The Annals of Statistics
1999, Vol. 27, No. 6, 1927–1947

BALANCE AND ORTHOGONALITY IN DESIGNS FOR MIXED CLASSIFICATION MODELS

By Dawn M. VanLeeuwen, David S. Birkes and Justus F. Seely

New Mexico State University, Oregon State University and Oregon State University

A classification model is easiest to analyze when it has a balanced design. Many of the nice features of balanced designs are retained by error-orthogonal designs, which were introduced in a recent paper by the authors. The present paper defines a kind of "partially balanced" design and shows that this partial balance is sufficient to ensure the error-orthogonality of a mixed classification model. Results are provided that make the partial balance condition easy to check. It is shown that, for a maximal-rank error-orthogonal design, the Type I sum of squares for a random effect coincides with the Type II sum of squares.

1. Introduction. The analysis of a classification model is relatively straightforward when the data are balanced, that is, when there are equal numbers of observations for all combinations of levels of the factors. Then the sums of squares for an ANOVA table are easily calculated and are unambiguous; for example, in a balanced two-way model, the sum of squares for the first factor is the same whether it is adjusted for the second factor or only for the mean. Under the assumption that the observations have a joint normal distribution, the sums of squares are independent and exact F-tests are available for many hypotheses of interest. Moreover, a complete sufficient statistic exists, which implies that uniformly minimum variance unbiased estimators can be obtained for all unbiasedly estimable functions of the model parameters.

A design does not necessarily have to be balanced to enjoy these desirable properties. Extending Houtman and Speed's (1983) definition of an orthogonal design, which is based on Nelder's (1965) concept of orthogonal block structure, VanLeeuwen, Seely and Birkes (1998) (herein abbreviated as VSB) defined the notion of an "error-orthogonal design," which is a design in which the least-squares estimator of the mean vector is a uniformly best linear unbiased estimator (UBLUE) and the covariance matrix of the vector of least-squares residuals has orthogonal block structure. The class of error-orthogonal designs includes all orthogonal designs, all balanced classification designs (provided they are "proper" as defined in Section 5), all fixed-effects models, and many other models; other examples are given in later sections. (See the remark at the end of this section regarding the terms "design" and

Received November 1996; revised July 1999.
AMS 1991 *subject classifications*. Primary 62J10; secondary 62K99.
Key words and phrases. Mixed linear model, ANOVA, sums of squares, orthogonal block structure, orthogonal design.

"model.") Error-orthogonal designs are similar to balanced designs in many ways.

The purposes of this paper are (1) to review the desirable features of error-orthogonal designs that have been established to date, (2) to prove another desirable feature of most error-orthogonal designs, namely, that Type I and Type II sums of squares coincide for a random effect, (3) to show that a certain type of balance in the incidence matrix implies error-orthogonality and (4) to provide convenient methods for checking the balance requirements.

For purpose (1) we list the following properties of an error-orthogonal design:

(a) By definition, the mean vector has a UBLUE. This implies that a UBLUE exists for every estimable linear function of the fixed effects. When there is no UBLUE, there is no clear choice for an estimator—one might choose the least-squares estimator or the BLUE corresponding to some particular prior guess for the covariance matrix or the BLUE corresponding to an estimate of the covariance matrix or the maximum likelihood estimator (MLE). For an error-orthogonal design, all these choices coincide [Szatrowski (1980), Theorem 2].

(b) Also by definition, the covariance matrix of the least-squares residuals has orthogonal block structure. This implies that the sum of the squares of the least-squares residuals has a "canonical" ANOVA decomposition into a sum of component sums of squares. The component sums of squares, under the assumption of normality, are independent and distributed as scalar multiples of chi-squared distributions. [In VSB see the paragraph that includes display (2) and the paragraph following. Also see Brown (1983) for ANOVA decompositions of the total sum of squares.] This leads to exact F-tests of variance components. These tests are uniformly most powerful unbiased (UMPU) [El-Bassiouni and Seely (1980)].

(c) The least-squares estimator of the mean vector together with the ANOVA sums of squares in (b), under normality, jointly form a complete sufficient statistic (Lemma 2.5 in VSB). This implies that the UBLUEs in (a) are actually uniformly minimum variance unbiased estimators (UMVUEs). Also, UMVUEs for variance components can be obtained as linear combinations of the ANOVA sums of squares.

(d) The residual maximum likelihood estimators (REMLEs) of the variance components have explicit expressions [El-Bassiouni (1983)], so that they can be computed without iteration.

(e) If an error-orthogonal classification design has maximal rank, which would usually be true, then the sums of squares for the random effects in the model are unambiguous in the sense that Type I and Type II sums of squares coincide (Theorem 6.3 below). Thus the Type I sums of squares, which have the appealing property that they add up to the total sum of squares, also have the property that each sum of squares for a random effect is adjusted for all effects not including it.

These are some of the benefits of using an error-orthogonal design. It should be noted we are not claiming that error-orthogonal designs are the only designs having these benefits. It is also worth noting that the optimal properties of being UBLU, UMVU, and UMPU, which are mentioned in the list of benefits, are optimal only in the context of a given design and that other designs may possibly provide better estimators and tests. The established advantage of error-orthogonal designs is not necessarily that they are optimal with respect to some numerical optimality criterion (although we conjecture that they are) but that they provide relatively "easy," "straightforward" analysis. The analysis is relatively straightforward because there is a clear choice for estimating a fixed effect [properties (a) and (c)], for estimating a variance component [property (c)], for testing a variance component [property (b)] and for forming an ANOVA table for the random effects [property (e)]. The analysis is relatively easy because there are noniterative formulas for the MLE of a fixed effect [property (a)] and for the REMLE of a variance component [property (d)].

It can be difficult to check error-orthogonality directly from the definition, and so it is helpful to have more convenient conditions to work with. In the VSB paper, some conditions for orthogonality and for error-orthogonality were derived, but the conditions apply primarily to models in which the random effects are either all nested or are all main effects. In the present paper, attention is restricted to classification models, but weaker conditions are obtained that can be applied to all types of classification models. The conditions here seem more convenient to verify than those in VSB in that they are expressed in terms of patterns in the incidence matrix [similar to conditions given in Brown (1983)].

Further facts about error-orthogonality from the VSB paper are reviewed in Section 2. Notation and terminology are presented in Section 3. In Section 4 we introduce some notions of "partially balanced" designs. Corollary 5.2 states that a certain kind of partial balance implies error-orthogonality. A simple illustration of this result is provided by a random main-effects-only model. Consider the two-way designs obtained by ignoring all but two factors. If, for all pairs of factors, these two-way designs are balanced, then the model has an error-orthogonal design.

The equality of Type I and Type II sums of squares for random effects in error-orthogonal designs of maximal rank is shown in Section 6. The results in Section 7 facilitate the process of checking the partial balance condition. Proofs of lemmas and theorems are collected in Section 8.

REMARK. For brevity we sometimes abuse terminology by saying "a such-and-such design" or "a such-and-such model" to mean "a model with a such-and-such design." For example, the properties listed above for error-orthogonal designs would be more precisely described as properties of models having error-orthogonal designs. And the phrase "balanced model" means "model with a balanced design." We hope this will cause no confusion.

2. Review of error-orthogonal designs. This section summarizes some basic facts from VSB on error-orthogonal designs. First, we need some terminology and notation concerning matrices. A symmetric idempotent matrix is called an *orthogonal projection matrix*, or in this paper, since no other kind of projection matrix occurs, simply a *projection matrix*. The range (or column space) of a matrix A is denoted by $\mathscr{R}(A)$, and the unique projection matrix whose range is $\mathscr{R}(A)$ is denoted by P_A. A set of matrices is said to be *pairwise orthogonal* if $A'B = 0$ for all distinct A and B in the set. The set is said to be *commutative* if $AB = BA$ for all A and B in the set. The *linear span* of the set consists of all linear combinations of the matrices in the set; for a set $\mathscr{A}$, the linear span is denoted by $\operatorname{sp}\mathscr{A}$.

Consider the following general mixed linear model (not necessarily a classification model) for a random vector Y:

$$(2.1)\qquad E(Y) = X\beta \quad \text{and} \quad \operatorname{Cov}(Y) = V_\theta = \theta_1 V_1 + \cdots + \theta_r V_r + \sigma^2 I,$$

where $X, V_1, \ldots, V_r$ are known matrices, β is an unconstrained vector of unknown fixed effects and $\theta = (\theta_1, \ldots, \theta_r, \sigma^2)$ is a vector of unknown variance–covariance parameters. The only conditions on the variance–covariance parameters are that the set Θ of possible vectors θ contains a nonempty open subset of $(r+1)$-dimensional Euclidean space (this will be called the *open set assumption*) and that the set $\mathscr{V} = \{V_\theta \colon \theta \in \Theta\}$ of all possible covariance matrices of Y consists of positive-definite matrices. One implication of the open set assumption is that $\operatorname{sp}\mathscr{V} = \operatorname{sp}\{V_1, \ldots, V_r, I\}$.

Error-orthogonality involves the notion of orthogonal block structure, which was introduced by Nelder (1965). The definition below is a slight alteration of the one given by Houtman and Speed [(1983), page 1070] and is applied to the vector of least-squares residuals rather than to the data vector Y. The least-squares estimator of $X\beta$ is $P_X Y$ and the vector of least-squares residuals is MY where $M = I - P_X$. Note that the set of possible covariance matrices of MY can be expressed as $M\mathscr{V}M = \{MVM \colon V \in \mathscr{V}\}$.

Definition. $\operatorname{Cov}(MY)$ is said to have *orthogonal block structure* (OBS) if $M\mathscr{V}M = \{\sum_{i=1}^{s} \pi_i F_i \colon (\pi_1, \ldots, \pi_s) \in \Pi\}$ for some s where $F_1, \ldots, F_s$ are nonzero pairwise-orthogonal projection matrices and Π contains a nonempty open subset of s-dimensional Euclidean space.

OBS can also be characterized in terms of the concept of a quadratic subspace (or Jordan algebra), which has been a useful tool in linear model theory for studying complete sufficiency [Seely (1971), (1972), (1977)] and maximum likelihood [Rao and Kleffe (1988)] under normality, for nonnegative unbiased estimation of variance components [Pukelsheim (1981)], and for optimal design [Pukelsheim (1993)].

