a, citation and similar papers at core.ac.uk
brought to
provided by Elsevier -

ELSEVIER Linear Algebra and its Applications 321 (2000) 239–280
www.elsevier.com/locate/laa

Reparametrisation of interest in non-uniform factorial designs

A. Kobilinsky

Laboratoire de Biométrie, INRA Versailles, Route de Saint-Cyr, F 78026 Versailles Cedex, France

Received 5 March 1997; accepted 10 October 1999

Submitted by G.P.H. Styan

Abstract

A definition of factorial effects relying on the treatment structure defined by the hierarchies is proposed. It applies to a non-uniform situation, where the number of levels of a nested factor within the classes defined by each set of levels of its nesting factors may vary. A reparametrisation whose parameters belongs to these factorial effects is obtained. The development is based on the notion of reference treatment design, a conceptual design that can be used as a basis of comparison to assess the properties of any factorial design. © 2000 Elsevier Science Inc. All rights reserved.

Keywords: Analysis of variance; ANOVA; Orthogonal designs; Block structure; Projective limit; Partially ordered set; POSET; Hierarchy; Nesting factor; Lattice

1. Introduction

Consider a study to determine the influence on a response y of two crossed factors A, B. We denote by T_A and T_B their respective sets of levels. The set T of feasible treatments is the cartesian product $T = T_A \times T_B$. The expectation of the response when treatment $(a, b) \in T$ is experimented is denoted by $\tau(a, b)$ and is called the effect of treatment (a, b). Marginal means of these treatment effects are usually introduced. These means may be weighted and are denoted with the usual dot notation. They are:

E-mail address: kobi@versailles.inra.fr (A. Kobilinsky).

0024-3795/00/$ - see front matter © 2000 Elsevier Science Inc. All rights reserved.
PII: S 0 0 2 4 - 3 7 9 5 (9 9) 0 0 2 3 7 - 2

the general mean: $\tau(\bullet,\bullet) = \sum_a \sum_b W(a,b)\tau(a,b)$,
the means by level of A: $\tau(a,\bullet) = \sum_b W_B(b)\tau(a,b)$,
the means by level of B: $\tau(\bullet,b) = \sum_a W_A(a)\tau(a,b)$,

where the weights $W(a,b)$, $W_A(a)$, $W_B(b)$ satisfy

$$\sum_a W_A(a) = 1, \quad \sum_b W_B(b) = 1, \quad W(a,b) = W_A(a)W_B(b). \tag{1}$$

The use of a system of weights $W_B(b)$ independent of a to define the means by level of A guarantees that the differences $\tau(a,\bullet) - \tau(a',\bullet)$ can be attributed to the factor A and not to the factor B.

The general mean, main effects and interaction of factors A and B are defined from these means as indicated in Table 1.

In most cases, the weights $W_A(a)$ are chosen equal to $1/|T_A|$, the weights $W_B(b)$ equal to $1/|T_B|$ and the weights $W(a,b)$ are then all equal to $1/|T|$. But it can be natural in some circumstances to use unequal weights. Scheffe [15] gives such an example. Factor A is the variety of cotton, B is the location in California. If a single variety is to be selected for all of California, it may be reasonable to weight the different locations with weights $W_B(b)$ proportional to the total acreages of cotton in the corresponding regions.

In non-uniform cases, when the number of levels of a nested factor within the classes defined by each set of levels of its nesting factors may vary, the weights cannot generally be chosen equal.

Consider the following very simple example. There are three treatments, a control and two other variants of a new treatment to be compared to the control. A possible way to deal with that situation is to introduce a factor A whose levels are 0 for the control, 1 for the new treatments, then a factor B nested within A, with levels 0 for the control, 1 and 2 for the two other treatments. We denote by T_A and T_B the set of levels of the two factors, by $\phi_{AB} : T_B \to T_A$ the mapping defined by $\phi_{AB}(0) = 0$, $\phi_{AB}(1) = 1$, $\phi_{AB}(2) = 1$ which gives for each level of B the corresponding level of A.

The treatments can be represented by the pairs $(a,b) \in T_A \times T_B$ which satisfy $\phi_{AB}(b) = a$. We denote as previously by T the set of these treatments and by $\tau(a,b)$ the effect of treatment $(a,b) \in T$. Table 2 gives the corresponding means and factorial effects. The weights $W(a,b)$, $W_A(a)$, $W_B(b)$ must satisfy in that hierarchical case the following constraints:

$$\sum_a W_A(a) = 1, \quad \sum_{b\in\phi_{AB}^{-1}(a)} W_B(b) = 1, \quad W(a,b) = W_A(a)W_B(b). \tag{2}$$

If $a = \phi_{AB}(b)$, we say that b is *nested within a*, or more simply is *within a*. It is natural to choose the weights $W_B(b)$ equal within each level a of A. This leads to $W_B(0) = 1$, $W_B(1) = W_B(2) = 1/2$. The weights $W_A(a)$ may then be chosen equal to $1/3$ for $a = 0$ and $2/3$ for $a = 1$, which makes the $W(a,b)$ all equal to $1/3$. Alternatively they may be chosen equal to $1/2$, which gives $W(0,0) = 1/2$ and

Table 1
Definition of factorial effects in the two-way layout

General mean	$\mu = \tau(\bullet, \bullet)$
Main effect of a	$\alpha_a = \tau(a, \bullet) - \mu$
Main effect of b	$\beta_b = \tau(\bullet, b) - \mu$
Interaction effect of (a, b)	$\gamma_{ab} = \tau(a, b) - (\mu + \alpha_a + \beta_b)$

Table 2
Definition of factorial effects in the two-way nested layout

General mean	$\tau(\bullet, \bullet) = \sum_{(a,b)\in T} W(a, b)\tau(a, b)$
Means by level of A	$\tau(a, \bullet) = \sum_{b\in\phi_{AB}^{-1}(a)} W_B(b)\tau(a, b)$
General mean	$\mu = \tau(\bullet, \bullet)$
Main effect of a	$\alpha_a = \tau(a, \bullet) - \mu$
Main effect of b within $a = \phi_{AB}(b)$	$\beta_{ab} = \tau(a, b) - \tau(a, \bullet)$

$W(1, 1) = W(1, 2) = 1/4$. In that latter case, the control is given twice the weight of the two other treatments in the general mean. Of course any other intermediate choice is possible.

It is in general not very difficult to define similarly the factorial effects of interest in a given more complex situation involving both nesting and crossing. However general softwares must be able to deal with any system of weights and any kind of treatment structure. There is thus a need to have a clear and general process to define the factorial effects from this structure even when it is not uniform.

1.1. Reference design

Such a general process has been clearly described for orthogonal designs [24]. Whatever nature, orthogonal or not, has the actual design under consideration, this process can be used to define the factorial effects provided the set T of all feasible treatments, with suitable weight function W and model $\mathscr{E}$, itself defines an orthogonal design. The latter is called the *reference design*. It is a conceptual one, used to define factorial effects, study the aliasing or assess, by comparison with it, the quality of any actual design under investigation.

In the first example with two crossed factors, the orthogonality of the reference design $T = T_A \times T_B$ follows from condition (1) imposed to the weights. More generally, assume there are n crossed factors with sets of levels $T_1, \ldots, T_n$, and that the weight function W is a product of marginal weights:

$$W(t_1, \ldots, t_n) = W_1(t_1) \cdots W_n(t_n) \quad \text{with} \quad \sum_{t_i\in T_i} W_i(t_i) = 1 \text{ for all } i. \tag{3}$$

Let $I = \{1, \ldots, n\}$ and for each subset J of I, denote by ϕ_J the canonical projection $(t_i)_{i \in I} \mapsto (t_i)_{i \in J}$ of index J. Let then $\mathscr{E}$ be the family of subsets of I containing, besides the empty set associated with the constant factor and the sets $\{1\}, \ldots, \{n\}$ associated with the main effects, all the subsets associated with non-zero interactions. The family $\mathscr{E}$, possibly completed in a suitable way, can be assumed to be closed for the intersection. Then the triplet $(T, W, \mathscr{E})$ defines a reference orthogonal design and thus induces a decomposition into meaningful factorial effects.

Note that this kind of reference design can also be used when there are nested factors, provided each factor can be identified with a canonical projection ϕ_J. In that case, if J is a subset in $\mathscr{E}$ and $i \in J$, any factor j nesting the factor i must also belong to J. Therefore if i is nested within some other factor j, the singleton $\{i\}$ does not pertain to $\mathscr{E}$.

That kind of reference design was used to study aliased effects and derive principal factor efficiencies in several contexts [8,10,12]. The corresponding block structure, formed by the partitions induced on T by the factors, has been studied under the name *poset block structure* [4,6]. If the weights are equal, the associated factorial effects are those which are generally taken into account by variance analysis software in the uniform case. The associated linear functions of the parameters are known, when they are estimable, as the estimable functions of type III [16,18].

However, the structure associated with this kind of reference design is necessarily uniform. Section 4 shows how an orthogonal reference design can be deduced from the knowledge of nesting relations in a very general, possibly non-uniform, context. Section 5 gives then a process leading to a reparametrisation whose parameters belongs to the factorial effects induced by this orthogonal reference design.

The reference design can also be used in variance analysis to provide a rigorous and easy definition of adjusted means, hence of most interesting non-standard linear functions of the parameters (Section 6).

To motivate this rather technical development on non-uniform designs, we first introduce in Section 2 some considerations on the different strategies nowadays used in ANOVA.

In Section 3, we then recall the main notions needed to define and check design orthogonality. The notations take the weight function into account.

2. Factorial effects, tests of hypothese in ANOVA

The definition of factorial effects and associated sum of squares in unbalanced design is the matter of a long controversy, which clearly appears in the article with discussion [14] and is well summed up in [17]. It is still alive today [3,9,19].

As written in [17], the linear modelers can be divided into two camps, the R-notationers and the R*-notationers. To test a factorial effect, main effect or interaction, the R-notationers use the reduction R of the residual sum of square due to the introduction of this factorial effect in the model. They do not reparameterise the model nor introduce constraints on the parameters. Hence to test a factorial effect, they have to exclude other effects imbedding it from the model. For instance, let *A*, *B*, *C* be three factors such that *C* is nested in *A*, and *B* is crossed with *A* and *C*. If the model is $A + B + AB + AC + ABC$, R-notationers usually compute the *AB* sum of square in the model without *ABC*, the *A* sum of squares in the model without *AB*, *AC*, *ABC* that is in the additive model $A + B$.

On the contrary, $R^\star$ notationers define and test all factorial effects in the same unique whole model, using marginal means as in Table 1 to define factorial effects imbedded in other effects of the model. To do so, they have to introduce a system of weights satisfying relations like those in (1) and (2), or the equivalent system of constraints on the parameters.

In uniform situations, a natural uniquely defined system is the uniform weighting which is generally the only one adopted by ANOVA softwares. We show in Section 2.2 that this uniform weighting can be completely inadequate to analyse some very useful designs even in a case including only crossed factors.

In non-uniform situations with nested factors, the example in the introduction shows that things are far more complicated. Section 2.3 considers two other simple examples with non-uniform data. Analyses of variance performed on these examples give results which vary from one software to the other in an incomprehensible manner. The fact had already been noticed by Searle [19] who concluded that it is better not to use the $R^\star$-approach (i.e. type III sum of squares) until things are clarified.

This article clarifies the situation by showing how to define a suitable system of weights in every situation. To study the properties of the associated reparametrisation in the more general case, we need some notions of algebra which may appear quite sophisticated for the problem considered. But the results are in fact very simple and allow to propose a clear and coherent way to perform ANOVA in non-uniform situations.

However, to prompt R-practitioners to read what follows, we first show in Section 2.1 all the difficulties raised by the R-approach even in the simple case of an unbalanced two-way layout.

2.1. Difficulties with the R-approach

At first sight, the R-approach may appear simpler than the $R^\star$ one because it does not require the somewhat subjective choice of a system of weights to select which sums of squares and associated contrasts are inspected. However, in the R-approach, the expectation within the whole model of the contrasts or sum of square associated with a non-maximal factorial effect is design dependent. This generally makes these contrasts or sum of square uneasy to interpret, and forbids comparison between homologous effects coming from designs with different numbers of replications.

To illustrate this point, let us consider again a study with one response *y* and two crossed factors *A*, *B*. We assume that *A* and *B* have two levels coded -1 and 1 and that the number of replications of the treatments is as given in Table 3. There is only

Table 3
An unbalanced design with two two-levels factors

		B	
		–1	1
A	–1	1	m
	1	m	m

one observation for treatment $(-1, -1)$ and m for each of the other treatments. As m increases, the design is increasingly non-orthogonal and unbalanced. Of course no one would use such a design when $m \gg 1$, but this simple situation makes it possible to understand what can occur in a much more less trivial way when the number of factors exceeds 2.

We denote by y_{abj} the jth response for treatment (a, b), where (a, b) is one of the four treatments $(-1, -1)$, $(-1, 1)$, $(1, -1)$, $(1, 1)$, and let $\tau(a, b) = E(y_{abj})$. The factorial effects are defined as in Table 1, with constant weights $W(a, b) = 1/4$. Since there are only two levels for each factor, it is easy to check that $\alpha_a = a\alpha$, $\beta_b = b\beta$, $\gamma_{ab} = ab\gamma$ where

$$\begin{aligned}
\alpha &= \tfrac{1}{4}(\tau(1,1) + \tau(1,-1) - \tau(-1,1) - \tau(-1,-1)) = \tfrac{1}{2}(\tau(1,\bullet) - \tau(-1,\bullet)),\\
\beta &= \tfrac{1}{4}(\tau(1,1) - \tau(1,-1) + \tau(-1,1) - \tau(-1,-1)) = \tfrac{1}{2}(\tau(\bullet,1) - \tau(\bullet,-1)),\\
\gamma &= \tfrac{1}{4}(\tau(1,1) - \tau(1,-1) - \tau(-1,1) + \tau(-1,-1)).
\end{aligned} \tag{4}$$

The equality $\gamma_{ab} = \tau(a, b) - (\mu + \alpha_a + \beta_b)$ in the last row of Table 1 can be written as

$$\tau(a, b) = \mu + a\alpha + b\beta + ab\gamma. \tag{5}$$

It leads to the linear model

$$E(y) = X\theta = X_1\theta_1 + X_2\gamma,$$

where y is the vector of $3m + 1$ responses, $\theta = (\mu, \alpha, \beta, \gamma)'$, $\theta_1 = (\mu, \alpha, \beta)'$ and X is the matrix in Table 4 which is decomposed for further use into the submatrices X_1 including the three columns associated with μ, α, β and the one column matrix X_2 associated with γ.

