一种新的设计优化框架

张爱军 方开泰 李润泽 Agus Sudjianto
(香港浸会大学) (宾州州立大学) (福特汽车公司)

Abstract

以Majorization不等式理论为基础，本文提出一种新的设计优化框架，进而归纳和统一了正交设计、均匀设计和超饱和设计中的一些测度。该优化框架有张有驰，包含严苛的Majorization检验和灵活的Schur凸比较。通过选择可分离的组合核函数和指数核函数，经典的正交强度、广义最小低阶混杂、中心化L_2-偏差等准则被囊括进我们的这个优化框架。

关键词：混杂，偏差，正交设计，均匀设计，Majorization，Schur凸核函数。

1. 介绍

本文着重讨论三种平衡格子点设计（也叫U类型设计）：正交设计、均匀设计和超饱和设计。它们被广泛应用于工农业、科研开发以及近来倍受关注的电脑仿真实验，除了提供更加有效的参数估计，还在很大程度上减少实验成本。如何选择和构建设计是同行们长期研究的课题。针对正规的部分因子设计，由Box，Hunter and Hunter (1978) 提出的最大分辨力以及由Fries and Hunter (1980) 提出的最小低阶混杂准则被人们长期用来评价正交设计。到了上个世纪九十年代，实验设计者们开始意识到非正规因子设计其实不容忽略，如二水平的Plackett-Burman设计(Plackett and Burman, 1946) 除了可估主效应，还提供部分交互效应的信息；然而早期流行的最小低阶混杂准则只能衡量正规因子设计，后来Ma and Fang (2001) 和Xu and Wu (2001) 不谋而合，都成功地给出了推广，即广义最小低阶混杂。针对均匀设计，早期的星偏差准则（参见Fang and Wang, 1994)）逐渐被中心化、可卷型L_2-偏差（参见Hickernell, 1998 和Fang and Mukerjee, 2000）以及近年新提出的离散偏差（参见Hickernell and Liu, 2002 和Fang, Lin and Liu, 2003b）所取代。至于超饱和设计，我们从众多的准则里选出三个代表，$E(s^2)$ (Booth and Cox, 1962)、Ave(χ^2) (Yamada and Lin, 1999) 和Ave(f^2) (Fang, Lin and Ma, 2000)，分别评估二水平、三水平和多水平超饱和设计里两两因子之间的正交程度。

上面回顾的这众多准则，每个都有自己的出发点，运用几何、组合抑或统计推断，所谓百家争鸣的美丽。一般用户然却哀愁起来，在具体的实验现场考虑用什么测度来评选设计时，隐约可见他们困惑的表情。这不禁给我们提出了两个很自然的问题：1）这些测度之间有无相互关联？2）它们能否更进一步被被统一到单一框架？在这篇文章我们带着这样的问题上路，运用Majorization不等式理论(Mashall and Olkin, 1979)对试验设计给出一种新的优化框架。

2. 优化框架

Majorization不等式理论起源于全体居民的收入测量，后来被广泛应用并产生很多理论结果。详情参见Marshall and Olkin (1979)，我们在这边稍作介绍。对任意m长度的非负向量$\mathbf{x} \in \mathbb{R}_+^m$，它的（递增）次序统计量可被写成$x_{[1]} \le x_{[2]} \le \cdots \le x_{[m]}$. 对$\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^m$，如果

$$\sum_{r=1}^{k} x_{[r]} \ge \sum_{r=1}^{k} y_{[r]}, \quad k = 1, 2, \cdots, m-1 \quad \text{并且} \quad \sum_{r=1}^{m} x_r = \sum_{r=1}^{m} y_r, \tag{1}$$

我们则说$\mathbf{x}$被$\mathbf{y}$涵盖，表达成$\mathbf{x} \preceq \mathbf{y}$。对于$1 \le k \le m-1$，当至少存在一个$\sum_{r=1}^{k} x_{[r]} > \sum_{r=1}^{k} y_{[r]}$，我们可以写成严格的涵盖形式$\mathbf{x} \prec \mathbf{y}$。定义在$\mathbb{R}_+^m$上的一个实函数$\psi$，如果对每一对$\mathbf{x} \preceq \mathbf{y}$都满足$\psi(\mathbf{x}) \le \psi(\mathbf{y})$，那么此函数被称为是Schur凸的。Schur凸函数$\psi(\mathbf{x})$具备很多优良性质，首先它关于$\mathbf{x}$的分量对称，几个Schur凸函数相加或者相乘仍然是Schur凸的。一类特殊的Schur凸函数值得一提，叫做可分离凸函数，其一般形式为$\psi(\mathbf{x}) = \sum_{r=1}^{m} h(x_r)$，其中$h(x)$的二阶导数$h''(x) \ge 0$。可分离凸函数相当普遍，譬如$\sum_{r=1}^{m} x_r^p$ 和$\sum_{r=1}^{m} x_r \log x_r$ 都是很显著的例子。

记$\mathbf{X}(n, q^s)$为n次和s因子的试验，每个因子含有q水平。一个平衡设计的每个因子的q水平出现的次数相等，即我们常叫的U类型设计，记所有的平衡设计集合为$\mathcal{U}(n, q^s)$。经典的正交表$OA(n, s, q, t)$，当正交强度$t \ge 2$时，隶属于$\mathcal{U}(n, q^s)$的子集。考虑$t = 2$，Rao不等式要求$s(q-1) \le n-1$；当$s(q-1) = n-1$，该设计饱和；反之则达不到两两因子之间的正交性，这样的设计叫做超饱和设计。长期以来，优良的正交设计和均匀设计选自于平衡设计集$\mathcal{U}(n, q^s)$，即便以$E(s^2)$, Ave(χ^2) 和Ave(f^2)筛选出来的超饱和设计也不例外。我们在本文提出的优化框架也是以平衡设计为研究对象。更且，记$\mathcal{D}(n, q^s)$为平衡设计筛选空间，它可以是$\mathcal{U}(n, q^s)$本身，也可以是$\mathcal{U}(n, q^s)$的某个子集。

Hamming距离在因子设计的研究中举足轻重。设计里任何两个实验点$\mathbf{x}_i, \mathbf{x}_j$之间的Hamming距离$\beta(\mathbf{x}_i, \mathbf{x}_j)$ 定义为两者之间不同分量的个数。对每一个设计$\mathbf{X}(n, q^s)$，我们定义它的全息距离向量为

$$\boldsymbol{\beta}(\mathbf{X}) = (\beta_1, \beta_2, \ldots, \beta_{n(n-1)/2})'.$$

为了简便起见，记$m = n(n-1)/2$为该向量的长度。注意$\beta(\mathbf{x}_i, \mathbf{x}_k)$ $(1 \le i < k \le n)$唯一对应于全息距离向量中的$\beta_{n(i-1)+k-i(i+1)/2}$元素。任何两个同构的设计（即其中一个通过适当的行变换、列变换及水平置换变成另一个）的必要条件是它们的全息距离向量的次序统计量相等。更且，我们有下面一个重要的引理。

