# SUPPLEMENT TO "SPACE-FILLING DESIGNS WITH KRONECKER PRODUCT STRUCTURES UNDER KERNEL-BASED CRITERIA"

Ruonan Zheng, Xinran Zhang, Jian-Feng Yang, Min-Qian Liu*

*Nankai University*

This supplementary material discusses the corresponding results about extending the mean squared correlation criterion and distance variance criterion to Equation (2.1) respectively, provides the proofs of some theoretical results, some additional comparisons and simulations, and a large table.

# S1 Discussion About Two Criteria

When extending the mean squared correlation criterion and distance variance criterion to the form of Equation (2.1), our conclusion is that they are not suitable for discussion within the framework of this paper.

## S1.1 Mean squared correlation criterion

Correlation is often used for evaluating the space-filling property of a design. Let $\rho_{ave}$ be the mean squared correlation of a design $D$, that is, $\rho_{ave}(D) = \frac{2}{m(m-1)} \sum_{i<j} \rho_{ij}^2$, where $\rho_{ij}$ is the correlation of the $i$th and $j$th columns of $D$. It is obvious that $D$ is orthogonal if and only if $\rho_{ave}(D) = 0$. Designs with smaller values of $\rho_{ave}$ are considered

*Corresponding author: Min-Qian Liu, NITFID, LPMC & KLMDASR, School of Statistics and Data Science, Nankai University, Tianjin 300071, China. E-mail: mqliu@nankai.edu.cn.

as more space-filling designs.

The mean squared correlation criterion is an intuitive criterion, as orthogonality can be seen as the stepping stone of space-filling property. Although mean squared correlation criterion seems unable to be written in the form of Equation (2.1), Chen and Tang (2022) proves that it can be represented as the mean of $\phi$ in any two columns. The specific conclusion is as follows. Let $f(x,y)=xy$, then we have

$$\rho_{ave}(D)=\frac{1}{m(m-1)}\sum_{u\neq v} q\left(D_{uv}\right),$$

where $D_{uv}$ is an $N\times 2$ matrix formed by the $u$th and $v$th columns of $D$ and

$$q\left(D_{uv}\right)=\frac{144}{N^2(s^2-1)^2}\sum_{i=1}^{N}\sum_{j=1}^{N} f\left(x_{iu},x_{ju}\right) f\left(x_{iv},x_{jv}\right).$$

In fact, this is equivalent to the result in Equation (2.1) with $m=2$ and $f(x,y)=xy$. This inspires us to directly take $f(x,y)=xy$ in Equation (2.1) and we call the resulting criterion generalized mean squared correlation criterion. It is obvious that such a $f(x,y)$ is symmetric both in the variables and about the origin. The fact that $f(x,y)$ is a kernel function on $\mathcal{X}$ can also be directly inferred from the following equation

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j f\left(x_i,x_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j x_i x_j=\left(\sum_{i=1}^{n} c_i x_i\right)^2\geq 0.$$

Thus, we have the following two corollaries.

**Corollary S1.** *Let $f(x,y)=xy$, then we have $C(D,g,q)\geq 0$ for any $D$, $q$ and $g=0,1,\ldots,q$.*

**Corollary S2.** *For $D_0$ defined in Equation (3.1), we have*

$$\phi(D_0)=\begin{cases} 1/N^2\sum_{i=1}^{N}\sum_{j=1}^{N}\prod_{k=1}^{m} x_{ik}^t x_{jk}^t, \text{ if } m \text{ is even;}\\ 1/N^2\sum_{i=1}^{N}\sum_{j=1}^{N}\prod_{k=1}^{m} x_{ik}^t x_{jk}^t A_t(H), \text{ if } m \text{ is odd.}\end{cases}$$

According to Corollary S2, it is surprising that when $m$ is even, the value of $\phi(D_0)$ is independent of $H$. For example, a design that repeats a point twice and a design composed of a point and its mirror image have the same value when $m$ is even. This goes against common sense. Therefore, the generalized mean squared correlation criterion is not pertinent for discussion within the context of this paper.

## S1.2 Distance variance criterion

Similar to the mean squared correlation criterion, the distance variance criterion appears to be inexpressible in the form of Equation (2.1). However, Chen and Tang (2022) also proved that it can be represented as the mean of $\phi$ in any two columns. The specific conclusion is as follows. Let $f(x, y) = |x - y|^p$, then we have

$$V(D) = \frac{1}{m(m-1)} \sum_{u \neq v} q\left(D_{uv}\right),$$

where $D_{uv}$ is an $N \times 2$ matrix formed by the $u$th and $v$th columns of $D$ and

$$\begin{aligned} q\left(D_{uv}\right) =& \frac{Nm \sum_{x=1}^{s} \sum_{y=1}^{s} |x-y|^{2p}}{s^2(N-1)} - \left(\frac{Nm \sum_{x=1}^{s} \sum_{y=1}^{s} |x-y|^{p}}{s^2(N-1)}\right)^2 \\ &+ \frac{m(m-1)}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} f\left(x_{iu}, x_{ju}\right) f\left(x_{iv}, x_{jv}\right). \end{aligned}$$

This inspires us to directly take $f(x, y) = |x - y|^p$ in Equation (2.1). However, it is not appropriate to discuss this criterion within the scope of this paper, and we explain it from two aspects.

Firstly, $f(x, y) = |x - y|^p$ is not a kernel function on $\mathcal{X}$. For example, if we take $(w_1, \ldots, w_s) = (1, -1, 0, \ldots, 0)$, then we have $\sum_{i=1}^{s} \sum_{j=1}^{s} w_i w_j f\left(x_i, x_j\right) = -2 < 0$ where $x_i = i - (s+1)/2$. Secondly, when $D$ is $N$ repetitions of a point, $\phi(D)$ reaches its minimum value of zero. This obviously does not conform to the intuitive understanding of a space-filling criterion. Therefore, it is necessary to add the requirement of $D$ being a U-type design when using the criterion by taking $f(x,y) = |x-y|^p$ in Equation (2.1).

## S2 Proofs of Theoretical Results

To prove Theorem 1, we need the following lemma from Tang, Xu and Lin (2012). One difference is that the following lemma replaces $z > 1$ with any real number $z$.

**Lemma S1.** *For two rows $\boldsymbol{h}_i$ and $\boldsymbol{h}_j$ in $H$, let $\tilde{\delta}_{ij}$ denote the number of places where they take the same value. Then for any real number $z$, we have*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} z^{\tilde{\delta}_{ij}} = \frac{n^2}{2^t}\sum_{i=0}^{q}(z-1)^i(z+1)^{q-i}A_i(H).$$

*Proof of Theorem 1.* We first note that

$$\begin{aligned}F(\boldsymbol{h}_p \otimes \boldsymbol{x}_i, \boldsymbol{h}_k \otimes \boldsymbol{x}_j) &= \prod_{c=1}^{q}\prod_{d=1}^{m} f(h_{pc}x_{id}, h_{kc}x_{jd}) = \prod_{d=1}^{m} f(x_{id}, x_{jd})^{\tilde{\delta}_{pk}} f(x_{id}, -x_{jd})^{q-\tilde{\delta}_{pk}} \\ &= [F(\boldsymbol{x}_i, -\boldsymbol{x}_j)]^q \left[\frac{F(\boldsymbol{x}_i, \boldsymbol{x}_j)}{F(\boldsymbol{x}_i, -\boldsymbol{x}_j)}\right]^{\tilde{\delta}_{pk}}.\end{aligned}$$

According to Lemma S1, we have

$$\begin{aligned}&\phi(H \otimes D) \\ =&\frac{1}{N^2n^2}\sum_{p=1}^{n}\sum_{i=1}^{N}\sum_{k=1}^{n}\sum_{j=1}^{N}[F(\boldsymbol{x}_i, -\boldsymbol{x}_j)]^q \left[\frac{F(\boldsymbol{x}_i, \boldsymbol{x}_j)}{F(\boldsymbol{x}_i, -\boldsymbol{x}_j)}\right]^{\tilde{\delta}_{pk}} \\ =&\frac{1}{N^2n^2}\sum_{i=1}^{N}\sum_{j=1}^{N}[F(\boldsymbol{x}_i, -\boldsymbol{x}_j)]^q \sum_{p=1}^{n}\sum_{k=1}^{n}\left[\frac{F(\boldsymbol{x}_i, \boldsymbol{x}_j)}{F(\boldsymbol{x}_i, -\boldsymbol{x}_j)}\right]^{\tilde{\delta}_{pk}} \\ =&\frac{1}{N^2n^2}\sum_{i=1}^{N}\sum_{j=1}^{N}[F(\boldsymbol{x}_i, -\boldsymbol{x}_j)]^q \frac{n^2}{2^q}\left[1+\frac{F(\boldsymbol{x}_i, \boldsymbol{x}_j)}{F(\boldsymbol{x}_i, -\boldsymbol{x}_j)}\right]^q \sum_{g=0}^{q}\left[\frac{\dfrac{F(\boldsymbol{x}_i, \boldsymbol{x}_j)}{F(\boldsymbol{x}_i, -\boldsymbol{x}_j)}-1}{\dfrac{F(\boldsymbol{x}_i, \boldsymbol{x}_j)}{F(\boldsymbol{x}_i, -\boldsymbol{x}_j)}+1}\right]^g A_g(H) \\ =&\sum_{g=0}^{q} C(D,g,q)A_g(H).\end{aligned}$$

