Fachbereich II – Mathematik - Physik - Chemie

BEUTH HOCHSCHULE FÜR TECHNIK BERLIN
University of Applied Sciences

02/2016

Ulrike Grömping

Interaction Contributions as Coding Invariant Single Degree of Freedom Contributions to Generalized Word Counts

Interaction Contributions als kodierungs-invariante ein-Freiheitsgrad Beiträge zu verallgemeinerten Wort-Anzahlen

Reports in Mathematics, Physics and Chemistry
Berichte aus der Mathematik, Physik und Chemie

ISSN (print): 2190-3913

Reports in Mathematics, Physics and Chemistry

Berichte aus der Mathematik, Physik und Chemie

The reports are freely available via the Internet:
http://www1.beuth-hochschule.de/FB_II/reports/welcome.htm

02/2016, October 2016

Interaction Contributions as Coding Invariant Single Degree of Freedom Contributions to Generalized Word Counts
Interaction Contributions als kodierungs-invariante ein-Freiheitsgrad Beiträge zu verallgemeinerten Wort-Anzahlen (englischsprachig)

Update history
March 22 2016: Changed "Table 1" to "Table 2" at the bottom of p. 14, and changed "1" to "4" on p.16 directly above Example 4

Editorial notice / Impressum

Published by / Herausgeber:
Fachbereich II
Beuth Hochschule für Technik Berlin
Luxemburger Str. 10
D-13353 Berlin
Internet: http://public.beuth-hochschule.de/FB_II/
E-Mail: fbiireports@beuth-hochschule.de

Responsibility for the content rests with the author(s) of the reports.
Die inhaltliche Verantwortung liegt bei den Autor/inn/en der Berichte.

ISSN (print): 2190-3913

Interaction Contributions as Coding Invariant Single Degree of Freedom Contributions to Generalized Word Counts

Ulrike Grömping

Beuth University of Applied Sciences, Berlin

Abstract

This paper proposes “interaction contributions” (ICs), tabulated in interaction frequency tables, for a coding invariant single degree of freedom decomposition of generalized word counts for factorial designs with qualitative factors. The ICs are based on singular value decomposition and relate to bias contributions by interaction degrees of freedom on the estimation of the intercept. Fontana, Rapallo and Rogantin’s (2016) work on mean aberrations has given rise to this work. The paper introduces ICs and their tabulations and illustrates their behavior in various examples. Some of these compare the proposed ICs to mean aberration tables for situations where both are applicable.

Keywords: experimental design; qualitative factors; combinatorial equivalence; mean aberrations; generalized word length pattern

1. Introduction

Xu and Wu (2001) introduced the generalized word length pattern (GWLP) which is now widely used as the basis of generalized minimum aberration (GMA). For a design with n factors, the GWLP can be written as $(A_0,A_1,\ldots,A_n)$, where $A_0=1$ generally holds. For any $j>0$, let $\mathcal{S}_j = \{S \subseteq \{1,\ldots,n\}: |S|=j\}$ denote the set of all j-factor sets. A_j can be written as the sum of contributions $a_j(S)$ from all such sets, i.e. as $A_j = \sum_{S\in\mathcal{S}_j} a_j(S)$. The $a_j(S)$ are called projected a_j values in the sequel. In many applications, the GWLP is applied to orthogonal array designs, which implies that $A_1=A_2=0$, so that the first interesting entry is A_3. In this paper, $A_1=0$ is assumed (i.e., level balance of all factors), and the number R with $A_1=\ldots=A_{R-1}=0$ and $A_R>0$ is called the resolution of the design; this is in line with the conventional understanding of resolution (e.g. in Hedayat, Sloane and Stufken 1999 p.280) for $R\geq3$ and extends the concept to $R=2$, e.g. for supersaturated designs. Based on the projected a_3 values, Xu, Cheng and Wu (2004) proposed “minimum projection aberration” for resolution III designs with 3-level factors, and Schoen (2009) used frequency tables of projected a_3 values for the ranking of 18 run factorial designs.

Recently, Grömping and Xu (2014) introduced two decompositions of projected a_R values (with R the resolution)

- into sums of R^2 values from linear models for explaining all columns of an orthogonally coded main effects model matrix for one factor in the R factor set S by a full model in the other R–1 factors from S
- or into sums of squared canonical correlations between an arbitrarily coded main effects model matrix for one factor in the R factor set S and a full model matrix in the other R–1 factors from S.

Both decompositions work for full resolution sets S only; they are motivated by considering the impact of confounding in the R factor set S on estimation of each factor's main effects coefficients; Grömping (in press) investigated the use of tabulations of those decompositions (average R^2 frequency tables (ARFTs) and squared canonical correlation frequency tables (SCFTs)) for ranking designs and checking design equivalence.

Xu and Wu (2001) discussed the statistical meaning of the A_j in terms of the confounding between j-factor interactions and the overall mean. Following a similar logic, Fontana, Rapallo and Rogantin (2016) decomposed the projected a_j values into "aberrations" for individual interaction degrees of freedom, using the complex contrasts that were introduced by Bailey (1982). They attempted to render these decompositions coding invariant by permuting the levels of the individual interaction columns, which leads to tabulations of "mean aberrations"; however, their approach does not achieve coding invariance for factors with more than 3 levels, as can e.g. be verified by calculating the pattern of mean aberrations for the cyclic 5-level Latin square (their design $\mathcal{F}_2$) after swapping the first two levels for the third factor. Fontana et al.'s work motivated the author to develop the truly coding invariant decompositions of projected a_j values that are presented in this article. These are based on singular value decomposition (SVD); ambiguities arising from singular values with multiplicity larger than one are resolved in two different ways, which leads to two types of so-called interaction contribution frequency tables (ICFTs). The tabulated interaction contributions have a statistical interpretation in terms of bias contributions of the interaction to estimation of the overall mean, and the ICFTs can be used for assessing combinatorial equivalence of designs.

Section 2 of this paper introduces notation and basic concepts, and the brief Section 3 provides a fundamental coding invariance theorem. Section 4 develops the bias on the intercept from wrongly omitting the highest order interaction in a model with R factors based on a design of resolution R. Section 5 presents the decomposition results, relating them to previous insights on the bias. Section 6 provides some examples, and the final section discusses connections to further related work and reasonable future steps.

2. Notation and basic concepts

Before discussing the basics for factorial designs, some matrix products are defined and rules for them established. In the following, the superscript T denotes transposition, $\mathbf{1}_N$ and $\mathbf{0}_N$ denote column vectors of N ones or zeroes, respectively, and $\mathbf{e}_i$ denotes a unit vector with the value "1" in position i and zeroes everywhere else.

Definition 1 (matrix products)

(i) For an $m\times n$ matrix $\mathbf{A}$ and an $r\times s$ matrix $\mathbf{B}$, the Kronecker product is defined as the $mr\times ns$ matrix

$$\mathbf{A}\otimes\mathbf{B}=\begin{pmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{pmatrix}.$$

(ii) For an $n_a\times N$ matrix $\mathbf{A}$ and an $n_b\times N$ matrix $\mathbf{B}$, the column wise Khatri-Rao product is defined as the $n_a n_b\times N$ matrix

$$\mathbf{A}\odot_c\mathbf{B}=\left(\mathbf{a}_1\otimes\mathbf{b}_1 \quad \cdots \quad \mathbf{a}_N\otimes\mathbf{b}_N\right),$$

where $\mathbf{a}_i$, $\mathbf{b}_i$, i=1,…,N denote the i-th columns of $\mathbf{A}$ and $\mathbf{B}$, respectively, and $\otimes$ denotes the Kronecker product.

