

# An evaluation of a new approach for deconvolution of liquid chromatography-diode array detection data



LUNDS TEKNISKA  
HÖGSKOLA  
Lunds universitet

Bachelor thesis:  
Jasmina Zdralic

© Copyright Jasmina Zdralic

LTH School of Engineering at Campus Helsingborg  
Lund University  
Box 882  
SE-251 08 Helsingborg  
Sweden

LTH Ingenjörshögskolan vid Campus Helsingborg  
Lunds Universitet  
Box 882  
251 08 Helsingborg

Printed in Sweden  
Media-Tryck  
Biblioteksdirektionen  
Lund University  
Lund 2004

## **Abstract**

An evaluation of a new approach for deconvolution of liquid chromatography-diode array detection data

This bachelor thesis describes the evaluation of a new software for deconvolution of liquid chromatography-diode array detection (LC-DAD). The principal component analysis (PCA) based algorithm, Archimedes, is under development and there are still some issues that need to be resolved. However, the results are encouraging and potentially this algorithm will significantly extend and improve the applicability of LC-DAD.

There are a number of parameters that affects the stability of the UV-signal, the baseline, and thereby the accuracy and precision of the LC-analysis, like solvent absorbance, temperature- and pressure fluctuations, flow rate changes, etc. In addition, different analytes often have different spectra and thus different UV-response. Archimedes deconvolutes the DAD-matrix and determines spectra and concentration profiles for both analytes and parameters resulting in a background absorbance.

By eliminating the background absorbance perfect baselines are obtained, i.e. no background. Furthermore, since the concentration profile is wavelength independent Archimedes makes the DAD a more general detector. In theory Archimedes should even allow the deconvolution of partially co-eluting analytes.

Keywords: LC-DAD, PCA, spectral background, background subtraction

## **Sammanfattning**

En utvärdering av en ny metod för dekonvolering av vätskekromatografi-diode array detektions data

Detta examensarbete beskriver utvärderingen av en ny mjukvara för dekonvolering av vätskekromatografi-diode array detektion (LC-DAD). Den principal komponent analys (PCA) baserade algoritmen, Archimedes, är under utveckling och där finns fortfarande vissa problem som behöver klaras upp. Likväl är resultaten uppmuntrande och potentiellt kommer denna algoritm att signifikant utöka och förbättra användbarheten av LC-DAD.

Det finns ett antal parametrar som påverkar stabiliteten av UV-signalen, baslinjen, och därigenom också noggrannheten och precisionen av LC-analysen. Dessa är lösningsmedlets absorption, temperatur-och tryckfluktuationer, flödesförändringar, etc. Dessutom har olika analyter ofta olika spektra och sålunda olika UV-responser. Archimedes dekonvolverar DAD-matrisen och bestämmer spektra och koncentrationsprofiler för både analyter och de parametrar som resulterar i en bakgrundsabsorbans.

Genom att eliminera bakgrundsabsorptionen erhålls perfekta baslinjer, dvs. ingen bakgrundsabsorbans. Dessutom blir DAD en mer generell detektor genom Archimedes eftersom koncentrationsprofilen är våglängdsberoende. Teoretiskt sett skulle Archimedes även kunna göra det möjligt att dekonvolera delvis överlappande analyter.

