_t)\prod_{t=j+1}^{J-1}g_t\\
&+g_j'\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(y_{j-(J-2l+1)}I_{j\ge J-2(l-1)}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1}g_t\right)\\
&\times(1-g_j)\prod_{t=1}^{j-1}(1-g_t)\prod_{t=j+1}^{J-1}g_t\Bigg]\Delta_J^{-2}\\
=&-\Bigg[g_j'\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(I_{j=J-2l+1}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1}g_t\right)\prod_{t=1}^{j-1}(1-g_t)\prod_{t=j+1}^{J-1}g_t
\end{aligned}
$$

$$+g_j' \sum_{l=1}^{\lceil \frac{J}{2} \rceil} \left( y_{j-(J-2l+1)} I_{j \geq J-2(l-1)} \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{J-1} g_t \right) \prod_{t=1}^{j-1} (1-g_t) \prod_{t=j+1}^{J-1} g_t \Bigg] \Delta_J^{-2}$$

$$= - \begin{cases} \left[ g_j' \prod_{t=1}^{j-1}(1-g_t) \prod_{t=j+1}^{J-1} g_t \sum_{l=\frac{J-j+1}{2}}^{\lceil \frac{J}{2} \rceil} \left( \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1), t \neq j}^{J-1} g_t \right) \right] \Delta_J^{-2}, \\ J-j \text{ is odd} \\ \left[ g_j' \prod_{t=1}^{j-1}(1-g_t) \prod_{t=j+1}^{J-1} g_t \sum_{l=\frac{J-j+2}{2}}^{\lceil \frac{J}{2} \rceil} \left( \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1), t \neq j}^{J-1} g_t \right) \right] \Delta_J^{-2}, \\ J-j \text{ is even} \end{cases}$$

$$= - \begin{cases} \left[ g_j' \prod_{t=1}^{j-1}(1-g_t) \prod_{t=j+1}^{J-1} g_t \sum_{l=\frac{J-j+1}{2}}^{\lceil \frac{J}{2} \rceil} \left( \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{j-1} g_t \right) \right] \\ \times \prod_{t=j+1}^{J-1} g_t \Delta_J^{-2}, J-j \text{ is odd} \\ \left[ g_j' \prod_{t=1}^{j-1}(1-g_t) \prod_{t=j+1}^{J-1} g_t \sum_{l=\frac{J-j+2}{2}}^{\lceil \frac{J}{2} \rceil} \left( \prod_{t=1}^{J-2l} (1-g_t) \prod_{t=J-2(l-1)}^{j-1} g_t \right) \right] \\ \times \prod_{t=j+1}^{J-1} g_t \Delta_J^{-2}, J-j \text{ is even} \end{cases}$$

$$= - \begin{cases} \left[ g_j' \prod_{t=1}^{j-1}(1-g_t) \prod_{t=j+1}^{J-1} g_t \sum_{l=1}^{\lceil \frac{j+1}{2} \rceil} \left( \prod_{t=1}^{j+1-2l} (1-g_t) \prod_{t=j+1-2(l-1)}^{j-1} g_t \right) \right] \\ \times \prod_{t=j+1}^{J-1} g_t \Delta_J^{-2} \\ \left[ g_j' \prod_{t=1}^{j-1}(1-g_t) \prod_{t=j+1}^{J-1} g_t \sum_{l=1}^{\lceil \frac{j}{2} \rceil} \left( \prod_{t=1}^{j-2l} (1-g_t) \prod_{t=j-2(l-1)}^{j-1} g_t \right) \right] \\ \times \prod_{t=j+1}^{J-1} g_t \Delta_J^{-2} \end{cases}$$

$$= - \prod_{t=1}^{j-1} (1-g_t) \prod_{t=j+1}^{J-1} g_t \Delta_j g_j' \prod_{t=j+1}^{J-1} g_t \Delta_J^{-2}$$

$$= u_{j+1,j},$$

which implies Equation (S3) holds for $j = 1, \dots, J-1$.

Thirdly, we prove that the following equation

$$\frac{\partial \pi_l}{\partial \theta_j} = u_{lj} = u_{l-1,j}\frac{1-g_{l-1}}{g_{l-1}}, \tag{S4}$$

holds for $l = 2, \dots, J$, $j = 1, \dots, J-1$, $j \neq l-1$. For each $j$, we prove Equation (S4) holds for $l = 2, \dots, J$ by induction.

(1) When $l = 2$, $j \neq 1$, from Section 2.2,

$$\pi_2 = \frac{(1-g_1)\prod_{t=2}^{J-1} g_t}{\Delta_J}.$$

Based on the previously proven facts, we have

$$\begin{aligned}
\frac{\partial \pi_2}{\partial \theta_j} &= \left[\left(\frac{\partial}{\partial \theta_j}(1-g_1)\prod_{t=2}^{J-1} g_t\right)\Delta_J - \frac{\partial \Delta_J}{\partial \theta_j}(1-g_1)\prod_{t=2}^{J-1} g_t\right]\Delta_J^{-2} \\
&= \left[\left(\frac{\partial}{\partial \theta_j}\prod_{t=1}^{J-1} g_t\right)\Delta_J - \frac{\partial \Delta_J}{\partial \theta_j}\prod_{t=1}^{J-1} g_t\right]\Delta_J^{-2}\frac{1-g_1}{g_1} \\
&= u_{1,j}\frac{1-g_1}{g_1} \\
&= u_{2,j},
\end{aligned}$$

which implies Equation (S4) holds for $l = 2$.

(2) Suppose Equation (S4) holds for $3, \dots, l-1$ ($l < J$). By

$$\pi_l = \frac{\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t}{\Delta_J},$$

when $j < l-1$, we have

$$\begin{aligned}
\frac{\partial \pi_l}{\partial \theta_j} &= \left[\frac{\partial}{\partial \theta_j}\left(\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right)\Delta_J - \frac{\partial \Delta_J}{\partial \theta_j}\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right]\Delta_J^{-2} \\
&= \left[-\frac{\partial g_j}{\partial \theta_j}\prod_{t=1,t\neq j}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\Delta_J - \frac{\partial \Delta_J}{\partial \theta_j}\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right]\Delta_J^{-2} \\
&= \left[-\frac{\partial g_j}{\partial \theta_j}\prod_{t=1,t\neq j}^{l-2}(1-g_t)\prod_{t=l-1}^{J-1} g_t\frac{1-g_{l-1}}{g_{l-1}}\Delta_J - \frac{\partial \Delta_J}{\partial \theta_j}\prod_{t=1}^{l-2}(1-g_t)\prod_{t=l-1}^{J-1} g_t\frac{1-g_{l-1}}{g_{l-1}}\right]\Delta_J^{-2} \\
&= \left[\frac{\partial}{\partial \theta_j}\left(\prod_{t=1}^{l-2}(1-g_t)\prod_{t=l-1}^{J-1} g_t\right)\Delta_J - \frac{\partial \Delta_J}{\partial \theta_j}\prod_{t=1}^{l-2}(1-g_t)\prod_{t=l-1}^{J-1} g_t\right]\Delta_J^{-2}\frac{1-g_{l-1}}{g_{l-1}} \\
&= \frac{\partial \pi_{l-1}}{\partial \theta_j}\frac{1-g_{l-1}}{g_{l-1}} = u_{l-1,j}\frac{1-g_{l-1}}{g_{l-1}} = u_{l,j}.
\end{aligned}$$

When $j \geq l$, we have

$$\begin{aligned}
\frac{\partial \pi_l}{\partial \theta_j} &= \left[\frac{\partial}{\partial \theta_j}\left(\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right)\Delta_J - \frac{\partial \Delta_J}{\partial \theta_j}\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right]\Delta_J^{-2} \\
&= \left[\frac{\partial g_j}{\partial \theta_j}\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l,t\neq j}^{J-1} g_t\Delta_J - \frac{\partial \Delta_J}{\partial \theta_j}\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right]\Delta_J^{-2} \\
&= \left[\frac{\partial g_j}{\partial \theta_j}\prod_{t=1}^{l-2}(1-g_t)\prod_{t=l-1,t\neq j}^{J-1} g_t\frac{1-g_{l-1}}{g_{l-1}}\Delta_J - \frac{\partial \Delta_J}{\partial \theta_j}\prod_{t=1}^{l-2}(1-g_t)\prod_{t=l-1}^{J-1} g_t\frac{1-g_{l-1}}{g_{l-1}}\right]\Delta_J^{-2} \\
&= \left[\frac{\partial}{\partial \theta_j}\left(\prod_{t=1}^{l-2}(1-g_t)\prod_{t=l-1}^{J-1} g_t\right)\Delta_J - \frac{\partial \Delta_J}{\partial \theta_j}\prod_{t=1}^{l-2}(1-g_t)\prod_{t=l-1}^{J-1} g_t\right]\Delta_J^{-2}\frac{1-g_{l-1}}{g_{l-1}} \\
&= \frac{\partial \pi_{l-1}}{\partial \theta_j}\frac{1-g_{l-1}}{g_{l-1}} = u_{l-1,j}\frac{1-g_{l-1}}{g_{l-1}} = u_{l,j},
\end{aligned}$$

which implies Equation (S4) holds for $l = 2, \ldots, J$, $j = 1, \ldots, J-1$, $l \neq j+1$.

Furthermore, for the case $l = J$, using the fact $\pi_1 + \cdots + \pi_J = 1$ and the facts already proved above, we have

$$\frac{\partial \pi_J}{\partial \theta_j} = -\sum_{l=1}^{J-1}\left(\frac{\partial \pi_l}{\partial \theta_j}\right) = -\sum_{l=1}^{J-1} u_{l,j} = -\sum_{l=1}^{J-1} u_{l,j} = u_{J,j}.$$

Finally, we need to prove that the following equation

$$\frac{\partial \pi_l}{\partial \beta_s} = x_s \sum_{k=1}^{J-1} u_{l,k}, \tag{S5}$$

holds for $l = 1, \ldots, J$.

