Models for Order-of-Addition Screening Experiments

Jing-Wen Huang[1] and Hongquan Xu[2]

[1] *Academia Sinica* and [2] *University of California, Los Angeles*

.

S1. Proofs

To facilitate the proofs, we first show the full design is D-optimal under the following component-indicator model

$$y = \beta_0 + \sum_{k=1}^{m-1} S_k(\boldsymbol{x})\alpha_k + \epsilon, \tag{S1.1}$$

where $S_k(\boldsymbol{x})$ are component indicators defined in (3.1). Let $\boldsymbol{\xi}$ be the full design consisting of $n = \binom{m}{q}q!$ possible permutations of q components. Let $f_1(\boldsymbol{x}) = (1, S_1(\boldsymbol{x}), \ldots, S_{m-1}(\boldsymbol{x}))^T$ and $M_1(\boldsymbol{\xi}) = n^{-1}\sum_{\boldsymbol{x}\in\boldsymbol{\xi}} f_1(\boldsymbol{x})f_1(\boldsymbol{x})^T$. The following lemma shows that the equality in (2.9) holds for any run $\boldsymbol{x} \in \boldsymbol{\xi}$ so that the full design is D-optimal under the component-indicator model (S1.1).

Lemma 4. *For any* $\boldsymbol{x} \in \boldsymbol{\xi}$, $f_1(\boldsymbol{x})^T M_1^{-1}(\boldsymbol{\xi}) f_1(\boldsymbol{x}) = m$.

Proof. It is straightforward to verify that the information matrix for the full design under the model (S1.1) is

$$M_1(\boldsymbol{\xi}) = \begin{pmatrix} 1 & \frac{q}{m}\mathbf{1}_{m-1}^T \\ \frac{q}{m}\mathbf{1}_{m-1} & \frac{q}{m}\left(\frac{m-q}{m-1}\boldsymbol{I}_{m-1} + \frac{q-1}{m-1}\boldsymbol{J}_{m-1}\right) \end{pmatrix},$$

where $\mathbf{1}_{m-1}$ is a vector of ones, $\boldsymbol{I}_{m-1}$ is the identity matrix of order $m-1$ and $\boldsymbol{J}_{m-1}$ is the matrix of ones of order $m-1$. Therefore, the inverse of the information matrix is

$$M_1^{-1}(\boldsymbol{\xi}) = \begin{pmatrix} 1 + \frac{q(m-1)^2}{m-q} & -\frac{m(m-1)}{m-q}\mathbf{1}_{m-1}^T \\ -\frac{m(m-1)}{m-q}\mathbf{1}_{m-1}^T & \frac{m(m-1)}{q(m-q)}(\boldsymbol{I}_{m-1} + \boldsymbol{J}_{m-1}) \end{pmatrix}.$$

Because $(S_1(\boldsymbol{x}), \ldots, S_{m-1}(\boldsymbol{x}))$ consists of either q 1's and $m-q-1$ 0's or $q-1$ 1's and $m-q$ 0's, it is straightforward to verify that $f_1(\boldsymbol{x})^T M_1^{-1}(\boldsymbol{\xi}) f_1(\boldsymbol{x}) = m$. This completes the proof. □

Proof of Theorem 1. For convenience, we arrange the terms of the PWOi model (3.2) as follows:

$$f(\boldsymbol{x}) = (1, S_1(\boldsymbol{x}), \ldots, S_{m-1}(\boldsymbol{x}), Z_{1,2}(\boldsymbol{x}), \ldots, Z_{m-1,m}(\boldsymbol{x}))^T = (f_1(\boldsymbol{x})^T, f_2(\boldsymbol{x})^T)^T,$$

