Journal of the Korean Statistical Society ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

Minimum projection uniformity for computer experiment with quantitative factors

Tingxun Gou [a,*], Hong Qin [a], Kashinath Chatterjee [b]

[a] School of Mathematics and Statistics, Central China Normal University, Wuhan, China
[b] Department of Statistics, Visva-Bharati University, Santiniketan, India

ARTICLE INFO

Article history:
Received 2 June 2017
Accepted 29 January 2018
Available online xxxx

AMS 2000 subject classifications:
62K15
62K10
62K99

Keywords:
Wrap-around L_2-discrepancy
Computer experiment
Lower bound
Uniformity pattern

ABSTRACT

Computer experiments involving quantitative factors at high levels are becoming more and more important in the study of complex experiments arising in the area of science and engineering. Uniform designs are found to be widely applicable in computer experiments in the form of space-filling designs. In this paper, the projection uniformity for quantitative designs is studied under wrap-around L_2-discrepancy. A lower bound of uniformity pattern for general asymmetric designs is provided, which can be used to serve as a benchmark for both comparing different designs and also to determine the optimal design. As a byproduct, a lower bound of wrap-around L_2-discrepancy measure for the asymmetric design is also obtained. Some illustrative examples and numerical comparisons are also provided for supporting our theoretical results.

1. Introduction

Computer experiments are becoming an indispensable tool for the study of complex phenomenon or system (Santner, Williams, & Notz, 2003) and are adopted widely because of its flexibility (Wu, 2015). A lot of new challenges of modeling and designing for computer experiment arose, and space-filling property is a well known essential demand for designing computer experiment (Bates, Buck, Riccomagno, & Wynn, 1996). In case of simulation studies where a little knowledge is known about the function to be modeled, the uniform design which seeks experimental points to be uniformly scattered in the experimental domain has a wide application. Its practical success is due to its economical and flexible experimental runs to study many factors with high levels simultaneously. In this regard, mention may be made to Fang, Li, and Sudjianto (2006).

Recently, more and more attentions are paid to the space-filling property of lower dimension projection of design for computer experiment. Orthogonality or its relaxation, such as correlation, is believed to be a useful stepping stone for this purpose. He and Tang (2013) introduced strong orthogonal array which has better lower dimension uniformity than orthogonal array with same strength. Chen and Qian (2014) and Naveau, Guillou, Cooley, and Diebolt (2009) provided a class of Latin hypercube designs which can control the orthogonality or correlation for projection designs. In particular, Joseph, Gul, and Ba (2015) improved the traditional maxmin distance criterion by introducing maximum projection design and taking the projection space-filling property into account, and compared empirically it with Latin hypercube design and existing uniform design.

* Corresponding author.
E-mail addresses: weixiaoluzhu@126.com, txgou@mails.ccnu.edu.cn (T. Gou).

https://doi.org/10.1016/j.jkss.2018.01.002
1226-3192/

Please cite this article in press as: Gou, T., et al., Minimum projection uniformity for computer experiment with quantitative factors. Journal of the Korean Statistical Society (2018), https://doi.org/10.1016/j.jkss.2018.01.002.

For considering the projection uniformity of a design, Fang and Qin (2005) established a link among uniformity pattern, aberration and orthogonality for two-level designs. Zhang and Qin (2006) provided further results on minimum projection uniformity criterion so as to rank different two-level designs. For a high-level design, Hickernell and Liu (2002) pioneered the notion of projection discrepancy pattern. Qin, Wang, and Chatterjee (2012) suggested a uniformity pattern for q-level designs, which is generalized to asymmetric designs by Balakrishnan, Qin, and Chatterjee (2016). It should be noted that the projection uniformity measure considered in Balakrishnan et al. (2016) and Qin et al. (2012) is different from that of Fang and Qin (2005).

Computer experiment with high-level quantitative factors may be encountered often in practice. However, the uniformity pattern mentioned above is not appropriate for such a computer experiment due to the following reason. The uniformity pattern proposed in the existing works only considers whether the levels for each factor are equal or not, but the distance among them is ignored. For a function with quantitative variables, the distinct two inputs(levels) with different distances can lead to different responses, such as response surface methodology. The main objective of this paper is to study the projection uniformity pattern of designs for computer experiment with quantitative factors at general levels in the line of Fang and Qin (2005).

The paper is organized as follows. Section 2 describes in brief the notion of uniformity pattern with reference to computer experiment and the minimum projection uniformity criterion. In Section 3, we provide the derivation of a lower bound of uniformity pattern for general asymmetric case. Some illustrative examples are provided in Section 4. Finally, Section 5 gives the concluding remark.

2. Uniformity pattern of design for computer experiment

Uniformity is an important kind of space-filling property and is applied widely to evaluate the performance of competitive designs. Hickernell (1998) provided many modified L_2-type discrepancies based on different reproducing kernel Hilbert spaces for measuring uniformity of designs. Among these discrepancy measures, the centered L_2- and wrap-around L_2-discrepancies are found to possess good properties. But, because of the fact that the centered L_2-discrepancy is affected significantly by the center point of the experimental domain, here we consider the wrap-around L_2-discrepancy to measure the uniformity of designs.

