Biometrika (2009), **96**, 1, *pp*. 67–82
doi: 10.1093/biomet/asn070

D-optimal design of split-split-plot experiments

BY BRADLEY JONES

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513, U.S.A.
bradley.jones@jmp.com

AND PETER GOOS

Faculty of Applied Economics, Universiteit Antwerpen,
Prinsstraat 13, 2000 Antwerpen, Belgium
peter.goos@ua.ac.be

SUMMARY

In industrial experimentation, there is growing interest in studies that span more than one processing step. Convenience often dictates restrictions in randomization in passing from one processing step to another. When the study encompasses three processing steps, this leads to split-split-plot designs. We provide an algorithm for computing D-optimal split-split-plot designs and several illustrative examples.

Some key words: Coordinate-exchange algorithm; D-optimality; Hard-to-change factor; Multi-stratum design; Split-plot design; Split-split-plot design.

1. INTRODUCTION

Split-plot designs are commonplace in industrial applications because there are often system level, set-up or processing factors that are difficult, expensive or time-consuming to change between successive processing runs. When this happens there is a natural inclination to perform all the runs of a particular level combination of such hard-to-change factors in succession. Such an ordering of the runs is clearly not random and many completely randomized designs have been reordered after the randomization to become split-plot designs inadvertently. Clearly it is preferable to design the few changes of the hard-to-change factors rather than to discover an after the fact run reordering or, even worse, fail to notice the reordering at all. Including this grouping of runs as a part of the design problem allows the researcher to maximize the information obtained about the statistical model given this restriction in randomization. Webb et al. (2004) have shown the cost in efficiency due to inadvertent split plotting.

The split-plot structure divides the experimental runs into two strata. The top stratum contains the whole plots. A whole plot is a group of runs where the hard-to-change factor combinations remain constant. The lower stratum contains the individual subplot runs.

The practical need for more than two strata in a design arises when experimenting on processes with multiple steps or stages. If the experimental units can be reordered between stages, then a split-lot or strip-plot design results; see for example Mee & Bates (1998). In many cases, it is either too complicated or even impossible to reorder the experimental units between strata. This leads to the split-split-plot design structure.

Schoen (1999) provides an example of multi-stage processing, leading to a split-split-plot design in his case study involving the production of cheese. Cheese processing starts with milk

storage. Typically milk from one storage facility provides the raw material for several curds processing units, the second processing stage. Then the curds are further processed to yield individual cheeses.

Split-split-plot designs are a three-stratum extension of split-plot designs. They divide into whole plots, subplots within whole plots and individual runs within each subplot. The top stratum remains the whole-plot stratum. Over the course of the experiment, the whole-plot or very-hard-to-change factors' levels are changed the fewest times. The levels of subplot factors or hard-to-change factors are changed more frequently. Also the levels of these factors must change whenever the whole-plot factor levels change in order to preserve the nested unit structure that characterizes the split-split-plot design. The sub-subplot factors or easy-to-change factors should be reset between each run regardless of whether their level changes.

Typically, the number of subplots is an integer multiple of the number of whole plots and the total number of runs is an integer multiple of the number of subplots. In this article, we assume this kind of structure, although it is not a requirement in general.

The extension from two-stratum to three-stratum experiments is not trivial. There is little literature on the design of such studies. Edmondson (1991) pointed out that various levels of splitting experimental plots or units are often required in agricultural and horticultural studies, but did not go beyond split-plot designs, with one level of splitting, in his article. Trinca & Gilmour (2001) considered the design and analysis of multi-stratum experiments including non-orthogonal designs. Their design approach attempts to orthogonalize each stratum of the design as much as possible with respect to the higher strata. Schoen (1999) constructed an orthogonal two-level split-split-plot design by joining fractional factorial designs in order to create the desired nesting structure. Brien & Bailey (2006) provided a diagrammatic method for describing complex nesting and crossed structures with many practical examples.

This article describes an algorithm for creating D-optimal split-split-plot designs and provides several specific examples of these designs to demonstrate their utility for screening experimentation. It builds on Goos (2002, 2006) who introduced an optimal design approach to construct split-plot designs and provided algorithms for finding optimal split-plot designs that exchange points from a starting design with points from a candidate set, and on Jones & Goos (2007), who showed how to avoid the construction of a candidate set in the search for optimal split-plot designs. Their candidate-set-free algorithm runs in polynomial time in the number of factors thus allowing the construction of designs with many more factors and runs than was previously feasible.

2. MODEL AND DESIGN CRITERION

For a split-split-plot experiment with b_1 whole plots, b_2 subplots per whole plot and k runs per subplot, and thus sample size $n = b_1 b_2 k$, the model can be written as

$$Y = X\beta + Z_1\gamma_1 + Z_2\gamma_2 + \varepsilon, \tag{1}$$

where Y is the n-dimensional vector of the responses arranged per whole plot and per subplot, X represents the $n \times p$ model matrix containing the settings of the very-hard-to-change factors, w, the hard-to-change factors, s, the easy-to-change factors, t, their model expansions, β is a p-dimensional vector containing the p fixed effects in the model, and $Z_1 = I_{b_1} \otimes 1_{b_2 k}$ is an $n \times b_1$ matrix of zeroes and ones. A one in row j of column i of Z_1 means that run j is in the ith whole plot. Furthermore, $Z_2 = I_{b_1} \otimes I_{b_2} \otimes 1_k = I_{b_1 b_2} \otimes 1_k$ is an $n \times b_1 b_2$ matrix of zeroes and ones indicating how the n runs have been assigned to the b_2 subplots within each of the b_1 whole plots. The b_1- and $b_1 b_2$-dimensional vectors γ_1 and γ_2 are the random effects associated with the

whole plots and the subplots, respectively. Finally, ε is the n-dimensional vector containing the random errors. It is assumed that

$$E(\varepsilon) = 0_n, \operatorname{cov}(\varepsilon) = \sigma_\varepsilon^2 I_n, \tag{2}$$

$$E(\gamma_1) = 0_{b_1}, \ \operatorname{cov}(\gamma_1) = \sigma_{\gamma_1}^2 I_{b_1}, \tag{3}$$

$$E(\gamma_2) = 0_{b_1 b_2}, \ \operatorname{cov}(\gamma_2) = \sigma_{\gamma_2}^2 I_{b_1 b_2}, \tag{4}$$

and

$$\operatorname{cov}(\gamma_1, \varepsilon) = 0_{b_1 \times n}, \operatorname{cov}(\gamma_2, \varepsilon) = 0_{b_1 b_2 \times n}, \operatorname{cov}(\gamma_1, \gamma_2) = 0_{b_1 \times b_1 b_2}. \tag{5}$$

The variances $\sigma_{\gamma_1}^2$, $\sigma_{\gamma_2}^2$ and σ_ε^2 in these expressions are referred to as the whole-plot variance, the subplot variance and the error variance, respectively. Under the assumptions (2)–(5), the covariance matrix of the responses, $\operatorname{var}(Y)$, is

$$V = \sigma_\varepsilon^2 I_n + \sigma_{\gamma_1}^2 Z_1 Z_1' + \sigma_{\gamma_2}^2 Z_2 Z_2'. \tag{6}$$

This matrix is of the form $V = \operatorname{diag}(V_1, \ldots, V_{b_1})$, where each

$$V_i = \sigma_\varepsilon^2 I_{b_2 k} + \sigma_{\gamma_1}^2 1_{b_2 k} 1_{b_2 k}' + \sigma_{\gamma_2}^2 (I_{b_2} \otimes 1_k 1_k') = \sigma_\varepsilon^2 \{ I_{b_2 k} + \eta_1 1_{b_2 k} 1_{b_2 k}' + \eta_2 (I_{b_2} \otimes 1_k 1_k') \},$$

and $\eta_1 = \sigma_{\gamma_1}^2 / \sigma_\varepsilon^2$ and $\eta_2 = \sigma_{\gamma_2}^2 / \sigma_\varepsilon^2$. These two variance ratios measure the extent to which observations are correlated. The correlation between two observations in the same subplot is equal to $(\eta_1 + \eta_2)/(1 + \eta_1 + \eta_2)$. The correlation between two observations from the same whole plot but a different subplot is $\eta_1/(1 + \eta_1 + \eta_2)$. It is clear that larger values for η_1 and η_2 result in more correlated observations within the whole plots and subplots.