By

$$\pi_l = \frac{\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t}{\Delta_J},$$

and the facts already proved above, we have

$$\begin{aligned}
\frac{\partial \pi_l}{\partial \beta_s} =& \left[\frac{\partial}{\partial \beta_s}\left(\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right)\Delta_J\right.\\
&\left.-\frac{\partial}{\partial \beta_s}\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1} g_t\right)\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right]\Delta_J^{-2}\\
=& \left[\sum_{k=1}^{J-1}\frac{\partial g_k}{\partial \beta_s}\left(f_{k-l}\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right)\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1} g_t\right)\right.\\
&\left.-\sum_{l=1}^{\lceil \frac{J}{2}\rceil}\left(\sum_{k=1}^{J-1}\frac{\partial g_k}{\partial \beta_s} y_{k-(J-2l+1)}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1} g_t\right)\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right]\Delta_J^{-2}\\
=& \left[\sum_{k=1}^{J-1} g_k' x_s\left(f_{k-l}\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right)\Delta_J\right.\\
&\left.-\sum_{k=1}^{J-1}\left(\sum_{l=1}^{\lceil \frac{J}{2}\rceil} g_k' x_s y_{k-(J-2l+1)}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1} g_t\right)\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right]\Delta_J^{-2}\\
=& x_s\sum_{k=1}^{J-1}\left[g_k'\left(f_{k-l}\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right)\Delta_J\right.\\
&\left.-\left(\sum_{l=1}^{\lceil \frac{J}{2}\rceil} g_k' y_{k-(J-2l+1)}\prod_{t=1}^{J-2l}(1-g_t)\prod_{t=J-2(l-1)}^{J-1} g_t\right)\prod_{t=1}^{l-1}(1-g_t)\prod_{t=l}^{J-1} g_t\right]\Delta_J^{-2}
\end{aligned}$$

$$=x_s \sum_{k=1}^{J-1} u_{l,k},$$

which implies Equation (S5) holds for $l = 1, \ldots, J$.

Combining Equations (S2), (S3),(S4) and (S5), Lemma S2 is proved.

□

*Proof of Theorem 1.* Combining the results of Lemmas S1 and S2, we have

$$\begin{aligned} M(\xi) &= \sum_{i=1}^{m} \omega_i M(\boldsymbol{x}_i) \\ &= \sum_{i=1}^{m} \omega_i \left( \frac{\partial \boldsymbol{\pi}(\boldsymbol{x}_i)}{\partial \boldsymbol{\gamma}^\top} \right)^\top D^{-1}(\boldsymbol{x}_i) \left( \frac{\partial \boldsymbol{\pi}(\boldsymbol{x}_i)}{\partial \boldsymbol{\gamma}^\top} \right) \\ &= \sum_{i=1}^{m} \omega_i H^\top(\boldsymbol{x}_i) U^\top(\boldsymbol{x}_i) D^{-1}(\boldsymbol{x}_i) U(\boldsymbol{x}_i) H(\boldsymbol{x}_i), \end{aligned}$$

which completes the proof. □

## S1.2 Proof of Corollary 1

*Proof.* The main task is to identify the complete class. Let $c = \beta x$(where $\beta \neq 0$), then there is a bijection between $x$ and $c$, and $x = c/\beta$, denote the mapping space as $\mathcal{X}'$. For a complete class $\Xi^*$, define two designs $\xi \notin \Xi^*$ and $\tilde{\xi} \in \Xi^*$ on $\mathcal{X}'$,

$$\begin{aligned} \xi &= \left\{ (c_i, \omega_i), c_i \in \mathcal{X}', \sum_{i=1}^{m} \omega_i = 1 \right\}, \\ \tilde{\xi} &= \left\{ (\tilde{c}_i, \tilde{\omega}_i), \tilde{c}_i \in \mathcal{X}', \sum_{i=1}^{k} \tilde{\omega}_i = 1 \right\}. \end{aligned} \tag{S6}$$

For the case $J = 3$, Theorem 3.2 in Hao and Yang (2020) has proved that at most 4 support points form a complete class.

For the case $J = 4$, Model (2.2) is essentially

$$\log\left(\frac{\pi_{i,j}}{\pi_{i,j+1}}\right) = \alpha_j + \beta x_i = \alpha_j + c_i, j = 1, 2, 3.$$

For the support point $c$, let $a_j = e^{\alpha_j + c}$, $j = 1, 2, 3$. The information matrix at $c$ is

$$M(x) = \Lambda^{-2} \begin{pmatrix} a_1a_2a_3(a_2a_3 + a_3 + 1) & a_1a_2a_3(a_3 + 1) & a_1a_2a_3 & \frac{c}{\beta}a_1a_2a_3(a_2a_3 + 2a_3 + 3) \\ a_1a_2a_3(a_3 + 1) & a_1a_2a_3(a_3 + 1) + a_2a_3(a_3 + 1) & a_1a_2a_3 + a_2a_3 & \frac{c}{\beta}[a_1a_2a_3(2a_3 + 3) + a_2a_3(a_3 + 2)] \\ a_1a_2a_3 & a_1a_2a_3 + a_2a_3 & a_1a_2a_3 + a_2a_3 + a_3 & \frac{c}{\beta}(3a_1a_2a_3 + 2a_2a_3 + a_3) \\ \frac{c}{\beta}a_1a_2a_3(a_2a_3 + 2a_3 + 3) & \frac{c}{\beta}[a_1a_2a_3(2a_3 + 3) + a_2a_3(a_3 + 2)] & \frac{c}{\beta}(3a_1a_2a_3 + 2a_2a_3 + a_3) & (\frac{c}{\beta})^2[a_1a_2a_3(a_2a_3 + 4a_3 + 9) + a_2a_3(a_3 + 4) + a_3] \end{pmatrix},$$

where $\Lambda = a_1a_2a_3 + a_2a_3 + a_3 + 1$.

To prove the complete class result,

**Step 1: (Selection)** Among the first three columns, select the following set as maximal linear independent nonconstant functions:

$$\Psi_1(c) = a_1a_2a_3(a_2a_3 + a_3 + 1)\Lambda^{-2} = \frac{e^{\alpha_1+\alpha_2+\alpha_3+3c}\left(e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c} + 1\right)}{\left(e^{\alpha_1+\alpha_2+\alpha_3+3c} + e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c} + 1\right)^2},$$

$$\Psi_2(c) = a_1a_2a_3(a_3 + 1)\Lambda^{-2} = \frac{e^{\alpha_1+\alpha_2+\alpha_3+3c}\left(e^{\alpha_3+c} + 1\right)}{\left(e^{\alpha_1+\alpha_2+\alpha_3+3c} + e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c} + 1\right)^2},$$

$$\begin{aligned}\Psi_3(c) &= \left[a_1a_2a_3(a_3 + 1) + a_2a_3(a_3 + 1)\right]\Lambda^{-2} \\ &= \frac{e^{\alpha_1+\alpha_2+\alpha_3+3c}\left(e^{\alpha_3+c} + 1\right) + e^{\alpha_2+\alpha_3+2c}\left(e^{\alpha_3+c} + 1\right)}{\left(e^{\alpha_1+\alpha_2+\alpha_3+3c} + e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c} + 1\right)^2},\end{aligned}$$

$$\Psi_4(c) = a_1a_2a_3\Lambda^{-2} = \frac{e^{\alpha_1+\alpha_2+\alpha_3+3c}}{\left(e^{\alpha_1+\alpha_2+\alpha_3+3c} + e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c} + 1\right)^2},$$

$$\Psi_5(c) = (a_1a_2a_3 + a_2a_3)\Lambda^{-2} = \frac{e^{\alpha_1+\alpha_2+\alpha_3+3c} + e^{\alpha_2+\alpha_3+2c}}{\left(e^{\alpha_1+\alpha_2+\alpha_3+3c} + e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c} + 1\right)^2},$$

$$\Psi_6(c) = (a_1a_2a_3 + a_2a_3 + a_3)\Lambda^{-2} = \frac{e^{\alpha_1+\alpha_2+\alpha_3+3c} + e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c}}{\left(e^{\alpha_1+\alpha_2+\alpha_3+3c} + e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c} + 1\right)^2},$$

$$\Psi_7(c) = \frac{c}{\beta}a_1a_2a_3(a_2a_3 + 2a_3 + 3)\Lambda^{-2} = \frac{ce^{\alpha_1+\alpha_2+\alpha_3+3c}\left(e^{\alpha_2+\alpha_3+2c} + 2e^{\alpha_3+c} + 3\right)}{\beta\left(e^{\alpha_1+\alpha_2+\alpha_3+3c} + e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c} + 1\right)^2},$$

$$\begin{aligned}\Psi_8(c) &= \frac{c}{\beta}\left[a_1a_2a_3(2a_3 + 3) + a_2a_3(a_3 + 2)\right]\Lambda^{-2} \\ &= \frac{c\left[e^{\alpha_1+\alpha_2+\alpha_3+3c}\left(2e^{\alpha_3+c} + 3\right) + e^{\alpha_2+\alpha_3+2c}(e^{\alpha_3+c} + 2)\right]}{\beta\left(e^{\alpha_1+\alpha_2+\alpha_3+3c} + e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c} + 1\right)^2},\end{aligned}$$

$$\Psi_9(c) = \frac{c}{\beta}(3a_1a_2a_3 + 2a_2a_3 + a_3)\Lambda^{-2} = \frac{c\left(3e^{\alpha_1+\alpha_2+\alpha_3+3c} + 2e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c}\right)}{\beta\left(e^{\alpha_1+\alpha_2+\alpha_3+3c} + e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c} + 1\right)^2},$$

and let

$$\begin{aligned}\Psi_{10}(c) &= \left(\frac{c}{\beta}\right)^2\left[a_1a_2a_3(a_2a_3 + 4a_3 + 9) + a_2a_3(a_3 + 4) + a_3\right] \\ &= \frac{c^2\left[e^{\alpha_1+\alpha_2+\alpha_3+3c}\left(e^{\alpha_2+\alpha_3+2c} + 4e^{\alpha_3+c} + 9\right) + e^{\alpha_2+\alpha_3+2c}(e^{\alpha_3+c} + 4) + e^{\alpha_3+c}\right]}{\beta^2\left(e^{\alpha_1+\alpha_2+\alpha_3+3c} + e^{\alpha_2+\alpha_3+2c} + e^{\alpha_3+c} + 1\right)^2}.\end{aligned}$$