For any positive integer s, let $C^s = [0, 1)^s$ and $\mathcal{H}_n = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n\}$ be a set of n points in C^s. For any $\boldsymbol{u}, \boldsymbol{v} \in C^s$, let $J(\boldsymbol{u}, \boldsymbol{v}) = J_1(u_1, v_1) \times J_2(u_2, v_2) \times \cdots \times J_s(u_s, v_s)$, where the symbol "$\times$" represents the Cartesian product and, for $1 \le i \le s$,

$$J_i(u_i, v_i) = \begin{cases} [u_i, v_i], & \text{if } u_i \le v_i; \\ [0, v_i) \bigcup (u_i, 1), & \text{if } u_i > v_i. \end{cases}$$

For any $\boldsymbol{u}, \boldsymbol{v} \in C^s$, let $N(\mathcal{H}_n, J(\boldsymbol{u}, \boldsymbol{v}))$ and $Vol(J(\boldsymbol{u}, \boldsymbol{v}))$ be, respectively, the cardinality of the set $\mathcal{H}_n \cap J(\boldsymbol{u}, \boldsymbol{v})$ and the volume of the region $J(\boldsymbol{u}, \boldsymbol{v})$. It is to be noted that

$$Vol(J(u_i, v_i)) = \begin{cases} v_i - u_i, & \text{if } u_i \le v_i; \\ 1 - u_i + v_i, & \text{if } u_i > v_i, \end{cases}$$

$$Vol(J(\boldsymbol{u}, \boldsymbol{v})) = \prod_{i=1}^{s} Vol(J(u_i, v_i)).$$

Let $\mathcal{G}$ be the set of all possible nonempty subsets of the set $\{1, 2, \ldots, s\}$ and $\mathcal{G}_g = \{G \in \mathcal{G} : |G| = g\}$. For any $G \in \mathcal{G}_g$, let $C^g = [0, 1)^g$ be the g-dimensional unit hypercube. According to Hickernell (1998), the wrap-around L_2-discrepancy measure for any set $\mathcal{H}_n$ of n points from C^s is defined as

$$[WD_2(\mathcal{H}_n)]^2 = \sum_{g=1}^{s} \sum_{G \in \mathcal{G}_g} \iint_{C^g \times C^g} \left[\frac{N(\mathcal{H}_n^G, J(\boldsymbol{u}^G, \boldsymbol{v}^G))}{n} - Vol(J(\boldsymbol{u}^G, \boldsymbol{v}^G)) \right]^2 d\boldsymbol{u}^G d\boldsymbol{v}^G, \tag{1}$$

where $\mathcal{H}_n^G$ is the projection of $\mathcal{H}_n$ on C^g and the remaining symbols are with reference to G.

It is to be noted that if one considers only uniformity of design points on C^g corresponding to some design $G \in \mathcal{G}_g$, then the corresponding measure of uniformity can be defined as

$$[WD_{2,G}(\mathcal{H}_n)]^2 = \iint_{C^g \times C^g} \left[\frac{N(\mathcal{H}_n^G, J(\boldsymbol{u}^G, \boldsymbol{v}^G))}{n} - Vol(J(\boldsymbol{u}^G, \boldsymbol{v}^G)) \right]^2 d\boldsymbol{u}^G d\boldsymbol{v}^G. \tag{2}$$

Let $\mathcal{U}(n, q_1 q_2 \cdots q_s)$ denote a class of designs with n runs and s factors each at the levels $q_1, q_2, \ldots, q_s$, respectively. A design $d \in U(n, q_1 q_2 \cdots q_s)$ can be viewed as an $n \times s$ array with entries of the jth, $1 \le j \le s$, column from the set $\{0, 1, \ldots, q_j - 1\}$. Moreover, for $1 \le j \le s$, the entries of the jth column appear equally often. Here each row represents a run (point) and each column represents a factor. A typical level combination of the s factors is represented as $\boldsymbol{x} = (x_1, x_2, \ldots, x_s)$, where $x_j = 0, 1, \ldots, q_j - 1,\ 1 \le j \le s$. It is to be noted that through the mapping $f : x_j \to y_j = (2x_j + 1)/2q_j$, the n level combinations can be transferred into n points in $[0, 1)^s$.

Please cite this article in press as: Gou, T., et al., Minimum projection uniformity for computer experiment with quantitative factors. Journal of the Korean Statistical Society (2018), https://doi.org/10.1016/j.jkss.2018.01.002.

According to Hickernell (1998), for any design $d \in \mathcal{U}(n, q_1q_2\cdots q_s)$, the wrap-around L_2-discrepancy measure of uniformity, as given in (1), can be expressed as

$$[WD_2(d)]^2 = -\left(\frac{4}{3}\right)^s + \frac{1}{n^2}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\prod_{j=1}^{s}\left[\frac{3}{2} - |y_{i_1j} - y_{i_2j}|(1 - |y_{i_1j} - y_{i_2j}|)\right], \tag{3}$$

where $y_{i_1j} = (2x_{i_1j} + 1)/(2q_j)$, $y_{i_2j} = (2x_{i_2j} + 1)/(2q_j)$, x_{i_1j} and x_{i_2j} are the levels of the jth factor of the design d corresponding to the i_1th and i_2th runs, respectively, $1 \le i_1 \ne i_2 \le n$, $1 \le j \le s$.

It is noticed by many researchers that even though two designs have the same wrap-around L_2-discrepancy, they may have different uniformities when projected to a sub-design. Fang and Qin (2005) initiated the work on uniformity pattern to rank two-level designs. Following Fang and Qin (2005), we can define for any design $d \in \mathcal{U}(n, q_1q_2\cdots q_s)$

$$P_g(d) = \sum_{G\in\mathcal{G}_g}[WD_{2,G}(d)]^2, \tag{4}$$

where $[WD_{2,G}(d)]^2$ is as defined in (2). It is to be noted that $P_g(d)$ provides a measure of overall uniformity of the design d when it is projected on to the sets belonging to $\mathcal{G}_g$. Moreover, it also follows from (1) and (2) that, for any design $d \in \mathcal{U}(n, q_1q_2\cdots q_s)$,

$$[WD_2(d)]^2 = \sum_{g=1}^{s} P_g(d). \tag{5}$$

For any design $d \in \mathcal{U}(n, q_1q_2\cdots q_s)$, the following theorem provides an analytical expression of $[WD_{2,G}(d)]^2$ for $G \in \mathcal{G}_g$, which will be the basis of rest of the paper.

Theorem 1. *For any design $d \in \mathcal{U}(n, q_1q_2\cdots q_s)$ and for any $G \in \mathcal{G}_g$, $WD_{2,G}(d)$ can be analytically expressed as*

$$[WD_{2,G}(d)]^2 = -\left(\frac{1}{3}\right)^g + \frac{1}{n^2}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\prod_{j\in G}\left[\frac{1}{2} - |y_{i_1j} - y_{i_2j}|(1 - |y_{i_1j} - y_{i_2j}|)\right], \tag{6}$$

where the symbols are as defined in (3).