In the R*-strategy, θ is estimated by $\tilde{\theta} = (X'X)^{-1}X'y$ (we use a tilde to denote an R* estimate). It is equivalent to estimating each $\tau(a, b)$ by the mean $y_{ab\bullet}$ of the responses to treatment (a, b) and then to get the estimates of α, β, γ by replacing each $\tau(a, b)$ in (4) by its estimate $y_{ab\bullet}$. Thus

$$\begin{aligned}
\tilde{\alpha} &= \tfrac{1}{4}(y_{1,1,\bullet} + y_{1,-1,\bullet} - y_{-1,1,\bullet} - y_{-1,-1,\bullet}),\\
\tilde{\gamma} &= \tfrac{1}{4}(y_{1,1,\bullet} - y_{1,-1,\bullet} - y_{-1,1,\bullet} + y_{-1,-1}),
\end{aligned} \tag{6}$$

Users of the R-strategy estimate $\theta_1 = (\mu, \alpha, \beta)'$ only in the model with $\gamma = 0$, that is by $\hat{\theta}_1 = (X_1'X_1)^{-1}X_1'y$. The matrix $(X_1'X_1)^{-1}$ is given in Table 4. Using it, it is easy to check that the estimate of α in this context is

Table 4
Matrices X, X_1, $X'X$, $(X_1'X_1)^{-1}$ for the example of Table 3

$$X = \overbrace{\begin{array}{ccc|c} \multicolumn{3}{c|}{\overbrace{\qquad\qquad}^{X_1}} & \overbrace{\quad}^{X_2} \\ \mu & \alpha & \beta & \gamma \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & -1 & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & -1 \\ \cdots & \cdots & \cdots & \cdots \end{array}}
\begin{array}{l} \\ \\ \\ \updownarrow m \text{ rows} \\ \\ \updownarrow m \text{ rows} \\ \\ \updownarrow m \text{ rows} \end{array}$$

$$X'X = \begin{bmatrix} 3m+1 & m-1 & m-1 & -m+1 \\ m-1 & 3m+1 & -m+1 & m-1 \\ m-1 & -m+1 & 3m+1 & m-1 \\ -m+1 & m-1 & m-1 & 3m+1 \end{bmatrix}$$

$$(X_1'X_1)^{-1} = \begin{bmatrix} \frac{m+1}{2m(m+3)} & -\frac{m-1}{4m(m+3)} & -\frac{m-1}{4m(m+3)} \\ -\frac{m-1}{4m(m+3)} & \frac{m+1}{2m(m+3)} & \frac{m-1}{4m(m+3)} \\ -\frac{m-1}{4m(m+3)} & \frac{m-1}{4m(m+3)} & \frac{m+1}{2m(m+3)} \end{bmatrix}$$

$$\hat{\alpha} = \frac{1}{m+3}\left[\frac{m+1}{2}(y_{1,1,\bullet} - y_{-1,1,\bullet}) + (y_{1,-1,\bullet} - y_{-1,-1,\bullet})\right]. \tag{7}$$

The estimate of β is similar. The variances of $\tilde{\alpha}$ and $\hat{\alpha}$ can be deduced from those of the means. Under the usual assumption $\mathrm{Var}(y) = \sigma^2\mathbf{I}$, we have since $y_{-1,-1,\bullet} = y_{-1,-1}$

$$\mathrm{var}(y_{-1,-1,\bullet}) = \sigma^2, \quad \mathrm{var}(y_{1,1,\bullet}) = \mathrm{var}(y_{-1,1,\bullet}) = \mathrm{var}(y_{1,-1,\bullet}) = \frac{\sigma^2}{m},$$

hence

$$\mathrm{var}\big(\hat{\alpha}\big) = \frac{m+1}{2m(m+3)}\sigma^2, \quad \mathrm{var}(\tilde{\alpha}) = \frac{\sigma^2}{16}\left(1 + \frac{3}{m}\right). \tag{8}$$

If $\gamma = 0$, both $\hat{\alpha}$ and $\tilde{\alpha}$ are unbiased estimates of α and (8) then shows that $\hat{\alpha}$ is a better estimate of α than $\tilde{\alpha}$. Note however that the ratio

$$\frac{\mathrm{var}(\tilde{\alpha})}{\mathrm{var}(\hat{\alpha})} = \frac{1}{8}\frac{(m+3)^2}{m+1}$$

increases with m, but remains smaller than 2 if $m \leqslant 10$ so that the superiority of the R-estimate over the R*-one becomes decisive for $\gamma = 0$ only for very large values of m.

But in such an experiment, one can never assume $\gamma = 0$. Even if the test of the interaction failed to reject this hypothesis, this does not mean that $\gamma = 0$, but only that γ is too small to detect if it is greater or smaller than 0. To take this into account, there are two possible attitudes.

(1) Choose the R-approach, but carefully look at the expectation of $\hat{\alpha}$ and $\hat{\beta}$ for the interpretation. In the example, the expectation of $\hat{\alpha}$:

$$E(\hat{\alpha}) = \frac{1}{m+3}\left[\frac{m+1}{2}\big(\tau(1,1) - \tau(-1,1)\big) + \big(\tau(1,-1) - \tau(-1,-1)\big)\right],$$

gives, when m is large, nearly all the weight to the A-effect for $b = 1$. Note that if the number of replications in cells $(1, 1)$ and $(-1, -1)$ were interchanged, the A-effect would on the contrary give all the weight to level $b = -1$. Thus if $\gamma \neq 0$, the definition of the A-effect strongly depends on the experiment. Provided one is aware of that and does not try to compare estimates $\hat{\alpha}$ coming from different experiments, it may seem sensible to adapt in this way the definition of the A-effect to the data.

But continuation of this logic, which selects the contrasts examined according to the data to make the better use of the available information, should also lead to the examination of the A-effect in the model excluding β as well as γ. In this model, $\tau(a, b) = \mu + a\alpha$, α is estimated by

$$\check{\alpha} = \frac{1}{2}\left(\frac{y_{1,1,\bullet} + y_{1,-1,\bullet}}{2} - \frac{my_{-1,1,\bullet} + y_{-1,-1,\bullet}}{m+1}\right)$$

with a variance

$$\operatorname{var}(\check{\alpha}) = \frac{1}{4}\left(\frac{1}{2m} + \frac{1}{m+1}\right)\sigma^2,$$

which is even lower than $\operatorname{var}(\hat{\alpha})$. The expectation of this $\check{\alpha}$ under the whole model becomes even more difficult to interpret as it is a function of the three parameters of model (5) which can be non-zero even when β is the only non-zero parameter.

Such an approach using nested models to explore the data has thus the advantage of adapting itself to the data to make the contrasts examined more precise. But it leads to contrasts that are data dependent, difficult to interpret, the more so as the model becomes more complex, involving more factors, more interactions and possibly a mixture of qualitative and quantitative factors. This approach should therefore be avoided unless a strong non-orthogonality induces a drastic increase of variance on some parameters. An extreme case is when the columns X_δ and X_η associated with two parameters δ and η are equal: $X_\delta = X_\eta$. Let then X_0 be the submatrix made up with the other columns of X and θ_0 the corresponding vector of parameters. The model is

$$E(y) = X_0\theta_0 + X_\delta\delta + X_\eta\eta = X_0\theta_0 + X_\delta(\delta + \eta).$$

In the whole model, δ and η cannot be estimated. But if X_η is suppressed from the model and X_δ is not in the space generated by X_0, $\delta + \eta$ can be estimated as the parameter associated with X_δ. If δ and η pertain to single factorial effects, the sum cannot generally be given any simple interpretation. But if its estimate has an important absolute value, it indicates that either δ of η or both have important values. This can prompt the experimenter to go on with the experimentation to get separate estimates of them. In some cases, consideration making use of past knowledge or of the other estimates in θ_0 makes it possible to decide which of δ or θ accounts for the importance of the sum without further information.

It may therefore be appropriate when examining a factorial effect to drop the terms that are highly non-orthogonal with it in the model. But they should be the only terms dropped, because dropping terms makes the contrasts examined depend on the hazard of the data and therefore complicates the interpretation. In particular, there is generally no reason while examining some effects to drop all the terms imbedding it.

A final argument against the systematic use of R-approach is the impossibility to compare with it data coming from different designs. This approach is therefore of no use for the design of experiment and never appears in the literature on factorial designs.

(2) The second attitude is to adopt the R-approach as a way to get good biased estimates of the parameters in model (5). When $\gamma = 0$, the R-approach leads to a better estimate of α than the R*-approach. So it can be hoped that when γ is not significantly different from 0, the R-estimate $\hat{\alpha}$ has a better MSE (mean square error) than the R* estimate $\tilde{\alpha}$. Unfortunately, we show below that this is wrong in many contexts.

The estimate $\tilde{\alpha}$ is by construction unbiased and it therefore follows from (8) that

$$\mathrm{MSE}(\tilde{\alpha}) = \mathrm{var}(\tilde{\alpha}) = \frac{\sigma^2}{16}\left(1 + \frac{3}{m}\right).$$

The bias for $\hat{\theta}_1$ is $(X_1'X_1)^{-1}X_1'X_2\gamma$. The α coordinate of this vector is:

$$\mathrm{Bias}(\hat{\alpha}) = \frac{m-1}{m+3}\gamma.$$

So

$$\mathrm{MSE}(\hat{\alpha}) = \sigma^2\left[\frac{m+1}{2m(m+3)} + \frac{(m-1)^2}{(m+3)^2}\left(\frac{\gamma}{\sigma}\right)^2\right].$$

The ratio of these two MSEs is

$$\begin{aligned}\frac{\mathrm{MSE}\left(\hat{\alpha}\right)}{\mathrm{MSE}\left(\tilde{\alpha}\right)} &= \frac{8(m+1)}{(m+3)^2} + \frac{16m(m-1)^2}{(m+3)^3}\left(\frac{\gamma}{\sigma}\right)^2\\ &= v + b\left(\frac{\gamma}{\sigma}\right)^2,\end{aligned}$$

where

$$v = \frac{8(m+1)}{(m+3)^2}, \quad b = \frac{16m(m-1)^2}{(m+3)^3}.$$

The R-estimate is better than the R* one if $v + b(\gamma/\sigma)^2 < 1$, that is,

$$\mathrm{MSE}\left(\hat{\alpha}\right) < \mathrm{MSE}\left(\tilde{\alpha}\right) \quad \Longleftrightarrow \quad (\gamma/\sigma)^2 < \frac{1-v}{b} = \frac{m+3}{16m}.$$

Thus when γ/σ is greater than

$$S = \sqrt{(m+3)/16m}, \tag{9}$$

the R*-estimate $\tilde{\alpha}$ is better than the R-estimate $\hat{\alpha}$. Table 5 gives the threshold S for each $m \leqslant 10$. A question which naturally arises is then: what is the probability to reject the hypothesis $\gamma = 0$ of no interaction when γ/σ is equal to S?

The estimate of γ in the interactive model is given by (6). Its variance is

$$\operatorname{var}(\tilde{\gamma}) = \frac{m+3}{16m}\sigma^2 = k\sigma^2,$$

where

$$k = \frac{m+3}{16m}. \tag{10}$$

The test F of the hypothesis $\gamma = 0$ is thus

$$F = \frac{\tilde{\gamma}^2/k}{\tilde{\sigma}^2},$$

where $\tilde{\sigma}^2$ denotes the residual variance, computed with $M = 3(m-1)$ degrees of freedom. Under the usual normality assumptions, we have

$$\frac{\tilde{\gamma}}{\sqrt{k}\sigma} \sim \mathcal{N}\left(\frac{\gamma}{\sqrt{k}\sigma}, 1\right),$$

$$\frac{\tilde{\sigma}^2}{\sigma^2} \sim \frac{\chi^2_M}{M}$$

and thus

$$F = \frac{\tilde{\gamma}^2/k\sigma^2}{\tilde{\sigma}^2/\sigma^2} \sim F_{1,M}\left(\frac{\gamma^2}{k\sigma^2}\right), \tag{11}$$

where $F_{1,M}(\lambda)$ denotes the non-central F-distribution with 1 and M degrees of freedom and non-centrality parameter λ.

If γ/σ is equal to the threshold S given by (9), it follows from (10) that the non-centrality parameter on the right-hand side of (11) is 1. The probability P_1 to reject the hypothesis $\gamma = 0$ at level 5% with this non-centrality parameter is given in Table 5. We also give in this table the probability P_{10} to reject the hypothesis $\gamma = 0$ at the 5% level if γ/σ is 10 times the threshold S (the non-centrality parameter is then equal to 10). As this table shows, there are a wide range of values of γ/σ where the estimate $\tilde{\alpha}$ of the A-effect in the model with interaction has a better MSE than the estimate $\hat{\alpha}$ in the additive model although there is very little chance to detect the interaction.

Indeed, even if the interaction is found significantly different from 0, looking at the mean A-effect α defined in (4) still makes sense. If this A-effect is found much

Table 5
Comparison of R and R* estimates

	m								
	2	3	4	5	6	7	8	9	10
S	0.4	0.35	0.33	0.32	0.31	0.3	0.29	0.29	0.29
P_1	0.11	0.14	0.15	0.15	0.16	0.16	0.16	0.16	0.16
P_{10}	0.57	0.75	0.8	0.83	0.84	0.85	0.85	0.86	0.86

larger than the interaction, then it can be sensible from a practical point of view to neglect the interaction even if it is statistically significant. On the contrary, if this *A*-effect is of the same order or even smaller than the interaction, then this indicates that the two factors cannot be considered separately and that the four means have to be examined and compared as if they were the levels of the same four-level factor.

2.2. An example with crossed factors and unequal weights

As already mentioned, though in most uniform circumstances it is natural to use equal weights to define marginal means, unequal weights may sometimes be more appropriate or even essential. Kobilinsky [11] gives an example where choosing the classical uniform weights makes the results very difficult to use.

The example comes from a study on the influence of cheese making conditions on the texture and quality of the Arzúa-Ulloa cheese, a traditional Galician cheese [1]. In this study, six 2-level and one 3-level process factors are taken into account in a design with 32 units. The units are structured in eight blocks of size 4 (factor *j*) corresponding to the sets of four cheeses made the same day with the same milk. The 3-level factor, denoted by *A*, is the salting conditions: the salt can be added either in the milk, or in the curd, or in the brine which receives the fresh cheese.

To find a suitable design, it can be first done as if the salting conditions – factor *A* – had four levels defined by two pseudofactors A_1, A_2. It is easy to find the two possible sets of defining relations ensuring resolution IV and then, by backtrack search, to find for each of these two sets three 2-level block pseudofactors j_1, j_2, j_3 defining a system of eight blocks orthogonal to main effects. Table 6 gives the definitions and properties of the two corresponding regular fraction.

Table 6
The two regular $4 \times 2^6/8$ fractions of resolution 4

Definition	First fraction $E=A_1BCD,\ F=A_2BC,\ G=A_2BD$	Second fraction $E=A_1BC,\ F=A_1BD,\ G=A_1CD$
Blocks	$j_1=A_2B,\ j_2=A_2C,\ j_3=A_2D$	$j_1=A_1B,\ j_2=A_1C,\ j_3=A_1D$
Whole set of defining contrasts	A_1BEFG, A_1BCDE, A_1A_2DEF, A_1A_2CEG, A_2BDG, A_2BCF, $CDFG$	A_1BCE, A_1BDF, A_1CDG, A_1EFG, $BCFG$, $BDEG$, $CDEF$
Aliased factorial effects	$([j_2];\ A_2C;\ BF)$, $([j_2j_3];\ CD;\ FG)$, $([j_3];\ A_2D;\ BG)$, $([j_1j_3];\ A_2G;\ BD)$, $([j_1j_2j_3];\ A_1A_2E;\ DF;\ CG)$, $([j_1j_2];\ A_2F;\ BC)$, $([j_1];\ CF;\ DG;\ A_2B)$, $(A_1A_2C;\ EG)$, $(A_1A_2D;\ EF)$, $(CE;\ A_1A_2G)$, $(DE;\ A_1A_2F)$	$([j_3], CG, A_1D, BF)$, $([j_1j_3], EG, A_1F, BD)$, $([j_1], CE, DF, A_1B)$, $([j_2j_3], EF, A_1G, CD)$, $([j_2], A_1C, DG, BE)$, $([j_1j_2], FG, A_1E, BC)$, $([j_1j_2j_3], CF, DE, BG)$
Unaliased factorial effects	$A_1, A_2, A_1A_2, B, C, D, E, F, G$, $A_1B, A_1C, A_1D, A_1E, A_1F$, A_1G, A_2E, A_1A_2B, BE	$A_1, A_2, A_1A_2, B, C, D, E, F, G$, $A_2B, A_2C, A_2D, A_2E, A_2F, A_2G$, $A_1A_2B, A_1A_2C, A_1A_2D$, $A_1A_2E, A_1A_2F, A_1A_2G$
Residual degrees of freedom	2	3

To give three instead of four levels to factor A, the levels $(-1, 1)$ and $(1, -1)$ defined by A_1, A_2 are collapsed, in the way defined by Addelman [2], to one unique level which therefore appears twice as often as the two other levels, that is 16 times instead of 8. It is easy to derive the properties of the resulting design from those of the initial regular fraction and to show that the collapsing of levels preserves the resolution IV, provided one gives to the level resulting from the collapse twice the weight of the other two levels when defining the main effects and interactions.

