Calcutta Statistical Association Bulletin
Vol. 65 (Special 8th Triennial Symposium Proceedings Volume) 2013,
Nos. 257-260

A UNIFIED APPROACH TO FACTORIAL DESIGNS WITH RANDOMIZATION RESTRICTIONS

PRITAM RANJAN
and
NEIL SPENCER
Department of Mathematics and Statistics,
Acadia University, Wolfville, Canada

ABSTRACT : Factorial designs are commonly used to assess the impact of factors and factor combinations in industrial and agricultural experiments. Though preferred, complete randomization of trials is often infeasible, and randomization restrictions are imposed. In this paper, we discuss a finite projective geometric (PG) approach to unify the existence, construction and analysis of multistage factorial designs with randomization restrictions using *randomization defining contrast subspaces* (or flats of a PG). Our main focus will be on the construction of such designs, and developing a word length pattern scheme that can be used for generalizing the traditional design ranking criteria for factorial designs. We also present a novel isomorphism check algorithm for these designs.

Keywords and phrases : Collineation; design ranking criteria; isomorphism; spreads; stars.

1. Introduction

Full factorial and fractional factorial designs with randomization restrictions (e.g., block designs, split-plot designs, strip-plot designs, split-lot designs, and combinations thereof) are often used for designing industrial experiments when complete randomization of the trials is impractical (Miller 1997; Mee and Bates 1998; Vivacqua and Bisgaard 2004; Bingham et al. 2008). Although these designs maintain the same factorial treatment structure, their randomization structures are different. The existence, construction and analysis of a few of these designs are well known, whereas, for complicated randomization structures (e.g., in multistage split-lot experiments), even the existence of a desirable design can be difficult to determine; for instance, Bingham et al. (2008) were forced to search over the set of all possible designs to find those satisfying the required restrictions.

Ranjan et al. (2009) used the equivalence of a 2^n factorial experiment and the finite projective geometry $\mathcal{P}_n = PG(n-1, 2)$ to develop a unified theory for the existence and construction of such designs. This formulation is characterized by the features of randomization defining contrast subspaces (RDCSSs) for each stage of randomization, which correspond to the $(t-1)$-flats of $\mathcal{P}_n$. That is, the randomization structure of a multistage factorial experiment is completely characterized by a set of RDCSSs (flats). Ranjan et al. used projective geometric structures like spreads, covers and stars to establish existence results and construction algorithms for such multistage factorial experiments.

From an analysis viewpoint, a common strategy for assessing the significance of factorial effects in unreplicated factorial designs is to use half-normal plots with the restriction that the effects appearing on the same plot must have the same error variance. To ensure the significance assessment of all effects in an experiment, Ranjan et al. (2009) recommended that the RDCSSs be disjoint and large enough to construct useful half-normal plots. Schoen (1999) suggests maintaining a minimum of six - seven points per half-normal plot for meaningful analysis. As a result, the existence and construction of the desired design is equivalent to the availability of a set of disjoint large flats of $\mathcal{P}_n$. Ranjan et al. (2009) used spreads of $\mathcal{P}_n$, set of flats that partition $\mathcal{P}_n$, for such designs. In many practical situations (e.g., the plutonium alloy experiment of Bingham et al. 2008), overlap among RDCSSs cannot be avoided, thereby making the significance assessment of all factorial effects difficult. For such cases, Ranjan et al. (2010) proposed a new geometric structure called a *star* - a set of distinct flats of $\mathcal{P}_n$ that share a common overlap.

Though not straightforward, as illustrated in this paper, more than one candidate designs can be constructed in some cases. These designs can then be ranked using optimality criteria like maximum resolution, minimum abberation, number of clear effects and so on. In the spirit of regular fractional factorial designs, we develop a word length pattern scheme that can be used to generalize the traditional design ranking criteria. In order to sort through a class of candidate designs, it is important to check whether two designs are isomorphic. Spencer (2013) developed several isomorphism check algorithms for RDCSS-based designs.

The remainder of this paper is organized as follows: Section 2 establishes the finite projective geometry representation of such multistage factorial designs. We briefly review some important existence results of Ranjan et al. (2009) and Ranjan et al. (2010) in Section 3. The main focus of this paper are Sections 4 and 5. In Section 4, we discuss the construction of such designs using spreads and stars. The word length pattern scheme for RDCSS-based designs is developed in Section 5. In this section, we also briefly review the novel isomorphism check algorithm by Spencer (2013) for sorting through these multistage factorial designs. We also point out several interesting open problems in this area. Finally, Section 6 presents a few concluding remarks.

2. RDCSSs and projective Geometries

In recent years, considerable attention has been devoted to factorial and fractional factorial designs with specific randomization restrictions, such as blocked designs, split-plot designs, blocked split-plot designs, strip-plot designs and split-lot designs. The treatment structure of these designs are identical, however, they differ in the randomization structures. In this section, we compare the randomization structure of a few such designs.

Ranjan et al. (2009) used a finite projective geometry representation to unify the multistage factorial designs with randomization restrictions. Let $\mathcal{P}_n = PG(n-1, 2)$ be the $(n-1)$-dimensional finite projective space defined by the set of all non-null n-dimensional pencils over $GF(2)$. Equivalently, let $F_1, ..., F_n$ be the main effects (or n linearly independent two-level factors), then $\mathcal{P}_n = \langle F_1, ..., F_n \rangle$ denotes the set of all $2^n - 1$ factorial effects (excluding the grand mean), where $\langle \cdots \rangle$ denotes the span. Any n-dimensional pencil in $\mathcal{P}_n$ with r nonzero elements uniquely corresponds to an r-factor interaction. We use the generic term *randomization restriction factors* to refer to the blocking factors, whole-plot factors, sub-plot factors, sub-sub-plot factors, nesting factors, and so on. For a multistage design, let $\delta_l^{(j)}$ be the l-th randomization restriction factor for stage j. A set S of all non-null pencils formed as linear combinations of t independent randomization restriction factors in $\mathcal{P}_n$ constitutes a $(t-1)$-dimensional subspace (or, $(t-1)$-flat) of $\mathcal{P}_n$ (here $|S| = 2^t - 1$). We call such a $(t-1)$-flat a t-dimensional randomization defining contrast subspace (RDCSS).

A common strategy for assessing the significance of factorial effects in unreplicated factorial designs is to use half-normal plots with the restriction that the effects appearing on the same plot must have the same error variance. In the current setup, the RDCSSs (or equivalently, $(t-1)$-flats of $\mathcal{P}_n$) indicate which effects have the same variance and thus can appear on the same half-normal plot. That is, the overlapping structure of RDCSSs characterizes how to analyze such experiments. To ensure the significance assessment of all effects in an experiment, it is recommended that the RDCSSs be disjoint and large enough to construct useful half-normal plots (i.e., at least six-seven effects per plot).

Example 2.1 In a 2^4 factorial experiment if only batches of 4 homogeneous experimental units are available, one can run a blocked factorial design with 4 blocks of 4 units in each block. That is, two independent blocking factors ($\delta_1^{(1)}$ and $\delta_2^{(1)}$) have to be selected from $\mathcal{P}_4$ to divide 2^4 treatment combinations into 2^2 blocks. In this case, the RDCSS is, $S_1 = \langle \delta_1^{(1)}, \delta_2^{(1)} \rangle$, commonly known as block defining contrast subgroup/subspace. Typically, higher order interactions are considered for defining blocking factors, and their assessment is sacrificed.