(iii) For an $N\times n_a$ matrix $\mathbf{C}$ and an $N\times n_b$ matrix $\mathbf{D}$, the row wise Khatri-Rao product is the transpose of the column wise Khatri-Rao product of their transposes:

$$\mathbf{C}\odot_r\mathbf{D}=\left(\mathbf{C}^\mathrm{T}\odot_c\mathbf{D}^\mathrm{T}\right)^\mathrm{T}=\left(\mathbf{c}^1\otimes\mathbf{d}^1 \quad \cdots \quad \mathbf{c}^N\otimes\mathbf{d}^N\right)^\mathrm{T},$$

where $\mathbf{c}^i$ and $\mathbf{d}^i$ denote the transposed i-th rows of matrices $\mathbf{C}$ and $\mathbf{D}$, respectively.

(iv) For two $m\times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, the Hadamard or Schur or element wise product is defined as

$\mathbf{A}*\mathbf{B} = (a_{ij}\, b_{ij})_{i=1,\ldots,m,\, j=1,\ldots,n}$.

Lemma 1

For an $N\times n_a$ matrix $\mathbf{A}$ and an $N\times n_b$ matrix $\mathbf{B}$:

(i) $\left(\mathbf{A}\odot_r\mathbf{B}\right)\left(\mathbf{A}\odot_r\mathbf{B}\right)^\mathrm{T}=\left(\mathbf{A}\mathbf{A}^\mathrm{T}\right)*\left(\mathbf{B}\mathbf{B}^\mathrm{T}\right).$

(ii) $\left(\mathbf{A}\odot_r\mathbf{B}\right)^+=\left(\mathbf{A}\odot_r\mathbf{B}\right)^\mathrm{T}\left(\left(\mathbf{A}\mathbf{A}^\mathrm{T}\right)*\left(\mathbf{B}\mathbf{B}^\mathrm{T}\right)\right)^+,$

where the superscript "+" denotes the Moore Penrose inverse.

Lemma 1 (i) follows from $\left(\mathbf{A}\odot_c\mathbf{B}\right)^\mathrm{T}\left(\mathbf{A}\odot_c\mathbf{B}\right)=\left(\mathbf{A}^\mathrm{T}\mathbf{A}\right)*\left(\mathbf{B}^\mathrm{T}\mathbf{B}\right)$, which is a known result for the column wise Khatri Rao product and the Hadamard product (see e.g. Kolda and Bader 2009, Section 2.6), by applying it to $\mathbf{A}^\mathrm{T}$ and $\mathbf{B}^\mathrm{T}$ instead of $\mathbf{A}$ and $\mathbf{B}$. Lemma 1 (ii) follows from the Moore Penrose inverse for the column wise Khatri Rao product given in Kolda and Bader, or by checking the well-known Moore-Penrose conditions.

We now consider a factorial design with n factors in N runs. The i-th factor has s_i levels, i=1,…,n, which occur equally often, i.e. all factors are level-balanced. The model matrix is given as

$$\mathbf{M} = (\mathbf{M}_0\ \mathbf{M}_1\ \mathbf{M}_2\ \ldots\ \mathbf{M}_n)$$

with $\mathbf{M}_0=\mathbf{1}_N$, $\mathbf{M}_1$ the matrix of all main effects model matrices $\mathbf{X}_1, \ldots, \mathbf{X}_n$, $\mathbf{M}_2$ the matrix of all 2-factor interaction model matrices $\mathbf{X}_{\mathcal{I}(\{1,2\})}, \ldots, \mathbf{X}_{\mathcal{I}(\{n-1,n\})}$, and so forth, with $\mathbf{X}_{\mathcal{I}(S)}$ denoting the interaction model matrix for the interaction involving all factors from set $S\subseteq\{1,\ldots,n\}$.

Definition 2 (normalized orthogonal coding)

The model matrix **M** is said to be in normalized orthogonal coding, if

(i) the columns of $\mathbf{X}_i$ have mean 0, are orthogonal to each other and have squared length N,

(ii) for $S \in \mathcal{S}_j$, the interaction matrix $\mathbf{X}_{\mathcal{I}(S)}$ is the row wise Khatri-Rao product of the j main effects model matrices $\mathbf{X}_i$, $i \in S$.

Lemma 2

If $\mathbf{X}_i$ and $\tilde{\mathbf{X}}_i$ are both $N \times (s_i - 1)$ main effects model matrices in normalized orthogonal coding for factor i, there is an orthogonal $(s_i - 1) \times (s_i - 1)$ matrix **Q** such that

$$\tilde{\mathbf{X}}_i = \mathbf{X}_i\mathbf{Q} \Leftrightarrow \tilde{\mathbf{X}}_i\mathbf{Q}^{\mathrm{T}} = \mathbf{X}_i\,.$$

Lemma 2 is obvious from noting that different orthogonal bases for the main effect with all columns of the same squared length N can only be obtained from each other by rotation and reflection operations. Note that results generalize to complex coding by changing the transpose to conjugate transpose.

Lemma 3

For $S \in \mathcal{S}_j$, $N^2 a_j(S) = \mathbf{1}_N^{\mathrm{T}}\mathbf{X}_{\mathcal{I}(S)}\mathbf{X}_{\mathcal{I}(S)}^{\mathrm{T}}\mathbf{1}_N$.

Lemma 3 follows directly from Xu and Wu's (2001) definition of the A_j and is stated as a lemma only for ease of reference.

Finally, the paper makes use of SVD: an $m\times n$ matrix **A** can be written as $\mathbf{UDV}^{\mathrm{T}}$ with matrices **U** ($m\times m$) and **V** ($n\times n$) such that $\mathbf{U}^{\mathrm{T}}\mathbf{U}=\mathbf{UU}^{\mathrm{T}}=\mathbf{I}_m$ and $\mathbf{V}^{\mathrm{T}}\mathbf{V}=\mathbf{VV}^{\mathrm{T}}=\mathbf{I}_n$, and an $m\times n$ diagonal matrix **D** of min(m,n) singular values $\zeta_i = \zeta_i(\mathbf{A}) \geq 0$. The columns of **U** and **V** are called left and right singular vectors, respectively. The non-zero squared singular values coincide with the non-zero eigen values of the positive semidefinite matrices $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$. If all singular values are distinct, the first min(m,n) columns of matrices **U** and **V** are unique, up to sign switches of corresponding column pairs $\mathbf{u}_i$ and $\mathbf{v}_i$. Ambiguities can arise from multiple singular values of the same size, which lead to non-unique groups of singular vectors: if $N \times r$ sub matrices $\mathbf{U}_{sub}$ and $\mathbf{V}_{sub}$ correspond to identical singular values, these can be replaced by the pair $\mathbf{L}_{sub}$ and $\mathbf{M}_{sub}$ with $\mathbf{L}_{sub} = \mathbf{U}_{sub}\mathbf{Q}$ and $\mathbf{M}_{sub} = \mathbf{V}_{sub}\mathbf{Q}$ with any suitable $r \times r$ matrix **Q** for which $\mathbf{Q}^T\mathbf{Q}=\mathbf{Q}\mathbf{Q}^T=\mathbf{I}_r$.