Lemma 1 平衡设计集$\mathcal{U}(n, q^s)$的任何设计的全息距离向量的和是唯一确定的，等于$\frac{q-1}{2q} sn^2$。

这个引理让我们可以顺利应用Majorization理论，我们记全息距离向量的均值为$\bar{\beta} = \frac{ns(q-1)}{q(n-1)}$，并且记平均全息距离向量$(\bar{\beta}, \bar{\beta}, \ldots, \bar{\beta})$为$\bar{\boldsymbol{\beta}}$. 参照决策论，新的优化框架定义如下。

Definition 1 给定任意的平衡设计空间，如果其中的两个设计$\mathbf{X}_1, \mathbf{X}_2$存在关系$\boldsymbol{\beta}(\mathbf{X}_2) \prec \boldsymbol{\beta}(\mathbf{X}_1)$，则$\mathbf{X}_1$被称为不容许的，反之则为可容许估计。如果存在$\mathbf{X}^*$,

$$\boldsymbol{\beta}(\mathbf{X}^*) \preceq \boldsymbol{\beta}(\mathbf{X}), \text{ 对任何 } \mathbf{X} \in \mathcal{D}(n, q^s),$$

Table 1: 转置的$U_{27}(3^8)$均匀设计，选自http://www.math.hkbu.edu.hk/UniformDesign/

$U(27,3^8)$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	2	3	1	3	1	1	1	2	3	2	1	3	3	2	1	3	1	3	1	2	2	2	3	2	2	1	3
2	2	3	1	1	1	3	3	3	2	1	3	2	3	1	2	1	2	2	2	2	3	3	3	2	1	1	1
3	3	3	2	2	2	2	1	2	1	1	3	2	1	1	1	3	3	2	1	2	3	2	1	3	1	3	3
4	2	2	3	1	2	3	2	1	3	2	1	3	3	3	1	2	1	2	2	1	3	2	1	3	1	3	1
5	2	3	2	2	3	3	1	1	3	1	3	1	2	3	3	3	1	1	2	3	2	2	2	1	1	1	2
6	1	1	1	2	1	3	2	1	2	3	2	2	1	2	1	3	3	3	3	3	3	2	3	1	1	2	2
7	2	3	3	3	1	2	3	1	1	1	1	1	2	3	2	2	3	2	1	3	1	2	3	3	2	2	1
8	3	1	3	1	2	2	3	1	2	3	3	3	3	1	1	3	2	1	1	3	1	2	2	2	2	1	2

则$\mathbf{X}^*$是$\mathcal{D}(n,q^s)$中的*Majorum*设计。对预先给定的*Schur*凸核函数$\psi(\cdot)$，定义函数值$\psi(\boldsymbol{\beta}(\mathbf{X}))$ 为设计$\mathbf{X}(n,q^s)$的*Schur*效率。如果一个设计的*Schur*效率在$\mathcal{D}(n,q^s)$上达到最小，我们称它为关于*Schur*凸核函数$\psi(\cdot)$的*Schur*最优设计。

在这个定义里，我们引入了四个新名词：容许估计，Majorum设计，Schur效率和Schur最优设计。我们可以将它们细分为两个层次：1）严苛的Majorization检验；2）灵活的Schur凸比较。在第一层次，按照(1)用Majorization序来比较$\mathcal{D}(n,q^s)$中设计的全息距离向量，首先排除不容许设计。如果找到Majorum设计，则立即被选用；而更多情况下我们只能找到很多的容许设计，进入第二层次作凸比较。由于对任意的Schur凸核函数，每个设计的Schur效率是一个数值，所以任何不同的容许设计都可以清楚地比较大小。Schur凸核函数的选择很自由，下面我们会给一个例子来具体说明。

Example 1 *考虑27个实验点，8个三水平因子的实验设计，选自Fang, Ma and Winder (2002)的均匀设计表$U_{27}(3^8)$，重列于Table 1。假如我们有下列先验知识：8个因子里的4个比较显著，另外4个则相对无关。实验者欲在$U(27,3^8)$的4因子子设计集里找出理想的设计方案。*

总共有$\binom{8}{4}=70$个平衡的子设计，构成本例的设计空间$\mathcal{D}(27,3^4)$。为了演示目的，我们从中选出四个子设计：$\mathbf{X}_a=\{1,3,7,8\}$, $\mathbf{X}_b=\{2,3,7,8\}$, $\mathbf{X}_c=\{1,2,4,6\}$ and $\mathbf{X}_d=\{1,4,5,6\}$。

- *第一层次：**Majorization检验***
  每个子设计的全息距离向量长度为$\binom{27}{2}=351$、和为972, 将它按升序排列，计算它的累积和向量，并且用Majorization的方法加以比较。结果是

  $$\boldsymbol{\beta}(\mathbf{X}_a)\prec\boldsymbol{\beta}(\mathbf{X}_c)\prec\boldsymbol{\beta}(\mathbf{X}_d),\quad \boldsymbol{\beta}(\mathbf{X}_b)\prec\boldsymbol{\beta}(\mathbf{X}_c)\prec\boldsymbol{\beta}(\mathbf{X}_d).$$

  因此，$\mathbf{X}_c$和$\mathbf{X}_d$都是不容许设计，尽管$\mathbf{X}_c$显著超过$\mathbf{X}_d$。比较$\mathcal{D}(27,3^4)$内的候选设计后，$\mathbf{X}_a$和$\mathbf{X}_b$ 都是可容许设计。

- *第一层次：**Schur凸比较***
  我们举三个Schur凸核函数为例，分别为L_p赋范函数（$p\geq 1$），标准方差以及负熵函数。它们的表达式，以及对应于设计$\mathbf{X}_a,\mathbf{X}_b,\mathbf{X}_c,\mathbf{X}_d$的Schur效率的数值结果列

Table 2: Schur凸比较的数值结果。注：L_p赋范函数中$p = 1.5$。最后一列的$\psi(\overline{\boldsymbol{\beta}})$值为Schur效率的下界。

Schur-convex kernel	$\psi(\boldsymbol{\beta})$	$\mathbf{X}_a$	$\mathbf{X}_b$	$\mathbf{X}_c$	$\mathbf{X}_d$	$\psi(\overline{\boldsymbol{\beta}})$
L_p-norm function	$\left(\sum_{r=1}^{m} \beta_r^p\right)^{1/p}$	140.69	140.77	140.91	140.94	137.79
Standard deviation	$\sqrt{\frac{1}{n}\sum_{r=1}^{m}\left(\beta_r - \bar{\beta}\right)^2}$	0.7994	0.7994	0.8205	0.8240	0
Negative entropy	$\sum_{r=1}^{m} \beta_r \log \beta_r$	1032.44	1035.26	1037.31	1037.83	990.05

在Table 2。我们发现，对这三种不同的Schur凸核函数，四个设计的Schur效率遵循一致的排列结果:

$$\psi(\boldsymbol{\beta}(\mathbf{X}_a)) \leq \psi(\boldsymbol{\beta}(\mathbf{X}_b)) < \psi(\boldsymbol{\beta}(\mathbf{X}_c)) < \psi(\boldsymbol{\beta}(\mathbf{X}_d)).$$

