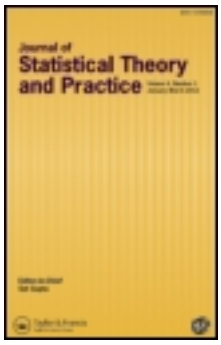

Journal of Statistical Theory and Practice
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ujsp20

Two factor saturated designs: cycles, Gini index and state polytopes

Roberto Fontana [a], Fabio Rapallo [b] & Maria-Piera Rogantin [c]
[a] Department of Mathematical Sciences, Politecnico di Torino, Turin, Italy
[b] Department DISIT, Università del Piemonte Orientale, Alessandria, Italy
[c] Dipartimento di Matematica, Università di Genova, Genova, Italy
Accepted author version posted online: 20 Sep 2013.

To cite this article: Journal of Statistical Theory and Practice (2013): Two factor saturated designs: cycles, Gini index and state polytopes, Journal of Statistical Theory and Practice, DOI: 10.1080/15598608.2014.840518

To link to this article: http://dx.doi.org/10.1080/15598608.2014.840518

Two factor saturated designs:

cycles, Gini index and state polytopes

Roberto Fontana, Department of Mathematical Sciences, Politecnico di Torino, Turin, Italy (roberto.fontana@polito.it)

Fabio Rapallo, Department DISIT, Università del Piemonte Orientale, Alessandria, Italy (fabio.rapallo@unipmn.it)

Maria-Piera Rogantin, Dipartimento di Matematica, Università di Genova, Genova, Italy (rogantin@dima.unige.it)

Abstract

In this paper we analyze and characterize the saturated fractions of two-factor designs under the simple effect model. Using Li et al.ear algebra, we define a criterion to check whether a given fraction is saturated or not. We also compute the number of saturated fractions, providing an alternative proof of the Cayley's formula. Finally we show how, given a list of saturated fractions, Gini indexes of their margins and the associated state polytopes could be used to classify them.

AMS Subject Classification: 62K15, 15B34, 05B20, 62H17

Key words: Estimability, Gini index, State polytope, Universal Markov basis.

1 Introduction

The study of estimable designs is an investigated problem in Design of experiments. Given a model, saturated fractions are subsets of the factorial design with as many points as the number

of the model parameters and such that all the parameters are estimable, i.e., the corresponding model matrix is full rank. The books Raktoe et al. (1981) Raktoe, Hedayat, and Federer and Bailey (2008) are general references for the theory of Design of experiments where the issue of saturated fractions is discussed.

In this paper, we present a new approach to lay the foundations for studying the fractions with the minimal number of points. As a first step in this direction, we analyze and characterize the saturated fractions of two-factor designs under the simple effect model. We point out that this theory allows also the generation of saturated fractions of multi-factor designs, as shown in Remark 1. The characterization of saturated fractions is a relevant question, as a randomly chosen minimal set of design points yields a singular model matrix with very high probability when the number of the factor levels becomes large, even under the simple effect model. Moreover, we study two methodologies that can support the classification of a list of saturated fractions. The first one makes use of the Gini index, see Gini (1912) , while the second one is based on state polytopes, see Sturmfels (1996) .

In order to characterize saturated fractions, our approach is based on two main ingredients. First, we apply tools from Li et al.ear algebra and Combinatorics to characterize the saturated fractions. Some notions, and in particular the definition of *k*-cycle used in Section 2, has already been considered in the framework of contingency tables in Kuhnt et al. (2013) Kuhnt, Rapallo, and Rehage for the definition of robust procedures for outliers detection in contingency tables. Second, we identify a factorial design with a contingency table whose entries are the indicator function of the fraction, i.e., they are equal to 1 for the fraction points and 0 otherwise. This

implies that a fraction can also be considered as a subset of cells of the table. In Design of experiments, the application of techniques from contingency table analysis has been introduced in Fontana et al. (2012) Fontana, Rapallo, and Rogantin for special designs coming from Sudoku problems. Moreover, in Aoki and Takemura (2010) and Aoki and Takemura (2012) such techniques are used for experiments with discrete response variable.

With respect to other approaches to the problem of describing saturated fractions, see e.g. Krafft and Schaefer (1997) and Dey et al. (1995) Dey, Shah, and Das, the techniques used here are mainly based on the notion of Markov bases and were originally developed for contingency tables in the field of Algebraic Statistics. As a general definition, Algebraic Statistics concern the application of polynomial algebra techniques to Statistics. This theory was first presented in Fontana et al. et al. (2001) Fontana et al., Riccomagno, and Wynn, and a recent account can be found in Drton et al. (2009) Drton, Sturmfels, and Sullivant. In the present paper, we show that the *k*-cycles defines a special Markov basis, named Universal Markov basis, introduced in Rapallo and Rogantin (2007) . This new approach has also interesting extensions to multi-factor designs, see Fontana et al. (2013) Fontana, Rapallo, and Rogantin, but in the two factor setting we are able to provide simpler proofs and to explicitly describe the saturated fractions. This issue leads us to a complete analysis of the saturated fractions with respect to several criteria.

We also develop a methodology to classify a given set of saturated designs. It is based on the computation of the Gini indexes of the univariate margins of each fraction. Then, for polynomial models, we show how the state polytope of each fraction could be used to compare different saturated fractions from a graphical point of view.

In this work we adopt different perspectives to analyze two-factor saturated designs. In the first part we focus on their algebraic characterization, including counting issues, while in the second part we study some methodologies for their classification. The paper is organized as follows. In Section 2 we set some notations, we state the problem, we define the *k*-cycles in a fraction and we characterize them in terms of orthogonal arrays. In Section 3 we prove the main result, showing that the absence of a *k*-cycle is a necessary and sufficient condition for obtaining a saturated fraction, while in Section 4, we enumerate the saturated fractions showing that their proportion over the whole number of fractions tends to zero as the number of levels increases. Section 5 is devoted to the classification of saturated fractions using several criteria, such as the word-length pattern, the Gini index and the state polytope.

2 Saturated designs and cycles

2.1 Notations and basic definitions

Let D be a full factorial design with 2 factors, A and B, with I and J levels, respectively $(I, J \geq 2), D = [I] \times [J] = \{1, \ldots, I\} \times \{1, \ldots, J\}$. We consider a Linear model on D:

$$Y_{i,j} = \mu_{i,j} + \varepsilon_{i,j} \quad for\ i \in [I], j \in [J],$$

where $Y_{i,j}$ are random variables with means $\mu_{i,j}$ and $\varepsilon_{i,j}$ are centered random variables that represent the error terms. In this paper we always consider the simple effect model, i.e.:

$$\mathbb{E}(Y_{i,j}) = \mu_{i,j} = \mu + \alpha_i + \beta_j \quad for\ i \in [I], j \in [J], \tag{2.1}$$

where μ is the mean parameter, and α_i and β_j are the main effects of *A* and *B*, respectively.