3. Coding invariance

With the tools from the previous section, it is straightforward to identify coding invariant matrices related to a linear model in normalized orthogonal coding (see Def. 2 above and equations (1) and (2) below) that will later be used for deriving a coding-invariant decomposition of the projected a_j values: Theorem 1 shows the coding invariance of $\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^T$; subsequently, a unique decomposition of $a_j(S)$ will be derived based on the eigen value decomposition of that matrix, or – equivalently – based on the SVD of the matrix $\mathbf{X}_{\mathcal{J}(S)}$. A corollary to the theorem shows that some aspects of the SVD are coding invariant, while others depend on the coding. Note that, in spite of also using singular values, the approach of the present paper is quite different from the proposal by Katsaounis, Dean and Jones (2013) of using singular values for checking design equivalence for 2-level designs.

<u>Theorem 1</u>

The matrix $\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^T$ does not depend on the choice of normalized orthogonal coding.

Proof: According to Lemma 1, $\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^T$ can be written as the Schur product of matrices $\mathbf{X}_i\mathbf{X}_i^T$, $i \in S$, with $\mathbf{X}_i$ a main effects model matrix in normalized orthogonal coding. Because of Lemma 2, $\tilde{\mathbf{X}}_i\tilde{\mathbf{X}}_i^T = \mathbf{X}_i\mathbf{X}_i^T$ for $\tilde{\mathbf{X}}_i$ with an arbitrary choice of normalized orthogonal coding for factor i. ///

<u>Corollary</u>

For the SVD $\mathbf{X}_{\mathcal{J}(S)}=\mathbf{U}\mathbf{D}\mathbf{V}^T$, the matrices **U** and **D** do not depend on the choice of normalized orthogonal coding, while **V** depends on that choice.

Proof: $\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^T = \mathbf{U}\mathbf{D}\mathbf{V}^T\mathbf{V}\mathbf{D}^T\mathbf{U}^T = \mathbf{U}\mathbf{D}\mathbf{D}^T\mathbf{U}^T$ is the eigen value decomposition of the coding invariant matrix $\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^T$; $\mathbf{X}_{\mathcal{J}(S)}^T\mathbf{X}_{\mathcal{J}(S)} = \mathbf{V}\mathbf{D}^T\mathbf{U}^T\mathbf{U}\mathbf{D}\mathbf{V}^T = \mathbf{V}\mathbf{D}^T\mathbf{D}\mathbf{V}^T$ is the eigen value decomposition of the coding dependent matrix $\mathbf{X}_{\mathcal{J}(S)}^T\mathbf{X}_{\mathcal{J}(S)}$. ///

4. The bias of the intercept estimate

Consider a resolution R design with R factors, $R \geq 2$. The linear model in normalized orthogonal coding can be written as

$$E(\mathbf{Y}) = \mu + \sum_{i=1}^{R} \mathbf{X}_i \boldsymbol{\beta}_i + \sum_{\substack{S \subseteq \{1,\ldots,R\}, \\ |S| \geq 2}} \mathbf{X}_{\mathcal{J}(S)} \boldsymbol{\beta}_{\mathcal{J}(S)} \tag{1}$$

with $\mathbf{Y}$ denoting the random $N\times 1$ vector of response values.

<u>Remark 1</u>

The model matrices in equation (1) depend on the choice of normalized orthogonal coding, the entire summands do not, since the corresponding coefficient vectors are modified accordingly. This is also true for interaction model matrices, which are row-wise Khatri-Rao products: denoting the interaction model matrix with changed coding of one or more factors in S as $\mathbf{W}_{\mathcal{J}(S)}$, the corresponding parameter vector as $\gamma_{\mathcal{J}(S)}$, and the Moore-Penrose inverse of $\mathbf{W}_{\mathcal{J}(S)}$ as $\mathbf{W}_{\mathcal{J}(S)}^{+}=\mathbf{W}_{\mathcal{J}(S)}^{\mathrm{T}}(\mathbf{W}_{\mathcal{J}(S)}\mathbf{W}_{\mathcal{J}(S)}^{\mathrm{T}})^{+}$, we get $\gamma_{\mathcal{J}(S)}=\mathbf{W}_{\mathcal{J}(S)}^{+}\mathbf{X}_{\mathcal{J}(S)}\boldsymbol{\beta}_{\mathcal{J}(S)}= \mathbf{W}_{\mathcal{J}(S)}^{\mathrm{T}}(\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^{\mathrm{T}})^{+}\mathbf{X}_{\mathcal{J}(S)}\boldsymbol{\beta}_{\mathcal{J}(S)}$, and $\mathbf{W}_{\mathcal{J}(S)}\gamma_{\mathcal{J}(S)}= \mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^{\mathrm{T}}(\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^{\mathrm{T}})^{+}\mathbf{X}_{\mathcal{J}(S)}\boldsymbol{\beta}_{\mathcal{J}(S)} = \mathbf{X}_{\mathcal{J}(S)}\boldsymbol{\beta}_{\mathcal{J}(S)}$ because of Theorem 1 and the properties of Moore Penrose inverses. Furthermore, note that for $\mathbf{X}_{\mathcal{J}(S)} = \mathbf{UDV_X}^{\mathrm{T}}$, $\mathbf{W}_{\mathcal{J}(S)} = \mathbf{UDV_W}^{\mathrm{T}}$, i.e., the SVDs only differ by the matrices of right singular vectors (see the corollary to Theorem 1).

Now, assume that model (1) is correct and we wrongly fit the smaller model

$$E(\mathbf{Y}) = \mu + \sum_{i=1}^{R} \mathbf{X}_i \boldsymbol{\beta}_i + \sum_{\substack{S \subseteq \{1,\ldots,R\}, \\ 2 \leq |S| \leq R-1}} \mathbf{X}_{\mathcal{J}(S)} \boldsymbol{\beta}_{\mathcal{J}(S)}\,, \tag{2}$$

omitting the highest order interaction. The estimate for μ is the average of the $\mathbf{Y}$ components, with expectation $\mu + \mathbf{1}_N^{\mathrm{T}}\mathbf{X}_{\mathcal{J}(\{1,\ldots,R\})}\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}/N$, i.e., bias $\mathbf{1}_N^{\mathrm{T}}\mathbf{X}_{\mathcal{J}(\{1,\ldots,R\})}\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}/N$. Note that, because of the design's resolution, the omission of main effects or lower order interactions with the any factor would not bias the intercept estimate, i.e., the bias would remain the same if we would, e.g., omit an entire factor instead of omitting only the R-factor interaction. Of course, this bias strongly depends on the sizes of the unknown coefficients in $\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}$.