This completes the proof. □

To prove Theorem 2, we need the following definition and lemma. For a symmetric function $f : \mathcal{X} \times \mathcal{X} \to R$, which is symmetric in the variables, let the $s \times s$ matrix $(f(x_i, x_j))_{1\le i,j\le s}$ be denoted as $M(f)$, where $x_i, x_j \in \mathcal{X}$. Let $\mathcal{X}^m$ be the set of all $s^m$ $s$-level points in the $m$-dimensional space. Similarly, for a symmetric function $F : \mathcal{X}^m \times \mathcal{X}^m \to R$, let the $s^m \times s^m$ matrix $(F(\boldsymbol{x}_i, \boldsymbol{x}_j))_{1\le i,j\le s^m}$ be denoted as $M(F)$, where $\boldsymbol{x}_i, \boldsymbol{x}_j \in \mathcal{X}^m$. Lemma S2 is from Chapter 3 in Van Den Berg, Christensen and Ressel (2012).

**Lemma S2.** *A symmetric function $f : \mathcal{X} \times \mathcal{X} \to R$ is a kernel function if and only if $M(f)$ is positive semidefinite. Similarly, a symmetric function $F : \mathcal{X}^m \times \mathcal{X}^m \to R$ is a kernel function if and only if $M(F)$ is positive semidefinite.*

*Proof of Theorem 2.* Given that the product of kernel functions is still a kernel function, our objective is to demonstrate that both $F(\boldsymbol{x}, \boldsymbol{y}) - F(\boldsymbol{x}, -\boldsymbol{y})$ and $F(\boldsymbol{x}, \boldsymbol{y}) + F(\boldsymbol{x}, -\boldsymbol{y})$ are kernel functions. As $f(x, y)$ is a kernel function, then $M(f)$ is positive semidefinite, as per Lemma S2. Consequently, by noting that $F(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^m f(x_i, y_i)$, we have $M(F) = \otimes^m M(f)$. Because $M(f)$ is positive semidefinite, $M(F)$ is also positive semidefinite.

Let $I_k$ be the $k \times k$ identity matrix and $\tilde{I}_k$ be a $k \times k$ square matrix where the elements on the anti-diagonal are 1, and the rest of the elements are 0. Then we have $M(F(\boldsymbol{x}, \boldsymbol{y}) - F(\boldsymbol{x}, -\boldsymbol{y})) = M(F)(I_{s^m} - \tilde{I}_{s^m})$ and $M(F(\boldsymbol{x}, \boldsymbol{y}) + F(\boldsymbol{x}, -\boldsymbol{y})) = M(F)(I_{s^m} + \tilde{I}_{s^m})$, respectively.

We first prove that $I_{s^m} - \tilde{I}_{s^m}$ and $I_{s^m} + \tilde{I}_{s^m}$ are positive semidefinite. Actually, we have

$$(I_{s^m} - \tilde{I}_{s^m})^2 = I_{s^m} - 2\tilde{I}_{s^m} + \tilde{I}_{s^m}^2 = 2(I_{s^m} - \tilde{I}_{s^m}),$$

$$(I_{s^m} + \tilde{I}_{s^m})^2 = I_{s^m} + 2\tilde{I}_{s^m} + \tilde{I}_{s^m}^2 = 2(I_{s^m} + \tilde{I}_{s^m}).$$

So the eigenvalues of $I_{s^m} - \tilde{I}_{s^m}$ and $I_{s^m} + \tilde{I}_{s^m}$ are only 0 or 2, which means that they are positive semidefinite.

Then we prove $M(F)(I_{s^m} - \tilde{I}_{s^m}) = (I_{s^m} - \tilde{I}_{s^m})M(F)$. The proof of $M(F)(I_{s^m} + \tilde{I}_{s^m}) = (I_{s^m} + \tilde{I}_{s^m})M(F)$ is similar. Let $I_{s^m} - \tilde{I}_{s^m} = (q_{ij})_{1\le i,j\le s^m}$ and $M(F) = (a_{ij})_{1\le i,j\le s^m}$. According to the properties of $I_{s^m} - \tilde{I}_{s^m}$, we have $q_{ij} = 1$ for $i = j$, $q_{ij} = -1$ for $i = s^m + 1 - j$ and $q_{ij} = 0$ for other $i$. Then the $(i, j)$th elements of $M(F)(I_{s^m} - \tilde{I}_{s^m})$ and $(I_{s^m} - \tilde{I}_{s^m})M(F)$ are $a_{ij} - a_{i,s^m+1-j}$ and $a_{ij} - a_{s^m+1-i,j}$ respectively. Because $f(x, y)$ is symmetric both in the variables and about the origin, $M(F)$ is symmetric with respect to the diagonal and anti-diagonal, which means that $a_{ij} - a_{i,s^m+1-j} = a_{ij} - a_{s^m+1-i,j}$. So we have $M(F)(I_{s^m} - \tilde{I}_{s^m}) = (I_{s^m} - \tilde{I}_{s^m})M(F)$. According to the algebraic knowledge, for two symmetric positive semidefinite matrices $A$ and $B$, if $AB = BA$, then $AB$ is positive semidefinite. Thus, $M(F)(I_{s^m} - \tilde{I}_{s^m})$ and $M(F)(I_{s^m} + \tilde{I}_{s^m})$ are positive semidefinite. Then $F(\boldsymbol{x}, \boldsymbol{y}) - F(\boldsymbol{x}, -\boldsymbol{y})$ and $F(\boldsymbol{x}, \boldsymbol{y}) + F(\boldsymbol{x}, -\boldsymbol{y})$ are kernel functions according to Lemma S2. This completes the proof. □

*Proof of Corollary 2.* In fact, $D_1$ can be regarded as the design obtained by taking the Kronecker product of $H = (1, -1)^T$ and $D$. Based on the expressions of $L_2$-type discrepancies and simple calculations, $\mathrm{Disc}^2(D_1, F)$ and $\mathrm{Disc}^2(D, F)$ differ only in their double summation terms. In other words, we only need to prove $\phi(H \otimes D) \le \phi(D)$, where $H = (1, -1)^T$ and $A_1(H) = 0$.

According to Theorem 1, we have

$$\begin{aligned}
\phi(D) - \phi(H \otimes D) &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(F(\boldsymbol{x}_i, \boldsymbol{x}_j) - \frac{F(\boldsymbol{x}_i, \boldsymbol{x}_j) + F(\boldsymbol{x}_i, -\boldsymbol{x}_j)}{2}\right) \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{F(\boldsymbol{x}_i, \boldsymbol{x}_j) - F(\boldsymbol{x}_i, -\boldsymbol{x}_j)}{2} \\
&= C(D, 1, 1).
\end{aligned}$$

According to Theorem 2, it is obvious that $C(D,1,1)\geq 0$. This completes the proof. $\square$

*Proof of Corollary 4.* (a) When $z_l = z$, we have

$$\sum_{k=0}^{mq} A_k(D_0)z^k = \sum_{g=0}^{q} C(D,g,q)A_g(H).$$

So $A_1(D_0) = \sum_{g=0}^{q} \frac{\partial C(D,g,q)}{\partial z}|_{z=0}A_g(H)$. According to the expression of $C(D,g,q)$, we have

$$\begin{aligned}\frac{\partial C(D,g,q)}{\partial z} =& \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left[(q-g)\left(\frac{F(\boldsymbol{x}_i,\boldsymbol{x}_j)+F(\boldsymbol{x}_i,-\boldsymbol{x}_j)}{2}\right)^{q-g-1}\right.\\ &\times\left(\frac{\partial F(\boldsymbol{x}_i,\boldsymbol{x}_j)}{\partial z}+\frac{\partial F(\boldsymbol{x}_i,-\boldsymbol{x}_j)}{\partial z}\right)\times\left(\frac{F(\boldsymbol{x}_i,\boldsymbol{x}_j)-F(\boldsymbol{x}_i,-\boldsymbol{x}_j)}{2}\right)^{g}\\ &+g\left(\frac{F(\boldsymbol{x}_i,\boldsymbol{x}_j)+F(\boldsymbol{x}_i,-\boldsymbol{x}_j)}{2}\right)^{q-g}\left(\frac{\partial F(\boldsymbol{x}_i,\boldsymbol{x}_j)}{\partial z}-\frac{\partial F(\boldsymbol{x}_i,-\boldsymbol{x}_j)}{\partial z}\right)\\ &\left.\times\left(\frac{F(\boldsymbol{x}_i,\boldsymbol{x}_j)-F(\boldsymbol{x}_i,-\boldsymbol{x}_j)}{2}\right)^{g-1}\right],\ \text{and} \qquad (S1)\\ \frac{\partial F(\boldsymbol{x}_i,\boldsymbol{x}_j)}{\partial z} =& \sum_{k'=1}^{m}\left(\sum_{l=1}^{s-1}p_l(x_{ik'})p_l(x_{jk'})\right)\prod_{k\neq k'}\left(1+\sum_{l=1}^{s-1}p_l(x_{ik})p_l(x_{jk})z\right). \qquad (S2)\end{aligned}$$

