UNIVERSITY OF PÉCS
Doctoral School of Chemistry

THE APPLICATION OF TWO-DIMENSIONAL CORRELATION IN CHROMATOGRAPHY AND THE INTRODUCTION TO ALTERATION ANALYSIS

PhD thesis

József Simon
Department of Analytical and
Environmental Chemistry

Supervisor:
Dr. Attila Felinger
professor of chemistry

PÉCS, 2018

Contents

1

Motivation and Problem Statement

> *Science, my lad, is made up of mistakes, but they are mistakes which it is useful to make, because they lead little by little to the truth.*
>
> **— Jules Verne**

In the last few decades chemometrics had a major role in the progress of analytical chemistry. Besides the advancement of lab equipment and instruments, there is a huge demand for constantly creating more sophisticated methods to analyse the ever increasing amount of data. The evolution of computer science allowed the data, generated in the labs, to rapidly evolve from simple graphs to complex, multi-dimensional sets, which rendered the day-to-day use of evaluating software not just recommended but mandatory.

The expectations for chemometric methods have two – in some way contradictory – cornerstones: be able to analyse the more and more advanced data sets; and create simple, easily understandable results. There is, however, a field in chemometrics which seems to incorporate these goals effortlessly, called two-dimensional correlation analysis (2DCOR). This method was built to extend the possibilities of infrared spectroscopy (IR) with some of the two-dimensional concept used in nuclear magnetic resonance spectroscopy (NMR). It was a great success and later proved to be useful with many probes other than IR. Despite the rapid growth of the field, 2DCOR operated mostly in spectroscopy, only a few attempts have been made outside that, however we strongly believe that its basic concept can be advantageous in chromatography as well.

The main aims of this work were:

- exploration of the possibilities, attributes and details of 2DCOR in the field of chromatography;
- extension of 2DCOR to three-dimensional datasets where a series of two-dimensional measurements - instead of the original one-dimensional - are evaluated;
- building a new method which uses the properties of 2DCOR, but is able to work with three-dimensional datasets and is fine-tuned to use in chromatography;
- comparison of 2DCOR and our new method on two-dimensional datasets where a series of one-dimensional measurements are used;
- proving the capability of 2DCOR in chromatography by applying it to measured chromatograms and comparing it to other chemometric methods;
- demonstrate the practical use of our new method on measured chromatograms and highlight its advantages in the evaluation of complex chromatographic problems.

In the next sections we will describe 2DCOR's origins, how it became a versatile, widely used chemometric method and its way from IR through other probes to chromatography. We will give details of the mathematical background for both 2DCOR and our new method. Furthermore our studies will be presented, where 2DCOR's abilities are being tested on computer generated series of chromatograms; our new method is introduced, in silico experiments will detail its properties with 3D datasets and it will be compared to 2DCOR in 2D. In the last chapter two practical examples will show the advantages of methods used in chromatography. The first one will present 2DCOR's performance compared to principal component analysis (PCA) in a reproducibility study of chromatographic columns, whereas the other focuses on concentration changes in supercritical fluid chromatography (SFC) measurements.

2 Literature review

> *You cannot teach a man anything; you can only help him discover it in himself.*
>
> **— Galileo Galilei**

2.1 2DCOR

The concept of 2D spectroscopy was first introduced in the field of NMR in 1976 [1]. It was a breakthrough in analytical chemistry, and opened a number of ways for promising possibilities. Ten years later, 2D spectroscopy outgrew the boundaries of resonance spectroscopy, when Noda introduced the perturbation based two-dimensional correlation spectroscopy (2DCOS) [2]. With that method a number of chemical phenomena became available for analysis and interpretation with vibrational spectroscopy. 2D correlation spectroscopy gained large interest and became very popular in IR, ultraviolet spectroscopy (UV) and visible spectroscopy (VIS). It only took a few years to realize that 2D spectroscopy was clearly a volatile technique that could be applied in a wider perspective.

The construction of the generalized method was the next step in the evolution of 2DCOS [3]. The fundamental step of the calculation was based on Fourier transform, i.e. the intensity variations had to be built by sinusoidal components.

In 2000 Noda published a study [4], which introduced an entirely new calculation method for 2D correlation, the discrete Hilbert transform. This procedure replaced the Fourier transform with matrix operations, lifted all the boundaries of the previous technique, but provided the same results (see Fig. 2.1). With that, an ultimate tool was made available, which could be applied to any kind of measured data [5]. By 2DCOR we refer to this general calculation method regardless what kind of data is used. Not surprisingly, 2D correlation has spread

rapidly from vibrational spectroscopy to new analytical methods, such as mass spectrometry [6–8], X-ray [9–12] or chromatography (detailed in Section 2.2).

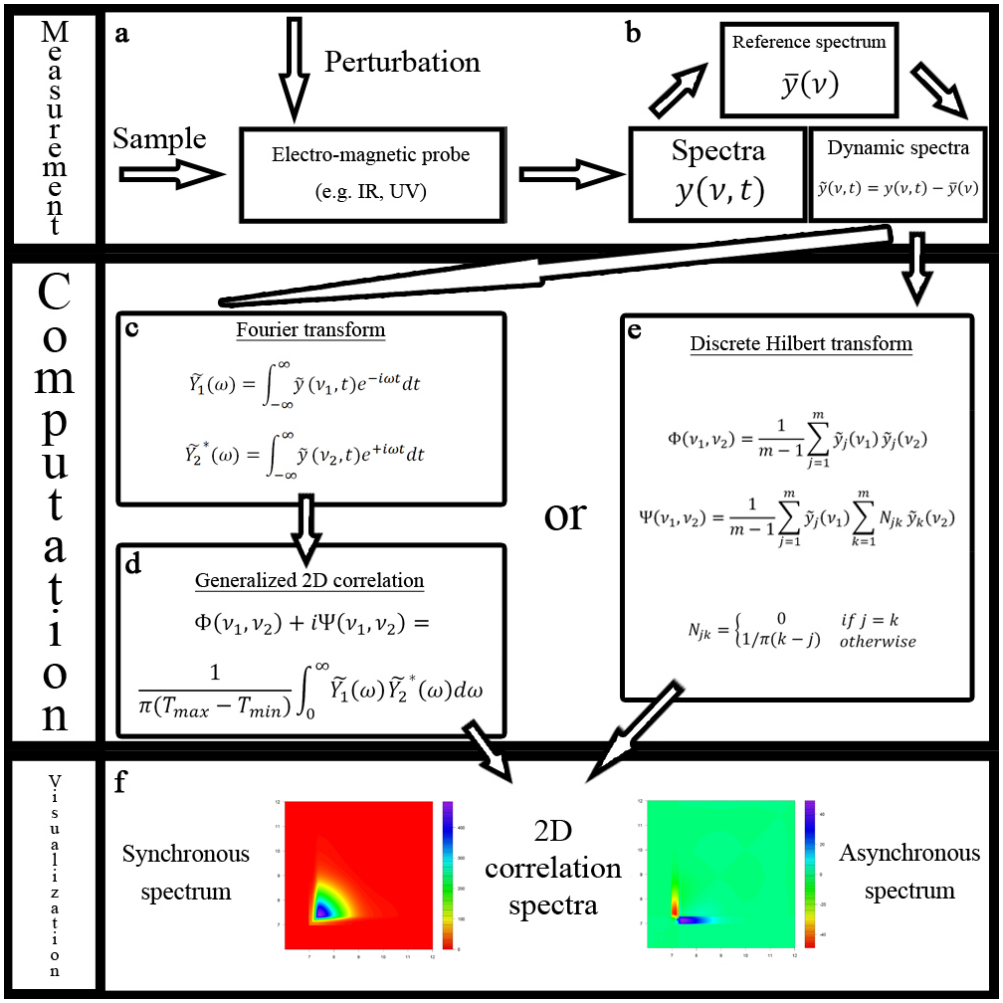

Fig. 2.1.: The process of 2D correlation: (a) measurement system, (b) dynamic spectra, (c) Fourier transform, (d) the generalized correlation function, (e) discrete Hilbert transform, (f) synchronous and asynchronous spectra [5]

The variety of data used by 2DCOR shows its earned place in chemometric methods, however, its growth is far more adequate than just leaping to other analytical probes [13–19].

2.2 2DCOR in chromatography

Despite its wide range of usefulness, the application of 2D correlation calculation in chromatography is still very rare. The first breakthrough was the work of Izawa et al. [20]. In their study, the foundations of the two-dimensional correlation gel permeation chromatography (2D-GPC) were established. Another

milestone was the application in gas chromatography (GC) published by Hyde et al. [21]. They used hetero-correlation analysis with GC-IR spectroscopy. There are several publications in this field, but 2D correlation has not really gained popularity in chromatography, and there may be much more potential in it.

One of the greatest benefits of 2DCOR is that it is able to handle data from any kind of measurements. The main interest is still in spectroscopy, but it slowly starts to gain popularity in chromatography as well [20–33].

2.3 Multi-way data analysis

Using three-way data is not a common procedure in analytical practice, despite the fact that its methods have a long history. The mathematics of multi-way analysis was first applied in psychometrics in 1970. The same method was developed independently as PARAFAC by Harshman [34] and as CANDECOMP by Carrol and Chang [35]. There are a few decomposition methods for multi-way data, but the main three are the PARAFAC, the Tucker3 [36] and performing a two-way PCA by unfolding the multi-way array. As always in chemometrics, there is no universal law to pick the best method; it highly depends on the actual task. PCA has the most degrees of freedom, so it is the most complex and flexible, and PARAFAC with the least degrees of freedom is the simplest and most restricted [37]. In the dissertation we do not go into more details regarding those methods, but excellent summaries can be found in the literature [38, 39], or a tutorial for PARAFAC [40] can give a good outlook on the field. All the mentioned studies are from the 20th century, but the methods are still being used and further developed [41, 42].

These decomposition methods can be used in image analysis too [43]. Such data are for example satellite-, spectroscopic- (fluorescence, IR, etc.) and microscopic images. As expected it is a very complex field, because of the numerous possibilities. One of them is the diversity of the applicable data, how they are interpreted. A multi-way array can be sorted by the properties of its ways (modes, dimensions). A way can be seen as an object (O) or a variable (V). In a two-way data it is rather simple, but in a three-way (multi-way) array it is not so straightforward. ALA on 2D chromatographic measurements can be considered an image analysis tool as a case of OOV (object-object-

variable), because the dimensions of the measurements - the points of the chromatographic plane - are the objects, and the perturbation is the variable.

The examples discussed so far for multi-way methods are from various fields of science, but chromatography and 2DCOR are no exceptions. There are plenty of examples in chromatography [44–47], however it is still rare in 2DCOR [48].

3 Mathematical background

> *Life is and will ever remain an equation incapable of solution, but it contains certain known factors.*
>
> **— Nikola Tesla**

3.1 2DCOR, original concept

First we want to show the original concept for 2DCOR. Its calculation is a grueling mathematic procedure with Fourier-transform [3] which was later replaced by a simpler approach with matrix algebra which we also use and it is presented in the next section. However, it is appropriate to present this concept as well, because it can lead to a better understanding of the roots of 2DCOR.

In this section we stuck to spectroscopic nomenclature and later switched to chromatographic interpretation, because we wanted to distance it from the other formulas, which were actually used in practice, and emphasize that it has only historical meaning in the boundaries of this thesis.

The basis of this formula is the so-called dynamic spectrum ($\tilde{y}(\nu, t)$) which is calculated from spectral intensities $y(\nu, t)$ at an appropriate spectral index (ν) and an external variable (t) which runs from T_{min} to T_{max}, with the help of a reference spectrum ($\bar{y}(\nu)$), that is usually the average.

$$\tilde{y}(\nu, t) = \begin{cases} y(\nu, t) - \bar{y}(\nu), & \text{if } T_{min} \leq t \leq T_{max} \\ 0, & \text{otherwise} \end{cases} \tag{3.1}$$

$$\bar{y}(\nu) = \frac{1}{T_{max} - T_{min}} \int_{T_{min}}^{T_{max}} y(\nu, t) \mathrm{d}t \tag{3.2}$$

The concept of 2DCOR is to quantitatively compare the patterns of spectral intensity variations. It is summarized in this equation:

$$\mathbf{X}(\nu_1, \nu_2) = \langle \tilde{y}(\nu_1, t) \cdot \tilde{y}(\nu_2, t') \rangle \tag{3.3}$$

The brackets ($\langle\rangle$) symbolize a cross-correlation function. To expound further we consider $\mathbf{X}(\nu_1, \nu_2)$ as a complex function:

$$\mathbf{X}(\nu_1, \nu_2) = \mathbf{\Phi}(\nu_1, \nu_2) + i\mathbf{\Psi}(\nu_1, \nu_2) \tag{3.4}$$

where i is the imaginary unit. $\mathbf{X}$ combines two orthogonal components which are the synchronous $\mathbf{\Phi}$ and asynchronous $\mathbf{\Psi}$ 2D correlation intensities respectively.

If we go one step further we can define now the previously introduced terms by the generalized two-dimensional correlation function:

$$\mathbf{\Phi}(\nu_1, \nu_2) + i\mathbf{\Psi}(\nu_1, \nu_2) = \frac{1}{\pi(T_{max} - T_{min})} \int_0^\infty \tilde{\mathbf{Y}}_1(\omega) \cdot \tilde{\mathbf{Y}}_2^*(\omega) \mathrm{d}\omega \tag{3.5}$$

where $\tilde{\mathbf{Y}}_1(\omega)$ is the Fourier transform of $\tilde{y}(\nu_1, t)$:

$$\begin{aligned} \tilde{\mathbf{Y}}_1(\omega) &= \int_{-\infty}^{\infty} \tilde{y}(\nu_1, t) \mathrm{e}^{-i\omega t} \mathrm{d}t \\ &= \tilde{\mathbf{Y}}_1^{\mathrm{Re}}(\omega) + i\tilde{\mathbf{Y}}_1^{\mathrm{Im}}(\omega) \end{aligned} \tag{3.6}$$

and $\tilde{\mathbf{Y}}_2^*(\omega)$ is the conjugate formula with $\tilde{y}(\nu_2, t)$:

$$\begin{aligned} \tilde{\mathbf{Y}}_2^*(\omega) &= \int_{-\infty}^{\infty} \tilde{y}(\nu_2, t) \mathrm{e}^{+i\omega t} \mathrm{d}t \\ &= \tilde{\mathbf{Y}}_2^{\mathrm{Re}}(\omega) - i\tilde{\mathbf{Y}}_2^{\mathrm{Im}}(\omega) \end{aligned} \tag{3.7}$$

where Re is the real and Im is the imaginary part.

3.2 Data

As stated in the previous chapter, 2DCOR can work with data from any kind of measurements. Alteration Analysis (ALA) also shares this property, so the application examples could be from any analytical field. In the present study, we focus only on chromatograms in order to keep the discussion simple, and we do not focus on comparing the methods in different fields of analytical chemistry. The goal is to present the basic concepts of 2DCOR and ALA as well as their relations.

The two methods are comparable if data from 1D measurements are used. For instance, when n chromatograms – each containing m digitalized points – are put in a matrix's rows ($\mathbf{X}$).

$$\mathbf{X} = \left(x_{i,j}\right) \qquad \begin{matrix} i = 1, 2, ..., n \\ j = 1, 2, ..., m \end{matrix} \tag{3.8}$$

To be clear in the nomenclature, in that case from the measurement's point of view, the first dimension is the time, and the second dimension is the perturbation we applied on the analytical system. From the two-way matrix's point of view, the first dimension (the rows) is the chromatograms (perturbation), and the second dimension is the time.

3.3 2DCOR

The field of 2DCOR has become diversified over the past twenty years, and the employed mathematical background behind the results is sometimes unclear. When using the term 2DCOR we refer to the discrete Hilbert transform [4] which uses simple matrix algebra. We present the basic calculations, but further details can be found in our paper [49], or there is an excellent book dedicated to the field of 2DCOR [50].

3.3.1 Pretreatment

The pretreatment of the data is also an essential part of 2DCOR. A reference signal is subtracted from the chromatograms. We use the most common method, the mean signal of the chromatograms (centering), see Eq. 3.9.

$$\begin{array}{lll} \overline{\mathbf{x}} = \left(\overline{x}_j\right) & \overline{x}_j = \dfrac{\sum_{i=1}^{n} x_{i,j}}{n} & i = 1, 2, ..., n \\ \mathbf{Y} = \left(y_{i,j}\right) & y_{i,j} = x_{i,j} - \overline{x}_j & j = 1, 2, ..., m \end{array} \tag{3.9}$$

Other choices for the reference signal can be a zero baseline – the original data are retained and used – or any given signal along the perturbation, for example the initial or the last data set in the series. There is no general rule on how to choose the reference signal, but centering is used in most cases, probably it is the most beneficial. When we use the average chromatogram ($\overline{\mathbf{x}}$) as the reference signal, each element that is present in the measurement but is not changing will be cut out. For example if there are two peaks in the chromatograms but only the first one is changing throughout the measurement, the treated data ($\mathbf{Y}$) will contain information about the first peak only, the second one will disappear because its values in every chromatogram are equal (or very close) to the average. Another major benefit is that baseline-correction is not needed. Since data pretreatment removes all the static elements from the data, so if we have the same baseline over the experiments, all data will be corrected for the baseline.

Please note that the effect of smoothing is not discussed in this thesis.

3.3.2 Synchronous 2D Correlation

The concept of 2DCOR comes from the cross-correlation function, which provides a complex value (Equation 3.4). The real ($\mathbf{\Phi}$) and the imaginary ($\mathbf{\Psi}$) parts refer to the synchronous and asynchronous components, respectively. The deduction is not very complicated, however, the matrix algebra approach is the

easiest and proven to be the most useful in practice [50]. The formula for the calculation of the synchronous correlation matrix ($\mathbf{\Phi}$) becomes quite simple:

$$\mathbf{\Phi} = \frac{1}{n-1}\mathbf{Y}^T\mathbf{Y} \tag{3.10}$$

We have to mention that in 2DCOR, the employed terms can be misguiding and this section is a great example to illustrate it. Both parts of the term "synchronous correlation" need some explanation. The term *synchronous* is stuck with the method, because in the original correlation concept time was the variable behind the perturbation, hence the term syn-chronous = the synergy of time (chrónos, *greek*) dependent variables. The method has evolved immensely since its introduction, and now any reasonable physical meaning (temperature, pressure, etc.) can be the basis of perturbation. On the other hand, the term *correlation* also originates from the concept and not from equation (Eq. 3.10) itself, because it is in fact the covariance of the retention times. Thus, a newcomer to this field is advised to study the origins and backgrounds of 2DCOR in order to fully understand the terminology.

3.3.3 Asynchronous 2D Correlation

The imaginary part ($\mathbf{\Psi}$) of the cross-correlation function (Equation 3.4) is a little more complicated to implement in matrix algebra, but with the discrete Hilbert transform it is more than possible [4]. With matrix algebra, the calculation for the asynchronous correlation matrix ($\mathbf{\Psi}$) is rather similar to its synchronous counterpart:

$$\mathbf{\Psi} = \frac{1}{n-1}\mathbf{Y}^T\mathbf{N}\mathbf{Y} \tag{3.11}$$

where $\mathbf{N}$ is the Hilbert–Noda transform matrix with the dimension of $n \times n$:

$$\mathbf{N} = \frac{1}{\pi} \begin{bmatrix} 0 & 1 & 1/2 & 1/3 & \dots \\ -1 & 0 & 1 & 1/2 & \dots \\ -1/2 & -1 & 0 & 1 & \dots \\ -1/3 & -1/2 & -1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \tag{3.12}$$

The elements of this matrix are generated as follows:

$$N_{i,j} = \begin{cases} 0, & \text{if } i = j \\ 1/\pi(j-i), & \text{otherwise} \end{cases} \tag{3.13}$$

3.4 Alteration Analysis

Originally, we developed ALA for three-way arrays, but its concepts allow that it can be used in lower dimensions as well. That is the reason why the formulas presented here are the simplifications of the ones found in our first paper in the topic [51]. For three-way arrays, ALA is advantageous, because it produces matrices instead of higher dimensional data sets, which cannot be visualized properly. This means that if we want to examine the changes in an n dimensional measurement with an additional dimension for the perturbation, the ALA maps will be n-dimensional, while 2DCOR maps will be $2n$-dimensional. In traditional one-dimensional measurements, the benefits are the same, the generated simple graphs are easier to interpret and to connect to the dimension of the measurement (wavelength, frequency, retention time, etc.)

3.4.1 Basic Altertion Map

The basic alteration map (BAM) shows the overall changes in the series of the chromatograms. It is needed because the synchronous and asynchronous alteration maps are maximum scaled, and the BAM has to show the magnitude of these changes. The basic alteration vector's (**b**) formula is very simple:

$$\mathbf{b} = \left(b_j\right) \qquad b_j = \max(x_j) - \min(x_j) \qquad j = 1, 2, ..., m \tag{3.14}$$

3.4.2 Difference Data Matrix

The first step in ALA is to calculate the differences (**D**) between every two adjacent points along the perturbation. This is the fundamental aspect. It shows the straightforward point of view of this method: it strictly focuses on the individual points of the measurement and investigates the changes at those points.

$$\mathbf{D} = \left(d_{i,j}\right) \qquad d_{i,j} = x_{i+1,j} - x_{i,j} \qquad \begin{matrix} i = 1, 2, ..., n-1 \\ j = 1, 2, ..., m \end{matrix} \tag{3.15}$$

3.4.3 Synchronous Alteration Map

At this point we can show all the changes that occur throughout the perturbation. The next two maps demonstrate how to separate those changes in some ways. The main goal of ALA is to investigate the properties of the changes of the huge data set and to illustrate them on simple visual graphs. To describe the functions beneath the changes without actual curve-fitting – which would take a lot of time and energy – some statistical parameters are emphasized, like the average ($\bar{d}_j$) and the standard deviation (σ_{d_j}) of the change.

In the synchronous alteration map, only the monotone changes will be shown, because they have a high average and a low deviation. If we multiply this fraction with the value of overall change, then we have the synchronous alteration

matrix ($\mathbf{s}'_j$). The term $+1$ has to be added, because if there were a perfect case, and the function of the change were linear with no noise, then σ_{d_j} becomes 0.

$$\mathbf{s}' = \left(s'_j\right) \qquad s'_j = \frac{b_j \bar{d}_j}{\sigma_{d_j} + 1} \qquad j = 1, 2, ..., m \tag{3.16}$$

As it can be seen below, the asynchronous alteration matrix has a different formula, but the two maps have to be comparable, that is why the values have to be scaled. We chose maximum scaling:

$$\mathbf{s} = (s_j) \qquad s_j = \frac{s'_j}{\max(|\mathbf{s}'|)} \qquad j = 1, 2, ..., m \tag{3.17}$$

3.4.4 Asynchronous Alteration Map

In the asynchronous alteration matrix ($\mathbf{a}'$) noise-like patterns and non-monotone, complex functions are dominant. We accomplished that with cutting out all the monotone functions, because functions like that have the same value for b_j and for $\left|\sum_{i=1}^{n-1} d_{i,j}\right|$, that is why in this case a'_j is going to be 0. But if a function is still present, although it belongs to the synchronous map, it is going to have a small value, because it has a low deviance in the changes (σ_{d_j}). The third part of the equation is simply present because it is often important to know the direction of some non-monotone functions, such as one single change in the series of points. This is the simplest way to implement this property to the asynchronous map.

$$\begin{gathered} \mathbf{a}' = \left(a'_j\right) \qquad j = 1, 2, ..., m \\ a'_j = \left(b_j - \left|\sum_{i=1}^{n-1} d_{i,j}\right|\right) \sigma_{d_j} \left(\max(x_j) + \min(x_j) - 2\bar{x}_j\right) \end{gathered} \tag{3.18}$$

The asynchronous matrix ($\mathbf{a}'$) is also maximum scaled.

$$\mathbf{a} = (a_j) \qquad a_j = \frac{a'_j}{\max(|\mathbf{a}'|)} \qquad j = 1, 2, ..., m \tag{3.19}$$

3.5 Extension of Alteration Analysis to 3D datasets

The formulas in this section are for the extensions of previous ones to 3D datasets. ALA was originally built for this purpose, so the development was the other way around. It is mathematically possible to do the same on 2DCOR, however it has little practical significance. In 2D data environments alteration formulas provide vectors which can be plotted in 2D graphs, in 3D data they become matrices which need 3D to plot. 2DCOR has matrices already in 2D data, so the upper dimensional maps are four-way arrays, not just three-way because we have 2D chromatograms to correlate so it is 2x2 dimensions. They can be plotted in 5D space, but the foundation of 2DCOR is to create relatively simple visual representation of the examined chemical system, that is why this approach is futile.

3.5.1 Three-way Data Array

The data frame ($\underline{\mathbf{X}}$) is a 3D array, and it contains n number of 2D chromatograms, each having $m \times p$ points as their dimensions.

$$\underline{\mathbf{X}} = (x_{i,j,k}) \qquad \begin{matrix} i = 1, 2, ..., n \\ j = 1, 2, ..., m \\ k = 1, 2, ..., p \end{matrix} \tag{3.20}$$

3.5.2 Basic Alteration Map

The basic alteration matrix (**B**):

$$\mathbf{B} = \left(b_{j,k}\right) \qquad b_{j,k} = \max(x_{i,j,k}) - \min(x_{i,j,k}) \qquad \begin{matrix} i = 1,2,...,n \\ j = 1,2,...,m \\ k = 1,2,...,p \end{matrix} \tag{3.21}$$

3.5.3 Difference Data Array

The difference data array ($\underline{\mathbf{D}}$), where the values are the differences between every next point in the third dimension of the data frame.

$$\underline{\mathbf{D}} = \left(d_{i,j,k}\right) \qquad d_{i,j,k} = x_{i+1,j,k} - x_{i,j,k} \qquad \begin{matrix} i = 1,2,...,n-1 \\ j = 1,2,...,m \\ k = 1,2,...,p \end{matrix} \tag{3.22}$$

3.5.4 Synchronous Alteration Map

The synchronous alteration matrix (**S**′):

$$\mathbf{S}' = \left(s'_{j,k}\right) \qquad s'_{j,k} = \frac{b_{j,k}\bar{d}_{j,k}}{\sigma_{d_{j,k}} + 1} \qquad \begin{matrix} j = 1,2,...,m \\ k = 1,2,...,p \end{matrix} \tag{3.23}$$

The matrix still has to be normalized, just like in 2D:

$$\mathbf{S} = \left(s_{j,k}\right) \qquad s_{j,k} = \frac{s'_{j,k}}{\max(|\mathbf{S}'|)} \qquad \begin{matrix} j = 1,2,...,m \\ k = 1,2,...,p \end{matrix} \tag{3.24}$$

3.5.5 Asynchronous Alteration Map

The asynchronous alteration matrix ($\mathbf{A}'$):

$$\begin{gathered}\mathbf{A}' = \left(a'_{j,k}\right) \qquad \begin{matrix} j = 1,2,...,m \\ k = 1,2,...,p \end{matrix} \\ a'_{j,k} = \left(b_{j,k} - \left|\sum_{i=1}^{n-1} d_{i,j,k}\right|\right) \sigma_{d_{j,k}} \left(\max(x_{j,k}) + \min(x_{j,k}) - 2\bar{x}_{j,k}\right)\end{gathered} \tag{3.25}$$

The asynchronous matrix is also normalized, the same way as its synchronous counterpart:

$$\mathbf{A} = \left(a_{j,k}\right) \qquad a_{j,k} = \frac{a'_{j,k}}{\max(|\mathbf{A}'|)} \qquad \begin{matrix} j = 1,2,...,m \\ k = 1,2,...,p \end{matrix} \tag{3.26}$$

3.5.6 Correlation Coefficient

In 2DCOR – besides the synchronous and asynchronous correlation – the so-called sample-sample correlation is often used. In this technique, we calculate the covariances not for the retention times (second dimension of the data matrix) but for the samples (first dimension). Combining that with the other correlation maps, we have the correlation information for both the retention times and the samples through the perturbation.

Although we use correlation coefficient maps, which are sometimes also used in 2DCOR, because despite the similarity with the covariance map, in most cases it is more informative.

For a matrix (**M**), the calculation of the correlation coefficient matrix (**P**) is very simple:

$$\begin{aligned}
&\mathbf{M} = \left(m_{i,j}\right) && \begin{matrix} i = k = l = 1, 2, ..., n \\ j = 1, 2, ..., m \end{matrix} \\
&\widetilde{\mathbf{M}} = \left(\widetilde{m}_{i,j}\right) && \widetilde{m}_{i,j} = m_{i,j} - \bar{m}^{j} \\
&\mathbf{P} = \left(\rho_{k,l}\right) && \rho_{k,l} = \frac{1}{m-1} \frac{\sum_{j=1}^{m} \widetilde{m}_{k,j} \widetilde{m}_{l,j}}{\sigma_{\widetilde{m}_k} \sigma_{\widetilde{m}_l}}
\end{aligned} \tag{3.27}$$

First the average of columns are subtracted from the proper values, then the covariances of rows are calculated and divided by the combined standard variations.

We can make the formula work on a three-way array ($\underline{\mathbf{X}}$) with some simple alterations.

$$\begin{aligned}
&\underline{\mathbf{X}} = \left(x_{i,j,k}\right) && \begin{matrix} i = l = q = 1, 2, ..., n \\ j = 1, 2, ..., m \\ k = 1, 2, ..., p \end{matrix} \\
&\underline{\widetilde{\mathbf{X}}} = \left(\widetilde{x}_{i,j,k}\right) && \widetilde{x}_{i,j,k} = x_{i,j,k} - \bar{x}_{j,k} \\
&\mathbf{P} = \left(\rho_{l,q}\right) && \rho_{l,q} = \frac{1}{mp-1} \frac{\sum_{j=1}^{m} \sum_{k=1}^{p} \widetilde{x}_{l,j,k} \widetilde{x}_{q,j,k}}{\sigma_{\widetilde{x}_l} \sigma_{\widetilde{x}_q}}
\end{aligned} \tag{3.28}$$

4 Experimental and methods

> *Science is what we understand well enough to explain to a computer. Art is everything else we do.*
>
> **— Donald Knuth**

4.1 Programming

An extensive part of our work was taken up by creating the computational background for the experiments. This thesis does not focus on the programming aspect, however we want to emphasize here that no commercial software was used, every script was written by the author. All calculations were executed in the programming language R [52] with the help of Rstudio software [53].

4.2 Chromatographic data simulation

Chromatographic peaks were generated with the exponentially modified gaussian peak (EMG). It has five parameters: t - time (abscissa), t_R - retention time, A - area under the peak, σ - width and τ - time constant (asymmetry) [54]:

$$\begin{aligned} \alpha &= \exp\left(\frac{\sigma^2}{2\tau^2} - \frac{t-t_R}{\tau}\right) \\ \beta &= \operatorname{erf}\left(\frac{t-t_R}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}\tau}\right) \end{aligned} \qquad y(t) = \begin{cases} \frac{A}{2\tau}\alpha(\beta+1), & \tau > 0 \\ \frac{A}{2\tau}\alpha(\beta-1), & \tau < 0 \end{cases} \tag{4.1}$$

In some cases baseline with a simple formula was added to the chromatograms:

$$y(t) = at^b + c \tag{4.2}$$

The noise was generated with normally distributed random number. The chromatograms were simply build by adding their elements – peaks, baseline and noise – together.

4.3 Changes along the perturbation

The perturbation-induced changes in the series of chromatograms were simulated by changing one parameter of the EMG peaks (Eq. 4.1) along the second dimension of the dataset - between the chromatograms - where the perturbation occurs. The functions which were used in the experiments are summarized in Table 4.1. The emphasis was on the comparison of monotonous and non-monotonous functions and to cover a wide variety of changes, but keeping it simple with relatively few examples.