We denote by p the number of estimable parameters. Therefore, $p = I + J - 1$ for the model of Equation (2.1). Under a suitable parametrization, the matrix of this model is a full-rank matrix with dimensions $IJ \times (I + J - 1)$. In this paper we will use the following *model matrix*:

$$X = \left(m_0 \,|\, a_1 \,|\, \ldots \,|\, a_{I-1} \,|\, b_1 \,|\, \ldots \,|\, b_{J-1} \right), \tag{2.2}$$

where m_0 is a column vector of 1's, $a_1, \ldots, a_{I-1}$ are the indicator vectors of the first $(I-1)$ levels of the factor *A*, and $b_1, \ldots, b_{J-1}$ are the indicator vectors of the first $(J-1)$ levels of the factor *B*. It is known that this matrix corresponds to the following reparametrized model:

$$\mu_{i,j} = \tilde{\mu} + \tilde{\alpha}_i + \tilde{\beta}_j, \quad for\ i \in [I], j \in [J]; \quad \tilde{\mu} = \mu + \alpha_I + \beta_J, \quad \tilde{\alpha}_i = \alpha_i - \alpha_I, \quad \tilde{\beta}_j = \beta_j - \beta_J.$$

A subset $\mathcal{F}$, or fraction, of a full design D, with minimal cardinality $\#\mathcal{F} = p$, that allows us to estimate the model parameters, is a *main-effect saturated design*. By definition, the model matrix $X_{\mathcal{F}}$ of a saturated design is non-singular.

Example 1. Let us consider the case $I = 3$, $J = 4$ and the fraction

$$\mathcal{F} = \{(1,1),(1,2),(2,2),(2,3),(3,3),(3,4)\}.$$

The model matrix *X* of the full design and the model matrix $X_{\mathcal{F}}$ of the fraction are given in Figure 1. Is this case, $\det(X_{\mathcal{F}}) = 1$.

Remark 1. Notice that the design matrix $X_{\mathcal{F}}$ in Example 1 has an immediate generalization to multi-factor designs. In fact, it is easy to see that $X_{\mathcal{F}}$ is the design matrix of the following saturated fraction of a 2^5 design under the simple effect model, where the fraction is*:*

$$\{(1,2,1,2,2),(1,2,2,1,2),(2,1,2,1,2),(2,1,2,2,1),(2,2,2,2,1),(2,2,2,2,2)\}.$$

In the same way, $X_{\mathcal{F}}$ can be considered as the design matrix of a saturated fraction of a $3\times3\times2$ design under the same model, where the fraction is:

$$\{(1,1,2),(1,2,2),(2,2,2),(2,3,1),(3,3,1),(3,3,2)\}.$$

2.2 *k*-cycles and orthogonal arrays

As mentioned in the Introduction, in general, the problem of selecting saturated designs is non trivial and this is true also in the simple case of two factor design. The key ingredient to characterize a saturated design for two-factor designs is the notion of *cycle*, coming from Li et al.ear algebra and Combinatorics. Here we give a definition in terms of Design of experiments.

Definition 1. A k-cycle ($k \geq 2$) is a subset with cardinality $2k$ of a factorial design $I \times J$ with $I, J \geq k$ where each of the k selected levels (among the *I*'s and *J*'s*)* of each factor has exactly two replications.

Example 2. Some examples of fractions with k-cycles are given in Figure *2,* where the cell (i, j) has a bullet if the design point (i, j) belongs to the fraction. As described above, we identify the

design points with the cells of a contingency table in order to simplify the presentation. The corresponding fractions are:

$$\mathcal{F}_1 = \{(1,1),(1,3),(2,1),(2,2),(3,2),(3,3),(4,3)\},$$

$$\mathcal{F}_2 = \{(1,1),(1,3),(2,2),(2,4),(3,2),(3,3),(4,1),(4,4)\},$$

$$\mathcal{F}_3 = \{(1,1),(1,3),(2,1),(2,3),(3,2),(3,4),(4,2),(4,4)\}.$$

Notice that $\mathcal{F}_1$ contains a 3-cycle, $\mathcal{F}_2$ and $\mathcal{F}_3$ contain a 4-cycle. In $\mathcal{F}_3$ the 4-cycle can be decomposed into two sub-cycles.

In a natural way, the coordinate points of the design D can be considered as the rows of a $\#D \times d$ matrix, where *d* is the number of factors. With a slight abuse of notation, we still call this matrix *design*. The same holds for fractions. In order to analyze the role of the *k*-cycles within the framework of Design of experiments we recall here a combinatorial definition of orthogonal array, see Hedayat et al. (1999) and Fontana (2013) .

Definition 2. A fraction $\mathcal{F}$ of a design $I_1 \times \cdots \times I_d$ with $\#\mathcal{F} = n$ is an orthogonal array of size n and strength t if, for all t-tuples of its factors $\mathcal{F}_{i_1},\ldots,\mathcal{F}_{i_t}$, all possible combinations of levels in $[I_{i_1}]\times\cdots\times[I_{i_t}]$ appear equally often. We denote such an orthogonal array with $OA(n;(I_1,\ldots,I_d);t)$.

Proposition 1. A k-cycle ($k \geq 2$) is:

1. an $OA(2k;(k,k);t)$ where $t=2$ if $k=2$, and $t=1$ if $k \geq 3$;

2. the union of two disjoint orthogonal arrays $OA(k;(k,k);1)$.

PROOF. 1. This fact follows from Definition 1. In particular, for $k=2$ the fraction F coincides with the full factorial design 2^2.

2. We construct two disjoint fractions OA_1 and OA_2, $OA_1 \bigcup OA_2 = F$, iteratively. Starting from a given point of the fraction, we assign alternatively to OA_1 and OA_2 the points of the fraction, choosing the first or the second factor.

• Choose a point of $\mathcal{F}$, say (i_1, j_1), and assign it to OA_1.

• Consider the unique point of F with the same level for the first factor, (i_1, j_2), with $j_2 \neq j_1$, and assign it to OA_2.

• Consider the unique point of F with the same level for the second factor, (i_2, j_2), with $i_2 \neq i_1$, and assign it to OA_1.

• Consider the unique point of $\mathcal{F}$ with the same level for the first factor, (i_2, j_3) with $j_3 \neq j_2 \neq j_1$ and assign it to OA_2.

• And so on, until the unique point to choose is already assigned.