As was already pointed out by Xu and Wu (2001), $a_R(\{1,\ldots,R\})$ is an indicator of the bias for the intercept from the R-factor interaction. In an overall way, according to Lemma 3, $a_R(\{1,\ldots,R\})$ is the sum of squares of the multipliers $\mathbf{1}_N^{\mathrm{T}}\mathbf{X}_{\mathcal{J}(\{1,\ldots,R\})}/N$ with which the unknown interaction parameters in $\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}$ enter the bias; it is thus a Frobenius norm and as such provides an upper bound for the sum of squares of the bias vector: $||\mathbf{1}_N^T\mathbf{X}_{\mathcal{J}(\{1,\ldots,R\})}\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}/N||_2^2 \le a_R(\{1,\ldots,R\})\ ||\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}||_2^2$. The bound is exact for the worst-case $\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}$ which is collinear to $\mathbf{X}_{\mathcal{J}(\{1,\ldots,R\})}^T\mathbf{1}_N$ (because this is collinear to the first eigen vector of matrix $\mathbf{X}_{\mathcal{J}(\{1,\ldots,R\})}^T\mathbf{1}_N\mathbf{1}_N^T\mathbf{X}_{\mathcal{J}(\{1,\ldots,R\})}$). That worst-case $\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}$ depends on the coding, while the worst-case bias does not. Denote the matrices of right singular vectors for two different codings $\mathbf{X}_{\mathcal{J}(\{1,\ldots,R\})}$ and $\mathbf{W}_{\mathcal{J}(\{1,\ldots,R\})}$ as $\mathbf{V}_\mathbf{X}$ and $\mathbf{V}_\mathbf{W}$, respectively. If we write $\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}=\mathbf{V}_\mathbf{X}\mathbf{c}$ and $\boldsymbol{\gamma}_{\mathcal{J}(\{1,\ldots,R\})}=\mathbf{V}_\mathbf{W}\mathbf{c}$, these two vectors correspond to the same contribution to the expected value of $\mathbf{Y}$ in (1) and to the same bias for the intercept in model (2), i.e. the coding invariant way of representing the parameter vector is through the vector $\mathbf{c}$ of linear combination coefficients for the right singular vectors:

$$\mathbf{1}_N^T\mathbf{X}_{\mathcal{J}(\{1,\ldots,R\})}\mathbf{V}_\mathbf{X}\mathbf{c}/N = \mathbf{1}_N^T\mathbf{W}_{\mathcal{J}(\{1,\ldots,R\})}\mathbf{V}_\mathbf{W}\mathbf{c}/N = \mathbf{1}_N^T\mathbf{U}\mathbf{D}\mathbf{c}/N = \sum_{i=1}^{\min(N,\mathrm{df}(\{1,\ldots,R\}))} c_i\zeta_i\bar{u}_i$$

with $\bar{u}_i$ the average of the i-th column of matrix $\mathbf{U}$. Consequently, in terms of $\mathbf{c}$, the squared bias of $\mathbf{1}_N^T\mathbf{Y}/N$ as an estimator for μ can be written as

$$\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}^T\mathbf{X}_{\mathcal{J}(\{1,\ldots,R\})}^T\mathbf{1}_N\mathbf{1}_N^T\mathbf{X}_{\mathcal{J}(\{1,\ldots,R\})}\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}/N^2 = \left(\sum_{i=1}^{\min(N,\mathrm{df}(\{1,\ldots,R\}))} c_i\zeta_i\bar{u}_i\right)^2. \tag{3}$$

In considerations like this, it is customary to consider length 1 vectors; here, $\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})} = \mathbf{V}\mathbf{c}$ is normalized to length 1, if and only if $\mathbf{c}^T\mathbf{c}=||\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}||_2^2=1$. The simplest such vectors are those with $\mathbf{c}=\mathbf{e}_i$, for which $\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}$ is exactly the i-th right singular vector of $\mathbf{X}_{\mathcal{J}(\{1,\ldots,R\})}$ (depending on the choice of normalized orthogonal coding), and the squared bias with this choice of $\boldsymbol{\beta}_{\mathcal{J}(\{1,\ldots,R\})}$ becomes $(\zeta_i\bar{u}_i)^2$ (independent of the choice of normalized orthogonal coding). The next section will show that these $(\zeta_i\bar{u}_i)^2$ yield a coding invariant decomposition of $a_R(\{1,\ldots,R\})$. The case of singular values with multiplicity $r > 1$ will be discussed separately, since such singular values imply non-unique singular vectors (see Section 5.2).

5. Coding invariant decomposition of $a_j(S)$ and its relation to the bias

This section points out, how a coding invariant decomposition of $a_j(S)$ can be obtained, and how this relates to the bias of the intercept estimate from wrongly omitting the highest order interaction $\mathcal{J}(S)$.

5.1. The decomposition

According to Lemma 3, $N^2a_j(S) = \mathbf{1}_N^T\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^T\mathbf{1}_N/N^2 = \mathbf{1}_N^T\mathbf{U}\mathbf{D}\mathbf{D}^T\mathbf{U}^T\mathbf{1}_N/N^2$. Denoting as $\bar{\mathbf{u}}$ the row vector of column averages of $\mathbf{U}$ and as $\mathbf{D}^2$ the diagonal matrix of squared non-zero singular values, augmented with zeroes as needed, this can be written as $\bar{\mathbf{u}}\mathbf{D}^2\bar{\mathbf{u}}^\mathrm{T}$, or in the form given in Theorem 2.

Theorem 2

Let $\zeta_i=\zeta_i(\mathbf{X}_{\mathcal{J}(S)})$ denote the i-th singular value of the matrix $\mathbf{X}_{\mathcal{J}(S)}$, $\bar{u}_i$ the column average of the corresponding i-th left singular vector, and df(S) the degrees of freedom from the interaction $\mathcal{J}$(S).

(i) Then the projected a_j value $a_j(S)$ can be decomposed as

$$a_j(S) = \sum_{i=1}^{\min(N,df(S))} \left(\zeta_i \bar{u}_i\right)^2 . \tag{4}$$

(ii) If all non-zero singular values have multiplicity one, the decomposition is unique.

(iii) Assuming there is at least one non-zero singular value ζ_i with multiplicity $r_i>1$ and corresponding $N\times r_i$ matrix $\mathbf{U}_{\mathrm{sub},i}$ of left singular vectors, the decomposition (4) is unique if and only if $\mathbf{1}_N^\mathrm{T}\mathbf{U}_{\mathrm{sub},i}=\mathbf{0}_{r_i}^\mathrm{T}$ for all such pairs ζ_i and $\mathbf{U}_{\mathrm{sub},i}$.

Proof:

ad (i): The decomposition is obvious from the properties of the SVD:

If $\mathbf{X}_{\mathcal{J}(S)} = \mathbf{UDV}^\mathrm{T}$, then $\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^\mathrm{T} = \mathbf{UDD}^\mathrm{T}\mathbf{U}^\mathbf{T} = \mathbf{UD}^2\mathbf{U}^\mathrm{T}$, with $\mathbf{D}^2$ defined as the diagonal matrix of squared singular values ζ_i^2 (augmented with zeroes, if necessary). It follows that

$$a_j(S) = \frac{1}{N^2}\mathbf{1}_N^\mathrm{T}\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^\mathrm{T}\mathbf{1}_N = \sum_{i=1}^{\min(N,df(S))} \zeta_i^2 \left(\frac{1}{N}\sum_{b=1}^{N} u_{bi}\right)^2 = \sum_{i=1}^{\min(N,df(S))} \left(\zeta_i \bar{u}_i\right)^2$$

ad (ii): If all non-zero singular values are unique, all corresponding columns of the matrix $\mathbf{U}$ are unique up to sign changes; sign changes do not affect the squared column averages.

ad (iii): A non-zero singular value ζ_i with multiplicity $r_i>1$ has a corresponding $N\times r_i$ matrix $\mathbf{U}_{\mathrm{sub},i}$ of left singular vectors whose columns are non-unique, as they can be rotated or reflected in arbitrary ways; however, if $\mathbf{1}_N^\mathrm{T}\mathbf{U}_{\mathrm{sub},i}=\mathbf{0}_{r_i}^\mathrm{T}$, the same is also true for all rotated versions $\mathbf{L}_{\mathrm{sub},i}\mathbf{Q}$, i.e. $\mathbf{1}_N^\mathrm{T}\mathbf{L}_{\mathrm{sub},I} = \mathbf{0}_{r_i}^\mathrm{T}\,\mathbf{Q} = \mathbf{0}_{r_i}^\mathrm{T}$. Thus, all the corresponding summands in (4) are zero, regardless of the choice of columns. If this is the case for all matrices of left-singular vectors corresponding to non-zero singular values with multiplicity $r_i > 1$, (4) yields a unique decomposition. Otherwise, the decomposition will change, depending on the arbitrary choice of left singular vectors. ///

Corollary

(i) It follows from Theorem 2 and equation (3) that the summand $(\zeta_i\bar{u}_i)^2$ of (4) is interpretable as the squared bias for the intercept from the interaction $\mathcal{J}(S)$ in case $\beta_{\mathcal{J}(S)}=\mathbf{v}_i$.

