Isomorphism Check for $2^n$ Factorial Designs with Randomization Restrictions

Neil A. Spencer $^1$ · Pritam Ranjan $^2$ · Franklin Mendivil $^3$

Published online: 2 October 2019 © Grace Scientific Publishing 2019

Abstract

Factorial designs with randomization restrictions are often used in industrial experiments when a complete randomization of trials is impractical. In the statistics literature, the analysis, construction, and isomorphism of factorial designs have been extensively investigated. Much of the work has been on a case-by-case basis—addressing completely randomized designs, randomized block designs, split-plot designs, etc. separately. In this paper, we take a more unified approach, developing theoretical results and an efficient relabeling strategy to both construct and check the isomorphism of multistage factorial designs with randomization restrictions. The examples presented in this paper particularly focus on split-lot designs.

Keywords Finite projective geometry · Multistage factorial designs · Split-lot designs · $(t-1)$-Spread · Stars

Part of special issue guest edited by Pritam Ranjan and Min Yang—Algorithms, Analysis and Advanced Methodologies in the Design of Experiments.

Pritam Ranjan
pritamr@iimidr.ac.in

Neil A. Spencer
nspencer@andrew.cmu.edu

Franklin Mendivil
franklin.mendivil@acadiau.ca

1 Department of Statistics and Data Science, Carnegie Mellon University, Pittsburgh, USA
2 Operations Management and Quantitative Techniques Area, Indian Institute of Management Indore, Indore, India
3 Department of Mathematics and Statistics, Acadia University, Wolfville, Canada
1 Introduction

Factorial designs are common in a wide variety of applications; however, complete randomization of trials is often impractical. This could be because some factors may be more expensive to change than others, the trials may need to be partitioned into homogeneous batches at each stage, and the experimental units may need to be processed multiple times under different settings, and so on. Popular factorial designs with randomization restrictions include blocked designs, split-plot designs, strip-plot designs, split-lot designs, and combinations thereof. Though both the construction and the analysis of such designs have been active areas of research for decades, most of the literature focuses on a case-by-case basis. The literature started with the exploration of completely randomized designs (CRDs) and randomized block designs (RBDs) by Fisher and Yates. Later, [1, 5] and many others investigated split-plot designs, [28] pioneered strip-plot designs, and [9, 27] presented some fundamental results on split-lot designs. See [15, 18, 29, 42], and [10] for detailed references. In this paper, we focus on multistage factorial designs with randomization restrictions under a unified framework.

For easier understanding and more concise notation, we concentrate on two-level factorial designs, and however, several results and algorithms presented in this paper can easily be extended for \( q \)-levels. Consider a factorial experiment investigating the significance of \( n \) basic factors and all of their interactions. Each \( r \)-factor interaction can be expressed as an \( n \)-dimensional vector composed of exactly \( r \) ones and \( n - r \) zeros, with the \( r \) ones indicating which basic factors are present in the interaction term. We employ the following shorthand. Let the first \( n \) uppercase letters \( A, B, C, \ldots \) denote the \( n \) basic factors in the experiment, and denote any interaction as a string of letters composed of the basic factors that are involved. For instance, \( \{ A, B, AB, C, \ldots, ABCDE \} \) denotes all basic factors and their interactions for a \( 2^5 \) factorial experiment, with \( AE = (1, 0, 0, 0, 1) \) representing a two-factor interaction between the first and the fifth basic factors.

Multistage factorial experiments are common in industrial applications, where all experimental units are processed at each stage, and the observations are taken at the end after the final stage. Some traditional designs like a RBD and a split-plot design can be thought of as multistage factorial designs with only one stage. A non-trivial example is a split-lot design (also referred to as multiway split-unit design), which consists of multiple processing stages with each stage using a split-plot design to partition the experimental units [36]. Popular applications include the laundry experiment for measuring wrinkles in [26] and [28], the fabrication of integrated circuits using silicon wafers in [27], and the plutonium alloy experiment in [6]. See Sect. 2.1 for more details on some of these examples.

Over the decades, a handful of unifying methodologies have been developed for studying different factorial designs with randomization restrictions. For instance, [30, 31] developed the notion of the simple block structure; [39] and [40] used association schemes to extend the simple block structure idea to orthogonal block structures which are more powerful and can characterize a wide
variety of designs (see [3] and [12] for more details). In this paper, we use the unified theory proposed by [33] which is inspired from the randomization group idea of [6]. The comparison of different unified frameworks is outside the purview of this paper, and we do not claim that the unified theory considered here is more general or powerful.

The key idea behind the unified theory of [33] is to realize that the set of all factorial effects (main effects and all possible interactions) of a \(2^n\) factorial design constitute an \((n-1)\)-dimensional finite projective geometry \(\mathcal{P}_n := PG(n-1,2)\) over \(GF(2)\), which is the same as the \(n\)-dimensional vector space \(V(2^n)\) over \(GF(2)\) without the zero element [7]. Furthermore, the randomization restrictions for any given stage of a multistage factorial design can be characterized by a projective subspace of \(\mathcal{P}_n\). Such a subspace is referred to as a *randomization defining contrast subspace* (RDCSS). Note that a \((t-1)\)-dimensional projective subspace (also referred to as a \((t-1)\)-flat) of \(PG(n-1,2)\) is the same as a \(t\)-dimensional vector subspace of \(V(2^n)\) excluding the zero element. In this approach, the overlapping pattern of the flats was exploited to construct useful split-lot designs. Ranjan et al. [34] formalized the construction and analysis of split-lot designs that are derived from the set of disjoint flats of \(\mathcal{P}_n\). In several real-life applications, for instance, in the plutonium alloy experiment of [6], the overlap between two RDCSSs (or flats) cannot be avoided. Subsequently, [35] proposed a new class of split-lot designs which is based on flats with a common overlap. Though the following work, [33–35], has been frequently cited in this article, this is a standalone paper, and we have presented all necessary results for completeness and easy readability.

The construction and ranking of designs become important whenever there are potentially multiple candidate designs meeting the design requirements of a particular factorial experiment. In the design of experiments literature, one can find a plethora of research articles that focus on innovative techniques for constructing good designs, a variety of design ranking criteria (e.g., maximum resolution and minimum aberration), and methods of sorting through the candidates to find different or non-isomorphic designs. However, most of the literature operates on a case-by-case basis (e.g., [5, 11, 22, 23]).

This paper concentrates on the problem of checking the isomorphism of \(2^n\) multistage factorial designs under the unified framework characterized by the RDCSS structure. As per [23], two fractional factorial designs are said to be isomorphic if one can be obtained from the other by relabeling the factors, reordering the runs, and switching the levels of factors. For a multistage factorial experiment, [6] introduces an update to this definition by adding stage-wise restrictions.

In this paper, we present a formal definition of isomorphism using the RDCSS-based unified framework. If performed naively, checking for the isomorphism of designs can involve iterating over all possible relabelings and reorderings, which quickly becomes computationally infeasible for large designs. We present a bitstring representation of \(\mathcal{P}_n\) which helps in developing an efficient search algorithm. A new search strategy is also proposed which exploits the geometric structure of RDCSSs to significantly reduce the search space. We further apply known results from projective geometry (e.g., [24, 38, 41]) to completely classify the isomorphism properties for several RDCSS-based designs (with small runsizes) that are useful from a
practical standpoint. We also provide a new result which establishes that all RDCSS-based designs constructed using the cyclic method of [19] are isomorphic. These results are useful for determining when several isomorphism classes of designs must be considered. Furthermore, all proposed algorithms and important functions have been implemented in R and available on GitHub for easy access.

After reviewing the background theory, important existing results, and motivating examples for the RDCSS-based multistage factorial experiments in Sect. 2, the formal definition of equivalence and isomorphism is presented in Sect. 3. Sections 4 and 5 presents new theories and algorithms for reducing the search space and efficiently iterating through all possible relabelings for an isomorphism check. Section 6.1 reviews the classification of such designs from a practical standpoint, and Sect. 6.2 presents a theoretical result on the cyclic construction of RDCSSs. Finally, the concluding remarks are summarized in Sect. 7.

2 Background Review

A good design is often expected to facilitate efficient analysis. Daniel [14] suggested that for unreplicated factorial experiments, the significance of factorial effects can be assessed using half-normal plots with the restriction that the effects appearing on the same plot must have the same error variance. Moreover, each half-normal plot should contain at least six or seven factorial effects for a meaningful inference. For an RDCSS-based multistage design, the variance of an estimator of a factorial effect is characterized by its presence in different RDCSSs. That is, if a multistage factorial design is defined by \(m\) RDCSSs and all RDCSSs are disjoint, then at most \(m + 1\) half-normal plots are required for the significance assessment of factorial effects. On the other hand, if some of the \(m\) RDCSSs overlap, then more than \(m + 1\) separate half-normal plots would have to be used for the identification of significant factorial effects. As a result, it is desirable to construct RDCSSs that are big enough (with at least six or seven effects per half-normal plot) and disjoint. Section 2.1 presents a quick recap of two popular examples of multistage split-lot designs, and Sect. 2.2 reviews some relevant results from the finite projective geometry literature, [33, 34] and [35], that are helpful for our discussion on isomorphism.

2.1 Examples of Multistage Experiments

In this section, we recap the silicon wafer example [27] and plutonium alloy experiment [6]).

