Searching for D-efficient Equivalent-Estimation Second-Order Split-Plot Designs

Nam-Ky Nguyen [*]& Tung-Dinh Pham [†]

August 4, 2014

Abstract

Several industrial experiments are set up using second-order split-plot designs (SPDs). These experiments have two types of factors: whole-plot (WP) factors and sub-plot (SP) factors. WP factors, also called hard-to-change factors are factors whose levels are hard or expensive to change. SP factors, also called easy-to-change factors are factors whose levels are easy or less expensive to change. In a split-plot experiment, the WP factors are confounded with blocks. Certain SPDs possess the equivalent-estimation property. For SPDs with this property, ordinary least-squares estimates of the model parameters are equivalent to the generalized least-squares estimates.

This paper describes a fast and simple algorithm which produce D-efficient equivalent-estimation SPDs by interchanging the levels of the SP factors within each WP. The performance of this algorithm is evaluated against the 111 SPD scenarios reported in Macharia & Goos (2010) and Jones & Goos (2012).

Keywords: Box-Behnken designs; Central-composite designs; D-optimality; Interchange-algorithm; Response surface model.

[*]Vietnam Institute for Advanced Study in Mathematics & VNU International School, Hanoi, Vietnam; e-mail: nknam@viasm.edu.vn.

[†]VNU University of Science, Hanoi, Vietnam; e-mail: tungpd@vnu.edu.vn.

1 Introduction

Split-plot designs were originally developed in the early 1920's by R. A. Fisher and first applied to agricultural experiments where factors are mostly qualitative. In recent years, split-plot designs have been applied to industrial experiments where factors are mostly quantitative and the second-order response surface model is used. In this paper, we simply call these second-order split-plot designs SPDs.

The following is an example of an experiment set up using an SPD. The aim of this experiment is to optimize the productivity and quality of lima beans as affected by plant density and NPK applications. In this experiment, plant density (8, 12 and 16 plants/m^2) is the WP factor and N (110, 150 and 190 kg/ha), P (50, 70 and 90 kg/ha) and K (160, 180 and 200 kg/ha) are SP factors. Each WP is sub-divided into six SPs and each SP receives one out of 3^3 NPK combinations. Additional examples of the use and analysis of SPDs in the industrial settings can be found in Vining et al. (2005), Kowalski et al. (2007) and Jones & Nachtsheim (2009).

In the recent SPD literature, one type of SPD which received considerable attention is the equivalent-estimation second-order SPD (hereafter abbreviated as EE-SPD) pioneered by the work of Vining et al. (2005). This type of design derived its name from the fact that the ordinary least-squares (OLS) estimates of the model parameters are equivalent to the generalized least-squares (GLS) ones. Therefore, the computation of these estimates does not require estimation of the variance components and the estimates of this type of design can be done by any multiple regression program. Macharia & Goos (2010) associated EE-SPDs to orthogonally blocked designs, for which the OLS estimates are equivalent to the GLS and intra-block ones.

The first EE-SPDs were introduced by Vining et al. (2005). These EE-SPDs are based on Box-Behnken designs or BBDs (Box & Behnken, 1960) and central-composite designs or CCDs (Box & Wilson, 1951). This work is extended further by Parker et al. (2006, 2007). Table 1 displays a BBD-based SPD and a CCD-based SPD for one WP factor and two SP factors in three WPs of size five.

Table 1: Two EE-SPDs for one WP factor and two SP factors in WPs of size five†

BBD-based w	s_1	s_2	CCD-based w	s_1	s_2
-1	-1	0	-1	-1	-1
-1	1	0	-1	-1	1
-1	0	-1	-1	1	-1
-1	0	1	-1	1	1
-1	0	0	-1	0	0
1	-1	0	1	-1	-1
1	1	0	1	-1	1
1	0	-1	1	1	-1
1	0	1	1	1	1
1	0	0	1	0	0
0	-1	-1	0	-1	0
0	-1	1	0	1	0
0	1	-1	0	0	-1
0	1	1	0	0	1
0	0	0	0	0	0

†WPs are separated by a blank line.

Although all BBD- and CCD-based SPDs are EE-SPD, they might not be attractive to experimenters as they are highly inefficient (see Goos (2006)). Macharia & Goos (2010) explored the relationship between D-optimal SPDs and EE-SPDs for the second-order response surface model and proposed the first algorithm for generating D-efficient EE-SPDs for a flexible choice of the number of WP and SP factors, the number of WPs and the run size of each WP. For 86 out of 111 scenarios they studied, their algorithm was able to produce EE-SPDs. Jones & Goos (2012) (hereafter abbreviated as JG) recently reported on a more successful algorithm. They list 60 scenarios out of the mentioned 111 scenarios for which no EE-SPDs had been found by Macharia & Goos (2010) or where they found more D-efficient EE-SPDs.

During the revision of this paper, we came across the work of Mylona et al. (2013) which introduced two families of EE-SPDs, one based on subset designs and the other based on supplementary set designs.

The aim of this paper is to introduce a fast and simple algorithm for searching EE-SPDs which are not only D-efficient but also more appealing to the experimenters. Section 2 reviews the general SPD model. Section 3 proposes a desired structure for the information matrices of SPDs. Section 4 outlines the new algorithm and Section 5 evaluates the performance of the algorithm against the 111 mentioned scenarios.

