PSYCHOMETRIKA—VOL. 61, NO. 1, 133–154
MARCH 1996

# UNIQUENESS PROOF FOR A FAMILY OF MODELS SHARING FEATURES OF TUCKER'S THREE-MODE FACTOR ANALYSIS AND PARAFAC/CANDECOMP

RICHARD A. HARSHMAN AND MARGARET E. LUNDY

UNIVERSITY OF WESTERN ONTARIO

Some existing three-way factor analysis and MDS models incorporate Cattell's "Principle of Parallel Proportional Profiles". These models can—with appropriate data—empirically determine a unique best fitting axis orientation without the need for a separate factor rotation stage, but they have not been general enough to deal with what Tucker has called "interactions" among dimensions. This article presents a proof of unique axis orientation for a considerably more general parallel profiles model which incorporates interacting dimensions. The model, $\mathbf{X}_k = \mathbf{A}\ {}^{A}\mathbf{D}_k\ \mathbf{H}\ {}^{B}\mathbf{D}_k\ \mathbf{B}'$, does not assume symmetry in the data or in the interactions among factors. A second proof is presented for the symmetrically weighted case (i.e., where ${}^{A}\mathbf{D}_k = {}^{B}\mathbf{D}_k$). The generality of these models allows one to impose successive restrictions to obtain several useful special cases, including PARAFAC2 and three-way DEDICOM.

Key words: Parallel proportional profiles, intrinsic axes, DEDICOM, PARAFAC2, Cattell, trilinear models, quadrilinear models, factor rotation problem, multidimensional scaling, principal components, oblique confactor.

The difference between independent and "interacting" dimensions in analysis of three-way data is illustrated by the following two familiar models,

$$x_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} + e_{ijk}, \text{ and}$$

$$x_{ijk} = \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{t=1}^{T} a_{ir} b_{js} c_{kt} g_{rst} + e_{ijk}.$$

The first is the CANDECOMP model, which forms the basis for the INDSCAL multidimensional scaling procedure (Carroll & Chang, 1970). It is also the PARAFAC model that is the basis of a three-way factor analysis procedure (Harshman, 1970; Harshman & Lundy, 1984, 1994). The second is Tucker's Three-Mode Factor Analysis model (Tucker, 1963, 1964, 1966), the basis for Kroonenberg and de Leeuw's (1980) TUCKALS programs. In both models, a, b and c are factor weights associated with the three ways of the data (which we call Modes A, B and C, respectively). In the first model, factor contributions to each data point are independent: the weights for

We want to express appreciation for the contributions of several colleagues: Jos M. F. ten Berge and Henk A. L. Kiers carefully went through more than one version of this article, found an important error, and contributed many improvements. J. Douglas Carroll and Shizuhiko Nishisato acted with unusual editorial preserverance and flexibility, thereby making possible the successful completion of a difficult assessment and revision process. Joseph B. Kruskal has long provided crucial mathematical insights and inspiration to those working in this area, but this is particularly true for us. His contributions to this specific article include early discussion of basic questions and careful examination of some earlier attempted proofs, finding them to be invalid. The anonymous reviewers also made useful suggestions. Some portions of this work were supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

Requests for reprints should be sent to Richard A. Harshman, Psychology Department, University of Western Ontario, London, Ontario, CANADA N6A 5C2.

0033-3123/96/0300-92066$00.75/0

factor r in Mode A are multiplied only by the weights for factor r in Modes B and C, with the summation over all R factors. In the second, however, the factors interact in their contributions to the data: the summation is over all possible combinations of dimensions (with g weighting the contributions of each combination). This gives the second model greater generality and a great deal of "rotational freedom", which can make interpretation of results more challenging but allows the study of complex three-way variation that cannot be fit by simpler models such as PARAFAC.

*Uniqueness.* A singular characteristic of the first model is its "intrinsic axis property". That is, a best-fitting solution not only determines a configuration of points in a reduced dimensional space, it also defines a unique best-fitting orientation of axes or dimensions in that space; axis orientation is thus an "intrinsic" part of the model. This uniqueness of axis orientation can offer an empirical solution to the "rotation problem" of factor analysis and MDS, when the data are appropriate. Uniqueness is obtained by using the third mode of the data to seek "Parallel Proportional Profiles" of factor loadings or scores (Cattell, 1944; Cattell & Cattell, 1955). The potential importance and empirical validity of such dimensions has been argued elsewhere by Cattell (e.g., 1944, 1978) and by the present authors (e.g., Harshman, 1970; Harshman & Lundy, 1984, pp. 147–151, 163–168). Carroll and Chang (1970) also pointed out the uniqueness property of the trilinear model and the important advantage of INDSCAL over two-way MDS when appropriate data are available.

Mathematical proof that Parallel Profiles can determine a unique axis orientation was first provided by Cattell and Cattell (1955). Jennrich (in Harshman, 1970) devised the first proof of uniqueness for a Parallel Profiles three-way factor model, and non-equivalent proofs (e.g., providing different sufficient conditions for uniqueness) were subsequently presented by Harshman (1972a), and de Leeuw and Pruzansky (1978). The most complete mathematical study of the rank of three-way arrays and the uniqueness of three-way decompositions was done by Kruskal (1976, 1977, 1981, 1989). More recently, interesting contributions have been made by Leurgans (1991) and Leurgans, Ross, and Abel (1993).

*Synthesis of models.* It would be useful in some situations to obtain intrinsic axes solutions for models incorporating interactions among dimensions. The uniqueness could help "tie down" empirically the wide range of solutions that the interaction terms otherwise make possible, and facilitate the discovery of new generalizations about interacting processes. One simple example is the issue of oblique versus orthogonal axes and the role of "g" in the factorial models of intellectual abilities. Another model of considerable potential interest is an oblique axis version of INDSCAL. It would allow for stimulus characteristics on two different dimensions to interact when contributing to overall stimulus (dis)similarities (see, for example, MDS analysis of the Size-Weight illusion, Dunn & Harshman, 1982).

A long-standing unresolved question has been whether models can combine interacting or oblique dimensions with the Proportional Profiles principle in a way that preserves dimensional uniqueness. One discouraging finding was Meredith's (1964) demonstration that if one permits unrestricted individual variation in the angles among oblique or interacting dimensions, nonunique proportional profiles solutions result. Harshman (1972b) subsequently conjectured, however, that proportional profiles models of intermediate generality, where some constraints are imposed on the variation of the factor correlations or interactions, could retain the uniqueness property. The PARAFAC2 model was offered as a case in point. Its uniqueness in the general case was questioned by Carroll and Wish (1974), and they reported finding a proof that with only two parallel occasions the model was not unique. The question of its uniqueness or nonuniqueness with three or more parallel occasions has resisted solution for over 20 years. Cattell has pursued essentially the same problem viewed in a two-mode context (which he calls the "Oblique Confactor rotation" problem; Cattell, 1978) for more than 30 years without success.

*The PARATUCK PT2 Model*

We present here a very general Proportional Profiles model which incorporates interacting dimensions. Because it integrates two lines of current three-mode research, one based on Cattell's Parallel Profiles principle and the other on Tucker's three-way interaction approach, we have adopted a suggestion by Kroonenberg (personal communication in September of 1989) that this type of model be called "PARATUCK". In fact, we use "PARATUCK2" (abbreviated to "PT2") for the model given below, since it resembles the T2 variant of Tucker's Three Mode Factor Analysis (see Kroonenberg, 1983; Tucker 1972). We have also formulated several natural extensions—a PARATUCK3 or PT3 model which more closely approximates the Tucker T3 structure, and a four mode PT4 model—but restrict ourselves to PT2 here.

*Notation and Model*

Consider an I by J by K three-way array $\underline{\mathbf{X}}$ for which the three "ways" or "modes" are labeled (measurement or classification) Mode A, Mode B and Mode C and are indexed by i, j and k respectively ($i = 1, \ldots, I$; $j = 1, \ldots, J$; and $k = 1, \ldots, K$). For notational convenience, we consider the array to be a stack of K successive I by J "slices" or "frontal slabs", with $\mathbf{X}_k$ being the k-th two-way slice or slab, and give the general structure of $\underline{\mathbf{X}}$ using an arbitrary $\mathbf{X}_k$. Thus the PT2 model is written

$$\mathbf{X}_k = \mathbf{A}\ {}^{A}\mathbf{D}_k\ \mathbf{H}\ {}^{B}\mathbf{D}_k\ \mathbf{B}',$$

where $\mathbf{A}$ is the I by R factor loading or factor pattern matrix for Mode A; ${}^{A}\mathbf{D}_k$ (pronounced "sup A D sub k") is an R by R diagonal matrix containing the weights for the columns of $\mathbf{A}$ at level k of the third mode. Similarly, ${}^{B}\mathbf{D}_k$ ("sup B D sub k") is an S by S diagonal matrix giving the weights at level k for the columns of $\mathbf{B}$, the J by S factor loading or pattern matrix for Mode B of the data. Note that the R diagonal elements of ${}^{A}\mathbf{D}_k$ are the R elements in row k of ${}^{A}\mathbf{C}$, the K by R Mode C weight matrix associated with $\mathbf{A}$. Thus, ${}^{a}d_{rr,k} = {}^{a}c_{kr}$; likewise, ${}^{B}\mathbf{D}_k$ has diagonal elements taken from row k of ${}^{B}\mathbf{C}$, the K by S Mode C weight matrix associated with $\mathbf{B}$, and so ${}^{b}d_{ss,k} = {}^{b}c_{ks}$. The elements of the R by S matrix $\mathbf{H}$ give the relative strength of different "interactions" between the factors of Mode A and those of Mode B. By convention, designations "A" and "B" are assigned so that $R \geq S$.

It can be seen that PT2 is general in that it requires neither symmetry of the data or model nor the same dimensionality for Modes A and B; in these ways, it resembles Tucker's T2 and T3 models. Nonetheless, it is a true Proportional Profiles model because the three-way variations in structure arise solely from changes in the weights on dimensions, and so it is more restricted than Tucker's T2 or T3. PT2 can also be viewed as a three-way version of asymmetrically weighted Dual Domain DEDICOM (e.g., Harshman, Green, Wind & Lundy, 1982). The $\mathbf{H}$ matrix would then be the asymmetric relationship matrix (designated $\mathbf{R}$ in the DEDICOM model).

A family of PT2 models, some of which are three-way generalizations of the DEDICOM models discussed by Harshman et al. (1982), can be generated by putting various constraints on the above model. Considering only cases where $R = S$, we get the following:

1. Asymmetrically weighted Dual Domain DEDICOM, where there are no added constraints.
2. Symmetrically weighted Dual Domain DEDICOM, where ${}^{A}\mathbf{D}_k = {}^{B}\mathbf{D}_k$.
3. Asymmetrically weighted Single Domain DEDICOM, where $\mathbf{A} = \mathbf{B}$.
4. Symmetrically weighted Single Domain DEDICOM, where $\mathbf{A} = \mathbf{B}$ and ${}^{A}\mathbf{D}_k = {}^{B}\mathbf{D}_k$.
5. PARAFAC2, where $\mathbf{A} = \mathbf{B}$, ${}^{A}\mathbf{D}_k = {}^{B}\mathbf{D}_k$ and $\mathbf{H}$ is symmetric; also essentially the scalar product form of oblique INDSCAL.
6. Symmetrically weighted Skew-symmetric DEDICOM (see Dawson & Harshman, 1986 for the two-way case), where $\mathbf{A} = \mathbf{B}$, ${}^{A}\mathbf{D}_k = {}^{B}\mathbf{D}_k$, $\mathbf{H}$ is skew-symmetric, R is an even number and the factors are pairwise equal in Mode C (i.e., the diagonals are pairwise equal in $\mathbf{D}_k$).

