Multiway Multidimensional Analysis for Pharmaco-EEG Studies

FMRIB Technical Report TR00DL2
A related paper has been submitted to JASA

Didier G. Leibovici

Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB),
Department of Clinical Neurology, University of Oxford, John Radcliffe Hospital,
Headley Way, Headington, Oxford, UK
didier@fmrib.ox.ac.uk

Abstract

Current pharmaco-dynamic studies involve electro-encephalographic signal recording (EEG) of subjects according to a crossover design wherein each period includes repeated measures of EEG (in days/times). Spatio-temporal distributions of parameters (frequency bands of the signal) are of interest for differences between placebo and verum doses. After a brief description of shortcomings mainly in using parametric univariate or classical multi-variate analysis, this paper aims to introduce some multiway multidimensional approaches to take into account the structure of the data, a table with $k > 2$ entries. A generalisation of the Singular Value Decomposition to multi-tables enables a quantification of each set of linear components describing a spatio-temporal effect of the bands but also subject dispersion, which achieve data reduction and description. The flexibility of the method, including preprocessing and metric choices are issued for pharmaco-EEG analysis, introducing a generalisation of Correspondence Analysis for more than two variables.

keywords: Pharmaco-dynamic studies, quantified EEG, Multiway Analysis, SVD, Correspondence analysis, Principal Tensor Analysis on k modes.

1 Introduction

When a compound is expecting to show some central nervous system (CNS) activity, its potentials still need to be well established for the drug to be classified properly before thinking about therapeutic effect. For that purpose pharmaco-dynamic (PDY) studies are required and currently involve electro-encephalographic signal recording (EEG) of healthy subjects according to a crossover design wherein each period includes repeated measures of EEG (days, times). Spatio-temporal distributions of parameters (frequency bands of the signal) are of interest for differences between placebo and verum doses. Links with additional variables such as neurocognitive variables (psychometric tests) can also be explored and will be addressed briefly in the discussion as well as a current interest on looking at pharmaco-kinetic parameters conjointly.
The data-recording methodology and the quantification of the EEG-signal used for the dataset analysed thereafter is fully described in [17]. A collection and quantification of EEG-data, for each of the 28 leads (international 10/20 system is complemented to 28 leads with B1, FC1, FC2, B2, W1, PC1, PC2, W2). At each time of measurement, EEGs are taken under 3 minutes vigilance controlled (VC) recording condition (subjects push two knobs with their eyes closed), followed by 3 minutes resting (R) recording condition (subjects relax with their eyes closed). After filtering and digitisation and artefact removal procedure completed, energy spectra (μV^2) is calculated, for each 2 second period over a frequency range of 0.5 to 32Hz, using the Fast Fourier Transform (FFT), and then averaged for each subject and each recording condition. Each mean energy spectra is averaged by standard frequency EEG bands : δ (0.5-3.5Hz), θ(4-7.5Hz), α_1 (8-9.5Hz), α_2 (10-12.5Hz), β_1 (13-17.5Hz), β_2 (18-20.5Hz), β_3 (21-32Hz) and *Total* (0.5-32Hz). Absolute energies and relative energies (percentage of the *Total* band energy) are considered. The alpha slow wave index ($ASI = \frac{\alpha}{(\delta+\theta)}$), the mean frequency (GMF) and the mean complexity (GCO) of the EEG spectrum are also calculated.
The whole process will then analyse at each time of measurement(typically 10 not regularly spaced measures): 28 leads$\times$(2 absolute or % $\times$(7 bands)+1 total+3 synthetic variables) $\times$2 conditions =28 locations$\times$18

parameters×2 conditions =1008 variables measured say 10 times on say 12 subjects. In fact only 7 parameters generated the 18.
This methodology was conducted for the following pharmaco-dynamic study (PDY),a placebo-controlled, double-blind trial, with randomisation of 12 healthy male subjects into a 4 periods and 4 treatments cross-over design. Each received a single morning dose of 10, 30, 90mg of compound or placebo and wake-EEG was performed on day 1 before administration and 0.5, 1, 1.5, 2, 2.5, 3, 4, 6, hours post-dosing and on day 2: 24 and 36 hours post-dosing. Blood was sampled for determination of drug plasma concentration and endocrinological assessments on day 1 before administration and then 1, 2, 4, 6 hours post-dosing and on day 2: 24 and 36 hours post-dosing.
The interest is in knowing if the compound has an effect? which dose? (dose effect?), at what point in time does it happen? where is it located on the scalp? for which frequency band or pattern of frequency band does this affect? To answer these questions, parametric and non-parametric testing methods have been routinely implemented. In the first place some weaknesses of these mainly univariate methods will be pointed out before introducing our proposed method involving multiway data analysis methods. The main method applied here was theoretically exposed in [16]. The purpose of this paper is show how to modify and apply it in this context. This involves different methods which are related to existing approaches in multidimensional analysis (two-way analysis), and thereby extending them to multiway data. An important generalisation is about Correspondence Analysis extended from the analysis of 2 variables to k variables enabling to break down the lack of complete independence into additive components relating to different level of interactions between the variables.

2 Shortcomings of current statistical methods

EEG mapping data analysis gives rise to some statistical problems when looking for treatments effects with inferential tests, mainly because of the structure of the data leading to multiple comparisons: in time points, in variables (EEG bands) and in locations (electrodes).
One must first notice that multiple testing correction such as Bonferroni adjustment is not applicable here as there is a large amount of highly correlated variables to be considered: 1008 measures for each time point (among about 10) and for each dose (say 3 doses). Multivariate parametric methods such as MANOVA are not to be recommended either mainly because of the small size of the sample (here 12 subjects) making the assumptions difficult to assess and models not practical to apply (*e.g.* covariance structure estimation).
Based on these points to analyse a wake-EEG (pharmaco-EEG data) the statistical method often shows a two level analysis as found in [3]. Their method or at least their approach is widely used, so comments will only be based on this one. The first level called *Statistical Decision Tree*, an *exploratory* level, uses non-parametric testing methods ; significant results (at the leaves of the tree) are submitted to the second level, a *confirmatory* level using qualitative criteria called Descriptive Data Analysis (DDA) and a quantitative confirmation using Principal Component Analysis (PCA). After the first level, two kinds of series of maps, one at each time point, are then produced: coloured maps of p-values for dose versus placebo comparisons (each map is 28 points which can be interpolated), and the SDT-maps which plots at each electrode the value of the "decision" of the direction of drug effect [3]). At the second level DDA intends to give confirmation of SDT results meeting a few qualitative criteria *e.g.* spatial coherence and time coherence of significant results (*i.e.* clusters of electrodes that lasts in time), then SDT/DDA results are confirmed or not by using a t-test on the principal component (PCA of the electrodes) "overlapping" the spatial SDT/DDA result (*e.g.* a frontal effect FP1, FP2, Fz if overlapped by the back/front principal component).
First of all, the term "Statistical Decision" in SDT is misleading as the method does not give any level of confidence on the final conclusion. The SDT procedure is applied electrode by electrode and time after time (in comparison to baseline), therefore multiple testing problem is still present. Also problematic is the comparison of probability maps from time to time or from EEG-band to EEG-band. In particular, a probability map does not enable quantification of the effect observed especially when produced from non-parametric tests. The only possible comparison is in fact the spread, if it is a true spread (multiple testing). Certainly the DDA procedure provides some control on false positives. It introduces qualitative criteria on coherence of the results. If these criteria seem common sense they might not support all the designs or drug analysed, and they bring an interpretation or conclusive step into the statistical analysis from which interpretations and conclusions are made. The PCA confirmation plays a similar role concerning multiple testing, but this time in an acceptable way as it is a statistical control. Notice it would seem more logical to start with PCA as an *exploratory* tool, then select electrodes most contributing to PC and/or meeting DDA criteria, and then to *confirm* hypothesis with SDT.
The main problem with the current method seems to be the multiple testing issue. This has been an issue in the medical imaging literature when looking at activation with different imaging techniques such as PET.

Solutions to exhibit the distribution of local maxima can be derived using random fields theory as in [22] or permutation testing procedure as in [7]. Permutation testing would be better in this context as here only 28 "pixels" forms the "image", making the smooth random field approximation difficult to keep. The approach is univariate in a highly multivariate context with three particular features : time correlation, spatial correlation and frequency-band correlation. Using univariate models (or separate models) does not account for spatial, temporal, and frequency structures of the data, which are discussed only *a posteriori* and qualitatively when making conclusions from the results.
The idea of using data reduction techniques is promising because it can at first reduce the problem of multiple testing, and secondly provide some modelisation of the structure. PCA is not fully appropriate for the multi-entries array as only bilinear modelisation is possible, one must use a data reduction method allowing multi-linear modelisation. Multiway data reduction techniques have been already used on multichannel evoked potentials data as in [6] with the PARAFAC method. In this later paper the signal was in input instead of its Fourier transform (summarised on bands) as in our case (*i.e. EEG analysis instead of qEEG analysis*). PARAFAC method seems more appropriate in EEG analysis as the focus is more on modelisation of the time course (long) than on decomposition of the variability. Choosing a tradeoff between *modelisation* and *decomposition* more focused on this later, the purpose of this paper is to describe an other method handling multi-entries data which conserves most of the properties of the PCA method, with therefore easier understanding.

3 Multiway multidimensional data reduction

The aim of this section is to explain the basics of the SVD-*k*modes method as an extension of SVD. The presentation can be limited to $k = 2$ and $k = 3$ as the case $k > 3$ is then a straightforward extension. Further details are in [16]. First of all the development of this generalisation of the SVD (Singular Value Decomposition) is described within tensor algebra framework in finite dimension. It enables us to extend matrix algebra calculus in an easy way. A tensor of order one is a vector, a tensor of order two is a matrix, a tensor of order three is three-way array etc...
Let $\{e_i\}_{i=1,n}$, $\{f_j\}_{j=1,p}$, and $\{g_k\}_{k=1,t}$ be the canonical bases respectively of $E = \mathbb{R}^n$, $F = \mathbb{R}^p$ and $G = \mathbb{R}^t$; with $a \in E$ and $b \in F$ let us define the bilinear map $a \otimes b$ by:

$$a \otimes b(x, y) =< a, x >_E < b, y >_F, \tag{1}$$

(where $< ., . >_E$ is the inner product in E) ; consider now the canonical bilinear maps built with the e_i and f_j, they constitute a base of a space noted $E \otimes F$ the tensor product of the spaces E and F. Without going further into algebraic concepts, notice that because of symmetry in equation (1) $x \otimes y$ can be considered as a linear map onto $E \otimes F$, so that one has the universal property of the tensor product: transforming a bilinear map (multilinear in general) into a linear map. An $n \times p$ matrix A of elements $A_{ij} \in \mathbb{R}$, can be written algebraically,

$$A = \sum_{ij} A_{ij} \quad e_i \otimes f_j \quad , \tag{2}$$

and A is said to belong to the space $E \otimes F$, tensorial product of the spaces E and F[1]. Notice that the array, the linear map associated, the tensor are noted A because of isomorphisms. In the same manner a three-way array $n \times p \times t$ A of elements $A_{ijk} \in \mathbb{R}$, can be written algebraically,

$$A = \sum_{ijk} A_{ijk} \quad e_i \otimes f_j \otimes g_k, \text{ and } A \in E \otimes F \otimes G. \tag{3}$$

The vectors of the space $E \otimes F \otimes G$ with the form $d = \alpha \otimes \beta \otimes \gamma$ (where $\alpha = \sum_i \alpha_i e_i, \beta = \sum_j \beta_j f_j, \gamma = \sum_k \gamma_k g_k$) are called *decomposed tensors*, and are said to be of rank one - a sum of r linearly independent decomposed tensors would give a rank r tensor-[2]. To finish with basic tools of tensor algebra, let us also introduce the generalisation of *a product of a vector by a matrix*: the product of vector (or a tensor) by a tensor, also called contraction and noted "..". For example let $A \in E \otimes F \otimes G, \gamma \in G$, then $A..\gamma \in E \otimes F$ with:

$$\begin{aligned} A..\gamma &= \sum_{ijk} A_{ijk} \quad e_i \otimes f_j < g_k, \gamma >_G \\ &= \sum_{ijk} A_{ijk}\gamma_k \quad e_i \otimes f_j. \end{aligned} \tag{4}$$

[1] One may see the $e_i \otimes f_j$ as the canonical basis elements of the tensor space and represented by the matrices $e_i \, {}^t f_j$.
[2] Note that for tensors of order two (*i.e.* matrices) it coincides with the rank definition of linear maps.