2 The General Split-Plot Design model

The general model for data from a split-plot experiment with m_W WP factors, m_S SP factors in b WPs of run size k ($bk = n$) is given by

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}, \tag{1}$$

where $\mathbf{Y}_{n\times 1}$ is the vector of responses, $\mathbf{X}_{n\times p}$ is the expanded design matrix of the second-order model for m ($= m_W + m_S$) factors, $p = (m+1)(m+2)/2$ is the number of parameters in (1), $\boldsymbol{\beta}_{p\times 1}$ is the coefficient vector containing p fixed effects. $\mathbf{Z}_{n\times b}$ is the (0,1)-matrix containing b dummy variables which are associated with the b WPs, $\boldsymbol{\gamma}_{b\times 1}$ is the vector containing the b random effects of the b WPs, and finally, $\boldsymbol{\epsilon}_{n\times 1}$ is the vector of the random errors. It is assumed that $\boldsymbol{\gamma}$ and $\boldsymbol{\epsilon}$ are uncorrelated, have zero mean and variance-covariance matrices $\sigma_\gamma^2\mathbf{I}_b$ and $\sigma_\epsilon^2\mathbf{I}_n$, respectively. Here, $\mathbf{I}_b$ and $\mathbf{I}_n$ are identity matrices of sizes b and n respectively. The assumed variance-covariance matrix of the model is thus of the form

$$\mathbf{V} = \sigma_\epsilon^2\mathbf{I}_n + \sigma_\gamma^2\mathbf{Z}\mathbf{Z}' = \sigma_\epsilon^2\mathbf{I}_n + \sigma_\gamma^2\mathbf{J}_b, \tag{2}$$

where

$$\mathbf{J}_b = \begin{pmatrix} \mathbf{J}_{k\times k} & \mathbf{0}_{k\times k} & \cdots & \mathbf{0}_{k\times k} \\ \mathbf{0}_{k\times k} & \mathbf{J}_{k\times k} & \cdots & \mathbf{0}_{k\times k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{k\times k} & \mathbf{0}_{k\times k} & \cdot & \mathbf{J}_{k\times k} \end{pmatrix}.$$

Here, $\mathbf{J}_{k\times k}$ is a $k \times k$ matrix of ones and $\mathbf{0}_{k\times k}$ is an $k \times k$ zero matrix.

An EE-SPD is an SPD whose GLS and OLS estimates are the same, i.e. $\hat{\boldsymbol{\beta}}_{\text{GLS}} = \hat{\boldsymbol{\beta}}_{\text{OLS}}$ where $\hat{\boldsymbol{\beta}}_{\text{GLS}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$ and $\hat{\boldsymbol{\beta}}_{\text{OLS}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$. Parker et al. (2007) showed that the following condition must hold for the equivalence of GLS and OLS estimates in the SPD setting:

$$\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{J}_b\mathbf{X} = \mathbf{J}_b\mathbf{X}. \tag{3}$$

JG remarked that the condition in (3) is the same as trace($\mathbf{C}'\mathbf{C}$) $= 0$ where

$$\mathbf{C} = (\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{J}_b\mathbf{X}. \tag{4}$$

Let $\mathbf{B} = \mathbf{X}'\mathbf{J}_b\mathbf{X}$ and note that $\mathbf{A} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ is an idempotent matrix, i.e. $\mathbf{A}^2 = \mathbf{A}$, it is not difficult to show that

$$\mathbf{C}'\mathbf{C} = k\mathbf{B} - \mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B} \tag{5}$$

and (5) can be used to speed up the calculation of trace($\mathbf{C}'\mathbf{C}$).

3 A Desired Structure for the Information Matrices of SPDs

Let $\mathbf{x}'_i$ be the i-th row of $\mathbf{X}$ in (1) and written as $(1, x_{i1}^2, x_{i2}^2, \ldots,\ x_{i1},\ x_{i2}, \ldots,\ x_{i1}x_{i2},\ x_{i1}x_{i3}, \ldots)$ and partition the information matrix $\mathbf{M} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ as

$$\left(\begin{array}{c|c} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \hline \mathbf{M}_{21} & \mathbf{M}_{22} \end{array}\right). \tag{6}$$

where $\mathbf{M}_{11}$ is a $(1 + m + m_W) \times (1 + m + m_W)$ sub-matrix of $\mathbf{M}$. Assume that (6) can be written as

$$\left(\begin{array}{c|c} \mathbf{M}_{11} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{D} \end{array}\right). \tag{7}$$

where $\mathbf{D}$ is a diagonal matrix. We denote any SPD whose information matrix is of the form (7) SPD*. It can be seen from (7) that for SPD*'s, m_S SP main effects and $\binom{m}{2}$ interactions can be estimated orthogonally. Note that in (7) we do not require that all elements corresponding to correlations between the intercept estimate and the estimates of the quadratic effects in be equal. Also, we do not require that every pair of quadratic effect estimates has the same correlation. This is different from CCDs or BBDs, or the designs of

Großmann & Gilmour (2013).

In this paper, we aim to develop an algorithm which searches for D-efficient EE-SPDs among the class of SPDs whose information matrices have the structure in (7) or are close to this structure.

4 The SPLIT algorithm

Before introducing SPLIT, our algorithm for searching D-efficient EE-SPDs with m_W WP factors, m_S SP factors in b WPs of run size k, we show how the information matrix $\mathbf{M} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ can be computed sequentially and updated when a WP is removed from or added to an SPD.