When the random error terms as well as the whole-plot and subplot effects are normally distributed, the maximum likelihood estimator of the unknown model parameter β in (1) is the generalized least-squares estimator

$$\hat{\beta} = (X' V^{-1} X)^{-1} X' V^{-1} Y,$$

with covariance matrix

$$\operatorname{var}(\hat{\beta}) = (X' V^{-1} X)^{-1}.$$

The use of $\hat{\beta}$ requires the estimation of the variance components $\sigma_{\gamma_1}^2$, $\sigma_{\gamma_2}^2$ and σ_ε^2, which can be substituted in V. This leads to the feasible generalized least-squares estimator. For the variance component estimation, we recommend restricted maximum likelihood estimation because of its generality. The unbiasedness and variance of the resulting estimator are discussed in Kackar & Harville (1984) and Harville & Jeske (1992). For the purpose of statistical inference, we advocate the use of the method of Kenward & Roger (1997) for determining the standard errors and the denominator degrees of freedom for the hypothesis tests concerning the fixed effects. When limitations on the number of whole plots, subplots or runs do not allow the variance components to be estimated, we would use either the method of Lenth (1989) or the permutation test approach of Loughin & Noble (1997), adapted for split-split-plot experiments in a way similar to that in which Loeppky & Sitter (2002) adapted these methods for split-plot experiments.

Under the model assumptions made, the information matrix on the unknown fixed parameters β is given by $M = X' V^{-1} X$. A commonly used criterion to select experimental designs is the D-optimality criterion which seeks designs that maximize the determinant of the information matrix, $|M| = |X' V^{-1} X|$. The D-optimality criterion has been used for constructing split-plot designs by Goos & Vandebroek (2001, 2003, 2004) and Goos & Donev (2007) and it is also

the criterion that is implemented in the candidate-set-free algorithm described in Jones & Goos (2007). We use the D-efficiency $(|M_1|/|M_2|)^{1/p}$ to compare the quality of two designs with information matrices M_1 and M_2. In general, the optimal split-plot design will depend on the variance ratios η_1 and η_2 through V. The sensitivity of the D-optimal designs to the choice of η_1 and η_2 is discussed in § 5.

3. DESIGN CONSTRUCTION ALGORITHM

3·1. *Algorithm outline*

This section first provides a rough general description of a modified coordinate-exchange algorithm for generating D-optimal split-split-plot designs. Next, we show how to substantially reduce the computational work by using a fast procedure to evaluate the change in the D-criterion value when making changes to a design. In the Appendix, a fast update for the inverse of the information matrix after exchanging one point for another is also provided.

The algorithm requires the prior specification of the following:

1. for each factor whether it is continuous, categorical or a mixture ingredient;
2. designation of the factors that are very hard to change, i.e. the factors applied to the whole-plot stratum;
3. designation of the factors that are hard to change, i.e. the factors applied to the subplot stratum;
4. any additional constraints on factor combinations;
5. the number, b_1, of independent resettings of the very-hard-to-change factors, i.e. the number of whole plots;
6. the number, b_2, of independent resettings of the hard-to-change factors for each setting of the very-hard-to-change factors, i.e. the number of subplots within each whole plot;
7. the number of observations, k, in each subplot;
8. the ratio η_1 of the variance associated with the very-hard-to-change factors, $\sigma^2_{\gamma_1}$, to the error variance, σ^2_ε;
9. the ratio η_2 of the variance associated with the hard-to-change factors, $\sigma^2_{\gamma_2}$, to the error variance, σ^2_ε;
10. the a priori model;
11. the number of random starting designs or tries, n_T, to consider.

Given this information, the body of the algorithm has two parts. The first is the creation of a starting design. The second is the iterative improvement of this design until no further improvement is possible. Improvements are measured by increases in the objective function, $|M| = |X'V^{-1}X|$. The two parts are performed n_T times. Each time the final value of $|M|$ found in the current iterate is compared to the maximum value of $|M|$ from all the previous iterates. If the current value is higher, then it becomes the new maximum and the current design is stored.

The starting design is formed column by column. For whole-plot factor columns, b_1 random numbers are chosen. For subplot factor columns, $b_1 b_2$ random numbers are chosen. For sub-subplot factor columns, the values for each of the n rows are chosen randomly. All the rows in a given whole plot have the same value for each whole-plot factor. If there is more than one whole-plot factor, however, the factors may have different values. Similarly, all the rows in a given subplot have the same value for each subplot factor. Different subplot factors may also have different values inside a subplot. This procedure gives the starting design the desired split-split-plot structure.

Improvements are made to the starting design by considering changes in the design on an element-by-element basis. This is inspired by the coordinate-exchange algorithm of Meyer & Nachtsheim (1995). The procedure for changing any given element depends on whether that element is an easy-to-change factor, a hard-to-change factor or a very-hard-to-change factor.

For an element in a sub-subplot factor column, the objective function is evaluated over a discrete number of values spanning the range of that factor. If the maximal value of the objective function is larger than the current maximum, then the current maximum is replaced and the current element in the design is replaced by the factor setting corresponding to the maximal value.

The procedure is more involved for an element in a subplot factor column. If such an element changes, then all the corresponding elements for that column in the same subplot must also change. A discrete number of values spanning the range of the subplot factor are evaluated. If the maximal value of the objective function is larger than the current maximum, then the current maximum is replaced and all elements in the subplot factor column in the subplot under consideration are replaced by the factor setting corresponding to the maximal value.

For an element in a whole-plot factor column, the procedure is the most computationally expensive. If such an element changes, then all the corresponding elements for that column in the same whole plot must also change. A discrete number of values spanning the range of the factor are evaluated. Again, if the maximal value of the objective function is larger than the current maximum, then the current maximum is replaced and all elements in the whole-plot factor column in the whole plot under consideration are replaced by the factor setting corresponding to the maximal value.

This element-by-element procedure continues until a complete cycle through the entire design has been performed. Then, another complete cycle through the design is performed, checking to see if any element has been changed in the current pass. This continues until no changes are made in a whole pass or until a specified maximum number of passes have been executed.

3·2. *Fast update procedures*

Fast procedures can be used in the algorithm for evaluating the impact of a change of the design on the objective function, $|M|$. Also, the inverse M^{-1} of the information can be updated at a relatively low computational cost. The update procedures all build on the following theorem which gives a simple analytical expression for the inverse of the covariance matrix V in (6).