Nyckelord: LC-DAD, PCA, spektral bakgrund, bakgrunds subtraktion

## Table of contents:

<b>1 Introduction .....</b>	<b>1</b>
<b>2 Experimental.....</b>	<b>4</b>
<b>2. 1 Equipment.....</b>	<b>4</b>
<b>2. 2 Chemicals .....</b>	<b>4</b>
<b>2. 3 Columns.....</b>	<b>4</b>
<b>2. 4 Softwares .....</b>	<b>5</b>
<b>2. 5 Representative experimental conditions .....</b>	<b>5</b>
<b>3 Results and discussion.....</b>	<b>8</b>
<b>3. 1 Workflow and user interface.....</b>	<b>8</b>
<b>3. 1. 1 Extracting effect vectors.....</b>	<b>9</b>
<b>3. 1. 2 Evaluating analyte data .....</b>	<b>11</b>
<b>3. 1. 3 Suggested improvements of the user interface.....</b>	<b>12</b>
<b>3. 2 Experiments made to investigate the capability of Archimedes.....</b>	<b>12</b>
<b>3. 2. 1 Temperature .....</b>	<b>13</b>
<b>3. 2. 2 Flow/Pressure.....</b>	<b>14</b>
<b>3. 2. 3 A more general detection- “concentration” chromatograms.....</b>	<b>16</b>
<b>3. 2. 4 UV-absorbing buffers.....</b>	<b>17</b>
<b>3. 2. 4. 1 Ammonia .....</b>	<b>17</b>
<b>3. 2. 4. 2 Trifluoroacetic acid .....</b>	<b>18</b>
<b>3. 2. 4. 3 Formic acid.....</b>	<b>19</b>
<b>3. 2. 5 A comparison with other methods for background subtraction .....</b>	<b>20</b>
<b>3. 2. 6 Bleeding columns .....</b>	<b>23</b>
<b>3. 3. 6. 1 Symmetry C18.....</b>	<b>23</b>
<b>3. 3. 6. 2 Polaris Amide C18.....</b>	<b>24</b>
<b>3. 3. 6. 3 Discovery HS F5.....</b>	<b>25</b>
<b>3. 2. 7 Improved quality of UV spectra-“error-free spectra” .....</b>	<b>26</b>
<b>3. 2. 8 Elimination of ghost peaks .....</b>	<b>29</b>
<b>3. 2. 9 Improvement of signal-to-noise ratio .....</b>	<b>35</b>
<b>3. 2. 10 Influence of number of effect vectors and wavelength range.....</b>	<b>36</b>
<b>3. 2. 11 Does the integration parameters of the residual curve matter? .....</b>	<b>40</b>
<b>3. 3 Potential application areas .....</b>	<b>42</b>

3. 4 “Bugs” and concerns .....	42
3. 4. 1 <i>Bugs</i> .....	42
3. 4. 2 <i>Concerns</i> .....	42
4 Conclusions.....	44
5 Acknowledgments .....	45
6 References.....	46
Appendix 1: Data files .....	48

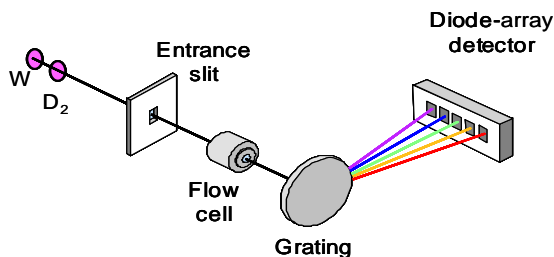
# 1 Introduction

Liquid chromatography in combination with UV detection (LC-UV) is the most widely used analytical technique [1] within the pharmaceutical industry. This combination is the main technique for the determination of dose strength, degradation products, enantiomeric purity, impurities, etc.

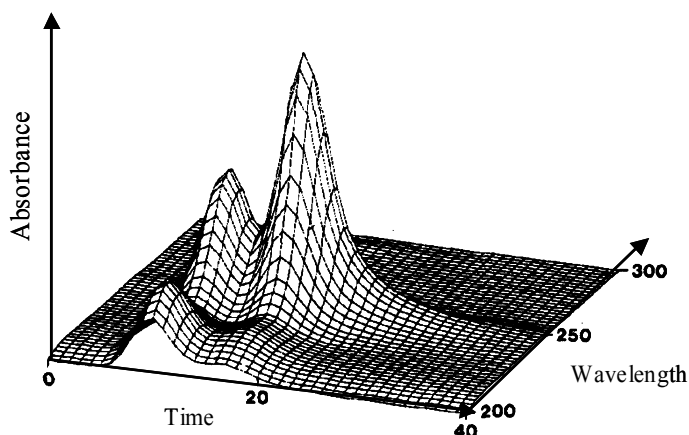
Although UV-detection is the most widely used detection principle for LC it has some important drawbacks and limitations. The most serious drawback is the large differences in the extinction coefficient and therefore also detector response that many compounds display. Another problem is the difficulty to define how much of the signal, which is related to an analyte and how much which is, related to the background, i.e. it is often difficult to “draw the baseline”.

The background signal and thereby the accuracy and precision of the analysis is affected by a number of uncontrolled physical factors, e.g. changes in mobile phase absorbance, temperature and refraction index (RI). The RI-effects are caused by mobile phase gradients and flow rate changes. Fluctuations in lamp intensity or temperature are other factors that also influence the signal. Since the influence of these effects neither can be predicted nor extracted from the measured data, the only way to eliminate background signals is to “draw a baseline” [2]. It is, however, not always evident how the baseline should be drawn. Is it a peak or is it just noise? Where does the peak start and where does it end? The definition of the baseline and the integration of chromatographic peaks is a demanding as well as time consuming process.