For $g\geq 2$, it is obvious that $\frac{\partial C(D,g,q)}{\partial z}|_{z=0}=0$. By substituting Equation (S2) into (S1) and let $z=0$, we can obtain

$$\begin{aligned}\frac{\partial C(D,1,q)}{\partial z}|_{z=0} &= \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{m}\sum_{l=1}^{s-1}\left(p_l(x_{ik})p_l(x_{jk})-p_l(x_{ik})p_l(-x_{jk})\right),\ \text{and}\\ \frac{\partial C(D,0,q)}{\partial z}|_{z=0} &= \frac{t}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{m}\sum_{l=1}^{s-1}\left(p_l(x_{ik})p_l(x_{jk})+p_l(x_{ik})p_l(-x_{jk})\right).\end{aligned}$$

Then $A_1(D_0)$ can be obtained immediately.

(b) When $z_l = z^l$, similar to (a), we have

$$\sum_{k=0}^{mq(s-1)} \beta_k(D_0)z^k = \sum_{g=0}^{q} C(D,g,q)A_g(H).$$

So $\beta_1(D_0) = \sum_{g=0}^{q} \frac{\partial C(D,g,q)}{\partial z}|_{z=0} A_g(H)$. For $g \geq 2$, it is obvious that $\frac{\partial C(D,g,q)}{\partial z}|_{z=0} = 0$. According to Equation (S1) and

$$\frac{\partial F(\boldsymbol{x}_i, \boldsymbol{x}_j)}{\partial z}|_{z=0} = \sum_{k=1}^{m} p_1(x_{ik}) p_1(x_{jk}),$$

we can obtain

$$\begin{aligned} \frac{\partial C(D,1,q)}{\partial z}|_{z=0} &= \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{m} p_1(x_{ik}) p_1(x_{jk}) \\ &= \frac{1}{N^2} \sum_{k=1}^{m} \left( \sum_{i=1}^{N} p_1(x_{ik}) \right)^2 = \beta_1(D), \text{ and} \\ \frac{\partial C(D,0,q)}{\partial z}|_{z=0} &= \frac{q}{2N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{m} (p_1(x_{ik}) p_1(x_{jk}) + p_1(x_{ik}) p_1(-x_{jk})) = 0. \end{aligned}$$

Then $\beta_1(D_0)$ can be obtained immediately. This completes the proof. □

Before proving Theorem 3, we need the following definition and lemma from Bhatia and Jain (2015).

**Definition S1.** Let $A = (a_{ij})$ be an $n \times n$ real symmetric matrix. It is said to be conditionally negative definite (CND for short), if $x^T A x \leq 0$ for all $x$ in the $(n-1)$ dimensional space $\mathcal{H} = \left\{ x = (x_1, \ldots, x_n)^T : \sum_{j=1}^{n} x_j = 0 \right\}$.

**Lemma S3.** *If $A = (a_{ij})_{n \times n}$ is CND, then its Hadamard inverse matrix $(1/a_{ij})_{n \times n}$ is positive semidefinite.*

*Proof of Theorem 3.* Let $\tilde{f}(x,y) = (\gamma + |x-y|)^l$. According to Lemmas S2 and S3, we only need to prove that $M(\tilde{f})$ is CND. Let $J_s$ be an $s \times s$ matrix with all elements unity, and let $x_1, \ldots, x_s$ represent all the elements in $\mathcal{X}$.

(a) When $l = 1$, according to Corollary 2.3 in Bhatia and Jain (2015), $M(\tilde{f})$ is CND.

(b) When $l = 2$, we have $M(\tilde{f}) = \gamma^2 J_s + 2\gamma M(\tilde{f}_1) + \mathrm{diag}(x_1^2, \ldots, x_s^2)J_s - 2\boldsymbol{x}\boldsymbol{x}^T + J_s\mathrm{diag}(x_1^2, \ldots, x_s^2)$, where $\tilde{f}_1 = |x - y|$ and $\boldsymbol{x} = (x_1, \ldots, x_s)^T$. It is obvious that $\gamma^2 J_s + \mathrm{diag}(x_1^2, \ldots, x_s^2)J_s + J_s\mathrm{diag}(x_1^2, \ldots, x_s^2)$ is CND and $\boldsymbol{x}\boldsymbol{x}^T$ is positive semidefinite. $M(\tilde{f}_1)$ is also CND according to (a). So $M(\tilde{f})$ is CND.

According to Lemma S3, $M(f)$ is positive semidefinite. Then $f(x, y)$ is a kernel function on $\mathcal{X}$ according to Lemma S2. □

Before proving Theorem 4, we need the following lemma. Lemma S4 is from Theorem 2.2 in Van Den Berg, Christensen and Ressel (2012).

**Lemma S4.** *If $A = (a_{ij})_{n\times n}$ is CND, then $(e^{-ta_{ij}})_{n\times n}$ is positive semidefinite for $t > 0$.*

*Proof of Theorem 4.* Let $f(x, y) = \rho^{d_p(x,y)}$. We only need to prove that $f(x, y)$ is a kernel function on $\mathcal{X}$. It is equivalent to proving that $M(f)$ is positive semidefinite according to Lemma S2. Next, we prove that $\det(M(f)) > 0$ holds for any $s \geq 2$, $\rho \leq 1/3$ and $p \geq 1$. Since $\rho^{d_p(x,y)}$ is a monotonically decreasing function with respect to $p$, we only need to consider the case when $p = 1$.

Let $M(f) = (m_{ij})_{s\times s}$. According to algebraic knowledge, if $m_{ii} > \sum_{j\neq i} |m_{ij}|$ for any $i = 1, 2, \ldots, s$, then we have $\det(M(f)) > 0$. Now we consider the following two cases.

(a) When $s$ is odd, the maximum value of $\sum_{j\neq i} |m_{ij}|$ is $2(\rho + \rho^2 + \cdots + \rho^{(s-1)/2})$. Because $2(\rho+\rho^2+\cdots+\rho^{(s-1)/2}) = 2\left(\rho - \rho^{(s+1)/2}\right)/(1-\rho)$ and $2\rho^{(s+1)/2} - 3\rho + 1 > 0$ for $\rho \leq 1/3$, we have $\sum_{j\neq i} |m_{ij}| < 1$ for any $i = 1, 2, \ldots, s$. Thus, we have $\det(M(f)) > 0$.

(b) When $s$ is even, the maximum value of $\sum_{j\neq i} |m_{ij}|$ is $2(\rho+\rho^2+\cdots+\rho^{s/2-1})+\rho^{s/2}$. Because $2(\rho + \rho^2 + \cdots + \rho^{s/2-1}) + \rho^{s/2} = \left(2\rho - \rho^{s/2} - \rho^{s/2+1}\right)/(1 - \rho)$ and $3\rho - \rho^{s/2} - \rho^{s/2+1} < 1$ for $\rho \leq 1/3$, we have $\sum_{j\neq i} |m_{ij}| < 1$ for any $i = 1, 2, \ldots, s$. Thus, we have $\det(M(f)) > 0$.