Definition. A linear subspace $\mathscr{L}$ of symmetric matrices is said to be a *quadratic subspace* if $A^2 \in \mathscr{L}$ for all $A \in \mathscr{L}$.

LEMMA 2.2 (See Lemma 2.1 in VSB). Cov(MY) *has OBS if and only if* sp $M\mathscr{V}M$ *is a commutative quadratic subspace.*

OBS pertains only to the covariance structure of the model. To ensure a straightforward analysis, a suitable relationship between the covariance structure and the mean vector must also be required. The existence of a UBLUE for the mean vector provides such a relationship. An estimator of a linear parametric function $a'\beta$ is a *uniformly best linear unbiased estimator* (UBLUE) if it is a linear function of Y, is unbiased for $a'\beta$ and, among all linear unbiased estimators of $a'\beta$, has the minimum variance for all possible values of the variance–covariance parameters. A vector-valued estimator is a UBLUE for a vector of linear parametric functions $A'\beta$ if each component of the estimator is a UBLUE for the corresponding component of $A'\beta$. If $X\beta$ has a UBLUE, we say that E(Y) has a UBLUE.

LEMMA 2.3 (See Lemma 2.3 in VSB). *The following statements are equivalent*:

(a) $P_X Y$ *is a UBLUE for* $X\beta$.
(b) *Every estimable linear function of* β *has a UBLUE.*
(c) $\mathscr{R}(V_i X) \subset \mathscr{R}(X)$ *for all* $i = 1, \dots, r$.
(d) M *commutes with* $\mathscr{V}$.

DEFINITION. A linear model for Y is said to have an *error-orthogonal design* (or, simply, to be *error-orthogonal*) if E(Y) has a UBLUE and Cov(MY) has OBS.

Typically the easiest way to establish the UBLUE property is to verify statement (c) of Lemma 2.3. To establish the OBS property, the following lemma can be handy.

LEMMA 2.4 (See Lemma 2.6 in VSB). *Suppose* E(Y) *has a UBLUE. Then,* Cov(MY) *has OBS if and only if* $V_i V_j = W_{ij} + Z_{ij}$ *for all* $1 \le i \le j \le r$, *where* $W_{ij} \in$ sp $\mathscr{V}$ *and either* $\mathscr{R}(Z_{ij}) \subset \mathscr{R}(X)$ *or* $\mathscr{R}(Z'_{ij}) \subset \mathscr{R}(X)$.

3. Notation and terminology. The focus of this paper is on classification models. In order to deal with general classification models having an arbitrary number of factors and possibly including some interactions and nested effects, we need notation that is general but not too cumbersome. For specific examples of the general notation introduced in this section, see the examples in Sections 5, 6 and 7.

The incidence matrix. Consider data $Y_{i_1 \cdots i_p k}$ that is classified according to p factors. The factors can be labeled by the integers $1, \dots, p$. For each factor g the index i_g ranges from 1 to t_g, which denotes the number of levels of factor g. The range of the index k is from 1 to $n_{i_1 \cdots i_p}$, which denotes the number of observations at level i_1 of factor $1, \dots,$ and level i_p of factor p. The

p-dimensional $t_1 \times \cdots \times t_p$ matrix of $n_{i_1 \cdots i_p}$'s is called the *incidence matrix* and is denoted by N. Each position, or *cell*, in the incidence matrix corresponds to a combination of levels of all the factors.

For a factor that is nested, we are requiring that the number of levels of the factor be the same within each combination of levels of the factors in which it is nested. However, note that even if the numbers of observed levels of a nested factor are different within the different combinations of the nesting factors, the numbers of levels can be made the same by introducing unobserved levels with corresponding entries 0 in the incidence matrix.

If $n_{i_1 \cdots i_p}$ is the same positive integer for all $i_1, \dots, i_p$, the design is said to be *balanced*, or for emphasis, since we will be dealing with some notions of "partially balanced" designs, *completely balanced*.

To every subset $\mathscr{F} = \{g_1, \dots, g_m\}$ of factors there corresponds a *marginal incidence matrix*, which is the m-dimensional matrix obtained from the incidence matrix N by summing over the indices for the other $p - m$ factors. We denote the marginal incidence matrix by $N^{(v)}$ and its entries by either $n_w^{(v)}$ or $n(v)[w]$, where $v = (g_1, \dots, g_m)$ is the subset of factors put into vector form, usually in increasing order, and $w = (i_1, \dots, i_m)$ varies over all combinations of levels of the factors. Thus $n(v)[w]$ is the number of observations at level i_1 of factor $g_1, \dots,$ and level i_m of factor g_m. Usually the notation for w indicates what v must be and we write simply $n[w]$. The matrix $N^{(v)}$ can be regarded, at least when none of the factors in v is nested in factors not in v, as the incidence matrix of a model obtained by dropping the other $p - m$ factors from the original model.

The model. A mixed classification model for the data can be expressed as

$$Y_{i_1 \cdots i_p k} = \mu + (\text{a sum of unknown fixed effects}) + (\text{a sum of unobserved random effects}) + e_{i_1 \cdots i_p k},$$

where μ is a fixed overall mean effect and $e_{i_1 \cdots i_p k}$ is a random error term. We usually use Greek letters for fixed effects and roman letters for random effects. The fixed and random effects may be main effects, interaction effects or nested effects. In matrix notation, a mixed classification model has the form

$$Y = \mathbf{1}\mu + G_1\tau_1 + \cdots + G_q\tau_q + H_1 d_1 + \cdots + H_r d_r + e, \tag{3.1}$$

where $\mathbf{1}$ denotes a column of 1's and each τ_j (respectively, d_j) is a vector of fixed (respectively, random) effects for the observed combinations of levels of the factors in a particular subset of factors. If τ_j is associated with a subset of factors $\mathscr{F}$, then its effects are called *$\mathscr{F}$-effects*. For example, if τ_j is associated with $\mathscr{F} = \{g\}$, then it is a vector of length t_g, and each of its entries is a parameter representing the effect of one of the levels of the gth factor. If τ_j is associated with $\mathscr{F} = \{g, h\}$, then it is a vector of length $t_g t_h$, at least when all combinations of levels of the two factors are observed. When some combinations are not observed, then we can choose either to remove or retain entries of τ_j (or d_j) that correspond to unobserved combinations of levels, and hence either to remove or retain columns of zeros in the matrix G_j (or H_j). If τ_j is associated with $\mathscr{F} = \{g, h\}$ and the two factors are cross-classified, then each of its entries is an interaction effect for a combination of levels of the gth and hth factors. If the two factors are nested, say with the hth factor nested in the gth factor, then each entry is the effect of a particular level of the hth factor within a particular level of the gth factor.

It makes no difference whether we remove or retain columns of zeros in the matrices G_j and H_j, because it does not affect $\mathscr{R}(X)$ nor V_j and hence does not affect the conditions for OBS and UBLUEs given in Section 2.

If a model includes $\mathscr{F}$-effects, then we call $\mathscr{F}$ an *included* subset of factors. For the empty set $\varnothing$, the $\varnothing$-effect is defined to be the overall mean μ. We always include μ in our classification models and so the empty set $\varnothing$ is an included subset of factors.

When $\text{Cov}(Y) = \sigma^2 I$, model (3.1) is called a *fixed-effects* model. When $X = \mathbf{1}$ and $r \geq 1$, the model is called a *random-effects* model. When Y has a multivariate normal distribution, the model is called a *normal* model.

ASSUMPTIONS. The remainder of the paper is concerned with the mixed classification model (3.1) under the following assumptions. The fixed effects are unknown parameters and no constraints are imposed on them. (A model with linear constraints on the fixed effects can be reparametrized as a model with no constraints on the fixed effects.) The random effects, that is, the entries of $d_1, \dots, d_r, e$, are unobservable random variables. These random variables are assumed to be uncorrelated with one another and to have mean 0. The entries of d_j are assumed to have a common unknown variance σ_j^2 ($j = 1, \dots, r$) and the entries of e are assumed to have a common unknown variance σ^2. With these assumptions, model (3.1) can be expressed in the form of model (2.1) with $X = (\mathbf{1}, G_1, \dots, G_q)$, $\beta = (\mu, \tau_1', \dots, \tau_q')'$, $V_j = H_j H_j'$, $\theta_j = \sigma_j^2$. As in Section 2, it is assumed that the set of possible vectors of variance parameters satisfies the open set assumption and that all possible covariance matrices of Y are positive-definite.

Single-effect matrices. Single-effect matrices are the basic components for expressing a classification model in matrix form. A matrix K is said to be a *single-effect matrix* if the entries of K are all 0's and 1's with exactly one 1 in each row. (In VSB these were called classification matrices and the definition was slightly less general.) Such a matrix is the model matrix for a one-way classification model with each column corresponding to a classification group. Each matrix G_j or H_j in (3.1) is a single-effect matrix.

4. Balance. In this section several notions of "balance" are defined for incidence matrices and classification models. For examples of these concepts and the notation, see the examples in Sections 5, 6 and 7.

DEFINITIONS. Let $\mathscr{F}$ and $\mathscr{G}$ be two subsets of factors. The following definitions do not require that the effects corresponding to $\mathscr{F}$ and $\mathscr{G}$ are actually included in the model.

(a) The incidence matrix N is said to be *balanced* with respect to $\mathscr{F}$ if all entries in the marginal incidence matrix $N^{(v)}$ are the same (and positive), where v is the vector form of $\mathscr{F}$. This property is denoted by $\mathrm{Bal}(\mathscr{F})$.

(b) N is *conditionally balanced* with respect to $\mathscr{F}$ given $\mathscr{G}$ if, for every given combination w_g of levels of the factors in $\mathscr{G}$, the number of observations $n[w_f^* w_g]$ is the same for all combinations w_f^* of levels of the factors in $\mathscr{F}^* = \mathscr{F} \setminus \mathscr{G}$. (The notation $\mathscr{A} \setminus \mathscr{B}$ designates the subset of members of the set $\mathscr{A}$ that are not in the set $\mathscr{B}$.) This property is denoted by $\mathrm{Bal}(\mathscr{F} \mid \mathscr{G})$. Note that the phrases "every given combination" and "all combinations" refer to all possible combinations, including combinations that are not observed.

(c) A design or model is said to be $\mathrm{Bal}(\mathscr{F})$ or $\mathrm{Bal}(\mathscr{F} \mid \mathscr{G})$ if its incidence matrix has that property.

