Received: 12 September 2019 | Revised: 14 February 2020 | Accepted: 19 February 2020

DOI: 10.1002/cem.3232

RESEARCH ARTICLE

WILEY Journal of CHEMOMETRICS

LiMM-PCA: Combining ASCA$^+$ and linear mixed models to analyse high-dimensional designed data

Manon Martin | **Bernadette Govaerts**

Institute of Statistics, Biostatistics and Actuarial Sciences (ISBA), Louvain Institute of Data Analysis and Modeling in Economics and Statistics (LIDAM), UCLouvain, Voie du Roman Pays 20, bte L1.04.01, Louvain-la-Neuve, B-1348, Belgium

Correspondence
Manon Martin, Institute of Statistics, Biostatistics and Actuarial Sciences (ISBA), Louvain Institute of Data Analysis and Modeling in economics and statistics (LIDAM), UCLouvain, Voie du Roman Pays 20, bte L1.04.01, Louvain-la-Neuve, B-1348, Belgium.
Email: manon.martin@uclouvain.be

Abstract

Nowadays, life science experiments—and especially "omics" fields—often imply a high volume of information from high throughput technologies that is gathered in the form of a wide and short multivariate response. These data are intrinsically correlated and generally produced by another multivariate set of factors or continuous variables, collected in what is defined as the design matrix. Such design factors usually involve the presence of a treatment, but other sources of biological or technical variability in the data are often measured as well. The ASCA framework, based on ANOVA and PCA, leads to promising results. By combining dimension reduction projection methods and classic statistical modelling, it enables to decipher the main sources of variability in the produced response and offers attractive graphical representations of the factors' effect. However, this approach has not yet been extended to more advanced designs involving random factors, being typically involved in longitudinal, hierarchical, or repeatability/reproducibility studies. This paper has its roots in the GLM version of ASCA, called ASCA$^+$, that leads to unbiased estimators of the factors' effects for unbalanced data. It is here extended by replacing GLM by LMM and adapting the methodology. Taking into account the error structure of the data indeed leads to more accurate data modelling and more generalisable results. The suggested methodology is applied to two experimental case studies that highlight the benefits of this approach as it leads to a refined data analysis with interesting inferential properties, while keeping the powerful visualisation outputs produced by ASCA.

KEYWORDS

ASCA, chemometrics, linear mixed models, PCA, random effects

1 | INTRODUCTION

1.1 | Analysis of designed experiments with high dimensional responses

In the field of experimental studies in life sciences, and especially for the "omics" sciences (eg, genomics, transcriptomics, and metabolomics), a large amount of information is commonly produced and results in multivariate quantitative data tables with an explosive number of highly correlated variables, often much larger than the number of samples available. Besides, in a large portion of these studies, the data are acquired on the basis of a multifactorial experimental design, where there is a need to decipher all the factors' impact on the generated multivariate dataset. The inclusion of possibly

advanced data structures (such as nesting, crossing, or unbalance), either related to the objectives or to the constraints of the study, ensures to model the data as accurately and completely as possible. Both aspects—high-dimensional correlated variables and their advanced experimental design structure—are important aspects that need to be taken into account during their analysis.

Principal component analysis (PCA) and partial least squares regression (PLS) are two popular multivariate dimension reduction techniques that are particularly suited to analyse such correlated and high-dimensional data: They simultaneously model the relationship between the variables and reduce the dataset dimensionality. Both methods provide powerful visualisation tools to ease the interpretation of the results. However, PCA is unsupervised; hence, there is no insurance that it will give a clear insight on the influence of the factors from the design if unwanted nuisance factors are the main sources of variability. By contrast, PLS[1] is a supervised method and is, in addition, built on the basis of the correlation between two blocks of data. But it needs to be tuned for model complexity to prevent overfitting. It is also essentially meant for very simple experimental designs, often including only one factor of interest. Furthermore, for both of these multivariate methods, statistical properties and inferential possibilities are limited, apart from the jack-knife approach by Martens et al.[2]

Regarding the advanced design issue, analysis of variance (ANOVA)[3,4] is a particularly well-established and reliable tool for the analysis of a univariate quantitative response. MANOVA[5] is its direct multidimensional generalisation that tests the significance of the factors on a multivariate response matrix. However, the latter cannot handle data tables with more variables than observations because of the covariance matrix singularity. Moreover, it provides no further interpretation tools and its different related test statistics can give conflicting results.[6]

Given the lack of tools to deal with such designed high-dimensional collinear data, several methods, that are typically based on a combination of statistical modelling and multivariate analysis, have emerged from the literature and are currently receiving the attention they deserve. Originally applied to metabolomic data, ANOVA–simultaneous components analysis (ASCA)[7,8] was developed as another multivariate generalisation of ANOVA to study the impact of a combination of fixed categorial predictors on a multivariate response. It combines the advantages of statistical modelling (ANOVA) and multivariate analysis (PCA). The first step of ASCA applies ANOVA in parallel on each response vector to decompose the response matrix into a sum of effect matrices. Then, the influence of each model effect on the multivariate response is analysed by applying PCA on the effect matrices. Concurrently developed on proteomic data, ANOVA-PCA (APCA)[6] differs from ASCA in the multivariate step, when adding the residuals to the effect matrices prior to the PCA.

Several other methods, similar or derived from ASCA, are also available, all based on ANOVA and each with specific features. PC-ANOVA[9,10] that permutes PCA and ANOVA, AoV-PLS[11] and ANOVA-TP[12] are two other alternatives that, respectively, apply PLS and PLS with target projection (TP) after the modelling step. Depending on the method, different PLS response and predictor matrices are built in the multivariate step. Additionally, multivariate methods designed to globally analyse all the effect matrices have been proposed, such as PARAFASCA,[13] AComDim,[14] or AMOPLS.[15] See Guisset et al[16] for a comparison between those methods.

Most of these integrated approaches were developed to study classic fixed and crossed factors designs. But in the case of particular design structures (nested factors, multilevel, repeated measurements), the researchers should consider the need to rely on more advanced modelling possibilities. Exceptions in the literature include Marini et al[12] with ANOVA–Target Projection (ANOVA-TP) and Luciano and Næs[10] with ASCA and PC-ANOVA that both applied mixed ANOVA with multiple crossed or nested factors in the presence of interactions. ASCA variants for multilevel data have also been developed by Timmerman[17] with the multilevel SCA (MSCA) which was applied to a time-resolved metabolomics dataset[18] and by de Noord and Theobald[19] with the multilevel PLS (MLPLS).

However, inferential properties of such models (eg, effect tests) and their potential generalisations to more advanced design structures are not dealt in depth within these papers. On the other hand, extensions of ASCA to unbalanced designs have been developed by Stanimirova et al[20] based on type III sum of squares (SS) and, more generally, by Thiel et al,[21] who suggests the use of the general linear model (GLM) notation and subsequent adaptations as an alternative to ANOVA-based effect matrices estimation. This approach, called ASCA$^+$, has the important advantage to offer a general framework to treat all possible experimental designs involving fixed (categorical) factors.

As a complement to ANOVA and GLM, the linear mixed model (LMM)[22-24] has nowadays become the gold standard statistical tool to model advanced balanced or unbalanced designs that includes fixed but also *random* effects. An effect is usually considered as fixed when the differences between its levels are of direct interest (eg, treatment effect), and we want to compare their mean effect. In contrast, the values of a random effect are considered as unreproducible random realisations of a larger population (eg, the patient effect). They affect the error structure of the data and are able to model complex dependence structures present in them. This class of modelling is extensively used for designs where observations

cannot be considered as independent between each other. In this context, the LMM is therefore an ideal candidate to generalise the modelling step of ASCA-related methodologies to more advanced designs.

1.2 | Motivations and objectives

Accordingly, this work aims to suggest a generic framework to extend the ASCA$^+$ methodology to LMM, called LiMM-PCA, taking into account the possible unbalance of the data as suggested in Snijders and Bosker.[21] The entire framework has to be adapted consequently and new methodological aspects developed to enable the comparison of both fixed and random sources of variation (test the significance of the effects, effect matrix augmentation, etc). In this paper, we will focus on an extension of ASCA, but the elaborated procedure could be applied to other multivariate methods, such as those mentioned in the background section hereabove.

In the field of chemometrics, this methodology has a direct and broad range of potential applications. For example, a repeatability/reproducibility study could indicate which of the measured biological or technical sources of variability are prominent and affect the measured signal; one could also be interested in longitudinal data where the dynamics of a system need to be integrated in the modelling process. Other potential and typical applications would be cross-over studies, eg, where several treatments are administered consecutively to a same group of subjects.

1.3 | Content

After the contextualisation and rationale of this research, Section 2 will describe the initial ASCA$^+$ methodology, followed by a reminder of the basics of LMM in this context in Section 3. Section 4 will then describe our suggested LiMM-PCA methodology with emphasis on new developments and adjustments compared with ASCA, and Section 5 will apply LiMM-PCA to two typical applications before closing by concluding remarks and perspectives in Section 6.

2 | ASCA$^+$ AND RELATED METHODS

The general scheme of ASCA is summarised in broad terms here. In a first step, the original response data matrix $\boldsymbol{Y}$ is decomposed into effect matrices $\hat{\boldsymbol{M}}_f$ according to a model linked to the experimental design. In the standard ASCA/APCA method, an ANOVA model is used, but this approach can be generalised using a GLM framework. In a second step, the decomposed effect matrices are analysed using a dimension reduction multivariate approach (PCA). In order to give it a concise presentation, this section is limited to elements that are necessary for its generalisation to mixed models thereafter.

2.1 | Fixed effects ANOVA in the GLM notation

Since fixed effects ANOVA can be seen as a special case of linear regression, the GLM extension of ASCA, called ASCA$^+$,[21] will serve here as a basis for further generalisations. The main motivation is that ANOVA estimators are biased in the case of an unbalanced design, and the total variance decomposition does not anymore produce orthogonal SS summing to total SS (SS_{Total}). GLM thus provides a general solution to analyse unbalanced and designed data with fixed effects which is equivalent to ANOVA in balanced settings.

The fixed effects ANOVA can be rewritten in the form of a GLM matrix notation as follows:

$$\boldsymbol{y} = \boldsymbol{X}\,\boldsymbol{\beta} + \boldsymbol{\epsilon}\,, \tag{1}$$

where $\boldsymbol{y}$ is the $(n \times 1)$ response vector, $\boldsymbol{X}$ is the $(n \times (p+1))$ model matrix, $\boldsymbol{\beta}$ is the $((p+1) \times 1)$ vector of parameters, and $\boldsymbol{\epsilon} \sim N(0, \sigma^2_\epsilon \boldsymbol{I_n})$ is the $(n \times m)$ vector of residuals. When the model contains F effects of interest (main factor effects, interactions, etc), $\boldsymbol{X}$ will be decomposed into $F+1$ blocs that include the constant term, and one for each model effect: $\boldsymbol{X} = (\boldsymbol{X}_0|\boldsymbol{X}_1| \ldots |\boldsymbol{X}_F)$, coded with a *sum* or *deviation* coding of the categorical variables.[21] The model matrices for the interaction terms are accordingly built by a column-wise multiplication of the corresponding main factors model matrices.

2.2 | Effect matrix decomposition

For a multivariate response matrix $\boldsymbol{Y}$ of size $(n \times m)$, the GLM can be written in matrix notation as $\boldsymbol{Y} = \boldsymbol{X}\Theta + \boldsymbol{E}$, where Θ is the $((p+1) \times m)$ parameter matrix and $\boldsymbol{E}$ is the $(n \times m)$ error matrix. By splitting the ordinary least squares estimator of

$\Theta : \hat{\Theta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$ with respect to each model effect: $\hat{\Theta}' = (\hat{\Theta}'_0|\hat{\Theta}'_1| \ \dots \ |\hat{\Theta}'_F)$, the ASCA$^+$ decomposition of the response matrix into effect matrices is $\hat{\boldsymbol{Y}} = \hat{\boldsymbol{M}}_0 + \sum_{f=1}^{F} \hat{\boldsymbol{M}}_f + \hat{\boldsymbol{E}}$, where $\hat{\boldsymbol{M}}_f = \boldsymbol{X}_f\hat{\Theta}_f$ and $\hat{\boldsymbol{E}} = \boldsymbol{Y} - \boldsymbol{X}\hat{\Theta}$.