The fabrication of integrated circuits on silicon wafers goes through a sequence of processing steps. Mee and Bates [27] discussed the construction and analysis of several split-lot designs for analyzing this process. Here, we present two designs for a 64-wafer experiment with nine processing stages and six basic factors. At each stage, all experimental units (i.e., 64 wafers) are processed and then passed on to the next stage. The measurements are taken at the end after the final stage, and the randomization of trials at each processing stage
is guided by a set of restrictions defined by carefully chosen factors and factor interactions. The designs are given by $IC_1 = \{\langle A, EF, BCE \rangle, \langle B, AF, CDF \rangle, \langle C, AB, ADE \rangle, \langle D, BC, BEF \rangle, \langle E, CD, ACF \rangle, \langle F, DE, ABD \rangle, \langle BD, BF, ACE \rangle, \langle AC, CE, BDF \rangle, \langle AD, BE, CF \rangle\}$, and $IC_2 = \{\langle A, BD, CF \rangle, \langle B, AF, CE \rangle, \langle C, BF, DE \rangle, \langle D, AC, BE \rangle, \langle E, AB, DF \rangle, \langle F, AE, CD \rangle, \langle AD, BC, EF \rangle, \langle ACE, ADF, BEF \rangle, \langle ABC, ADE, CEF \rangle\}$, where $\langle \ldots \rangle$ denotes the span of the vectors/effects within, e.g., $\langle A, EF, BCE \rangle = \{A, EF, AEF, BCE, ABCE, BCF, ABCF\}$. For both $IC_1$ and $IC_2$, the nine RDCSSs are disjoint and each of size seven. Thus, the significance of all 63 factorial effects (excluding the null) can easily be assessed by nine half-normal plots. The question we address here is whether or not the two designs are isomorphic. Of course, the ranking of designs is a different question and we leave it for future research.

The so-called plutonium alloy experiment (in [6]) took place at Los Alamos National Laboratory (LANL), where the objective was to identify the significant factors and factor combinations involved in the process of manufacturing a plutonium alloy cookie which was to be used further for some classified experiments. This cookie-making process involved five basic factors and had to go through three processing stages: casting and two different types of heat-treatments. Bingham et al. [6] considered a $2^5$ factorial split-lot design with 32 runs. Letting $f^*_i$ denote the flat in $P_5$ that characterizes the randomization of trials for the $i$-th processing stage, then as per [6], the restrictions are: $A, B \in f^*_1, C \in f^*_2$ and $D, E \in f^*_3$. Using a computer search, the authors found it impossible to construct disjoint RDCSSs which could facilitate meaningful half-normal plots. The design suggested at the end was, $PA_1 = \{\langle A, B, CDE \rangle, \langle C, AD, BE \rangle, \langle D, E, ABC \rangle\}$, which required four half-normal plots for the significance assessment of effects in $f^*_1 \cdot f^*_2 \cdot f^*_3$ (excluding the common $ABCDE$) and $P_5 \{f^*_1 \cdot f^*_2 \cdot f^*_3\}$. Later on, [35] recommended an alternative design $PA_2 = \{\langle A, B, DE, ACD \rangle, \langle C, AB, DE, ACD \rangle, \langle D, E, AB, ACD \rangle\}$ for this experiment. As with the previous example, can we check if the two designs $PA_1$ and $PA_2$ are isomorphic?

### 2.2 Projective Geometric Structures

The questions of the existence and construction of a pre-specified number of disjoint flats of $P_n = PG(n-1, 2)$ with given sizes are non-trivial. The combinatorics literature contains some results on the existence and construction of spreads and maximal partial spreads of $P_n$.

A spread of $P_n$ is a set of disjoint flats that includes every element of $P_n$. That is, a spread of $P_n$ is also a cover of $P_n$. A balanced $(t - 1)$-spread $\psi$ of $P_n$ consists of only $(t - 1)$-flats of $P_n$. For simplicity, we do not consider the unbalanced spread case in this paper. A $(t - 1)$-spread $\psi$ of $P_n$ contains $|\psi| = (2^n - 1)/(2^t - 1) = \sum_{i=1}^{n/t} 2^{(i-1)n}$ distinct $(t - 1)$-flats that can be used for constructing RDCSSs for different stages of randomization. For instance, in the silicon wafers example, $IC_1$ and $IC_2$ are two distinct 2-spreads of $P_6$. A necessary and sufficient condition for the existence of a balanced $(t - 1)$-spread of $P_n$ is that $t$ divides $n$ [2]. For instance, in the silicon wafer example, the existence of a $(3 - 1)$-spread of $P_6$ is ensured as 3 divides 6.

© Springer
If \( t \neq n \) (as in the plutonium example), then either a partial \((t - 1)\)-spread or a non-overlapping set of RDCSSs have to be used for design construction. A partial spread is simply a set of disjoint flats of \( P_n \). Lemma 1 discusses the existence of a partial \((t - 1)\)-spread of \( P_n \).

**Lemma 1** [16] Let \( P_n \) be a finite projective space \( PG(n - 1, 2) \), with \( n = kt + r \) for \( 0 < r < t < n \). Then, there exists a partial \((t - 1)\)-spread \( \psi \) of \( P_n \) with \(|\psi| = 2^n \frac{2^h - 1}{2^r - 1} - 2^r + 1 \).

Assuming the overlap between two RDCSSs cannot be avoided, [35] proposed a new geometric structure called a star\(^1\) which requires all constituent flats to have a common overlap.

**Definition 1** A balanced star, denoted by \( \Omega = St(n, \mu, t, t_0) \), of \( P_n \) is a set of \( \mu \) rays ((\(t - 1\))-flats) and a nucleus (one \((t_0 - 1)\)-flat) in \( P_n \), such that the intersection of any two of the \( \mu \) rays is the nucleus (so, \( 0 \leq t_0 < t < n \)).

A star \( \Omega \) is said to cover \( P_n \) if the combined set of elements of all rays in \( \Omega \) is equal to \( P_n \). For the purposes of this paper, all stars are balanced and covering. Thus, a star \( \Omega = St(n, \mu, t, t_0) \) provides \( \mu = \frac{(2^n - t_0) - 1}{2^t - 1} \) overlapping RDCSSs of size \( 2^r - 1 \) each. For instance, in the plutonium alloy experiment, \( PA_2 \) represents a \( St(5, 3, 4, 3) \). Lemma 2 is taken from Ranjan et al. [35, Lemma 3] which establishes the relationship between a spread and a star.

**Lemma 2** The existence of a balanced covering star \( \Omega = St(n, \mu, t, t_0) \) of \( P_n = PG(n - 1, 2) \) is equivalent to the existence of an \((h - 1)\)-spread \( \psi \) of \( P_u \), where \( u = n - t_0 \), and \( h = t - t_0 \).

The proof of Lemma 2 easily follows from the following construction steps. Let \( \{f_1, \ldots, f_\mu\} \) be the constituents of an \((h - 1)\)-spread \( \psi \) of \( P_u \). Then, there exists a \((t_0 - 1)\)-flat \( \pi \) (referred to as the nucleus) in \( P_n \setminus P_u \) such that \( f_1^* = \langle f_1, \pi \rangle \) and \( \{f_1^*, \ldots, f_\mu^*\} \) form a covering star \( \Omega \) of \( P_n \). For convenience, we denote such stars as \( \Omega = \psi \times \pi \). In the plutonium alloy experiment, \( PA_2 \) satisfies this structure with \( \pi = \langle AB, DE, ACD \rangle \) and \( \psi = \{\{A\}, \{C\}, \{AC\}\} \). Here, \( \psi \) corresponds to a 0-spread of \( P_2 = \langle A, C \rangle \).

One of the most popular methods of constructing a balanced spread of a finite projective space in \( GF(2) \) is the cyclic approach of [19]. Suppose \( h \) and \( u \) are positive integers such that \( h \) divides \( u \), and we wish to construct an \((h - 1)\)-spread \( \psi = \{f_1, \ldots, f_\mu\} \) of \( P_u \), where \( \mu = \frac{(2^u - 1)}{(2^h - 1)} \). The cyclic method for constructing \( \psi \) starts by writing the \( 2^u - 1 \) nonzero elements of \( GF(2^u) \) in cycles of length \( \mu \) (Table 1). The nonzero elements of \( GF(2^u) \) are written as \( \{\omega^0, \omega^1, \ldots, \omega^{2^{u-2}}\} \), where

\(^1\) Stars were recently reinvented in a collection of works [25, 37] where they are referred to as book spreads.
$\omega$ is a primitive element, and $\omega^i = \alpha \omega^0 + \alpha_1 \omega^1 + \cdots + \alpha_{u-1} \omega^{u-1}$, for $0 \leq i \leq 2^u - 2$, correspond to the vector representation $(\alpha_0, \ldots, \alpha_{u-1})$ of elements in $P_u$. 

[19] showed that the $f_i$’s are $(h - 1)$-flats and $\psi = \{f_1, \ldots, f_9\}$ partitions the set of all nonzero elements of $GF(2^u)$, i.e., $\psi$ is an $(h - 1)$-spread of $P_u$. For a quick reference, Table 2 shows the 2-spread $\psi$ of $P_6$ generated using the primitive polynomial $\omega^6 + \omega + 1$ and primitive element $\omega$. Here, $\mu = (2^6 - 1)/(2^3 - 1) = 9$ disjoint 2-flats $\{f_1, \ldots, f_9\}$ can be used to construct different RDCSSs as required.