Complete balance is equivalent to balance with respect to the set of all p factors. As an example of conditional balance, if an incidence matrix is conditionally balanced with respect to factors $\{1, 2, 3, 4\}$ given factors $\{3, 4, 5\}$, then $n[i_1 i_2 i_3 i_4 i_5]$ is constant over all levels i_1, i_2 for any given levels i_3, i_4, i_5; that is, we can write $n[i_1 i_2 i_2 i_4 i_5] = m[i_3 i_4 i_5]$.

LEMMA 4.1. *Let $\mathscr{E}$, $\mathscr{F}$, $\mathscr{G}$ and $\mathscr{H}$ be subsets of factors.*

(a) *For the empty set* $\varnothing$, $\mathrm{Bal}(\mathscr{H} \mid \varnothing) \Leftrightarrow \mathrm{Bal}(\mathscr{H})$.
(b) *If* $\mathscr{G} \subset \mathscr{F} \cup \mathscr{E}$ *and* $\mathscr{H} \setminus \mathscr{G} \subset \mathscr{F} \setminus \mathscr{E}$, *then* $\mathrm{Bal}(\mathscr{F} \mid \mathscr{E}) \Rightarrow \mathrm{Bal}(\mathscr{H} \mid \mathscr{G})$.
(c) *If* $\mathscr{H} \subset \mathscr{F}$, *then* $\mathrm{Bal}(\mathscr{F}) \Rightarrow \mathrm{Bal}(\mathscr{H})$.
(d) *If* $\mathscr{H} \subset \mathscr{F} \setminus \mathscr{E}$, *then* $\mathrm{Bal}(\mathscr{F} \mid \mathscr{E}) \Rightarrow \mathrm{Bal}(\mathscr{H})$.
(e) *If* $\mathscr{H} \cup \mathscr{G} \subset \mathscr{F}$, *then* $\mathrm{Bal}(\mathscr{F}) \Rightarrow \mathrm{Bal}(\mathscr{H} \mid \mathscr{G})$.
(f) *If* $\mathscr{H} \setminus \mathscr{G} \subset \mathscr{F}$, *then* $\mathrm{Bal}(\mathscr{F} \mid \mathscr{G}) \Rightarrow \mathrm{Bal}(\mathscr{H} \mid \mathscr{G})$.
(g) *If* $\mathscr{H} \setminus \mathscr{G} \subset \mathscr{F} \subset \mathscr{H} \cup \mathscr{G}$, *then* $\mathrm{Bal}(\mathscr{F} \mid \mathscr{G}) \Leftrightarrow \mathrm{Bal}(\mathscr{H} \mid \mathscr{G})$.
(h) *If* $\mathscr{G} \subset \mathscr{E}$ *and* $\mathscr{H} \cap \mathscr{E} = \mathscr{H} \cap \mathscr{G}$, *then* $\mathrm{Bal}(\mathscr{H} \mid \mathscr{E}) \Rightarrow \mathrm{Bal}(\mathscr{H} \mid \mathscr{G})$.
(i) *If* $\mathscr{E} \subset \mathscr{G}$ *and* $\mathscr{H} \cup \mathscr{E} = \mathscr{H} \cup \mathscr{G}$, *then* $\mathrm{Bal}(\mathscr{H} \mid \mathscr{E}) \Rightarrow \mathrm{Bal}(\mathscr{H} \mid \mathscr{G})$.
(j) $\mathrm{Bal}(\mathscr{H} \cup \mathscr{G}) \Leftrightarrow \mathrm{Bal}(\mathscr{H} \mid \mathscr{G})$ *and* $\mathrm{Bal}(\mathscr{G})$.

Part (a) is trivial but it is worth noting that unconditional balance can be regarded as a special case of conditional balance. Parts (c) through (i) are special cases of (b). Parts (b) and (j) are proved in Section 8.

The following lemma shows how balance in the incidence matrix with respect to subsets of factors is reflected in properties of the corresponding single-effect matrices. Note that the effects to which the matrices correspond are not required to be included effects.

LEMMA 4.2. *Let $\mathscr{F}$ and $\mathscr{G}$ be subsets of factors and let $\mathscr{L} = \mathscr{F} \cap \mathscr{G}$. Let F, G and L be the corresponding single-effect matrices.* (*If $\mathscr{L}$ is empty*, *let* $L = \mathbf{1}$.)

(a) *If* $\mathscr{F} \subset \mathscr{G}$, *then* $\mathscr{R}(F) \subset \mathscr{R}(G)$.
(b) *If the design is* $\mathrm{Bal}(\mathscr{F})$, *then* $FF' = mP_F$ *for some positive integer* m.
(c) *If the design is* $\mathrm{Bal}(\mathscr{F} \mid \mathscr{G})$, *then* $\mathscr{R}(FF'G) = \mathscr{R}(L)$.

(d) *If the design is* $\mathrm{Bal}(\mathscr{F})$ *and* $\mathrm{Bal}(\mathscr{F} \mid \mathscr{G})$, *then* $P_F P_G = P_L$.

Part (d) says that with suitable "partial" balance in the incidence matrix, the projection matrices associated with the effects in the model behave just as they would with complete balance. The proof is given in Section 8.

The concepts of balance and conditional balance for an incidence matrix depend only on the pattern of observations relative to the factorial structure of the data and do not depend on the particular model that is chosen. Next we define some types of balance for a mixed classification model. These definitions depend on which effects are included in the model and on the status of the effects as fixed or random. If $\mathscr{F}$-effects are included and are fixed (respectively, random), then $\mathscr{F}$ is called a *fixed-effect* (respectively, *random-effect*) subset.

DEFINITIONS. Consider a mixed classification model.

(a) The model is said to be *b-balanced* (or *BLUE-balanced*) if it is $\mathrm{Bal}(\mathscr{H} \mid \mathscr{G})$ for all random-effect subsets $\mathscr{H}$ and all fixed-effect subsets $\mathscr{G}$.

(b) The model is *r-balanced* (or *random-pairwise balanced*) if it is $\mathrm{Bal}(\mathscr{H} \cup \mathscr{L})$ for all random-effect subsets $\mathscr{H}$ and $\mathscr{L}$.

(c) The model is *b & r-balanced* if it is b-balanced and r-balanced.

(d) The model is *weakly b-balanced* if, for every random-effect subset $\mathscr{H}$ and fixed-effect subset $\mathscr{G}$, there exists a subset of factors $\mathscr{F}$ (not necessarily included) such that $\mathscr{G} \subset \mathscr{F}$, $\mathscr{H} \cap \mathscr{F}$ is contained in a fixed-effect subset, and the model is $\mathrm{Bal}(\mathscr{H} \mid \mathscr{F})$.

(e) The model is *p-balanced* (or *pairwise balanced*) if it is $\mathrm{Bal}(\mathscr{F} \cup \mathscr{G})$ for all included subsets $\mathscr{F}$ and $\mathscr{G}$.

For instance, consider a main-effects-only classification model. The model is b-balanced provided that, for every fixed factor g and random factor h, the entries within each row of $N^{(gh)}$ are all the same. If the model has only a single random factor h, then it is r-balanced provided that the entries of $N^{(h)}$ are all the same. If the model has more than one random factor, then the model is r-balanced provided that, for each distinct pair of random factors g and h, the entries of $N^{(gh)}$ are all the same. In the terminology of VSB, this is the same as requiring that the corresponding single-effect matrices G and H be J-balanced.

In a general mixed classification model the following facts are not hard to prove.

LEMMA 4.3. *Consider a mixed classification model.*

(a) *If the model is* b-*balanced*, *then it is weakly* b-*balanced*.

(b) *If the model is* p-*balanced*, *then it is* b & r-*balanced*.

(c) *A random-effects model is* p-*balanced if and only if it is* b & r-*balanced*.

(d) *If the model includes two random effects such that the union of the two subsets of factors is the set of all factors*, *then the model is* r-*balanced if and only if it is completely balanced*.

A model that is weakly b-balanced and not b-balanced is given in Example 5.4.

5. Error-orthogonality in mixed classification models. In this section sufficient conditions are established for a mixed classification model to have an error-orthogonal design.

Most (if not all) classification models that occur in practice satisfy the following property. A classification model is said to be *proper* if, whenever $\mathscr{F}$ and $\mathscr{G}$ are both random-effect subsets, then either $\mathscr{F} \cap \mathscr{G}$ is a random-effect subset or it is contained in a fixed-effect subset. A proper classification model may include $\{1,2\}$-effects without necessarily including $\{2\}$-effects. This happens when factor 2 is nested within factor 1. If the model includes both random $\{1,2\}$-effects and random $\{2,3\}$-effects, then properness requires that either $\{2\}$-effects must be included or else there must exist fixed $\mathscr{F}$-effects with $2 \in \mathscr{F}$.

THEOREM 5.1. (a) *If a mixed classification model is weakly* b-*balanced, then* $\mathrm{E}(Y)$ *has a UBLUE.*

(b) *If a proper mixed classification model is* r-*balanced and* $\mathrm{E}(Y)$ *has a UBLUE, then the model is error-orthogonal.*

The proof is given in Section 8. Weak b-balance is sufficient to ensure that $\mathrm{E}(Y)$ has a UBLUE, but it is not necessary; see Example 6.4 below.

COROLLARY 5.2. *If a proper mixed classification model is weakly* b-*balanced and* r-*balanced, then it is error-orthogonal.*

Therefore, by Lemma 4.3(a),(b), if a proper model is b & r-balanced, or in particular if it is p-balanced, then it is error-orthogonal.

If a classification model is not proper, then it may not have an error-orthogonal design even when the incidence matrix is completely balanced. This follows from Seifert [(1979), Theorem 2] in which necessary and sufficient conditions were given, in terms of subsets of factors, for a normal mixed classification model with a completely balanced incidence matrix to admit a complete sufficient statistic. For example, because it does not satisfy Seifert's condition (i), the balanced random-effects model

$$Y_{ijku} = \mu + a_i + b_j + c_{ik} + d_{jk} + e_{ijku}$$

has no complete sufficient statistic and hence does not have an error-orthogonal design [see property (c) in Section 1 above]. Note that the model is not proper because $\{1,3\}$ and $\{2,3\}$ are included subsets of factors but $\{1,3\} \cap \{2,3\} = \{3\}$ is not an included subset nor is it contained in $\varnothing$, the only fixed-effect subset.

