Statistics and Probability Letters 181 (2022) 109267

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

On definitive screening designs using Paley's conference matrices

Yaping Wang [a], Sixu Liu [b,*], Dennis K.J. Lin [c]

[a] KLATASDS–MOE, School of Statistics, East China Normal University, Shanghai 200062, China
[b] Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China
[c] Department of Statistics, Purdue University, West Lafayette, IN 47907, USA

A R T I C L E I N F O

Article history:
Received 20 January 2021
Received in revised form 8 August 2021
Accepted 27 September 2021
Available online 7 October 2021

MSC:
62K15
62K10

Keywords:
Correlation
Fold-over
Generalized aberration criterion
J_4-characteristic
Two-factor interactions

A B S T R A C T

Definitive screening designs (DSDs) are widely used for studying quantitative factors. However, DSDs constructed from different conference matrices are not equally good. We show DSDs using Paley's conference matrices guarantee desirable performance (either optimal or near-optimal) under several important criteria.

1. Introduction

Screening designs aim to find the active factors from a large number of factors involved in the early stages of experiments. They are typically two-level or three-level designs with economic run sizes. Jones and Nachtsheim (2011) introduced a new class of three-level designs called definitive screening designs (DSDs) and showed that they outperform standard screening designs in many aspects. DSDs have been widely used and investigated, see Yang et al. (2014), Phoa and Lin (2015), Schoen et al. (2019) and Vazquez et al. (2020) among others. A DSD has $2m+1$ runs consisting of m fold-over pairs and a center point for investigating m quantitative factors with three levels. Here, two runs x and y are called a fold-over pair, if $y=-x$. The original DSDs in Jones and Nachtsheim (2011) were constructed by an algorithm. Xiao et al. (2012) proposed to use conference matrices to construct such designs, which requires no computer search and yields orthogonal main effects. Throughout, we focus on DSDs based on conference matrices. For an even number m, a conference matrix of order m, denoted by $C=(c_{i,j})$, is an $m\times m$ matrix with diagonal entries $c_{i,i}=0$, $i=1,\ldots,m$, and off-diagonal entries $c_{i,j}\in\{1,-1\}$, $i\neq j$, such that $C^{\mathrm{T}}C=(m-1)I_m$, where I_m is an $m\times m$ identity matrix. Given a conference matrix C, a DSD can be constructed as $D=(C^{\mathrm{T}},-C^{\mathrm{T}},0_m)^{\mathrm{T}}$, where 0_m is a vector of m zeros. For D, assume that the response y_i follows the second-order linear model considered in Jones and Nachtsheim (2011),

$$y_i=\beta_0+\sum_{j=1}^{m}\beta_j d_{i,j}+\sum_{i=1}^{m}\beta_{jj}d_{i,j}^2+\sum_{j=1}^{m-1}\sum_{k=j+1}^{m}\beta_{jk}d_{i,j}d_{i,k}+\epsilon_i,\quad i=1,\ldots,2m+1, \tag{1}$$

* Corresponding author.
E-mail addresses: ypwang@fem.ecnu.edu.cn (Y. Wang), liusixu@mail.tsinghua.edu.cn (S. Liu), dkjlin@purdue.edu (D.K.J. Lin).

https://doi.org/10.1016/j.spl.2021.109267
0167-7152/

where $d_{i,j}$ is the (i,j)th entry of D, ϵ_i's are i.i.d zero-mean errors, and β_0, β_j, β_{jk}, $j = 1, \ldots, m$, $k = j, \ldots, m$ are unknown parameters representing the intercept, all main, quadratic and two-factor interaction effects, respectively.