(ii) For cases with non-unique summands (see Theorem 2 (iii)), the statement from (i) holds for any possible combination of left and right singular vectors.

Definition 3

(i) For a set $S \in \mathcal{S}_j$ with interaction model matrix $\mathbf{X}_{\mathcal{J}(S)}$, the terms $(\zeta_i \bar{u}_i)^2$, i=1,…, df(S) are called the interaction contributions for the set. For N < df(S), the last df(S) – N interaction contributions are defined as zeroes.

(ii) For an entire design in $n \geq j$ factors, the interaction contributions of all j-factor sets $S \in \mathcal{S}_j$ are called the interaction contributions of order j.

The interaction contributions of Definition 3 are coding invariant, but may be non-unique in cases with several identical singular values.

5.2. Resolving ambiguities

In the following, two different but related ways of obtaining unique summands in equation (4) for ambiguous cases are presented, namely a *concentrated* way (*conc*) that concentrates the entire sum of all ambiguous summands from a particular singular value ζ with multiplicity r in a single summand and leaves r-1 zero summands, and an *even* way that distributes the sum evenly over r summands.

Remark 2 (rotations)

In case of r identical singular values ζ for $\mathbf{X}_{\mathcal{J}(S)}$, denoting the $N \times r$ matrix $\mathbf{U}_{\text{sub}}$ and the df(S)$\times r$ matrix $\mathbf{V}_{\text{sub}}$ as the corresponding columns of matrices $\mathbf{U}$ and $\mathbf{V}$, consider an orthogonal $r \times r$ matrix $\mathbf{Q}$ ($\mathbf{Q}^\text{T}\mathbf{Q}=\mathbf{Q}\mathbf{Q}^\text{T}=\mathbf{I}_r$) with $\mathbf{L}_{\text{sub}} = \mathbf{U}_{\text{sub}}\mathbf{Q}$ and $\mathbf{M}_{\text{sub}} = \mathbf{V}_{\text{sub}}\mathbf{Q}$.

(i) $\bar{\mathbf{l}} := \mathbf{1}_N^\text{T}\mathbf{L}_{\text{sub}}/N = \mathbf{1}_N^\text{T}\mathbf{U}_{\text{sub}}\mathbf{Q}/N =: \bar{\mathbf{u}}\mathbf{Q}$,

where $\bar{\mathbf{l}}$ and $\bar{\mathbf{u}}$ denote 1×r vectors of column means of $\mathbf{L}_{\text{sub}}$ and $\mathbf{U}_{\text{sub}}$, respectively.

(ii) $\bar{\mathbf{l}} = \mathbf{0}_r$ if and only if $\bar{\mathbf{u}} = \mathbf{0}_r$.

(iii) For $\bar{\mathbf{u}} \neq \mathbf{0}_r$, the total contribution $\zeta^2 \sum_{i=1}^{r} \bar{l}_i^2 = \zeta^2 \bar{\mathbf{l}}^\text{T}\bar{\mathbf{l}} = \zeta^2 \bar{\mathbf{u}}^\text{T}\bar{\mathbf{u}}$ to (4) is unaffected by the choice of $\mathbf{Q}$, while the individual summands can (strongly) depend on $\mathbf{Q}$.

(iv) $\mathbf{X}_{\mathcal{J}(S),\text{sub}} = \zeta \mathbf{U}_{\text{sub}}\mathbf{V}_{\text{sub}}^\text{T} = \zeta \mathbf{L}_{\text{sub}}\mathbf{M}_{\text{sub}}^\text{T}$ is the coding invariant $N \times$df(S) summand of $\mathbf{X}_{\mathcal{J}(S)}$ that corresponds to the singular space for the singular value ζ.

Remark 3 (the concentrated case)

Consider the case of r identical singular values for $\mathbf{X}_{\mathcal{J}(S)}$, applying all notations as given in Remark 2.

(i) $\mathbf{Q}$ can be chosen such that $\bar{\mathbf{l}} = \|\bar{\mathbf{u}}\|_2\, \mathbf{e}_1^{\mathrm{T}}$; this is called the concentrated case. It can be obtained by determining $\mathbf{Q}=\mathbf{H}^{\mathrm{T}}$ with $\mathbf{H}$ the Householder transformation matrix that changes the direction of $\frac{\bar{\mathbf{u}}^{\mathrm{T}}}{\|\bar{\mathbf{u}}\|_2}$ to collinearity with $\mathbf{e}_1$.

The corresponding summands of (4) are $\zeta^2 \|\bar{\mathbf{u}}\|_2^{\,2}$ and r–1 zeroes.

(ii) For the coefficient vector $\boldsymbol{\beta}_{\mathcal{J}(S)}$, the bias contribution of the group is

$$\tfrac{1}{N}\mathbf{1}_N^{\mathrm{T}}\mathbf{X}_{\mathcal{J}(S),\mathrm{sub}}\boldsymbol{\beta}_{\mathcal{J}(S)} = \zeta\,\bar{\mathbf{l}}\,\mathbf{M}_{\mathrm{sub}}^{\mathrm{T}}\boldsymbol{\beta}_{\mathcal{J}(S)} = \zeta\,\|\bar{\mathbf{u}}\|_2\,\mathbf{m}_{\mathrm{sub},1}^{\mathrm{T}}\boldsymbol{\beta}_{\mathcal{J}(S)},$$

where $\mathbf{m}_{\mathrm{sub},1}$ is the first right singular vector of the concentrated rotation $\mathbf{M}_{\mathrm{sub}}=\mathbf{V}_{\mathrm{sub}}\mathbf{H}^{\mathrm{T}}$.

(iii) Among length 1 vectors $\boldsymbol{\beta}_{\mathcal{J}(S)}$, $\boldsymbol{\beta}_{\mathcal{J}(S)}=\pm\mathbf{m}_{\mathrm{sub},1}$ maximizes the squared bias contribution of the group, and the maximum is $\zeta^2\|\bar{\mathbf{u}}\|_2^{\,2}$.

The even case does the opposite of the concentrated case: instead of concentrating the entire contribution on one degree of freedom, it distributes it as evenly as possible. The rotation matrix for achieving the even distribution can be obtained by making use of a rectangular r-simplex, which is the generalization of a tri-rectangular tetrahedron to r dimensions: the r “legs” (=edges neighboring the apex) all meet at right angles at the apex; for the special case used for the even case, the apex is the origin $\mathbf{0}_r$, and all legs have length 1, which implies that the base of the r-simplex is equilateral, e.g., an equilateral triangle in case r=3. The altitude of such a rectangular r-simplex has the same angle acos(1/√r) to all legs and has length 1/√r, and the average of all legs equals the altitude. The following remark will be based on a matrix $\mathbf{R}$ whose columns consist of the r vertices except for the apex; these also define the r legs of length 1, and they yield an orthogonal matrix. That matrix is initially chosen such that the base point of the altitude is the point 1/√r $\mathbf{e}_1$ (implying that this is also the average of the columns of $\mathbf{R}$), and then rotated such that the altitude becomes collinear with the vector $\bar{\mathbf{u}}^{\mathrm{T}}$. Note that there are infinitely many possible matrices $\mathbf{R}$ belonging to a rectangular r-simplex with length 1 legs, apex $\mathbf{0}_r$ and altitude collinear to $\mathbf{e}_1$, because there are infinitely many possible rotations. This implies that – though the even interaction contributions are unique – the corresponding pairs of singular vectors are not.