The estimated effect matrices are recovered from the estimated model parameters and the design matrix with the following formula: $\hat{\boldsymbol{M}}_f = \boldsymbol{X}_f\hat{\Theta}_f$ for an effect f.

2.3 | Outputs

Two classes of outputs are usually derived: (a) diagnostic measures to evaluate the importance and significance of model terms and (b) graphical outputs to interpret and visualise the model effects in terms of observations and variables of the original dataset.

For balanced designs, the importance of a given effect is measured with the Frobenius norm of its effect matrix.

For unbalanced cases, the analysis of variance decomposition equation being not more valid, Thiel et al[21] proposed an adapted solution based on the type III SS commonly used in GLM.

Permutation tests are usually applied to assess the statistical significance of an effect (cf Thiel et al[21]).

One of the main purposes of ASCA is the visual inspection of the results. To be able to interpret each model effect, a PCA is first performed on the "pure" effect matrix: $\hat{\boldsymbol{M}}_f = \boldsymbol{T}_f\boldsymbol{P}'_f$. Plotting the first loadings vectors of $\boldsymbol{P}_f$ directly highlights which response variables are most affected by a model effect.

Zwanenburg et al[25] suggest to project the *augmented* effect matrix $\hat{\boldsymbol{M}}_f + \hat{\boldsymbol{E}}$ to derive the PCA scores. It leads to a better visualisation of the size of each model effect with respect to the variability of the residuals. The "augmented" scores matrix is calculated as $\boldsymbol{T}^a_f = (\hat{\boldsymbol{M}}_f + \hat{\boldsymbol{E}}) \times \hat{\boldsymbol{P}}_f$. This approach is referred as ASCA-E in Guisset et al[16] and will be used to augment the scores in LiMM-PCA.

2.4 | Limitations when random factors/effects are included in the model

A huge amount of literature is available on ANOVA and GLM models that explains extensively the usefulness of random effects in statistical models and the associated methodological difficulties. We suggest to the more applied readers to refer to Neter et al[3] for a very didactic presentation of many of these models or to Govaerts et al[26] for a synthesis to the attention of the chemists. These texts provide insights into the main reasons why ASCA is not well adapted to models including random factors.

Let us first start with balanced designs. In this case, the ANOVA theory is very well developed for models including only random effects (random models) or including both random and fixed effects (mixed models). The principle of total variance decomposition by an ANOVA equation remains correct regardless of the fixed or random nature of the factors. Consider, for example, the LMM with two factors, one fixed A and one random B. This model will be applied in Case Study 1 (Section 5.1) to analyse how a candies (factor A) are rated by b assessors (factor B) through several candies attributes in a sensory experiment, where all ratings are repeated c times. The model equation takes the following form:

$$y_{ijk} = \mu_{...} + \alpha_{i..} + \beta_{.j.} + (\alpha\beta)_{ij.} + \epsilon_{ijk} \ \text{ with } \ i = 1, \dots a; j = 1, \dots b; k = 1, \dots c\,, \tag{2}$$

where y_{ijk} is the observed response for the ith level of factor A, jth level of factor B and replicate k. $\mu_{...}$ is the overall mean, $\alpha_{i..}$ is the effect of level i of factor A, $\beta_{.j.}$ is the (random) effect of level j of factor B, and $(\alpha\beta)_{ij.}$ is the (random) interaction between i and j. In a mixed model, the $\beta_{.j.}$'s are supposed to be independent normal random variables $\sim iN(0, \sigma^2_B)$, the $(\alpha\beta)_{ij.} \sim iN(0, \sigma^2_{AB})$, and the errors $\epsilon_{ijk} \sim iN(0, \sigma^2_\epsilon)$. In this model, the main objectives are to understand how the mean response is affected when the level of A changes (eg, the candy) and what amount of variability is produced by the levels of factor B (eg, the assessors).

The related analysis of variance equation, which is the cornerstone of the ASCA decomposition, is identical to the one of the fixed model:

$$\begin{aligned} SS_T &= \sum_{ijk} (y_{ijk} - \bar{y}_{...})^2 \\ &= \sum_{i} (\bar{y}_{i..} - \bar{y}_{...})^2 + \sum_{j} (\bar{y}_{.j.} - \bar{y}_{...})^2 + \sum_{ij} (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 + \sum_{ijk} (y_{ijk} - \bar{y}_{ij.})^2 \\ &= SS_A + SS_B + SS_{AB} + SS_E. \end{aligned} \tag{3}$$

Effect	SS	df_x	MS_x	$E(MS_x)$	F Test
A	SS_A	$a-1$	SS_A/df_A	$\sigma^2_\epsilon + \frac{cb\sum\alpha_i^2}{(a-1)} + c\frac{a}{a-1}\sigma^2_{\alpha\beta}$	MS_A/MS_{AB}
B	SS_B	$b-1$	SS_B/df_B	$\sigma^2_\epsilon + ca\sigma^2_\beta$	MS_B/MS_E
AB	SS_{AB}	$(a-1)(b-1)$	SS_{AB}/df_{AB}	$\sigma^2_\epsilon + c\frac{a}{a-1}\sigma^2_{\alpha\beta}$	MS_{AB}/MS_E
Error	SS_E	$ab(c-1)$	SS_E/df_E	σ^2_ϵ	

Note. In this table, x represents the effect name (A, B or AB) or the error (E).

TABLE 1 ANOVA table and $E(MS)$ for the crossed 2-factor mixed ANOVA model

Until now, nothing has been altered by the randomness of factor B, but in what follows, adaptations are necessary. First, the sums of squares cannot be directly used to estimate the variances for random effects as they must be first corrected by other sums of squares. The consequences in ASCA are that the effect matrix $\hat{\boldsymbol{M}}_r$ of a random effect r, derived from Equation (3), cannot be directly used to interpret this effect.

In the ANOVA framework, the correct estimation is obtained on the basis of the so-called *Expected Mean Squares* ($E(MS)$) provided in Table 1 for this model. This table shows, for example, that $E(MS_B) = \sigma^2_\epsilon + ca\sigma^2_\beta$. The variance linked to factor B should therefore be estimated as $\hat{\sigma}^2_\beta = (MS_B - MS_E)/ca$.

In mixed models, the F test hypothesis to test for the presence of a fixed factor, say A, is written as $H_0 : \alpha_1, \ldots, \alpha_a = 0$ *versus* H_1 : not all $\alpha_i = 0$ and alternatively as $H_0 : \sigma^2_\beta = 0$ *versus* $H_1 : \sigma^2_\beta > 0$ when testing for the presence of a random factor, say B. The classic F tests for the significance test of an effect that were based on the ratio MS_f/MS_E cannot be used anymore. One should rather refer to the $E(MS)$ to decide which MS should be assigned to the denominator of each test statistic. For example, to test if fixed factor A is significant, the F-ratio MS_A/MS_{AB} is the one that should be used because the ratio of the corresponding $E(MS)$ is equal to 1 when A has no effect ($\alpha_i = 0$). In a multivariate setting, Marini et al[12] also uses this justification to decide which source of random variation to use in the case of mixed effects in order to obtain the reduced model residuals and perform permutation tests in ANOVA-TP. Hence, this has an incidence on ASCA since the denominator of these tests should naturally be used to visualise the noise in the ASCA-E effect matrix augmentation, for the computation of confidence bands, or in testing procedures.

Finally, this question is becoming more complex in the case of unbalanced designs since, in this case, the ANOVA matrix decomposition equation is no longer valid. Behind this issue lies the real motivation of introducing the LMM as a GLM extension. The aim is indeed to derive a generalisation of ASCA for the analysis of unbalanced experimental designs including random factors and a large number of responses. The main principles of LMM are presented in the next section before introducing LiMM-PCA, the suggested generalisation of ASCA to this class of models.

3 | THE LINEAR MIXED MODEL

3.1 | Model terms definition and justification

This chapter gives a succinct description of LMM (refer to Govaerts et al[26] for more information about mixed models in this context). LMM is a versatile regression toolkit in statistics when the observations cannot be assumed to be independent between each other (multiple measurements on the same patient, hierarchical data coming from different labs, repeated measures, etc). The model is built on the differentiation between the *fixed* and the *random* effects.[4,27] Typically, an effect is considered as fixed when the differences between its mean levels effects are of direct interest (eg, treatment effect), whereas the values of a random effect are considered as random realisations drawn from a larger population (eg, the patient effect). The levels are indeed a priori unknown and cannot be reproduced. Key aspects advocating the use of mixed models are that it leads to more valid inferences by taking into account the complex error structure of the data, the inferential results are generalised at the population of the random effect(s) and not just limited to their sampled levels; the random sources of variation can be compared with each other. Compared with ANOVA, LMM based on restricted maximum likelihood (REML) estimates offers more benefits, especially in the case on unbalanced data: In this context, consistency, asymptotic normality of the estimators, and the asymptotic sampling dispersion matrix of the estimators are known and enable to establish confidence intervals, or test hypotheses about derived parameters. Therefore, for advanced designs and models, LMM offers a flexible and generic framework for parameter estimation and inference compared with ANOVA.

3.2 | General notations and assumptions

The matrix notation of the LMM can be written as

$$\boldsymbol{y} = \boldsymbol{X}\,\boldsymbol{\beta} + \boldsymbol{Z}\,\boldsymbol{\alpha} + \boldsymbol{\epsilon}, \tag{4}$$

with $\boldsymbol{Z}$ the model matrix for random effects of size $(n \times q)$ and $\boldsymbol{\alpha}$ the random parameters vector of size $(q \times 1)$.

As in GLM, linear constraints apply between levels of a same effect for every fixed effect. Hence, the model matrix $\boldsymbol{X}$ is built with a *factor effects* coding as in ASCA$^+$. A dummy coding is used for the random effects design matrix $\boldsymbol{Z}$: The number of columns in $\boldsymbol{Z}$ is the sum of the levels' numbers for all the random effects.

The general assumptions for the random effects are $\boldsymbol{\alpha} \sim N(0, \boldsymbol{G})$ and $\boldsymbol{\epsilon} \sim N(0, \Sigma)$ with $\boldsymbol{G}$ of size $(q \times q)$ and Σ of size $(n \times n)$, the respective covariance matrices.

Here, we will assume the traditional *variance components* model with only categorial fixed and random effects as predictors. Moreover, there is no covariance between each pair of random effects such that $var(\boldsymbol{\alpha}) = \boldsymbol{G} = \left\{_d\ \sigma_r^2 \mathbf{I}_{q_r}\right\}_{r=1}^{R}$ for a random effect r having q_r levels, with $\mathbf{I}_{q_r}$ the identity matrix and $var(\boldsymbol{\epsilon}) = \Sigma = \sigma_\epsilon^2 \mathbf{I}_n$. The variance components for the random effects are contained in $\boldsymbol{\gamma} = (\sigma_1^2, \ \dots\ , \sigma_R^2)$.

The variance-covariance matrix of $\boldsymbol{y}$ is then $V = var(\boldsymbol{y}) = var(\boldsymbol{X}\,\boldsymbol{\beta} + \boldsymbol{Z}\,\boldsymbol{\alpha} + \boldsymbol{\epsilon}) = \boldsymbol{Z} var(\boldsymbol{\alpha}) \boldsymbol{Z}' + var(\boldsymbol{\epsilon}) = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' + \Sigma = \sum_{r=1}^{R} \boldsymbol{Z}_r \boldsymbol{Z}_r' \sigma_r^2 + \sigma_\epsilon^2 \mathbf{I}_n$. Thus, since $\mathbb{E}[\boldsymbol{y}] = \boldsymbol{X}\boldsymbol{\beta}$, we have $\boldsymbol{y} \sim N(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' + \Sigma)$.

