Construction of uniform U designs

Bochuan Jiang[a,b], **Mingyao Ai**[a*]

[a] *LMAM, School of Mathematical Sciences and Center for Statistical Science, Peking University, Beijing 100871, China*

[b] *Science and Technology on Complex Land Systems Simulation Laboratory, Beijing 100012, China*

Abstract. Orthogonal array based-Latin hypercubes, also called U designs, have popularly been adopted for designing a computer experiment. The relationship between the averaged squared discrepancy of all U designs generated from a given orthogonal array and the generalized wordtype pattern of the orthogonal array is established. Motivated by the relationship, we define a weighted wordtype pattern and a minimum weighted aberration criterion to compare orthogonal arrays of the same parameters. U designs generated from an orthogonal array with less weighted aberration are shown to have low discrepancies on average. Then, an algorithm to construct uniform U designs is proposed. It begins with a minimum weighted aberration orthogonal array and its advantage is illustrated by comparing with another two methods.

MSC: Primary 62K15; Secondary 62K10.

Keywords: discrepancy, generalized wordtype pattern, U design, stochastic algorithm

[*] Corresponding author. E-mail address: myai@math.pku.edu.cn (Ai).

1 Introduction

Latin hypercubes (LHs), proposed by McKay, Beckman and Conover (1979), have been widely adopted in the design of computer experiments with quantitative factors because they spread the points uniformly in any one-dimensional projection.

To improve the space-filling property of LHs in two- and more-dimensional projections, Tang (1993) constructed LHs based on an orthogonal array (OA) as follows. Let D be an OA of strength t, N runs and n factors with its i-th factor taking s_i levels from $R_{s_i} = \{0, \ldots, s_i - 1\}$. For each column of D, randomize its symbols and replace the Ns_i^{-1} positions of entry j by a random permutation of $\{jNs_i^{-1}, jNs_i^{-1} + 1, \ldots, (j+1)Ns_i^{-1} - 1\}$, for all $j \in R_{s_i}$. After the procedure is done for all columns of D, the resulting matrix, denoted by L_D, is a D-based LH. It is also called a U design by Tang (1993). Though L_D achieves stratification in all t-dimensional margins when an OA of strength t is employed, it may not be guaranteed to be uniform over the experimental domain.

For comparing the uniformity of designs, various discrepancies have been proposed. Among them, the centered L_2-discrepancy (CD) [Hickernell (1998a)] and wrap-around L_2-discrepancy (WD) [Hickernell (1998b)] are popularly applied. Under the CD and WD, Jiang and Ai (2012) introduce a stochastic algorithm to construct uniform designs without replications. Since there are some unreasonable phenomena associated with the CD and WD, the mixture discrepancy (MD) is proposed by Zhou et al. (2013). A U design is said to be uniform under a discrepancy if it has the smallest discrepancy value among all U designs of the same size.

In this paper, we show that the averaged squared discrepancy of all U designs generated from an OA is a linear combination of the generalized wordtype pattern [Ai and He (2006)] of the OA. Enlightened by this expression, the weighted wordtype pattern is proposed and the corresponding minimum weighted aberration (MWA) criterion is defined to sequentially minimize the components of the weighted wordtype pattern of an OA. U designs generated from an OA with less weighted aberration are shown to have lower discrepancies on average.

The rest of the article is organized as follows. Section 2 introduces the definitions of generalized wordtype pattern and discrepancies. Section 3 develops the relationship between averaged squared discrepancy of all U designs generated from an OA and the generalized wordtype pattern of the OA. Section 4 proposes an algorithm to construct uniform U designs. Comparisons with another two methods are given in Section 5. Section 6 concludes this article with some remarks.

2 Generalized wordtype pattern and discrepancies

Let $R_s = \{0, \ldots, s-1\}$ be the integer ring with modulus s. An orthogonal array (OA), denoted by $OA(N, s_1^{n_1} \cdots s_g^{n_g}, t)$, with N runs, $n = \sum_{i=1}^{g} n_i$ factors and strength t ($n \geq t \geq 1$), is an $N \times n$ matrix in which the first n_1 columns have s_1 levels from a set of s_1 elements, say R_{s_1}, the next n_2 columns have s_2 levels from a set of s_2 elements, say R_{s_2}, and so on, such that every $N \times t$ submatrix contains all possible level combinations as rows with the same frequency. When $s_1 = \cdots = s_g = s$, in particular, this special case is called a symmetric OA and denoted by $OA(N, s^n, t)$; otherwise, it is an asymmetric OA. Typically, an $OA(N, N^n, 1)$ is called a Latin hypercube (LH).

For an $OA(N, s_1^{n_1} \cdots s_g^{n_g}, t)$ D with $n = \sum_{i=1}^{g} n_i$, let $\{\chi_{u_i}^{(i)}, u_i \in R_{s_i}\}$ be the orthonormal contrast coefficients for the s_i-level factors, that is,

$$\sum_{x_i \in R_{s_i}} \chi_{u_i}^{(i)}(x_i)\overline{\chi_{v_i}^{(i)}(x_i)} = s_i \delta_{u_i,v_i}, \text{ for any } u_i, v_i \in R_{s_i}, \tag{1}$$

where $\overline{\chi_{v_i}^{(i)}(x_i)}$ is the complex conjugate of $\chi_{v_i}^{(i)}(x_i)$, $\delta_{u,v}$ equals 1 if $u = v$ and 0 otherwise. Let $\chi_0^{(i)}(x_i) = 1$, for any $x_i \in R_{s_i}$. Let $R = R_{s_1}^{n_1} \times \cdots \times R_{s_g}^{n_g}$. Then we consider the following contrast coefficients defined by tensor products:

$$\chi_{\boldsymbol{u}}(\boldsymbol{x}) = \prod_{i=1}^{n} \chi_{u_i}^{(i)}(x_i), \text{ for } \boldsymbol{u} = (u_1, \ldots, u_n) \in R, \text{ and } \boldsymbol{x} = (x_1, \ldots, x_n) \in R. \tag{2}$$

It can be verified that $\sum_{\boldsymbol{x} \in R} \chi_{\boldsymbol{u}}(\boldsymbol{x})\overline{\chi_{\boldsymbol{v}}(\boldsymbol{x})} = \prod_{i=1}^{g} s_i^{n_i} \delta_{\boldsymbol{u},\boldsymbol{v}}$ for any $\boldsymbol{u}, \boldsymbol{v} \in R$. So $\{\chi_{\boldsymbol{u}}, \boldsymbol{u} \in R\}$

are the orthonormal contrast coefficients.

Ai and He (2006) introduced the generalized wordtype pattern for comparing general factorial designs with multiple groups of factors. For any $\boldsymbol{u} \in R$, split $\boldsymbol{u}$ into g parts, i.e., $\boldsymbol{u} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_g)$, where $\boldsymbol{u}_i \in R_{s_i}^{n_i}$. For $0 \leq j_1 \leq n_1, \ldots, 0 \leq j_g \leq n_g$, define

$$B_{j_1,\ldots,j_g}(D) = N^{-2} \sum_{wt(\boldsymbol{u}_1)=j_1} \cdots \sum_{wt(\boldsymbol{u}_g)=j_g} |\chi_{\boldsymbol{u}}(D)|^2. \tag{3}$$

where $\chi_{\boldsymbol{u}}(D) = \sum_{\boldsymbol{x} \in D} \chi_{\boldsymbol{u}}(\boldsymbol{x})$ and $wt(\boldsymbol{u})$ is the number of nonzero elements of $\boldsymbol{u}$. Clearly, $B_{0,\ldots,0}(D) = 1$. The $B_{j_1,\ldots,j_g}(D)$'s are called the generalized wordtype pattern of the design D. For $j = 0, \ldots, n$, let

$$A_j(D) = \sum_{j_1+\cdots+j_g=j} B_{j_1,\ldots,j_g}(D). \tag{4}$$

Ai and He (2006) further verified that the vector $A(D) = (A_1(D), \ldots, A_n(D))$ is exactly the so called generalized wordlength pattern of design D in Xu and Wu (2001). The generalized minimum aberration (GMA) criterion is to sequentially minimize the components of $A(D)$.