If not all points of the fraction have been assigned, i.e. if the fraction contains sub-cycles, it is enough to start the procedure above on the remaining points, until all the points are assigned. In this way both OA_1 and OA_2 have, by construction, exactly one replicate for each of the k levels of the two factors.

Example 3. We show how the decomposition of a fraction into two orthogonal arrays works on a 4-cycle. Let $I = J = 4$, and consider the 4-cycle

$$\mathcal{F}_2 = \{(1,1),(1,3),(2,2),(2,4),(3,2),(3,3),(4,1),(4,4)\}$$

already considered in Example 3 and displayed in Figure 2. The relevant orthogonal arrays are:

$$OA_1 = \{(1,1),(2,2),(3,3),(4,4)\},$$

$$OA_2 = \{(1,3),(2,4),(3,2),(4,1)\}.$$

To determine the number of k-cycles we need the notion of derangement. A *derangement* is a permutation such that no element appears in its original position.

Proposition 2. The number of k-cycles is

$$\frac{k!\,!k}{2},$$

where $!k$ denotes the number of derangements of k elements.

PROOF. Let us consider OA_1 and OA_2 as in Proposition 1. OA_1 represents a permutation π_1 of $[k]$. The fraction OA_2 represents a derangement of $\pi_1([k])$.

To actually compute $!k$, recall that $!k$ can be approximated by $\lfloor k!/e+0.5 \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function. For more details on this theory, see for instance Hassani (2003) .

3 *k*-cycles and saturated fractions

As mentioned in the previous sections, the connections between saturated designs and cycles have been explored in a slightly different framework, in the study of robust estimators in contingency tables analysis, see Kuhnt et al. (2013) Kuhnt, Rapallo, and Rehage. Nevertheless, we restate here the relevant theorem within the language of Design of experiments and we give the proof, as its main algorithm will be used later in the paper.

Theorem 1. A fraction $\mathcal{F}$ with $p = I + J - 1$ points yields a saturated model matrix if and only if it does not contain cycles.

PROOF. In view of Proposition 1, a cycle can be decomposed into two disjoint orthogonal arrays, OA_1 and OA_2, of *k* points each.

$\Rightarrow$ Suppose that $\mathcal{F}$ contains a cycle. When we sum the rows of the model matrix $X_{\mathcal{F}}$ with coefficient +1 for the points in the OA_1 and with coefficient -1 for the points in OA_2, we produce a null Li et al.ear combination and therefore the determinant of $X_{\mathcal{F}}$ is zero.

$\Leftarrow$ Suppose that $X_{\mathcal{F}}$ is singular, i.e., there exists a null Li et al.ear combination of its rows with coefficients that are not all zero. Denote by $r_{(i,j)}$ the row of the model matrix corresponding to the point (i, j) of the fraction. Therefore, we have

$$\gamma_1 r_{(i_1,j_1)} + \cdots + \gamma_p r_{(i_p,j_p)} = 0 \tag{3.1}$$

and the coefficients $\gamma_1, \ldots, \gamma_p$ are non all zero. Without loss of generality, suppose that $\gamma_1 > 0$. As the indicator vector of the i_1-th level of the first factor is in the column span of X , and the same holds for the j_1-th level of the second factor, we must have: (a) a point with the same level for the first factor, say (i_1, j_2), with negative coefficient in Equation (3.1); (b) a point with the same level for the second factor, say (i_3, j_1), with negative coefficient in Equation (3.1). Therefore, there must be a point with level i_3 for the first factor and a point with level j_2 for the second factor with positive coefficients. Now, two cases can happen: If the point (i_3, j_2) is a chosen point and its coefficient in Equation (3.1) is positive, we have a 2-cycle; otherwise, we iterate the same argument, and we yield a *k*-cycle, with $k > 2$.

Remark 2. Proposition 1 and Theorem 1 lead to an interesting connection with the theory of Markov bases for this kind of experimental designs. In Algebraic Statistics, a Markov basis is an important object associated to a model for contingency tables under Li et al.ear constraints. Although a detailed presentation of such connection is beyond the goal of this paper, nevertheless it is interesting to provide a sketch of this issue. In fact, we have already identified

the two-factor design with a binary $I \times J$ contingency table, and therefore the search of its corresponding Markov basis is a key question in Algebraic Statistics.

The theory in Rapallo and Rogantin (2007) , Sections 4 and 5, states that the relevant Markov basis for our problem can be computed from the complete bipartite graph of the design. The *complete bipartite graph* has one vertex for each level of *A*, one vertex for each level of *B* and one edge connects each *A*-vertex with each *B*-vertex. A circuit of degree *k* is a closed path with $2k$ vertices, and with edges

$$(i_1, j_1), (j_1, i_2), (i_2, j_2), \ldots, (i_k, j_k), (j_k, i_1), \tag{3.2}$$

where $i_1, \ldots, i_k$ are distinct *A*-indices and $j_1, \ldots, j_k$ are distinct *B*-indices. A concise presentation of the theory of bipartite graphs can be found in Sturmfels (1996) . The *complete bipartite graph* for the 3×4 designs is depicted in Figure 3.

Then, the Markov basis for our problem is defined by associating a Markov move to each circuit of the complete bipartite graph. Such move has entry 1 for each edge in even position in the sequence (3.2), and has entry -1 for each edge in odd position.

Proposition 3. The *k*-cycles, decomposed into two orthogonal arrays as in Proposition 1, form a Markov basis.

PROOF. It is enough to observe that each circuit of degree *k* naturally defines a *k*-cycle decomposed as in Proposition 1.

4 The number of saturated designs

In this section we study the structure of the saturated fractions described in Section 2.

Definition 3. Given a fraction $\mathcal{F}$, we define its margins $m_A = (m_{A,1},\ldots,m_{A,I})$ and $m_B = (m_{B,1},\ldots,m_{B,J})$ where:

$$m_{A,i} = \sum_{(d_1,d_2)\in} (d_1 = i) \quad \text{for } i \in [I],$$

$$m_{B,j} = \sum_{(d_1,d_2)\in} (d_2 = j) \quad \text{for } j \in [J],$$

where $(\cdot)$ denotes the indicator function.

Notice that $m_{A,i}$ is the number of the occurrence in of the i-th level of the factor A. For example, the following saturated design

$$\mathcal{F} = \{(1,1),(1,2),(2,1),(2,4),(3,2),(3,3),(4,4)\}.$$

has margins $m_A = (2,2,2,1)$ and $m_B = (2,2,1,2)$.

The following lemmas for $I \times I$ designs will be used later in the proof of the main result of this section.