Arithmetically this can be seen considering A as a matrix with np rows and t columns, then calculating the image of γ by this matrix gives a representation of $A..\gamma$. Note that $A..B$ is the inner product between the tensors $< A, B >_{E\otimes F\otimes G}$.

3.1 SVD within tensor algebra framework

The first singular value of a data matrix A is :

$$\begin{aligned} \sigma_1 &= \max_{\substack{\|\psi\|_E=1 \\ \|\varphi\|_F=1}} A..(\psi\otimes\varphi) \\ &= A..(\psi_1\otimes\varphi_1) \quad (\text{ in tensor form}) \\ &= {}^t\psi_1 A\varphi_1 \quad (\text{ in matrix form}). \end{aligned} \tag{5}$$

ψ_1 is termed first principal component, φ_1 first principal axis, $(\psi_1\otimes\varphi_1)$ will be called first principal tensor. Solving the problem associated with this maximisation leads to transition formulae and then to the classical eigenequations where ψ_1, φ_1 and σ_1^2 are the first eigenvectors and eigenvalue of the respective symmetric operators:

$$\left\{\begin{array}{ll} X..\varphi &= \sigma\psi \\ X..\psi &= \sigma\varphi \end{array}\right. \text{ or } \left\{\begin{array}{ll} X\varphi &= \sigma\psi \\ {}^tX\psi &= \sigma\varphi \end{array}\right. \Rightarrow \left\{\begin{array}{ll} X\,{}^tX\psi &= \sigma^2\psi \\ {}^tXX\varphi &= \sigma^2\varphi \end{array}\right. . \tag{6}$$

To compute the solutions one can either use the eigenequations or execute an iterative algorithm using the transition formulae. To find the second and further solutions it is added an orthogonal constraints (uncorrelated vectors) onto the ψ and φ :

$$\begin{aligned} \sigma_2 &= \max_{\substack{\|\psi\|_E=1 \\ \|\varphi\|_F=1 \\ \psi\perp\psi_1 \text{ and } \varphi\perp\varphi_1}} A..(\psi\otimes\varphi) \\ &= A..(\psi_2\otimes\varphi_2). \end{aligned} \tag{7}$$

Here with $k=2$, the orthogonality constraint can be written either $\psi\perp\psi_1$ and $\varphi\perp\varphi_1$, or $(\psi\otimes\varphi)\in(\psi_1\otimes\varphi_1)^\perp$, or with the subspace termed *orthogonal-tensorial* of the first principal tensor $(\psi\otimes\varphi)\in\psi_1^\perp\otimes\varphi_1^\perp$.
The SVD is then written as an orthogonal decomposition of A,

$$A=\sum_{s=1}^{rank(A)}\sigma_s\psi_s\otimes\varphi_s=\sum_{s=1}^{rank(A)}\sigma_s\psi_s\,{}^t\varphi_s=\psi\Lambda^1/2\,{}^t\varphi \tag{8}$$

in tensor form, or in vector form, or in matrix form. Note that here (for 2 modes) the collection of ψ_s and the collection of φ_s give also orthogonal systems within the respective spaces ; that will not be generally the case for $k>2$ modes.

3.2 SVD-$k$modes for $k=3$ and $k\geq 3$

Following a similar expression of a singular value one can write the maximisation problem to find the first singular value of $A\in E\otimes F\otimes G$:

$$\begin{aligned} \sigma_1 &= \max_{\substack{\|\psi\|_E=1 \\ \|\varphi\|_F=1 \\ \|\phi\|_G=1}} A..(\psi\otimes\varphi\otimes\phi) \\ &= A..(\psi_1\otimes\varphi_1\otimes\phi_1). \end{aligned} \tag{9}$$

Solving the Lagrange problem allows to compute the first solution using the following iterative algorithm (the iteration $(n+1)$ has three steps) where one can recognise a generalisation of the transition formula (6) :

$$\begin{aligned} (X..\varphi_{(n)})..\phi_{(n)} &= {}^1\sigma_{(n+1)}\psi_{(n+1)} \\ (X..\phi_{(n)})..\psi_{(n)} &= {}^2\sigma_{(n+1)}\varphi_{(n+1)} \\ (X..\psi_{(n)})..\varphi_{(n)} &= {}^3\sigma_{(n+1)}\phi_{(n+1)} \end{aligned} \tag{10}$$

For the second and other solution an orthogonality constraint is added, but unlike for two modes we do not have only the constraint of belonging to the *orthogonal-tensorial* of the first principal tensor. For example one can put the constraint of belonging to the subspace $(\psi_1 \otimes \varphi_1^{\perp} \otimes \phi_1^{\perp})$; these solutions associated to ψ_1 are easily obtained from a SVD (SVD-*(k-1)*modes in general) after contracting the tensor A by ψ_1. The straightforward generalisation of (9), (10) and of the second solution aspect, to the $k \geq 3$ case can be found in [16].
Through this recursive algorithm, two types of principal tensors can be found: the *k-modes solutions* when the constraint is expressed with an *orthogonal-tensorial*, and their *associated k-modes solutions* obtained by SVD-*(k-1)*modes.
For example if $k = 3$ one has for each *3*-modes solutions, 3 sets of associated *3*-modes solutions: one set for each component of the *3*-modes solution. If $k = 4$ there are two levels of associations: each *4*-modes principal tensor will have 4 sets of associated principal tensors, each set being obtained by the tensor product of a component s_h of this *4*-modes principal tensor and the SVD-*3*modes solutions of $X..s_h$ (where X is the initial tensor to analyse, and s_h is the component in question), and then each SVD-*3*modes will also have associated solutions as described before.
One can write the SVD-*3*modes of A as an orthogonal decomposition :

$$A = \sum_s \sigma_s \psi_s \otimes \varphi_s \otimes \phi_s \tag{11}$$

Because of some good properties of this method, mainly a generalised Eckart-Young theorem [16](*i.e. nested model optimisation*[3], not usually found in other generalisation in the literature) we will confound the PTA-*k*modes and SVD-*k*modes like we do with PCA and SVD. The singular values obtained on *k*-modes solutions are treated in decreasing order, but for example, it happens often that a singular value obtained with an *associated solution* of the first (or m^{th}) k -modes solution is bigger than the singular value obtained with the second (or next one) k -modes solution. In the listings one must notice that for this reason, we kept the logical order of computation instead of the "true" decreasing order of the singular values.

3.3 Handling PTA-*k*modes method

SAS/IML programs running with macro facilities have been written by the author to compute the SVD-*k*modes of a tensor of any order [13] with or without non-identity metrics (R-functions are also available [14]). As stopping rule for the decomposition to finish, one can ask for a maximum number of *k*-modes solutions at each level of the algorithm, controlled by minimum amount of variability. Playing with these two sets of parameters allows either to get as close as wished to the full decomposition, or to pick up interesting Principal Tensors. Figure 1 shows what a PTA-*3*modes output listing looks like. The 3-modes solutions are noted `vs111`, `vs222`, ... the associated 3-modes solutions to the first mode (`X`) are noted `Xvs11`, `Xvs22`... One must notice that on the list of values the first associated solutions `Xvs11` (or `Yvs11`, or `Zvs11`) are to be discarded in the decomposition as it is a repeat of vs111 because of the general algorithm: let sx_1, sy_1, sz_1 the first solution (*i.e.* $X..(sx_1 \otimes sy_1 \otimes sz_1) = vs111$, the solutions associated to sx_1 are obtained by the SVD-*2*modes of $X..sx_1$, therefore one finds again vs111 as the first singular values with solutions sy_1, sz_1. Nonetheless it is interesting to keep these repetitions because of the information given by the local decomposition (PCTloc: local percent of variability). `PCT` (respectively `PCTloc`) are in the percent of sum of squares (equal to percent of variance if the tensor is overall centred), then equals to the squared of the singular values divided by the *total* (*local*) sum of squares. In a PTA-3modes `PCTloc` refers to the usual percent of variability for an SVD; in general *total* refers to the original tensor analysed and *local* to the tensor *currently* decomposed *i.e.* associated solutions at a given level.
The notations are adapted for PTA-*k*modes (notations are slightly different with the R functions [14]), for example for a tensor of order 5:

(a) *k*-modes solutions: vs11111, vs22222,...

(b) associated *k*-modes first level: 1vs1111, 2vs1111..., 5vs1111, 1vs2222,...

(c) associated *k*-modes second level: 1vs111,...,4vs666...,

(d) associated *k*-modes third level: 1vs11,...,3vs66...

[3]The least square approximation of A up to $r' > r$ *orthogonal decomposed tensors*, contains the approximation up to r, and is the truncation of 11 up to r' terms (with a decreasing order of singular values).

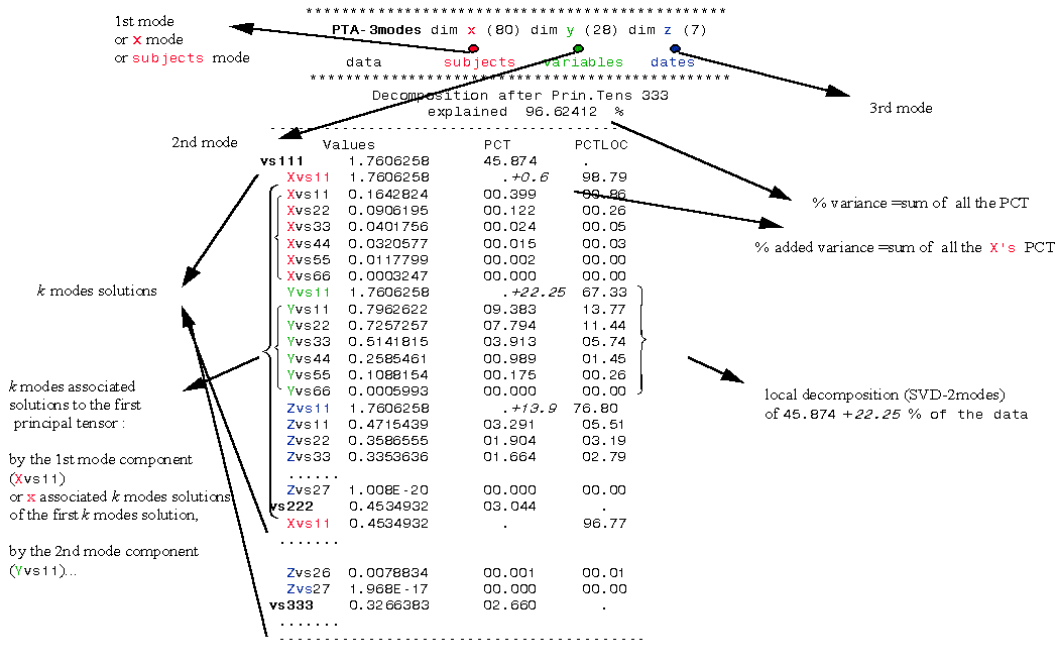


Figure 1: Example of output listing from PTA-3modes.

For each singular value, plots of components can then be produced using their normalised vector of coordinates. For the same Principal Tensor, plots of different components are read simultaneously as they correspond to the same singular value, but a basic rule must be kept to when interpreting the result. The sign of pairs of vectors are arbitrary, like in PCA, but unlike in PCA a solution is a triple of vectors (PTA-3modes) or a k-uple of vectors, then for example one has :

$$sx_1 \otimes sy_1 \otimes sz_1 = \tag{12}$$

$$(-sx_1) \otimes (-sy_1) \otimes sz_1 = (-sx_1) \otimes sy_1 \otimes (-sz_1) = sx_1 \otimes (-sy_1) \otimes (-sz_1).$$

So one must read the associations or oppositions of items from different components (*e.g.* modalities, variables, spatial configuration) considering the product of the signs of their coordinates, *i.e.* once the principal tensor has been mentally rebuilt. Plots of the same component for different Principal Tensors can be produced but one must be aware of possible non-orthogonality when they are not both k-modes Principal Tensors (remember the decomposition is orthogonal on the whole space not "completely" orthogonal in each space).