On the other hand, different types of discrepancies have been defined to measure the unifromity of a design. Hickernell (1998a,b) showed that the most widely used discrepancies can be defined by reproducing kernels with the form

$$\mathcal{K}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^{n} f(x_i, y_i), \tag{5}$$

where $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{y} = (y_1, \ldots, y_n)$ and $f(x, y)$ is defined on $[0, 1]^2$. For example, $f(x, y) = 1 + (|x - 0.5| + |y - 0.5| - |x - y|)/2$ for CD, $f(x, y) = 1.5 - |x - y| + |x - y|^2$ for WD and $f(x, y) = 1.875 - (|x - 0.5| + |y - 0.5| + 3|x - y| - 2|x - y|^2)/4$ for MD. For a U design $L = (l_{ij})$ of N runs and n factors, its squared discrepancy defined by the kernel $\mathcal{K}(\boldsymbol{x}, \boldsymbol{y})$ can be expressed as

$$Disc^2(L, \mathcal{K}) = c_0^n - \frac{2}{N} \sum_{i=1}^{N} \prod_{k=1}^{n} f_1(x_{ik}) + \frac{1}{N^2} \sum_{i,j=1}^{N} \prod_{k=1}^{n} f(x_{ik}, x_{jk}), \tag{6}$$

where $x_{ik} = (l_{ik} + 0.5)/N$, $f_1(x) = \int_0^1 f(x,y)dy$ and $c_0 = \int_{[0,1]^2} f(x,y)dxdy$.

Hickernell and Liu (2002) defined the projection discrepancy pattern to evaluate the low-dimensional projection uniformity of a design. Let v be any subset of $\{1,\ldots,n\}$, $\boldsymbol{x}_v$ be the elements of the vector $\boldsymbol{x}$ indexed by the elements of v, L_v be the projection of the design L. Let $\hat{f}(\cdot,\cdot) = f(\cdot,\cdot) - 1$. Define $\hat{\mathcal{K}}_v(\boldsymbol{x}_v, \boldsymbol{y}_v) = \prod_{j\in v} \hat{f}(x_j, y_j)$ and $\hat{\mathcal{K}}_\emptyset = 1$ by convention. For the kernels of form (5), we express

$$\mathcal{K}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^{n} f(x_i, y_i) = \sum_{\emptyset\subseteq v\subseteq R_s} \hat{\mathcal{K}}_v(\boldsymbol{x}_v, \boldsymbol{y}_v) \tag{7}$$

and

$$Disc^2(L, \mathcal{K}) = \sum_{\emptyset\subseteq v\subseteq R_s} Disc^2(L_v, \hat{\mathcal{K}}_v) = \sum_{j=0}^{n} Disc^2_{(j)}(L, \mathcal{K}), \tag{8}$$

where $Disc^2_{(j)}(L, \mathcal{K}) = \sum_{\#v=j} Disc^2(L_v, \hat{\mathcal{K}}_v)$ and $\#v$ is the cardinality of v. Since $\hat{\mathcal{K}}_\emptyset = 1$, it follows that $Disc^2_{(0)}(L, \mathcal{K}) = 0$. Then the projection discrepancy pattern was given by Hickernell and Liu (2002) as follows

$$PD(L, \mathcal{K}) = (Disc^2_{(1)}(L, \mathcal{K}), \ldots, Disc^2_{(n)}(L, \mathcal{K})).$$

For any two U designs L_1 and L_2, L_1 has smaller projection discrepancy pattern than L_2, or equivalently $PD(L_1, \mathcal{K}) < PD(L_2, \mathcal{K})$, if and only if the first, from the left, nonzero component of $PD(L_1, \mathcal{K}) - PD(L_2, \mathcal{K})$ is negative. Throughout, $Disc^2_{(j)}(L, \mathcal{K})$ is called the squared j-dimensional projection discrepancy of the design L.

3 Relationships between generalized wordtype pattern and discrepancies

Tang, Xu and Lin (2012) derived an expression of averaged CD^2 in terms of generalized wordlength pattern for three-level designs. Later, Tang and Xu (2013) and Zhou and Xu (2014) generalized their result to symmetric designs. Here we establish the relationship between averaged squared discrepancy of all U designs generated from an OA D and the generalized wordtype pattern of D.

Suppose D is an $OA(N, s_1^{n_1} \cdots s_g^{n_g}, t)$ with $n = \sum_{k=1}^{g} n_k$ and $\mathcal{K}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^{n} f(x_i, y_i)$. Let $\mathcal{L}_D$ be the collection of all U designs generated from D. Let

$$\overline{Disc^2}(\mathcal{L}_D, \mathcal{K}) = \prod_{k=1}^{g} [((N/s_k)!)^{s_k} s_k!]^{-n_k} \sum_{L \in \mathcal{L}_D} Disc^2(L, \mathcal{K}) \tag{9}$$

be the averaged squared discrepancy of all U designs generated from D. Let

$$c_1(s) = \sum_{l_1, l_2 \in R_s, l_1 \neq l_2} \sum_{z_1 \in T_{l_1,s}, z_2 \in T_{l_2,s}} f\left(\frac{z_1 + 0.5}{N}, \frac{z_2 + 0.5}{N}\right) \tag{10}$$

and

$$c_2(s) = N(s-1)(N-s)^{-1} (c_1(s))^{-1} \sum_{l \in R_s} \sum_{z_1, z_2 \in T_{l,s}, z_1 \neq z_2} f\left(\frac{z_1 + 0.5}{N}, \frac{z_2 + 0.5}{N}\right), \tag{11}$$

where s is a divisor of N and $T_{l,s} = \{lN/s, lN/s + 1, \cdots, (l+1)N/s - 1\}$. Then we obtain the following theorem which shows that the averaged squared discrepancy of all U designs generated from D can be linearly expressed in terms of the generalized wordtype pattern of D. The detailed proof is postponed to the Appendix.