Diode array detection (DAD) or photodiode array detection (PDA) is one version of UV-detection where the UV absorbance is recorded for a whole range of wavelengths simultaneously, i.e. for each data point in time a complete spectrum is recorded [3]. Figures 1 and 2 illustrate the principle for DAD and the data matrix generated from it.



**Figure 1.** Simplified diagram of diode array spectrophotometer with a tungsten lamp (350-950 nm) and a deuterium lamp (160-600 nm) [2].



**Figure 2.** An illustration of the data obtained by DAD, i.e. spectra at each time point in the chromatogram [4].

Since the introduction of DAD, in the late 1980, several attempts have been made to deconvolute DAD data in order to separate partially co-eluting peaks as well as assess peak purity (i.e. determine if there is a small peak hiding under a peak of interest). Different chemometric methods i.e. mathematical and statistical methods applied to chemical problems [5], have been used for this purpose. One of the more commonly used chemometrical tools is principle component analysis (PCA). The use of this technique for analysis of DAD data has been thoroughly studied, a search in Chemical Abstract on LC and DAD and PCA gave, approximately 5000 hits.

PCA is one of the most widespread multivariate chemometric techniques [6]. It allows a simple graphical interpretation of any large data tables but it can also be used to deconvolute a DAD data matrix into concentration profile and a spectrum for each significant contribution to the absorbance. The principle is briefly described in Section 3. 1.

Several studies show how PCA can be used to resolve partially overlapping peaks and evaluate peak purity. To our knowledge there have, however, not been any publications showing that the factors, which generates the background absorbance, can be modelled and subsequently subtracted.

Some chemometric methods used today are, as mentioned above, PCA [7, 9-12], multiple linear regression (MLR) [6-8, 13-14], evolving factor analysis (EFA) [10, 15-19], functional principle component analysis (FPCA) [20], multivariate curve resolution [1, 21], etc.

The purpose of this bachelor thesis was to evaluate a prototype of a PCA based software, “Archimedes”, that currently is under development. This software is, in contrast to previously published algorithms, claimed to be able



to deconvolute the DAD data matrix and subsequently model the signals from the different factors that contribute to the background absorbance as well as the absorbance of the analytes.

## 2 Experimental

### 2.1 Equipment

LC-DAD: Series 1100 (Agilent Technologies, Waldbronn, Germany)  
Autosampler: G1313A and G1329A  
Degasser: G1322A  
Quaternary pump: G1311A  
Thermostatted column compartment: G1316A  
Diode array detector: G1315A

A Mistral Type 880 column oven (Spark, Emmen, Holland)

Light induced degradation: An Atlas Suntest CPS+ accelerated lightning unit with xenon lamp and windowglassfilter (Atlas Material Testing Technology BV, Linsengericht, Germany)

### 2.2 Chemicals

Following chemicals, nicotine free base 98%, benzylamine HCL, diphenhydramine, nortriptyline HCL and phenol 99.5% were obtained from Sigma-Aldrich Chemie GmbH, Steinheim, Germany.

Remacemide HCL, terbutaline sulphate 99.5%, salbutamol sulphate 99.7%, ARC68397 HCL 98.4%, ARD12495 HCL, budesonide 99.6% and sample X were obtained from AstraZeneca R&D Lund, Lund, Sweden.

Potassiumdihydrogenphosphate *pro analysi*, orthophosphoric acid 85%, acetic acid 100% *pro analysi*, acetonitrile *gradient grade*, ammonium acetate 98% *pro analysi* were obtained from Merck, Darmstadt, Germany.

Methanol *HPLC-gradient grade* (J. T. Baker, Malinckrodt, Deventer, Holland), ethanol 99.5% (Kemetyl AB, Haninge, Sweden), formic acid 100% *analytical reagent* (Sigma-Aldrich Sweden AB, Stockholm, Sweden), ammonia 25% *pro analysi* (Merck VWR International AB, Stockholm, Sweden), TFA 100% *sequanal grade* (Pierce, Rockford, Illinois, USA), bambuterol 99.6% (Astra Draco AB, Lund, Sweden), procainamide HCL 99% (Aldrich Chemical company Inc., Milwaukee, USA) and reagent water was purified using Milli-Q (Millipore) equipment.