Following Macharia & Goos (2010) and JG, assume that $\sigma_\gamma^2 = \sigma_\epsilon^2 = 1$. Thus, $\mathbf{V}$ in (2) becomes $\mathbf{I}_n + \mathbf{J}_b$ and $\mathbf{V}^{-1}$ becomes $\mathbf{I}_n - \frac{1}{k+1}\mathbf{J}_b$. Let $\mathbf{A} = \mathbf{X}'\mathbf{X}$. Matrix $\mathbf{A}$ can be computed sequentially as $\sum_r^b \mathbf{A}_r$ where $\mathbf{A}_r = \sum_i^k \mathbf{x}_i\mathbf{x}_i'$ and $\mathbf{x}_i'$ is the i-th row of $\mathbf{X}_r$, the $k \times p$ sub-matrix of $\mathbf{X}$ associated with WP r. Similarly, matrix $\mathbf{B} = \mathbf{X}'\mathbf{J}_b\mathbf{X}$ first used in (5) can be computed sequentially as $\sum_r^b \mathbf{B}_r$ where $\mathbf{B}_r = \mathbf{w}_r\mathbf{w}_r'$ and $\mathbf{w}_r'$ is the $1 \times p$ vector containing the sum of each column of $\mathbf{X}_r$. Thus $\mathbf{M}$ can be computed sequentially as

$$\mathbf{M} = \sum_{r=1}^{b}(\mathbf{A}_r - \frac{1}{k+1}\mathbf{B}_r) \tag{8}$$

If the WP r is removed from or added to an SPD, $\mathbf{M}$ can simply be updated as

$$\mathbf{M}_{\text{updated}} = \mathbf{M} \mp (\mathbf{A}_r - \frac{1}{k+1}\mathbf{B}_r). \tag{9}$$

The update formulas such as the one for $\mathbf{M}$ in (9) is crucial in any design construction program. Without the update formulas, matrices have to be recomputed from scratch each time a change is made in the design matrix. See Arnouts & Goos (2010) for additional examples of update formulas in the context of split-plot and block designs.

Our SPLIT algorithm based on the previous matrix results involves the following steps:

1. Construct D_0, the $n \times m$ *input* design matrix with m_W WP factors

and m_S SP factors in n runs. Details of the method for constructing D_0 is in Remark 1.

2. Construct D, the $n \times m$ the *starting* design matrix by randomizing the levels of each SP factor in each WP of D_0. Calculate $\mathbf{M}$ using (8) and the objective function f which is the sum of squares of the elements in $\mathbf{M}_{12}$ and the upper-diagonal elements of $\mathbf{M}_{22}$ (See eq. (6)).

3. Remove WP r, $r = 1, \ldots, b$, from D and update $\mathbf{M}$ using (9). Among SP factors j, $j = 1 \ldots, m_S$ in the removed WP, search for a pair of levels for which the swap or interchange of these two levels in this WP will result in the biggest reduction in f when this WP is returned to D. If the search is successful, update f, D, and $\mathbf{M}$ using (9). If f cannot be further reduced, repeat this step with the next WP.

Step 3 is repeated until $f = 0$ (i.e. D becomes an SPD*) or f cannot be reduced further by any further level-swaps.

Remarks:

1. The main difference between the JG's EE-SPDs and SPLIT's is that in the latter, the setting of the WP factor levels and the distribution of the levels of each SP factor in each WP are made by the experimenter and not the computer and these tasks are accomplished via the input design. The construction of input designs sometimes involves trial and error and this appears to be a drawback of SPLIT. At the same time, the experimenter does have more control over his/her experiment.

When there is a single WP factor, the setting of the levels of this WP factor in each WP is straight forward. When there are two WP factors, depending on the number of WPs, we form each WP by replicating each of the following points k times: (i) $b = 7$: a 2^2 factorial plus three points (1, 0), (0, 1) and (-1, -1) or (0, 0); (ii) $b = 8$: a 2-factor CCD without a center point, i.e. a 2^2 factorial plus four axial points (0, -1), (0, 1), (-1, 0), (1, 0); (iii) $b = 9$: a 2-factor CCD with one center point, i.e. the points for $b = 8$ plus (0, 0); (iv) $b = 10$: the points for $b = 8$ plus two points (-1, 1) and (1, -1) or two center points; (v) $b = 11$: the points for $b = 8$ plus three points (-1, 1), (1, -1) and (0, 0); and (vi) $b = 12$: the points for $b = 8$ plus an additional 2^2

factorial. For three WP factors in 12 WPs such as the scenario 111 we use a 2^3 factorial plus (1, 0, 0), (0, 1, 0), (0, 0, 1) and (-1, -1, -1) or (0, 0, 0). This recommended method of setting the levels of the WP factors in each WP ensures that the main effects of each WP factor and their interactions are orthogonal or near-orthogonal to one another and to the quadratic effect of each WP factor.

Depending on the size of each WP, the distribution of the levels of each SP factor in each WP are made as follows: (i) $k = 2$: either -1 and 1 or 0 and 0; (ii) $k = 3$: -1, 0, 1; (iii) $k = 4$: either -1, -1, 1, 1 or -1, 1, 0, 0; (iv) $k = 5$: -1, -1, 0, 1, 1; and (v) $k = 6$: -1, -1, 0, 0, 1, 1. This distribution of SP levels in each WP ensures that the sum of the levels of each SP factor in each WP is zero. As a result, the main effects of each SP factor and the interactions between this SP factor and a WP factor are orthogonal to the mean and to the main and quadratic effects of each WP factor.

2. Steps 2 and 3 make up a *try* and at least one thousand tries might be required to find an EE-SPD. At the end of each try, we compute trace($\mathbf{C}'\mathbf{C}$) using the following formula:

$$\text{trace}(\mathbf{C}'\mathbf{C}) = k\sum_{i}^{p}\mathbf{B}_{ii} - \sum_{l}^{p}\sum_{i}^{p}\sum_{j}^{p}\mathbf{A}_{ij}^{-1}\mathbf{B}_{il}\mathbf{B}_{jl} \qquad (10)$$

where $\mathbf{A}_{ij}^{-1}$ and $\mathbf{B}_{ij}$ are the (i, j)-th elements of $\mathbf{A}^{-1}$ and $\mathbf{B}$ respectively (see eq. (5)). Among all tries which produce EE-SPDs, i.e. SPDs with trace($\mathbf{C}'\mathbf{C}$) = 0, the one with the highest value of $|\mathbf{M}|$ is selected. The D-efficiency of this EE-SPD is computed as

$$D_{\text{eff}} = (|\mathbf{M}|/|\mathbf{M}_{\text{opt}}|)^{1/p}, \qquad (11)$$

where $|\mathbf{M}_{\text{opt}}|$ is the determinant of the information matrix of the D-optimal design using model (1) and can be obtained from JG's supplementary material. Note that this D-optimal design is not restricted to 3-level.