用经典的正交性和均匀性准则评估，$\mathbf{X}_a$和$\mathbf{X}_b$都是分辨力为3的正交设计，广义字长型为

$$\mathbf{X}_a : (0, 0, 10/9, 8/9); \quad \mathbf{X}_b : (0, 0, 46/27, 20/27)$$

可卷型L_2偏差为$WL_2(\mathbf{X}_a) = 0.4242$和$WL_2(\mathbf{X}_b) = 0.4245$。不仅如此，全部比较$\mathcal{D}(27, 3^4)$候选空间中的70个设计后，$\mathbf{X}_a$不仅是此空间内具有广义最小低阶混杂的正交设计，同时也是拥有最小可卷型L_2偏差的均匀设计。

一般来说，不管预先给定什么Schur凸核函数，一个Majorum设计肯定也是Schur最优设计，但反之不一定成立。此间的充要关系由下面的定理阐述。

Theorem 1 对$\mathcal{U}(n, q^s)$中的两个设计$\mathbf{X}_1, \mathbf{X}_2$, $\boldsymbol{\beta}(\mathbf{X}_1) \preceq \boldsymbol{\beta}(\mathbf{X}_2)$ 当且仅当对任意的*Schur*凸核函数$\psi(\cdot)$都成立$\psi(\boldsymbol{\beta}(\mathbf{X}_1)) \leq \psi(\boldsymbol{\beta}(\mathbf{X}_2))$。

根据Majorization不等式理论，我们还能很自然地推出Schur效率的下界。

Theorem 2 对任意预先给定的*Schur*凸核函数$\psi(\cdot)$，平衡设计$X(n, q^s)$的*Schur*效率具有下界$\psi(\overline{\boldsymbol{\beta}})$，其中$\overline{\boldsymbol{\beta}} = (\bar{\beta}, \bar{\beta}, \ldots, \bar{\beta})$，$\bar{\beta} = \frac{ns(q-1)}{q(n-1)}$.

定理中的下界能被$\mathcal{U}(n, q^s)$中的等距设计$\overline{\mathbf{X}}$达到，$\overline{\mathbf{X}}$的每两个不同实验点的Hamming距离为整数$\frac{ns(q-1)}{q(n-1)}$。饱和的正交表$OA(n, s, q, 2)$便是一类典型的等距设计任意两行之间的Hamming距离为$\lambda = n/q$。一些优良的超饱和设计也是等距设计，譬如Lin (1993)用半Hadamard方法和Fang, et. al. (2003a)用可分解的平衡不饱和区组设计方法构建的超饱和设计都是等距设计。至于上面的例子$\mathcal{U}(27, 3^4)$，它们的下界可参见Table 2的最后一列。然而，由于$\bar{\beta} = \frac{972}{351}$ 并非整数，因此达不到定理里面的绝对下界。

3. 组合型Schur凸比较

本节我们讨论以可分离的组合核函数为基础的Schur凸比较，来考察因子设计的正交性。一个重要的理论结果是建立了广义最小低阶混杂和组合型Schur凸比较的等价关联。由于广义最小低阶混杂准则是最大分辨力、最小低阶混杂和针对二水平设计的最小G_2混杂等诸准则的推广，我们的组合型Schur凸比较也相应地将这些准则联系起来。

广义最小低阶混杂（简称GMA）准则由Ma and Fang (2001) 和Xu and Wu (2001) 提出，基于因子间的效应等级原则，假定低阶效应（如主效应、二阶交互效应）比高阶交互效应重要。所以我们常选用一些低阶的线性模型（如主效应模型、二阶合成模型)。然而一些特殊的未知高阶交互效应有时候不可忽略，从而产生低阶效应估计的混淆。GMA准则意在从低维到高维逐渐降低所有可能的混淆。考虑ANOVA模型，

$$\mathbf{y} = \alpha_0 \mathbf{1}_n + \mathbf{X}_{(1)}\boldsymbol{\alpha}_1 + \ldots + \mathbf{X}_{(s)}\boldsymbol{\alpha}_s + \boldsymbol{\varepsilon},$$

其中$\mathbf{y}$代表n个观察值，$\mathbf{1}_n$是n个全是1的向量，α_0是截距，$\boldsymbol{\alpha}_j$是所有的j阶交互效应的参数，$\mathbf{X}_{(j)}$矩阵则代表对应于$\boldsymbol{\alpha}_j$的系数。$\boldsymbol{\varepsilon}$是相互独立的系统误差向量。定义

$$A_j(\mathbf{X}) = \frac{1}{n^2}\left\|\mathbf{X}_{(j)}\right\|_F^2 = \mathrm{tr}\Big(\mathbf{X}_{(j)}^H \mathbf{X}_{(j)}\Big), \quad j = 1, \ldots, s, \tag{2}$$

其中$\|\cdot\|_F$代表Frobenius矩阵范数，详情可参见Xu and Wu (2001)。行向量$(A_1(\mathbf{X}), \ldots, A_s(\mathbf{X}))$被称为广义字长型（简称GWP）。这里我们引进Fang and Zhang (2003)的一种部分排序。对$\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^s$, 如果$\mathbf{x} - \mathbf{y}$的首个非零分量是负数，我们记为$\mathbf{x} \vdash \mathbf{y}$；如果$\mathbf{x} \vdash \mathbf{y}$ or $\mathbf{x} = \mathbf{y}$，我们记为$\mathbf{x} \models \mathbf{y}$。于是，$(A_1(\mathbf{X}_1), \ldots, A_s(\mathbf{X}_1)) \vdash (A_1(\mathbf{X}_2), \ldots, A_s(\mathbf{X}_2))$ 代表$\mathbf{X}_1$的混杂长度比$\mathbf{X}_2$小。GMA设计$\mathbf{X}^*(n, q^s)$与不同构的其他设计$\mathbf{X}(n, q^s)$比较时都有$(A_1(\mathbf{X}^*), \ldots, A_s(\mathbf{X}^*)) \models (A_1(\mathbf{X}), \ldots, A_s(\mathbf{X}))$。Ma and Fang (2001) 和Xu and Wu (2001) 还参考了MacWilliams等式(MacWilliams and Sloane, 1977) 进而给出了设计理论和编码理论的联系，

$$A_j(\mathbf{X}) = \frac{1}{n}\sum_{l=0}^{s} E_l(\mathbf{X}) P_j(l; s, q), \quad j = 1, \ldots, s, \tag{3}$$

其中$P_j(x; s, q) = \sum_{w=0}^{j}(-1)^w (q-1)^{j-w}\binom{l}{w}\binom{s-l}{j-w}$ 是Krawtchouk 多项式，$(E_0(\mathbf{X}), E_1(\mathbf{X}), \ldots, E_s(\mathbf{X}))$是$\mathbf{X}(n, q^s)$距离分布，它的每个分量为

$$E_l(\mathbf{X}) = \frac{1}{n}\#\Big\{(\mathbf{x}_i, \mathbf{x}_k) : \beta(\mathbf{x}_i, \mathbf{x}_k) = l, \quad i, k = 1, \ldots, n\Big\}, \quad l = 0, 1, \ldots, s.$$