4 Using PTA-$k$modes for PDY studies

Based on the previously described drug experiment our purpose is now to show some practical examples of using PTA-$k$modes as a method of extracting the main results for pharmaco-EEG studies. The first consideration when using a multiway method is to define the structure of the data or the structure of interest. The whole structure of the data contains:

(a) a spatial dimension,

(b) a time dimension,

(c) a variable dimension (*e.g.* frequency bands in absolute energy),

(d) a dose dimension,

(e) a subject dimension,

(f) a condition dimension (Rest or Vigilance Control)

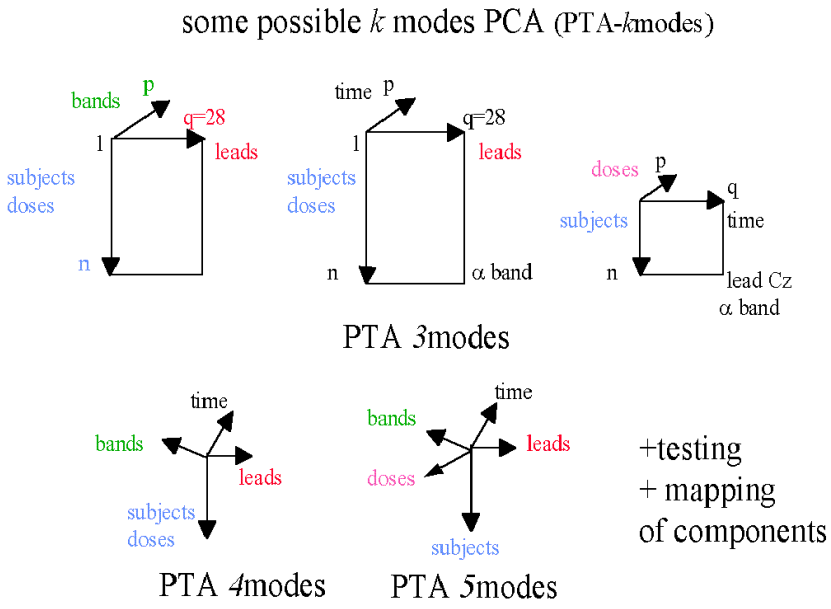


Figure 2: choices of modes for PTA-*k*modes of pharmaco-EEG data.

When performing a PTA-*k*modes we are looking for links between these dimensions through optimisation of linear modelisations of each in order to maximise variability. From geometrical point of view one looks for decomposed tensors or rank one tensors (*i.e.* tensor product of linear modelisations on each mode) giving the best projection of the data according to least squares error. The sum of squares explained is the sum of the squared singular values obtained, and the decomposed tensors are the corresponding principal tensors. Dimensions can be directly taken as modes or combined depending on focuses chosen for analysis. Figure fig.2 illustrates the main possible ways to organise the data to perform PTA-*k*modes for PDY studies, where each arrow means a mode of the tensor.
After choosing the *tensor to analyse* (choices of modes), *preprocessing* (*e.g.* centring, reducing) can be done using or not *metrics on each mode* (choice of a global tensorial metric), added *linear constraints* on some modes is also possible. All of those offer flexibility.
Pharmaco-EEG designs usually are cross-over designs, this makes possible the PTA-*5*modes shown on figure fig.2 which fully takes into account the repeated measure aspect of the design (one could also separate the sequence), but does not necessarily give a greater interest (see discussion). Secondly it also makes possible to build the *subject*dose* mode preferably with comparison to placebo (*i.e.* dose is in fact dose versus placebo). This will reduce subject variability effect. In the same manner *Time* mode is in fact often *Time* versus baseline. There after *dose* and *Time* will be considered respectively *versus* placebo and baseline.

4.1 Plotting Principal Tensor components for PDY studies

Special plots used here for pharmaco-EEG studies when using PTA-*k*modes with the choice of modes devised above are described as follows. Plot of *lead* mode (also called electrode mode or spatial mode) is the map of the leads with size of characters proportional to coordinates, in red (dark) for positive coordinates in cyan (clear) for negative coordinates. When appropriate a **one dimensional (horizontal)** plot of *subject*dose* mode is artificially split vertically according to dose membership joining each different the same subject across in order to describe the profiles, the mean at each dose is also plotted with a character size proportional to the standard deviation of the dose group. This last plot is sometimes stacked with the plot of bands (again artificially spread vertically). As the components are normalised to one, only the relative differences within the components are of interest and the sign of the values: the indications on the axe may be limited to positioning the zero (vertical line).

4.2 A first analysis

Before going further the investigation of interesting choices to perform the analysis, let us have a glimpse at the kind of results the PTA-*k*modes provides on our actual data.

Table 1: PTA-3modes of *total* energy for *verum versus placebo versus* 1^{st} *baseline*: listing of the decomposition up to the second *k*modes and associated solutions.

```
*******************************************
  PTA-3modes  dim x (36) dim y (28) dim z (9)
  data    dose*sujets    electrodes   time
  pdy2833
*******************************************
total band day 1 vs bl: verum vs plb
-----------------------------------
  Decomposition after Prin.tens 222
          explained 96.314621 %
-----------------------------------
Values                PCT      PCTloc
 vs111    10.264229   59.128 %   .     vs222    2.2504552   02.842     .
  Xvs11   10.264229     .       98.65    Xvs11   2.2504552     .      82.49
  Xvs11   1.0391565   00.606    01.01    Xvs11   0.8390228   00.395   11.47
  Xvs22   0.3847897   00.083    00.14    Xvs22   0.4463147   00.112   03.24
  Xvs33   0.2981522   00.050    00.08    Xvs33    0.319963   00.057   01.67
  Xvs44   0.2686759   00.041    00.07    Xvs44   0.1950296   00.021   00.62
  Xvs55   0.1569165   00.014    00.02    Xvs55   0.1361017   00.010   00.30
  Xvs66   0.1193496   00.008    00.01    Xvs66   0.1151387   00.007   00.22
  Xvs77   0.1046065   00.006    00.01    Xvs77   4.105E-17   00.000   00.00
  Yvs11   10.264229     .       74.28    Yvs11   2.2504552     .      59.11
  Yvs11   3.8071438   08.135    10.22    Yvs11   1.0813549   00.656   13.65
  Yvs22   2.7134542   04.132    05.19    Yvs22   0.9753945   00.534   11.10
  Yvs33   2.3470325   03.092    03.88    Yvs33   0.7212513   00.292   06.07
  Yvs44   2.2247245   02.778    03.49    Yvs44   0.6557279   00.241   05.02
  Yvs55   1.4559578   01.190    01.49    Yvs55   0.5062663   00.144   02.99
  Yvs66   1.1157027   00.699    00.88    Yvs66   0.4203873   00.099   02.06
  Yvs77   0.8935416   00.448    00.56    Yvs77    1.46E-16   00.000   00.00
  Zvs11   10.264229     .       86.81    Zvs11   2.2504552     .      65.33
  Zvs11   2.3062862   02.985    04.38    Zvs11   0.7975893   00.357   08.21
  Zvs22   2.1126748   02.505    03.68    Zvs22   0.6559268   00.241   05.55
  Zvs33   1.3551495   01.031    01.51    Zvs33   0.5516294   00.171   03.93
  Zvs44   1.0262505   00.591    00.87    Zvs44   0.5009633   00.141   03.24
  Zvs55   0.8132626   00.371    00.54    Zvs55   0.4359773   00.107   02.45
  Zvs66   0.7534717   00.319    00.47    Zvs66   0.4117442   00.095   02.19
  Zvs77   0.6554737   00.241    00.35    Zvs77   0.3741479   00.079   01.81
  Zvs88   0.5877239   00.194    00.28    Zvs88   0.3511398   00.069   01.59
  Zvs99   0.5511592   00.170    00.25    ...
                                        ---------------------------------
```

Using the statistical method of [3], briefly described in section 2, results reported from an independent body were summarised as follow:

> "The most consistent treatment effects observed are a **dose dependent increase of the** *total* **energy** with a **peak measured around time 2.5-3h** post-dosing, and a spectral redistribution of energy in **favour of the** δ **frequency band**. These effects are detected already after 0.5-2h and are maximal around 3h post-dosing, as all the other significant effects observed. In addition to the increase observed for the δ band, a **selective increase of the** β_1 band and **a reduction of the** α_1 band is observed around 3h post-dosing."

In order to try to replicate these results, a first analysis has been made only on *total* band with a *time* dimension, an *electrodes* dimension, and a *subject*dose* dimension. The conclusion of the official report seems to be reflected in this analysis: dose-dependent increase of total energy with a peak around 12h (3h post dosing). But on fig.3 it is possible to see that the main variation describing this conclusion, 59.128%, is explained by subject 11 at dose 30mg. Nonetheless this conclusion is confirmed (see table 2) by two tensors explaining respectively 4.13% and 3.09% of the variability (could they be increased if subject 11 was discarded?). Notice a slight left occipital (O2, T6) spatial preference. Complete analysis of the data is not the purpose here (see [11, 12]), but the method enabled to extract similar results, especially when using 4-modes analysis to describe the redistribution of the

Table 2: Dose differences tests: Friedman's test (expected sum of ranks E(Sr)=24) for fig.3 and table 1 ; *(WSR)*: p-value of Wilcoxon Signed-Rank test for *xx.mg vs plb*.

Source		*Friedman's* p-value	*10mg vs plb* Sr	*(WSR)*	*30mg vs plb* Sr	*(WSR)*	*90mg vs plb* Sr	*(WSR)*
sx111	59.128	0.33	20	(0.42)	25	(0.67)	27	(0.09)
sxsy111	8.13%	0.33	27	(0.96)	21	(0.33)	24	(0.79)
sxsy111	4.13%	0.01	18	(0.79)	22	(0.20)	32	(0.01)
sxsy111	3.09%	0.07	19	(0.62)	23	(0.46)	30	(0.07)
sx222	2.84%	0.77	22	(0.23)	25	(0.38)	25	(0.15)

total energy (see further as well). For example the redistribution in favour of δ was found but seemed as strong as for θ.
This method for PDY studies can provide results and conclusions in a more concise manner and offers a more descriptive aspect of the result leading for example to a better understanding and criticism of it *e.g.* outliers. As the sample sizes are usually relatively small (here 12 subjects) possible outliers have a greater impact on the results.
Subject 2 seems to particularly affect the 2^{nd} Principal Tensor (not significant for dose differences in table 2). Notice the *back front* opposition seen on this tensor, more consistent as the dose increase (comparing spread for each dose), reversing after the peak of activity. This was something expected by the neuro-pharmacologist. The problem of outliers may be handled in different ways. In the next section preprocessing of the data is investigated as a way of "targeting" the analysis but also of minimising outliers effects.

5 Preprocessing before a PTA-$k$modes

Preprocessing such as centring and/or scaling enables the analysis to focused on a chosen variation of interest. It can sometimes be a solution to the problem of *subject* effect and outlier effects (effects occurring in the pharmaco-EEG data, see [11]). Centring and/or reducing variables before analysis is common in multivariate analysis such as PCA. Usually centring is seen as a simple statistical model focussing on residuals from a regression model and has connections with algebraic and geometrical properties, such as being the projection onto the orthogonal of the subspace generated by the regressors. With this interpretation and leaving, aside the statistical sampling who generated the data, it is possible in fact to centre and/or reduce on any mode or combined mode of our multi-entries data.
Remembering the structure of our data described at the beginning of section 4, the question is now "what are we looking for?" which should guide our decisions with respect to centring/reducing. The question could be formulated as "what are we not interested in?". At first sight the answer to this one is *subject* differences, but also the interactions of *subject* and the other "dimensions". If the data is firstly whole centred, centring and removing effects (as in ANOVA) are linked. To understand this point consider the problem in two modes with x whole centred, *i.e.* $\bar{x} = x_{..} = \sum_{ij} x_{ij} = 0$:

$$x^c_{ij} = x_{ij} - x_{.j}$$

is called in [20] centring *across* the first mode, and is equivalent in ANOVA language to remove the effect of the second mode. This is the reason why one usually centres across the *mode of interest*. It is possible to give an algebraic expression to this transformation : with X a tensor of order k when centring across say the second mode, X is transformed to :

$$X^c = (Id_{n_1} \otimes (Id_{n_2} - P_{\Delta_{n_2}}) \otimes \cdots \otimes Id_{n_k})X \tag{13}$$

where $P_{\Delta_{n_2}}$ is the orthogonal projector onto the vector $\Delta_{n_2} =^t (1, 1, \ldots, 1)$ of length n_2, and Id_{n_u} is the identity operator onto $\mathbb{R}^{n_u}$. The expression (13) is using the tensor product of linear operators (see also next section) which is isomorphic to the Kronecker product of their matrices (for any given basis choices). Performing a double centring say across the first mode and across the second mode can be written :

$$X^c = ((Id_{n_1} - P_{\Delta_{n_1}}) \otimes (Id_{n_2} - P_{\Delta_{n_2}}) \otimes \cdots \otimes Id_{n_k})X \tag{14}$$

It easy to show that it is equivalent to perform the two single centring one after the other. Care is needed in multiple centring and /or reducing involving centring across *slices* (2 modes varying) as well, as one can cancel or slightly modify the other centring. This is because doing successively *mode* centring and *slice* centring may