In practice, the effect sparsity assumption is generally warranted. In such a case, after a few active factors are identified using a screening design, we often fit a multiple regression model containing the main effects, possibly some quadratic effects, and two-factor interactions involving these factors. Therefore, the properties of a screening design projected onto small subsets of columns are crucial. Investigating such properties and designs ranking has received much attention in the literature of two-level fractional factorial designs (Deng and Tang, 1999; Bulutoglu and Cheng, 2003 and Shi and Tang, 2018). The same issue exists for DSDs, i.e., some of them are better than others. This is because, for $m \geq 24$, there are many nonisomorphic conference matrices of the same order. Conference matrices can differ significantly in terms of their projections and thus lead to DSDs with different qualities. In a seminal work, Schoen et al. (2019) studied isomorphism classes of two-, three- and four-factor projections of conference matrices and proposed a generalized aberration criterion for ranking DSDs. But determining the minimum generalized aberration design among all nonisomorphic DSDs is a challenging problem. This paper reveals that applying a specific type of conference matrix, the Paley's conference matrix (Paley, 1933), to construct DSD guarantees desirable design properties. The main contribution of this work is twofold. First, a general lower bound on the maximum absolute correlation among two-factor interaction columns in model (1) for DSD is established, which provides a benchmark for any DSD. Second, an upper bound on the maximum absolute correlation among two-factor interaction columns in model (1) for DSDs constructed from Paley's conference matrices is obtained, which provides theoretical support for such Paley-based DSDs. Numerical results show that the Paley-based DSDs are either optimal or nearly optimal under several criteria.

The remainder of this paper is organized as follows. Section 2.1 presents some preliminaries. Section 2.2 establishes a lower bound on the maximum correlation among two-factor interaction columns, which is of order $\mathcal{O}(1/\sqrt{m})$, for any DSD with m factors. Section 2.3 shows an upper bound on the maximum correlation among two-factor interaction columns, again of order $\mathcal{O}(1/\sqrt{m})$, for a DSD constructed from Paley's conference matrix. This upper bound is numerically shown to be tight for small and medium m. Section 3 gives some comparisons and a discussion. The online supplementary material contains detailed comparison results and four Paley's conference matrices for practical use.

2. Results

2.1. Preliminaries

Two conference matrices are isomorphic if one can be obtained from the other by some row permutations, column permutations, sign switches in columns, or sign switches in rows, where the permutation applied to rows should be the same as the permutation applied to columns. A conference matrix $C = (c_{i,j})$ of order m is called normalized if $c_{1,1} = 0$, $c_{i,1} = 1$ and $c_{1,i} = 1$, $i = 2, \ldots, m$. Conference matrices of the same order can be grouped into different isomorphism classes (Schoen et al., 2019). In each isomorphism class, a normalized conference matrix can be used as a representative.

Paley (1933) proposed a method of constructing conference matrices which we call Paley's conference matrices; see also Ionin and Kharaghani (2007) for more details. Denote by $GF(s) = \{\alpha_1(=0), \alpha_2, \ldots, \alpha_s\}$ the Galois field with $s \geq 3$ elements and define the quadratic character $\chi(\cdot)$ of $GF(s)$ by $\chi(0) = 0$;

$$\chi(\alpha) = \begin{cases} 1, & \text{if } \alpha = \beta^2 \text{ for some } \beta \in GF(s); \\ -1, & \text{otherwise,} \end{cases}$$

where $\alpha \in GF(s)\backslash\{0\}$. Let Q be the $s \times s$ matrix with the (i,j)th entry being $\chi(\alpha_i - \alpha_j)$, $i, j = 1, \ldots, s$. Then the (normalized) Paley's conference matrix of order $m = s + 1$ is

$$C = \begin{pmatrix} 0 & 1_s^{\mathrm{T}} \\ 1_s & Q \end{pmatrix},$$

where 1_s is a vector of s ones. This method is valid for all odd prime powers s. Xiao et al. (2012) tabulated conference matrices of sizes $m = 2, 4, \ldots, 18$. Except for $m = 2$ and 16, all their matrices are members of Paley's conference matrices (see Section 3). In the supplementary material, we display Paley's conference matrices of order $m = 20, 24, 26$ and 28 for practical use.

The J-characteristic defined by Deng and Tang (1999) is a useful tool for studying fractional factorial designs. For a four-column design, its j_4-characteristic is the sum of the elementwise products of the four columns and the J_4-characteristic is the absolute value of the j_4-characteristic. The following lemma rephrases Theorems 3 and 4 of Schoen et al. (2019).