Table 4.1.: The formulas of the changes

#	Type	Formula	#	Type	Formula
1	linear	$ax + b$	5	sine	$a \sin x + b$
2	quadratic	$ax^2 + b$	6	cosine	$a \cos x + b$
3	exponential	$ae^x + b$	7	single	$\begin{cases} ax + b, & \text{if } x = c \\ 0, & \text{otherwise} \end{cases}$
4	EMG	Eq. 4.1			

4.4 Properties of generated examples

The parameters of generated chromatograms are presented as tables in the appendix which were the bases of plots in Chapter 5. In the tables only the names of change-types appear, the exact formulas can be seen in Table 4.1. The parameters which were changed are highlighted in red, the others were kept throughout the perturbation.

4.5 Reproducibility of chromatographic columns

4.5.1 Reproducibility of chromatographic retention data

Felinger et al. studied the reproducibility of chromatographic retention data on reversed-phased high performance liquid chromatography (HPLC) columns, and identified the factors that influenced the reproducibility [55]. The systematically measured data were provided by Kele et al. [56, 57]. From the huge amount of experimental data, the results of a test mixture (mixture #3 in Refs. [56, 57]) on five Symmetry C18 (Waters, Milford, MA, USA) columns were available for the comparison. The test mixture contained thiourea, theobromine, theophylline, caffeine, pyridine, phenol, 2,2-dipyridyl, and 1,3-dihydroxynaphtalene eluted in methanol-water (30:70 v/v). The details of the experiments are given by Kele et al. [56, 57].

4.5.2 Isotherm reproducibility

Continuing the work of Kele et al. [56–61], Gritti et al. studied the reproducibility of HPLC columns under nonlinear conditions, and determined the isotherms of overloaded band profiles [62, 63]. Felinger et al. in turn examined the reproducibility of the equilibrium isotherms with principal component analysis [64]. The examination included seven samples, each containing one component: aniline, ethylbenzene, phenol, caffeine, propranolol with or without buffer, and theophylline.

The isotherms were determined for every analyte, respectively, on ten Kromasil C18 (Eka Nobel, Bohus, Sweden) columns with the inverse method. PCA [65] was then performed on the isotherm parameters to compare the HPLC columns [64].

4.5.3 Smoothing and baseline correction

When one works with real measured chromatograms, it is necessary to integrate a data pretreatment step to the calculations, otherwise the correlation maps

will contain false information or the expected results may be hidden by the artifacts of the baseline drift or noise. In this work the asymmetric least squares approach [66] is selected for baseline correction. For smoothing, the Savitzky–Golay algorithm [67] was used, which is the most common technique in this field [68].

4.5.4 Scaling

It is often a problem that some features in the chromatograms stay hidden and therefore valuable information is lost. That occurs because of the great difference between the magnitudes of the individual peaks. The solution to this problem is a properly chosen scaling. Scaling may appear in many forms; in this study we use the generalized scaling, described by Noda [15]. After the treatment, also the least pronounced patterns of the maps become visible. The scaling parameters, however, should be chosen carefully, because the drawback of the method is that it produces artificial baseline- and noise-like disturbing patterns.

4.5.5 Top view of chromatograms

During the comparison of chromatograms, the correlation maps provide the differences in the data as expected. But for a better understanding, we have to know which sample is responsible for a given type of difference, and the synchronous and asynchronous maps are unable to tell us. We chose the simplest solution to this issue. The data were already arranged in a matrix. We simply plotted this chromatomatrix in the same manner as the correlation maps, in the form of contour maps. By this way we obtained a pseudo 3D graph, where the abscissa represents the retention time and the ordinate represents the number of the sample. With that representation, every difference in the chromatograms can be clearly paired with the given sample.

4.6 ALA in practice

The thesis presents two experiments with measured data aimed to highlight the advantages of ALA. The first one is constructed by changing the concentration of seven compounds, listed in Table 4.3, throughout a series of chromatograms and then comparing the results to computer generated data. The parameters of this data are presented in Table A.26 where A is equivalent to the concentrations of the measured chromatograms. The experimental conditions are listed below (Table 4.2). The second experiment has the same conditions except the sample concentrations were permanently 0.4 mg/ml and the solvent composition was changed. This perturbation was started at 0:100% methanol:acetonitrile ratio and ended at 100:0% with 10% steps.

Table 4.2.: Parameters of the chromatographic (SFC) system.

instrument:	Waters ACQUITY UPC^2 System
column:	Supelcosil ABZ+Plus (alkylamide, 3 μm, 4.6x150 mm)
column temperature:	60°C
mobile phase:	100% CO_2
flow rate:	1.0 ml/min
solvent:	acetonitrile
sample temperature:	25°C
injection volume:	2 μl
detection:	192, 200 and 260 nm
backpressure:	150 bar

The parameters of peak height changes are in Table A.26 in the appendix. There were seven compounds listed below:

Table 4.3.: The compounds and their retention times in the experiments linked to Fig. 5.69 b), 5.70 b)

#	t_R (min)	Compound
1	1.68	ethylbenzene
2	1.91	butylbenzene
3	2.22	hexylbenzene
4	2.64	octylbenzene
5	3.2	decylbenzene
6	3.92	dodecylbenzene
7	4.88	tetradecylbenzene

5 Results and discussions

> *Science does not know its debt to imagination.*
>
> **— Ralph Waldo Emerson**

5.1 2DCOR in chromatography

Our first goal is to show the properties and benefits of 2DCOR with chromatographic examples. 2DCOS has a long history and detailed background, however the few papers working in chromatography do not pay attention to the fundamental workings. In theory there is no difference if the type of data was changed (in this case from spectra to chromatograms), but it has to be clarified and also a description is inevitable for the reader to have a glimpse at what 2DCOR is capable of.

5.1.1 Linear changes

The simplest case is where only the size of peaks – area under the peak – is changing and just with a linear function. In the first example four peaks are changing along the perturbation with four different rates and different directions (Fig. 5.1 and 5.2). Two of them have positive changes, the others have negative (Table A.1).

From the synchronous correlation map (SCM) (Fig. 5.2 a)) we can see three peaks in the diagonal called auto-peaks and many off-diagonal peaks called cross-peaks. Auto-peaks give information about the magnitude of change on the given peak while cross-peaks tell the correlations between the changes. As mentioned there are three auto-peaks meaning three changing peaks, but we can see cross-peaks with another peak at 2.0 min, meaning there are not three but four changing peaks in this system. The heights of the auto-peaks show

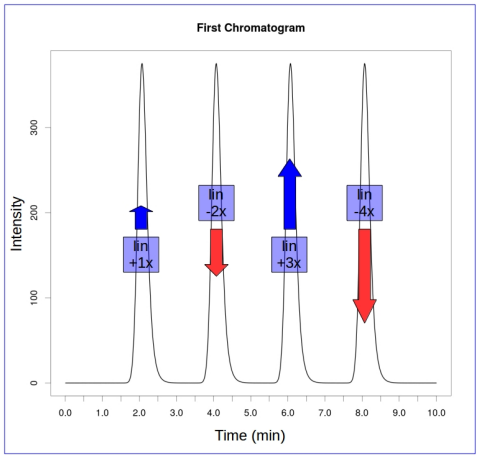

Fig. 5.1.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four linear changes with different rates and different directions.

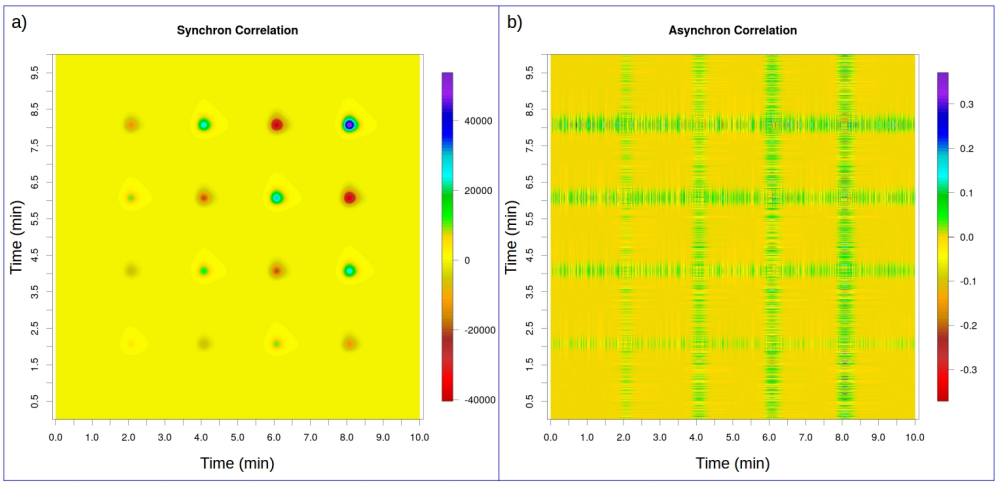

Fig. 5.2.: a) synchronous and b) asynchronous correlation maps with four linear changes.

four different rates (as expected). The first (2.0 min) has the smallest change, the fourth (4.0 min) has largest.

The presence of large cross-peaks means that the changes are much correlated, meaning they have similarly executed changes. As we know, they have the same linear function with different rates, so these patterns just show that. The heights of cross-peaks are a combination of the two corresponding auto-peaks. The directions of cross-peaks indicate the directions of the two corresponding peaks. The SCM is always symmetrical to the diagonal, so the coherent peaks have the same direction. If the cross-peaks are positive that means the two peaks are changing in the same direction, otherwise one is positive while the other is negative, but we can not tell which is which. The auto-peaks are

always positive. Thus from this patterns we can say that first (2.0 min) and third (4.0 min) peak and second (6.0 min) and fourth (8.0 min) have the same directions.

Only noise can be seen on the asynchronous correlation map (ACM) (Fig. 5.2 b)) which means there are no asynchronous changes present. Later examples will show the details of this map.

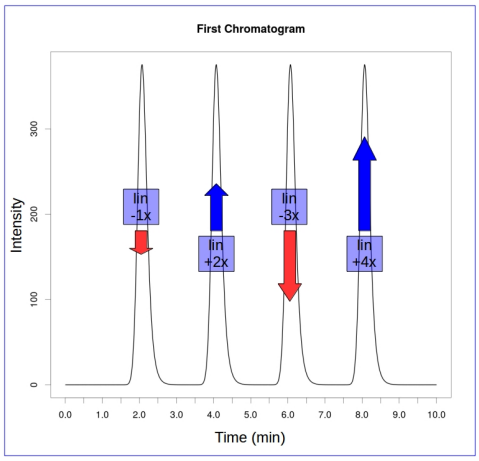

Fig. 5.3.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four linear changes with different rates and different directions. It is present to illustrate the directional anomaly compared to Fig. 5.1.

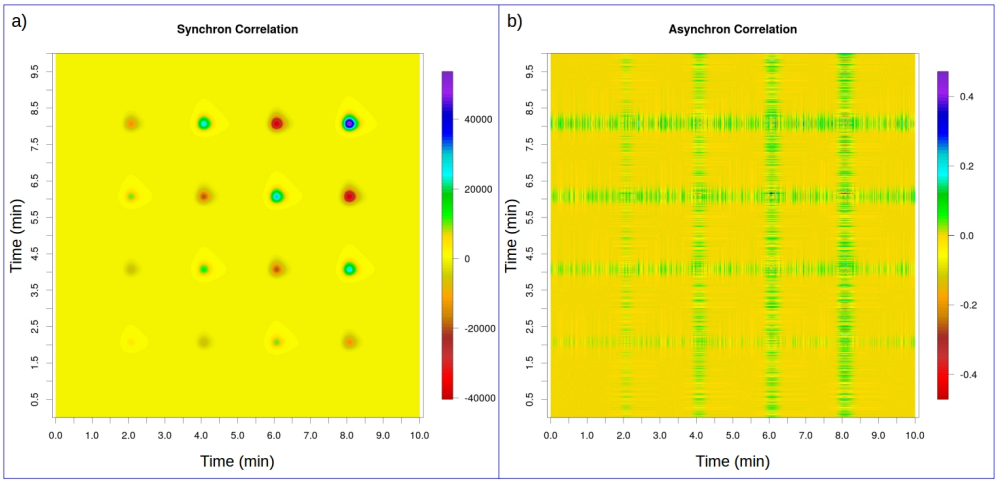

Fig. 5.4.: a) synchronous and b) asynchronous correlation maps with four linear changes to illustrate the directional anomaly compared to Fig. 5.2.

The next figures (Fig. 5.3 and 5.4) are almost identical to the previous ones (Fig. 5.1 and 5.2), the difference can only be found in the parameters (Table A.2). In this example the two pair of peaks are still changing in the same direction but between them it has shifted. As discussed above SCM shows the similarity or

difference between the directions of two changes, however the exact way is kept hidden.

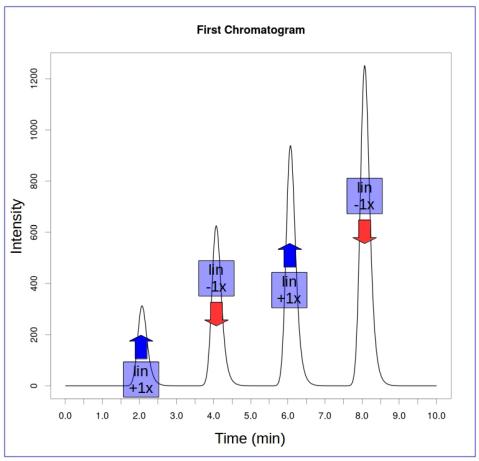

Fig. 5.5.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four linear changes with same rates but different original values.

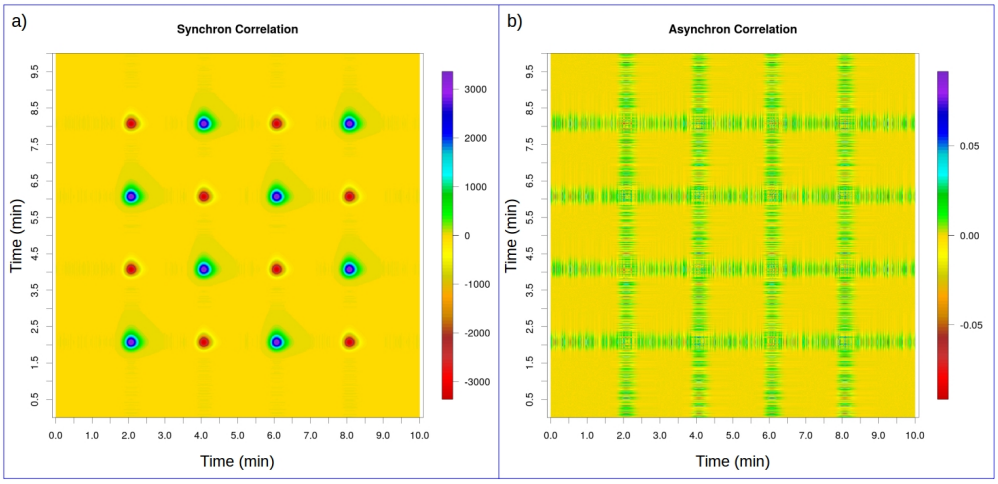

Fig. 5.6.: a) synchronous and b) asynchronous correlation maps with four linear changes with different bases.

Figs. 5.5 and 5.6 show that the heights of correlation peaks only depend on the magnitudes of changes and do not consider the initial heights of chromatographic peaks (Table A.3). That means if the compound of mixture we want to attend is negligible to other peaks in the initial chromatogram, it will not bother our correlation study, because if we can generate considerable change selectively on that peak, the original handicap becomes obsolete, we will see clearly our sign of interest on the correlation maps without any other interfering peaks.

The patterns of cross-peaks still show the coupled directions of changes with positive and negative peaks. The rates are similar thus every peak in SCM have the same heights.

5.1.2 Other monotonous changes

Linear change is the ideal case in correlation experiments, however in practice much more complex situations will appear. Now we move on step by step to these cases. First quadratic (Fig. 5.8) then exponential (Fig. 5.10) changes are present on the maps, similar to the previous arrangement (Fig. 5.7, Table A.4; Fig. 5.9, Table A.5).

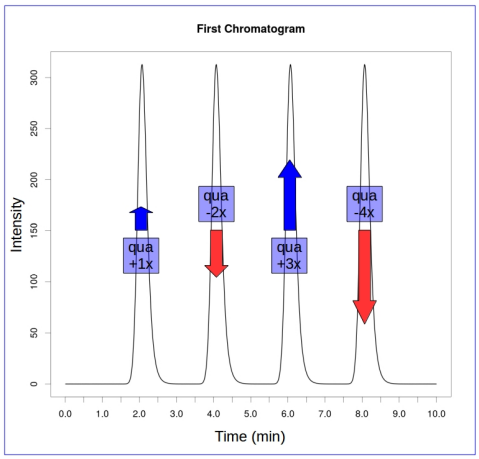

Fig. 5.7.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four quadratic changes with different rates and different directions.

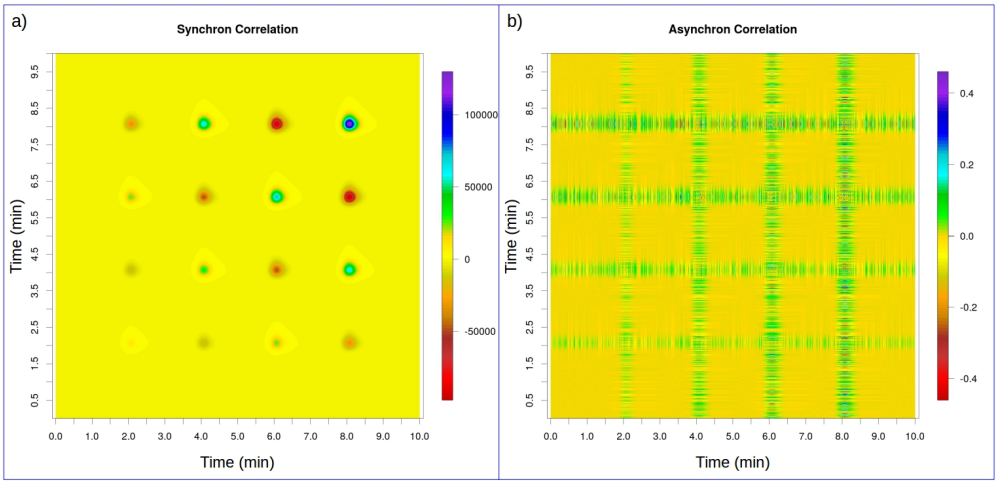

Fig. 5.8.: a) synchronous and b) asynchronous correlation maps with four quadratic changes.

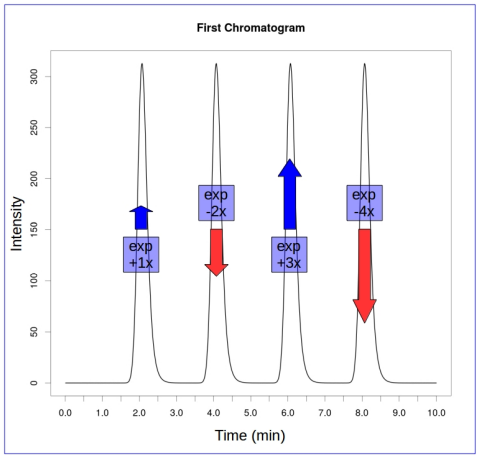

Fig. 5.9.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four exponential changes with different rates and different directions.

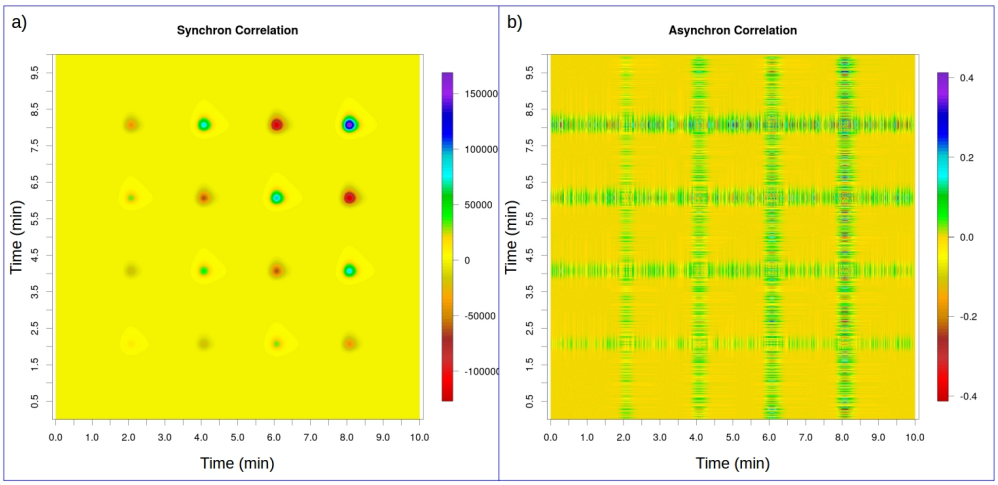

Fig. 5.10.: a) synchronous and b) asynchronous correlation maps with four exponential changes.

The asynchronous maps (Fig. 5.8 b), 5.10 b)) continue to have only noise. The synchronous maps (Fig. 5.8 a), 5.10 a)) are almost identical to Fig. 5.2 a), only the magnitudes differ.

The next two figures (Fig. 5.12, 5.15) have more than one kind of change (Fig. 5.11, Table A.6; Fig. 5.14, Table A.7) and this is the point where 2DCOR starts to reveal its real treasures. The ACM becomes relevant and the separation of different types appears.

We can see from the SCM (Fig. 5.12 a)) that there are six changing peaks (six auto-peaks) with the same direction (all positive cross-peaks), but with different magnitudes (variant heights for the peaks). The interesting part is the ACM

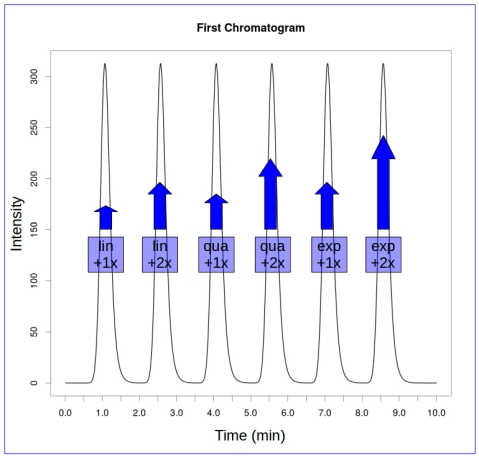

Fig. 5.11.: Visual representation of the given changes in chromatograms, projected on the first point in the series: three distinguishable changes with two different rates respectively.

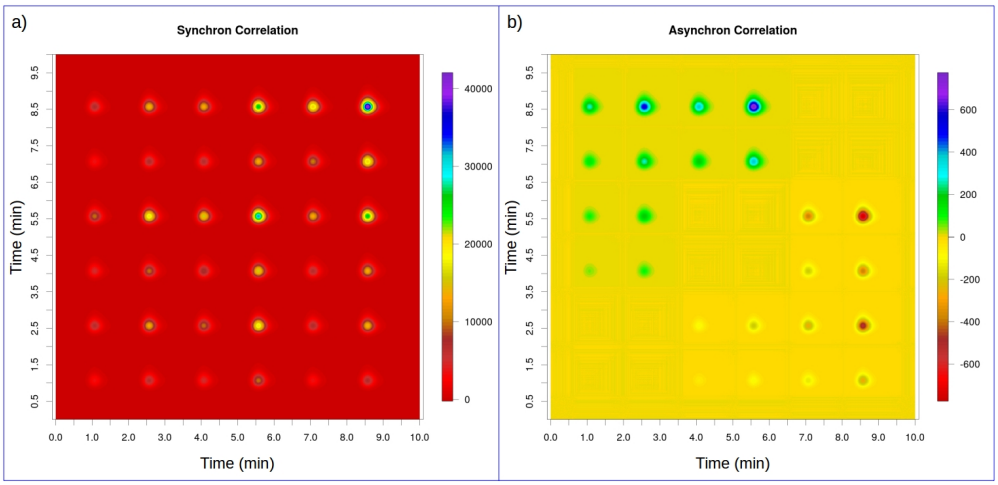

Fig. 5.12.: a) synchronous and b) asynchronous correlation maps with different monotonous changes.

(Fig. 5.12 b)). Now we have peaks in this map as well but only cross-peaks. ACM never has auto-peaks. Unlike SCM this map is not symmetrical to the diagonal. The presence of asynchronous peaks means that there are changes with different runs. Their functions in the dimension of perturbation differ. Please note that not all possible cross-peak has appeared, meaning every two peaks are synchronous and asynchronity has happened between these pairs. Arguably the most useful feature of 2DCOR is to tell the sequence of events in the chemical system. In Fig. 5.12 b) the peaks in the lower half are negative, the others are positive, this indicates that the first two changes happened first, the two in the middle of chromatogram were next and the two at the end were last. As we know they had linear, quadratic and exponential functions respectively

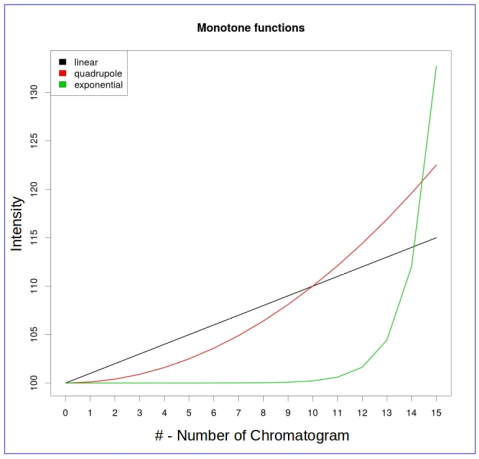

Fig. 5.13.: The monotonous functions used in the experiments.

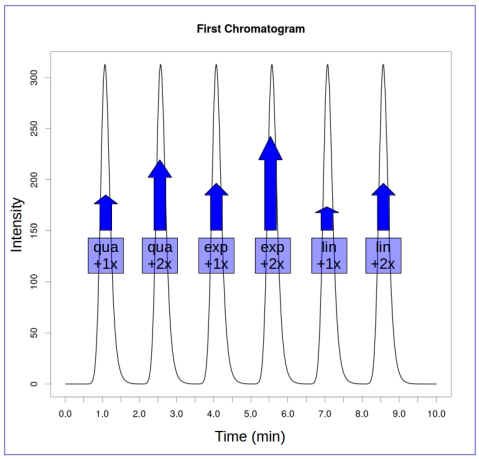

Fig. 5.14.: Visual representation of the given changes in chromatograms, projected on the first point in the series: three distinguishable changes with two different rates respectively, but with different order from the previous one (Fig. 5.11)

and if we look at the graph of these functions maybe it becomes clear why is that order (Fig. 5.13).

The second graphs (Fig. 5.15) show similar features, however the pattern of the ACM has been mixed up. The order is as follows: last, first and second pair of peaks. Compared to the parameters of chromatograms in Table A.7 it is in harmony with our intentions.

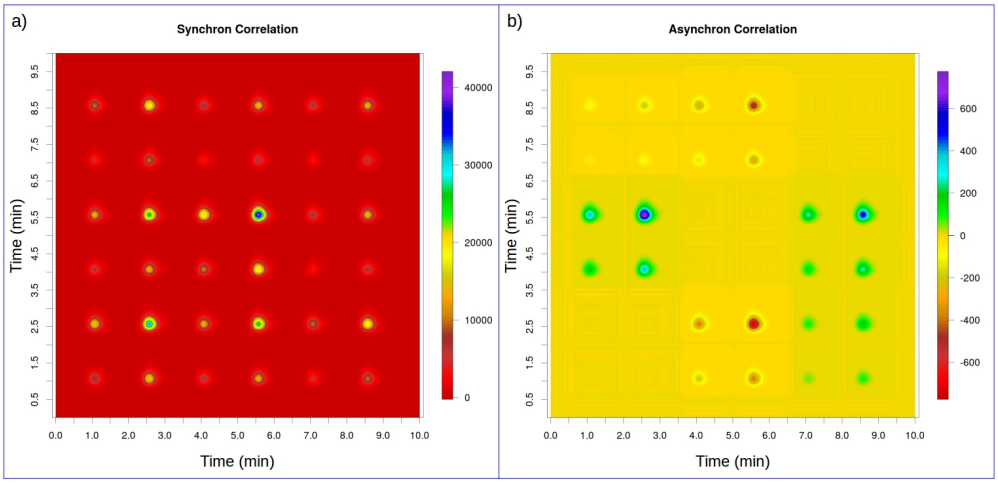

Fig. 5.15.: a) synchronous and b) asynchronous correlation maps with different monotonous changes.

5.1.3 Non-monotonous changes

In this chapter non-monotonous changes are considered. The last section dealt with in some ways simpler cases, what we mean is monotonous, especially linear, changes are easier to interpret. Theoretically there is no difference in the sense that 2DCOR can handle both, the patterns appearing on the maps are not much harder to explain. The physical meaning behind these changes can be tricky to grasp on, however real-life examples surely contain at least portions of non-monotonous changes.

What we see in the next example (Fig. 5.16 and 5.17) is very similar to where we started, the simple linear change (Fig. 5.2). The features discussed there are applicable here as well. However from the chromatograms perspective it is different (Table A.8) and still the maps are almost identical. This is supposed to illustrate that 2DCOR is not concerned with the source of changes or their individual properties of them. It emphasises the connections or correlations between them. This becomes clearer when later we compare 2DCOR with ALA.

Although the maps have not changed much compared to the previous chapter, with this kind of non-monotonous change we can set a third parameter (c, Table 4.1) which shifts this single change on the dimension of perturbation. There is never an asynchronous peak with same kind of monotonous changes, but here (Figs. 5.18 and 5.19) with different parameter c (Table A.9) the ACM is filled with peaks.

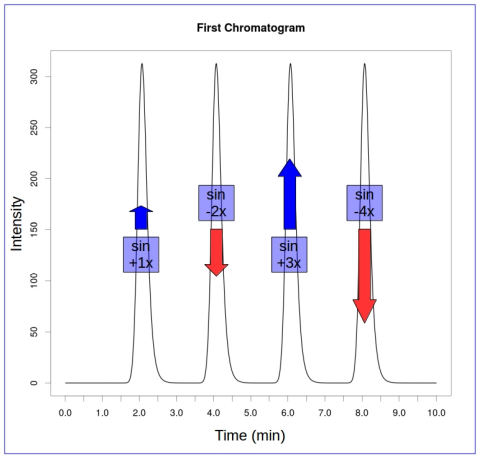

Fig. 5.16.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four single changes with different rates and different directions.

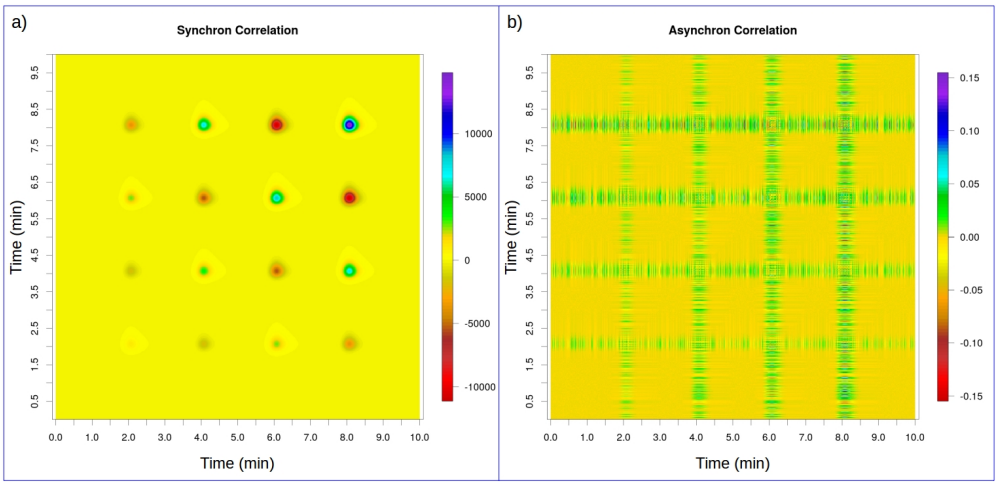

Fig. 5.17.: a) synchronous and b) asynchronous correlation maps with four single changes with different ratios.