Proof. The proof of this theorem follows from the following two facts as well as from Eqs. (3)–(5):

(I) for any positive integer s, we can write

$$\left(\frac{4}{3}\right)^s = 1 + \sum_{i=1}^{s}\binom{s}{i}\left(\frac{1}{3}\right)^i;$$

(II) from the distribution law, we get

$$\begin{aligned}
&\prod_{j=1}^{s}\left[\frac{3}{2} - |y_{i_1j} - y_{i_2j}|(1 - |y_{i_1j} - y_{i_2j}|)\right] \\
&= \prod_{j=1}^{s}\left[1 + \left\{\frac{1}{2} - |y_{i_1j} - y_{i_2j}|(1 - |y_{i_1j} - y_{i_2j}|)\right\}\right] \\
&= 1 + \sum_{g=1}^{s}\sum_{G\in\mathcal{G}_g}\prod_{j\in G}\left[\frac{1}{2} - |y_{i_1j} - y_{i_2j}|(1 - |y_{i_1j} - y_{i_2j}|)\right].
\end{aligned}$$

For a given $G \in \mathcal{G}_g$, let d^G be the design whose factors are that of d just fell into G, that is, d^G is the projection design of d indexed by G. It is trivial to check

$$[WD_{2,G}(d^G)]^2 = [WD_{2,G}(d)]^2. \tag{7}$$

According to Eqs. (4) and (5), if $g = 1$, one can check that

$$\begin{aligned}
[WD_{2,G}(d)]^2 &= [WD_{2,G}(d^G)]^2 = [WD_2(d^G)]^2 \\
&= -\frac{1}{3} + \frac{1}{n^2}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\left[\frac{1}{2} - |y_{i_1} - y_{i_2}|(1 - |y_{i_1} - y_{i_2}|)\right].
\end{aligned} \tag{8}$$

Please cite this article in press as: Gou, T., et al., Minimum projection uniformity for computer experiment with quantitative factors. Journal of the Korean Statistical Society (2018), https://doi.org/10.1016/j.jkss.2018.01.002.

If $g = 2$, we have $[WD_2(d^G)]^2 = \sum_{G^* \subseteq G, |G^*|=1} [WD_{2,G^*}(d^G)]^2 + [WD_{2,G}(d^G)]^2$, where all $[WD_{2,G^*}(d^G)]^2$'s have an analytical formula based on Eq. (8). Thus, depending on Eqs. (4), (5) and (8),

$$[WD_{2,G}(d)]^2 = [WD_{2,G}(d^G)]^2 = [WD_2(d^G)]^2 - \sum_{G^* \subseteq G, |G^*|=1} [WD_{2,G^*}(d^G)]^2$$

$$= -\left(\frac{1}{3}\right)^2 + \frac{1}{n^2} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \prod_{j \in G} \left[\frac{1}{2} - |y_{i_1 j} - y_{i_2 j}|(1 - |y_{i_1 j} - y_{i_2 j}|)\right]. \quad (9)$$

Similarly, if $g = 3$, we have

$$[WD_2(d^G)]^2 = \sum_{G^* \subseteq G, |G^*|=1} [WD_{2,G^*}(d^G)]^2 + \sum_{G^* \subseteq G, |G^*|=2} [WD_{2,G^*}(d^G)]^2 + [WD_{2,G}(d^G)]^2,$$

where all $[WD_{2,G^*}(d^G)]^2$'s have an analytical formula based on Eqs. (8) and (9). Therefore,

$$[WD_{2,G}(d)]^2 = [WD_{2,G}(d^G)]^2 = [WD_2(d^G)]^2 - \sum_{G^* \subseteq G, |G^*|=2} [WD_{2,G^*}(d^G)]^2$$

$$- \sum_{G^* \subseteq G, |G^*|=1} [WD_{2,G^*}(d^G)]^2$$

$$= -\left(\frac{1}{3}\right)^2 + \frac{1}{n^2} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \prod_{j \in G} \left[\frac{1}{2} - |y_{i_1 j} - y_{i_2 j}|(1 - |y_{i_1 j} - y_{i_2 j}|)\right]. \quad (10)$$

Apply the same method above repeatedly till $g = s$, the proof is completed.

Definition 1. For any design $d \in \mathcal{U}(n, q_1 q_2 \cdots q_s)$, the vector $(P_1(d), P_2(d), \ldots, P_s(d))$ is called the uniformity pattern of d. Given any two designs $d_1, d_2 \in \mathcal{U}(n, q_1 q_2 \cdots q_s)$, d_1 will be said to have less projection uniformity if, for some integer r, we have $P_i(d_1) = P_i(d_2)$ for $1 \le i \le r - 1$ and $P_r(d_1) < P_r(d_2)$. If no design has less projection uniformity than a design d, then d will be said to have minimum projection uniformity.

Minimum projection uniformity criterion, essentially, asks for the minimization of uniformity pattern sequentially over a given class of designs. In the next section, we provide a lower bound of uniformity pattern for general asymmetric designs.

3. A lower bound of uniformity pattern for asymmetrical designs

In this section we concentrate on the derivation of a lower bound of uniformity pattern for designs belonging to the class $\mathcal{U}(n, q_1 q_2 \cdots q_s)$. In this regard the following definition will be helpful.

Definition 2. A multiset is a generalization of the concept of a set that, unlike a set, allows multiple instances of the multiset's elements. In other words an indexed family, (a_i), where i is in some index-set, may define a multiset, sometimes written as $\{a_i\}$, in which the multiplicity of any element x is the number of indices i such that $a_i = x$.