Lemma 1. *For any four-factor projection of a conference matrix of order m, (1) if $m = 0$ (mod 4), there are $m/4$ possible values of the J_4-characteristics which are $m-4$, $m-8$, …, 0; (2) if $m = 2$ (mod 4), there are $(m-2)/4$ possible values of the J_4-characteristics which are $m-4$, $m-8$, …, 2.*

2.2. A general lower bound on the maximum correlation among two-factor interaction columns

For an m-factor DSD, the $(2m+1)\times(1+2m+\binom{m}{2})$ model matrix under model (1) has the following nice properties (see Jones and Nachtsheim, 2011 and Xiao et al., 2012):

1. All main effect columns are orthogonal to each other;
2. All main effect columns and quadratic effect columns are orthogonal;
3. All main effect columns and two-factor interaction columns are orthogonal;
4. The correlation of any pair of quadratic effect columns is $1/3-1/(m-1)$;
5. The correlation of any quadratic effect column involving a given factor with a two-factor interaction column involving that factor is 0;
6. The correlation of any quadratic effect column involving a factor with a two-factor interaction column not involving that factor is $\pm[(2m+1)/\{3(m-1)(m-2)\}]^{1/2}$;
7. The correlation of any pair of two-factor interaction columns involving a common factor is $\pm 1/(m-2)$.

The above properties always hold, and are irrelevant to which conference matrix of order m is chosen. However, the correlations among two-factor interaction columns without common factors are extremely complicated (Xiao et al., 2012). They vary considerably when nonisomorphic conference matrices are used. To select a DSD with less aliasing among two-factor interactions, we adopt a minimax criterion that minimizes the maximum absolute correlation among two-factor interaction columns of the model matrix. Let $\lceil\cdot\rceil$ and $\lfloor\cdot\rfloor$ denote the ceiling and floor functions, respectively. Theorem 1 below establishes a lower bound of such maximum absolute correlation, which provides a benchmark for any DSD.

Theorem 1. *For any m-factor DSD involving $2m+1$ runs constructed from a conference matrix, the maximum absolute correlation between pairs of all two-factor interaction columns in model (1) is greater than or equal to*

$$(m-4\lambda^*)/(m-2), \text{ where } \lambda^* = \left\lfloor m/4 - \left\lceil \{m-1-3/(m-3)\}^{1/2}\right\rceil/4 \right\rfloor. \quad (2)$$

Proof of Theorem 1. Let D be an m-factor DSD constructed from an m-order conference matrix C. Denote the m columns of C by $c_1,\ldots,c_m$. As the absolute correlation of any pair of two-factor interaction columns involving a common factor is $1/(m-2)$, a constant, we only need to consider the absolute correlations of pairs of two-factor interaction columns involving four distinct factors. Observe that such an absolute correlation equals $J_4/(2m-4)$ with the J_4-characteristic calculated for the four-factor projection of D where the pair of two-factor interactions involve, it suffices to obtain a lower bound of $\max(J_4)$ for D. Since $\max(J_4)$ for D is twice $\max(J_4)$ for C, below we only consider $\max(J_4)$ for C.

Now fix the first three columns of C and calculate the sum of squared $J_4(c_1,c_2,c_3,c_j)$, $j=1,\ldots,m$, where $J_4(c_1,c_2,c_3,c_j)$ denotes the J_4-characteristic of the four-column design (c_1,c_2,c_3,c_j). We have

$$\sum_{j=1}^{m} J_4^2(c_1,c_2,c_3,c_j) = \sum_{j=1}^{m}\Big(\sum_{i=1}^{2m+1} c_{i,1}c_{i,2}c_{i,3}c_{i,j}\Big)^2 = \sum_{j=1}^{m}(c_1\odot c_2\odot c_3)^{\mathrm{T}} c_j c_j^{\mathrm{T}}(c_1\odot c_2\odot c_3),$$

which simplifies to $(c_1\odot c_2\odot c_3)^{\mathrm{T}}CC^{\mathrm{T}}(c_1\odot c_2\odot c_3)=(m-1)(m-3)$ since $CC^{\mathrm{T}}=C^{\mathrm{T}}C=(m-1)I_m$, where $\odot$ represents for the elementwise product of vectors. Also, by the property of C, $J_4(c_1,c_2,c_3,c_j)$ must equal one, $j=1,2,3$. Therefore, the average of squared J_4-characteristic values of all the four-factor projections of C indexed by column numbers 1, 2, 3 and j $(4\le j\le m)$ is $(m-3)^{-1}[(m-1)(m-3)-\sum_{j=1}^{3}J_4^2(c_1,c_2,c_3,c_j)]=(m-1)-3/(m-3)$. Because the squared value of $\max(J_4)$ for C cannot be smaller than the above value and a J-characteristic must be an integer, we have