Theorem 1. *For an $OA(N, s_1^{n_1} \cdots s_g^{n_g}, t)$ D with $n = \sum_{k=1}^{g} n_k$ and a discrepancy kernel $\mathcal{K}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^{n} f(x_i, y_i)$, if $c_2(s_k) > 1$ for $k = 1, \ldots, g$, we have*

$$\begin{aligned} &\overline{Disc^2}(\mathcal{L}_D, \mathcal{K}) = \alpha(N, n_1, \ldots, n_g, s_1, \ldots, s_g) \\ &+ \prod_{k=1}^{g} \left(\frac{c_1(s_k)(c_2(s_k) + s_k - 1)}{N^2(s_k - 1)}\right)^{n_k} \sum_{i_1=0}^{n_1} \cdots \sum_{i_g=0}^{n_g} \prod_{k=1}^{g} \left(\frac{c_2(s_k) - 1}{c_2(s_k) + s_k - 1}\right)^{i_k} B_{i_1,\ldots,i_g}(D), \end{aligned} \tag{12}$$

where $\alpha(N, n_1, \ldots, n_g, s_1, \ldots, s_g) = c_0^n - 2\left(\frac{1}{N}\sum_{l=0}^{N-1} f_1\left(\frac{l+0.5}{N}\right)\right)^n + \frac{1}{N}\left(\frac{1}{N}\sum_{l=0}^{N-1} f\left(\frac{l+0.5}{N}, \frac{l+0.5}{N}\right)\right)^n - \frac{1}{N}\prod_{k=1}^{g}\left(\frac{s_k c_1(s_k) c_2(s_k)}{N^2(s_k - 1)}\right)^{n_k}$, c_0, $f_1(\cdot)$, $c_1(\cdot)$ and $c_2(\cdot)$ are defined in (6),(10) and (11), respectively.

Note that the commonly used discrepancies such as CD, WD and MD satisfy the condition $c_2(s) > 1$ when s is a divisor of N. For example, for the WD, $c_1(s) = N^2(1 - \frac{1}{s})(\frac{4}{3} - \frac{1}{6s} + \frac{1}{6s^2} - \frac{1}{6N^2})$ and $c_2(s) = (9 - \frac{2}{s} + \frac{1}{s^2} - \frac{2}{N} + \frac{1}{Ns})/(8 - \frac{1}{s} + \frac{1}{s^2} - \frac{1}{N^2})$.

Now we consider the projection uniformity of U designs generated from an $OA(N, s_1^{n_1} \cdots s_g^{n_g}, t)$ D. For $j = 1, \ldots, n$, let

$$\overline{Disc^2_{(j)}}(\mathcal{L}_D, \mathcal{K}) = \prod_{k=1}^{g} [((N/s_k)!)^{s_k}\, s_k!]^{-n_k} \sum_{L \in \mathcal{L}_D} Disc^2_{(j)}(L, \mathcal{K}) \tag{13}$$

be the averaged squared j-dimensional projection discrepancy of all U designs generated from D. Let

$$\hat{c}_1(s) = \sum_{l_1, l_2 \in R_s, l_1 \neq l_2} \sum_{z_1 \in T_{l_1,s}, z_2 \in T_{l_2,s}} \hat{f}\left(\frac{z_1 + 0.5}{N}, \frac{z_2 + 0.5}{N}\right) \tag{14}$$

and

$$\hat{c}_2(s) = N(s-1)(N-s)^{-1}(\hat{c}_1(s))^{-1} \sum_{l \in R_s} \sum_{z_1, z_2 \in T_{l,s}, z_1 \neq z_2} \hat{f}\left(\frac{z_1 + 0.5}{N}, \frac{z_2 + 0.5}{N}\right), \tag{15}$$

where $\hat{f}(x, y) = f(x, y) - 1$. For any $v \subseteq \{1, \ldots, n\}$, let $n_k^{(v)}$ be the number of s_k-level factors in the projection D_v. By noting the fact that $\overline{Disc^2_{(j)}}(\mathcal{L}_D, \mathcal{K}) = \sum_{\#v=j} \overline{Disc^2}(\mathcal{L}_{D_v}, \hat{\mathcal{K}}_v)$, we immediately obtain the following result from Theorem 1.

Corollary 1. *For an $OA(N, s_1^{n_1} \cdots s_g^{n_g}, t)$ D with $n = \sum_{i=1}^{g} n_i$ and a discrepancy kernel $\mathcal{K}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^{n} f(x_i, y_i)$, if $\hat{c}_2(s_k) > 1$ for $k = 1, \ldots, g$, we have*

$$\overline{Disc^2_{(j)}}(\mathcal{L}_D, \mathcal{K}) = \sum_{\#v=j} \hat{\alpha}(N, j, n_1^{(v)}, \ldots, n_g^{(v)}, s_1, \ldots, s_g)$$
$$+ \sum_{\#v=j} \prod_{k=1}^{g} \left(\frac{\hat{c}_1(s_k)(\hat{c}_2(s_k) + s_k - 1)}{N^2(s_k - 1)}\right)^{n_k^{(v)}} \sum_{i_1=0}^{n_1^{(v)}} \cdots \sum_{i_g=0}^{n_g^{(v)}} \prod_{k=1}^{g} \left(\frac{\hat{c}_2(s_k) - 1}{\hat{c}_2(s_k) + s_k - 1}\right)^{i_k} B_{i_1,\ldots,i_g}(D_v),$$

where $j = 1, \ldots, n$, $\hat{\alpha}(N, j, n_1^{(v)}, \ldots, n_g^{(v)}, s_1, \ldots, s_g) = \hat{c}_0^j - 2\left(\frac{1}{N}\sum_{l=0}^{N-1} \hat{f}_1\left(\frac{l+0.5}{N}\right)\right)^j + \frac{1}{N}\left(\frac{1}{N}\sum_{l=0}^{N-1} \hat{f}\left(\frac{l+0.5}{N}, \frac{l+0.5}{N}\right)\right)^j - \frac{1}{N}\prod_{k=1}^{g}\left(\frac{s_k \hat{c}_1(s_k)\hat{c}_2(s_k)}{N^2(s_k-1)}\right)^{n_k^{(v)}}$, $\hat{f}_1(x) = \int_0^1 \hat{f}(x, y)dy$ and $\hat{c}_0 =$

$\int_{[0,1]^2} \hat{f}(x,y)dxdy$.

Note that the commonly used discrepancies such as CD, WD and MD also satisfy the condition $\hat{c}_2(s) > 1$ when s is a divisor of N. For example, for the WD, $\hat{c}_1(s) = N^2(1-\frac{1}{s})(\frac{1}{3}-\frac{1}{6s}+\frac{1}{6s^2}-\frac{1}{6N^2})$ and $\hat{c}_2(s) = (3-\frac{2}{s}+\frac{1}{s^2}-\frac{2}{N}+\frac{1}{Ns})/(2-\frac{1}{s}+\frac{1}{s^2}-\frac{1}{N^2})$.

When D is a symmetric $OA(N, s^n, t)$, combining equation (4), Theorem 1 and Corollary 1 with the fact that $\sum_{\#v=j} A_i(D_v) = \binom{n-i}{j-i} A_i(D)$, we obtain the following result which shows that the averaged squared discrepancy and averaged squared j-dimensional projection discrepancy of all U designs generated from D are linearly expressed in terms of the generalized wordlength pattern of D.