Lemma 1. Let $\mathcal{F}_I$ be a saturated $I \times I$ design. Its margins satisfy the following conditions:

1. $m_{A,+} = m_{B,+} = 2I - 1$ where $+$ denotes the summation over the corresponding index;

2. $m_{A,i} \geq 1$, $m_{B,j} \geq 1$ for all $i, j \in [I]$;

3. there exist $i, j \in [I]$ such that $m_{A,i} = m_{B,j} = 1$;

4. let $i_\star \in [I]$ be an index such that $m_{A,i_\star} = 1$. Let $(i_\star, j_\star)$ be the only point (d_1, d_2) of $\mathcal{F}_I$ such that $d_1 = i_\star$. Then $m_{B,j_\star} > 1$.

PROOF. 1. It is immediate to see that $m_{A,+} = \sum_{i=1}^{I} \sum_{(d_1,d_2)\in\mathcal{F}} (d_1 = i) = \#\mathcal{F}_I = 2I - 1$ and the same holds for *B*.

2. Refer to the matrix representation in Equation 2.2. By absurd, suppose that there exists an index i such that $m_{A,i} = 0$. We distinguish two cases: if $i \leq I - 1$ then $X_{\mathcal{F}_I}$ has a null column, corresponding to a_i; if $i = I$, the sum of the columns $a_1, \ldots, a_{I-1}$ is equal to m_0. In both cases, $X_{\mathcal{F}_I}$ is singular. The same applies to *B*.

3. This point follow immediately from items 1 and 2.

4. For sake of readability, we suppose that $i_\star = I$ and $j_\star = I$. Thus, the sum of $a_1, \ldots, a_{I-1}$ is equal to the sum of $b_1, \ldots, b_{I-1}$. Hence, $X_{\mathcal{F}_I}$ is singular.

*Remark 3***.** The four conditions in Lemma 1 are not sufficient for characterizing the saturated fractions. A simple counterexample is the following fraction of a 5×5 design*:*

$$\{(1,1),(1,2),(2,1),(2,3),(3,2),(3,3),(4,4),(4,5),(5,4)\},$$

that is not saturated, as it contains a 3-cycle.

In the following result we analyze how the saturation property is preserved when we add one level to each of the two factors, moving from an $I \times I$ design to an $(I+1)\times(I+1)$ design.

*Lemma 2***.** Let $\mathcal{F}_I$ be a saturated $I \times I$ design, and define an $(I+1)\times(I+1)$ design containing $\mathcal{F}_I$ as*:*

$$\mathcal{F}_{I+1} =_I \bigcup E_{I+1} \text{ with } \mathcal{F}_I \bigcap E_{I+1} = \varnothing.$$

Then $\mathcal{F}_{I+1}$ is saturated if and only if E_{I+1} has exactly two points, chosen in the union of the $(I+1)$-th row with the $(I+1)$-th column of $\mathcal{F}_{I+1}$, with the conditions $m_{A,I+1} \geq 1$ and $m_{B,I+1} \geq 1$.

PROOF. E_{I+1} must contain exactly two design points, as $\#\mathcal{F}_I = 2I-1$ and $\#\mathcal{F}_{I+1} = 2I+1$. If one or two points are not in the union of the $(I+1)$-th row with the $(I+1)$-th column of $_{I+1}$, there is a contradiction with Lemma 1. If both points are chosen in the $(I+1)$-th row and in the first *I* columns, the margin $m_{B,I+1} = 0$, a contradiction. With an analogous proof, the two points can not be chosen in the $(I+1)$-th column and in the first *I* rows. All the remaining cases are valid choices, as one among $m_{A,I+1}$ and $m_{B,I+1}$ is equal to 1 and therefore no cycles can appear.

We are now ready to approach the problem of computing the number of saturated fractions.

Proposition 4. Given $m_A = (m_{A,1},\ldots,m_{A,I})$ and $m_B = (m_{B,1},\ldots,m_{B,J})$ with $m_{A,+} = m_{B,+} = I + J - 1$, $m_{A,i} \geq 1, i \in [I]$ and $m_{B,j} \geq 1, j \in [J]$, the number of saturated designs with margins equal to m_A and m_B is

$$\binom{I-1}{m_{B,1}-1,\ldots,m_{B,I}-1}\binom{J-1}{m_{A,1}-1,\ldots,m_{A,I}-1}. \tag{4.1}$$

PROOF. First we consider $J = I$. Without loss of generality, we can assume that the margins of the fraction are arranged in the form:

$$m_{A,1} \geq m_{A,2} \geq \cdots \geq m_{A,I} = 1,$$

$$m_{B,1} \geq m_{B,2} \geq \cdots \geq m_{B,I} = 1.$$

Since $m_{A,I} = 1$, we can choose a point for the last row, but we have to exclude all design points (I,h) with $m_{B,h} = 1$, in order to satisfy the condition in Lemma 1, item 4. In the same way, we choose a point in the last column.

We repeat the same argument on the $(I-1)\times(I-1)$ design obtained by deletion of the last row and of the last column. It is immediate to see that both margins of such design have a component equal to 1. Hence, we iterate $(I-2)$ times the procedure above, until we have a degenerate 1×1 design with 1 as its unique element.

If we analyze this algorithm backward, we note that at each step we add two points according to the rule in Lemma 2, and therefore the constructed fraction is saturated. Thus, the procedure generates

$$\binom{I-1}{m_{B,1}-1,\ldots,m_{B,I}-1}\binom{I-1}{m_{A,1}-1,\ldots,m_{A,I}-1},$$

as each row can be chosen until the margin decreases to 1, and the same holds for columns.

Now, consider $J > I$. The *B*-margin can be arranged in the form:

$$m_{B,1} \geq m_{B,2} \geq \cdots \geq m_{B,I} = 1 = m_{B,I+1} = 1 = \cdots m_{B,J} = 1.$$

With this ordering of the margins, the square $I \times I$ table on the left can be analyzed as in the previous case, while for the last $J - I$ columns there is only the constraint given by the *A*-margin, because the *B*-margin is always equal to 1, and therefore no *k*-cycles can appear. Hence, the formula in (??) is proved.

*Theorem 2***.** The number of saturated $I \times J$ designs is $I^{(J-1)}J^{(I-1)}$.

PROOF. It follows from Proposition 4 and the classical multinomial theorem, by summation of all possible terms (see Equation 4.1) corresponding to all possible margins.

We notice that the key consequence of the above enumerations is that the proportion of singular design matrices is not negligible, and it becomes as large as *I* and *J* increase.

Corollary 1. Let us randomly choose $\mathcal{F} \subset I \times J$ with $\#\mathcal{F} = I + J - 1$. The probability that is a saturated $I \times J$ design is

$$\frac{I^{(J-1)} J^{(I-1)}}{\binom{IJ}{I+J-1}}$$

and it tends to 0 as I and J goes to infinity.