The model fitting process is separated into three distinct parts: estimate the fixed effects $\boldsymbol{\beta}$ by maximum likelihood (ML), the variance parameters by REML, and predict the random effects by ML. More information about the estimation procedure is provided in the Supporting Information.

3.3 | Test of effects significance

In order to test the significance of one or several fixed effects $\boldsymbol{\beta}_f$ for $f = 1, \ \dots\ , F$ or one or several variance parameters σ_r^2 for $r = 1, \ \dots\ , R$, the Wald F statistic and the log likelihood ratio (LLR) are the two most well-known test procedures.[28] The latter compares two nested models and is based on the (R)LLR statistic: $2[\log(L_{H_1}) - \log(L_{H_0})]$, where L_{H_1} and L_{H_0} are the likelihoods of two nested models, with (under H_1) and without (under H_0) the tested effect(s). Given its versatility and the multivariate testing procedure explained thereafter in Section 4.5, LiMM-PCA will rely on the (R)LLR approach to test the effects significance.

For fixed effects, maximum likelihood estimates must be used in the test statistic but for variance parameters, the REML likelihood values should be used to compare different variance structures through the LLR statistic.[29]

The (R)LLR statistic follows asymptotically a (mixture of) chi-squared distribution(s). However, several limitations (anti-conservative P values for moderate to small sample sizes for fixed effects[30]; H_0 lying at the boundary of the parameter space for random effects[31]) make its use difficult to assess the statistical significance in practice.

An alternative solution for significance tests that can be applied both for fixed effects and variance parameters is to determine the full and null distributions of the (R)LLR test statistic by a parametric bootstrap approach.[32,33] This approach has been adapted to a multivariate setting in LiMM-PCA.

4 | LIMM-PCA: COMBINING ASCA$^+$ AND LMM FOR ADVANCED DESIGNS AND MODELLING

The aim of this paper is to extend the current ASCA$^+$ methodology to more advanced designs and models by the use of a multivariate mixed model that includes both fixed and random effects. The proposed method is called LMMs-PCA or LiMM-PCA. A summary of its workflow is illustrated in Figure 1, where the yellow boxes highlight new developments compared to ASCA$^+$.

4.1 | Step 1: PCA orthogonalisation and dimension reduction of the response matrix

In LiMM-PCA, before assessing links between the multivariate responses and the factors of interest, the response matrix $\boldsymbol{Y}$ should be prepared in order to comply as much as possible with the statistical properties that are expected in LMM.

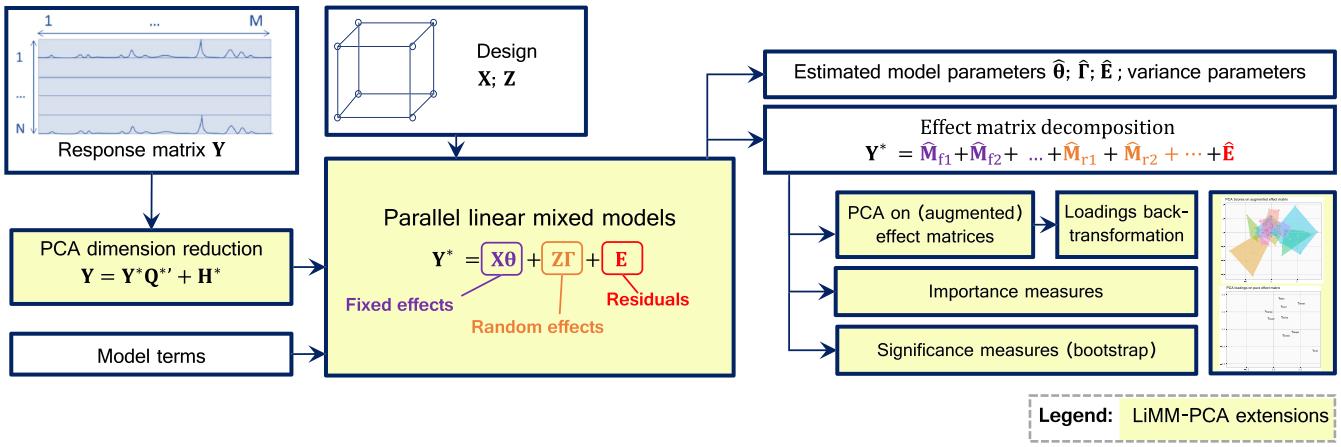


FIGURE 1 Combination of ASCA$^+$ and LMM to analyse high-dimensional designed data

First, a mathematical transformation can be applied to each response to ensure that its distribution is close to a normal one (eg, $\log(\boldsymbol{Y})$, $\sqrt{\boldsymbol{Y}}$, or any other Box-Cox transformations). Then, a linear transformation should be applied to the data in order to get the more independent responses as possible. This second step can be achieved by PCA orthogonalisation of the response matrix because it produces a set of uncorrelated responses, which are independent if the original responses do follow a multivariate normal distribution.

This transformation is motivated by the fact that linear mixed modelling of multiple responses benefits from much simpler inferential properties (eg, it enables to factorise the likelihood) when the responses are independent (cf Section 4.2).

After these transformations, the methodology developed subsequently implies that the hypotheses of independence and normality are almost verified for the transformed responses. If these properties cannot be achieved for the data of interest, most of the methodology remains still valid, but some inferential results may be affected.

This PCA transformation should be associated with a dimension reduction by dropping the very last components, especially when the number of variables exceeds that of the observations. This will have a positive impact on the computational burden without losing the important information contained in the data. This transformation—PCA and dimension reduction—is not affecting the final results interpretation since the back-transformation, applied to the loadings, will allow to project them onto the original m-dimensional response space.

More formally, the normalised and centred $(n \times m)$ response matrix $\boldsymbol{Y}$ will be transformed by PCA into a $(n \times m^*)$ matrix $\boldsymbol{Y}^*$ defined as $\boldsymbol{Y} = \boldsymbol{Y}^*\boldsymbol{Q}^{*'} + \boldsymbol{H}^*$, where m^*, the number of PC retained, should be high enough to retain most of the information contained in the data. The choice of m^* can be assessed by a scree plot or any other classic tool. The $(m \times m^*)$ loadings matrix $\boldsymbol{Q}^*$ is kept for later use to revert to the original variables space.

4.2 | Step 2: parallel mixed modelling

As in ASCA—where an ANOVA/GLM model is first fitted to each response variable—an LMM is fitted, by maximum likelihood estimation, to each column $\boldsymbol{y}_j^*$ of the new response matrix $\boldsymbol{Y}^*$:

$$\boldsymbol{y}_j^* = \boldsymbol{X}\boldsymbol{\beta}_j + \boldsymbol{Z}\boldsymbol{\alpha}_j + \boldsymbol{\epsilon}_j \quad \text{for } j = 1, \dots, m^*. \tag{5}$$

The model should to be built according to the underlying experimental design, and we will assume, as in ASCA, that the model matrix $\boldsymbol{X}$ and $\boldsymbol{Z}$ are organised in blocks corresponding to the F fixed and R random effects of the model: $\boldsymbol{X} = (\boldsymbol{X_0}|\boldsymbol{X_1}|\boldsymbol{X_2}|...\boldsymbol{X_F})$ and $\boldsymbol{Z} = (\boldsymbol{Z_1}|\boldsymbol{Z_2}|...\boldsymbol{Z_R})$. The first column of $\boldsymbol{X}$ is a constant term which is not always necessary if the design in balanced.

For each response j, the fixed parameters $\boldsymbol{\beta}_j$ and random predictors $\boldsymbol{\alpha}_j$ are estimated as $\hat{\boldsymbol{\beta}}_j = (\boldsymbol{X}'\hat{\boldsymbol{V}}_j^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\hat{\boldsymbol{V}}_j^{-1}\boldsymbol{y}_j^*$, $\hat{\boldsymbol{\alpha}}_j = \hat{\boldsymbol{G}}_j\boldsymbol{Z}'\hat{\boldsymbol{V}}_j^{-1}(\boldsymbol{y}_j^* - \boldsymbol{X}\hat{\boldsymbol{\beta}}_j)$, where $\hat{\boldsymbol{G}}_j$ is the estimated random effects covariance matrix for response j. These estimators are completed by the (RE)ML estimators of random parameters: A variance $\hat{\sigma}_{rj}^2$ is associated to each random effect r and the variance $\hat{\sigma}_{\epsilon j}^2$ to the residuals.

The estimation of these m^* models results in an estimated multivariate model of the following form:

$$\boldsymbol{Y}^* = \boldsymbol{X}\hat{\Theta} + \boldsymbol{Z}\hat{\Gamma} + \hat{\boldsymbol{E}}\,, \tag{6}$$

where $\hat{\Theta}$ is the $((p+1)\times m^*)$ matrix of estimated fixed parameters and $\hat{\Gamma}$ is the $(q\times m^*)$ matrix of random predictors.

According to the decomposition of matrices $\boldsymbol{X}$ and $\boldsymbol{Z}$ in blocks, $\hat{\Theta}'$ and $\hat{\Gamma}'$ can be decomposed into $\hat{\Theta}' = (\Theta_0'|\Theta_1'|\Theta_2'|\ \ldots\ |\Theta_F')$ and $\hat{\Gamma}' = (\Gamma_1'|\Gamma_2'|\ \ldots\ |\Gamma_R')$, respectively.

$\hat{\boldsymbol{E}}$, the $(n\times m^*)$ residual matrix, is deduced from previous calculations as $\hat{\boldsymbol{E}} = \boldsymbol{Y}^* - (\boldsymbol{X}\hat{\Theta} + \boldsymbol{Z}\hat{\Gamma})$.

Note that the hypothesis of independence assumed in the first step of the LiMMPCA procedure becomes handy at this stage because it enables to neglect complicated covariance structures between the random effects of the m^* parallel models.

4.3 | Step 3: effect matrix decomposition

Similarly to ASCA$^+$, the response matrix can then be decomposed into a sum of fixed and random effect matrices: $\boldsymbol{Y}^* = \hat{\boldsymbol{M}}_0 + \sum_{f=1}^{F}\hat{\boldsymbol{M}}_f + \sum_{r=1}^{R}\hat{\boldsymbol{M}}_r + \hat{\boldsymbol{E}}$.

For a fixed effect f, $\hat{\boldsymbol{M}}_f$ is simply calculated as $\boldsymbol{X}_f\hat{\Theta}_f$, and for a random effect r, $\hat{\boldsymbol{M}}_r$ is calculated as $\boldsymbol{Z}_r\hat{\Gamma}_r$.

Note that these effect matrices are orthogonal to each other for balanced cases and when an appropriate coding for the factors is chosen for the model matrices (cf Thiel et al[21] for more details).

4.4 | Step 4a: quantification of effects importance

Quantifying the effects importance is a common output derived in ASCA-like methods applications through the literature. Before the visual interpretation of the effect matrices content, it is indeed important to get an idea of what portion of the total variance of the data is explained by each model term. Another, more formal, approach consists in evaluating which of them have a statistically significant effect on the responses. The graphical interpretation of the effect matrices should therefore be guided by this prior information at hand.

Compared with ANOVA and GLM, the quantification of effects importance is not as straightforward for mixed models since it should be able to compare the fixed and random sources of variation based on a common scale. It should further solve the nonorthogonality issue between the effect matrices when the design is not balanced. Hence, a simple calculation of their Frobenius norm, as recommended in the traditional ASCA approach, is no longer appropriate.

The suggested approach is based on the work of Nakagawa and Schielzeth[34] that offers a general definition of marginal and conditional R^2 in the univariate LMM and GLMM framework for independent random effects. Their approach provides a direct solution to quantify fixed and random effects importance in the balanced case.