THEOREM 1. *The inverse of the covariance matrix V is equal to*

$$V^{-1} = \sigma_\varepsilon^{-2} I_n - c_1 Z_1 Z_1' - c_2 Z_2 Z_2', \tag{7}$$

where

$$c_1 = \sigma_\varepsilon^{-2} \frac{\eta_1 - \frac{\eta_1 \eta_2 k}{1+\eta_2 k}}{1 + \eta_1 b_2 k + \eta_2 k}, \quad c_2 = \sigma_\varepsilon^{-2} \frac{\eta_2}{1 + \eta_2 k}.$$

A proof of this result is obtained by multiplying the right-hand side of (7) by the right-hand side of (6) and observing that this produces the identity matrix. When doing so, the following matrix results prove to be useful: $Z_1' Z_1 = b_2 k I_{b_1}$, $Z_2' Z_2 = k I_{b_2}$ and $Z_1 Z_1' Z_2 Z_2' = k Z_1 Z_1'$.

The theorem can be used to derive an alternative expression for the information matrix. As a matter of fact, because of (7), we have that $V^{-1} = \mathrm{diag}(V_1^{-1}, \ldots, V_{b_1}^{-1})$, where each

$$V_i^{-1} = \sigma_\varepsilon^{-2} \{ I_{b_2 k} - c_1 1_{b_2 k} 1_{b_2 k}' - c_2 (I_{b_2} \otimes 1_k 1_k') \}.$$

This makes it possible to rewrite the information matrix as follows:

$$\begin{aligned} M &= \sum_{i=1}^{b_1} X_i' V_i^{-1} X_i, \\ &= \sigma_\varepsilon^{-2} X'X - c_1 \sum_{i=1}^{b_1} \left(X_i' 1_{b_2 k} 1_{b_2 k}' X_i\right) - c_2 \sum_{i=1}^{b_1} \sum_{j=1}^{b_2} (X_{ij}' 1_k 1_k' X_{ij}), \\ &= \sigma_\varepsilon^{-2} X'X - c_1 \sum_{i=1}^{b_1} \left(X_i' 1_{b_2 k}\right)\left(X_i' 1_{b_2 k}\right)' - c_2 \sum_{i=1}^{b_1} \sum_{j=1}^{b_2} (X_{ij}' 1_k)(X_{ij}' 1_k)', \end{aligned} \tag{8}$$

where X_i is the part of X corresponding to the ith whole plot and X_{ij} is the part of X corresponding to the jth subplot within the ith whole plot. If the model expansion of the lth design point in the jth subplot of whole plot i is denoted by $f(w_i, s_{ij}, t_{ijl})$, then $X'X$ can be written as $\sum_{i=1}^{b_1} \sum_{j=1}^{b_2} \sum_{l=1}^{k} f(w_i, s_{ij}, t_{ijl}) f'(w_i, s_{ij}, t_{ijl})$. As a result, (8) can be written as a sum of outer products of vectors. This opens the prospect of fast updates of the information matrix, its determinant and its inverse.

Updating the determinant of the information matrix after changing the level of an easy-to-change factor, a hard-to-change factor or a very-hard-to-change factor can be done using a formula of the form $|M^*| = |M||I_{d_i} + D_i U_i' M^{-1} U_i|$, where M and M^* represent the information matrix before and after the change, respectively, D_i is a d_i-dimensional diagonal matrix and U_i is a $p \times d_i$ matrix. This is shown in detail in the Appendix. The matrices D_i and U_i as well as the integer d_i all depend on whether the factor level that is modified corresponds to an easy-to-change, a hard-to-change or a very-hard-to-change factor.

For example, after a change in the level of an easy-to-change factor in the lth run of the jth subplot within whole plot i, the matrices D_i and U_i needed for the update are given by

$$D_1 = \operatorname{diag}\left(-\sigma_\varepsilon^{-2}, c_1, c_2, \sigma_\varepsilon^{-2}, -c_1, -c_2\right) \tag{9}$$

and

$$U_1 = \begin{bmatrix} f(w_i, s_{ij}, t_{ijl}) & X_i' 1_{b_2 k} & X_{ij}' 1_k & f(w_i, s_{ij}, t_{ijl}^*) & X_i^{*'} 1_{b_2 k} & X_{ij}^{*'} 1_k \end{bmatrix}'. \tag{10}$$

In this last expression,

$$X_i^{*'} 1_{b_2 k} = X_i' 1_{b_2 k} - f(w_i, s_{ij}, t_{ijl}) + f\left(w_i, s_{ij}, t_{ijl}^*\right) \tag{11}$$

and

$$X_{ij}^{*'} 1_{b_2 k} = X_{ij}' 1_{b_2 k} - f(w_i, s_{ij}, t_{ijl}) + f\left(w_i, s_{ij}, t_{ijl}^*\right), \tag{12}$$

with $f'(w_i, s_{ij}, t_{ijl})$ the original row of X where the change took place, $f'(w_i, s_{ij}, t_{ijl}^*)$ the modified row, and X_i^* and X_{ij}^* the updated versions of X_i and X_{ij}, respectively. Note that $f'(w_i, s_{ij}, t_{ijl})$ is different from $f'(w_i, s_{ij}, t_{ijl}^*)$ only in the elements corresponding to the main effect, the interactions and the higher-order effects the factor whose level was changed is involved in. This simplifies the updating of $X_i^{*'} 1_{b_2 k}$ and $X_i^{*'} 1_{b_2 k}$.

For a change in the level of an easy-to-change factor, the value of d_i is six. As a result, computing the new determinant using the update formula requires calculating the determinant of a 6×6 matrix instead of the determinant of a $p \times p$ matrix. As split-split-plot designs involve at least three factors, the number of model parameters, p, will often be substantially larger than six, so that the update formula will lead to substantial savings in computational effort.

Table 1. *D-optimal* 16*-run split-split-plot design with two whole plots, each consisting of two subplots for estimating a main-effects model with one very-hard-to-change factor* w*, one hard-to-change factor* s *and twelve easy-to-change factors* t_1–t_{12}

Whole plot	Subplot	w	s	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
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	2	−1	1	1	−1	−1	−1	−1	1	−1	−1	1	−1	−1	−1
1	2	−1	1	−1	−1	1	1	1	1	1	−1	−1	1	1	−1
1	2	−1	1	−1	1	−1	−1	−1	−1	1	1	1	1	1	1
1	2	−1	1	1	1	1	1	1	−1	−1	1	−1	−1	−1	1
2	3	1	1	1	1	−1	1	1	1	1	−1	1	−1	1	1
2	3	1	1	−1	1	1	−1	−1	1	−1	−1	−1	1	−1	1
2	3	1	1	1	−1	1	−1	−1	−1	1	1	−1	−1	1	−1
2	3	1	1	−1	−1	−1	1	1	−1	−1	1	1	1	−1	−1
2	4	1	−1	−1	1	1	−1	1	−1	1	−1	1	−1	−1	−1
2	4	1	−1	1	−1	−1	1	−1	−1	1	−1	−1	1	−1	1
2	4	1	−1	1	−1	1	−1	1	1	−1	1	1	1	1	1
2	4	1	−1	−1	1	−1	1	−1	1	−1	1	−1	−1	1	−1

Updating the determinant of the information matrix after a change in either a hard-to-change factor or a very-hard-to-change factor is more involved. It requires the computation of the determinant of a $2(k+2) \times 2(k+2)$ matrix and a $2(b_2k + b_2 + 1) \times 2(b_2k + b_2 + 1)$ matrix, respectively. Therefore, the update formulae for changes in hard-to-change and very-hard-to-change factor levels are advantageous when k and b_2 are small compared to the number of model parameters, p.