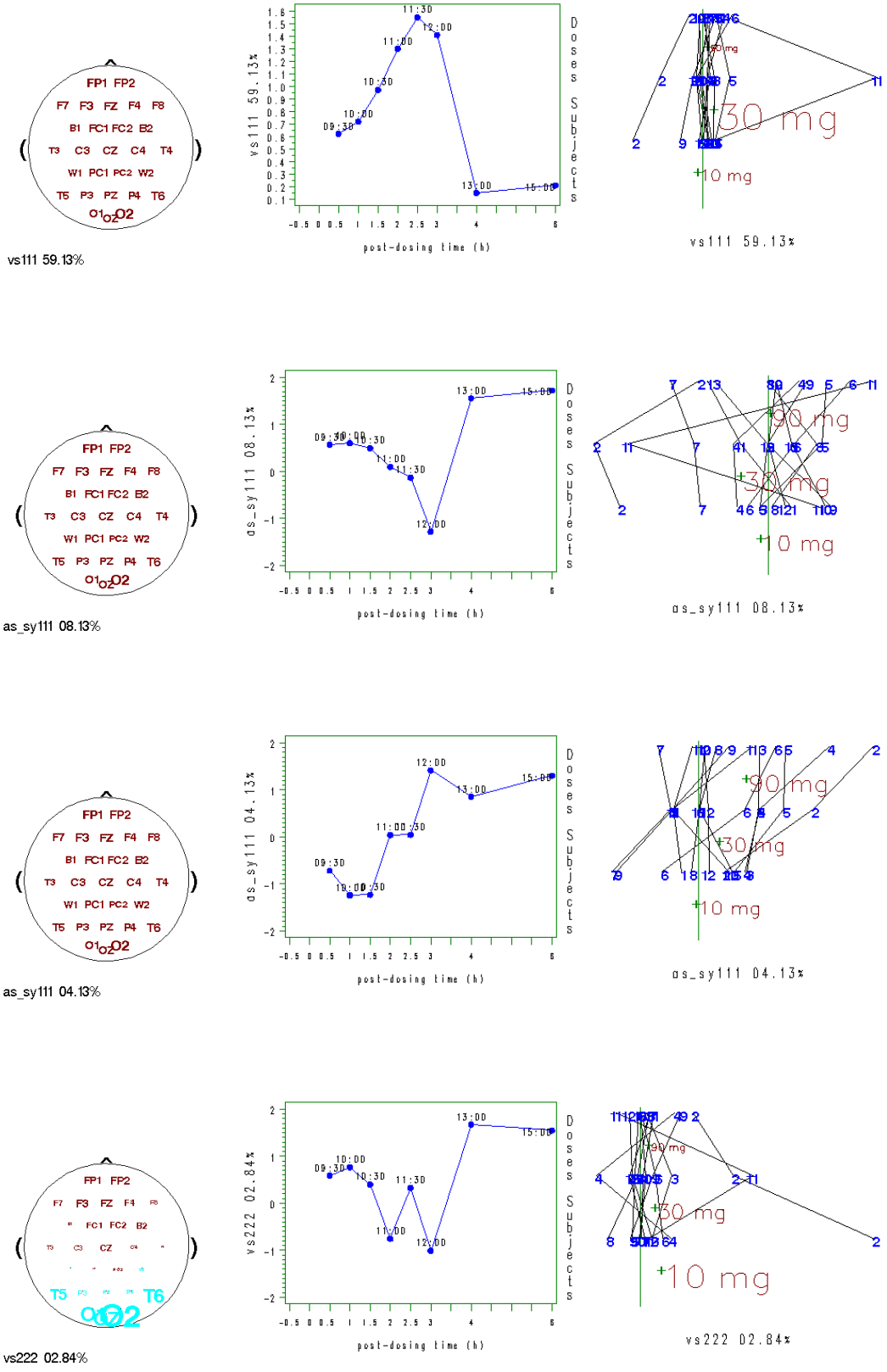
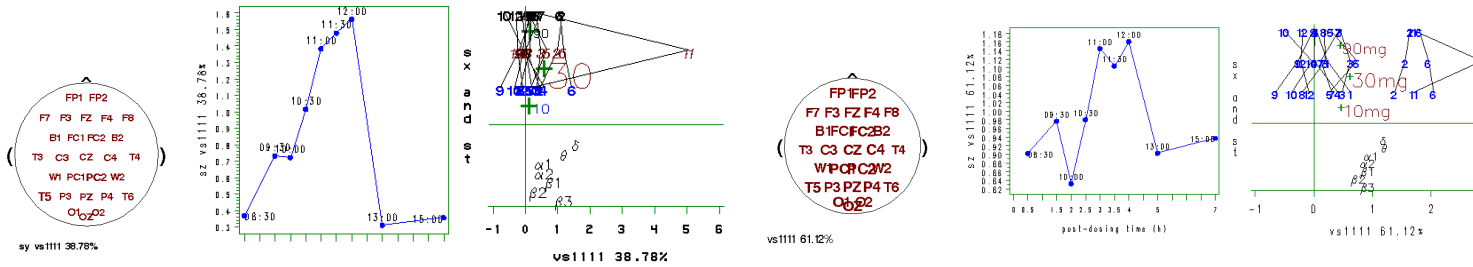
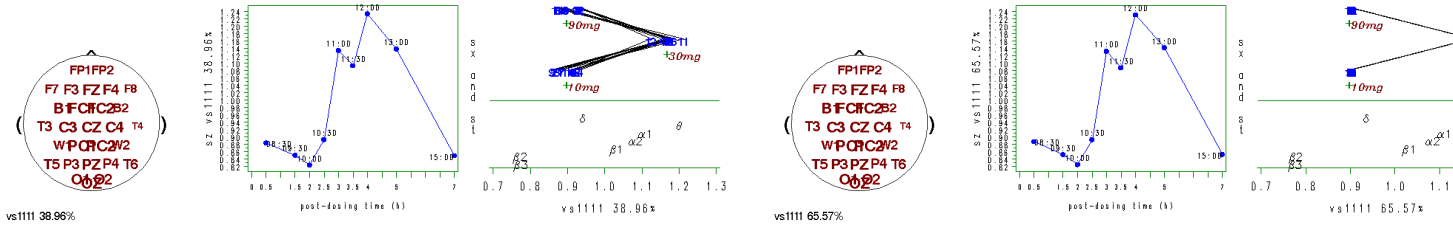


Figure 3: PTA-3modes *total* energy for verum versus placebo versus 1^{st} baseline: tensors of table 2, analysis table1



(a) Original data

(b) Subject scaled data



(c) *globally-modified data*

(d) *levels-modified data*

Figure 4: PTA-4modes of all bands (absolute energies) for verum versus placebo versus first baseline: 1^{st} Principal tensor (a) original data, (b) subjects scaled to unit, (c) *globally-modified data* (d) *levels-modified data*; (preprocessing b, c, and d explained in the text).

break the tensorial structure. A simple example involving only slice centring and showing a broken tensorial structure would be to centre across slice [mode 1 and mode 2] and then across slice [mode 2 and mode 3]! For such situations, iterative centring and scaling can be also be thought as for the PARAFAC method in [20] but this preprocessing algorithm may then become a true modelling part of the analysis for which interpretation and analysis of the *explained part* $(X - X^c)$ may be needed.

For EEG data it is possible to have this ANOVA like approach for all the entries (at least when looking at absolute energies) as one can consider the data as a measure of the EEG amplitude on the subjects at the repeated conditions : electrodes, frequency bands, time, and doses. Note the structure of the data ensures a balanced design and so orthogonality of factors for the ANOVA. The problem in using this ANOVA approach is that if distributions are not normal with small variances, a factor effect measured with *means* will not be completely removed, *i.e.* the variation left may be still important. Others structures than the one giving the mean model can be considered for an entry h, the formula (14) can be reformulated replacing the model Δ_{n_h} by the appropriate G_{n_h} (a "design" describing the structure usually including Δ_{n_h}) .

Scaling (or reducing) variables is commonly used when the variable units are not the same. For EEG measurements *subjects* can be thought to have their own units. Global *subject* differences in location and in variability are not of interest for our purpose, so reducing their variability to the same unit would also improve the analysis. Unfortunately centring or removing effects does not insure vanishing outliers, but sometimes successfully diminishes the variability induced by their presence so that it appears in a less important (lower singular value) Principal Tensor. A complete illustration of this fact is shown in [12]. On fig.4 a comparison of the first principal tensor obtained from the data *[dose*subject* $\times$*lead* $\times$ *time* $\times$*band]* with different preprocessing is shown .

To modify the data using the ANOVA approach it is possible to remove interactions of each factor (mode) with subjects *globally* or by *levels* of other factors. For example on the subject scaled data was removed : *subject.dose* by time and band, *subject.band* by time and dose, *subject.time* by band dose, and, *subject.electrode* by band dose. Notice the first interactions are then computed on the electrode units, and the last one is computed on time units. This way of proceeding can make more sense than computing these interactions globally on the rest of the units, and actually provided better results. A verification was done in comparing full ANOVA models (the *subject* factor being the experimental units) respectively on the *subject scaled data*, the *globally-modified data* (interactions and main subject effect removed), the *levels-modified data* (as before but by levels). We obtained as explained respectively : 5%, 13%, and 35%, which means some unwanted variation in the data were

successfully removed.

6 Analysing summaries and PTAIV-*k*modes

Analysing a summary statistic of subjects changes obviously the data analysed, and clearly it means that this summary "sufficiently" informs on the distributions. Analysing means, medians or trimeans can be also a solution to subjects outliers which put into question the choice of the location parameter. Looking at only one location parameter implicitly suppose the distributions to be unimodal, which was a sensible assumption here but PTA-*k*modes could be done with a mode representing different location parameters of the empirical subject distributions. This has not been done here as we focused on comparing different main location parameters using a PTA-3modes.

On fig.5 the first Principal Tensor of the different PTA-3modes on *means*, *medians* or *trimeans* over the 12 subjects for each dose , band, time, and electrode, as a tensor of order three, is shown. Each analysis constitutes a *dose profile analysis*, each left plot of figure fig.5 representing a dose-effect curve (*versus* time) for the corresponding principal tensor of the given profile summary. For the 1^{st} mode (*dose* $\times$ *time*) a major difference between these three summaries can be seen for the *30mg* curve in comparison to the other doses: no apparent dose effect (around peak) is observed with the *mean*). The *median* and *trimean* give similar results, the *10mg* curve becomes flatter for *trimean*. For the 2^{nd} mode (*electrode*) a slight gradient towards the back is seen for *median* and *trimean* but the three plots are very similar. The 3^{rd} mode (*band*) representation for *mean* differs from the two others mainly on δ, β_2 and α_2.

With small samples the trimean ($0.25q1 + 0.5median + 0.25q3$) seems to be a good compromise between the two extremes of mean and median either too sensible to outliers or not all. It has been already used in pharmaco-EEG studies for example by [8]. Analysing *means* of the subjects is in fact equivalent to performing a PTAIV-3modes on the three-ways arranged data (*i.e.* tensor of order three) X with the following modes : ($dose \times subjects \times time$) as the first mode, *electrodes* as the second mode, and *bands* as the third mode. PTAIV means Principal Tensor Analysis with Instrumental Variables and refers to an extension of PCAIV, [18] or [19], to multiway data, [10]. In the optimisation procedure one considers linear constraints on the solution defined by the Instrumental Variables which are usually linked to the design. In our context the optimisation becomes to maximise $X..(\psi \otimes \varphi \otimes \phi)$ with a linear constraint on ψ as belonging to the subspace generated by the indicator matrix of *dose* $\times$ *time* structure S_{dt}, $\psi \in \mathcal{I}m(S_{dt})$. S_{dt} is a matrix of $3 \times 9 = 27$ columns, each one identifying entries of the first mode as in the current *dose* and *time* by a value 1, 0 otherwise). This means that the values in ψ will be equal for all the units with the same *dose* and *time*. Writing the maximisation to find a singular value gives (denoting $x, y, z \in \mathbf{S}_1$ for $\|x\|_{E_1} = 1$, $\|y\|_{E_2} = 1$ and $\|z\|_{E_3} = 1$:

$$
\begin{aligned}
\sigma &= \max_{x,y,z \in \mathbf{S}_1 \text{ and } x \in \mathcal{I}m(S_{dt})} X..(x \otimes y \otimes z) \\
&= \max_{x,y,z \in \mathbf{S}_1 \text{ and } x \in \mathcal{I}m(S_{dt})} < X, x \otimes y \otimes z >_{E_1 \otimes E_2 \otimes E_3} \\
&= \max_{x,y,z \in \mathbf{S}_1} < X, P_{S_{dt}} x \otimes y \otimes z >_{E_1 \otimes E_2 \otimes E_3} \\
&= \max_{x,y,z \in \mathbf{S}_1} < (P_{S_{dt}} \otimes Id_{E_2} \otimes Id_{E_3})^* X, x \otimes y \otimes z >_{E_1 \otimes E_2 \otimes E_3} \\
&= \max_{x,y,z \in \mathbf{S}_1} ((P_{S_{dt}} \otimes Id_{E_2} \otimes Id_{E_3})X)..(x \otimes y \otimes z) \\
&= ((P_{S_{dt}} \otimes Id_{E_2} \otimes Id_{E_3})X)..(\psi \otimes \varphi \otimes \phi). \qquad (15)
\end{aligned}
$$

Equality in(15) means that PTAIV-*k*modes is performed as a PTA-*k*modes of the projected data $(P_{S_{dt}} \otimes Id_{E_2} \otimes Id_{E_3})X$ which in that case will be equivalent to analyse the means data by *dose* and *time* for each *band* and *electrode*. Note that analysing $(P_{S_{dt}}^{\perp} \otimes Id_{E_2} \otimes Id_{E_3})X$ corresponds to the residual analysis (projection on the orthogonal of the structure). Thus one has a double decomposition of the sum of squares : explained by the structure plus its residuals, and within each part with the SVD-*k*modes decomposition.