In the following sections, we present proofs of the uniqueness of axis orientation for both the general PT2 model and for the symmetrically weighted version (i.e., where ${}^{A}\mathbf{D}_k = {}^{B}\mathbf{D}_k$), subject to the restrictions that R = S and $\mathbf{H}$ has no zeros. (These restrictions simplify certain steps in the argument, but we believe that they are not always required for uniqueness; see discussion.) Under these conditions, uniqueness is proven for all of the above special cases except for Skew-symmetric DEDICOM.

PT2 Theorem 1 (Nonsymmetric Weights, R = S Case)

*Theorem*

Consider an $I \times J \times K$ three-way array $\underline{\mathbf{X}}$ with the following (PT2) structure

$$\mathbf{X}_k = \mathbf{A}\ {}^{A}\mathbf{D}_k\ \mathbf{H}\ {}^{B}\mathbf{D}_k\ \mathbf{B}', \tag{1}$$

where $\mathbf{A}$ is $I \times R$, $\mathbf{B}$ is $J \times S$, $\mathbf{H}$ is $R \times S$, ${}^{A}\mathbf{D}_k$ is $R \times R$ diagonal and ${}^{B}\mathbf{D}_k$ is $S \times S$ diagonal. Suppose there is an alternative representation of $\underline{\mathbf{X}}$ with matrices of the same size and structural form

$$\mathbf{X}_k = \overset{*}{\mathbf{A}}\ {}^{A}\overset{*}{\mathbf{D}}_k\ \overset{*}{\mathbf{H}}\ {}^{B}\overset{*}{\mathbf{D}}_k\ \overset{*}{\mathbf{B}}'. \tag{2}$$

We show that, under the assumptions given below, the two representations are necessarily related as follows:

$$\overset{*}{\mathbf{A}}({}^{A}\mathbf{P}\ {}^{A}\boldsymbol{\Delta}) = \mathbf{A}, \tag{3}$$

$$\overset{*}{\mathbf{B}}({}^{B}\mathbf{P}\ {}^{B}\boldsymbol{\Delta}) = \mathbf{B}, \tag{4}$$

$$({}^{A}\bar{\boldsymbol{\Delta}}\ {}^{A}\boldsymbol{\Delta}^{-1}\ {}^{A}\mathbf{P}')\overset{*}{\mathbf{H}}({}^{B}\mathbf{P}\ {}^{B}\boldsymbol{\Delta}^{-1}\ {}^{B}\bar{\boldsymbol{\Delta}}) = \mathbf{H}, \tag{5}$$

and that for any $\mathbf{X}_k \neq \mathbf{0}$

$$(z_k\ {}^{A}\mathbf{P}')\,{}^{A}\overset{*}{\mathbf{D}}_k({}^{A}\mathbf{P}\ {}^{A}\bar{\boldsymbol{\Delta}}^{-1}) = {}^{A}\mathbf{D}_k, \text{ and} \tag{6}$$

$$(z_k^{-1}\ {}^{B}\mathbf{P}')\,{}^{B}\overset{*}{\mathbf{D}}_k({}^{B}\mathbf{P}\ {}^{B}\bar{\boldsymbol{\Delta}}^{-1}) = {}^{B}\mathbf{D}_k, \tag{7}$$

where the $\boldsymbol{\Delta}$ and $\bar{\boldsymbol{\Delta}}$ matrices are diagonal, the $\mathbf{P}$ are permutation matrices, and the $z_k$ are nonzero scalars. In other words, the representation in this form of any such $\underline{\mathbf{X}}$ is unique up to trivial permutation and/or rescaling of columns.

*Assumptions.* We assume that **A**, **B** and **H** have full column rank, **H** has no zero elements, and that there is "adequate variation in the dimension weights". We assume for this theorem that R = S (but for clarity, and to facilitate subsequent possible generalizations, we maintain the two symbols R and S).

*Adequate variation in dimension weights.* Let us define MAX ≡ RS(R + 1)(S + 1)/4. This is the maximum number of unique d-weight combinations that can be composed using two diagonal elements (possibly the same one twice) from each of ${}^{A}\mathbf{D}_k$ and ${}^{B}\mathbf{D}_k$. We also define a vector $\gamma_k$ whose MAX components are the d-weight combinations so constructed. That is, the elements of $\gamma_k$ are as follows:

$$
\begin{aligned}
&{}^{a}d^2_{11,k}\ {}^{b}d^2_{11,k}, \ldots, {}^{a}d^2_{RR,k}\ {}^{b}d^2_{SS,k}, {}^{a}d_{11,k}\ {}^{a}d_{22,k}\ {}^{b}d^2_{11,k}, \ldots, \\
&{}^{a}d_{(R-1)(R-1),k}\ {}^{a}d_{RR,k}\ {}^{b}d^2_{SS,k}, {}^{a}d^2_{11,k}\ {}^{b}d_{11,k}\ {}^{b}d_{22,k}, \ldots, \\
&{}^{a}d^2_{RR,k}\ {}^{b}d_{(S-1)(S-1),k}\ {}^{b}d_{SS,k}, {}^{a}d_{11,k}\ {}^{a}d_{22,k}\ {}^{b}d_{11,k}\ {}^{b}d_{22,k}, \ldots, \\
&{}^{a}d_{(R-1)(R-1),k}\ {}^{a}d_{RR,k}\ {}^{b}d_{(S-1)(S-1),k}\ {}^{b}d_{SS,k},
\end{aligned} \tag{8}
$$

where ${}^{a}d_{rr,k}$ refers to diagonal r of ${}^{A}\mathbf{D}_k$ and ${}^{b}d_{ss,k}$ to diagonal s of ${}^{B}\mathbf{D}_k$.

The factor weight variation across k is "adequate" to uniquely determine the axis orientation of the solution, given the following two conditions:

i. There are at least MAX distinct k for which the associated $\gamma_k$ are linearly independent (so that the matrix formed by these vectors has rank MAX) and
ii. All of the $\mathbf{D}_k$ are full rank for the MAX Mode C levels referred to in condition (i), that is, all diagonal elements of these $\mathbf{D}_k$ are nonzero, and thus the elements of the matrix from condition (i) are all nonzero.

(When R = S = 2, for example, at least 9 levels of Mode C are required to meet the above conditions; when R = S = 3, one needs at least 36.)

*Proof*

We first establish some rank and nonzero properties of several matrices and some equivalence relationships that follow easily when R = S. In particular, $\overset{*}{\mathbf{H}}$ and another matrix we call $\tilde{\mathbf{H}}_k$ are shown to have the same properties as **H** (i.e., full rank and hence no proportional rows, since R = S; no zero entries). This restricts the values possible for a "cross-product ratio", which we then set up. This ratio makes explicit the requirement of invariance in certain ratios of parameters that arises from the parallel profiles property of the model. This invariance requirement restricts the possible transformation matrices for **A** and **B**, and so allows us to first prove (3) and (4), and then (5), (6) and (7).

For the case of R = S = 1, the theorem follows immediately. The **P** matrices equal 1, the **D**, **H**, and **Δ** matrices reduce to scalars, MAX = 1, and **A** and **B** are the unique vectors determined (up to column rescaling) by the rank-1 decomposition of any $\mathbf{X}_k$. Hence, the remainder of the proof deals only with the case of R = S ≥ 2. □

*Rank and linear relations between alternative representations*

Equation (1) implies that $\mathbf{X}_k$ is rank R (= S) for the MAX or more levels of Mode C with full rank $\mathbf{D}_k$, and so all of the constituent matrices in (2) must be rank R (= S) for these levels as well. The $\mathbf{D}_k$ and $\overset{*}{\mathbf{D}}_k$ may vary in rank across the other levels of Mode C.

Since $\overset{*}{\mathbf{A}}$ and $\overset{*}{\mathbf{B}}$ have the same rank and size as $\mathbf{A}$ and $\mathbf{B}$, they have full column rank, and since $\mathbf{A}$ and $\overset{*}{\mathbf{A}}$ span the same rank-R space, as do $\mathbf{B}$ and $\overset{*}{\mathbf{B}}$, nonsingular linear transformations ${}^{A}\mathbf{T}$ and ${}^{B}\mathbf{T}$ must exist such that

$$\overset{*}{\mathbf{A}}\ {}^{A}\mathbf{T} = \mathbf{A}, \text{ and} \tag{9}$$

$$\overset{*}{\mathbf{B}}\ {}^{B}\mathbf{T} = \mathbf{B}. \tag{10}$$

Since $\overset{*}{\mathbf{A}}$ and $\overset{*}{\mathbf{B}}$ are full column rank, their Moore-Penrose generalized inverses $\overset{*}{\mathbf{A}}{}^{+}$ and $\overset{*}{\mathbf{B}}{}^{+}$ are full rank as well. Using this, we substitute (9) and (10) into (1), equate it with (2), and simplify to get two forms of a matrix that we call $\tilde{\mathbf{H}}_k$:

$${}^{A}\mathbf{T}\ {}^{A}\mathbf{D}_k\ \mathbf{H}\ {}^{B}\mathbf{D}_k\ {}^{B}\mathbf{T}' = {}^{A}\overset{*}{\mathbf{D}}_k\ \overset{*}{\mathbf{H}}\ {}^{B}\overset{*}{\mathbf{D}}_k \equiv \tilde{\mathbf{H}}_k. \tag{11}$$

Clearly, $\tilde{\mathbf{H}}_k$ must also be rank R for the MAX k with full rank $\mathbf{D}_k$. The scalar equivalents for any element of $\tilde{\mathbf{H}}_k$ are

$$\tilde{h}_{rs,k} = {}^{a}\overset{*}{d}_{rr,k}\ \overset{*}{h}_{rs}\ {}^{b}\overset{*}{d}_{ss,k}, \text{ and} \tag{12}$$

$$\tilde{h}_{rs,k} = \sum_{u=1}^{R} \sum_{w=1}^{S} {}^{a}t_{ru}\, {}^{a}d_{uu,k}\, h_{uw}\, {}^{b}d_{ww,k}\, {}^{b}t_{sw}. \tag{13}$$

These alternative forms of $\tilde{\mathbf{H}}_k$ will be useful below.

*Implications of all nonzero* **H**

We have assumed $\mathbf{H}$ is all nonzero and now want to show that $\overset{*}{\mathbf{H}}$ is also all nonzero. We start by supposing that any element of $\overset{*}{\mathbf{H}}$ is zero, say $\overset{*}{h}_{rs} = 0$. Then (12) implies that $\tilde{h}_{rs,k} = 0$ for all k, and (13) implies that

$$\begin{aligned}&[{}^{a}d_{11,k}\,{}^{b}d_{11,k}]({}^{a}t_{r1}h_{11}\,{}^{b}t_{s1}) + ,\ldots, + [{}^{a}d_{11,k}\,{}^{b}d_{SS,k}]({}^{a}t_{r1}h_{1S}\,{}^{b}t_{sS})\\ &\quad + [{}^{a}d_{22,k}\,{}^{b}d_{11,k}]({}^{a}t_{r2}h_{21}\,{}^{b}t_{s1}) + ,\ldots, + [{}^{a}d_{RR,k}\,{}^{b}d_{SS,k}]({}^{a}t_{rR}h_{RS}\,{}^{b}t_{sS}) = 0.\end{aligned} \tag{14}$$

The d-weights are nonzero by assumption for MAX levels of k. Suppose that some t-h-t products are also nonzero. Squaring both sides of (14) and then combining terms with equal d-weights gives a weighted sum of some elements of $\gamma_k$ which is set equal to zero for all MAX levels of k; this contradicts our linear independence assumption for the MAX $\gamma_k$. Now suppose instead that all t-h-t products are zero. This implies that all the ${}^{a}t\,{}^{b}t$ products are zero, since $\mathbf{H}$ is nonzero. Equation (14) involves all pairwise combinations of elements from row r of ${}^{A}\mathbf{T}$ and row s of ${}^{B}\mathbf{T}$ (i.e., the outer product of the two vectors). Hence, for all the products to be zero, either one or both rows must be zero, but this would violate the full rank of the $\mathbf{T}$ matrices.