Example 2.2 Consider a 2^4 split-plot design, where A, B are whole-plot factors and C, D are sub-plot factors (e.g., cheese making experiment in Bingham, Schoen and Sitter 2004). Similar to Example 2.1, there is only one stage of randomization restriction and is characterized by $S_1 = \langle A, B\rangle$ (i.e., $\delta_1^{(1)} = A$ and $\delta_2^{(1)} = B$). Though the treatment structure, randomization structure and analysis may appear to be the same as in Example 2.1, it is worth noting that the randomization factors $\{\delta_1^{(1)}, \delta_2^{(1)}\}$ in Example 2.1 can be chosen by the statistician to benefit the analysis, whereas, in this case, the randomization factors are selected beforehand by the experimenter and assumed to be significant.

Example 2.3 Consider a 2^6 split-split-plot design with two whole-plot factors (A, B), two sub-plot factors (C, D), and two sub-sub-plot factors (E, F). The randomization restrictions are defined by $S_1 = \langle A, B\rangle$ and $S_2 = \langle A, B, C, D\rangle$. In this case, the RDCSSs satisfy the following relationship: $S_1 \subset S_2 \subset \mathcal{P}_6$. The effects in S_1 are assumed to be important (too few points for a half-normal plot), and the significance of the remaining factorial effects are assessed via two half-normal plots for $\mathcal{P}_6 \backslash S_2$ and $S_2 \backslash S_1$.

Example 2.4 In a 2^5 strip-plot design suppose the row configurations are defined by A, B, and the column configurations by C, D, E (e.g., washer-dryer example in Miller 1997). Then, two RDCSSs, $S_1 = \langle A, B\rangle$ and $S_2 = \langle C, D, E\rangle$, are required to characterize this design. Note that S_1 and S_2 are disjoint (i.e., $S_1 \cap S_2 = \phi$). From an analysis viewpoint, the effects in S_1 cannot be tested and assumed to be important, whereas, the remaining effects can be tested using two half-normal plots of $\mathcal{P}_5 \backslash \{S_1 \cup S_2\}$ and S_2.

Example 2.5 Consider a three-stage 2^5 split-lot design with restrictions $S_1 \supset \{A, B\}$, $S_2 \supset \{C\}$ and $S_3 \supset \{D, E\}$, for instance, in the plutonium alloy experiment (Bingham et al. 2008). They used $S_1 = \langle A, B, ABCDE\rangle$, $S_2 = \langle C, AD, ABCDE\rangle$ and $S_3 = \langle D, E, ABCDE\rangle$, and tested the significance of all effects except $ABCDE$ using four half-normal plots. Note that S_i's are not disjoint in this case, and $S_i \cap S_j = \{ABCDE\}$ for all $i \neq j$.

In some cases, the construction of such factorial designs is straightforward, whereas in other cases even the existence is unknown. For instance, in Example 2.5, it is ideal to construct three disjoint RDCSSs of size seven each, which turns out to be impossible. Bingham et al. (2008) used a computer search to conclude this, and find the best possible design presented in Example 2.5. Ranjan et al. (2009) and Ranjan et al. (2010) established the existence and construction of such factorial designs using geometric structures like spreads, covers, and stars of $\mathcal{P}_n$ (a few important results are presented in the next section).

3. Theory of RDCSS-based Designs

Ranjan et al. (2009) established the existence of a set of disjoint RDCSSs in a 2^n experiment via the existence of a *spread* of $\mathcal{P}_n = PG(n-1,2)$. A *spread* of $\mathcal{P}_n$ is a set of disjoint flats of $\mathcal{P}_n$ that cover $\mathcal{P}_n$.

Definition 3.1 A $(t_1-1,...,t_\mu-1)$-spread ψ of $\mathcal{P}_n = PG(n-1,2)$ is a collection of pairwise disjoint subspaces $S_i, i=1,...,\mu$, such that $|S_i| = 2^{t_i}-1$ and $\mathcal{P}_n = \cup_{i=1}^{\mu} S_i$, where $1 \le t_1,\ldots,t_\mu \le n$.

Without loss of generality, we assume that $t_1 \le t_2 \le \cdots \le t_\mu$, and non-trivial spreads correspond to $1 < t_i < n$. If all t_i's are same, the spread ψ is said to be *balanced*, otherwise, we call ψ a *mixed spread* of $\mathcal{P}_n$. Though the definition of a mixed spread is simple, its existence for the general case is not guaranteed.

Lemma 3.1 For the existence of a $(t_1-1,...,t_\mu-1)$-spread ψ of $PG(n-1,2)$, the following conditions are necessary:

(i) $2^n - 1 = \sum_{i=1}^{\mu}(2^{t_i}-1)$,

(ii) $t_i + t_j \le n$ for every $i \ne j$ $(i,j = 1,...,\mu)$.

A necessary and sufficient condition for the existence of a $(t-1)$-spread is that t divides n (André 1954). If exists, the size of a balanced $(t-1)$-spread ψ of $\mathcal{P}_n$ is $|\psi| = (2^n-1)/(2^t-1)$. From factorial design perspective, the desired number of RDCSSs is typically less than μ (the total number of flats in the spread). In that case, if there does not exist a "full" spread with desired flat sizes one can perhaps use a "partial" spread of $\mathcal{P}_n$, i.e., a set of disjoint flats of $\mathcal{P}_n$, that is not a cover of $\mathcal{P}_n$. See Ranjan et al. (2009) for more results on the existence of balanced/mixed full/partial spreads of $\mathcal{P}_n$.

In many practical situations like Example 2.5 (same as the plutonium alloy experiment of Bingham et al. 2008), the overlap among the RDCSSs cannot be avoided. For such cases, Ranjan et al. (2010) proposed designs based on a new geometric structure called a *star* - a set of distinct rays (flats of $\mathcal{P}_n$) that share a common overlap (the nucleus). Let $\Omega = St(n;t_1,...,t_\mu;t_0)$ denote a mixed or unbalanced star that is also a cover (referred to as a *covering star*) of $\mathcal{P}_n$ with μ rays ((t_i-1)-flats) and a (t_0-1)-dimensional nucleus, such that $t_1 \le \cdots \le t_\mu$. A covering star can be thought of as a generalization of a spread of $\mathcal{P}_n$.

Lemma 3.2 The existence of a covering star $\Omega = St(n; t_1, ..., t_\mu; t_0)$ of $\mathcal{P}_n = PG(n-1, 2)$ is equivalent to the existence of a $(h_1 - 1, ..., h_\mu - 1)$-spread ψ of $\mathcal{P}_u = PG(u-1, 2)$, where $u = n - t_0$, and $h_i = t_i - t_0$ for each i.