where $f_1(\boldsymbol{x})$ is the vector of the intercept and component indicators and $f_2(\boldsymbol{x})$ is the vector of the PWO factors. By Lemma 1, the PWO factors $Z_{i,j}(\boldsymbol{x})$'s are orthogonal to the intercept and component indicators $S_k(\boldsymbol{x})$'s under the full design; therefore, $f_1(\boldsymbol{x}) f_2(\boldsymbol{x})^T$ is a matrix of zeros. Thus the information matrix $M(\boldsymbol{\xi}) = n^{-1} \sum_{\boldsymbol{x} \in \boldsymbol{\xi}} f(\boldsymbol{x}) f(\boldsymbol{x})^T = \mathrm{diag}(M_1(\boldsymbol{\xi}), M_2(\boldsymbol{\xi}))$ is a block diagonal matrix, where $M_1(\boldsymbol{\xi})$ is the information matrix corresponding to the intercept and the component indicators, and $M_2(\boldsymbol{\xi})$ is the information matrix corresponding to the PWO factors. Then $f(\boldsymbol{x})^T M(\boldsymbol{\xi})^{-1} f(\boldsymbol{x}) = f_1(\boldsymbol{x})^T M_1(\boldsymbol{\xi})^{-1} f_1(\boldsymbol{x}) + f_2(\boldsymbol{x})^T M_2(\boldsymbol{\xi})^{-1} f_2(\boldsymbol{x})$. By Lemma 4, we have $f_1(\boldsymbol{x})^T M_1(\boldsymbol{\xi})^{-1} f_1(\boldsymbol{x}) = m$. On the other hand, from the result in Stokes and Xu (2024), we have $f_2(\boldsymbol{x})^T M_2(\boldsymbol{\xi})^{-1} f_2(\boldsymbol{x}) = 1 + \binom{m}{2} - 1$. Therefore, $f(\boldsymbol{x})^T M(\boldsymbol{\xi})^{-1} f(\boldsymbol{x}) = m + \binom{m}{2} = p$ for every $\boldsymbol{x}$ in the full design. By the equivalence theorem checking condition (2.9), the full design is D-optimal under the PWOi model (3.2). □

Proof of Lemma 2. We consider the case for component 1, and the remaining cases for components $2, \ldots, m-1$ can be argued in the same manner. For any run $\boldsymbol{x}$, it has q components so that $\{b_1(\boldsymbol{x}), \ldots, b_m(\boldsymbol{x})\}$ consists of a permutation of $1, \ldots, q$ and $(m-q)$ 0s. Then by the definition of $p_1(x)$ in (3.4) and the property (3.6), we have $\sum_{i=1}^m p_1(b_i(\boldsymbol{x})) = \sum_{k=1}^q p_1(k) = 0$. Therefore, $p_1(b_1(\boldsymbol{x}))\sum_{i=2}^m p_1(b_i(\boldsymbol{x})) = -p_1(b_1(\boldsymbol{x}))^2$. In addition, by (3.4) and (3.5), we have $\frac{1}{c_2}p_2(b_1(\boldsymbol{x})) = \frac{1}{c_1^2}p_1(b_1(\boldsymbol{x}))^2 - \frac{q^2-1}{12}S_1(\boldsymbol{x})$. Then,

$$\begin{aligned}\frac{q^2-1}{12}S_1(\boldsymbol{x}) &= \frac{1}{c_1^2}p_1(b_1(\boldsymbol{x}))^2 - \frac{1}{c_2}p_2(b_1(\boldsymbol{x})) \\ &= -\frac{1}{c_1^2}p_1(b_1(\boldsymbol{x}))\sum_{i=2}^m p_1(b_i(\boldsymbol{x})) - \frac{1}{c_2}p_2(b_1(\boldsymbol{x})).\end{aligned}$$

So the component indicator $S_1(\boldsymbol{x})$ is a linear combination of $p_2(b_1(\boldsymbol{x}))$ and interactions $p_1(b_1(\boldsymbol{x}))p_1(b_2(\boldsymbol{x})), \ldots, p_1(b_1(\boldsymbol{x}))p_1(b_m(\boldsymbol{x}))$, and the coefficients are independent of $\boldsymbol{x}$. The proof is complete. □

Proof of Lemma 3. We consider two cases: (i) $i = k$ and (ii) $i \neq k$.

(i) When $i = k$, for $j = 1$ or 2, the inner product between $p_j(b_i(\boldsymbol{x}))$ and $S_k(\boldsymbol{x})$ is the same as the inner product between $p_j(b_i(\boldsymbol{x}))$ and the intercept, which is 0 for the full design; therefore, $p_j(b_i(\boldsymbol{x}))$ is orthogonal to $S_i(\boldsymbol{x})$.