3. The SPLIT algorithm is an example of the *interchange* algorithm, used extensively in different design settings, e.g. super-saturated designs (see e.g. Nguyen, 1996 and Trinca & Gilmour 2000, 2002). The algorithms of Jones &

Goos (2007) for constructing D-optimal SPDs, Macharia & Goos (2010) and JG for constructing EE-SPDs (and Jones & Nachtsheim (2011) and Nguyen & Stylianou (2013)) for constructing definitive screening designs) are examples of the *coordinate-exchange* algorithm (see Meyer & Nachtsheim, 1995). This algorithm allows D_{ij}, the (i, j)-th element of D to take an integer value in the range $[-1,\ 1]$ in the case of Macharia & Goos (2010) or a value which can be an integer or non-integer in the range $[-1,\ 1]$ in the case of JG. The algorithms of Goos & Vandebroek (2001, 2003, 2004) for constructing D-optimal SPDs are examples of the *point-exchange* algorithm. Unlike the interchange and coordinate-exchange algorithms, this algorithm requires a *candidate set*, whose construction can be problematic when the number of factors is large.

Table 2: Steps of SPLIT to produce an EE-SPD for scenario 25†
(one WP factor and two SP factors in five WPs of size three)

(1)			(2)			(3)			(3′)		
w	s_1	s_2	w	s_1	s_2	w	s_1	s_2	w	s_1	s_2
1	1	1	1	0	0	1	0	$\underline{-1}$	1	0	-1
1	0	0	1	-1	-1	1	-1	$\underline{0}$	1	-1	$\underline{1}$
1	-1	-1	1	1	1	1	1	1	1	1	$\underline{0}$
1	1	1	1	-1	0	1	-1	0	1	-1	0
1	0	0	1	1	-1	1	1	-1	1	1	-1
1	-1	-1	1	0	1	1	0	1	1	0	1
-1	1	1	-1	0	-1	-1	0	-1	-1	0	-1
-1	0	0	-1	-1	1	-1	-1	1	-1	-1	1
-1	-1	-1	-1	1	0	-1	1	0	-1	1	0
-1	1	1	-1	-1	0	-1	-1	0	-1	-1	0
-1	0	0	-1	1	-1	-1	1	-1	-1	1	-1
-1	-1	-1	-1	0	1	-1	0	1	-1	0	1
0	1	1	0	0	0	0	0	0	0	0	0
0	-1	-1	0	1	1	0	1	1	0	1	1
0	0	0	0	-1	-1	0	-1	-1	0	-1	-1

†See Table 5 of JG. WPs are separated by a blank line.
‡Swapped levels in columns (3) and (3′) are underlined.

Table 2 shows the steps of a single try of SPLIT in constructing an SPD with one WP factor and two SP factors in five WPs of size two (scenario 25 of Table 5 of JG). Column (1) of this Table which corresponds to Step 1 shows D_0. Column (2) which corresponds to Step 2 shows D, i.e. D_0 after randomization. At this step, the f value is 29.1875. Columns (3) and (3′) correspond to two iterations of Step 3 after which f has been reduced to 28.5 and 23.25 respectively. Since f cannot be reduced further, trace($\mathbf{C}'\mathbf{C}$) is computed. This value turns out to be 0 showing that the constructed design in column (3′) is in fact an EE-SPD. The D-efficiency of this EE-SPD is 91.7% which indicates a substantial gain in efficiency over JG's which had a D-efficiency of only 80.5%

5 Results and Discussion

This Section presents some proof-of-concept examples documenting the evidence of the potential usefulness of our method of constructing EE-SPDs. This is to be followed by detailed computational results.

5.1 Proof-of-Concept Examples

Table 3 shows a CCD-based EE-SPD* of Vining et al. (2005) in 12 WPs of size four for a split-plot experiment on the strength of ceramic pipe with the temperatures in zone 1 and zone 2 of the furnace as WP factors and the amount of binder in the formulation and the grinding speed of the batch as SP factors. Out of 48 runs, this design involves 24 runs at the zero level of each WP factor and 30 runs at the zero level of each SP factor. The D-efficiency of this EE-SPD* relative to the D-optimal SPD in Table 10 of Jones & Nachtsheim (2009) and reproduced in Table 3 is 58.2%. The second design in Table 3 is an alternative EE-SPD* constructed by SPLIT. Each factor of the SPLIT design is set to zero only eight times. The D-efficiency of the SPLIT design relative to the D-optimal design by Jones & Nachtsheim (2009) is 88.9%. It is worth mentioning that the settings of the WP factor levels in each WP of the SPLIT design and the D-optimal one are identical.