### Table 2

| $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $f_6$ | $f_7$ | $f_8$ | $f_9$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| F     | E     | D     | C     | B     | A     | EF    | DE    | CD    |
| BC    | AB    | AEF   | DF    | CE    | BD    | AC    | BEF   | ADE   |
| CDEF  | BCD   | ABCD  | ABCEF | ADBF  | ACF   | BF    | AE    | DEF   |
| CDE   | BCD   | ABC   | ABCEF | ADBF  | ACF   | BF    | AE    | DEF   |
| 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    |

### 3 Isomorphism of RDCSS-Based Designs

In this paper, we only consider the spread- and star-based designs, and for simplicity, we also assume that all RDCSSs of a design are of the same size (so the resulting structure is balanced). Let $d_1$ and $d_2$ be two $2^n$ multistage factorial designs with $\mu$ stages of randomization, and let the respective RDCSSs be represented by $d_1 = \{f_1, \ldots, f_\mu\}$ and $d_2 = \{g_1, \ldots, g_\mu\}$. In spirit of [6], the two designs are said to be isomorphic (denoted by $d_1 \cong d_2$) if one can be obtained from the other by applying some sort of relabeling of factors and factor levels and/or reordering of effects within the RDCSSs. We formalize this definition by first bundling up the reordering and rearrangement-type operations together in one concept called “equivalence” and then addressing the relabeling step.

**Definition 2** Two $2^n$ RDCSS-based factorial designs $d_1 = \{f_1, \ldots, f_\mu\}$ and $d_2 = \{g_1, \ldots, g_\mu\}$ are said to be equivalent (denoted by $d_1 \equiv d_2$) if and only if,
for every $f_i \in d_1$, there is a unique $g_j \in d_2$ such that $\{f_i\} = \{g_j\}$ (set equality), for $1 \leq i, j \leq \mu$.

This notion of equivalence will not only take care of rearrangement of factor combinations within a given RDCSS, but also account for reordering of the RDCSSs themselves. Let $\mathcal{E}(d_1)$ be the set of all such $2^n$ designs in $\mu$-stages that are equivalent to $d_1$ (i.e., $\mathcal{E}(d_1)$ denotes the equivalence class of $d_1$). If $|f_i| = 2^t - 1$ for each $f_i \in d_1$, then

$$|\mathcal{E}(d_1)| = \mu \cdot (2^t - 1)!^\mu. \quad (1)$$

Assuming the existence of a $(t - 1)$-spread or a covering star $St(n, \mu, t, t_0)$ of $\mathcal{P}_n$ involved in the construction of the $2^n$ design, the maximum value of $\mu$ is $(2^n - 1)/(2^t - 1)$ or $(2^n - t_0 - 1)/(2^t - t_0 - 1)$, respectively. When $n$ is large, checking the equivalence of two designs by naively iterating through the entire equivalence class of one of them is too computationally intensive, e.g., in the plutonium alloy experiment, $|\mathcal{E}(PA_2)| = 6(15!)^3 \approx 1.37 \times 10^{37}$. However, the computational burden can be reduced through a combination of sorting and the following bitstring representation scheme.

Each element of $\mathcal{P}_n$ can be represented as a unique binary string of $2^n - 1$ bits with exactly one nonzero entry. For instance, following the Yates Order [8] of $\mathcal{P}_3$, the bitstring representations of the elements of $\mathcal{P}_3 = \{A, B, AB, C, AC, BC, ABC\}$ are $A \rightarrow 10000000, B \rightarrow 01000000, \ldots, ABC \rightarrow 00000001$. In this representation, the contents of any RDCSS $f$ can now be uniquely identified by the sum of the bitstring representations of its elements. For instance, $f = \{AB, AC, BC\}$ is now uniquely identified by the representation 0010110. After converting to this representation, checking the equivalence of two RDCSSs amounts to checking the equality of two bitstrings. Furthermore, checking the equivalence of two RDCSS-based designs becomes equivalent to checking the equality of two sets of bitstrings, which is straightforward if one sorts them first. Note that this new bitstring representation is typically more advantageous for spreads with larger values of $t$. For smaller $t$ (say, 2 or 3), simply sorting the elements in each RDCSS is sufficient.

We use a collineation of $\mathcal{P}_n$ to express the relabelings of factors and factor combinations [13]. A collineation of $\mathcal{P}_n$ is a mapping of the points from $\mathcal{P}_n$ to $\mathcal{P}_n$ such that $(t - 1)$-flats gets mapped to $(t - 1)$-flats for all $1 \leq t \leq n$. A collineation of $\mathcal{P}_n$ can be characterized by a full rank $n \times n$ matrix $C$ over $GF(2)$, referred to as the collineation matrix [4], where the $j$-th column of $C$ is the image of the $j$-th basic factor (say) $F_j$ (i.e., the factorial effect in $\mathcal{P}_n$ that $F_j$ gets mapped to). For instance, the $3 \times 3$ collineation matrix

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

relabels the basic factors as $A \rightarrow A, B \rightarrow B$ and $C \rightarrow BC$. See Appendix C for easy implementation in $\mathbb{R}$. We interchangeably use the terms “collineation” and
“collineation matrix” to refer to the same linear mapping. Let \( C_n \) be the set of all collineations of \( P_n = PG(n - 1, 2) \), then the size of \( C_n \) is given by

\[
|C_n| = \prod_{j=1}^{n} (2^n - 2^{j-1}).
\]

(2)

The proof of (2) follows by simply counting the total number of linearly independent images of \( F_j \). That is, given that the images of \( F_1, \ldots, F_j \) have already been selected, the total number of possible images for \( F_{j+1} \) is \( 2^n - 2^j \).

Suppose we wish to construct a collineation \( C \) which defines the mapping between \( \{x_1, \ldots, x_n\} \) and \( \{y_1, \ldots, y_n\} \), i.e., \( C(x_i) = y_i \text{, for } i = 1, \ldots, n \). One intuitive method of constructing such a collineation matrix is to solve a system of \( n \) equations with \( n \) unknowns over \( GF(2) \). As discussed in Algorithm 1 of Sect. 4, this equation solving approach could be computationally expensive for checking isomorphism. Thus, we propose a two-step alternative approach. First construct a collineation matrix \( C_{x,B} \) that canonicalize \( \{x_1, \ldots, x_n\} \), i.e., \( x_1 \rightarrow A, x_2 \rightarrow B \), and so on, and then construct \( C_{B,y} \) that maps the canonical basis elements to \( y_j \)'s, i.e., \( A \rightarrow y_1, B \rightarrow y_2 \), etc. As a result, the desired collineation matrix is \( C = C_{B,y} \cdot C_{x,B} \).

For instance, for the silicon wafers example, let \( x_1 = A, x_2 = EF, x_3 = BCE, x_4 = B, x_5 = AF, x_6 = CDF \) be the effects from RDCSSs in \( IC_1 \), and \( y_1 = A, y_2 = BD, y_3 = CF, y_4 = B, y_5 = AF, y_6 = CE \) be chosen from RDCSSs in \( IC_2 \). Then, \( C_{x,B} \) can be constructed by writing \( x_i \)'s as column vectors and then inverting the matrix, i.e.,

\[
C_{x,B} = \begin{bmatrix}
1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 1
\end{bmatrix}^{-1} = \begin{bmatrix}
1 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 & 0 & 0
\end{bmatrix}.
\]

From an implementation standpoint, the inversion of a matrix \( M \) over \( GF(2) \) can easily be done in R [32] using the code: \texttt{solve(M) %% 2} As highlighted in Algorithm 1, this inversion is required much less often as compared to the naive method (via solving a system of \( n \) equations in \( n \) unknowns). Next, \( C_{B,y} \) is constructed by simply writing \( y_j \)'s as column vectors. As a result, the desired collineation matrix that defines the mapping from \( x_i \)'s to \( y_j \)'s is given by

\[
C = C_{B,y} \cdot C_{x,B} = \begin{bmatrix}
1 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 0
\end{bmatrix} \cdot \begin{bmatrix}
1 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 & 0 & 0
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1
\end{bmatrix}.
\]
Applying this collineation on RDCSSs of $IC_1$ gives $C(IC_1) = \{\langle A, BD, EF \rangle, \langle B, AF, CE \rangle, \langle CD, AB, ABE \rangle, \langle DEF, BCD, D \rangle, \langle BDF, CEF, ACDF \rangle, \langle ADEF, D, CDF \rangle, \langle ADEF, DF, CDF \rangle, \langle BDEF, BF, ABCF \rangle, \langle ADEF, DF, CDF \rangle\}.$

We now combine the notion of equivalence and collineation to formally define the isomorphism of multistage factorial designs characterized by the RDCSSs.

**Definition 3** Two $2^n$ RDCSS-based factorial designs $d_1 = \{f_1, \ldots, f_n\}$ and $d_2 = \{g_1, \ldots, g_n\}$ are said to be isomorphic (denoted by $d_1 \cong d_2$) if and only if there exists a collineation $C$ over $P_n$ such that $C(d_1) \equiv d_2$. In this case, we say that $C$ is an isomorphism establishing collineation (IEC) from $d_1$ to $d_2$.

For any pair of RDCSS-based designs in $P_n$, if there exists one IEC, then there are typically many more-obtained by multiplying with the collineation matrices in $C_n$. However, determining if one IEC exists is the most difficult part. For instance, in the silicon wafers example, $C(IC_1)$ (discussed above) is not equivalent to $IC_2$. This is apparent when comparing their bitstring representations. Let $a_{(1)}(1), \ldots, a_{(9)} \in \{0, 1\}^{63}$ denote the bitstring representations of the RDCSSs of $C(IC_1)$ (sorted according to their first nonzero entries) and let $b_{(1)}(1), \ldots, b_{(9)} \in \{0, 1\}^{63}$ denote the analogous sorted representation for $IC_2$. Note that each bitstring, here, would be of length 63 with only seven 1’s and the rest zeros. Determining these bitstring representations reveals that

$$a_{(1)} = 100000000110000000000000000000000000000000000000000000000,$$
$$b_{(1)} = 1000000001100000000000000000000000000000000000000000000000.$$

Clearly, $a_{(1)} \neq b_{(1)}$, which is sufficient to conclude $C(IC_1) \not\cong IC_2$. Therefore, $C$ is not an IEC from $IC_1$ to $IC_2$. However, all viable collineations must be checked before we could conclude that $IC_1$ and $IC_2$ are non-isomorphic. If we assume that $IC_1$ and $IC_2$ are isomorphic, how do we find an IEC?