4. Designs for main-effects models

Using the modified coordinate-exchange algorithm sketched in § 3, we have been able to construct orthogonal split-split-plot designs with diagonal information matrices for many nesting structures for which the numbers of whole plots and subplots within the whole plots are powers of two or multiples of four. Table 1 shows such a design with 16 runs, two whole plots and two subplots per whole plot, accommodating one very-hard-to-change factor w, one hard-to-change factor s, and twelve easy-to-change factors, t_1–t_{12}. The information matrix of that design, which is a projection of a Hall-type IV orthogonal array (Sun and Wu, 1993), has the information matrix diag(1·2308 I_2, 3·2, 16 I_{12}) when $\sigma^2_{\gamma_1}$, $\sigma^2_{\gamma_2}$ and σ^2_ε are all one. This demonstrates that no information on the easy-to-change factors is lost because of the nested unit structure of the design. However, this design does not allow for the estimation of the whole plot error variance, $\sigma^2_{\gamma_1}$, because it only has two whole plots. Also, it is impossible to estimate the error variance, σ^2_ε, because the twelve degrees of freedom at the sub-subplot level are used up for estimating the main effects of the twelve easy-to-change factors.

A main-effects design that allows $\sigma^2_{\gamma_1}$ to be estimated but not σ^2_ε is displayed in Table 2. The design has six whole plots each consisting of two subplots with two observations and, like the design in Table 1, accommodates one very-hard-to-change factor w, one hard-to-change factor s and twelve easy-to-change factors, t_1–t_{12}. For each pair of runs in a subplot, the levels of the easy-to-change factors are each other's opposites. The information matrix of that design equals

Table 2. *D-optimal 24-run split-split-plot design with six whole plots, each consisting of two subplots for estimating a main-effects model with one very-hard-to-change factor w, one hard-to-change factor s and twelve easy-to-change factors t_1–t_{12}*

Whole plot	Subplot	w	s	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
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	2	−1	−1	−1	1	−1	1	1	1	−1	1	1	−1	1	1
1	2	−1	−1	1	−1	1	−1	−1	−1	1	−1	−1	1	−1	−1
2	3	1	−1	1	1	1	1	1	1	−1	1	−1	1	−1	−1
2	3	1	−1	−1	−1	−1	−1	−1	−1	1	−1	1	−1	1	1
2	4	1	1	−1	1	1	−1	−1	−1	−1	1	−1	−1	1	−1
2	4	1	1	1	−1	−1	1	1	1	1	−1	1	1	−1	1
3	5	1	1	1	1	1	1	−1	−1	1	1	1	1	1	1
3	5	1	1	−1	−1	−1	−1	1	1	−1	−1	−1	−1	−1	−1
3	6	1	−1	1	−1	−1	1	−1	−1	−1	1	1	−1	−1	−1
3	6	1	−1	−1	1	1	−1	1	1	1	−1	−1	1	1	1
4	7	1	1	1	−1	−1	−1	−1	1	−1	1	−1	1	1	1
4	7	1	1	−1	1	1	1	1	−1	1	−1	1	−1	−1	−1
4	8	1	−1	−1	1	−1	1	−1	−1	−1	−1	−1	1	−1	1
4	8	1	−1	1	−1	1	−1	1	1	1	1	1	−1	1	−1
5	9	−1	−1	1	−1	1	1	1	−1	−1	−1	−1	−1	1	1
5	9	−1	−1	−1	1	−1	−1	−1	1	1	1	1	1	−1	−1
5	10	−1	1	1	1	1	−1	−1	1	−1	−1	1	−1	−1	1
5	10	−1	1	−1	−1	−1	1	1	−1	1	1	−1	1	1	−1
6	11	−1	−1	1	1	−1	−1	1	−1	−1	−1	1	1	1	−1
6	11	−1	−1	−1	−1	1	1	−1	1	1	1	−1	−1	−1	1
6	12	−1	1	1	1	−1	−1	1	−1	1	1	−1	−1	−1	1
6	12	−1	1	−1	−1	1	1	−1	1	−1	−1	1	1	1	−1

diag(3·4286 I_2, 8, 24 I_{12}) when $\sigma^2_{\gamma_1}$, $\sigma^2_{\gamma_2}$ and σ^2_ε are all one, so that, here too, no information is lost for estimating the easy-to-change factor effects due to the nested unit structure of the design.

D-optimal split-split-plot designs with identical nesting structures and fewer easy-to-change factors can easily be obtained from Tables 1 and 2 by removing columns corresponding to easy-to-change factors. Dropping one or more easy-to-change factor columns from the design makes the error variance, σ^2_ε, estimable, so that it is possible to test the significance of the remaining sub-subplot factors.

5. DESIGNS FOR INTERACTION MODELS

In completely randomized two-level designs, a diagonal information matrix guarantees an optimal design. In this section, we provide an example with a counter-intuitive result indicating that, for two-level split-split-plot designs, a diagonal information matrix may not be optimal.

Consider the D-optimal 32-run split-split-plot design with eight whole plots consisting of two subplots each in Table 3. This design is the best one obtained using our algorithm for estimating all the parameters of an interaction model including two very-hard-to-change factors w_1 and w_2, one hard-to-change factor s and three easy-to-change factors t_1, t_2 and t_3 when $\eta_1 = \eta_2 = 1$. It has a D-criterion value of $4{\cdot}80132 \times 10^{26}$ when $\sigma^2_{\gamma_1}$, $\sigma^2_{\gamma_2}$ and σ^2_ε are all one.

Although the design is orthogonal, its information matrix is not diagonal, i.e. $X'X$ is diagonal but $X'V^{-1}X$ is not. The design, however, has a lot of attractive features. First, all main effects

Table 3. *D-optimal 32-run split-split-plot design with eight whole plots, each consisting of two subplots for estimating an interaction model with two very-hard-to-change factors w_1 and w_2, one hard-to-change factor s and three easy-to-change factors t_1, t_2 and t_3*

Whole plot	Subplot	w_1	w_2	s	t_1	t_2	t_3	Whole plot	Subplot	w_1	w_2	s	t_1	t_2	t_3
1	1	1	1	1	−1	−1	1	5	9	−1	−1	1	1	1	−1
1	1	1	1	1	1	1	−1	5	9	−1	−1	1	−1	−1	−1
1	2	1	1	−1	1	−1	−1	5	10	−1	−1	−1	1	−1	1
1	2	1	1	−1	−1	1	1	5	10	−1	−1	−1	−1	1	1
2	3	−1	1	−1	−1	−1	1	6	11	1	−1	−1	1	1	−1
2	3	−1	1	−1	1	1	−1	6	11	1	−1	−1	−1	−1	1
2	4	−1	1	1	1	−1	−1	6	12	1	−1	1	1	−1	1
2	4	−1	1	1	−1	1	1	6	12	1	−1	1	−1	1	−1
3	5	1	−1	−1	−1	−1	−1	7	13	−1	1	1	1	−1	1
3	5	1	−1	−1	1	1	1	7	13	−1	1	1	−1	1	−1
3	6	1	−1	1	1	−1	−1	7	14	−1	1	−1	−1	−1	−1
3	6	1	−1	1	−1	1	1	7	14	−1	1	−1	1	1	1
4	7	−1	−1	−1	−1	1	−1	8	15	1	1	−1	−1	1	−1
4	7	−1	−1	−1	1	−1	−1	8	15	1	1	−1	1	−1	1
4	8	−1	−1	1	1	1	1	8	16	1	1	1	1	1	1
4	8	−1	−1	1	−1	−1	1	8	16	1	1	1	−1	−1	−1