For the random effects, the estimated variance for one response $\boldsymbol{y}_j^*$ conditional on $\boldsymbol{X}$ is defined as the sum of the estimated variance components linked to all random effects of the model (residual included): $\widehat{var(\boldsymbol{y}_j|\boldsymbol{X})} = \sum_{r=1}^{R}\hat{\sigma}_{rj}^2 + \hat{\sigma}_{\epsilon j}^2$. Regarding the fixed effects, Nakagawa and Schielzeth[34] and Rousseau[35] both suggest to quantify the variance of an effect f over a response j by the population variance of the estimated effect matrices: $\hat{\sigma}_{fj}^2 = var(\hat{\boldsymbol{M}}_{fj})$ where $\hat{\boldsymbol{M}}_{fj}$ is the jth column of $\hat{\boldsymbol{M}}_f$. When the design is balanced and the factors are encoded with a sum coding, the population variance is basically the Frobenius norm of $\hat{\boldsymbol{M}}_{fj}$ divided by n. The total variance of $\boldsymbol{y}_j^*$ is later defined as the sum of all variance components linked to fixed and random effects.

At the multivariate stage, and thanks to the orthogonal property between the responses $\boldsymbol{y}_j^*$, all these individual variances can be aggregated over all responses. That way, one can obtain the global variance for the multivariate response and the portion of this variance that is linked to each effect included in the model. The global variance of $\boldsymbol{Y}^*$ is estimated as

$$\widehat{var(\boldsymbol{Y}^*)} = \sum_{j=1}^{m^*}\widehat{var(\boldsymbol{y}_j^*)} = \sum_{j=1}^{m^*}\left(\sum_{f=1}^{F}\hat{\sigma}_{fj}^2 + \sum_{r=1}^{R}\hat{\sigma}_{rj}^2 + \hat{\sigma}_{\epsilon j}^2\right)\,,$$

and the variance linked to one model term/effect $g = \{r, f\}$ is $\widehat{Var\ g} = \sum_{j=1}^{m^*}\hat{\sigma}_{ij}^2$. The latter can also be expressed as a percentage of the total variance as $\%\widehat{Var\ g} = \widehat{Var(g)} * 100/\widehat{var(\boldsymbol{Y}^*)}$.

Note that if the design is balanced, and the model contains only fixed effects, this approach is equivalent to what is usually done in ASCA. In contrast, if the design is unbalanced, the formula for the fixed effect variances can be adapted

on the basis of type III SS, as suggested in Thiel et al.[21] In the multivariate LMM framework, $\widehat{Var\ f}$ can be calculated as follows: $\hat{\Theta}'\boldsymbol{L}'(\boldsymbol{L}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{L}')^{-1}\boldsymbol{L}\hat{\Theta}/n$, where $\boldsymbol{L}$ is the contrast matrix to select in $\boldsymbol{X}$ and $\hat{\Theta}$ the columns/lines linked to an effect f.

4.5 | Step 4b : testing effects significance

This section provides a methodology to test for the statistical significance of one or a group of model terms over the response variables to complement the computation of the effects importance. In single response LMM models, the more general testing procedure is the LLR test. It is extended here to a multivariate test procedure able to deal with the multiplicity issue of univariate tests. Once more, this test relies on the hypothesis of near independence between the individual transformed responses $\boldsymbol{y}_j^*$ thanks to the orthogonalisation of the initial response matrix $\boldsymbol{Y}$.

The global LLR test statistics for (a group of) effect(s) $\tilde{g} = \{\tilde{f}, \tilde{r}\}$ GLLR$_{\tilde{g}}$is defined as the sum of the independent single response LLR statistics:

GLLR$_{\tilde{g}}$ statistic: $\Lambda_{\tilde{g}}^{obs} = 2[\sum_{j=1}^{m^*}(\log(L_{H_1,j\tilde{g}}) - \log(L_{H_0,j\tilde{g}}))]$,

where $L_{H_1,j\tilde{g}}$ and $L_{H_0,j\tilde{g}}$ are the likelihood values of two nested models with response j: the one under H_1, where the tested effect(s) $\tilde{g}$ is (are) included in the model, and the other under H_0, without the tested effect(s) $\tilde{g}$ of interest.

When testing for the fixed effects, the asymptotic distribution under H_0 is a $\chi^2_{m^*\times k}$ where k corresponds to the number of parameter's restrictions between the two univariate models under H_0 and under H_1.

Recall that testing for the random effects is not as straightforward when testing the nullity of variance components at the border of the parameter space. In this context and as explained in Section 3.3, the distribution of the RLLR under standard conditions (independent and identically distributed [i.i.d.] random terms) is indeed a mixture of χ^2 distributions, and we will illustrate through the case study in Section 5.1, that this approximation is providing good results.

However, due to difficulties in the derivation of finite sample distributions of the (R)LLR, the effects significance are rather assessed by a parametric bootstrap procedure using the scheme suggested in Davison et al,[32] Sinha,[36] or Halekoh et al.[33] It is extended in the multivariate case to evaluate the significance of (a) given model effect(s) $\tilde{g}$ as follows:

1. Define the model under H_0 obtained by excluding the effect(s) $\tilde{g}$ from the full model: $\boldsymbol{Y}^* = \tilde{\boldsymbol{X}}\tilde{\Theta} + \tilde{\boldsymbol{Z}}\tilde{\Gamma} + \boldsymbol{E}$ where the $\sim$ notation represents the model elements without the effect(s) $\tilde{g}$ to be tested.
2. Fit the multivariate mixed model under H_1 and H_0 by, respectively, including and excluding the effect(s) $\tilde{g}$.
3. Calculate the observed GLLR statistic $\Lambda_{\tilde{g}}^{\text{obs}}$.
4. Keep the following estimated parameters of the model under H_0 to use them in the bootstrap loop: the fixed parameters matrix $\hat{\tilde{\Theta}}$, the variances of the random effects $\hat{\sigma}^2_{rj}$, and the residual variances $\hat{\sigma}^2_{\epsilon j}$ $\forall j = 1, \ldots, m$.
5. Repeat a high number of times ($b = 1, \ldots, B$):
   - Generate a bootstrap response matrix $\boldsymbol{Y}_b^*$ as $\boldsymbol{Y}_b^* = \tilde{\boldsymbol{X}}\hat{\tilde{\Theta}} + \tilde{\boldsymbol{Z}}\tilde{\Gamma}_b + \boldsymbol{E}_b$ where the matrices $\tilde{\Gamma}_b$ and $\boldsymbol{E}_b$ are randomly generated from independent normal distributions with zero mean and variances issued from the estimated model under H_0.
   - Estimate the full and restricted models on this new response matrix $\boldsymbol{Y}_b^*$ and calculate the GLLR statistic $\Lambda_{\tilde{g}}^b$. Note that this statistic should be close to 0 since the responses have been generated based on the restricted model.
6. Calculate the P value for the test statistic as the tail probabilities defined in Davison and Hinkley[32], Chapter 4: $p_{\tilde{g}}^{\text{boot}} = \frac{\sum_{b=1}^{B} I(\Lambda_{\tilde{g}}^b \geq \Lambda_{\tilde{g}}^{\text{obs}}) + 1}{B+1}$.

This procedure may be applied individually to each effect r or f of the mixed model in order to get a measure of their significance.

4.6 | Step 5 : visual representation of the effect matrices

The last and probably the more informative step of LiMM-PCA aims at offering to the scientist a way to visualise the modelling results though a set of graphics. LiMM-PCA's philosophy is close to what is nowadays used in the ASCA framework.

In ASCA, each pure effect matrix $\boldsymbol{M}_g$ for an effect of interest g (ie, those which have been shown to be important in Step 4) is first decomposed by PCA. The loadings are plotted to illustrate which variables are affected by the effect g and the pure or augmented scores are then used to visualise the behaviour of this effect on the experimental points. In LiMM-PCA, this PCA decomposition of $\hat{\boldsymbol{M}}_g$ is written as $\hat{\boldsymbol{M}}_g = \boldsymbol{T}_g\boldsymbol{P}_g^{*\prime}$, where c_g is the rank of $\hat{\boldsymbol{M}}_g$, $\boldsymbol{T}_g$ is the "pure" scores matrix of

size ($n \times c_g$), and $\boldsymbol{P}_g^*$ is the loadings matrix of size ($c_g \times m^*$). As the original responses $\boldsymbol{Y}$ were orthogonalised into $\boldsymbol{Y}^*$ by PCA before being introduced in the mixed models, the $\boldsymbol{P}_g^*$ loadings must be back-transformed in order to be interpreted in their initial variable space. The resulting ($c_g \times m$) matrix is obtained as follows: $\boldsymbol{P}_g = \boldsymbol{P}_g^* \boldsymbol{Q}^{*\prime}$.

In ASCA$^+$, the effect matrices were augmented by the residuals to allow the graphical visualisation of the effects significance via a scores plot. However, the scores augmentation with random effects requires particular attention in LMM as more than one source of random variation is now present in the model. Recall from Table 1 that if effects B and AB are random in the two-way mixed model described in Equation (2), the MS_{AB} is the denominator in the F value test statistic when testing for fixed effect A. In LiMM-PCA, we suggest then to use the corresponding effect matrix $\boldsymbol{M}_{AB}$ to augment $\boldsymbol{M}_A$. In addition, a correction factor is used here to take into account the degrees of freedom (df) and the F distribution quantile that are used in the formula of the associated F test.

For the fixed effect A, we propose concretely to calculate the augmented effect matrix as $\boldsymbol{M}_A + C * \boldsymbol{M}_{AB}$ with C calculated as $\sqrt{df_A / df_{AB} * F_{df_A, df_{AB}, 1-\alpha}}$, where $F_{(df_A, df_{AB}, 1-\alpha)}$ is the $(1-\alpha)$th (eg, 95th) percentile of the $F_{(df_A, df_{AB})}$ distribution. Indeed, the correction factor C can be considered as an empirical choice of correction, as the distribution is only following a *pseudo* F distribution. The degrees of freedom in this formula should however be calculated with care because the random effects in LMM are not estimated from the simple ANOVA decomposition given in Equation (3) but with the predictors of the random effects in the LMM. These estimators are "shrunk" compared with the ANOVA SS and actual degrees of freedom must be calculated accordingly. In the statistical literature, the "degrees of freedom" for each model effect are renamed as "effective dimensions" (ED) and can be calculated following a methodology described in Eilers[37] and in the Supporting Information. An ED for an effect takes the same value as the degrees of freedom of the ANOVA model (ie, the number of parameters) when the effect is fixed. For a random effect r, the ED can be any value between 0 and the total number of coefficients related to this effect, depending on the value of its variance σ_r^2 estimated from the data.

5 | CASE STUDIES

This section illustrates step by step the results of the application of LiMM-PCA to two datasets, issued from sensory sciences and metabolomics, both already published in the literature.

5.1 | Candy sensory dataset

5.1.1 | Dataset description

The first dataset is issued from a sensory experiment where a panel of 11 judges rated five different candies with respect to a series of nine attributes: transparency, acidity, sweet taste, raspberry flavour, sugar-coated texture tested with a spoon, biting strength in the mouth, hardness, elasticity in the mouth, and stick to teeth in the mouth. Each candy was evaluated in triplicate by each assessor and the resulting balanced design is presented here in an unfolded data table of dimensions (165×9) where the ($165 = 11 \times 5 \times 3$) observations are identified by four digits: judge-candy-replicate, eg, 0712 corresponds to judge 7, candy 1, and replicate 2.

This dataset was first published in Luciano and Næs[10] where they compared the PC-ANOVA and the ASCA approaches to model this dataset with a fixed and crossed 2-factor ANOVA model. This dataset is also present in Liland et al,[38] where the authors suggest a new analytical solution to estimate the multivariate classification uncertainty based on confidence ellipsoids in the balanced case, and extending their results for multiple testing of effect's levels differences.

Different questions of interest for this dataset will be answered in this section, the main ones being: *What is the mean value of each attribute for the different tested candies? Are these attributes significantly different across the candies? Is there an agreement between the judges? Do some judges appreciate certain candies? Does the dataset contain outlying judges or observations?*

5.1.2 | Statistical model

In this article, the statistical model chosen to analyse the design results is a typical 2-factor mixed ANOVA model where the candy is a fixed factor and the subject is considered as random in order to allow the generalisation of the results to the *population* of potential consumers from which the 11 assessors are issued. This sensory profiling modelling in discussed in details in Næs and Langsrud[39] and, in particular, the consequence of defining an assessor as fixed or random in such studies.