We arrive at a contradiction regardless of our assumption about the value of the t-h-t products in (14). Thus we must conclude that all elements of $\overset{*}{\mathbf{H}}$ and hence $\tilde{\mathbf{H}}_k$ are nonzero, just as for $\mathbf{H}$. This has implications for the "cross-product ratio", which is introduced below.

*The "Cross-Product Ratio"*

The "parallel profiles" characteristic of the model implies that certain ratios of parameters remain constant across levels of Mode C. We now consider one such invariant relationship, a ratio of two products of weighted $\overset{*}{\mathbf{H}}$ elements, and hence of corresponding elements of $\tilde{\mathbf{H}}_k$, which we call the "cross-product ratio". One way to define it is

$$\frac{(\overset{*}{h}_{r's}\ {}^{a}\overset{*}{d}_{r'r',k}\ {}^{b}\overset{*}{d}_{ss,k})(\overset{*}{h}_{rs'}\ {}^{a}\overset{*}{d}_{rr,k}\ {}^{b}\overset{*}{d}_{s's',k})}{(\overset{*}{h}_{rs}\ {}^{a}\overset{*}{d}_{rr,k}\ {}^{b}\overset{*}{d}_{ss,k})(\overset{*}{h}_{r's'}\ {}^{a}\overset{*}{d}_{r'r',k}\ {}^{b}\overset{*}{d}_{s's',k})} = \frac{\overset{*}{h}_{r's}\ \overset{*}{h}_{rs'}}{\overset{*}{h}_{rs}\ \overset{*}{h}_{r's'}} = q_{rr'ss'}, \tag{15}$$

where r, r′ = 1, . . . , R and r ≠ r′; s, s′ = 1, . . . , S and s ≠ s′; and k is any level of Mode C for which the Theorem assumptions hold (and hence where ${}^{A}\overset{*}{\mathbf{D}}_k$ and ${}^{B}\overset{*}{\mathbf{D}}_k$ are full rank). We see that the ratio is independent of k because the $\overset{*}{d}$-weights cancel out.

Using the equivalence in (12), (15) can also be expressed as

$$\frac{\tilde{h}_{r's,k}\ \tilde{h}_{rs',k}}{\tilde{h}_{rs,k}\ \tilde{h}_{r's',k}} = q_{rr'ss'}, \tag{16}$$

where r, r′, s, and s′ are as defined above for (15); k is any level of Mode C for which the Theorem assumptions hold; and $q$ is a nonzero scalar value. Where the values of either r and r′ or s and s′ are reversed, we have the equally valid reciprocal form of (16). For example, $q_{2112} = q_{1212}^{-1}$.

As defined, each cross-product ratio involves four elements of $\tilde{\mathbf{H}}_k$ (or $\overset{*}{\mathbf{H}}$) that are the intersection points of two rows r and r′ and two columns s and s′, indicated by the subscripts of $q$. If the matrices are 2 × 2, only one ratio (or its reciprocal) exists; otherwise, more can be set up. Each ratio has a unique value that remains constant for the MAX values of k that meet the Theorem assumptions. For these values of k, the fact that $\tilde{\mathbf{H}}_k$ is all nonzero means that $q$ can never be zero, infinity or undefined. Except when R = S = 2, it may sometimes happen that $q = 1$, but the requirement that $\tilde{\mathbf{H}}_k$ has full rank restricts the number of such $q$. The mutual independence of the rows of $\tilde{\mathbf{H}}_k$ implies that for any pair of rows r and r′, there exist at least two columns s and s′ such that $q_{rr'ss'} \neq 1$. In other words, $q \neq 1$ occurs at least once for ratios constructed within a given pair of rows, although one cannot predict which columns will be involved. Equivalently, the column independence of the $\tilde{\mathbf{H}}_k$ implies that for any columns s and s′, there are at least two rows r and r′ such that $q_{rr'ss'} \neq 1$.

By substituting (13) into (16), we can derive an alternative representation of the cross-product ratio $q_{rr'ss'}$:

$$\frac{\sum_{u=1}^{R}\sum_{v=1}^{R}\sum_{w=1}^{S}\sum_{x=1}^{S} {}^{a}t_{ru}\,{}^{a}t_{r'v}\,{}^{a}d_{uu,k}\,{}^{a}d_{vv,k}\,h_{uw}\,h_{vx}\,{}^{b}d_{ww,k}\,{}^{b}d_{xx,k}\,{}^{b}t_{s'w}\,{}^{b}t_{sx}}{\sum_{u=1}^{R}\sum_{v=1}^{R}\sum_{w=1}^{S}\sum_{x=1}^{S} {}^{a}t_{ru}\,{}^{a}t_{r'v}\,{}^{a}d_{uu,k}\,{}^{a}d_{vv,k}\,h_{uw}\,h_{vx}\,{}^{b}d_{ww,k}\,{}^{b}d_{xx,k}\,{}^{b}t_{sw}\,{}^{b}t_{s'x}}. \tag{17}$$

We see that the numerator and denominator differ by the interchange of s and s′. Also, the fixed subscripts in (17) correspond to the subscripts of $q$, with r and r′ denoting the rows of ${}^{A}\mathbf{T}$ and s and s′ denoting the rows of ${}^{B}\mathbf{T}$ that are used in the alternative representation of the ratio. For example, $q_{1345}$ would involve elements from rows 1 and 3 of ${}^{A}\mathbf{T}$ and rows 4 and 5 of ${}^{B}\mathbf{T}$.

*Using the Cross-Product Ratio to find the form of* ${}^{A}\mathbf{T}$ *and* ${}^{B}\mathbf{T}$

The quadruple summation in (17) yields a sum of $R^2S^2$ terms in both the numerator and denominator of the ratio, some of which are identical in both places (but may arise in a different order) and some of which are not. The terms in the numerator that have identical matches in the denominator (and vice versa) will be called "ratio-matched" terms. In the remaining, unmatched terms, a given set of d-weights is multiplied by one set of t and h values in the numerator and a different set of t and h values in the denominator. The ratio-matched terms always involve at least one squared d-element (i.e., u = v and/or w = x) while the nonmatched ones do not (i.e., u ≠ v and w ≠ x). The ratio-matched terms will play a key role in the proof.

All d-weights (i.e., d-element combinations) represented in (8) occur in both the top and bottom of the ratio, some arising in only one term in each place, some in two terms and others in four terms in both the numerator and denominator.

*Example terms.* We will now look at some specific terms in the numerator and denominator of (17). First, consider terms with two squared d-elements. For example, if uvwx = 1111 (u = v and w = x) we obtain ${}^{a}t_{r1}\,{}^{a}t_{r'1}\,{}^{a}d_{11,k}^{2}h_{11}^{2}\,{}^{b}d_{11,k}^{2}\,{}^{b}t_{s'1}\,{}^{b}t_{s1}$. Each such combination of two squared d-elements occurs only once in the numerator and once in the denominator and is ratio matched (i.e., is associated with the same set of t and h elements in both places). Next, consider terms with one squared d-element. For example, when uvwx = 1112 (u = v, w ≠ x) in the numerator, we have ${}^{a}t_{r1}\,{}^{a}t_{r'1}\,{}^{a}d_{11,k}^{2}h_{11}h_{12}\,{}^{b}d_{11,k}\,{}^{b}d_{22,k}\,{}^{b}t_{s'1}\,{}^{b}t_{s2}$ and when uvwx = 1121 (as before, except wx values have been reversed), we obtain the same term except the ${}^{b}$t-elements are now ${}^{b}t_{s'2}\,{}^{b}t_{s1}$. (The reversal of order of the h elements has no effect and so will be ignored.) The identical (ratio-matched) terms occur in the denominator but in the reverse order—the first when uvwx = 1121 and the second when uvwx = 1112. Likewise, when uvwx = 1211 (u ≠ v, w = x) or uvwx = 2111 in the numerator, we obtain two terms with the same d-elements—${}^{a}d_{11,k}\,{}^{a}d_{22,k}\,{}^{b}d_{11,k}^{2}$. These two uvwx combinations produce ratio-matching terms in the denominator, and this time they arise in the same order (for the same uvwx combinations) as in the numerator.

Finally, we consider terms with no squared d-elements. For example, the term produced by uvwx = 1212 (u ≠ v and w ≠ x) has a d-weight of ${}^{a}d_{11,k}\,{}^{a}d_{22,k}\,{}^{b}d_{11,k}\,{}^{b}d_{22,k}$. The three other "reversals" of uv and wx values—2112, 1221 and 2121—also yield terms with this d-weight in the numerator and denominator. However, none of these terms have identical tt-hh-tt products in both places, as can be verified by the reader.

All possible uvwx values in (17) produce terms with d-weight characteristics in both the numerator and denominator that are similar to those above: either a unique combination of two squared d-elements; or a combination of one squared and two unsquared d-elements that is replicated in two terms; or a weight consisting of four unsquared d-elements that occurs in four different terms. And, as noted earlier, any term in the numerator containing at least one squared d-element has a "ratio-matched" or identical term in the denominator, and vice versa. This fact will be used to help determine the form of ${}^{A}\mathbf{T}$ and ${}^{B}\mathbf{T}$.

*Factoring by d-weights.* Suppose we now combine terms with the same d-weights. Those with one squared d-element can be collected into pairs (since, as we saw above, any such d-weight occurs in two terms) and those with no squared elements can be collected into groups of four. Factoring out the common d-weights results in MAX terms in the numerator and in the denominator, with each of the d-weight combinations given in (8) occurring once in each place.

We will call the sums of tt-hh-tt products associated with a given d-weight that weight's "t-h-t coefficient." For example, the t-h-t coefficient associated with the d-weight ${}^{a}d_{11,k}^{2}\,{}^{b}d_{11,k}^{2}$ is ${}^{a}t_{r1}\,{}^{a}t_{r'1}h_{11}^{2}\,{}^{b}t_{s'1}\,{}^{b}t_{s1}$ (where uvwx = 1111), and the one associated with ${}^{a}d_{11,k}^{2}\,{}^{b}d_{11,k}\,{}^{b}d_{22,k}$ is $({}^{a}t_{r1}\,{}^{a}t_{r'1}h_{11}h_{12}\,{}^{b}t_{s'1}\,{}^{b}t_{s2} + {}^{a}t_{r1}\,{}^{a}t_{r'1}h_{11}h_{12}\,{}^{b}t_{s'2}\,{}^{b}t_{s1})$ (from uvwx = 1112 and 1121). Note that these t-h-t coefficients are identical in the numerator and denominator of (17); hence these particular composite terms resulting from the factoring by d-weights are also ratio-matched. In contrast, the t-h-t coefficient for ${}^{a}d_{11,k}\,{}^{a}d_{22,k}\,{}^{b}d_{11,k}\,{}^{b}d_{22,k}$ (uvwx = 1212, 2112, etc.) is a sum of four products, and, while this d-weight occurs in both the top and bottom of (17), the t-h-t coefficients in the two places differ.