### 2.3 Columns

XTerra MS C18 150 x 3.0 mm, 3.5  $\mu\text{m}$  (Waters, Milford, Massachusetts, USA)

Symmetry C18 150 x 2 mm, 3  $\mu\text{m}$  (Waters)

Discovery HS F5 150 x 2.0 mm, 3  $\mu\text{m}$  (Supelco, Bellefonte, PA, USA)

Polaris Amide C18 150 x 2 mm, 3  $\mu\text{m}$  (Varian, Torrance, CA, USA)

HyPurity C18 150 x 3 mm, 3 µm (Thermo Electron Corporation, ChromTech, Hägersten, Sweden)

Zorbax SB C18 150 x 3 mm, 3 µm (Agilent Technologies, Waldbronn, Germany)

## 2. 4 Softwares

ACD/ChromManager by Advanced Chemistry Development Inc. Toronto, Ontario, Canada.

ChemStation for LC 3D version B. 01. 01 [113] with Archimedes version Arch. 02. 01 [7] by Agilent Technologies.

## 2. 5 Representative experimental conditions

The DAD parameters were the same for all the experiments:

Peakwidth	0.05 min
Slit	2 nm
Spectrum store	All
Spectrum range	200-950 nm
Spectrum step	1 nm
Spectrum threshold	N/A
Lamps	UV and VIS on
Time	Stop time as pump Post time off
Autobalance	Pre run on and post run off
Margin for negative absorbance	1000 mAU

See Appendix 1 for the experiment's data files.

For the **temperature induced baseline disturbances** the mobile phases were water and acetonitrile (50:50 v/v) with a flow of 0.40 mL/min. Signals were stored at 210,8 and 254,8 nm (reference off).

For the **flow-and pressure disturbances** the mobile phase was water and signals were stored at 210,8 and 254,8 nm (reference off).

Three different **UV-absorbing buffers** were used as mobile phases to obtain different baseline disturbances:

- A) 10 mM formic acid in water and 10 mM formic acid in 95% acetonitrile
- B) 10 mM TFA in water and 10 mM TFA in 95% acetonitrile
- C) 20 mM ammonia in water and 20 mM ammonia in 95% acetonitrile

All three gradients were 5-100% acetonitrile for 20 min, post time 10 min with a flow of 0.64 mL/min. The injection volume was 2  $\mu$ L and column temperature was set to 20 °C with an XTerra MS C18 column. Signals were stored at 254 and 210 nm (reference off).

To generate **different types and amount of bleeding from columns** the following columns were used:

- A) Symmetry C18
- B) Discovery HS F5
- C) Polaris Amide C18

Since these columns have an inner diameter (i.d.) of 2 mm the detector was equipped with a semi micro cell to minimize band broadening and the column temperature was set to 60 °C. The mobile phases were 20 mM  $\text{KH}_2\text{PO}_4$  (pH 2.7) in water and 20 mM  $\text{KH}_2\text{PO}_4$  (pH 2.7) in methanol/water (65:35 v/v) with a flow of 0.21 mL/min.

The gradient was:

Time (min)	% $\text{KH}_2\text{PO}_4$ in MeOH/ $\text{H}_2\text{O}$
0	5
20	100
25	100
26	5

with a post time of 15 min and the injection volume was 1  $\mu$ L. Signals were stored at 210,16 (reference off) and 254,16 nm (reference 450,80 nm).

The data generated for the evaluation if Archimedes could give a **more general detection** the mobile phases were 10 mM formic acid and acetonitrile. The acetonitrile gradient was 5-95% for 20 min with a flow of 0.64 mL/min. The injection volume was 2  $\mu$ L and a Zorbax SB C18 column was used with the temperature 40 °C. Signals were stored at 254,8 and 210,8 (reference 300,100).

For all the other investigations the mobile phases were 10 mM NH<sub>4</sub>Ac in water and 10 mM NH<sub>4</sub>Ac in acetonitrile/water (95:5 v/v) with a flow of 0.64 mL/min. The injection volume was 7 μL and a HyPurity C18 column was used with the temperature 36 °C. Signals were stored at 224 nm (reference off).