Note that a star with empty nucleus (i.e., $t_0 = 0$) simplifies to a spread. Lemma 3.2 can be combined with Lemma 3.1 to establish necessary conditions for the existence of a covering star $\Omega = St(n; t_1, ..., t_\mu; t_0)$ of $\mathcal{P}_n$. Let $St(n; \mu; t; t_0)$ be a balanced covering star with μ rays consisting of $(t-1)$-flats and a $(t_0 - 1)$-dimensional nucleus. Then, the existence of a balanced covering star $\Omega = St(n; \mu; t; t_0)$ of $\mathcal{P}_n$ is equivalent to the existence of a $(t - t_0 - 1)$-spread ψ of $\mathcal{P}_{n-t_0}$, and depends on the condition $(t - t_0)$ divides $(n - t_0)$. If such a covering star exists, the total number of rays is $|\Omega| = |\psi| = \mu = (2^{n-t_0} - 1)/(2^{t-t_0} - 1)$. In such a case if a star is denoted by $\Omega = \{S_1^*, ..., S_\mu^*\}$ and the corresponding spreads by $\psi = \{S_1, ..., S_\mu\}$, then $S_i^* = \langle S_i, \pi \rangle$, where π is the nucleus of Ω. For convenience, we denote this as $\Omega = \psi \times \pi$. Unlike balanced spreads, the existence of balanced covering stars rely on less stringent conditions, and the existence can be guaranteed for every t and n $(t < n)$.

Lemma 3.3 For every t $(2 \leq t < n)$ and $t_0 = t - 1$, there exists a balanced covering star $St(n; \mu; t; t_0)$ of $\mathcal{P}_n = PG(n-1, 2)$ with $\mu = 2^{n-t+1} - 1$.

See Ranjan et al. (2010) for more results on the existence of balanced and mixed covering stars of $PG(n-1, q)$ for prime or prime power q. Next we discuss how to construct such factorial designs using spreads and stars of $PG(n-1, 2)$.

4. Construction of RDCSS-based Designs

Our primary objective is to construct a set of RDCSSs (or flats) with predetermined restrictions on the size and their elements, for instance, $S_1 \supset \{A, B\}$, $S_2 \supset \{C, D, E, F\}$ and so on. In some cases designs with randomization restrictions are not too difficult to construct (e.g., block designs or split-plot designs), whereas for designs like split-lot designs (in Example 2.5), existence and construction can be challenging. This becomes much harder if we want at least six - seven effects per half-normal plot. Ranjan et al. suggest the following rule:

1. If *there exists a set of disjoint flats* that satisfy the size restriction, construct a *full/partial balanced/mixed spread*. Then relabel the points of $\mathcal{P}_n$ (i.e., all $2^n - 1$ factorial effects) to meet the restriction on the elements of RDCSSs.
2. If *there does not exist a set of disjoint flats* that satisfy the size restriction, construct a *full/partial balanced/mixed star.* Then relabel

the points of $\mathcal{P}_n$ (i.e., all 2^n factorial effects) to meet the restriction on the elements of RDCSSs.

In this paper, we will only discuss the construction of balanced (full) spreads and balanced covering stars of $PG(n-1,2)$. First we briefly review the construction of spreads and stars, then we present a tool for relabeling the factorial effects.

Recall that a balanced $(t-1)$-spread ψ of $\mathcal{P}_n$ exists if and only if $t|n$, and in that case, $\mu = |\psi| = (2^n-1)/(2^t-1)$. The construction of such a spread ψ starts with writing the 2^n-1 nonzero elements of $GF(2^n)$ in cycles of length μ (see Table 4.1). This cyclic construction is ensured by the condition $t|n$ (Hirschfeld 1998).

Table 4.1 Cycles of length μ for a $(t-1)$-spread of $PG(n-1,2)$.

S_1	S_2	......	S_μ
0	1	......	$\mu-1$
μ	$\mu+1$	......	$2\mu-1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(2^t-2)\mu$	$(2^t-2)\mu+1$	......	$(2^t-1)\mu-1$

The elements of this cycle are combined with a root of a primitive polynomial over $GF(2)$ of degree n. Let w be a root of a primitive polynomial of degree n for $GF(2^n)$. Then, the 2^n-1 elements of $\mathcal{P}_n$ are w^i, $i = 0, ..., 2^n-2$, where

$$w^i = \alpha_0 w^{n-1} + \alpha_1 w^{n-2} + \cdots + \alpha_{n-2} w + \alpha_{n-1}$$

represents an r-factor interaction $\delta = (\alpha_0, \alpha_1, ..., \alpha_{n-1})$, for $\alpha_i \in GF(2)$, if exactly r elements in δ are nonzero. For instance, a 1-spread of $PG(3,2)$ consists of five 1-flats (see Table 4.2). Note that the non-trivial value of t that divides $n=4$ is $t=2$ (i.e., a 1-spread).

Table 4.2 A 1-spread ψ_1 of $PG(3,2)$ obtained using the cyclic construction.

S_1	S_2	S_3	S_4	S_5
D	C	B	A	CD
BC	AB	ACD	BD	AC
BCD	ABC	$ABCD$	ABD	AD

For $n=5$ (or any prime n), there does not exist any non-trivial balanced spread of $\mathcal{P}_n$. For $n=6$, both 1-spread and 2-spread exist. Table 4.3 presents the cyclic 2-spread of $\mathcal{P}_6$.

We now discuss the construction of a balanced covering star $St(n;\mu;t;t_0)$ of $\mathcal{P}_n$. From Theorem 3 of Ranjan et al. (2009), one can construct two disjoint flats, π (a (t_0-1)-flat) and $\mathcal{U}$ (a $(n-t_0-1)$-flat) of $\mathcal{P}_n$. Assuming

Table 4.3 A 2-spread ψ_2 of $PG(5,2)$ using the cyclic construction.

S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
F	E	D	C	B	A	EF	DE	CD
BC	AB	AEF	DF	CE	BD	AC	BEF	ADE
CDEF	BCDE	ABCD	ABCEF	ABDF	ACF	BF	AE	DEF
CDE	BCD	ABC	ABEF	ADF	CF	BE	AD	CEF
BDE	ACD	BCEF	ABDE	ACDEF	BCDF	ABCE	ABDEF	ACDF
BCF	ABE	ADEF	CDF	BCE	ABD	ACEF	BDF	ACE
BDEF	ACDE	BCDEF	ABCDE	ABCDEF	ABCDF	ABCF	ABF	AF

$(t-t_0)|(n-t_0)$, let $\psi = \{S_1,...,S_\mu\}$ be a $(t-t_0-1)$-spread of $\mathcal{U}$, then a balanced covering star $St(n;\mu;t;t_0)$ of $\mathcal{P}_n$ is $\Omega = \{S_1^*,...,S_\mu^*\}$ with μ rays $S_i^* = \langle S_i, \pi\rangle$ and nucleus π. For instance, a covering star $St(5;4;3;1)$ of $\mathcal{P}_5$ can be constructed using ψ_1 in Table 4.2. Let $\pi = \{E\}$ and $\mathcal{U} = \langle A,B,C,D\rangle$, then the resulting star is $\Omega_1 = \{S_1^*,...,S_5^*\}$, where $S_i^* = \langle S_i, E\rangle$ and $S_i \in \psi_1$, $i = 1,...,5$ (see Table 4.4).

Table 4.4 A balanced covering star $\Omega_1 = St(5;5;3;1)$ of $PG(4,2)$ based on ψ_1 (Table 4.2).