For example, $\{a, a, b\}$ and $\{a, b\}$ are different multisets although they are the same set.
For $1 \le j \le s$, $1 \le i_1 \ne i_2 \le n$, let us define

$$\alpha^j_{i_1 i_2} = |y_{i_1 j} - y_{i_2 j}|(1 - |y_{i_1 j} - y_{i_2 j}|),$$

where $y_{i_1 j}, y_{i_2 j} \in \left\{(2l + 1)/(2q_j) : 0 \le l \le q_j - 1\right\}$. For $1 \le j \le s$, let B_j be the set of possible values of $\alpha^j_{i_1 i_2}$ and $A_j = \{\alpha^j_{i_1 i_2} : 1 \le i_1 \ne i_2 \le n\}$. Then, it is to be noted that

$$B_j = \left\{ \frac{l(q_j - l)}{q_j^2} : 0 \le l \le m_j \right\},$$

where m_j is some positive integer such that either $q_j = 2m_j$ or $q_j = 2m_j + 1$, and that for each j, A_j is a multiset.
The following lemma will be helpful in presenting the main result of this section.

Lemma 1. *For any design $d \in \mathcal{U}(n, q_1 q_2 \cdots q_s)$ and for $1 \le j \le s$, the multiset A_j can be expressed as*

$$A_j = \bigcup_{l=0}^{m_j} A_{jl},$$

where, when $q_j = 2m_j + 1$ with some positive integer m_j, A_{jl} is a multiset consisting of a single element $l(q_j - l)/q_j^2$ with multiplicity t_{jl}, here

$$t_{jl} = \begin{cases} \dfrac{n(n-q_j)}{q_j}, & \text{if } l = 0, \\ \dfrac{2n^2}{q_j}, & \text{if } l = 1, 2, \ldots, m_j; \end{cases}$$

when $q_j = 2m_j$ with some positive integer m_j, A_{jl} is a multiset consisting of a single element $l(q_j - l)/q_j^2$ with multiplicity t_{jl}, here

$$t_{jl} = \begin{cases} \dfrac{n(n-q_j)}{q_j}, & \text{if } l = 0, \\ \dfrac{2n^2}{q_j}, & \text{if } l = 1, 2, \ldots, m_j - 1, \\ \dfrac{n^2}{q_j}, & \text{if } l = m_j. \end{cases}$$

Proof. We first consider the proof of case $q_j = 2m_j + 1$. For any j, $1 \le j \le s$, here we have $q_j = 2m_j + 1$, where m_j is some positive integer, and for $1 \le i_1 \ne i_2 \le n$, $\alpha^j_{i_1 i_2}$ takes value from the set B_j. It is to be noted that the cardinality of the set B_j is $(q_j + 1)/2$. Moreover, the function $\alpha^j_{i_1 i_2}$ is symmetric about the center point of the interval $(0, 1)$. Let $V = \{0, 1, \ldots, q_j - 1\}$, considering a fixed number l, where $0 \le l \le m_j$, for any level $u \in V$, there exist an unique level $v = u + l \pmod{q_j} \in V$ corresponding to u. Therefore, for any given element $l (\ne 0)$ the value $l(q_j - l)/q_j^2$ belonging to B_j, will appear $2n^2/q_j$ times in A_j. If $l = 0$, the value 0 will appear $n(n - q_j)/q_j$ times in A_j. This proves that A_j is a multiset and, for each l, $1 \le l \le m_j$, A_{jl} is a multiset with multiplicity $2n^2/q_j$ and A_{j0} is a multiset with multiplicity $n(n - q_j)/q_j$. This completes the proof of case $q_j = 2m_j + 1$. The proof of case $q_j = 2m_j$ will follow in a similar manner.

Let d be any design in $\mathcal{U}(n, q_1 q_2 \cdots q_s)$ and G be any set belonging to $\mathcal{G}_g$, $1 \le g \le s$. Let G_O and G_E be two subsets of G which are such that $G_O \cup G_E = G$ and $G_O \cap G_E = \phi$. Moreover, suppose, for $j \in G_O$, there exists positive integers m_j such that $q_j = 2m_j + 1$ and, for $j \in G_E$, there exists positive integers m_j such that $q_j = 2m_j$. Then we have the following lemma which provides a lower bound to $[WD_{2,G}(d)]^2$.

Lemma 2. *If the given design d is projected on to the set G, then we have*

$$[WD_{2,G}(d)]^2 \ge LB_G(d), \tag{11}$$

where

$$LB_G(d) = C_g + \frac{n-1}{n} e^S, \; C_g = -\left(\frac{1}{3}\right)^g + \frac{1}{n}\left(\frac{1}{2}\right)^g,$$

and

$$n(n-1)S = \sum_{j \in G} \frac{n(n-q_j)}{q_j} \cdot \ln\left(\frac{1}{2}\right) + \sum_{j \in G} \sum_{l=1}^{m_j - 1} \left[\frac{2n^2}{q_j} \cdot \ln\left(\frac{1}{2} - \frac{l(q_j - l)}{q_j^2}\right)\right] + \sum_{j \in G_O} \left[\frac{2n^2}{q_j} \cdot \ln\left(\frac{1}{2} - \frac{m_j(m_j+1)}{q_j^2}\right)\right] + \sum_{j \in G_E} \frac{n^2}{q_j} \cdot \ln\left(\frac{1}{4}\right).$$

Proof. Based on Eq. (6), when the design d is projected on to G, we get

$$\begin{aligned} [WD_{2,G}(d)]^2 &= C_g + \frac{1}{n^2} \sum_{i_1=1}^{n} \sum_{i_2(\ne i_1)=1}^{n} \prod_{j \in G} \left[\frac{1}{2} - |y_{i_1 j} - y_{i_2 j}|(1 - |y_{i_1 j} - y_{i_2 j}|)\right] \\ &= C_g + \frac{1}{n^2} \sum_{i_1=1}^{n} \sum_{i_2(\ne i_1)=1}^{n} e^{\Sigma_{j \in G} \ln\left(\frac{1}{2} - |y_{i_1 j} - y_{i_2 j}|(1 - |y_{i_1 j} - y_{i_2 j}|)\right)}. \end{aligned}$$