EXAMPLE 5.3. Consider a three-way main-effects-only model:

$$Y_{ijku} = \mu + \alpha_i + \beta_j + \gamma_k + e_{ijku},$$

where $i = 1, 2, 3$, $j = 1, 2, 3$, $k = 1, 2$ and $u = 1, \dots, n_{ijk}$. Whether or not the design is error-orthogonal depends on which effects are random and on the balance properties of the incidence matrix.

(a) Suppose all three factors are random. Then by Lemma 4.3(c), b & r-balance is equivalent to p-balance, that is, Bal({1, 2}), Bal({1, 3}) and Bal({2, 3}). For example, the following design is p-balanced and hence error-orthogonal:

$$\text{Design I:} \qquad (n_{ij1}) = \begin{bmatrix} 0 & 2 & 4 \\ 2 & 4 & 0 \\ 4 & 0 & 2 \end{bmatrix}, \qquad (n_{ij2}) = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 0 & 4 \\ 0 & 4 & 2 \end{bmatrix}.$$

The p-balance property can be verified by forming the three marginal two-factor incidence matrix and noting that the entries of $N^{(12)}$ are all 4's, the entries of $N^{(13)}$ are all 6's, and the entries of $N^{(23)}$ are all 6's. This design is error-orthogonal regardless of which factors are random or fixed.

(b) Suppose factor 1 is fixed and factors 2 and 3 are random. Then, the design is b & r-balanced if and only if it is Bal({2} | {1}), Bal({3} | {1}) and Bal({2, 3}). Design I is b & r-balanced. Also the following design is b & r-balanced and hence error-orthogonal:

$$\text{Design II:} \qquad (n_{ij1}) = \begin{bmatrix} 0 & 1 & 2 \\ 2 & 4 & 0 \\ 4 & 1 & 4 \end{bmatrix}, \qquad (n_{ij2}) = \begin{bmatrix} 2 & 1 & 0 \\ 2 & 0 & 4 \\ 2 & 5 & 2 \end{bmatrix}.$$

The b & r-balance can be verified by forming the marginal incidence matrices:

$$N^{(12)} = \begin{bmatrix} 2 & 2 & 2 \\ 4 & 4 & 4 \\ 6 & 6 & 6 \end{bmatrix}, \qquad N^{(13)} = \begin{bmatrix} 3 & 3 \\ 6 & 6 \\ 9 & 9 \end{bmatrix}, \qquad N^{(23)} = \begin{bmatrix} 6 & 6 \\ 6 & 6 \\ 6 & 6 \end{bmatrix}.$$

Since the entries within each row of $N^{(12)}$ and $N^{(13)}$ are the same, the design is Bal({2} | {1}) and Bal({3} | {1}). Since the entries of $N^{(23)}$ are all 6's, the design is Bal({2, 3}).

(c) Suppose factors 1 and 2 are fixed and factor 3 is random. Then, the design is b & r-balanced if and only if it is Bal({3} | {1}), Bal({3} | {2}) and Bal({3}). Designs I and II are b & r-balanced. Also the following design is b & r-balanced and hence error-orthogonal:

$$\text{Design III:} \qquad (n_{ij1}) = \begin{bmatrix} 0 & 2 & 3 \\ 1 & 1 & 3 \\ 2 & 3 & 2 \end{bmatrix}, \qquad (n_{ij2}) = \begin{bmatrix} 0 & 4 & 1 \\ 2 & 1 & 2 \\ 1 & 1 & 5 \end{bmatrix}.$$

The b & r-balance can be verified by forming the marginal incidence matrices $N^{(13)}$ and $N^{(23)}$ and noting that the entries within each row are the same. This verifies that the design is Bal({3} | {1}) and Bal({3} | {2}). To show that the design is Bal({3}), we can invoke Lemma 4.1(d) or simply note that $N^{(3)} = [17 \;\; 17]$.

(d) Suppose all three factors are fixed. Then all designs are error-orthogonal.

EXAMPLE 5.4. Consider a three-way mixed model with one two-factor interaction,

$$Y_{ijku} = \mu + \alpha_i + b_j + (\alpha b)_{ij} + \gamma_k + e_{ijku},$$

where $i = 1, 2$, $j = 1, 2$, $k = 1, 2$, $u = 1, \ldots, n_{ijk}$, α_i is the fixed effect of the ith level of factor 1, b_j is the random effect of the jth level of factor 2, $(\alpha b)_{ij}$ is a random interaction and γ_k is the fixed effect of the kth level of factor 3. Suppose the incidence matrix is the following:

$$(n_{ij1}) = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}, \qquad (n_{ij2}) = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}.$$

The fixed-effect subsets of factors are $\emptyset$, $\{1\}$ and $\{3\}$ and the random-effect subsets are $\{2\}$ and $\{1, 2\}$. The design is not b-balanced because it is not $\mathrm{Bal}(\{1, 2\} \mid \{3\})$; that is, n_{ijk} does not depend only on k. Nevertheless, the design is error-orthogonal because it is weakly b-balanced and r-balanced, as is shown in Example 7.3.

6. Unambiguous sums of squares. Analysis of a classification model is often based on an ANOVA table which associates a sum of squares to each of the effects in the model. The sum of squares for an effect is usually adjusted for other effects, but it is not always clear exactly which other effects to adjust for. Before investigating this ambiguity, we need some notation and terminology.

NOTATION AND TERMINOLOGY. Consider a mixed classification model. Let $\mathbb{C} = \{\mathscr{F}_1, \ldots, \mathscr{F}_s\}$ be a collection of subsets of factors and let $F_1, \ldots, F_s$ be the corresponding single-effect matrices.

(a) An effect is called a $\mathbb{C}$-effect if it is an $\mathscr{F}_i$-effect for some $i = 1, \ldots, s$.
(b) The *model matrix corresponding to the* $\mathbb{C}$-*effects* is $C = (F_1, \ldots, F_s)$.
(c) The *unadjusted sum of squares for* $\mathbb{C}$-*effects* is $\mathrm{SS}(\mathbb{C}) = Y'P_C Y$.
(d) Let $\mathbb{D}$ be another collection of subsets of factors. The *sum of squares for* $\mathbb{C}$-*effects adjusted for* $\mathbb{D}$-*effects* is $\mathrm{SS}(\mathbb{C} \mid \mathbb{D}) = \mathrm{SS}(\mathbb{C} \cup \mathbb{D}) - \mathrm{SS}(\mathbb{D})$.
(e) Let $\mathbb{G}$ denote the collection of all fixed-effect subsets.
(f) If $\mathscr{F}$ and $\mathscr{G}$ are two subsets of factors with $\mathscr{F} \subset \mathscr{G}$, we say that $\mathscr{F}$-effects are *contained* in $\mathscr{G}$-effects.
(g) Given a random-effect subset of factors $\mathscr{H}$, let $\mathbb{H}_1$ (respectively, $\mathbb{H}_2$, and $\mathbb{H}_3$) denote the collection of all random-effect subsets, not including $\mathscr{H}$ itself, that are contained in $\mathscr{H}$ (respectively, neither are contained in nor contain $\mathscr{H}$, and contain $\mathscr{H}$).

Consider a random-effect subset of factors $\mathscr{H}$. In forming the sum of squares for $\mathscr{H}$-effects, which effects should be adjusted for? Usually the sum of squares for a random effect is adjusted for all the fixed effects, that is, for all $\mathbb{G}$-effects. It is also usual to adjust for all effects that are contained in the $\mathscr{H}$-effects, that is, for all $\mathbb{H}_1$-effects. The sum of squares would certainly not be adjusted for any $\mathbb{H}_3$-effects because $\mathrm{SS}(\mathscr{H} \mid \mathscr{H}_3) = 0$ for all $\mathscr{H}_3 \in \mathbb{H}_3$. However, it is not clear which, if any, $\mathbb{H}_2$-effects should be adjusted for. Theorem 6.3 below states that in a maximal-rank error-orthogonal model, the same sum of squares is obtained no matter which $\mathbb{H}_2$-effects are adjusted for. In this sense the sum of squares for a random effect can be said to be "unambiguous."

LEMMA 6.1. *In an error-orthogonal linear model, let* $A, B \in \operatorname{sp} \mathscr{V}$. (a) $MP_A = P_A M = P_{MA}$, (b) $\mathscr{R}(P_{X,A} P_{X,B}) \subset \mathscr{R}(X) + \mathscr{R}(P_A P_B)$, (c) $P_{X,A}$ *and* $P_{X,B}$ *commute*.

The matrix $(\mathbf{1}, G_1, \dots, G_q, H_1, \dots, H_r)$ is called the *all-included-effects model matrix*. A classification model is said to have *maximal rank* if the rank of the all-included-effects model matrix is the same as it would be if all the entries in its incidence matrix were nonzero.

LEMMA 6.2. *Consider a proper mixed classification model. Let H be the single-effect matrix corresponding to a random-effect subset* $\mathscr{H}$, *let U be the model matrix corresponding to the* $\mathbb{H}_1$*-effects, and let W be a model matrix corresponding to some or all of the* $\mathbb{H}_2$*-effects. If either* (*a*) *the model has maximal rank or* (*b*) *the model is r-balanced and* $E(Y)$ *has a UBLUE, then* (*c*) $\mathscr{R}(X, H) \cap \mathscr{R}(X, W) \subset \mathscr{R}(X, U)$.

THEOREM 6.3. *In a proper mixed classification model, let* $\mathscr{H}$ *be a random-effect subset, let* $\mathbb{G}$, $\mathbb{H}_1$ *and* $\mathbb{H}_2$ *be as defined at the beginning of this section and let* $\mathbb{L}_2 \subset \mathbb{H}_2$. *If the model is error-orthogonal and either* (*a*) *has maximal rank or* (*b*) *is r-balanced, then* (*c*) $SS(\mathscr{H} \mid \mathbb{G}, \mathbb{H}_1, \mathbb{L}_2) = SS(\mathscr{H} \mid \mathbb{G}, \mathbb{H}_1)$.

For $\mathbb{L}_2 = \mathbb{H}_2$, $\mathrm{SS}(\mathscr{H} \mid \mathbb{G}, \mathbb{H}_1, \mathbb{H}_2)$ is called the *Type II* sum of squares for $\mathscr{H}$-effects. The *Type I* sum of squares for $\mathscr{H}$-effects is $\mathrm{SS}(\mathscr{H} \mid \mathbb{G}, \mathbb{H}_1, \mathbb{L}_2)$ where $\mathbb{L}_2$ consists of all random-effect subsets in $\mathbb{H}_2$ that precede $\mathscr{H}$ in some particular ordering of the included subsets of factors. The theorem implies that the Type I (no matter what ordering of the included subsets of factors is chosen) and Type II sums of squares for $\mathscr{H}$-effects coincide.