Computer-generated EE-SPDs were reported by Macharia & Goos (2010) and JG to provide experimenters with more flexible and D-efficient EE-SPDs. The algorithms used by these authors to produce EE-SPDs are well documented in their papers. The improvement in D-efficiencies of new EE-SPDs

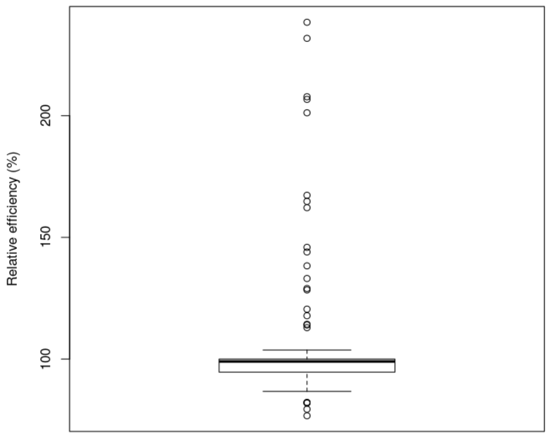

Figure 1: Box-plot containing the relative D-efficiencies of the SPLIT designs and JG's designs.

over JG's is most noticeable when the latter take non-integer values for factor levels. Tables 4-6 show JG's EE-SPDs and ours for scenarios 48, 94 and 109 in Table 5 of JG. For these three scenarios, the D-efficiencies of JG designs are 37.8, 44.5 and 44.9% and ours are 90.2, 92.5 and 90.3% respectively. Note that our designs for scenarios 48 and 109 are also EE-SPD*'s. The new design for scenario 48 could be a good choice for the lima bean experiment mentioned in Section 1.

5.2 Detailed Computational Results

JG managed to find the EE-SPD solutions for all 111 scenarios. However, 24 of these solutions involve non-integer values for factor levels (11 are listed in Table 5 of JG). We were able to find EE-SPD solutions for 105 scenarios. These scenarios include the 25 scenarios which JG failed to offer integer solutions. Five of the six scenarios for which we failed to find EE-SPD solutions involve WPs of size two.

To see how efficient our EE-SPDs are relative to JG's for each scenario, we compute the relative efficiency or RE (%) of each new EE-SPD which is the ratio of the D-efficiency of the new design to JG's. The box-plot for the RE of 105 new EE-SPDs is in Figure 1. There are two new EE-SPDs with REs less than 80% but usually the REs are above 90%. Table 7 lists 25 scenarios where the new EE-SPDs have a RE larger than 100%. The scenarios in this table are a subset of the 60 scenarios in Table 5 of JG for which no EE-SPDs had been found by Macharia & Goos (2010) or where more D-efficient EE-SPDs are found by JG. Note that 16 JG solutions in this table involve non-integer values.

A referee is keen to know how good the EE-SPD*'s are with respect to the D-optimality. All together, 34 solutions of ours are EE-SPD*'s: five are listed in Table 7, seven have D-efficiencies matching those of the corresponding JG designs (and it turns out that these JG designs are also EE-SPD*'s) and the rest have D-efficiencies slightly smaller than the corresponding JG designs. For certain scenarios, using different input designs, SPLIT produces different solutions. Starting from an input design with more zero-levels gives a higher chance to obtain an EE-SPD* and less efficient design.

Since SPLIT only involves a matrix inversion at the end of each try, it is a fairly fast algorithm. For scenario 25 (Table 2), SPLIT gets an EE-SPD instantaneously. For scenario 94 (Table 5) SPLIT finds about 10 EE-SPDs in 1000 tries which take less than a second on a laptop with Intel Core 2 Duo CPU T9500 @ 2.60GHz $\times$ 2 processor. For scenario 48 (Table 4), SPLIT finds about ten EE-SPDs in 100,000 tries which takes five minutes on the same laptop and only two out of these ten are EE-SPD*. For scenario 109 (Table 6) SPLIT gets 275 EE-SPDs in 100,000 tries which take about ten minutes on the same laptop but many more tries are required to obtain an EE-SPD*.

6 Concluding remarks

EE-SPDs have been discussed by Vining et al. (2005), Parker et al. (2006, 2007) or more recently by Macharia & Goos (2010) and JG. Readers, however, should note that D-optimal SPDs (particularly those that do not involve non-integer values) are preferable to EE-SPDs unless the EE-SPDs are also D-optimal or very close to D-optimal. Experimenters having access to decent software such as JMP, SAS or R, etc. which can handle GLS should not resort to EE-SPDs habitually as it is only the computation of point estimates which does not depend on the variance components. The computation of other statistics such as standard errors, confidence intervals, t-test, etc. all still depend on the variance components, for which a good software package is required.

This paper introduce a fast and simple algorithm called SPLIT for constructing 3-level EE-SPDs in most situations. The main difference between the EE-SPDs of JG and SPLIT's is that for the latter, the setting of the WP factor levels and the distribution of the levels of each SP factor in each WP require the input of the experimenter. SPLIT's EE-SPDs are often D-efficient and possess information matrices having a simple structure very close the one in (7). This leads to a simpler interpretation of the results.

The SPLIT algorithm was implemented in a Java program. This Java program as well as the input and output design matrices for the scenarios discussed in this paper are available from the first author.

Acknowledgement

The authors are grateful to the Editor and a referee for many helpful comments, suggestions and corrections which improved the performance of the algorithm as well as the readability of the paper.