In other words, the determinant of arbitrary sequential principal minor of matrix $M(f)$ is greater than 0. Hence, $M(f)$ is positive definite. Further results can be obtained from Lemma S4 and Theorem 3. This completes the proof. $\square$

*Proof of Theorem 5.* When $H_t$ is a two-level full factorial design with $t$ factors, we have

$$\begin{aligned}\phi(D_0) = C(D,0,q) &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{F(\boldsymbol{x}_i,\boldsymbol{x}_j)+F(\boldsymbol{x}_i,-\boldsymbol{x}_j)}{2}\right)^q \\ &= \frac{1}{N^2}\sum_{i=1}^{N}\left(\frac{F(\boldsymbol{x}_i,\boldsymbol{x}_i)+F(\boldsymbol{x}_i,-\boldsymbol{x}_i)}{2}\right)^q + \frac{1}{N^2}\sum_{i\neq j}\left(\frac{F(\boldsymbol{x}_i,\boldsymbol{x}_j)+F(\boldsymbol{x}_i,-\boldsymbol{x}_j)}{2}\right)^q.\end{aligned}$$

According to the Holder's inequality, we have

$$\begin{aligned}\sum_{i=1}^{N}(F(\boldsymbol{x}_i,\boldsymbol{x}_i)+F(\boldsymbol{x}_i,-\boldsymbol{x}_i))^q\cdot N^{q-1} &\geq \left(\sum_{i=1}^{N}F(\boldsymbol{x}_i,\boldsymbol{x}_i)+\sum_{i=1}^{N}F(\boldsymbol{x}_i,-\boldsymbol{x}_i)\right)^q \\ &\geq \left(N\prod_{i=1}^{N}F(\boldsymbol{x}_i,\boldsymbol{x}_i)^{\frac{1}{N}}+N\prod_{i=1}^{N}F(\boldsymbol{x}_i,-\boldsymbol{x}_i)^{\frac{1}{N}}\right)^q\end{aligned}$$

and

$$\begin{aligned}&\sum_{i\neq j}(F(\boldsymbol{x}_i,\boldsymbol{x}_j)+F(\boldsymbol{x}_i,-\boldsymbol{x}_j))^q\cdot(N(N-1))^{q-1} \\ \geq&\left(\sum_{i\neq j}F(\boldsymbol{x}_i,\boldsymbol{x}_j)+\sum_{i\neq j}F(\boldsymbol{x}_i,-\boldsymbol{x}_j)\right)^q \\ \geq&(N(N-1))^q\left(\prod_{i\neq j}F(\boldsymbol{x}_i,\boldsymbol{x}_j)^{\frac{1}{N(N-1)}}+\prod_{i\neq j}F(\boldsymbol{x}_i,-\boldsymbol{x}_j)^{\frac{1}{N(N-1)}}\right)^q.\end{aligned}$$

The equality holds if all $F(\boldsymbol{x}_i,\boldsymbol{x}_j)$ for $i\neq j$, $i,j=1,\ldots,N$, all $F(\boldsymbol{x}_i,-\boldsymbol{x}_j)$ for $i\neq j$, $i,j=1,\ldots,N$, all $F(\boldsymbol{x}_i,\boldsymbol{x}_i)$ for $i=1,\ldots,N$, and all $F(\boldsymbol{x}_i,-\boldsymbol{x}_i)$ for $i=1,\ldots,N$ are equal, respectively. Because $D$ is a U-type design, $\prod_{i=1}^{N}F(\boldsymbol{x}_i,\boldsymbol{x}_i)$, $\prod_{i=1}^{N}F(\boldsymbol{x}_i,-\boldsymbol{x}_i)$, $\prod_{i\neq j}F(\boldsymbol{x}_i,\boldsymbol{x}_j)$, and $\prod_{i\neq j}F(\boldsymbol{x}_i,-\boldsymbol{x}_j)$ are all constants which only depend on $N$, $m$, $s$ and $f(x,y)$. Thus, by carrying out some tedious calculations, we can obtain the value of $LW_1$. This completes the proof. $\square$

*Proof of Theorem 7.* Since $D_0 = H \otimes D$, the $L_2$-distance between any two rows of $D_0$ has the following three situations: (a) $qd_2(\boldsymbol{x}_i, \boldsymbol{x}_j)$; (b) $4\delta(\boldsymbol{h}_p, \boldsymbol{h}_q)d_2(\boldsymbol{x}_i, 0)$; (c) $\delta(\boldsymbol{h}_p, \boldsymbol{h}_q)d_2(\boldsymbol{x}_i, -\boldsymbol{x}_j) + (q - \delta(\boldsymbol{h}_p, \boldsymbol{h}_q))d_2(\boldsymbol{x}_i, \boldsymbol{x}_j)$. The minimum distances for the first two situations are (a) $qd_2(D)$; (b) $4\delta(H)d_o(D)$. So we have $d_2(D_0) \leq 4\delta(H)d_o(D)$. Because $D_0$ is a U-type design, we have $d_o(D) \leq m(s^2 - 1)/12$.

We consider the following two cases.

*Case 1.* If $\delta(H) \leq q/2$, we have $d_2(D_0) \leq 4\delta(H)d_o(D) \leq 2q \cdot m(s^2 - 1)/12 = mq(s^2 - 1)/6$.

*Case 2.* If $\delta(H) > q/2$, for $qd_2(\boldsymbol{x}_i, \boldsymbol{x}_j)$, we have

$$qd_2(\boldsymbol{x}_i, \boldsymbol{x}_j) = q(d_2(\boldsymbol{x}_i, 0) + d_2(\boldsymbol{x}_j, 0)) - 2q\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle; \tag{S3}$$

for $(q - \delta(\boldsymbol{h}_p, \boldsymbol{h}_q))d_2(\boldsymbol{x}_i, \boldsymbol{x}_j) + \delta(\boldsymbol{h}_p, \boldsymbol{h}_q)d_2(\boldsymbol{x}_i, -\boldsymbol{x}_j)$, we have

$$\begin{aligned} &(q - \delta(\boldsymbol{h}_p, \boldsymbol{h}_q))d_2(\boldsymbol{x}_i, \boldsymbol{x}_j) + \delta(\boldsymbol{h}_p, \boldsymbol{h}_q)d_2(\boldsymbol{x}_i, -\boldsymbol{x}_j) \\ =&q(d_2(\boldsymbol{x}_i, 0) + d_2(\boldsymbol{x}_j, 0)) + (4\delta(\boldsymbol{h}_p, \boldsymbol{h}_q) - 2q)\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle. \end{aligned} \tag{S4}$$

Because $D_0$ is a U-type design, we have

$$\sum_{i \neq j}(d_2(\boldsymbol{x}_i, 0) + d_2(\boldsymbol{x}_j, 0)) = mN(N-1)(s^2 - 1)/6,$$

which means that there exist $i'$ and $j'$ such that

$$d_2(\boldsymbol{x}_{i'}, 0) + d_2(\boldsymbol{x}_{j'}, 0) \leq m(s^2 - 1)/6.$$

From Equations (S3) and (S4), and $\delta(H) > q/2$, we have $qd_2(\boldsymbol{x}_{i'}, \boldsymbol{x}_{j'}) \leq mq(s^2 - 1)/6$ if $\langle \boldsymbol{x}_{i'}, \boldsymbol{x}_{j'} \rangle \geq 0$ and $(q - \delta(\boldsymbol{h}_p, \boldsymbol{h}_q))d_2(\boldsymbol{x}_{i'}, \boldsymbol{x}_{j'}) + \delta(\boldsymbol{h}_p, \boldsymbol{h}_q)d_2(\boldsymbol{x}_{i'}, -\boldsymbol{x}_{j'}) \leq mq(s^2 - 1)/6$ if $\langle \boldsymbol{x}_{i'}, \boldsymbol{x}_{j'} \rangle < 0$. So we know that $d_2(D_0) \leq mq(s^2 - 1)/6$ immediately. This completes the proof. $\square$

*Proof of Proposition 1.* It is obvious that $G^{(k)}$ satisfies Conditions (i) and (iii), so we only prove that it satisfies Condition (ii). Let $\mathbf{1}$ be a vector of all ones with appropriate length. Denote the columns of $A^{(k)}$ and $H^{(k)}$ as $a_1^{(k)}, a_2^{(k)}, \ldots, a_{2^k}^{(k)}$ and $h_1^{(k)}, h_2^{(k)}, \ldots, h_{2^k}^{(k)}$, respectively. In addition, it is also obvious that $H^{(k)}$ is column orthogonal.

We next prove it in three steps. In the first step, we use mathematical induction to prove that the following two equations,

$$\mathbf{1}^T \left[ (a_i^{(k)} + a_j^{(k)}) \circ h_i^{(k)} \circ h_j^{(k)} \right] = 0, \text{ and} \tag{S5}$$

$$\mathbf{1}^T \left[ (a_i^{(k)} + a_j^{(k)}) \circ h_{2^k+1-i}^{(k)} \circ h_{2^k+1-j}^{(k)} \right] = 0, \tag{S6}$$

hold for any positive integer $k$ and $1 \le i \ne j \le 2^k$.

When $k = 1$, it is obvious that Equations (S5) and (S6) hold. Assume that when $k = n$, Equations (S5) and (S6) hold. When $k = n+1$, we consider the following three cases.

*Case 1.* For $1 \le i \ne j \le 2^n$, we have

$$\begin{aligned}
&\mathbf{1}^T \left[ (a_i^{(n+1)} + a_j^{(n+1)}) \circ h_i^{(n+1)} \circ h_j^{(n+1)} \right] \\
=&\mathbf{1}^T \left[ (a_i^{(n)} + a_j^{(n)}) \circ h_i^{(n)} \circ h_j^{(n)} \right] + \mathbf{1}^T \left[ (a_i^{(n)} + 2^n + a_j^{(n)} + 2^n) \circ h_i^{(n)} \circ h_j^{(n)} \right] = 0, \text{ and} \\
&\mathbf{1}^T \left[ (a_i^{(n+1)} + a_j^{(n+1)}) \circ h_{2^{n+1}+1-i}^{(n+1)} \circ h_{2^{n+1}+1-j}^{(n+1)} \right] \\
=&\mathbf{1}^T \left[ (a_i^{(n)} + a_j^{(n)}) \circ h_i^{(n)} \circ h_j^{(n)} \right] + \mathbf{1}^T \left[ (a_i^{(n)} + 2^n + a_j^{(n)} + 2^n) \circ -h_i^{(n)} \circ -h_j^{(n)} \right] = 0.
\end{aligned}$$

The last equality is obtained from the inductive assumption and the orthogonality of $H^{(n)}$.