It was the second fraction which was in this case selected because it leads after the collapse to a fraction which can estimate, besides main effects, all two-factor interactions involving A in the model including all two-factor interactions and the block effects. It turns out that the corresponding design is of resolution IV even if the levels of A are uniformly weighted. But this is not true of the first fraction. For this fraction, given explicitly in Table 7, Table 8 gives the linear estimable combination of parameters for two reparametrisations. The weighted one uses the adequate unequal weights preserving the resolution IV, while the classical one based on uniform weights loses it. In this second parametrisation, some main effects are confounded with two factor interactions which makes the results extremely difficult to interpret.

2.3. Analysis of variance of non-uniform data: the puzzle

Known softwares offering an R* approach only propose equal weights. They are thus unable to give a proper analysis for resolution IV designs as the one mentioned in the previous section. But they can correctly analyse most cases where factors are either crossed or nested, provided nesting relationship are *uniform*. Following Speed and Bailey [21], we say that a factor B nested in A is uniformly nested if the number of levels of B is the same within each of the classes defined by the levels of A.

Whenever there are non-uniform nestings, most softwares still produce a result, but the results may differ from one software to another.

Consider again the situation with three factors used to illustrate the R-notation in the beginning of Section 2. Assume that A and B have two levels and that C has three levels for $A = 1$, but only two for $A = 2$. Factor B is completely crossed with C and A. The design is given on the leftside of Table 9 together with a simulated observed variate y. Some treatments have been repeated twice in order to get residual degrees of freedom. Table 10 gives the sum of squares obtained with the model $A + B + AB + AC + ABC$ by different softwares. For three of these softwares, the corresponding programs are given in Table 11.

Most results are identical, except for the main effect of B. With the software Splus, there are some differences between the UNIX version 3.2 and the Windows version 4.5 that were used. In the UNIX version, the function drop1.aov was used to drop terms from the model in the hope of getting some R* type sums of squares. But this version of Splus [22] does not cope with non-uniformity and considers that C should have a third level within level 2 of A. It therefore adds two supplementary columns in the X matrix of the linear model and produces the diagnostic that 2 out of 12 effects are not estimable. It consequently produces a lot of zeros in the analysis of variance “with drop1.aov”. The Windows version allows us to obtain the same type III sums of squares as in SAS with the statement “summary(result, ssType = 3)” applied to the result of “aov”. The SPSS windows version [23] also provides the type III sums of squares of SAS in a standard way. However Drton [9] found with the unique sum of squares of SPSS release 6.1 and the same data a different result which we reported on the rightside of Table 10. SPSS warns the user that “UNIQUE sum of squares are obtained assuming the redundant effects (possibly caused by missing cells) are actually null” and that “The hypothesis tested may not be the hypothesis of interest”.

Table 7
The first fraction defined in Table 6

A_1	A_2	j_1	j_2	j_3	A	B	C	D	E	F	G	j
0	0	0	0	0	2	1	1	1	1	1	1	7
1	0	0	0	0	1	1	1	1	0	1	1	7
0	1	1	1	1	1	1	1	1	1	0	0	0
1	1	1	1	1	0	1	1	1	0	0	0	0
0	0	1	0	0	2	0	1	1	0	0	0	3
1	0	1	0	0	1	0	1	1	1	0	0	3
0	1	0	1	1	1	0	1	1	0	1	1	4
1	1	0	1	1	0	0	1	1	1	1	1	4
0	0	0	1	0	2	1	0	1	0	0	1	5
1	0	0	1	0	1	1	0	1	1	0	1	5
0	1	1	0	1	1	1	0	1	0	1	0	2
1	1	1	0	1	0	1	0	1	1	1	0	2
0	0	1	1	0	2	0	0	1	1	1	0	1
1	0	1	1	0	1	0	0	1	0	1	0	1
0	1	0	0	1	1	0	0	1	1	0	1	6
1	1	0	0	1	0	0	0	1	0	0	1	6
0	0	0	0	1	2	1	1	0	0	1	0	6
1	0	0	0	1	1	1	1	0	1	1	0	6
0	1	1	1	0	1	1	1	0	0	0	1	1
1	1	1	1	0	0	1	1	0	1	0	1	1
0	0	1	0	1	2	0	1	0	1	0	1	2
1	0	1	0	1	1	0	1	0	0	0	1	2
0	1	0	1	0	1	0	1	0	1	1	0	5
1	1	0	1	0	0	0	1	0	0	1	0	5
0	0	0	1	1	2	1	0	0	1	0	0	4
1	0	0	1	1	1	1	0	0	0	0	0	4
0	1	1	0	0	1	1	0	0	1	1	1	3
1	1	1	0	0	0	1	0	0	0	1	1	3
0	0	1	1	1	2	0	0	0	0	1	1	0
1	0	1	1	1	1	0	0	0	1	1	1	0
0	1	0	0	0	1	0	0	0	0	0	0	7
1	1	0	0	0	0	0	0	0	1	0	0	7

Table 8
Aliased effects with two different parametrisations

Weighted parametrisation	Classical parametrisation
A	A
A^2	A^2
B	B
C	$C + E.G/3$
D	$D + E.F/3$
E	$E + (C.G + D.F + j^7)/3$
F	$F + D.E/3$
G	$G + C.E/3$
$A.B$	$A.B$
$A^2.B$	$A^2.B$
$A.C$	$A.C$
$A^2.C + E.G$	$A^2.C + 2\sqrt{2}\,E.G/3$
$A.D$	$A.D$
$A^2.D + E.F$	$A^2.D + 2\sqrt{2}\,E.F/3$
$A.E$	$A.E$
$A^2.E + C.G + D.F + j^7$	$A^2.E + 2\sqrt{2}\,(C.G + D.F + j^7)/3$
$A.F$	$A.F$
$A^2.F + D.E$	$A^2.F + 2\sqrt{2}\,D.E/3$
$A.G$	$A.G$
$A^2.G + C.E$	$A^2.G + 2\sqrt{2}\,C.E/3$
$B.C + j^4$	$B.C + j^4$
$B.D + j^5$	$B.D + j^5$
$B.E$	$B.E$
$B.F + j^2$	$B.F + j^2$
$B.G + j^3$	$B.G + j^3$
$C.D + F.G + j^6$	$C.D + F.G + j^6$
$C.F + D.G + j$	$C.F + D.G + j$
Residual degrees of freedom: 4	

It is also possible using the “difference contrasts” in SPSS to get the sums of squares corresponding to the weights W_1 [9].

Since there is a term ABC in the model, marginal means can be computed from the cell means which are given on the rightside of Table 9. The marginal means for B are given at the bottom of the table. There are two natural ways to compute them and hence the main effect for B. In the first way, equal weights are given to the five levels of factor C (weight W_2). This gives the unequal weights $3/5$, $2/5$ to the levels 1, 2 of A, respectively. In the second way, equal weights $1/2$ are given to the two levels of A and consequently unequal weights ($1/6$, $1/6$, $1/6$, $1/4$, $1/4$) to the five levels of C (weight W_1). The third weight W_p introduced is the one leading to the SAS type III mean squares in that case.

It is easy to deduce the mean square for B from these marginal means m_{B1}, m_{B2} and from the number of replications r_{abc} in the cells:

$$\mathrm{MS}(B) = \frac{(m_{B1} - m_{B2})^2}{\sum_{a,c} W_{ac}^2 \left(\frac{1}{r_{a1c}} + \frac{1}{r_{a2c}}\right)}.$$

For instance if $W = W_1$, the denominator is:

$$\begin{aligned}0.2673611111 = {} & (1/6)^2(1+1) + (1/6)^2(0.5+0.5) + (1/6)^2(0.5+0.5) \\ & + (1/4)^2(0.5+0.5) + (1/4)^2(0.5+1)\end{aligned}$$

and thus

$$\mathrm{MS}(B) = (23.5 - 62/3)^2/0.2673611111 = 30.02597403.$$

The SAS type III sums of squares are defined [17] by an orthogonalisation process in the dual of the parameter space, where the vector θ of parameters is defined in the usual way:

$$\begin{aligned}\theta' = (&\mu, \alpha_1, \alpha_2, \beta_1, \beta_2, \alpha\beta_{11}, \alpha\beta_{12}, \alpha\beta_{21}, \alpha\beta_{22},\ \alpha\gamma_{11}, \alpha\gamma_{12}, \alpha\gamma_{13}, \alpha\gamma_{21}, \alpha\gamma_{22}, \\ &\alpha\beta\gamma_{111}, \alpha\beta\gamma_{112}, \alpha\beta\gamma_{113}, \alpha\beta\gamma_{121}, \alpha\beta\gamma_{122}, \alpha\beta\gamma_{123}, \alpha\beta\gamma_{211}, \alpha\beta\gamma_{212}, \\ &\alpha\beta\gamma_{221}, \alpha\beta\gamma_{222}).\end{aligned}$$

It has dimension 24 and orthogonality is with respect to the usual scalar product of R^{24}. In the non-uniform case, it seems difficult to give a sense to this scalar product,

Table 9
Example with C nested in A and B crossed with A and C

Design

A	C	B	y
1	1	1	54
1	1	2	14
1	2	1	21
1	2	1	17
1	2	2	36
1	2	2	28
1	3	1	24
1	3	1	25
1	3	2	18
1	3	2	15
2	1	1	17
2	1	1	12
2	1	2	21
2	1	2	25
2	2	1	15
2	2	1	14
2	2	2	18

Cell means and B-marginal means

					B			
					1		2	
W_1	W_2	W_p	A	C	mean	nb.rep	mean	nb.rep
1/6	1/5	$p/3$	1	1	54	(1)	14	(1)
1/6	1/5	$p/3$	–	2	19	(2)	32	(2)
1/6	1/5	$p/3$	–	3	24.5	(2)	16.5	(2)
1/4	1/5	$(1-p)/2$	2	1	14.5	(2)	23	(2)
1/4	1/5	$(1-p)/2$	–	2	14.5	(2)	18	(1)
Marginal means for W_1					23.5		62/3	
Marginal means for W_2					25.3		20.7	
Marginal means for W_p, $p = 0.5294117647$[a]					24.029		20.676	

[a] SAS type III mean square for B can be computed from the B-means obtained with the weight W_p, where $p = 0.5294117647$

Table 10
Mean squares for example of Table 9

Factorial effect	d.f.[a]	Mean squares					
		Weights W_1	Weights W_2	SAS type III	Splus UNIX	MINITAB	SPSS 6.1
A	1	314.29	314.29	314.29	0	314.29	223.21
B	1	30.03	81.39	42.75	0	30.03	34.30
$A.B$	1	291.84	291.84	291.84	0	291.84	118.30
$A.C$	3	84.53	84.53	84.53	84.53	84.53	84.53
$A.B.C$	3	317.67	317.67	317.67	317.67	317.67	317.67

[a] d.f.: degrees of freedom.

hence to the mean squares thus defined. In the example however, it can easily be seen that the B type-III sum of squares is associated with the B-effect computed with the weight W_p given in Table 9. Note that the means computed with the LSMEANS statement are different: they are in fact the B-means associated with the weight W_1. So there is no coherence between sum of squares and adjusted means in that case.

In Splus under Windows, we unsuccessfully tried to get the adjusted means by asking for them in the menu: Statistics > Analysis of variance > fixed effects. This produced the following diagnostic: "Error in model.means.lm(x, estimable.functions = F): computataions failed because of term $(c\%$ in $\%a) : b$".

The adjusted mean squares in MINITAB [13] are those obtained with the weights W_1 giving the same weight to the two levels of A.

The computation of sums of squares in this example relies on the definition of the weights W_A, W_B, W_C associated with the three factors. It seems natural in this context to give the same weight to the two levels of B and similarly to give equal weights to all the levels of C within some level of A, that is to take

$$\begin{aligned}
&W_B(1) = W_B(2) = 1/2,\\
&W_C(1,1) = W_C(1,2) = W_C(1,3) = 1/3,\\
&W_C(2,1) = W_C(2,2) = 1/2,
\end{aligned}$$

where $W_C(a,c)$ is the weight associated to the level c of C within the level a of the nesting factor A.

For the factor A, we have introduced two natural choices:

$$\begin{aligned}
&W_A(1) = W_A(2) = 1/2,\\
&W_A(1) = 3/5, \quad W_A(2) = 2/5.
\end{aligned}$$

Let $\mathscr{T}$ be a term in the model. The weights on which the corresponding factorial effect depends are easy to find (see Proposition 5.4). They are the weights associated to factors which appear in a term including $\mathscr{T}$ but not in $\mathscr{T}$ itself.

In the example, the factorial effects A, AB, AC, ABC do not depend on W_A since A appears in their definition. But B is dependent on W_A since A appears in the term AB which includes B.

Another small example with four factors A, B, C, D and the hierarchies

$$A \geqslant B, \quad C \geqslant D$$

Table 11
Programs used to compute the MS in Table 10

```
SAS

data d;
infile 'nonunif1.don';
input A C B V;
run;
proc glm data=d;
class A C B;
model V=A C(A)  A*B B C*B(A)/ ss3 e3;
lsmeans A C(A) A*B B C*B(A);
run;

Splus

d<-read.table("nonunif1.don",header=T)
d$a<-factor(d$a)
d$b<-factor(d$b)
d$c<-factor(d$c)
result<-aov(v~a/c*b,d)
drop1.aov(result,scope=result$call)
summary(result,ssType=3)             (Windows version only)

SPSS                                (release 6.1)

MANOVA
  y BY a(1 2) c(1 3) b(1 2)
  /N0PRINT PARAM(ESTIM)
  /METHOD=UNIQUE
  /ERROR WITHIN
  /DESIGN = a, b, c WITHIN a, a BY b, b BY c WITHIN a .
```

is detailed in Table 12. As in Table 10, each column of mean squares corresponds either to a given system of weight, or to the output of a particular software. We have introduced four systems of weight given besides the data. The fourth one W_4 was selected because it corresponds to some of the SAS type III sum of squares.

Note that the systems of weights only differ by the weights associated with A and C. For the nested factors B and D, the standard natural weights have been selected in each case, that is

Table 12
Example with four factors satisfying $A \geqslant B, C \geqslant D$

Design				
A	B	C	D	V
1	1	1	1	3.3
1	1	2	2	6.6
1	1	2	2	7.5
1	1	2	3	13.6
2	2	1	1	6.3
2	2	1	1	8.9
2	2	2	2	11.4
2	2	2	3	17.9
2	2	2	3	15.5
2	3	1	1	11.9
2	3	1	1	11.9
2	3	2	2	14.9
2	3	2	2	14.5
2	3	2	3	19.9
2	3	2	3	20.4

System of weights

	$W_A(1)$	$W_A(2)$	$W_C(1)$	$W_C(2)$
W_1	1/2	1/2	1/2	1/2
W_2	1/3	2/3	1/2	1/2
W_3	1/3	2/3	1/3	2/3
W_4	0.45	0.55	0.45	0.55

Factorial effect	ddl	Mean Squares					
		W_1	W_2	W_3	W_4	SAS t-III	MINITAB
A	1	79.18	79.18	88.93	83.80	83.80	79.18
C	1	95.29	121.15	121.15	104.16	104.16	95.29
$A.C$	1	0.62	0.62	0.62	0.62	0.62	0.62
$A.B$	1	36.96	36.96	36.11	37.59	36.96	36.96
$C.D$	1	67.89	77.01	77.01	72.03	67.89	67.89
$A.C.D$	1	0.64	0.64	0.64	0.64	0.64	0.64
$A.B.C$	1	0.52	0.52	0.52	0.52	0.52	0.52

$$W_B(1,1) = 1, \quad W_B(2,2) = W_B(2,3) = 1/2,$$
$$W_D(1,1) = 1, \quad W_D(2,2) = W_D(2,3) = 1/2.$$

The model is

$$\mathscr{E} = \{A, C, AC, AB, CD, ACD, ABC\}.$$

It does not include the interaction *ABCD* between *B* and *D*.

The rule previously mentioned shows that *AC*, *ABC* and *ACD* are independent of the weights W_A, W_C while *A*, *AB* are depending on W_C and *C*, *CD* on W_A. This explains the difference between the columns of mean squares. In that example, the SAS type III sums of squares for *A*, *C* correspond to the system of weight W_4 and those for *AB*, *CD* to the system of weight W_1. As in the preceding example, the sums of squares for MINITAB correspond to the first system W_1 of weights.