Similar patterns are expected as in Fig. 5.17, however there is a small but essential detail: here the peaks are changing in different directions. The maps are still reliable and we get the same information, only our statements about the patterns need to be extended, because the sequence is not determined only by the asynchronous peaks but the combination of synchronous *and* asynchronous cross-peaks, with negative synchronous cross-peaks the outcomes explained in Section 5.1.2 are reversed. With this in mind we can conclude that the sequence is still from the first peak (2.0 min) to the last (8.0 min).

Please also pay attention to the smaller synchronous cross-peaks compared to monotonous changes. These indicate a far-from-linear behaviour of the changes.

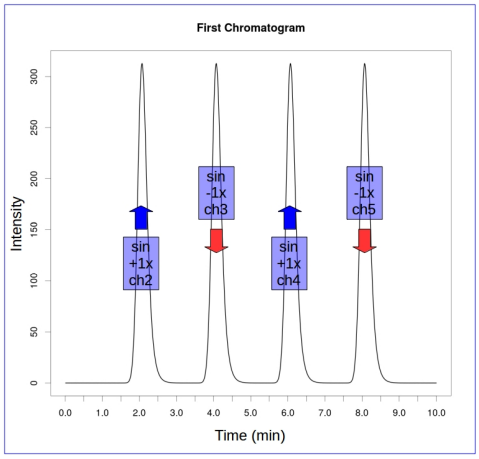

Fig. 5.18.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four single changes with same rates but different directions and at different point in the series.

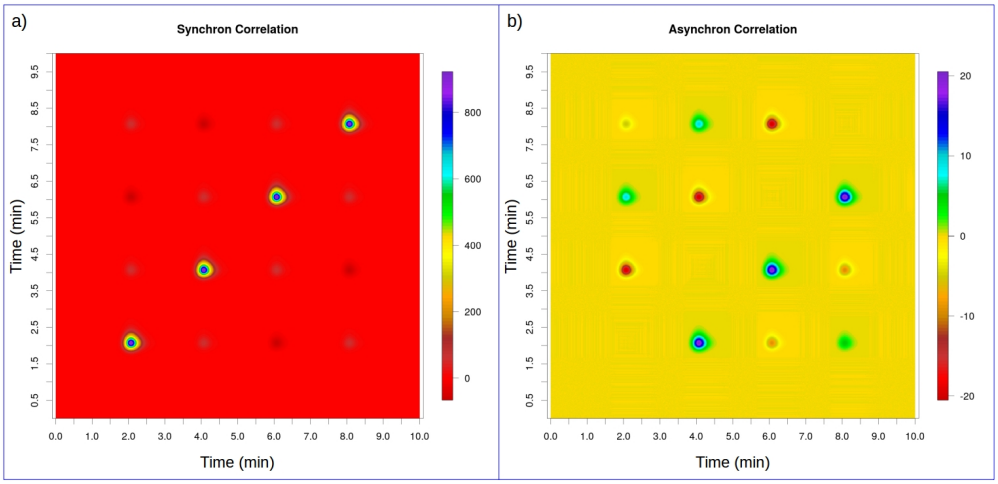

Fig. 5.19.: a) synchronous and b) asynchronous correlation maps with four single changes with different points in the perturbation.

Furthermore we widened the function in the perturbation's dimension so it influences more points in the series of chromatograms, now it has EMG function just like the chromatographic peaks as well. Fig. 5.20 shows the EMGs we used for the next three experiments as seen along the perturbation.

In the first step in this series we simulated four peaks with similar parameters, only the area under the peak (A) is different in the perturbation's function (Table A.10). Fig. 5.21 shows that the principles of evaluation are the same with this kind of change: SCM highlights the four similar changes with the same direction and different ratios and ACM contains only noise for no asynchronity.

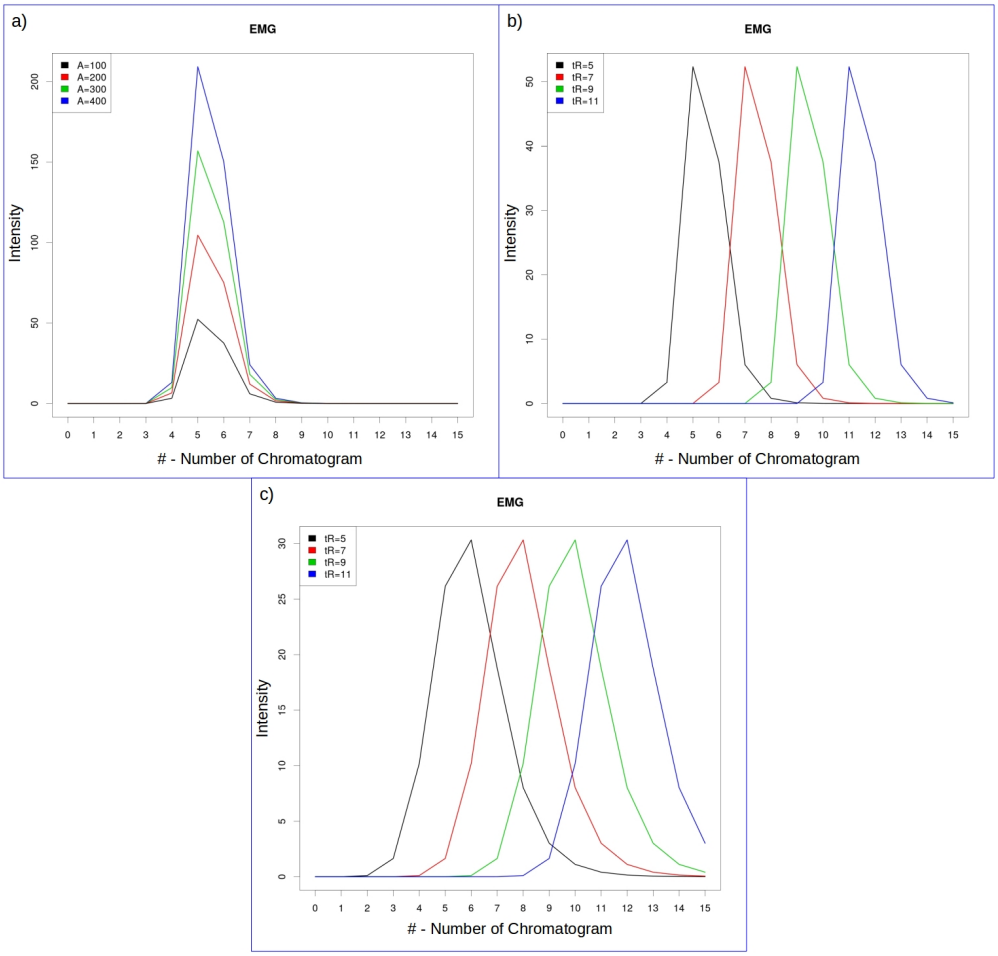

Fig. 5.20.: The EMG functions of perturbation used in the experiments. a) is for Fig. 5.21, b) is for Fig. 5.22, c) is for Fig. 5.23,

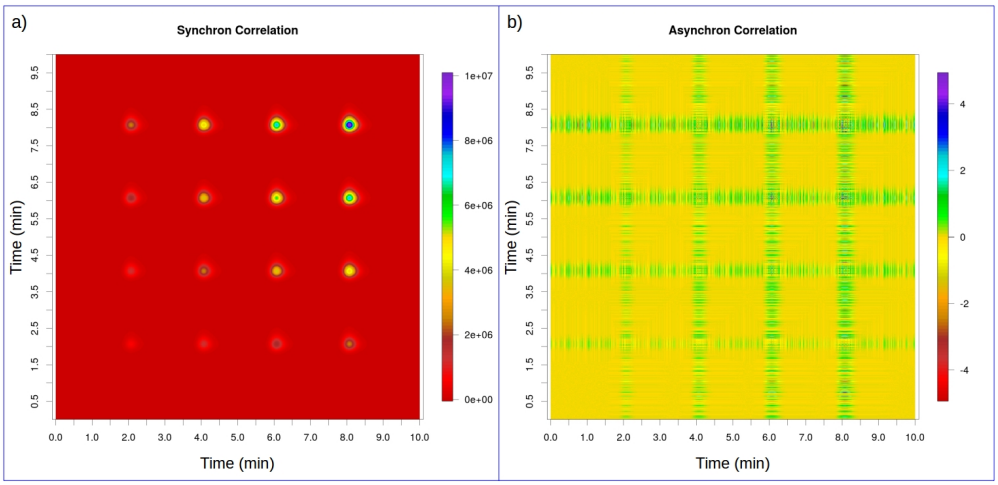

Fig. 5.21.: a) synchronous and b) asynchronous correlation maps with four EMG changes with different ratios.

With EMG function we can set the location of change just like with the single function, technically retention time parameter (t_R) determines where the

change will occur in the points of series. Fig. 5.22 shows what happens when we mix up this detail. At first glance the patterns seem familiar, but soon it can be realized that something is odd. In SCM the closest cross-peaks are positive without exception, indicating that every change is going in the same direction. However the other cross-peaks are negative, creating an impossibility and also mess up the well-founded rules of sequential order examination. Because ACM has a simple anti-symmetrical pattern, but it could meet our expectation (Table A.11) only if the synchronous cross-peaks were positive. In conclusion it is an extreme case and an exception and it can be explained by the function unique shape and the overlaps between the four changes (Fig. 5.20 b)). The start of every curve is in order but while one slope declines the other grows so there is a good chance this phenomenon appears because of this.

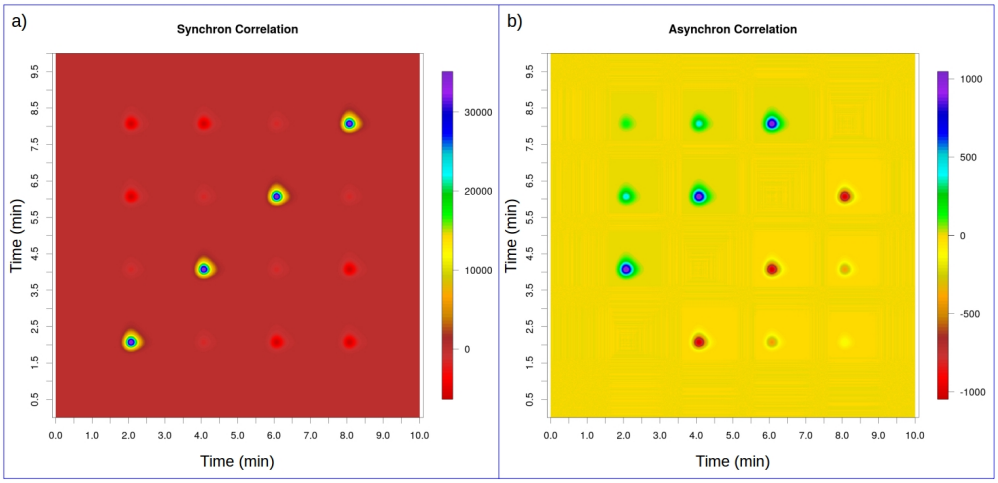

Fig. 5.22.: a) synchronous and b) asynchronous correlation maps with four EMG changes with different maximum points in the perturbation.

We continued to widen the EMGs, now with twice the width (Table A.12). Basically the patterns are kept the same, only the cross-peaks became more prominent (Fig. 5.23). One detail worth mentioning is that the fourth change is not contained within the boundaries of series (Fig. 5.20 c)) and the smaller cross-peaks between the first and fourth peaks are showing this.

The last kinds of function we want to examine is sine and cosine. Handling these are again different from the latter. We saw monotonous changes, others had one extreme point but here the changes constantly fluctuate their direction, naturally it results in unique patterns. As seen in Fig. 5.25 a) there are four synchronous auto-peaks and four (2x2) cross-peaks, but between the two groups there are no cross-peaks. So the first two and last two peaks are synchronous and going in opposite directions. The two groups are, however, totally asyn-

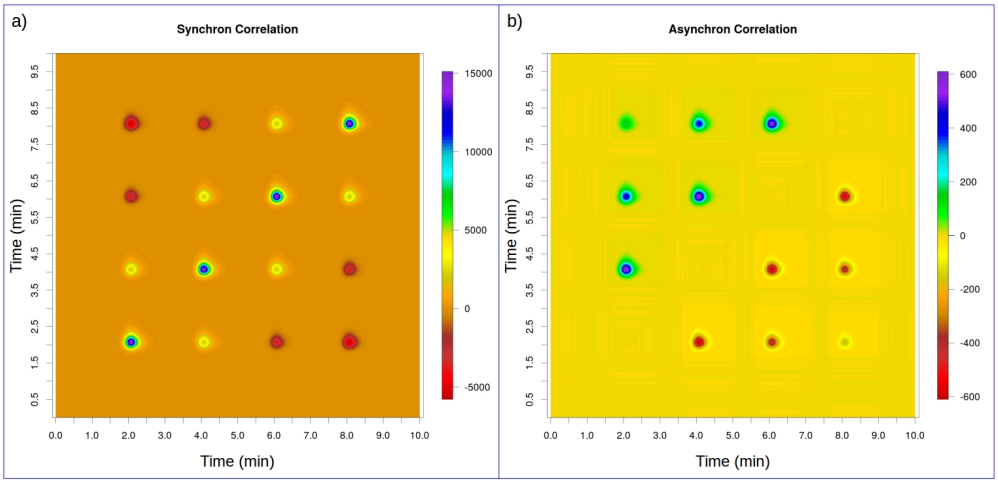

Fig. 5.23.: a) synchronous and b) asynchronous correlation maps with four wider EMG changes with different maximum points in the perturbation.

chronous. From the visual representation (Fig. 5.24) and the parameters in Table A.13 we can see it covers the anticipated.

The ACM (Fig. 5.25 b)) in this case is hard to explain, because there are no synchronous cross-peaks where asynchronous ones are present, but it clearly confirms the results from the previous map. Also, from the functions' view it is hard to point out which change precedes the other because of fluctuation.

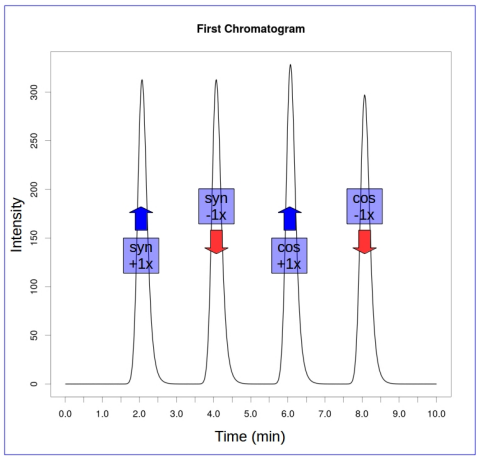

Fig. 5.24.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four wavelike changes, two sinusoidal and two cosinusoidal with the same amplitudes but different directions.

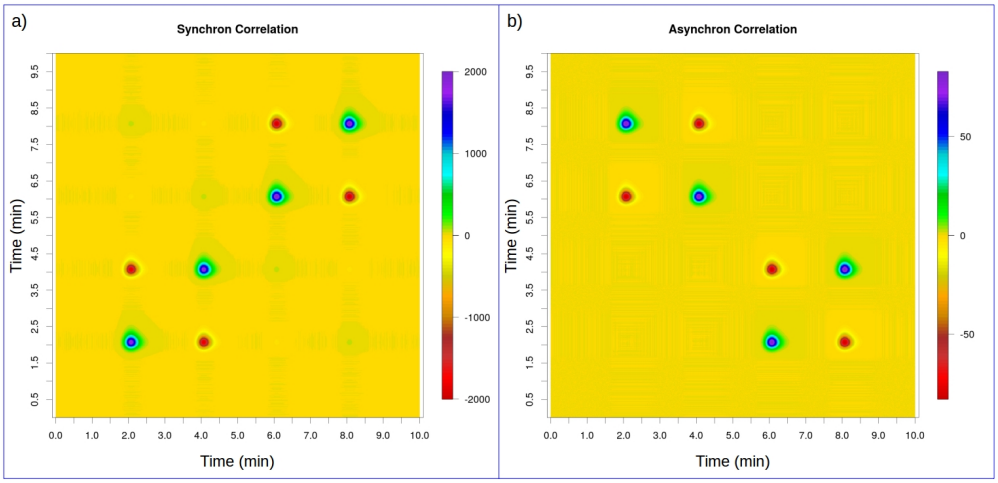

Fig. 5.25.: a) synchronous and b) asynchronous correlation maps with two sine and two cosine changes.

5.1.4 Changes in shape

One can argue that the previously detailed cases can become hard to follow quite fast. All of those changes occurred only on the height of peaks and that is the simplest perturbation which can occur, any other change which involves the shape or location of peak makes the evaluation much more difficult. We tried to cover a wide range of occurrences with height changes and still got dazzled by their extensive complexity. In the next two sections we do not want to touch every aspect but to give an overview in order to keep the discussion relatively simple.

In the original field, spectroscopic peak shape changes are rare and there is no report of migrating peaks. On the contrary in chromatography it is almost unavoidable to have at least minimal distortion and displacement in correlation systems. So the next few examples are very important for us because they diverge from 2DCOS and fundamental in chromatographic correlation study.

The first example in peak shape changes (Figs. 5.26 and 5.27) shows right away the problems with this kind of changes. First it seems even auto-peaks become complex, they have a center and surrounding hills (technically cross-peaks) with different shapes and directions. It is because when σ is changing, the intensities of peak's center and sides move oppositely.

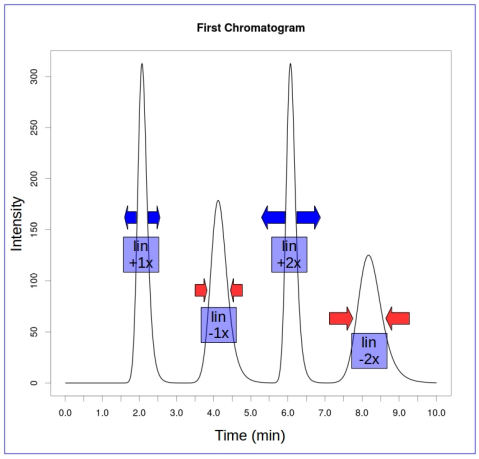

Fig. 5.26.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four linear shape changes with two rates and different directions.

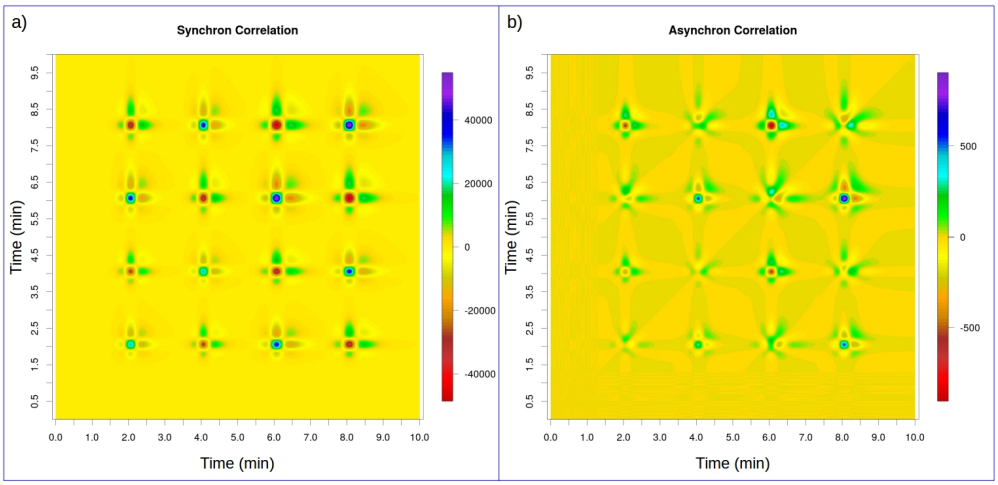

Fig. 5.27.: a) synchronous and b) asynchronous correlation maps with four linear changes in the peak shape.

Furthermore, everything has synchronous and asynchronous values as well despite the fact that the changes on peak shapes were linear everywhere (Table A.14). The solution is very simple: the maps consider only the intensities at certain points and not the functions how they change and the values do not change linearly.

The basic rules, however, still apply to the patterns. We can see from SCM that the first and third peak change in the same direction just like the second and fourth, the first two have the same magnitude which is smaller than the last pair. ACM is again harder to interpret, because it tells a sequence but it has no physical meaning, because we know they are the same changes just in every second cases reversed. It is only good for telling which peak is narrowing and

which is widening, but for that the maps have to contain both. In conclusion we will see unmatched patterns when peak shape changes occur so they can be identified.

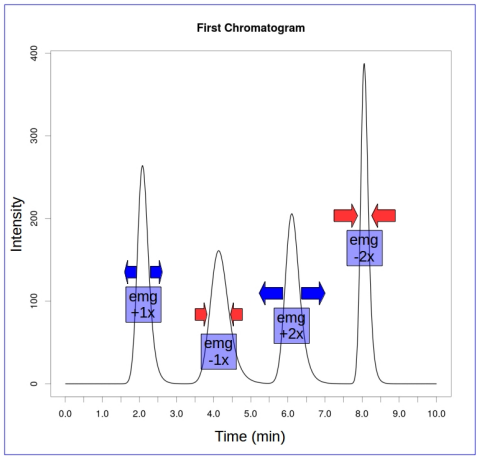

Fig. 5.28.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four EMG changes with two rates and different directions.

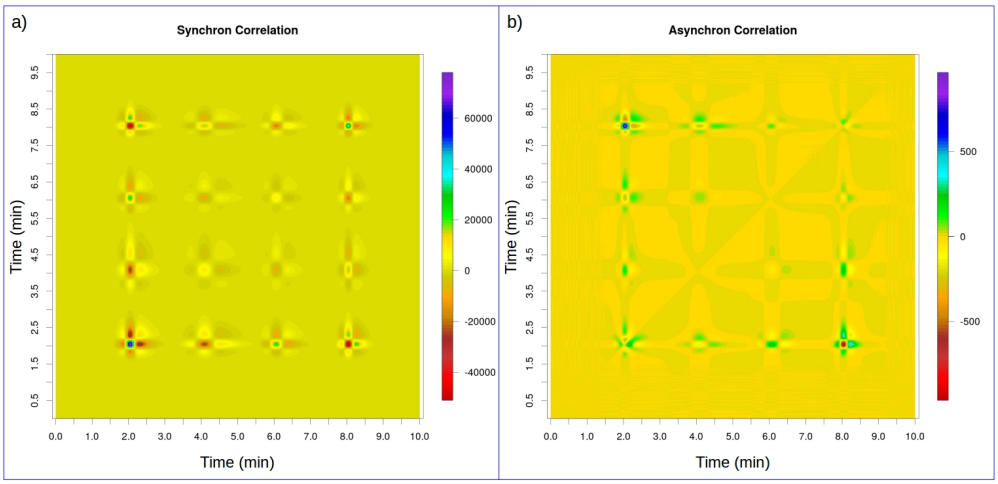

Fig. 5.29.: a) synchronous and b) asynchronous correlation maps with four EMG changes in the peak shape.

The second example has EMG functions in the changes of peak shape to demonstrate how non-monotonous changes appear (Fig. 5.28, Table A.15). The patterns (Fig. 5.29) have no new features here because of the above detailed phenomenon. The EMG's own properties makes the two directions differ from what we saw with linear functions, but we can not distinguish the two. The peak shape changes are more complex than to see deeper the functions what caused them.

5.1.5 Peak migration

The patterns seen in the next three examples also contain complex peaks like the previous ones, because during the migration of peak the maximum point and the sides behave differently. They also have a distinct look which got the nickname "butterfly" in the field due to the ACM's strong resemblance to that insect.

The first example has four linear peak shifts in the same direction (Fig. 5.30, Table A.16), but the first two peaks have longer trajectory than the rest. The mentioned butterfly patterns can be seen accordingly to the parameters (Fig. 5.31). What we see becomes more clear when we change the direction of every second peak and look at those maps in Fig. 5.33 with the parameters of Table A.17 (Fig. 5.32). The auto-peaks have not changed but the cross-peaks have shifted. Now we can conclude that 2DCOR can show peak shifts with unmistakable patterns and even tell their directions.

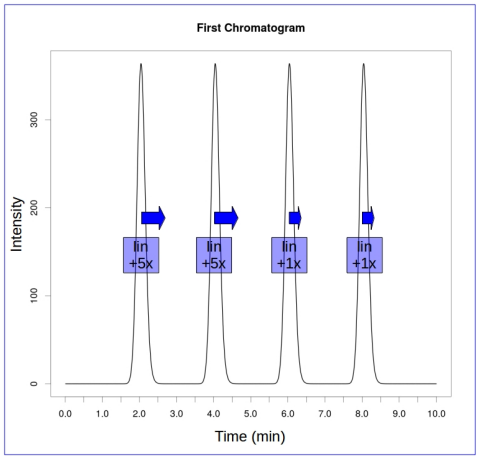

Fig. 5.30.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four linear peak location changes with two rates but same direction.

Non-monotonous peak migrations are just a little different from their monotonous counterparts. We chose EMG functions to change the t_R parameters with, and like in Fig. 5.33 there are two rates and directions. In Fig. 5.35 The butterfly patterns can be seen and we can tell the type of change, only here it is clear that these shifts were not linear, they were more complex (Figs. 5.34 and Fig. 5.35). Similarly to the width changes, ACM is not really informative about sequential order, its sole purpose is to show the directions of changes.

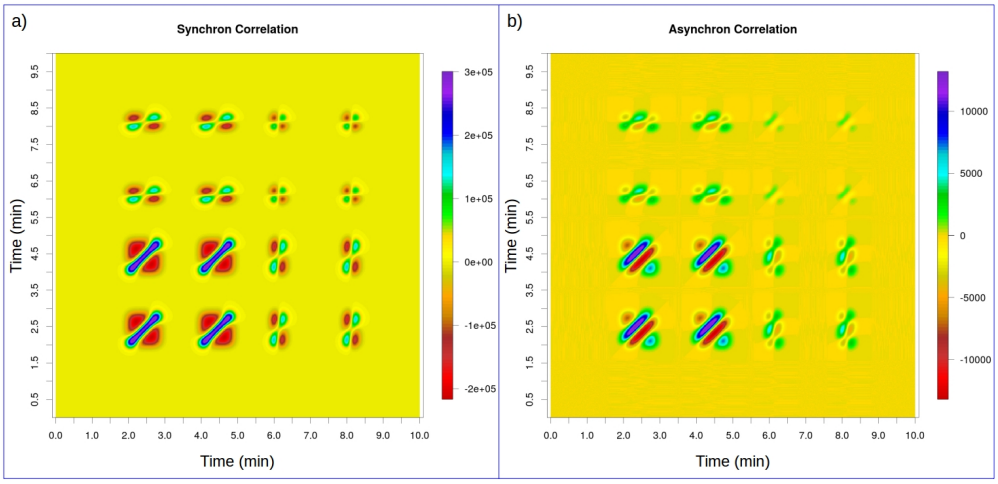

Fig. 5.31.: a) synchronous and b) asynchronous correlation maps with four linearly migrating peaks in the same direction.

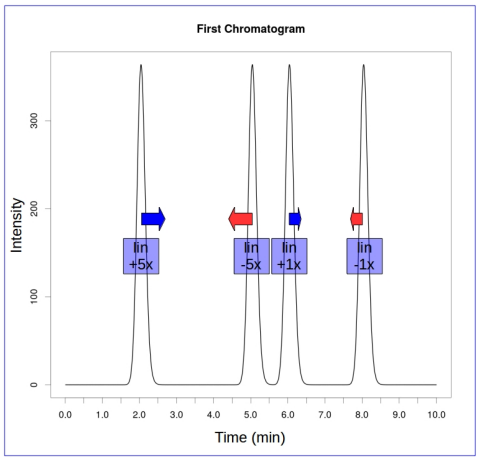

Fig. 5.32.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four linear peak location changes with two rates and different directions.

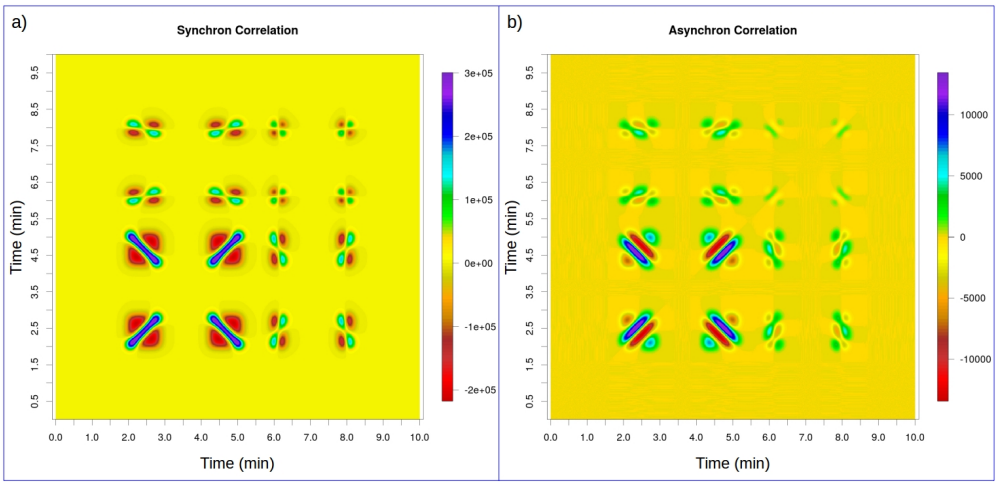

Fig. 5.33.: a) synchronous and b) asynchronous correlation maps with four linearly migrating peaks in two directions.

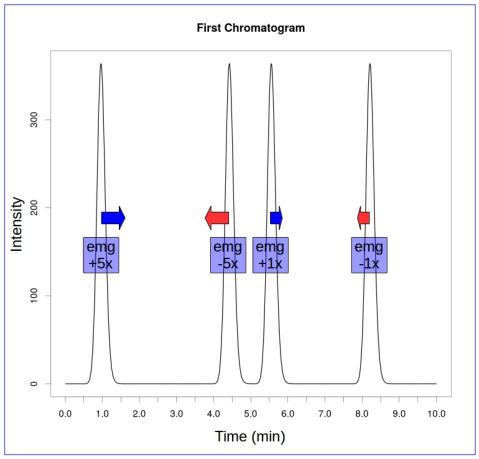

Fig. 5.34.: Visual representation of the given changes in chromatograms, projected on the first point in the series: four EMG location changes with two different rates and different directions.

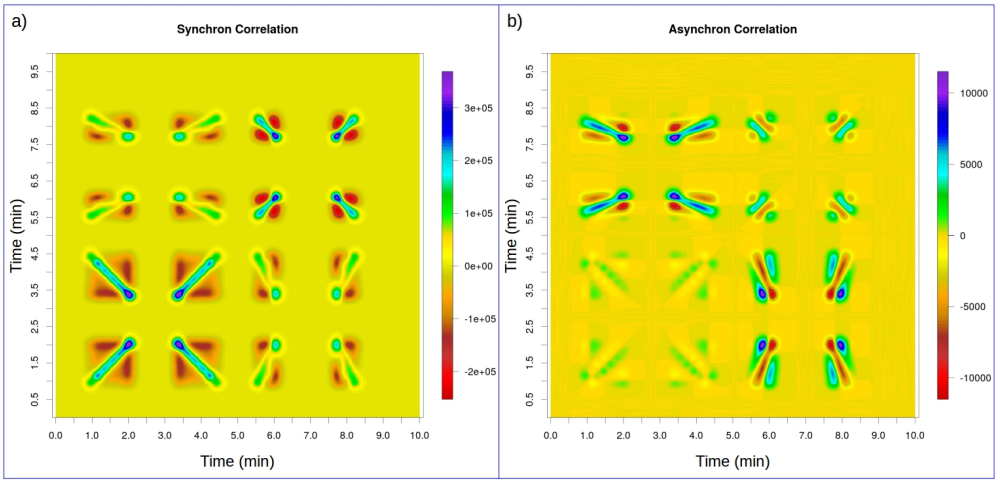

Fig. 5.35.: a) synchronous and b) asynchronous correlation maps with four migrating peaks in two directions with EMG functions.