Remark 4 (the even case)

For the case of r identical singular values for $\mathbf{X}_{\mathcal{J}(S)}$, apply all notations as given in Remark 2.

(i) $\mathbf{Q}$ can be chosen such that $\bar{\mathbf{l}} = \frac{1}{\sqrt{r}}\|\bar{\mathbf{u}}\|_2 \, \mathbf{1}_r^{\mathrm{T}}$; this is called the even case. It can be obtained by multiplying the inverse $\mathbf{H}^{\mathrm{T}}$ of the Householder transformation matrix that changes the direction of $\frac{\bar{\mathbf{u}}^{\mathrm{T}}}{\|\bar{\mathbf{u}}\|_2}$ to collinearity with $\mathbf{e}_1$ (see Remark 3) from the left with an orthogonal matrix $\mathbf{R}$ composed of the length 1 legs of a rectangular r-simplex with apex $\mathbf{0}_r$ and altitude collinear to $\mathbf{e}_1$. This achieves an equal angle of $\bar{\mathbf{u}}^{\mathrm{T}}$ to all columns of $\mathbf{Q}$, and an equal angle of $\mathbf{1}_N$ to all columns of $\mathbf{L}_{\mathrm{sub}}=\mathbf{U}_{\mathrm{sub}}\mathbf{Q} = \mathbf{U}_{\mathrm{sub}}\mathbf{H}^{\mathrm{T}}\mathbf{R}$.

(ii) For the corresponding coefficient vector $\boldsymbol{\beta}_{\mathcal{J}(S)}$, the bias contribution of the group is

$\frac{1}{N}\mathbf{1}_N^{\mathrm{T}}\mathbf{X}_{\mathcal{J}(S),\mathrm{sub}}\boldsymbol{\beta}_{\mathcal{J}(S)} = \zeta\bar{\mathbf{l}}\,\mathbf{M}_{\mathrm{sub}}^{\mathrm{T}}\boldsymbol{\beta}_{\mathcal{J}(S)} = \frac{\zeta}{\sqrt{r}}\|\bar{\mathbf{u}}\|_2\,\mathbf{1}_r^{\mathrm{T}}\mathbf{M}_{\mathrm{sub}}{}^{\mathrm{T}}\boldsymbol{\beta}_{\mathcal{J}(S)}$.

(iii) Among length 1 vectors $\boldsymbol{\beta}_{\mathcal{J}(S)}$, any vector $\boldsymbol{\beta}_{\mathcal{J}(S)} = \frac{1}{\sqrt{r}}\mathbf{M}_{\mathrm{sub}}\,\mathrm{diag}_r(\pm 1)\mathbf{1}_r$ maximizes the squared bias contribution of the group, and the maximum is $\zeta^2\|\bar{\mathbf{u}}\|_2^2$; $\mathrm{diag}_r(\pm 1)$ denotes a diagonal matrix whose diagonal elements are arbitrary combinations of "+1" or "-1" values.

Remark 5

(i) For the case of r identical singular values for $\mathbf{X}_{\mathcal{J}(S)}$, apply all notations as given in Remark 2, and add suffixes c and e for the concentrated and even case, respectively. The group's normalized average right singular vector $\frac{1}{\sqrt{r}}\mathbf{M}_{\mathrm{sub,e}}\mathbf{1}_r$ of the even case (see Remark 4 (iii), where all permissible sign changes are included through the diagonal matrix) coincides with the group's first right right singular vector $\mathbf{m}_{\mathrm{sub,c,1}}$ of the concentrated case (see Remark 3).

(ii) Part (i) of the remark holds for all corresponding pairs of $\mathbf{M}_{\mathrm{sub,e}}$ and $\mathbf{m}_{\mathrm{sub;c,1}}$; note in particular that non-identity diagonal matrices in Remark 4 (iii) correspond to a modified $\mathbf{M}_{\mathrm{sub,e}}$ with an identity matrix instead, implying modified $\mathbf{L}_{\mathrm{sub,e}}$, $\mathbf{M}_{\mathrm{sub,c}}$ and $\mathbf{L}_{\mathrm{sub,c}}$, and the identical maximum $\zeta^2\|\bar{\mathbf{u}}\|_2^2$.

Proof of part (i): Let $\mathbf{H}$ denote the Householder transformation of Remarks 3 and 4, $\mathbf{R}$ a matrix derived from a rectangular r-simplex as stated in Remark 4. Then, according to Remarks 3 and 4,

$\mathbf{L}_{\mathrm{sub,c}} = \mathbf{U}\mathbf{H}^{\mathrm{T}}$, $\qquad \mathbf{L}_{\mathrm{sub,e}} = \mathbf{U}\mathbf{H}^{\mathrm{T}}\mathbf{R} = \mathbf{L}_{\mathrm{sub,c}}\mathbf{R}$,

$\mathbf{M}_{\mathrm{sub,c}} = \mathbf{V}\mathbf{H}^{\mathrm{T}}$, $\qquad \mathbf{M}_{\mathrm{sub,e}} = \mathbf{V}\mathbf{H}^{\mathrm{T}}\mathbf{R} = \mathbf{M}_{\mathrm{sub,c}}\mathbf{R}$.

Furthermore, the coincidences

$$\mathbf{L}_{\text{sub,e}}\mathbf{1}_r/\sqrt{r} = \mathbf{l}_{\text{sub,c,1}},$$

and $$\mathbf{M}_{\text{sub,e}}\mathbf{1}_r/\sqrt{r} = \mathbf{m}_{\text{sub,c,1}}$$

result from the fact that $\mathbf{l}_{\text{sub,c,1}}=\mathbf{L}_{\text{sub,c}}\mathbf{e}_1$ and $\mathbf{m}_{\text{sub,c,1}}= \mathbf{M}_{\text{sub,c}}\mathbf{e}_1$, together with the fact that $\mathbf{R}\mathbf{1}_r/\sqrt{r}$ is $\sqrt{r}$ times the average of the legs of the r-simplex that defined $\mathbf{R}$; this average is the altitude $1/\sqrt{r}\ \mathbf{e}_1$. ///

5.3. Interaction contribution frequency tables

The interaction contributions of Definition 3 lend themselves to tabulation and can be used for assessing how the bias depends on the vector of interaction coefficients as well as for distinguishing non-isomorphic qualitative fixed level designs with the same projected a_j values. It will be most interesting to consider such tables for projected a_R values, with R the resolution of the design. Contrary to the decomposition results from Grömping and Xu (2014), however, this decomposition works for projected a_j values with arbitrary j; the statistical interpretation as a bias contribution works as well, if it is acknowledged that this is not the only contribution towards the bias of the estimate for μ.