(ii) When $i \neq k$, $\sum_{\boldsymbol{x}\in\boldsymbol{\xi}} p_j(b_i(\boldsymbol{x}))S_k(\boldsymbol{x}) = \binom{m-2}{q-2}(q-1)!\sum_{x=1}^q p_j(x) = 0$ by (3.6), where $\boldsymbol{\xi}$ is the full design; therefore, $p_j(b_i(\boldsymbol{x}))$ and $S_k(\boldsymbol{x})$ are orthogonal for $j = 1, 2$ when $i \neq k$. □

Proof of Theorem 2. We only show the proofs for the QCP and QCPi models, as the proofs for the FOCP and FOCPi models are simpler due to the block diagonal forms of the information matrices explicitly given below.

(i) For any run $\boldsymbol{x}$ in the full design $\boldsymbol{\xi}$, let

$$f(\boldsymbol{x}) = (1, p_1(b_1(\boldsymbol{x})), \ldots, p_1(b_{m-1}(\boldsymbol{x})), p_2(b_1(\boldsymbol{x})), \ldots, p_2(b_{m-1}(\boldsymbol{x})))^T$$

and $M(\boldsymbol{\xi})$ be the information matrix for the full design $\boldsymbol{\xi}$ under the QCP model (2.7). By the definitions of $b_i(\boldsymbol{x})$, $p_1(x)$, $p_2(x)$, and the property (3.6), we have for any $\boldsymbol{x} \in \boldsymbol{\xi}$,

$$\sum_{i=1}^{m} p_j(b_i(\boldsymbol{x})) = 0 \text{ and } \sum_{i=1}^{m} p_j(b_i(\boldsymbol{x}))^2 = q \text{ for } j = 1, 2. \tag{S1.2}$$

It can be verified that

$$M(\boldsymbol{\xi}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{q}{m}(\frac{m}{m-1}\boldsymbol{I}_{m-1} + \frac{-1}{m-1}\boldsymbol{J}_{m-1}) & 0 \\ 0 & 0 & \frac{q}{m}(\frac{m}{m-1}\boldsymbol{I}_{m-1} + \frac{-1}{m-1}\boldsymbol{J}_{m-1}) \end{pmatrix},$$

and the inverse of $\frac{q}{m}(\frac{m}{m-1}\boldsymbol{I}_{m-1} + \frac{-1}{m-1}\boldsymbol{J}_{m-1})$ is $\frac{m-1}{q}(\boldsymbol{I}_{m-1} + \boldsymbol{J}_{m-1})$. Then

$$\begin{aligned} & f(\boldsymbol{x})^T M(\boldsymbol{\xi})^{-1} f(\boldsymbol{x}) \\ = \; & 1 + \frac{2(m-1)}{q} \sum_{k=1}^{m-1} p_1(b_k(\boldsymbol{x}))^2 + \frac{2(m-1)}{q} \sum_{k=1}^{m-2} \sum_{l>k}^{m-1} p_1(b_k(\boldsymbol{x})) p_1(b_l(\boldsymbol{x})) \\ + \; & \frac{2(m-1)}{q} \sum_{k=1}^{m-1} p_2(b_k(\boldsymbol{x}))^2 + \frac{2(m-1)}{q} \sum_{k=1}^{m-2} \sum_{l>k}^{m-1} p_2(b_k(\boldsymbol{x})) p_2(b_l(\boldsymbol{x})). \end{aligned}$$

By (S1.2), for $j = 1, 2$,

$$\sum_{k=1}^{m-1} p_j(b_k(\boldsymbol{x}))^2 + 2 \sum_{k=1}^{m-2} \sum_{l>k}^{m-1} p_j(b_k(\boldsymbol{x})) p_j(b_l(\boldsymbol{x})) = \left(\sum_{k=1}^{m-1} p_j(b_k(\boldsymbol{x})) \right)^2 = p_j(b_m(\boldsymbol{x}))^2,$$

and

$$
\begin{aligned}
f(\boldsymbol{x})^T M(\boldsymbol{\xi})^{-1} f(\boldsymbol{x}) &= 1 + \frac{(m-1)}{q}\sum_{k=1}^{m} p_1(b_k(\boldsymbol{x}))^2 + \frac{(m-1)}{q}\sum_{k=1}^{m} p_2(b_k(\boldsymbol{x}))^2 \\
&= 1 + (m-1) + (m-1) = 1 + 2(m-1).
\end{aligned}
$$

By the equivalence theorem checking condition (2.9), the full design $\boldsymbol{\xi}$ is D-optimal under the QCP model (2.7).