For instance, let us consider $I = J$. For $I = 3$ we obtain a saturated design in 64% of cases, for $I = 4$ in 36% of cases, for $I = 5$ in 19% of cases, while for $I = 6$ in 10% of cases. Hence, the characterization of non-singular designs, as given in Theorem 1, is useful from an algorithmic point of view, because the random choice of a subset with $I + J - 1$ points does is not an efficient procedure.

*Example 4***.** We discuss here the case $I = J = 4$ extensively. The number of saturated designs is $4^6 = 4096$, corresponding to 36% of designs with 7 points. The possible configurations of margins, up to the permutation of the levels, are: $(4,1,1,1)$, $(3,2,1,1)$ and $(2,2,2,1)$. The table below shows the number of saturated design with such margins for one factor, see Proposition 4, the number of multiset permutations of such configurations of margins and the product of them.

Margin	# from Prop. 4	Multiset permutations	Total
(4,1,1,1)	$\binom{3}{3,0,0,0}$	$\binom{4}{1,3}$	4
(3,2,1,1)	$\binom{3}{2,1,0,0}$	$\binom{4}{1,1,2}$	36
(2,2,2,1)	$\binom{3}{1,1,1,0}$	$\binom{4}{3,1}$	24

The table below shows the number of saturated design with respect to the two margins.

Margins	(4,1,1,1)	(3,2,1,1)	(2,2,2,1)
(4,1,1,1)	16	144	96
(3,2,1,1)	144	1296	864
(2,2,2,1)	96	864	576

Finally, Figure 4 shows the saturated 4×4 designs identified by contingency tables. As in the previous figures, a cell (i, j) has a bullet if the design point (i, j) belongs to the fraction. For each margin configuration, a representative of each equivalence class of tables is displayed. The equivalence is up to permutation of rows, permutation of columns, and transposition. We notice that in two cases there is more than one equivalence class.

5 Classification of saturated fractions

Many criteria and methodologies exist to evaluate a given fraction. In this section we consider the word-length pattern and the state polytope of a saturated fraction.

5.1 Generalized word-length pattern and Gini index for two-level factors

We focus here on the word-length pattern of a fraction as stated in terms of indicator function in Li et al. (2003) Li, Li et al., and Ye. We consider *s* factors, each with 2 levels, coded as $+1$ and -1. The full factorial design $\mathcal{L}$ is $\{1,-1\}^s$. We denote by $\mathcal{L}$ the set $\{0,1\}^s$ and by X^α, $\alpha \in \mathcal{L}$, a simple term or an interaction. Given a single-replicate fraction $\mathcal{F} \subseteq \mathcal{D}$, the indicator function of $\mathcal{F}$ written in polynomial form is $F(x) = \sum_{\alpha \in \mathcal{L}} b_\alpha X^\alpha(x)$, with $x \in \mathcal{D}$. We have $b_\alpha = \frac{1}{2^s} \sum_{x \in \mathcal{F}} X^\alpha(x)$. The reader can refer to Fontana et al. (2000) for a detailed presentation of the indicator function.

The generalized word-length pattern of $\mathcal{F}$ is the *s*-tuple $\left(\gamma_1(\mathcal{F}),\ldots,\gamma_s(\mathcal{F})\right)$, where each element $\gamma_j(\mathcal{F})$ measures the degree of aliasing of the interactions of order *j* (if $j=1$ the simple terms are considered). More precisely:

$$\gamma_j(\mathcal{F}) = \sum_{\|\alpha\|=j, \alpha \in \mathcal{L}} \left(\frac{b_\alpha}{b_0}\right)^2, \quad j = 1,\ldots,s$$

and $\| \alpha \| = \sum_{k=1}^{s} \alpha_k$ is the order of interaction.

Let us denote by $n_{\alpha,e}$ the number of points *x* of $\mathcal{F} \subseteq \mathcal{D}$ for which $X^\alpha(x) = e$, with $e \in \{+1,-1\}$:

$$n_{\alpha,e} = \sum_{x \in} (X^\alpha(x) = e),$$

where $(X^{\alpha}(x)=e)=1$ if $X^{\alpha}(x)=e$, and $(X^{\alpha}(x)=e)=0$ otherwise. According to Proposition 5, Fontana (2013) , the coefficients b_{α}, $\alpha \in \mathcal{L}$, of the indicator function *F* of a fraction $\mathcal{F}$ can be computed using $n_{\alpha,e}$ as stated in the following result.

Proposition 5. Let $n_{\alpha,e}=\sum_{x\in}(X^{\alpha}(x)=e)$ where $e\in\{+1,-1\}$. It holds

$$b_{\alpha}=\frac{1}{2^{s}}(n_{\alpha,1}-n_{\alpha,-1}),\ \alpha\in\mathcal{L}.$$

*Remark 4***.** For $\alpha=(0,\ldots,0)$, we get $b_0=\#\mathcal{F}/2^{s}$ and therefore

$$b_0=\frac{n_{\alpha,1}+n_{\alpha,-1}}{2^{s}},\ \forall\alpha\in\mathcal{L}.$$

For $\|\alpha\|=1$ we obtain the univariate margins of $\mathcal{F}$ For example $\alpha=(1,0,\ldots,0)$ provides

$$n_{(1,0,\ldots,0),e}=\sum_{x=(x_1,\ldots,x_s)\in}(x_1=e).$$

Now we observe that there is a strong relation between the generalized word-length pattern of a fraction $\mathcal{F}$ and the Gini indexes computed on $n_{\alpha,1}$ and $n_{\alpha,-1}$, Gini (1912) and Xu (2003) . The normalized Gini index *G* measures the inequality among the values of a frequency distribution. If we denote by $q_1,\ldots,q_n$, $q_i\le q_{i+1}$, the ordered frequencies corresponding to *n* units, it can be computed as

$$G(q_1,\ldots,q_n)=1-\frac{2}{n-1}\sum_{i=1}^{n-1}Q_i\ ,\qquad where Q_i=\frac{1}{q_1+\ldots+q_n}\sum_{k=1}^{i}q_k\ .$$

Proposition 6 specifies this relation.