Note that the results discussed thus far in this section do not assume any specific overlapping pattern of the RDCSSs. Next, we discuss the IEC search algorithms separately for spread-based designs (with disjoint RDCSSs), and the star-based designs, where all RDCSSs have a common overlap.

## 4 Search Algorithm for Spread-Based Designs

When undertaken naively, determining if two RDCSS-based designs on $P_n$ are isomorphic involves exhaustively searching over the entire space $C_n$— as enumerated by (2)— to check if any are IECs. Depending on the value of $n$, this full search space can be prohibitively large, meaning an exhaustive search would be computationally intractable. We propose a strategy that exploits the structure of IECs to reduce the size of this search space. Here, we motivate and describe our strategy as it applies to RDCSS-based designs with no overlap. In Sect. 5, we extend this approach to RDCSSs that share a common overlap (i.e., star-based designs).
From the definition of isomorphism (Definition 3), a collineation \( C \) can be considered as an IEC if and only if, for all \( i \in \{1, 2, \ldots, \mu\} \), the RDCSS \( f_j \) in \( d_i \) is relabeled by \( C \) to \( g_j \) in \( d_1 \) for some \( j \in \{1, 2, \ldots, \mu\} \). We refer to this as the full \textit{RDCSS mapping property} (FRMP) of an IEC. The FRMP characterizes which collineations to consider when searching for an IEC. For example, if \( C \) is an IEC from \( IC_1 \) to \( IC_2 \) for the silicon wafers example in Sect. 2.1, then we know that \( C(A), C(EF), \) and \( C(BCE) \) must all fall within the same RDCSS of \( IC_2 \); we need not consider any \( C \) for which \( C(A) \) and \( C(EF) \) belong to different RDCSSs, such as \( C(A) = B \) and \( C(EF) = C \). Furthermore, because \( A \) and \( B \) are in different RDCSSs in \( IC_1 \), we know that \( A \) and \( B \) must be mapped to different RDCSSs in \( IC_2 \); we can rule out any other collineation, such as those for which \( A \rightarrow BD \) and \( B \rightarrow A \), when searching for an IEC.

Unfortunately, we are not aware of any constructive approaches for isolating collineations that completely satisfy the FRMP. For this reason, we introduce a relaxed version of the FRMP—called the \textit{partial RDCSS mapping property} (PRMP)—that does admit a constructive approach. A collineation \( C \) satisfies the PRMP for points \( x_1, \ldots, x_n \in \mathcal{P}_n \) if it meets the following requirement: For all \( i, j : 1 \leq i, j \leq n \), there exist unique \( k, l \) \((1 \leq k \neq l \leq \mu)\) such that, \( x_i, x_j \in f_k \) if and only if \( C(x_i), C(x_j) \in g_l \). That is, a collineation \( C \) satisfies the PRMP for \( x_1, \ldots, x_n \) if \( x_i, x_j \) being co-members of an RDCSS in \( d_1 \) occurs if and only if \( C(x_i) \) and \( C(x_j) \) are co-members of an RDCSS in \( d_2 \).

Our reasons for defining the PRMP in this manner are twofold. Firstly, the constraints imposed by the PRMP are a strict subset of the constraints imposed by the FRMP. As a result, any collineation satisfying the FRMP (i.e., IECs) will also satisfy the PRMP for any given \( x_1, \ldots, x_n \); our reduction in the search space does not ignore any possible IECs. Secondly, because the constraints in the PRMP involve just \( n \) points, the collineations satisfying the PRMP are now straightforward to construct.

Recall that for linearly independent \( x_1, \ldots, x_n \), any collineation \( C \) is characterized by the images \( y_1 = C(x_1), \ldots, y_n = C(x_n) \) with the corresponding collineation matrix being easily determined using the algorithm provided in Sect. 3. Thus, the entire search space for IECs given by a PRMP can be constructed by iterating through the possible options for \( y_1, \ldots, y_n \) in \( d_2 \) that match the RDCSS co-membership structure of \( x_1, \ldots, x_n \) in \( d_1 \). The remainder of this section describes an approach that efficiently iterates through the collineation matrices in this class to look for IECs. We can then check the equivalence of \( C(d_1) \) and \( d_2 \)—via the bitstring representation—for every \( C \) in this class.

We now develop some theory to provide guidance on how to choose the basis set \( x_1, \ldots, x_n \in \mathcal{P}_n \). Proposition 1 presents an upper bound on the number of collineations satisfying the PRMP for a basis \( \{x_1, \ldots, x_n\} \). Appendix A provides a proof of this result, as well as a comment explaining why it is only an upper bound.

**Proposition 1** Let \( d_1 = \{f_1, \ldots, f_\mu\} \) and \( d_2 = \{g_1, \ldots, g_\mu\} \) be two \( 2^n \) RDCSS-based designs obtained from balanced \((t-1)\)-spreads of \( \mathcal{P}_n \). Let \( \{x_1, \ldots, x_n\} \) be a basis of \( \mathcal{P}_n \), \( m_i = |f_i \cap \{x_1, \ldots, x_n\}| \) for \( i = 1, \ldots, \mu \), and \( \ell \) be the number of nonzero \( m_i \)’s. Then, the number of collineations \( C \) from \( d_1 \) to \( d_2 \) which satisfy the PRMP for \( x_1, \ldots, x_n \) is bounded above by

\[ C \text{ Springer} \]
Recall that a balanced \((t-1)\)-spread of \(P_n\) exists if and only if \(t\) divides \(n\). Therefore, the upper bound (3) from Proposition 1 is minimized for \(\ell = n/t\), which further implies either \(m_i = t\) or zero. Proposition 2 guarantees the existence of such a basis for any set of RDCSSs obtained from a balanced \((t-1)\)-spread of \(P_n\) (see Appendix A for the proof).

**Proposition 2** For any balanced \((t-1)\)-spread-based multistage design \(d_1 = \{f_1, \ldots, f_\mu\}\) in \(P_n\), there exists \(\ell_0 = n/t\) distinct RDCSSs \(f_{u_1}, \ldots, f_{u_{\ell_0}}\) from \(d_1\), with \(1 \leq u_1, \ldots, u_{\ell_0} \leq \mu\), such that \(\bigcup_{i=1}^{\ell_0} f_{u_i} = P_n\).

Subsequently, there exist a set \(x_1, \ldots, x_n\) such that the maximum number of collinearations that satisfy PRMP can be reduced to

\[
\frac{\mu!}{(\mu - \ell)!} \prod_{i=1}^{\ell} \left( \prod_{j=1}^{m_i} \left( 2^{j-1} - 2^{i-1} \right) \right).
\]

As compared to the naive approach (enumerated in (2)), the proposed approach corresponds to a reduction in the search space by 7 orders of magnitude for checking the isomorphism of 2-spreads of \(P_6\). For 1-spreads and 4-spreads of \(P_{10}\), the search space is reduced by 13 orders and 12 orders of magnitude, respectively. For larger \(n\), the improvements are even greater.

For checking the isomorphism of \(d_1 = \{f_1, \ldots, f_\mu\}\) and \(d_2 = \{g_1, \ldots, g_\mu\}\) in \(P_n\), we follow a systematic approach to search for the IEC by iterating through the candidate collineations. First, choose \(\ell_0 = n/t\) out of \(\mu\) RDCSSs in \(d_1\) that can generate a basis \(\{x_1, \ldots, x_n\}\) for \(P_n\) such that each of the \(\ell_0\) RDCSSs contributed \(t\) points to the basis (as in Proposition 2). Then, construct a collineation matrix \(C_{x,B}\) to transform the \(x\)'s to a canonical basis with basic factors, \(\{A, B, \ldots\}\) (as demonstrated in Sect. 3). Note that the isomorphism of \(C_{x,B}(d_1)\) and \(d_2\) implies the isomorphism of \(d_1\) and \(d_2\). Now, iterate through all possible sets of \(\ell_0\) RDCSSs from \(d_2\) to construct the basis set \(\{y_1, \ldots, y_n\}\) in \(d_2\) and the corresponding collineation matrix \(C_{B,Y}\) that maps the canonical basis to \(y\)'s. If \(C_{B,Y}(C_{x,B}(d_1))\) is equivalent to \(d_2\), then we have found the IEC. The step-by-step algorithm is summarized in Algorithm 1. Note that two subscripts for the \(x\)'s and \(y\)'s are introduced to keep track of their RDCSS membership in \(d_1\) and \(d_2\), respectively. Let \(F_1, \ldots, F_\mu\) be an alternative notation for the basic factors, with the assumption that \(F_1 := A, F_2 := B, \) and so on.
A few quick remarks are worth noting. From an implementation standpoint, the stages in Step 4 are nested, making it straightforward to parallelize at any stage of the hierarchy. The matrix inversion is required only for Step 3, and the factorization $C = C_{B} \cdot C_{x,B}$ avoids the need to repeatedly solve that systems of linear equations associated with $C(x_{i,j}) = y_{\sigma_{k}(i),j}$. Moreover, Algorithm 1 iterates through all possible collineations satisfying the PRMP for a particular chosen set $\{x_{1}, \ldots, x_{n}\}$. As discussed above, the set of collineations satisfying a PRMP is a superset of the collineations satisfying the FRMP (which contains all possible IECs). Therefore, by exhausting the set of collineations for one chosen set $\{x_{1}, \ldots, x_{n}\}$, we are guaranteed to have already visited all possible IECs.