For one attribute (ie, response), the model corresponds to the one already presented in Section 2.4. In this model, y_{ijk} represents the value of the attribute of interest for candy i, judge j, and replicate k; and $a = 5$, $b = 11$ and $k = 3$. μ

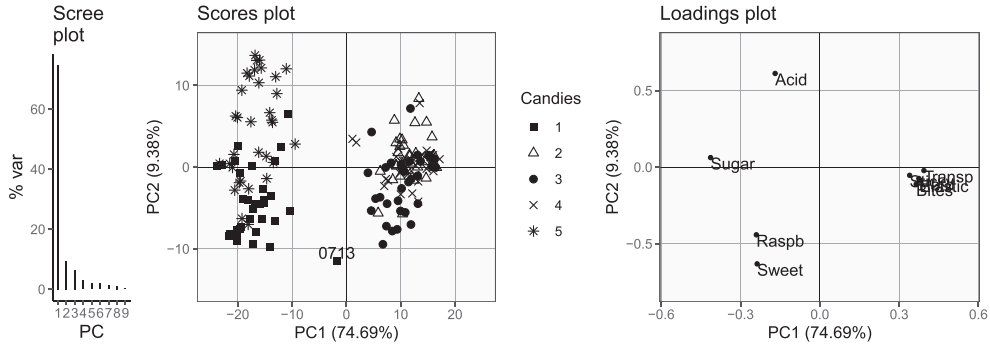


FIGURE 2 PCA scree plot, plots of PC1-2 scores and loadings for the sensory attributes response matrix

represents the overall mean, α_i is the (fixed) effect of candy i, β_j is the (random) effect of judge j (β_j are independent and normally distributed, ie, $\sim iN(0, \sigma_J^2)$), and $(\alpha\beta)_{ij}$ is the (random) interaction between candy i and j ($(\alpha\beta)_{ij} \sim iN(0, \sigma_{CJ}^2)$). The errors ϵ_{ijk} are $iN(0, \sigma_\epsilon^2)$.

5.1.3 | Step 1: PCA orthogonalisation, data exploration, and dimension reduction

The first step of LiMM-PCA is the PCA dimension reduction and orthogonalisation of the response matrix $\boldsymbol{Y}$. This section also discusses the quality of the data.

The PC1-2 scores and loadings of PCA on the raw attributes matrix are illustrated in Figure 2. The first PC captures almost three quarters of the variance present in the response matrix and clearly discriminates between the two previously observed groups of candies, based on clusters of distinct attributes values. According to PC1, candies {2,3,4} (*{1,5}*) have a higher (*lower*) attribute value for transparency, biting strength, hardness, elasticity, and stick to teeth and a lower (*higher*) one for acidity, sweet taste, sugar coated texture, and raspberry flavour. The second PC mainly captures the negative correlation between acidity and {sweet, raspberry} grades: Candies {2, 5} are mainly rated having more acidity whereas candies {1, 3} have a sweeter, more raspberry taste. Note that an outlying observation from candy 1 is observed in the PC1-2 scores plot.

A diagnostic plot measuring Hotelling T^2 vs the squared residual Q distances for a chosen PCA model (here with four PCs retained)[40] was also built to detect potential outliers (cf Supporting Information). T^2 represents the distance of one observation to the center of the points in the PCA model whereas the Q statistic measures the residual variation not taken into account in the PCA model. At this stage, this graph does not point out important outliers.

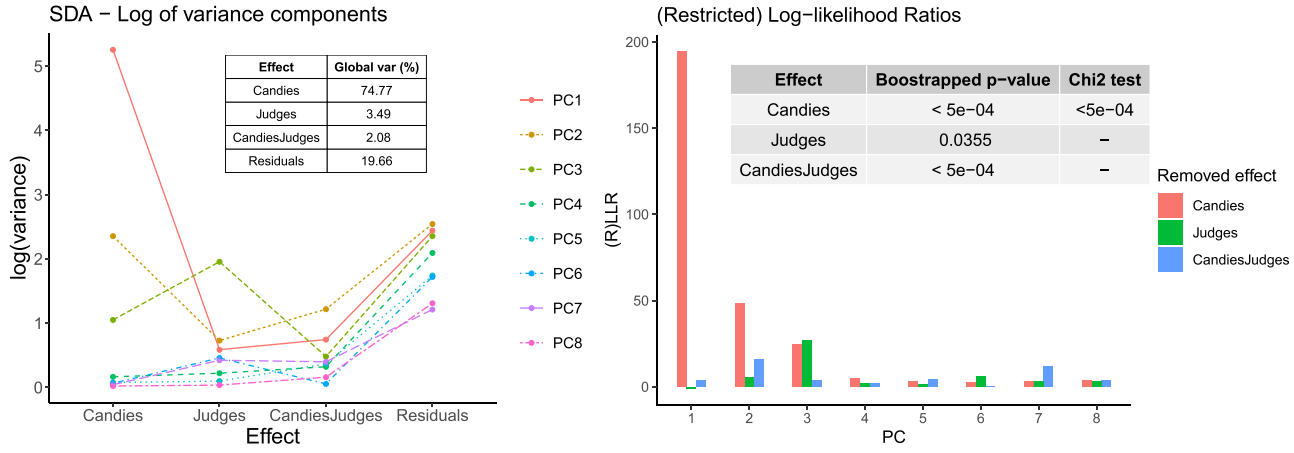
In Luciano and Næs,[10] PC-ANOVA was applied to these data, and their results will serve here as a reference for the subsequent analysis. In their paper, the candy effect was assessed by applying univariate F tests on each scores vector of the PCA applied to the attributes matrix. However, the authors used the mean squares (MS) ratios of fixed effects ANOVA model for their test. For this analysis, we prefer to adopt the more conservative ratios, as advocated in mixed models, along with a Bonferroni correction of the P values to correct for test multiplicity (P values were multiplied by $3 \times 8 = 24$, the number of tests performed). The adapted results (given in Supporting Information) show that the effects of the three model terms are deemed significant at least once in the first three PCs. Hence, they can all be considered as significant in the present analysis accordingly. The candy effect is the predominant one in the data since it is significant in the two first PCs, which are capturing most of the attributes variability.

Only the last PC is discarded in the following steps of LiMM-PCA in order to retain most of the information contained in the data, leaving a total of eight PCs retained in the new response matrix. The new matrix accounts for an overall of 99.67% of the total variance.

5.1.4 | Steps 2 and 3: mixed modelling and effect matrix decomposition

The multivariate LMM for this case study corresponds to the 2-factor mixed ANOVA model with crossed factors already discussed in Section 5.1.2. It is written as $\boldsymbol{Y}^* = \boldsymbol{X}_0\Theta_0 + \boldsymbol{X_C}\Theta_C + \boldsymbol{Z}_J\Gamma_J + \boldsymbol{Z}_{CJ}\Gamma_{CJ} + \boldsymbol{E}$, where $0, C, J$, and CJ, respectively, account for the fixed intercept, the candy, the judge, and the interaction effects. The fitted model residuals are approximately normal (cf Supporting Information).

After the model estimation by REML, the response matrix can be further decomposed as a sum of effect matrices: $\boldsymbol{Y}^* = \hat{\boldsymbol{M}}_0 + \hat{\boldsymbol{M}}_C + \hat{\boldsymbol{M}}_J + \hat{\boldsymbol{M}}_{CJ} + \hat{\boldsymbol{E}}$.



(A) Variance components. (B) (R)LLR.

FIGURE 3 Variance components and (R)LLR issued from the mixed models applied on the sensory dataset

5.1.5 | Step 4: effects importance and significance testing

According to the maximum likelihood methodology presented in Section 4.4, the percentage of variance for each PC linked to each model effect has been calculated and is illustrated in Figure 3A.

Figure 3A shows that the first two PCs are mainly linked to the candies and the interaction effects whereas the third one is mostly linked to the judge main effect, but also to the candies one. The remaining PCs do not account for a large amount of variability and are not interpreted further. In this figure are also provided the global variance, accounting for the effects importance. It is split into the model effects as expected in sensory profiling; ie, the fixed candy effect is by far the main source of variability in the attributes rating, whereas the random effects (the assessor and the interaction) represent a minor part of the variability present in the data. Note that the residual variability (or the intra-assessor variability) is large compared with the other random effects.

Prior to the assessment of the effects significance, the graph of the LLR test statistics per PC (Figure 3B) already gives a good indication of which effect leads to a larger increase of the global likelihood of the mixed models. The figure further provides the P values of the GLLR test statistics obtained by the bootstrap strategy presented in Section 4.5 and by a χ^2 with 32 (4×8) degrees of freedom for the fixed effect. They confirm that all the effects included in the model are highly statistically significant.

Both graphs in Figure 3 show that the candy main effect leads to an important increase of the LLR in the first PC, but also in PC2 and 3. The judge and interaction effects are in turn responsible only for a more moderate increase of some of the first PCs and only a formal test is able to conclude to their true significance.

Histograms of the bootstrapped G(R)LLR under H_0 are presented for each model effect in Supporting Information, where they are compared with the observed test statistics. It shows that for the fixed effect, the χ^2_{32} works well to approximate the small sample distribution of the test statistic. Other derived bootstrap results are also depicted in the Supporting Information where it is shown based on histograms and q-q plots that the χ^2_4 distribution matches the LLR under H_0 for the fixed effect as well as mixtures of χ^2 do better approximate RLLR under H_0 for the random effects.

5.1.6 | Step 5: visual representation of the effect matrices

In order to visualise the resulting effect matrices, PCA is now applied on each of them: $\hat{\boldsymbol{M}}_g$ is reconstructed into $\boldsymbol{T}_g\boldsymbol{P}_g^*$. The loadings are then back-transformed to visualise the variables and the augmented effect matrices projected on the space to the first PC's to visualise the observations.

Based on $E(MS)$ and F tests from the 2-factor mixed ANOVA model presented in Section 2.4 (cf Table 1), Table 2 presents how the effect matrices should be augmented in the ASCA-E fashion: The interaction effect matrix, corrected with the correction factor as explained in Section 4.6, is added to the candy effect matrix, and the assessors and the interaction effect matrices are augmented by the corrected residuals matrix.

The resulting scores and loadings plots are provided in Figure 4. Note that to enhance the visual interpretation of scores plots, ellipses were added for the candy and judge effects, and points were grouped and linked by the candy-judge centroid value of those scores. They lead to the conclusions that follow.

Effect	Type	Associated ED	Effect Matrix Added	Correction Factor
Candy	F	$ED_C = a - 1$	$\hat{\boldsymbol{M}}_{CJ}$	$\sqrt{F_{df_C,df_{CJ}} \times \frac{ED_C}{ED_{CJ}}}$
Judge	R	ED_J	$\hat{\boldsymbol{E}}$	$\sqrt{F_{df_J,df_E} \times \frac{ED_J}{ED_E}}$
CJ	R	ED_{CJ}	$\hat{\boldsymbol{E}}$	$\sqrt{F_{df_{CJ},df_E} \times \frac{ED_{CJ}}{ED_E}}$
Error	R	$ED_E = n - (ED_C + ED_J + ED_{CJ})$	-	-

TABLE 2 Effect matrix and correction factor added to augment each effect matrix for the sensory dataset

Note. Type is either fixed (F) or random (R). a is the number of levels of the candy effect (5). F_{ED_1,ED_2} represents the 95% quantile of the F distribution with ED_1, ED_2 the partial effective dimensions.

As already observed in the PCA scores of the raw data (Figure 2), the candies (A) are well separated. Differences between the candies {1,5} and {2,3,4} are mainly driven along PC1 by the previously described attributes clustering in this figure. Furthermore, the loadings indicate that candies having an important sugar texture are likely to have a lower rating for transparency, biting strength, hardness, elasticity, and stick to teeth attributes. Similarly, sweet candies with raspberry flavour will have a lower acidity taste.

In the scores plot for the judge effect (B), the first PC captures most of the variance linked to this effect. Clearly, some judges are less homogeneous in their rating than others (eg, 2, 5, and 7) and tendencies of underrate or overrate the attributes can be observed: eg, assessors 7 and 5 *{resp. 2 and 4}* tend to overrate *{resp. underrate}* the attributes, and mainly the acidity.