Similarly to *Case 1*, for *Case 2.* $2^n < i \ne j \le 2^{n+1}$ and *Case 3.* $1 \le i \le 2^n < j \le 2^{n+1}$, we have

$$\mathbf{1}^T \left[ (a_i^{(n+1)} + a_j^{(n+1)}) \circ h_i^{(n+1)} \circ h_j^{(n+1)} \right] = 0, \text{ and}$$

$$\mathbf{1}^T\left[(a_i^{(n+1)}+a_j^{(n+1)})\circ h_{2^{n+1}+1-i}^{(n+1)}\circ h_{2^{n+1}+1-j}^{(n+1)}\right]=0.$$

In the second step, we use mathematical induction to prove that the following equation,

$$\mathbf{1}^T\left[(a_i^{(k)}-a_j^{(k)})\circ h_i^{(k)}\circ h_{2^k+1-j}^{(k)}\right]=0, \tag{S7}$$

holds for any positive integer $k$ and $1\leq i\neq j\leq 2^k$.

When $k=1$, it is obvious that Equation (S7) holds. Assume that when $k=n$, Equation (S7) holds. When $k=n+1$, we similarly consider the three cases discussed in the first step. Since the computational process is similar to that in the first step, the details are omitted here.

In the third step, we use mathematical induction to prove that the following two equations,

$$\mathbf{1}^T\left[a_i^{(k)}\circ a_j^{(k)}\circ h_i^{(k)}\circ h_j^{(k)}\right]=0,\ \text{and} \tag{S8}$$

$$\mathbf{1}^T\left[a_i^{(k)}\circ a_j^{(k)}\circ h_{2^k+1-i}^{(k)}\circ h_{2^k+1-j}^{(k)}\right]=0, \tag{S9}$$

hold for any positive integer $k$ and $1\leq i\neq j\leq 2^k$. Equation (S8) implies that $G^{(k)}$ satisfies Condition (ii).

When $k=1$, it is obvious that Equations (S8) and (S9) hold. Assume that when $k=n$, Equations (S8) and (S9) hold. When $k=n+1$, we consider the following three cases.

*Case 1.* For $1\leq i\neq j\leq 2^n$, we have

$$\begin{aligned}&\mathbf{1}^T\left[a_i^{(n+1)}\circ a_j^{(n+1)}\circ h_i^{(n+1)}\circ h_j^{(n+1)}\right]\\ =&\mathbf{1}^T\left[a_i^{(n)}\circ a_j^{(n)}\circ h_i^{(n)}\circ h_j^{(n)}\right]+\mathbf{1}^T\left[(a_i^{(n)}+2^n)\circ(a_j^{(n)}+2^n)\circ h_i^{(n)}\circ h_j^{(n)}\right]=0,\ \text{and}\\ &\mathbf{1}^T\left[a_i^{(n+1)}\circ a_j^{(n+1)}\circ h_{2^{n+1}+1-i}^{(n+1)}\circ h_{2^{n+1}+1-j}^{(n+1)}\right]\end{aligned}$$

$$=\mathbf{1}^T\left[a_i^{(n)}\circ a_j^{(n)}\circ h_i^{(n)}\circ h_j^{(n)}\right]+\mathbf{1}^T\left[(a_i^{(n)}+2^n)\circ(a_j^{(n)}+2^n)\circ(-h_i^{(n)})\circ(-h_j^{(n)})\right]=0.$$

The last equality is obtained from the inductive assumption, the orthogonality of $H^{(n)}$ and Equation (S5).

In a way similar to *Case 1*, for *Case 2.* $2^n < i \neq j \leq 2^{n+1}$ and *Case 3.* $1 \leq i \leq 2^n < j \leq 2^{n+1}$, we have

$$\mathbf{1}^T\left[a_i^{(n+1)}\circ a_j^{(n+1)}\circ h_i^{(n+1)}\circ h_j^{(n+1)}\right]=0,\ \text{and}$$
$$\mathbf{1}^T\left[a_i^{(n+1)}\circ a_j^{(n+1)}\circ h_{2^{n+1}+1-i}^{(n+1)}\circ h_{2^{n+1}+1-j}^{(n+1)}\right]=0.$$

This completes the proof. □

In order to prove Lemma 3, we first give a lemma from Rankin (1955).

**Lemma S5.** *For $N$ points $\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N$ in an $m$-dimensional space with $N \geq m+2$, there exists some $i \neq j$ such that $\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle \geq 0$.*

*Proof of Lemma 3.* For a U-type $D(N,s^m)$ with $N \geq m+2$, there are two situations.

*Case 1.* If $d_o(D)=m(s^2-1)/12$, since $N \geq m+2$, there definitely exist two points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ such that $\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle \geq 0$ according to Lemma S5. So we have

$$\begin{aligned}d_2(\boldsymbol{x}_i,\boldsymbol{x}_j)&=d_2(\boldsymbol{x}_i,0)+d_2(\boldsymbol{x}_j,0)-2\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle\\&=\frac{m(s^2-1)}{6}-2\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle \leq \frac{m(s^2-1)}{6},\end{aligned}$$

that is $d_2(D) \leq m(s^2-1)/6$.

*Case 2.* If there exists an $i$ such that $d_2(\boldsymbol{x}_i,0) < m(s^2-1)/12$, by noting that $m$ is even and $s=4$, we have

$$d_2(\boldsymbol{x}_i,0) \leq \frac{m(s^2-1)}{12}-2=\frac{5m}{4}-2.$$

Because

$$\begin{aligned}\sum_{j=1}^{N} d_2(\boldsymbol{x}_j, \boldsymbol{x}_i) &= \frac{Nm(s^2-1)}{12} + Nd_2(\boldsymbol{x}_i, 0)\\ &\leq \frac{Nm(s^2-1)}{12} + \frac{5Nm}{4} - 2N\\ &= \frac{5Nm}{2} - 2N,\end{aligned}$$

there exists some $i \neq j$, such that

$$d_2(\boldsymbol{x}_i, \boldsymbol{x}_j) \leq \frac{5m}{2}\frac{N}{N-1} - \frac{2N}{N-1},$$

which implies $d_2(\boldsymbol{x}_i, \boldsymbol{x}_j) \leq 5m/2 = m(s^2-1)/6$ since $N \geq 5m/4$. This completes the proof. □

# S3 More Comparisons and Simulations

In this section, we present some additional comparisons and simulations to demonstrate the effectiveness of the KTA algorithm.

## S3.1 Comparison under the maximum projection criterion

In this subsection, we compare the performance of designs generated by the KTA algorithm with those obtained by other methods under the maximum projection (MaxPro) criterion. Since the KTA method generates designs with repeated levels, we set $\gamma = 0.1$ and $l = 2$ in Equation (4.1). In addition, for the sake of fairness and to eliminate the interference of constant terms on the comparison of final results, we scale the levels of all designs to $[0, 1]$ and adopt the MaxPro criterion as follows

$$\psi(D) = \left\{\frac{1}{N(N-1)}\sum_{i\neq j}\prod_{k=1}^{m}\frac{1}{(0.1 + |x_{ik} - x_{jk}|)^2}\right\}^{1/m}. \tag{S10}$$