Since the function e^x is strictly convex, from Jensen's inequality, we know that

$$[WD_{2,G}(d)]^2 \ge C_g + \frac{n-1}{n} e^S,$$

where

$$S=\frac{1}{n(n-1)}\sum_{j\in G}\sum_{i_1=1}^{n}\sum_{i_2(\neq i_1)=1}^{n}\ln\left(\frac{1}{2}-|x_{ik}-y_{jk}|(1-|x_{ik}-y_{jk}|)\right).$$

From Lemma 1, it follows that

$$n(n-1)S=\sum_{j\in G}\frac{n(n-q_j)}{q_j}\cdot\ln\left(\frac{1}{2}\right)+\sum_{j\in G}\sum_{l=1}^{m_j-1}\left[\frac{2n^2}{q_j}\cdot\ln\left(\frac{1}{2}-\frac{l(q_j-l)}{q_j^2}\right)\right]+\sum_{j\in G_O}\left[\frac{2n^2}{q_j}\cdot\ln\left(\frac{1}{2}-\frac{m_j(m_j+1)}{q_j^2}\right)\right]+\sum_{j\in G_E}\frac{n^2}{q_j}\cdot\ln\left(\frac{1}{4}\right).$$

This completes the proof of Lemma 2.

The following theorem provides a lower to bound to the measure of uniformity pattern of any design when projected on a subdesign involving g factors.

Theorem 2. *For a given design $d\in\mathcal{U}(n,q_1q_2\cdots q_s)$ and for any $1\le g\le s$, we have*

$$P_g(d)\ge LBP_g, \tag{12}$$

where $LBP_g=\sum_{G\in\mathcal{G}_g}LB_G(d)$ and $LB_G(d)$ is as defined in Lemma 2.

Proof. Based on Eq. (4) and Lemma 2, the proof of this theorem follows immediately.

Corollary 1. *Consider a design $d\in\mathcal{U}(n,q^s)$ and a set $G\in\mathcal{G}_g$. Then we have*

$$[WD_{2,G}(d)]^2\ge LB_G(d), \tag{13}$$

where $LB_G(d)=C_g+\left(\frac{n-1}{n}\right)e^S$, $C_g=-\left(\frac{1}{3}\right)^g+\frac{1}{n}\left(\frac{1}{2}\right)^g$, and for some positive integer m,

$$S=\begin{cases}\frac{g(n-q)}{(n-1)q}\ln\left(\frac{1}{2}\right)+\frac{2ng}{(n-1)q}\sum_{l=1}^{m}\left[\frac{1}{2}-\frac{l(q-l)}{q^2}\right], & \text{if } q=2m+1; f\\ \frac{g(n-q)}{(n-1)q}\ln\left(\frac{1}{2}\right)+\frac{ng}{(n-1)q}\ln\left(\frac{1}{4}\right)+\frac{2ng}{(n-1)q}\sum_{l=1}^{m-1}\left[\frac{1}{2}-\frac{l(q-l)}{q^2}\right], & \text{if } q=2m.\end{cases}$$

Based on Corollary 1, a theorem similar to Theorem 2 can be established related to $P_g(d)$ values.

With reference to Eq. (5), Chatterjee, Fang, and Qin (2005) obtained the following lower bound of wrap-around L_2-discrepancy measure for a design d belonging to $\mathcal{U}(n,2^{m_1}3^{m_2})$,

$$[WD_2(d)]^2\ge LB_1(d), \tag{14}$$

where

$$LB_1(d)=-\left(\frac{4}{3}\right)^{m_1+m_2}+\frac{1}{n^2}\left(\frac{5}{4}\right)^{m_1}\left(\frac{23}{18}\right)^{m_2}\sum_{i=0}^{m_1}\sum_{j=0}^{m_2}\binom{m_1}{i}\binom{m_2}{j}\left(\frac{1}{5}\right)^i\left(\frac{4}{23}\right)^j\theta_{ij},$$

$\theta_{ij}=n\theta_{ij}^*+(n-2^i3^j\theta_{ij}^*)(\theta_{ij}^*+1)$ and θ_{ij}^* is the largest integer contained in $\frac{n}{2^i3^j}$ for $i=0,1,\ldots,m_1$ and $m_2=0,1,\ldots,m_2$.

Along the line of Theorem 2, we obtain a lower bound of wrap-around L_2-discrepancy, stated in Eq. (5), for general asymmetrical design. Let $R=\{1,2,\ldots,s\}$. Let R_O and R_E be two subsets of R which are such that $R_O\cup R_E=R$ and $R_O\cap R_E=\phi$. Moreover, suppose, for $j\in R_O$, there exists positive integers m_j such that $q_j=2m_j+1$ and, for $j\in R_E$, there exists positive integers m_j such that $q_j=2m_j$. Then we have the following theorem which provides a lower bound to $[WD_2(d)]^2$.

Theorem 3. *Given any design $d\in\mathcal{U}(n,q_1q_2\cdots q_s)$, we have*

$$[WD_2(d)]^2\ge LB_2(d), \tag{15}$$

Please cite this article in press as: Gou, T., et al., Minimum projection uniformity for computer experiment with quantitative factors. Journal of the Korean Statistical Society (2018), https://doi.org/10.1016/j.jkss.2018.01.002.

Table 1
Two four-level designs d_1 and d_2 with 9 factors and 32 runs.