3. Orthogonal design

Let T be a set of treatments. A factor A on T can be identified with a mapping $\phi_A : T \to T_A$ giving for each treatment its corresponding level. The range T_A of ϕ_A is the set of levels of the factor A.

If A and B are factors on T, we adopt the convention that $A \geqslant B$ if A nests B, that is if for every t, s in T

$$\phi_B(t) = \phi_B(s) \quad \Longrightarrow \quad \phi_A(t) = \phi_A(s),$$

or equivalently if there exists a mapping $\phi_{AB} : T_B \to T_A$ such that $\phi_A = \phi_{AB} \circ \phi_B$. If $a = \phi_{AB}(b)$ is then the level of A corresponding to a given level b of B, a is said to *nest* b.

The factors A and B are said to be equivalent, and we write $A \sim B$, if $A \leqslant B$ and $B \leqslant A$. This occurs iff they induce the same partition of T. The partition induced by a factor A is formed by the reciprocal images $\phi_A^{-1}(a)$ of its levels a in T_A.

With each factor A and corresponding mapping ϕ_A from T into T_A is associated the *contravariant* linear mapping $\phi_A^* : x_A \mapsto x_A \circ \phi_A$ from $\mathbb{R}^{T_A}$ into $\mathbb{R}^T$ and its image $S_A = \phi_A^*\left(\mathbb{R}^{T_A}\right)$, subspace of functions from T into $\mathbb{R}$ which are constant on each class $\phi_A^{-1}(a)$. The correspondence $A \mapsto S_A$ is such that A nests B ($A \geqslant B$) iff $S_A \subset S_B$, and A and B are equivalent iff $S_A = S_B$. Moreover, any two factors A and B have a supremum $A \vee B$ which is the smaller factor nesting both of them and $S_{A\vee B} = S_A \cap S_B$.

A model is a family $\mathscr{E}$ of factors.

Assume the experimenter wishes to study n *primary* factors, numbered $1, \ldots, n$. For each i in the set $I = \{1, \ldots, n\}$ of these factors, we denote by T_i its set of levels and by ϕ_i the corresponding mapping from T into T_i. The model $\mathscr{E}$ generally includes the constant factor, the primary factors and the *product factors* associated with the non-zero interactions.

If $J \subset I$ is the subset of primary factors defining such an interaction, the associated product factor, denoted by ϕ_J, is defined by

$$\phi_J(t) = (\phi_i(t))_{i \in J}. \tag{12}$$

It coincides with the product mapping $\phi_J = \prod_{i\in J} \phi_i$ and is for this reason called the product of the family of factors $(\phi_i)_{i\in J}$. Its set of levels T_J is a subset of $\prod_{i\in J} T_i$. We shall generally refer to it as *the factor J*, though it will sometimes be more convenient to denote it ϕ_J to distinguish it from the subset. For instance we shall write sometimes $\phi_J \leqslant \phi_K$ rather than $J \leqslant K$.

When J is reduced to a single element i, we assume that $T_J = T_i$ and identify ϕ_J with ϕ_i.

In what follows, a *design* is a triplet $(T, W, \mathscr{E})$ where T is a set of treatments, W a weight function on T and $\mathscr{E}$ a model. The weight function W is a function from T into the set $\mathbb{R}^{+*}$ of strictly positive real numbers satisfying $\sum_{t\in T} W(t) = 1$. It induces the following scalar product on $\mathbb{R}^T$:

$$\langle x, z\rangle = \sum_{t\in T} W(t)x(t)z(t). \tag{13}$$

Orthogonality being defined with respect to this scalar product, two factors A and B are said to be *geometrically orthogonal* if the orthogonal supplementary subspaces of $S_A \cap S_B$ in S_A and S_B respectively are orthogonal:

$$S_A \cap (S_A \cap S_B)^{\perp} \perp S_B \cap (S_A \cap S_B)^{\perp}. \tag{14}$$

Definition 3.1 (*Orthogonal design*). The design $(T, W, \mathscr{E})$ is orthogonal if:
(i) the factors in $\mathscr{E}$ are surjective, non-equivalent and geometrically orthogonal,
(ii) $\mathscr{E}$ is closed under the formation of maxima.

Let $(T, W,\mathscr{E})$ be an orthogonal design. For A in $\mathscr{E}$, define $\overline{S}_A$ as the subspace of vectors in S_A orthogonal to each subspace S_B for $B > A$. Then it is clear from their definition that the subspaces $\overline{S}_A$, $A \in \mathscr{E}$, are orthogonal and that for each A, S_A is the direct sum of the subspaces $\overline{S}_B$ for $B \geqslant A$.

In fact the model $\mathscr{E}$ is used for two things. First to define the subspace S of $\mathbb{R}^T$ to which the vector τ of treatment effects must belong: it is the sum of the S_A for $A \in \mathscr{E}$. Then to provide a decomposition of τ into meaningful components by projection onto the orthogonal subspaces $\overline{S}_A$:

$$\tau = \sum_{A\in\mathscr{E}} Q_A\, \tau, \tag{15}$$

where Q_A is the operator of orthogonal projection onto $\overline{S}_A$.

Assume $\mathscr{E}$ includes the constant factor. If $\tau \in \mathbb{R}^T$ is the vector of treatment effects, the set of linear forms $\{\tau \mapsto \langle x, \tau\rangle \mid x \in \overline{S}_A\}$ is, when A is different from the constant factor, the space of contrasts traditionally associated with the term A of the model. Note that the weight function must be taken into account in the definition of contrasts. The linear form $\langle x, \tau\rangle$ is a contrast if x is orthogonal to the constant vector $\mathbf{1}$, that is, if

$$\sum_{t\in T} W(t)x(t) = 0.$$

The weight $W(S)$ of a subset S of T is defined as the sum of the weights of its elements

$$W(S) = \sum_{s\in S} W(s), \tag{16}$$

and the weight function W_A induced by A on T_A by

$$W_A(a) = W\left(\phi_A^{-1}(a)\right). \tag{17}$$

Assume that ϕ_A is a surjection onto T_A. If x_A, z_A are two vectors in $\mathbb{R}^{T_A}$, let $\langle x_A, z_A\rangle_A = \langle x_A \circ \phi_A, z_A \circ \phi_A\rangle$ be the scalar product induced by the scalar product (13) of $\mathbb{R}^T$. Then

$$\langle x_A, z_A\rangle_A = \sum_{a\in T_A} W_A(a)x_A(a)z_A(a) \tag{18}$$

and ϕ_A^* is an isomorphism of $\mathbb{R}^{T_A}$ equipped with the scalar product (18) onto S_A equipped with the scalar product (13).

We denote by P_A the operator of orthogonal projection from $\mathbb{R}^T$ onto S_A. Since the canonical basis $(e_a)_{a\in T_A}$ of $\mathbb{R}^{T_A}$ is orthogonal for the scalar product (18), so is its image $(e_a \circ \phi_A)_{a\in T_A}$ by ϕ_A^* for the scalar product (13). Hence

$$\begin{aligned} P_A\, x &= \sum_{a\in T_A} \frac{\langle x, e_a\circ\phi_A\rangle}{\langle e_a\circ\phi_A, e_a\circ\phi_A\rangle} e_a\circ\phi_A \\ &= \sum_{a\in T_A} \frac{\sum_{\phi_A(t)=a} W(t)x(t)}{\sum_{\phi_A(t)=a} W(t)} e_a\circ\phi_A. \end{aligned} \tag{19}$$

Thus the projection $P_A\, x$ is obtained by replacing for every $a\in T_A$ all the coordinates of index t in $\phi_A^{-1}(a)$ by their weighted mean

$$\overline{x}_a = \frac{\sum_{t\in\phi_A^{-1}(a)} W(t)x(t)}{W_A(a)}. \tag{20}$$

If $\overline{x}_A = (\overline{x}_a)_{a\in T_A}$ is the vector of these means, then

$$P_A\, x = \phi_A^*(\overline{x}_A). \tag{21}$$

Let $\tilde{P}_A$ be the mapping sending x onto $\overline{x}_A$:

$$\tilde{P}_A\, x = \overline{x}_A. \tag{22}$$

Equality (21) gives the equality

$$P_A = \phi_A^* \circ \tilde{P}_A, \tag{23}$$

which shows that $\tilde{P}_A$ is the mapping corresponding to P_A when S_A is identified to $\mathbb{R}^{T_A}$ through the isomorphism ϕ_A^*.

Equality (21) can be expressed in a more familiar way. We let D, D_A be the diagonal matrices with the weights $W(t)$, $W_A(a)$ on the diagonal and X_A be the matrix of ϕ_A^* with respect to the canonical basis of $\mathbb{R}^{T_A}$ and $\mathbb{R}^T$. Then

$$D_A = X_A' D X_A, \quad \overline{x}_A = D_A^{-1} X_A' D\, x, \quad P_A\, x = X_A D_A^{-1} X_A' D\, x\ . \tag{24}$$

Let $(T, W, \mathscr{E})$ be an orthogonal design and A a given factor of $\mathscr{E}$. Each factor B nesting A induces a factor on T_A, that is, the mapping ϕ_{BA} from T_A into T_B which satisfies $\phi_B = \phi_{BA}\circ\phi_A$. The family of factors thus induced by the factors $B \geqslant A$ in $\mathscr{E}$ is denoted by $\mathscr{E}_A$ and called the family induced by $\mathscr{E}$ on T_A. The design $(T_A, W_A, \mathscr{E}_A)$ is called the design induced on T_A by the design $(T, W, \mathscr{E})$.

With each factor ϕ_{BA} in $\mathscr{E}_A$ is associated the *contravariant* linear mapping ϕ_{BA}^* : $x_B \mapsto x_B\circ\phi_{BA}$ from $\mathbb{R}^{T_B}$ into $\mathbb{R}^{T_A}$ and the subspace ${}_AS_B = \phi_{BA}^*(\mathbb{R}^{T_B})$ of $\mathbb{R}^{T_A}$. It is clear that $\phi_B^* = \phi_A^*\circ\phi_{BA}^*$. Consequently, $S_B = \phi_A^*\,({}_AS_B)$. The subspaces S_B,

$B \geqslant A$, of $\mathbb{R}^T$ are thus the images by ϕ_A^* of the corresponding subspaces ${}_AS_B$ of $\mathbb{R}^{T_A}$. Since ϕ_A^* is an isomorphism from $\mathbb{R}^{T_A}$ with the scalar product (13) onto the subspace S_A with the scalar product (18), it respects the orthogonality and hence the following proposition.

Proposition 3.1. *Let $(T, W, \mathscr{E})$ be an orthogonal design and A a factor in $\mathscr{E}$. Then the design $(T_A, W_A, \mathscr{E}_A)$ induced by $(T, W, \mathscr{E})$ on T_A is orthogonal. The decomposition into sums of orthogonal subspaces*

$$\mathbb{R}^{T_A} = \bigoplus_{B \geqslant A} {}_A\overline{S}_B, \quad S_A = \bigoplus_{B \geqslant A} \overline{S}_B$$

induced by these two designs correspond to each other by the linear injective mapping ϕ_A^*.

Let Q_B be the operator of orthogonal projection onto $\overline{S}_B$. When S_B is identified to $\mathbb{R}^{T_B}$ through ϕ_B^*, Q_B is identified to the mapping $\tilde{Q}_B$ such that

$$Q_B = \phi_B^* \circ \tilde{Q}_B. \tag{25}$$

If $B \geqslant A$, $\phi_B^* = \phi_A^* \circ \phi_{BA}^*$ and therefore

$$Q_B = \phi_A^* \circ \phi_{BA}^* \circ \tilde{Q}_B, \tag{26}$$

which shows that $\phi_{BA}^* \circ \tilde{Q}_B$ is the mapping corresponding to Q_B when S_A and $\mathbb{R}^{T_A}$ are identified through ϕ_A^*. From the decomposition of S_A given by Proposition 3.1, it follows that $P_A = \sum_{B \geqslant A} Q_B$, hence $\tilde{P}_A = \sum_{B \geqslant A} \phi_{BA}^* \circ \tilde{Q}_B$ and

$$\tilde{Q}_A = \tilde{P}_A - \sum_{B > A} \phi_{BA}^* \circ \tilde{Q}_B. \tag{27}$$

This equality can be used to compute recurrently $\tilde{Q}_A$.

The following proposition, weighted equivalent of Proposition 1 of Tjur [24], gives a practical condition of geometrical orthogonality.

Proposition 3.2. *Let A, B be two factors defined on T and $H = A \vee B$. Then A and B are geometrically orthogonal if and only if for every couple $(a, b) \in T_A \times T_B$ such that a and b are both nested into the same level h of T_H*

$$W_{A \times B}(a, b) W_H(h) = W_A(a) W_B(b).$$

The factor $A \times B$ is the mapping $t \mapsto (\phi_A(t), \phi_B(t))$ from T into $T_A \times T_B$. Consequently, $W_{A \times B}(a, b)$ is the sum of the weights of the elements having respectively a and b as levels of A and B. Note that the product $A \times B$ is equivalent to $A \wedge B$. If $A = \phi_J$ and $B = \phi_K$, it is moreover equivalent to $\phi_{J \cup K}$.

4. Reference design in the non-uniform case

We now show how to define a suitable reference orthogonal design in the general case. We let $I = \{1, \dots, n\}$ be the set of primary factors studied by the experimenter. Any treatment can be defined by the family $t = (t_i)_{i \in I}$ of corresponding levels of these factors. However, any such vector in $\prod_{i \in I} T_i$ does not necessarily define a feasible treatment. If factor i is compelled by the nature of things to nest another factor j, then the levels t_i and t_j must be compatible, that is must satisfy $t_i = \phi_{ij}(t_j)$. We shall assume here that these are the only constraints to be satisfied.

More precisely, it is assumed that I is partially ordered by the nesting relation and that for each couple i, j in I such that $i \geqslant j$, there is a mapping $\phi_{ij} : T_j \to T_i$ giving for each level t_j of j the nesting level $t_i = \phi_{ij}(t_j)$ of i. These mappings must clearly satisfy the following two conditions:
1. if $i \geqslant j \geqslant k$, then $\phi_{ik} = \phi_{ij} \circ \phi_{jk}$ and
2. for each i, ϕ_{ii} is the identity of T_i.

The feasible treatments are assumed to be all the families $t = (t_i)_{i \in I}$ of $\prod_{i \in I} T_i$ satisfying $t_i = \phi_{ij}(t_j)$ when $i \geqslant j$. Thus, the set T of treatments of the reference design is

$$T = \{(t_i)_{i \in I} \mid t_i = \phi_{ij}(t_j) \text{ for } i, j \text{ in } I \text{ and } i \geqslant j\}. \tag{28}$$

This set is known as the *projective limit* of the family $(T_i)_{i \in I}$ [7]. The projective limit T_J of any subfamily $(T_i)_{i \in J}$ is defined similarly:

$$T_J = \{(t_i)_{i \in J} \mid t_i = \phi_{ij}(t_j) \text{ for } i, j \text{ in } J \text{ and } i \geqslant j\}. \tag{29}$$

If $J = \emptyset$, we adopt the convention that T_J is a set with one element.

The *factor* i on T is then the projection ϕ_i of index i, which sends a treatment $t = (t_i)_{i \in I}$ in T on the corresponding level t_i in T_i. For each subset J of I, the factor J is the mapping $\phi_J = \prod_{i \in J} \phi_i$ defined by (12). It coincides on T with the canonical projection of index J:

$$\phi_J\Big((t_i)_{i \in I}\Big) = (t_i)_{i \in J}. \tag{30}$$

It is clear that ϕ_J sends T into the projective limit T_J.

If $J \subset K$, the factor J nests the factor K. More precisely, let ϕ_{JK} be the projection of index J from T_K into T_J defined by

$$\phi_{JK}\Big((t_i)_{i \in K}\Big) = (t_i)_{i \in J}. \tag{31}$$

Then

$$\phi_J = \phi_{JK} \circ \phi_K. \tag{32}$$

However, even if J is strictly included in K, the mappings ϕ_J and ϕ_K may be equivalent. Assume indeed that for each $k \in K$, there is a $j \in J$ such that $j \leqslant k$. Then the coordinates on K of an element $t \in T$ are completely determinated by its coordinates on J. Consequently $\phi_K \sim \phi_J$. As a particular case, we get the following proposition.