Corollary 2. *For an $OA(N, s^n, t)$ D and a discrepancy kernel $\mathcal{K}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^n f(x_i, y_i)$, if $\hat{c}_2(s) > 1$, we have*

$$\overline{Disc^2}(\mathcal{L}_D, \mathcal{K}) = \alpha(N, n, s) + \left(\frac{c_1(s)(c_2(s)+s-1)}{N^2(s-1)}\right)^n \sum_{i=0}^n \left(\frac{c_2(s)-1}{c_2(s)+s-1}\right)^i A_i(D)$$

and

$$\begin{aligned}\overline{Disc^2_{(j)}}(\mathcal{L}_D, \mathcal{K}) &= \binom{n}{j} \hat{\alpha}(N, n, s, j) \\ &+ \left(\frac{\hat{c}_1(s)(\hat{c}_2(s)+s-1)}{N^2(s-1)}\right)^j \sum_{i=0}^j \left(\frac{\hat{c}_2(s)-1}{\hat{c}_2(s)+s-1}\right)^i \binom{n-i}{j-i} A_i(D),\end{aligned}$$

where $\alpha(N, n, s) = c_0^n - 2\left(\frac{1}{N}\sum_{l=0}^{N-1} f_1\left(\frac{l+0.5}{N}\right)\right)^n + \frac{1}{N}\left(\frac{1}{N}\sum_{l=0}^{N-1} f\left(\frac{l+0.5}{N}, \frac{l+0.5}{N}\right)\right)^n - \frac{1}{N}\left(\frac{sc_1(s)c_2(s)}{N^2(s-1)}\right)^n$, c_0 and $f_1(\cdot)$ are defined in (6), and for $j = 1, \ldots, n$, $\hat{\alpha}(N, n, s, j) = \hat{c}_0^j - 2\left(\frac{1}{N}\sum_{l=0}^{N-1} \hat{f}_1\left(\frac{l+0.5}{N}\right)\right)^j + \frac{1}{N}\left(\frac{1}{N}\sum_{l=0}^{N-1} \hat{f}\left(\frac{l+0.5}{N}, \frac{l+0.5}{N}\right)\right)^j - \frac{1}{N}\left(\frac{s\hat{c}_1(s)\hat{c}_2(s)}{N^2(s-1)}\right)^j$, $\hat{c}_0$ and $\hat{f}_1(\cdot)$ are defined in Corollary 1.

4 An algorithm for constructing uniform U designs

Suppose D is an $OA(N, s_1^{n_1}\cdots s_g^{n_g}, t)$ with $n=\sum_{i=1}^g n_i$, $\mathcal{K}(\boldsymbol{x},\boldsymbol{y})=\prod_{i=1}^n f(x_i,y_i)$. Inspired by Theorem 1, we define the weighted wordtype pattern as

$$WB_j(D)=\frac{1}{C_j}\sum_{i_1+\ldots+i_g=j}\prod_{k=1}^{g}\left(\frac{c_2(s_k)-1}{c_2(s_k)+s_k-1}\right)^{i_k}B_{i_1,\ldots,i_g}(D), \tag{16}$$

where $j=0,\ldots,n$, $C_j=\frac{1}{m_j}\sum_{i_1+\ldots,i_g=j}\prod_{k=1}^g\left(\frac{c_2(s_k)-1}{c_2(s_k)+s_k-1}\right)^{i_k}$, $m_j=\#\{(i_1,\ldots,i_g):\sum_{k=1}^g i_k=j, i_k=0,\ldots,n_k, k=1,\ldots,g\}$. Clearly, $WB_0(D)=1$. The minimum weighted aberration (MWA) criterion is to sequentially minimize the vector $WB(D)=(WB_1(D),\ldots,WB_n(D))$. It should be mentioned that when D is symmetric, $WB_j(D)=A_j(D)$ for $j=1,\ldots,n$, and the MWA criterion reduces to the GMA criterion.

Assume that $c_2(s_k)>1$ for $k=1,\ldots,g$. Then equation (12) in Theorem 1 can be rewritten as

$$\begin{aligned}\overline{Disc^2}(\mathcal{L}_D,\mathcal{K})&=\alpha(N,n_1,\ldots,n_g,s_1,\ldots,s_g)\\&+\prod_{k=1}^{g}\left(\frac{c_1(s_k)(c_2(s_k)+s_k-1)}{N^2(s_k-1)}\right)^{n_k}\sum_{i=0}^{n}C_i\cdot WB_i(D).\end{aligned} \tag{17}$$

The coefficient C_i of $WB_i(D)$ decreases geometrically as i increases, thus to minimize the averaged squared discrepancy of all U designs based on an OA tends to agree with the MWA criterion for the corresponding OA.

Example 1. *Suppose the discrepancy measure CD is used. Let D_0 be the $OA(36, 2^{11}3^{12}, 2)$ on Sloane's website (http:// neilsloane.com) whose design matrix is given in transpose in Table 1. Consider selecting a subarray D with parameters $OA(36, 2^3 3^3, 2)$ from D_0. For each subarray D, let $\overline{CD^2}(\mathcal{L}_D)$ be the averaged CD^2 values of all U designs generated from D. There are* 36300 *subarrays with parameters $OA(36, 2^3 3^3, 2)$ in D_0. Among the generalized wordtype patterns corresponding to these subarrays,* 674 *are distinct with each other. We sort these* 674 *representative subarrays based on $\overline{CD^2}(\mathcal{L}_D)$, MWA and GMA, and denote the ranks* by r_{CD}, r_{WB} and r_A, respectively. Table 2 tabulates the values r_{CD}, $\overline{CD^2}(\mathcal{L}_D)$, r_{WB}, r_A, as well as one representative of the columns, and part of the generalized wordtype patterns. Here $B_{0,1} = B_{1,0} = B_{0,2} = B_{1,1} = B_{2,0} = 0$ for all subarrays. It can be seen from Table 2 that compared with r_A, the r_{WB}'s are more consistent with r_{CD}'s, that is, less weighted aberration leads to lower $\overline{CD^2}(\mathcal{L}_D)$ values.

Let D_1 be subarray of D_0 with $r_{WB} = 1$ and D_2 be the subarray of D_0 with $r_{WB} = 600$. By randomly generating 10^5 U designs based on D_1 and D_2, respectively, two histograms of the CD^2 values of U designs in each case are obtained and shown in Figure 1. It can be seen that U designs based on D_1 tend to have lower CD^2 values, which agrees with the above conclusion.