$$\max(J_4) \ge \lceil\{m-1-3/(m-3)\}^{1/2}\rceil$$

for C. Then the desired result follows by the above inequality and Lemma 1. □

By the proof of Theorem 1, the lower bound $(m-4\lambda^*)/(m-2) \ge \lceil\sqrt{m-1-3/(m-3)}\rceil/(m-2)$, which is of order $\mathcal{O}(1/\sqrt{m})$. Theorem 1 can also be stated in form of the J_4-characteristic, see Corollary 1.

Corollary 1. *For all four-factor projections of an m-factor DSD involving $2m+1$ runs constructed from any conference matrix, $\max(J_4)\ge 2m-8\lambda^*$, where λ^* is defined in (2).*

2.3. Results on definitive screening designs using Paley's conference matrices

Next, we show in Theorem 2 that the DSDs using Paley's conference matrices have an upper bound on the maximum absolute correlation between two-factor interaction columns.

Theorem 2. *When $m\ge 6$, for an m-factor DSD involving $2m+1$ runs constructed from Paley's conference matrix, the maximum absolute correlation between pairs of all two-factor interaction columns in model (1) is less than or equal to*

$$(m-4\lambda_*)/(m-2), \text{ where } \lambda_* = \left\lceil (m-\lfloor 2\sqrt{m-1}\rfloor)/4 \right\rceil. \quad (3)$$

Table 1
λ^*, λ_*, the lower bound (for general DSDs), the upper bound (for Paley-based DSDs) and true value (for Paley-based DSDs) of maximum absolute correlation (denoted as $|r|^{LB}_{\max}$, $|r|^{UB}_{\max}$, $|r|^{true}_{\max}$, respectively) between pairs of all two-factor interaction columns.

m	λ^*	λ_*	$\|r\|^{LB}_{\max}$	$\|r\|^{UB}_{\max}$	$\|r\|^{true}_{\max}$	m	λ^*	λ_*	$\|r\|^{LB}_{\max}$	$\|r\|^{UB}_{\max}$	$\|r\|^{true}_{\max}$	m	λ^*	λ_*	$\|r\|^{LB}_{\max}$	$\|r\|^{UB}_{\max}$	$\|r\|^{true}_{\max}$
6	1	1	0.5	0.5	0.5	30	6	5	0.214	0.357	0.357	68	14	13	0.181	0.242	0.242
8	1	1	0.667	0.667	0.667	32	6	6	0.267	0.267	0.267	72	15	14	0.171	0.229	0.229
10	1	1	0.75	0.75	0.75	38	7	7	0.278	0.278	0.278	74	16	15	0.139	0.194	0.194
12	2	2	0.4	0.4	0.4	42	8	8	0.25	0.25	0.25	80	17	16	0.154	0.205	0.205
14	2	2	0.5	0.5	0.5	44	9	8	0.19	0.286	0.286	82	18	16	0.125	0.225	0.225
18	3	3	0.375	0.375	0.375	48	10	9	0.174	0.261	0.261	84	18	17	0.146	0.195	0.195
20	3	3	0.444	0.444	0.444	50	10	9	0.208	0.292	0.292	90	20	18	0.113	0.205	0.205
24	4	4	0.364	0.364	0.364	54	11	10	0.192	0.269	0.269	98	22	20	0.104	0.188	0.188
26	5	4	0.25	0.417	0.417	60	13	12	0.138	0.207	0.207						
28	5	5	0.308	0.308	0.308	62	13	12	0.167	0.233	0.233						

To prove Theorem 2, the following lemma is crucial. It is a result similar to Theorem 2.1 of Bulutoglu and Cheng (2003) and Lemma 3 of Shi and Tang (2018) for two-level orthogonal arrays, which are applications of the Hasse–Weil bound (Hasse, 1936; Weil, 1948) in analytical number theory.