Proposition 4.1. *Let J be a subset of I and K the ancestral subset generated by J, that is, the set of elements greater or equal than an element of J. Then ϕ_J and ϕ_K are equivalent factors.*

A subset J of I is said to be *ancestral* if

$$j \in J \quad \text{and} \quad k > j \quad \Rightarrow \quad k \in J. \tag{33}$$

In view of Proposition 4.1, we consider from now on only factors ϕ_J associated to ancestral subsets J of I.

For $i \in I$, we denote by $]i$ the set of factors in I strictly greater than i and by $[i$ the set of those which are greater or equal to i

$$]i = \{j \in I \mid j > i\}, \quad [i = \{j \in I \mid j \geqslant i\}. \tag{34}$$

We let ρ_i be the mapping from T_i into the projective limit $T_{]i}$ of the family $(T_j)_{j>i}$ defined by

$$\rho_i(t_i) = \Big(\phi_{ji}(t_i)\Big)_{j \in]i}. \tag{35}$$

If $]i$ is empty, $T_{]i}$ is reduced to one element and ρ_i is the constant mapping. Note that

$$\phi_{]i} = \rho_i \circ \phi_i. \tag{36}$$

The following assumption is needed to avoid constraints other than those induced by nesting relations and to guarantee that no primary factor reduces to the product of the factors nesting it.

Assumption 4.1. Each mapping ρ_i is surjective but not injective.

The projective limit $T_{]i}$ of the family $(T_j)_{j>i}$ will be called the *precursor set* of T_i. We shall say of an element t_i such that $\rho_i(t_i) = v$ that it has v as precursor. The assumption tells that for each i, the sets $\rho_i^{-1}(v)$ for v in $T_{]i}$ are not empty and that at least one of them has two or more elements.

For each i in I, let W_i be a weight function from T_i into the set $\mathbb{R}^{+*}$ of strictly positive real numbers satisfying

$$\sum_{t_i \in \rho_i^{-1}(v)} W_i(t_i) = 1 \quad \text{for every } v \in T_{]i}. \tag{37}$$

Define then the weight $W(t)$ of an element $t = (t_i)_{i \in I}$ in T as the product of the weights of its coordinates t_i:

$$W(t) = \prod_{i \in I} W_i(t_i). \tag{38}$$

We will see that the set T and the weight function W provide two basic ingredients of the searched reference orthogonal design. The third ingredient is the model whose factors are here the projections ϕ_J associated to the elements of a family $\mathscr{E}$ of ancestral subsets J of I.

The geometrical orthogonality of these projections will follow from the following proposition.

Proposition 4.2. *Let J be an ancestral subset of I. Then for each $t_J = (t_i)_{i\in J}$ in the projective limit T_J,*

$$W_J(t_J) = \prod_{i\in J} W_i(t_i).$$

The weight function W_J induced by factor J is defined as in (17) by $W_J(t_J) = W\left(\phi_J^{-1}(t_J)\right)$.

Proof. The result is proved by descending recurrence on the number $|J|$ of elements in J. It is clearly true for $J = I$ by the definition of W. Assume it is true for $|J| > m$ and consider a subset J such that $|J| = m$ and a fixed $t_J = (t_i)_{i\in J}$ in T_J. Select a maximal element j in $I \backslash J$ and let $K = J \cup \{j\}$. It follows from (32) that

$$\phi_J^{-1}(t_J) = \phi_K^{-1}\left(\phi_{JK}^{-1}(t_J)\right) = \bigsqcup_{t_K\in\phi_{JK}^{-1}(t_J)} \phi_K^{-1}(t_K),$$

where $\bigsqcup$ indicates a disjoint union. Thus

$$W_J(t_J) = W\left(\phi_J^{-1}(t_J)\right) = \sum_{t_K\in\phi_{JK}^{-1}(t_J)} W\left(\phi_K^{-1}(t_K)\right) = \sum_{t_K\in\phi_{JK}^{-1}(t_J)} W_K(t_K).$$

The set $\phi_{JK}^{-1}(t_J)$ contains all the elements $t_K = (t_i)_{i\in K}$ which have the same coordinates as t_J for $i \in J$ and a coordinate t_j satisfying $\phi_{ij}(t_j) = t_i$ for each $i > j$ in J (the case $j > i \in J$ has not to be considered since J is ancestral). This condition on t_j is equivalent to $\rho_j(t_j) = v$ where $v = \phi_{]j,J} t_J = (t_i)_{i\in]j}$. The use of the recurrence hypothesis and of (37) then gives

$$\begin{aligned} W_J(t_J) &= \sum_{t_K\in\phi_{JK}^{-1}(t_J)} \prod_{i\in K} W_i(t_i) = \prod_{i\in J} W_i(t_i) \sum_{t_j\in\rho_j^{-1}(v)} W_j(t_j) \\ &= \prod_{i\in J} W_i(t_i). \qquad \square \end{aligned}$$

The following corrollary follows immediately from the strict positivity of the weights $W_i(t_i)$.

Corollary 4.1. *The mapping ϕ_J associated to an ancestral subset J of I sends T onto the projective limit T_J.*

Thus T_J is the set of levels of the product factor $\phi_J = \prod_{i \in J} \phi_i$. This corollary also implies in conjunction with the next easily proved proposition that the mappings ϕ_i associated with the primary factors i in I are surjective.

Proposition 4.3. *The canonical projection $\phi_{i,[i}$ from $T_{[i}$ into T_i is an isomorphism whose inverse is the mapping $t_i \mapsto (\phi_{ji}(t_i))_{j \in [i}$.*

This proposition allows to identify $T_{[i}$ with T_i and for any $j \geqslant i$ the mapping $\phi_{[j,[i}$ with ϕ_{ji}. The spaces $\mathbb{R}^{T_{[i}}$ and $\mathbb{R}^{T_i}$ can consequently be identified, but it must be noted that the scalar product induced on the latter space by the scalar product of $\mathbb{R}^T$ is associated with $W_{\{i\}} = W_{[i}$ and not with W_i.

Proposition 4.4. *The mapping sending an ancestral subset J on the partition induced by ϕ_J is a lattice isomorphism. That is, if J and K are both ancestral, the equivalence $\phi_J \sim \phi_K$ occurs if and only if $J = K$. If $J \subset K$, then $\phi_J \geqslant \phi_K$ and*

$$\phi_{J \cap K} \sim \phi_J \vee \phi_K, \quad \phi_{J \cup K} \sim \phi_J \wedge \phi_K.$$

Proof. Assume $J \backslash K$ is not empty and select a minimal element j in it. Note that j is also minimal in $J \cup K$, otherwise there is an element k in K such that $k \leqslant j$ and the ancestrality of K implies $j \in K$ which is in contradiction with the choice of j.

Since ρ_j is not injective, there exists a precursor $v = (t_i)_{i \in]j}$ in $T_{]j}$ such that $\rho_j^{-1}(v)$ contains at least to distinct elements t_j and t'_j. Let $u = (t_i)_{i \in [j}$ be the element obtained by adding the coordinate t_j to v. Then u clearly belongs to the projective limit $T_{[j}$ of the family $(T_i)_{i \geqslant j}$. Hence by Corollary 4.1 there is an element $t = (t_i)_{i \in I}$ having the same coordinates as u for each $i \geqslant j$. In its projection $(t_i)_{i \in J \cup K}$ by $\phi_{J \cup K}$, substitute t_j by t'_j. The resulting element clearly belongs to $T_{J \cup K}$, hence is the projection by $\phi_{J \cup K}$ of an element $s \in T$. Then t and s have the same image by ϕ_K but not by ϕ_J which proves that these two factors are not equivalent.

If $J \subset K$, (32) implies $\phi_J \geqslant \phi_K$.

Let K and J be arbitrary ancestral subsets and $H = J \cap K$. The mapping ϕ_H nests both ϕ_J and ϕ_K, hence $\phi_H \geqslant \phi_J \vee \phi_K$. To prove the opposite inequality, consider two elements s, t such that $\phi_H(s) = \phi_H(t)$, that is, such that $s_i = t_i$ for $i \in H$. Let $u_i = s_i$ for $i \in J$ and $u_i = t_i$ for $i \in K \backslash J$. The family $(u_i)_{i \in J \cup K}$ clearly belongs to the projective limit $T_{J \cup K}$. By Corollary 4.1, it is the projection by $\phi_{J \cup K}$ of an element u of T. Then $\phi_J(s) = \phi_J(u)$ and $\phi_K(u) = \phi_K(t)$ so that s and t are equivalent for $\phi_J \vee \phi_K$. This proves $\phi_J \vee \phi_K \geqslant \phi_H$.

The proof of the other equality $\phi_{J \cup K} \sim \phi_J \wedge \phi_K$ is immediate. $\square$

We can now prove the geometrical orthogonality of any pair of product factors ϕ_J and ϕ_K. Assume the levels t_J in T_J and t_K in T_K are both nested into the same level of $\phi_J \vee \phi_K \sim \phi_{J \cap K}$. Then their coordinates in $J \cap K$ are equal and there are elements t_i for $i \in J \cup K$ such that $t_J = (t_i)_{i \in J}$, $t_K = (t_i)_{i \in K}$.

Let then $h = (t_i)_{i \in J \cap K}$ be the common nesting level of $\phi_J \vee \phi_K$ and $g = (t_i)_{i \in J \cup K}$. Then the treatments with (t_J, t_K) as level of $\phi_J \times \phi_K$ are the same as those with level g of $\phi_{J \cup K}$, hence by Proposition 4.2

$$\begin{aligned} W_{\phi_J \times \phi_K}(t_J, t_K) W_{J \cap K}(h) &= W_{J \cup K}(g) \times W_{J \cap K}(h) \\ &= \prod_{i \in J \cup K} W(t_i) \prod_{i \in J \cap K} W(t_i) \\ &= \prod_{i \in J} W(t_i) \prod_{i \in K} W(t_i) = W_J(t_J) W_K(t_K). \end{aligned}$$

By Proposition 3.2, we therefore have the following proposition.

Proposition 4.5. *The projection ϕ_J for $J \subset I$ are geometrically orthogonal.*

We now assume that $\mathscr{E}$ is a family of ancestral subsets of I which is closed for the intersection. The corresponding family of projections ϕ_J, $J \in \mathscr{E}$, is then closed under the formation of maxima and thus defines, together with T and W, an orthogonal design and orthogonal subspaces $\overline{S}_J$.

The next section gives a useful process to get basis of these subspaces.

5. Full rank meaningful reparametrisation for the orthogonal reference design

Let Q_J denote the operator of orthogonal projection onto $\overline{S}_J$. The replacement of Q_A by Q_J in (15) gives

$$\tau = \sum_{J \in \mathscr{E}} Q_J \, \tau. \tag{39}$$

To handle this decomposition in practice, it is convenient to have for each J a basis $\mathscr{X}_J$ of $\overline{S}_J$, so that $Q_J \, \tau$ is a linear combination of the vectors x in $\mathscr{X}_J$:

$$Q_J \, \tau = \sum_{x \in \mathscr{X}_J} \alpha_x \, x. \tag{40}$$

The parameters α_x in (40), uniquely determined as linear forms of $Q_J \, \tau$, span the space of contrasts associated with J. Note that when the basis $\mathscr{X}_J$ is orthogonal, they take the following simple form:

$$\alpha_x = \langle x, \tau \rangle \, / \, \langle x, x \rangle \, . \tag{41}$$

Together, (39) and (40) lead to the model

$$\tau = \sum_{J \in \mathscr{E}} \sum_{x \in \mathscr{X}_J} \alpha_x \, x, \tag{42}$$

which provides the expectation $\tau(t)$ of the response in function of the parameters α_x for every feasible treatment t :

$$\tau(t) = \sum_{J\in\mathscr{E}} \sum_{x\in\mathscr{X}_J} \alpha_x\, x(t). \tag{43}$$

At least for the reference design T, this leads to a full rank model whose parameters belong to the factorial effects of interest and which is therefore very convenient to perform an analysis of variance [11]. We now describe a simple way to get such a basis $\mathscr{X}_J$ from which model (43) can be derived.

For our aim, the model $\mathscr{E}$ is first completed so that if J and K are ancestral subsets of I,

$$J \in \mathscr{E} \text{ and } K \subset J \quad \Longrightarrow \quad K \in \mathscr{E}. \tag{44}$$

This can be done by adding every ancestral subset K included in a subset of the initial family $\mathscr{E}$. Note that this completion does not change the sum S of the space S_J, that is, the subspace containing τ, and simply leads to a finer decomposition into orthogonal subspaces $\overline{S}_J$.

If $J = \emptyset$, T_J is a set with one element and $\overline{S}_J = S_J$ is the one-dimensional subspace generated by the constant vector $\mathbf{1}$ of $\mathbb{R}^T$.

Consider now an arbitrary ancestral subset $J \neq \emptyset$. The process described hereafter leads to a basis $\mathscr{X}_J$ of ${}_J\overline{S}_J$ which can be immediately transformed in a basis of $\overline{S}_J$ by the isomorphism ϕ_J^*.

Denote by $m(J)$ a set of minimal elements in J and $M(J) = J\backslash m(J)$ (later $m(J)$ will be *the* set of *all* minimal elements of J). Note that $M(J)$ is also ancestral.

Let $\pi_J = \phi_{M(J)J}$ be the canonical projection from T_J onto $T_{M(J)}$. Then T_J is the disjoint union of the $\pi_J^{-1}(v)$ for v in $T_{M(J)}$. Consequently, if $F_J(v)$ denotes the subspace of vectors in $\mathbb{R}^{T_J}$ with zero coordinates outside $\pi_J^{-1}(v)$, then

$$\mathbb{R}^{T_J} = \bigoplus_{v\in T_{M(J)}} F_J(v). \tag{45}$$

It is clear that the subspaces $F_J(v)$, $v \in T_{M(J)}$, are orthogonal to each other:

$$x \in F_J(v), \quad z \in F_J(v') \quad \text{and} \quad v \neq v' \;\Rightarrow\; \langle x, z\rangle_J = 0. \tag{46}$$

For each $i \in m(J)$, let δ_i be the canonical projection from $M(J)$ onto $]i$,

$$\delta_i = \phi_{]i,M(J)}. \tag{47}$$

Consider then a fixed element v in $T_{M(J)}$. The subspace $F_J(v)$ can be identified with $\mathbb{R}^{\pi_J^{-1}(v)}$ by simply dropping the 0 outside $\pi_J^{-1}(v)$. Then each element t_J in $\pi_J^{-1}(v)$ has the same coordinates as v on $M(J)$ and, for each $i \in m(J)$, its coordinate t_i of index i can be any element in $\rho_i^{-1}(\delta_i v)$. Thus $\pi_J^{-1}(v)$ can be identified with the Cartesian product $\prod_{i\in m(J)} \rho_i^{-1}(\delta_i v)$ and this identification induces an isomorphism between $\mathbb{R}^{\pi_J^{-1}(v)}$, hence $F_J(v)$, and $\bigotimes_{i\in m(J)} \mathbb{R}^{\rho_i^{-1}(\delta_i v)}$:

$$F_J(v) \sim \mathbb{R}^{\pi_J^{-1}(v)} \sim \bigotimes_{i\in m(J)} \mathbb{R}^{\rho_i^{-1}(\delta_i v)}. \tag{48}$$

For each $i \in m(J)$, let z_i be a vector of $\mathbb{R}^{\rho_i^{-1}(\delta_i v)}$. When identified to an element of $F_J(v) \subset \mathbb{R}^{T_J}$, that is to a function from T_J into $\mathbb{R}$, the tensor product $\bigotimes_{i \in m(J)} z_i$ is defined by

$$\left(\bigotimes_{i \in m(J)} z_i\right)(t_J) = \begin{cases} \prod_{i \in m(J)} z_i(t_i) & \text{for } t_J = (t_i) \in \pi_J^{-1}(v), \\ 0 & \text{for } t_J \notin \pi_J^{-1}(v). \end{cases} \tag{49}$$

The images of this tensor product by ϕ_J^*, or by ϕ_{JK}^* where K is an ancestral subset containing J, are defined quite similarly. For instance, if $t_K = (t_i)_{i \in K}$ belongs to the projective limit T_K,

$$\begin{aligned} \phi_{JK}^* \left(\bigotimes_{i \in m(J)} z_i\right)(t_K) &= \left(\bigotimes_{i \in m(J)} z_i\right)(\phi_{JK}(t_K)) \\ &= \begin{cases} \prod_{i \in m(J)} z_i(t_i) & \text{if } v = \phi_{M(J)K}(t_K) \\ 0 & \text{if } v \neq \phi_{M(J)K}(t_K). \end{cases} \end{aligned} \tag{50}$$

To simplify notations, it is therefore possible to omit the mapping ϕ_J^*, or ϕ_{JK}^*, and to consider the tensor product $\bigotimes_{i \in m(J)} z_i$ as defined directly on T or T_K.