Table 1: Design matrix of an $OA(36, 2^{11}3^{12}, 2)$

col	Design matrix in transpose
1	0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1
2	0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0
3	0 1 0 1 0 0 1 1 1 0 1 0 0 1 0 1 0 0 1 1 1 0 1 0 0 1 0 1 0 0 1 1 1 0 1 0
4	0 0 1 1 1 0 0 1 0 1 1 0 0 0 1 1 1 0 0 1 0 1 1 0 0 0 1 1 1 0 0 1 0 1 1 0
5	0 0 1 1 0 1 0 0 1 0 1 1 0 0 1 1 0 1 0 0 1 0 1 1 0 0 1 1 0 1 0 0 1 0 1 1
6	0 0 1 0 0 0 1 1 1 1 0 1 0 0 1 0 0 0 1 1 1 1 0 1 0 0 1 0 0 0 1 1 1 1 0 1
7	0 1 1 1 1 0 1 0 0 0 0 1 0 1 1 1 1 0 1 0 0 0 0 1 0 1 1 1 1 0 1 0 0 0 0 1
8	0 0 0 0 1 1 1 1 0 0 1 1 0 0 0 0 1 1 1 1 0 0 1 1 0 0 0 0 1 1 1 1 0 0 1 1
9	0 1 0 0 1 0 0 0 1 1 1 1 0 1 0 0 1 0 0 0 1 1 1 1 0 1 0 0 1 0 0 0 1 1 1 1
10	0 0 0 1 1 1 1 0 1 1 0 0 0 0 0 1 1 1 1 0 1 1 0 0 0 0 0 1 1 1 1 0 1 1 0 0
11	0 1 1 0 1 1 0 1 1 0 0 0 0 1 1 0 1 1 0 1 1 0 0 0 0 1 1 0 1 1 0 1 1 0 0 0
12	1 1 2 2 1 1 2 2 1 1 2 2 2 2 0 0 2 2 0 0 2 2 0 0 0 0 1 1 0 0 1 1 0 0 1 1
13	1 1 2 2 2 2 1 1 2 2 1 1 2 2 0 0 0 0 2 2 0 0 2 2 0 0 1 1 1 1 0 0 1 1 0 0
14	2 2 1 1 1 1 2 2 2 2 1 1 0 0 2 2 2 2 0 0 0 0 2 2 1 1 0 0 0 0 1 1 1 1 0 0
15	2 2 1 1 2 2 1 1 1 1 2 2 0 0 2 2 0 0 2 2 2 2 0 0 1 1 0 0 1 1 0 0 0 0 1 1
16	1 2 1 2 2 0 2 0 0 1 0 1 2 0 2 0 0 1 0 1 1 2 1 2 0 1 0 1 1 2 1 2 2 0 2 0
17	1 2 1 2 0 2 0 2 1 0 1 0 2 0 2 0 1 0 1 0 2 1 2 1 0 1 0 1 2 1 2 1 0 2 0 2
18	2 1 2 1 2 0 2 0 1 0 1 0 0 2 0 2 0 1 0 1 2 1 2 1 1 0 1 0 1 2 1 2 0 2 0 2
19	2 1 2 1 0 2 0 2 0 1 0 1 0 2 0 2 1 0 1 0 1 2 1 2 1 0 1 0 2 1 2 1 2 0 2 0
20	1 2 2 1 0 1 1 0 2 0 0 2 2 0 0 2 1 2 2 1 0 1 1 0 0 1 1 0 2 0 0 2 1 2 2 1
21	1 2 2 1 1 0 0 1 0 2 2 0 2 0 0 2 2 1 1 2 1 0 0 1 0 1 1 0 0 2 2 0 2 1 1 2
22	2 1 1 2 0 1 1 0 0 2 2 0 0 2 2 0 1 2 2 1 1 0 0 1 1 0 0 1 2 0 0 2 2 1 1 2
23	2 1 1 2 1 0 0 1 2 0 0 2 0 2 2 0 2 1 1 2 0 1 1 0 1 0 0 1 0 2 2 0 1 2 2 1

Table 2: Comparison between MWA and GMA

r_{CD}	$\overline{CD^2}(\mathcal{L}_D)$	r_{WB}	r_A	columns	$(B_{03}, B_{12}, B_{21}, B_{30}, B_{13}, B_{22}, B_{31}, B_{23}, B_{32}, B_{33})$
1	0.015006	1	2	1 2 4 17 18 19	0.50, 0, 0, 0.11, 1.50, 1.33, 0, 0.17, 1.33, 0.06
2	0.015032	2	1	2 4 7 16 17 20	0.13, 0.33, 0, 0.11, 1.54, 1.33, 0, 0.54, 1.00, 0.01
3	0.015079	3	3	1 2 8 12 15 20	0.13, 0.50, 0, 0.11, 0.88, 1.67, 0, 1.21, 0.50, 0.01
4	0.015101	4	7	1 2 3 13 14 15	0.50, 0.33, 0, 0.11, 0.17, 2.00, 0, 1.50, 0.33, 0.06
5	0.015108	5	11	1 2 6 12 13 14	0.50, 0.33, 0, 0.11, 0.67, 2.00, 0, 0.50, 0.67, 0.22
6	0.015115	6	10	1 2 3 12 13 14	0.50, 0.33, 0, 0.11, 0.17, 2.67, 0, 0.83, 0.33, 0.06
7	0.015119	8	9	1 7 9 12 13 16	0.13, 0.67, 0, 0.11, 1.21, 1.00, 0, 0.88, 1.00, 0.01
8	0.015129	7	12	1 2 3 12 13 15	0.50, 0.33, 0, 0.11, 0.17, 3.33, 0, 0.17, 0.33, 0.06
9	0.015130	9	13	2 4 5 12 13 20	0.13, 0.67, 0, 0.11, 1.21, 1.50, 0, 0.88, 0.50, 0.01
10	0.015130	10	14	6 7 9 15 17 21	0.13, 0.67, 0, 0.11, 1.21, 1.50, 0, 0.38, 0.83, 0.18
11	0.015130	11	15	2 4 6 12 13 20	0.13, 0.67, 0, 0.11, 0.96, 1.67, 0, 0.71, 0.67, 0.10
12	0.015133	13	4	2 4 11 12 13 20	0.13, 0.67, 0, 0.11, 0.96, 1.83, 0, 0.54, 0.50, 0.26
13	0.015133	12	5	6 8 9 15 17 21	0.13, 0.67, 0, 0.11, 1.21, 1.67, 0, 0.21, 1.00, 0.01
14	0.015137	14	6	2 5 6 12 13 20	0.13, 0.67, 0, 0.11, 0.71, 2.17, 0, 0.71, 0.50, 0.01
15	0.015147	15	8	2 6 11 12 13 20	0.13, 0.67, 0, 0.11, 0.46, 2.83, 0, 0.54, 0.17, 0.10
16	0.015155	16	19	1 2 5 17 18 19	0.50, 0.50, 0, 0.11, 1.50, 1.33, 0, 0.17, 0.83, 0.06
17	0.015156	18	21	1 2 7 12 16 20	0.88, 0.33, 0, 0.11, 0.79, 0.67, 0, 0.46, 1.67, 0.10
18	0.015169	17	22	1 2 7 12 13 14	0.50, 0.50, 0, 0.11, 0, 3.00, 0, 0.50, 0.17, 0.22
19	0.015173	20	25	2 5 7 12 13 20	0.13, 0.83, 0, 0.11, 0.79, 1.50, 0, 0.88, 0.67, 0.10
20	0.015177	19	24	1 2 4 12 16 20	0.88, 0.33, 0, 0.11, 0.79, 1.67, 0, 0.46, 0.67, 0.10

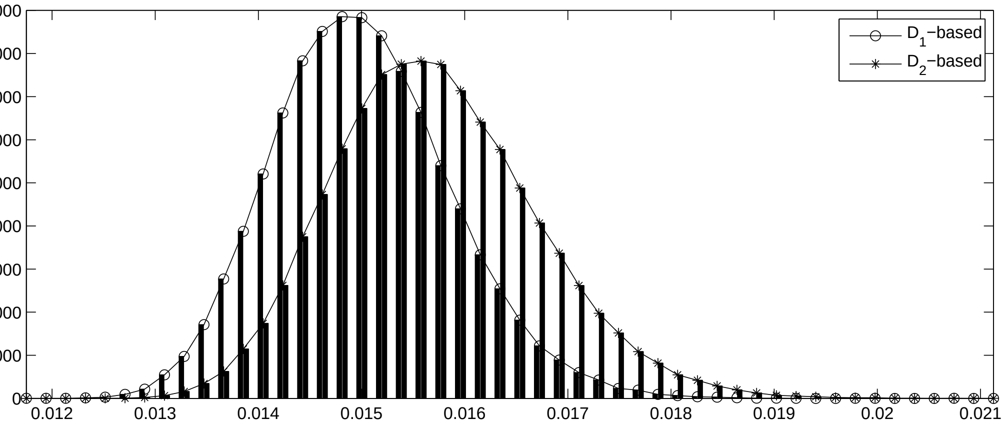

Figure 1: Two histograms of CD^2 values of U designs generated from D_1 and D_2.