S_1^*	S_2^*	S_3^*	S_4^*	S_5^*
D	C	B	A	CD
BC	AB	ACD	BD	AC
BCD	ABC	$ABCD$	ABD	AD
E	E	E	E	E
DE	CE	BE	AE	CDE
BCE	ABE	$ACDE$	BDE	ACE
$BCDE$	$ABCE$	$ABCDE$	$ABDE$	ADE

Note that the spreads and stars obtained in this manner may not satisfy the experimenter's requirement on the containment of specific randomization restrictions factors. Despite using a different primitive polynomial or perhaps alternative construction methods (e.g., transversal approach), the resulting spread (or star) may require relabeling of the factorial effects to meet the needs for the experimental design. The relabeling of factorial effects or points of $\mathcal{P}_n$ can be done using a *collineation* (Coxeter 1974; Batten 1997).

Definition 4.1 A collineation of $\mathcal{P}_n = PG(n-1,2)$ is a permutation of its points such that $(t-1)$-flats are mapped to $(t-1)$-flats, for $1 \le t \le n$. Such a permutation can be characterized by an $n \times n$ matrix referred to as the collineation matrix (denoted by $\mathcal{C}$).

This collineation matrix $\mathcal{C}$ can be used to map any element $\delta \in \mathcal{P}_n$ to another element $\delta' = \mathcal{C}\cdot\delta \in \mathcal{P}_n$, where δ and δ' are column vectors. Such a collineation matrix can be constructed by mapping a set of n linearly independent basis elements, typically, the canonical basis (i.e., the main effects or n basic factors), to another set of n linearly independent basis elements (or factorial effects) of $\mathcal{P}_n$. Let $\mathcal{C}_n$ be the set of all distinct $n \times n$ collineation matrices for $\mathcal{P}_n = PG(n-1,2)$, then the total number of collineation matrices for $\mathcal{P}_n$ is $|\mathcal{C}_n| = \prod_{i=1}^{n}(2^n - 2^{i-1})$. The result follows from a counting argument based on the total number of possible options for mapping the i-th independent basis element given that $i-1$ linearly independent basis elements have already been mapped. We now present a few examples to illustrate how to construct such collineation matrices, and how they are used in constructing factorial designs with randomization restrictions.

Example 4.1 Suppose we wish to construct a two-stage 2^6 split-lot experiment in 64 runs with randomization restrictions defined by $S_1 \supset \{A, B, C\}$ and $S_2 \supset \{D, E, F\}$. This is a toy example, as one can simply construct $S_1 = \langle A, B, C \rangle$ and $S_2 = \langle D, E, F \rangle$ to meet all the desired features (the factor assignment based on the randomization restrictions and the size requirement for useful half-normal plots). Nonetheless, we use this simple example to illustrate the steps of construction (more complicated examples will follow).

Since $t = 3$ divides $n = 6$, there exists a 2-spread of $\mathcal{P}_6$ (e.g., ψ_2 in Table 4.3). Clearly, all S_i's meet the size requirement, however, the elements of ψ_2 will have to be relabeled to meet the factor assignment or randomization restrictions. That is, we need to construct a collineation matrix $\mathcal{C}$ for the desired relabeling.

1. *Identify the flats (S_i's in ψ_2) to be mapped.* Since n independent effects (or pencils) are required for constructing the collineation matrix, and each S_i contains t independent effects, we need n/t (here, $n/t = 6/3 = 2$) flats from ψ_2. Suppose we choose S_1 and S_2.
2. *Identify the set of n linearly independent effects from the chosen flats that will be mapped.*

   Suppose we wish to establish the following mapping:

$$\begin{array}{llll} S_1: & F \to A, & BC \to B, & CDEF \to C, \\ S_2: & E \to D, & AB \to E, & BCDE \to F. \end{array} \tag{4.1}$$

If the mapping in (4.1) is defined by $\mathcal{C}$, the inverse mapping $\mathcal{C}^{-1}$ is characterized by

$$\begin{array}{lll} A \to F, & B \to BC, & C \to CDEF, \\ D \to E, & E \to AB, & F \to BCDE. \end{array}$$

The corresponding matrices are

$$\mathcal{C}^{-1} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{C} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \end{pmatrix},$$

and the transformed spread $\psi_2' = \mathcal{C}^{-1} \cdot \psi_2$, which satisfied the experimenter's requirement, is shown in Table 4.5.

Table 4.5 The transformed 2-spread $\psi_2' = \mathcal{C}^{-1} \cdot \psi_2$ of $PG(5, 2)$ for Example 4.1.

S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
A	D	BDF	ABCF	ACF	ACEF	AD	BF	ACD
B	E	CDEF	ABDF	ABCDF	ABCD	BE	CDF	ABCE
C	F	ACDE	BCDEF	ABDEF	ABE	CF	ACDEF	ABF
AC	DF	ABCEF	ADE	BCDE	BCF	ACDF	ABCDE	BCDF
ABC	DEF	ABD	BEF	AEF	ADF	ABCDEF	ABEF	ADEF
AB	DE	BCE	CD	BD	BDEF	ABDE	BCD	BDE
BC	EF	AF	ACE	CE	CDE	BCEF	AE	CEF

Example 4.2 Consider a three-stage 2^6 factorial experiment with the randomization restrictions defined by $S_1 \supset \{A, B, C\}$, $S_2 \supset \{D, E\}$ and $S_3 \supset \{F\}$. In this case, $S_1 = \langle A, B, C \rangle$, $S_2 = \langle D, E \rangle$ and $S_3 = \langle F \rangle$ are all pairwise disjoint, but the minimum size requirement would be violated, and the assessment of several potentially important effects in $\langle D, E, F \rangle$ will have to be sacrificed.

Since there exists a 2-spread of $\mathcal{P}_6$, one can construct a collineation matrix $\mathcal{C}$ to transform ψ_2 (Table 4.3) to $\psi_2'' = \mathcal{C}^{-1} \cdot \psi_2$ that meets the experiment's restriction on the randomization of the trials and meets the size requirement for efficient analysis. The desired collineation matrix $\mathcal{C}$ can be obtained using the following mapping:

$$\begin{array}{llll} S_1: & F \to A, \quad BC \to B, \quad CDEF \to C, & & \\ S_2: & E \to D, \quad AB \to E, & S_3: & D \to F. \end{array}$$

Example 4.3 Consider a three-stage 2^5 split-lot design in 32 experimental units with restrictions $S_1 \supset \{A, B\}$, $S_2 \supset \{C\}$ and $S_3 \supset \{D, E\}$ (e.g., in the plutonium alloy experiment of Bingham et al. 2008). Similar to Example 4.2, the obvious set of disjoint RDCSSs $S_1 = \langle A, B \rangle$, $S_2 = \langle C \rangle$ and $S_3 = \langle D, E \rangle$ violate the minimum size requirement for efficient analysis via half-normal plots. It is desired to construct RDCSSs with $t_i \geq 3$. However, according to Theorem 3 of Ranjan et al. (2009), since $t(=3) > n/2(=2.5)$, there does not exist even two disjoint 2-flats in $PG(4,2)$, and an overlap cannot be avoided.