Proposition 6. Let G_α the normalized Gini index corresponding to $n_{\alpha,1}$ and $n_{\alpha,-1}$, $G_\alpha = G(n_{\alpha,1}, n_{\alpha,-1})$. The following equality holds

$$(G_\alpha)^2 = \left(\frac{b_\alpha}{b_0}\right)^2 .$$

PROOF. From $b_\alpha = \frac{1}{2^s}(n_{\alpha,1} - n_{\alpha,-1})$ we get

$$\frac{b_\alpha}{b_0} = \frac{n_{\alpha,1} - n_{\alpha,-1}}{n_{\alpha,1} + n_{\alpha,-1}}.$$

Now let us consider the case $n_{\alpha,1} \le n_{\alpha,-1}$. From the definition of normalized Gini index we get

$$G_\alpha = G(n_{\alpha,1}, n_{\alpha,-1}) = 1 - 2Q_1 , \qquad where Q_1 = \frac{1}{n_{\alpha,1} + n_{\alpha,-1}} n_{\alpha,1} .$$

It follows

$$G_\alpha = \frac{n_{\alpha,-1} - n_{\alpha,1}}{n_{\alpha,1} + n_{\alpha,-1}} = -\frac{b_\alpha}{b_0}.$$

The case $n_{\alpha,-1} < n_{\alpha,1}$ gives $G_\alpha = +\frac{b_\alpha}{b_0}$. This completes the proof.

Corollary 2. The generalized word-length pattern $(\gamma_1(\mathcal{F}),\ldots,\gamma_s(\mathcal{F}))$ of a two level single replicate fraction of can be computed as

$$\gamma_j(\mathcal{F}) = \sum_{\|\alpha\|=j,\alpha\in\mathcal{L}} G_\alpha^2,\ j=1,\ldots,s, \qquad where G_\alpha = G(n_{\alpha,1}, n_{\alpha,-1}).$$

For two level designs we have proved that Gini indexes G_α are sufficient to compute the generalized word-length pattern of a fraction $\mathcal{F}$. Now we experiment the use of Gini indexes for saturated $I \times J$ designs.

We denote by m_A and m_B the margins of an $I \times J$ saturated design. As an example we consider two cases: $I = J = 7$ and $I = J = 4$. We compute $G_{(1,0)} \equiv G(m_A)$ and $G_{(0,1)} \equiv G(m_B)$ for all the possible margins, up to permutations of the levels of the factors.

The following table summarizes the results.

$m_X, X \in \{A, B\}$	$G(m_X)$	2 cm	$m_X, X \in \{A, B\}$	$G(m_X)$	$G^{(2)}(m_X)$
4-6 4,1,1,1	0.43		7,1,1,1,1,1,1	0.46	0.45
3,2,1,1	0.33		6,2,1,1,1,1,1	0.44	0.44
2,2,2,1	0.14		5,3,1,1,1,1,1	0.41	0.43
et al.e1-2			4,4,1,1,1,1,1	0.38	0.42
			5,2,2,1,1,1,1	0.38	0.41
			4,3,2,1,1,1,1	0.36	0.41
			3,3,3,1,1,1,1	0.31	0.38
			4,2,2,2,1,1,1	0.31	0.36
			3,3,2,2,1,1,1	0.28	0.35

	3,2,2,2,2,1,1	0.21	0.26
	2,2,2,2,2,2,1	0.08	0.12

$G^{(2)}$ is the second order Gini index and is used to distinguish between configurations that give the same Gini index G. It appears that Gini indexes are suitable to distinguish between *unbalanced designs*, like those with margins $m_A = m_B = (4,1,1,1)$ or $m_A = m_B = (7,1,1,1,1,1,1)$ and *more balanced* designs, like those with margins $m_A = m_B = (2,2,2,1)$ or $m_A = m_B = (2,2,2,2,2,2,1)$, respectively.

5.2 State polytopes

We introduce this topic by a couple of simple examples. Let us consider two factors, both with 3 levels. Let us make the hypothesis that they are quantitative variables x and y, so that it makes sense to consider polynomial models, with a different parametrization with respect to our previous setting. In particular, we consider the hierarchical model

$$\mathbb{E}(Y) = \mu + \alpha_{1,0}\, x + \alpha_{2,0}\, x^2 + \alpha_{0,1}\, y + \alpha_{0,2}\, y^2 \qquad (5.1)$$

and the fraction $\mathcal{F}_1 = \{(1,1),(1,2),(2,2),(3,2),(3,3)\}$. The corresponding design matrix $X_1^{(1)}$ is

$$\mathcal{F}_1 = \begin{matrix}(1,1)\\(1,2)\\(2,2)\\(3,2)\\(3,3)\end{matrix} \Rightarrow X^{(1)}_{\mathcal{F}1} = \begin{pmatrix} 1 & x & x^2 & y & y^2 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 4 \\ 1 & 2 & 4 & 2 & 4 \\ 1 & 3 & 9 & 2 & 4 \\ 1 & 3 & 9 & 3 & 9 \end{pmatrix}$$

We observe that the fraction $\mathcal{F}_1$ with respect to the model (5.1) is saturated because the determinant of the associated design matrix $\det(X^{(1)}_{\mathcal{F}1})$ is different from zero.

Let us now consider the same fraction $\mathcal{F}_1$ but with a different hierarchical model:

$$\mathbb{E}(Y) = \mu + \alpha_{1,0}\, x + \alpha_{2,0}\, x^2 + \alpha_{0,1}\, y + \alpha_{1,1}\, xy. \qquad (5.2)$$

The fraction $\mathcal{F}_1$ is saturated also for this new model (5.2) because $\det(X^{(2)}_{\mathcal{F}1}) \neq 0$.

Let us now consider a different design $_2 = \{(1,1),(1,2),(1,3),(2,1),(3,1)\}$ with the model (5.1). We get

$$\mathcal{F}_2 = \begin{matrix}(1,1)\\(1,2)\\(1,3)\\(2,1)\\(3,1)\end{matrix} \Rightarrow X^{(1)}_{\mathcal{F}2} = \begin{pmatrix} 1 & x & x^2 & y & y^2 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 4 \\ 1 & 1 & 1 & 3 & 9 \\ 1 & 2 & 4 & 1 & 1 \\ 1 & 3 & 9 & 1 & 1 \end{pmatrix}$$

The fraction $\mathcal{F}_2$ is saturated with respect to the model (5.1) because $\det(X^{(1)}_{\mathcal{F}_2}) \neq 0$.

Finally we consider the same design $\mathcal{F}_2$ but with the model (5.2). In this case we obtain that $\mathcal{F}_2$ is not a saturated design with respect to the model (5.2) because $\det(X^{(2)}_{\mathcal{F}_2}) = 0$.

This example shows that saturated designs are different with respect to the hierarchical polynomials models that they allow to estimate. The fraction $\mathcal{F}_1$ looks richer than the fraction $\mathcal{F}_2$ because it allows to estimate both the models (5.1) and (5.2). State polytopes are an efficient tool to deal with this kind of comparison. In qualitative terms state polytopes allow to associate to each design a polytope that contains all the polynomial models that can be estimated using it. The description of this methodology is out of the scope of this paper. We invite the interested reader to refer to the literature, Sturmfels (1996) , Berstein et al. (2010) Berstein, Maruri-Aguilar et al. (2012).