By using the MWA criterion, we modify the threshold accepting (TA) algorithm in Fang, Maringer, Tang and Winker (2006) to search uniform or nearly uniform U designs. For an $OA(N, s_1^{n_1} \cdots s_g^{n_g}, t)$ D, let $\mathcal{N}_D(L_D)$ be the neighborhood set of the U design L_D. Let M be any discrepancy measure including CD, WD and MD. Let τ be the number of iterations, and $T_1, \ldots, T_\tau$ be the sequence of the thresholds. As done in Fang, Maringer, Tang and Winker (2006), choose $\tau = u\lambda$, where u, λ are positive integers. Then the sequence of the thresholds are determined as follows: $T_1 = \cdots = T_u$, $T_{u+1} = \cdots = T_{2u} = \gamma T_1$, $T_{2u+1} = \cdots = T_{3u} = \gamma^2 T_1$, and so on, where $\gamma = (T_\tau/T_1)^{\frac{1}{\lambda-1}}$. The values $T_1 = 0.01$ and $T_\tau = 10^{-6}$ are chosen for the initial and terminal thresholds, respectively. The modified TA algorithm, named MWA-based method, for construction of uniform U designs is summarized in Algorithm 1.

Algorithm 1 The MWA-based construction method of uniform U designs based on an $OA(N, s_1^{n_1} \cdots s_g^{n_g}, t)$

1: Initialize τ and the sequence of thresholds $T_1, \ldots, T_\tau \in (0, 1)$.
2: Use the MWA criterion to select a subarray D with parameters $OA(N, s_1^{n_1} \cdots s_g^{n_g}, t)$ from an existing OA.
3: Randomly generate a U design L_D based on D. Let $L_D^{\min} := L_D$.
4: **for** $i = 1$ to τ **do**
5: Generate a new design $\widetilde{L_D} \in \mathcal{N}_D(L_D)$ and compute $\nabla_M = M^2(\widetilde{L_D}) - M^2(L_D)$.
6: **if** $\nabla_M < M^2(L_D) \cdot T_i$, **then** $L_D = \widetilde{L_D}$.
7: **if** $M^2(L_D) < M^2(L_D^{\min})$, **then** $L_D^{\min} := L_D$.
8: **end for**

Algorithm 1 differs from the TA algorithm in Fang, Maringer, Tang and Winker (2006) in two aspects. The first difference lies in the selection of starting designs. Instead of randomly generating an LH, the proposed algorithm randomly chooses a U design based on an MWA OA D to conduct the subsequential iterations. The second difference is that the neighborhood set of the current solution, i.e., $\mathcal{N}_D(L_D)$, is defined based on D. Here, two U designs are called neighboring based on D if they are both U designs generated from D and one can be obtained from the other by simply exchanging two entries within the same column. Consequently, all the solutions are U designs based on D during the iteration. This appears to be a crucial item, as equation (17) implies that U designs generated from an OA with less weighted aberration tend to have smaller squared discrepancies. Moreover, this helps reduce the computational complexity, in spite of the restriction on the structure of the solutions.

5 Performance analysis

If the $OA(N, s_1^{n_1} \cdots s_g^{n_g}, t)$ D in Step 2 of Algorithm 1 is generated randomly from an existing OA instead, then Algorithm 1 becomes another construction method called non-MWA-based method in this article. To illustrate the performance of the MWA-based method, the comparisons of the MWA-based method with the non-MWA-based method and the TA algorithm are performed. The calculation is done by using the Matlab program on a personal computer with 3.20 GHz CPU processor and 4 Gb memory.

Here, we construct uniform U designs based on asymmetric $OA(36, 2^{n-1}3, t)$'s for $n = 3$–6. Let D_0 be the $OA(36, 2^{11}3^{12}, 2)$ given in Example 1. When the discrepancy measure CD, WD or MD is used, it can be verified that for $n = 3$–6, the first $n-1$ columns of columns labeled by $1, 2, 3, 4, 5$ together with the column labeled by 12 form an MWA subarray with parameters $OA(36, 2^{n-1}3, 2)$. As explained earlier, we choose an MWA subarray and a random subarray of D_0 to be the initial D's in Algorithm 1 for the MWA-based and non-MWA-based methods, respectively.

As shown in Winker and Fang (1997), for improving the efficiency of stochastic algorithm, the number of inner iterations u may follow in 10^4–10^5 and the number of outer iterations λ may follow in 10–100. In Tables 3, 4 and 5, the MWA-based, non-MWA-based and TA methods are performed with the set $u = 10^4$ and $\lambda = 100$. The detailed formula of ∇_{CD} can be found in Fang, Maringer, Tang and Winker (2006), and that of ∇_{WD} or ∇_{MD} can be similarly derived.

For each case, each method is repeated 30 times. Let ACD^2, AWD^2 and AMD^2 denote the averaged CD^2, WD^2 and MD^2, respectively, of the 30 uniform U designs searched by each method. Tables 3, 4 and 5 tabulate the values of ACD^2, AWD^2 and AMD^2, respectively, and total running time in seconds for the MWA-based, non-MWA-based and TA methods.

It can be seen from Tables 3, 4 and 5 that with the same number of iterations, the

Table 3: Comparison under the CD when $N = 36$

n	MWA-based		Non-MWA-based		TA	
	ACD^2	Time(s)	ACD^2	Time(s)	ACD^2	Time(s)
3	0.000740	786.3	0.000741	767.0	0.000742	534.8
4	0.001822	760.2	0.001822	767.2	0.001824	577.8
5	0.003864	771.3	0.003886	761.1	0.003872	539.9
6	0.007453	764.1	0.007579	811.4	0.007365	602.7

Table 4: Comparison under the WD when $N = 36$

n	MWA-based		Non-MWA-based		TA	
	AWD^2	Time(s)	AWD^2	Time(s)	AWD^2	Time(s)
3	0.001830	635.81	0.001832	631.91	0.001838	532.74
4	0.004959	641.73	0.004968	636.73	0.004953	529.28
5	0.011862	640.20	0.011882	638.18	0.011856	533.49
6	0.025920	643.52	0.025926	639.87	0.025916	539.25

Table 5: Comparison under the MD when $N = 36$

n	MWA-based		Non-MWA-based		TA	
	AMD^2	Time(s)	AMD^2	Time(s)	AMD^2	Time(s)
3	0.001818	787.01	0.001819	778.96	0.001821	559.51
4	0.005802	774.64	0.005814	773.56	0.005812	559.73
5	0.016471	777.31	0.016476	775.65	0.016497	562.96
6	0.042811	797.99	0.042831	784.97	0.042734	566.63

TA takes less running time. However, the MWA-based method tends to get the smallest averaged squared discrepancies, especially under the CD and MD. For example, the MWA-based method obtains the smallest ACD^2 and AWD^2 values in three of all four cases. For the WD criterion, the three methods perform similarly.

When the number of outer iterations λ or the number of inner iterations u for the TA is increased, the TA and the MWA-based methods run for similar length of time. In this case, the TA method performs slightly better than the MWA method, as is shown in Table 6. However, the results obtained by the MWA-based method achieve stratification in two- or more-dimensional projections, which can not be guaranteed by the TA method.