Let us now rewrite (17) more compactly to emphasize the structure of the ratio after factoring by d-weights. Suppose we reorder our terms so that all ratio-matched terms occur before any nonmatched terms. We can then define $[f_i(d_k)]$ as that nonlinear function of d-elements corresponding to the i-th d-weight in the numerator and denominator of (17) after factoring by d-weights and reordering. (Note that this can also be defined as the i-th element of an equivalently reordered version of the $\gamma_k$ vector given in (8).) Let $\alpha_i$ or $\beta_j$ represent the corresponding t-h-t coefficient (i.e., a sum of one, two, or four tt-hh-tt products, as demonstrated above). Suppose also that M is the number of terms that are ratio-matched (identical in the numerator and denominator) and N is the number of unmatched terms. (Thus, M = RS(R + S)/2, N = RS(R − 1)(S − 1)/4 and M + N = MAX). Then (17) becomes

$$\frac{([f_1(d_k)]\alpha_1 + \cdots + [f_M(d_k)]\alpha_M) + [f_{M+1}(d_k)]\alpha_{M+1} + \cdots + [f_{M+N}(d_k)]\alpha_{M+N}}{([f_1(d_k)]\alpha_1 + \cdots + [f_M(d_k)]\alpha_M) + [f_{M+1}(d_k)]\beta_1 + \cdots + [f_{M+N}(d_k)]\beta_N} = q. \tag{18}$$

Relating (18) to (17), we see that the first M terms in (18) are ratio-matched, with d-weights comprised of at least one squared d-element and $\alpha$-weights comprised of tt-hh-tt products from one or two terms in (17). The subsequent N terms have matching d-weights (4 unsquared d-elements) but different t-h-t coefficients (the sum of tt-hh-tt products from 4 terms). While the d-weights vary with k, the $\alpha$- and $\beta$-weights do not, since they are composed of elements from fixed matrices.

When we rewrite (18) as

$$(q - 1)([f_1(d_k)]\alpha_1 + \cdots + [f_M(d_k)]\alpha_M) + [f_{M+1}(d_k)](q\beta_1 - \alpha_{M+1}) + \cdots + [f_{M+N}(d_k)](q\beta_N - \alpha_{M+N}) = 0, \tag{19}$$

it becomes clear that to maintain our linear independence assumption the weights for every $[f_i(d_k)]$ must be zero (recall the conditions associated with (8), and that $f_i(d_k)$ is the i-th element of $\gamma_k$). If we choose $\tilde{\mathbf{H}}_k$ elements such that $q \neq 1$, we then must have $\alpha_1, \ldots, \alpha_M = 0$. As will be demonstrated below, the zero property of these M $\alpha$-weights (i.e., the t-h-t coefficients for the ratio-matched terms) is sufficient to determine the form of ${}^{A}\mathbf{T}$ and ${}^{B}\mathbf{T}$.

*A simplifying convention for* ${}^{A}\mathbf{T}$ *and* ${}^{B}\mathbf{T}$. The full rank property of ${}^{A}\mathbf{T}$ and ${}^{B}\mathbf{T}$ implies that at least one nonzero element occurs in each row and column. Without loss of generality, we can adopt the convention that the columns of $\overset{*}{\mathbf{A}}$ and $\overset{*}{\mathbf{B}}$ will be ordered such that there are nonzero elements on the diagonals of ${}^{A}\mathbf{T}$ and ${}^{B}\mathbf{T}$, which is always possible. This does not limit the proof, since the Theorem claims equivalence only up to a possible permutation and rescaling of columns.

*Filling in* ${}^{A}\mathbf{T}$ *and* ${}^{B}\mathbf{T}$. For any two rows i and j ($i \neq j$) of $\tilde{\mathbf{H}}_k$, we choose two columns s, s′ ($s \neq s'$) such that $q_{ijss'} \neq 1$. This is always possible, since the rows of $\tilde{\mathbf{H}}_k$ are nonproportional (because $\tilde{\mathbf{H}}_k$ has full rank). We modify (17) for $q_{ijss'}$ (i.e., replacing the fixed subscripts r and r′ with i and j, respectively) and use the zero property of $\alpha$-weights for certain ratio-matched terms to show that ${}^{a}t_{ij} = 0$. (To get ${}^{a}t_{21} = 0$, for example, we would work with ratio-matched terms from $q_{12ss'} \neq 1$, and in the equations below, would substitute 1 for i and 2 for j, respectively.)

We begin by examining the term with two squared d-weights (and hence only one tt-hh-tt product in the corresponding $\alpha$) obtained from uvwx = iiss. As noted earlier, such terms are ratio-matched, and so the $\alpha$-weight is zero, giving us

$$ {}^{a}t_{ii}\,{}^{a}t_{ji}\,h_{is}\,h_{is}\,{}^{b}t_{s's}\,{}^{b}t_{ss} = 0, \tag{20} $$

which implies that ${}^{a}t_{ji}\,{}^{b}t_{s's} = 0$ (since diagonal elements are nonzero by convention and $\mathbf{H}$ is all nonzero by assumption).

Next we examine ratio-matched terms with one squared d-weight (and hence with a sum of two tt-ht-tt products in the associated $\alpha$). We consider the $\alpha$ derived from the two terms where uvwx = iiss′ and uvwx = iis′s. The tt-hh-tt product from uvwx = iiss′ immediately vanishes because ${}^{a}t_{ji}\,{}^{b}t_{s's} = 0$, leaving

$$ {}^{a}t_{ii}\,{}^{a}t_{ji}\,h_{is'}\,h_{is}\,{}^{b}t_{s's'}\,{}^{b}t_{ss} = 0, \tag{21} $$

which implies ${}^{a}t_{ji} = 0$. This holds for every i, j where $i \neq j$, and so ${}^{A}\mathbf{T}$ is shown to be diagonal.

We can similarly show that ${}^{B}\mathbf{T}$ is also diagonal, except now it is the nonproportionality of the *columns* of $\tilde{\mathbf{H}}_k$ that allows us to find $q_{rr'ij} \neq 1$ ($i \neq j$). Using the zero property of the $\alpha$-weights for the ratio-matched terms where uvwx = rrii, and uvwx = rr′ii and r′rii, we find that ${}^{b}t_{ji} = 0$.

*Consequences of diagonal* $\mathbf{T}$ *matrices: Proof of (3) and (4)*

We have shown ${}^{A}\mathbf{T}$ and ${}^{B}\mathbf{T}$ to be diagonal matrices when we initially assume that the order of columns in $\overset{*}{\mathbf{A}}$ and $\overset{*}{\mathbf{B}}$ is such that no diagonal elements of ${}^{A}\mathbf{T}$ or ${}^{B}\mathbf{T}$ are zero. To determine the form of the transformation matrices when all permutations of column order of $\overset{*}{\mathbf{A}}$ and $\overset{*}{\mathbf{B}}$ are allowed (hence allowing the nonzero elements of the $\mathbf{T}$ matrices to occur off the diagonal), we define a diagonal matrix ${}^{A}\mathbf{\Delta}$ and a permutation matrix ${}^{A}\mathbf{P}$ (either or both of which might be identity matrices) such that

$$ {}^{A}\mathbf{T} = {}^{A}\mathbf{P}\ {}^{A}\mathbf{\Delta}. \tag{22} $$

Similarly, for ${}^{B}\mathbf{T}$ we define a diagonal matrix ${}^{B}\mathbf{\Delta}$ and a permutation matrix ${}^{B}\mathbf{P}$ such that

$$ {}^{B}\mathbf{T} = {}^{B}\mathbf{P}\ {}^{B}\mathbf{\Delta}. \tag{23} $$

Substituting (22) and (23) into (9) and (10), respectively, yields

$$ \overset{*}{\mathbf{A}}\ {}^{A}\mathbf{P}\ {}^{A}\mathbf{\Delta} = \mathbf{A}, \text{ and} \tag{24} $$

$$ \overset{*}{\mathbf{B}}\ {}^{B}\mathbf{P}\ {}^{B}\mathbf{\Delta} = \mathbf{B}, \tag{25} $$

which proves (3) and (4) in the Theorem. □

*Relationship of* $\mathbf{H}$ *to* $\overset{*}{\mathbf{H}}$*: Proof of (5)*

Substituting (22) and (23) into (11), we get

$$ {}^{A}\mathbf{P}\ {}^{A}\boldsymbol{\Delta}\ {}^{A}\mathbf{D}_k\ \mathbf{H}\ {}^{B}\mathbf{D}_k\ {}^{B}\boldsymbol{\Delta}\ {}^{B}\mathbf{P}' = {}^{A}\overset{*}{\mathbf{D}}_k\ \overset{*}{\mathbf{H}}\ {}^{B}\overset{*}{\mathbf{D}}_k. \tag{26} $$

Now we choose some level with full rank $\mathbf{D}_k$, say $k = 1$. We pre- and post-multiply $\overset{*}{\mathbf{H}}$ by the identity matrices $({}^{A}\mathbf{P}\ {}^{A}\mathbf{P}')$ and $({}^{B}\mathbf{P}\ {}^{B}\mathbf{P}')$, respectively, and rewrite (26) somewhat, commuting diagonal matrices in the process, to obtain

$$ \mathbf{H} = {}^{A}\mathbf{D}_1^{-1}({}^{A}\mathbf{P}'\ {}^{A}\overset{*}{\mathbf{D}}_1\ {}^{A}\mathbf{P}){}^{A}\boldsymbol{\Delta}^{-1}\ {}^{A}\mathbf{P}'\ \overset{*}{\mathbf{H}}\ {}^{B}\mathbf{P}\ {}^{B}\boldsymbol{\Delta}^{-1}({}^{B}\mathbf{P}'\ {}^{B}\overset{*}{\mathbf{D}}_1\ {}^{B}\mathbf{P}){}^{B}\mathbf{D}_1^{-1}. \tag{27} $$

Note that ${}^{A}\boldsymbol{\Delta}^{-1}$ and ${}^{B}\boldsymbol{\Delta}^{-1}$ exist because ${}^{A}\boldsymbol{\Delta}$ and ${}^{B}\boldsymbol{\Delta}$ are nonsingular as defined. The expressions in parentheses are versions of ${}^{A}\overset{*}{\mathbf{D}}_1$ and ${}^{B}\overset{*}{\mathbf{D}}_1$ in which diagonals have been reordered if necessary to match the ordering in ${}^{A}\mathbf{D}_1$ and ${}^{B}\mathbf{D}_1$, respectively.

Let us now define the diagonal matrices

$$ {}^{A}\bar{\boldsymbol{\Delta}} \equiv {}^{A}\mathbf{D}_1^{-1}\ {}^{A}\mathbf{P}'\ {}^{A}\overset{*}{\mathbf{D}}_1\ {}^{A}\mathbf{P} \text{ and} $$

$$ {}^{B}\bar{\boldsymbol{\Delta}} \equiv {}^{B}\mathbf{P}'\ {}^{B}\overset{*}{\mathbf{D}}_1\ {}^{B}\mathbf{P}\ {}^{B}\mathbf{D}_1^{-1}, \tag{28} $$

and rewrite (27) as

$$ \mathbf{H} = ({}^{A}\bar{\boldsymbol{\Delta}}\ {}^{A}\boldsymbol{\Delta}^{-1}\ {}^{A}\mathbf{P}')\overset{*}{\mathbf{H}}({}^{B}\mathbf{P}\ {}^{B}\boldsymbol{\Delta}^{-1}\ {}^{B}\bar{\boldsymbol{\Delta}}). \tag{29} $$

Although the definition of ${}^{A}\bar{\Delta}$ and ${}^{B}\bar{\Delta}$ is somewhat arbitrary (i.e., any level with full rank $\mathbf{D}_k$ can be chosen), the form of (29) is not. Hence (5) is proven. □

$\mathbf{D}$*-weight relationships: Proof of (6) and (7)*

*A scalar indeterminacy.* There is a trivial scalar indeterminacy basic to the model. Since symmetry in the parameters on the left versus right of $\mathbf{H}$ is not required, a scalar can be applied to any matrix in (1), so long as its reciprocal is applied to some other matrix in the product. Of course, the same scalar must be applied through all slices for $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{H}$, since these matrices are fixed for all k. Thus, the scaling indeterminacy for these matrices can be eliminated by adopting conventions such as unit length factors and unit $\mathbf{H}$ diagonals, for example, with compensatory scaling of ${}^{A}\mathbf{C}$ and/or ${}^{B}\mathbf{C}$ (and hence uniform rescaling of all the ${}^{A}\mathbf{D}_k$ and/or ${}^{B}\mathbf{D}_k$ matrices) as needed. However, there is an additional scalar indeterminacy of the two $\mathbf{D}_k$ matrices at any given value of k. A distinct scalar multiplication of ${}^{A}\mathbf{D}_k$ can be done at any k, provided ${}^{B}\mathbf{D}_k$ at the same level is multiplied by the reciprocal scalar. We will use $z_k$ and $z_k^{-1}$ in the following section to represent the scalar indeterminacy primarily with respect to its possible effect on the relationships between ${}^{A}\mathbf{D}_k$ and ${}^{B}\mathbf{D}_k$ weights.