5.1.6 Overlapping peaks

So far the examples covered separated peaks, however the this section deals with overlapping peaks where the changes interfere with each other or from the other way around the peaks are physically connected. The question is how well 2DCOR can separate those signals. And the answer is: it can exceptionally well. Actually, this has been a strong argument for 2DCOR, because of the 2D environment the resolution is higher compared to traditional 1D projections which leads to easier detection of overlapping peaks.

We set up a chromatographic system with five overlapping peaks, so ten peaks in general, the parameters can be seen in Table A.19. These peaks are narrow and their retention times have little difference, so they are seemingly inseparable as Fig. 5.36 shows with every chromatogram projected to each other. The changes are very small too, they can be barely seen by the naked eye let alone evaluate efficiently.

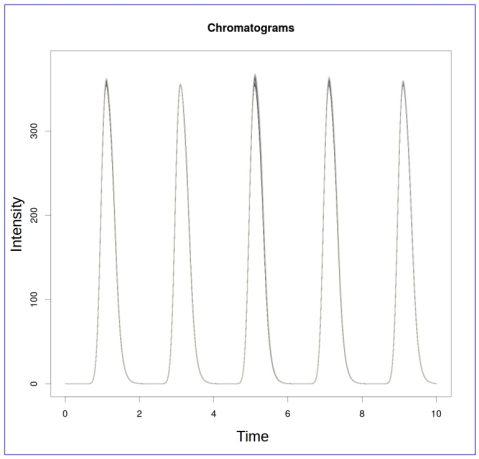

Fig. 5.36.: The series of chromatograms with five (a total of ten) seemingly inseparable overlapping peaks.

Fig. 5.37 contains the correlation maps of those changes. We can see immediately that they are much more informative, they represent those features very extensively. Let us now focus on the separation of overlaps. Only one peak can be distinguished from its pair in SCM (a)), the second group where one peak is changing positively and the other in negative direction. ACM (b)) is more suitable in this case, two more overlaps can be revealed. At the end only the first and the third group looks like one peak. The differences of changes in them are really small, but it still shows the limitations of this feature. In conclusion we can say that 2DCOR enhances our chances to spot overlap-

ping peaks in all cases, however if we manage to evoke very different kind of changes on the overlapping peaks, like different directions or monotonous versus non-monotonous changes, it can be a powerful tool.

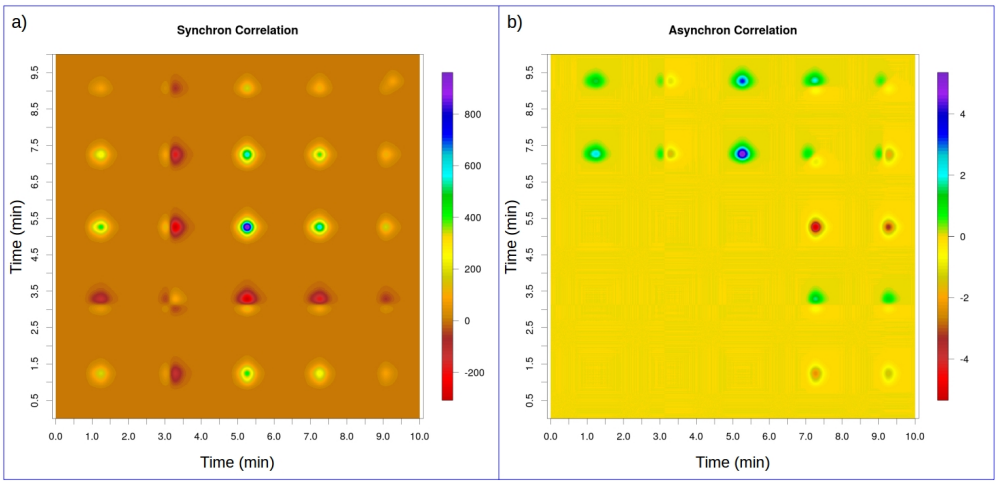

Fig. 5.37.: a) synchronous and b) asynchronous correlation maps with five (a total of ten) seemingly inseparable overlapping peaks.

A second example is presented too because the first one was something like an extreme case and a more usual scenario is expedient to a closure. Here three pair of peaks are present, they are wider, the inflexion points are visible but they are still highly overlapped. However the true nature of changes are again hard to tell, if not impossible, from the projection of the chromatograms (Fig. 5.38). Parameters are in Table A.20.

The patterns are more prominent on every group, no peak or information stays hidden, but remember, they are still highly overlapped peaks with small changes, other visual evaluations can be troublesome. The most interesting change from the previous example is that the two peaks in the first group can be seen clearly despite it is the same change.

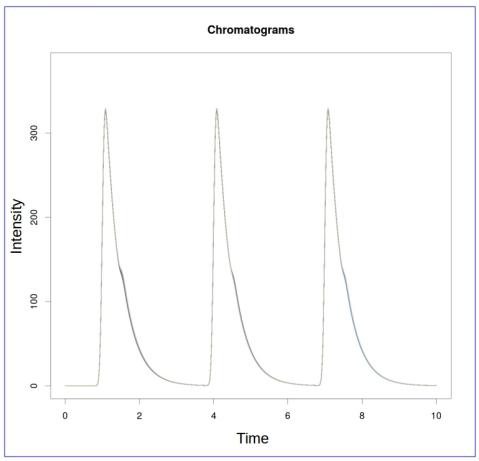

Fig. 5.38.: The series of chromatograms with three (a total of six) highly overlapping peaks.

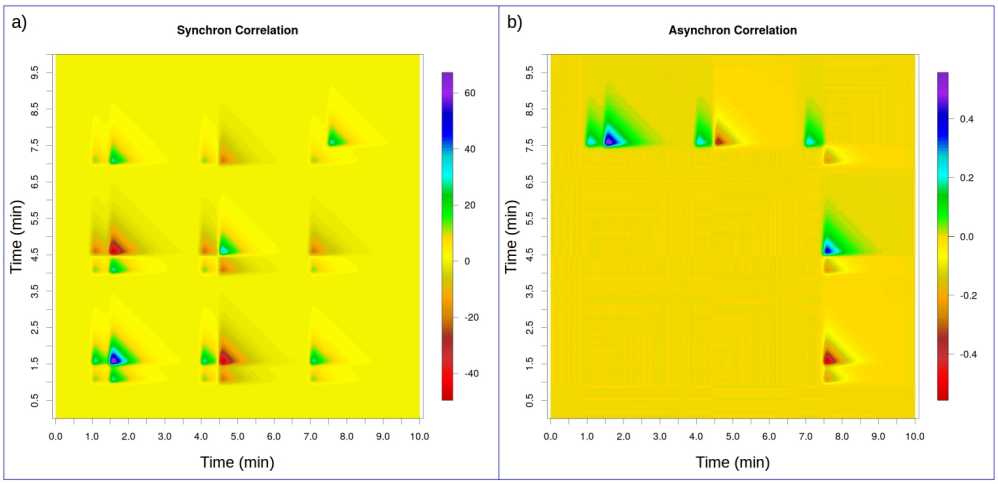

Fig. 5.39.: a) synchronous and b) asynchronous correlation maps with three (a total of six) highly overlapping peaks.

5.2 Introduction to ALA and comparison to 2DCOR

ALA was made to extend the possibilities of 2DCOR to higher dimensions, to successfully make correlation maps on 3D datasets. However it has been proved that because of the simplified mathematics ALA is just as useful in 2D as in 3D and provides a different but in some ways simpler and more practical alternative to 2DCOR.

First we will introduce our method through traditional datasets where perturbation is applied to 1D measurements forming a 2D dataset. The examples will be parallel to the ones in the previous section (Section 5.1). In that part we detailed the properties of correlation maps and the reader could get familiar with the right mindset for correlation experiments. Alteration maps have the same fundamentals so now we can focus on ALA's unique features. Chapter 3 described the calculations for both methods. ALA seems very different there, here we also want to show how the two are still connected.

5.2.1 Linear changes

Before we begin with the examples we should take a look at the alteration maps. Instead of two – as in 2DCOR – there are three maps: the synchronous and asynchronous maps have the same purpose as in the original method, i. e. showing the different aspects of changes triggered by the perturbation, the addition of basic alteration map was necessary because the latter two are normalized and the overall magnitudes of changes have to be shown and only the three together give complete information about the properties of changes as we will see.

Fig. 5.40 shows the alteration maps side by side for the same example as Fig. 5.2 (Table A.1). Compare the two figures and the first and arguably the most important advantage of ALA is shown immediately. We have simple linear plots instead of the 3D maps plotted in pseudo-3D graphs which's two halves are not even necessary, the SCM is always symmetrical while ACM is anti-symmetrical to the diagonal. That means ALA highly improves the simplicity of information gathering. Having three maps is definitely a disadvantage but

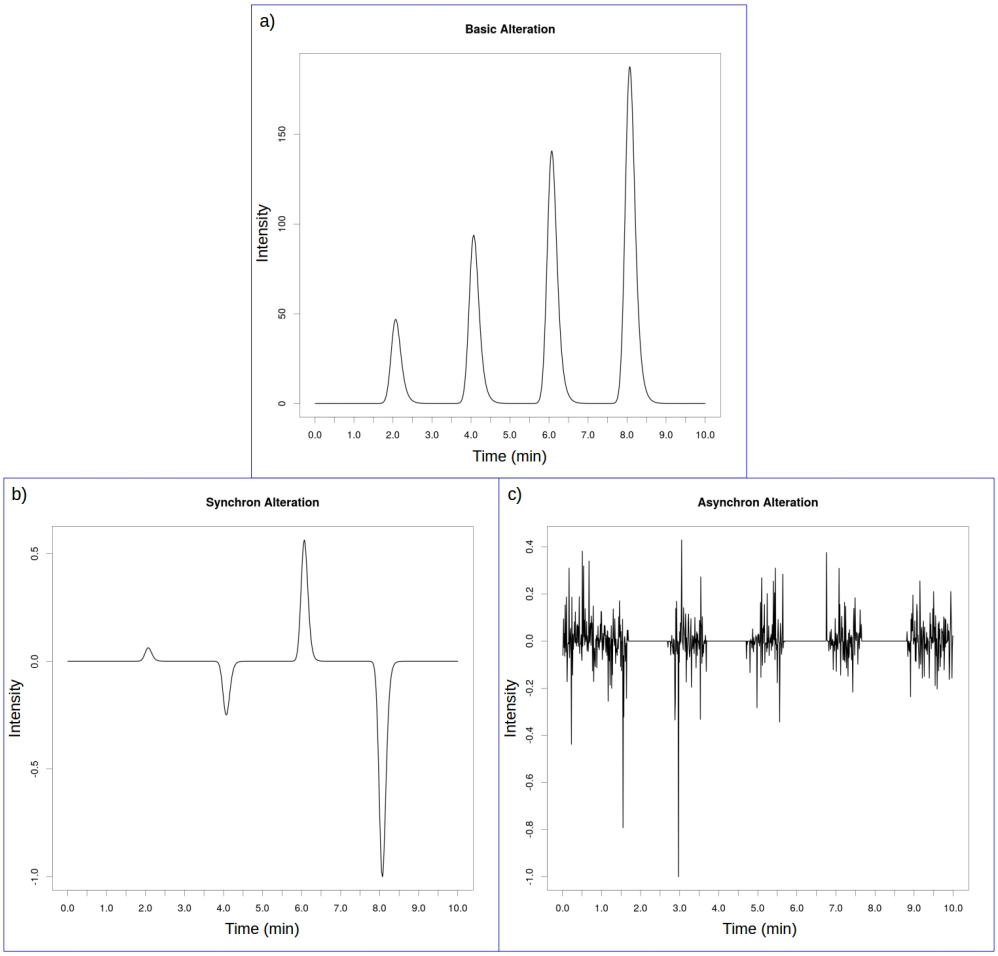

Fig. 5.40.: Example for the a) basic, b) synchronous and c) asynchronous alteration maps.

it is dwarfed by the fact that the nature of these graphs allows us to plot them in only one figure. The combined plot for Fig. 5.40 is shown in Fig. 5.41 a).

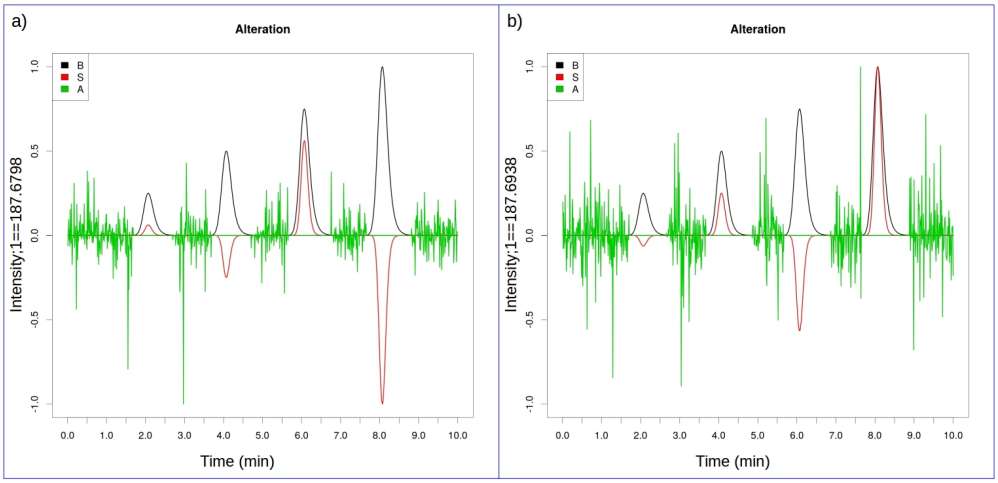

Fig. 5.41.: Alteration maps of four linear changes with different directions. a) the initial set up and b) the reversed.

On the BAM (black) we can see all changes irrespectively of their type. It is always positive and shows only the magnitudes. synchronous alteration map (SAM) (red) shows monotonous changes almost exclusively and asynchronous alteration map (AAM) shows non-monotonous changes. This differs from 2DCOR because the alteration maps separate different types of changes while correlation maps emphasise different behaviours of peak pairs. The concepts of the two methods are maybe not matched, but we will show the parallel outcomes of their evaluations.

Here we can see four changes with four magnitudes. We can not tell if they are linear or not, but they have no asynchronous value, so it is certain that they are monotonous. The AAM has an interesting feature, it is made up only of noise except for the locations where the peaks should be present. This phenomenon has no practical use at this point, maybe in future cases it can be used for peak detection. SAM shows the directions of changes and it can not be emphasised enough that it is the real direction the changes took not just some relative attribute. ALA focuses on the peaks individually and from that it gains the ability to give more detailed information. Meanwhile it lacks direct information on the correlations between changes, which is the main focus of 2DCOR. However this information can be gathered through the comparison of peaks, if two changes have the same values for both three alteration maps, we can safely say they are highly correlated in the sense of 2DCOR. Although we do not want to use that term in this field, because there is no mathematical background to back it.

Thus the directions are additional information in the alteration maps compared to their correlation counterparts. The difference between the two graphs in Fig. 5.41 is that the direction are reversed (Table A.1, A.2). The two ALA plots show exactly this change while the 2DCOR maps remain the same (Fig. 5.2 and 5.4), because the relative directions are not changed.

The next graph (Fig. 5.42) shows the alteration maps where the linear changes have same magnitudes, only the initial heights of the peaks are different (Table A.3). Compared to the correlation maps (Fig. 5.6) they deliver the same information and additionally the exact directions are only shown here.

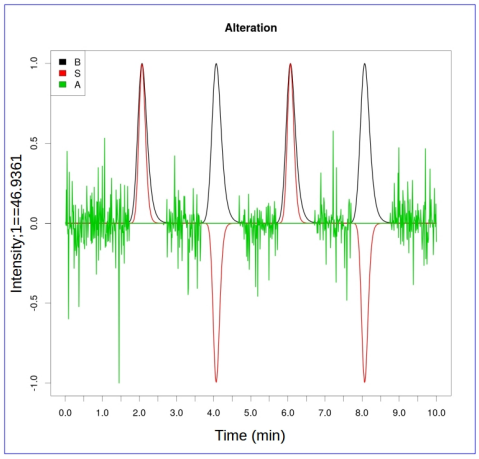

Fig. 5.42.: Alteration maps of four linear changes with different bases.

5.2.2 Other monotonous changes

Besides the linear function, other monotonous changes were studied by correlation maps (Fig. 5.8, 5.10) with the parameters from Table A.4, A.5. The alteration maps for the same quadratic and exponential changes are in Fig. 5.43 a) and b) respectively. The properties are similar to what we discussed with linear changes and still alteration maps have the same information as correlation maps and more.

The interesting feature is to look at the ratios of BAM and SAM together in both graphs and compare them to linear changes (Fig. 5.41). The quadratic and exponential SAM peaks follow the BAM, the synchronous values have linear response to the magnitude of change while linear change has a quadratic-like response in SAM. This occurs because of the calculation of SAM (Eq. 3.16) with the properties of these functions. It makes a bit harder to evaluate more complex alteration maps, but the next example will show that actually it can be managed quite easily and also the noise patterns in AAM will be of help: linear changes have no noise at the peak's location, quadratic changes have noise on the margins of peaks and exponential changes have high noise at the peaks.

The inclusion of more than one types of monotonous changes are revealing another basic difference between 2DCOR and ALA. Both shows the differences of changes, but for that the original method has the asynchronous correlation cross-peaks (Fig. 5.12, 5.15), ALA uses the BAM and SAM ratios (Fig. 5.44). For similar basic peaks the linear change has the highest synchronous peaks and

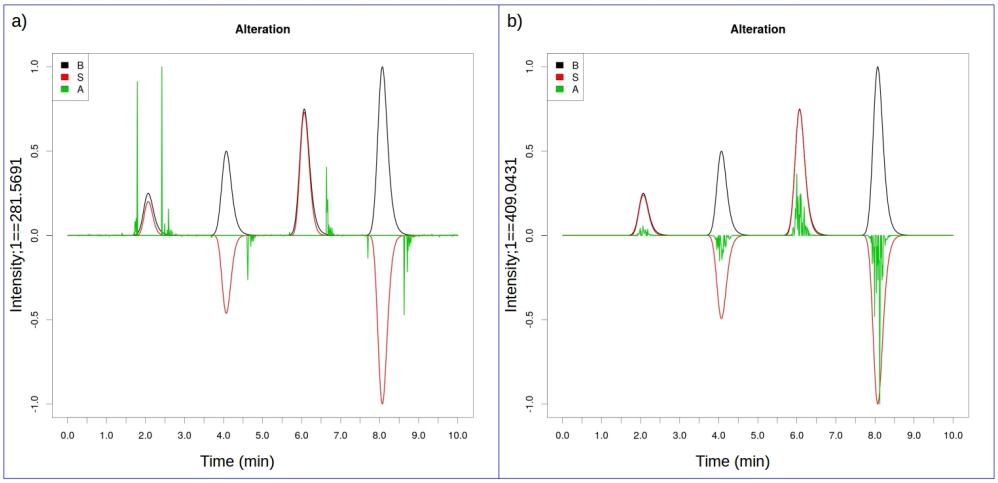

Fig. 5.43.: Alteration maps of a) four quadratic changes and b) four exponential changes.

the exponential has the lowest. In this case the AAM is not really necessary, but it makes it clear which are the exponential peaks. From this information, the sequential order told by 2DCOR can be set.

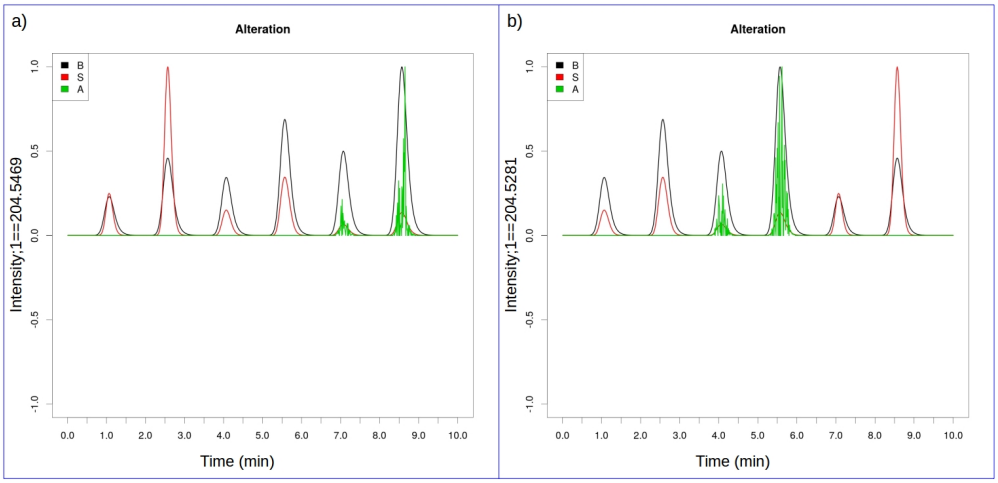

Fig. 5.44.: Alteration maps with different monotonous changes.

5.2.3 Non-monotonous changes

We started with one single change along the perturbation, because it has a unique style in ALA. It has only AAM peaks, so in this way it is the opposite of monotonous changes. They have some similarities in the form of single change's AAM peaks are not going with BAM (Fig. 5.45 a)), they do not have a linear response while the rates of changes certainly are linear (Table A.8). Furthermore we can see the directions too.

Fig. 5.45 b) has the same type of changes only this time the rates are the same, the difference is that these single changes are realized at different points in the series of chromatograms (Table A.9). Correlation maps show the sequential order in an elaborate manner, through comparing synchronous and asynchronous cross-peaks (Fig. 5.19). Unfortunately alteration maps do not have this feature. The sequence of monotonous changes can be obtained, but with this type it lacks this information. We have to mention that this is the only disadvantage of our method compared to 2DCOR, however later we will show a possible solution.

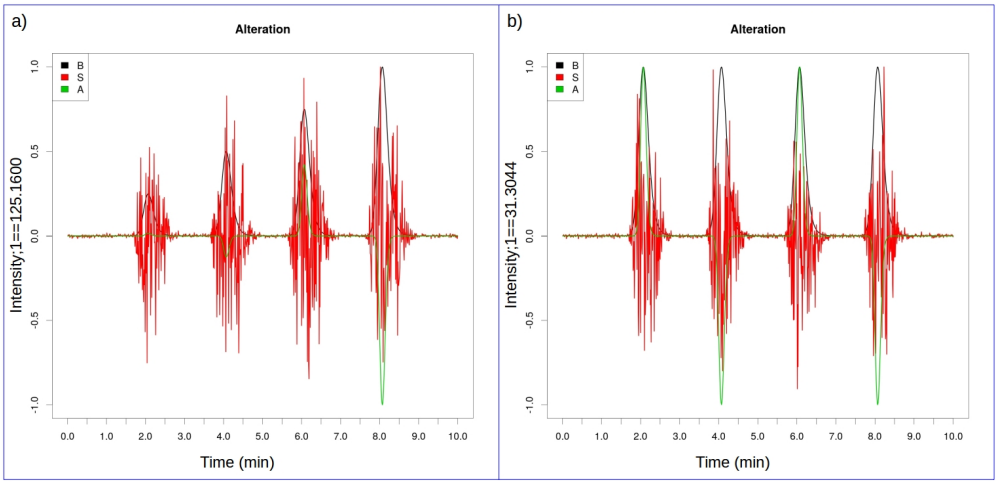

Fig. 5.45.: Alteration maps with four single changes with different rates.

EMG is another non-monotonous function we examined. Its formula is nothing like the single change yet from the perspective of perturbation they are quite similar. The change is localized to one extreme point, only the EMG is affecting more than one. The patterns of alteration maps show this similarity (Fig. 5.46). The ratios of BAM and AAM are identical and SAM has only noise exaggerated at the peaks. This example has the same t_R parameter for every peak (Table A.10), but for the next figure it is different for every peak (Table A.11, A.12). Fig. 5.47 a) and b) has the same retention times for the perturbation function, but they are wider in b). The difference is that a recognizable synchronous peak appears in a) at the last peak and b) has all four. It can be explained by the properties of alteration maps. SAM was build to show monotonous changes, not exclusively but to emphasise them, one single change is the opposite of that so SAM shuts it out entirely. However when monotonous behaviour gets more prominent, SAM notices, thus the sequential order is appearing. It is a reversed order: the latest change has the highest value, but it is shown nevertheless. Our studies suggest that at least half of the points along the perturbation has to be involved for SAM to identify the sequence. If it was not the case, the noise

vales the synchronous peaks. However, further improvements on the method may be able to include this feature.

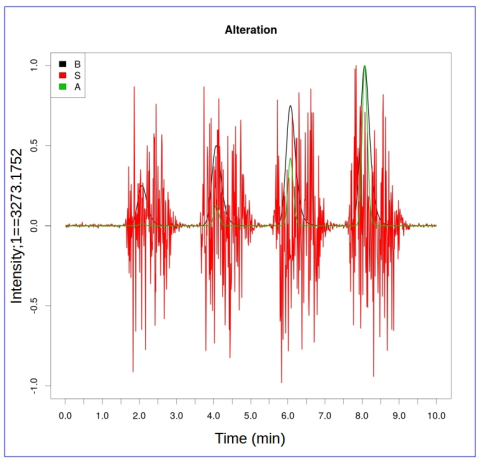

Fig. 5.46.: Alteration maps with four EMG changes with different rates.

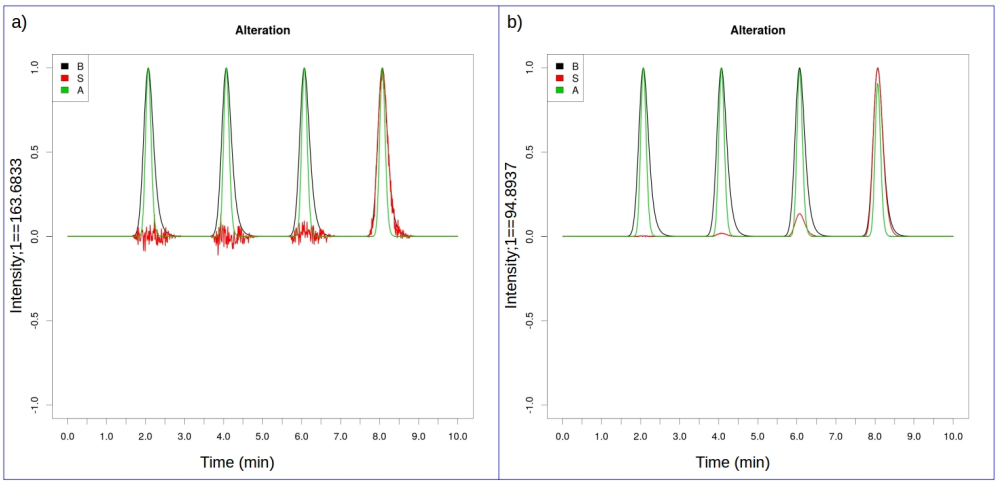

Fig. 5.47.: Alteration maps with four EMG changes with different retention time parameters.

A very interesting case is when waveforms are the functions behind the perturbation. The example has four changes with sine and cosine run and both are added and subtracted from the initial peak (Table A.13). Correlation maps (Fig. 5.25) show that there is two different changes through the absolute out of phase behaviour of the two types with no synchronous cross peaks. ALA (Fig. 5.48) goes an extra mile again, because if we know that waves are involved in the experiment the two species can be identified: sine changes have opposite SAM and AAM peaks while cosines have peaks with identical directions.

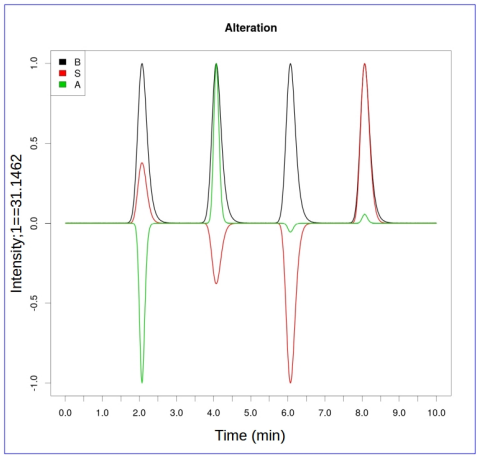

Fig. 5.48.: Alteration maps with four EMG changes with different retention time parameters.

5.2.4 Changes in shape

Section 5.1.4 dealt with how informative correlation maps are in peak shape changes. There we discussed the difficulty with these cases, the complex patterns we have to evaluate (Figs. 5.27, 5.29), but if we are familiar with these patterns, 2DCOR can be a very effective tool for deciphering what happens in the chemical system. Alteration maps neither have trouble bringing the information (Fig. 5.49). They have the luxury to deliver it in 1D graphs, so it is much easier to understand them, and again the real directions can only be seen here. For example, the first peak in Fig. **??** a) has positive sides and negative middle part in SAM, that means during the perturbation the sides were increasing and the middle was decreasing, which leads to the fact that the peak was widening, the peak shape parameters were increased. The second peak has the opposite look which means narrowing. If we look at the parameters, what the graph was based on, we can see exactly that (Table A.14).

Furthermore AAM in **??** a) has interesting peaks. Only some distorted looking peaks appear, two at every change while in graph b) three unambiguous peaks developed. The parameters attached to the second graph hold non-monotonous changes contrary to the previous linear ones and this difference is what shown in the two graphs. 2DCOR also informs us about this with the small cross-peaks, but ALA is more straightforward.

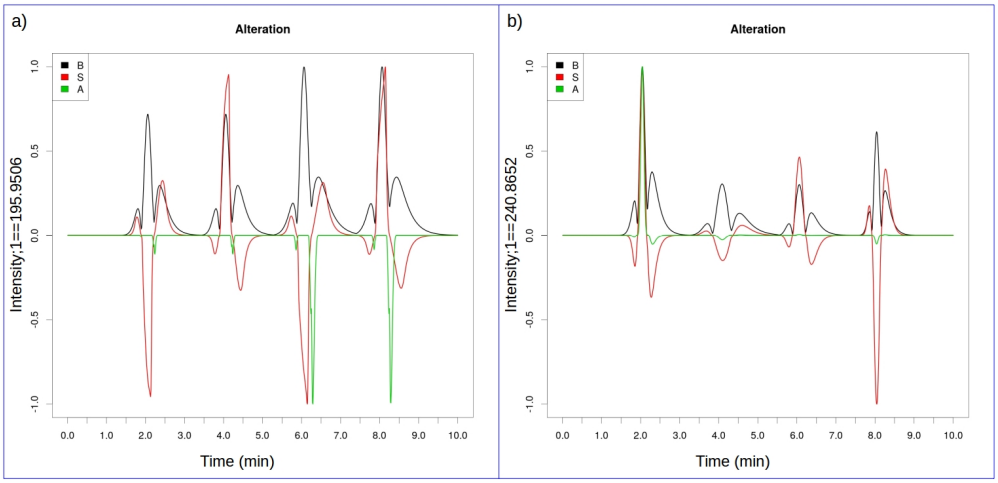

Fig. 5.49.: Alteration maps with peak shape changes.