Lemma 2. *For all four-factor projections of a Paley's conference matrix of order m,* $\max(J_4) \le \lfloor 2\sqrt{m-1} \rfloor$.

Proof of Lemma 2. Let C be a normalized Paley's conference matrix of order m. We consider the j_4-characteristic of the four-factor projection of C indexed by column numbers i_1, i_2, i_3 and i_4, where $1 \le i_1 < i_2 < i_3 < i_4 \le m$.

Case I: $i_1 > 1$. Denote $p(x) = (x-\alpha_{i_1})(x-\alpha_{i_2})(x-\alpha_{i_3})(x-\alpha_{i_4})$, which is a quartic polynomial on $GF(s)$. By the Paley's construction and the fact that $\chi(\alpha)\chi(\beta) = \chi(\alpha\beta)$ for all $\alpha, \beta \in GF(s)$, we have $j_4 = 1 + \sum_{x \in GF(s)} \chi(p(x))$. Consider the equation $z^2 = p(x)$ and let N_s denote the number of solutions (z, x) over $GF(s)$. We have four solutions with $z = 0$, i.e., $(0, \alpha_{i_j})$, $j = 1, 2, 3, 4$. The other $N_s - 4$ solutions come in pairs with (z, x) and $(-z, x)$ where $z \ne 0$. These facts reveal that for $x \in GF(s)$, $x \ne \alpha_{i_1}, \alpha_{i_2}, \alpha_{i_3}, \alpha_{i_4}$, $\chi(p(x))$ equals 1 for $N_s/2 - 2$ times and -1 for $s - 4 - (N_s/2 - 2)$ times. Therefore,

$$j_4 = 1 + N_s/2 - 2 - (s - 4 - (N_s/2 - 2)) = N_s - s + 1.$$

On the other hand, as $p(x)$ is a quartic polynomial without double roots, the Hasse–Weil bound (see Hasse, 1936; Weil, 1948 and Bulutoglu and Cheng, 2003) implies $|N_s - s + 1| \le 2\sqrt{s}$, which proves that $J_4 = |j_4| \le 2\sqrt{s} = 2\sqrt{m-1}$.

Case II: $i_1 = 1$. In this case the column $(0, 1_s^{\mathrm{T}})^{\mathrm{T}}$ is chosen. Denote $q(x) = (x-\alpha_{i_2})(x-\alpha_{i_3})(x-\alpha_{i_4})$, which is a cubic polynomial on $GF(s)$ without double roots. Then we have $j_4 = \sum_{x \in GF(s)} \chi(q(x))$. Let N'_s be number of solutions (z, x) to the equation $z^2 = q(x)$ over $GF(s)$. We have

$$j_4 = N'_s/2 - 2 - (s - 4 - (N'_s/2 - 2)) = N'_s - s.$$

It again follows from the Hasse–Weil bound that $J_4 = |N'_s - s| \le 2\sqrt{s} = 2\sqrt{m-1}$.

Combining both cases and noting that j_4 is an integer, we conclude that $J_4 \le \lfloor 2\sqrt{m-1} \rfloor$. □

Proof of Theorem 2. Theorem 2 follows from Lemmas 1 and 2, the fact that $\max(J_4)$ for $D = (C^{\mathrm{T}}, -C^{\mathrm{T}}, 0_m)^{\mathrm{T}}$ is twice $\max(J_4)$ for C and the property that the absolute correlation of any pair of two-factor interaction columns involving a common factor is $1/(m-2)$. □

When $m = 4$, $(m - 4\lambda_*)/(m-2) = 0$, but the maximum absolute correlation between pairs of all two-factor interaction columns is 0.5, which occurs when a pair of two-factor interactions involve a common factor. The upper bound in Theorem 2 performs well. We have numerically verified that it is tight for any $6 \le m < 500$ such that a Paley's conference matrix of order m exists.