Rigourously when analysing other summaries such as *medians* (*idem* for *trimean*), one does not perform a PTAIV, nonetheless defining a structure by putting a 1 only for the median point (for each *dose* and *time*), which depends on the *band* and *electrode* (the structure is on the whole tensor space, would provide a PTA on a projected data.

Comparisons *versus* baseline for *time* mode and comparison *versus* placebo for *dose* mode (the modes have been understood that way all through), are in fact already preprocessed summary measures and can be seen as

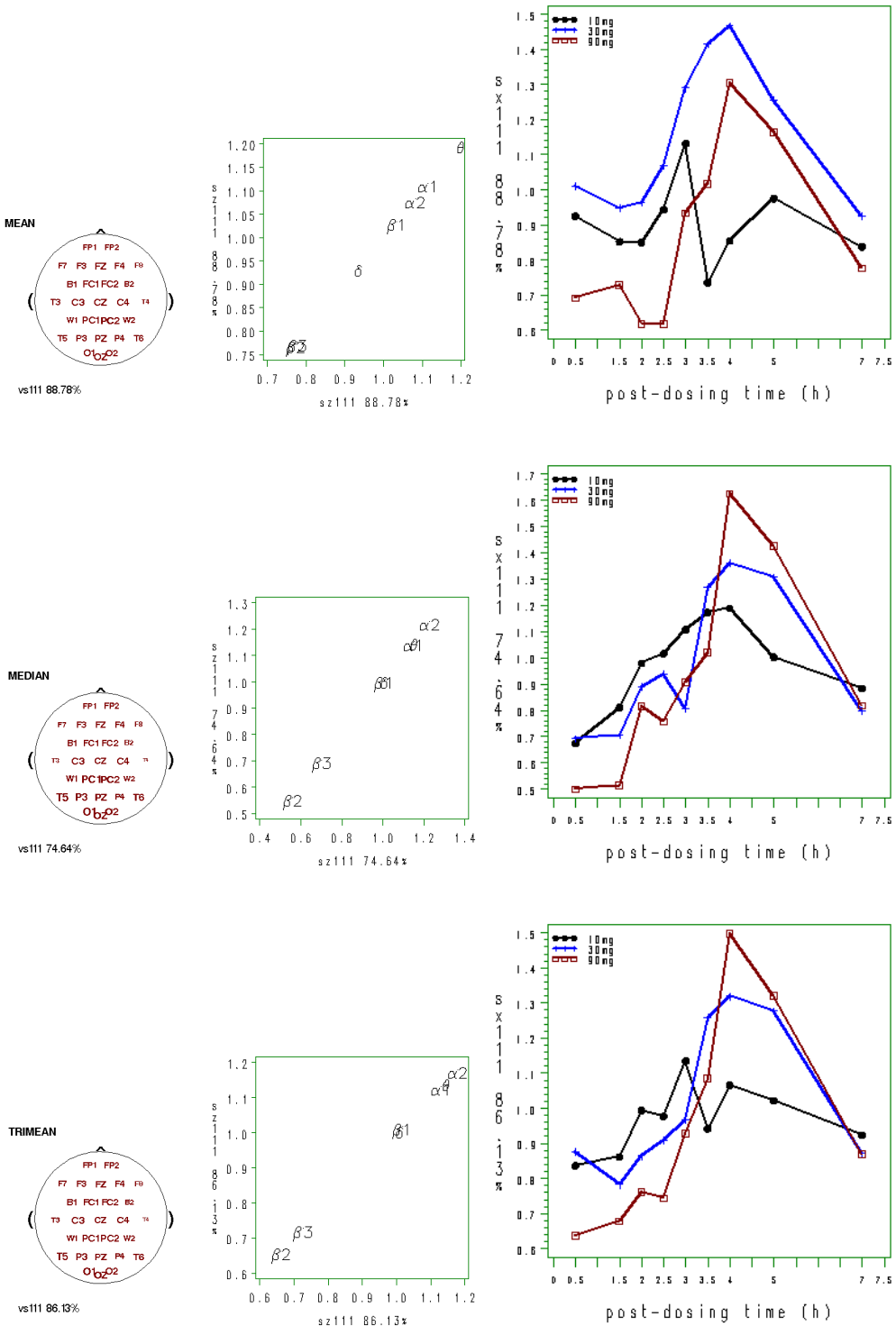


Figure 5: PTA-3modes dose means, medians or trimeans (*data subject scaled*) for all bands (absolute energy) for verum *versus* placebo *versus* first baseline : 1^{st} Principal Tensors.

Table 3: PTAIV Variability explained for different summaries of X. (* =rounded; last column: % relatively to the corresponding source ≡ %relatively to the original (data X).)

Source	Sum of Squares	% explained*	1^{st} P.Tensor %
data X	2.28	100%	-
Means projected	0.49	22%	88.78% ≡ 19%
Medians projected	0.30	13%	74.64% ≡ 10%
Trimeans projected	0.37	16%	86.13% ≡ 14%

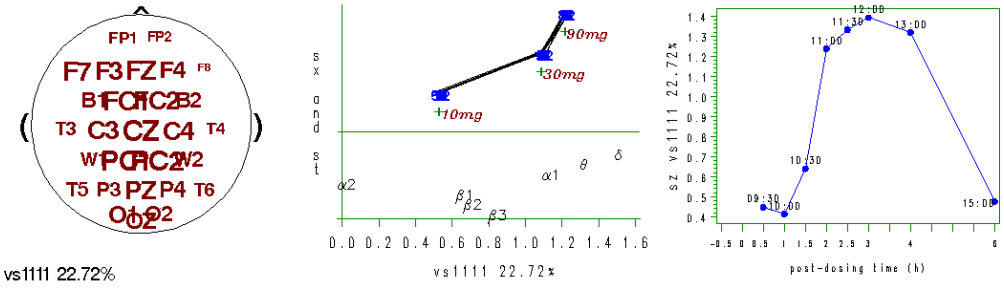


Figure 6: PTA-4modes *levels-modified data* (see page.12) for all bands (absolute energy) for verum *versus* placebo *versus second* baseline : 1^{st} Principal Tensor.

projected data. It is likely that when performing pharmaco-EEG experiments, designs contain two baselines as in the methodology described in introduction. Apparently the second baseline is usually taken as the reference in statistical analysis: a greater stability is often observed with this measure(less subject variability). So far in this paper the analysis has been done *versus* first baseline (08h00), and further analysis will be done *versus* second baseline (08h30), the injection was in fact done at 09h00. In all the previous analysis the second baseline (*versus* the first) was in fact included in the analysis, this means a true *post dosing time* is to be decreased by one in the graphics.The only interest of including a (second) baseline in the *analysis* is when studying the possible *initial drop of activity* just after injection. Nonetheless to achieve less random subject effect, and a better comparison with reported results, thee foregoing analysis will be done with the second baseline. It is reassuring that similar results were obtained concerning the peak time activity but closer results, to the officially reported ones, concerning the dose dependent δ band favoured for the total band effect seen the *mean* analysis with nonetheless θ as much important. These results were also confirmed by a PTA-4modes analysis, seen on fig.6.
Notice the importance of this second baseline choice towards the dose effect seen on fig.7 and fig.6 not observed with the first baseline on the same analysis (*c.f.* fig.5 fig.4(d)). The band components are quite different as well pointing now the δ redistribution but also θ. One could see a gradient in the slow waves and after the fast waves, but α_2 (mid-range) is now out of pattern in these first Principal Tensors and as matter of fact does not contribute to this Principal Tensor. The spatial components are very similar but seem more central with the analysis *versus* second baseline.

6.1 Supplementary points

Performing a multiway analysis on qEEG data, based on the subjects summary, somehow gives a robust *description* . A posteriori validation using subject measures can be done using supplementary points technic to compute subjects scores. Supplementary points is similar to "prediction" as one supposes the *description* or the model known and want to compute the outcome given the new observation. By analogy of its use in PCA: let $\sigma_h(s_1 \otimes s_2 \otimes s_3)$ be a Principal Tensor of the PTA-3modes of the data,
$X = P_S Y$ [dose*time ×lead ×band], *i.e.* $X..(s_1 \otimes s_2 \otimes s_3) = \sigma_h$, one then compute the *subject*dose*time*'s supplementary points of the data Y [subject*dose*time × lead × band] by $\tilde{s}_1 = \frac{1}{\sigma_h}(Y..(s_2 \otimes s_3))$. If $P_S = P_{S_{dt}} \otimes Id_l \otimes Id_b$ performed the mean over subjects summary, one would have *in most of the cases* $\tilde{s}_1^{dt} = 1/n_s \sum_s \tilde{s}_1^{sdt} = s_1^{dt}$, *i.e.* one retrieves the component s_1. It is then possible to plot standard error of means (SEM) for example, as on the figure fig.8. It is analogue as to do a multivariable regression (*subjects*dose*time* are the variables) onto

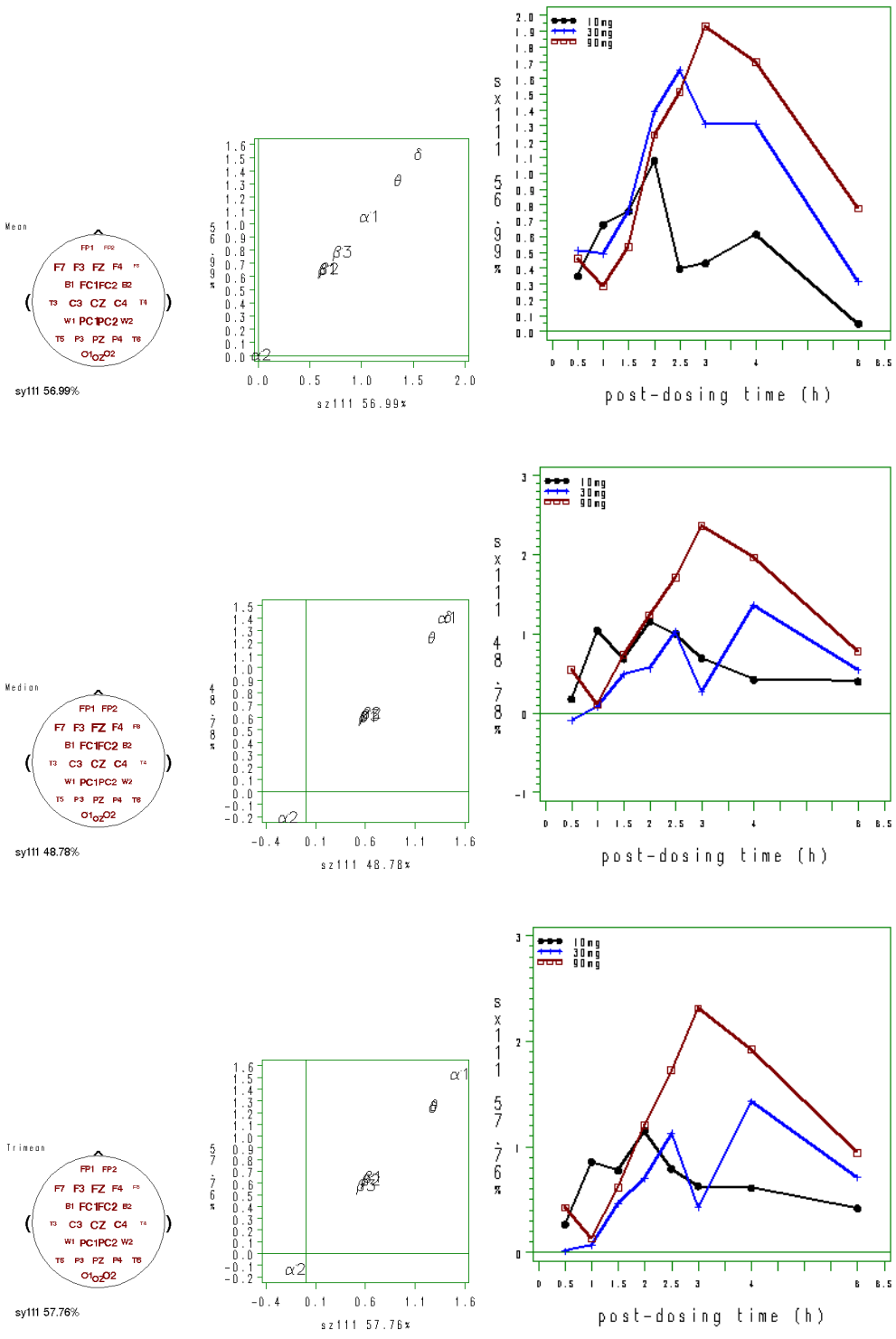


Figure 7: Same as for figure 5 but *versus second* baseline : 1^{st} Principal Tensors.