下面我们讨论如何建立GMA准则和组合型Schur凸比较的关联。预先定义s个可分离的组合核函数如下，

$$\psi_j(\boldsymbol{\beta}(\mathbf{X})) = \frac{2}{q^j}\sum_{r=1}^{m}\binom{s-\beta_r}{j} + \binom{s}{j}\left(\frac{n}{q^j} - \frac{n^2}{q^{2j}}\right), \quad j = 1, \ldots, s. \tag{4}$$

注意其中的分量函数

$$\binom{s-x}{j} = \begin{cases} 0 & \text{if } s - j < x \le s \\ \frac{1}{j!}(s-x)\cdots(s-x-j+1) & \text{if } 0 \le x \le s - j \end{cases}$$

是凸函数。定义$D_j(\mathbf{X})=\sqrt{\psi_j(\boldsymbol{\beta}(\mathbf{X}))}$，并称由此组成的向量$(D_1(\mathbf{X}),\dots,D_s(\mathbf{X}))$为设计$\mathbf{X}(n,q^s)$的偏异型(deviation pattern)。

Theorem 3 偏异型检测设计$\mathbf{X}(n,q^s)$偏离j阶（$j=1,\dots,s$）正交性的程度：

$$D_j(\mathbf{X})=\sqrt{\frac{1}{q^j}\sum\nolimits_u\sum\nolimits_d\left(N_{\tau_1,\dots,\tau_j}^{(l_1,\dots,l_j)}-\frac{n}{q^j}\right)^2}, \tag{5}$$

其中的求和符号$\sum_u$ 与$\sum_d$ 分别代表$\sum\limits_{1\le l_1<\dots<l_j\le s}$ 与 $\sum\limits_{1\le\alpha_1,\dots,\alpha_j\le q}$；$N_{\tau_1,\dots,\tau_j}^{(l_1,\dots,l_j)}$指代第$\{l_1,\dots,l_j\}$因子们取分别取$\{\tau_1,\dots,\tau_j\}$水平的实验点数目.

这个定理说明，偏异型中的每个分量都是非负的，并且$D_j(\mathbf{X})=0$ （$j\le t$）当且仅当设计$\mathbf{X}(n,q^s)$ 是强度为t的正交表。由于$D_j(\mathbf{X})=\sqrt{\psi_j(\boldsymbol{\beta}(\mathbf{X}))}$是关于$\boldsymbol{\beta}(\mathbf{X})$的Schur凸函数，我们可以根据定理1和2得到下面的推论。

Corollary 1 对$\mathcal{U}(n,q^s)$中的设计$\mathbf{X}_1,\mathbf{X}_2$, $\boldsymbol{\beta}(\mathbf{X}_1)\preceq\boldsymbol{\beta}(\mathbf{X}_2)$ 意味着偏异型中的每个分量$D_j(\mathbf{X}_1)\le D_j(\mathbf{X}_2)$对$j=1,\dots,s$同时成立。并且，每个分量的下界是

$$D_j^*=\sqrt{n(n-1)q^{-j}\binom{s-\bar{\beta}}{j}+nq^{-j}(1-nq^{-j})\binom{s}{j}}.$$

通过一些组合上的处理技巧，我们可以推导出下面的定理。

Theorem 4 对任意设计$\mathbf{X}(n,q^s)$，它的偏异型和广义字长型关联如下

$$D_j^2(\mathbf{X})=\frac{n^2}{q^{2j}}\sum_{k=1}^{j}\binom{s-k}{j-k}A_k(\mathbf{X}),\quad for\ j=1,\dots,s. \tag{6}$$

和GMA准则类似，我们也可假设因子间的效应等级原则，用部分序'$\models$'比较不同设计的偏异型。通过定理4可知，偏异型和广义字长型通过一个全是整数的下三角矩阵以线性方式关联；并且，设计$\mathbf{X}_1$的混杂程度从k阶比$\mathbf{X}_2$小，当且仅当对$j<k$有$D_j(\mathbf{X}_1)=D_j(\mathbf{X}_2)$ 并且$D_k(\mathbf{X}_1)<D_k(\mathbf{X}_2)$。这样的等价关系可以重申为：

> 在因子间的效应等级原则下，一个设计拥有广义最小低阶混杂当且仅当它在组合型Schur凸比较下是Schur最优。

同样，我们可以推导广义字长型的下界从而得到关于广义最小低阶混杂的基准（benchmark）。

Theorem 5 对$\mathcal{U}(n,q^s)$中的设计$\mathbf{X}_1,\mathbf{X}_2$，$\boldsymbol{\beta}(\mathbf{X}_1)\preceq\boldsymbol{\beta}(\mathbf{X}_2)$意味着

$$(A_1(\mathbf{X}_1),\dots,A_s(\mathbf{X}_1))\models(A_1(\mathbf{X}_2),\dots,A_s(\mathbf{X}_2)).$$

该广义字长型的下界为

$$(A_1^*,\dots,A_s^*)\models(A_1(\mathbf{X}),\dots,A_s(\mathbf{X})),$$

其中每个分量（$j=1,\dots,s$）为$A_j^*=\left(1-\frac{1}{n}\right)\sum_{w=0}^{j}(-1)^w(q-1)^{j-w}\binom{\bar{\beta}}{w}\binom{s-\bar{\beta}}{j-w}+\frac{(q-1)^j}{n}\binom{s}{j}$。这样的$(A_1^*,\dots,A_s^*)$ 可以看成是广义最小低阶混杂的一个基准（*benchmark*）。

4. 指数型Schur凸比较

均匀设计自从Fang and Wang 一九七八年提出以来，已被应用于很多领域并产生经济效益，具体可参见Fang (1980), Wang and Fang (1981) 和Fang and Wang (1994)。最近，均匀设计更被应用于电脑仿真试验，譬如用于福特汽车公司的引擎设计。下面我们用本文新提出的优化框架来分析设计的均匀性，尤其是近几年才发展成熟的一些偏差测度。

Fang and Wang (1994) 建议采用星偏差来测量试验点之间的均匀性。记$F(\mathbf{x})$为实验空间$\mathcal{X}$上的均匀分布，记$F_n(\mathbf{x})$为设计$X(n, q^s)$的经验分布函数。伪蒙特卡洛方法中的偏差是某一种范数$||F_n(\mathbf{x}) - F(\mathbf{x})||$（参见Hickernell, 1998）。星偏差、中心化L_2偏差以及可卷型L_2偏差定义在单位立方体上，而离散偏差(Hickernell and Liu, 2002; Fang, Lin and Liu, 2003b)定义在格子点集$\mathcal{X} = \{1, \ldots, q\}^s$上。此节我们着重讨论离散偏差，下面用再生核希尔伯特空间 (Hickernell, 1998)的定义方法简单回顾一下。再生核$\mathbb{K}(\mathbf{x}, \mathbf{w})$是$\mathcal{X} \times \mathcal{X}$上的实函数，关于$(\mathbf{x}, \mathbf{w})$参数对乘且非正定（参见 Wahba (1990)）。我们选择下列的核函数，对$a > 0$和$0 < \rho < 1$，

$$\mathbb{K}_d(\mathbf{x}, \mathbf{w}) = \prod_{j=1}^{s} K_d(x_j, w_j) \quad \text{with} \quad K_d(x_j, w_j) = \begin{cases} a & \text{if } x_j = w_j \\ \rho a & \text{otherwise} \end{cases} .$$