**Step 2: (Simplification)** The task is to show the following system for any two designs $\xi$ and $\tilde{\xi}$ in (S6),

$$\begin{aligned} \sum_{i=1}^{m} \omega_i \Psi_j\left(c_i\right) &= \sum_{i=1}^{k} \tilde{\omega}_i \Psi_j\left(\tilde{c}_i\right), j=1,\ldots,9, \\ \sum_{i=1}^{m} \omega_i \Psi_{10}\left(c_i\right) &\leq \sum_{i=1}^{k} \tilde{\omega}_i \Psi_{10}\left(\tilde{c}_i\right), \end{aligned} \tag{S7}$$

and it is sufficient to show either

$$\begin{aligned} &\{1,\Psi_1,\Psi_2,\Psi_3,\Psi_4,\Psi_5,\Psi_6,\Psi_7,\Psi_8,\Psi_9\} \text{ and } \{1,\Psi_1,\Psi_2,\Psi_3,\Psi_4,\Psi_5,\Psi_6,\Psi_7,\\ &\Psi_8,\Psi_9,\Psi_{10}\} \text{ are Chebyshev Systems, or}\\ &\{1,\Psi_1,\Psi_2,\Psi_3,\Psi_4,\Psi_5,\Psi_6,\Psi_7,\Psi_8,\Psi_9\} \text{ and } \{1,\Psi_1,\Psi_2,\Psi_3,\Psi_4,\Psi_5,\Psi_6,\Psi_7,\\ &\Psi_8,\Psi_9,-\Psi_{10}\} \text{ are Chebyshev Systems.} \end{aligned} \tag{S8}$$

Because the denominators present in $\Psi(c)$, the recursive construction of $F(c)$ as detailed in Theorem 2 of Yang and Stufken (2012) is likely to be complex and cumbersome. As a result, the resulting function $F(c)$ can become quite complicated. To address this, we simplify the process through a series of steps that retain either the equality in (S7) or the Chebyshev System in (S8), but involve simpler functions.

First, we omit $\beta$ in $\Psi_7$, $\Psi_8$ and $\Psi_9$ which does not change the equality in (S7). Then multiply all $\Psi$ functions including the constant $\Psi_0 = 1$ by the denominator and conduct row or column operations that do not change the sign of matrix determinant. At last we get rid of positive constants like

$e^{\alpha_1}, e^{\alpha_2}, e^{\alpha_3}$ and $\beta^2$ which preserve the Chebyshev System. Eventually, a set of functions $\Psi$ is simplified to

$$\left\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c}, ce^{4c}, e^{5c}, ce^{5c}, e^{6c},\right.$$
$$\left. c^2e^c\left(e^{\alpha_1+2\alpha_2+\alpha_3+4c} + 4e^{\alpha_1+\alpha_2+\alpha_3+3c} + 9e^{\alpha_1+\alpha_2+2c} + e^{\alpha_2+\alpha_3+2c} + 4e^{\alpha_2+c} + 1\right)\right\}.$$

To show (S8) is equivalent to verifying either those following claims hold

$\left\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c}, ce^{4c}, e^{5c}, ce^{5c}, e^{6c}\right\}$ and $\left\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c}, ce^{4c}, e^{5c}, ce^{5c}, e^{6c}, c^2e^c\left(e^{\alpha_1+2\alpha_2+\alpha_3+4c} + 4e^{\alpha_1+\alpha_2+\alpha_3+3c} +9e^{\alpha_1+\alpha_2+2c} + e^{\alpha_2+\alpha_3+2c} + 4e^{\alpha_2+c} + 1\right)\right\}$ are Chebyshev Systems,

$\left\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c}, ce^{4c}, e^{5c}, ce^{5c}, e^{6c}\right\}$ and $\left\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c}, ce^{4c}, e^{5c}, ce^{5c}, e^{6c}, -c^2e^c\left(e^{\alpha_1+2\alpha_2+\alpha_3+4c} + 4e^{\alpha_1+\alpha_2+\alpha_3+3c} +9e^{\alpha_1+\alpha_2+2c} + e^{\alpha_2+\alpha_3+2c} + 4e^{\alpha_2+c} + 1\right)\right\}$ are Chebyshev Systems.

**Step 3: (Calculation)** The sequence of $f_{ii}$ functions can be easily calculated according to Theorem 2 of Yang and Stufken (2012). Here $f_{11} = e^c$, $f_{22} = 1$, $f_{33} = 2e^c$, $f_{44} = 1$, $f_{55} = 6e^c$, $f_{66} = 1$, $f_{77} = 12e^c$, $f_{88} = 1$, $f_{99} = 20e^c$, $f_{10,10} = 1$, $f_{11,11} = 30e^c$, $f_{12,12} = -(\frac{1}{15}e^{\alpha_1+2\alpha_2+\alpha_3-c} + \frac{2}{75}e^{\alpha_1+\alpha_2+\alpha_3-2c} + \frac{3}{100}e^{\alpha_1+\alpha_2-3c} + \frac{1}{300}e^{\alpha_2+\alpha_3-3c} + \frac{2}{75}e^{\alpha_2-4c} + \frac{1}{15}e^{-5c})$, and $F(c) = \prod_{i=1}^{12} f_{ii}(c) = -288e^c(20e^{\alpha_1+2\alpha_2+\alpha_3+4c} + 8e^{\alpha_1+\alpha_2+\alpha_3+3c} + 9e^{\alpha_1+\alpha_2+2c} + e^{\alpha_2+\alpha_3+2c} + 8e^{\alpha_2+c} + 20) < 0$. Then designs with at most $6 = 2(4-1)$ support points form a complete class is a direct consequence of the case (d) of Theorem 2 in Yang and Stufken (2012).

For $J = 3$, we need to verify either of those following claims hold

$\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c}\}$ and $\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c},$

$c^2e^c\left(e^{\alpha_1+\alpha_2+2c} + 4e^{\alpha_1+c} + 1\right)\}$ are Chebyshev Systems,

$\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c}\}$ and $\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c},$

$-c^2e^c\left(e^{\alpha_1+\alpha_2+2c} + 4e^{\alpha_1+c} + 1\right)\}$ are Chebyshev Systems.

Direct calculation shows $F(c) = \prod_{i=1}^{8} f_{ii}(c) = -8e^c(3e^{\alpha_1+\alpha_2+2c}+4e^{\alpha_1+c}+3) < 0$. Then according to case (d) of Theorem 2 in Yang and Stufken (2012), designs with at most $4 = 2(3-1)$ points form a complete class.

Similarly, for $J = 5$, we need to verify either of two following claims hold

$\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c}, ce^{4c}, e^{5c}, ce^{5c}, e^{6c}, ce^{6c}, e^{7c}, ce^{7c}, e^{8c}\}$ and

$\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c}, ce^{4c}, e^{5c}, ce^{5c}, e^{6c}, ce^{6c}, e^{7c}, ce^{7c}, e^{8c}, c^2e^c(\triangle)\}$

are Chebyshev Systems,

$\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c}, ce^{4c}, e^{5c}, ce^{5c}, e^{6c}, ce^{6c}, e^{7c}, ce^{7c}, e^{8c}\}$ and

$\{1, e^c, ce^c, e^{2c}, ce^{2c}, e^{3c}, ce^{3c}, e^{4c}, ce^{4c}, e^{5c}, ce^{5c}, e^{6c}, ce^{6c}, e^{7c}, ce^{7c}, e^{8c}, -c^2e^c(\triangle)\}$

are Chebyshev Systems,

where $\triangle$ is the sum of some functions. And $F(c) = \prod_{i=1}^{16} f_{ii}(c) < 0$, then according to case (d) of Theorem 2 in Yang and Stufken (2012), designs with at most $8 = 2(5-1)$ points form a complete class.