Bingham et al. (2008) used a computer search to come up with $S_1 = \langle A, B, ABCDE\rangle$, $S_2 = \langle C, AD, ABCDE\rangle$ and $S_3 = \langle D, E, ABCDE\rangle$. From an analysis viewpoint, this design is acceptable because the significance of all effects except $ABCDE$ can be assessed. Note that the RDCSSs, S_1, S_2, S_3, are rays of a balanced covering star $\Omega_1' = St(5;5;3;1) = \{S_1^*, S_2^*, S_3^*, S_4^*, S_5^*\}$ of $PG(4,2)$, with $S_i^* = S_i$ for $i = 1, ..., 3$, and Ω_1' is obtained by applying the following mapping to Ω_1 in Table 4.4 :

$$\begin{array}{llll} S_1^*: & D \to A, \quad BC \to B, & & \\ S_2^*: & C \to C, & S_3^*: & ACD \to D, \quad BE \to E. \end{array}$$

This begs the question whether or not there exist other stars that can be used to obtain a better design in this case. Recall that a balanced covering star $St(n;\mu;t;t_0)$ of $PG(n-1,2)$ exists if and only if $(t-t_0)|(n-t_0)$.

- For $t_0 = 1$ and $n = 5$, the feasible options for constructing stars are: $t = 2$ and $t = 3$. (Bingham et al. used a star with $t = 3$).
- For $t_0 = 2$ and $n = 5$, only $t = 3$ leads to a feasible star with $\mu = (2^3-1)(2^1-1) = 7$ rays (the existence follows from Lemma 3.3).
- For $t_0 = 3$ and $n = 5$, the only possible star is with $t = 4$ and $\mu = (2^2-1)(2^1-1) = 3$ (the existence follows from Lemma 3.3).

Note that the last case with $t_0 = 3$ and $t = 4$ leads to the best design from an analysis viewpoint as all factorial effects can be assessed using four half-normal plots, one for the nucleus π (with seven effects) and three plots for $S_i \backslash \pi$, for $i = 1, 2, 3$ (each with eight effects). We now implement the systematic procedure discussed above for constructing such a star-based design. We will construct this design in two steps: first construct a $St(5;3;4;3)$, and then use an appropriate collineation to match the randomization restrictions.

- Similar to the spread-to-star construction shown in Table 4.4, one can construct $\Omega = St(5;3;4;3)$ as follows: Let $\pi = \langle A, B, C\rangle$ and $\mathcal{U} = \langle D, E\rangle$ be two disjoint flats of $\mathcal{P}_5$. Then, the star obtained from this construction is $\Omega = \{\langle D, \pi\rangle, \langle E, \pi\rangle, \langle DE, \pi\rangle\}$ as $\psi = \{\{D\}, \{E\}, \{DE\}\}$ is the only 0-spread of $\mathcal{U}$. That is,

$$
\begin{aligned}
S_1^* &= \{A, B, AB, C, AC, BC, ABC, D, AD, BD, ABD, CD, ACD, BCD,\\
&\quad ABCD\},\\
S_2^* &= \{A, B, AB, C, AC, BC, ABC, E, AE, BE, ABE, CE, ACE, BCE,\\
&\quad ABCE\},\\
S_3^* &= \{A, B, AB, C, AC, BC, ABC, DE, ADE, BDE, ABDE, CDE,\\
&\quad ACDE, BCDE, ABCDE\}.
\end{aligned}
$$

- Clearly, the RDCSSs characterized by Ω do not satisfy the experimenter's restriction on the randomization of factors ($S_1 \supset \{A, B\}$, $S_2 \supset \{C\}$ and $S_3 \supset \{D, E\}$), and a relabeling is needed. The desired collineation matrix $\mathcal{C}$ (for relabeling $\mathcal{P}_5$) can be obtained by the following mapping:

$$
\begin{array}{lllll}
S_1: & AD \to A, & BD \to B, & S_2: & E \to C,\\
S_3: & DE \to D, & CDE \to E. & &
\end{array}
$$

The resulting star $\Omega' = \mathcal{C}^{-1} \cdot \Omega$ is given by

$$
\begin{aligned}
S_1^* &= \{ACD, BCD, AB, ACE, DE, ABDE, BCE, CD, A, B, ABCD,\\
&\quad ADE, CE, ABCE, BDE\},\\
S_2^* &= \{ACD, BCD, AB, ACE, DE, ABDE, BCE, C, AD, BD, ABC,\\
&\quad AE, CDE, ABCDE, BE\},\\
S_3^* &= \{ACD, BCD, AB, ACE, DE, ABDE, BCE, D, AC, BC, ABD,\\
&\quad ACDE, E, ABE, BCDE\},
\end{aligned}
$$

where the common overlap (nucleus) is $\pi = \langle ACD, AB, DE\rangle$. The main idea behind choosing these relabelings was to make sure that the main effects are not in the overlap.

Although the choice of n independent mappings for constructing a collineation matrix $\mathcal{C}$ is not difficult, one may have to try a few combinations to make sure that the set of n mappings are linearly independent. Assuming one can construct several spread/star - based designs that satisfy the experimenter's requirements, the next section presents techniques for finding good candidates from a class of acceptable designs.

5. Ranking of RDCSS-based Designs

The construction methods for spreads (stars) and different choices of collineation matrices may lead to several acceptable designs that satisfy the randomization restrictions imposed by the experimenter, and the size restriction for half-normal plot analysis. In such a case, one might want to rank the designs based on some design criteria, and/or check whether the designs obtained are distinct or isomorphic. In this section, we briefly discuss a few ranking criteria, and an algorithm for checking isomorphism of such designs.

5.1 Design criteria

Several design criteria have been proposed thus far for ranking regular and non-regular fractional factorial designs, for instance, *maximum resolution* criterion (Box and Hunter 1961), *minimum aberration* (Fries and Hunter 1980), *maximum number of clear effects* (Wu and Chen 1992) and *generalized minimum aberration* (Deng and Tang 1999). Though most of the literature on design ranking criteria focused on fractions of completely randomized 2^k factorial designs, some of these criteria have also been generalized for designs with randomization restrictions like split-plot designs (e.g., Bingham and Sitter 1999; Mukerjee and Fang 2002). Note that most of these criteria are based on the word length pattern of the underlying *fraction defining contrast subgroup* (FDCSG).

It is straightforward to draw parallelism between FDCSGs and RDCSSs, however it is important to note that a typical design criterion is based on only one FDCSG, whereas in the current framework, we have multiple RDCSSs even in the unreplicated full factorial case. For fractional factorial designs with randomization restrictions, we will have to combine stage-wise FDCSGs with RDCSSs to account for the variance distribution of the new factors. To avoid the additional complexity due to fractionation, we only discuss the word length pattern for unreplicated full factorial RDCSS-based designs.

Alternatively, one can use the V-criterion developed for multistage factorial designs with randomization restrictions (Bingham et al. 2008). The V-criterion is computed as

$$V = \sum_{j=1}^{m^*} (p_j - \bar{p})^2,$$

where m^* is the total number of half-normal plots required for the significance assessment of factorial effects, p_j is the proportion of main effects and 2fi's in the j-th half-normal plot, and $\bar{p}$ the average of p_j's. Bingham et al. recommend minimizing the value of V.