离散L_2偏差可以表达为

$$DL_2(\mathbf{X}) \quad = \quad \left\{ \int_{\mathcal{X}^2} \mathbb{K}_d(\mathbf{x}, \mathbf{w}) d[F(\mathbf{x}) - F_n(\mathbf{x})] d[F(\mathbf{w}) - F_n(\mathbf{w})] \right\}^{\frac{1}{2}} .$$

通过简单的计算，我们有

$$(DL_2(\mathbf{X}))^2 = -\mu^s + \frac{1}{n^2} \sum_{i,k=1}^{n} \prod_{j=1}^{s} K_d(x_{ij}, x_{kj}), \tag{7}$$

其中$\mu = \sum_{\alpha=1}^{q} K_d(1, \alpha)/q = a[1 + (q-1)\rho]/q$. 由此可见，离散偏差$DL_2(\mathbf{X})$可以表达成全息距离向量的函数。

Theorem 6 对U类型设计$\mathbf{X}(n, q^s)$，定义指数型$Schur$效率为下列的可分离的指数函数

$$\xi_\rho(\boldsymbol{\beta}(\mathbf{X})) = \sum_{r=1}^{m} \rho^{\beta_r}, \quad 0 < \rho < 1. \tag{8}$$

则离散偏差的平方可以表达成

$$(DL_2(\mathbf{X}))^2 = -\mu^s + \frac{a^s}{n} + \frac{2a^s}{n^2} \xi_\rho(\boldsymbol{\beta}(\mathbf{X})),$$

并且它的下界为

$$(DL_2(\mathbf{X}))^2 \geq \left(\frac{1 + (n-1)\rho^{\bar{\beta}}}{n} - \left[\frac{1 + (q-1)\rho}{q} \right]^s \right) a^s.$$

中心化L_2偏差以及可卷型L_2偏差的解析表达式如下:

$$\begin{aligned} CL_2(\mathbf{X}) &= \left\{\left(\frac{13}{12}\right)^s - \frac{2}{n}\sum_{i=1}^n\prod_{j=1}^s\left[1+\frac{1}{2}|\tilde{x}_{ij}-\frac{1}{2}|-\frac{1}{2}|\tilde{x}_{ij}-\frac{1}{2}|^2\right]\right. \\ &\quad \left.+\frac{1}{n^2}\sum_{i,k=1}^n\prod_{j=1}^s\left[1+\frac{1}{2}|\tilde{x}_{ij}-\frac{1}{2}|+\frac{1}{2}|\tilde{x}_{kj}-\frac{1}{2}|-\frac{1}{2}|\tilde{x}_{ij}-\tilde{x}_{kj}|\right]\right\}^{\frac{1}{2}} \\ WL_2(\mathbf{X}) &= \left\{-\left(\frac{4}{3}\right)^s+\frac{1}{n^2}\sum_{i,k=1}^n\prod_{j=1}^s\left[\frac{3}{2}-|\tilde{x}_{kj}-\tilde{x}_{kj}|\Big(1-|\tilde{x}_{kj}-\tilde{x}_{kj}|\Big)\right]\right\}^{\frac{1}{2}} \end{aligned}$$

其中$\mathbf{X}=(x_{kj})$，伪水平数$\tilde{x}_{ij}=\frac{2x_{ij}-1}{2q}$。对低水平设计，这两种偏差具有与离散偏差类似的性质：再生核的函数值可被Hamming距离确定。选择(8)中特殊的指数基ρ，我们可以将低水平设计的中心化和可卷型L_2偏差囊括到Majorization的优化框架。

Theorem 7 对二水平U类型设计$\mathbf{X}(n,2^s)$，

$$\begin{aligned} (CL_2(\mathbf{X}))^2 &= \frac{1}{n^2}\left(\frac{5}{4}\right)^s\Big[2\xi_{\frac{4}{5}}(\boldsymbol{\beta}(\mathbf{X}))+n\Big]+\left(\frac{13}{12}\right)^s-2\left(\frac{35}{32}\right)^s \\ &\geq \frac{n-1}{n}\left(\frac{5}{4}\right)^{s-\bar{\beta}}+\frac{1}{n}\left(\frac{5}{4}\right)^s+\left(\frac{13}{12}\right)^s-2\left(\frac{35}{32}\right)^s, \\ (WL_2(\mathbf{X}))^2 &= \frac{1}{n^2}\left(\frac{3}{2}\right)^s\Big[2\xi_{\frac{5}{6}}(\boldsymbol{\beta}(\mathbf{X}))+n\Big]-\left(\frac{4}{3}\right)^s \\ &\geq \frac{n-1}{n}\left(\frac{3}{2}\right)^s\left(\frac{5}{6}\right)^{\bar{\beta}}+\frac{1}{n}\left(\frac{3}{2}\right)^s-\left(\frac{4}{3}\right)^s. \end{aligned}$$

对三水平U类型设计$\mathbf{X}(n,3^s)$，

$$\begin{aligned} (WL_2(\mathbf{X}))^2 &= \frac{1}{n^2}\left(\frac{3}{2}\right)^s\Big[2\xi_{\frac{23}{27}}(\boldsymbol{\beta}(\mathbf{X}))+n\Big]-\left(\frac{4}{3}\right)^s \\ &\geq \frac{n-1}{n}\left(\frac{3}{2}\right)^s\left(\frac{23}{27}\right)^{\bar{\beta}}+\frac{1}{n}\left(\frac{3}{2}\right)^s-\left(\frac{4}{3}\right)^s. \end{aligned}$$

上述定理中关于CL_2, WL_2 and DL_2给出的下界和文献Fang and Mukerjee (2000)以及Fang, Lin and Liu (2003)中的结果是一致的；这些下界能被等距设计达到。

5. 超饱和设计的一些测度

最近几年超饱和设计（Booth and Cox, 1962; Lin, 1993）受到了越来越多的关注，很多测度也应运而生。本节我们主要考虑三种准则：$E(s^2)$、Ave(χ^2) 和Ave(f^2), 它们广泛地被分别用来检测二水平、三水平以及多水平的超饱和设计。我们的研究表明，单一的二次多项式型的Schur凸核函数就可以充分表达这三种准则，并使之统一。

超饱和设计$\mathbf{X}(n,q^s)$的参数满足$n\le s(q-1)$。记s列为$\mathbf{x}_{(1)},\ldots,\mathbf{x}_{(s)}$。当$q=2$，Booth and Cox (1962) 介绍了$E(s^2)$准则：

$$E(s^2)=\frac{1}{\binom{s}{2}}\sum_{1\le j<l\le s}\hat{\mathbf{x}}_{(j)}^T\hat{\mathbf{x}}_{(l)},$$