Here we give the main concepts to understand the two-factor case. Consider a hierarchical polynomial model

$$\mathbb{E}(Y) = \sum_{(i,j)\in L_{1,2}\subset L_1\times L_2} \alpha_{i,j}\, x^i y^j$$

where $L_1 = \{0,\ldots,I-1\}$, $L_2 = \{0,\ldots,J-1\}$ and $L_{1,2} \subset L_1 \times L_2$ is such that if $(i,j) \in L_{1,2}$ then $(i',j') \in L_{1,2}$ for all $i' \le i$, $j' \le j$. Let $h = \sum_{(i,j)\in L_{1,2}} i$ and $k = \sum_{(i,j)\in L_{1,2}} j$ be the sums of the variable degrees appearing in the model. For instance, when $I = J = 4$, for the saturated model without interactions, $\mathbb{E}(Y) = \mu + \alpha_{1,0}\, x + \alpha_{2,0}\, x^2 + \alpha_{3,0}\, x^3 + \alpha_{0,1}\, y + \alpha_{0,2}\, y^2 + \alpha_{0,3}\, y^3$, we have $h = k = 6$.

Given a fraction, for each permutation of the projections we can compute the points (h,k) pertaining to all hierarchical estimable models. The convex hull of such points is the *state polytope* of the fraction. All hierarchical models in the state polytope are estimable. The Minkowsky sum of the state polytope and the positive quadrant $\mathbb{R}_+ \times \mathbb{R}_+$ is named *state polyhedron* of the fraction.

In Figure 5 we report the state polyhedra corresponding to the six 4×4 non-equivalent saturated fractions coming from our description of the no interaction model in Section 2; the one-factor projections, $(4,1,1,1)$, $(3,2,1,1)$ and $(2,2,2,1)$, are displayed in the upper-right corner. The red points (h,k) represent saturated models. Notice that the point $(6,6)$, displayed in green, represent an estimable model in all the six configurations.

It is evident that the bottom-right design, whose margins are both $(2,2,2,1)$ is richer, in terms of models that can be estimated, than the top-left designs, whose margins are both $(4,1,1,1)$.

When I and J increase, the number of configurations increases as well and the complete study becomes complicated. As an example, in Figure 6 we report the state polyhedra corresponding to two 7×7 saturated designs.

6 Conclusions

In this paper we studied two factor saturated designs for main effect models from different perspectives. The results that have been obtained suggest several new directions and applications.

In the first part of this work, we provided the algebraic characterization of saturated designs, basing on *k*-cycles. Using such characterization, we proved a formula to compute the number of designs with given margins. As a consequence, we derived an explicit expression for the number of all the saturated designs. The extension to *m*-factor designs, $m \geq 2$ for models with main effects and interactions is under study. It is evident that such extension to any kind of saturated designs have a strong impact on the statistical practice. The characterization of *m*-factor designs in terms of cycles is described in Fontana et al. (2013).

Finally we studied the classification of saturated designs. Given a list of fractions, we showed how the Gini indexes of their univariate margins could be useful to highlight the most balanced designs. The connections between our classification and other classical statistical criteria, like *D*-optimality and minimum aberration, have to be explored. *D*-optimality for two factor designs has been already studied in Mukerjee et al. (1986). Minimum aberration is a classical criterion in Design of experiments and the reader can refer to Fries and Hunter (1980) and Li et al. (2003) for the basic definitions and results.

The classification of saturated designs with respect to their univariate margins provides a set of non-isomorphic designs. Their use for nonparametric statistical testing, analogously to Basso et al. (2004) and Arboretti Giancristofaro et al. (2012) for orthogonal fractions, should be evaluated.

Acknowledgment

The authors would like to thank Professor Bernd Sturmfels (U.C. Berkeley) for his useful suggestions. RF acknowledges SAS institute for providing software. FR is partially supported by the PRIN2009 grant number 2009H8WPX5.

References

Aoki, S., Takemura, A., 2010. Markov chain Monte Carlo tests for designed experiments. J. Statist. Plann. Inference 140 (3), 817–830.

Aoki, S., Takemura, A., 2012. Design and analysis of fractional factorial experiments from the viewpoint of computational algebraic statistics. J. Stat. Theory Pract. 6 (1), 147–161.

Arboretti Giancristofaro, R., Fontana, R., Ragazzi, S., 2012. Construction and nonparametric testing of orthogonal arrays through algebraic strata and inequivalent permutation matrices. Comm. Statist. Theory Methods 41 (16-17).

Bailey, R. A., 2008. Design of comparative experiments. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge.

Basso, D., Salmaso, L., Evangelaras, H., Koukouvinos, C., 2004. Nonparametric testing for main effects on inequivalent designs. In: mODa 7—Advances in model-oriented design and analysis. Contrib. Statist. Physica, Heidelberg, pp. 33–40.

Berstein, Y., Maruri-Aguilar, H., Onn, S., Riccomagno, E., Wynn, H., 2010. Minimal average degree aberration and the state polytope for experimental designs. Ann. Inst. Statist. Math. 62 (4), 673–698.

Dey, A., Shah, K. R., Das, A., 1995. Optimal block designs with minimal and nearly minimal number of units. Statist. Sinica 5 (2), 547–558.

Drton, M., Sturmfels, B., Sullivant, S., 2009. Lectures on Algebraic Statistics. Birkhauser, Basel.

Fontana, R., 2013. Algebraic generation of minimum size orthogonal fractional factorial designs: an approach based on integer linear programming. Comput. Statist., 1–13Online first, doi:10.1007/s00180-011-0296-7.

Fontana, R., Pistone, G., Rogantin, M. P., 2000. Classification of two-level factorial fractions. J. Statist. Plann. Inference 87 (1), 149–172.

Fontana, R., Rapallo, F., Rogantin, M. P., 2012. Markov bases for sudoku grids. In: Di Ciaccio, A., Coli, M., Angulo Ibanez, J. M. (Eds.), Advanced Statistical Methods for the Analysis of Large Data-Sets. Studies in Theoretical and Applied Statistics. Springer, Berlin, pp. 305–315.

Fontana, R., Rapallo, F., Rogantin, M. P., 2013. A characterization of saturated designs for factorial experiments. arXiv:1304.7914v1.

Fries, A., Hunter, W. G., 1980. Minimum aberration 2^{k-p} designs. Technometrics 22 (4), 601–608.

Gini, C., 1912. Variabilit`a e mutabilit`a: contributo allo studio delle distribuzioni e delle relazioni statistiche. Tipogr. di P. Cuppini.

Hassani, M., 2003. Derangements and applications. J. Integer Seq. 6 (1), Article 03.1.2, 8 pp. (electronic).