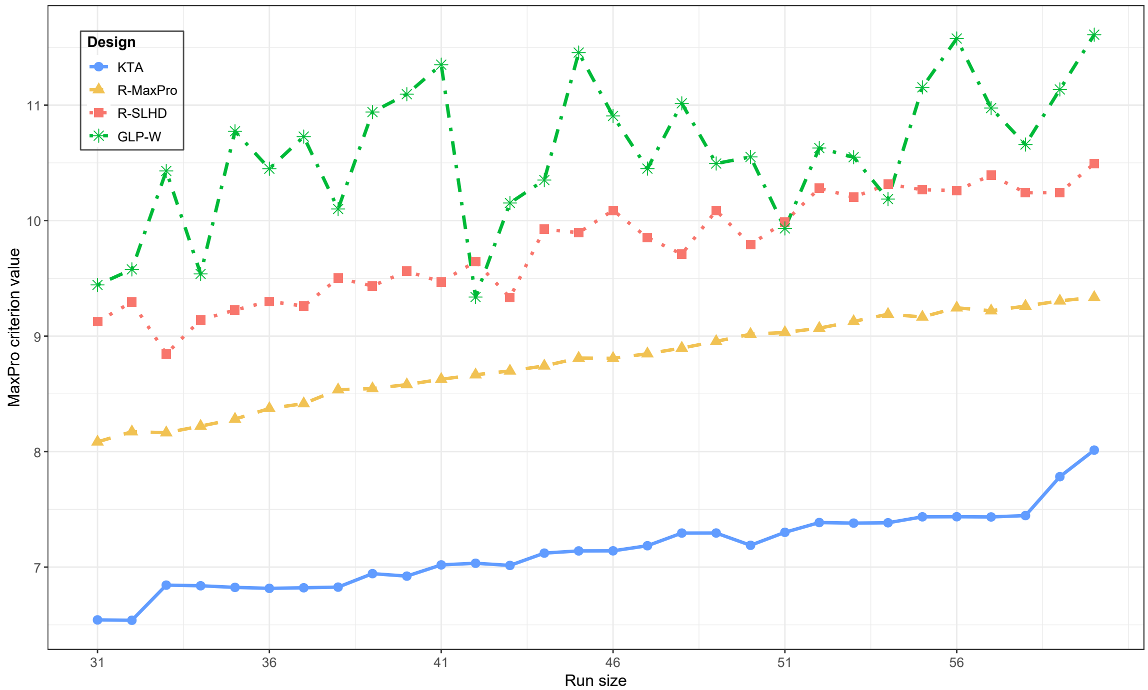


Figure S1: The $\psi$ values of the designs with 8 factors generated by different methods.

Figures S1 and S2 show the performance of different designs under the MaxPro criterion with the number of factors being 8 and 12, respectively, where the run size ranges from 31 to 60 and the number of levels is equal to the run size. KTA designs (KTA) are generated by the KTA algorithm with $q = 1$. "R-MaxPro" denotes the designs obtained by using the function MaxProLHD from the R package "MaxPro". Sliced LHDs (R-SLHD) are obtained by using the function maximinSLHD from the R package "SLHD". GLP-W denotes the designs constructed by selecting columns from the designs generated from Algorithm 1 in Ye, Yuan and Wang (2025), as the designs are obtained from the good lattice point method based on the Williams transformation. For the sake of fairness, the maximum number of iterations for the first three methods is set to $10^7$, with all other parameters at their default settings. For the "R-MaxPro"

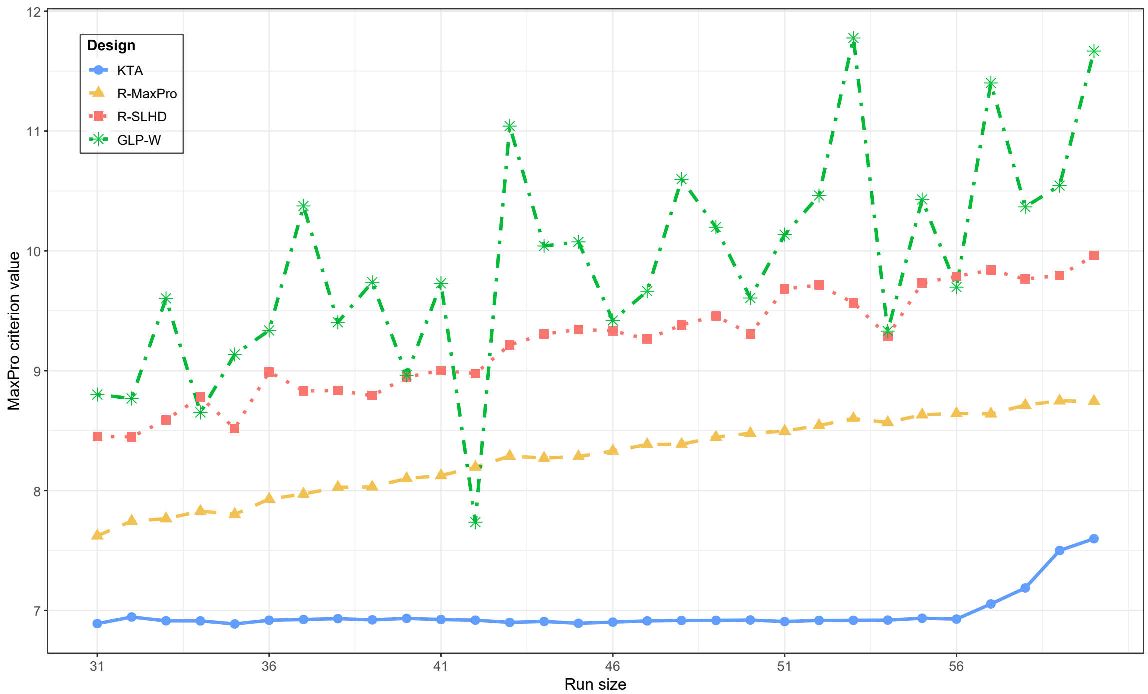


Figure S2: The $\psi$ values of the designs with 12 factors generated by different methods.

and "R-SLHD", designs are repeatedly generated 100 times, and the best-performing one is selected for comparison purposes. As can be seen from the figures, the designs generated by the KTA algorithm outperform the other designs, which indicates that the KTA algorithm also exhibits excellent performance under the MaxPro criterion.

## S3.2 Comparison under the MD

Figure S3 compares the MD$^2$ of designs generated by KTA with $t = 2$ and UD, where UD are generated by the R package "UniDOE" with default parameters. These designs all have 8 levels and 40 factors. When MD is selected, we replace $\tilde{\phi}(D, q)$ with $\tilde{\phi}_{MD}(D, q)$ in Algorithm 1. As we can see, designs obtained by KTA are superior than UD under the MD. This indicates that the KTA algorithm is also effective under the

MD.

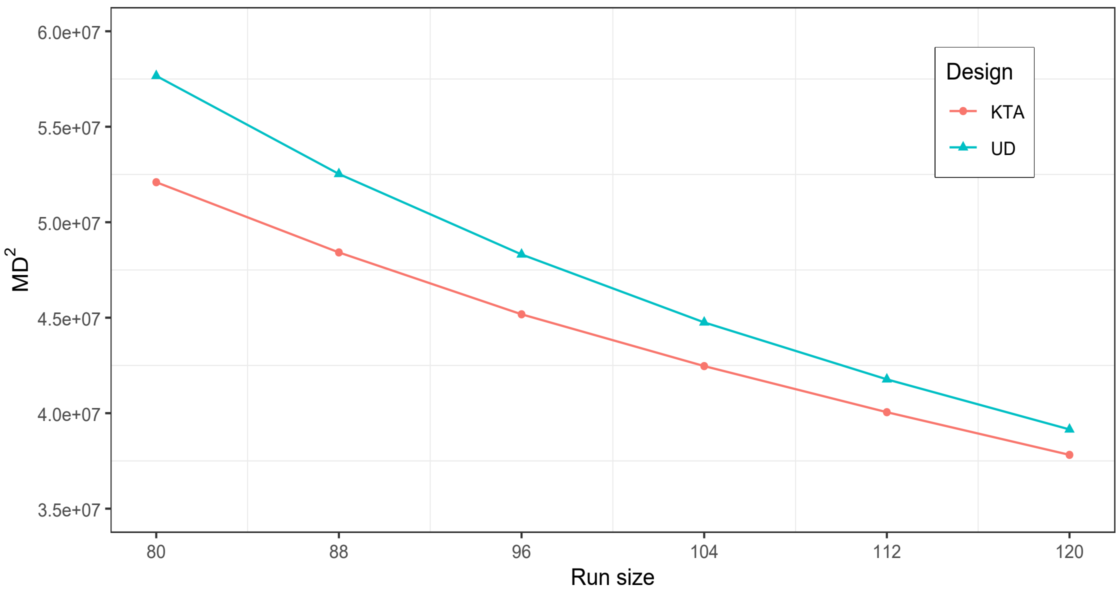


Figure S3: The $\mathrm{MD}^2$ of designs with 40 factors and 8 levels generated by KTA and UD.

## S3.3 Comparison under the CD

To further demonstrate the effectiveness of the KTA algorithm under the CD, we compare the designs generated by KTA algorithm with several strong orthogonal arrays (SOAs). An $N \times m$ matrix with entries from $\{0, 1, \ldots, s^t - 1\}$ is called an SOA of $N$ runs, $m$ factors, $s^t$ levels and strength $t$, denoted by $SOA(N, m, s^t, t)$, if any of its $N \times g$ $(1 \leq g \leq t)$ submatrix can be collapsed into an $OA(N, g, s^{u_1} \times \cdots \times s^{u_g}, g)$ for any positive integers $u_1, \ldots, u_g$ satisfying $u_1 + \cdots + u_g = t$, where collapsing $s^t$ levels into $s^{u_j}$ levels is done by $\lfloor a/s^{t-u_j} \rfloor$ for $a = 0, 1, \ldots, s^t - 1$.

Table S1 compares some SOAs and the KTA-generated designs with the same run size and number of factors under the CD, where all designs have 8 levels and the sources of SOAs are also listed. We set $\tilde{\phi}(D, q) = \tilde{\phi}_{CD}(D, q)$, with all other parameters kept at their default values. The third column of Table S1 presents the value of $q$ set in the KTA algorithm.