5.2.5 Peak migration

Peak shift has a unique pattern in 2DCOR (Fig. 5.31, 5.33, 5.35), we can see from Fig. 5.50 that ALA is no exception, but the meaning behind it can be gathered without much effort. We have three examples: the first one has four linear peak shifts with the same direction (Table A.16), in the second, two of them are reversed (Table A.17), the third one has non-monotonous peak shifts (Table A.18).

The patterns vary with the distance taken by the peak, just as in 2DCOR. If this distance is small we can see two peaks in BAM and their synchronous pair. From the direction of synchronous peaks, the direction of migration can be told: from negative to positive. They do not have asynchronous value, only a tiny peak in the middle. On the other hand, if long distance is taken by the peak the BAM's peak become a plateau. Synchronous peaks still fulfil their purpose, but here asynchronous values arise too. We can differentiate monotonous and non-monotonous shifts, one has symmetrical patterns (a), b)), the other does not (c)). As we can see in this case, 2DCOR and ALA are very similar to each other.

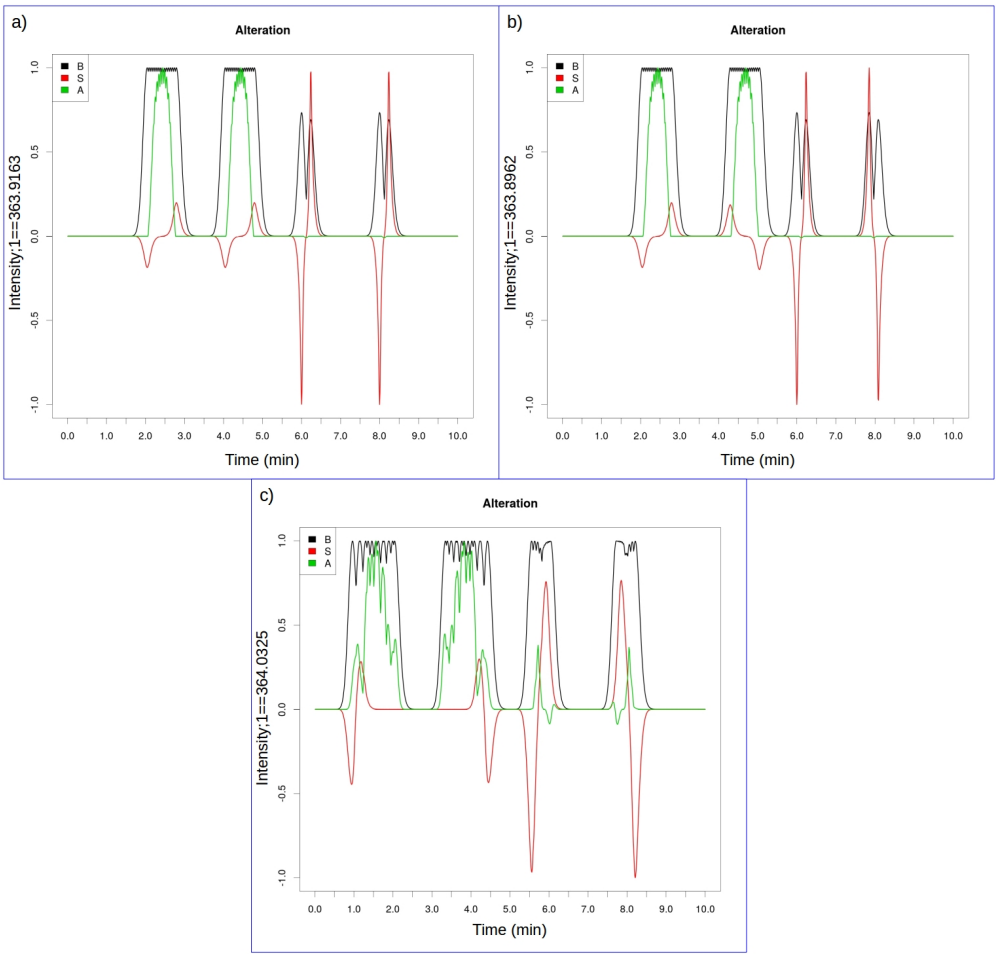

Fig. 5.50.: Alteration maps with shifting peaks.

5.2.6 Overlapping peaks

2DCOR proved to be successful in detecting overlapping peaks, described in Section 5.1.6. ALA is no exception. Fig. 5.51 a) and b) shows the alteration maps for the experiments of correlation maps in Fig. 5.37 and Fig. 5.39. The parameters are presented in Table A.19 and Table A.20.

In the first example (a)) the two overlapping peaks can be easily determined in the groups, where the changes are very dissimilar, different directions in the second group and different types in the fifth. Unfortunately, the other groups seem like just one combined peak, the separation is not prominent. 2DCOR neither could tell the overlap in the first and third groups, but through asynchronous cross peaks it could show that in the fourth. It is because of the separable sequential behavior of linear and quadratic changes. In ALA these are more similar. They can be identified because with quadratic change the ratio gap between BAM and SAM is bigger compared to linear change, but in this case it is not helping with separate the overlapping peaks.

In the second example (b)) the situation is identical to what we could see in correlation maps, every last detail is visible on the maps, even the small difference in changes is showcased (first group). So we can conclude that both methods can be highly efficient to separate overlapping changes, only 2DCOR can be slightly more susceptible in extreme cases.

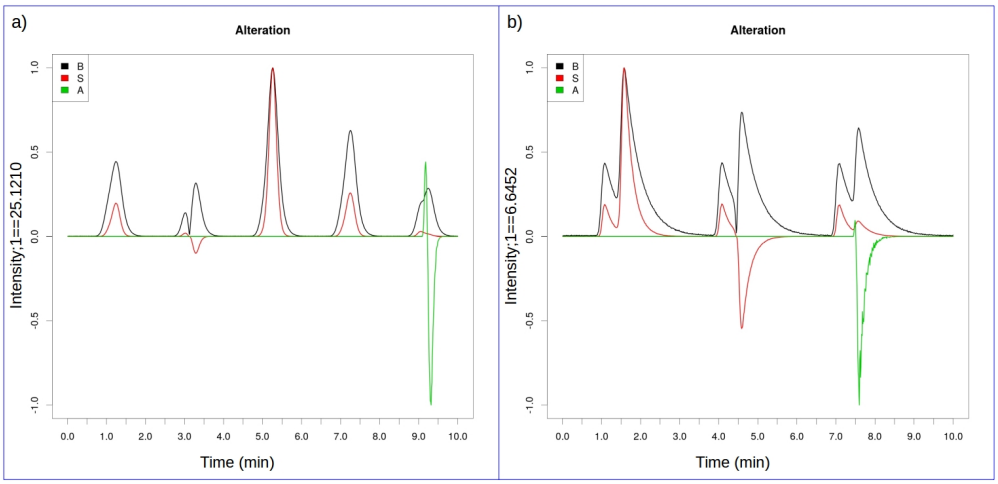

Fig. 5.51.: Alteration maps with a) five (a total of ten) seemingly inseparable overlapping and b) three (a total of six) highly overlapping peaks.

5.3 ALA in 3D environments

The previous section dealt with comparing the two methods, however 2DCOR's reach ends with 1D measurements, only ALA is capable to adapt to 3D datasets built from a series of 2D chromatograms. The simple graphs for alteration maps become more complex. They will be 3D plots visualized from an above perspective as contour-plots just like correlation maps. So alteration maps in 3D will have a resemblance with 2D correlation maps. There are major differences as well: There will be three maps, basic, synchronous and asynchronous, just as discussed in 2D, but here they have to be plotted separately. The most important one, however, is that they are not symmetric in any way (only in extreme cases, but that would come from the arrangement of 2D chromatograms and not from the properties of the map), the diagonal has no large part in their evaluation. Because unlike the correlation maps what we see is not an abstract second dimension but a chromatographic plane where every point is of interest.

5.3.1 Properties of alteration maps

Chapter 3 states that the formulas of 2D and 3D ALA are the same, only in 3D we examine every dot in a plane rather than a line, the points of interest leap into higher dimension, but the calculations in the perturbation's dimension remain the same. So, we expect no less from the maps, but to have the same properties like their previous counterparts.

The next experiment is designed to summarize the working of alteration maps, it contains a few representative functions of perturbation we mentioned before (Table A.21). Fig. 5.52 shows four plots, the first one (a)) represents the starting point of the perturbation, the first chromatogram in the series. The second is the BAM (b)) which as we know signals every change and their magnitudes. Comparing the two informs us that only those peaks appear in the alteration map where some change occurs, the rest is cut out, static elements of the system, which are permanent throughout the perturbation, do not disturb the evaluation. Thus the first major advantage of 2DCOR and ALA is secured.

Without further details we see peaks at different locations in SAM (c)) and AAM (d)), so clearly some form of separation is apparent. The two maps have

only four peaks while BAM has eight. This mystery is solved when the maps are magnified, as it is shown in Fig. 5.53. Now we can see all the synchronous or asynchronous peaks associated with the basic peaks. This gives us the difficulty with 3D ALA, it is hard to plot them efficiently with every detail visible, while in linear graphs one can usually see wider scale of peaks. However it is not really the method's fault, even today with recent technology, 3D plotting is still not easy in either field. 2DCOR has to deal with this problem too, in fewer dimensions. Maybe virtual or augmented reality will enhance this aspect in the near future.

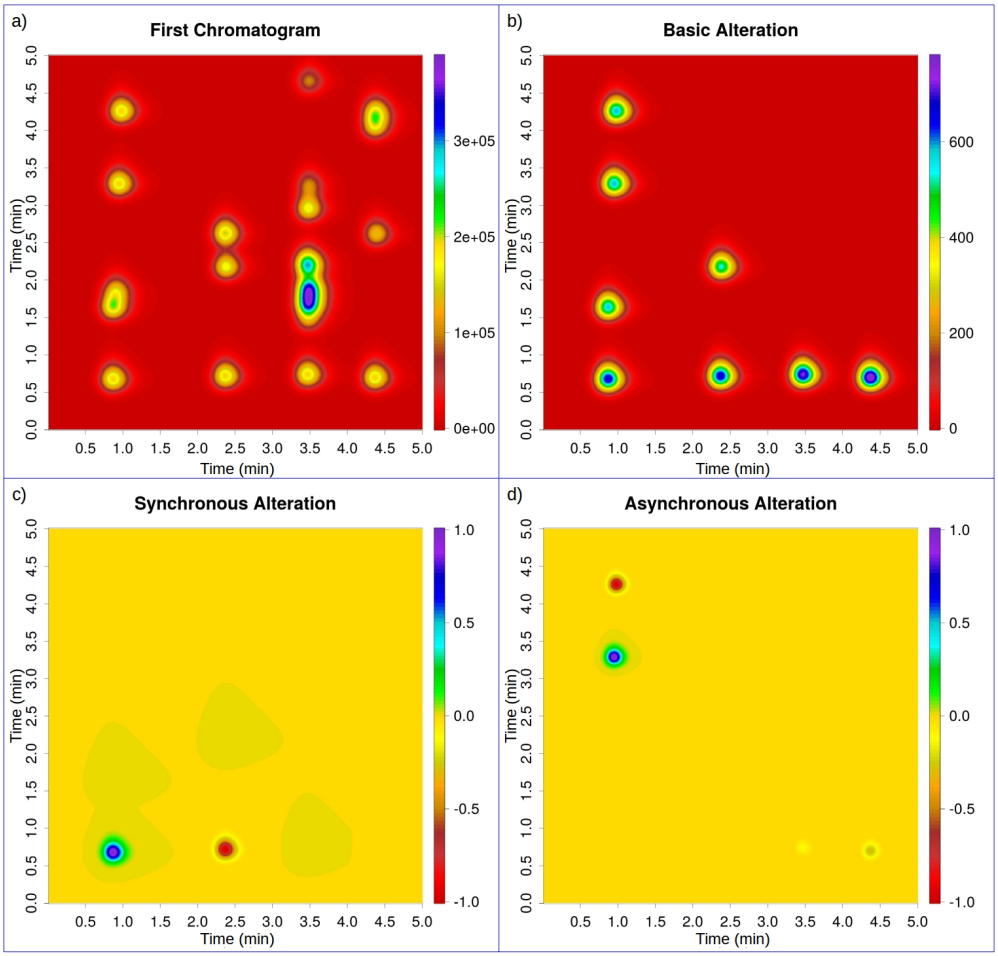

Fig. 5.52.: Example for exhibit the properties of alteration maps. The a) first chromatogram of the series, b) basic, c) synchronous and d) asynchronous alteration maps.

If we have a closer look on the alteration maps, we can see the regular tendencies we discussed in 2D experiments. Namely the monotonous changes (#1-4) only have synchronous peaks, single changes (#5,6) only have asynchronous and wavelike changes (#7,8) have both. The monotonous changes can be separated too, he ratios of basic and synchronous peaks reveal the linear (#1,2),

quadratic (#3) and exponential changes (#4). We already isolated single changes from wavelike ones, but also sine (#7) has opposite facing peaks while cosine has two positive peaks. The direction of certain changes can also be pointed out, #2 and #6 is negative while #1,3-5 is positive.

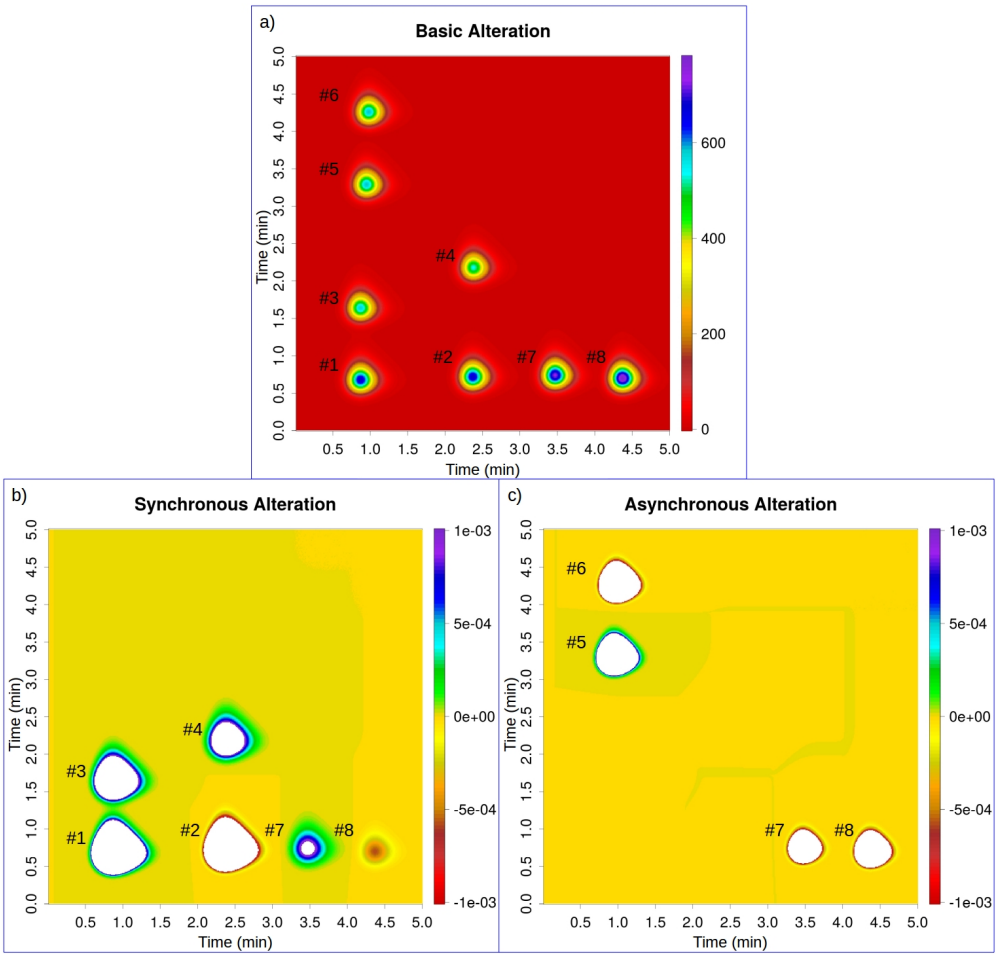

Fig. 5.53.: More detailed version of the maps from Fig. 5.52. a) basic, b) synchronous and c) asynchronous alteration maps.

The changes in peak height are demonstrated above, but what about other changes that affect peak shape and location? An instructional experiment is made to clue us in (Table A.22). Fig. 5.54 shows a few examples of what we can see on alteration maps with these changes. There are twelve patterns associated with them, the first seven (#1-7) belong to shape changes. A significant update from 2D is that now there are two ways where the peak can change its shape. In the lower line (#1-4) only one dimension is affected. #1 and #3 is widened, #2 and #4 got thinner, indicated by SAM (b)) with negative center and positive sides and reversed respectively. The difference between pairs is that the change happens either in the first or second dimension.

The middle line (#5-7) is where both dimensions have shape distortion. #5's width is getting larger with the same magnitude in both directions. #6 and #7 is widening in both ways too, but significantly more in one dimension.

The AAM seems to have no peaks for peak shape changes, that is not the case, magnifying the map will reveal them, but the peaks of shift are much higher than them, so they disappear in the base colour. However for more complex examples they are absolutely essential, but for this case they are not really necessary.

Finally, the upper line (#8-12) represents peak shifts. There are examples for two dimensions (#8,9 and #10,11) and two directions (#8,10 positive and #9,11 negative). The last one (#12) shows what happens when the movement is not aligned with any dimension, from the parameters perspective: the t_R of both dimensions are changing.

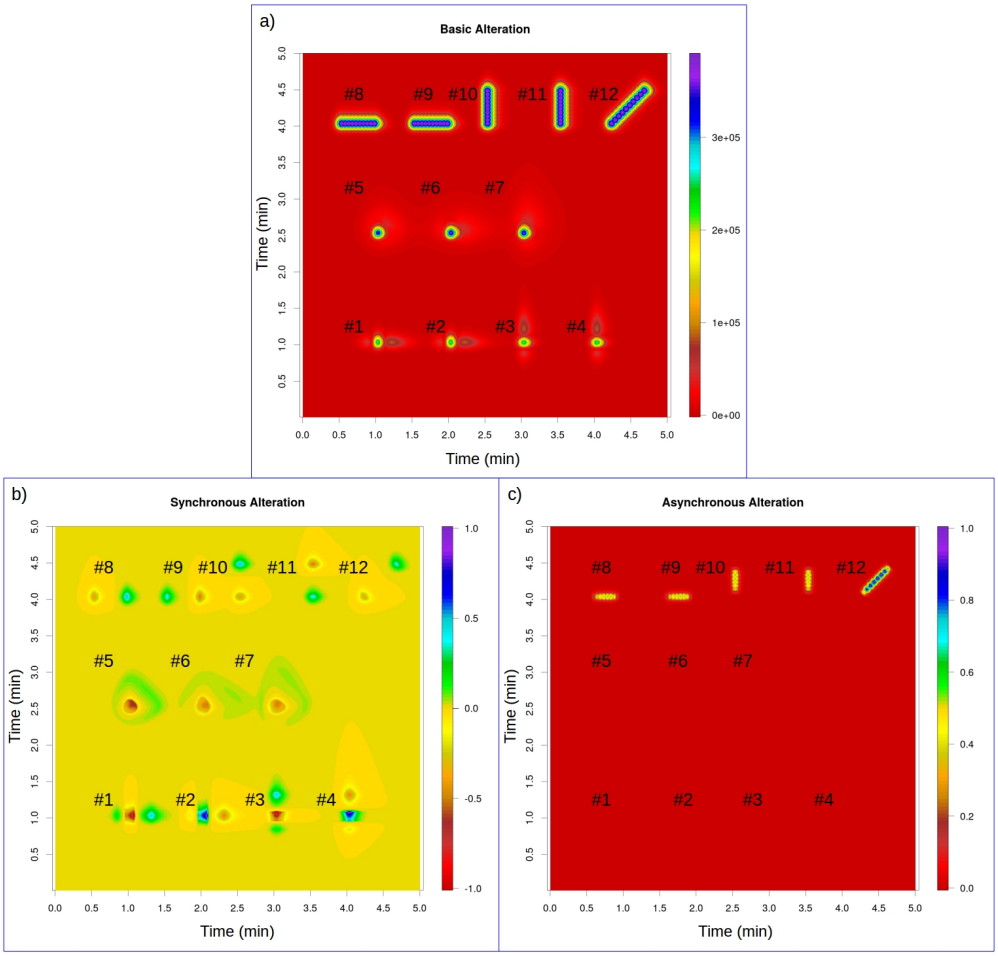

Fig. 5.54.: Examples for shape change and peak shift in 3D. a) basic, b) synchronous and c) asynchronous alteration maps.

2DCOR and ALA are proved to be effective in detecting overlapping peaks (Sections 5.1.6, 5.2.6). Now we will see how 3D ALA is up to the task. The overlapping peaks, we will see, are identical to the ones studied before, only this time they were fitted in one 2D chromatogram due to the expanded space. The first chromatogram of the series (Fig. 5.55) shows how they look on the chromatographic plane. The peaks were changed in only one dimension the second was untouched, the exact parameters can be found in Table A.25. There are eight peak groups with two-two overlapping peaks, five seemingly inseparable (#1-5) and three recognizable but still highly overlapped (#6-8).

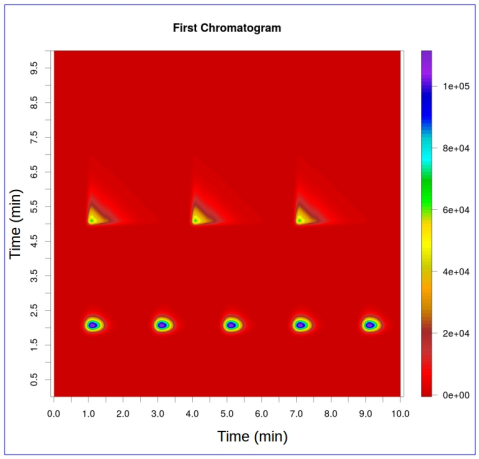

Fig. 5.55.: Examples for overlapping peaks in 3D. The first chromatogram of the series.

The alteration maps with default options are shown in Fig. 5.56. And clearly scaling issues are present here as well, only a few peaks can be seen. We chose the magnification approach against scaling ones due to its simplicity. We adjusted the options of z axis to show the smaller peaks, and this time every change is featured in the maps (Fig. 5.57). The patterns are reminiscent of 2D ALA maps, in fact they are basically identical, only expanded with another dimension. To start with disadvantages #1,3 can still not be separated, but other than that only advantages left. SAM is enough to separate #2,7, the changes with opposite directions. The two peaks in groups #5,8 divided between SAM and AAM on the basis of their monotonicity, linear showed in the first, sine in the latter. #6 has the same setup as #1, but because of the bigger difference of retention time, the two peaks are appeared separately. The greatest mystery is #4 where there is two different types of monotonous changes, but only one peak appears, however it is a strangely shaped one. We can not tell unequivocally its overlapped background, but due to the nature of 2D separation and 3D visualization the shape distortions of two co-eluting peaks can be seen more

clearly, it is also especially conspicuous on #6-8. So, in this case we have a higher chance to spot this anomaly on #4 too.

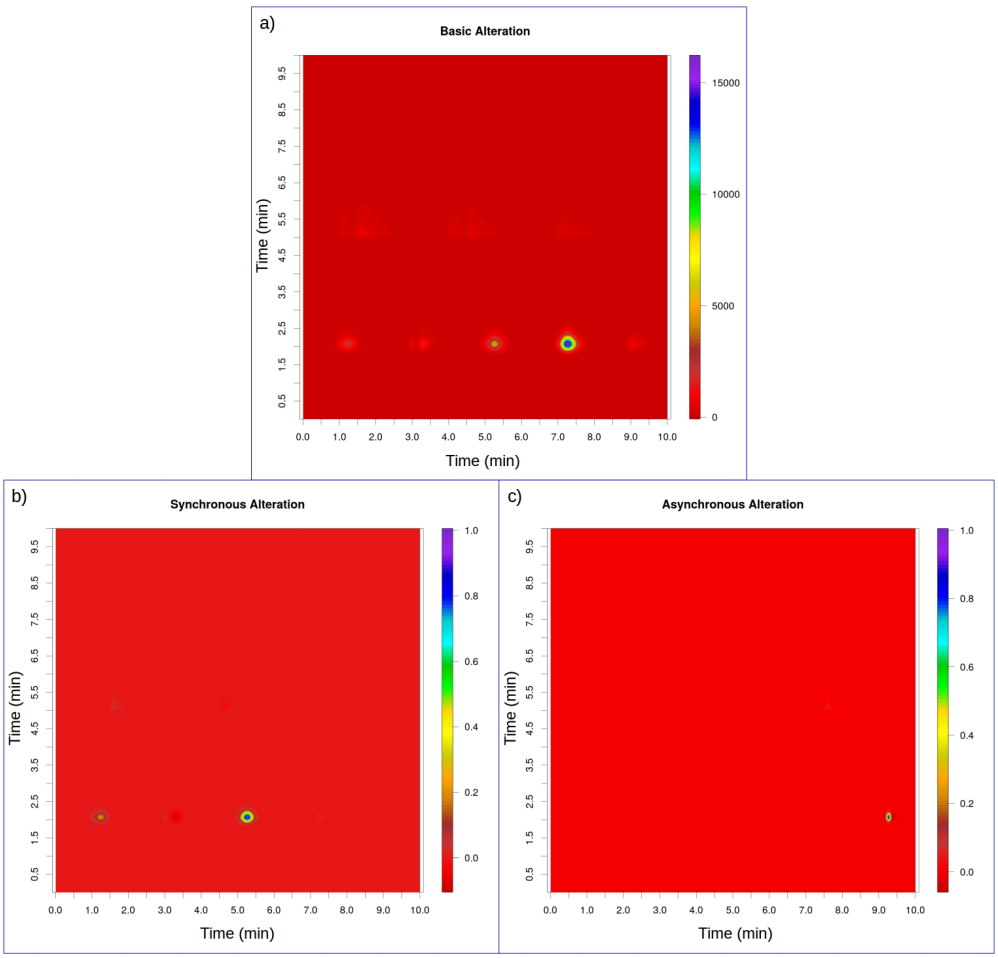

Fig. 5.56.: Examples for overlapping peaks in 3D. a) basic, b) synchronous and c) asynchronous alteration maps.

In conclusion we can rightfully say that ALA in 3D delivers the same information as in 2D. The singular differences come from the difficulties of higher dimensions. On one hand, visualizing the maps becomes significantly more troublesome, the sophisticated details are harder to detect with the naked eye and the maps have to be plotted separately, making it harder to compare their values. On the other hand, changes (height not included) now have another dimension to be realized which makes more combinations possible leading to more complex patterns. Except all that the desired information still can be gathered within reasonable conditions, because the representation of perturbation's dimension, the changes in the system are reminiscent to 2D ALA.

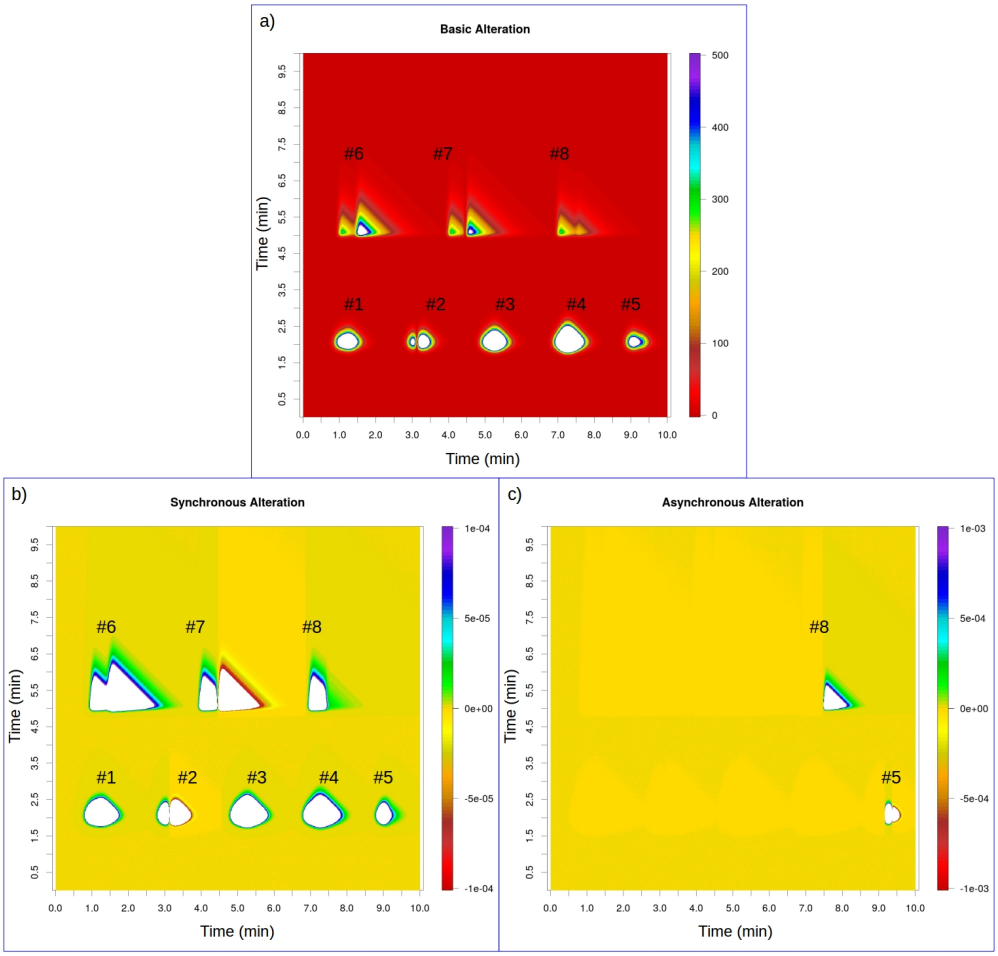

Fig. 5.57.: Examples for overlapping peaks in 3D. Magnified correlation maps: a) basic, b) synchronous and c) asynchronous alteration maps.

5.3.2 Practical examples

In this section we will show two cases where alteration analysis can become handy. These are simple examples, because the exploitation of all the properties of ALA in action will not fit in the boundaries of the present study. These are just two of the many practical approaches with more realistic chromatograms than the examples before, because in measured systems it is more likely to have only one or two kinds of changes.

If one wants to evaluate a simple chromatogram to unveil the usual properties of the peaks in it, one will not need alteration analysis, there are already sophisticated methods included in almost every chromatographic software to do that. Alteration analysis becomes useful when there is a series of measurements and one needs to know the changes between them. Also, if there are only two or three chromatograms, their analysis can be done easily visually,

by placing them next to each other. The difficulty arises when there are several or even more chromatograms, or there are only minor changes between chromatograms. The comparison still can be done the same way visually with the default software, but it will be hard and time-consuming. In contrary, alteration analysis can be done in a few minutes with just a few simple and straightforward graphs.

One of the cases where alteration analysis can be very attractive is when the chromatograms are crowded (Fig. 5.58 a)) and there are only a few small changes between chromatograms (Table A.23). It is almost impossible to notice the differences visually in the plotted chromatograms. The alteration maps, however (Fig. 5.58 b), c), d)) provide a simple solution. In the basic map (Fig. 5.58 b)) we can see the changes, the few spots where the chromatograms differ. It is also shown that there are two types of changes. Three of the four changes are synchronous (Fig. 5.58 c)), probably all of them linear, because the ratios on the basic and asynchronous maps are the same. But there is a smaller asynchronous change (Fig. 5.58 d)). It has a peak only on the asynchronous map, so we can assume that it is a single difference. And that is one of the reasons to include the correlation coefficient map (Fig. 5.58 e)) into the method, because it shows the exact sample where this single change has occurred. It can be seen by the break of the pattern at sample #6 of 8.