are estimated independently. Only one of the six main effects, namely one corresponding to an easy-to-change factor, is not estimated independently of the two-factor interaction effects. Second, only six of the 462 off-diagonal elements of the information matrix and of the variance-covariance matrix of the parameter estimates are nonzero, so the information matrix is very nearly diagonal. The six nonzero off-diagonal elements of the variance-covariance matrix of the parameter estimates are all equal to $\pm 1/96 = \pm 0{\cdot}01042$ when $\sigma^2_{\gamma_1}$, $\sigma^2_{\gamma_2}$ and σ^2_ε are all one. They correspond to the covariances between the estimates of the main effect of one of the easy-to-change factors and its interactions with the two very-hard-to-change factors w_1 and w_2. These covariances are small compared to the variances of the fixed parameter estimates, which are displayed in the column labelled D-optimal in Table 4. A third attractive feature of the design is that eight of the fifteen effects involving easy-to-change factors are estimated with variance $1/32 = 0{\cdot}03125$, which is the best possible variance for a design involving 32 runs. Only the interactions involving pairs of easy-to-change factors are estimated substantially less precisely. The least precise of these interaction estimates has variance $3/32 = 0{\cdot}09375$, just like the main effect of the hard-to-change factor s and the two interactions between that factor and the very-hard-to-change factors w_1 and w_2.

The literature on minimum aberration two-level split-plot designs provides no ready-to-use alternative to the D-optimal 32-run split-split-plot design, but it does provide building blocks for generating 32-run designs with the desired split-split-plot structure with eight whole plots consisting of two subplots of size two. Some of these possess the attractive feature that their information matrix is diagonal, so that, unlike with the D-optimal design, the main effects and the two-factor interaction effects can be estimated independently. One such design can be constructed starting from a minimum aberration 32-run two-level split-plot design with two whole-plot factors, four subplot factors and eight whole plots of size four given in Bingham et al. (2004). The design's defining relation is $t_3 = w_1 w_2 s t_1 t_2$ and the contrast column $w_1 s t_2$ is used to obtain the desired number of eight whole plots. The minimum aberration design can be used as a 32-run split-split-plot design by using one of its four subplot columns for the hard-to-change factor s.

Table 4. *Variance of estimates of fixed model parameters for the D-optimal 32-run split-split-plot design in Table 3 and a combinatorially constructed alternative when* $\sigma^2_{\gamma_1} = \sigma^2_{\gamma_2} = \sigma^2_{\varepsilon} = 1$

Stratum	Effect	D-optimal	Alternative
WP	Intercept	0·21875	0·21875
WP	w_1	0·21875	0·21875
WP	w_2	0·21875	0·21875
WP	w_1w_2	0·21875	0·21875
SP	s	0·09375	0·09375
SP	w_1s	0·09375	0·09375
SP	w_2s	0·09375	0·09375
SSP	t_1	0·03125	0·03125
SSP	t_2	0·03125	0·03125
SSP	t_3	0·04167	0·03125
SSP	w_1t_1	0·03125	0·03125
SSP	w_1t_2	0·03125	0·03125
SSP	w_1t_3	0·04167	0·03125
SSP	w_2t_1	0·03125	0·03125
SSP	w_2t_2	0·03125	0·03125
SSP	w_2t_3	0·04167	0·03125
SSP	st_1	0·03125	0·03125
SSP	st_2	0·03125	0·03125
SSP	st_3	0·03977	0·03125
SSP	t_1t_2	0·09375	0·09375
SSP	t_1t_3	0·07721	0·21875
SSP	t_2t_3	0·06908	0·09375

A better alternative, however, can be constructed by arranging a half fraction of a factorial design with defining relation $t_2 = w_1w_2st_1$ in the desired split-split-plot structure using the contrast columns w_1, w_2 and $w_2t_1t_3$ to partition the 32 runs into eight whole plots. The variances of the parameter estimates for this combinatorially constructed alternative design are displayed in Table 4. Desirable features of the design, which has a diagonal information matrix, are that all whole-plot and subplot effects are estimated with maximum precision, and that all but three sub-subplot effects are estimated with the best possible variance, $1/32 = 0{\cdot}03125$. However, this is at the expense of the three two-factor interaction effects between the easy-to-change factors, which are raised to one of the higher strata. Two of these interaction effects are raised to the subplot stratum, and thus estimated with variance $3/32 = 0{\cdot}09375$, while the third one is even raised to the whole-plot stratum. This results in a variance of $7/32 = 0{\cdot}21875$. As a consequence of all this, the D-criterion value of the alternative design, when $\sigma^2_{\gamma_1}$, $\sigma^2_{\gamma_2}$ and σ^2_{ε} are all one, is $3{\cdot}17836 \times 10^{26}$. The D-efficiency of the combinatorially constructed design, relative to the D-optimal design, is 98·14%. Thus, the algorithmically constructed design has a higher determinant than the alternative despite its having a nondiagonal information matrix. This is achieved by sacrificing some of the precision of the estimates of three subplot effects and the independence between these estimates in order to obtain reasonably small variances for the estimates of the two-factor interaction effects between the easy-to-change factors.

The larger variances for the estimates of the main and interaction effects of some of the easy-to-change factors seem very difficult to avoid in many split-plot design problems, especially when the number of runs in every subplot is as small as two. Optimal design construction algorithms like ours, however, attempt to limit that increase by introducing some imbalance in the levels of the easy-to-change factors within the subplots. This is illustrated by the design in Table 3,

Table 5. *D-optimal* 12*-run split-split-plot designs with three whole plots, each consisting of two subplots with two runs for estimating a main-effects model with one very-hard-to-change factor* w*, one hard-to-change factor* s *and one easy-to-change factor* t*. The three factors are categorical*

				t		
Whole plot	Subplot	w	s	$\eta_2 = 0{\cdot}1$	$\eta_2 = 1$	$\eta_2 = 10$
1	1	A	a	1	1	1
1	1	A	a	2	2	2
1	2	A	b	2	2	2
1	2	A	b	3	3	1
2	3	B	b	1	1	2
2	3	B	b	2	3	3
2	4	B	c	2	2	1
2	4	B	c	3	3	3
3	5	C	a	2	2	1
3	5	C	a	3	3	3
3	6	C	c	1	1	2
3	6	C	c	2	2	3

where the low and the high level of the easy-to-change factor t_3 are unbalanced in each of the subplots 7–10. This explains the small variance inflation for the main-effect estimate of that factor when compared to the main-effect estimates for t_1 and t_2. In all other subplots, the levels of the easy-to-change factors are balanced.

6. Sensitivity to η_1 and η_2

As mentioned in § 2, the D-optimal split-split-plot designs depend on the two variance ratios, η_1 and η_2. Goos (2002) studied the effect of changing the ratio of the whole-plot variance to the error variance on D-optimal split-plot designs. He showed that, for given numbers and sizes of whole plots, D-optimal split-plot designs for first-order models are in many standard cases not a function of this ratio. Split-plot response surface designs did, however, show some sensitivity to changes in this ratio. Sometimes as many as three different designs were found to be D-optimal, each over mutually exclusive intervals of the variance ratio. Over practical ranges of the variance ratio, the D-optimal design usually did not change.