Table 3: CCD-based EE-SPD*, SPLIT's EE-SPD* and D-optimal SPD for an experiment on the strength of ceramic pipe.†

CCD-based				SPLIT				D-optimal			
w_1	w_1	s_1	s_2	w_1	w_2	s_1	s_2	w_1	w_2	s_1	s_2
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	1	-1	-1	-1	1
-1	-1	-1	1	-1	-1	1	-1	-1	-1	1	-1
-1	-1	1	1	-1	-1	-1	1	-1	-1	1	1
-1	1	-1	-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	1	-1	-1	1	1	1	-1	1	-1	1
-1	1	-1	1	-1	1	-1	1	-1	1	1	-1
-1	1	1	1	-1	1	1	-1	-1	1	1	1
1	-1	-1	-1	1	-1	1	-1	1	-1	-1	-1
1	-1	1	-1	1	-1	1	1	1	-1	-1	1
1	-1	-1	1	1	-1	-1	-1	1	-1	1	-1
1	-1	1	1	1	-1	-1	1	1	-1	1	1
1	1	-1	-1	1	1	1	-1	1	1	-1	-1
1	1	1	-1	1	1	-1	-1	1	1	-1	1
1	1	-1	1	1	1	-1	1	1	1	1	-1
1	1	1	1	1	1	1	1	1	1	1	1
-1	0	0	0	-1	-1	0	-1	-1	-1	-1	1
-1	0	0	0	-1	-1	-1	0	-1	-1	0	-1
-1	0	0	0	-1	-1	1	0	-1	-1	1	0
-1	0	0	0	-1	-1	0	1	-1	-1	1	1
1	0	0	0	-1	1	0	-1	-1	1	-1	-1
1	0	0	0	-1	1	-1	0	-1	1	-1	1
1	0	0	0	-1	1	0	1	-1	1	0	0
1	0	0	0	-1	1	1	0	-1	1	1	-1
0	-1	0	0	1	-1	1	0	1	-1	-1	-1
0	-1	0	0	1	-1	0	1	1	-1	-1	1
0	-1	0	0	1	-1	-1	0	1	-1	0	1
0	-1	0	0	1	-1	0	-1	1	-1	1	0
0	1	0	0	1	1	1	0	1	1	-1	-1
0	1	0	0	1	1	-1	0	1	1	-1	1
0	1	0	0	1	1	0	1	1	1	0	-1
0	1	0	0	1	1	0	-1	1	1	1	0
0	0	-1	0	-1	0	-1	-1	-1	0	-1	-1
0	0	1	0	-1	0	-1	1	-1	0	-1	1
0	0	0	-1	-1	0	1	1	-1	0	0	0
0	0	0	1	-1	0	1	-1	-1	0	1	1
0	0	0	0	1	0	-1	-1	1	0	-1	0
0	0	0	0	1	0	1	1	1	0	0	-1
0	0	0	0	1	0	1	-1	1	0	1	-1
0	0	0	0	1	0	-1	1	1	0	1	1
0	0	0	0	0	-1	1	-1	0	-1	-1	-1
0	0	0	0	0	-1	1	1	0	-1	-1	0
0	0	0	0	0	-1	-1	-1	0	-1	0	1
0	0	0	0	0	-1	-1	1	0	-1	1	-1
0	0	0	0	0	1	1	-1	0	1	-1	0
0	0	0	0	0	1	1	1	0	1	0	1
0	0	0	0	0	1	-1	-1	0	1	1	-1
0	0	0	0	0	1	-1	1	0	1	1	1

†See Vining et al. (2005). WPs are separated by a blank line.

Table 4: JG's EE-SPD and SPLIT's EE-SPD* for scenario 48 (one WP factor and three SP factors in six WPs of size six).†

JG				SPLIT			
w_1	s_1	s_2	s_3	w_1	s_1	s_2	s_3
0	0.5391	-0.7994	-0.9241	0	0	0	0
0	0.7618	0.8948	0.8876	0	1	1	1
0	0.8512	0.4249	-0.5734	0	0	0	0
0	-0.6599	0.6393	0.6646	0	1	-1	-1
0	-0.9426	0.887	0.2522	0	-1	1	-1
0	0.9664	-0.3771	-0.0301	0	-1	-1	1
-1	0.4785	0.8823	-0.2468	-1	-1	1	1
-1	0.5373	-0.7627	0.1032	-1	1	1	-1
-1	0.3261	0.1672	0.8003	-1	1	0	1
-1	-0.819	-0.7452	-0.7282	-1	0	-1	0
-1	-1	-0.5025	1	-1	-1	0	-1
-1	-0.9274	0.9008	-0.6651	-1	0	-1	0
1	0.5022	-0.1115	-0.7528	1	1	1	-1
1	-0.1962	0.4626	0.7031	1	0	-1	-1
1	-0.6058	0.8575	-0.6065	1	0	1	1
1	1	-0.0288	0.8948	1	1	-1	1
1	1	0.0074	0.9046	1	-1	0	0
1	0.0892	0.3581	1	1	-1	0	0
-1	-0.5807	-0.9361	-0.3964	-1	1	-1	1
-1	-0.1699	1	0.1213	-1	0	1	0
-1	-0.9472	-0.1092	-0.6196	-1	1	0	-1
-1	-0.9603	-0.2004	0.9879	-1	-1	0	1
-1	0.9297	-0.6358	0.8395	-1	-1	-1	-1
-1	0.3239	0.8214	-0.6694	-1	0	1	0
0	0.9228	0.3958	0.2069	0	1	1	1
0	-1	0.6561	0.3573	0	0	0	0
0	0.8757	-1	-0.5076	0	0	0	0
0	-0.4536	1	1	0	-1	-1	1
0	0.1711	0.0088	-1	0	1	-1	-1
0	1	0.6087	0.2202	0	-1	1	-1
-1	-0.4073	0.9048	-0.0211	1	0	-1	1
-1	-0.8253	0.282	0.767	1	0	1	-1
-1	-1	-0.5412	-0.7986	1	-1	-1	-1
-1	-0.5722	-0.7117	-0.0072	1	1	0	0
-1	0.7043	-0.8063	1	1	1	0	0
-1	0.6959	0.8122	-0.6769	1	-1	1	1

†See Table 5 of JG. WPs are separated by a blank line.