In the second example (Fig. 5.59, Table A.24) all the peaks of the chromatograms change synchronously with the perturbation. We can say so, because the asynchronous map (Fig. 5.59 d)) has no peaks, it is close to zero almost everywhere, only the noise can be seen on it. This feature can be noticed from the plotted chromatograms, but one little detail may stay hidden. Four of the peaks change slightly differently from the others. Their ratios are kept in the basic and synchronous maps (Fig. 5.59 b), c)), this suggests that they are somehow connected, at least more than the other peaks. The correlation coefficient map (Fig. 5.59 e)) verifies linear synchronicity with its symmetric pattern.

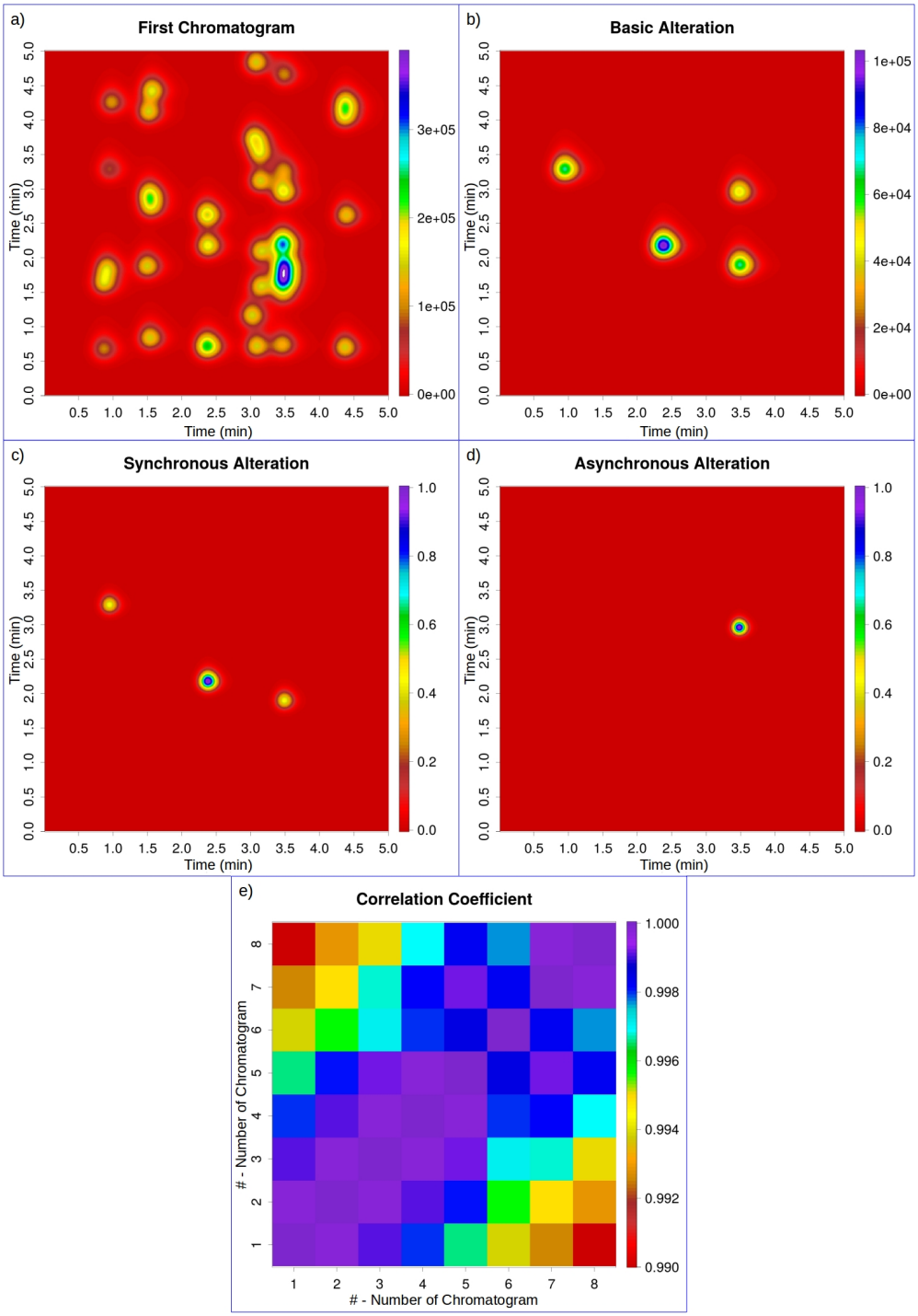

Fig. 5.58.: Practical example of the usefulness of ALA with only a few small changes. The a) first chromatogram of the series, b) basic, c) synchronous, d) asynchronous alteration and e) correlation coefficient maps.

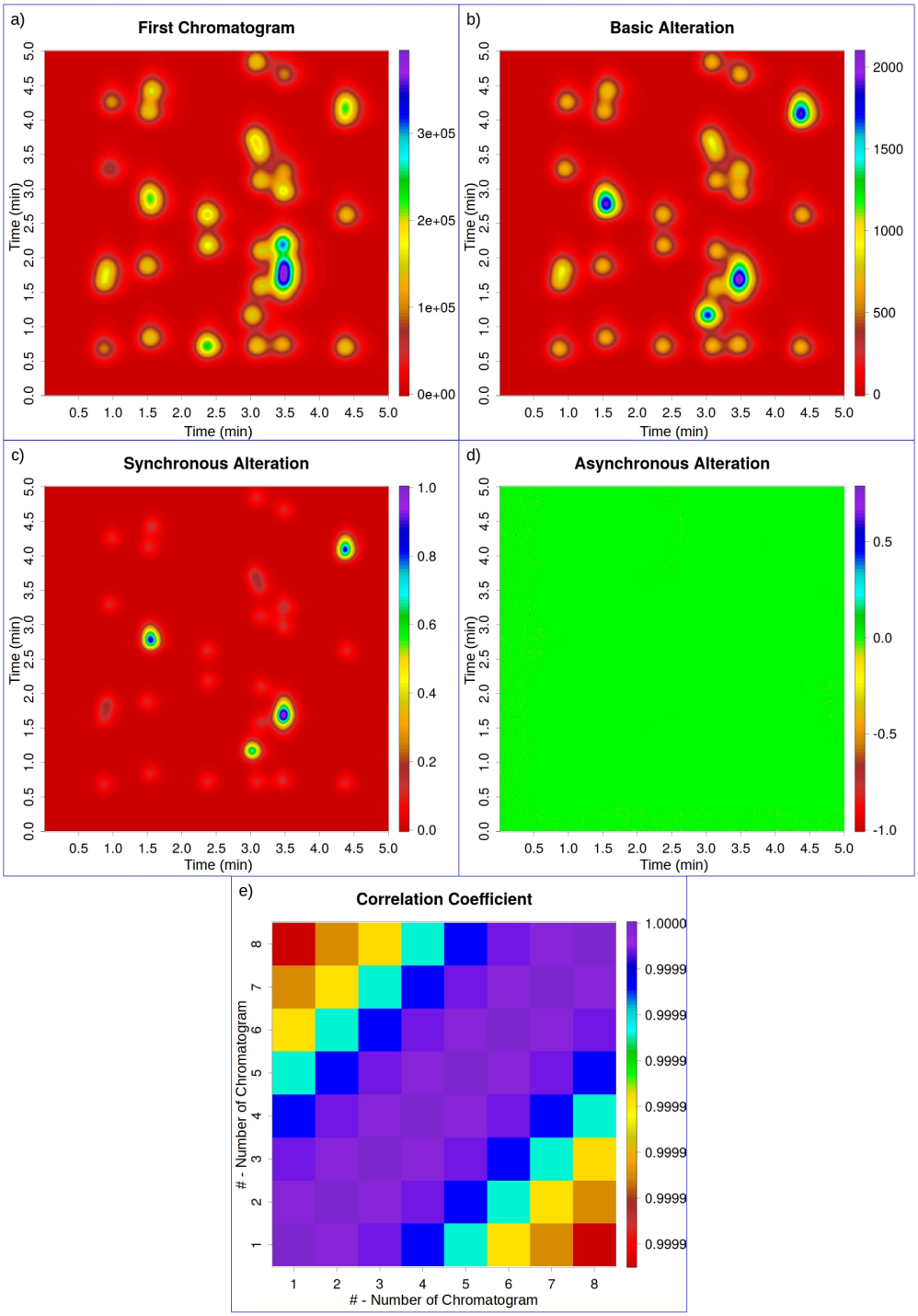

Fig. 5.59.: Practical example of the usefulness of ALA with many but two different changes. The a) first chromatogram of the series, b) basic, c) synchronous, d) asynchronous alteration and e) correlation coefficient maps.

5.4 Experiments on measured data

After establishing fundamental properties via computer generated data 2DCOR and ALA has to be tested on measured chromatograms. First we provide an example to how 2DCOR can be useful in chromatography. This technique is compared with PCA in a reproducibility study. After that ALA gets into the highlight. Its workings in practice will be demonstrated by two experiments, in the first one peak height changes will be the focus, in the second one methanol's role in detection will be studied.

5.4.1 Reproducibility of chromatographic columns

After the establishment of the method by a number of computer experiments, the first study on measured data was the comparison of the retention properties of the components of a test mixture (Fig. 5.60) on five HPLC columns. The same sample was injected five times on each column. This resulted in 24 chromatograms (on the first column only 4 injections were made) with approximately 27 minutes of elution time, digitalized with more than sixteen-thousand data points respectively. Without the proper calculation technique, the evaluation of this data matrix can be quite challenging. However the 2D-correlation can fairly simplify that procedure, providing valuable information within a short time.

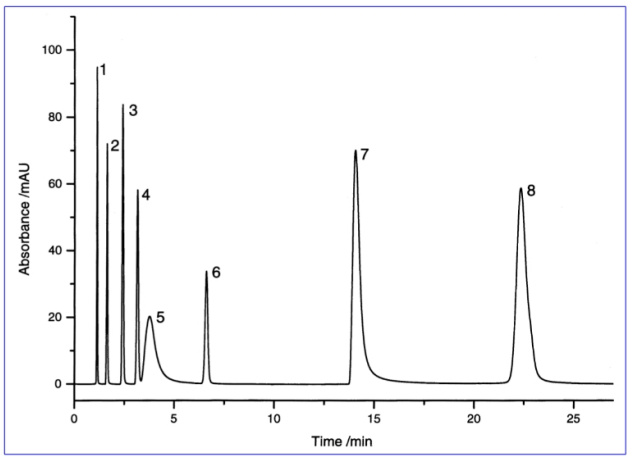

Fig. 5.60.: The test mixture of eight compounds used in the reproducibility study. The experimental conditions are detailed in Section 4.5.1 and in its original source: Kele et al. [56]

Comparing the five-five chromatograms measured on the same columns, the conclusion is that, as expected, there is no significant difference between them. Surely there are minor fluctuations in the retention times of the peaks, but the values of the asynchronous maps are two orders of magnitude lower than the synchronous maps, indicating the presence of only negligible differences. There are also no noticeable distinctions on the top-views, same as the sample–sample correlation coefficient maps.

The results are represented with the correlation maps of the data obtained on column #3 (Fig. 5.61). If we examine the synchronous and asynchronous correlation maps in this example, the patterns only show the consequences of retention time changes. However the top-view of the chromatograms show a straight line for each peak in Fig. 5.61 c), meaning the retention time changes are practically negligible in that instance. In the fourth map (sample–sample correlation coefficient map) there is no correlation coefficient lower than 0.998, verifying that only minor differences can be observed between the chromatograms.

Studying all the 24 chromatograms together, similar results can be concluded, although this time the differences are more distinct (Fig. 5.62). As the top-view shows, the variation of the retention times comes from the dissimilarities of the HPLC columns, because the chromatograms are settled in five groups, each representing an HPLC column. The sample–sample correlation coefficient map also indicates high correlations inside the groups (with values no less than 0.98), differences only appear between the groups. The lowest correlation coefficients (0.57–0.63) are found between columns #2 and #4. For all the other pairs of HPLC columns, the values are always greater than 0.8. That explains the chessboard-like pattern of the map, because the 5×5 surfaces are obtained for almost identical chromatograms, measured on the same column.

Furthermore, the conclusion for the correlation between the HPLC columns is that column #1 and #3 are the most similar, whereas the chromatograms measured on columns #2 and #4 are the most different. On the whole, however, there is a good correlation for the entire data matrix, so the reproducibility is far above plausible.

In this case, the perfect reproducibility of the chromatograms on the same columns are also clearly shown, therefore there is no need to make the compari-

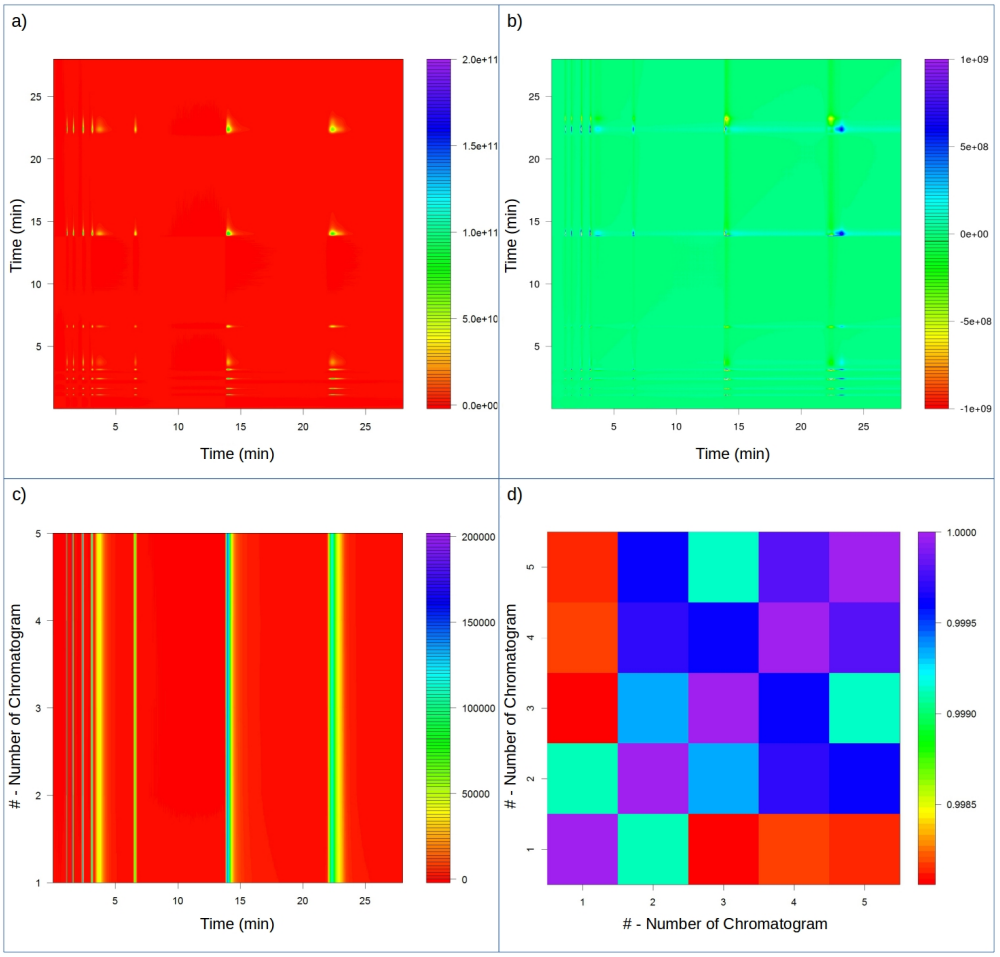

Fig. 5.61.: a) synchronous and b) asynchronous correlation maps, c) top-view of the chromatograms measured on column #3 and d) the sample–sample correlation coefficient map.

son on every column respectively. These four graphs contain all the information we need.

At this point we know that the 2D-correlation method provides information about the data in a logical manner, but we have not yet verified its results. To achieve that, we compare the 2D correlation to the previously published results of PCA calculation on the same data set [55].

The score plot of the PCA (Fig. 5.63) on the retention times reveals that HPLC columns #2 and #4 differ the most in their properties, and there is almost no difference between the retentions measured on columns #1 and #3. This confirms the results of sample–sample correlation. Hence, one can conclude that the two methods – principal component analysis and 2D correlation – provide identical conclusions.

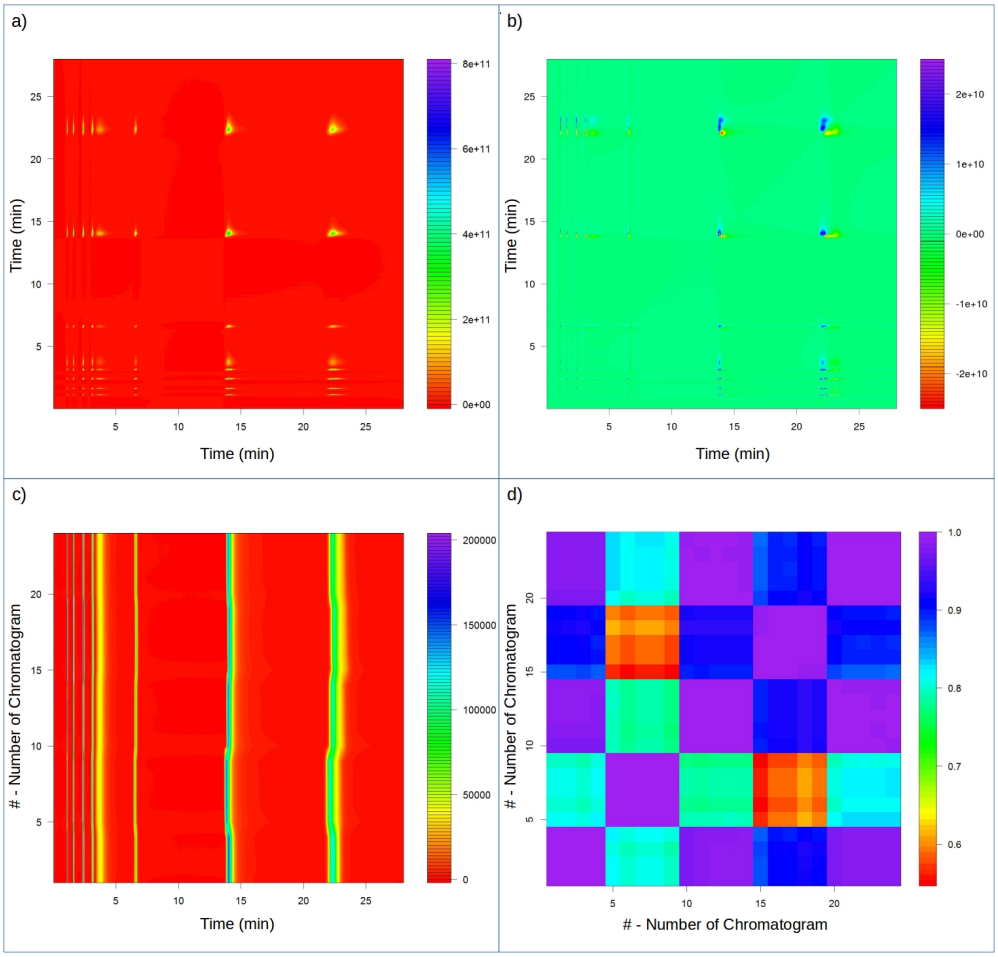

Fig. 5.62.: a) synchronous and b) asynchronous correlation maps, c) top-view of all the chromatograms measured on every column and d) the sample–sample correlation coefficient map.

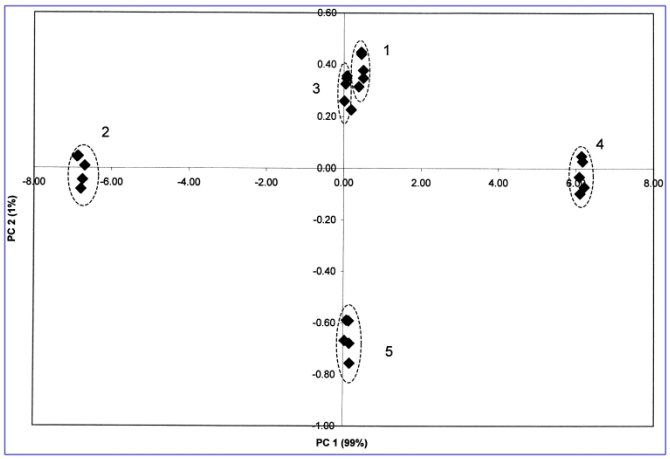

Fig. 5.63.: The score plot of PCA performed on retention times of five HPLC columns [55].

Moreover, 2D correlation gives not only the quantity of agreement or disagreement between samples, but also an insight of the quality of difference via the

synchronous and asynchronous correlation maps. With PCA, first the retention times had to be determined, then the calculation was performed only on the set of retention time values. With 2D correlation, there is no need for this step; the raw data, the entire digitized chromatogram can be directly used in the evaluation. Therefore, with the calculation of 2D correlation maps, we can also identify subtle peak shape changes even if the retention times remain unchanged.

5.4.2 Isotherm reproducibility

The 2D correlation calculation was also applied to overloaded band profiles to study the reproducibility of HPLC columns under nonlinear conditions. Seven analytes were separately injected on ten HPLC columns. Then isotherm parameters were determined for each analyte on each column, and the results were subjected to 2D correlation calculations.

The synchronous, asynchronous, and sample–sample correlation coefficient maps, as well as the top-view of the chromatograms for aniline are presented in Fig. 5.64. In this case, the synchronous and asynchronous correlation maps are rather simple, because the chromatograms contain only one peak. Nevertheless, they clarify that the differences are in the retention times. The top-view of the chromatograms indicates the same. A slight retention time change can be seen on columns #1 to #4, and between #1 and #2 in particular. For columns #5 to #9 a similar conclusion can be drawn, then column #10 shows a radical difference. The sample–sample correlation coefficient map shows a similar pattern. Columns #2 to #9 have great correlation coefficients among them (>0.91). Values for column #1 are smaller, but still above 0.82. Column #10 diverges more from the group with correlation coefficients between 0.56 and 0.72.

The 2D correlation studies of the other samples have yielded quite similar results. First of all the conclusion is that columns #2 to #9 are rather similar, they show a very good correlation, although for #4 or #5 value is a bit lower. It is, however, obvious that columns #1 and #10 differ from the rest of the group and from each other the most.

Once again, the usefulness of the 2D correlation analysis is confirmed by the comparison with the results of PCA obtained on the same data set. It can-

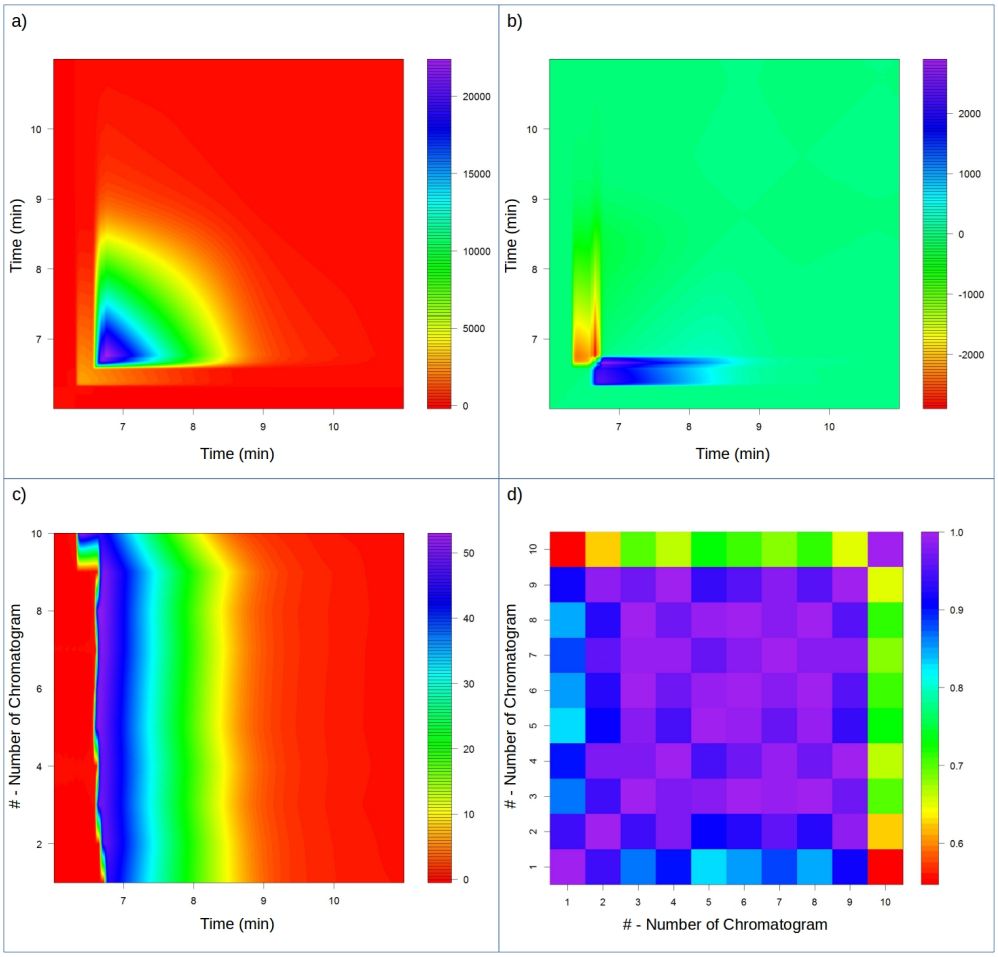

Fig. 5.64.: a) synchronous and b) asynchronous correlation maps, c) top-view for the chromatograms of aniline and d) the sample–sample correlation coefficient map.

not be emphasized enough, that for the PCA calculations, the determination of the isotherms and isotherm parameters is crucial, which is a massive additional step in the calculations with its possible errors. On the other hand, 2D-correlation utilizes the raw chromatograms; therefore the entire band profile is involved in the calculation. In this study pretreatment, such as baseline correction or smoothing was not necessary.

PCA was applied to the isotherm parameters of each sample separately, then the results were summarized for the HPLC columns. The correlation calculation is demonstrated here for aniline, and the results obtained with PCA for the same sample are found in Fig. 5.65. That figure reveals the score plots of the PCA performed on the Jovanović and Langmuir isotherm parameters determined for aniline on each HPLC column. The graphs show that the properties of columns #2 to #9 are rather similar. The high-contrast differences are only in

columns #1 and #10. Thus recalling the results of 2D correlation plotted in Fig. 5.64, we can conclude that the validity of the correlation method has been proven once again.

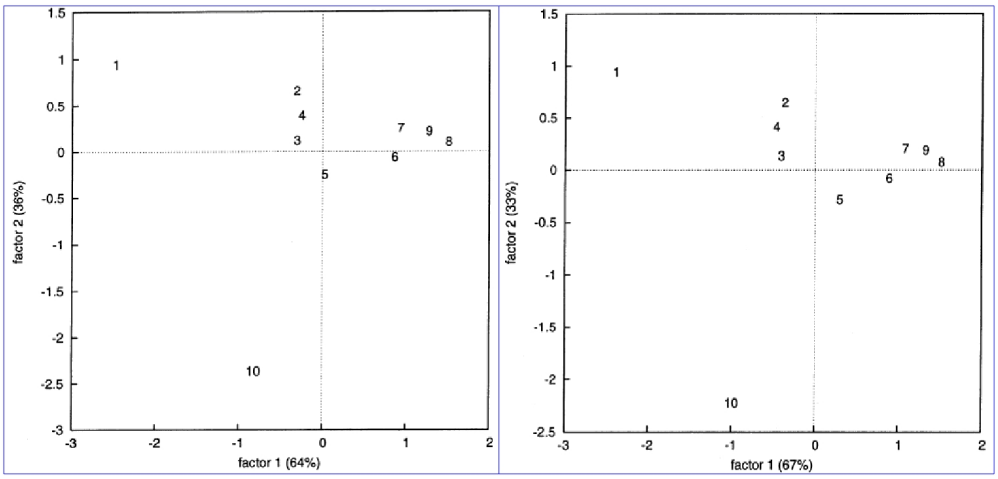

Fig. 5.65.: The score plots of PCA performed on the Jovanović and Langmuir isotherm parameters of aniline. [64].

There is, however, a minor difference between the results of the two methods. In the score plots (Fig. 15 of Ref. [64]), column #5 is located between two groups, indicating lesser similarity, while in the sample–sample correlation coefficient map (Fig. 5.64 d)) column #4 has the lower correlation value. Examining all the results it turns out that almost every column has a stable place on the score plots, except for #4 or #5. Their behavior depends on the employed model of isotherms, but this alteration can not be compared with the other distances on the plot, same holds for the values of the sample–sample correlation map. That is why this difference can be considered irrelevant, and does not change the fact that the correlation method is efficiently comparable to PCA.

Felinger et al. compared the reproducibility of the HPLC columns under overloaded conditions by summing up the results of the PCAs obtained for each analyte, respectively [64]. Up to this point we followed the same approach in 2D correlation too, but with a simple manipulation of the data, the comparison of the HPLC columns can be evaluated directly including the chromatograms of all analytes. To do so, the chromatograms of the various analytes measured on the same columns are combined together (Fig. 5.66). Analyzing those appended chromatograms, information can be obtained directly for the HPLC columns.

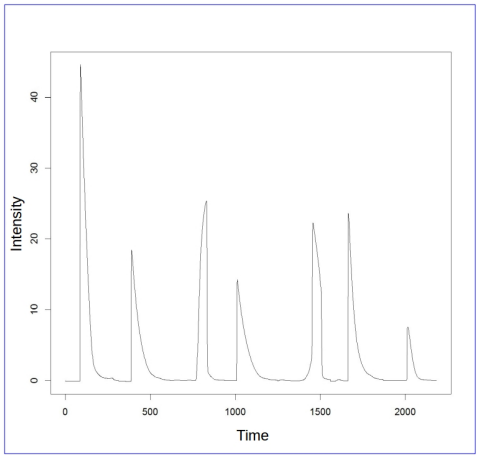

Fig. 5.66.: Combined chromatograms measured on column #1.