Let $z = \bigotimes_{i \in m(J)} z_i$ and $x = \bigotimes_{i \in m(J)} x_i$ be two such tensor products in $F_J(v)$. Then (18), with J instead of A, gives

$$\langle x, z \rangle_J = \sum_{t_J \in T_J} W_J(t_J) x(t_J) z(t_J) = \sum_{t_J \in \pi_J^{-1}(v)} W_J(t_J) x(t_J) z(t_J).$$

It follows from Proposition 4.2 that

$$W_J(t_J) = W_{M(J)}(v) \prod_{i \in m(J)} W_i(t_i) \quad \text{for } t_J = (t_i) \in \pi_J^{-1}(v).$$

Hence

$$\begin{aligned} \langle x, z \rangle_J &= \sum_{(t_i) \in \prod_{i \in m(J)} \rho_i^{-1}(\delta_i v)} W_{M(J)}(v) \prod_{i \in m(J)} W_i(t_i) x_i(t_i) z_i(t_i) \\ &= W_{M(J)}(v) \prod_{i \in m(J)} \left(\sum_{t_i \in \rho_i^{-1}(\delta_i v)} W_i(t_i) x_i(t_i) z_i(t_i) \right). \end{aligned}$$

Let $\langle x, z \rangle_i$ denote the scalar product on $\mathbb{R}^{\rho_i^{-1}(\delta_i v)}$ associated with the weight function W_i, that is

$$\langle x, z \rangle_i = \sum_{t_i} W_i(t_i) x(t_i) z(t_i), \tag{51}$$

where t_i varies over $\rho_i^{-1}(\delta_i v)$. Then the previous equality gives the following proposition.

Proposition 5.1. *If* $z = \bigotimes_{i\in m(J)} z_i$ *and* $x = \bigotimes_{i\in m(J)} x_i$ *are two tensor products in* $F_J(v)$ *defined as in* (49), *then* $\langle x, z\rangle_J = W_{M(J)}(v)\prod_{i\in m(J)}\langle x_i, z_i\rangle_i$.

For each $i \in m(J)$, let $\mathscr{Z}_i(\delta_i v)$ be a basis of $\mathbb{R}^{\rho_i^{-1}(\delta_i v)}$. Then it is well known that

$$\mathscr{Z}_J(v) = \bigotimes_{i\in m(J)} \mathscr{Z}_i(\delta_i v), \tag{52}$$

which is by definition the set of all tensor products $\bigotimes_{i\in m(J)} z_i$ between elements $z_i \in \mathscr{Z}_i(\delta_i v)$, is a basis of the tensor product given in (48), hence of $F_J(v)$. It follows from (45) that the union $\mathscr{Z}_J$ over $v \in T_{M(J)}$ of these bases:

$$\mathscr{Z}_J = \bigcup_{v\in T_{M(J)}} \mathscr{Z}_J(v) \tag{53}$$

is a basis of $\mathbb{R}^{T_J}$. The following proposition sums up this result and the preceding definitions.

Proposition 5.2. *Let J be an ancestral subset of I*, $m(J)$ *a set of minimal element of J and* $M(J) = J\backslash m(J)$. *For each* $v \in T_{M(J)}$ *and* $i \in m(J)$, *define* $\delta_i v$ *as the canonical projection of* v *onto* $T_{]i}$. *Let* $\mathscr{Z}_i(\delta_i v)$ *be a basis of* $\mathbb{R}^{\rho_i^{-1}(\delta_i v)}$ *and* $\mathscr{Z}_J(v)$ *be the set of tensor products* $z = \bigotimes_{i\in m(J)} z_i$ *defined by* (49). *Then the union* $\mathscr{Z}_J = \bigcup_v \mathscr{Z}_J(v)$ *is a basis of* R^{T_J}.

It is now assumed that $m(J)$ is the set of all minimal elements of J. Each basis $\mathscr{Z}_i(\delta_i v)$ is selected so that its first element is the vector **1** having all its coordinates equal to 1 and its other elements are orthogonal to **1** for the scalar product $\langle,\rangle_i$ associated with W_i:

$$x_i \in \mathscr{Z}_i(\delta_i v),\quad x_i \neq \mathbf{1} \quad\Rightarrow\quad \langle x_i, \mathbf{1}\rangle = \sum_{t_i\in\rho_i^{-1}(\delta_i v)} W_i(t_i)x_i(t_i) = 0. \tag{54}$$

Denote by $\mathscr{X}_i(\delta_i v)$ the set of these other elements, that is, $\mathscr{X}_i(\delta_i v) = \mathscr{Z}_i(\delta_i v)\backslash\mathbf{1}$. Let $\mathscr{X}_J(v)$ be the tensor product between these sets:

$$\mathscr{X}_J(v) = \bigotimes_{i\in m(J)} \mathscr{X}_i(\delta_i v) \tag{55}$$

and finally $\mathscr{X}_J$ the union over v of these tensor product:

$$\mathscr{X}_J = \bigcup_{v\in T_{M(J)}} \mathscr{X}_J(v). \tag{56}$$

Proposition 5.3. $\mathscr{X}_J$ *is a basis of* ${}_J\overline{S}_J$. *It is orthogonal if each basis* $\mathscr{X}_i(\delta_i v)$ *is orthogonal.*

As indicated after (50), the tensor products in $\mathscr{X}_J$ can be considered as defined directly on T and $\mathscr{X}_J$ can thus be identified with its image by ϕ_J^* which provides the basis of $\overline{S}_J$ requested for decomposition (40).

Proof. $\mathscr{X}_J$ is made up of all tensor products $\bigotimes_{i\in m(J)} z_i$ in $\mathscr{Z}_J$ whose components z_i are distinct from **1**, hence orthogonal to **1**. From (46) and Proposition 5.1, these tensor products are orthogonal to the other elements of $\mathscr{Z}_J$, that is to the tensor products having at least one component z_i equal to **1**. It remains to show that these last tensor products generate the sum of the spaces ${}_JS_L$ associated to ancestral sets L strictly included in J.

If L is such a set, there is at least one minimal element j in J not belonging to L. Thus $L \subset J\backslash\{j\}$ and consequently ${}_JS_L \subset {}_JS_{J\backslash\{j\}}$. It is therefore enough to consider sets L of the form $L = J\backslash\{j\}$ for some $j \in m(J)$.

Assume therefore that $L = J\backslash\{j\}$. Since $m(J)\backslash\{j\}$ is a set of minimal elements of L, Proposition 5.2 can be used. It shows that $\mathscr{Z}_L = \bigcup_{v\in T_{M(J)}} \mathscr{Z}_L(v)$ generates $\mathbb{R}^{T_L}$. Here, $\mathscr{Z}_L(v)$ is the set of tensor products $\bigotimes_{i\in m(J)\backslash\{j\}} z_i$ such that $z_i \in \mathscr{Z}_i(\delta_i v)$. Such a tensor product is defined as in (49) by

$$\left(\bigotimes_{i\in m(J)\backslash\{j\}} z_i\right)(t_L) = \prod_{i\in m(J)\backslash\{j\}} z_i(t_i)$$

if $t_L \in \pi_L^{-1}(v)$, where $\pi_L = \phi_{M(J)L}$, and by 0 otherwise.

The image by ϕ_{LJ}^* of $\mathscr{Z}_L$ thus generates ${}_JS_L$. If $t_J = (t_i)$ and $z = \bigotimes_{i\in m(J)\backslash\{j\}} z_i$, then

$$\phi_{LJ}^*(z)(t_J) = z\left(\phi_{LJ}(t_J)\right) = \prod_{i\in m(J)\backslash\{j\}} z_i(t_i).$$

If we let $z_j = \mathbf{1}$, the last product is also equal to $\prod_{i\in m(J)} z_i(t_i)$ and therefore

$$\phi_{LJ}^*(z) = \bigotimes_{i\in m(J)} z_i.$$

Thus the tensor product $\bigotimes_{i\in m(J)} z_i$ with $z_j = \mathbf{1}$ generates ${}_JS_L$ and the whole set of tensor products having a component equal to **1** generates the sum of the spaces ${}_JS_L$.

If the $\mathscr{X}_i(\delta_i v)$ are orthogonal, the orthogonality of $\mathscr{X}_J$ follows from (46) and Proposition 5.1. $\square$

Consider now a model $\mathscr{E}$ satisfying (44). Let $\overline{J}$ be the set of indices which are not in J but belong to some set K in $\mathscr{E}$ including J:

$$\overline{J} = \left(\bigcup_{K/K\in\mathscr{E},J\subset K} K\right) \Big\backslash J, \tag{57}$$

Proposition 5.4. *The space of contrasts* $\{\langle x,\tau\rangle, x\in \overline{S}_J\}$ *associated with the factorial effect J only depends on the weight* W_j *such that* $j\in\overline{J}$.

Corollary 5.1. *If there is no K strictly including J in* $\mathscr{E}$, *the space of contrasts associated with J is independent of the chosen weights.*

The proof closely follows that given by Kobilinsky [11] in the simpler case of uniform reference designs.

Proof. We denote by $\{\mathscr{V}\}$ the subspace generated by a family $\mathscr{V}$ of vectors.

Proposition 5.3 shows that ${}_J\overline{S}_J$ is the sum of the spaces $\{\mathscr{X}_J(v)\}$, hence $\overline{S}_J$ the sum of the spaces $\phi_J^*(\{\mathscr{X}_J(v)\})$ for $v\in T_{M(J)}$. It is therefore sufficient to show the result when $x\in\phi_J^*(\{\mathscr{X}_J(v)\})$.

From (55), we have

$$\{\mathscr{X}_J(v)\}=\bigotimes_{i\in m(J)}\{\mathscr{X}_i(\delta_i v)\}.$$

and $\{\mathscr{X}_i(\delta_i v)\}$ is the subspace of $\{\mathscr{Z}_i(\delta_i v)\}$ orthogonal to $\mathbf{1}$, that is, the subspace of vectors x_i in $\mathbb{R}^{T_i}$ such that

1. $x_i(t_i)$ is zero when $\rho_i(t_i)\neq\delta_i v$ (i.e. when t_i is not compatible with v).
2. x_i is orthogonal to $\mathbf{1}$: $\langle x_i,\mathbf{1}\rangle_i=\sum_{t_i}W_i(t_i)x(t_i)=0$.

Thus the tensor products $\bigotimes_{i\in m(J)}x_i$ with $x_i\in\{\mathscr{X}_i(\delta_i v)\}$ span $\{\mathscr{X}_J(v)\}$ and their images by ϕ_J^* span $\phi_J^*(\{\mathscr{X}_J(v)\})$. Let x be one of these images:

$$x=\phi_J^*\left(\bigotimes_{i\in m(J)}x_i\right),\qquad x_i\in\{\mathscr{X}_i(\delta_i v)\}.$$

Then (50) applied with $K=I$ gives for $t=(t_i)$

$$x(t)=\begin{cases}\displaystyle\prod_{i\in m(J)}x_i(t_i) & \text{if } v=\phi_{M(J)}(t),\\ 0 & \text{if } v\neq\phi_{M(J)}(t)\ .\end{cases}$$

Hence

$$\begin{aligned}\langle x,\tau\rangle&=\sum_{t\in T}W(t)x(t)\tau(t)\\&=\sum_{t\in\phi_{M(J)}^{-1}(v)}\left(\prod_{i\in I}W_i(t_i)\right)\left(\prod_{i\in m(J)}x_i(t_i)\right)\tau(t)\\&=\sum_{t\in\phi_{M(J)}^{-1}(v)}\left(\prod_{i\in M(J)}W_i(t_i)\right)\left(\prod_{i\in m(J)}W_i(t_i)x_i(t_i)\right)\left(\prod_{i\notin J}W_i(t_i)\right)\tau(t).\end{aligned}$$

Using Proposition 4.2 we get

$$\langle x, \tau \rangle = \sum_{t \in \phi_{M(J)}^{-1}(v)} W_{M(J)}(v) \left(\prod_{i \in m(J)} z_i(t_i) \right) \left(\prod_{i \notin J} W_i(t_i) \right) \tau(t),$$

where z_i is the coordinatewise product of W_i and z_i defined by

$$z_i(t_i) = W_i(t_i) x_i(t_i).$$

The conditions 1 and 2 on x_i are equivalent to similar conditions on z_i:
1. $z_i(t_i) = 0$ if $\rho_i(t_i) \neq \delta_i v$,
2. $\langle z_i, \mathbf{1} \rangle = \sum_{t_i} z(t_i) = 0$.

In the second condition, the scalar product is the standard one on $\mathbb{R}^{T_i}$. It does not depend on W_i. Hence the space of contrasts $\langle x, \tau \rangle$ for x in $\phi_J^* (\{\mathscr{X}_J(v)\})$ is independant of the weights W_i such that $i \in m(J)$. Since this space is also generated by the ratios $\langle x, \tau \rangle / W_{M(J)}(v)$, it is moreover independant of the W_j for $i \in M(J)$. It remains to show that it is also independant of W_j if j does not belong to any K strictly including J.

Since τ belongs to the sum S of the spaces S_K for $K \in \mathscr{E}$, we have $\tau = \sum_{K \in \mathscr{E}} \delta_K$ where for each K, $\delta_K \in S_K$. We can therefore consider $\langle x, \delta_K \rangle$ instead of $\langle x, \tau \rangle$.

If K does not include J, this contrast is 0 because S_K is orthogonal to $\overline{S}_J$ by Proposition 4.5. It is therefore not dependant on any W_i.

Consider then a K including J. Since $\overline{S}_J \subset S_J \subset S_K$, x belongs to S_K as well as δ_K. There are therefore elements x_K and τ_K in $\mathbb{R}^{T_K}$ such that

$$x = \phi_K^*(x_K), \quad \delta_K = \phi_K^*(\tau_K).$$

In view of the remark following (18), we have

$$\langle x, \tau \rangle = \langle x_K, \tau_K \rangle_K = \sum_{t_K} W_K(t_K) x_K(t_K) \tau_K(t_K).$$

It then follows from Proposition 4.2 that $W_K(t_K)$ only depends of the W_k for $k \in K$.

So $\langle x, \delta_K \rangle$ only depends on W_j if $J \subset K$ and $j \in K$. Hence $\langle x, \tau \rangle$ only depends on the W_j such that $j \in \bigcup_{K/J \subset K} K$. The result follows since we know from the first part of the proof that $\langle x, \tau \rangle$ is independant of the weights W_j for $j \in J$. $\square$

Example 5.1. There are four primary factors *A*, *B*, *C*, *D*, with non-trivial order relations

$$D \leqslant A, \quad C \leqslant A, \quad C \leqslant B.$$

The model is

$$\mathscr{E} = \{\emptyset, A, B, A.B, A.D, A.B.C, A.B.D\}$$

A term like $A.B.D$ denotes the subset $\{A, B, D\}$. Thus this model includes all ancestral subsets except the whole set $I = \{A, B, C, D\}$.

The number of levels are:

$$A : 2, \quad B : 2, \quad D(A = 1) : 3, \quad D(A = 2) : 2,$$

$$C(A=1,B=1):3,\quad C(A=1,B=2):2,$$
$$C(A=2,B=1):2,\quad C(A=2,B=2):3.$$

By $C(A=a,B=b)$ we denote the subset of levels of C such that the nesting factors A, B in $]C$ have levels a, b, respectively, that is, the subset $\rho_C^{-1}(v)$ associated with the precursor $v=(a,b)$ of C.