Table 5: JG's EE-SPD and SPLIT's for scenario 94 (two WP factors and two SP factors in ten WPs of size three).†

JG				SPLIT			
w_1	w_2	s_1	s_2	w_1	w_2	s_1	s_2
0	0	0	0	1	1	1	-1
0	0	0	0	1	1	-1	0
0	0	0	0	1	1	0	1
1	-1	0	1	1	0	-1	-1
1	-1	0	1	1	0	0	0
1	-1	0	1	1	0	1	1
1	1	0	0	1	-1	0	-1
1	1	0	0	1	-1	1	0
1	1	0	0	1	-1	-1	1
0	0	0	0	0	1	0	0
0	0	0	0	0	1	1	1
0	0	0	0	0	1	-1	-1
-1	1	1	0	0	-1	0	0
-1	1	1	0	0	-1	1	1
-1	1	1	0	0	-1	-1	-1
1	1	1	0	-1	1	0	-1
1	1	1	0	-1	1	1	0
1	1	1	0	-1	1	-1	1
-1	-1	0	-1	-1	0	-1	-1
-1	-1	0	-1	-1	0	0	0
-1	-1	1	1	-1	0	1	1
0	1	1	1	-1	-1	-1	0
0	1	1	-1	-1	-1	0	1
0	1	-1	0	-1	-1	1	-1
-1	1	-1	-1	1	-1	0	-1
-1	1	-1	1	1	-1	1	0
-1	1	-1	-1	1	-1	-1	1
1	-0.2242	-1	-1	-1	1	0	-1
1	-0.2242	-1	1	-1	1	1	0
1	-0.2242	1	-1	-1	1	-1	1

†See Table 5 of JG. WPs are separated by a blank line.

Table 6: JG's EE-SPD and SPLIT's EE-SPD* for scenario 109 (two WP factors and three SP factors in eight WPs of size six).†

JG					SPLIT				
w_1	w_2	s_1	s_2	s_3	w_1	w_2	s_1	s_2	s_3
-1	-1	0.5608	-0.9285	-0.9249	-1	-1	0	-1	-1
-1	-1	-0.0636	0.101	0.9757	-1	-1	-1	0	0
-1	-1	1	-1	0.08	-1	-1	1	1	-1
-1	-1	0.5156	-1	-0.5614	-1	-1	-1	0	0
-1	-1	-0.0674	0.1807	0.9794	-1	-1	1	-1	1
-1	-1	0.7043	-0.9039	-0.9384	-1	-1	0	1	1
0	1	-0.1102	-0.9291	0.9659	1	-1	0	1	-1
0	1	-0.9597	-0.1577	-0.25	1	-1	1	0	0
0	1	-0.2042	-0.956	-0.9116	1	-1	0	-1	1
0	1	-0.3594	0.9096	-0.8927	1	-1	-1	1	1
0	1	0.9958	0.637	0.9963	1	-1	-1	-1	-1
0	1	-0.9995	-0.2368	0.9935	1	-1	1	0	0
1	-1	1	0.9105	-0.9913	-1	1	-1	1	1
1	-1	-0.8644	-0.5947	-0.7272	-1	1	1	0	0
1	-1	-0.857	-0.9702	-0.9746	-1	1	0	-1	1
1	-1	-0.0942	-0.5554	-0.3228	-1	1	1	0	0
1	-1	-0.9222	0.7405	0.9	-1	1	0	1	-1
1	-1	0.8851	-0.8975	0.755	-1	1	-1	-1	-1
0	0	0.5699	-1	-0.922	1	1	-1	0	0
0	0	-0.2167	-1	0.58	1	1	0	-1	-1
0	0	-0.0737	-0.9989	0.5278	1	1	-1	0	0
0	0	0.9151	0.8265	-0.8835	1	1	0	1	1
0	0	0.9102	0.8611	-0.8444	1	1	1	1	-1
0	0	0.8668	1	0.9505	1	1	1	-1	1
0	-1	1	-1	-0.0059	0	-1	-1	0	-1
0	-1	1	-1	0.0603	0	-1	1	-1	-1
0	-1	1	-1	0.0274	0	-1	0	1	0
0	-1	0.0985	0.9988	1	0	-1	1	0	1
0	-1	0.1579	-1	0.3202	0	-1	-1	-1	1
0	-1	0.1761	-1	0.1066	0	-1	0	1	0
0	1	-0.1269	0.9603	-1	1	0	0	-1	0
0	1	-0.9037	-0.8422	0.906	1	0	-1	0	1
0	1	-0.6638	0.6229	0.8773	1	0	1	0	-1
0	1	0.1012	-0.4593	0.119	1	0	0	-1	0
0	1	0.9293	-0.0913	0.9991	1	0	1	1	1
0	1	-0.9733	-0.9233	-1	1	0	-1	1	-1
1	-1	-0.7834	-0.7823	0.5188	-1	0	0	-1	0
1	-1	-1	0.8476	-1	-1	0	0	-1	0
1	-1	0.1593	-0.9999	-0.8963	-1	0	-1	1	-1
1	-1	0.759	0.7333	1	-1	0	1	1	1
1	-1	-0.9708	-0.9286	-0.0502	-1	0	-1	0	1
1	-1	0.9832	-0.2368	-0.9331	-1	0	1	0	-1
-1	1	-0.4832	-0.4272	1	0	1	1	0	1
-1	1	0.6825	0.9896	-0.7524	0	1	-1	0	-1
-1	1	-0.2976	0.5491	0.9729	0	1	0	1	0
-1	1	-0.4314	-0.4299	0.9966	0	1	-1	-1	1
-1	1	1	0.8786	-1	0	1	1	-1	-1
-1	1	0.5791	0.9992	-0.7558	0	1	0	1	0

†See Table 5 of JG. WPs are separated by a blank line.

Table 7. The D-efficiencies of JG's EE-SPDs and SPLIT's for 25 scenarios where SPLIT produces more efficient EE-SPDs.