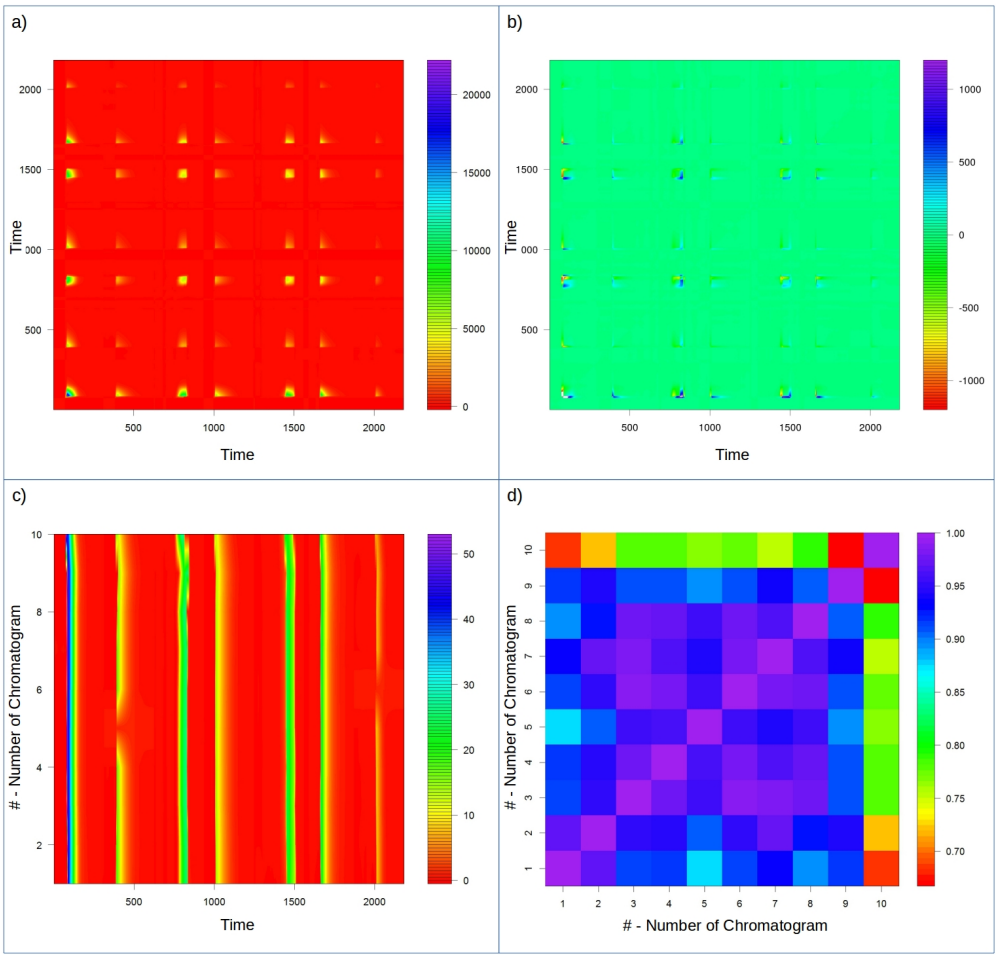

Fig. 5.67.: a) synchronous and b) asynchronous correlation maps, c) top-view of the generated chromatograms and d) the sample–sample correlation coefficient map.

The results of this approach are rather similar to those described above. The difference is that, instead of seven times four graphs (four graphs for every

sample) we can get to the conclusion with only four graphs (Fig. 5.67). This time the sample–sample correlation coefficient map shows full conformity with the PCA, because in the highly correlated group, column #5 has the lowest correlation values.

5.4.3 ALA in practice

This experiment was designed simply to put the theoretical properties of ALA, described in Section 5.2, into practice. The focus was peak height changes, later examples will delve more into other kind of changes. A series of SFC measurements were executed with seven compounds in the sample and their concentration was changed according to different functions from sample to sample (Fig. 5.68), details are in Section 4.6.

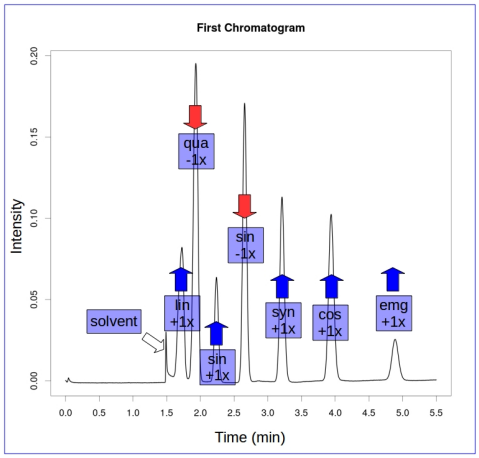

Fig. 5.68.: Visual representation of the given changes in measured chromatograms, projected on the first point in the series: two linear and single changes with opposite directions, sine and cosine and an EMG change.

Computer simulation was also made with the peak parameters in the measured chromatograms in order to see what we should get in the alteration maps. So Fig. 5.69 a) shows this generated data and b) is the measured counterpart. At a first glance the two images are different, but closer examination reveals that these differences are negligible on most peaks. The major difference is the BAM peak heights and their ratios. This issue comes from the fact that simulation does not take into account the differences of absorption at 192 nm.

After this statement we can see that the patterns of three peaks are identical, the first two and the sixth (in order of the generated maps, Fig. 5.69 a)). On the other peaks the difference comes from appearing SAM values. It is because

there is a small peak shift which is triggered by the reproducibility issues of the chromatographic system. Unfortunately, in real life chromatographic measurements we can not expect perfect alignment in all cases, retention times are naturally not as stable as the locations of spectroscopic peaks, and these relatively small migrations are affecting the alteration and correlation maps. However there are still plenty of information which is delivered by ALA.

The monotone changes (first two peaks) are easily separable, they have no visible asynchronous value, only the magnified maps (Fig. 5.70 b)) show small peaks, but they are clearly just the result of some noise. Also quadratic change (second) has different BAM and SAM rate, not as much as in the generated data, but still noticeable. Single changes (third, fourth) have the highest AAM peaks, so this feature is also reliable and just like monotonous peaks, the directions are clear. The properties of wavelike changes (fifth, sixth) are intact as well. We already mentioned the similarity of peak #6 which is the cosine change, but the sine change has the expected pattern with the added peak shift: SAM and AAM peaks are opposites, only the AAM peak is shattered into three because of the movement of the peak's tailing.

The biggest differences are ironically at the two ends of the chromatogram. One small peak emerged before the first expected one. It is from the solvent and because of that it is highly unlikely that it will disturb the evaluation. On one hand it is not overlapping any of the peaks, on the other hand usually we have some knowledge about the chromatographic system, especially when we are the ones developing the method. The other difference is the big existing SAM value, relatively to AAM, on the last peak. It is likely to be added because of the reasons like the other instances and also as mentioned EMG changes are hard to grasp on.

The corresponding correlation maps are presented in Fig. 5.71. We do not want to go into detail, but it has to be mentioned that it is a great example why the evaluation process is far more cumbersome with 2DCOR. Similar information can be gathered at the end for sure, but there are scaling issues, which are to be honest apparent in ALA as well, but can be more easily overcome, and the patterns are unnecessary overcomplicated. The connection between maps of generated and measured data is similar to ALA. They are alike overall with a few differences only caused by peak shifts.

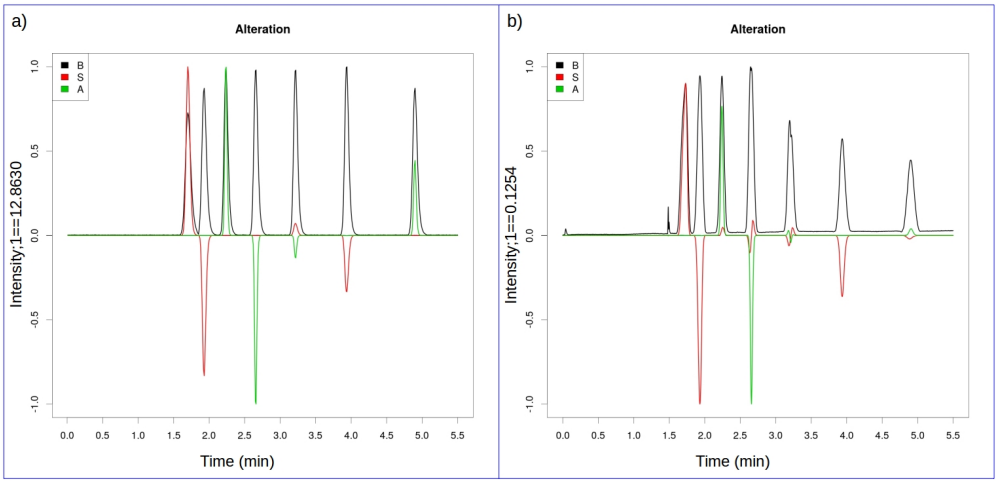

Fig. 5.69.: Alteration maps for peak height changes in measured chromatograms. a) computer simulated alternative, b) measured data.

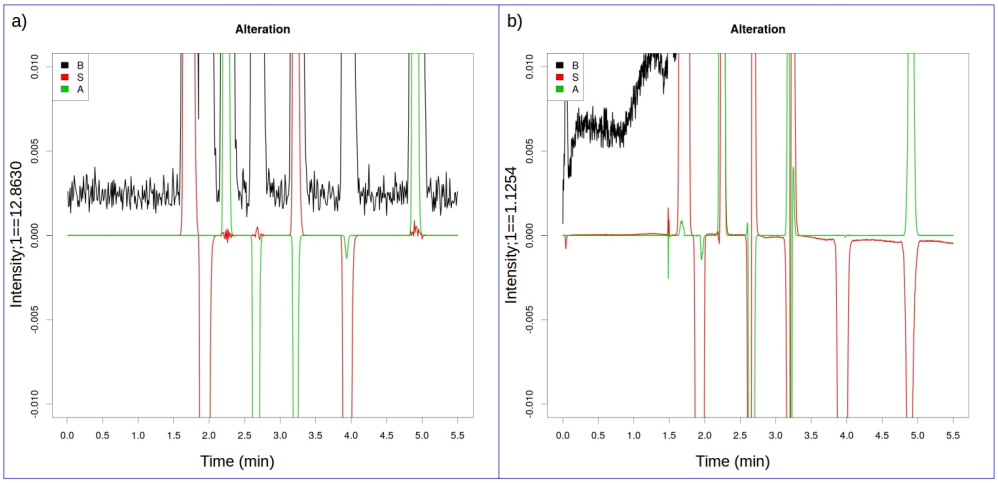

Fig. 5.70.: Magnified alteration maps for peak height changes in measured chromatograms. a) computer simulated alternative, b) measured data.

The last example comes from a series of experiments where the effect of methanol was studied in SFC (Fig. 5.72). Eleven samples were made with mixed solvent of acetonitrile and methanol by varying percentage of the latter from 0 to 100. Figure 5.73 shows the resulting alteration maps. We made two variations with detection on different wavelength, 192 nm a) and 260 nm b). The reason behind this is that methanol has little absorption on 260 nm so the compounds can be studied here separately from its sign.

First of all, there are far more signals than what would be expected. Every peak has comparable changes not just the ones which are affected directly by the peak of methanol. In theory the last peak (at about 7.5 min) can not be influenced by the perturbation, but alteration patterns can still be seen.

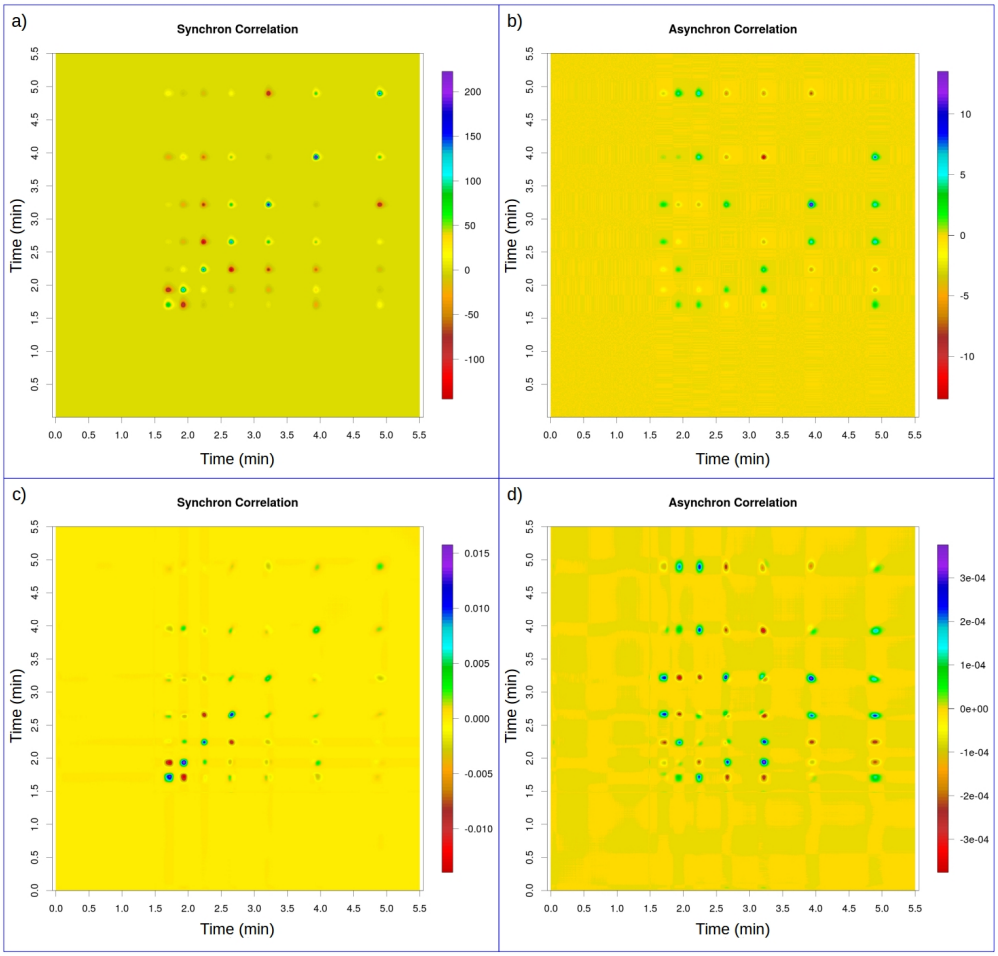

Fig. 5.71.: Correlation maps for peak height changes in measured chromatograms. a) synchronous map of computer simulated alternative, b) asynchronous map of computer simulated alternative, c) synchronous map of measured data, d) asynchronous map of measured data.

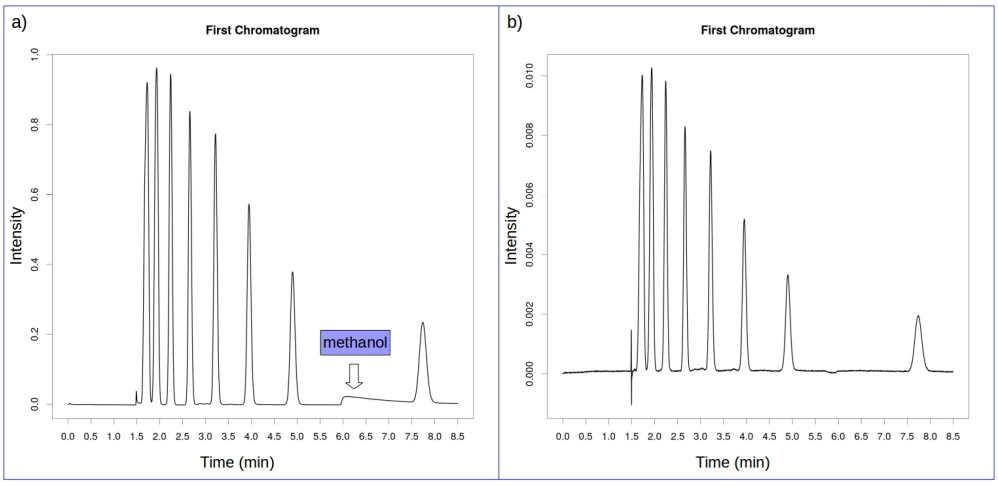

Fig. 5.72.: First chromatograms for the effects of methanol in a SFC separation. a) detection at 192 nm, b) detection at 260 nm.

We assume that it is due to the poor reproducibility of SFC system and not an intended result, because its pattern is different than any other peak's. It features high asynchronous value concentrated at the middle of the peak and the maximum is split into two in AAM and BAM. This leads us to the assumption that it is only a relatively small unintentional peak shift.

The main feature is the big basic peak of methanol on a) from about 2.5 min to 6.0 min. The direction is not clear for the first glance, because the linked synchronous peak is positive everywhere, but other peaks along its way show, with positive-negative synchronous peak combinations, that its migration started at the back of the chromatogram to earlier elution times. Also, asynchronous peaks are getting bigger in this direction in b), which is the result of methanol being absent at that point for more time in the series of chromatograms. The peak at 4.7 min is mostly synchronous because methanol is a constant factor at that retention time, but it is arriving late in the series at 2.5 min. The peak at 2.2 min is also influenced by methanol but not as strongly as the previously mentioned ones.

The first peak (at around 1.6 min) is the exact opposite of the peaks from 2.2 to 5.0 min and there are some irregular shapes before it. This means a different effect causing its change. Further investigation showed that acetonitrile was the source of it and the shapes were also belonged to it. The most interesting part is the second peak because seemingly no mentioned effect has contact with it and clearly it has a unique alteration pattern which shows peak shape changes in contrary to any other peaks' shifting tendency. At this point we are unsure of the exact explanation, we assume that the described causes and the perturbation overall have this indirect effect on it.

Figure 5.74 shows the correlation maps of the same experiment. The first impression is, especially with the maps of 192 nm (a) and b)), that they are overcrowded and with this set up it is hard to evaluate the minor details. They need scaling and/or magnifications and focusing on several parts separately to do it rightly. Thus we do not aim to further discuss them, they already made their purpose, which is to show how easier it is to evaluate ALA maps.

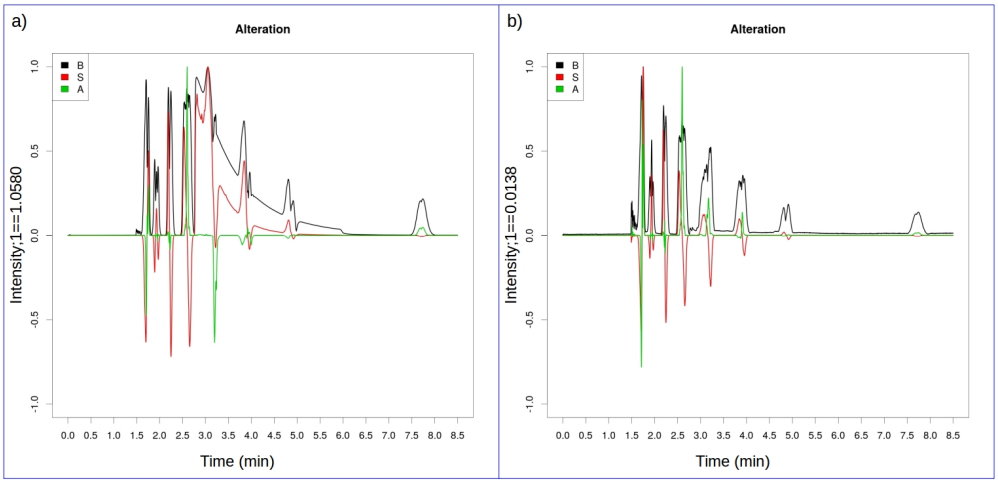

Fig. 5.73.: Alteration maps for the effects of methanol in a SFC separation. a) detection at 192 nm, b) detection at 260 nm.

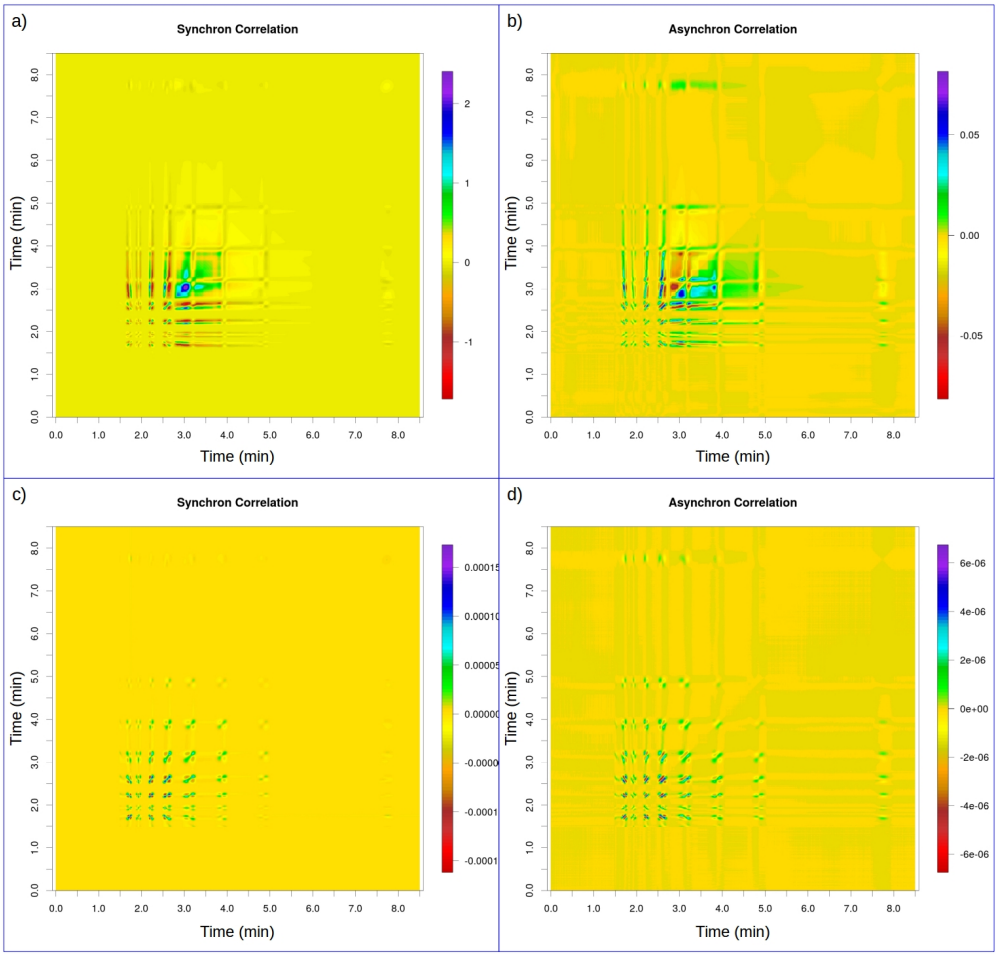

Fig. 5.74.: Correlation maps for the effects of methanol in a SFC separation. a) synchronous map of detection at 192 nm, b) asynchronous map of detection at 192 nm, c) synchronous map of detection at 260 nm, d) asynchronous map of detection at 260 nm.

6 Conclusion

> *Science cannot solve the ultimate mystery of nature. And that is because, in the last analysis, we ourselves are a part of the mystery that we are trying to solve.*
>
> **— Max Planck**

2DCOR's features were inspected in perspective of using it in chromatography instead of its original field, spectroscopy. In a series of chromatograms where some kind of perturbation is applied to the system throughout the measurements, it can show the triggered changes and their relations on simple graphs with easy usage, from them we can gather essential information about the chemical system. There are endless possible cases what can occur in the correlation maps, it would be a rigorous task to cover them all, in this document a wide variety of representative changes were studied and the unique patterns they form in the correlation maps were presented and discussed. The results show that 2DCOR can be applied to this field efficiently, its advantages are surely beneficial.

2DCOR, however, has a few flaws: The chromatographic peaks in most cases correlate less than spectroscopic peaks, they change mostly individually, so it would make more sense to study them separately, while 2DCOR emphasises the correlations between peaks and fails to deliver basic attributes about them like the individual direction of changes. Also, chromatographic peaks tend to have a slight retention time shifts during the measurements which can disturb the detection of peak height changes. It is true, that the method can show accurately the peak shifts if that is the interest of study, but usually that is only a by-product of the system. Furthermore 2DCOR is limited to 1D measurements because the maps are two-dimensional and need 3D to plot them and 2D chromatography is becoming more and more popular with every year. Higher dimensional correlation maps are mathematically possible to calculate, but they are not useful in practice. Finally, correlation maps can

become crowded very quickly with auto- and cross-peaks, which makes the evaluation no easier.

ALA was made to fine-tune 2DCOR's fundamental properties to chromatographic usage. It creates 1D alteration maps from 1D measurements instead of 2D, which means similar information can be get from much simpler presentations and it can be applied to 2D measurement systems as well. The properties of our novel method were demonstrated by comparing it with its parent method. It is proven that it can deliver the same information as 2DCOR while has some additional features. Alteration maps are simpler to evaluate because they have no additional dimensions and cross-peaks. Although this can lead to the argument that they lack direct information about the correlations of peaks, which is arguably true, but it does not mean that this feature was left out, only simplified, because instead of cross-peaks we have to compare the individual peaks and the relevant information is revealed, so indirectly it is present. After the comparison of two methods in 2D experiments the nature of 3D ALA is presented and showed that its features are parallel with the lower dimensional counterpart, only the visualization bears a few difficulties comparably.

All the experiments about the basic properties of methods were carried out on computer generated data. After those establishments experimentally recorded chromatograms were used to illustrate the benefits of such techniques. First 2DCOR was applied in a study about reproducibility of chromatographic columns. Then two experiments were carried out with ALA in the focus. On one hand simple peak height changes were implemented with SFC in order to verify the link between simulation and measured data. On the other hand, the effect of methanol on other compounds in chromatograms was studied in SFC too. ALA was able to provide the necessary information about the chemical systems through the evaluation of perturbation induced changes and opened the way for further applications and developments.

Acknowledgement

I would like to give my deepest gratitude for my supervisor, Dr. Attila Felinger for his guidance, patience and advices. Without him this work would never be made.

Special thanks to Annamária Sepsey who helped with the format of this dissertation, my daily struggles and many other; Csanád Rédei for the SFC measurements and also my co-workers at the Department of Analytical and Environmental Chemistry who have lent their help on my path throughout the years.

Last but not least I would like to thank my beloved wife, Judit Simonné Dr. Tenk, for the years of unmeasurable support and also the lots of (necessary) goading. My mother and father for smoothing my way from the very beginning. Gábor Mikle because I needed a close friend who actually understands chemistry. Also many thanks to every friend and family member.

Thank you!

7 Thesis points

1. We introduced a novel method called Alteration Analysis (ALA), in order to extend two-dimensional correlation analysis (2DCOR) to three-dimensional datasets where a series of two-dimensional measurements – instead of the original one-dimensional – are evaluated.

2. Our method, ALA was built to use the properties of 2DCOR, but it made able to work with three-dimensional datasets and fine-tuned to use in chromatography, because of its focus on individual changes rather than the connection between them. These benefits were explored by a series of examples which showed that ALA can work equally on one and two-dimensional measurements, but more importantly that it made possible for the field of 2DCOR the analysis of higher dimensional datasets.

3. We made an intensive comparison of our method and 2DCOR on two-dimensional datasets where a series of one-dimensional measurements are used. We proved that ALA is more than capable of competing with its predecessor, because it has many advantages to overcome its one disadvantage. It can work with 3D datasets, it can show the absolute direction of changes instead of only relative ones and it can be understood much easier with its simpler graphs.

4. We tested the capability of 2DCOR in chromatography by applying it to measured chromatograms and comparing it to other chemometric methods. We clarified that it can be applied well in this field and it delivered the same informations about the reproducibility of HPLC columns as PCA.

5. We demonstrated the practical use of our new method, ALA on measured chromatograms and its advantages were highlighted in the evaluation of complex chromatographic problems. The information in ALA maps of measured chromatograms were compared to in silico experiments and only negligible differences arose, meaning the theory behind ALA can be put into good use in practical experiments. Furthermore, the effect of methanol was shown on ALA maps in SFC measurements.

Publications

Publications related to this thesis

1. **Simon, J.**, Felinger, A., Two-dimensional correlation analysis of the reproducibility of high-performance liquid chromatography columns, J. Chromatogr. A, 2015, 1384, 115–123

2. **Simon, J.**, Felinger, A., Correlation analysis on 3D data – Introducing the alteration analysis, Chemometr. Intell. Lab. Syst., 2016, 158, 54–60

3. **Simon, J.**, Felinger, A., Exploring the changes in a series of measurements – The comparison of the two-dimensional correlation analysis and the alteration analysis, Chemometr. Intell. Lab. Syst., 2017, 168, 28–37

Posters and presentations related to this thesis

1. **Simon, J.**, Felinger, A., Two-dimensional correlation in chromatography, 9th Balaton Symposium on High-Performance Separation Methods, 2013. 09.04–09.06., Siófok

2. **Simon, J.**, Felinger, A., Kétdimenziós korreláció alkalmazása a kromatográfiában, XXXVI. Kémiai Előadói Napok, 2013. 10.28–10.30., Szeged

3. **Simon, J.**, Felinger, A., Two-dimensional correlation in chromatography, 10th János Szentágothai Transdisciplinary Conference and Student Competition, 2013. 11.04–11.05., Pécs

4. **Simon, J.**, Felinger, A., Correlation on 3D data – Alteration Analysis, The 8th International Symposium on Two-Dimensional Correlation Spectroscopy, 2015. 07.08–07.11., Vienna

5. **Simon, J.**, Felinger, A., Correlation on 3D data – Alteration Analysis, 10th Balaton Symposium on High-Performance Separation Methods, 2015. 09.02–09.04., Siófok

6. **Simon, J.**, Felinger, A., Correlation on 3D data – Alteration Analysis, Conferentia Chemometrica, 2015. 09.13–09.16., Budapest

7. **Simon, J.**, Felinger, A., Alternation Analysis: An alternative to 2DCOR, The 9th International Symposium on Two-Dimensional Correlation Spectroscopy, 2017. 06.07–06.10., Victoria

8. **Simon, J.**, Felinger, A., An alternative to two-dimensional correlation analysis: Alteration Analysis, Conferentia Chemometrica, 2017. 09.03–09.06., Gyöngyös, Farkasmály

9. **Simon, J.**, Felinger, A., Correlation Analysis on a Series of 1D and 2D Chromatograms, 11th Balaton Symposium on High-Performance Separation Methods, 2017. 09.06–09.08., Siófok

Posters and presentations not related to this thesis

1. **Simon, J.**, Felinger, A., Modelling the 2D correlation in chromatography with exponentially modified Gaussian peak shape, III. Interdisciplinary Conference, 2014. 04.15–04.17., Pécs

2. **Simon, J.**, Felinger, A., Modelling the 2D correlation in chromatography with exponentially modified Gaussian peak shape, 30th International Symposium on Microscale Bioseparations, 2014. 04.27–05.01., Pécs

3. **Simon, J.**, Felinger, A., Enhancement of chromatographic peak determinations by two-dimensional correlation, 30th International Symposium on Chromatography, 2014. 09.14–09.18., Salzburg

4. **Simon, J.**, Felinger, A., A kromatográfiás kétdimenziós korrelációs térképek leírása módosított Gauss-görbékkel, Elválasztástudományi Vándorgyűlés, 2014. 11.12–11.14., Egerszalók

5. **Simon, J.**, Lambert, N.,Felinger, A., Kromatográfiás hatékonyság vizsgálata alterációs analízissel, Elválasztástudományi Vándorgyűlés, 2016. 11.09–11.11., Kecskemét

List of Symbols

Greek Symbols

α first part of EMG function

β second part of EMG function

$\mathbf{\Phi}$ synchronous correlation matrix

$\mathbf{\Psi}$ asynchronous correlation matrix

σ standard deviation

Other Variables

$\underline{\mathbf{X}}$ raw data array

$\overline{\mathbf{x}}$ average chromatogram vector

$\underline{\mathbf{D}}$ difference data array

$\mathbf{A}'$ raw asynchronous alteration matrix

$\mathbf{a}'$ raw asynchronous alteration vector

$\mathbf{A}$ maximum scaled asynchronous alteration matrix

$\mathbf{a}$ maximum scaled asynchronous alteration vector

$\mathbf{B}$ basic alteration matrix

$\mathbf{b}$ basic alteration vector

$\mathbf{D}$	difference matrix
$\mathbf{M}$	example data matrix
$\mathbf{N}$	Hilbert–Noda transform matrix
$\mathbf{P}$	correlation coefficient matrix
$\mathbf{S}'$	raw synchronous alteration matrix
$\mathbf{s}'$	raw synchronous alteration vector
$\mathbf{S}$	maximum scaled synchronous alteration matrix
$\mathbf{s}$	maximum scaled synchronous alteration vector
$\mathbf{X}$	raw data matrix
$\mathbf{Y}$	centered data matrix
τ	time constant
A	area under the peak
t	time
t_R	retention time

Glossary

2D-GPC two-dimensional correlation gel permeation chromatography

2DCOR two-dimensional correlation analysis

2DCOS two-dimensional correlation spectroscopy

AAM asynchronous alteration map

ACM asynchronous correlation map

ALA Alteration Analysis

BAM basic alteration map

EMG exponentially modified gaussian peak

GC gas chromatography

HPLC high performance liquid chromatography

IR infrared spectroscopy

NMR nuclear magnetic resonance spectroscopy

PCA principal component analysis

SAM synchronous alteration map

SCM synchronous correlation map

SFC supercritical fluid chromatography

UV ultraviolet spectroscopy

VIS visible spectroscopy

Bibliography

(1) Aue, W. P.; Bartholdi, E.; Ernst, R. R. *J. Chem. Phys.* **1976**, *64*, 2229–2246.