In terms of algorithmic complexity, the speed-up of moving from the naive search method to Algorithm 1 is proportional to the reduction in the search space size from (2) to (4). The complexity of the equivalence check algorithm (as described in Sect. 3) is $O(n^{2}2^{n})$. We have also implemented this algorithm in R (see Appendix C for illustration).

Example: To demonstrate Algorithm 1, we will illustrate a run to check the isomorphism of $IC_{1}$ and $IC_{2}$ (in the silicon wafers experiment) below. Recall that $IC_{1} = \{f_{1} = \langle A, EF, BCE \rangle, f_{2} = \langle B, AF, CDF \rangle, f_{3} = \langle C, AB, ADE \rangle, f_{4} = \langle D, BC, BEF \rangle, f_{5} = \langle E, CD, ACF \rangle, f_{6} = \langle F, DE, ABD \rangle, f_{7} = \langle BD, BF, ACE \rangle, f_{8} = \langle AC, CE, BDF \rangle, f_{9} = \langle AD, BE, CF \rangle\}$, and $IC_{2} = \{g_{1} = \langle A, BD, CF \rangle, g_{2} = \langle B, AF, CE \rangle, g_{3} = \langle C, BF, DE \rangle, g_{4} = \langle D, AC, BE \rangle, g_{5} = \langle E, AB, DF \rangle, g_{6} = \langle F, AE, CD \rangle, g_{7} = \langle AD, BC, EF \rangle, g_{8} = \langle ACE, ADF, BEF \rangle, g_{9} = \langle ABC, ADE, CEF \rangle\}$. The following represents an iteration of Step 4 for which an ICE from $IC_{1}$ to $IC_{2}$ is found.

---

**Algorithm 1** Isomorphism check between two $(t-1)$-spread based designs in $P_{n}$.

1. Choose $t_{0} = n/t$ out of $\mu$ RDCSSs from $d_{1} = \{f_{1}, \ldots, f_{\mu}\}$, that satisfy Proposition 2. Let these be $\{f_{u_{1}}, \ldots, f_{u_{t_{0}}}\}$ for $1 \leq u_{1}, \ldots, u_{t_{0}} \leq \mu$.
2. For $i = 1, \ldots, t_{0}$, specify $\{x_{i,1}, \ldots, x_{i,t}\} \in f_{u_{i}}$ such that $f_{u_{i}} = \langle x_{i,1}, \ldots, x_{i,t}\rangle$.
3. Construct the collineation matrix $C_{B} \cdot C_{x,B}$ which maps the $x$’s in Step 2 to the canonical basis $\{A, B, \ldots\}$ such that each $x_{i,j}$ is mapped to $F_{(i-1)+j}$.
4. (a) Choose $t_{0}$ out of $\mu$ RDCSSs from $d_{2} = \{g_{1}, \ldots, g_{\mu}\}$ for mapping $f_{u_{i}}$’s (note that ordering is important). Let that be $\{g_{v_{1}}, \ldots, g_{v_{t_{0}}}\}$ for $1 \leq v_{1}, \ldots, v_{t_{0}} \leq \mu$. If $|\{g_{v_{1}} \cup \ldots \cup g_{v_{t_{0}}}\}| < 2^{n} - 1$, proceed to the next choice for $\{g_{v_{1}}, \ldots, g_{v_{t_{0}}}\}$.
   (b) For $i = 1, \ldots, t_{0}$, choose a $\{y_{i,1}, \ldots, y_{i,t}\} \in g_{v_{i}}$ such that $g_{v_{i}} = \langle y_{i,1}, \ldots, y_{i,t}\rangle$.
   (c) Choose one of the $(t_{0})!$ permutations of the elements $1, \ldots, t_{0}$, say $\sigma_{k}$, for $k = 1, \ldots, (t_{0})!$.
   (d) Construct $C_{B,y}$ which maps the canonical basis elements to $y$’s such that $F_{(i-1)+j}$ is mapped to $y_{\sigma_{k}(i), j}$.
   (e) If $C_{B,y}(C_{x,B}(d_{1}))$ is equivalent to $d_{2}$ (as per the bitstring method in Section 3), then $d_{1} \cong d_{2}$, and report $C = C_{B,y} \cdot C_{x,B}$ as an IEC and exit; otherwise, continue.
   (f) Go to Step 4(c) and choose another ordering $\sigma_{k}$ if possible, otherwise, continue.
   (g) Go to Step 4(b) and choose another basis if possible, otherwise, continue.
   (h) Go to Step 4(a) and choose another set of RDCSSs if possible, otherwise report that $d_{1}$ and $d_{2}$ are non-isomorphic.
1. Here, $\ell_0 = 2$ with $f_{u_1} = \langle A, EF, BCE \rangle$ and $f_{u_2} = \langle B, AF, CDF \rangle$.
2. Let $x_{1,1} = A, x_{1,2} = EF, x_{1,3} = BCE$, and $x_{2,1} = B, x_{2,2} = AF, x_{2,3} = CDF$.
3. The collineation matrix $\mathcal{C}_{x,B}$ that canonicalize these $x$’s has been presented in Sect. 3.
4. Here, we calculate the total number of options at each stage, and then demonstrate their values when the first IEC found.

(a) There are $9!/ (9-2)! = 72$ choices for $\{v_1, v_2\} \subset \{1, \ldots, 9\}$. Our first choice of $\{v_1, v_2\} = \{1, 2\}$ leads to an IEC.

(b) There are $(2^3 - 1) \cdot (2^3 - 2) \cdot (2^3 - 2^2) = 7 \cdot 6 \cdot 4 = 168$ choices for linearly independent $y_{1,1}, y_{1,2}, y_{1,3} \in g_1$, and 168 choices for linearly independent $y_{2,1}, y_{2,2}, y_{2,3} \in g_2$. Out of $(168)^2$ iterations, we found an IEC in the 359th step of our search. The $y$’s that led to the first IEC are $y_{1,1} = A$, $y_{1,2} = BD$, $y_{1,3} = BCDF$, $y_{2,1} = B$, $y_{2,2} = ABCEF$, $y_{2,3} = ABF$.

(c) There are two options for the permutation: $\sigma_1 = (1, 2)$ or $\sigma_2 = (2, 1)$. Our first IEC was found using $\sigma_1$.

(d) Here,

\[
    C_{B,y} = \begin{bmatrix}
    1 & 0 & 0 & 0 & 1 & 1 \\
    0 & 1 & 1 & 1 & 1 & 1 \\
    0 & 0 & 1 & 0 & 1 & 0 \\
    0 & 1 & 1 & 0 & 0 & 0 \\
    0 & 0 & 0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 & 1 & 1 
\end{bmatrix}.
\]

(e) A run of the equivalence check algorithm verifies that $C_{B,y}(C_{x,B}(d_1))$ is equivalent to $d_2$ (details omitted for space). Thus, $d_1 \cong d_2$, and

\[
    C = C_{B,y} \cdot C_{x,B} = \begin{bmatrix}
    1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\
    0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
    0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
    0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
    0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
    0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
    0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 
\end{bmatrix} = \begin{bmatrix}
    1 & 0 & 0 & 1 & 0 & 0 \\
    0 & 1 & 0 & 0 & 0 & 1 \\
    0 & 0 & 0 & 1 & 1 & 1 \\
    0 & 0 & 0 & 0 & 1 & 0 \\
    0 & 0 & 1 & 0 & 1 & 1 \\
    0 & 0 & 0 & 1 & 1 & 1 
\end{bmatrix}
\]

is an IEC. We can exit the algorithm.

5 Search Algorithm for Star-Based Designs

Recall that a balanced star-based design refers to a multistage factorial design with equal sized RDCSSs that share a common overlap. As per Lemma 1, a star $\Omega = St(n, \mu, t, t_0)$ can be expressed as $\Omega = \psi \times \pi$, where $\psi$ is a $((t - t_0) - 1)$-spread
of $P_{n-t_0}$, and $\pi$ is $(t_0 - 1)$-dimensional subspace in $P_n \setminus P_{n-t_0}$. As earlier, let $u = n - t_0$ and $h = t - t_0$. Thus, the isomorphism check between two star-based designs $d_1$ and $d_2$ (with $(t_0 - 1)$-dimensional nuclei) can be reduced to checking isomorphism between two $(h - 1)$-spreads of $P_{n-t_0}$ by iterating through the elements of $C_{n-t_0}$ instead of $C_n$. Even for small nuclei (e.g., $t_0 = 1$ or 2), this corresponds to a large reduction in the search space.

Algorithm 2 summarizes the steps of how Algorithm 1 can be used to search for an IEC between $d_1$ and $d_2$ based on $\psi_1 \times \pi_1$ and $\psi_2 \times \pi_2$, respectively.

Algorithm 2: Isomorphism check between two $St(n, t, t_0)$-based designs $d_1$ and $d_2$, which correspond to stars $\Omega_1 = \psi_1 \times \pi_1$ and $\Omega_2 = \psi_2 \times \pi_2$, respectively.