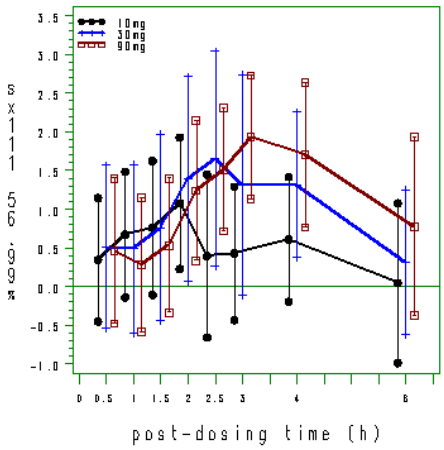


Figure 8: Use of supplementary points for SEM plots on the first Principal Tensor (fig.7).

one variable $(s_2 \otimes s_3)$.
It is mentioned "*in most of the cases*" because $X..(s_2 \otimes s_3)$ may not be equal to $\sigma_h s_1$, as the orthogonal decomposition is on the whole space. Then the "supplementary points" denomination become misleading and will be called *pseudo-supplementary points*. They will generate a different summary component, sum of the original one plus some orthogonal to it. Looking at the orthogonality constraints in the decomposition it is relatively easy to check that to first Principal Tensor will always generate true supplementary points and every $k$modes Principal Tensor could generate pseudo-components mixing only with previous associated solutions (in the order of $k$modes solutions). Those pseudo-components generated from supplementary points can be also generated from the decomposition itself and may be worthwhile to consider for post-analysis. Another way of performing supplementary points is as follow: let $\sigma_h(s_1 \otimes s_2 \otimes s_3)$ be a Principal Tensor of the PTA-3modes of the data,
X [dose*time ×lead ×band], *i.e.* $X..(s_1 \otimes s_2 \otimes s_3) = \sigma_h$, one then compute the *subject*'s supplementary points with the data Z [dose*time × lead × band × subjects] considering $Z = (\sigma_h s_1 \otimes s_2 \otimes s_3) \otimes z_4 + \epsilon$ gives $z_4 = 1/\sigma_h Z..((s_1 \otimes s_2 \otimes s_3)$. Notice that this way every subject "profile" are proportional, as it is a rank one approximation of the previous method, *i.e.* $\tilde{s}_1 = z_4 \otimes s_1 + e$.

7 Non-Identity metrics in PTA-$k$modes

The general method of SVD-$k$modes can be performed with non-identity metrics. This means that inner products can be considered weighted (diagonal metrics) or cross-weighted (non-diagonal metrics). The whole algebraic setup used at the beginning of the paper is the same if one understands the contracted product (operation ..) as containing the metrics. For example, equations (5) expressing the classic SVD become :

$$\begin{aligned} \sigma_1 &= \max_{\substack{\|\psi\|_E=1 \\ \|\varphi\|_F=1}} A..(\psi \otimes \varphi) \\ &= A..(\psi_1 \otimes \varphi_1) \quad (\text{ in tensor form}) \\ &= {}^t\psi_1 D_E A D_F \varphi_1 \quad (\text{ in matrix form}); \end{aligned} \tag{16}$$

where spaces E and F have now respectively the metrics D_E and D_F (instead of the identity metrics). The only change is in the last expression (matrix form) because the metrics are "included" in the contracted product operation as well as in the norms (as defined from the inner product) ; equation (6) is as well :

$$\left\{ \begin{array}{ll} X..\varphi & = \sigma\psi \\ X..\psi & = \sigma\varphi \end{array} \right. \text{ or } \left\{ \begin{array}{ll} XD_F\varphi & = \sigma\psi \\ {}^tXD_E\psi & = \sigma\varphi \end{array} \right. \Rightarrow \left\{ \begin{array}{ll} XD_F^tXD_E\psi & = \sigma^2\psi \\ {}^tXD_EXD_F\varphi & = \sigma^2\varphi \end{array} \right. . \tag{17}$$

Consideration of metrics offers flexibility to the analysis. For example with two modes it is classical to recognise a discriminant analysis as a SVD on the projected data with Σ^{-1} (the inverse of the covariance matrix) as metric on the space defined by the variables or as a SVD of the original data with W^{-1}, the inverse of the within covariance matrix. Roughly speaking, when looking for "best directions", directions of high *within* variation have a lower weight than the other directions (see also [2] when the group structure is not known). Another good

example of non-identity metrics is also in the next section, correspondence analysis. Nonetheless the method offers the possibility only of decomposed whole space metrics, *i.e.* of the form:

$$M_{E_1} \otimes M_{E_2} \otimes M_{E_3} \ldots \otimes M_{E_k} \tag{18}$$

where every metrics as algebraic object (self-adjoint linear operator) is confounded with its definite or semi-definite positive matrix representation (like X as tensor and array or vector). The tensor product operation is the one for linear operators (see also [4]). It is left with the same notation (as for vectors) because it is possible to confound the algebraic notation and arithmetic, as well . This is because it becomes the Kronecker tensor product sometimes called the outer product which operates either on vectors or matrices. One must note that (18) is a linear operator onto the whole space,which operates separately onto every space defining the tensor space (this is in fact the definition of the tensor product of linear operators). Arithmetically and computationally this can be written:

$$\begin{aligned}(M_E \otimes M_F \otimes M_G)(Y) &= (M_E \otimes M_F \otimes M_G)(Y) \\ &= (M_E \otimes M_F \otimes M_G)(\sum_u e_u \otimes f_u \otimes g_u) \\ &= \sum_u (M_E e_u) \otimes (M_F f_u) \otimes (M_G g_u)\end{aligned} \tag{19}$$

Without knowing the decomposition of Y, this last expression cannot be used, nonetheless isomorphism properties within multilinear maps (tensor) can be used to perform successively the different operators (*e.g.* $Y \in E \otimes F \otimes G \sim \mathcal{L}(E^*; F \otimes G)$ then 19 is equivalent to $(M_F \otimes M_G)(Y \circ M_E)$ where $\circ$ stands for composition of applications or matrix multiplication). The contraction product includes the metrics using this property and could also have been understood as a canonical contraction product (without metrics) of the transformed tensor (the contracting one) by the metric operators, *i.e.* the canonical contraction would be using only the dual product instead of the inner product:

$$Y..z = Y..{}_c(M_E z) = [Y \circ M_E]..{}_c z \tag{20}$$

What is a good choice of metrics for pharmaco-EEG studies ? Generally Choices are geared towards "elimination" of unwanted variation, such as in discriminant analysis one do not want to relate the within group variation. For pharmaco-EEG data it would be interesting to eliminate natural variation of bands and electrodes as well as within dose variation. To achieve estimation of natural variation, enough placebo observations or better some "null" data on the subjects studied are needed. Metric choice and their estimation is a key issue in multidimensional and/or multiway analysis particularly for this kind of data and deserves more attention (see [15]).

8 k-modes Correspondence Analysis

Choice of metrics offer the possibility to perform generalisations of established multivariate methods. An interesting one for pharmaco-EEG data is a generalisation of correspondence analysis. The purpose of this analysis of a multiple contingency table is to break down the whole χ^2 statistic as the sum of squared singular values which are associated with Principal Tensors giving a description of lack of independence. Although usually applied to contingency data, a correspondence analysis approach is valid here as each cell is a measure of energy amplitude according to a particular frequency band, a particular lead, a particular time, etc..., so the whole count and marginals have a meaning of energy amplitude. The literature has been abundant regarding correspondence analysis methods for more than two variables but usually looks at two by two lack of independence and not the lack of complete independence. Using the PTA-$k$modes framework the extension from 2 to k variables is straightforward the analysis of the lack of complete independence.

8.1 FCA-2modes

Correspondence analysis of a two-way contingency table with cells $n_{ij}, \quad i = 1 \cdots I, \quad j = 1 \cdots J$ can be described as follows. The usual notations are:

$$n_{i.} = \sum_j n_{ij}, \quad n_{.j} = \sum_i n_{ij}, \quad n_{..} = N = \sum_{ij} n_{ij}$$

and then the observed proportions are defined as $p_{ij} = n_{ij}/N$. Diagonal metrics containing vector margins $P_{I.} = {}^t(\cdots p_{i.} \cdots)$ and $P_{.J}$ used thereafter are noted D_I and D_J. Correspondence analysis provides a decomposition of the measure of lack of independence between the two categorical variables indexed respectively by i and j in performing the PCA (or generalised PCA) of the following triple ([5]):

$$(D_I^{-1} P D_J^{-1} \quad - 1\!\!1_{IJ}, \quad D_I, \quad D_J) \tag{21}$$

where the triple is defined as *(data, metric on $\mathbb{R}^I$, metric on $\mathbb{R}^J$)*. The measure of lack of independence can be written :

$$\frac{\chi^2}{N} = \sum_{ij} \frac{(p_{ij} - p_{i.}p_{.j})^2}{p_{i.}p_{.j}} = \sum_{ij} p_{i.}p_{.j}(\frac{p_{ij}}{p_{i.}p_{.j}} - 1)^2 = \sum_s \sigma_s^2 \tag{22}$$

where the σ_s are the singular values of the PCA of the triple given above. From the data reconstruction formula, one can write for $r \leqslant min(I-1, J-1)$:

$$\frac{\widehat{p}_{ij}}{p_{i.}p_{.j}} = 1 + \sum_{s=1}^{r} \sigma_s \psi_{is} \varphi_{sj} \tag{23}$$

or equivalently in a tensor form:

$$D_I^{-1} \widehat{P} D_J^{-1} \quad = 1\!\!1_{IJ} + \sum_{s=1}^{r} \sigma_s \psi_s \otimes \varphi_s = 1\!\!1_I \otimes 1\!\!1_J + \sum_{s=1}^{r} \sigma_s \psi_s \otimes \varphi_s = \sum_{s=0}^{r} \sigma_s \psi_s \otimes \varphi_s \tag{24}$$

where $\sigma_0 = 1$, $\psi_0 = 1\!\!1_I$, and $\varphi_0 = 1\!\!1_J$. If $r = min(I-1, J-1)$ the approximation is exact *i.e.* $\widehat{P}$ is P. From equation (24) and $\sum_{ij} p_{ij} = 1$ (which implies the solution $s = 0$) it is possible to perform the PCA of the triple:

$$(D_I^{-1} P D_J^{-1}, \quad D_I, \quad D_J) \text{ or in tensor form } ((D_I^{-1} \otimes D_J^{-1})P, \quad D_I, \quad D_J). \tag{25}$$

This last equation generalised for $k > 2$ enables to look at lack of marginal independence through associated solutions of the first Principal Tensor ([10]).