For an m-stage 2^n unreplicated full factorial design, m^* is typically greater than m. In the case of a spread-based design, most likely $m^* = m+1$

(the additional half-normal plot corresponds to $\mathcal{P}_n \backslash \{S_1 \cup \cdots \cup S_m\}$). For a star-based design, m^* can be $m+1$ or $m+2$. If the number of rays $\mu = m$, then $m^* = m+1$ (m half-normal plots correspond to $S_j \backslash \pi$, and one half-normal plot is required for the effects in the nucleus). On the other hand, if $\mu > m$, then $m^* = m+2$, where the effects in the rays (excluding nucleus) not used for the RDCSSs are assessed separately.

Suppose the RDCSS at stage j ($1 \leq j \leq m$) is characterized by $S_j = \langle \delta_1^{(j)}, \delta_2^{(j)}, ..., \delta_{t_j}^{(j)} \rangle$, where $\{\delta_1^{(j)}, \delta_2^{(j)}, ..., \delta_{t_j}^{(j)}\}$ is the set of t_j linearly independent *randomization factors* (or factorial effects) of $\mathcal{P}_n$. Let $\mathcal{A}_d$ be the set of all admissible designs that satisfy the randomization restrictions imposed by the experimenter, and meet the size requirement for efficient analysis (i.e., six-seven effects per half-normal plot). Our objective is to rank the designs in $\mathcal{A}_d$. Note that the randomization restrictions imposed by the experimenter (e.g., $S_1 \supset \{A, B\}$ and $S_2 \supset \{C\}$) characterizes parts of a few RDCSSs, and the remaining randomization factors ($\delta_l^{(j)}$) can be varied to choose an optimal design under a specific criterion.

Example 5.1 Consider the setup in Example 4.2, a three-stage 2^6 factorial experiment in 64 runs with randomization restrictions defined by $S_1 \supset \{A, B, C\}$, $S_2 \supset \{D, E\}$ and $S_3 \supset \{F\}$. From an analysis viewpoint, we should construct S_i's such that $|S_i| \geq 2^3 - 1$, i.e., each S_i must have three linearly independent randomization factors. We also know that $t(=3)|n(=6)$, which guarantees the existence of a 2-spread of $\mathcal{P}_6$ (e.g., in Tables 4.3 and 4.5). In this case, the set of all admissible designs is $\mathcal{A}_d = \{S_1 = \langle A, B, C \rangle, S_2 = \langle D, E, \delta_3^{(2)} \rangle, S_3 = \langle F, \delta_2^{(3)}, \delta_3^{(3)} \rangle$, where $\delta_3^{(2)}, \delta_2^{(3)}$ and $\delta_3^{(3)}$ are chosen so that S_i's are disjoint. All eligible options for $\delta_3^{(2)}$ and $\delta_2^{(3)}$ can be found systematically in the following manner:

- $\delta_3^{(2)} \in \mathcal{P}_6 \backslash \{\langle D, E, S_1 \rangle \cup \langle D, E, F \rangle\}$, and
- assuming $\delta_3^{(2)}$ is known, $\delta_2^{(3)} \in \mathcal{P}_6 \backslash \{\langle F, S_1 \rangle \cup \langle F, S_2 \rangle\}$.

There are more than one hundred designs in $\mathcal{A}_d$. Let $d_1 \in \mathcal{A}_d$ be defined by $\delta_3^{(2)} = AF, \delta_2^{(3)} = ACD$ and $\delta_3^{(3)} = ABE$. Then,

$$\begin{aligned} S_1 &= \{A, B, AB, C, AC, BC, ABC\}, \\ S_2 &= \{D, E, DE, AF, ADF, AEF, ADEF\}, \\ S_3 &= \{F, ACD, ACDF, ABE, ABEF, BCDE, BCDEF\}, \end{aligned}$$

and $\mathcal{P}_6 \backslash \{S_1 \cup S_2 \cup S_3\}$ contains the remaining 42 factorial effects (10- 2fi's, 15- 3fi's, 11-4fi's, 5- 5fi's and 1- 6fi). In this case $m^* = 4$ half-normal plots are required for significance assessment of all factorial effects. In terms of V-criterion, the distribution of p_i is: $p_1 = 6/7, p_2 = 4/7, p_3 = 1/7, p_4 = 10/42$, and hence $V_{d_1} = 0.1065$. Table 5.1 compares the values of V-criterion for four candidate designs (including d_1) from $\mathcal{A}_d$.

Table 5.1 Ranking on a few three-stage 2^6 factorial designs using V-criterion

	$\delta_3^{(2)}, \delta_2^{(3)}, \delta_3^{(3)}$	V-criterion
d_1	AF, ACD, ABE	0.1065
d_2	$AF, ACD, ABCE$	0.1065
d_3	ABF, ACD, CDE	0.0793
d_4	$ABF, ACD, ABCDEF$	0.0976

Clearly, d_3 is better than d_1, d_2 and d_4 in terms of V-criterion. We now look at full word length pattern of the RDCSSs of d_1, d_2, d_3 and d_4 for information like resolution and minimum aberration. For any $d_i \in \mathcal{A}_d$ and $j = 1, ..., m^*$, let $W^{d_i}_{*,j} = (w^{d_i}_{1,j}, ..., w^{d_i}_{n,j})$ be the word length pattern for j-th half-normal plot, where $w^{d_i}_{k,j}$ denotes the number of words of length k (or k-factor interactions) on j-th half-normal plot. Then the word length pattern for all four designs are as follows:

$$W^{d_1}_{*,1} = (3,3,1,0,0,0), W^{d_1}_{*,2} = (2,2,2,1,0,0), W^{d_1}_{*,3} = (1,0,2,3,1,0),$$
$$W^{d_1}_{*,4} = (0,10,15,11,5,1),$$
$$W^{d_2}_{*,1} = (3,3,1,0,0,0), W^{d_2}_{*,2} = (2,2,2,1,0,0), W^{d_2}_{*,3} = (1,0,2,3,1,0),$$
$$W^{d_2}_{*,4} = (0,10,15,11,5,1),$$
$$W^{d_3}_{*,1} = (3,3,1,0,0,0), W^{d_3}_{*,2} = (2,1,1,2,1,0), W^{d_3}_{*,3} = (1,1,3,2,0,0),$$
$$W^{d_3}_{*,4} = (0,10,15,11,5,1),$$
$$W^{d_4}_{*,1} = (3,3,1,0,0,0), W^{d_4}_{*,2} = (2,1,1,2,1,0), W^{d_4}_{*,3} = (1,0,3,1,1,1),$$
$$W^{d_4}_{*,4} = (0,11,15,12,4,0).$$

This example shows that if the word length patterns of two designs d and d' are same, then $V_d = V_{d'}$, for instance, $w^{d_1}_{k,j} = w^{d_2}_{k,j}$ for all k. Note that the reverse is not necessarily true.