1. Determine two bases $\{p_1, \ldots, p_{t_0}\}$ and $\{p_2, \ldots, p_{t_0+1}, \ldots, p_n\}$ of the nuclei $\pi_1$ and $\pi_2$, respectively.
2. Construct a collineation matrix $C_{\pi_1, B}$ mapping $\{p_1, \ldots, p_{t_0}\}$ to the $t_0$ trailing basic factors $F_{n-t_0+1}, \ldots, F_n$. The pre-images of $F_1, \ldots, F_{n-t_0}$ can be chosen as an arbitrary linearly independent set from $P_n \setminus \pi_1$.
3. Similarly, construct a collineation matrix $C_{\pi_2, B}$ mapping $\{p_{t_0+1}, \ldots, p_n\}$ to the $t_0$ trailing basic factors $F_{t_0+1}, \ldots, F_n$.
4. Extract designs $d_1^*$ and $d_2^*$ on $P_{n-t_0}$ corresponding to the spreads $C_{\psi_1, B}(\psi_1)$ and $C_{\psi_2, B}(\psi_2)$.
5. Run Algorithm 1 on $d_1^*$ and $d_2^*$. If we come across a $C^*$ that is an IEC, then an IEC for $d_1$ and $d_2$ is given by $C = C_{\psi_1, B} \cdot C^* \cdot C_{\psi_2, B}$. Otherwise, $d_1$ and $d_2$ are non-isomorphic.

Note that in Algorithm 2, Steps 1-4 relabel the $(h - 1)$-spreads $\psi_1$ and $\psi_2$ of $P_{n-t_0}$ such that their points are within the span of the first $n - t_0$ canonical factors $F_1, \ldots, F_{n-t_0}$ (or $A, B, \ldots$). This relabeling allows Algorithm 1 to be run using a search space of collineations over $P_{n-t_0}$ rather than over $P_n$, as all effects involving the last $t_0$ basic factors are discarded along with the nuclei. This reduction in the search space decreases both the number of collineations that need to be considered as well as the dimension of the structures involved.

Similar to Algorithm 1, the computational complexity of Algorithm 2 improves over the naive method proportionally to the reduction in the search space from (2) evaluated at $n$ to (4) evaluated at $n - t_0$. Furthermore, the complexity of each equivalence check decreases from $O(n^2 2^n)$ to $O(n^2 2^{n-t_0})$ because the objects being compared are smaller after reducing to spreads. Appendix C illustrates the usage of our $R$ implementation.

In the plutonium alloy example, $PA_1 = \{\langle A, B, CDE \rangle, \langle C, AD, BE \rangle, \langle D, E, ABC \rangle\}$ is derived from a balanced covering star $St(5, 5, 3, 1)$ of $P_5$, whereas, $PA_2 = \{\langle A, B, DE, ACD \rangle, \langle C, AB, DE, ACD \rangle, \langle D, E, AB, ACD \rangle\}$ represents a covering star $St(5, 3, 4, 3)$ of $P_5$. Since the sizes of the two nuclei are different, the two designs are trivially non-isomorphic.

6 Results on Special Cases

Thus far we have developed theoretical results and algorithms for checking whether or not two RDCSS-based designs are isomorphic. However, we often want to compare all possible admissible designs and find the optimal one as per some ranking
criterion. For this purpose, we need to construct all possible different (or non-isomorphic) admissible designs. This is a much bigger challenge for RDCSS-based designs because formal construction methods for all possible spreads or stars are not known yet. In this section, we discuss some results from the combinatorics literature on the complete classification of spreads that can be used for small RDCSS-based designs in $\mathcal{P}_n$.

### 6.1 Complete Classification

A balanced $(t-1)$-spread of $\mathcal{P}_n$ is referred to as *trivial* for $t = 1$ and $t = n$. Of course, we are more interested in the non-trivial cases. As expected, complete classification of non-isomorphic balanced spreads is known for only small $n$.

- **A1.** $[n \leq 3]$: All balanced spreads are trivial, meaning for every given $n \leq 3$ all $(t-1)$-spreads of $\mathcal{P}_n$ are isomorphic.
- **A2.** $[n = 4]$: [2] ensures the existence of non-trivial spreads for $t = 2$. [38] show that all 1-spreads of $\mathcal{P}_4$ are isomorphic.
- **A3.** $[n = 5]$: All covering spreads are trivial, as $n = 5$ is prime. However, a complete classification of partial spreads of $\mathcal{P}_5$ is available in [17], indicating that there are 4 isomorphism classes of maximal 1-spreads of $\mathcal{P}_5$ consisting of 9 RDCSSs.
- **A4.** $[n = 6]$: There exist non-trivial balanced spreads for $t = 2$ and $t = 3$. [41] showed that all 2-spreads ($t = 3$) of $\mathcal{P}_6$ are isomorphic, and [24] used exhaustive search to show that there exist 131044 mutually non-isomorphic 1-spreads ($t = 2$) of $\mathcal{P}_6$.
- **A5.** $[n = 7]$: All covering spreads are trivial, as $n = 7$ is prime. However, [20] have shown that there exist 715 isomorphism classes of maximal partial 2-spreads of $\mathcal{P}_7$ consisting of 17 RDCSSs.

For $n \geq 8$, we are unaware of any such results. For an example of two non-isomorphic 1-spreads of $\mathcal{P}_6$, consider $d_1 = \{f_1, \ldots, f_{21}\}$, shown in Table 3, and $d_2 = \{g_1, \ldots, g_{21}\}$, where $g_i = f_i$ for $i = 1, \ldots, 18$, and $g_{19} = \{ACE, AF, CEF\}$, $g_{20} = \{BCDF, CF, BD\}$, and $g_{21} = \{ABCD, AEF, BCDEF\}$. Here, the spread for $d_2$ was obtained via partitioning the RDCSSs $f_{19}, f_{20},$ and $f_{21}$ in $d_1$ into 3 new RDCSSs. The non-isomorphism of the spreads was verified using Algorithm 1.

Admittedly, multistage factorial experiments with a small number of basic factors have been typically more common in industrial experiments. However, some of the modern experiments, for instance using computer simulation models, involve a large number of factors. We now argue that the classification of balanced spreads presented above generates a rich class of star-based designs.

One may not find 0-spreads to be useful for a spread-based design, but such spreads can play a crucial role in constructing useful star-based designs, e.g., via Lemma 2. For instance, in the plutonium alloy example, a 0-spread of $\mathcal{P}_2$ is used to construct the balanced covering star $St(5, 3, 4, 3)$ of $\mathcal{P}_5$ which generates $PA_2$—a three-stage $2^5$ split-lot design with four randomization factors at each stage. This
|    | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $f_6$ | $f_7$ | $f_8$ |
|----|-------|-------|-------|-------|-------|-------|-------|-------|
| F  | E     | D     | C     | B     | A     | EF    | DE    |
| ABCEF | ABDF  | ACF   | BF    | AE    | DEF   | CDE   | BCD   |
| ABCE  | ABDEF | ACDF  | BCF   | ABE   | ADEF  | CDF   | BCE   |
| $f_9$ | $f_{10}$ | $f_{11}$ | $f_{12}$ | $f_{13}$ | $f_{14}$ | $f_{15}$ | $f_{16}$ |
| CD  | BC    | AB    | DF    | CE    | AC    | BEF   | ADE   |
| ABC | ABEF  | ADF   | BE    | AD    | BDE   | ACD   | BCEF  |
| ABD | ACEF  | BDF   | BDEF  | ACDE  | ABCDE | ABCDEF | ABCDF |
| $f_{17}$ | $f_{18}$ | $f_{19}$ | $f_{20}$ | $f_{21}$ |
| CDEF | BCDE  | ABCD  | AEF   | BD    |
| ABDE | ACDEF | BCDF  | CF    | CEF   |
| ABCF | ABF   | AF    | ACE   | BCDEF |
design is desirable as the significance of all effects can be assessed using four half-normal plots (see [35] for details). Similarly, 1-spreads and 2-spreads can also be augmented with different sized nuclei to form a variety of stars, providing flexibility for constructing small to large star-based designs.

We now use the results from A1–A5 and the theoretical results presented in Sects. 3 and 5 to classify the non-isomorphic balanced covering stars $St(n, \mu, t, t_0)$ of $P_n$. For convenience, we follow the same notation, $u = n - t_0$ and $h = t - t_0$, as in Lemma 2.

B1. For any given $0 \leq u \leq 5, t_0 \geq 0$ and $h$ that divides $u$, all balanced covering stars $St(u + t_0, \mu, h + t_0, t_0)$ of $P_{u+t_0}$ are isomorphic to each other. The proof follows from Lemma 1 and A1–A3.

B2. For $t_0 \geq 0$, all balanced covering stars $St(6 + t_0, \mu, 2 + t_0, t_0)$ of $P_{6+t_0}$ are isomorphic to each other. The result follows from A4.

B3. For every $t_0 \geq 0$, there exist 131044 mutually non-isomorphic balanced stars $St(6 + t_0, \mu, 3 + t_0, t_0)$ of $P_{6+t_0}$. The result follows from A4.

The categories defined by B1 and B2 contain most of the popular stars used to obtain RDCSS-based designs. For cases falling outside of these categories, it may be necessary to search over representatives from all isomorphism classes when searching for designs. This is a difficult problem because other than the exhaustive search methods used by [41] and [24], there is no known strategy for obtaining representatives of all isomorphism classes.

6.2 Cyclic-Spread-Based Designs

In this section, we provide results demonstrating that the most popular approach of constructing a balanced spread—the cyclic approach of [19]—only accesses a single isomorphism class, even if many exist. The algebraic results used in the proofs (presented in Appendix B) are mostly based on [21].