For the interaction effect (C), the first PC is driven by the distinction between the acidity and {sweet, raspberry taste}, already spotted by the PCA on the candy main effect matrix. From the scores plot, the main judges responsible for the interaction effect are 2 and 7. It means that they are the one differing the most from the average profile of attributes rating. The main candy responsible for the interaction is candy 5. It can be interpreted as a disagreement over the ratings of candy 5 attributes by the judges, which is also visible from the greater scattering of scores linked to this candy in the scores plot of the candies main effect matrix.

A comparison of these LiMM-PCA results with the ASCA results in Luciano and Næs[10] shows clearly the same results for the loadings of candy effect but only a similar pattern for the scores, that are augmented in our case.

Scores and loadings for the judge effect are comparable along PC1 with these ASCA results except that here PC1 captures a larger part of the effect matrix variability. Nevertheless, results differ along PC2. For the interaction, PC1-2 loadings are similar, although more informative than ASCA. However, the scores augmentation enables to spot the main observations responsible for this interaction.

Finally, PCA scores and loadings on the residuals matrix (D) represent the variability of the three replicates for each volunteer and candy. What stands out first is that there is no emerging structure from those residuals. The variability in the PC1-2 space is mainly driven by two outlying observations on candy 1: one from assessor 7 and the other from assessor 2. Note that removal of those outliers would lead to a slight transfer of the global variance distribution between the residuals and all the model effects. Compared with the ones obtained, global variances would, respectively, be 76.39%, 3.65%, 2.72%, and 17.24%. The loadings plot shows that, along PC1, the variability is driven by the outlying observation from assessor 2, mainly regarding the attributes sweet and raspberry taste and transparency, which are clearly misattributed when looking at the data table. The outlying observation from assessor 7 is the main source of variability that built PC2 with the attributes acidity, biting strength, hardness, elasticity, and stick to teeth being erroneously rated given the other replicates from that judge. Clearly, the two judges provided abnormal scores for these two evaluations. Note that scores and loadings of the interaction effect and the residuals identify the principal outliers already highlighted in the Hotelling T^2 vs squared residuals distance Q (cf Supporting Information). However, the LiMM-PCA analysis is richer in terms of interpretation as these outlying observations observed in the scores plot can be directly related to the attributes rating represented in the corresponding loadings plot.

5.2 | Variance components analysis of metabolomics data

5.2.1 | Dataset description

The metabolomics human serum dataset[35] is meant to study how serum NMR metabolomics spectra are influenced by various sources of variability linked to either biological (volunteer, sampling) or analytical (tube, time replication) factors. This dataset was mainly designed to quantify and compare: (a) the variability between and within volunteers and (b) the biological and analytical sources of variability. It also enables to test if these sources of variability must be considered as statistically significant with respect to pure repeatability caused by analytical noise.

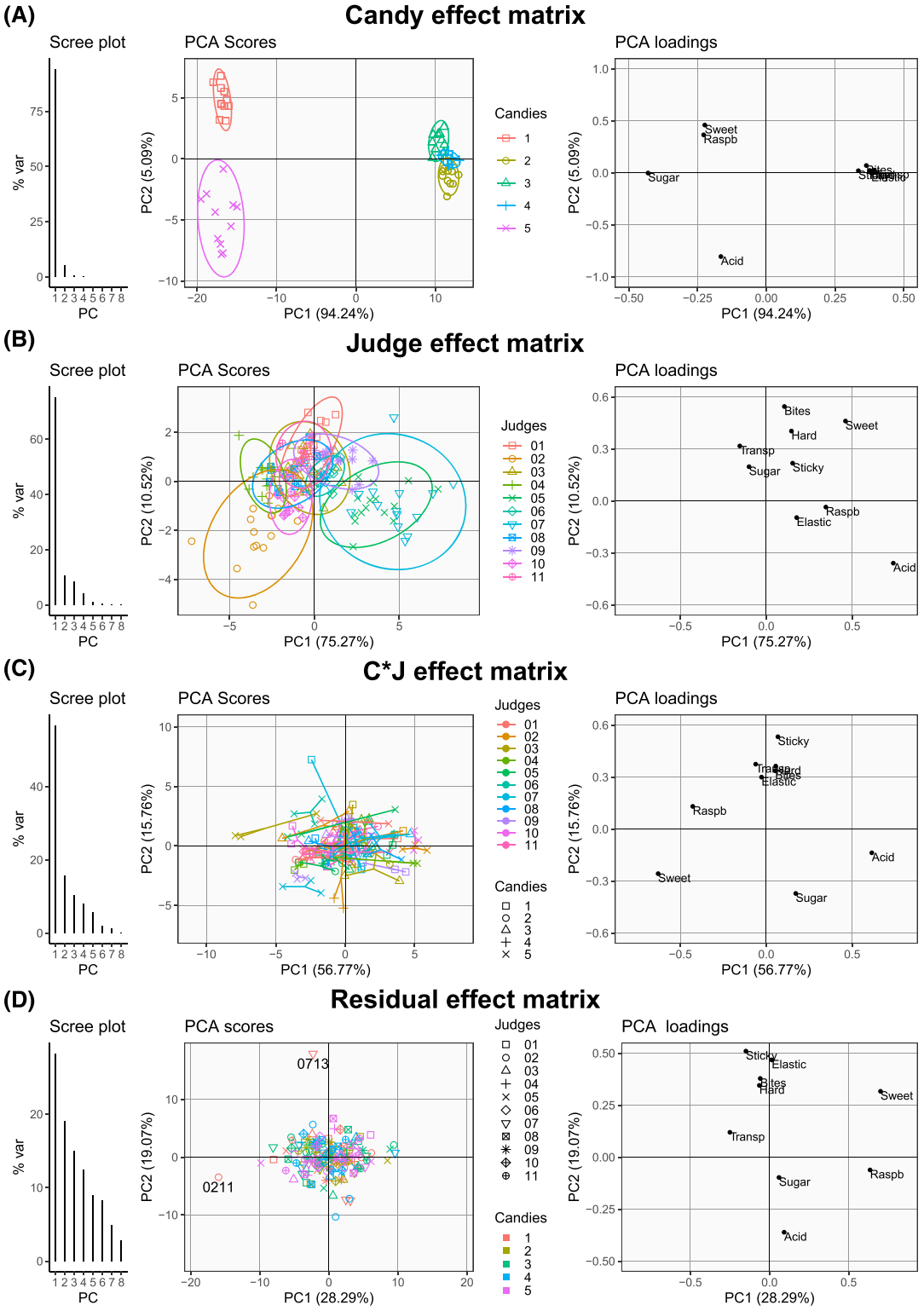


FIGURE 4 Individual PCA on the effect matrices for the sensory dataset. From left to right: scree plots, scores and loadings of {PC1,PC2}. The scree plots and loadings plot are from the PCA on the pure effect matrices whereas the scores are augmented

In this study, 12 healthy volunteers were enrolled. For each volunteer, three blood samplings were performed on non-consecutive days and under similar conditions. For each sampling, two tubes were collected. The serum was then extracted from each sample and frozen. Each serum sample was then analysed twice by ^{1}H NMR spectroscopy the same day with a couple of hours between the two measurements in order to evaluate whether the samples can be conserved after defrost-

ing. Each spectrum is available at a resolution of 750 buckets (ie, variables) and observations are named with a series of five digits as follows: volunteer-sampling-tube-time; eg, 01312 corresponds to volunteer 01, sampling 3, tube 1, and time 2. For more information on data preparation and preprocessing, please refer to Rousseau.[35]

Four observations had an acquisition problem and were removed beforehand (04221, 08222, 11121, and 11122), rendering the design unbalanced with a total of 140 observations instead on 144.

5.2.2 | Statistical model

In Rousseau,[35] the 750 responses of this four factors experiment were already analysed in parallel by univariate LMM, including the main effect of the three random effects—volunteer (V), sampling (S) and tube (U)—and the fixed factor of this study—time (T). Note that the time effect should indeed be introduced as fixed in the model since the same time laps was applied to all samples. All other factors were considered as random because their levels represent specific instances among the population of all possible factor's levels. All the interactions were considered as negligible in the model.

In this paper, we will proceed with the same model but rather consider the tube as a fixed factor because the data shows an unexpected but clear shift between the first and second tubes collected for each patient.

The univariate mixed ANOVA model can then be written as

$$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_{l(k)} + \epsilon_{ijkl} \text{ with } i = 1, \dots a; j = 1, \dots b; k = 1, \dots c; l = 1, \dots d\,, \tag{7}$$

where y_{ijkl} is the value of the response of interest for time i, tube j, volunteer k, and sample l; and $a = 2$, $b = 2$, $c = 12$ and $d = 3$. μ is the overall mean, α_i is the fixed effect of time i, β_j is the fixed effect of tube j, γ_k is the random effect of volunteer k ($iN(0, \sigma_V^2)$) and $\delta_{l(k)}$ is the random effect of sampling l nested in volunteer k ($iN(0, \sigma_S^2)$). The errors ϵ_{ijkl} are supposed to be $iN(0, \sigma_\epsilon^2)$.

In this quite intricate model, it can be shown that the following ANOVA ratios should be used to test the significance of each effect: MS_T/MS_E for time effect, MS_U/MS_E for tube effect, MS_V/MS_S for volunteer effect and MS_S/MS_E for sample effect. A detailed explanation for effect matrix augmentation in the multivariate case is provided in Section 5.2.6 below.

5.2.3 | Step 1: PCA orthogonalisation, data exploration and dimension reduction

The prior PCA applied to the raw spectral matrix shows that 15 PCs of the 140 available cumulate 99.12% of the total variance present in the spectral profiles and are retained in what follows. PCA scores and loadings of the first 2 PCs are illustrated in Figure 5. They gather more than 85% of the total variability. The scores plot shows already a clear volunteer effect among the four factors of interest.

A graph diagnostic for a PCA model with four PCs retained (cf Supporting Information) shows potential outliers in this dataset. Some of them will be more clearly identified during the LiMM-PCA modelling.

5.2.4 | Steps 2 and 3: mixed modelling and effect matrix decomposition

The multivariate mixed model corresponding to the ANOVA model defined in Section 5.2.2 can be written as $\boldsymbol{Y}^* = \boldsymbol{X}_0\Theta_0 + \boldsymbol{X}_T\Theta_T + \boldsymbol{X}_U\Theta_U + \boldsymbol{Z}_V\Gamma_V + \boldsymbol{Z}_{S[V]}\Gamma_{S[V]} + \boldsymbol{E}$, where the 0 index accounts for the fixed intercept and "[]" describes the nested model structure. The fitted model residuals are approximately normal (cf Supporting Information).

After the parallel (RE)ML estimation of the model for the 15 PCs, the effect matrix decomposition equation is the following: $\boldsymbol{Y}^* = \hat{\boldsymbol{M}}_0 + \hat{\boldsymbol{M}}_T + \hat{\boldsymbol{M}}_U + \hat{\boldsymbol{M}}_V + \hat{\boldsymbol{M}}_{S[V]} + \hat{\boldsymbol{E}}$. Each effect matrix is of size (140×15). Note that since the design is slightly unbalanced, these matrices are not anymore perfectly orthogonal.