Table 6: Comparison when $N = 36$ and TA takes more iterations

	MWA-based $\lambda = 100, u = 1e4$		TA $\lambda = 100, u = 1.4e4$		TA $\lambda = 140, u = 1e4$	
n	ACD^2	Time(s)	ACD^2	Time(s)	ACD^2	Time(s)
3	0.000740	786.3	0.000742	767.0	0.000743	769.1
4	0.001822	760.2	0.001821	766.0	0.001815	777.3
5	0.003864	771.3	0.003853	766.2	0.003856	773.1
6	0.007453	764.1	0.007337	767.2	0.007333	778.6

Table 7: Improvement of existing uniform LHs under the CD

N	n	Prev.CD^2	CD^2	N	n	Prev.CD^2	CD^2	N	n	Prev.CD^2	CD^2
16	3	0.003194	0.003172	20	4	0.004860	0.004804	20	18	1.519361	1.516423
16	4	0.007071	0.006770	20	5	0.009783	0.009687	24	3	0.001527	0.001512
16	5	0.013909	0.013640	20	6	0.017808	0.017723	24	4	0.003582	0.003543
16	6	0.025049	0.024809	20	7	0.030373	0.030191	24	5	0.007313	0.007225
16	7	0.042079	0.041950	20	8	0.048886	0.048647	24	6	0.013532	0.013384
16	10	0.154901	0.154838	20	9	0.076302	0.075579	24	7	0.023162	0.023063
16	11	0.225217	0.224228	20	13	0.336067	0.335083	24	8	0.037764	0.037361
16	12	0.321534	0.318356	20	14	0.464123	0.463347	24	9	0.058905	0.058083
16	13	0.450431	0.449135	20	15	0.634151	0.632837	24	10	0.089478	0.087653
16	14	0.622102	0.620562	20	16	0.857695	0.855298	24	11	0.131500	0.130771
20	3	0.002109	0.002102	20	17	1.146821	1.143971				

Table 7 tabulates the best CD^2 values obtained by the MWA-based method and the results of existing LHs on the UD website (http://www.uic.edu.hk/isci/). The columns "Prev.CD^2" indicate the CD^2 values previously listed on the UD website, while the columns "CD^2" indicate the results obtained by our code. Here, the best CD^2 values are obtained by using the MWA-based method with the set $u = 10^5$ and $\lambda = 100$ and the OAs "oa.16.15.2.2.0","oa.20.19.2.2.toniii" and "oa.24.12.2.3" on Sloane's website.

6 Conclusions

We show in this paper that the averaged squared discrepancy of all U designs based on an OA is a linear combination of the generalized wordtype pattern of the OA. Based on this expression, we define a weighted wordtype pattern and a corresponding MWA criterion to compare OAs of the same parameters. For symmetric OAs in particular, the MWA reduces to the GMA criterion. U designs generated from an MWA design are shown to have low discrepancies on average. Thus, an algorithm to construct uniform U designs is proposed. Subsequently it is possible to obtain many new uniform U designs, which have smaller squared discrepancies than the existing LHs on the UD website.

Appendix

To prove Theorem 1, we need the following lemma.

Lemma 1. *For an $OA(N, s_1^{n_1} \cdots s_g^{n_g}, t)$ D with $n = \sum_{k=1}^g n_k$, denote $\boldsymbol{x}_i = (\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{ig})$ as its ith design point, where $\boldsymbol{x}_{ik} \in R_{s_k}^{n_k}$ for $k = 1, \ldots, g$. Let δ_{ijk} be the number of positions where $\boldsymbol{x}_{ik}$ and $\boldsymbol{x}_{jk}$ take the same value. Then for any numbers $z_1, \ldots, z_g$ greater than 1,*

$$\sum_{i,j=1}^{N} \prod_{k=1}^{g} z_k^{\delta_{ijk}} = N^2 \prod_{k=1}^{g} \left(\frac{z_k + s_k - 1}{s_k} \right)^{n_k} \sum_{i_1=0}^{n_1} \cdots \sum_{i_g=0}^{n_g} \prod_{k=1}^{g} \left(\frac{z_k - 1}{z_k + s_k - 1} \right)^{i_k} B_{i_1,\ldots,i_g}(D).$$

Proof. The multiple distance distribution of D is defined by

$$B'_{l_1,\ldots,l_g}(D) \;=\; N^{-1} \#\{(\boldsymbol{x}_i, \boldsymbol{x}_j) : d_H(\boldsymbol{x}_{ik}, \boldsymbol{x}_{jk}) = l_k, \text{ for } k = 1, \ldots, g.\},$$

where for $k = 1, \ldots, g$, $l_k \in \{0, \ldots, n_k\}$ and $d_H(\boldsymbol{x}_{ik}, \boldsymbol{x}_{jk}) = n_k - \delta_{ijk}$ is the Hamming distance of $\boldsymbol{x}_{ik}$ and $\boldsymbol{x}_{jk}$. Sloane and Stufken (1996) showed that the generalized wordtype pattern $B_{j_1,\ldots,j_g}(D)$ equals the MacWilliams transforms of the multiple distance distribution, that is,

$$B_{j_1,\ldots,j_g}(D) = N^{-1} \sum_{i_1=0}^{n_1} \cdots \sum_{i_g=0}^{n_g} B'_{i_1,\ldots,i_g}(D) \prod_{k=1}^{g} P_{j_k}(i_k; n_k, s_k), \tag{18}$$

where $j_i \in \{0, \ldots, n_i\}$ for $i = 1, \ldots, g$, $P_j(k; n, s) = \sum_{i=0}^{j} (-1)^i (s-1)^{j-i} \binom{k}{i} \binom{n-k}{j-i}$ are the Krawtchouk polynomials. By the orthogonality of the Krawtchouk polynomials, it is easy to show that

$$B'_{j_1,\ldots,j_g}(D) = N \left(\prod_{k=1}^{n} s_k^{-n_k} \right) \sum_{i_1=0}^{n_1} \cdots \sum_{i_g=0}^{n_g} B_{i_1,\ldots,i_g}(D) \prod_{k=1}^{g} P_{j_k}(i_k; n_k, s_k),$$

for $j_i \in \{0, \ldots, n_i\}$, $i = 1, \ldots, g$. According to the definition of distance distribution and Lemma A.1 in Tang, Xu and Lin (2012), we have

$$\begin{aligned}
\sum_{i,j=1}^{N} \prod_{k=1}^{g} z_k^{\delta_{ijk}} &= N \sum_{j_1=0}^{n_1} \cdots \sum_{j_g=0}^{n_g} B'_{j_1,\ldots,j_g}(D) \prod_{k=1}^{g} z_k^{n_k - j_k} \\
&= N^2 \prod_{k=1}^{g} \left(z_k s_k^{-1} \right)^{n_k} \sum_{j_1=0}^{n_1} \cdots \sum_{j_g=0}^{n_g} \sum_{i_1=0}^{n_1} \cdots \sum_{i_g=0}^{n_g} B_{i_1,\ldots,i_g}(D) \prod_{k=1}^{g} P_{j_k}(i_k; n_k, s_k) z_k^{-j_k} \\
&= N^2 \prod_{k=1}^{g} \left(z_k s_k^{-1} \right)^{n_k} \sum_{i_1=0}^{n_1} \cdots \sum_{i_g=0}^{n_g} B_{i_1,\ldots,i_g}(D) \prod_{k=1}^{g} \left[\left(1 + \frac{s_k - 1}{z_k} \right)^{n_k - i_k} \left(1 - \frac{1}{z_k} \right)^{i_k} \right] \\
&= N^2 \prod_{k=1}^{g} \left(\frac{z_k + s_k - 1}{s_k} \right)^{n_k} \sum_{i_1=0}^{n_1} \cdots \sum_{i_g=0}^{n_g} B_{i_1,\ldots,i_g}(D) \prod_{k=1}^{g} \left(\frac{z_k - 1}{z_k + s_k - 1} \right)^{i_k}.
\end{aligned}$$