Runs	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1
3	0	1	1	0	1	2	2	2	2	0	1	1	0	1	2	2	2	2
4	0	1	1	1	0	3	3	3	3	0	1	1	1	0	3	3	3	3
5	0	2	2	2	2	0	0	2	3	0	2	2	2	2	0	0	2	3
6	0	2	2	3	3	1	1	3	2	0	2	2	3	3	1	1	3	2
7	0	3	3	2	3	2	2	0	1	0	3	3	2	3	2	2	0	1
8	0	3	3	3	2	3	3	1	0	0	3	3	3	2	3	3	1	0
9	1	0	1	2	2	1	3	0	2	1	0	1	2	3	0	3	3	1
10	1	0	1	3	3	0	2	1	3	1	0	1	3	2	1	2	2	0
11	1	1	0	2	3	3	1	2	0	1	1	2	0	3	3	1	0	0
12	1	1	0	3	2	2	0	3	1	1	1	2	1	2	2	0	1	1
13	1	2	3	0	0	1	3	2	1	1	2	3	0	1	0	3	1	2
14	1	2	3	1	1	0	2	3	0	1	2	3	1	0	1	2	0	3
15	1	3	2	0	1	3	1	0	3	1	3	0	2	1	3	1	2	3
16	1	3	2	1	0	2	0	1	2	1	3	0	3	0	2	0	3	2
17	2	0	2	0	2	3	2	3	1	2	0	2	2	0	3	2	1	2
18	2	0	3	1	3	3	0	2	2	2	0	3	0	2	2	1	3	3
19	2	1	2	2	0	1	2	1	0	2	1	0	3	3	0	2	1	3
20	2	1	3	3	1	1	0	0	3	2	1	3	2	1	1	0	3	0
21	2	2	0	0	3	2	3	1	3	2	2	0	1	3	2	3	2	0
22	2	2	1	1	2	2	1	0	0	2	2	1	3	1	3	0	0	1
23	2	3	0	2	1	0	3	3	2	2	3	1	1	2	0	1	0	2
24	2	3	1	3	0	0	1	2	1	2	3	2	0	0	1	3	2	1
25	3	0	2	3	1	2	3	2	0	3	0	2	3	1	2	3	0	3
26	3	0	3	2	0	2	1	3	3	3	0	3	1	3	3	0	2	2
27	3	1	2	1	3	0	3	0	1	3	1	0	2	2	1	3	0	2
28	3	1	3	0	2	0	1	1	2	3	1	3	3	0	0	1	2	1
29	3	2	0	3	0	3	2	0	2	3	2	0	0	2	3	2	3	1
30	3	2	1	2	1	3	0	1	1	3	2	1	2	0	2	1	1	0
31	3	3	0	1	2	1	2	2	3	3	3	1	0	3	1	0	1	3
32	3	3	1	0	3	1	0	3	0	3	3	2	1	1	0	2	3	0

where $LB_2(d) = C_s + \frac{n-1}{n}e^S$, $C_s = -\left(\frac{1}{3}\right)^s + \frac{1}{n}\left(\frac{1}{2}\right)^s$ *and*

$$(n-1)S = \sum_{j=1}^{s} \frac{n-q_j}{q_j} \cdot \ln\left(\frac{3}{2}\right) + \sum_{j=1}^{s}\sum_{l=1}^{m_j-1} \left(\frac{2n}{q_j}\right) \ln\left[\frac{3}{2} - \frac{l(q_j-l)}{q_j^2}\right] + \sum_{j\in R_O}\left[\frac{2n}{q_j} \cdot \ln\left(\frac{3}{2} - \frac{m_j(m_j+1)}{q_j^2}\right)\right] + \sum_{j\in R_E} \frac{n}{q_j} \cdot \ln\left(\frac{5}{4}\right).$$

4. Some illustrative examples

In this section, some illustrative examples are provided as an application of our theoretical findings.

For designs belonging to the class $\mathcal{U}(n, 2^{m_1}3^{m_2})$, we found through examples that for most of the cases the lower bound $LB_2(= LB_2(d))$ is tighter than $LB_1(= LB_1(d))$. Fig. 1 displays few numerical comparisons between these two lower bounds for some values of m_1 and m_2. From Fig. 1, for different values of m_1, we find that the new lower bound LB_2 performs better than the existing lower bound LB_1 when m_2 is big and n is small, and the advantage of LB_1 becomes more apparent when the ratio m_2/n goes bigger. It means our lower bound is tighter for supersaturated designs. When n is much bigger than m_2, LB_1 will be better. On this occasion, the experiment needs more runs so that the design is not economical and may not be acceptable. The similar phenomenon also occurs if we fix m_2 and consider the difference $LB_2 - LB_1$ by changing m_1 and n.

Example 1. Consider two four-level U-type designs d_1 and d_2 involving 9 factors and 32 runs given in Table 1. The wrap-around L_2-discrepancy measures of d_1 and d_2 are 1.055390 and 1.055353, respectively. This means the design d_2 is better than d_1 based on wrap-around L_2-discrepancy. However, following Definition 1, it follows from Table 2 that design d_1 is preferable than d_2.

Example 2. Two four-level designs, $d_1 \in \mathcal{U}(8, 4^6)$ and $d_2 \in \mathcal{U}(8, 4^{13})$ given in Table 3, are investigated to illustrate the advantage of the lower bound of uniformity pattern. The uniformity pattern and lower bounds are tabulated in Tables 4 and 5 for these two designs, respectively.

Please cite this article in press as: Gou, T., et al., Minimum projection uniformity for computer experiment with quantitative factors. Journal of the Korean Statistical Society (2018), https://doi.org/10.1016/j.jkss.2018.01.002.