其中的$\hat{\mathbf{x}}_{(j)}$将$\mathbf{x}_{(j)}$的水平转换为$(-1,1)$，$\hat{\mathbf{x}}_{(j)}^T\hat{\mathbf{x}}_{(l)}$代列与列的内积。当$q=3$, Yamada and Lin (1999) 定义了

$$\mathrm{Ave}(\chi^2)=\frac{1}{\binom{s}{2}}\sum_{1\le j<l\le s}\chi^2(\mathbf{x}_{(j)},\mathbf{x}_{(l)}),$$

其中的

$$\chi^2(\mathbf{x}_{(j)},\mathbf{x}_{(l)})=\frac{9}{n}\sum_{\tau_1,\tau_2=1}^{3}(N_{\tau_1,\tau_2}^{(j,l)}-n/9)^2.$$

至于多水平超饱和设计，Fang, Lin and Ma (2000) 将上述的准则推广为

$$\mathrm{Ave}(f^2)=\frac{1}{\binom{s}{2}}\sum_{1\le j<l\le s}f_{NOD}^{jl},$$

其中

$$f_{NOD}^{jl}=\sum_{\tau_1,\tau_2=1}^{q}(N_{\tau_1,\tau_2}^{(j,l)}-n/q^2)^2.$$

请注意$\hat{\mathbf{x}}_{(j)}^T\hat{\mathbf{x}}_{(l)}=4\sum_{\tau_1,\tau_2=1}^{2}(N_{\tau_1,\tau_2}^{(j,l)}-n/4)^2$. 于是，我们发现这些准则都跟第三节定义的偏异型的第二个分量D_2非常类似。根据定理3，我们不难得到下列关系式：

$$E(s^2)=\frac{32D_2^2}{s(s-1)},\quad \mathrm{Ave}(\chi^2)=\frac{162D_2^2}{ns(s-1)},\quad \mathrm{Ave}(f^2)=\frac{2q^2D_2^2}{s(s-1)}.$$

因此我们可以套用偏异型的思路来研究超饱和设计的这些准则。在$j=2$时，相应的核函数可以写成二次多项式型的形式：

$$\psi_2(\boldsymbol{\beta}(\mathbf{X}))=\sum_{r=1}^{m}\beta_r^2-(2s-1)\beta_r+s(s-1). \tag{9}$$

顺其自然我们可以得到下面的定理。

Theorem 8 超饱和设计$\mathbf{X}(n,q^s)$的准则$E(s^2)$!$\mathrm{Ave}(\chi^2)$和$\mathrm{Ave}(f^2)$，采用二次多项式型的*Schur*凸核函数*(9)*，我们有

$$\begin{aligned}
&\textit{at } q=2: && E(s^2)=\frac{16}{s(s-1)}\psi_2(\boldsymbol{\beta}(\mathbf{X}))+4n-n^2\ge\frac{n^2(s-n+1)}{(s-1)(n-1)}\\
&\textit{at } q=3: && Ave(\chi^2)=\frac{36}{ns(s-1)}\psi_2(\boldsymbol{\beta}(\mathbf{X}))+9-n\ge\frac{2n(2s-n+1)}{(s-1)(n-1)}\\
&\textit{generally:} && Ave(f^2)=\frac{4}{s(s-1)}\psi_2(\boldsymbol{\beta}(\mathbf{X}))+n-\frac{n^2}{q^2}\ge\frac{n^2(q-1)[(q-1)s-n-1]}{q^2(s-1)(n-1)}.
\end{aligned}$$

这里得到的下界跟经典的理论结果也是一致的，参见Nguyen (1996), Yamada and Lin (1999)和Fang, Lin and Ma (2000)。很多用来安排试验的超饱和设计都能达到这些下界。

6. 结束语

用Majorization方法来比较全息距离向量其实是很简单的一个概念。从几何上看，这是驱使实验点间的相互距离尽可能地分布。我们并且可以很灵活地选择单值的代理准则，包括检测设计的正交性和均匀性。我们打算按照这种思路作更深层次的理论研究，譬如研究定量的多水平均匀设计的偏差性质；同时期待此类优化框架下最优设计的构造，并使之应用于实际问题。本文中一些重要定理的证明摆在附录里。

附录

Proof of Lemma 1: Each factor in a U-type design $\mathbf{X}(n, q^s)$ has q different levels all appearing n/q times, so we have $\sum_{k=1}^n \delta(X_{iw}, X_{kw}) = n - n/q$ and

$$\sum_{i=1}^n \sum_{k=1}^n \beta(\mathbf{x}_i, \mathbf{x}_k) = \sum_{w=0}^s \sum_{i=1}^n \left(\sum_{k=1}^n \delta(X_{iw}, X_{kw}) \right) = sn\left(n - \frac{n}{q}\right).$$

Note that $\beta(\mathbf{x}_i, \mathbf{x}_i) = 0$. Then for the PDV that collects $\beta(\mathbf{x}_i, \mathbf{x}_j)$ for $i < j$, its elements sum up to $\frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \beta(\mathbf{x}_i, \mathbf{x}_j)$, i.e. $\frac{q-1}{2q} sn^2$.

Proof of Theorem 3: Define the indicator variable

$$\delta_{ik}^{(l_1,\ldots,l_j)} = \begin{cases} 1 & \text{if the column group } \{l_1, \ldots, l_j\} \text{ takes the same} \\ & \text{level combination on runs } \mathbf{x}_i, \mathbf{x}_k \\ 0 & \text{otherwise} \end{cases}$$

and it can be verified that

$$\sum\nolimits_u \delta_{ik}^{(l_1,\ldots,l_j)} = \binom{s - \beta(\mathbf{x}_i, \mathbf{x}_j)}{j}; \quad \sum_{i=1}^n \sum_{k=1}^n \delta_{ik}^{(l_1,\ldots,l_j)} = \sum\nolimits_d \left[N_{\tau_1,\ldots,\tau_j}^{(l_1,\ldots,l_j)} \right]^2.$$

Note that $\sum_u 1 = \binom{s}{j}, \sum_d 1 = q^j$ and $\sum_u \sum_d N_{\tau_1,\ldots,\tau_j}^{(l_1,\ldots,l_j)} = n\binom{s}{j}$. Then we have

$$\begin{aligned} D_j^2(\mathbf{X}) &= \frac{2}{q^j} \sum_{r=1}^{n(n-1)/2} \binom{s-\beta_r}{j} + \frac{n}{q^j}\binom{s}{j} - \frac{n^2}{q^{2j}}\binom{s}{j} \\ &= \frac{1}{q^j} \sum_{i=1}^n \sum_{k=1}^n \binom{s - \beta(\mathbf{x}_i, \mathbf{x}_j)}{j} - \frac{n^2}{q^{2j}}\binom{s}{j} \\ &= \frac{1}{q^j} \sum_{i=1}^n \sum_{k=1}^n \sum\nolimits_u \delta_{ik}^{(l_1,\ldots,l_j)} - \frac{n^2}{q^{2j}}\binom{s}{j} \\ &= \frac{1}{q^j} \sum\nolimits_u \sum\nolimits_d \left[N_{\tau_1,\ldots,\tau_j}^{(l_1,\ldots,l_j)} \right]^2 - \frac{n^2}{q^{2j}}\binom{s}{j}. \end{aligned}$$

By the method of variance decomposition

$$\begin{aligned} \sum\nolimits_u \sum\nolimits_d \left(N_{\tau_1,\ldots,\tau_j}^{(l_1,\ldots,l_j)} - \frac{n}{q^j} \right)^2 &= \sum\nolimits_u \sum\nolimits_d \left(\left[N_{\tau_1,\ldots,\tau_j}^{(l_1,\ldots,l_j)} \right]^2 - \frac{2n}{q^j} N_{\tau_1,\ldots,\tau_j}^{(l_1,\ldots,l_j)} + \frac{n^2}{q^{2j}} \right) \\ &= \sum\nolimits_u \sum\nolimits_d \left[N_{\tau_1,\ldots,\tau_j}^{(l_1,\ldots,l_j)} \right]^2 - \frac{n^2}{q^j}\binom{s}{j}, \end{aligned}$$

equation (5) follows and the theorem is proved.