Example 5.2 Consider the setup of Example 4.3, where the objective was to construct a three-stage 2^5 split-lot design in 32 runs with randomization restriction defined by $S_1 \supset \{A, B\}$, $S_2 \supset \{C\}$ and $S_3 \supset \{D, E\}$. A design that meets the experimenter's restrictions and the size requirements is a balanced covering star $St(5;3;4;3)$ given by three rays

$$\begin{aligned} S_1 &= \{ACD, BCD, AB, ACE, DE, ABDE, BCE, CD, A, B, ABCD, ADE, CE, \\ &\quad ABCE, BDE\}, \\ S_2 &= \{ACD, BCD, AB, ACE, DE, ABDE, BCE, C, AD, BD, ABC, AE, CDE, \\ &\quad ABCDE, BE\}, \\ S_3 &= \{ACD, BCD, AB, ACE, DE, ABDE, BCE, D, AC, BC, ABD, ACDE, E, \\ &\quad ABE, BCDE\}, \end{aligned}$$

and the nucleus $\pi = \{ACD, BCD, AB, ACE, DE, ABDE, BCE\}$. For this design, $m^* = 4$ (three half-normal plots for $S_j \backslash \pi, j = 1, 2, 3$, and the fourth one for the nucleus). The distribution of p_j is: $p_1 = 4/8, p_2 = 5/8, p_3 = 4/8$ and $p_4 = 2/7$, and thus $V = 0.0198$. The word length pattern for this design is

$$W_{*,1} = (2,2,2,2,0),\ W_{*,2} = (1,4,2,0,1),\ W_{*,3} = (2,2,2,2,0),\ W_{*,4} = (0,2,4,1,0).$$

It turns out that there is a unique design based on $St(5;3;4;3)$ that meets the experimenter's restriction: $S_1 \supset \{A, B\}$, $S_2 \supset \{C\}$ and $S_3 \supset \{D, E\}$, and allow the significance assessment of all factorial effects in $\mathcal{P}_5$.

In this paper we have only presented the generalization of the word length pattern for RDCSS-based designs, and not for specific design criteria like maximum resolution, generalized minimum aberration, maximum estimation capacity etc. There are several ways to combine the information in these word length patterns for coming up with design criteria similar to the usual fractional factorial designs. However, there are a few important points worth noting. For any given design d_i, all $W_{*,j}^{d_i}$'s are not independent. This dependence is introduced due to the total number of k-factor interactions: $\sum_{j=1}^{m^*} w_{k,j}^{d_i} = \binom{n}{k}$ for all $1 \leq k \leq n$. More importantly, the randomization restrictions are often imposed on the main effects (or basic factors) which leads to non-zero values of $w_{1,j}^{d_i}$ that correspond to $S_1, ..., S_m$. Consequently, it is undesirable to sequentially minimize $w_{k,j}^{d_i}, k = 1, ..., n$ (a basis for most of the efficient design criteria based on FDCSGs).

5.2 Isomorphism check

Since stars are generalizations of spreads, we will focus on the isomorphism of stars (particularly the balanced case). Similar results on spreads can easily be deduced, and are also generalizable for mixed cases. Suppose we wish to check isomorphism of two balanced covering stars $St(n; \mu; t; t_0)$ of $\mathcal{P}_n = PG(n-1, 2)$. Similar to regular factorial designs, the isomorphism check algorithm for stars and spreads has two main parts: (a) check for the relabeling of factors and factor combinations (i.e., points of $\mathcal{P}_n$), and (b) check for the rearrangement (or equivalence) of the factorial effects in

RDCSSs (or rays). We use collineation matrices to relabel the factorial effects of $\mathcal{P}_n$.

Definition 5.1 Two balanced covering stars $St(n;\mu;t;t_0)$ of $\mathcal{P}_n = PG(n-1,2)$ (denoted by Ω_1 and Ω_2) are said to be equivalent (i.e., $\Omega_1 \equiv \Omega_2$) if and only if, for every $f_i^* \in \Omega_1$, $\exists g_j^* \in \Omega_2$ such that $f_i^* = g_j^*$. Furthermore, Ω_1 and Ω_2 are said to be isomorphic (i.e., $\Omega_1 \cong \Omega_2$) if and only if, there exists a collineation matrix $\mathcal{C}$ of $\mathcal{P}_n$ such that $\mathcal{C} \cdot \Omega_1 \equiv \Omega_2$.

Rearrangement of the factorial effects within a ray and changing the order of the rays generate equivalent stars. From Definition 5.1, equivalence check is an important step in checking isomorphism of stars. Let $\mathcal{E}(\Omega)$ denote the equivalence class of a star Ω. Then, for a balanced covering star $\Omega = St(n;\mu;t;t_0)$ of $\mathcal{P}_n$, the size of the equivalence class of Ω is $|\mathcal{E}(\Omega)| = \mu! \cdot [(2^t-1)!]^{\mu}$, where $\mu = (2^{n-t_0}-1)/(2^{t-t_0}-1)$. That is, for every collineation matrix in $\mathcal{C}_n$ (with $|\mathcal{C}_n| = \prod_{i=1}^{n}\left(2^n - 2^{i-1}\right)$), one may have to make $|\mathcal{E}(\Omega_1)|$ many comparisons to verify $\mathcal{C} \cdot \Omega_1 \equiv \Omega_2$, which can be computationally very intensive. We propose several techniques for reducing the computational burden.

For checking equivalence of stars (or spreads) we propose a prime representation (Λ) of the assigning points of $\mathcal{P}_n$. It starts with sorting the points in Yates order, and then assigning the i-smallest prime number to the i-th point of $\mathcal{P}_n$. As a result, a $(t-1)$-flat f^* of $\mathcal{P}_n$ is represented by the product of the primes that corresponds to the elements of f^*. For example, the sorted elements of $\mathcal{P}_3$ are $\{A, B, AB, C, AC, BC, ABC\}$, with the prime representation: $A \to 2$, $B \to 3$, ..., $ABC \to 17$. Furthermore, for $f^* = \{AB, AC, BC\}$, $\Lambda(f^*) = 5 \cdot 11 \cdot 13 = 715$. Note that the *unique-prime-factorization theorem* facilitates the reverse identification of elements of f^* given the value of $\Lambda(f^*)$. This approach is efficient for equivalence check because we only need a sequence of μ scalars, instead of an array of vectors, to characterize a star.

The number of collineation checks can be reduced using Lemma 3.2 which establishes a one-to-one correspondence between the existence of a balanced covering star $St(n;\mu;t;t_0)$ of $\mathcal{P}_n$ and a $(t-t_0-1)$-spread of $\mathcal{P}_{n-t_0}$. This correspondence can be further strengthened for isomorphism as well. For checking the isomorphism of Ω_1 and Ω_2, we can assume $\pi_1 \equiv \pi_2$ without loss of generality (otherwise, a relabeling can be used to enforce $\pi_1 \equiv \pi_2$).

Lemma 5.1 Let Ω_1 and Ω_2 be two balanced stars $St(n;\mu;t;t_0)$ of $PG(n-1,2)$ with the same nucleus π. Then, $\Omega_1 \cong \Omega_2$ if and only if $\psi_1 \cong \psi_2$, where ψ_1 and ψ_2 are $(h-1)$-spreads of $PG(u-1,2)$, such that $\Omega_j = \psi_j \times \pi$ for $j = 1,2$, $h = t - t_0$ and $u = n - t_0$.