The upper bound in Theorem 2 is again of order $\mathcal{O}(1/\sqrt{m})$, since $(m-4\lambda_*)/(m-2) \le \lfloor 2\sqrt{m-1} \rfloor/(m-2)$. It is close to the lower bound in Theorem 1, thus it guarantees the good performance of Paley-based DSDs. Table 1 lists the values of λ^* in (2), λ_* in (3), the lower bound in Theorem 1, the upper bound in Theorem 2 and the true value of maximum absolute correlation between pairs of all two-factor interaction columns for Paley-based DSDs with $6 \le m < 100$. By Table 1, the two bounds in Theorems 1 and 2 are identical for $m = 6, 8, 10, 12, 14, 18, 20, 24, 28, 32, 38, 42$, and therefore in these cases a Paley-based DSD minimizes the maximum absolute correlation between pairs of all two-factor interaction columns among all DSDs. For any other m the gap between λ_* and λ^* is small. The true values of maximum absolute correlations (in Table 1) always attain the upper bound in Theorem 2 and decrease approximately at a rate of $1/\sqrt{m}$.

Theorem 2 can be rephrased using J_4-characteristic given in the following corollary, which supports the good performance of Paley-based DSDs under the criterion of minimizing $\max(J_4)$.

Corollary 2. *For all four-factor projections of an m-factor DSD involving $2m+1$ runs constructed from Paley's conference matrix,* $\max(J_4) \le 2m - 8\lambda_*$, *where λ_* is defined in (3).*

For a DSD, let $F_{4,\lambda}$ be the frequency of the J_4-characteristics of $2m - 8\lambda$ for $\lambda = 1, \ldots, m/4$ when $m = 0 \pmod 4$, or $\lambda = 1, \ldots, (m-2)/4$ when $m = 2 \pmod 4$. Schoen et al. (2019) proposed a generalized aberration criterion that sequentially minimizes $F_{4,\lambda}$, $\lambda = 1, 2, \ldots$ for ranking nonisomorphic DSDs. For D_1 and D_2 of the same size, D_1 has less generalized aberration than D_2, if there exists a λ_0 such that $F_{4,\lambda_0}(D_1) < F_{4,\lambda_0}(D_2)$ and $F_{4,\lambda}(D_1) = F_{4,\lambda}(D_2)$ for $1 \leq \lambda < \lambda_0$. D_1 is called a minimum generalized aberration design if no other designs have less generalized aberration than it. By Corollaries 1 and 2, any DSD satisfies $\max(J_4) \geq 2m - 8\lambda^*$, and a DSD based on Paley's conference matrix, denoted by D_P, satisfies $2m - 8\lambda^* \leq \max(J_4) \leq 2m - 8\lambda_*$, where λ^* and λ_* are defined in (2) and (3), respectively. If a DSD based on a non-Paley's conference matrix, denoted by D, has $\max(J_4) \geq 2m - 8(\lambda_* - 1)$, then there must be at least one $F_{4,\lambda'} > 0$, $1 \leq \lambda' \leq \lambda_* - 1$. Hence D_P must have less generalized aberration than D since D_P has $F_{4,\lambda} = 0$ for all $1 \leq \lambda \leq \lambda'$. The numerical comparisons in the next section and the supplementary material show that D_P often has minimum generalized aberration when m is small.

3. Comparison and discussion

We compare some DSDs using Paley's conference matrices with designs using other conference matrices. When $m \leq 26$, all nonisomorphic conference matrices of order m, denoted as C_i, $i \geq 1$, are available in Schoen et al. (2019) or their website (http://www.pietereendebak.nl/oapackage/conference.html). We also refer the readers to Elster and Neumaier (1995) and Ionin and Kharaghani (2007) for good reviews of existing construction methods of conference matrices. By Schoen et al. (2019), when $m \leq 18$, there is only one isomorphism class of conference matrices. Therefore all DSDs with $m = 6, 8, 10, 12, 14$ and 18 are Paley-based DSDs, and they satisfy Theorem 2 and Corollary 2. In the supplementary material, we present detailed comparison results for the cases $m = 20, 24, 26, 28$ and 82 where there exist non-Paley's conference matrices. The maximum absolute correlation between pairs of all two-factor interaction columns, $\max(J_4)$, and the frequency vectors $F_{4,\lambda}$, $\lambda = 1, 2, \ldots$ for the generalized aberration criterion are calculated and compared. From these comparisons and other numerical studies not reported here, among all nonisomorphic conference matrices, Paley's conference matrices are recommended for constructing DSDs if they are available, especially when $m < 100$. Such DSDs have small $\max(J_4)$ and sometimes are the minimum generalized aberration designs.