Although the spreads obtained via the cyclic construction method (outlined in Sect. 2.2) may vary with the choice of the primitive element and primitive polynomial of $GF(2^u)$, the next two results establish that such spreads are equivalent or isomorphic.

**Theorem 1** Let $\psi_1 = \{f_1, \ldots, f_\mu\}$ and $\psi_2 = \{g_1, \ldots, g_\mu\}$ be two $(h - 1)$-spreads of $P_u$ constructed using the cyclic method with two different roots $\alpha$ and $\beta$ of the same primitive polynomial $P(\omega)$. Then $\psi_1$ is equivalent to $\psi_2$.

**Theorem 2** Let $\psi_1 = \{f_1, \ldots, f_\mu\}$ and $\psi_2 = \{g_1, \ldots, g_\mu\}$ be two $(h - 1)$-spreads of $P_u$ constructed using the cyclic method with two different primitive polynomials $P_1(\omega)$ and $P_2(\omega)$, respectively. Then, $\psi_1$ is isomorphic to $\psi_2$.

See Appendix B for the proofs of Theorems 1 and 2. Although the cyclic construction method for $(h - 1)$-spreads of $P_u$ is widely-used, it accesses only a fraction...
of all possible spread/star-based designs. For example, only one of the 131044 isomorphism classes of 1-spreads of $P_6$ is obtained using the cyclic method. As per our knowledge, one may have to rely on exhaustive search to find non-isomorphic spread/star-based designs.

7 Concluding Remarks

In this paper, we formalize the definition of isomorphism of multistage factorial designs under the unified framework based on randomization defining contrast subspace (RDCSS), developed by [33]. Focussing on the RDCSS-based designs that are derived from balanced spreads and balanced covering stars, we have developed isomorphism check algorithms that are more efficient than the naive approach of iterating through all possible relabelings and reorderings. We have also provided a complete classification of small designs that are typically assumed to be important from practical standpoint. Furthermore, the proposed algorithms and relevant functions are implemented in R for easy access.

A few remarks are as follows. Both the theoretical results and the algorithms can easily be generalized for unbalanced spreads and stars; however, the construction and complete classification of such designs require more work. Some of the theoretical results will also hold when generalized to $q$-level multistage factorial designs with randomization restrictions. The proposed relabeling approach can also be used to find a design that meets the pre-specified randomization restrictions.

Acknowledgements Ranjan’s research was partially supported by the IIM Indore’s Grant for External Research Collaboration. Mendivil’s research was funded in part by NSERC 2012:238549.

Compliance with Ethical Standards

Conflicts of interest The authors declare that they have no conflict of interest.

Appendix A: Proofs of Results in Sect. 4

Proof of Proposition 1 Without loss of generality, suppose that $f_1, \ldots, f_{\ell}$ are the RDCSSs for $d_1$ that contain at least one of $x_1, \ldots, x_n$. Then, by the partial RDCSS mapping property, the elements of distinct $f_1, \ldots, f_{\ell}$ must be mapped to distinct RDCSSs in $d_2$. There are $\mu!/\mu - \ell)!$ different ways to choose a correspondences between these $\ell$ RDCSSs in $d_1$ and $\ell$ of the $\mu$ RDCSSs which comprise $d_2$. Subsequently, there are $\prod_{j=1}^{m} \left(2^j - 2^{j-1}\right)$ distinct choices of linearly independent points in each RDCSS of $d_2$ to which we can map the $m_i$ points $f_i$. Combining these counts as a product yields the result. $\square$

Comment on Proposition 1. The upper bound given in Proposition 1 is not necessarily tight— it is possible that some RDCSS correspondences do not yield full rank solutions. For example, if $n = 6$, $m_1 = m_2 = m_3 = 2$ and $g_1 = \langle A, B, \rangle$,
Proof of Proposition 2 When \( t = 1 \), the result is trivial, so we limit our consideration to \( t > 1 \). Suppose that within \( \psi \) there exists \( k (t - 1) \)-flats \( f_{u_1}, \ldots, f_{u_t} \) of \( \mathcal{P}_n \) such that \( |\langle \cup_{i=1}^k f_{u_i} \rangle| = 2^k - 1 \) for some integer \( k \leq n/t \). This is guaranteed to at least hold for \( k = 1 \) by definition of a spread. If \( k = n/t \), then the result is immediate. Otherwise, \( k \leq n/t - 1 \), and of the \( 2^n - 2k^t \) points not contained within \( \cup_{i=1}^k f_{u_i} \), \( 2^n - 2^k \) are not contained by \( \langle \cup_{i=1}^k f_{u_i} \rangle \) leaving \( 2^k - 2^t \) that do fall within that span. Recall that all \((t-1)\)-spreads of \( \mathcal{P}_n \) contain \( \mu = (2^n - 1)/(2^t - 1) (t - 1) \)-flats. Then, \( |\psi| - k \) is given by

\[
\mu - k = \frac{2^n - 1}{2^t - 1} - k = \sum_{i=1}^{n/t} 2^{(t-1)i} - k \geq 2^k - k.
\]

The pigeonhole principle guarantees that at least one \((t-1)\)-flat \( f_{u_{i+1}} \) shares no elements with \( \langle \cup_{i=1}^k f_{u_i} \rangle \). This flat can be appended to the list \( f_{u_1}, \ldots, f_{u_t} \) without introducing any linear dependence. Proceeding inductively, additional flats can be included until \( k = n/t \).

\[\square\]

Appendix B: Proofs of Results in Sect. 6.2

First we give a technical lemma and then the proofs of the two theorems.

Lemma 3 Let \( \psi = \{f_1, \ldots, f_\mu\} \) be an \((h - 1)\)-spread of \( \mathcal{P}_\mu \) constructed with the cyclic method using a primitive polynomial \( P(\omega) \) and root \( \omega \). Then,

(a) \( x_1 = \omega^a \) and \( x_2 = \omega^b \) are in the same \((t - 1)\)-flat \( f \in \psi \) if and only if \( a \equiv b \text{ mod } \mu \);

(b) the set of all nonzero roots of \( \omega^{2^h} - \omega \) is equal to the set of all elements of the form \( \omega^a \) where \( a \equiv 0 \text{ mod } \mu \). Thus, the first \((h - 1)\)-flat, \( f_1 \in \psi \), corresponds to the set of all nonzero elements of \( GF(2^h) \);

(c) \( f_2, \ldots, f_\mu \) are multiplicative cosets of \( f_1 \) in the group \( GF(2^\mu) \) of nonzero elements of \( GF(2^\mu) \).

Proof (a) follows trivially from the cyclic structure in Table 1. For part (b), since \( \mu(2^h - 1) = 2^\mu - 1 \), or \( \mu 2^h \equiv \mu \text{ mod } 2^\mu - 1 \),

\[
(\omega^{\mu})^{2^h} = (\omega^{\mu 2^h})^\mu \equiv (\omega^\mu)^\mu = \omega^\mu.
\]
and hence, \(\omega^2\mu\) is a root of \(\omega^{2^k} - \omega\). Part (c) follows from noting that the elements of \(f_i\) are of the form \(\omega^{k\mu+i} = \omega^i\omega^{k\mu}\), where \(0 \leq i < \mu\).

**Proof of Theorem 1** We need to show that for every \(g_j \in \psi_2\), there exists a unique \(f_j \in \psi_1\) such that the elements in \(g_j\) are in \(f_j\). Let \(e_1\) and \(e_2\) be two distinct effects in \(g_j\), then from Lemma 3(a), \(e_1 = \beta^a, e_2 = \beta^b\) and \(a \equiv b \pmod{\mu}\). From Theorem 2.14 of [21], there exists \(0 \leq k \leq u\) such that \(\beta = a^{2^k}\). Thus, \(e_1 = \beta^a = (a^{2^k})^a = a^{2^ka}\) and \(e_2 = \beta^b = a^{2^kb}\). Note that \(a \equiv b \pmod{\mu}\) implies \(2^ka \equiv 2^kb \pmod{\mu}\), as \(\gcd(2^k, \mu) = 1\). Consequently, \(e_1\) and \(e_2\) must belong to the same flat in \(\psi_1\) (from Lemma 3(a)).

**Proof of Theorem 2** We establish the existence of an IEC by constructing one. Our isomorphism will be a field isomorphism, which makes it easier to show that it is an IEC.

Let \(a\) be the primitive root of \(P_1(\omega)\) which is used to construct \(\psi_1\) and let \(\beta\) be the primitive root of \(P_2(\omega)\) which is used to construct \(\psi_2\). By Lidl and Niederreiter [21, Thm 2.40], there is a primitive polynomial \(Q(x)\) of degree \(u\) whose roots form a basis for \(GF(2^u)\) over \(\mathbb{Z}_2\). Note that if \(\omega\) is one of these roots then the other \(u-1\) roots are all of the form \(\omega^{2i}\) for \(i = 1, \ldots, u-1\). There are \(a, b \in \{1, \ldots, 2^u - 2\}\) with both \(a^a\) and \(b^b\) roots of \(Q(x)\). We define our IEC \(\Phi\) by first setting

\[
\Phi((a^a)^z) = (\beta^b)^{2i} \quad \text{for } i = 0, 1, \ldots, u-1,
\]

and then extending \(\Phi\) to all of \(GF(2^u)\) by linearity. Since the roots of \(Q(x)\) form a basis, this uniquely defines \(\Phi\).