Lemma 5.1 implies that the maximum number of collineation checks for checking the isomorphism of Ω_1 and Ω_2 is $|\mathcal{C}_{n-t_0}|$, which is often much smaller than $|\mathcal{C}_n|$ (by a factor of $2^{t_0 \cdot (n-t_0)} \cdot \prod_{i=1}^{t_0}(2^n - 2^{i-1})$). For instance, if $n = 6$, $t = 4$ and $t_0 = 2$, $|\mathcal{C}_{n-t_0}| = 20160$, whereas $|\mathcal{C}_n| = 20158709760$. For Example 5.1, $n = 5$, $t = 4$, $t_0 = 3$, and thus $|\mathcal{C}_{n-t_0}| = 6$ and $|\mathcal{C}_n| = 9999360$. Next, we propose a technique of reducing this search space further.

A necessary condition for $\mathcal{C} \in \mathcal{C}_{n-t_0}$ that establishes $\psi_1 \cong \psi_2$ (i.e., $\mathcal{C} \cdot \psi_1 \equiv \psi_2$) is that every $f \in \psi_1$ maps to a unique $g \in \psi_2$ such that $\mathcal{C} \cdot f = g$. Since ψ_1 is a $(h-1)$-spread of $PG(u-1, 2)$, every flat in ψ_1 contains h linearly independent basis elements (or factorial effects). Thus, a set of u/h linearly independent flats of ψ_1 is sufficient for defining the required $\mathcal{C}$ that would establish $\psi_1 \cong \psi_2$. Let $\mathcal{D} \subseteq \mathcal{C}_{n-t_0}$ be the set of collineations that map u/h flats of ψ_1 to ψ_2. Then $\mathcal{D}$ contains all collineations that establish $\psi_1 \cong \psi_2$, and

$$|\mathcal{D}| = \binom{\mu}{u/h} \cdot (u/h)! \cdot |\mathcal{C}_h|^{(u/h)}.$$

For the example, $n = 6, t = 4$ and $t_0 = 2$, $|\mathcal{C}_{n-t_0}| = 20160$, whereas $|\mathcal{D}| = 720$ (another substantial reduction in the number of checks). In general, $|\mathcal{D}| < |\mathcal{C}_{t-t_0}|$ if $h \geq 2$. For Example 5.2, the size of $\mathcal{D}$ is the same as that of $\mathcal{C}_{n-t_0}$ as $h = 1$.

The isomorphism check algorithm presented here checks whether or not two balanced covering stars (or spreads if $t_0 = 0$) of $\mathcal{P}_n$ are isomorphic to each other. Finding all non-isomorphic stars would require generating all possible stars $St(n; t_1, ..., t_\mu; t_0)$, which is a non-trivial problem. It is known that all balanced spreads of $PG(n-1, 2)$ for $n \leq 5$ are isomorphic to each other, there exist 131044 non-isomorphic 1-spreads of $\mathcal{P}_6$ (Mateva and Topalova 2009), and all 2-spreads of $\mathcal{P}_6$ are isomorphic to each other (Topalova and Zhelezova 2009).

6. Concluding Remarks

In this paper we discussed an RDCSS-based approach for unifying factorial designs with multiple randomization restrictions. For the simplicity of examples and popularity in industrial experiments the results and algorithms presented here focus on $PG(n-1, 2)$, however, the theoretical results can easily be generalized for $PG(n-1, q)$ (more precisely, all existence results would be the same except the sizes of the spreads and stars).

To date, there have been a few other attempts to unify such multi-stage factorial designs, for instance, Patterson and Bailey (1978) used design keys, Nelder (1965a, b) developed simple block structure, Speed and Bailey (1982) developed it further to orthogonal block structure and used association schemes, and Butler (2004) proposed grid representations. One can establish parallelism between many of these approaches with the spread-based design, however, we believe that the star-based designs are more powerful.

When checking equivalence of stars (or spreads), a bitstring representation of $\mathcal{P}_n$ can be used as an alternative to the prime representation (Λ), which can be more efficient in some cases. We believe that the number of collineation checks can also be reduced further when checking isomorphism of mixed spreads and unbalanced stars of $\mathcal{P}_n$.

REFERENCES

Addelman, S. (1964). Some two-level factorial plans with split plot confounding *Technometrics*, **6** 253-258.

Batten, L. M. (1997). *Combinatorics of finite geometries.* Cambridge University Press, 2nd edition, Cambridge, UK.

Bingham, D.R., Schoen, E.D. and Sitter, R.R. (2004). Designing Fractional Factorial Split-Plot Experiments with Few Whole-Plot Factors, *J. Royal Statist. Soci.* **53** 325-339.

Bingham, D. and Sitter, R. R. (1999). Minimum-aberration two-level fractional factorial split-plot designs. *Technometrics* **41** 62-70.

Bingham, D., Sitter, R., Kelly, E., Moore, L., and Olivas, J. D. (2008). Factorial designs with multiple levels of randomization. *Statist. Sinica* **18** 493-513.

Bisgaard, S. (1994). A note on the defnition of resolution for blocked 2^{k-p} designs. *Technometrics* **36** 308-311.

Box, G. E. P. and Hunter, J. S. (1961). The 2^{p-k} fractional factorial designs. Part II. *Technometrics* **3** 449-458.

Coxeter, H. S. M. (1974). *Projective geometry.* University of Toronto Press, Toronto, Ont., second edition.

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

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

Hirschfeld, J. W. P. (1998). Projective geometries over finite fields. *Oxford Univ. Press.*

Mateva, Z. and Topalova, S. (2009). Line spreads of $PG(5, 2)$, *J. Combinatorial Designs* **17** 90-102.

Mee, R. W. and Bates, R. L. (1998). Split-lot designs: experiments for multistage batch processes. *Technometrics* **40** 127-140.

Miller, A. (1997). Strip-plot configurations of fractional-factorials. *Technometrics* **39** 153-161.

Mukerjee, R. and Fang, K. T. (2002). Fractional factorial split-plot designs with minimum aberration and maximum estimation capacity. *Statist. Sinica* **12** 885-903.

Ranjan, P., Bingham, D. and Dean, A. (2009). Existence and Construction of Randomization Defining Contrast Subspaces for Regular Factorial Designs, *The Annals of Statist.* **37** 3580-3599.

Ranjan, P., Bingham, D. and Mukerjee, R. (2010), Stars and Regular Fractional Factorial Designs with Randomization Restrictions, *Statistica Sinica* **20** 1637-1653.

Schoen, E. D. (1999). Designing fractional two-level experiments with nested error structures. *J. Appli. Statist.* **26** 495-508.

Spencer, N. (2013). Isomorphism of Spreads and Stars, B.Sc. Thesis, *Acadia University.*

Topalova, S. and Zhelezova, S. (2008). Classification of the 2-spreads of $PG(5, 2)$, *Math. and Edu. Math.* 163-169.

Vivacqua, C. A. and Bisgaard, S. (2004). Strip-block experiments for process improvement and robustness. *Quality Engineering* **16** 495-500.

Wu, C. F. J. and Chen, Y. (1992). A graph-aided method for planning two-level experiments when certain interactions are important. *Technometrics* **34** 162-175.