(ii) For any run $\boldsymbol{x}$ in the full design $\boldsymbol{\xi}$, now $f(\boldsymbol{x})$ in the QCPi model is $f(\boldsymbol{x}) = (f_1(\boldsymbol{x})^T, f_2(\boldsymbol{x})^T)^T$, where $f_1(\boldsymbol{x}) = (1, S_1(\boldsymbol{x}), \cdots, S_{m-1}(\boldsymbol{x}))^T$ is the vector of the intercept and component indicators, and $f_2(\boldsymbol{x}) = (p_1(b_1(\boldsymbol{x})), \ldots, p_1(b_{m-1}(\boldsymbol{x})), p_2(b_1(\boldsymbol{x})), \ldots, p_2(b_{m-1}(\boldsymbol{x})))^T$ is the vector of extended QCP factors. We have $f_1(\boldsymbol{x})^T M_1(\boldsymbol{\xi})^{-1} f_1(\boldsymbol{x}) = m$ by Lemma 4, and $f_2(\boldsymbol{x})^T M_2(\boldsymbol{\xi})^{-1} f_2(\boldsymbol{x})$ $= 2(m-1)$ from the proof of part (i). By Lemma 3, $f_1(\boldsymbol{x}) f_2(\boldsymbol{x})^T$ is a matrix of zeros, and therefore, $f(\boldsymbol{x})^T M(\boldsymbol{\xi})^{-1} f(\boldsymbol{x}) = f_1(\boldsymbol{x})^T M_1(\boldsymbol{\xi})^{-1} f_1(\boldsymbol{x}) + f_2(\boldsymbol{x})^T M_2(\boldsymbol{\xi})^{-1} f_2(\boldsymbol{x}) = m + 2(m-1) = 3m - 2$ for every $\boldsymbol{x} \in \boldsymbol{\xi}$. By the equivalence theorem checking condition (2.9), the full design $\boldsymbol{\xi}$ is D-optimal under the QCPi model (3.9). □

Proof of Theorem 3. Instead of working on the SOCP model (3.7), we consider the following model:

$$
\begin{aligned}
y = \beta_0 + \sum_{i=1}^{m} p_1(b_i(\boldsymbol{x}))\beta_i + \sum_{i=1}^{m} p_2(b_i(\boldsymbol{x}))\beta_{ii} \\
+ \sum_{1 \le i < j \le m} p_1(b_i(\boldsymbol{x})) p_1(b_j(\boldsymbol{x}))\beta_{ij} + \epsilon,
\end{aligned} \tag{S1.3}
$$

where the position variables are exchangeable, that is, the model remains the same when we exchange component labels. The models (S1.3) and

(3.7) are equivalent because the columns in both models span the same linear space under the full design.

Let π be a permutation of the m components $\{1, \ldots, m\}$. For $b = (b_1, \ldots, b_m)$, let $\pi(b) = (b_{\pi(1)}, \ldots, b_{\pi(m)})$ and $g(b) = (1, p_1(b_1), \ldots, p_1(b_m), p_2(b_1), \ldots, p_2(b_m), p_1(b_1)p_1(b_2), \ldots, p_1(b_{m-1})p_1(b_m))^T$. It is easy to see that there exists a permutation matrix $\boldsymbol{R}_\pi$ such that $g(\pi(b))^T = g(b)^T \boldsymbol{R}_\pi$ as the position variables in the model (S1.3) are exchangeable.

For a run $\boldsymbol{x}$ in the full design $\boldsymbol{\xi}$, let $b(\boldsymbol{x}) = (b_1(\boldsymbol{x}), \ldots, b_m(\boldsymbol{x}))$ and $f(\boldsymbol{x}) = g(b(\boldsymbol{x}))$, where $b_i(\boldsymbol{x})$ is defined in (3.3). Let $\pi(b(\boldsymbol{x})) = (b_{\pi(1)}(\boldsymbol{x}), \ldots, b_{\pi(m)}(\boldsymbol{x}))$. For any $\boldsymbol{x} \in \boldsymbol{\xi}$, $b(\boldsymbol{x})$ is a row of the position matrix, so is $\pi(b(\boldsymbol{x}))$. Let $\boldsymbol{Z}$ be the model matrix of the full design under the exchangeable model (S1.3). Then