Hedayat, A. S., Sloane, N. J. A., Stufken, J., 1999. Orthogonal arrays. Theory and applications. Springer Series in Statistics. Springer-Verlag, New York.

Krafft, O., Schaefer, M., 1997. A-optimal connected block designs with nearly minimal number of observations. J. Statist. Plann. Inference 65 (2), 375–386.

Kuhnt, S., Rapallo, F., Rehage, A., 2013. Outlier detection in contingency tables based on minimal patterns. Stat. Comput.Online first, doi:10.1007/s11222-013-9382-8.

Li, W., Lin, D. K. J., Ye, K. Q., 2003. Optimal foldover plans for two-level nonregular orthogonal designs. Technometrics 45 (4), 347–351.

Maruri-Aguilar, H., S´aenz-de Cabez´on, E., Wynn, H. P., 2012. Betti numbers of polynomial hierarchical models for experimental designs. Ann. Math. Artif. Intell. 64 (4), 411–426.

Mukerjee, R., Chatterjee, K., Sen, M., 1986. D-optimality of a class of saturated main-effect plans and allied results. Statistics 17 (3), 349–355.

Pistone, G., Riccomagno, E., Wynn, H. P., 2001. Algebraic Statistics: Computational Commutative Algebra in Statistics. Chapman&Hall/CRC, Boca Raton.

Raktoe, B. L., Hedayat, A., Federer, W. T., 1981. Factorial designs. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York.

Rapallo, F., Rogantin, M. P., 2007. Markov chains on the reference set of contingency tables with upper bounds. Metron 65 (1), 35–51.

Sturmfels, B., 1996. Gr¨obner bases and convex polytopes. Vol. 8 of University lecture series (Providence, R.I.). American Mathematical Society.

Xu, K., 2003. How has the literature on Gini's index evolved in the past 80 years? Dalhousie University, Economics Working Paper.

Figure 1: The model matrix X of the full factorial 3×4 design and the model matrix $X_{\mathcal{F}}$ of the fraction in Example 1.

$$X = \begin{array}{c} \\ (1,1) \\ (1,2) \\ (1,3) \\ (1,4) \\ (2,1) \\ (2,2) \\ (2,3) \\ (2,4) \\ (3,1) \\ (3,2) \\ (3,3) \\ (3,4) \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & a_1 & a_2 & b_1 & b_2 & b_3 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \qquad X_{\mathcal{F}} = \begin{array}{c} \\ (1,1) \\ (1,2) \\ (2,2) \\ (2,3) \\ (3,3) \\ (3,4) \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & a_1 & a_2 & b_1 & b_2 & b_3 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

Figure 2: The complete bipartite graph for a 3×4 design.

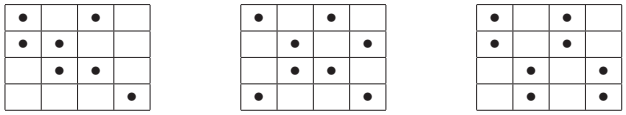

Figure 3: State polyhedra for six non-equivalent 4×4 designs.

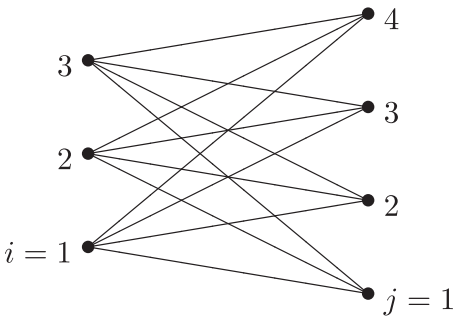

Figure 4: State polyhedra for two non-equivalent 7×7 designs.

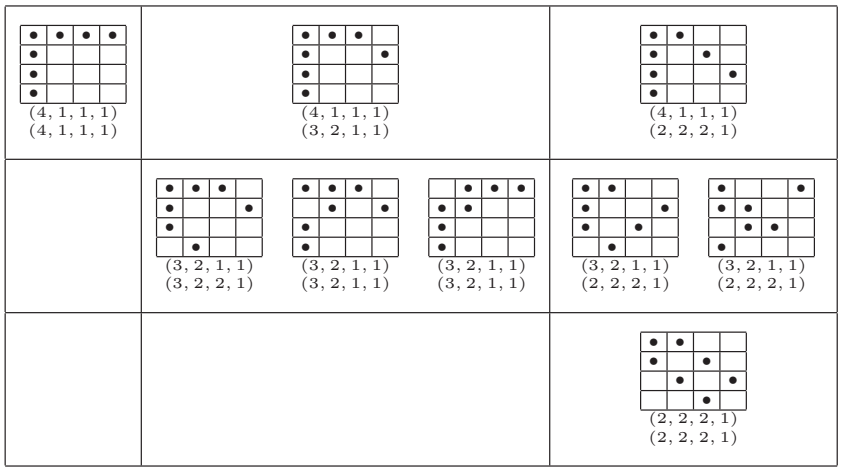

Figure 5: State polyhedra for six non-equivalent 4 × 4 designs.

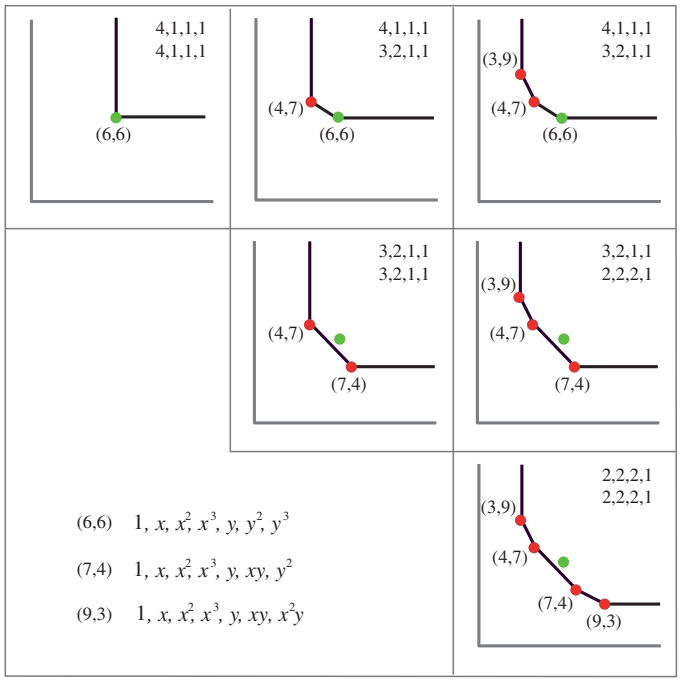

Figure 6: State polyhedra for two non-equivalent 7 × 7 designs.