The rest is similar and omitted here. □

### S1.3 Proof of Theorem 2

*Proof.* Note that the transformation of design point does not change the complete class result, because of the following factorization of the information matrix. Known $A(c) = U^\top(c)D^{-1}(c)U(c) = U^\top(\boldsymbol{x})D^{-1}(\boldsymbol{x})U(\boldsymbol{x})$, then for a design point $\boldsymbol{x}$ and its transformed design point $\boldsymbol{s}$,

$$M(\boldsymbol{x}) = H^\top(\boldsymbol{x})A(c)H(\boldsymbol{x}) = Q^\top H^\top(\boldsymbol{s})A(c)H(\boldsymbol{s})Q, \tag{S9}$$

where $H(\boldsymbol{x})$ is the design matrix corresponding to $\boldsymbol{x} = (x_1, \ldots, x_q)^\top$, $H(\boldsymbol{s})$ is the design matrix for $\boldsymbol{s} = (x_1, \ldots, x_{q-1}, c)^\top$, and $H(\boldsymbol{s})Q = H(\boldsymbol{x})$,

$$H(\boldsymbol{s}) = \begin{pmatrix} 1 & \cdots & 0 & x_1 & \cdots & c \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & x_1 & \cdots & c \end{pmatrix},$$

$$Q = \begin{pmatrix} I_{J-1} & 0_{(J-1)\times(q-1)} & 0_{(J-1)\times 1} \\ 0_{(q-1)\times(J-1)} & I_{q-1} & T(\boldsymbol{\beta}) \\ 0_{1\times(J-1)} & 0_{1\times(q-1)} & 1/\beta_q \end{pmatrix},$$

where $T(\boldsymbol{\beta}) = (-\beta_1/\beta_q, \ldots, -\beta_{q-1}/\beta_q)^\top$. Let $M(\boldsymbol{s})$ stand for the information matrix at $\boldsymbol{s} = (x_1, \ldots, x_{q-1}, c)^\top$, one can easily obtain $M(\boldsymbol{s})$ from $M(\boldsymbol{x})$ by (S9). The structures of them are identical. For convenience,

$H(\boldsymbol{s}_i)$ and $A(c_i)$ are abbreviated as $H_i$ and $A_i$. Then for a given design $\xi$, the information matrix of $\xi_s$ is

$$M(\xi_s) = \sum_{i=1}^{m} \omega_i H_i^\top A_i H_i.$$

First of all, define following weights $r_j = (1 - x_{ij})/2$ such that

$$\begin{aligned} &r_j(-1) + (1 - r_j)\,(1) = x_{ij}, \\ &r_j(-1)^2 + (1 - r_j)\,(1)^2 \geq x_{ij}^2,\ j = 1, \ldots, q-1. \end{aligned} \tag{S10}$$

The first equality is easy to verify, and the second inequality is due to the fact that the function $f(x) = x^2$ is convex. For an arbitrary design point, say $\boldsymbol{s}_i = (x_{i1}, \ldots, x_{i,q-1}, c_i)^\top$, consider the following two design points, $\boldsymbol{s}_{i1} = (-1, x_{i2}, \ldots,\ x_{i,q-1}, c_i)^\top$ and $\boldsymbol{s}_{i2} = (1, x_{i2}, \ldots,\ \ x_{i,q-1}, c_i)^\top$, and their design matrices are $H_{i1}$ and $H_{i2}$. Let $\omega_{i1} = r_1\omega_i$ and $\omega_{i2} = \omega_i - \omega_{i1}$, then $\omega_i H_i^\top A_i H_i$ and $\sum_{l=1}^2 \omega_{il} H_{il}^\top A_i H_{il}$ are exactly the same except the $J$th diagonal element.

This is true due to two facts. First the (S10) holds. Second, entries in $M(\xi_s)$ are linear in $x_{i1}$ except the $J$th diagonal components are quadratic in $x_{i1}$. As a result,

$$\omega_i H_i^\top A_i H_i \leq \sum_{l=1}^{2} \omega_{il} H_{il}^\top A_i H_{il}.$$

Repeat the procedures until $x_{i,q-1}$, and we have the following

$$\omega_i H_i^\top A_i H_i \leq \sum_{l=1}^{2^{q-1}} \tilde{\omega}_{il} \tilde{H}_{il}^\top A_i \tilde{H}_{il},$$

$$\sum_{i=1}^{m} \omega_i H_i^\top A_i H_i \leq \sum_{i=1}^{m} \sum_{l=1}^{2^{q-1}} \tilde{\omega}_{il} \tilde{H}_{il}^\top A_i \tilde{H}_{il}. \tag{S11}$$

$\tilde{H}_{il}$ is the design matrix for $\tilde{\boldsymbol{s}}_{il} = (b_{l1}, \ldots, b_{l,q-1}, c_i)$, $b_{lj} = -1$ or 1, and $(b_{l1}, \ldots, b_{l,q-1})$ are all combinations of them for $l = 1, \ldots, 2^{q-1}$. Thus $(b_{l1}, \ldots, b_{l,q-1})$, $l = 1, \ldots, 2^{q-1}$ form an OA$(2^{q-1}, q-1, 2, q-1)$. Note that the right hand side of (S11) only depends on $c_i$, and they have the same set of linear independent non-constant functions. Then following Lemma 1, there exist at most $(K+2)/2$ points $\tilde{c}_i$ such that

$$\sum_{i=1}^{m} \sum_{l=1}^{2^{q-1}} \tilde{\omega}_{il} \tilde{H}_{il}^\top A_i \tilde{H}_{il} \leq \sum_{i=1}^{(K+2)/2} \sum_{l=1}^{2^{q-1}} \tilde{\omega}_{il} \tilde{H}_{il}^\top A_i \tilde{H}_{il}.$$

These points $\tilde{c}_i$ form a single-factor design $\xi_c = \{(\tilde{c}_i, \tilde{\omega}_i), i = 1, \ldots, (K+2)/2, \sum_{i=1}^{(K+2)/2} \tilde{\omega}_i = 1\}$ and $\tilde{\omega}_i = \sum_{l=1}^{2^{q-1}} \tilde{\omega}_{il}$ and $\tilde{\xi}_s$ is only relevant to $\xi_c$ and $\tilde{\omega}_{il}$. □

### S1.4 Proof of Theorem 3

*Proof.* Let $\widetilde{H} = \left(H^\top(\boldsymbol{x}_1), \ldots, H^\top(\boldsymbol{x}_m)\right)$, and

$\widetilde{W} = \operatorname{diag}\left(\omega_1 U^\top(\boldsymbol{x}_1) D^{-1}(\boldsymbol{x}_1) U(\boldsymbol{x}_1), \ldots, \omega_m U^\top(\boldsymbol{x}_m) D^{-1}(\boldsymbol{x}_m) U(\boldsymbol{x}_m)\right)$.

According to Theorem 1, the Fisher information matrix can be written as $M(\xi) = \widetilde{H}\widetilde{W}\widetilde{H}^\top$. Because $0 < \pi_j(\boldsymbol{x}_i) < 1, j = 1, \ldots, J$, $D(\boldsymbol{x}_i)$ is positive definite. After performing an elementary row transformation on $U(\boldsymbol{x})$, the first $J-1$ rows can be changed into a diagonal matrix with $v_{11} = g_1' \prod_{t=2}^{J-1} g_t \Delta_J^{-1}(1-g_1)^{-1}, \ldots, v_{ll} = \prod_{t=1}^{l-1}(1-g_t) g_l' \prod_{t=l+1}^{J-1} g_t \Delta_J^{-1}(1-g_l)^{-1}, \ldots, v_{J-1,J-1} = \prod_{t=1}^{J-2}(1-g_t) g_{J-1}' \Delta_J^{-1}(1-g_{J-1})^{-1}$ as diagonal elements, denoted as $V(\boldsymbol{x})$. Evidently, $0 < g_l < 1, 0 < 1-g_l < 1, l = 1, \ldots, J-1$. By Assumption 2, since $g_l' > 0$ for $l = 1, \ldots, J-1$, it follows that $v_{ll} > 0$ for $l = 1, \ldots, J-1$. Therefore, $V(\boldsymbol{x})$ has full column rank, which implies that $U(\boldsymbol{x})$ has full column rank. Therefore, $M(\xi)$ is positive definite if and only if $\widetilde{H}$ has full row rank.

Let $X = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m)$. After some elementary column transformations for the matrix $\left(H^\top(\boldsymbol{x}_1), \ldots, H^\top(\boldsymbol{x}_m)\right)$, we obtain a new matrix

$$H_{\text{new}} = \begin{pmatrix} 1_m^\top & 0 & \cdots & 0 \\ 0 & 1_m^\top & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ X & X & \cdots & X \end{pmatrix},$$

where $1_m$ is an $m$-dimensional vector all of 1. In order to keep $H_{\text{new}}$ full row rank, $X$ is full row rank and $C(1_m) \cap C\left(X^\top\right) = \{\mathbf{0}\}$, $C\left(X^\top\right)$ denotes the column space of $X^\top$. Thus $m \geq q$ and the rank of the matrix $\left(1_m, X^\top\right)$ is at least $q+1$. In summary, $M(\xi)$ is positive definite only if $m \geq q+1$.

As a direct conclusion of Theorem 2.2 of Fedorov and Leonov (2014), under regularity conditions, there exists a D-optimal design that contains no more than $p(p+1)/2$ support points, which are the design points with positive weights. $\square$

## S1.5 Proof of Theorem 4

*Proof.* Following the proof of Theorem 2 in Yang et al. (2017), we have the following result for the determinant of the Fisher information matrix $M(\xi)$.

The Fisher information matrix under the design $\xi$ is

$$M(\xi) = \sum_{i=1}^{m} \omega_i M(\boldsymbol{x}_i).$$

To facilitate the subsequent derivation, we denote the $(t, l)$th entry of $M(\boldsymbol{x}_i)$ by $m_{tl}^i$.