□

Proof of Theorem 1. For any $OA(N, s_1^{n_1} \cdots s_g^{n_g}, t)$ D with $n = \sum_{i=1}^{g} n_i$ and any $L \in \mathcal{L}_D$, let x_{ik} and l_{ik} denote their (i,k)-th entries, respectively. First,

$$
\begin{aligned}
& \sum_{L \in \mathcal{L}_D} \sum_{i=1}^{N} \prod_{k=1}^{n} f_1\left(\frac{l_{ik}+0.5}{N}\right) = \sum_{i=1}^{N} \sum_{L \in \mathcal{L}_D} \prod_{k=1}^{n} f_1\left(\frac{l_{ik}+0.5}{N}\right) \\
= \;& N\left[\sum_{l=0}^{N-1} f_1\left(\frac{l+0.5}{N}\right)\right]^n \prod_{k=1}^{g}\left[\left(\frac{N}{s_k}!\right)^{s_k-1}\left(\frac{N}{s_k}-1\right)!(s_k-1)!\right]^{n_k} \qquad (19)
\end{aligned}
$$

only depends on $N, n_1, \ldots, n_g$ and $s_1, \ldots, s_g$. Similarly, we obtain

$$
\begin{aligned}
& \sum_{L \in \mathcal{L}_D} \sum_{i=1}^{N} \prod_{k=1}^{n} f\left(\frac{l_{ik}+0.5}{N}, \frac{l_{ik}+0.5}{N}\right) \\
= \;& N\left[\sum_{l=0}^{N-1} f\left(\frac{l+0.5}{N}, \frac{l+0.5}{N}\right)\right]^n \prod_{k=1}^{n}\left[\left(\frac{N}{s_k}!\right)^{s_k-1}\left(\frac{N}{s_k}-1\right)!(s_k-1)!\right]^{n_k}. \qquad (20)
\end{aligned}
$$

At last,

$$
\begin{aligned}
& \sum_{L \in \mathcal{L}_D} \sum_{i,j=1, i \neq j}^{N} \prod_{k=1}^{n} f\left(\frac{l_{ik}+0.5}{N}, \frac{l_{jk}+0.5}{N}\right) \\
= \;& \sum_{i,j=1, i\neq j}^{N} \sum_{L \in \mathcal{L}_D}\left[\prod_{k: x_{ik}=x_{jk}} f\left(\frac{l_{ik}+0.5}{N}, \frac{l_{jk}+0.5}{N}\right)\right]\left[\prod_{k: x_{ik}\neq x_{jk}} f\left(\frac{l_{ik}+0.5}{N}, \frac{l_{jk}+0.5}{N}\right)\right] \\
= \;& \sum_{i,j=1, i\neq j}^{N} \prod_{k=1}^{g}\left[a_{k1}(s_k-1)! \sum_{l \in R_{s_k}} b_{k,l,l}\right]^{\delta_{ijk}}\left[a_{k2}(s_k-2)! \sum_{l_1 \neq l_2 \in R_{s_k}} b_{k,l_1,l_2}\right]^{n_k-\delta_{ijk}} \\
= \;& \prod_{k=1}^{g}[a_{k2}(s_k-2)!c_1(s_k)]^{n_k} \sum_{i,j=1}^{N} \prod_{k=1}^{g} c_2(s_k)^{\delta_{ijk}} - N\prod_{k=1}^{g}[a_{k2}(s_k-2)!c_1(s_k)c_2(s_k)]^{n_k}, \qquad (21)
\end{aligned}
$$

where for $k = 1, \ldots, g$, $a_{k1} = \left(\frac{N}{s_k}!\right)^{s_k-1}\left(\frac{N}{s_k}-2\right)!$, $a_{k2} = \left(\frac{N}{s_k}!\right)^{s_k-2}\left(\frac{N}{s_k}-1\right)!\left(\frac{N}{s_k}-1\right)!$, $b_{k,l_1,l_2} = \sum_{z_1 \in T_{l_1,s_k}, z_2 \in T_{l_2,s_k}, z_1 \neq z_2} f(\frac{z_1+0.5}{N}, \frac{z_2+0.5}{N})$, $T_{l,s} = \{lN/s, lN/s+1, \cdots, (l+1)N/s-1\}$, $c_1(\cdot)$ and $c_2(\cdot)$ are defined in (10) and (11), respectively. Combining equations (19), (20), (21) with Lemma 1, the desired result follows.

References

Ai, M. and He, S. (2006). Generalized wordtype pattern for nonregular factorial designs with multiple groups of factors. *Metrika*, **64**, 95-108.

Fang, K., Maringer, D., Tang, Y. and Winker, P. (2006) Lower bounds and stochastic optimization algorithms for uniform designs with three or four levels. *Math. Comp.*, **75**, 859-878.

Hickernell, F. J. (1998a) A generalized discrepancy and quadrature error bound. *Math. Comp.*, **67**, 299-322.

Hickernell, F. J. (1998b) Lattice rules: How well do they measure up?, in P. Hellekalek and G. Larcher, *Random and Quasi-Random Point Sets*, Springer-Verlag, New York (1998b) 109-166.

Hickernell, F. J. and Liu, M. (2002). Uniform designs limit aliasing. *Biometrika*, **89**, 893-904.

Jiang, B. and Ai, M. (2012). Construction of uniform designs without replications. *Journal of Complexity*, **30**, 98-110.

McKay, M. D., Beckman, R. J. and Conover, W. J. (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Thechnometrics*, **21**, 239-245.

Sloane, N. J. A. and Stufken, J. (1996). A linear programming bound for orthogonal arrays with mixed levels. *J. Stat. Plan. Inference*, **56**, 295-305.

Tang, B. (1993) Orthogonal array-based Latin hypercubes. *Journal of the American statistical Association*, **88**, 1392-1397.

Tang, Y., and Xu, H. (2013) An effective construction method for multi-level uniform designs. *J. Stat. Plan. Inference*, **143**, 1583-1589.

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

Winker, P. and Fang, K. (1997) Application of threshold-accepting to the evaluation of the discrepancy of a set of points. *SIAM Numer. Analysis*, **34**, 2028-2042.

Xu, H. and Wu, C. F. J. (2001) Generalized minimum aberration for asymmetrical fractional factorial designs. *Ann. Stat.*, **29**, 1066-1077.

Zhou, Y., Fang, K. and Ning, J. (2013) Mixture discrepancy for quasi-random point sets, *J. complexity*, **29**, 283-301.

Zhou, Y. and Xu, H. (2014). Space-filling fractional factorial designs. *J. Amer. Statist. Assoc.*, **109**, 1134-1144.