Of course, the split-split-plot structure is more complicated than the split-plot structure, so it is necessary to address this question again. We investigated the effect of changing the two variance ratios η_1 and η_2 over broad ranges, and report the results obtained for two different design problems that are illustrative for the sensitivity of the D-optimal designs to η_1 and η_2.

The first design problem involved a main-effects model in three categorical factors each at three levels. The first scenario had three whole plots, six subplots and a total sample size of twelve. We computed designs for a 3 by 3 grid of η_1 and η_2 values from 0·1 to 1 to 10. We found three different D-optimal designs, one for each value of η_2. The three designs are shown in Table 5. For each design, the whole plot and subplot structure is the same. There are four runs at each level of both the whole-plot factor and the subplot factor. The differences in the designs come in the sub-subplot factor levels. For $\eta_2 = 0{\cdot}1$, there are three runs at two levels and six runs at the other. For $\eta_2 = 1$, there are three runs at one level, four runs at another and five runs at the third. For $\eta_2 = 10$, there are four runs at each level. In this case, the arrangement of the three levels of the easy-to-change factor in six subplots of two runs takes the form of a duplicated balanced incomplete block design for three treatments with three blocks of size two.

Table 6. *Comparative D-efficiencies for misspecifications of η_1 and η_2 assuming σ_ε^2 is one for a design problem involving one very-hard-to-change factor, one hard-to-change factor and three easy-to-change factors and a model with main effects and two-factor interactions*

	$\sigma_{\gamma_1}^2 = \sigma_{\gamma_2}^2 = 0{\cdot}1$			$\sigma_{\gamma_1}^2 = \sigma_{\gamma_2}^2 = 1$			$\sigma_{\gamma_1}^2 = \sigma_{\gamma_2}^2 = 10$		
	η_1			η_1			η_1		
η_2	0·1	1·0	10	0·1	1·0	10	0·1	1·0	10
0·1	100·00	100·00	100·00	100·00	100·00	100·00	89·63	89·63	89·63
1	100·00	100·00	100·00	100·00	100·00	100·00	89·63	89·63	89·63
10	96·36	96·37	96·42	95·69	95·95	96·05	99·88	99·99	100·00

In order to investigate whether the information content of the three designs in Table 5 is very different, we compare the objective function, $|M|$, for each of them assuming $\sigma_{\gamma_1}^2$, $\sigma_{\gamma_2}^2$ and σ_ε^2 are all one. The determinant of the information matrix of the D-optimal design for $\eta_2 = 1$ is 3978·7, while the design with four runs at each level, obtained for $\eta_2 = 10$, has a determinant of 3944·7. The relative efficiency of this design is 99·88%. The determinant of the design with 3, 3 and 6 runs per level, obtained for $\eta_2 = 0{\cdot}1$, is 3672·6 with a relative efficiency of 98·86%. While the easy-to-change factor levels look very different for the three designs, there is thus no substantial difference among them for the purpose of estimating the parameters of the model.

The second design problem considered five continuous factors, with one very-hard-to-change and one hard-to-change factor, and a model including all main effects and two-factor interactions. There were six whole plots, 12 subplots and 24 runs in total. We computed nine D-optimal designs for this setting using the same 3 by 3 design in $\log \eta_1$ and $\log \eta_2$.

Here the story is more involved as the designs vary depending on both η_1 and η_2. We found four distinct designs over the nine possibilities where we distinguish between designs based on the determinant of their information matrix while holding the covariance matrix, V, fixed at one value of η_1 and η_2. One of the designs was optimal for all six scenarios where $\eta_2 < 10$. When $\sigma_{\gamma_1}^2 = \sigma_{\gamma_2}^2 = 0{\cdot}1$ and $\sigma_\varepsilon^2 = 1$, the D-efficiencies of the four distinct designs found, relative to the optimal design for $\eta_1 = \eta_2 = 0{\cdot}1$, range from 96·36% to 100%. When $\sigma_{\gamma_1}^2 = \sigma_{\gamma_2}^2 = \sigma_\varepsilon^2 = 1$, the D-efficiencies of the designs found, relative to the optimal design for $\eta_1 = \eta_2 = 1$, range from 95·69% to 100%. The six designs that were the same all had full efficiency for these values of the variance components. When $\sigma_{\gamma_1}^2 = \sigma_{\gamma_2}^2 = 10$ and $\sigma_\varepsilon^2 = 1$, the D-efficiencies of the designs found, relative to the optimal design for $\eta_1 = \eta_2 = 10$, range from 89·63% to 100%. Detailed results are in Table 6.

For the second design problem in this sensitivity study, we also investigated whether a design could be found that is more robust to the values of the variance ratios η_1 and η_2 than the ones generated using our algorithm, which assumes a point prior for their values. To this end, we implemented a Bayesian approach in which we used independent χ^2 distributions with two degrees of freedom as priors for η_1 and η_2. The 2·5%, 50% and 97·5% percentiles of these prior distributions are 0·05, 1·39 and 7·38, respectively, so that we allowed for considerable uncertainty about the variance ratios. It turns out that the Bayesian D-optimal split-split-plot design is equivalent to the non-Bayesian design obtained for $\eta_2 = 0{\cdot}1$ and $\eta_2 = 1$, so that the Bayesian approach did not provide the desired robustness.

The concern is that misspecifying η_1 and η_2 for the purpose of designing a split-split-plot experiment with our algorithm might lead to the use of a design that is very inefficient. The last study shows that it is better to choose large values for η_2 since designs that were optimal assuming small values did not perform well if the true values were larger, whereas designs that were optimal assuming large values still perform reasonably well if the true values are smaller.

In both of the sensitivity studies reported here, it is more important to correctly specify the ratio of the subplot variance to the error variance, η_2, than the ratio of the whole-plot variance to the error variance, η_1, because the quality of the optimal split-split-plot designs obtained changes more with η_2 than with η_1.

7. Discussion

We have provided an algorithmic approach to the construction of D-optimal split-split-plot designs. We have shown how to use this approach to create screening designs. We have also considered the effect that changing the two relevant variance ratios has on the D-optimal design.

For more examples of screening designs, some examples of split-split-plot response surface designs and for a discussion of several attractive design options for the cheese production experiment in Schoen (1999), we refer the reader to an unpublished working paper by the authors, where it is also discussed what approach can be used when economic considerations dictate a number of whole plots that does not allow for the estimation of the whole-plot variance.

In this article, we have assumed that the number of subplots is an integer multiple of the number of whole plots, and that the total number of runs is an integer multiple of the number of subplots, $n = b_1 b_2 k$. This scenario is perhaps the most common one in practice, where these parameters are usually dictated by the logistics of the experiment and by time and cost constraints. Nevertheless, there exist experimental situations where there are no hard constraints on the number of whole plots, the number of subplots within whole plots and/or the number of runs within subplots. In such cases, experimenters may want to deviate from the scenario discussed here and still use the algorithm we outlined above to generate a D-optimal split-split-plot design for the unit structure that they have in mind. However, in such cases the time-saving update formulae for the information matrix described in § 3·2 can no longer be used because Theorem 1 is valid only for the scenario we focused on in this article. Another approach would be to modify our algorithm so that it generates the D-optimal numbers of whole plots, subplots within whole plots and runs within subplots, in addition to the optimal design points. It seems likely that allowing for different sizes of whole plots and subplots might improve the efficiency of D-optimal response surface designs. Similar research was done by Goos & Vandebroek (2004), who presented an algorithm for determining the optimal split-plot structure of an experimental design. The modified split-split-plot algorithm could also be extended for computing D-optimal designs that do not just focus on the precise estimation of the factor effects contained within β, but also on that of the variance components $\sigma^2_{\gamma_1}$, $\sigma^2_{\gamma_2}$ and σ^2_{ε}.