Proof of Theorem 4: Denote as $\mathbf{X}_v$ the j-factor sub-design corresponding to the subset $v = \{l_1, \ldots, l_j\}$ of $\{1, \ldots, s\}$. For $\mathbf{X}_v(n, q^j)$,

$$\begin{aligned} D_j^2(\mathbf{X}_v) &= \frac{2}{q^j}\sum_{r=1}^{m}\binom{j-\beta_r}{j} + \binom{j}{j}\left(\frac{n}{q^j} - \frac{n^2}{q^{2j}}\right) \\ &= \frac{n}{q^j}(E_0(\mathbf{X}_v) - 1) + \frac{n}{q^j} - \frac{n^2}{q^{2j}} \\ &= \frac{n^2}{q^{2j}}\sum_{k=1}^{j} A_k(\mathbf{X}_v), \end{aligned}$$

where $E_0(\mathbf{X}_v)$ is defined right below (3) and it can be expressed through the summation of $A_i(\mathbf{X}_v)$'s

$$n\left[1 + \sum\nolimits_{k=1}^{j} A_k(\mathbf{X}_v)\right] = q^s E_0(\mathbf{X}_v), \qquad (10)$$

which will be verified at the end of this proof.

From Theorem 3, we can derive that $D_j(\mathbf{X})$ can be expressed through the root mean square of $D_j(\mathbf{X}_v)$ for all the j-factor sub-designs, i.e.,

$$D_j(\mathbf{X}) = \sqrt{\sum_{|v|=j} D_j^2(\mathbf{X}_v)}.$$

With reference to the definition of GWP (2) through contrasts and further to Xu and Wu (2001), let us write $A_k(\mathbf{X}_v) = \frac{1}{n^2}\sum\limits_{wt(\mathbf{u})=k}\left|\chi_{\mathbf{u}}(\mathbf{X}_v)\right|^2$ where $\{\chi_{\mathbf{u}}, \mathbf{u} \in \mathrm{GF}^j(q)\}$ are given orthonormal contrasts and $wt(\mathbf{u})$ is the number of non-zero elements of $\mathbf{u}$. Then, we can show the deviation pattern more explicitly through contrasts.

$$D_j^2(\mathbf{X}) = \frac{n^2}{q^{2j}}\sum_{|v|=j}\sum_{k=1}^{j} A_k(\mathbf{X}_v) = \frac{1}{q^{2j}}\sum_{|v|=j}\sum_{k=1}^{j}\sum_{wt(\mathbf{u})=k}\left|\chi_{\mathbf{u}}(\mathbf{X}_v)\right|^2$$

Please note an important fact that each contrast $\chi_{\mathbf{u}}(\mathbf{X}_v)$ of the sub-design $\mathbf{X}_v$ is also a contrast of $\mathbf{X}(n, q^s)$. It corresponds to such $\chi_{\mathbf{w}}(\mathbf{X})$ ($\mathbf{w} \in \mathrm{GF}^s(q)$) that $\mathbf{w}$ coincides with $\mathbf{u}$ at the positions of $v = \{1, \ldots, s\}$ and has null elements elsewhere. For each subset v of $\{1 \ldots, s\}$, let us denote by Ω_v the set of such specific $\mathbf{w}$'s. Then, we have

$$\begin{aligned} D_j^2(\mathbf{X}) &= \frac{1}{q^{2j}}\sum_{k=1}^{j}\sum_{|v|=j}\sum_{\mathbf{w}\in\Omega_v}\left|\chi_{\mathbf{w}}(\mathbf{X})\right|^2 \\ &= \frac{1}{q^{2j}}\sum_{k=1}^{j}\binom{s-j}{j-k}\sum_{wt(\mathbf{w})=k}\left|\chi_{\mathbf{w}}(\mathbf{X})\right|^2 \end{aligned}$$

By writing $\frac{1}{n^2}\sum\limits_{wt(\mathbf{w})=k}\left|\chi_{\mathbf{w}}(\mathbf{X})\right|^2$ back to $A_k(\mathbf{X})$, equation (6) is proved.

Let us now verify (10) for $\mathbf{X}(n, q^s)$ through Krawtchouk polynomials. For an integer l ($0 \leq l \leq s$) and any real number y, Krawtchouk polynomials $P_j(l; s, q)$ have the following property

$$\sum_{j=0}^{s} P_j(l; s, q) y^j = [1 + (q-1)y]^{s-l}(1-y)^l.$$

By setting $y = 1$, we have

$$\sum_{j=0}^{s} P_j(0; s, q) = q^s; \quad \sum_{j=0}^{s} P_j(l; s, q) = 0, \text{ for } l = 1, \ldots, s.$$

By $P_0(l; s, q) = 1$ and MacWilliams expression (3) for A_1 to A_s:

$$\begin{aligned} n\left[1 + \sum_{j=1}^{s} A_j(\mathbf{X})\right] &= n + \sum_{j=1}^{s}\sum_{l=0}^{s} E_l(\mathbf{X}) P_j(l; s, q) \\ &= n + \sum_{l=0}^{s} E_l(\mathbf{X}) \left(\sum_{j=0}^{s} P_j(l; s, q) - 1\right) \\ &= n + q^s E_0(\mathbf{X}) - \sum_{l=0}^{s} E_l(\mathbf{X}) \\ &= q^s E_0(\mathbf{X}). \end{aligned}$$

Proof of Theorem 6: As the Schur-convex kernel chosen separable exponential (8), the squared discrete L_2-discrepancy (7) can be expressed by

$$\begin{aligned} [DL_2(\mathbf{X})]^2 &= -\mu^s + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{k=1}^{n} (\rho a)^{\beta(\mathbf{x}_i,\mathbf{x}_k)} a^{s-\beta(\mathbf{x}_i,\mathbf{x}_k)} \\ &= -\mu^s + \frac{a^s}{n^2}\sum_{i=1}^{n}\sum_{k=1}^{n} \rho^{\beta(\mathbf{x}_i,\mathbf{x}_k)} \\ &= -\mu^s + \frac{a^s}{n} + \frac{2a^s}{n^2}\xi_\rho(\boldsymbol{\beta}(\mathbf{X})). \end{aligned}$$

Its lower bound can be obtained from Theorem 2.