$$\boldsymbol{Z}^T\boldsymbol{Z} = \sum_{\boldsymbol{x}\in\boldsymbol{\xi}} f(\boldsymbol{x})f(\boldsymbol{x})^T = \sum_{\boldsymbol{x}\in\boldsymbol{\xi}} g(b(\boldsymbol{x}))g(b(\boldsymbol{x}))^T = \sum_{\boldsymbol{x}\in\boldsymbol{\xi}} g(\pi(b(\boldsymbol{x})))g(\pi(b(\boldsymbol{x})))^T$$

because both $\{b(\boldsymbol{x}) : \boldsymbol{x} \in \boldsymbol{\xi}\}$ and $\{\pi(b(\boldsymbol{x})) : \boldsymbol{x} \in \boldsymbol{\xi}\}$ represent the same set of position vectors of the full design. Because $g(\pi(b(\boldsymbol{x})))^T = g(b(\boldsymbol{x}))^T\boldsymbol{R}_\pi$ and $g(\pi(b(\boldsymbol{x})))g(\pi(b(\boldsymbol{x})))^T = \boldsymbol{R}_\pi^T g(b(\boldsymbol{x}))g(b(\boldsymbol{x}))^T\boldsymbol{R}_\pi$, we have

$$\boldsymbol{Z}^T\boldsymbol{Z} = \sum_{\boldsymbol{x}\in\boldsymbol{\xi}} \boldsymbol{R}_\pi^T g(b(\boldsymbol{x}))g(b(\boldsymbol{x}))^T\boldsymbol{R}_\pi = \boldsymbol{R}_\pi^T(\boldsymbol{Z}^T\boldsymbol{Z})\boldsymbol{R}_\pi.$$

Because $\boldsymbol{R}_\pi$ is a permutation matrix, we have $\boldsymbol{R}_\pi^{-1} = \boldsymbol{R}_\pi^T$ and

$$(\boldsymbol{Z}^T\boldsymbol{Z})^- = \boldsymbol{R}_\pi^T(\boldsymbol{Z}^T\boldsymbol{Z})^-\boldsymbol{R}_\pi,$$

where $(\boldsymbol{Z}^T\boldsymbol{Z})^-$ is the Moore-Penrose generalized inverse. We use the generalized inverse here because the model (S1.3) is over-parameterized. Then

$$\begin{aligned} & g(\pi(b(\boldsymbol{x})))^T(\boldsymbol{Z}^T\boldsymbol{Z})^- g(\pi(b(\boldsymbol{x}))) \\ = \;& g(b(\boldsymbol{x}))^T\boldsymbol{R}_\pi(\boldsymbol{Z}^T\boldsymbol{Z})^-\boldsymbol{R}_\pi^T g(b(\boldsymbol{x})) \\ = \;& g(b(\boldsymbol{x}))^T(\boldsymbol{Z}^T\boldsymbol{Z})^- g(b(\boldsymbol{x})), \end{aligned}$$

which holds for all $\boldsymbol{x} \in \boldsymbol{\xi}$ and any permutation π.

For any $\boldsymbol{x}, \boldsymbol{x}' \in \boldsymbol{\xi}$, there exists at least one permutation π of $\{1, \ldots, m\}$ such that $b(\boldsymbol{x}') = \pi(b(\boldsymbol{x}))$. Then

$$\begin{aligned} & f(\boldsymbol{x}')^T(\boldsymbol{Z}^T\boldsymbol{Z})^- f(\boldsymbol{x}') \\ = \;& g(b(\boldsymbol{x}'))^T(\boldsymbol{Z}^T\boldsymbol{Z})^- g(b(\boldsymbol{x}')) = g(\pi(b(\boldsymbol{x})))^T(\boldsymbol{Z}^T\boldsymbol{Z})^- g(\pi(b(\boldsymbol{x}))) \\ = \;& g(b(\boldsymbol{x}))^T\boldsymbol{R}_\pi(\boldsymbol{Z}^T\boldsymbol{Z})^-\boldsymbol{R}_\pi^T g(b(\boldsymbol{x})) = g(b(\boldsymbol{x}))^T(\boldsymbol{Z}^T\boldsymbol{Z})^- g(b(\boldsymbol{x})) \\ = \;& f(\boldsymbol{x})^T(\boldsymbol{Z}^T\boldsymbol{Z})^- f(\boldsymbol{x}). \end{aligned}$$