Definition 4

The table of the $\left(\zeta_i\,\bar{u}_i\right)^2$ obtained from all sets $S \in \mathcal{S}_j$ – with uniqueness enforced as indicated in Remarks 3 or 4, if necessary – is called the Interaction Contribution Frequency Table of order j, or ICFT_j; it comes in the versions $\text{ICFT}_{j,\text{conc}}$ (with conc short for concentrated) and $\text{ICFT}_{j,\text{even}}$.

6. Examples

This section gives several examples. Where possible (i.e., for symmetric designs), mean aberrations by Fontana et al. (2016) are calculated in addition to $\text{ICFT}_{\text{conc}}$ and $\text{ICFT}_{\text{even}}$. For some smaller designs, the worst case interaction parameter vectors (i.e. the first right singular vectors of the concentrated rotation) are also given. Some of the designs are at least "generalised regular" or "Abelian group regular" (see Kobilinski, Monod and Bailey in press; Grömping and Bailey 2016). They are simply called "regular" in the sequel.

The first worked example uses the design given in Table 2 of Grömping and Xu (2014), which is given again here for convenience.

Table 1: Example design in two 2-level factors and one 4-level factor

A	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1
C	0	2	1	3	3	1	2	0

Example 1: $A_3 = a_3(\{1,2,3\} = 1$ for the regular design of Table 1 (the AB interaction is completely confounded with the 0/2 vs 1/3 contrast of factor C). Factors A and B are coded as -1/+1 (unique normalized orthogonal coding), factor C is coded with normalized Helmert coding. With S={1,2,3}, the coding-dependent matrix $\mathbf{X}_{\mathcal{J}(S)}$ and its coding invariant cross product $\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^{\mathrm{T}}$ are given as

$$\mathbf{X}_{\mathcal{J}(S)} = \begin{pmatrix} -\sqrt{2} & -\sqrt{2/3} & -\sqrt{1/3} \\ 0 & \sqrt{8/3} & -\sqrt{1/3} \\ -\sqrt{2} & \sqrt{2/3} & \sqrt{1/3} \\ 0 & 0 & -\sqrt{3} \\ 0 & 0 & -\sqrt{3} \\ -\sqrt{2} & \sqrt{2/3} & \sqrt{1/3} \\ 0 & \sqrt{8/3} & -\sqrt{1/3} \\ -\sqrt{2} & -\sqrt{2/3} & -\sqrt{1/3} \end{pmatrix}, \quad \mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^{\mathrm{T}} = \begin{pmatrix} 3 & -1 & 1 & 1 & 1 & 1 & -1 & 3 \\ -1 & 3 & 1 & 1 & 1 & 1 & 3 & -1 \\ 1 & 1 & 3 & -1 & -1 & 3 & 1 & 1 \\ 1 & 1 & -1 & 3 & 3 & -1 & 1 & 1 \\ 1 & 1 & -1 & 3 & 3 & -1 & 1 & 1 \\ 1 & 1 & 3 & -1 & -1 & 3 & 1 & 1 \\ -1 & 3 & 1 & 1 & 1 & 1 & 3 & -1 \\ 3 & -1 & 1 & 1 & 1 & 1 & -1 & 3 \end{pmatrix}.$$

The non-zero eigen values of $\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^{\mathrm{T}}$, equal to the non-zero squared singular values of $\mathbf{X}_{\mathcal{J}(S)}$, are $\zeta_1^2= \zeta_2^2= \zeta_3^2=8$, i.e. there are three non-unique pairs of singular vectors. With the SVD algorithm used in R for Windows, the initial squared column means of the matrix **U** are 1/48, 1/16 and 1/24, respectively; the contributions to $a_3(S)$ are thus 8 times these values, i.e., 1/6, 1/2 and 1/3. Concentrating the entire overlap on the first vector, $\mathrm{ICFT}_{3,\mathrm{conc}}$ shows two zeroes and one "1" from these non-zero singular values. Distributing the overlap evenly, $\mathrm{ICFT}_{3,\mathrm{even}}$ shows three 1/3 values instead. The right singular vector related to the only non-zero singular value in the concentrated case is $\mathbf{v}_1=(-\sqrt{1/2}, \sqrt{1/6}, -\sqrt{1/3})^{\mathrm{T}}$, i.e. with this coding, the largest bias on the intercept resulting from the three factor interaction occurs for coefficient vectors proportional to this $\mathbf{v}_1$. The even vectors are non-unique, but their normalized average coincides with the above $\mathbf{v}_1$.

Example 2: Consider a regular design in 9 runs with three 3-level factors and an interaction model matrix as given in Example 1 of Grömping and Xu (2014), for which $A_3 = a_3(\{1,2,3\} = 2$. With S={1,2,3}, the coding invariant cross product $\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^{\mathrm{T}}$ is given as

$$\mathbf{X}_{\mathcal{J}(S)}\mathbf{X}_{\mathcal{J}(S)}^{\mathrm{T}} = \begin{pmatrix} 8 & 2 & 2 & 2 & -1 & 2 & 2 & 2 & -1 \\ 2 & 8 & 2 & 2 & 2 & -1 & -1 & 2 & 2 \\ 2 & 2 & 8 & -1 & 2 & 2 & 2 & -1 & 2 \\ 2 & 2 & -1 & 8 & 2 & 2 & 2 & -1 & 2 \\ -1 & 2 & 2 & 2 & 8 & 2 & 2 & 2 & -1 \\ 2 & -1 & 2 & 2 & 2 & 8 & -1 & 2 & 2 \\ 2 & -1 & 2 & 2 & 2 & -1 & 8 & 2 & 2 \\ 2 & 2 & -1 & -1 & 2 & 2 & 2 & 8 & 2 \\ -1 & 2 & 2 & 2 & -1 & 2 & 2 & 2 & 8 \end{pmatrix}.$$

The non-zero eigen values of $\mathbf{X}_{\mathcal{I}(S)}\mathbf{X}_{\mathcal{I}(S)}^T$, equal to the non-zero squared singular values of $\mathbf{X}_{\mathcal{I}(S)}$, are ζ_1^2=18, ζ_2^2= ζ_3^2=…=ζ_7^2=9, ζ_8^2=0, i.e. there are two unique and six non-unique pairs of singular vectors. In this case,

$\bar{u}_1 = -1/3,$

$\bar{u}_2 = ... = \bar{u}_8 = 0.$

Thus, the non-uniqueness of the second to seventh pairs of singular vectors is irrelevant (see Remark 2 (ii)), and we obtain a unique $ICFT_3$ that shows one interaction contribution 18/9=2 and seven zeroes. The right singular vector corresponding to the non-zero contribution is proportional to (√3 -1 -1 -√3 1 √3 √3 -1), i.e. the most harmful coefficient vectors in terms of bias for the intercept are proportional to this vector for the coding used. For this symmetric 3-level design, the mean aberrations by Fontana et al. (2016) are well-defined and coding invariant; they consist of two ones and six zeroes.

Table 2: Two resolution II designs d_1=(A,B_1) and d_2=(A,B_2) in two 4-level factors (transposed)

A	0	0	1	1	2	2	3	3
B_1	0	1	2	3	0	1	2	3
B_2	0	1	2	3	0	3	1	2

Example 3: Table 2 shows the two non-isomorphic GMA designs for two 4-level factors in 8 runs; both have A_2=1, and the first one is regular. They can be distinguished by their $ICFT_{2,even}$ (d_1 has three "1/3" and 6 zeroes, d_2 has five "1/5" and four zeroes) but not by their $ICFT_{2,conc}$ (both have one "1" and eight zeroes). The worst case length 1 parameter vectors are again obtainable as the first right singular vectors (not shown). The mean aberrations are three "1/3"s and six zeroes for d_1 and four "1/12"s, two "1/6"s and one "1/3" for d_2 for the level allocations given in Table 2; after swapping levels 1 and 2 in factor A, the mean aberrations change to one "1/3", four "1/6"s and four zeroes for d_1 and to three "1/3"s and six zeroes for d_2.