Table S1: The $\mathrm{CD}^2$ of SOAs and designs generated by KTA.

| SOA | Source | $q$ | $\mathrm{CD}^2$ of SOA | $\mathrm{CD}^2$ of KTA |
|---|---|---|---|---|
| $SOA(64, 15, 8^3, 3)$ | Shi and Tang (2020) | 2 | 0.2666 | 0.2285 |
| $SOA(64, 15, 8^3, 3)$ | Shi and Tang (2020) | 1 | 0.2666 | 0.2099 |
| $SOA(128, 31, 8^3, 3)$ | Shi and Tang (2020) | 3 | 14.5874 | 3.8233 |
| $SOA(128, 31, 8^3, 3)$ | Shi and Tang (2020) | 2 | 14.5874 | 4.1640 |
| $SOA(256, 63, 8^3, 3)$ | Shi and Tang (2020) | 3 | $2.6247 \times 10^5$ | $2.8447 \times 10^3$ |
| $SOA(256, 63, 8^3, 3)$ | Shi and Tang (2020) | 2 | $2.6247 \times 10^5$ | $3.4552 \times 10^3$ |
| $SOA(64, 20, 8^3, 3)$ | Shi and Xu (2024) | 2 | 1.1635 | 0.7117 |
| $SOA(64, 20, 8^3, 3)$ | Shi and Xu (2024) | 1 | 1.1635 | 0.6741 |
| $SOA(128, 40, 8^3, 3)$ | Shi and Xu (2024) | 3 | 174.7867 | 22.1975 |
| $SOA(128, 40, 8^3, 3)$ | Shi and Xu (2024) | 2 | 174.7867 | 17.9238 |
| $SOA(256, 80, 8^3, 3)$ | Shi and Xu (2024) | 3 | $6.2808 \times 10^7$ | $1.2826 \times 10^5$ |
| $SOA(256, 80, 8^3, 3)$ | Shi and Xu (2024) | 2 | $6.2808 \times 10^7$ | $1.6591 \times 10^5$ |

As can be seen from Table S1, the designs generated by the KTA algorithm outperform the SOAs under the CD, and the gap becomes more pronounced as the number of rows and columns increases.

## S3.4 Case studies

We further demonstrate the practical value of the KTA algorithm via two case studies. In these case studies, we evaluate different space-filling designs by building surrogate models using these designs and comparing their predictive performance on the test set.

The first case study uses the borehole function to generate response values, which is a commonly used function for simulation in computer experiments (Chen et al., 2016; Shi and Xu, 2024). The borehole function is

$$f_1(x_1,\ldots,x_8)=\frac{2\pi x_3(x_4-x_6)}{\ln(x_2/x_1)(1+\frac{2x_3x_7}{\ln(x_2/x_1)x_1^2x_8}+x_3/x_5)},$$

where $x_1\in[0.05,0.15]$ is the radius of the borehole $(m)$, $x_2\in[100,50000]$ is the radius of the influence $(m)$, $x_3\in[63070,115600]$ is the transmissivity $(m^2/yr)$ of the upper aquifer, $x_4\in[990,1110]$ is the potentiometric head $(m)$ of the upper aquifers, $x_5\in[63.1,116]$ is the transmissivity $(m^2/yr)$ of the lower aquifers, $x_6\in[700,820]$ is the potentiometric head $(m)$ of the lower aquifers, $x_7\in[1120,1680]$ is the length $(m)$ of the borehole, $x_8\in[9855,12045]$ is the hydraulic conductivity $(m/yr)$ of the borehole, and the response $f_1(x_1,\ldots,x_8)$ is the rate of waterflow $(m^3/yr)$.

The second case study uses the OTL circuit function and the aircraft wing function to generate response values, that is, $f_2(x_1,\ldots,x_{16})=y_1^*(x_1,\ldots,x_6)+y_2^*(x_7,\ldots,x_{16})$, where $y_i^*=(y_i-\min y_i)/(\max y_i-\min y_i)$, $i=1,2$,

$$\begin{aligned}
y_1(x_1,\ldots,x_6)=&\frac{\left(\frac{12x_2}{x_1+x_2}+0.74\right)x_6(x_5+9)+11.35x_3}{x_6(x_5+9)+x_3}+\frac{0.74x_3x_6(x_5+9)}{[x_6(x_5+9)+x_3]\,x_4},\\
y_2(x_7,\ldots,x_{16})=&0.036x_7^{0.758}x_8^{0.0035}\left(\frac{x_9}{\cos^2 x_{10}}\right)^{0.6}x_{11}^{0.006}x_{12}^{0.04}\left(\frac{100x_{13}}{\cos x_{10}}\right)^{-0.3}(x_{14}x_{15})^{0.49}\\
&+x_7x_{16}.
\end{aligned}$$

Here, $y_1$ is the OTL circuit function with 6 factors and $y_2$ is the aircraft wing function with 10 factors. Table S2 presents the physical meanings and ranges of the factors in $f_2(x_1,\ldots,x_{16})$. For further details, one can refer to Ben-Ari and Steinberg (2007); Forrester, Sobester and Keane (2008); Moon, Dean and Santner (2012).

Table S2: Parameters and ranges of the factors in $f_2(x_1, \ldots, x_{16})$.

| Factor | Parameter | Lower bound | Upper bound |
|---|---|---|---|
| $x_1$ | Resistance (K-Ohms) | 50 | 150 |
| $x_2$ | Resistance (K-Ohms) | 25 | 70 |
| $x_3$ | Resistance (K-Ohms) | 0.5 | 3 |
| $x_4$ | Resistance (K-Ohms) | 1.2 | 2.5 |
| $x_5$ | Resistance (K-Ohms) | 0.25 | 1.2 |
| $x_6$ | Current gain (Amperes) | 50 | 300 |
| $x_7$ | Wing area (ft$^2$) | 150 | 200 |
| $x_8$ | Weight of fuel in the wing (lb) | 220 | 300 |
| $x_9$ | Aspect ratio | 6 | 10 |
| $x_{10}$ | Quarter-chord sweep (deg) | $-10$ | 10 |
| $x_{11}$ | Dynamic pressure at cruise (lb/ft$^2$) | 16 | 45 |
| $x_{12}$ | Taper ratio | 0.5 | 1 |
| $x_{13}$ | Aerofoil thickness to chord ratio | 0.08 | 0.18 |
| $x_{14}$ | Ultimate load factor | 2.5 | 6 |
| $x_{15}$ | Flight design gross weight (lb) | 1700 | 2500 |
| $x_{16}$ | Paint weight (lb/ft$^2$) | 0.025 | 0.08 |

In both case studies, we use the four designs (KTA, R-MaxPro, R-SLHD, and GLP-W) to generate data, with 32 runs in the first case study and 40 runs in the second. The KTA designs are generated under the MaxPro criterion with $q = 1$. All factors are scaled to $[0, 1]$. We employ the GP_fit function from the R package "GPfit" to fit a Gaussian process model with a Gaussian correlation function, and use the predict.GP function from the same package for predictions, with all other parameters set to their default values.

To evaluate the performance of different designs, we calculate the normalized root mean squared error (RMSE) of the Gaussian process models built using these designs on the same test set, which is defined as

$$\text{Normalized RMSE} = \left\{ \frac{\sum_{X_i \in X_{test}} (\hat{f}_j(X_i) - f_j(X_i))^2}{\sum_{X_i \in X_{test}} (\bar{y} - f_j(X_i))^2} \right\}^{1/2}$$

for $j = 1, 2$. Here, $X_{test}$ is the test set, $\hat{f}_j(X_i)$ denotes the predictions from the fitted Gaussian process models, and $\bar{y}$ is the average response for the training data for building the model. We generate the test set with 10000 points using the randomLHS function from the R package "lhs", repeat this process to randomly generate 1000 test sets, and then plot the results as boxplots as shown in Figures S4 and S5.

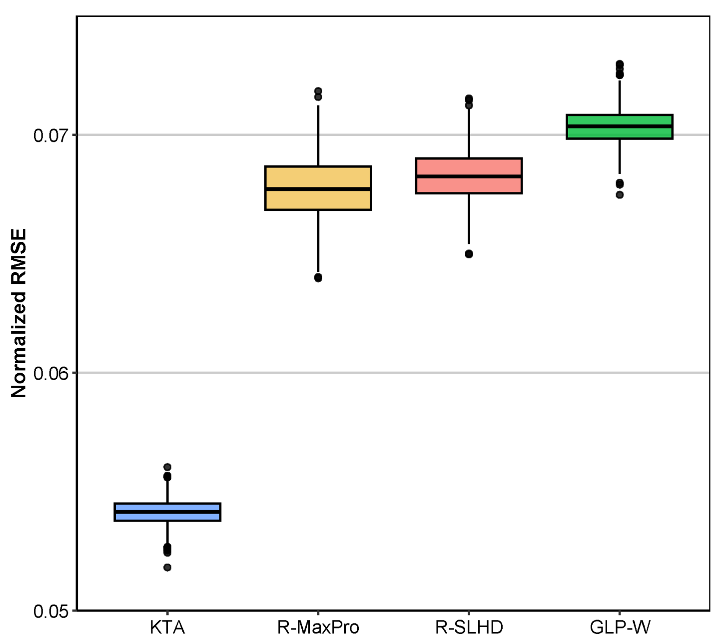


Figure S4: The normalized RMSEs of different designs with 32 runs under $f_1$.