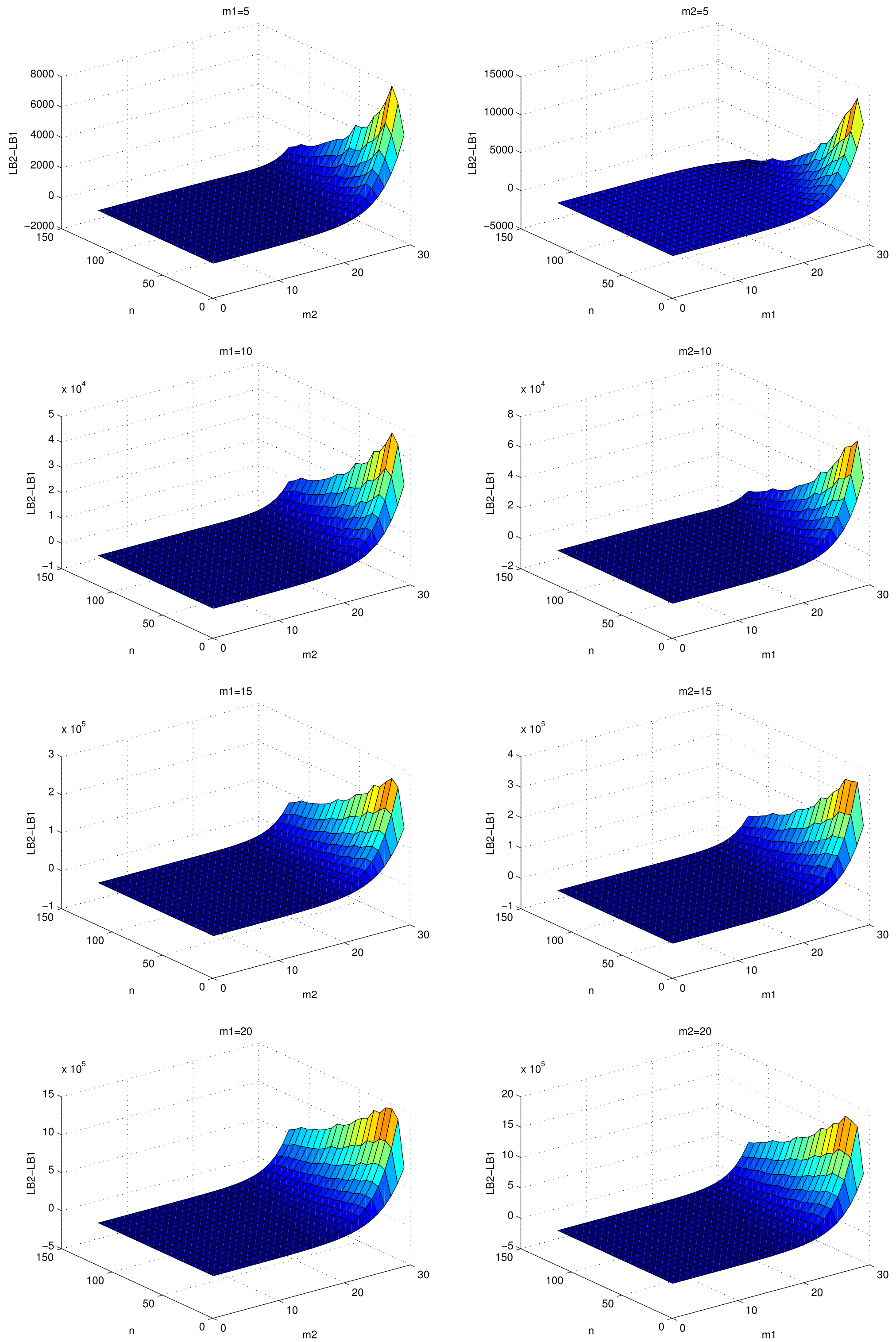

Fig. 1. Comparison between LB_1 and LB_2.

Example 3. Consider two five-level designs with 10 runs and 9 factors, denoted by d_1 and d_2. The designs d_1 and d_2 are provided in Table 6, and their uniformity patterns are depicted in Tables 7 and 8, respectively. Based on Tables 7 and 8, one

Table 2
The uniformity pattern of design d_1 and d_2.

i	1	2	3	4	5	6	7	8	9
$P_i(d_1)$	0.093750	0.253906	0.314006	0.233264	0.114029	0.037404	0.007975	0.001001	0.000056
$P_i(d_2)$	0.093750	0.253906	0.314010	0.233257	0.114012	0.037391	0.007970	0.001000	0.000056

Table 3
Two four-level designs d_1 and d_2 with 8 runs.

Runs	d_1						d_2												
	1	2	3	4	5	6	1	2	3	4	5	6	7	8	9	10	11	12	13
1	2	3	1	3	3	1	1	3	1	1	2	1	3	1	1	1	1	1	0
2	3	0	2	2	2	2	2	1	2	3	2	2	0	0	3	2	0	1	2
3	0	2	0	2	1	1	0	3	1	0	0	0	1	3	3	0	0	2	1
4	0	3	3	1	3	3	3	0	2	2	1	1	1	1	2	0	3	3	3
5	3	1	3	3	0	0	0	1	3	2	0	3	3	0	0	3	1	3	0
6	1	1	1	1	2	0	2	2	0	3	1	0	2	3	0	1	2	0	3
7	1	2	2	0	0	2	1	0	0	0	3	3	0	2	1	3	3	0	2
8	2	0	0	0	1	3	3	2	3	1	3	2	2	2	2	2	2	2	1

Table 4
The uniformity pattern and its lower bound of design d_1.

i	1	2	3	4	5	6
$P_i(d_1)$	0.062500	0.142904	0.134137	0.063468	0.015004	**0.001413**
$LB_i^P(d_1)$	0.021193	0.092536	0.111273	0.058880	0.014660	**0.001413**

Table 5
The uniformity pattern and its lower bound of design d_2.

i	1	2	3	4	5	6	7
$P_i(d_2)$	0.135417	0.788411	2.069700	3.253040	3.421790	2.545180	1.375960
$LB_i^P(d_2)$	0.045918	0.481189	1.591200	2.806610	3.144500	2.424750	1.338630
i	8	9	10	11	12	13	
$P_i(d_2)$	0.545759	0.157849	0.032501	0.004525	**0.000383**	**0.000015**	
$LB_i^P(d_2)$	0.537493	0.156567	0.032368	0.004518	**0.000383**	**0.000015**	

Table 6
Two five-level designs d_1 and d_2 with 10 runs.