The weights are given in Table 13. The levels in this table are numbered sequentially, and for a nested factor i, independantly within each subset $\rho_i^{-1}(v)$ determined by the levels of the nesting factors. In fact, the numbers on the lines beginning by C or D are *pseudolevels* that cannot be considered independantly of the levels of the nesting factors. The true levels are therefore the combinations of pseudolevels of the factors nesting or equal to the given factor. For instance, the true levels of D are the five pairs of values of (A,D), that is, $(1,1)$, $(1,2)$, $(1,3)$, $(2,1)$, $(2,2)$. The mapping ρ_D is then the projection $(A,D)\mapsto A$ on the first coordinate. Similarly, the true levels of C are the 10 triples $(1,1,1)$ to $(2,2,3)$ of values of (A,B,C) and ρ_C is the projection $(A,B,C)\mapsto(A,B)$ onto the first two coordinates.

Table 13 also gives for each i in $\{A,B,C,D\}$ and each precursor v_i in $T_{]i}$ an orthonormal basis $\mathscr{X}_i(v_i)$, for the scalar product (51), of the orthogonal of $\mathbf{1}$ within $\mathbb{R}^{\rho_i^{-1}(v_i)}$. Again, the notation $A=a$, $B=b$ following $\mathscr{X}_C$ refers to the element $v_i=(a,b)$ in the precursor set $T_{]C}$ of C, that is, $\mathscr{X}_C(A=a,B=b)=\mathscr{X}_C(a,b)$.

The vectors of $\mathscr{X}_i(v_i)$ appear as row vectors and are denoted sequentially $x_{i1}(v_i)$, $x_{i2}(v_i)$, …or more simply x_{i1}, x_{i2}, …when the precursor v_i involved is made clear by the context. Thus for $i=C$, $A=2$, $B=2$, that is, $v_i=(2,2)$, the basis is made up of $x_{C1}=[\sqrt{3/2},-\sqrt{3/2},0]$ and $x_{C2}=[1/\sqrt{2},1/\sqrt{2},-2/\sqrt{2}]$.

The weight W on T appears in Table 14 where the marginal weights W_i are also reported. Within the table, there is one cell per element in the projective limit T.

Table 13
Weight functions W_i and basis $\mathscr{X}_i(v)$

A	1	2
W_A	1/2	1/2
	$\mathscr{X}_A$	
x_{A1}	[1	−1]

B	1	2
W_B	1/2	1/2
	$\mathscr{X}_B$	
x_{B1}	[1	−1]

A	1			2	
D	1	2	3	1	2
W_D	1/3	1/3	1/3	1/2	1/2
	$\mathscr{X}_D(A=1)$			$\mathscr{X}_D(A=2)$	
x_{D1}	[$\sqrt{3/2}$	$-\sqrt{3/2}$	0]	[1	−1]
x_{D2}	[$1/\sqrt{2}$	$1/\sqrt{2}$	$-2/\sqrt{2}$]		

A	1			1		2		2		
B	1			2		1		2		
C	1	2	3	1	2	1	2	1	2	3
W_C	1/3	1/3	1/3	1/2	1/2	1/2	1/2	1/3	1/3	1/3
	$\mathscr{X}_C(A=1,B=1)$			$\mathscr{X}_C(A=2,B=1)$		$\mathscr{X}_C(A=1,B=2)$		$\mathscr{X}_C(A=2,B=2)$		
x_{C1}	[$\sqrt{3/2}$	$-\sqrt{3/2}$	0]	[1	−1]	[1	−1]	[$\sqrt{3/2}$	$-\sqrt{3/2}$	0]
x_{C2}	[$1/\sqrt{2}$	$1/\sqrt{2}$	$-2/\sqrt{2}$]					[$1/\sqrt{2}$	$1/\sqrt{2}$	$-2/\sqrt{2}$]

Table 14
The weight W induced on the projective limit T by the W_i

				B		1			2	
				W_B		1/2			1/2	
A	W_A	D	W_D							
				C	1	2	3	C	1	2
				W_C	1/3	1/3	1/3	W_C	1/2	1/2
		1	1/3		1/36	1/36	1/36		1/24	1/24
1	1/2	2	1/3		1/36	1/36	1/36		1/24	1/24
		3	1/3		1/36	1/36	1/36		1/24	1/24
				C	1	2	C	1	2	3
				W_C	1/2	1/2	W_C	1/3	1/3	1/3
2	1/2	1	1/2		1/16	1/16		1/24	1/24	1/24
		2	1/2		1/16	1/16		1/24	1/24	1/24

Since $\mathscr{E}$ satisfies condition (44), Proposition 5.3 can be used to get the vectors x appearing in (42). These vectors are divided by their norm, given by Proposition 5.1, to get an orthonormal basis. They are numbered sequentially $x_0, x_1, \ldots$ and given explicitly in Table 17. To simplify, the bases $\mathscr{X}_i(\delta_i v)$ used to define $\mathscr{X}_J(v)$ in (55) have always been selected to be those of Table 13, though it would have been possible to select them differently for each $J \in \mathscr{E}$ and $v \in T_{M(J)}$.

We give in what follows some more indications on how to get the vectors x_i of $\mathscr{X}_J$ for each J in $\mathscr{E}$.

- $J = \emptyset$. The only associated vector is $x_0 = \mathbf{1}$.
- $J = \{A\}$. There is just one vector $x_1 = x_{A1}$ defined on T_A by $x_{A1}(1) = 1$, $x_{A1}(2) = -1$ and therefore on T by $x_{A1}(1, b, c, d) = 1$, $x_{A1}(2, b, c, d) = -1$.
- $J = \{B\}$. As for $J = \{A\}$, there is only one vector $x_2 = x_{B1}$.
- $J = \{A, B\}$. The set of minimal elements is $m(J) = \{A, B\}$ and thus $M(J) = \emptyset$. The only vector in $\mathscr{X}_J$ is $x_3 = x_{A1} \otimes x_{B1}$ which is defined on T by $(x_{A1} \otimes x_{B1})\,(a, b, c, d) = x_{A1}(a)x_{B1}(b)$ (it is the coordinatewise product of x_1 and x_2).
- $J = \{A, D\}$. Then $m(J) = \{D\}$ and $M(J) = \{A\}$. The orthogonal basis $\mathscr{X}_J$ includes two vectors x_{D1}, x_{D2} for $A = 1$, one x_{D1} for $A = 2$. Since $W_{M(J)}(v) = 1/2$ for $v = 1, 2$, their norms given by Proposition 5.1 are $1/\sqrt{2}$ and we can take $x_4 = \sqrt{2}x_{D1}$, $x_5 = \sqrt{2}x_{D2}$ for $A = 1$, $x_6 = \sqrt{2}x_{D1}$ for $A = 2$ as orthonormal basis. The values of these vectors, which depend only on A and D, are given in Table 15.
- $J = \{A, B, C\}$. Then $m(J) = \{C\}$ and $M(J) = \{A, B\}$. The norm given by Proposition 5.1 is $\sqrt{W_{M(J)}(v)} = 1/2$ for each of the four couples $v = (a, b)$. The orthonormal basis $\mathscr{X}_J$ includes six vectors, two for $A = 1$, $B = 1$ ($x_7 = 2x_{C1}$, $x_8 = 2x_{C2}$), one for $A = 1$, $B = 2$ ($x_9 = 2x_{C1}$), one for $A = 2$, $B = 1$ ($x_{10} = 2x_{C1}$) and finally two for $A = 2$, $B = 2$ ($x_{11} = 2x_{C1}$, $x_{12} = 2x_{C2}$).

- $J = \{A, B, D\}$. Then $m(J) = \{B, D\}$ and $M(J) = \{A\}$. There are two tensor products $\sqrt{2}x_{B1} \otimes x_{D1}$, $\sqrt{2}x_{B1} \otimes x_{D2}$ to consider for $A = 1$ and one $\sqrt{2}x_{B1} \otimes x_{D1}$ for $A = 2$. Their values which depend only on the levels of A, B, D are given on the rightside of Table 16.
- If $J = \{A, B, C, D\}$ had also be in $\mathscr{E}$, we would have also introduced four vectors for $A = 1$, $B = 1$ ($x_{16} = 2x_{C1} \otimes x_{D1}, x_{17} = 2x_{C2} \otimes x_{D1}, x_{18} = 2x_{C1} \otimes x_{D2}, x_{19} = 2x_{C2} \otimes x_{D2}$), two for $A = 1$, $B = 2$ ($x_{20} = 2x_{C1} \otimes x_{D1}$, $x_{21} = 2x_{C1} \otimes x_{D2}$), one for $A = 2$, $B = 1$ ($x_{22} = 2x_{C1} \otimes x_{D1}$) and finally two for $A = 2$, $B = 2$ ($x_{23} = 2x_{C1} \otimes x_{D1}, x_{24} = 2x_{C2} \otimes x_{D1}$).

  To link this with the previous notation, consider an element $v = (a, b)$ in $T_{M(J)}$. Since $]C = \{A, B\}$ and $]D = \{A\}$, the projections δ_C and δ_D are defined by $\delta_C(a, b) = (a, b)$, $\delta_D(a, b) = a$ and thus $\mathscr{X}_J(a, b) = \mathscr{X}_C(a, b) \otimes \mathscr{X}_D(a)$. Let $n_C(a, b)$ be the number of levels of C for $A = a$, $B = b$, that is within $\rho_C^{-1}(a, b)$ and similarly $n_D(a)$ the number of levels of D within $\rho_D^{-1}(a)$. The vectors in $\mathscr{X}_J(a, b)$ are the $(n_C(a, b) - 1)(n_D(a) - 1)$ products $x_{Cj}(a, b) \otimes x_{Dk}(a)$.

The 25 vectors $x_0, \ldots, x_{24}$ make up an orthogonal basis of $\mathbb{R}^T$ for the scalar product associated with the weight W given in Table 14. The 16 vectors $x_0, \ldots, x_{15}$ associated with the model $\mathscr{E}$ are explicited in Table 17, which also gives on its left the weight W and the levels of the four factors. The arrows on the left point to a fraction considered in Section 7.

6. Adjusted means

Let K be an ancestral subset of I. The mean response $\mu_K(t_K)$ at level t_K of K is defined as the weighted mean

$$\mu_K(t_K) = \sum_{t, \phi_K(t) = t_K} W(t)\tau(t)/W_K(t_K). \tag{58}$$

The replacement of $\tau(t)$ by its expression (43) in function of the parameters α_x gives

Table 15
The orthonormal basis of $\mathscr{X}_{AD}$

		$A = 1$		$A = 2$
A	D	$\sqrt{2}\, x_{D1}$	$\sqrt{2}\, x_{D2}$	$\sqrt{2}\, x_{D1}$
1	1	$\sqrt{3}$	1	0
1	2	$-\sqrt{3}$	1	0
1	3	0	-2	0
2	1	0	0	$\sqrt{2}$
2	2	0	0	$-\sqrt{2}$
		x_4	x_5	x_6

$$\mu_K(t_K) = \sum_{J \in \mathscr{E}} \sum_{x \in \mathscr{X}_J} \lambda_x(t_K)\alpha_x, \tag{59}$$

where

$$\lambda_x(t_K) = \sum_{t, \phi_K(t)=t_K} W(t)x(t)/W_K(t_K). \tag{60}$$

The mean responses $\mu_K(t_K)$ have been seen in (20) to be the coordinates of the orthogonal projection $P_K\tau$ of τ on S_K. More precisely, let $\tilde{P}_K$ be the mapping such that $P_K = \phi_K^* \tilde{P}_K$, that is, the mapping replacing P_K when S_K is identified to $\mathbb{R}^{T_K}$ by ϕ_K^*. Then

$$\mu_K(t_K) = \left(\tilde{P}_K\tau\right)(t_K)$$

and similarly

$$\lambda_x(t_K) = \left(\tilde{P}_K x\right)(t_K).$$

If $x \in \mathscr{X}_J$ and $J \not\subset K$, then $\tilde{P}_K x = 0$ and consequently $\lambda_x(t_K) = 0$. If $x \in \mathscr{X}_J$ and $J \subset K$, then since $\mathscr{X}_J \subset S_J \subset S_K$, x has the same coordinates for all t such that $\phi_K(t) = t_K$ and consequently $\lambda_x(t_K) = x(t)$ for any such t. Moreover if $x \in \mathscr{X}_J(v)$ but $v \neq \phi_{M(J)K}(t_K)$, then $x(t) = 0$ for all t such that $\phi_K(t) = t_K$ and $\lambda_x(t_K) = 0$. Hence the following proposition.

Proposition 6.1. *Let x be a vector in $\mathscr{X}_J$. If $J \not\subset K$, then $\lambda_x(t_K) = 0$. If $J \subset K$, then $\lambda_x(t_K) = x(t)$ for any t such that $\phi_K(t) = t_K$. In particular, $\lambda_x(t_K) = 0$ if $x \in \mathscr{X}_J(v)$ but $v \neq \phi_{M(J)K}(t_K)$.*

Table 16
The orthonormal basis of $\mathscr{X}_{ABD}$

				$A=1$		$A=2$	$A=1$		$A=2$
A	B	D	x_{B1}	x_{D1}	x_{D2}	x_{D1}	$\sqrt{2}\,x_{B1}\otimes x_{D1}$	$\sqrt{2}\,x_{B1}\otimes x_{D2}$	$\sqrt{2}\,x_{B1}\otimes x_{D1}$
1	1	1	1	$\sqrt{3}$	1	0	$\sqrt{3}$	1	0
1	1	2	1	$-\sqrt{3}$	1	0	$-\sqrt{3}$	1	0
1	1	3	1	0	-2	0	0	-2	0
1	2	1	-1	$\sqrt{3}$	1	0	$-\sqrt{3}$	-1	0
1	2	2	-1	$-\sqrt{3}$	1	0	$\sqrt{3}$	-1	0
1	2	3	-1	0	-2	0	0	2	0
2	1	1	1	0	0	$\sqrt{2}$	0	0	$\sqrt{2}$
2	1	2	1	0	0	$-\sqrt{2}$	0	0	$-\sqrt{2}$
2	2	1	-1	0	0	$\sqrt{2}$	0	0	$-\sqrt{2}$
2	2	2	-1	0	0	$-\sqrt{2}$	0	0	$\sqrt{2}$
							x_{13}	x_{14}	x_{15}