According to the Leibniz formula for the determinant,

$$|M(\xi)| = \left| \sum_{i=1}^{m} \omega_i M(\boldsymbol{x}_i) \right| = \sum_{\sigma \in S_p} (-1)^{\mathrm{sgn}(\sigma)} \prod_{t=1}^{p} \sum_{i=1}^{m} \omega_i m_{t,\sigma(t)}^i,$$

where $\sigma$ is a permutation of $\{1, 2, \ldots, p\}$, $S_p$ is the set of all $p!$ permutations and $\mathrm{sgn}(\sigma)$ is the sign or signature of $\sigma$. Therefore,

$$|M(\xi)| = \sum_{\substack{\alpha_1 \geq 0, \ldots, \alpha_m \geq 0 \\ \alpha_1 + \cdots + \alpha_m = p}} \sum_{\sigma \in S_p} (-1)^{\mathrm{sgn}(\sigma)} \sum_{\tau \in \Phi(\alpha_1, \ldots, \alpha_m)} \prod_{t=1}^{p} \omega_{\tau(t)} m_{t,\sigma(t)}^{\tau(t)}$$

$$= \sum_{\substack{\alpha_1 \geq 0, \ldots, \alpha_m \geq 0 \\ \alpha_1 + \cdots + \alpha_m = p}} \left( \sum_{\tau \in \Phi(\alpha_1, \ldots, \alpha_m)} \sum_{\sigma \in S_p} (-1)^{\mathrm{sgn}(\sigma)} \prod_{t=1}^{p} m_{t,\sigma(t)}^{\tau(t)} \right) \omega_1^{\alpha_1} \cdots \omega_m^{\alpha_m}$$

$$= \sum_{\substack{\alpha_1 \geq 0, \ldots, \alpha_m \geq 0 \\ \alpha_1 + \cdots + \alpha_m = p}} \left( \sum_{\tau \in \Phi(\alpha_1, \ldots, \alpha_m)} |M_\tau| \right) \omega_1^{\alpha_1} \cdots \omega_m^{\alpha_m}.$$

Denote $\sum_{\tau \in \Phi(\alpha_1, \ldots, \alpha_m)} |M_\tau|$ as $c_{\alpha_1, \ldots, \alpha_m}$. This proves Equations (3.5) and (3.6) and shows that $|M(\xi)|$ is an order-$p$ homogeneous polynomial of $\omega_1, \ldots, \omega_m$.

Next, we need to show that the coefficients calculated in Equation (3.6) in conditions (1) or (2) are zero.

(1) We have known $M(\boldsymbol{x}_i) = H^\top(\boldsymbol{x}_i) U^\top(\boldsymbol{x}_i) D^{-1}(\boldsymbol{x}_i) U(\boldsymbol{x}_i) H(\boldsymbol{x}_i)$ and $U^\top(\boldsymbol{x}_i)$ has full column rank, i.e., has rank $J-1$, thus $\mathrm{rank}(M(\boldsymbol{x}_i)) \leq J-1$, $i = 1, \ldots, m$. Since $\max_{1 \leq i \leq m} \alpha_i \geq J$, without loss of generality, we assume $\alpha_1 \geq J$. Then for any $\tau \in \Phi(\alpha_1, \ldots, \alpha_m)$, $M_\tau$ has at least $J$ rows that are the same as the corresponding rows of $M(\boldsymbol{x}_1)$, these rows must be linearly correlated. Thus $|M_\tau| = 0$, which implies $c_{\alpha_1, \ldots, \alpha_m} = 0$.

(2) Let $\bar{H} = (H^\top(\boldsymbol{x}_1) U^\top(\boldsymbol{x}_1), \ldots, H^\top(\boldsymbol{x}_m) U^\top(\boldsymbol{x}_m))$ and $\bar{W} = \mathrm{diag}(\omega_1 D^{-1}(\boldsymbol{x}_1), \ldots, \omega_m D^{-1}(\boldsymbol{x}_m))$, then $M(\xi) = \bar{H}\bar{W}\bar{H}^\top$. By Cauchy-Binet formula in Section 0.8.7 of Horn and Johnson (2012), it follows

$$c_{\alpha_1, \ldots, \alpha_m} = \sum_{(v_1, \ldots, v_p) \in \Lambda(\alpha_1, \ldots, \alpha_m)} |\bar{H}[v_1, \ldots, v_p]|^2 \prod_{i:\alpha_i > 0} \prod_{l:(i-1)J < v_l \leq iJ} \pi_{i, v_l - (i-1)J}^{-1} \geq 0,$$

where $\Lambda(\alpha_1, \ldots, \alpha_m) = \{(v_1, \ldots, v_p) \mid 1 \leq v_1 < \cdots < v_p \leq mJ; \sum_{l=1}^{p} I\{(i-$

$1)J < v_l \le iJ\} = \alpha_i, i = 1, \dots, m\}$, $I\{(i-1)J < v_l \le iJ\}$ is 1 if $(i-1)J < v_l \le iJ$ is true, and 0 otherwise, and $\bar{H}[v_1, \dots, v_p]$ is the submatrix consisting of the $v_1$th$, \dots, v_p$th columns of $\bar{H}$. Without loss of generality, we assume $\alpha_1 \ge \dots \ge \alpha_n > 0 = \alpha_{n+1} = \dots = \alpha_m$, where $n \le q$. Suppose there exist some $(\alpha_1, \dots, \alpha_m)$ such that $c_{\alpha_1,\dots,\alpha_m} > 0$. This implies that there exist $(v_1, \dots, v_p) \in \Lambda(\alpha_1, \dots, \alpha_m)$ such that $\text{rank}(\bar{H}[v_1, \dots, v_p]) = p$ and $1 \le v_1 < \dots < v_p \le nJ$. Let $\bar{\bar{H}} = \bar{H}[1, \dots, nJ]$ be a $p \times nJ$ matrix, then $\bar{\bar{H}}$ is full row rank. Let $\bar{\bar{W}} = n^{-1}\,\text{diag}(D^{-1}(\boldsymbol{x}_1), \dots, D^{-1}(\boldsymbol{x}_n))$, then $\bar{\bar{H}}\bar{\bar{W}}\bar{\bar{H}}^\top$ is a positive definite matrix. We can regard $\bar{\bar{H}}\bar{\bar{W}}\bar{\bar{H}}^\top$ as the Fisher information matrix under uniform weighted design at $n$ design points, thus $n \ge q+1$ is obtained from Theorem 3. This creates a contradiction, so when $n = \sum_{i=1}^m I\{\alpha_i > 0\} \le q$, $c_{\alpha_1,\dots,\alpha_m} = 0$. □

### S1.6 Proof of Example 1

*Proof.* For convenience, denote $\pi_j(x_i)$ as $\pi_{ij}$. For model (3.7),

$$H(x_i) = \begin{pmatrix} 1 & 0 & x_i \\ 0 & 1 & x_i \end{pmatrix}, U(x_i) = \begin{pmatrix} \pi_{i1}\pi_{i2} + \pi_{i1}\pi_{i3} & \pi_{i1}\pi_{i3} \\ -\pi_{i1}\pi_{i2} & \pi_{i2}\pi_{i3} \\ -\pi_{i1}\pi_{i3} & -\pi_{i1}\pi_{i3} - \pi_{i2}\pi_{i3} \end{pmatrix}$$

and $D(x_i) = \text{diag}(\pi_{i1}, \pi_{i2}, \pi_{i3})$. Then

$$M(x_i) = H^\top(x_i)U^\top(x_i)D^{-1}(x_i)U(x_i)H(x_i)$$

$$
= \begin{pmatrix} \pi_{i1}\pi_{i2} + \pi_{i1}\pi_{i3} & \pi_{i1}\pi_{i3} & (\pi_{i1}\pi_{i2} + 2\pi_{i1}\pi_{i3})x_i \\ \pi_{i1}\pi_{i3} & \pi_{i1}\pi_{i3} + \pi_{i2}\pi_{i3} & (2\pi_{i1}\pi_{i3} + \pi_{i2}\pi_{i3})x_i \\ (\pi_{i1}\pi_{i2} + 2\pi_{i1}\pi_{i3})x_i & (2\pi_{i1}\pi_{i3} + \pi_{i2}\pi_{i3})x_i & (\pi_{i1}\pi_{i2} + 4\pi_{i1}\pi_{i3} + \pi_{i2}\pi_{i3})x_i^2 \end{pmatrix}.
$$

Assume that the weights of design $\xi$ at points $x_1$ and $x_2$ are $\omega_1$ and $\omega_2$, respectively. Then the determinant of the Fisher information matrix corresponding to design $\xi$ is

$$
|M(\xi)| = (b_1\omega_1 + b_2\omega_2)\omega_1\omega_2,
$$

where $b_1 = (\pi_{21}\pi_{22} + 4\pi_{21}\pi_{23} + \pi_{22}\pi_{23})\pi_{11}\pi_{12}\pi_{13}(x_1 - x_2)^2$, $b_2 = (\pi_{11}\pi_{12} + 4\pi_{11}\pi_{13} + \pi_{12}\pi_{13})\ \pi_{21}\pi_{22}\pi_{23}(x_1 - x_2)^2$. By Corollary 2 of Yang et al. (2017), it follows that $\omega_1 = \frac{b_1 - b_2 + \sqrt{b_1^2 - b_1 b_2 + b_2^2}}{2b_1 - b_2 + \sqrt{b_1^2 - b_1 b_2 + b_2^2}}$, $\omega_2 = \frac{b_1}{2b_1 - b_2 + \sqrt{b_1^2 - b_1 b_2 + b_2^2}}$ for maximizing $|M(\xi)|$. □

### S1.7 Proof of Theorem 5

*Proof.* According to the proof of Theorem 2, $M(\tilde{\xi}_s) = \sum_{i=1}^{(K+2)/2} \sum_{l=1}^{2^{q-1}} \tilde{\omega}_{il} \tilde{H}_{il}^\top \times A_i \tilde{H}_{il}$ and for a given $\boldsymbol{\gamma}$, $A_i$ is only relevant to $\tilde{c}_i$. For $J$ response categories, $A_i$ is a $(J-1) \times (J-1)$ matrix. Let the $(j,t)$th element of $A_i$ be denoted as $A_i^{j,t}$, the sum of the $j$th row as $A_i^j$, and the sum of all elements as $A_i^{sum}$. Then the $(j,t)$th $(j \le t)$ elements of $M(\tilde{\xi}_s)$ is

$$\begin{cases} \sum_{i=1}^{(K+2)/2} \tilde{\omega}_i A_i^{j,t}, & 1 \le j \le t \le J-1; \\ \sum_{i=1}^{(K+2)/2} \tilde{\omega}_i A_i^{sum}, & j = t, j = J, \ldots, J+q-2; \\ \sum_{i=1}^{(K+2)/2} \tilde{\omega}_i A_i^{sum} \tilde{c}_i^2, & j = t = J+q-1; \\ \sum_{i=1}^{(K+2)/2} \sum_{l=1}^{2^{q-1}} \tilde{\omega}_{il} A_i^j b_{l,t-J+1}, & 1 \le j \le J-1, J \le t \le J+q-2; \\ \sum_{i=1}^{(K+2)/2} \tilde{\omega}_i A_i^j \tilde{c}_i, & 1 \le j \le J-1, t = J+q-1; \\ \sum_{i=1}^{(K+2)/2} \sum_{l=1}^{2^{q-1}} \tilde{\omega}_{il} A_i^{sum} b_{l,j-J+1} b_{l,t-J+1}, & J \le j < t \le J+q-2; \\ \sum_{i=1}^{(K+2)/2} \sum_{l=1}^{2^{q-1}} \tilde{\omega}_{il} A_i^{sum} b_{l,j-J+1} \tilde{c}_i, & J \le j \le J+q-2, t = J+q-1. \end{cases}$$