After substituting (29) into (26) and simplifying, we have

$$ {}^{A}\mathbf{P}\ {}^{A}\mathbf{D}_k\ {}^{A}\bar{\boldsymbol{\Delta}}\ {}^{A}\mathbf{P}'\ \overset{*}{\mathbf{H}}\ {}^{B}\mathbf{P}\ {}^{B}\bar{\boldsymbol{\Delta}}\ {}^{B}\mathbf{D}_k\ {}^{B}\mathbf{P}' = {}^{A}\overset{*}{\mathbf{D}}_k\ \overset{*}{\mathbf{H}}\ {}^{B}\overset{*}{\mathbf{D}}_k. \tag{30} $$

Now we define

$$ {}^{A}\dot{\mathbf{D}}_k \equiv {}^{A}\mathbf{P}\ {}^{A}\mathbf{D}_k\ {}^{A}\bar{\boldsymbol{\Delta}}\ {}^{A}\mathbf{P}' \text{ and } {}^{B}\dot{\mathbf{D}}_k \equiv {}^{B}\mathbf{P}\ {}^{B}\bar{\boldsymbol{\Delta}}\ {}^{B}\mathbf{D}_k\ {}^{B}\mathbf{P}', \tag{31} $$

and (30) becomes

$$ {}^{A}\dot{\mathbf{D}}_k \overset{*}{\mathbf{H}} {}^{B}\dot{\mathbf{D}}_k = {}^{A}\overset{*}{\mathbf{D}}_k \overset{*}{\mathbf{H}} {}^{B}\overset{*}{\mathbf{D}}_k. \tag{32} $$

The $\dot{\mathbf{D}}_k$ matrices are just rescaled versions of the $\mathbf{D}_k$ with diagonals permuted to match the ordering in the $\overset{*}{\mathbf{D}}_k$, and so retain the rank properties of the $\mathbf{D}_k$.

Equation (32) is equivalent to RS scalar equations of the form

$$ {}^{a}\dot{d}_{uu,k} \overset{*}{h}_{uw} {}^{b}\dot{d}_{ww,k} = {}^{a}\overset{*}{d}_{uu,k} \overset{*}{h}_{uw} {}^{b}\overset{*}{d}_{ww,k} \text{ for } u = 1, \ldots, R \text{ and } w = 1, \ldots, S, \tag{33} $$

which, because $\overset{*}{\mathbf{H}}$ is all nonzero, can be rewritten as

$$ {}^{a}\dot{d}_{uu,k} {}^{b}\dot{d}_{ww,k} = {}^{a}\overset{*}{d}_{uu,k} {}^{b}\overset{*}{d}_{ww,k}. \tag{34} $$

The RS equations implied by (34) can be written explicitly as an equality of two outer products. If we let ${}^{b}\overset{*}{\mathbf{c}}_k = \text{diag}({}^{B}\overset{*}{\mathbf{D}}_k)$ (i.e., row k of ${}^{B}\overset{*}{\mathbf{C}}$, the Mode C weight or "loading" matrix applied to $\overset{*}{\mathbf{B}}$), let ${}^{b}\dot{\mathbf{c}}_k = \text{diag}({}^{B}\dot{\mathbf{D}}_k)$ (i.e., the vector with elements ${}^{b}\dot{d}_{11,k} \ldots {}^{b}\dot{d}_{uu,k}, \ldots {}^{b}\dot{d}_{RR,k}$), and define ${}^{a}\overset{*}{\mathbf{c}}_k$ and ${}^{a}\dot{\mathbf{c}}_k$ in parallel fashion, then we can rewrite (34) as

$$ {}^{a}\dot{\mathbf{c}}_k {}^{b}\dot{\mathbf{c}}'_k = {}^{a}\overset{*}{\mathbf{c}}_k {}^{b}\overset{*}{\mathbf{c}}'_k. \tag{35} $$

It is evident from (35) that

$$ {}^{a}\dot{\mathbf{c}}_k = z_k {}^{a}\overset{*}{\mathbf{c}}_k, \text{ and} $$

$$ {}^{b}\dot{\mathbf{c}}_k = z_k^{-1} {}^{b}\overset{*}{\mathbf{c}}_k, \tag{36} $$

provided all the $\mathbf{c}$ vectors are nonzero. Thus whenever $\mathbf{X}_k \neq \mathbf{0}$,

$$ {}^{A}\dot{\mathbf{D}}_k = z_k {}^{A}\overset{*}{\mathbf{D}}_k, \text{ and} $$

$$ {}^{B}\dot{\mathbf{D}}_k = z_k^{-1} {}^{B}\overset{*}{\mathbf{D}}_k. \tag{37} $$

To find the relationship between the $\mathbf{D}_k$ and $\overset{*}{\mathbf{D}}_k$, we now substitute from (37) into (31) and rearrange, obtaining

$$ {}^{A}\mathbf{D}_k = z_k {}^{A}\mathbf{P}' {}^{A}\overset{*}{\mathbf{D}}_k {}^{A}\mathbf{P} {}^{A}\bar{\mathbf{\Delta}}^{-1} \text{ and} $$

$$ {}^{B}\mathbf{D}_k = z_k^{-1} {}^{B}\mathbf{P}' {}^{B}\overset{*}{\mathbf{D}}_k {}^{B}\mathbf{P} {}^{B}\bar{\mathbf{\Delta}}^{-1}. \tag{38} $$

This proves (6) and (7) for $\mathbf{X}_k \neq \mathbf{0}$, and the proof of Theorem 1 is complete. □

## PT2 Theorem 2 (Symmetrically Weighted Case)

Consider the special case of PARATUCK2 in which ${}^{A}\mathbf{D}_k = {}^{B}\mathbf{D}_k$ for all k. Here we will denote the weight matrix simply as $\mathbf{D}_k$. For this case, rotational uniqueness does not automatically follow from Theorem 1, because the definition of "adequate variation" used there is based on the assumption of two independently varying $\mathbf{D}_k$. While we can use many of the same steps as in Theorem 1, a separate argument will be required, and this is provided below. (Because it is an independent case, we will start at (1) again for the equation sequence numbers.)

*Theorem*

Suppose there exists an I by J by K three-way array $\underline{\mathbf{X}}$ with the following structure

$$\mathbf{X}_k = \mathbf{A}\ \mathbf{D}_k\ \mathbf{H}\ \mathbf{D}_k\ \mathbf{B}', \tag{1}$$

where $\mathbf{A}$ is I × R, $\mathbf{B}$ is J × R, $\mathbf{H}$ is R × R, and the $\mathbf{D}_k$ are R × R diagonal. Now suppose there is an alternative representation of $\underline{\mathbf{X}}$ with matrices of the same size and structural form

$$\mathbf{X}_k = \overset{*}{\mathbf{A}}\ \overset{*}{\mathbf{D}}_k\ \overset{*}{\mathbf{H}}\ \overset{*}{\mathbf{D}}_k\ \overset{*}{\mathbf{B}}'. \tag{2}$$

We show that, under the assumptions given below, the two representations are necessarily related as follows:

$$\overset{*}{\mathbf{A}}(\mathbf{P}\ {}^{A}\boldsymbol{\Delta}) = \mathbf{A}, \tag{3}$$

$$\overset{*}{\mathbf{B}}(\mathbf{P}\ {}^{B}\boldsymbol{\Delta}) = \mathbf{B}, \tag{4}$$

$$(\bar{\boldsymbol{\Delta}}\ {}^{A}\boldsymbol{\Delta}^{-1}\ \mathbf{P}')\overset{*}{\mathbf{H}}(\mathbf{P}\ {}^{B}\boldsymbol{\Delta}^{-1}\ \bar{\boldsymbol{\Delta}}) = \mathbf{H}, \tag{5}$$

and that for any $\mathbf{X}_k \neq \mathbf{0}$

$$(z_k\ \mathbf{P}')\overset{*}{\mathbf{D}}_k(\mathbf{P}\ \bar{\boldsymbol{\Delta}}^{-1}) = \mathbf{D}_k, \tag{6}$$

where the $\boldsymbol{\Delta}$ and $\bar{\boldsymbol{\Delta}}$ matrices are diagonal, $\mathbf{P}$ is a permutation matrix, and $z_k = \pm 1$. In other words, the representation in this form of any such $\underline{\mathbf{X}}$ is unique up to trivial permutation and/or rescaling of columns.

*Assumptions.* We assume that $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{H}$ have rank R, that $\mathbf{H}$ has no zero elements, and that there is "adequate variation in the dimension weights".

*Adequate variation in dimension weights.* Define MAXS ≡ R(R + 1)(R + 2)(R + 3)/24, the maximum number of unique d-weight combinations that can be composed using four diagonal elements (with possible repetition) from $\mathbf{D}_k$. Let us also define a vector $\boldsymbol{\tau}_k$ whose MAXS components are so constructed. The elements of $\boldsymbol{\tau}_k$ are as follows:

$$d^4_{11,k}, \ldots, d^4_{RR,k}, d^3_{11,k}\, d_{22,k}, \ldots, d_{(R-1)(R-1),k}\, d^3_{RR,k}, d^2_{11,k}\, d^2_{22,k}, \ldots,$$
$$d^2_{(R-1)(R-1),k}\, d^2_{RR,k}, d^2_{11,k}\, d_{22,k}\, d_{33,k}, \ldots, d_{(R-2)(R-2),k}\, d_{(R-1)(R-1),k}\, d^2_{RR,k},$$
$$d_{11,k}\, d_{22,k}\, d_{33,k}\, d_{44,k}, \ldots, d_{(R-3)(R-3),k}\, d_{(R-2)(R-2),k}\, d_{(R-1)(R-1),k}\, d_{RR,k}. \tag{7}$$

The subscript notation "rr,k" is the same as was used in Theorem 1. Obviously, some elements do not occur if R ≤ 3 (e.g., ones with four first-power d-weights).

We say the factor weight variation across k is "adequate" to uniquely determine the axis orientation of the solution, given the following two conditions:

i. There are at least MAXS distinct k for which the associated $\boldsymbol{\tau}_k$ are linearly independent (so that the square matrix formed by these vectors has rank MAXS) and

For the MAXS Mode C levels in condition (i), none of the dimension weights is zero (i.e., all diagonal elements of the $\mathbf{D}_k$ are nonzero, and hence the components of the MAXS $\boldsymbol{\tau}_k$ are all nonzero).

(When R = 2, for example, at least 5 levels of Mode C are required to meet the above conditions; when R = 3 or R = 4, one needs 15 or 35, respectively.)

*Proof*

The argument proceeds as for Theorem 1 except that ${}^{a}d_{rr,k} = {}^{b}d_{rr,k} = d_{rr,k}$ and ${}^{a}\overset{*}{d}_{rr,k} = {}^{b}\overset{*}{d}_{rr,k} = \overset{*}{d}_{rr,k}$, and so ${}^{A}\mathbf{D}_k$ and ${}^{B}\mathbf{D}_k$ are replaced with $\mathbf{D}_k$, and ${}^{a}d$ and ${}^{b}d$ are replaced with d in all the equations, and the same for the $\overset{*}{\mathbf{D}}_k$ and $\overset{*}{d}$. Also, wherever the linear independence of the MAX $\boldsymbol{\gamma}_k$ was invoked, the linear independence of the MAXS $\boldsymbol{\tau}_k$ is now used. Necessary equations are given below, but the reader is sometimes left to fill in details of their derivation by referring to corresponding sections in the Theorem 1 proof.