Table 17
Matrix X of the linear model after reparametrisation

					A					1		2	1		1	2	2		1		2
					B								1		2	1	2				
	W	A	B	C	D		A	B	AB	D_1	D_2	D_1	C_1	C_2	C_1	C_1	C_1	C_2	BD_1	BD_2	BD_1
	1/36	1	1	1	1	1	1	1	1	$\sqrt{3}$	1	0	$\sqrt{6}$	$\sqrt{2}$	0	0	0	0	$\sqrt{3}$	1	0
→	1/36	1	1	2	1	1	1	1	1	$\sqrt{3}$	1	0	$-\sqrt{6}$	$\sqrt{2}$	0	0	0	0	$\sqrt{3}$	1	0
→	1/36	1	1	3	1	1	1	1	1	$\sqrt{3}$	1	0	0	$-2\sqrt{2}$	0	0	0	0	$\sqrt{3}$	1	0
→	1/24	1	2	1	1	1	1	−1	−1	$\sqrt{3}$	1	0	0	0	2	0	0	0	$-\sqrt{3}$	−1	0
→	1/24	1	2	2	1	1	1	−1	−1	$\sqrt{3}$	1	0	0	0	−2	0	0	0	$-\sqrt{3}$	−1	0
→	1/36	1	1	1	2	1	1	1	1	$-\sqrt{3}$	1	0	$\sqrt{6}$	$\sqrt{2}$	0	0	0	0	$-\sqrt{3}$	1	0
	1/36	1	1	2	2	1	1	1	1	$-\sqrt{3}$	1	0	$-\sqrt{6}$	$\sqrt{2}$	0	0	0	0	$-\sqrt{3}$	1	0
→	1/36	1	1	3	2	1	1	1	1	$-\sqrt{3}$	1	0	0	$-2\sqrt{2}$	0	0	0	0	$-\sqrt{3}$	1	0
	1/24	1	2	1	2	1	1	−1	−1	$-\sqrt{3}$	1	0	0	0	2	0	0	0	$\sqrt{3}$	−1	0
→	1/24	1	2	2	2	1	1	−1	−1	$-\sqrt{3}$	1	0	0	0	−2	0	0	0	$\sqrt{3}$	−1	0
	1/36	1	1	1	3	1	1	1	1	0	−2	0	$\sqrt{6}$	$\sqrt{2}$	0	0	0	0	0	−2	0
	1/36	1	1	2	3	1	1	1	1	0	−2	0	$-\sqrt{6}$	$\sqrt{2}$	0	0	0	0	0	−2	0
→	1/36	1	1	3	3	1	1	1	1	0	−2	0	0	$-2\sqrt{2}$	0	0	0	0	0	−2	0
	1/24	1	2	1	3	1	1	−1	−1	0	−2	0	0	0	2	0	0	0	0	2	0
→	1/24	1	2	2	3	1	1	−1	−1	0	−2	0	0	0	−2	0	0	0	0	2	0
→	1/16	2	1	1	1	1	−1	1	−1	0	0	$\sqrt{2}$	0	0	0	2	0	0	0	0	$\sqrt{2}$
	1/16	2	1	2	1	1	−1	1	−1	0	0	$\sqrt{2}$	0	0	0	−2	0	0	0	0	$\sqrt{2}$
→	1/24	2	2	1	1	1	−1	−1	1	0	0	$\sqrt{2}$	0	0	0	0	$\sqrt{6}$	$\sqrt{2}$	0	0	$-\sqrt{2}$
→	1/24	2	2	2	1	1	−1	−1	1	0	0	$\sqrt{2}$	0	0	0	0	$-\sqrt{6}$	$\sqrt{2}$	0	0	$-\sqrt{2}$
→	1/24	2	2	3	1	1	−1	−1	1	0	0	$\sqrt{2}$	0	0	0	0	0	$-2\sqrt{2}$	0	0	$-\sqrt{2}$
→	1/16	2	1	1	2	1	−1	1	−1	0	0	$-\sqrt{2}$	0	0	0	2	0	0	0	0	$-\sqrt{2}$
→	1/16	2	1	2	2	1	−1	1	−1	0	0	$-\sqrt{2}$	0	0	0	−2	0	0	0	0	$-\sqrt{2}$
	1/24	2	2	1	2	1	−1	−1	1	0	0	$-\sqrt{2}$	0	0	0	0	$\sqrt{6}$	$\sqrt{2}$	0	0	$\sqrt{2}$
	1/24	2	2	2	2	1	−1	−1	1	0	0	$-\sqrt{2}$	0	0	0	0	$-\sqrt{6}$	$\sqrt{2}$	0	0	$\sqrt{2}$
→	1/24	2	2	3	2	1	−1	−1	1	0	0	$-\sqrt{2}$	0	0	0	0	0	$-2\sqrt{2}$	0	0	$\sqrt{2}$
						x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}

Thus

$$\mu_K(t_K) = \sum_J \sum_{x \in \mathscr{X}_J} x(t)\alpha_x, \tag{61}$$

where t is any element such that $\phi_K(t) = t_K$ and J varies only among the subsets of K in $\mathscr{E}$. If $x \in \mathscr{X}_J$ and $v = \phi_{M(J)}(t)$, then $x(t) = 0$ for all x outside $\mathscr{X}_J(v)$. Thus the sum for $x \in \mathscr{X}_J$ can be restricted to the set $\mathscr{W}_J = \mathscr{X}_J(v) = \mathscr{X}_J\left(\phi_{M(J)K}(t_K)\right)$.

When K is the whole set of primary factors ($K = I$), (61) coincides with model (43). In the other cases, the form is similar but J varies only over subsets of K.

If α_x is estimable for each $x \in \cup\mathscr{W}_J$, where $J \in \mathscr{E}$ and $J \subset K$, the mean responses $\mu_K(t_K)$ associated with the levels $t_K \in T_K$ are estimable and their estimations, known as the *adjusted means* for factor K are obtained by adding hats on μ and α in (61).

If the factorial effect of K is significant, it is usual to carry on by the examination of these adjusted means or of some linear combinations of them. Of particular interest are the estimates of the coordinates of $Q_K\tau$, or equivalently the coordinates of $\tilde{Q}_K\tau$, which can be determined recurrently by formula (27). These coordinates are called the factorial effects of factor K. The factorial effect of index t_K is denoted by $\alpha_K(t_K)$.

Example 6.1. Consider again Example 5.1. The treatment in T are identified with the feasible quadruples (a, b, c, d) of levels of the four factors. We use the dot notation to denote a weighted mean like $\mu_K(t_K)$: the dots replace the indices of factors which are not in K. For instance $\tau(a, \bullet, \bullet, \bullet)$ is the weighted mean $\mu_A(a)$ of all treatment effects such that $\phi_A(t) = a$ and $\hat{\tau}(a, \bullet, \bullet, \bullet)$ the corresponding adjusted mean.

Using (27) and (22), we find the factorial effects of Table 18.

The corresponding estimates are obtained by adding hats on α and τ. The factorial effects are given in function of the mean responses which are themselves expressed in function of the parameters α_x in Table 19. In that last table, the x are indexed as

Table 18
Factorial effects in Example 6.1

$\alpha_\emptyset$	=	$\tau(\bullet, \bullet, \bullet, \bullet)$
$\alpha_A(a)$	=	$\tau(a, \bullet, \bullet, \bullet) - \tau(\bullet, \bullet, \bullet, \bullet)$
$\alpha_B(b)$	=	$\tau(\bullet, b, \bullet, \bullet) - \tau(\bullet, \bullet, \bullet, \bullet)$
$\alpha_{AB}(a, b)$	=	$\tau(a, b, \bullet, \bullet) - \alpha_A(a) - \alpha_B(b) - \alpha_\emptyset$
	=	$\tau(a, b, \bullet, \bullet) - \tau(a, \bullet, \bullet, \bullet) - \tau(\bullet, b, \bullet, \bullet) + \tau(\bullet, \bullet, \bullet, \bullet)$
$\alpha_{AD}(a, d)$	=	$\tau(a, \bullet, \bullet, d) - \alpha_A(a) - \alpha_\emptyset$
	=	$\tau(a, \bullet, \bullet, d) - \tau(a, \bullet, \bullet, \bullet)$
$\alpha_{ABC}(a, b, c)$	=	$\tau(a, b, c, \bullet) - \alpha_{AB}(a, b) - \alpha_A(a) - \alpha_B(b) - \alpha_\emptyset$
	=	$\tau(a, b, c, \bullet) - \tau(a, b, \bullet, \bullet)$
$\alpha_{ABD}(a, b, d)$	=	$\tau(a, b, \bullet, d) - \alpha_{AD}(a, d) - \alpha_{AB}(a, b) - \alpha_A(a) - \alpha_B(b) - \alpha_\emptyset$
	=	$\tau(a, b, \bullet, d) - \tau(a, \bullet, \bullet, d) - \tau(a, b, \bullet, \bullet) + \tau(a, \bullet, \bullet, \bullet)$

Table 19
Mean responses in Example 6.1

$\tau(\bullet,\bullet,\bullet,\bullet)$	$=$	$\mu_{\emptyset}$	$=$	α_0
$\tau(a,\bullet,\bullet,\bullet)$	$=$	$\mu_A(a)$	$=$	$\alpha_0+\alpha_1x_1(a)$
$\tau(\bullet,b,\bullet,\bullet)$	$=$	$\mu_B(a)$	$=$	$\alpha_0+\alpha_2x_2(b)$
$\tau(a,b,\bullet,\bullet)$	$=$	$\mu_{AB}(a,b)$	$=$	$\alpha_0+\alpha_1x_1(a)+\alpha_2x_2(b)+\alpha_3x_3(a,b)$
$\tau(a,\bullet,\bullet,d)$	$=$	$\mu_{AD}(a,d)$	$=$	$\alpha_0+\alpha_1x_1(a)+\alpha_4x_4(a,d)+\alpha_5x_5(a,d)+\alpha_6x_6(a,d)$
$\tau(a,b,c,\bullet)$	$=$	$\mu_{ABC}(a,b,c)$	$=$	$\alpha_0+\alpha_1x_1(a)+\alpha_2x_2(b)+\alpha_3x_3(a,b)+\sum_{i=7}^{12}\alpha_ix_i(a,b,c)$
$\tau(a,b,\bullet,d)$	$=$	$\mu_{ABD}(a,b,d)$	$=$	$\alpha_0+\alpha_1x_1(a)+\alpha_2x_2(b)+\alpha_3x_3(a,b)$
				$+\sum_{i=4}^{6}\alpha_ix_i(a,d)+\sum_{i=13}^{15}\alpha_ix_i(a,b,d)$

Table 20
Factor efficiencies for the arrow defined design of Table 17

Factorial effect k	A	B	AB	AD	ABC	ABD
Σ_k	$\frac{11}{6}$	$\frac{11}{6}$	$\frac{11}{6}$	$\begin{bmatrix}\frac{4}{3}&0&0\\0&\frac{4}{3}&0\\0&0&2\end{bmatrix}$	$\begin{bmatrix}\frac{8}{3}&0&0&0&0&0\\0&\frac{8}{9}&0&0&0&0\\0&0&2&0&0&0\\0&0&0&2&0&0\\0&0&0&0&\frac{4}{3}&0\\0&0&0&0&0&\frac{4}{3}\end{bmatrix}$	$\begin{bmatrix}\frac{4}{3}&0&0\\0&\frac{4}{3}&0\\0&0&2\end{bmatrix}$
Factor efficiencies	$\frac{6}{11}$	$\frac{6}{11}$	$\frac{6}{11}$	$\begin{bmatrix}\frac{3}{4}&\frac{3}{4}&\frac{1}{2}\end{bmatrix}$	$\begin{bmatrix}\frac{9}{8}&\frac{3}{4}&\frac{3}{4}&\frac{1}{2}&\frac{1}{2}&\frac{3}{8}\end{bmatrix}$	$\begin{bmatrix}\frac{3}{4}&\frac{3}{4}&\frac{1}{2}\end{bmatrix}$

at the bottom of Table 17, then α_{x_i} is replaced by α_i and finally, $x_i(t)$ is replaced by $x_i(t_J)$ whenever $x_i \in \mathscr{X}_J$ and $\phi_J(t) = t_J$.

7. Factor efficiencies

Factor efficiencies are obtained by comparing the variances of estimation in the design under consideration to those that would be obtained with the reference design [12]. To take into account the difference between the number of units in these two designs, the variances are first transformed to *per unit* variances by multiplying them by the corresponding numbers of units.

The comparison is made for each factorial effect separately. If a factorial effect includes several parameters, the comparison is between the associated per unit co-variance matrices. Their simultaneous diagonalisation leads to the *principal* factor efficiencies.

The computation of efficiencies is straightforward if the parametrisation is defined by (42), where the vectors x are an orthonormal basis such as the one provided by Proposition 5.3. The *per unit* information matrix of the reference design is then the identity matrix and the per unit associated covariance matrix is $\sigma^2\mathbf{I}$. If $\sigma^2\Sigma$ is the corresponding per unit covariance matrix in the design under consideration, the factor efficiencies are immediately deduced from the blocks associated to the factorial effects on the diagonal of Σ. If Σ_k is the block associated with the kth factorial effect, the corresponding factor efficiencies are just the inverses of the eigenvalues of Σ_k.

Example 7.1. We consider the saturated design with the 16 treatments indicated by arrows on the leftside of Table 17, which was obtained with a D-optimal exchange algorithm. The corresponding X matrix contains the 16 corresponding lines of the table. The per unit information matrix is $M = X'X/16$ and $\Sigma = M^{-1}$. Table 20 gives the blocks Σ_k associated with the six factorial effects, which happen to be diagonal in that example, and the corresponding efficiencies.

Acknowledgements

Many modifications and additions were prompted by the comments and recommendations of G. Styan. The author wishes to thank him for his invaluable help in the redaction of his article.

References

[1] M. Almena, Y. Noël, A. Kobilinsky, A. Cepeda, Texture of Arzúa–Ulloa cheese: I. Evaluation before ripening using a fractional design, J. Dairy Res., submitted.
[2] S. Addelman, Orthogonal main effects plans for asymmetrical factorial experiments, Technometrics 4 (1962) 21–46.
[3] J.M. Azaïs, Analyse de variance non-orthogonale, L'exemple de SAS/GLM, Rev. Statist. Appl. 42 (2) (1994) 27–41.
[4] R.A. Bailey, Discussion of paper by T. Tjur, Int. Statist. Rev. 52 (1984) 65–77.
[5] R.A. Bailey, Orthogonal partitions in designed experiments, Designs, Codes and Cryptography 8 (1996) 45–77.
[6] R.A. Bailey, C.E. Praeger, C.A. Rowley, T.P. Speed, Generalized wreath products of permutation groups, Proc. London Math. Soc. 3 (47) (1983) 69–82.
[7] N. Bourbaki, Elements de mathématiques, Théorie des ensembles, chapter III: Ensemble ordonnés, cardinaux, nombres entiers, Hermann, Paris, 1977.
[8] S. Cliquet, C. Durier, A. Kobilinsky, Principle of a fractional factorial design for qualitative and quantitative factors: Application to the production of *Bradyrhizobium japonicum* in culture media, Agronomie 14 (1994) 569–587.
[9] M. Drton, Analyse de variance dans des situations hiérarchiques non-equirépétées, Mémoire de DEA mathématiques appliquées, Labo. Stat. Proba., Univ. Paul Sabatier, Toulouse, 1999.
[10] A. Kobilinsky, Confounding in relation to duality of finite Abelian groups, Linear Algebra Appl. 70 (1985) 321–347.

[11] A. Kobilinsky, Les plans factorials, in: J.J. Droesbeke, J. Fine, I. Saporta (Eds.), Plans d'expériences: applications à l'entreprise, TECHNIP, Paris, 1997, Ch. 3, pp. 69–209.
[12] A. Kobilinsky, H. Monod, Juxtaposition of regular factorial designs and the complex linear model, Scand. J. Statist. 22 (2) (1995) 223–254.
[13] MINITAB Inc., Reference manual, US ISBN 0 92 5636 22 3 (1994).
[14] J.A. Nelder, A reformulation of linear models (with discussion), J.R. Statist. Soc. A 140 (1) (1977) 48–77.
[15] H. Scheffe, The Analysis of Variance, Wiley, New York, 1959, 477 pp.
[16] SAS Institute, Inc., The four types of estimable functions, in: SAS/STAT User's Guide, Reference Version 6, fourth edition, SAS Institute Inc., Cary, NC, USA, 1990.
[17] SAS Institute, Inc., Tests of hypotheses in fixed-effects linear models, Technical Report R-101, Cary, NC, USA, 1978.
[18] S.R. Searle, Linear Models for Unbalanced Data, Wiley, New York, 1987, 536 pp.
[19] S.R. Searle, Analysis of variance computing package output for unbalanced data from fixed-effects models with nested factors, Amer. Statist. 48 (2) (1994) 148–153.
[20] T.P. Speed, R.A. Bailey, Factorial dispersion models, Internat. Statist. Rev. 55 (3) (1987) 261–277.
[21] T.P. Speed, R.A. Bailey, On a class of association schemes derived from lattices of equivalence relations, in: P. Schultz, C.E. Praeger, R.P. Sullivan (Eds.), Algebraic Structures and Applications, Marcel-Dekker, New York, 1982.
[22] S-Plus4, Guide to Statistics, MathSoft, Inc., Seattle, Wahington, 1997.
[23] SPSS Inc., SPSS Advanced Statistics 7.5, SPSS Inc, Chicago, 1997.
[24] T. Tjur, Analysis of variance models in orthogonal designs (with discussion), Internat. Statist. Rev. 52 (1984) 33–65.