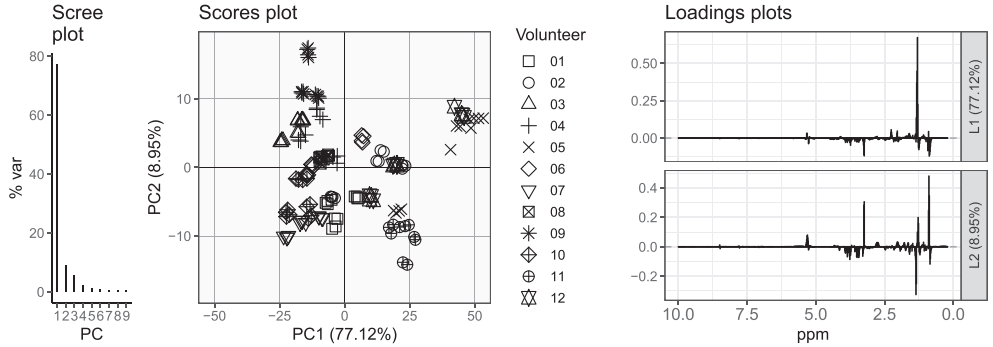
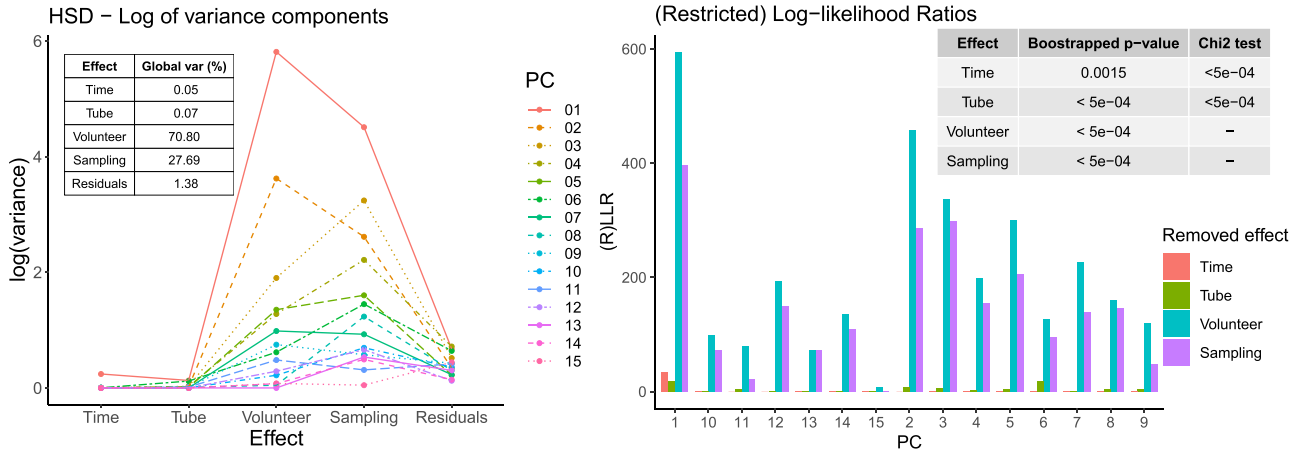


FIGURE 5 PCA scores and loadings of the raw metabolomics dataset



(A) Variance components. (B) (R)LLR.

FIGURE 6 Variance components and (R)LLR from the multivariate mixed model applied to the metabolomics dataset

TABLE 3 Effect matrix and correction factor added to augment each effect matrix for the metabolomics dataset

Effect	Type	Associated *df*	Effect Matrix Added	Correction Factor
Time	F	$ED_T = a - 1$	$\hat{\boldsymbol{E}}$	$\sqrt{F_{ED_T,ED_E} \times \frac{ED_T}{ED_E}}$
Tube	F	$ED_U = b - 1$	$\hat{\boldsymbol{E}}$	$\sqrt{F_{ED_U,ED_E} \times \frac{ED_U}{ED_E}}$
Volunteer	R	ED_V	$\hat{\boldsymbol{M}}_S$	$\sqrt{F_{ED_V,ED_S} \times \frac{ED_V}{ED_S}}$
Sampling	R	ED_S	$\hat{\boldsymbol{E}}$	$\sqrt{F_{ED_S,ED_E} \times \frac{ED_S}{ED_E}}$
Error	R	$ED_E = n - (ED_T + ED_U + ED_S + ED_V)$	-	-

Note. Type is either fixed (F) or random (R). a and b are the number of levels respectively of the time (2) and tube (2), and F_{ED_1,ED_2} represents the 95% quantile of the F distribution with ED_1, ED_2 the partial effective dimensions.

5.2.5 | Step 4: Effects importance and significance testing

For each retained PC $\boldsymbol{y}_j^*$, the total variance is decomposed by effect and aggregated over all PCs in Figure 6A. The volunteer effect captures more than two thirds of the total dataset variability, followed by the sampling effect. The remaining effects (tube and time) only account for a marginal portion of the total information extracted from the dataset. The global residual variance is also negligible. For all the effects but time, their variability contribution is widespread across multiple PCs. The time effect is only caught by PC1. Note that the first three PCs capture most of the volunteer's and sampling's variabilities and the tube effect is mainly contained in PC4 and 6.

For the statistical significance of the effects, the (R)LLR values in Figure 6B indicate to what extend the addition of an effect leads to an increase in the (restricted) log-likelihoods: Unsurprisingly, the volunteer and the sampling effects appear to be the prominent ones. As shown for variances components, the effects are spanned across all the PCs except for the time effect, that only affects PC1. The global bootstrap test results in this figure confirms that all the fixed and random effects are deemed significant in the multivariate mixed model, although the importance of the time and tube effects is rather small.

Histograms of the bootstrapped (R)LLR under H_0 confirmed that for the fixed effects, the χ^2_{15} (15 PCs × 1 degree of freedom) works well to approximate the small sample size distribution of the test statistic.

5.2.6 | Step 5: visual representation of the effect matrices

Based on the $E(MS)$ and F test calculations (cf Supporting Information), Table 3 tells which of the effect matrices must be added, corrected with the correction factor as explained in Section 4.6, to the effect matrix of interest in order to compare this effect against the appropriate random source of variation.

Results of PCA applied to the effect matrices with the augmented scores are given in Figure 7. Without going into the metabolites identification details here, a general remark concerning the loadings is that they reveal which of the spectral

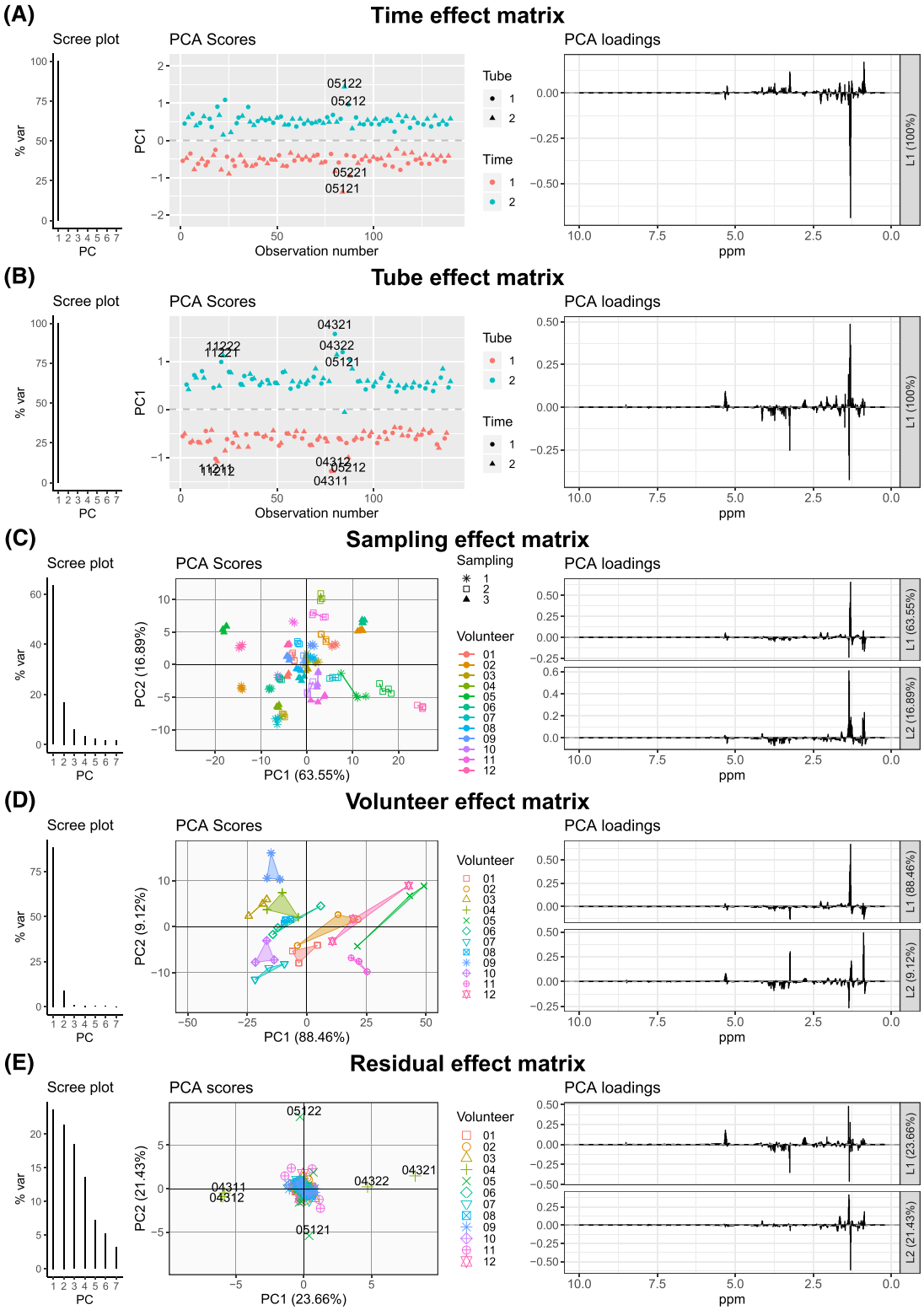


FIGURE 7 Individual PCA on the effect matrices for the metabolomics dataset. From left to right: scree plots, scores and loadings of {PC1,PC2}. The scree plots and loadings are from the PCA on the pure effect matrices whereas the scores are augmented

descriptors are most likely responsible for the variability observed in the scores between the groups of effect's levels. They can also tell if a shift between peaks of the same metabolites across spectra has been detected.

With regards to the time and tube effects (A and B), only one PC is built because these terms only have 1 degree of freedom. Therefore, PC1 captures 100% of the corresponding effect matrix variability. These two graphs show a very clear

separation of the observations between times 1 and 2 and tubes 1 and 2. These graphs also highlight several outliers for spectra coming from volunteers 4, 5, and 11.

The sampling scores plot (C) depicts observations coming from the same volunteer and sampling. All clusters of points are well separated, confirming the significant effect of the sample within each volunteer regarding model residuals. The differences between the three samplings of a same volunteer are particularly noticeable for volunteers 5 and 11. The volunteer scores plot (D), shows, as expected, a very clear separation in the PC1-2 space. These two scores plots also illustrate the random nature of the volunteer and sample factors. Even if they are highly significant and show a clear separation between their factor levels, the later behave randomly around (0,0) as expected from the underlying normality assumption.

A look into the PCA scores and loadings of the residuals (E)—which account for the excluded higher order interaction terms and other data randomness—spots mainly the "outliers" from volunteers 4, 5, and 11, already spotted in the diagnostic plot (cf Supporting Information) and the scores from the time and tube effect matrices. It seems that samples three of volunteer 4, one of volunteer 5, and two of volunteer 11 are especially concerned across all these graphs. The residual loadings indicate that the variability along these first two PCs is mainly driven by peak shifts in the spectra.

6 | CONCLUSIONS AND PERSPECTIVES

LMM has well-known inferential properties and is considered as a powerful and major toolbox for statisticians. Yet, they only appear occasionally in the chemometrics literature. The suggested methodology, called LiMM-PCA, manages to model advanced error structures and thus generalises ASCA to a broader framework, by using such LMM. This extension to model more advanced designs opens up to a more general modelling framework for future research in that topic. In this respect, new methodological developments have been suggested to appropriately augment the effect matrices, as well as being able to compare the contributions of the mixed effects and test their statistical significance.

The property to include random effects is indeed highly recommended to have a correct inference for specific designs such as in repeatability/reproducibility studies or for longitudinal data.