As in Theorem 1, the case of R = 1 follows immediately, and so the argument focuses on cases where R ≥ 2. □

*Rank, linear transformation, and nonzero parameter relationships*

Using the same reasoning as for Theorem 1, it is clear that $\overset{*}{\mathbf{A}}, \overset{*}{\mathbf{B}}$ and $\overset{*}{\mathbf{H}}$ are rank R, and the $\overset{*}{\mathbf{D}}_k$ are rank R for at least MAXS levels of Mode C. Furthermore, nonsingular transformations ${}^{A}\mathbf{T}$ and ${}^{B}\mathbf{T}$ exist for $\mathbf{A}$ and $\mathbf{B}$, respectively, such that

$$\overset{*}{\mathbf{A}}\ {}^{A}\mathbf{T} = \mathbf{A} \text{ and} \tag{8}$$

$$\overset{*}{\mathbf{B}}\ {}^{B}\mathbf{T} = \mathbf{B}. \tag{9}$$

As before, we can now derive two equivalent representations for $\tilde{\mathbf{H}}_k$:

$${}^{A}\mathbf{T}\ \mathbf{D}_k\ \mathbf{H}\ \mathbf{D}_k\ {}^{B}\mathbf{T}' = \overset{*}{\mathbf{D}}_k\ \overset{*}{\mathbf{H}}\ \overset{*}{\mathbf{D}}_k = \tilde{\mathbf{H}}_k. \tag{10}$$

Again it is evident that $\tilde{\mathbf{H}}_k$ is rank R for at least MAXS k. The corresponding scalar equations are

$$\tilde{h}_{rs,k} = \overset{*}{d}_{rr,k}\ \overset{*}{h}_{rs}\ \overset{*}{d}_{ss,k} \text{ and} \tag{11}$$

$$\tilde{h}_{rs,k} = \sum_{u=1}^{R} \sum_{w=1}^{R} {}^{a}t_{ru}\, d_{uu,k}\, h_{uw}\, d_{ww,k}\, {}^{b}t_{sw}, \tag{12}$$

for r and s = 1, . . . , R. We can also show that an all nonzero $\mathbf{H}$ implies an all nonzero $\overset{*}{\mathbf{H}}$ and $\tilde{\mathbf{H}}_k$, using the same argument as in Theorem 1.

*The "Cross-Product Ratio"*

As before, we consider a ratio of four elements from $\tilde{\mathbf{H}}_k$ (and equivalently, from $\overset{*}{\mathbf{H}}$) which remains constant for the (at least MAXS) levels of k where the theorem assumptions hold. As in Theorem 1, it is defined as

$$\frac{\tilde{h}_{r's,k}\ \tilde{h}_{rs',k}}{\tilde{h}_{rs,k}\ \tilde{h}_{r's',k}} = q_{rr'ss'} \tag{13}$$

where r, r′, s and s′ = 1, . . . , R; $r \neq r'$; $s \neq s'$; and k is any level of Mode C for which the theorem assumptions hold. All properties of the ratio described in Theorem 1 also hold here.

After substituting (12) into (13), the alternative form of the symmetrically-weighted ratio is

$$\frac{\sum_{u=1}^{R}\sum_{v=1}^{R}\sum_{w=1}^{R}\sum_{x=1}^{R} {}^{a}t_{ru}\,{}^{a}t_{r'v}\,d_{uu,k}\,d_{vv,k}\,h_{uw}\,h_{vx}\,d_{ww,k}\,d_{xx,k}\,{}^{b}t_{s'w}\,{}^{b}t_{sx}}{\sum_{u=1}^{R}\sum_{v=1}^{R}\sum_{w=1}^{R}\sum_{x=1}^{R} {}^{a}t_{ru}\,{}^{a}t_{r'v}\,d_{uu,k}\,d_{vv,k}\,h_{uw}\,h_{vx}\,d_{ww,k}\,d_{xx,k}\,{}^{b}t_{sw}\,{}^{b}t_{s'x}} = q_{rr'ss'}. \tag{14}$$

*Using the Cross-Product Ratio to find the form of* $^{A}\mathbf{T}$ *and* $^{B}\mathbf{T}$

Expanding (14), we obtain a sum of $R^4$ terms in the numerator and in the denominator of the ratio, some of which are identical in both places—ratio-matched—and some of which are not. In contrast to Theorem 1, here the d-weights do not always reflect the difference. Terms with either a fourth-power d-element or a cubed d-element always are ratio-matched, and those with four different d's never are. However, terms with one or two squared d's (i.e., when u = w and/or v = x or when u = x and/or v = w) are sometimes ratio-matched but more often are not.

*Example terms.* The terms differ from those in Theorem 1 because now no distinction is made between the d's arising from uv and those from wx, and so we again examine some specific ones. For example, uvwx = 1111 produces $d_{11,k}^4\,{}^{a}t_{r1}\,{}^{a}t_{r'1}\,h_{11}^2\,{}^{b}t_{s'1}\,{}^{b}t_{s1}$ in the numerator and denominator, and no other uvwx gives rise to the same d-weight. In contrast, uvwx = 2111, 1211, 1121, and 1112 all yield terms containing $d_{11,k}^3 d_{22,k}$. (Recall that two different d-weights resulted from these combinations in Theorem 1.) The terms are identical in the numerator and denominator, but as before, the last two occur in reverse order (i.e., the same one is derived from uvwx = 1121 in the top and uvwx = 1112 in the bottom, and vice versa).

Now consider the six permutations of 1122. The six corresponding terms all contain $d_{11,k}^2 d_{22,k}^2$, and the associated tt-hh-tt products are ratio-matched for uvwx = 1122 and 2211, but not for uvwx = 1212, 1221, 2112 and 2121. Similarly, of the twelve terms with a d-weight of $d_{11,k}^2 d_{22,k} d_{33,k}$ (from the permutations of 1123, assuming $R \geq 3$), only four are ratio-matched (for uvwx = 1123, 1132, 2311 and 3211). And, when $R \geq 4$, the d-weight $d_{11,k} d_{22,k} d_{33,k} d_{44,k}$ occurs in 24 terms in the top and bottom of the ratio (due to the 24 permutations of 1234) but none with the same tt-hh-tt product in both places.

*Factoring by d-weights.* All uvwx values in the top and bottom of (14) give rise to terms with similar d-weights to those described above, either (i) a d-element raised to the fourth power, that occurs in one term, (ii) two d-elements, one of which is cubed, that repeats in four terms, (iii) two squared d's, that occurs in six terms, (iv) three d's, one of which is squared, that repeats in 12 terms, or (v) four d's, that occurs in 24 terms. Only terms in the numerator whose d-weight contains a third- or fourth-power d-element can always be matched with identical ones in the denominator.

Factoring by d-weights, we combine four terms for each $d_{ii}^3 d_{jj}$ combination, six for each $d_{ii}^2 d_{jj}^2$, twelve for each one containing one squared d, and 24 for each combination of four different d's. This reduces the ratio to one with MAXS terms in the numerator and in the denominator, with all of the d-weight combinations in (7) represented. Now only terms with d-weights of the form $d_{ii}^4$ or $d_{ii}^3 d_{jj}$ are ratio-matched, the t-h-t coefficients of all the others differ between the top and bottom of the ratio. Note that the ratio-matched terms have relatively simple t-h-t coefficients—either one tt-hh-tt product or a sum of four—which is fortunate, since they play a key role in the following section.

We now reorder so that ratio-matched terms occur first, and rewrite the (factored) ratio to clarify its structure, as we did in Theorem 1. We define $[f_i(d_k)]$ as reordered component i of $\tau_k$ given in (7), and let $\alpha_i$ or $\beta_j$ represent the corresponding sum of tt-hh-tt products. Using $M = R^2$ as the number of terms that match in the numerator and denominator and $N = R(R-1)(R^2+7R-6)/24$ as the number of unmatched terms (i.e., M + N = MAXS), (14) becomes

$$(q-1)([f_1(d_k)]\alpha_1 + \cdots + [f_M(d_k)]\alpha_M) + [f_{M+1}(d_k)](q\beta_1 - \alpha_{M+1}) + \cdots + [f_{M+N}(d_k)](q\beta_N - \alpha_{M+N}) = 0. \quad (15)$$

Reasoning as in Theorem 1, in order for (15) to fulfill our linear independence assumptions with $q \neq 1$, we must have $\alpha_1, \ldots, \alpha_M = 0$; this allows us to find the form of the transformation matrices.

As in Theorem 1, we will work with $q \neq 1$ to determine the values in the transformation matrices. We again adopt a temporary convention that the diagonals of ${}^A\mathbf{T}$ and ${}^B\mathbf{T}$ are nonzero, and invoke the zero property of the $\alpha$-weights (t-h-t coefficients) for ratio-matched terms. This time however, we fix three subscripts when selecting the ratio.

*Filling in* ${}^A\mathbf{T}$ *and* ${}^B\mathbf{T}$. First we choose $q_{ijsi} \neq 1$ ($i \neq j$, $s \neq i$), which always exists since rows of $\tilde{\mathbf{H}}_k$ are nonproportional. We replace the fixed subscripts r, r′ and s′ in (14) with i, j, and i, respectively. Looking at the (ratio-matched) terms with d-weights of $d_{ii,k}^4$, $d_{ii,k}^3 d_{jj,k}$ and $d_{ii,k}^3 d_{ss,k}$, we can set the corresponding $\alpha$-weights to zero. From the first term (where uvwx = iiii), we get

$$ {}^a t_{ii}\, {}^a t_{ji}\, h_{ii}^2\, {}^b t_{ii}\, {}^b t_{si} = 0, \quad (16)$$

which implies

$$ {}^a t_{ji}\, {}^b t_{si} = 0. \quad (17)$$

The permutations of uvwx = iiij give us the four tt-hh-tt products that comprise the $\alpha$ of the second term. Two of these vanish upon substitution of (17), leaving

$$ {}^a t_{ii}\, {}^a t_{ji}\, h_{ii} h_{ij}\, {}^b t_{ii}\, {}^b t_{sj} + {}^a t_{ii}\, {}^a t_{jj}\, h_{ii} h_{ji}\, {}^b t_{ii}\, {}^b t_{si} = 0. \quad (18)$$

From this, we can conclude that ${}^a t_{ji} = 0$ implies ${}^b t_{si} = 0$.

Similarly, substituting (17) into the $\alpha$ derived from the permutations of iiis eliminates two of the tt-hh-tt products, leaving

$$ {}^a t_{ii}\, {}^a t_{ji}\, h_{ii} h_{is}\, {}^b t_{ii}\, {}^b t_{ss} + {}^a t_{ii}\, {}^a t_{js}\, h_{ii} h_{si}\, {}^b t_{ii}\, {}^b t_{si} = 0. \quad (19)$$

From this, we can conclude that ${}^b t_{si} = 0$ implies ${}^a t_{ji} = 0$. Combining these implications with (17), we conclude that ${}^a t_{ji} = 0$. Since this is true for every i, j ($i \neq j$), ${}^A\mathbf{T}$ is diagonal.

A parallel procedure shows ${}^B\mathbf{T}$ to be diagonal as well. As in Theorem 1, we use the nonproportionality of the columns of $\tilde{\mathbf{H}}_k$ to choose $q_{riij} \neq 1$ ($r \neq i$, $i \neq j$), and find that ${}^b t_{ji} = 0$ for every i, j ($i \neq j$).