Thus, the diagonal elements of the projection matrix $\boldsymbol{P} = \boldsymbol{Z}(\boldsymbol{Z}^T\boldsymbol{Z})^-\boldsymbol{Z}^T$ are a constant. As the projection matrix $\boldsymbol{P}$ is idempotent, its trace is equal to its rank, $p = (m+4)(m-1)/2$, which is the number of parameters in the identifiable model (3.7). So all of the diagonal elements of $\boldsymbol{P}$ are equal to p/n, where $n = \binom{m}{q}q!$ is the number of runs of the full design.

Let $\boldsymbol{X}$ and $M(\boldsymbol{\xi}) = \boldsymbol{X}^T\boldsymbol{X}/n$ be the model matrix and the information matrix of the full design under the identifiable model (3.7), respectively. Because the identifiable model (3.7) is equivalent to the exchangeable model (S1.3), the hat matrix $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$ is the same as the projection matrix $\boldsymbol{P}$, so all of its diagonal elements are p/n. Then for any $\boldsymbol{x} \in \boldsymbol{\xi}$,

$$f(\boldsymbol{x})^T M(\boldsymbol{\xi})^{-1} f(\boldsymbol{x}) = n f(\boldsymbol{x})^T(\boldsymbol{X}^T\boldsymbol{X})^{-1} f(\boldsymbol{x}) = n(p/n) = p.$$

By the equivalence theorem checking condition (2.9), the full design is D-optimal under the SOCP model (3.7). This completes the proof. □

Proof of Theorem 4. It can be verified that a COA(n, q, m) and the full design have the same information matrix, which is given in the proof of Theorem 2 under the QCP model; therefore, every COA(n, q, m) is D-optimal under the QCP model. By the same argument, every COA(n, q, m) is D-optimal under the FOCP, FOCPi, QCPi, and NCP models. □

Proof of Theorem 7. Stokes and Xu (2024) has shown the D-optimality of the design under the PWO model. To prove the D-optimality under the PWOi model, as done in the proof of Theorem 1, we need to verify the following three statements:

(i) The PWO factor $Z_{i,j}(\boldsymbol{x})$ and the component indicator $S_k(\boldsymbol{x})$ are orthogonal;

(ii) The information matrix corresponding to the component indicators is the same as that of the full design;

(iii) The information matrix corresponding to the PWO factors is the same as that of the full design.

First, $\mathbf{S}^{\text{PWO}}_{n,m,q}$ consists of $\binom{m}{q}$ q-component OofA-OAs, each of $n_1 = 12\lceil(\binom{q}{2}+1)/12\rceil$ runs and components $(i_1, \ldots, i_q)$. If both components i and j appear in an OofA-OA, there are $n_1/2$ runs with $Z_{i,j}(\boldsymbol{x}) = 1$ and the other $n_1/2$ runs with $Z_{i,j}(\boldsymbol{x}) = -1$ because the OofA-OA is balanced. Then regardless whether component k appears in that OofA-OA or not, the inner product between $Z_{i,j}(\boldsymbol{x})$ and $S_k(\boldsymbol{x})$ is 0 for that OofA-OA. If either component i or j does not appear in an OofA-OA, $Z_{i,j}(\boldsymbol{x}) = 0$; therefore, the inner product between $Z_{i,j}(\boldsymbol{x})$ and $S_k(\boldsymbol{x})$ is still 0 for that OofA-OA. This proves that the PWO factor $Z_{i,j}(\boldsymbol{x})$ and the component indicator $S_k(\boldsymbol{x})$ are orthogonal; the first statement holds. Second, because each of the $\binom{m}{q}$ q-component combinations $(i_1, \ldots, i_q)$ appears equally often (n_1 times each), the inner product between any two component indicators is $q(q-1)/[m(m-1)]$, which is the same as the full design under the component-indicator model in Lemma 4; therefore, the second statement holds. The third statement follows from the proof of Theorem 6 of Stokes and Xu (2024). The proof is complete.

□