Our next task is to show that \(\Phi\) is a field isomorphism. By our definition, \(\Phi\) is linear; we need to show that \(\Phi\) preserves multiplication. Since \(Q(x)\) is primitive and \(a^a, b^b\) are both roots of \(Q(x)\), it is enough to show \(\Phi((a^a)^k) = (\beta^b)^k\) for all \(k = 1, \ldots, 2^u - 1\). Fix \(k\). Since \(a^a, a^{2a}, \ldots, a^{2^{u-1}a}\) are the distinct roots of \(Q(x)\) and are a basis, there are constants \(c_i \in \mathbb{Z}_2\) so that

\[
a^{ak} = \sum_i c_i a^{2^k i}.
\]

Consider the polynomial \(H(x) = x^k - \sum_i c_i x^{2^k i}\). Then \(H(\omega^a) = 0\) by definition of \(c_i\). However, since \(x \mapsto x^{2^k}\) is a field automorphism for any \(j\), this means that \(H(\omega^{2^k}) = 0\) as well. Thus, all the roots of \(Q(x)\) are also roots of \(H(x)\). Since \(b^b\) is a root of \(Q(x)\), then \(H(b^b) = 0\) or \(b^{bk} = \sum_i c_i b^{2^k i}\). However, then

\[
\Phi(a^{ak}) = \Phi(\sum_i c_i a^{2^k i}) = \sum_i c_i \Phi(a^{2^k i}) = \sum_i c_i (\beta^b)^{2^k i} = (\beta^b)^k
\]

and so \(\Phi\) is also a field isomorphism. We claim that \(\Phi\) is an IEC. To see this, we first note that by is also a field isomorphism. We [21, Thm 2.21], \(\Phi\) maps the roots
of $x^{2^h} - x$ to roots of $x^{2^h} - x$. That is, it maps $GF(2^h) \subset GF(2^u)$ to itself. This indicates, by Lemma 3(b), that $\Phi(f_i) = g_i$. Additionally, since $\Phi$ is a field isomorphism, it maps any multiplicative coset of $GF(2^h)^*$ in $GF(2^u)^*$ to some multiplicative coset of $GF(2^h)^*$. By Lemma 3(c), each $(h - 1)$-flat $f_i$ of $\psi$ is mapped to a unique $(h - 1)$-flat $g_j$ in $\psi_2$. Thus $\Phi$ is an IEC for $\psi_1$ and $\psi_2$. □

Appendix C: R Codes for Easy Implementation

In this section, we discuss various functions that are used to implement Algorithm 1 and Algorithm 2 for checking the isomorphism of balanced spread- and star-based designs. These functions are implemented in R and have been uploaded to GitHub for easy access (see https://github.com/neilspencer/IsoCheck/). The usage and brief description of the key functions are as follows:

The isomorphism of two $(t - 1)$-spreads of $PG(n - 1, 2)$, spread1 and spread2, can be checked using the following R code:

```r
R> checkSpreadIsomorphism(spread1, spread2, returnfirstIEC = T).
```

The third argument "returnfirstIEC = T" specifies whether the algorithm searches until it finds the first IEC (might only take a few second) or if it continues to search for and returns all IECs (which can take a long time). For non-isomorphic spreads or stars, the run times are the same (none are found). However, for isomorphic spreads, stopping once we have found one IEC (which means they are isomorphic) is much faster.

Similar to spread-isomorphism, two stars star1 and star2 can be checked for isomorphism using the following R code

```r
R> checkStarIsomorphism(star1, star2, returnfirstIEC = T).
```

It is assumed that both spreads are $(t - 1)$-spreads, and both stars are $St(n, \mu, t, t_0)$ of $PG(n - 1, 2)$. The isomorphism check for stars is slightly different than for spread—it exploits the spread to star correspondence to reduce the dimension of the search space (as described by Algorithm 2). Both checkSpreadIsomorphism and checkStarIsomorphism call several important functions such as finding the bitstring representation of flats for checking equivalence, and applying collineations for relabeling of spreads and stars. The usage of these functions are illustated as follows:

```r
R> getBitstrings(spread1)
R> applyCollineation(C, spread1)
R> checkspreadEquivalence(spread1, spread2)
R> checkstarEquivalence(star1, star2)
```
Though the user can input the spreads of their choice in a specified format as discussed in the "readme" file and "exampleScript.R", we have coded several spreads and stars that are used in this paper (see the help manual of the R package "IsoCheck").

References

1. Addelman S (1964) Some two-level factorial plans with split plot confounding’. Technometrics 6(3):253–258
2. André J (1954) Über nicht-desarguessche ebenen mit transitiver translationsgruppe. Math Z 60(1):156–186
3. Bailey RA (2004) Association schemes: designed experiments, algebra and combinatorics, vol 84. Cambridge University Press, Cambridge
4. Batten LM (1997) Combinatorics of finite geometries. Cambridge University Press, Cambridge
5. Bingham D, Sitter RR (1999) Minimum-aberration two-level fractional factorial split-plot designs. Technometrics 41(1):62–70
6. Bingham D, Sitter R, Kelly E, Moore L, Olivas JD (2008) Factorial designs with multiple levels of randomization. Stat Sin 18(2):493–513
7. Bose RC (1947) Mathematical theory of the symmetrical factorial design. Sankhyā: Indian J Stat 8:107–166
8. Box GEP, Hunter WG, Hunter JS (1978) Statistics for experimenters. John Wiley & Sons, Inc., New York
9. Butler NA (2004) Construction of two-level split-lot fractional factorial designs for multistage processes. Technometrics 46:445–451
10. Cheng CS (2016) Theory of factorial design: single-and multi-stratum experiments. Chapman and Hall/CRC, London
11. Cheng CS, Tang B (2005) A general theory of minimum aberration and its applications. Ann Stat 33(2):944–958. https://doi.org/10.1214/009053604000001228
12. Cheng CS, Tsai PW (2011) Multistratum fractional factorial designs. Stat Sin 21:1001–1021
13. Coxeter HSM (1969) Introduction to geometry, 2nd edn. Wiley, New York
14. Daniel C (1959) Use of half-normal plots in interpreting factorial two-level experiments. Technometrics 1(4):311–341
15. Dean A, Voss D (1999) Design and analysis of experiments. Springer, New York
16. Eisfeld J, Storme L (2000) Partial t-spreads and minimal t-covers in finite projective spaces. Ghent University, Intensive Course on Finite Geometry and its Applications
17. Gordon NA, Shaw R, Soicher LH (2004) Classification of partial spreads in PG (4, 2). Available online as http://www.maths.qmul.ac.uk/~leonard/partialspreads/PG42new.pdf
18. Hedayat AS, Sloane NJA, Stufken J (2012) Orthogonal arrays: theory and applications. Springer, Berlin
19. Hirschfeld J (1998) Projective geometries over finite fields. Oxford mathematical monographs. Oxford University Press, New York
20. Honold T, Kiermaier M, Kurz S (2019) Classification of large partial plane spreads in PG(6, 2) and related combinatorial objects. J Geom 110(1):5
21. Lidl R, Niederreiter H (1994) Introduction to finite fields and their applications. Cambridge University Press, Cambridge
22. Lin C, Sitter R (2008) An isomorphism check for two-level fractional factorial designs. J Stat Plan Inference 134:1085–1101
23. Ma CX, Fang KT, Lin DK (2001) On the isomorphism of fractional factorial designs. J Complex 17(1):86–97
24. Mateva ZT, Topalova ST (2009) Line spreads of PG (5, 2). J Comb Des 17(1):90–102
25. McDonough T, Shaw R, Topalova S (2014) Classification of book spreads in PG (5, 2). Note di Mat 33(2):43–64
26. Mee R (2009) A comprehensive guide to factorial two-level experimentation. Springer, Berlin
27. Mee RW, Bates RL (1998) Split-lot designs: experiments for multistage batch processes. Technometrics 40(2):127–140
28. Miller A (1997) Strip-plot configurations of fractional factorials. Technometrics 39:153–161
29. Mukerjee R, Wu C (2006) A modern theory of factorial design. Springer, New York
30. Nelder J (1965a) The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. Proc R Soc Lond A 283:147–162
31. Nelder J (1965b) The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance. Proc R Soc Lond A 283:163–178
32. R Core Team (2014) R: A Language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/
33. Ranjan P (2007) Factorial and fractional factorial designs with randomization restrictions—a projective geometric approach. Ph.D. thesis, Department of Statistics and Actuarial Science-Simon Fraser University
34. Ranjan P, Bingham DR, Dean AM (2009) Existence and construction of randomization defining contrast subspaces for regular factorial designs. Ann Stat 37(6A):3580–3599
35. Ranjan P, Bingham DR, Mukerjee R (2010) Stars and regular fractional factorial designs with randomization restrictions. Stat Sin 20(4):1637–1653
36. Ryan TP (2007) Modern experimental design. Wiley, London
37. Shaw R, Topalova ST (2014) Book spreads in PG (7, 2). Discrete Math 330:76–86
38. Soicher L (2000) Computation of partial spreads, web preprint
39. Speed TP, Bailey RA (1982) On a class of association schemes derived from lattices of equivalence relations. Algebraic structures and applications. Marcel Dekker, New York, pp 55–74
40. Tjur T (1984) Analysis of variance models in orthogonal designs. Int Stat Rev 52:33–81
41. Topalova S, Zhelezova S (2010) 2-Spreads and transitive and orthogonal 2-parallelisms of PG (5, 2). Graphs Comb 26(5):727–735
42. Wu CJ, Hamada M (2009) Experiments: planning, analysis and optimization, 2nd edn. Wiley, London

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.