Runs	d_1									d_2								
	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
1	0	0	0	1	2	1	2	3	4	3	2	3	4	2	2	4	4	4
2	0	1	3	2	4	3	1	1	3	4	1	2	0	4	4	2	3	1
3	1	2	3	0	1	1	3	0	2	1	3	0	1	1	2	3	4	0
4	1	3	1	1	4	0	0	2	1	0	0	1	3	3	3	4	1	2
5	2	0	4	0	0	3	0	4	0	1	2	4	3	0	4	0	2	1
6	2	4	1	4	3	4	2	0	3	3	1	3	1	0	0	3	0	2
7	3	1	4	3	3	0	4	3	2	2	0	0	2	2	0	0	3	3
8	3	2	2	4	2	2	1	2	0	2	3	2	4	4	1	1	0	0
9	4	3	2	2	1	4	4	4	4	4	4	1	2	1	3	1	1	4
10	4	4	0	3	0	2	3	1	1	0	4	4	0	3	1	2	2	3

Table 7
The uniformity pattern and its lower bound of design d_1.

i	1	2	3	4	5	6	7	8	9
$P_i(d_1)$	0.060000	0.242240	0.420441	0.409698	0.245879	0.0933912	0.0219954	0.00294492	**0.000171988**
$LB_i^P(d_1)$	0.001490	0.114864	0.301804	0.348417	0.226919	0.0898777	0.0216344	0.00292904	**0.000171988**

can find that d_1 possesses less $P_i(\cdot)$ than d_2 for $i = 1, 2, \ldots, 9$. Depending on Definition 1, d_1 is better than d_2. Moreover, one component of uniformity pattern of d_1 achieves the lower bound provided in Theorem 2.

5. Conclusion and further discussion

Computer experiment (or simulation) becomes more and more important to investigate complicated systems, and space-filling property of design is an essential demand for such experiment. However, different designs with same space-filling

Table 8
The uniformity pattern and its lower bound of design d_2.

i	1	2	3	4	5	6	7	8	9
$P_i(d_2)$	0.060000	0.251584	0.440560	0.428227	0.255339	0.0962831	0.0225247	0.00299860	0.000174316
$LB_i^P(d_2)$	0.001490	0.114864	0.301804	0.348417	0.226919	0.0898777	0.0216344	0.00292904	0.000171988

property may possess different space-filling properties for lower dimension projection designs. Recently, more and more researchers are focusing on exploring designs which have good space-filling property of lower dimension projection, such as He and Tang (2013). And some new criteria are introduced to provide more abundant information of designs when lower dimension projection is considered, such as the work of Joseph et al. (2015) and Zhang and Qin (2006). In this paper, minimum projection uniformity designs are introduced and studied for computer experiment with quantitative factors. And we provide a lower bound of uniformity pattern for general symmetrical and asymmetrical designs, which is more appropriate to systematically rank and select good design for computer experiment.

We also note that a challenge problem, which refers to construct minimum projection design, is worthy studying further. The major challenge in constructing minimum projection designs lies in that this criterion asks for sequentially minimizing the uniformity pattern. Thus, it is a constraint sequential optimization problem. We should first optimize the first component of uniformity pattern, and keep its optimal value as a constraint, then we can optimize the second component, and so on. The difficulty is that when we are optimizing some component, the optimal components before it may become bad. One feasible way is to optimize the weighted sum of uniformity pattern with an appropriate weight, as shown in Xu (2005), and apply heuristic threshold accepting algorithm to avoid local minimizer. Moreover, a sharp lower bound can also be used to enhance the effect for searching minimum projection uniformity designs.

Acknowledgment

This work was partially supported by the National Natural Science Foundation of China (Nos. 11271147, 11471135, and 11471136).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jkss.2018.01.002.

References

Balakrishnan, N., Qin, H., & Chatterjee, K. (2016). Generalized projection discrepancy and its application in experimental designs. *Metrika*, *79*, 19–35.
Bates, R. A., Buck, R. J., Riccomagno, E., & Wynn, H. P. (1996). Experimental design and observation for large systems. *Journal of the Royal Statistical Society. Series B.*, *58*, 77–94.
Chatterjee, K., Fang, K. T., & Qin, H. (2005). Uniformity in factorial designs with mixed levels. *Journal of Statistical Planning and Inference*, *128*, 593–607.
Chen, J. J., & Qian, P. Z. G. (2014). Latin hypercube designs with controlled correlations and multi-dimensional stratification. *Biometrika*, *101*, 319–332.
Fang, K. T., Li, R. Z., & Sudjianto, A. (2006). *Design and modeling for computer experiments*. Boca Raton: Chapman & Hall/CRC.
Fang, K. T., & Qin, H. (2005). Uniformity pattern and related criteria for two-level factorials. *Science in China, Series A (Mathematics, Physics, Astronomy & Technological Sciences)*, *48*, 1–11.
He, Y. Z., & Tang, B. X. (2013). Strong orthogonal arrays and associated Latin hypercubes for computer experiments. *Biometrika*, *100*, 254–260.
Hickernell, F. J. (1998). *Lattice rules: how well do they measure up? random and quasi-random point sets*. (pp. 109–166). New York: Springer.
Hickernell, F. J., & Liu, M. Q. (2002). Uniform designs limit aliasing. *Biometrika*, *89*, 893–904.
Joseph, V. R., Gul, E., & Ba, S. (2015). Maximum projection designs for computer experiments. *Biometrika*, *102*, 371–380.
Naveau, P., Guillou, A., Cooley, D., & Diebolt, J. (2009). Orthogonal and nearly orthogonal designs for computer experiments. *Biometrika*, *96*, 51–65.
Qin, H., Wang, Z. H., & Chatterjee, K. (2012). Uniformity pattern and related criteria for q-level factorials. *Journal of Statistical Planning and Inference*, *142*, 1170–1177.
Santner, T. J., Williams, B. J., & Notz, W. I. (2003). *The design and analysis of computer experiments*. New York: Springer-Verlag.
Wu, C. F. J. (2015). Post-Fisherian experimentation: from physical to virtual. *Journal of the American Statistical Association*, *110*, 612–620.
Xu, H. Q. (2005). Some nonregular designs from the NordstromCRobinson code and their statistical properties. *Biometrika*, *92*, 385–397.
Zhang, S. L., & Qin, H. (2006). Minimum projection uniformity criterion and its application. *Statistics & Probability Letters*, *76*, 634–640.