Proof of Theorem 7: For two-level designs, $\tilde{X}_{ij} \in \{0.25, 0.75\}$, and

$$\begin{aligned} (CL_2(\mathbf{X}))^2 &= \left(\frac{13}{12}\right)^s - \frac{2}{n}\sum_{i=1}^{n}\prod_{j=1}^{s}\left[1 + \frac{1}{8} - \frac{1}{32}\right] + \frac{1}{n^2}\sum_{i,k=1}^{n}\prod_{j=1}^{s}\left[\frac{5}{4} - \frac{1}{2}|\tilde{X}_{ij} - \tilde{X}_{kj}|\right] \\ &= \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{k=1}^{n}\left(\frac{5}{4}\right)^{s-\beta(\mathbf{x}_i,\mathbf{x}_k)} \\ &= \frac{1}{n^2}\left(\frac{5}{4}\right)^s \sum_{i=1}^{n}\sum_{k=1}^{n}\left(\frac{4}{5}\right)^{\beta(\mathbf{x}_i,\mathbf{x}_k)} + \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s \end{aligned}$$

$$
\begin{aligned}
&= \frac{2}{n^2}\left(\frac{5}{4}\right)^s \xi_{\frac{4}{5}}(\boldsymbol{\beta}(\mathbf{X})) + \frac{1}{n}\left(\frac{5}{4}\right)^s + \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s; \\
(WL_2(\mathbf{X}))^2 &= -\left(\frac{4}{3}\right)^s + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{k=1}^{n}\left(\frac{3}{2}\right)^{s-\beta(\mathbf{x}_i,\mathbf{x}_k)}\left(\frac{5}{4}\right)^{\beta(\mathbf{x}_i,\mathbf{x}_k)} \\
&= \frac{1}{n^2}\left(\frac{3}{2}\right)^s \sum_{i=1}^{n}\sum_{k=1}^{n}\left(\frac{5}{6}\right)^{\beta(\mathbf{x}_i,\mathbf{x}_k)} - \left(\frac{4}{3}\right)^s \\
&= \frac{2}{n^2}\left(\frac{3}{2}\right)^s \xi_{\frac{5}{6}}(\boldsymbol{\beta}(\mathbf{X})) + \frac{1}{n}\left(\frac{3}{2}\right)^s - \left(\frac{4}{3}\right)^s.
\end{aligned}
$$

For three-level designs, $\tilde{X}_{ij} \in \{1/6, 3/6, 5/6\}$ and $|\tilde{X}_{kj} - \tilde{X}_{kj}|\Big(1 - |\tilde{X}_{kj} - \tilde{X}_{kj}|\Big)$ takes only two possible values of 0 and $2/9$. The exponential base can be calculated through

$$\rho = \frac{\frac{3}{2} - \frac{2}{9}}{\frac{3}{2}} = \frac{23}{27}.$$

Thus, proof for $q = 3$ is similar to $WL_2(\mathbf{X})$ at $q = 2$. The lower bounds for either CL_2- and WL_2-discrepancies can be readily derived by Theorem 2.

References

[1] Arnold, B.C. (1987). *Majorization and the Lorenz Order: A Brief Introduction*. Springer-Verlag, Berlin.

[2] Box, G.E.P., Hunter, W.G. and Hunter, J.S. (1978). *Statistics for Experimenters*. Wiley, New York.

[3] Booth, K.H.V. and Cox, D.R. (1962). Some systematic supersaturated designs. *Technometrics*, **4**: 489–495.

[4] Chen, J., Sun, D.X. and Wu, C.F.J. (1993). A catalogue of two-level and three-level fractional factorial designs with small runs. *Int. Stat. Review*, **61**: 131–145.

[5] Cheng, C.S. and Mukerjee, R. (1998). Regular fractional factorial designs with minimum aberration and maximum estimation capacity. *Ann. Statist.*, **26**: 2289–2300.

[6] Cheng, C.S., Steinberg, D.M. and Sun, D.X. (1999). Minimum aberration and model robustness for two-level fractional factorial designs. *J. R. Statist. Soc. B*, **61**: 85–93.

[7] Fang, K.T. (1980). Experimental design by uniform distribution. *Acta Mathematicae Applicatae Sinica*, **3**: 363–372.

[8] Fang, K.T., Ge, G.N., Liu, M.Q. and Qin, H. (2003a). Construction on minimum generalized aberration designs. *Metrika*, **57**: 37–50.

[9] Fang, K.T., Lin, D.K.J. and Liu, M.Q. (2003b). Optimal mixed-level supersaturated design. *Metrika*, to appear.

[10] Fang, K.T., Lin, D.K.J. and Ma, C.X. (2000). On the construction of multi-level supersaturated designs. *J. Statist. Plan. Inference*, **86**: 239–252.

[11] Fang, K.T., Ma, C.X. and Winker, P. (2002). Centered L_2-discrepancy of random sampling and Latin hypercube design, and construction of uniform designs, *Math. Computation*, **71**, 275–296.

[12] Fang, K.T. and Mukerjee, R. (2000). A connection between uniformity and aberration in regular fractions of two-level designs. *Biometrika*, **87**: 193–198.

[13] Fang, K.T. and Wang, Y. (1994). *Number-Theoretic Methods in Statistics*. Chapman and Hall, London.

[14] Fang, K.T. and Zhang, A. (2003). Minimum aberration majorization in non-isomorphic saturated designs. *J. Statist. Plan. Inference*, to appear.

[15] Fries, A. and Hunter, W.G. (1980). Minimum aberration 2^{k-p} designs. *Technometrics*, **22**: 601–608.

[16] Hickernell, F.J. (1998). A generalized discrepancy and quadrature error bound. *Math. Comp.*, **67**: 299–322.

[17] Hickernell, F.J. and Liu, M.Q. (2002). Uniform designs limit aliasing. *Biometrika*, **89**: 893–904.

[18] Lin, D.K.J. (1993). A new class of supersaturated designs. *Technometrics*, **35**: 28–31.

[19] Ma, C.X. and Fang, K.T. (2001). A note on generalized aberration in factorial designs. *Metrika*, **53**: 85–93.

[20] MacWilliams, F.J. and Sloane, N.J.A. (1977). *The Theory of Error-Correcting Codes*. North Holland, Amsterdam.

[21] Marshall, A.W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

[22] Nguyen, N.K. (1996). An algorithmic approach to constructing supersaturated designs. *Technometrics*, **38**: 69–73.

[23] Wahba, G. (1990). *Spline Models for Observational Data*. SIAM, Philadelphia.

[24] Wang, Y. and Fang, K.T. (1981). A note on uniform distribution and experimental design. *Kexue Tongbao (Chinese Science Bulletin)*, **26**: 485–489.

[25] Xu, H. (2003). Minimum moment aberration for nonregular designs and supersaturated designs. *Statist. Sinica*, **13**: 691–708.

[26] Xu, H. and Wu, C.F.J. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs. *Ann. Statist.*, **29**: 1066–77.

[27] Yamada, S. and Lin, D.K.J. (1999). Three-level supersaturated designs. *Statist. Prob. Letters*, **45**: 31–39.