Apart from these extensions, some interesting research questions remain. More work could be done to attempt to bound the effect of misspecification of the two variance ratios on which the D-optimal design depends. Another potentially interesting topic for future research is the use of Bayesian optimal design criteria. Such criteria could be used to prioritize the precise estimation of the main effects when constructing optimal split-plot and split-split-plot designs. The fast update formulae in the Appendix for the inverse of the information matrix also suggest that it is possible to compute A-optimal and V-optimal, also called I-optimal, split-split-plot designs at an acceptable computational cost. Finally, the design of experiments involving hard-to-change factors and nonnested unit structures would be a useful topic for further investigation.

A version of the algorithm presented in this article is available in the commercial software JMP. The algorithm can handle any scenario where the number of whole plots, the number of subplots and the number of runs are fixed by the experimenter. The datasets contained in the article are available from the authors as JMP data files or Excel files.

ACKNOWLEDGEMENT

The authors are grateful to Eric Schoen for his suggestions regarding the combinatorial construction of several designs that were used as benchmarks for the algorithm.

APPENDIX

The update formulae derived here extensively use two matrix-algebraic results for matrices of the form $R + STU$, where R and T are nonsingular $r \times r$ and $t \times t$ matrices, respectively, and S and U are $r \times t$ and $t \times r$ matrices, respectively. The first result is

$$|R + STU| = |R||T||T^{-1} + UR^{-1}S| = |R|\left|I_t + TUR^{-1}S\right|, \tag{A1}$$

while the second result is

$$(R + STU)^{-1} = R^{-1} - R^{-1}S(T^{-1} + UR^{-1}S)^{-1}UR^{-1}. \tag{A2}$$

Detailed proofs of these results can be found in Harville (1997), for example. The results are especially useful when $|R|$, $|T|$, R^{-1} and T^{-1} are easy to obtain. This is exactly the case in the construction of D-optimal split-split-plot designs because T is a diagonal matrix and because $|R|$ and R^{-1} are being stored during the entire operation of the design construction algorithm.

Changes to the level of an easy-to-change factor

A change to the level of an easy-to-change factor in the lth run of the jth subplot within the ith whole plot only affects the corresponding row in the model matrix X. Such a change does not require the information matrix to be recomputed from scratch. In order to see this, denote the original of the affected row by $f'(w_i, s_{ij}, t_{ijl})$ and its modified version by $f'(w_i, s_{ij}, t^*_{ijl})$. The updated versions of $X_i'1_{b_2k}$ and $X_{ij}'1_k$ can then be obtained using (11) and (12), and the information matrix (8) can be updated using

$$\begin{aligned} M^* = M &- \sigma_\varepsilon^{-2} f(w_i, s_{ij}, t_{ijl}) f'(w_i, s_{ij}, t_{ijl}) + c_1 \left(X_i'1_{b_2k}\right)\left(X_i'1_{b_2k}\right)' + c_2 (X_{ij}'1_k)(X_{ij}'1_k)' \\ &+ \sigma_\varepsilon^{-2} f\left(w_i, s_{ij}, t^*_{ijl}\right) f'\left(w_i, s_{ij}, t^*_{ijl}\right) - c_1 \left(X_i^{*'}1_{b_2k}\right)\left(X_i^{*'}1_{b_2k}\right)' - c_2 \left(X_{ij}^{*'}1_k\right)\left(X_{ij}^{*'}1_k\right)', \end{aligned}$$

where M and M^* represent the information matrix before and after the change, respectively, and X_i^* and X_{ij}^* represent the updated versions of X_i and X_{ij}, respectively. This can be rewritten as

$$M^* = M + \begin{bmatrix} f'(w_i, s_{ij}, t_{ijl}) \\ \left(X_i'1_{b_2k}\right)' \\ (X_{ij}'1_k)' \\ f'\left(w_i, s_{ij}, t^*_{ijl}\right) \\ \left(X_i^{*'}1_{b_2k}\right)' \\ \left(X_{ij}^{*'}1_k\right)' \end{bmatrix}' \begin{bmatrix} -\sigma_\varepsilon^{-2} & 0 & 0 & 0 & 0 & 0 \\ 0 & c_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & c_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_\varepsilon^{-2} & 0 & 0 \\ 0 & 0 & 0 & 0 & -c_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -c_2 \end{bmatrix} \begin{bmatrix} f'(w_i, s_{ij}, t_{ijl}) \\ (X_i'1_{b_2k})' \\ (X_{ij}'1_k)' \\ f'\left(w_i, s_{ij}, t^*_{ijl}\right) \\ \left(X_i^{*'}1_{b_2k}\right)' \\ \left(X_{ij}^{*'}1_k\right)' \end{bmatrix},$$

which is of the form $R + STU$ so that the results in (A1) and (A2) can be used for computing $|M^*|$ and M^{*-1}. Defining D_1 and U_1 as in (9) and (10), and substituting these into (A1) and (A2) leads to the update formulae for the determinant and the inverse of the information matrix

$$|M^*| = |M|\left|I_6 + D_1U_1'M^{-1}U_1\right| = |M||D_1|\left|D_1^{-1} + U_1'M^{-1}U_1\right|,$$

and

$$M^{*-1} = M^{-1} - M^{-1}U_1\left(D_1^{-1} + U_1'M^{-1}U_1\right)^{-1}U_1'M^{-1}.$$

The matrix product $M^{-1}U_1$ plays a key role in these update formulae. The second expression for updating the determinant is slightly less computationally involved than the first as D_1 is a constant diagonal matrix for given values of b_1, b_2, k and the three variance components.

Changes to the level of a hard-to-change factor

A change to the level of a hard-to-change factor cannot be made for a single run because the level of such a factor has to be constant for all the runs in a given subplot. Such a change can therefore only be made to all the runs in an entire subplot. Modifying the level of a hard-to-change factor level in the jth subplot within whole plot i therefore results in the following information matrix:

$$\begin{aligned} M^* = M &- \sigma_\varepsilon^{-2} X_{ij}' X_{ij} + c_1 \left(X_i' 1_{b_2 k}\right)\left(X_i' 1_{b_2 k}\right)' + c_2 (X_{ij}' 1_k)(X_{ij}' 1_k)' \\ &+ \sigma_\varepsilon^{-2} X_{ij}^{*'} X_{ij}^* - c_1 \left(X_i^{*'} 1_{b_2 k}\right)\left(X_i^{*'} 1_{b_2 k}\right)' - c_2 (X_{ij}^{*'} 1_k)(X_{ij}^{*'} 1_k)', \end{aligned} \tag{A3}$$

where

$$X_{ij}^* = \left[f\left(w_i, s_{ij}^*, t_{ij1}\right) \ \ldots \ f\left(w_i, s_{ij}^*, t_{ijk}\right) \right]'$$

is the modified version of

$$X_{ij} = [f(w_i, s_{ij}, t_{ij1}) \ \ldots \ f(w_i, s_{ij}, t_{ijk})]',$$

and

$$X_i^{*'} 1_{b_2 k} = X_i' 1_{b_2 k} - X_{ij}' 1_k + X_{ij}^{*'} 1_k.$$

Each X_{ij}^* and X_{ij} only differ in the columns corresponding to the factor whose level is changed, its interactions and higher-order terms involving it. Now, (A3) can be written as $M^* = M + U_2 D_2 U_2'$, where

$$D_2 = \operatorname{diag}\left(-\sigma_\varepsilon^{-2} I_k, c_1, c_2, \sigma_\varepsilon^{-2} I_k, -c_1, -c_2\right),$$

$$U_2 = \left[X_{ij}' \quad X_i' 1_{b_2 k} \quad X_{ij}' 1_k \quad X_{ij}^{*'} \quad X_i^{*'} 1_{b_2 k} \quad X_{ij}^{*'} 1_k \right]'.$$