Divide $M(\tilde{\xi}_s)$ into $\left(\begin{smallmatrix} A & B \\ B^\top & D \end{smallmatrix}\right)$, where $A$ is a $(J-1)\times(J-1)$ matrix related only to $\xi_c = \{(\tilde{c}_i, \tilde{\omega}_i), i = 1, \ldots, (K+2)/2, \sum_{i=1}^{(K+2)/2} \tilde{\omega}_i = 1\}$, $B$ is a $(J-1) \times q$ matrix, and $D$ is a $q \times q$ matrix. Then $|M(\tilde{\xi}_s)| = |D||A - BD^{-1}B^\top|$. Assume that the eigenvalues of $D$ are $\lambda_t, t = 1, \ldots, q$ and the diagonal elements of $D$ are $D_{tt}, t = 1, \ldots, q$. It is known that

$$\prod_{t=1}^{q} \lambda_t \le \prod_{t=1}^{q} D_{tt}.$$

The equality holds if $D$ is a diagonal matrix, then $\sum_{i=1}^{(K+2)/2} \sum_{l=1}^{2^{q-1}} \tilde{\omega}_{il} A_i^{sum}$ $\times b_{l,j-J+1} b_{l,t-J+1} = 0$, $J \le j < t \le J+q-2$ and $\sum_{i=1}^{(K+2)/2} \sum_{l=1}^{2^{q-1}} \tilde{\omega}_{il} A_i^{sum}$ $\times b_{l,j-J+1} \tilde{c}_i = 0$, $J \le j \le J+q-2, t = J+q-1$. Therefore, $\tilde{\omega}_{il} = \tilde{\omega}_i / 2^{q-1}, i = 1, \ldots, (K+2)/2, l = 1, \ldots, q-1$ can be derived. In addition,

$\sum_{i=1}^{(K+2)/2} \sum_{l=1}^{2^{q-1}} \tilde{\omega}_{il} A_i^j b_{l,t-J+1} = 0$, $1 \leq j \leq J-1, J \leq t \leq J+q-2$. Thus, the elements of the first $J-1$ columns of $B$ are all 0, and the last column is related to $\xi_c$ only. $D$ is a diagonal matrix related only to $\xi_c$. Then the information matrix of this design is only relevant to $\xi_c$ and is D-optimal under $\xi_c$. Denote this design as $\tilde{\xi}_s' = \{(\tilde{\boldsymbol{s}}_{il}, \tilde{\omega}_i/2^{q-1}), i = 1, \ldots, (K+2)/2, l = 1, \ldots, 2^{q-1}\}$, $\tilde{\boldsymbol{s}}_{il} = (b_{l1}, \ldots, b_{l,q-1}, \tilde{c}_i)$ and $b_{lj} = -1$ or 1.

For each $i$, $i = 1, \ldots, (K+2)/2$, select $2^{q'-1}$ $\tilde{\boldsymbol{s}}_{il}$ from $\tilde{\xi}_s'$ such that $(b_{l1}, \ldots, b_{l,q-1})$, $l = 1, \ldots, 2^{q'-1}$ form an $\text{OA}(2^{q'-1}, q-1, 2, 2)$. Let $\tilde{\xi}_s^* = \{(\tilde{\boldsymbol{s}}_{il}, \tilde{\omega}_i/2^{q'-1}), i = 1, \ldots, (K+2)/2, l = 1, \ldots, 2^{q'-1}\}$. It can be verified that $M(\tilde{\xi}_s^*) = M(\tilde{\xi}_s')$. Therefore, $\tilde{\xi}_s^*$ is D-optimal under $\xi_c$ and only relevant to $\xi_c$.

□

## S1.8 Proof of Theorem 6

*Proof.* Define $S_0 := \{\mathbf{w} = (\omega_1, \ldots, \omega_m)^\top \in \mathbb{R}^m | \omega_i \geq 0, i = 1, \ldots, m; \sum_{i=1}^m \omega_i = 1\}$. Obviously $S_0$ is a closed convex set. Since $f(\mathbf{w}) = |M(\xi)| = \sum_{\substack{\alpha_1 \geq 0, \ldots, \alpha_m \geq 0 \\ \alpha_1 + \cdots + \alpha_m = p}} c_{\alpha_1, \ldots, \alpha_m} \omega_1^{\alpha_1} \cdots \omega_m^{\alpha_m}$ is an order-$p$ homogeneous polynomial of $\omega_1, \ldots, \omega_m$, then it must be continuous on $S_0$. According to the Weierstrass theorem (see, for example, Theorem 3.1 in Sundaram (1996)), there must exist a $\mathbf{w}^* \in S_0$ such that $f(\mathbf{w})$ attains its maximum at $\mathbf{w}^*$. □

## S1.9 Proof of Theorem 7

*Proof.* (1)⇒(2) $\xi^*$ is the D-optimal design, then $\forall\, \xi \in \Xi$, $|M(\xi^*)| \geq |M(\xi)|$ holds, and $\log|M(\xi^*)| \geq \log|M(\xi)|$ also holds. Thus $\forall\, \varepsilon(0 \leq \varepsilon \leq 1)$ and $\boldsymbol{x} \in \mathcal{X}$, we have

$$\begin{aligned}&\log\left|M\left((1-\varepsilon)\xi^* + \varepsilon\boldsymbol{x}\right)\right| - \log\left|M\left(\xi^*\right)\right| \\ =& \log\left|(1-\varepsilon)M(\xi^*) + \varepsilon M(\boldsymbol{x})\right| - \log\left|M\left(\xi^*\right)\right| \leq 0.\end{aligned}$$

Then

$$\phi(\boldsymbol{x}, \xi^*) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\left(\log\left|(1-\varepsilon)M(\xi^*) + \varepsilon M(\boldsymbol{x})\right| - \log\left|M\left(\xi^*\right)\right|\right) \leq 0.$$

(2)⇒(1) Section 3.5.2 of Silvey (1980) gives many properties of the *Fréchet derivate*, e.g., $\phi(\xi, \xi^*) = \phi(\sum_{\boldsymbol{x}} \omega(\boldsymbol{x})\boldsymbol{x}, \xi^*) = \sum_{\boldsymbol{x}} \omega(\boldsymbol{x})\phi(\boldsymbol{x}, \xi^*)$, where $\omega(\boldsymbol{x})$ denotes the weight of design $\xi$ at design point $\boldsymbol{x}$ and $0 \leq \omega(\boldsymbol{x}) \leq 1$.

$\forall\, \boldsymbol{x} \in \mathcal{X}$, $\phi(\boldsymbol{x}, \xi^*) \leq 0$. Then $\forall\, \xi \in \Xi$, according to the above property, we have $\phi(\xi, \xi^*) = \sum_{\boldsymbol{x}} \omega(\boldsymbol{x})\phi(\boldsymbol{x}, \xi^*) \leq 0$. Concavity of $\log|M(\xi)|$ implies that

$$\frac{1}{\varepsilon}\left(\log\left|(1-\varepsilon)M(\xi^*) + \varepsilon M(\xi)\right| - \log\left|M\left(\xi^*\right)\right|\right)$$

is a non-increasing function of $\varepsilon$ in $0 < \varepsilon \leq 1$. By putting $\varepsilon = 1$, we have the result that

$$0 \geq \phi(\xi, \xi^*) \geq \log|M(\xi)| - \log|M(\xi^*)|.$$

So $\forall\ \xi \in \Xi$, $\log|M(\xi^*)| \geq \log|M(\xi)|$, i.e. $\xi^*$ is the D-optimal design.

(1)⇒(3) $\xi^*$ is the D-optimal design, thus $\forall\ \boldsymbol{x} \in \mathcal{X}$, $\phi(\boldsymbol{x}, \xi^*) \leq 0$. It is easy to derive

$$\phi(\xi^*, \xi^*) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left( \log|(1-\varepsilon)M(\xi^*) + \varepsilon M(\xi^*)| - \log|M(\xi^*)| \right) = 0,$$

and

$$\phi(\xi^*, \xi^*) = \sum_{\boldsymbol{x}} \omega(\boldsymbol{x}) \phi(\boldsymbol{x}, \xi^*).$$

Suppose $\boldsymbol{x}$ be the design point of $\xi^*$, then $\omega(\boldsymbol{x}) > 0$. Thus it follows that at each design point $\boldsymbol{x}$ of $\xi^*$, $\phi(\boldsymbol{x}, \xi^*)$ attains its maximum value, i.e. $\phi(\boldsymbol{x}, \xi^*) = 0$.