*Relationship of* $\mathbf{A}$ *to* $\overset{*}{\mathbf{A}}$ *and* $\mathbf{B}$ *to* $\overset{*}{\mathbf{B}}$

Just as in Theorem 1, we now allow for nonzero elements off the diagonals, and hence a more general result, by defining a diagonal matrix ${}^{A}\mathbf{\Delta}$ and a permutation matrix ${}^{A}\mathbf{P}$ (either or both of which might be identity matrices) such that

$$ {}^{A}\mathbf{T} = {}^{A}\mathbf{P}\ {}^{A}\mathbf{\Delta}. \tag{20}$$

Similarly, for ${}^{B}\mathbf{T}$ we define a diagonal matrix ${}^{B}\mathbf{\Delta}$ and a permutation matrix ${}^{B}\mathbf{P}$ such that

$$ {}^{B}\mathbf{T} = {}^{B}\mathbf{P}\ {}^{B}\mathbf{\Delta}. \tag{21}$$

*Proof of (3) through (6)*

*Proof of (3) and (4).* Before we show that ${}^{A}\mathbf{P} = {}^{B}\mathbf{P} = \mathbf{P}$, as we must to prove (3) and (4), we first obtain results about $\mathbf{H}$ and the $\mathbf{D}_k$ matrices that do not require equality of ${}^{A}\mathbf{P}$ and ${}^{B}\mathbf{P}$. We proceed just as in Theorem 1, Equations (26) through (38) (replacing ${}^{A}\mathbf{D}_k$ and ${}^{B}\mathbf{D}_k$ with $\mathbf{D}_k$, of course), eventually obtaining two expressions for $\overset{*}{\mathbf{D}}_k$:

$$\overset{*}{\mathbf{D}}_k = z_k^{-1}\ {}^{A}\mathbf{P}\ \mathbf{D}_k\ {}^{A}\bar{\mathbf{\Delta}}\ {}^{A}\mathbf{P}' = z_k\ {}^{B}\mathbf{P}\ {}^{B}\bar{\mathbf{\Delta}}\ \mathbf{D}_k\ {}^{B}\mathbf{P}', \tag{22}$$

where $z_k$ is a scalar multiplier expressing the left-right scalar indeterminacy, as before (see (38) in Theorem 1).

Using the right hand equality in (22), we substitute appropriately from (28) of Theorem 1, commute diagonal matrices to simplify, and obtain

$$\mathbf{D}_k\mathbf{D}_1^{-1} = z_k^2\ {}^{A}\mathbf{P}'\ {}^{B}\mathbf{P}\ \mathbf{D}_k\ \mathbf{D}_1^{-1}\ {}^{B}\mathbf{P}'\ {}^{A}\mathbf{P}. \tag{23}$$

*Showing that* $z_k = \pm 1$. Suppose the permutation in (23) does not change the position of the u-th diagonal. Then for such an element we can write $d_{uu,k}d_{uu,1}^{-1} = z_k^2 d_{uu,k}d_{uu,1}^{-1}$, which implies $z_k = \pm 1$.

Now suppose the permutation changes the positions of *all* the diagonal elements. Then the element in position u is moved to position v, the one in v is moved to w, and so on, with some element being moved into location u. Applying (23), this chain of relocations results in a set of equations

$$d_{uu,k}d_{uu,1}^{-1} = z_k^2 d_{vv,k}d_{vv,1}^{-1},$$

$$d_{vv,k}d_{vv,1}^{-1} = z_k^2 d_{ww,k}d_{ww,1}^{-1},$$

$$\cdots$$

$$= z_k^2 d_{uu,k}d_{uu,1}^{-1}.$$

Substituting backwards from the last equation to the first gives $d_{uu,k}d_{uu,1}^{-1} = (z_k^2 z_k^2 \ldots z_k^2)d_{uu,k}d_{uu,1}^{-1}$, which implies $z_k = \pm 1$ in this case as well.

*Showing that* ${}^{A}\mathbf{P} = {}^{B}\mathbf{P}$. Assuming ${}^{A}\mathbf{P} \neq {}^{B}\mathbf{P}$, some diagonal elements of $(\mathbf{D}_k\ \mathbf{D}_1^{-1})$ will change position on the right hand side of (23). Suppose the diagonal in position u is permuted to position v. With $z_k^2 = 1$, from (23) we have $d_{uu,k}d_{uu,1}^{-1} = d_{vv,k}d_{vv,1}^{-1}$ which implies that for all MAXS values for k $d_{uu,k} = y\ d_{vv,k}$, where $y = d_{uu,1}d_{vv,1}^{-1}$. But this would cause some of the $\boldsymbol{\tau}_k$ to violate the linear independence required for "adequate variation." Thus we must conclude instead that no elements are permuted, and so ${}^{A}\mathbf{P} = {}^{B}\mathbf{P} = \mathbf{P}$.

We now rewrite (20) and (21) as

$$ {}^{A}\mathbf{T} = \mathbf{P}\ {}^{A}\boldsymbol{\Delta} \quad \text{and} \tag{24} $$

$$ {}^{B}\mathbf{T} = \mathbf{P}\ {}^{B}\boldsymbol{\Delta}. \tag{25} $$

Substituting into (8) and (9), respectively, then gives

$$ \overset{*}{\mathbf{A}}\ \mathbf{P}\ {}^{A}\boldsymbol{\Delta} = \mathbf{A} \quad \text{and} \tag{26} $$

$$ \overset{*}{\mathbf{B}}\ \mathbf{P}\ {}^{B}\boldsymbol{\Delta} = \mathbf{B}, \tag{27} $$

which proves (3) and (4) in the Theorem. □

*Proof of (5).* By substituting (24) and (25) into (10), we obtain

$$ \mathbf{P}\ {}^{A}\boldsymbol{\Delta}\ \mathbf{D}_k\ \mathbf{H}\ \mathbf{D}_k\ {}^{B}\boldsymbol{\Delta}\ \mathbf{P}' = \overset{*}{\mathbf{D}}_k\ \overset{*}{\mathbf{H}}\ \overset{*}{\mathbf{D}}_k. \tag{28} $$

Then, after choosing a level with full rank $\mathbf{D}_k$, say $k = 1$, pre- and post-multiplying $\overset{*}{\mathbf{H}}$ by $(\mathbf{PP}')$ and rearranging as before, we get

$$ \mathbf{H} = \mathbf{D}_1^{-1}(\mathbf{P}'\ \overset{*}{\mathbf{D}}_1\ \mathbf{P})\,{}^{A}\boldsymbol{\Delta}^{-1}\ \mathbf{P}'\ \overset{*}{\mathbf{H}}\ \mathbf{P}\ {}^{B}\boldsymbol{\Delta}^{-1}(\mathbf{P}'\ \overset{*}{\mathbf{D}}_1\ \mathbf{P})\mathbf{D}_1^{-1}. \tag{29} $$

Now if we define a single permuted diagonal matrix

$$ \bar{\boldsymbol{\Delta}} \equiv \mathbf{D}_1^{-1}\ \mathbf{P}'\ \overset{*}{\mathbf{D}}_1\ \mathbf{P}, \tag{30} $$

we can rewrite (29) as

$$ \mathbf{H} = (\bar{\boldsymbol{\Delta}}\ {}^{A}\boldsymbol{\Delta}^{-1}\ \mathbf{P}')\overset{*}{\mathbf{H}}(\mathbf{P}\ {}^{B}\boldsymbol{\Delta}^{-1}\ \bar{\boldsymbol{\Delta}}). \tag{31} $$

Thus (5) is proven. □

*Proof of (6).* Substituting (31) into (28) and simplifying, we now obtain

$$ \mathbf{P}\ \underline{\mathbf{D}_k\ \bar{\boldsymbol{\Delta}}}\ \mathbf{P}'\ \overset{*}{\mathbf{H}}\ \mathbf{P}\ \bar{\boldsymbol{\Delta}}\ \mathbf{D}_k\ \mathbf{P}' = \overset{*}{\mathbf{D}}_k\ \overset{*}{\mathbf{H}}\ \overset{*}{\mathbf{D}}_k. \tag{32} $$

Defining

$$ \dot{\mathbf{D}}_k \equiv \mathbf{P}\ \underline{\mathbf{D}_k}\ \bar{\boldsymbol{\Delta}}\ \mathbf{P}' \tag{33} $$

allows us to rewrite (32) as

$$ \dot{\mathbf{D}}_k\ \overset{*}{\mathbf{H}}\ \dot{\mathbf{D}}_k = \overset{*}{\mathbf{D}}_k\ \overset{*}{\mathbf{H}}\ \overset{*}{\mathbf{D}}_k, \tag{34} $$

from which we derive the simplified scalar equivalent

$$ \dot{d}_{uu,k}\ \dot{d}_{ww,k} = \overset{*}{d}_{uu,k}\ \overset{*}{d}_{ww,k} \text{ for } u, w = 1, \ldots, R. \tag{35} $$

From cases where $u = w$, we get $\dot{d}_{uu,k} = \pm\overset{*}{d}_{uu,k}$ and from cases where $u \neq w$, we get $\dot{d}_{uu,k}\ \overset{*}{d}{}^{-1}_{uu,k} = \dot{d}^{-1}_{ww,k}\ \overset{*}{d}_{ww,k}$ for $u, w = 1, \ldots, R$; hence

$$ \dot{\mathbf{D}}_k = z_k\ \overset{*}{\mathbf{D}}_k \tag{36} $$

where $z_k = \pm 1$. After equating (36) and (33) and rearranging terms, we obtain

$$\mathbf{D}_k = z_k \, \mathbf{P}' \, \overset{*}{\mathbf{D}}_k \, \mathbf{P} \, \bar{\mathbf{\Delta}}^{-1}. \tag{37}$$

This establishes (6), and the proof is complete. □

## Discussion

Theorem 1 establishes that the orientation of PARATUCK2 axes is sometimes uniquely determined. That is, the factor pattern (or stimulus projection) matrices **A** and **B** are determined up to rearrangement of column order and/or proportional increase/decrease (or reflection) of all the elements in a column. Such alterations do not generally change the interpretation of a factor or dimension. Some simple conventions determining order of factors and scaling of modes can be adopted to remove the remaining indeterminacy, in which case the model is fully identifiable—given that the assumptions are fulfilled and the data variation is "adequate." Theorem 2 demonstrates that the argument is easily adapted to handle models with equal ${}^{A}\mathbf{D}_k$ and ${}^{B}\mathbf{D}_k$. Together, these theorems prove the rotational uniqueness *for sufficiently large K* of several psychometrically interesting special cases such as PARAFAC2 or oblique INDSCAL, and three-way versions of DEDICOM.

### *Plausibility of assumptions*

*Assumption 1.* Full rank assumptions: These assumptions are standard with multilinear models, and easy to fulfill.

*Assumption 2.* All nonzero **H** assumption: For some types of data it is unlikely that any factor interactions would be exactly zero, and so the assumption is plausible. For other types of data it is probably too restrictive (and for skew-symmetric DEDICOM, the model *requires* zeroes on the diagonal of **H**). The all nonzero assumption is useful for the ratio-based argument used here, but the actual effects of zeros in **H** is not well understood. We have both empirical and mathematical results which show partial non-uniqueness of 2 × 2 PT2 with exactly one zero. However, in most cases that we have examined, zeros in **H** do not interfere with uniqueness. For example, we have empirical results that suggest uniqueness still holds in skew-symmetric DEDICOM, 3 × 3 nonsymmetric PT2 with one zero, 3 × 3 PARAFAC2 with four zeroes, and symmetrically weighted PT2 or PARAFAC2 with diagonal **H**. These results provide motivation to develop further proofs in which the all nonzero **H** assumption is relaxed.