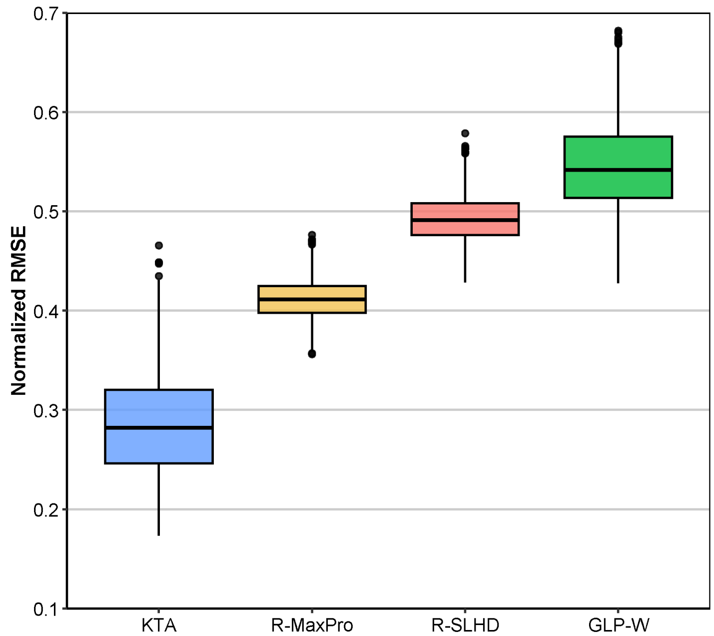


Figure S5: The normalized RMSEs of different designs with 40 runs under $f_2$.

As can be seen from the Figures S4 and S5, the Gaussian process models built using the designs generated by the KTA algorithm exhibit a smaller normalized RMSE. This indicates that the KTA algorithm can not only generate designs with excellent performance under a certain criterion but also ensure the resulting designs possess certain practical value. These results collectively demonstrate the effectiveness and practicality of the KTA algorithm.

# S4 A Large Table

Table S3: The $CD^2$ values of different designs.

| $N_0$ | $m_0$ | $s$ | $N$ | $m$ | $q$ | UD | ED | MM | KTA |
|---|---|---|---|---|---|---|---|---|---|
| 15 | 24 | 3 | 4 | 12 | 2 | 6.2059 | 6.1981 | **3.9038** | 3.9400 |
| 15* | 24 | 3 | 8 | 24 | 1 | 6.2059 | 6.1981 | 3.9038 | **3.7146** |
| 15 | 22 | 3 | 4 | 11 | 2 | 4.1402 | 4.1379 | **2.8674** | 2.9973 |
| 15* | 22 | 3 | 8 | 22 | 1 | 4.1402 | 4.1379 | 2.8674 | **2.8196** |
| 15 | 20 | 3 | 8 | 20 | 1 | 2.7296 | 2.7296 | 2.1366 | **2.0956** |
| 9 | 20 | 3 | 5 | 20 | 1 | 4.3353 | 4.3353 | 2.6129 | **2.5916** |
| 15 | 18 | 3 | 8 | 18 | 1 | 1.7773 | 1.7755 | 1.5452 | **1.5230** |
| 12 | 16 | 3 | 6 | 16 | 1 | 1.3704 | 1.3704 | 1.2128 | **1.1938** |
| 24 | 15 | 4 | 12 | 15 | 1 | 0.7046 | 0.7074 | 0.6832 | **0.6453** |
| 12 | 15 | 4 | 6 | 15 | 1 | 1.3778 | 1.3778 | 1.0806 | **1.0398** |
| 8 | 15 | 4 | 4 | 15 | 1 | 2.1020 | 2.1016 | 1.3790 | **1.3048** |
| 15 | 15 | 3 | 8 | 15 | 1 | 0.9105 | 0.9086 | 0.8952 | **0.8904** |
| 9 | 15 | 3 | 5 | 15 | 1 | 1.4030 | 1.4030 | **1.1805** | 1.2026 |
| 60 | 20 | 6 | 15 | 10 | 2 | 0.9225 | 0.9312 | 0.8604 | **0.8066** |
| 60 | 25 | 6 | 15 | 12 | 2 | 3.1070 | 3.1277 | 2.2785 | **2.0723** |
| 102 | 25 | 6 | 13 | 9 | 3 | 1.9106 | 1.9298 | **1.6282** | 1.7471 |
| 102* | 25 | 6 | 26 | 13 | 2 | 1.9106 | 1.9298 | 1.6282 | **1.5034** |
| 60 | 29 | 6 | 15 | 15 | 2 | 7.9773 | 8.0190 | 4.5032 | **3.8583** |
| 102 | 29 | 6 | 13 | 10 | 3 | 4.7800 | 4.8205 | 3.3380 | **3.3170** |
| 70 | 25 | 7 | 18 | 13 | 2 | 2.2898 | 2.3096 | 1.9159 | **1.7606** |
| 70 | 29 | 7 | 18 | 15 | 2 | 5.7675 | 5.7927 | 3.7370 | **3.3802** |
| 105 | 29 | 7 | 13 | 10 | 3 | 3.8919 | 3.9241 | **2.9659** | 3.1104 |
| 105* | 29 | 7 | 26 | 15 | 2 | 3.8919 | 3.9241 | 2.9659 | **2.7466** |
| 80 | 25 | 8 | 10 | 9 | 3 | 2.1842 | 2.2062 | **1.7424** | 1.8741 |
| 80* | 25 | 8 | 20 | 13 | 2 | 2.1842 | 2.2062 | 1.7424 | **1.5802** |
| 112 | 29 | 8 | 14 | 10 | 3 | 4.0666 | 4.1067 | **2.9069** | 2.9517 |
| 112* | 29 | 8 | 28 | 15 | 2 | 4.0666 | 4.1067 | 2.9069 | **2.6000** |
| 110 | 25 | 11 | 14 | 9 | 3 | 1.4375 | 1.4583 | **1.3011** | 1.4363 |
| 110* | 25 | 11 | 28 | 13 | 2 | 1.4375 | 1.4583 | 1.3011 | **1.1991** |
| 154 | 29 | 11 | 19 | 10 | 3 | 2.6798 | 2.7079 | **2.2306** | 2.3042 |
| 154* | 29 | 11 | 39 | 15 | 2 | 2.6798 | 2.7079 | 2.2306 | **2.0105** |

# References

Ben-Ari, E. N. and Steinberg, D. M. (2007). Modeling data from computer experiments: an empirical comparison of kriging with MARS and projection pursuit regression. *Qual. Eng.* **19**, 327–338.

Bhatia, R. and Jain, T. (2015). Mean matrices and conditional negativity. *Electron. J. Linear Algebra* **29**, 206–222.

Chen, G. and Tang, B. (2022). A study of orthogonal array-based designs under a broad class of space-filling criteria. *Ann. Statist.* **50**, 2925–2949.

Chen, H., Loeppky, J. L., Sacks, J. and Welch, W. J. (2016). Analysis methods for computer experiments: How to assess and what counts? *Statist. Sci.* **31**, 40–60.

Forrester, A., Sobester, A. and Keane, A. (2008). *Engineering Design via Surrogate Modelling: A Practical Guide.* Wiley, Chichester.

Moon, H., Dean, A. M. and Santner, T. J. (2012). Two-stage sensitivity-based group screening in computer experiments. *Technometrics* **54**, 376–387.

Rankin, R. A. (1955). The closest packing of spherical caps in $n$ dimensions. *Glasg. Math. J.* **2**, 139–144.

Shi, C. and Tang, B. (2020). Construction results for strong orthogonal arrays of strength three. *Bernoulli* **26**, 418–431.

Shi, C. and Xu, H. (2024). A projection space-filling criterion and related optimality results. *J. Amer. Statist. Assoc.* **119**, 2658–2669.

Tang, Y., Xu, H. and Lin, D. K. J. (2012). Uniform fractional factorial designs. *Ann. Statist.* **40**, 891–907.

Van Den Berg, C., Christensen, J. P. R. and Ressel, P. (2012). *Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions.* Springer Science & Business Media, New York.

Ye, Y., Yuan, R. and Wang, Y. (2025). Construction of maximum projection Latin hypercube designs using number-theoretic methods. *Scand. J. Statist.* **52**, 1899–1931.