(2) Noda, I. *Bull. Am. Phys. Soc.* **1986**, *31*, 520.

(3) Noda, I. *Appl. Spectrosc.* **1993**, *47*, 1329–1336.

(4) Noda, I. *Appl. Spectrosc.* **2000**, *54*, 994–999.

(5) Noda, I. *Vibr. Spectrosc.* **2012**, *60*, 146–153.

(6) Sigman, M. E.; Clark, C. D. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 3731–3736.

(7) Cozzolino, D.; Cynkar, W.; Dambergs, R.; Smith, P. *Sens. Actuators, B* **2010**, *145*, 628–634.

(8) Abdulla, H. A.; Sleighter, R. L.; Hatcher, P. G. *Anal. Chem.* **2013**, *85*, 3895–3902.

(9) Choi, H. C.; Ryu, S. R.; Ji, H.; Kim, S. B.; Noda, I.; Jung, Y. M. *J. Phys. Chem. B* **2010**, *114*, PMID: 20690586, 10979–10985.

(10) Smirnova, D. S.; Kornfield, J. A.; Lohse, D. J. *Macromolecules* **2011**, *44*, 6836–6848.

(11) Guo, L.; Spegazzini, N.; Sato, H., et al. *Macromolecules* **2012**, *45*, 313–328.

(12) Park, Y.; Kim, N. H.; Kim, J. M., et al. *Appl. Spectrosc.* **2011**, *65*, PMID: 21352653, 320–325.

(13) Noda, I. *Vibr. Spectrosc.* **2004**, *36*, 143–165.

(14) Noda, I. *J. Mol. Struct.* **2006**, *799*, 2–15.

(15) Noda, I. *J. Mol. Struct.* **2008**, *883–884*, 2–26.

(16) Noda, I. *J. Mol. Struct.* **2010**, *974*, 3–24.

(17) Noda, I. *J. Mol. Struct.* **2014**, *1069*, 3–22.

(18) Noda, I. *J. Mol. Struct.* **2014**, *1069*, 23–49.

(19) Noda, I. *Chin. Chem. Lett.* **2015**, *26*, 167–172.

(20) Izawa, K.; Ogasawara, T.; Masuda, H.; Okabayashi, H.; Noda, I. *PhysChemComm* **2001**, *4*, 57–59.

(21) Hyde, J. R.; Bourne, R. A.; Noda, I.; Stephenson, P.; Poliakoff, M. *Anal. Chem.* **2004**, *76*, 6197–6206.

(22) Izawa, K.; Ogasawara, T.; Masuda, H.; Okabayashi, H.; O'Connor, C. J.; Noda, I. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1053–1061.

(23) Izawa, K.; Ogasawara, T.; Masuda, H.; Okabayashi, H.; O'Connor, C. J.; Noda, I. *J. Phys. Chem. B* **2002**, *106*, 2867–2874.

(24) Izawa, K.; Ogasawara, T.; Masuda, H.; Okabayashi, H.; Noda, I. *Macromolecules* **2002**, *35*, 92–96.

(25) Suzuki, K.; Oku, J.-i.; Izawa, K.; Okabayashi, H.-F.; Noda, I.; O'Connor, C. J. *Colloid Polym. Sci.* **2004**, *283*, 306–316.

(26) Suzuki, K.; Oku, J.-i.; Izawa, K.; Okabayashi, H.-F.; Noda, I.; O/'Connor, C. J. *Polym. J.* **2004**, *36*, 959–970.

(27) Suzuki, K.; Oku, J.-I.; Izawa, K.; Okabayashi, H.-F.; Noda, I.; O'Connor, C. J. *J. Polym. Sci. Part B* **2004**, *42*, 3447–3460.

(28) Amandi, R.; Hyde, J. R.; Ross, S. K.; Lotz, T. J.; Poliakoff, M. *Green Chem.* **2005**, *7*, 288–293.

(29) Hyde, J. R.; Walsh, B.; Poliakoff, M. *Angew. Chem.* **2005**, *117*, 7760–7763.

(30) Suzuki, K.; Oku, J.-i.; Izawa, K.; Okabayashi, H.-F.; Noda, I.; O'Connor, C. J. *Colloid Polym. Sci.* **2005**, *283*, 551–558.

(31) Kamerzell, T. J.; Li, M.; Arora, S.; Ji, J. A.; Wang, Y. J. *J. Pharmaceut. Sci.* **2011**, *100*, 1341–1349.

(32) Andary, J.; Maalouly, J.; Ouaini, R.; Chebib, H.; Rutledge, D. N.; Ouaini, N. *Chemometr. Intell. Lab. Syst.* **2012**, *113*, 58–67.

(33) Lee, Y.-K.; Hur, J. *Environmental Pollution* **2017**, *227*, 490–497.

(34) Harshman, R. A. *UCLA Work. Pap. Phon.* **1970**, *16*, 1–84.

(35) Carroll, J. D.; Chang, J. *Psychometrika* **1970**, *35*, 283–319.

(36) Kroonenburg, P. M., *Three-mode principal component analysis: Theory and applications*; DSWO Press: Leiden, 1983.

(37) Kiers, H. A. L. *Psychometria* **1991**, *56*, 449–470.

(38) Geladi, P. *Chemom. Intell. Lab. Syst.* **1989**, *7*, 11–30.

(39) Smilde, A. K. *Chemom. Intell. Lab. Syst.* **1992**, *15*, 143–157.

(40) Bro, R. *Chemom. Intell. Lab. Syst.* **1997**, *38*, 149–171.

(41) Faber, N. K. M.; Bro, R.; Hopke, P. K. *Chemom. Intell. Lab. Syst.* **2003**, *65*, 119–137.

(42) phan, A. H.; Cichocki, A. *Neurocomputing* **2011**, *74*, 1970–1984.

(43) Huang, J.; Wium, H.; Qvist, K. B.; Esbensen, K. H. *Chemom. Intell. Lab. Syst.* **2003**, *66*, 141–158.

(44) Smilde, A. K.; Doornbos, D. A. *J. Chemom.* **1991**, *5*, 345–360.

(45) Arroyo, D.; Ortiz, M. C.; Sarabia, L. A.; Palacios, F. *J. Chromatogr. A* **2008**, *1187*, 1 –10.

(46) Soares, P. K.; Bruns, R. E.; Scarminio, I. S. *Anal. Chim. Acta* **2008**, *2012*, 36 –44.

(47) Prieto, N.; Rodriguez-Méndez, M.; Leardi, R.; Oliveri, P.; Hernando-Esquisabel, D.; Iñiguez-Crespo, M.; de Saja, J. *Anal. Chim. Acta* **2012**, *719*, 43 –51.

(48) Shinzawa, H.; Nishida, M.; Kanematsu, W.; Tanaka, T.; Suzuki, K.; Noda, I. *Analyst* **2012**, *137*, 1913–1921.

(49) Simon, J.; Felinger, A. *J. Chromatogr. A* **2015**, *1384*, 115–123.

(50) Noda, I.; Ozaki, Y., *Two-Dimensional Correlation Spectroscopy – Applications in Vibrational and Optical Spectroscopy*; John Wiley & Sons, Ltd: 2005.

(51) Simon, J.; Felinger, A. *Chemometr. Intell. Lab. Syst.* **2016**, *158*, 54–60.

(52) R Core Team R: A Language and Environment for Statistical Computing.; R Foundation for Statistical Computing, Vienna, Austria, 2013.

(53) RStudio Team RStudio: Integrated Development Environment for R.; RStudio, Inc., Boston, MA, 2016.

(54) Felinger, A., *Data Analysis and Signal Processing in Chromatography*; Data Handling in Science and Technology; Elsevier Science: 1998.

(55) Felinger, A.; Kele, M.; Guiochon, G. *J. Chromatogr. A* **2001**, *913*, 23–48.

(56) Kele, M.; Guiochon, G. *J. Chromatogr. A* **1999**, *830*, 41–54.

(57) Kele, M.; Guiochon, G. *J. Chromatogr. A* **1999**, *830*, 55–79.

(58) Kele, M.; Guiochon, G. *J. Chromatogr. A* **1999**, *855*, 423–453.

(59) Kele, M.; Guiochon, G. *J. Chromatogr. A* **2000**, *869*, 181–209.

(60) Kele, M.; Guiochon, G. *J. Chromatogr. A* **2001**, *913*, 89–112.

(61) Kele, M.; Guiochon, G. *J. Chromatogr. A* **2002**, *960*, 19–49.

(62) Gritti, F.; Guiochon, G. *J. Chromatogr. A* **2003**, *1003*, 43–72.

(63) Gritti, F.; Guiochon, G. *J. Chromatogr. A* **2003**, *1021*, 25–53.

(64) Felinger, A.; Gritti, F.; Guiochon, G. *J. Chromatogr. A* **2004**, *1024*, 21–38.

(65) S. Wold K. Esbensen, P. G. *Chemometr. Intell. Lab.* **1987**, *2*, 37–52.

(66) Boelens, H. F.; Dijkstra, R. J.; Eilers, P. H.; Fitzpatrick, F.; Westerhuis, J. A. *J. Chromatogr. A* **2004**, *1057*, 21 –30.

(67) A. Savitzky, M. J. E. G. *Anal. Chem.* **1964**, *36*, 1627–1639.

(68) J. Riordon E. Zubritsky, A. N. *Anal. Chem.* **2000**, *72*, 324A–329A.

(69) Shinzawa, H.; Morita, S.-I.; Awa, K.; Okada, M.; Noda, I.; Ozaki, Y.; Sato, H. *Appl. Spectrosc.* **2009**, *63*, 501–506.

(70) Shinzawa, H.; Hashimoto, K.; Sato, H.; Kanematsu, W.; Noda, I. *J. Mol. Struct.* **2014**, *1069*, 176–182.

(71) Noda, I. In *Handbook of Vibrational Spectroscopy*; John Wiley & Sons, Ltd: 2006.

(72) Noda, I.; Dowrey, A. E.; Marcoli, C.; Story, G. M.; Ozaki, Y. *Appl. Spectrosc.* **2000**, *54*, 236A–248A.

(73) Noda, I. *J. Mol. Struct.* **2010**, *974*, 108–115.

Appendix

A

Table A.1.: The parameters of chromatograms for Fig. 5.2, 5.1, 5.40, 5.41 a)

#	t_R	A			σ	τ
		Type	a	b		
1	2.0	linear	1	100	0.1	0.1
2	4.0	linear	-2	100	0.1	0.1
3	6.0	linear	3	100	0.1	0.1
4	8.0	linear	-4	100	0.1	0.1

Table A.2.: The parameters of chromatograms for Fig. 5.3, 5.4, 5.41 b)

#	t_R	A			σ	τ
		Type	a	b		
1	2.0	linear	-1	100	0.1	0.1
2	4.0	linear	2	100	0.1	0.1
3	6.0	linear	-3	100	0.1	0.1
4	8.0	linear	4	100	0.1	0.1

Table A.3.: The parameters of chromatograms for Fig. 5.5, 5.6, 5.42

#	t_R	A			σ	τ
		Type	a	b		
1	2.0	linear	1	100	0.1	0.1
2	4.0	linear	-1	200	0.1	0.1
3	6.0	linear	1	300	0.1	0.1
4	8.0	linear	-1	400	0.1	0.1

Table A.4.: The parameters of chromatograms for Fig. 5.7, 5.8, 5.43 a)

#	t_R	A			σ	τ
		Type	a	b		
1	2.0	quadratic	1	100	0.1	0.1
2	4.0	quadratic	-2	100	0.1	0.1
3	6.0	quadratic	3	100	0.1	0.1
4	8.0	quadratic	-4	100	0.1	0.1

Table A.5.: The parameters of chromatograms for Fig. 5.9, 5.10, 5.43 b)

#	t_R	A			σ	τ
		Type	a	b		
1	2.0	exponential	1	100	0.1	0.1
2	4.0	exponential	-2	100	0.1	0.1
3	6.0	exponential	3	100	0.1	0.1
4	8.0	exponential	-4	100	0.1	0.1

Table A.6.: The parameters of chromatograms for Fig. 5.11, 5.12, 5.44 a)

#	t_R	A			σ	τ
		Type	a	b		
1	1.0	linear	1	100	0.1	0.1
2	2.5	linear	2	100	0.1	0.1
3	4.0	quadratic	1	100	0.1	0.1
4	5.5	quadratic	2	100	0.1	0.1
5	7.0	exponential	1	100	0.1	0.1
6	8.5	exponential	2	100	0.1	0.1

Table A.7.: The parameters of chromatograms for Fig. 5.14, 5.15, 5.44 b)

#	t_R	A			σ	τ
		Type	a	b		
1	1.0	quadratic	1	100	0.1	0.1
2	2.5	quadratic	2	100	0.1	0.1
3	4.0	exponential	1	100	0.1	0.1
4	5.5	exponential	2	100	0.1	0.1
5	7.0	linear	1	100	0.1	0.1
6	8.5	linear	2	100	0.1	0.1

Table A.8.: The parameters of chromatograms for Fig. 5.16, 5.17, 5.45 a)

#	t_R	A				σ	τ
		Type	a	b	c		
1	2.0	single	1	100	5	0.1	0.1
2	4.0	single	-2	100	5	0.1	0.1
3	6.0	single	3	100	5	0.1	0.1
4	8.0	single	-4	100	5	0.1	0.1

Table A.9.: The parameters of chromatograms for Fig. 5.18, 5.19, 5.45 b)

#	t_R	A				σ	τ
		Type	a	b	c		
1	2.0	single	1	100	2	0.1	0.1
2	4.0	single	-1	100	3	0.1	0.1
3	6.0	single	1	100	4	0.1	0.1
4	8.0	single	-1	100	5	0.1	0.1

Table A.10.: The parameters of chromatograms for Fig. 5.21, 5.46

#	t_R	A					σ	τ
		Type	A	t_R	σ	τ		
1	2.0	EMG	100	5	0.5	0.5	0.1	0.1
2	4.0	EMG	200	5	0.5	0.5	0.1	0.1
3	6.0	EMG	300	5	0.5	0.5	0.1	0.1
4	8.0	EMG	400	5	0.5	0.5	0.1	0.1

Table A.11.: The parameters of chromatograms for Fig. 5.22, 5.47 a)

#	t_R	A					σ	τ
		Type	A	t_R	σ	τ		
1	2.0	EMG	100	5	0.5	0.5	0.1	0.1
2	4.0	EMG	100	7	0.5	0.5	0.1	0.1
3	6.0	EMG	100	9	0.5	0.5	0.1	0.1
4	8.0	EMG	100	11	0.5	0.5	0.1	0.1

Table A.12.: The parameters of chromatograms for Fig. 5.23, 5.47 b)

#	t_R	A					σ	τ
		Type	A	t_R	σ	τ		
1	2.0	EMG	100	5	1.0	1.0	0.1	0.1
2	4.0	EMG	100	7	1.0	1.0	0.1	0.1
3	6.0	EMG	100	9	1.0	1.0	0.1	0.1
4	8.0	EMG	100	11	1.0	1.0	0.1	0.1

Table A.13.: The parameters of chromatograms for Fig. 5.24, 5.25, 5.48

#	t_R	A			σ	τ
		Type	a	b		
1	2.0	sine	1	100	0.1	0.1
2	4.0	sine	-1	100	0.1	0.1
3	6.0	cosine	1	100	0.1	0.1
4	8.0	cosine	-1	100	0.1	0.1

Table A.14.: The parameters of chromatograms for Fig. 5.26, 5.27, 5.49 a)

#	t_R	A	σ			τ		
			Type	a	b	Type	a	b
1	2.0	100	linear	0.005	0.1	linear	0.005	0.1
2	4.0	100	linear	-0.005	0.175	linear	-0.005	0.175
3	6.0	100	linear	0.01	0.1	linear	0.01	0.1
4	8.0	100	linear	-0.01	0.25	linear	-0.01	0.25

Table A.15.: The parameters of chromatograms for Fig. 5.28, 5.29, 5.49 b)

#	t_R	A	σ					τ				
			Type	A	t_R	σ	τ	Type	A	t_R	σ	τ
1	2.0	100	EMG	3	1	6	6	EMG	3	1	6	6
2	4.0	100	EMG	-3	1	6	6	EMG	-3	1	6	6
3	6.0	100	EMG	6	-5	12	12	EMG	6	-5	12	12
4	8.0	100	EMG	-6	-5	12	12	EMG	-6	-5	12	12

Table A.16.: The parameters of chromatograms for Fig. 5.30, 5.31, 5.50 a)

#	t_R			A	σ	τ
	Type	a	b			
1	linear	0.05	2	100	0.1	0.05
2	linear	0.05	4	100	0.1	0.05
3	linear	0.01	6	100	0.1	0.05
4	linear	0.01	8	100	0.1	0.05

Table A.17.: The parameters of chromatograms for Fig. 5.32, 5.33, 5.50 b)

#	t_R			A	σ	τ
	Type	a	b			
1	linear	0.05	2	100	0.1	0.05
2	linear	-0.05	5	100	0.1	0.05
3	linear	0.01	6	100	0.1	0.05
4	linear	-0.01	8	100	0.1	0.05

Table A.18.: The parameters of chromatograms for Fig. 5.34, 5.35, 5.50 c)

#	t_R						A	σ	τ
	Type	A	t_R	σ	τ	s*			
1	EMG	30	4	5	5	0.13	100	0.1	0.05
2	EMG	-30	4	5	5	1.63	100	0.1	0.05
3	EMG	30	4	8	8	4.83	100	0.1	0.05
4	EMG	30	4	8	8	6.83	100	0.1	0.05

*s is an additional starting value, the EMG function was added to it.

Table A.19.: The parameters of chromatograms for Fig. 5.36, 5.37, 5.51 a)

#	t_R	A			σ	τ
		Type	a	b		
1	1.0	linear	0.1	100	0.1	0.1
2	1.2	linear	0.2	50	0.1	0.1
3	3.0	linear	0.1	100	0.1	0.1
4	3.2	linear	-0.2	50	0.1	0.1
5	5.0	linear	0.1	100	0.1	0.1
6	5.2	linear	0.5	50	0.1	0.1
7	7.0	linear	0.1	100	0.1	0.1
8	7.2	quadratic	0.2	100	0.1	0.1
9	9.0	linear	0.1	100	0.1	0.1
10	9.2	sine	0.2	50	0.1	0.1

Table A.20.: The parameters of chromatograms for Fig. 5.38, 5.39, 5.51 b)

#	t_R	A			σ	τ
		Type	a	b		
1	1.0	linear	0.1	170	0.05	0.4
2	1.5	linear	0.2	10	0.05	0.4
3	4.0	linear	0.1	170	0.05	0.4
4	4.5	linear	-0.2	10	0.05	0.4
5	7.0	linear	0.1	170	0.05	0.4
6	7.5	sine	0.2	10	0.05	0.4

Table A.21.: The parameters of 2D chromatograms for Fig. 5.52, 5.53

#	t_R1	$A1$				$\sigma1$	$\tau1$	t_R2	$A2$	$\sigma2$	$\tau2$
		Type	a	b	c						
1	0.8	linear	0.1	200	-	0.1	0.1	0.61	100	0.1	0.1
2	2.3	linear	-0.1	200	-	0.1	0.1	0.65	100	0.1	0.1
3	3.4	sine	0.4	200	-	0.1	0.1	0.67	100	0.1	0.1
4	4.3	cosine	0.4	200	-	0.1	0.1	0.63	100	0.1	0.1
5	0.8	quadratic	0.012	200	-	0.1	0.1	1.57	100	0.1	0.1
6	0.85	NO	0	120	-	0.1	0.1	1.79	100	0.1	0.1
7	0.88	single	0.6	100	5	0.1	0.1	3.22	100	0.1	0.1
8	0.91	single	-0.6	100	5	0.1	0.1	4.19	100	0.1	0.1
9	2.31	exponential	0.0005	180	-	0.1	0.1	2.11	100	0.1	0.1
10	2.3	NO	0	200	-	0.1	0.1	2.56	100	0.1	0.1
11	3.41	NO	0	110	-	0.1	0.1	1.41	100	0.1	0.1
12	3.41	NO	0	300	-	0.1	0.1	1.62	100	0.1	0.1
13	3.42	NO	0	260	-	0.1	0.1	1.83	100	0.1	0.1
14	3.4	NO	0	270	-	0.1	0.1	2.15	100	0.1	0.1
15	3.41	NO	0	180	-	0.1	0.1	2.89	100	0.1	0.1
16	3.42	NO	0	90	-	0.1	0.1	3.21	100	0.1	0.1
17	3.42	NO	0	100	-	0.1	0.1	4.59	100	0.1	0.1
18	4.32	NO	0	140	-	0.1	0.1	2.55	100	0.1	0.1
19	4.3	NO	0	160	-	0.1	0.1	3.99	100	0.1	0.1
20	4.31	NO	0	150	-	0.1	0.1	4.18	100	0.1	0.1

Table A.22.: The parameters of 2D chromatograms for Fig. 5.54

#	t_R1			$A1$	$\sigma1$, $\tau1$		
	Type	a	b		Type	a	b
1	NO	0	1	100	linear	0.01	0.05
2	NO	0	2	100	linear	-0.01	0.14
3	NO	0	3	100	NO	0	0.05
4	NO	0	4	100	NO	0	0.05
5	NO	0	1	100	linear	0.01	0.05
6	NO	0	2	100	linear	0.02	0.05
7	NO	0	3	100	linear	0.01	0.05
8	linear	0.05	0.5	100	NO	0	0.05
9	linear	-0.05	1.95	100	NO	0	0.05
10	NO	0	2.5	100	NO	0	0.05
11	NO	0	2.5	100	NO	0	0.05
12	linear	0.05	4.2	100	NO	0	0.05

#	t_R2			$A2$	$\sigma2$, $\tau2$		
	Type	a	b		Type	a	b
1	NO	0	1	100	NO	0	0.05
2	NO	0	1	100	NO	0	0.05
3	NO	0	1	100	linear	0.01	0.05
4	NO	0	1	100	linear	-0.01	0.14
5	NO	0	1	100	linear	0.01	0.05
6	NO	0	1	100	linear	0.01	0.05
7	NO	0	1	100	linear	0.02	0.05
8	NO	0	4	100	NO	0	0.05
9	NO	0	4	100	NO	0	0.05
10	NO	0	4	100	NO	0	0.05
11	NO	0	4	100	NO	0	0.05
12	linear	0.05	4	100	NO	0	0.05

Table A.23.: The parameters of 2D chromatograms for Fig. 5.58

#	t_R1	$A1$				$\sigma1$	$\tau1$	t_R2	$A2$	$\sigma2$	$\tau2$
		Type	a	b	c						
1	0.8	NO	0	100	-	0.1	0.1	0.61	100	0.1	0.1
2	2.3	NO	0	250	-	0.1	0.1	0.65	100	0.1	0.1
3	3.4	NO	0	120	-	0.1	0.1	0.67	100	0.1	0.1
4	4.3	NO	0	150	-	0.1	0.1	0.63	100	0.1	0.1
5	1.47	NO	0	150	-	0.1	0.77	0.77	100	0.1	0.1
6	1.43	NO	0	150	-	0.1	0.1	1.81	100	0.1	0.1
7	1.48	NO	0	150	-	0.1	0.1	2.68	100	0.1	0.1
8	1.46	NO	0	150	-	0.1	0.1	2.86	100	0.1	0.1
9	1.45	NO	0	150	-	0.1	0.1	4.05	100	0.1	0.1
10	1.5	NO	0	150	-	0.1	0.1	4.37	100	0.1	0.1
11	3.02	NO	0	150	-	0.1	0.1	0.65	100	0.1	0.1
12	2.95	NO	0	150	-	0.1	0.1	1.1	100	0.1	0.1
13	3.08	NO	0	150	-	0.1	0.1	1.52	100	0.1	0.1
14	3.08	NO	0	150	-	0.1	0.1	2.03	100	0.1	0.1
15	3.07	NO	0	150	-	0.1	0.1	3.05	100	0.1	0.1
16	3.06	NO	0	150	-	0.1	0.1	3.46	100	0.1	0.1
17	2.98	NO	0	150	-	0.1	0.1	3.67	100	0.1	0.1
18	3.01	NO	0	150	-	0.1	0.1	4.77	100	0.1	0.1
19	0.8	NO	0	150	-	0.1	0.1	1.57	100	0.1	0.1
20	0.85	NO	0	120	-	0.1	0.1	1.79	100	0.1	0.1
21	0.88	linear	10	80	-	0.1	0.1	3.22	100	0.1	0.1
22	0.91	NO	0	110	-	0.1	0.1	4.19	100	0.1	0.1
23	2.31	linear	15	180	-	0.1	0.1	2.11	100	0.1	0.1
24	2.30	NO	0	200	-	0.1	0.1	2.56	100	0.1	0.1
25	3.41	NO	0	110	-	0.1	0.1	1.41	100	0.1	0.1
26	3.41	NO	0	300	-	0.1	0.1	1.62	100	0.1	0.1
27	3.42	linear	10	260	-	0.1	0.1	1.83	100	0.1	0.1
28	3.4	NO	0	270	-	0.1	0.1	2.15	100	0.1	0.1
29	3.41	single	50	180	6	0.1	0.1	2.89	100	0.1	0.1
30	3.42	NO	0	90	-	0.1	0.1	3.21	100	0.1	0.1
31	3.42	NO	0	100	-	0.1	0.1	4.59	100	0.1	0.1
32	4.32	NO	0	140	-	0.1	0.1	2.55	100	0.1	0.1
33	4.30	NO	0	160	-	0.1	0.1	3.99	100	0.1	0.1
34	4.31	NO	0	150	-	0.1	0.1	4.18	100	0.1	0.1

Table A.24.: The parameters of 2D chromatograms for Fig. 5.59

#	t_R1	$A1$			$\sigma1$	$\tau1$	t_R2	$A2$	$\sigma2$	$\tau2$
		Type	a	b						
1	0.8	linear	0.1	100	0.1	0.1	0.61	100	0.1	0.1
2	2.3	linear	0.1	250	0.1	0.1	0.65	100	0.1	0.1
3	3.4	linear	0.1	120	0.1	0.1	0.67	100	0.1	0.1
4	4.3	linear	0.1	150	0.1	0.1	0.63	100	0.1	0.1
5	1.47	linear	0.1	150	0.1	0.77	0.77	100	0.1	0.1
6	1.43	linear	0.1	150	0.1	0.1	1.81	100	0.1	0.1
7	1.48	linear	0.25	150	0.1	0.1	2.68	100	0.1	0.1
8	1.46	linear	0.1	150	0.1	0.1	2.86	100	0.1	0.1
9	1.45	linear	0.1	150	0.1	0.1	4.05	100	0.1	0.1
10	1.5	linear	0.1	150	0.1	0.1	4.37	100	0.1	0.1
11	3.02	linear	0.1	150	0.1	0.1	0.65	100	0.1	0.1
12	2.95	linear	0.25	150	0.1	0.1	1.1	100	0.1	0.1
13	3.08	linear	0.1	150	0.1	0.1	1.52	100	0.1	0.1
14	3.08	linear	0.1	150	0.1	0.1	2.03	100	0.1	0.1
15	3.07	linear	0.1	150	0.1	0.1	3.05	100	0.1	0.1
16	3.06	linear	0.1	150	0.1	0.1	3.46	100	0.1	0.1
17	2.98	linear	0.1	150	0.1	0.1	3.67	100	0.1	0.1
18	3.01	linear	0.1	150	0.1	0.1	4.77	100	0.1	0.1
19	0.8	linear	0.1	150	0.1	0.1	1.57	100	0.1	0.1
20	0.85	linear	0.1	120	0.1	0.1	1.79	100	0.1	0.1
21	0.88	linear	0.1	80	0.1	0.1	3.22	100	0.1	0.1
22	0.91	linear	0.1	110	0.1	0.1	4.19	100	0.1	0.1
23	2.31	linear	0.1	180	0.1	0.1	2.11	100	0.1	0.1
24	2.30	linear	0.1	200	0.1	0.1	2.56	100	0.1	0.1
25	3.41	linear	0.1	110	0.1	0.1	1.41	100	0.1	0.1
26	3.41	linear	0.25	300	0.1	0.1	1.62	100	0.1	0.1
27	3.42	linear	0.1	260	0.1	0.1	1.83	100	0.1	0.1
28	3.4	NO	0	270	0.1	0.1	2.15	100	0.1	0.1
29	3.41	linear	0.1	180	0.1	0.1	2.89	100	0.1	0.1
30	3.42	linear	0.1	90	0.1	0.1	3.21	100	0.1	0.1
31	3.42	linear	0.1	100	0.1	0.1	4.59	100	0.1	0.1
32	4.32	linear	0.1	140	0.1	0.1	2.55	100	0.1	0.1
33	4.30	linear	0.1	160	0.1	0.1	3.99	100	0.1	0.1
34	4.31	linear	0.1	150	0.1	0.1	4.18	100	0.1	0.1

Table A.25.: The parameters of 2D chromatograms for Fig. 5.55, 5.56, 5.57

#	t_R1	$A1$			$\sigma 1$	$\tau 1$	t_R2	$A2$	$\sigma 2$	$\tau 2$
		Type	a	b						
1	1.0	linear	0.1	100	0.1	0.1	2.0	100	0.1	0.1
2	1.2	linear	0.2	50	0.1	0.1	2.0	100	0.1	0.1
3	3.0	linear	0.1	100	0.1	0.1	2.0	100	0.1	0.1
4	3.2	linear	-0.2	50	0.1	0.1	2.0	100	0.1	0.1
5	5.0	linear	0.1	100	0.1	0.77	2.0	100	0.1	0.1
6	5.2	linear	0.5	50	0.1	0.1	2.0	100	0.1	0.1
7	7.0	linear	0.1	100	0.1	0.1	2.0	100	0.1	0.1
8	7.2	quadratic	0.2	50	0.1	0.1	2.0	100	0.1	0.1
9	9.0	linear	0.1	100	0.1	0.1	2.0	100	0.1	0.1
10	9.0	sine	0.2	50	0.1	0.1	2.0	100	0.1	0.1
11	1.0	linear	0.1	170	0.05	0.4	5.0	100	0.05	0.4
12	1.5	linear	0.2	10	0.05	0.4	5.0	100	0.05	0.4
13	4.0	linear	0.1	170	0.05	0.4	5.0	100	0.05	0.4
14	4.5	linear	-0.2	10	0.05	0.4	5.0	100	0.05	0.4
15	7.0	linear	0.1	170	0.05	0.4	5.0	100	0.05	0.4
16	7.5	sine	0.2	10	0.05	0.4	5.0	100	0.05	0.4

Table A.26.: The parameters of chromatograms for Fig. 5.69 a), 5.70 a)

#	t_R	A					σ	τ
		Type	a	b	c	d		
1	1.68	linear	0.005	0.025	-	-	0.03	0.03
2	1.91	quadratic	-0.0005	0.07	-	-	0.025	0.025
3	2.22	single	0.045	0.025	5	-	0.022	0.022
4	2.64	single	-0.045	0.07	5	-	0.022	0.022
5	3.2	sine	0.0231	0.0471	-	-	0.022	0.022
6	3.92	cosine	0.0231	0.0471	-	-	0.022	0.022
7	4.88	EMG	$A = 0.086$	$t_R = 5$	$s = 0.5$	$\tau = 0.5$	0.025	0.025