8.2 FCA-3modes and FCA-$k$modes

As for PTA-$k$modes one will present only the case $k = 3$, the framework for $k > 3$ being the same. With similar notations for a three-way table $I \times J \times K$, one performs the PTA-3modes of the quadruple:

$$((D_I^{-1} \otimes D_J^{-1} \otimes D_K^{-1})P, \quad D_I, \quad D_J, \quad D_K) \tag{26}$$

This has similar properties as for FCA-2modes moreover if one notes:

$$\Pi_{ijk} = \Pi_{.jk} + \Pi_{i.k} + \Pi_{ij.} + \Delta_{ijk}$$

for

$$(\frac{p_{ijk} - p_{i..}p_{.j.}p_{..k}}{p_{i..}p_{.j.}p_{..k}}) = (\frac{p_{.jk} - p_{.j.}p_{..k}}{p_{.j.}p_{..k}}) + (\frac{p_{i.k} - p_{i..}p_{..k}}{p_{i..}p_{..k}}) + (\frac{p_{ij.} - p_{i..}p_{.j.}}{p_{i..}p_{.j.}}) + (\frac{p_{ijk} - \delta_{ijk}}{p_{i..}p_{.j.}p_{..k}}),$$

where $\delta_{ijk} = p_{ij.}p_{..k} + p_{i.k}p_{.j.} + p_{.jk}p_{i..} - 2p_{i..}p_{.j.}p_{..k}$, one has the following property:

$$\|\Pi_{ijk}\|^2 = \|\Pi_{.jk}\|^2 + \|\Pi_{i.k}\|^2 + \|\Pi_{ij.}\|^2 + \|\Delta_{ijk}\|^2, \tag{27}$$

where $\|\|$ is the norm on the tensor space, *i.e.* using the metric $D_I \otimes D_J \otimes D_K$. This result dating from Lancaster(1951, 1980) was reported recently in [1] where a particular generalisation of correspondence analysis based on [9]'s book was derived. Equation (27) means that deviation from three-way independence can be orthogonally decomposed into deviations from independence for the two-way margins of the three-way table, and a three-way interaction term. Each two-way margins deviation from independence is reminiscent of (simple) correspondence analysis. To be convinced of this point just rewrite equation (27) as below wherein terms as in equation (22) can be identified:

$$\begin{aligned} \frac{\chi^2}{N} &= \sum_{ijk} p_{i..}p_{.j.}p_{..k}(\frac{p_{ijk} - p_{i..}p_{.j.}p_{..k}}{p_{i..}p_{.j.}p_{..k}})^2 \\ &= \sum_{jk} p_{.j.}p_{..k}(\frac{p_{.jk} - p_{.j.}p_{..k}}{p_{.j.}p_{..k}})^2 + \sum_{ik} p_{i..}p_{..k}(\frac{p_{i.k} - p_{i..}p_{..k}}{p_{i..}p_{..k}})^2 + \sum_{ij} p_{i..}p_{.j.}(\frac{p_{ij.} - p_{i..}p_{.j.}}{p_{i..}p_{.j.}})^2 \\ &+ \sum_{ijk} p_{i..}p_{.j.}p_{..k}(\frac{p_{ijk} - \delta_{ijk}}{p_{i..}p_{.j.}p_{..k}})^2. \end{aligned} \tag{28}$$

When performing the PTA-3modes (26) one retrieves simply and naturally these lack of marginal independence. The inertia or sum of squares is :

$$\sum_{ijk} p_{i..}p_{.j.}p_{..k}(\frac{p_{ijk}}{p_{i..}p_{.j.}p_{..k}})^2 = \sum_{s=0}^{r}\sigma_s = 1 + \sum_{s=1}^{r}\sigma_s = 1 + \frac{\chi^2}{N};$$

the first ($s = 0$) principal tensor being $\mathbb{1}_I \otimes \mathbb{1}_J \otimes \mathbb{1}_K$ with $\sigma_0 = 1$, its associated principal tensors relate to two-way margins decompositions, *i.e.* each term of the second row of equation (28). One can write a reconstruction formula similar to expressions (23) or (24):

$$\begin{aligned}\widehat{P} &= (D_I \otimes D_J \otimes D_K)(1\mathbb{1}_I \otimes \mathbb{1}_J \otimes \mathbb{1}_K + \sum_{s=1}^{r}\sigma_s\psi_s \otimes \varphi_s \otimes \phi_s) \qquad (29)\\ &= (P_{I..} \otimes P_{.J.} \otimes P_{..K}) + (D_I \otimes D_J \otimes D_K)(\sum_{s=1}^{r}\sigma_s\psi_s \otimes \varphi_s \otimes \phi_s)\end{aligned}$$

and also achieve the full decomposition (or reconstruction). Though no explicit expression of the maximal rank r can be calculated beforehand and is a subject of research in multiway analysis.

8.3 FCA-$k$modes for pharmaco-EEG

In order to be able to perform the analysis on the data *versus* placebo and *versus* baseline, *versus* is now the ratio instead of the difference. For every cell, $(n_{ijk} - 1)$ can be interpreted as the increase (positive or negative) from baseline and placebo. The total increase was $(4938.5 - 28 \times 24 \times 7) = 234.5$. Results of FCA-3modes of the dose means data, completing those similar seen before e.g. figure fig.7, are given below.
From listing table 4 it is possible to summarise the χ^2 decomposition as in the table 5. First of all the solution corresponding to independence explains 95.574% of the variability, these can be related to marginal effects, *i.e.* multiplicative effect of the marginals. Of the 27% of lack of independence attributable to three-way interaction, 17.5% were concentrated on vs222 and its associated solutions, 4% to vs333 and associated solutions, 3% to vs444 and associated solutions, the remaining 2.5% being spread further. Notice the singular values 1 related to complete independence (first Principal Tensor in formula (30) and the repetition associated to two-way marginal decomposition. The reconstruction formula (30) can be written exhaustively:

$$\begin{aligned}\widehat{P} &= (D_I \otimes D_J \otimes D_K)(1\mathbb{1}_I \otimes \mathbb{1}_J \otimes \mathbb{1}_K + \sum_{s=1}^{r}\sigma_s\psi_s \otimes \varphi_s \otimes \phi_s) \qquad (30)\\ &= (D_I \otimes D_J \otimes D_K)(1\mathbb{1}_I \otimes \mathbb{1}_J \otimes \mathbb{1}_K \qquad (31)\\ &+ \sum_{s_I=1}^{r_I}\sigma_{s_I}\mathbb{1}_I \otimes \varphi_s^I \otimes \phi_s^I + \sum_{s_J=1}^{r_J}\sigma_{s_J}\psi_{s_J}^J \otimes \mathbb{1}_J \otimes \phi_{s_J}^J + \sum_{s_K=1}^{r_K}\sigma_{s_K}\psi_{s_K}^K \otimes \varphi_{s_K}^K \otimes \mathbb{1}_K \qquad (32)\\ &+ \sum_{s'=1}^{r'}\sigma_{s'}\psi_{s'} \otimes \varphi_{s'} \otimes \phi_{s'}) \qquad (33)\end{aligned}$$

The χ^2 distribution to test independence is difficult to apply here as n_{ijk} represents the "change" from baseline and placebo in a ratio form. The expected frequencies of *changes* given the contingency table is $n(n_{ijk} - 1)$ ($n = 12$ subjects) with possible negative values. The percent of χ^2 helps to assess the importance of the effect, as well as comparing the relative impact, this last is better seen with the χ^2/df. The *dose* $\times$ *band* interaction seems to be very dominant.

The marginal solution (independence) must not be discarded from reports of analysis as usually done in Correspondence Analysis, because our interest here is also on approximation and description of the effects. From formula (26) using only the marginal solution (s = 0) one can approximate the increase for a particular cell by the product of the average marginal increases:

$$\begin{aligned}\hat{\hat{n}}_{ijk} &= p_{i..}p_{.j.}p_{..k}N \qquad (34)\\ &= p_{i..}p_{.j.}p_{..k}N^3/N^2\\ &\approx p_{i..}p_{.j.}p_{..k}N^3/(IJK)^2 \qquad (35)\\ &= (\frac{p_{i..}N}{JK})(\frac{p_{.j.}N}{IK})(\frac{p_{..k}N}{IJ}) \qquad (36)\end{aligned}$$

Table 4: FCA-3modes on means by dose band electrode and time of absolute energies for verum versus placebo versus baseline, subject scaled data.

```
++++++++++++++++++++++++++++++++++++++++++++++++++
 FCA-3modes PTA-3modes   dim x (24) dim y (28) dim z (7)
  data              doses_time x electrodes x  bands
++++++++++++++++++++++++++++++++++++++++++++++++++
 PDY2833  day 1 vs bl vs plb   absolute energy
 means of subjects on the subject scaled data
-----------------------------------
     Decomposition after Prin.Tens 222
        explained  99.577455% (FCA 90.453056 %)
     ---------------------------------
VALUES                 PCGLO   PCLOC  PFCA
vs111            1    95.574 %   .      .        vs222    0.0506223    00.245     .    05.534
 Xvs11           1       .      99.51    .        Xvs11    0.0506223      .      82.30    .
 Xvs11   0.0470513    00.212   00.22  04.781     Xvs11    0.0153816    00.023   07.60  00.511
 Xvs22   0.0403039    00.155   00.16  03.508     Xvs22    0.0122856    00.014   04.85  00.326
 Xvs33   0.0218824    00.046   00.05  01.034     Xvs33     0.010821    00.011   03.76  00.253
 Xvs44   0.0173196    00.029   00.03  00.648     Xvs44    0.0059887    00.003   01.15  00.077
 Xvs55   0.0127685    00.016   00.02  00.352     Xvs55    0.0032374    00.001   00.34  00.023
 Xvs66   0.0109674    00.011   00.01  00.260     Xvs66    6.191E-19    00.000   00.00  00.000
 Yvs11           1       .      97.83    .        Yvs11    0.0506223      .      68.24    .
 Yvs11   0.1025283    01.005   01.03  22.700     Yvs11    0.0246204    00.058   16.14  01.309
 Yvs22   0.0746615    00.533   00.55  12.037     Yvs22     0.016941    00.027   07.64  00.620
 Yvs33   0.0568457    00.309   00.32  06.978     Yvs33     0.012839    00.016   04.39  00.356
 Yvs44    0.044454    00.189   00.19  04.267     Yvs44     0.009938    00.009   02.63  00.213
 Yvs55   0.0227088    00.049   00.05  01.114     Yvs55    0.0059792    00.003   00.95  00.077
 Yvs66   0.0196142    00.037   00.04  00.831     Yvs66    3.222E-18    00.000   00.00  00.000
 Zvs11           1       .      99.32    .        Zvs11    0.0506223      .      41.06    .
 Zvs11   0.0448294    00.192   00.20  04.340     Zvs11    0.0337258    00.109   18.22  02.456
 Zvs22   0.0421095    00.169   00.18  03.829     Zvs22     0.032129    00.099   16.54  02.229
 Zvs33   0.0323766    00.100   00.10  02.264     Zvs33     0.022515    00.048   08.12  01.095
 Zvs44   0.0216684    00.045   00.05  01.014     Zvs44     0.015803    00.024   04.00  00.539
 Zvs55   0.0211943    00.043   00.04  00.970     Zvs55    0.0143946    00.020   03.32  00.447
...                                              ....
                                                  -------------------------------------
```

Table 5: Decomposition of the lack of complete independence (see text).

source	χ^2	df	χ^2/df	% of χ^2
2-way interactions				
lead × *band*	24.2	162	0.149	10.5
dosetime × *band*	109.6	138	1.156	48
dosetime × *lead*	33.6	621	0.054	14.5
total 2-way	167.4	921	0.181	73
3-way interaction	61.4	2805	0.021	27
total	228.8	3726	0.061	100

where the approximation (35) will be better if $IJK \approx N$, so if globally only few changes occurred (making the marginal approximation (34) better as well. Each ratio in the formula has the form of a standardised marginal change ratio O/E where the expected value is related to the hypothesis of no change. The *average increase*$(O/E-1)$ are given on figure fig.9 giving a relative strength of influence on the increase. For example an approximate amount of increase using formula (36) for 90mg at 3h(40%) on δ(18%) at lead F7 (21%): $1.40 \times 1.18 \times 1.21 = 1.99$, so an increase of 99%. This is an estimation based firstly on the *model of independence* and secondly using *average margin increase*. Under independence only one would estimate (formula 34) 81%, and the observed value (without modelling) is 74%.

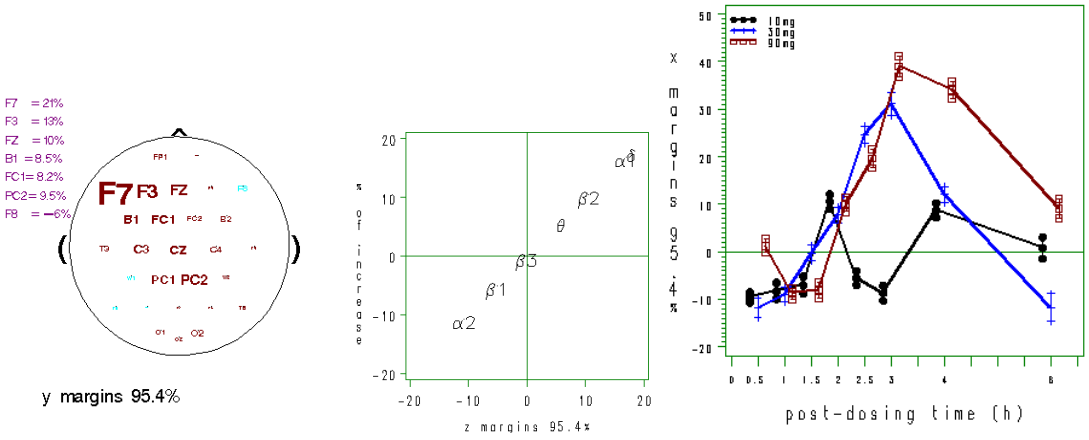


Figure 9: Average percentage of increase deducted from the Margins defining complete independence in FCA-3modes means(data subject scaled) for all bands (absolute energy) for verum *versus* placebo *versus* second baseline.