Scenario	m_W	m_S	b	k	JG	SPLIT	Relative efficiency (%)
25	1	2	5	3	80.5§	91.7	114.0
30	1	2	6	3	70.9	91.5	129.0
36	1	2	7	4	82.6	85.4‡	103.4
44	1	3	5	6	66.5	92.0	138.3
46	1	3	6	4	66.4	85.3	128.4
47	1	3	6	5	59.3†	96.2	162.2
48	1	3	6	6	37.8†§	90.2‡	238.4
66	2	1	10	4	90.8†	92.0	101.3
68	2	1	10	6	83.1†	97.9	117.8
69	2	1	11	2	90.9	91.5	100.6
71	2	1	11	4	91.2†	91.4	100.2
73	2	1	11	6	67.6†	97.4	144.0
74	2	1	12	2	80.3	91.8	114.3
76	2	1	12	4	94.2	94.7‡	100.5
78	2	1	12	6	47.5†	98.2‡	206.7
89	2	2	9	3	75.2†	90.6	120.5
94	2	2	10	3	44.5†§	92.5	207.8
95	2	2	10	4	50.2†	82.7	164.8
97	2	2	10	6	84.2†	95.1	112.9
99	2	2	11	3	38.5†	89.3	231.8
100	2	2	11	4	60.6†	88.4	145.9
101	2	2	11	5	66.3†	88.2	133.1
102	2	2	11	6	55.7†	93.2	167.3
109	2	3	8	6	44.9†	90.3‡	201.2
111	3	3	12	4	93.7	97.2	103.7

†JG design involves non-integer levels.

‡SPLIT design is also an EE-SPD*.

§The D-efficiency was recomputed using $|\mathbf{M}_{\text{opt}}|$ of the new D-optimal design constructed by the JMP software.

References

Arnouts, H. & Goos, P. (2010). Update formulas for split-plot and block designs, *Computational Statistics & Data Analysis*, 54, 3381-3391.

Box, G.E.P. & Wilson, K.B. (1951). On the experimental attainment of optimum conditions, *Journal of the Royal Statistical Society.* Series B 13, 1-45.

Box, G.E.P. & Behnken, D.W. (1960). Some new three level designs for the study of quantitative variables, *Technometrics*, 2, 455-477.

Goos, P. (2006). Optimal versus orthogonal and equivalent-estimation design of blocked and split-plot experiments, *Statistica Neerlandica*, 60, 361-378.

Goos, P. & Vandebroek, M. (2001). Optimal split-plot designs, *Journal of Quality Technology*, 33, 436-450.

Goos, P. & Vandebroek, M. (2003), D-optimal split-plot designs with given numbers and sizes of whole plots, *Technometrics*, 45, 235-245.

Goos, P. & Vandebroek, M. (2004). Outperforming completely randomized designs, *Journal of Quality Technology*, 36, 12-26.

Großmann H. & Gilmour S. G. (2013). Partially orthogonal blocked three-level response surface designs. Unpublished manuscript.

Jones, B. & Goos, P. (2007)., A candidate-set-free algorithm for generating D-optimal split-plot designs, *Journal of the Royal Statistical Society, Series C*, 56, 347-364.

Jones, B. & Goos, P. (2012). An algorithm for finding D-efficient equivalent-estimation second-order split-plot designs, *Journal of Quality Technology*, 44, 363-374.

Jones, B., & Nachtsheim, C. J. (2011). A class of three levels designs for definitive screening in the presence of second order effects, *Journal of Quality Technology*, 43, 115.

Jones, B. & Nachtsheim, C. J. (2009). Split-plot designs: what, why and how, *Journal of Quality Technology*, 41, 340-361.

Kowalski, S. M., Parker, P. A., & Vining, G. G. (2007). Tutorial: Industrial split-plot experiments. *Quality Engineering*, 19, 1-15.

Macharia, H. & Goos, P. (2010). D-optimal and D-efficient equivalent-estimation second-order split-plot designs, *Journal of Quality Technology*, 42, 358-372.

Meyer, R. K. & Nachtsheim, C. J. (1995). The co-ordinate exchange algorithm for constructing exact optimal experimental designs. *Technometrics*, 37, 6069.

Mylona, K., Macharia, H. & Goos, P. (2013). Three-level equivalent-estimation split-plot designs based on subset and supplementary difference set designs, *IIE Transactions*, 45, 1153-1165.

Nguyen, N.K. (1996). An algorithmic approach to constructing supersaturated designs, *Technometrics* 38, 69-73.

Nguyen, N.K. & Stylianou, S. (2013). Constructing definitive screening designs using cyclic generators, *Journal of Statistics Theory & Practice*, Accepted author version posted online: 14 Mar 2013.

Parker, P. A., Kowalski, S. M. & Vining, G. G. (2006). Classes of split-plot response surface designs for equivalent-estimation, *Quality and Reliability Engineering International* 22, 291-305.

Parker, P. A., Kowalski, S. M. & Vining, G. G. (2007). Construction of balanced equivalent-estimation second-order split-plot designs, *Technometrics* 49, 56-65.

Trinca, L. A. & Gilmour. S. G. (2000) An algorithm for arranging response surface designs in small blocks. *Comput. Statist. Data Anal.* 33, 25-43.

Trinca, L. A. & Gilmour. S. G. (2002) Erratum to "An algorithm for arranging response surface designs in small blocks [Comput. Statist. Data Anal. 33 (2000) 25-43]. *Comput. Statist. Data Anal.* 40, 475.

Vining, G. G., Kowalski, S. M. & Montgomery, D. C. (2005). Response surface designs within a split-plot structure. *Journal of Quality Technology* 37, 115-129.