*Assumption 3.* The "adequate variation" assumption: When the $\mathbf{D}_k$ elements are drawn randomly from a continuous distribution, violations of "adequate variation" should arise "with probability zero". However, there are situations (e.g., with certain externally specified variations in the $\mathbf{D}_k$) where more relaxed requirements would be useful, so alternatives should be investigated.

The definition of "adequate variation" may seem surprisingly complicated. The complexity arises partly because of the PT2 model itself and partly because of the cross-product rule. The model multiplies elements from two different $\mathbf{D}_k$ (or in symmetrically weighted cases, from the same $\mathbf{D}_k$ twice), while the cross-product ratio multiplies two entities derived from the model, thus combining four d-weights. Hence, the linear independence assumptions are stated in terms of all possible product combinations of four d-weights, two from ${}^{A}\mathbf{D}_k$ and two from ${}^{B}\mathbf{D}_k$ (or all from the same $\mathbf{D}_k$ in Theorem 2). Simpler assumptions, such as linear independence of the d variations, would block most but not all conditions that would allow nonuniqueness; the ones we use here are the only ones we have found so far that block all of the most subtle conditions.

In the current Theorems, "adequate variation" also requires that no dimensions have zero weights on MAX (or MAXS) of the levels showing independent variation. Such a restriction is not necessary for uniqueness. To drop it, however, would require adding another lengthy section to deal with various possible combinations of less than MAX full-rank levels and levels for which the $\mathbf{D}_k$ matrices show rank deficiencies of various kinds. This is better handled by a separate theorem and argument. We should point out, however, that the full rank case is the most difficult one, where the data at all levels are mixtures of all factors. The number of levels needed for uniqueness would generally be smaller when some factors have zero d-weights for some levels. Indeed, Harshman (letter to J. D. Carroll, 1972) pointed out that the PARAFAC2 dimensions were uniquely defined whenever there was, for each dimension, some level on which only that dimension had a nonzero weight. Thus, in such a case, only R levels of Mode C would be needed for a unique solution. Carroll and Wish (1974) subsequently refined this by noting that while **A** and **B** would be unique with R levels, **H** would be nonunique unless there were an (R + 1)-th level on which all dimensions had nonzero weights.

*Assumption 4.* The R = S assumption. This assumption is natural in a wide range of cases, such as PARAFAC2, but is sometimes undesirable. Recall that the PARATUCK2 model (like Tucker's T2 and T3) is more general, and allows a different number of columns for **A** versus **B**. In empirical tests, we have (as yet) found no counter examples to uniqueness when R = 3 with S = 2 and even when R = 5 with S = 2 (for data generated with random elements in all matrices); hence, we suspect that in at least some cases uniqueness may hold when R > S. However, ten Berge (p.c., 1995) has shown that a partial nonuniqueness will occur when two or more rows of **H** are proportional. (When R = S, such proportionality is ruled out by the full column rank of **H**, but when **R** > S it must be specifically ruled out by an additional assumption.) In itself, this is a modest assumption to add. But we have not yet been able to show that this would imply that $\overset{*}{\mathbf{H}}$ will also have no proportional rows, as required for the current form of the argument. Further study of the R > S case would seem worthwhile.

### *Theoretical and Practical Issues*

*Number of "occasions" (levels of Mode C) needed for uniqueness.* We do not necessarily consider MAX and MAXS to be the minimum number of levels for which uniqueness of the PT2 representation holds. Although we have not tried in these proofs to mathematically determine the minimum number, we have found empirically that in $2 \times 2$ cases with positivity constraints on the Mode C weights, the PARAFAC2 solution appears to be unique with as few as three levels of Mode C and the nonsymmetrically weighted PT2 model with six. Ten Berge and Kiers (1996), present a proof of uniqueness for PARAFAC2 when R = 2 which determines exactly the minimum number of levels needed for uniqueness in various conditions; it confirms that three is enough when Mode C positivity constraints are imposed, and shows that four is enough "with probability one" for real data. In general, the minimum number of levels needed for uniqueness of particular models at different dimensionalities is an issue that deserves further study, since as R increases our MAXS and particularly MAX values quickly become rather large. (Where R = S = 2, for example, MAX = 9 and where R = S = 4, MAX = 100.)

*Model fitting.* Widespread application of PARATUCK models may require further advances in model-fitting algorithms. When fit by alternating least squares, these models seem particularly subject to convergence difficulties (resembling the "swamps" described by Mitchell & Burdick, 1994; and sharing some characteristics of the "degeneracies" described in Kruskal, Harshman, & Lundy, 1989; Harshman & DeSarbo, 1984, etc.). In general, these are *not* local optima, and most of them are probably not true "degeneracies" (although degeneracies can also cause problems when the data contain certain kinds of systematic deviation from the model being fit). Many trials from different starting positions are sometimes needed to find a path to the optimum without severe convergence bottlenecks. It will be important to look for ways of avoiding these obstacles and preventing degeneracies when fitting PT2 and related models to data.

*Changing interactions.* One important kind of generalization that has not been proven unique by this proof is the case of proportional profiles when the central interaction matrix undergoes some (constrained) changes from one level of k to another. Consideration of one such case will be a natural consequence of developing the "PARATUCK3" model. In the PT3 model, **H** is replaced by $\sum_{t=1}^{T} {}^{g}c_{kt}\, \mathbf{G}_t$, a weighted sum of frontal slabs from a three way array $\underset{\sim}{\mathbf{G}}$. In Tucker's terminology, $\underset{\sim}{\mathbf{G}}$ would be the T3 "core matrix." The slab weights for level k are given in the k-th row of a K by T Mode C weight matrix ${}^{G}\mathbf{C}$. It would seem likely that with few enough slabs in the core, PT3 could have a unique solution.

There is, however, a different kind of model which has already been shown to have uniqueness despite some variation in the matrix of angles among factors or dimensions. A separate proof (Harshman, 1992)—using an indirect application of a theorem from Kruskal's (1976, 1977) analysis of PARAFAC/CANDECOMP uniqueness—shows uniqueness of axes for a three-mode model based on classical equations for hierarchical factor analysis; this model allows restricted changes in interactions or angles among factors when the higher order factors change their strength of influence across occasions. The nature of this model, and its basis for uniqueness, is somewhat different from any of the three-way factor models considered here.

*Exploring Tucker-type models.* Despite the limitations of this proof, it suggests some possibilities that have a certain appeal to the psychometric imagination. It should prove most interesting to investigate empirical domains where aspects of Tucker generality might be needed, to determine what patterns of factor interaction are discovered when one does not need to rely on more traditional rotation methods.

## References

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika, 35*, 283–319.

Carroll, J. D., & Wish, M. (1974). Models and methods for three-way multidimensional scaling. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology: Vol. 2* (pp. 57–105). San Francisco: W. H. Freeman.

Cattell, R. B. (1944). "Parallel Proportional Profiles" and other principles for determining the choice of factors by rotation. *Psychometrika, 9*, 267–283.

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences.* New York: Plenum.

Cattell, R. B., & Cattell, A. K. S. (1955). Factor rotation for proportional profiles: Analytical solution and an example. *British Journal of Statistical Psychology, 8*, 83–92.

Dawson, M. R. W., & Harshman, R. A. (1986). The multidimensional analysis of asymmetries in alphabetic confusion matrices: Evidence for global-to-local and local-to-global processing. *Perception & Psychophysics, 40*, 370–383.

de Leeuw, J., & Pruzansky, S. (1978). A new computational method to fit the weighted Euclidean distance model. *Psychometrika, 43*, 479–490.

Dunn, T. R., & Harshman, R. A. (1982). A multidimensional scaling model for the size-weight illusion. *Psychometrika, 47*, 25–45.

Harshman, R. A. (1970). Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis. *UCLA Working Papers in Phonetics, 16*, 1–84. (University Microfilms No. 10,085)

Harshman, R. A. (1972a). Determination and proof of minimum uniqueness conditions for PARAFAC1. *UCLA Working Papers in Phonetics, 22*, 111–117. (University Microfilms No. 10,085)

Harshman, R. A. (1972b). PARAFAC2: Mathematical and technical notes. *UCLA Working Papers in Phonetics, 22*, 30–44. (University Microfilms No. 10,085)

Harshman, R. A. (1992). *Unusual uniqueness properties of a three-way factor analysis model that incorporates "higher order" factors*. Unpublished manuscript.

Harshman, R. A., & DeSarbo, W. (1984). An application of PARAFAC to a small sample problem, demonstrating preprocessing, orthogonality constraints, and split-half diagnostic techniques. In H. G. Law, C. W. Snyder, Jr., J. A. Hattie, & R. P. McDonald (Eds.), *Research methods for multimode data analysis* (pp. 602–642). New York: Praeger.

Harshman, R. A., & Lundy, M. E. (1984). The PARAFAC model. In H. G. Law, C. W. Snyder, Jr., J. A. Hattie, & R. P. McDonald (Eds.), *Research methods for multimode data analysis* (pp. 122–215). New York: Praeger.

Harshman, R. A., & Lundy, M. E. (1994). PARAFAC: Parallel factor analysis. *Computational Statistics & Data Analysis, 18*, 39–72.

Harshman, R. A., Green, P. E., Wind, Y., & Lundy, M. E. (1982). A model for the analysis of asymmetric data in marketing research. *Marketing Science, 1*, 205–242.

Kroonenberg, P. M. (1983). *Three mode principal component analysis*. Leiden: DSWO Press.

Kroonenberg, P. M., & de Leeuw, J. (1980). Principal components analysis of three-mode data by means of alternating least-squares algorithms. *Psychometrika, 5*, 69–97.

Kruskal, J. B. (1976). More factors than subjects, tests than treatments: An indeterminacy theorem for canonical decomposition and individual differences scaling. *Psychometrika, 41*, 281–293.

Kruskal, J. B. (1977). Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra and Its Applications, 18*, 95–138.

Kruskal, J. B. (1981). Multilinear models for data analysis. *Behaviormetrika, 10*, 1–20.

Kruskal, J. B. (1989). Rank, decomposition, and uniqueness for 3-way and N-way arrays. In R. Coppi & S. Bolasco (Eds.), *Multiway data analysis* (pp. 7–18). Amsterdam: North Holland.

Kruskal, J. B., Harshman, R. A., & Lundy, M. E. (1989). How 3-MFA data can cause degenerate PARAFAC solutions, among other relationships. In R. Coppi & S. Bolasco (Eds.), *Multiway data analysis* (pp. 115–130). Amsterdam: North Holland.

Leurgans, S. E. (1991). *Multilinear models: Array formulas* (Technical Report 473). Columbus, OH: Ohio State University, Department of Statistics.

Leurgans, S. E., Ross, R. T., & Abel, R. B. (1993). A decomposition for 3-way arrays. *SIAM Journal on Matrix Analysis and Applications, 14*, 1064–1083.

Meredith, W. (1964). Notes on factorial invariance. *Psychometrika, 29*, 177–185.

Mitchell, B. C., & Burdick, D. S. (1994). Slowly converging PARAFAC sequences: Swamps and two-factor degeneracies. *Journal of Chemometrics, 8*, 155–168.

ten Berge, J. M. F., & Kiers, H. A. L. (1996). Some uniqueness results for PARAFAC2. *Psychometrika, 61*, 123–132.

Tucker, L. R. (1963). Implications of factor analysis of three-way matrices for measurement of change. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 122–137). Madison, Wisconsin: University of Wisconsin Press.

Tucker, L. R. (1964). The extension of factor analysis to three-dimensional matrices. In N. Frederiksen & H. Gulliksen (Eds.), *Contributions to mathematical psychology* (pp. 109–127). New York: Holt, Rinehart, & Winston.

Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika, 31*, 279–311.

Tucker, L. R. (1972). Relations between multidimensional scaling and three-mode factor analysis. *Psychometrika, 37*, 3–27.

*Manuscript received 10/20/92*
*Final version received 9/11/95*