Most examples of error-orthogonal models in our experience satisfy both (a) and (b). We are not aware of any error-orthogonal proper classification model that has maximal rank but is not r-balanced. Example 6.4 below presents an error-orthogonal proper classification model that is r-balanced but does not have maximal rank. The unambiguity property (c) holds for some error-orthogonal models that satisfy neither (a) nor (b), as is shown in Example 6.5 below. Note that any model with nested error structure has unambiguous sums of squares for its random effects, regardless of whether it is error-orthogonal or not.

EXAMPLE 6.4. Consider the following mixed three-way main-effects-only classification model:

$$Y_{ijku} = \mu + \alpha_i + \beta_j + c_k + e_{ijku},$$

where $i = 1,2$, $j = 1,2,3$, $k = 1,2,3$ and $u = 1,\dots,n_{ijk}$ with $n_{111} = n_{112} = n_{121} = n_{122} = 1$, $n_{233} = 2$ and all other n_{ijk} equal to zero. The fixed-effect subsets of factors are $\emptyset$, $\{1\}$ and $\{2\}$ and the only random-effect subset is $\{3\}$. Write the model in matrix form as $Y = \mathbf{1}\mu + A\alpha + B\beta + Cc + e$. The model does not have maximal rank, because $\operatorname{rank}(\mathbf{1}, A, B, C) = 4$ and the maximal rank is 6. The model is r-balanced because it is Bal($\{3\}$). The model is not weakly b-balanced but nevertheless E(Y) has a UBLUE, because $\mathscr{R}(CC'\mathbf{1}) \subset \mathscr{R}(CC'A) \subset \mathscr{R}(CC'B) \subset \mathscr{R}(A) \subset \mathscr{R}(X)$. By Theorem 5.1(b) the model is error-orthogonal.

EXAMPLE 6.5. For the model in the preceding example, suppose now that the second factor is regarded as random:

$$Y_{ijku} = \mu + \alpha_i + b_j + c_k + e_{ijku}.$$

Now the fixed-effect subsets are $\emptyset$ and $\{1\}$ and the random-effect subsets are $\{2\}$ and $\{3\}$. From above we know that the model does not have maximal rank. Also, the model is not r-balanced, because it is not Bal($\{2,3\}$). So neither condition (a) nor (b) in Theorem 6.3 is satisfied, but it can be shown that the model is error-orthogonal and satisfies (c) [see VanLeeuwen, Birkes and Seely (1997)].

The notion of unambiguous sums of squares can also be applied to fixed effects. Let $\mathscr{G}$ be a fixed-effect subset of factors. Let $\mathbb{F}_1$ (respectively, $\mathbb{F}_2$, and $\mathbb{F}_3$) denote the collection of all included subsets of factors, not including $\mathscr{G}$ itself, that are contained in $\mathscr{G}$ (respectively, neither are contained in nor contain $\mathscr{G}$, and contain $\mathscr{G}$). In forming the sum of squares for $\mathscr{G}$-effects, usually one would adjust for all $\mathbb{F}_1$-effects and would not adjust for any $\mathbb{F}_3$-effects. But it is not clear which, if any, of the $\mathbb{F}_2$-effects should be adjusted for. In Theorem 6.6 below it is seen that if the model is p-balanced then there is no ambiguity.

Suppose that a mixed classification model is p-balanced. Consider the model having the same included subsets of factor and the same incidence matrix but with all effects random (except μ). If the random-effects model is proper, then it is error-orthogonal and so Theorem 6.3 is applicable. This argument yields the following theorem.

THEOREM 6.6. *Suppose that a mixed classification model is p-balanced and that for every pair of included subsets of factors $\mathscr{E}$ and $\mathscr{F}$, the intersection $\mathscr{E} \cap \mathscr{F}$ is also included. Let $\mathscr{G}$ be a fixed-effect subset, let $\mathbb{F}_1$ and $\mathbb{F}_2$ be as defined in the second paragraph above, and let $\mathbb{L}_2 \subset \mathbb{F}_2$. Then $SS(\mathscr{G} \mid \mathbb{F}_1, \mathbb{L}_2) = SS(\mathscr{G} \mid \mathbb{F}_1)$.*

7. Shortcuts for checking balance.

Verification of b-balance or weak b-balance or r-balance directly from their definitions can be tedious, and so convenient shortcut methods are developed in this section.

Throughout this section, $\mathbb{F}$ (respectively, $\mathbb{G}$ and $\mathbb{H}$) denotes the collection of all included (respectively, fixed-effect and random-effect) subsets of factors.

For short, a subset of factors (not necessarily included) will sometimes be called an *f-set*.

DEFINITIONS. Let $\mathbb{C}$ be a collection of f-sets, let $\mathscr{F} \in \mathbb{C}$, and let $\mathscr{H}$ and $\mathscr{K}$ be f-sets (not necessarily in $\mathbb{C}$).

(a) We say that $\mathscr{F}$ is *maximal* (respectively, *minimal* in $\mathbb{C}$ if it is not contained in (respectively, does not contain) any other member of $\mathbb{C}$.

(b) Let $\mathbb{C}_{\text{Max}}$ (respectively, $\mathbb{C}_{\text{Min}}$) denote the collection of f-sets that are maximal (respectively, minimal) in $\mathbb{C}$.

(c) Let $\mathbb{C}_{\text{I}}(\mathscr{H};\mathscr{K}) = \{\mathscr{E} \in \mathbb{C}\colon\ \mathscr{E} \cap \mathscr{H} = \mathscr{K}\}$ and $\mathbb{C}_{\text{Max-I}}(\mathscr{H}) = \{\mathscr{D} \in \mathbb{C}\colon\ \mathscr{D} \in [\mathbb{C}_{\text{I}}(\mathscr{H};\mathscr{D} \cap \mathscr{H})]_{\text{Max}}\}$. In other words, an f-set is in $\mathbb{C}_{\text{Max-I}}(\mathscr{H})$ if it is in $\mathbb{C}$ and is not contained in any other f-set in $\mathbb{C}$ having the same intersection with $\mathscr{H}$.

(d) Let $\mathbb{C}_{\text{U}}(\mathscr{H};\mathscr{K}) = \{\mathscr{E} \in \mathbb{C}\colon\ \mathscr{E} \cup \mathscr{H} = \mathscr{K}\}$ and $\mathbb{C}_{\text{Min-U}}(\mathscr{H}) = \{\mathscr{D} \in \mathbb{C}\colon\ \mathscr{D} \in [\mathbb{C}_0(\mathscr{H};\mathscr{D} \cup \mathscr{H})]_{\text{Min}}\}$. In other words, an f-set is in $\mathbb{C}_{\text{Min-U}}(\mathscr{H})$ if it is in $\mathbb{C}$ and does not contain any other f-set in $\mathbb{C}$ having the same union with $\mathscr{H}$.

(e) Let $\mathbb{XN}(\mathscr{H}) = [\mathbb{F}_{\text{Min-U}}(\mathscr{H})]_{\text{Max-I}}(\mathscr{H})$. Thus, an f-set is in $\mathbb{XN}(\mathscr{H})$ if it is in $\mathbb{F}_{\text{Min-U}}(\mathscr{H})$ and is not contained in any other f-set in $\mathbb{F}_{\text{Min-U}}(\mathscr{H})$ having the same intersection with $\mathscr{H}$.

(f) Let $\mathbb{NX}(\mathscr{H}) = [\mathbb{F}_{\text{Max-I}}(\mathscr{H})]_{\text{Min-U}}(\mathscr{H})$. Thus, an f-set is in $\mathbb{NX}(\mathscr{H})$ if it is in $\mathbb{F}_{\text{Max-I}}(\mathscr{H})$ and does not contain any other f-set in $\mathbb{F}_{\text{Max-I}}(\mathscr{H})$ having the same union with $\mathscr{H}$.

LEMMA 7.1. *Let $\mathbb{C}$ be a collection of* f-*sets and let $\mathscr{F} \in \mathbb{C}$. There exist* f-*sets $\mathscr{N}$ and $\mathscr{M}$ such that $\mathscr{N}$ is minimal in $\mathbb{C}$, $\mathscr{M}$ is maximal in $\mathbb{C}$ and $\mathscr{N} \subset \mathscr{F} \subset \mathscr{M}$.*

LEMMA 7.2. *Consider a mixed classification model.*

(a) *The model is* b-*balanced if and only if it is* $\text{Bal}(\mathscr{H} \mid \mathscr{G})$ *for all* $\mathscr{H} \in \mathbb{H}_{\text{Max}}$ *and all* $\mathscr{G} \in \mathbb{G}$.

(b) *If the model is* $\text{Bal}(\mathscr{H} \mid \mathscr{G})$ *for all* $\mathscr{H} \in \mathbb{H}_{\text{Max}}$ *and all* $\mathscr{G} \in \mathbb{G}_{\text{Max}}$, *then it is weakly* b-*balanced*.

(c) *The model is weakly* b-*balanced if and only if, for each* $\mathscr{H} \in \mathbb{H}_{\text{Max}}$ *and* $\mathscr{G} \in \mathbb{G}_{\text{Max}}$, *there exists an* f-*set* $\mathscr{F}$ *such that* $\mathscr{G} \subset \mathscr{F}$, $\mathscr{H} \cap \mathscr{F}$ *is contained in a member of* $\mathbb{G}$, *and the model is* $\text{Bal}(\mathscr{H} \mid \mathscr{F})$.

(d) *The model is* r-*balanced if and only if it is* $\text{Bal}(\mathscr{H}_1 \cup \mathscr{H}_2)$ *for all* $\mathscr{H}_1, \mathscr{H}_2 \in \mathbb{H}_{\text{Max}}$.

(e) *The model is* p-*balanced if and only if it is* $\text{Bal}(\mathscr{F}_1 \cup \mathscr{F}_2)$ *for all* $\mathscr{F}_1, \mathscr{F}_2 \in \mathbb{F}_{\text{Max}}$.

The following two examples illustrate the application of Lemma 7.2.