(3)⇒(1) At each design point $\boldsymbol{x}$ of $\xi^*$, $\phi(\boldsymbol{x}, \xi^*)$ attains its maximum value 0, we have $\forall\ \boldsymbol{x} \in \mathcal{X}$, $\phi(\boldsymbol{x}, \xi^*) \leq 0$. Therefore, $\xi^*$ is the D-optimal design. □

## S2 Beyond Proportional Odds Assumption

This section is an extension of the AC po model.

Similarly, the general AC partial proportional odds (ppo) model is

$$g\left(\frac{\pi_j(\boldsymbol{x}_i)}{\pi_j(\boldsymbol{x}_i) + \pi_{j+1}(\boldsymbol{x}_i)}\right) = \eta_{ij} = \boldsymbol{x}_{i0}^\top \boldsymbol{\theta}_j + \boldsymbol{x}_{i1}^\top \boldsymbol{\beta}, \tag{S12}$$

for $j = 1, \ldots, J-1$, where $\boldsymbol{x}_i = (\boldsymbol{x}_{i0}^\top, \boldsymbol{x}_{i1}^\top)^\top \in \mathbb{R}^q$. $\boldsymbol{\theta}_j \in \mathbb{R}^{q_0}$ stands for the unknown parameters belong to the $j$th category only, $j = 1, \ldots, J-1$, $\boldsymbol{\beta} \in$

$\mathbb{R}^{q_1}$ stands for the unknown parameters that are common in all categories, $p = (J-1)q_0 + q_1$. AC models under po and nonproportional odds (npo) assumptions are two special cases with $\boldsymbol{x}_{i0} = 1, j = 1, \ldots, J-1$, and $\boldsymbol{x}_{i1} = 0$, respectively.

Here $\boldsymbol{\eta}_i = H(\boldsymbol{x}_i)\boldsymbol{\gamma}$, where $H(\boldsymbol{x}_i)$ is $(J-1) \times p$ design matrix with

$$H(\boldsymbol{x}_i) = \begin{pmatrix} \boldsymbol{x}_{i0}^\top & 0 & \cdots & 0 & \boldsymbol{x}_{i1}^\top \\ 0 & \boldsymbol{x}_{i0}^\top & \cdots & 0 & \boldsymbol{x}_{i1}^\top \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{x}_{i0}^\top & \boldsymbol{x}_{i1}^\top \end{pmatrix},$$

and $\boldsymbol{\gamma} = (\boldsymbol{\theta}_1^\top, \ldots, \boldsymbol{\theta}_{J-1}^\top, \boldsymbol{\beta}^\top)^\top$.

The Fisher information matrix for Model (S12) under the design $\xi$ is the same as Equation (2.4).

The following are modifications of some related theorems. The proofs are similar to those for the AC po model and are therefore omitted.

**Theorem S1.** *The Fisher information matrix $M(\xi)$ for Model (S12) calculated in Equation (2.4) is positive definite only if $m \geq q$. Furthermore, there exists a D-optimal design with $m \leq p(p+1)/2$.*

**Theorem S2.** *The determinant of the Fisher information matrix $M(\xi)$ is*

$$|M(\xi)| = \sum_{\substack{\alpha_1 \geq 0, \ldots, \alpha_m \geq 0 \\ \alpha_1 + \cdots + \alpha_m = p}} c_{\alpha_1, \ldots, \alpha_m} \omega_1^{\alpha_1} \cdots \omega_m^{\alpha_m}, \tag{S13}$$

*where*

$$c_{\alpha_1,\ldots,\alpha_m} = \sum_{\tau \in \Phi(\alpha_1,\ldots,\alpha_m)} |M_\tau| \,. \tag{S14}$$

*Furthermore, let* $n = \sum_{i=1}^{m} I\{\alpha_i > 0\}$*, where* $I\{\alpha_i > 0\}$ *is 1 if* $\alpha_i > 0$ *is true, and 0 otherwise. Then the coefficients (S14) are zero if the* $(\alpha_1, \ldots, \alpha_m)$ *satisfies one of the following conditions.*

*(1)* $\max_{1\le i\le m} \alpha_i \ge J$*.* *(2)* $n \le q-1$*.*

For the AC ppo models with $q$ continuous factors, we focus on the support points $\boldsymbol{s}_i = (x_{i1}, \ldots, x_{i,q_0}, x_{i,q_0+1}, \ldots, x_{i,q-1}, c_{i1}, \ldots, c_{i,J-1})^\top$, where $c_{ij} = \sum_{t=1}^{q_0} \theta_{jt} x_{it} + \sum_{t=q_0+1}^{q} \beta_{t-q_0} x_{it}$, $j = 1, \ldots, J-1$ and $\theta_{jt} \neq 0$, $\beta_{t-q_0} \neq 0$ for all possible $t$.

**Theorem S3.** *In the transformed design space of AC ppo model with q continuous factors, for an arbitrary design* $\xi_s = \{(\boldsymbol{s}_i, \omega_i),\ i = 1, \ldots, m;\ \sum_{i=1}^{m} \omega_i = 1\}$*, there exists a design* $\tilde{\xi}_s$ *such that the following inequality for information matrices hold:* $M(\xi_s) \le M(\tilde{\xi}_s)$*, where*

$$\tilde{\xi}_s = \left\{ (\tilde{\boldsymbol{s}}_{il}, \tilde{\omega}_{il}), i = 1, \ldots, m, l = 1, \ldots, 2^{q-1} \right\},$$

*and* $\tilde{\boldsymbol{s}}_{il} = (b_{l1}, \ldots, b_{l,q-1}, \tilde{c}_{i1}, \ldots, \tilde{c}_{i,J-1})$*. Here* $b_{lj} = -1$ *or* $1$*, and* $(b_{l1}, \ldots, b_{l,q-1})$*,* $l = 1, \ldots, 2^{q-1}$ *are all combinations of them, and* $\tilde{c}_{i1}, \ldots, \tilde{c}_{i,J-1}, i = 1, \ldots, m$ *are* $m(J-1)$ *numbers need to be solved.*

## S3 Simulation studies

**Example S1.** This example is a supplement to Example 3. In this example, we demonstrate the optimal design $\xi^*$ searched out by our method in Table S1. For comparison, we also report the results for $\xi_{2OA}$ and the corresponding $\xi$ in Table S2 and the D-optimal design $\xi^*_{For}$ constructed in Huang et al. (2024) in Table S3.

Table S1: D-optimal designs $\xi_c^*$ and $\xi^*$

| $\xi_c^*$ | $\omega_i$ | $\xi^*$ | | | | | | $\omega_{il}$ |
|---|---|---|---|---|---|---|---|---|
| −1.56997 | 0.51688 | 1 | 1 | 1 | 1 | 1 | 0.47668 | 0.06461 |
| | | 1 | 1 | 1 | −1 | −1 | −0.18999 | 0.06461 |
| | | 1 | −1 | −1 | 1 | −1 | −1.52332 | 0.06461 |
| | | 1 | −1 | −1 | −1 | 1 | 0.47668 | 0.06461 |
| | | −1 | 1 | −1 | 1 | −1 | −3.52332 | 0.06461 |
| | | −1 | 1 | −1 | −1 | 1 | −1.52332 | 0.06461 |
| | | −1 | −1 | 1 | 1 | 1 | 1.14335 | 0.06461 |
| | | −1 | −1 | 1 | −1 | −1 | 0.47668 | 0.06461 |
| −2.84867 | 0.33735 | 1 | 1 | 1 | 1 | 1 | 0.05044 | 0.04217 |
| | | 1 | 1 | 1 | −1 | −1 | −0.61623 | 0.04217 |
| | | 1 | −1 | −1 | 1 | −1 | −1.94956 | 0.04217 |
| | | 1 | −1 | −1 | −1 | 1 | 0.05044 | 0.04217 |
| | | −1 | 1 | −1 | 1 | −1 | −3.94956 | 0.04217 |
| | | −1 | 1 | −1 | −1 | 1 | −1.94956 | 0.04217 |
| | | −1 | −1 | 1 | 1 | 1 | 0.71711 | 0.04217 |
| | | −1 | −1 | 1 | −1 | −1 | 0.05044 | 0.04216 |
| −3.90903 | 0.14577 | 1 | 1 | 1 | 1 | 1 | −0.30301 | 0.01822 |
| | | 1 | 1 | 1 | −1 | −1 | −0.96968 | 0.01822 |
| | | 1 | −1 | −1 | 1 | −1 | −2.30301 | 0.01822 |
| | | 1 | −1 | −1 | −1 | 1 | −0.30301 | 0.01822 |
| | | −1 | 1 | −1 | 1 | −1 | −4.30301 | 0.01822 |
| | | −1 | 1 | −1 | −1 | 1 | −2.30301 | 0.01822 |
| | | −1 | −1 | 1 | 1 | 1 | 0.36366 | 0.01822 |
| | | −1 | −1 | 1 | −1 | −1 | −0.30301 | 0.01823 |

Table S2: Designs $\xi_{2OA}$ and $\xi$

| $\xi_{2OA}$ | $\omega_i$ | $\xi$ | | | | | | $\omega_{il}$ |
|---|---|---|---|---|---|---|---|---|
| $-1.61058$ | $0.56690$ | 1 | 1 | 1 | 1 | 1 | 0.46314 | 0.07086 |
| | | 1 | 1 | 1 | $-1$ | $-1$ | $-0.20353$ | 0.07086 |
| | | 1 | $-1$ | $-1$ | 1 | $-1$ | $-1.53686$ | 0.07086 |
| | | 1 | $-1$ | $-1$ | $-1$ | 1 | 0.46314 | 0.07086 |
| | | $-1$ | 1 | $-1$ | 1 | $-1$ | $-3.53686$ | 0.07086 |
| | | $-1$ | 1 | $-1$ | $-1$ | 1 | $-1.53686$ | 0.07086 |
| | | $-1$ | $-1$ | 1 | 1 | 1 | 1.12981 | 0.07087 |
| | | $-1$ | $-1$ | 1 | $-1$ | $-1$ | 0.46314 | 0.07087 |
| $-3.17952$ | $0.43310$ | 1 | 1 | 1 | 1 | 1 | $-0.05984$ | 0.05414 |
| | | 1 | 1 | 1 | $-1$ | $-1$ | $-0.72651$ | 0.05414 |
| | | 1 | $-1$ | $-1$ | 1 | $-1$ | $-2.05984$ | 0.05414 |
| | | 1 | $-1$ | $-1$ | $-1$ | 1 | $-0.05984$ | 0.05414 |
| | | $-1$ | 1 | $-1$ | 1 | $-1$ | $-4.05984$ | 0.05414 |
| | | $-1$ | 1 | $-1$ | $-1$ | 1 | $-2.05984$ | 0.05414 |
| | | $-1$ | $-1$ | 1 | 1 | 1 | 0.60683 | 0.05413 |
| | | $-1$ | $-1$ | 1 | $-1$ | $-1$ | $-0.05984$ | 0.05413 |