Finally, we conclude this paper with a discussion. Searching for minimum generalized aberration DSDs among all isomorphism classes is important but often numerically unaffordable, especially when the design size is large, as it requires enumerating and comparing all nonisomorphic conference matrices. In this work, we show that DSDs using Paley's conference matrices guarantee desirable properties: they have small maximum absolute correlations among all two-factor interaction columns, small maximum J_4-characteristics, and perform well under the generalized aberration criterion. The construction of such designs is rather straightforward; it only requires Paley's conference matrices and avoids computational problems.

It is worth mentioning some possible limitations of DSDs based on Paley's conference matrices. First, such a design only exists when $m - 1$ is an odd prime power. Second, even when Paley's conference matrix exists, there might exist a DSD based on non-Paley's conference matrix with smaller $\max(J_4)$ or less generalized aberration. In fact, from Table 1, in 15 cases for which a Paley's conference matrix with $30 \leq m \leq 98$ ($m \neq 32, 38, 42$) exists, we do not know whether that matrix gives the smallest $\max(J_4)$. However, in 12 cases, the best possible $\max(J_4)$ and the $\max(J_4)$ for Paley-based design can differ by at most 8. For $m \in \{82, 90, 98\}$, the best possible $\max(J_4)$ and the $\max(J_4)$ for Paley-based design can differ by at most 16, but the run sizes of the DSDs make that the corresponding difference in maximum correlations is rather small.

Acknowledgments

The authors sincerely thank the editor, the associate editor and two referees for their constructive comments and suggestions that significantly improved the paper. Sixu Liu and Yaping Wang are supported by NSFC (11901199, 11871487, 11971171), Shanghai Sailing Program (19YF1412800) and Shanghai Chenguang Program (19CG26). Dennis Lin is supported by the U.S. National Science Foundation (Grant No. DMS-18102925).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spl.2021.109267.

References

Bulutoglu, D.A., Cheng, C., 2003. Hidden projection properties of some nonregular fractional factorial designs and their applications. Ann. Statist. 31, 1012–1026.

Deng, L.Y., Tang, B., 1999. Generalized resolution and minimum aberration criteria for Plackett–Burman and other nonregular factorial designs. Statist. Sinica 9, 1071–1082.

Elster, C., Neumaier, A., 1995. Screening by conference designs. Biometrika 82, 589–602.

Hasse, H., 1936. Zur Theorie der abstrakten elliptischen Funktionenkörper. I–III. J. Reine Angew. Math. 175, 55–62, 69–88, 193–208.

Ionin, Y.J., Kharaghani, H., 2007. Balanced generalized weighing matrices and conference matrices. In: The CRC Handbook of Combinatorial Designs, second ed. Taylor and Francis, Boca Raton, pp. 306–313.

Jones, B., Nachtsheim, C.J., 2011. A class of three-level designs for definitive screening in the presence of second-order effects. J. Qual. Technol. 43, 1–15.
Paley, R.E.A.C., 1933. On orthogonal matrices. J. Math. Phys. 12, 311–320.
Phoa, F.K.H., Lin, D.K.J., 2015. A systematic approach for the construction of definitive screening designs. Statist. Sinica 25, 853–861.
Schoen, E.D., Eendebak, P.T., Goos, P., 2019. A classification criterion for definitive screening designs. Ann. Statist. 47, 1179–1202.
Shi, C., Tang, B., 2018. Designs from good Hadamard matrices. Bernoulli 24, 661–671.
Vazquez, A.R., Goos, P., Schoen, E.D., 2020. Projections of definitive screening designs by dropping columns: Selection and evaluation. Technometrics 62, 37–47.
Weil, A., 1948. Sur les courbes algébriques et les variétés qui s'en déduisent, No. 1041 in Actualités scientifiques et industrielles. Hermann, Paris.
Xiao, L., Lin, D.K.J., Bai, F., 2012. Constructing definitive screening designs using conference matrices. J. Qual. Technol. 44, 2–8.
Yang, J., Lin, D.K.J., Liu, M., 2014. Construction of minimal-point mixed-level screening designs using conference matrices. J. Qual. Technol. 46, 251–264.