As a result,

$$|M^*| = |M| \left| I_{2(k+2)} + D_2 U_2' M^{-1} U_2 \right| = |M| |D_2| \left| D_2^{-1} + U_2' M^{-1} U_2 \right|,$$

and

$$M^{*-1} = M^{-1} - M^{-1} U_2 \left(D_2^{-1} + U_2' M^{-1} U_2\right)^{-1} U_2' M^{-1}.$$

Changes to the level of a very-hard-to-change factor

Finally, after a change in one of the levels of a very-hard-to-change factor, the update is even more involved as such a change has an impact on all $b_2 k$ runs in that stratum. If the change is performed in the ith whole plot, then the new information matrix can be computed as

$$\begin{aligned} M^* = M &- \sigma_\varepsilon^{-2} X_i' X_i + c_1 \left(X_i' 1_{b_2 k}\right)\left(X_i' 1_{b_2 k}\right)' + c_2 \sum_{j=1}^{b_2} (X_{ij}' 1_k)(X_{ij}' 1_k)' \\ &+ \sigma_\varepsilon^{-2} X_i^{*'} X_i^* - c_1 \left(X_i^{*'} 1_{b_2 k}\right)\left(X_i^{*'} 1_{b_2 k}\right)' - c_2 \sum_{j=1}^{b_2} \left(X_{ij}^{*'} 1_k\right)\left(X_{ij}^{*'} 1_k\right)', \end{aligned} \tag{A4}$$

where

$$X_{ij}^* = \left[f\left(w_i^*, s_{ij}, t_{ij1}\right) \ \ldots \ f\left(w_i^*, s_{ij}, t_{ijk}\right) \right]'$$

is the modified version of X_{ij}, and

$$X_i^* = \left[f\left(w_i^*, s_{i1}, t_{i11}\right) \ \ldots \ f\left(w_i^*, s_{ib_2}, t_{ib_2 k}\right) \right]'$$

is the modified version of X_i. Each X_{ij}^* and X_{ij} only differ in the columns corresponding to the factor whose level is changed, its interactions and higher-order terms involving it. Now, (A4) can be written as

$M^* = M + U_3 D_3 U_3'$, where

$$D_3 = \text{diag}\left(-\sigma_\varepsilon^{-2} I_{b_2 k}, c_1, c_2 I_{b_2}, \sigma_\varepsilon^{-2} I_{b_2 k}, -c_1, -c_2 I_{b_2}\right),$$

$$U_3 = \left[\begin{array}{cccccc} X_i' X_i' 1_{b_2 k} & X_{i1}' 1_k \ldots X_{ib_2}' 1_k & X_i^{*'} X_i^{*'} 1_{b_2 k} & X_{i1}^{*'} 1_k \ldots X_{ib_2}^{*'} 1_k \end{array}\right]'.$$

As a result,

$$|M^*| = |M| \left| I_{2(b_2 k + b_2 + 1)} + D_3 U_3' M^{-1} U_3 \right| = |M||D_3| \left| D_3^{-1} + U_3' M^{-1} U_3 \right|,$$

and

$$M^{*-1} = M^{-1} - M^{-1} U_3 \left(D_3^{-1} + U_3' M^{-1} U_3\right)^{-1} U_3' M^{-1}.$$

REFERENCES

BINGHAM, D. R., SCHOEN, E. D. & SITTER, R. R. (2004). Designing fractional factorial split-plot experiments with few whole-plot factors. *Appl. Statist.* **53**, 325–39. Corrigendum, **54**, 955–8.

BRIEN, C. J. & BAILEY, R. A. (2006). Multiple randomizations. *J. R. Statist. Soc.* B **68**, 571–609.

EDMONDSON, R. N. (1991). Agricultural response surface experiments based on four-level factorial designs. *Biometrics* **47**, 1435–48.

GOOS, P. (2002). *The Optimal Design of Blocked and Split-Plot Experiments*. New York: Springer.

GOOS, P. (2006). The usefulness of optimal design for generating blocked and split-plot response surface experiments. *Statist. Neer.* **60**, 361–78.

GOOS, P. & DONEV, A. N. (2007). Tailor-made split-plot designs with mixture and process variables. *J. Qual. Technol.* **39**, 326–39.

GOOS, P. & VANDEBROEK, M. (2001). Optimal split-plot designs. *J. Qual. Technol.* **33**, 436–50.

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

GOOS, P. & VANDEBROEK, M. (2004). Outperforming completely randomized designs. *J. Qual. Technol.* **36**, 12–26.

HARVILLE, D. A. (1997). *Matrix Algebra from a Statistician's Perspective*. New York: Springer.

HARVILLE, D. A. & JESKE, D. R. (1992). Mean squared error of estimation or prediction under a general linear model. *J. Am. Statist. Assoc.* **87**, 724–31.

JONES, B. & GOOS, P. (2007). A candidate-set-free algorithm for generating D-optimal split-plot designs. *Appl. Statist.* **56**, 347–64.

KACKAR, R. N. & HARVILLE, D. A. (1984). Approximations for standard errors of estimators of fixed and random effects in mixed linear models. *J. Am. Statist. Assoc.* **79**, 853–62.

KENWARD, M. G. & ROGER, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* **53**, 983–97.

LENTH, R. V. (1989). Quick and easy analysis of unreplicated factorials. *Technometrics* **31**, 469–73.

LOEPPKY, J. & SITTER, R. R. (2002). Analyzing unreplicated blocked fractional factorial and fractional factorial split-plot designs. *J. Qual. Technol.* **34**, 229–43.

LOUGHIN, T. M. & NOBLE, W. (1997). A permutation test for effects in an unreplicated factorial design. *Technometrics* **39**, 180–90.

MEE, R. & BATES, R. L. (1998). Split-lot designs: experiments for multistage batch processes. *Technometrics* **40**, 127–40.

MEYER, R. K. & NACHTSHEIM, C. J. (1995). The coordinate-exchange algorithm for constructing exact optimal experimental designs. *Technometrics* **37**, 60–9.

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

SUN, D. X. & WU, C. F. J. (1993). Statistical properties of Hadamard matrices of order 16. In *Quality Through Engineering Design,* Ed. W. Kuo, pp. 169–79. New York: Elsevier.

TRINCA, L. A. & GILMOUR, S. G. (2001). Multi-stratum response surface designs. *Technometrics* **43**, 25–33.

WEBB, D., LUCAS, J. M. & BORKOWSKI, J. J. (2004). Factorial experiments when factor levels are not necessarily reset. *J. Qual. Technol.* **36**, 1–11.

[*Received May* 2008. *Revised October* 2008]