EXAMPLE 7.3. Consider a three-way mixed model with one two-factor interaction:

$$Y_{ijku} = \mu + \alpha_i + b_j + (\alpha b)_{ij} + \gamma_k + e_{ijku},$$

where $i = 1, \dots, t_1$, $j = 1, \dots, t_2$, $k = 1, \dots, t_3$, $u = 1, \dots, n_{ijk}$. This is the same model as in Example 5.4 except that here t_1, t_2, t_3 can be any positive integers. The collection of fixed-effect subsets is $\mathbb{G} = \{\emptyset, \{1\}, \{3\}\}$ and the collection of random-effect subsets is $\mathbb{H} = \{\{2\}, \{1,2\}\}$. The collections of maximal fixed-effect and random-effect subsets are $\mathbb{G}_{\text{Max}} = \{\{1\}, \{3\}\}$ and $\mathbb{H}_{\text{Max}} = \{\{1,2\}\}$. Suppose the design is $\text{Bal}(\{2\} \mid \{1,3\})$ and $\text{Bal}(\{1,2\})$; that is, n_{ijk} depends only on i and k and $n_{ij\cdot}$ is the same for all i and j. These properties, which hold for the particular design given in Example 5.4, can be shown to imply that the model is weakly b-balanced and r-balanced, and hence error-orthogonal. To show weak b-balance, apply Lemma 7.2(c) with $\mathscr{F} = \{1,3\}$. Note that both members of $\mathbb{G}_{\text{Max}}$ are contained in $\mathscr{F}$, that $\{1,2\} \cap \{1,3\} = \{1\} \in \mathbb{G}$, and that the model is $\text{Bal}(\{1,2\} \mid \{1,3\})$, which is equivalent to $\text{Bal}(\{2\} \mid \{1,3\})$. By Lemma 7.2(d), the design is r-balanced if and only if it is $\text{Bal}(\{1,2\})$.

EXAMPLE 7.4. Consider a random-effects classification model including all two-factor interaction effects and no higher-order effects. By Lemmas 4.3(c) and 7.2(e), the model is b & r-balanced if and only if its incidence matrix is balanced with respect to every subset of four factors. For example, if the model has five factors, then it is b & r-balanced if and only if $n[i_1 i_2 i_3 i_4]$, $n[i_1 i_2 i_3 i_5]$, $n[i_1 i_2 i_4 i_5]$, $n[i_1 i_3 i_4 i_5]$ and $n[i_2 i_3 i_4 i_5]$ are constant (possibly five different constants). To illustrate, suppose all five factors have two levels. Let a and b be two nonnegative integers and set $n[i_1 i_2 i_3 i_4 i_5] = a$ or b according as $i_1 + i_2 + i_3 + i_4 + i_5$ is even or odd. Then all five marginal four-factor incidence matrices have all entries equal to $a + b$. For $a = 0$ and $b = 1$, this is a half-fractional factorial design.

As the preceding examples demonstrate, Lemma 7.2 typically reduces the number of conditions that need to be checked in order to establish that a model is b-, weakly b-, r- or p-balanced. For b & r-balance, often the number of conditions can be reduced even further by the following lemma.

LEMMA 7.5. *Consider a mixed classification model. Recall the notation introduced at the beginning of this section. The following statements are equivalent*:

(a) *The model is b & r-balanced.*
(b) *The model is* $\text{Bal}(\mathscr{H} \mid \mathscr{F})$ *for all* $\mathscr{H} \in \mathbb{H}_{\text{Max}}$ *and all* $\mathscr{F} \in \mathbb{F}$.
(c) *The model is* $\text{Bal}(\mathscr{H} \mid \mathscr{F})$ *for all* $\mathscr{H} \in \mathbb{H}_{\text{Max}}$ *and all* $\mathscr{F} \in \mathbb{XN}(\mathscr{H})$.
(d) *The model is* $\text{Bal}(\mathscr{H} \mid \mathscr{F})$ *for all* $\mathscr{H} \in \mathbb{H}_{\text{Max}}$ *and all* $\mathscr{F} \in \mathbb{NX}(\mathscr{H})$.

EXAMPLE 7.6. Consider the following model for a split-plot experiment with subsampling:

$$Y_{ijku} = \mu + \alpha_i + b_j + (\alpha b)_{ij} + \gamma_k + (\alpha\gamma)_{ik} + e_{ijku},$$

which can be obtained from the model in Example 7.3 by adding the $(\alpha\gamma)$ interaction term. The included subsets of factors are $\emptyset$, $\{1\}, \{2\}, \{1,2\}, \{3\}, \{1,3\}$. The random-effect subsets are $\{2\}$, $\{1,2\}$ and so $\mathbb{H}_{\text{Max}} = \{\{1,2\}\}$. To check the model for b & r-balance, it is convenient to find either $\mathbb{XN}(\{1,2\})$ or $\mathbb{NX}(\{1,2\})$

and apply Lemma 7.5. Union of $\{1,2\}$ with the six included subsets yields $\{1,2\},\{1,2\},\{1,2\},\{1,2\},\{1,2,3\},\{1,2,3\}$. Among the four included subsets whose union with $\{1,2\}$ is $\{1,2\}$, the only minimal f-set is $\varnothing$. Among the two included subsets whose union with $\{1,2\}$ is $\{1,2,3\}$, the only minimal f-set is $\{3\}$. Thus $\mathbb{F}_{\text{Min-U}}(\{1,2\}) = \{\varnothing,\{3\}\}$. Intersection of $\{1,2\}$ with $\varnothing$, $\{3\}$ yields $\varnothing$, $\varnothing$. Among the two members of $\mathbb{F}_{\text{Min-U}}(\{1,2\})$, both of whose intersection with $\{1,2\}$ is $\varnothing$, the only maximal f-set is $\{3\}$. Thus $\mathbb{XN}(\{1,2\}) = \{\{3\}\}$. By Lemma 7.5(a $\Leftrightarrow$ c), the model is b & r-balanced if and only if it is $\text{Bal}(\{1,2\} \mid \{3\})$.

In the preceding example we could have used condition (d) of Lemma 7.5 rather than condition (c), but less work is involved in applying condition (c). It has been our experience that this is typically the case.

EXAMPLE 7.7. Consider the model obtained from the preceding split-plot model by dropping the γ_k term. The two models are merely reparametrizations of one another in the sense that they specify the same set of possible mean vectors and the same set of possible covariance matrices. This is because the range of the single-effect matrix corresponding to γ-effects is contained in the range of the single-effect matrix corresponding to $(\alpha\gamma)$-effects. As before, $\mathbb{H}_{\text{Max}} = \{\{1,2\}\}$. It can be shown that $\mathbb{XN}(\{1,2\}) = \mathbb{NX}(\{1,2\}) = \{\varnothing,\{1,3\}\}$. Therefore, this model is b & r-balanced if and only if the design satisfies the condition (C_2): $\text{Bal}(\{1,2\})$ and $\text{Bal}(\{1,2\} \mid \{1,3\})$. Recall that the condition for b & r-balance in Example 7.6 was (C_1): $\text{Bal}(\{1,2\} \mid \{3\})$. Lemma 4.1(d),(i) implies that $(C_1) \Rightarrow (C_2)$, but the design in Example 5.4 satisfies (C_2) and not (C_1). Thus we see that b & r-balance is not invariant under reparametrization of the model.

Error-orthogonality is invariant under reparametrization of the model. By Lemma 2.3, the property that $E(Y)$ has a UBLUE depends only on $\mathscr{R}(X)$, which coincides with the set of possible mean vectors of Y, and on the set $\mathscr{V}$ of possible covariance matrices of Y. By Lemma 2.2, the property that $\text{Cov}(MY)$ has OBS depends only on M, which depends only on $\mathscr{R}(X)$, and on $\mathscr{V}$. Examples 7.6 and 7.7 show that in determining whether or not a particular model is error-orthogonal, it can sometimes be advantageous to consider a different parametrization of the mean vector.

8. Proofs.

PROOF OF LEMMA 4.1(b), (j). For any subset of factors, say $\mathscr{C}$ (or $\mathscr{C}^*$ or $\mathscr{C}^\circ$), let w_c (or w_c^* or w_c°) denote an arbitrary combination of levels of the factors. Let $\mathscr{F}^* = \mathscr{F} \setminus \mathscr{E}$, $\mathscr{H}^* = \mathscr{H} \setminus \mathscr{G}$, $\mathscr{D} = (\mathscr{F} \cup \mathscr{E}) \setminus (\mathscr{H} \cup \mathscr{G})$, $\mathscr{G}^* = \mathscr{G} \cap \mathscr{F}^*$, $\mathscr{G}^\circ = \mathscr{G} \cap \mathscr{E}$, $\mathscr{D}^* = \mathscr{D} \cap \mathscr{F}^*$ and $\mathscr{D}^\circ = \mathscr{D} \cap \mathscr{E}$. The $\text{Bal}(\mathscr{F} \mid \mathscr{E})$ property says that the number $n[w_f^* w_e]$ depends only on w_e; that is, for every given w_e, $n[w_f^* w_e]$ is the same, say $m[w_e]$, for all w_f^*. Let the symbol $\dot\cup$ denote a disjoint union. The assumptions in (b) imply that $\mathscr{H} \cup \mathscr{G} = \mathscr{H}^* \,\dot\cup\, \mathscr{G} \subset \mathscr{F} \cup \mathscr{E}$. Then $\mathscr{F}^* \,\dot\cup\, \mathscr{E} = \mathscr{F} \cup \mathscr{E} = \mathscr{H}^* \,\dot\cup\, \mathscr{G} \,\dot\cup\, \mathscr{D} = \mathscr{H}^* \,\dot\cup\, \mathscr{G}^* \,\dot\cup\, \mathscr{G}^\circ \,\dot\cup\, \mathscr{D}^* \,\dot\cup\, \mathscr{D}^\circ$. Note that $\mathscr{E} = \mathscr{G}^\circ \,\dot\cup\, \mathscr{D}^\circ$. The $\text{Bal}(\mathscr{F} \mid \mathscr{E})$ property says that $n[w_h^* w_g^* w_g^\circ w_d^* w_d^\circ] = m[w_g^\circ w_d^\circ]$. Now $n[w_h^* w_g]$

$= n[w_h^* w_g^* w_g^\circ] = \sum_{w_d^*} \sum_{w_d^\circ} n[w_h^* w_g^* w_g^\circ w_d^* w_d^\circ] = \sum_{w_d^*} \sum_{w_d^\circ} m[w_g^\circ w_d^\circ]$, which depends only on w_g° and hence only on w_g. This verifies property Bal($\mathscr{H} \mid \mathscr{G}$) and so (b) is proved.