Two case studies, based on already published real experimental datasets, showed that it successfully leads to a comprehensive and fine-tuned analysis of such high-dimensional and designed data. On the one side, it keeps the advantages of ASCA$^+$ by combining dimension reduction techniques and classic modelling. Effect matrix projections are undoubtedly a powerful visualisation tool of the multivariate structures in the space of each effect of the statistical model linked to the experimental design. This approach further enables to model without bias an unbalanced dataset as in ASCA$^+$. On the other side, new proposed developments lead to correct inference when considering the inclusion of random effects, and new methodological aspects are established to quantify and compare the mixed variability sources, as well as to provide a global test procedure of effects significance.

Some obstacles or unconsidered aspects are still remaining, such as the influence of the initial PCA on the generalisation of the results. One solution would be to apply a resampling scheme on the whole procedure. Also, augmenting the matrices in the visualisation step, which is routinely used by chemometricians, here only serves for visual purposes as the general scope of LiMM-PCA is rather to model unbalanced data. As a matter of fact, the augmented scores cannot be used to formally assess the significance of the effects. Further research should lead to an adaptation of the methodology to enable a visual assessment of the effects significance (ie, visualise the model uncertainty and perform multiple tests of effect level differences), such as in Liland et al.[38]

One intricate aspect of ASCA-related methods that has not been debated here is the data pre-processing, and principally the scaling issue, already discussed in Timmerman et al.[41] The authors demonstrated that scaling of such methods influences the data analysis results and subsequent interpretation. In the future, more emphasis and research should be conducted in this non-trivial topic.

Other perspectives arising from this work are both methodological and practical: beyond the two presented applications, LiMM-PCA could be directly applied to other kinds of chemometric applications of interest. Its theoretical aspects could especially be extended to include continuous variables or modify the structure of the variance of the residuals in the model, eg, to analyse other kinds of longitudinal data.

ACKNOWLEDGEMENTS

The authors thank the UCLouvain and in particular the Institute of Statistics, Biostatistics and Actuarial Sciences (ISBA), and the Statistical Methodology and Computing Support (SMCS) for their support, Pascal de Tullio for fruitful discussions

on the metabolomics data, and Eli Lilly company for providing the metabolomics data used in this paper. They also especially thank Tormod Næs for providing them with the sensory dataset. Finally, the first author gratefully acknowledges funding from the Belgian Fund for Scientific Research (F.R.S.-FNRS) with a FRIA grant.

CONFLICT OF INTEREST

Authors declare that they have no conflict of interest.

ORCID

Manon Martin https://orcid.org/0000-0003-4800-0942

REFERENCES

1. Wold S, Sjöström M, Eriksson L. Pls-regression: a basic tool of chemometrics. *Chemom Intell Lab Syst*. 2001;58(2):109-130. http://www.sciencedirect.com/science/article/pii/S0169743901001551. PLS Methods.
2. Martens H, Høy M, Westad F, Folkenberg D, Martens M. Analysis of designed experiments by stabilised PLS regression and Jack-knifing. *Chemom Intell Lab Syst*. 2001;58(2):151-170. http://www.sciencedirect.com/science/article/pii/S0169743901001575. PLS Methods.
3. Neter J, Kutner MH, Nachtsheim CJ, Wasserman W. *Applied Linear Statistical Models*, Vol. 4. Chicago: Irwin; 1996.
4. Gelman A. Analysis of variance: why it is more important than ever. *Ann Stat*. 2005;33(1):1-31. http://www.jstor.org/stable/3448650
5. Mardia K, Bibby J, Kent J. *Multivariate analysis*. London: Academic Press; 1979.
6. de B. Harrington P, Vieira NE, Espinoza J, Nien JK, Romero R, Yergey AL. Analysis of variance-principal component analysis: a soft tool for proteomic discovery. *Anal Chim Acta*. 2005;544(1):118-127. http://www.sciencedirect.com/science/article/pii/S0003267005002692. Papers Presented at the 9th International Conference on Chemometrics in Analytical Chemistry.
7. Smilde AK, Jansen JJ, Hoefsloot HCJ, Lamers R-JAN, van der Greef J, Timmerman ME. Anova-simultaneous component analysis (ASCA): a new tool for analyzing designed metabolomics data. *Bioinformatics*. 2005;21(13):3043-3048. https://doi.org/10.1093/bioinformatics/bti476
8. Jansen JJ, Hoefsloot HCJ, van der Greef J, Timmerman ME, Westerhuis JA, Smilde AK. ASCA: analysis of multivariate data obtained from an experimental design. *J Chemom*. 2005;19(9):469-481. https://doi.org/10.1002/cem.952
9. Ellekjær MR, Ilseng MA, Næs T. A case study of the use of experimental design and multivariate analysis in product improvement. *Food Qual Prefer*. 1996;7(1):29-36. http://www.sciencedirect.com/science/article/pii/0950329395000186
10. Luciano G, Næs T. Interpreting sensory data by combining principal component analysis and analysis of variance. *Food Qual Prefer*. 2009;20(3):167-175. http://www.sciencedirect.com/science/article/pii/S0950329308001171
11. Ghaziri AE, Qannari EM, Moyon T, Alexandre-Gouabau M-C. AOV-PLS: a new method for the analysis of multivariate data depending on several factors. *Electron J Appl Stat Anal*. 2015;8(2):214-235. http://siba-ese.unisalento.it/index.php/ejasa/article/view/14988
12. Marini F, de Beer D, Joubert E, Walczak B. Analysis of variance of designed chromatographic data sets: the analysis of variance-target projection approach. *J Chromatogr A*. 2015;1405:94-102. http://www.sciencedirect.com/science/article/pii/S0021967315007839
13. Jansen JJ, Bro R, Hoefsloot HCJ, van den Berg FWJ, Westerhuis JA, Smilde AK. Parafasca: ASCA combined with PARAFAC for the analysis of metabolic fingerprinting data. *J Chemom*. 2008;22(2):114-121. https://onlinelibrary.wiley.com/doi/abs/10.1002/cem.1105
14. Bouveresse DJ-R, Pinto RC, Schmidtke LM, Locquet N, Rutledge DN. Identification of significant factors by an extension of ANOVA–PCA based on multi-block analysis. *Chemom Intell Lab Syst*. 2011;106(2):173-182. http://www.sciencedirect.com/science/article/pii/S016974391000081X. Chimiométrie 2009, Paris, France, 30 November - 1 December 2009.
15. Boccard J, Rudaz S. Exploring omics data from designed experiments using analysis of variance multiblock orthogonal partial least squares. *Anal Chim Acta*. 2016;920:18-28. http://www.sciencedirect.com/science/article/pii/S0003267016303920
16. Guisset S, Martin M, Govaerts B. Comparison of PARAFASCA, ACOMDIM, and AMOPLS approaches in the multivariate GLM modelling of multi-factorial designs. *Chemom Intell Lab Syst*. 2019;184:44-63. http://www.sciencedirect.com/science/article/pii/S0169743917307748
17. Timmerman ME. Multilevel component analysis. *Br J Math Stat Psychol*. 2006;59(2):301-320. https://onlinelibrary.wiley.com/doi/abs/10.1348/000711005X67599
18. Jansen JJ, Hoefsloot HCJ, van der Greef J, Timmerman ME, Smilde AK. Multilevel component analysis of time-resolved metabolic fingerprinting data. *Anal Chim Acta*. 2005;530(2):173-183. http://www.sciencedirect.com/science/article/pii/S0003267004013005
19. de Noord OE, Theobald EH. Multilevel component analysis and multilevel PLS of chemical process data. *J Chemom*. 2006;19(5-7):301-307. https://onlinelibrary.wiley.com/doi/abs/10.1002/cem.933
20. Stanimirova I, Michalik K, Drzazga Z, Trzeciak H, Wentzell PD, Walczak B. Interpretation of analysis of variance models using principal component analysis to assess the effect of a maternal anticancer treatment on the mineralization of rat bones. *Anal Chim Acta*. 2011;689(1):1-7. http://www.sciencedirect.com/science/article/pii/S0003267011000705
21. Thiel M, Féraud B, Govaerts B. ASCA+ and APCA+: extensions of ASCA and APCA in the analysis of unbalanced multifactorial designs. *J Chemom*. 2017;31(6):e2895-n/a. https://doi.org/10.1002/cem.2895. e2895 cem.2895.
22. Snijders TAB, Bosker RJ. *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: SAGE Publications; 1999.

23. McCulloch CE, Searle SR. *Generalized, linear and mixed models*, Wiley Series in Probability and Statistics. New York: John Wiley & Sons; 2001.
24. Brown H, Prescott R. *Applied mixed models in medicine*. Chichester: John Wiley & Sons; 2006.
25. Zwanenburg G, Hoefsloot HCJ, Westerhuis JA, Jansen JJ, Smilde AK. ANOVA–principal component analysis and ANOVA–simultaneous component analysis: a comparison. *J Chemom*. 2011;25(10):561-567. https://onlinelibrary.wiley.com/doi/abs/10.1002/cem.1400
26. Govaerts B, Francq B, Marion R, Martin M, Thiel M. The essentials on linear regression, ANOVA, general linear and linear mixed models for the chemist. Discussion Paper DP2020/01, Belgium, Institute of Statistics, Biostatistics and Actuarial Sciences, UCLouvain; 2020.
27. Searle SR, Casella G, McCulloch C. *Variance components: Wiley series in probability and mathematical statistics*. New York: John Wiley & Sons; 1992.
28. Gumedze FN, Dunne TT. Parameter estimation and inference in the linear mixed model. *Linear Algebra Appl*. 2011;435(8):1920-1944. http://www.sciencedirect.com/science/article/pii/S002437951100320X
29. Verbeke G, Molenberghs G. Springer series in statistics. Linear Mixed Models for Longitudinal Data; 2000.
30. Pinheiro JC, Bates DM. Linear mixed-effects models: basic concepts and examples. *Mixed-effects models in S and S-Plus*. New York, NY: Springer; 2000:3-56.
31. Self SG, Liang K-Y. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J Am Stat Assoc*. 1987;82(398):605-610. http://www.jstor.org/stable/2289471
32. Davison AC, Hinkley DV. *Bootstrap methods and their application* (CambridgeSeries in Statistical and Probabilistic Mathematics). Cambridge: Cambridge university press; 1997.
33. Halekoh U, Højsgaard S, et al. A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models—the R package pbkrtest. *J Stat Softw*. 2014;59(9):1-30.
34. Nakagawa S, Schielzeth H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol Evol*. 2013;4(2):133-142. https://doi.org/10.1111/j.2041-210x.2012.00261.x
35. Rousseau R. Statistical contribution to the analysis of metabonomics data in 1h NMR spectroscopy. *Ph.D. Thesis*: Institut de statistique, biostatistique et sciences actuarielles, Université catholique de Louvain, Belgium; 2011.
36. Sinha SK. Bootstrap tests for variance components in generalized linear mixed models. *Can J Stat*. 2009;37(2):219-234.
37. Eilers PHC. The truth about the effective dimension. *Statistica Neerlandica*. 2018;72(3):201-209. https://onlinelibrary.wiley.com/doi/abs/10.1111/stan.12131
38. Liland KH, Smilde A, Marini F, Næs T. Confidence ellipsoids for ASCA models based on multivariate regression theory. *J Chemom*. 2018;32(5):e2990. https://onlinelibrary.wiley.com/doi/abs/10.1002/cem.2990
39. Næs T, Langsrud Ø. . Fixed or random assessors in sensory profiling? *Food Qual Prefer*. 1998;9(3):145-152. http://www.sciencedirect.com/science/article/pii/S095032939600050X. Sensometric Workshop.
40. Jackson JE. *A user's guide to principal components*, Vol. 587, New York: John Wiley & Sons; 2005.
41. Timmerman ME, Hoefsloot HCJ, Smilde AK, Ceulemans E. Scaling in ANOVA-simultaneous component analysis. *Metabolomics*. 2015;11(5):1265-1276. https://doi.org/10.1007/s11306-015-0785-8

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Martin M, Govaerts B. LiMM-PCA: Combining ASCA$^+$ and linear mixed models to analyse high-dimensional designed data. *Journal of Chemometrics*. 2020;34:e3232. https://doi.org/10.1002/cem.3232