Notice the consistency between *versus* being either the difference or the ratio, at least for the dose time profiles (PTA-3modes on means (fig.5) and FCA-3modes on means (fig.9)). On figure fig.10 are displayed Principal Tensors relating the most of the lack of independence. The first four are related to deviation from *two-way margins independence i.e.* respectively *lead* $\times$ *band* interaction, *dosetime* $\times$ *band* interactions (two chosen), and *dosetime* $\times$ *lead* interaction. The last principal tensor relates to *three-way interaction*. Notice in deviations from *two-way margins independence* the component on the third is always 1 everywhere as it is there to build the two-way margins. For *dosetime* $\times$ *band* interaction, the first Principal Tensor (22%) relates to an opposition between δ and α_1 band waves associated to a time decrease for all doses (less important for 90mg), making δ increased (*versus* placebo) at the beginning of the experiment and progressively reversing this effect to finish to a decreased δ, and the reverse for α_1. The second Principal Tensor shows a different profile for 10mg comparatively mainly to 90mg towards fast and slow waves opposition. This "behaviour" can also be seen in the three-way interaction, this time opposing back and fronto-temporal activity.

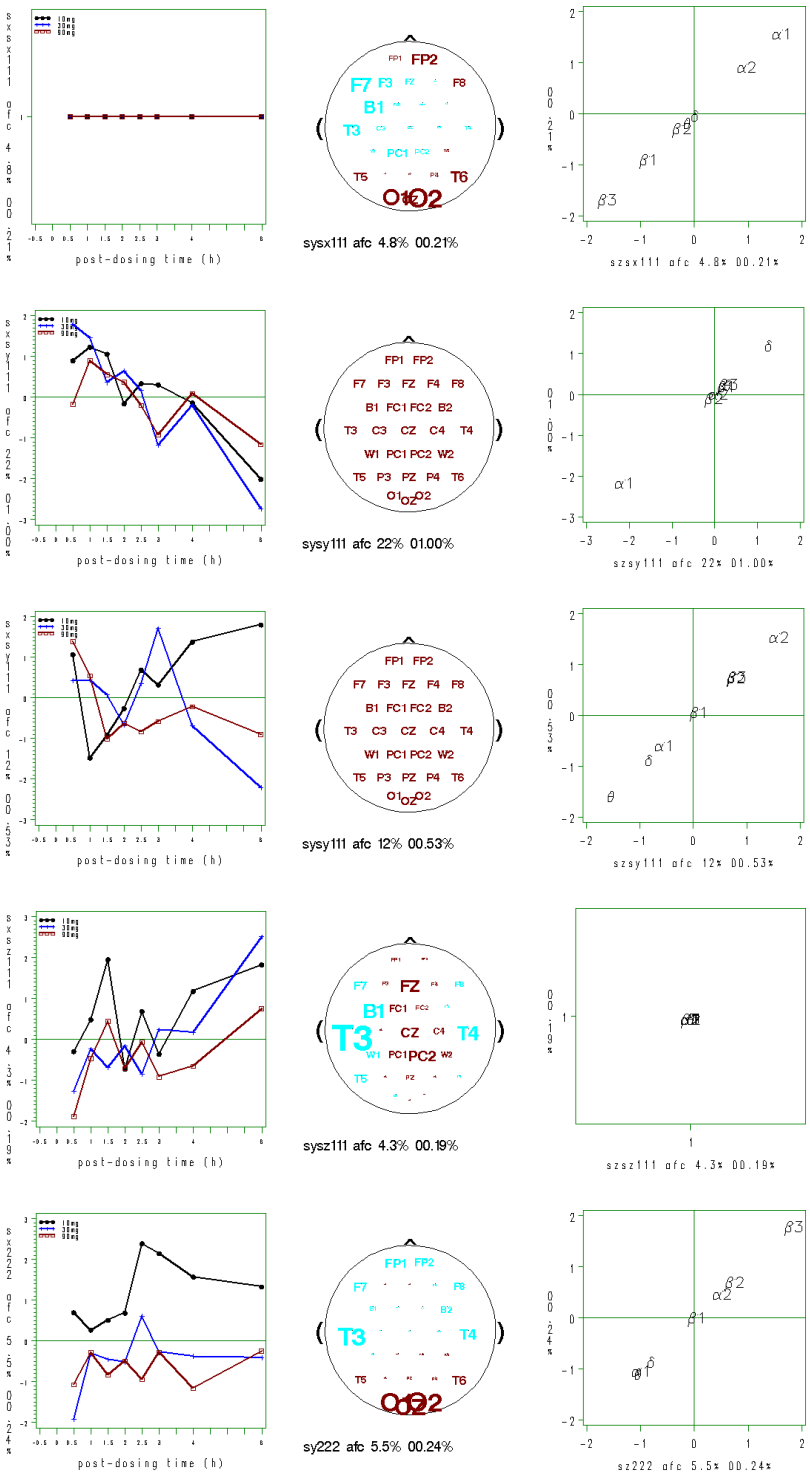


Figure 10: Some other Principal Tensors (PT) of the FCA-3modes on means(*data subject scaled*) for all bands (absolute energy) for verum *versus* placebo *versus* second baseline (*versus* is the ratio): marginal time two-way dependence (one PT), *idem* for marginal space (two PTs), *idem* for marginal bands (1 PT), and second Principal Tensor (second k-modes solution).

9 Conclusion

The aim of this paper was to describe a general framework of multiway multidimensional analysis as a tool to analyse pharmaco-EEG data. The multiway method chosen enables data reduction of the complex structure of the data, providing a quantified hierarchy of effects (sets of linear components on each mode). The *subject* $*$ *dose* component can then be used for testing the drug effect with any cross-over analysis. Even with simple preprocessing the method performed well in extracting main variability features of the data but also revealed outliers.
The existence of outliers seems to be an important aspect of this kind of data because of a strong subject variability of EEG recordings associated associated with drug intake. Without discussing sample size required for this type of analysis, (which would need to be fully addressed) this puts into question the use of subjects as a mode (or combined with doses) in the analysis and enforce to use summaries across subjects. Centring , scaling, and interactions removal were used to try to avoid outliers. Another approach was to represent doses by a summary measure across subjects (mean, median, trimean) and do the analysis with this mode. A larger sample would reduce outlier problem but would also improve estimation of the summaries. Analysing robust summaries may be interesting when comparing or classifying different drugs if the same design was used but not necessarily on the same subjects. Supplementary points technique can confirm and add more information on dispersion of the evolution of the dose-time profiles.
The method described here can involve the use of metrics as in generalised multidimensional analysis, *e.g.* discriminant analysis. Note that the problem in pharmaco-EEG analysis is not necessarily a discriminant one as in the first place the neuro-pharmacist wants to identify effects of the drug and only secondly dose pattern effects. Correspondence Analysis on k modes was introduced as a particular PTA-$k$modes of a $(k+1)$ uple. This method seem very well suited to pharmaco-EEG data as conditional independence can be analysed and quantified. Complete independence, two way interactions, three way interactions take part of the same analysis, and are in turn also decomposed as sum of Principal Tensors. Analysing the links between pharmaco-dynamic variables and pharmaco-kinetic variables has not been investigated in this paper, but implementing for example inter-battery analysis ideas ([21]) with PTA-$k$modes seems straightforward. An expending literature about multiway PLS (Partial Least Squares) is worth reading for this purpose.

Acknowledgments

This work was done while the author was located at SANOFI-Recherche Biostatistics Department in Montpellier, with a grant from Elf-Aquitaine, thanks to Dr.G.Derzko who also initiated the research and welcomed me in his team.

References

[1] A Carlier and P.M Kroonenberg. Decompositions and Biplots in Three-way Correspondence Analysis. *Psychometrika*, 61:355–373, 1996.

[2] H Caussinus and A Ruiz. Interesting Projections of Multidimensional Data by Means of Generalized Principal Component Analysis. In COMPSTAT 90, pages 121–126, Heidelberg, 1990. Physica-Verlag.

[3] K.T Dago, R Luthringer, R Lengelle, G Rinaudo, and J-P Macher. Statistical Decision Tree : A tool for Studying Pharmaco-EEG effects of CNS-active Drugs. *Neuropsychobiology*, 29:91–96, 1994.

[4] J Dauxois, Y Romain, and S Viguier-Pla. Tensor Products and Statistics. *Linear Algebra and its Applications*, 210:59–88, 1994.

[5] Y Escoufier. L'Analyse des Correspondances : ses proprits et ses extensions. In Proceeding of the 45^th session of the ISI, pages 28.2.1–28.2.16, 1985.

[6] A.S Field and D Graupe. Topographic Component (Parallel Factor) Analysis of Multichannel Evoked Potentials: Practical Issues in Trilinear Spatiotemporal Decomposition. *Brain Topography*, 3(4):407–423, 1991.

[7] A.P Holmes, R.C Blair, J.D.G Watson, and I Ford. Nonparametric Analysis of Statistic Images from Functional Mapping Experiments. *Journal of Cerebral Blood Flow and Metabolism*, 16:7–22, 1996.

[8] P Jhnig and M Jobert. Methodological Considerations for the Evaluation of EEG Mapping Data : A Practical Example Based on a Placebo/Diazepam Crossover Trial. *Neuropsychobiology*, 31:31–46, 1995.

[9] P.M Kroonenberg. Three-mode Principal Component Analysis: Theory and Applications. DSWO Press, Leiden, 1983.

[10] D Leibovici. Facteurs Mesures Rptes et Analyses Factorielles : applications un suivi pidmiologique. PhD Thesis in Mathmatiques et Applications, scholarship : Biostatistiques, Universit de Montpellier II, 1993.

[11] D Leibovici. Multiway Multidimensional Analysis for Pharmaco-EEG Studies : part I. Biostatistics Montpellier, SANOFI-RECHERCHE, 1998.

[12] D Leibovici. Multiway Multidimensional Analysis for Pharmaco-EEG Studies : part II. Biostatistics Montpellier, SANOFI-RECHERCHE, 1998.

[13] D Leibovici. SVD-*k*modes Program's Manual. *Macros SAS/IML* , SANOFI-RECHERCHE Biostatistics Montpellier, 1998.

[14] D Leibovici. An SVD-kmodes R-code, the package PTAk version 1.1-4. . *http://www.fmrib.ox.ac.uk/~didier/work/tensor.html*, CRAN, 2001.

[15] D Leibovici. Metric choices for fMRI Multiway Data Analysis. *In Preparation*, 2001.

[16] D Leibovici and R Sabatier. A Singular Value Decomposition of k-Way Array for a Principal Component Analysis of Multiway Data, PTA-k . *Linear Algebra and Its Applications*, 269:307–329, 1998.

[17] J-P Macher, R Luthringer, M Toussaint, and R Minot. Brain Mapping and Psychoactive Drugs. In Hindermarch I and Stonier P.D, editors, *Human Psychopharmacology : Methods and Measures*, volume Vol 3. Wiley, 1990.

[18] C.R Rao. The use and interpretation of Principal Component Analysis in applied research. *Sankhya*, A26:329–359, 1964.

[19] R Sabatier, J-D Lebreton, and D Chessel. Principal Component Analysis with Instrumental Variables as a tool for modelling composition data. In Coppi R Bolasco S, editor, *Multiway Data Analyses*, pages 341–352. Elsevier Science Pub., North-Holland, 1989.

[20] J.M.F Ten Berge. Convergence of PARAFAC preprocessing procedures and the Deming-Stephan method of iterative proportional fitting. In Coppi R Bolasco S, editor, *Multiway Data Analyses*, pages 53–63. Elsevier Science Pub., North-Holland, 1989.

[21] L.R Tucker. An Inter-battery method of Factor Analysis. *psychometrika*, 23:111–136, 1958.

[22] K.J Worsley, S Marrett, P Neelin, A.C Vandal, K.J Friston, and A.C. Evans. A unified statistical approach for dtermining significant signals in images of cerebral activation. *Human Brain Mapping*, 4:58–73, 1995.