Table S3: D-optimal design $\xi^*_{For}$

| | $\xi^*_{For}$ | | | | | | $\omega_i$ |
|---|---|---|---|---|---|---|---|
| 1 | $-1$ | $-1$ | 1 | $-1$ | 1 | 2.00000 | 0.08322 |
| 2 | 1 | $-1$ | $-1$ | $-1$ | $-1$ | $-0.90398$ | 0.10898 |
| 3 | 1 | $-1$ | $-1$ | 1 | 1 | $-0.21869$ | 0.10204 |
| 4 | $-1$ | 1 | 1 | 1 | $-1$ | $-1.63117$ | 0.14459 |
| 5 | $-1$ | 1 | $-1$ | $-1$ | 1 | $-1.90297$ | 0.11757 |
| 6 | $-1$ | $-1$ | $-1$ | 1 | $-1$ | $-2.91807$ | 0.10720 |
| 7 | 1 | 1 | 1 | $-1$ | 1 | 0.33563 | 0.09275 |
| 8 | $-1$ | $-1$ | 1 | 1 | 1 | 0.77604 | 0.03485 |
| 9 | 1 | $-1$ | 1 | $-1$ | $-1$ | 0.74782 | 0.06714 |
| 10 | $-1$ | $-1$ | 1 | $-1$ | 1 | 1.45673 | 0.04920 |
| 11 | 1 | $-1$ | 1 | 1 | 1 | 1.40546 | 0.04781 |
| 12 | $-1$ | $-1$ | 1 | $-1$ | $-1$ | $-0.25447$ | 0.02166 |
| 13 | 1 | $-1$ | 1 | 1 | 1 | 1.84637 | 0.02299 |

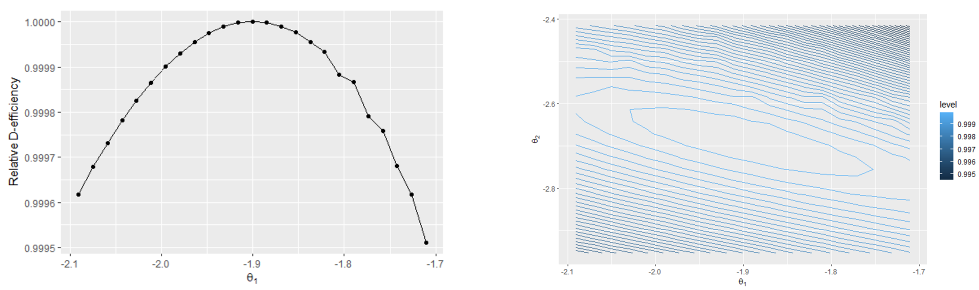
Due to randomness, each run of the code may result in a different $\xi^*_{For}$. We list only one of them. As can be seen in Table S3, $\xi^*_{For}$ contains 13 design points. And from Tables S1 and S2, $\xi^*$ and $\xi$ have 24 and 16 design points respectively. It is calculated that the relative efficiencies of $\xi^*$ against $\xi$ and $\xi^*_{For}$ are 0.99022 and 0.94754, respectively. Thus, compared to $\xi^*_{For}$, using $\xi$ can be significantly more efficient and does not add much cost to change the experimental settings.

Note that for Tables S1 and S2, the weights of the eight design points in each block should be equal; the slight differences shown are caused by rounding.

**Example S2.** As an example, take model (5.8) and consider the situation where the initial guesses of the parameters are incorrect. Suppose the pre-specified value of the parameter vector for the locally optimal design fluctuates in a moderate range (10% the magnitude of the true value).

For visualization purposes, we report the results for only one of the three parameters that is misspecified (we choose $\theta_1$ as an example) and two parameters that are misspecified (we choose $\theta_1$ and $\theta_2$ as an example) in Figure S1.

As can be seen in Figure S1(a), the relative D-efficiencies of the D-optimal designs are all greater than 99.95% when $\theta_1$ is misspecified. Based




(a) $\theta_1$ is misspecified

(b) both $\theta_1$ and $\theta_2$ are misspecified

Figure S1: Relative D-efficiencies when the parameters are misspecified.

on the contour plot of Figure S1(b), it can be shown that when $\theta_1$ and $\theta_2$ are misspecified, the relative D-efficiencies are also greater than 99.4%.

To give a comprehensive result, we also consider the case that all the three parameters are misspecified. The results are summarized in Table S4. The minimum efficiency is 98.7%, which indicates the D-optimal designs with moderately misspecified parameters are quite robust and still have satisfactory performances.

Table S4: Summary of relative D-efficiencies when all the parameters are misspecified

| Min | 1st Quartile | Median | 3st Quartile | Max |
|---|---|---|---|---|
| 0.98714 | 0.99673 | 0.99793 | 0.99894 | 1.00000 |

**Example S3.** Consider the experiment of the dose-response relationship in Chuang-Stein and Agresti (1997). The five ordered categories death, vegetative state, major disability, minor disability and good recovery describe the clinical outcome of patients who experienced trauma. In the literature on critical care, these five categories are often called the Glasgow Outcome Scale (GOS). The experiment includes four treatment groups, with placebo serving as the control. The three intravenous doses for the investigational medication are labelled as low, medium and high. Let $x$ denote the dose of medicine. Assume that the safe dose range of the medicine is 0-300mL and placebo, low, medium and high correspond to $x_1 = 0$, $x_2 = 100$, $x_3 = 200$ and $x_4 = 300$, respectively. From the above there are $m = 4$ design points and $J = 5$ categories. After simulation, it is found that probit link is the best among the five commonly used link functions in terms of the BIC criterion.

Then the AC po model under probit link function is given by

$$\Phi^{-1}\left(\frac{\pi_{i,j}}{\pi_{i,j} + \pi_{i,j+1}}\right) = \alpha_j + \beta x_i, i = 1, 2, 3, 4, j = 1, 2, 3, 4. \tag{S15}$$

The parameter estimates are $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4, \hat{\beta})^\top = (0.73748, -0.61707, -0.00838, 0.36878, -0.00042)^\top$.

Considering $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4, \hat{\beta})^\top$ as the assumed values, the locally D-optimal design of model (S15) is

$$\xi^* = \begin{pmatrix} 0 & 300 \\ 0.49775 & 0.50225 \end{pmatrix}.$$

The relative D-efficiency of the original allocation and the uniform allocation with respect to the locally D-optimal design are 89.34% and 89.20%, respectively.

To evaluate the performance of $\xi^*$ obtained by Algorithm 1, we compare it with the optimal design $\xi^*_{For}$ obtained by the ForLion algorithm proposed by Huang et al. (2024). They show equivalent performance in relative D-efficiency, with computational times for $\xi^*$ and $\xi^*_{For}$ presented in Figure S2.

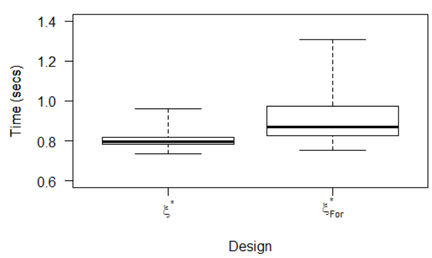



Figure S2: Computational time of $\xi^*$ and $\xi^*_{For}$

From Figure S2, it can be seen that $\xi^*$ is more advantageous in terms of computational time.

## Bibliography

Ai, M. Y., Z. Q. Ye and J. Yu (2023). Locally D-optimal designs for hierarchical response experiments. *Statistica Sinica* **33**, 381–399.

Chuang-Stein, C. and A. Agresti(1997). A review of tests for detecting a monotone dose-response relationship with ordinal response data. *Statistics in medicine* **16**, 2599–2618.

Fedorov, V. V. and S. L. Leonov (2014). *Optimal Design for Nonlinear Response Models*. Chapman & Hall/CRC.

Hao, S. and M. Yang (2020). Support point of locally optimal designs for multinomial logistic regression models. *Journal of Statistical Planning and Inference* **209**, 144–159.

Horn, R. A. and C. R. Johnson (2012). *Matrix Analysis*. 2nd Edition. Cambridge university press.

Huang, Y. F., K. Li, A. Mandal and J. Yang (2024). ForLion: a new algorithm for D-optimal designs under general parametric statistical models with mixed factors. *Statistics and Computing* **34:157**.

Silvey, S. D. (1980). *Optimal Design*. Chapman & Hall/CRC.

Sundaram, R. K. (1996). *A First Course in Optimization Theory*. Cambridge university press.

Yang, J., L. P. Tong and A. Mandal (2017). D-optimal designs with ordered categorical data. *Statistica Sinica* **27**, 1879–1902.

Yang, M. and J. Stufken (2012). Identifying locally optimal designs for nonlinear models: A simple extension with profound consequences. *Annals of Statistics* **40(3)**, 1665–1681.