The implication $\Rightarrow$ in (j) follows from (e) and (c). For the reverse implication, assume Bal($\mathscr{G}$) and Bal($\mathscr{H} \mid \mathscr{G}$). These properties say that $n[w_g] = m$ for all w_g and $n[w_h^* w_g] = m[w_g]$ for all w_h^*. Now $m = n[w_g] = \sum_{w_h^*} n[w_h^* w_g] = \sum_{w_h^*} m[w_g] = t_h^* m[w_g]$, where t_h^* is the product of the numbers of levels of the factors in $\mathscr{H}^*$. Therefore $n[w_h^* w_g] = m[w_g] = m/t_h^*$ for all $w_h^* w_g$, hence Bal($\mathscr{H} \cup \mathscr{G}$), since $\mathscr{H} \cup \mathscr{G} = \mathscr{H}^* \dot{\cup} \mathscr{G}$. □

PROOF OF LEMMA 4.2. For any subset of factors, say $\mathscr{F}$ (or $\mathscr{F}^*$), let w_f (or w_f^*) denote an arbitrary combination of levels of the factors. Each column of F corresponds to a combination w_f and the number of 1's in the column is $n[w_f]$. Bal($\mathscr{F}$) means that $n[w_f] = m$ for all w_f, which implies $F'F = mI$ and $P_F = F(F'F)^{-1}F' = m^{-1}FF'$, hence (b).

The proofs of (a) and (c) require that we carefully index the entries of the three single-effect matrices F, G and L. By ordering the factors suitably, we can write $w_f = w_f^* w_l$ and $w_g = w_l w_g^*$, where $\mathscr{F}^* = \mathscr{F} \setminus \mathscr{G}$ and $\mathscr{G}^* = \mathscr{G} \setminus \mathscr{F}$. A combination of levels of all the factors can be written as $w = w_f^* w_l w_g^* w_h$, where $\mathscr{H} = \{1, \dots, p\} \setminus (\mathscr{F} \cup \mathscr{G})$. The columns of F can be indexed by $w_f = w_f^* w_l$, the columns of G can be indexed by $w_g = w_l w_g^*$ and the columns of L can be indexed by w_l. The rows of F, G and L correspond to observations. Suppose that row i corresponds to an observation taken at levels $z = z_f^* z_l z_g^* z_h$ of the factors. The (i, w_l)-entry of L is 1 if $z_l = w_l$ and is 0 if not. The $(i, w_l w_g^*)$-entry of G is 1 if $z_l z_g^* = w_l w_g^*$ and is 0 if not. For (a) we assume $\mathscr{F} = \mathscr{L}$. From the description of the entries of L and G, we see that the w_l-column of L is the sum of the $w_l w_g^*$-columns of G as w_g^* varies. Thus $\mathscr{R}(F) = \mathscr{R}(L) \subset \mathscr{R}(G)$.

For (c) we assume Bal($\mathscr{F} \mid \mathscr{G}$), that is, $n[z_f^* w_l w_g^*] = m[w_l w_g^*]$. The $(i, w_l w_g^*)$-entry of $FF'G$ is the same as the $(z_f^* z_l, w_l w_g^*)$-entry of $F'G$, which is the dot product of the $z_f^* z_l$-column of F and the $w_l w_g^*$-column of G. The dot product is $n[z_f^* w_l w_g^*] = m[w_l w_g^*]$ if $z_l = w_l$ and is 0 if not. Thus the $w_l w_g^*$-column of $FF'G$ is equal to the w_l-column of L multiplied by $m[w_l w_g^*]$, hence $\mathscr{R}(FF'G) \subset \mathscr{R}(L)$. Conversely, the w_l-column of L is equal to the $w_l w_g^*$-column of $FF'G$ multiplied by $1/m[w_l w_g^*]$, provided we can find w_g^* such that $m[w_l w_g^*] \neq 0$. If no such w_g^* exists, then $n[w_l] = \sum_{w_f^*} \sum_{w_g^*} n[w_f^* w_l w_g^*] = \sum_{w_f^*} \sum_{w_g^*} m[w_l w_g^*] = 0$, and so the w_l-column of L is a column of 0's.

For (d) we assume Bal($\mathscr{F}$) and Bal($\mathscr{F} \mid \mathscr{G}$). Using parts (b), (c) and (a), we obtain $\mathscr{R}(P_F P_G) = \mathscr{R}(P_F G) = \mathscr{R}(FF'G) = \mathscr{R}(L) \subset \mathscr{R}(G)$. Now apply Lemma 8.1 below. □

LEMMA 8.1. *Let P and Q be two projection matrices. The following statements are equivalent*: (*a*) *P and Q commute*, (*b*) *PQ is symmetric*, (*c*) *$\mathscr{R}(PQ) \subset \mathscr{R}(Q)$*, (*d*) *$PQ$ is a projection matrix*, (*e*) *$P + Q - PQ$ is a projection* *matrix. If any, hence all, of the statements hold, then* $\mathscr{R}(PQ) = \mathscr{R}(P) \cap \mathscr{R}(Q)$ *and* $\mathscr{R}(P + Q - PQ) = \mathscr{R}(P) + \mathscr{R}(Q)$.

The proof of this lemma is left to the reader.

PROOF OF THEOREM 5.1. Lemma 2.3 implies that $\mathrm{E}(Y)$ has a UBLUE if and only if $\mathscr{R}(HH'G) \subset \mathscr{R}(X)$ for every pair of single-effect matrices G and H corresponding, respectively, to a fixed-effect subset $\mathscr{G}$ and a random-effect subset $\mathscr{H}$. Under the assumption of weak b-balance, given such subsets $\mathscr{G}$ and $\mathscr{H}$, there exist subsets $\mathscr{F}$ and $\mathscr{J}$ such that $\mathscr{G} \subset \mathscr{F}, \mathscr{J}$ is a fixed-effect subset, $\mathscr{L} \subset \mathscr{J}$ where $\mathscr{L} = \mathscr{H} \cap \mathscr{F}$, and the design is $\mathrm{Bal}(\mathscr{H} \mid \mathscr{F})$. Let F, L and J be the single-effect matrices corresponding to $\mathscr{F}$, $\mathscr{L}$ and $\mathscr{J}$, respectively. Lemma 4.2(a),(c) implies $\mathscr{R}(G) \subset \mathscr{R}(F)$, $\mathscr{R}(L) \subset \mathscr{R}(J)$ and $\mathscr{R}(HH'F) = \mathscr{R}(L)$. Now $\mathscr{R}(HH'G) \subset \mathscr{R}(HH'F) = \mathscr{R}(L) \subset \mathscr{R}(J) \subset \mathscr{R}(X)$.

To prove (b), it suffices to show that $\mathrm{Cov}(MY)$ has OBS, which will follow from Lemma 2.4. Recall that $V_j = H_j H_j'$ for $j = 1, \ldots, r$ and $\mathrm{sp}\,\mathscr{V} = \mathrm{sp}\{V_1, \ldots, V_r, I\}$. We will show that $V_i V_j = W + Z$ where $W \in \mathrm{sp}\,\mathscr{V}$ and $\mathscr{R}(Z) \subset \mathscr{R}(X)$. To reduce the use of subscripts, let $F = H_i$ and $H = H_j$. Let $\mathscr{F}$ and $\mathscr{H}$ be the subsets of factors corresponding to F and H. Let $\mathscr{L} = \mathscr{F} \cap \mathscr{H}$ and let L be the single-effect matrix corresponding to $\mathscr{L}$. The assumption of r-balance says that the design is $\mathrm{Bal}(\mathscr{F} \cup \mathscr{H})$. By Lemma 4.1(c),(e), the design is also $\mathrm{Bal}(\mathscr{F})$, $\mathrm{Bal}(\mathscr{H})$, $\mathrm{Bal}(\mathscr{L})$ and $\mathrm{Bal}(\mathscr{F} \mid \mathscr{H})$. Therefore, by Lemma 4.2(b),(d), $V_i V_j = FF'HH' = m_F m_H P_F P_H = m_F m_H P_L = cLL'$ where $c = m_F m_H / m_L$. Since the model is proper, either (1) $\mathscr{L}$ is a random-effect subset or (2) $\mathscr{L}$ is contained in a fixed-effect subset, say $\mathscr{G}$. In case (1), $LL' \in \mathscr{V}$, so let $W = cLL'$ and $Z = 0$. In case (2), $\mathscr{R}(L) \subset \mathscr{R}(G) \subset \mathscr{R}(X)$, so let $W = 0$ and $Z = cLL'$. □

PROOF OF LEMMA 6.1. By Lemma 2.3, error-orthogonality implies $MA = AM$, and so $\mathscr{R}(MP_A) = \mathscr{R}(MA) \subset \mathscr{R}(P_A)$. Now apply Lemma 8.1 to obtain (a). It is a general fact that $P_{X,A} = P_X + P_{MA}$. By using (a) we see that $P_{X,A} P_{X,B} = (P_X + MP_A)(P_X + MP_B) = P_X + P_A P_B M$, hence (b). To show (c) it suffices to show P_{MA} and P_{MB} commute. Note that $2MA = MAM \in \mathrm{sp}\, M\mathscr{V}M$. By Lemma 2.2, $\mathrm{sp}\, M\mathscr{V}M$ is a commutative quadratic subspace, and so, by (2.b) in Seely (1971), $P_{MA} \in \mathrm{sp}\, M\mathscr{V}M$. □

PROOF OF LEMMA 6.2. Assume (a). Let $T = (X, H_1, \ldots, H_r)$ be the all-included-effects model matrix. Note that statement (c) can be written as $\mathscr{R}(TC_1) \cap \mathscr{R}(TC_2) \subset \mathscr{R}(TC_3)$ for suitable C_1, C_2 and C_3. Let S be what the all-included-effects model matrix would be if there were exactly one observation for each cell. Each row of S corresponds to a cell and we can write $T = KS$ where the matrix K replicates each row according to the number of observations in the corresponding cell. Assumption (a) says that T has maximal rank, which implies that K is a one-to-one transformation from $\mathscr{R}(S)$ onto $\mathscr{R}(T)$. Hence, $\mathscr{R}(KSC_1) \cap \mathscr{R}(KSC_2) \subset \mathscr{R}(KSC_3)$ if and only if $\mathscr{R}(SC_1) \cap \mathscr{R}(SC_2) \subset \mathscr{R}(SC_3)$. Therefore, to prove (c) we can assume a completely balanced design. Since a completely balanced design satisfies (b), it suffices to prove that (b) ⇒ (c).