Example 3 underlines the fact that mean aberrations for more than three levels depend on the level coding. This is also the case for the two 5-level designs investigated by Fontana et al. For both of these, $ICFT_{even}$ and $ICFT_{conc}$ coincide with each other, and the two unique ICFTs are also identical (63 zeroes and one "4" each).

Example 4: The $ICFT_3$ for the classical L18 is given as

	0	1/6	1/2	2/3	1	2
ICFTconc	320	0	28	9	6	1
ICFTeven	287	36	40	0	0	1

Note that from concentrated to even ICFT, the 1 entries turn into 1/2 entries, while the 2/3 entries turn into 1/6 entries. As this is a mixed level design, the mean aberrations by Fontana et al. cannot be calculated.

Example 5: Table 3 gives $ICFT_{3,conc}$ and $ICFT_{3,even}$ as well as the mean aberrations by Fontana et al. for the three non-isomorphic L18 with seven 3-level factors. Here, the mean aberrations are well-defined and do not depend on level allocation. Note that, from concentrated to even ICFT, the “1/3” entries turn into “1/12” entries, and the “1” entries turn into “1/2” entries. From even ICFT to mean aberration, the “2” entries become twice as many “1” entries, the “1/2” entries become twice as many “1/4” entries, and the “1/12” entries remain unchanged. Thus, mean aberrations are closer to the even than to the concentrated ICs. In this context, note that the $ICFT_3$ for the unique GMA design for six 3-level factors in 18 runs is unique and consists of 120 zeroes and 20 “1/2”s, while there are 100 zero mean aberrations and 40 “1/4”s, which is completely analogous.

Table 3: ICFTs and mean aberration frequency tables (MAFT) for the three non-isomorphic orthogonal 18 run designs in seven 3-level factors

		0	1/12	1/4	1/3	1/2	1	2
1	ICFTconc	227	0	0	36	16	0	1
	ICFTeven	119	144	0	0	16	0	1
	MAFT	102	144	32	0	0	2	0
2	ICFTconc	233	0	0	24	20	2	1
	ICFTeven	159	96	0	0	24	0	1
	MAFT	134	96	48	0	0	2	0
3	ICFTconc	245	0	0	0	28	6	1
	ICFTeven	239	0	0	0	40	0	1
	MAFT	198	0	80	0	0	2	0

Example 6: For the twelve 3-level columns of the Taguchi 36 run design (see NIST / Sematech 2016), the unique $ICFT_3$ and the mean aberration frequency table (MAFT) are given as follows:

value	0	0.0625	0.125	0.201	0.25	0.4375	0.5	0.674
ICFT	1524	0	192	16	0	0	12	16
MAFT	1320	384	0	24	0	32	0	0

Again, some mean aberrations are halves of ICs, like in the previous example, while the 16 0.4375 mean aberrations are obtained as means of the 16 ICs of 0.201 and the 16 ICs of 0.674.

7. Discussion

This paper has introduced interaction contributions and the related ICFTs for a new coding invariant single degree of freedom decomposition of generalized word counts A_j according to Xu and Wu (2001). Like the mean aberrations by Fontana et al., these decompositions yield as many contributions as there are degrees of freedom for j-factor interactions. These decompositions have the potential to distinguish designs, for which projected a_j values are the same. They are independent of coding and level allocation, contrary to the mean aberrations, for which this property holds for designs with up to three levels only. However, calculation times for ICFTs are much worse than those of mean aberration tables, at least in the author's implementation. Thus, it would be helpful if coding invariant mean aberrations could be extended to designs with factors at more than three levels.

Grömping and Xu (2014) and Grömping (in press) considered decompositions of projected a_R values (R the resolution) based on the relation of main effects to R–1 factor interactions. In terms of design quality, the impact of R–1 factor interactions on main effect estimation is much more interesting than the impact of R factor interactions on estimation of the overall mean. Thus, ICFTs are not particularly interesting as a criterion for design quality. They may, however, be useful for distinguishing non-isomorphic designs with identical projected a_j values. Grömping and Bailey (2016) defined the lenient regularity criterion "CC regularity" based on the squared canonical correlation decomposition derived in Grömping (in press). So far, a compelling relation of interaction contributions to design regularity has not been found; however, it cannot entirely be precluded; for example, it is conceivable that regular designs only have integers in the list of concentrated interaction contributions.

For the cases for which the mean aberrations are uniquely defined, the even ICFT values seem to be related to the mean aberrations; the examples – especially Example 6 – have shown that the relation is not always straightforward. It will be interesting to investigate the connections between the two concepts; ideally, a better understanding of these may contribute to developing a coding invariant version of mean aberrations for dimensions larger than three, which could become a very useful tool in checking combinatorial equivalence of factorial designs.

References

Bailey, R.A. (1982). The decomposition of treatment degrees of freedom in quantitative factorial experiments. JRSS **B** 44, 63-70.

Fontana, R., Rapallo, F. and Rogantin, M.P. (2016). Aberration in qualitative multilevel designs. *Journal of Statistical Planning and Inference* **174**, 1-10.

Grömping, U. (2016). R Package DoE.base for Factorial Designs. Report 1/2016, *Reports in Mathematics, Physics and Chemistry*, Department II, Beuth University of Applied Sciences Berlin.

Grömping, U. (2016, online early). Frequency tables for the coding invariant quality assessment of factorial designs. *IIE Transactions*, DOI 10.1080/0740817X.2016.1241458.

Grömping, U. and Bailey, R.A. (2016). Regular fractions of factorial arrays. In: Kunert, J., Müller, C.H., Atkinson, A.C. (Eds.), MODA11—Advances in Model-Oriented Design and Analysis. Springer, pp. 143–151.

Grömping, U. and Xu, H. (2014). Generalized resolution for orthogonal arrays. *The Annals of Statistics* **42**, 918-939.

Hedayat, S., Sloane, N.J. and Stufken, J. (1999). *Orthogonal Arrays: Theory and Applications*. Springer, New York.

Katsaounis, T.I., Dean, A.M. and Jones, B. (2013). On equivalence of fractional factorial designs based on singular value decomposition. *Journal of Statistical Planning and Inference* **11**, 1950-1953.

Kobilinski, A., Monod, H. and Bailey, R.A. (in press). Automatic generation of generalised regular factorial designs. To appear in *Computational Statistics and Data Analysis*.

Kolda, T.G. and Bader, B.W. (2009). Tensor Decompositions and Applications. *SIAM Review* **51**, 455-500.

NIST/SEMATECH (2016). *e-Handbook of Statistical Methods*, http://www.itl.nist.gov/div898/handbook/, 21 August 2016.

Schoen, E. (2009). All orthogonal arrays with 18 runs. *Quality and Reliability Engineering International* **25** (3), 467-480.

Xu, H., Cheng, S.-W. and Wu, C.F.J. (2004). Optimal projective three-level designs for factor screening and interaction detection. *Technometrics* **46**, 280–292. MR2082498

Xu, H. and Wu, C.F.J. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs [corrected republication of MR1863969]. *Ann. Statist.* **29**, 1066–1077. MR1869240