Assume (b). Write $W = (F_1, \dots, F_k)$ where each F_i is a single-effect matrix corresponding to a random-effect subset $\mathscr{F}_i$ which neither is contained in nor contains $\mathscr{H}$. Let $A = HH' \in \mathscr{V}$ and $B = WW' = F_1F_1' + \cdots + F_kF_k' \in \text{sp}\,\mathscr{V}$. By Theorem 5.1(b) and Lemma 6.1(c), $P_{X,A}$ and $P_{X,B}$ commute. By Lemmas 8.1 and 6.1(b), $\mathscr{R}(X,H) \cap \mathscr{R}(X,W) = \mathscr{R}(P_{X,A}) \cap \mathscr{R}(P_{X,B}) = \mathscr{R}(P_{X,A}P_{X,B}) \subset \mathscr{R}(X) + \mathscr{R}(P_AP_B)$. Next note $\mathscr{R}(P_AP_B) = \mathscr{R}(P_HB) \subset \Sigma\mathscr{R}(P_HF_i) = \Sigma\mathscr{R}(P_HP_{F_i})$. It suffices to show $\mathscr{R}(P_HP_{F_i}) \subset \mathscr{R}(X,U)$. By r-balance and Lemmas 4.2(d) and 4.1(c),(e), $P_HP_{F_i} = P_{L_i}$ where L_i is the single-effect matrix corresponding to $\mathscr{H} \cap \mathscr{F}_i$. Since the model is proper, either $\mathscr{R}(L_i) \subset \mathscr{R}(U)$ or $\mathscr{R}(L_i) \subset \mathscr{R}(X)$. □

PROOF OF THEOREM 6.3. Let H, U and W be as defined in Lemma 6.2. We must show that $P_{X,H,U,W} - P_{X,U,W} = P_{X,H,U} - P_{X,U}$. Let $P = P_{X,H,U}$ and $Q = P_{X,U,W}$. Note that $P = P_{X,A}$ and $Q = P_{X,B}$ where $A = HH' \in \mathscr{V}$ and $B = UU' + WW' \in \text{sp}\,\mathscr{V}$, and so by Lemma 6.1(c), P and Q commute. By Lemma 8.1, $P + Q - PQ = P_{P,Q} = P_{X,H,U,W}$. It remains to show $PQ = P_{X,U}$. By Lemma 8.1, PQ is the projection matrix whose range is $\mathscr{R}(P) \cap \mathscr{R}(Q) = \mathscr{R}(X,H) \cap [\mathscr{R}(X,U) + \mathscr{R}(X,W)] = \mathscr{R}(X,U) + [\mathscr{R}(X,H) \cap \mathscr{R}(X,W)]$. By Lemma 6.2, this is simply $\mathscr{R}(X,U)$. □

PROOF OF LEMMA 7.2. For (a), (d) and (e), apply Lemmas 7.1 and 4.1(c),(f). Part (b) follows from (c) with $\mathscr{F} = \mathscr{G}$. The "only if" half of (c) is immediate. To prove the "if" half, take any $\mathscr{H} \in \mathbb{H}$ and $\mathscr{G} \in \mathbb{G}$. By Lemma 7.1 there exist $\mathscr{H}^* \in \mathbb{H}_{\text{Max}}$ and $\mathscr{G}^* \in \mathbb{G}_{\text{Max}}$ such that $\mathscr{H} \subset \mathscr{H}^*$ and $\mathscr{G} \subset \mathscr{G}^*$. There exist an f-set $\mathscr{F}$ and $\mathscr{J} \in \mathbb{G}$ such that $\mathscr{G}^* \subset \mathscr{F}$, $\mathscr{H}^* \cap \mathscr{F} \subset \mathscr{J}$, and the model is $\text{Bal}(\mathscr{H}^* \mid \mathscr{F})$. Now, $\mathscr{G} \subset \mathscr{G}^* \subset \mathscr{F}$, $\mathscr{H} \cap \mathscr{F} \subset \mathscr{H}^* \cap \mathscr{F} \subset \mathscr{J}$ and, by Lemma 4.1(f), the model is $\text{Bal}(\mathscr{H} \mid \mathscr{F})$. □

PROOF OF LEMMA 7.5. Assume (a). This implies $\text{Bal}(\mathscr{H} \mid \mathscr{G})$ for all $\mathscr{H} \in \mathbb{H}$ and all $\mathscr{G} \in \mathbb{G}$ and $\text{Bal}(\mathscr{H} \cup \mathscr{E})$ for all $\mathscr{H}, \mathscr{E} \in \mathbb{H}$. Lemma 4.1(e) implies (b). Now assume (b). Lemma 7.2(a) implies b-balance. To show r-balance, take any $\mathscr{H}, \mathscr{E} \in \mathbb{H}$. By Lemma 7.1 there exist $\tilde{\mathscr{H}}, \tilde{\mathscr{E}} \in \mathbb{H}_{\text{Max}}$ such that $\mathscr{H} \subset \tilde{\mathscr{H}}$ and $\mathscr{E} \subset \tilde{\mathscr{E}}$. Statement (b) implies $\text{Bal}(\tilde{\mathscr{H}} \mid \tilde{\mathscr{E}})$ and $\text{Bal}(\tilde{\mathscr{E}} \mid \phi)$, that is, $\text{Bal}(\tilde{\mathscr{E}})$, and hence, by Lemma 4.1(j),(c), $\text{Bal}(\tilde{\mathscr{H}} \cup \tilde{\mathscr{E}})$ and $\text{Bal}(\mathscr{H} \cup \mathscr{E})$. Thus (b) $\Rightarrow$ (a). Clearly (b) $\Rightarrow$ (c) and (b) $\Rightarrow$ (d). Now assume (c). Given $\mathscr{H} \in \mathbb{H}_{\text{Max}}$ and $\mathscr{F} \in \mathbb{F}$, we must show $\text{Bal}(\mathscr{H} \mid \mathscr{F})$. Consider the collection $\mathbb{D} = \{\mathscr{D} \in \mathbb{F}\colon \mathscr{H} \cup \mathscr{D} = \mathscr{H} \cup \mathscr{F}\}$. By Lemma 7.1, there exists $\mathscr{N} \in \mathbb{D}_{\text{Min}}$ such that $\mathscr{N} \subset \mathscr{F}$. Note that $\mathscr{N} \in \mathbb{F}_{\text{Min-U}}(\mathscr{H})$. Lemma 7.1 implies the existence of $\mathscr{M} \in [\mathbb{F}_{\text{Min-U}}(\mathscr{H})]_{\text{Max-I}}(\mathscr{H}) = \mathbb{XN}(\mathscr{H})$ such that $\mathscr{N} \subset \mathscr{M}$. By Lemma 4.1(h),(i), $\text{Bal}(\mathscr{H} \mid \mathscr{M}) \Rightarrow \text{Bal}(\mathscr{H} \mid \mathscr{N}) \Rightarrow \text{Bal}(\mathscr{H} \mid \mathscr{F})$. Thus (c) $\Rightarrow$ (b). The proof of (d) $\Rightarrow$ (b) is similar. □

Acknowledgment. We thank the referees for comments that improved the presentation of this paper.

REFERENCES

BROWN, K. G. (1983). Sub-balanced data and the mixed analysis of variance. *J. Amer. Statist. Assoc.* **78** 162–167.

EL-BASSIOUNI, M. Y. (1983). On the existence of explicit restricted maximum likelihood estimators in multivariate normal models. *Sankhyā Ser. B* **42** 1069–1085.

EL-BASSIOUNI, M. Y. and SEELY, J. (1980). Optimal tests for certain functions of the parameters in a covariance matrix with linear structure. *Sankhyā Ser. A* **42** 64–77.

HOUTMAN, A. M. and SPEED, T. P. (1983). Balance in designed experiments with orthogonal block structure. *Ann. Math. Statist.* **42** 1069–1085.

NELDER, J. A. (1965). The analysis of randomized experiments with orthogonal block structure I. Block structure and the null analysis of variance. *Proc. Roy. Soc. London Ser. A* **273** 147–162.

PUKELSHEIM, F. (1981). On the existence of unbiased nonnegative estimates of variance covariance components. *Ann. Statist.* **9** 293–299.

PUKELSHEIM, F. (1993). *Optimal Design of Experiments*. Wiley, New York.

RAO, C. R. and KLEFFE, J. (1988). *Estimation of Variance Components and Applications*. North-Holland, New York.

SEELY, J. (1971). Quadratic subspaces and completeness. *Ann. Math. Statist.* **42** 710–721.

SEELY, J. (1972). Completeness for a family of multivariate normal distributions. *Ann. Math. Statist.* **43** 1644–1647.

SEELY, J. (1977). Minimal sufficient statistics and completeness for multivariate normal families. *Sankhyā Ser. A* **39** 170–185.

SEIFERT, B. (1979). Optimal testing for fixed effects in general balanced mixed classification models. *Math. Operationsforsch. Statist. Ser. Statist.* **10** 237–255.

SZATROWSKI, T. H. (1980). Necessary and sufficient conditions for explicit solutions in the multivariate normal estimation problem for patterned means and covariances. *Ann. Statist.* **8** 802–810.

VANLEEUWEN, D. M., BIRKES, D. S. and SEELY, J. F. (1997). Balance and orthogonality in designs for mixed classification models. Technical Report 168, Dept. Statistics, Oregon State Univ.

VANLEEUWEN, D. M., SEELY, J. F. and BIRKES, D. S. (1998). Sufficient conditions for orthogonal designs in mixed linear models. *J. Statist. Plann. Inference* **73** 373–389.

AGRICULTURAL BIOMETRIC SERVICE
BOX 30003, MSC 3130
NEW MEXICO STATE UNIVERSITY
LAS CRUCES, NEW MEXICO 88003

DEPARTMENT OF STATISTICS
44 KIDDER HALL
OREGON STATE UNIVERSITY
CORVALLIS, OREGON 97331