The gradient was:

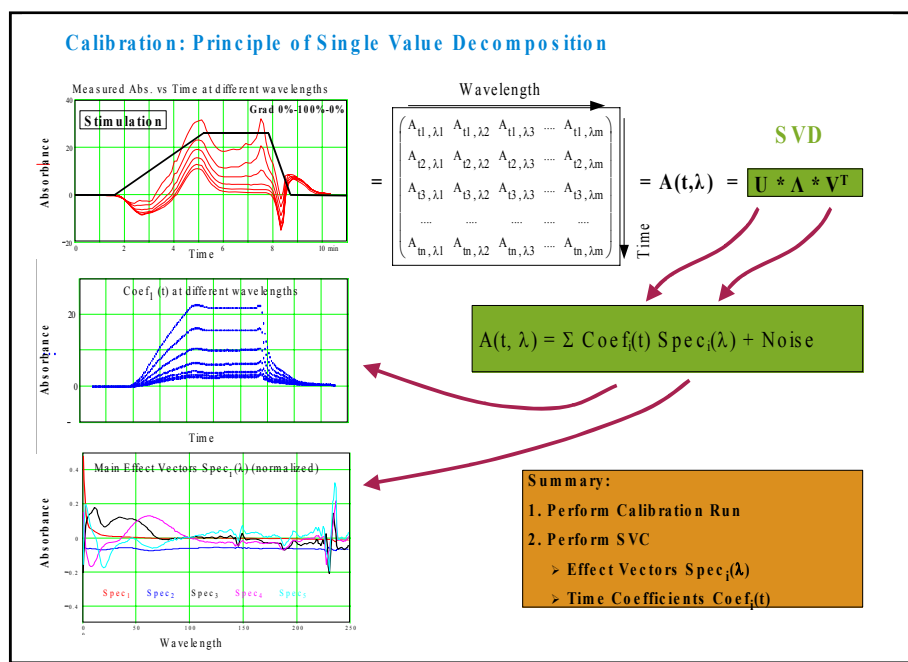
<b>Time (min)</b>	<b>% NH<sub>4</sub>Ac in acetonitrile/H<sub>2</sub>O</b>
0	5
2.5	5
17.5	100
25.5	100
25.6	5
33	5

For a more controlled experiment of error-free spectra and elimination of ghost peaks a “pure solvent run” (PSR) containing a “known ghost peak” (ARC68397) and two analytical runs with the ghost peak separated from and co-eluting with an analyte (ARD12495) was done. For the separated analytes the mobile phases were 65 % of 20 mM KH<sub>2</sub>PO<sub>4</sub> in methanol/water (65:35 v/v) and 35% of 20 mM KH<sub>2</sub>PO<sub>4</sub> in water. But for the partially overlapped peaks the mobile phases were 90% of 20 mM KH<sub>2</sub>PO<sub>4</sub> in methanol/water (65:35 v/v) and 10% of methanol. The flow was 0.64 mL/min, injection volume 2 μL and the column temperature was 60 °C. The analysis time was set to 20 min. A Zorbax SB C18 column was used and signals were stored at 210 and 254 nm (reference off).

### 3 Results and discussion

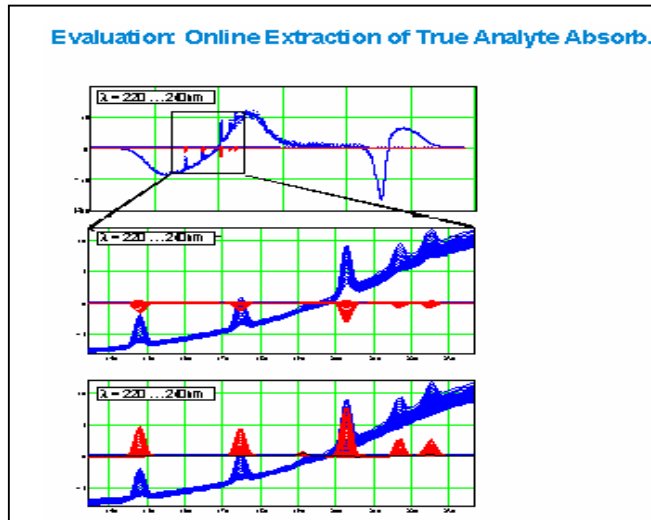
#### 3.1 Workflow and user interface

As a first step the spectra and concentration profiles for the background effects are determined from a PSR, i.e. a blank run where only the sample solvent is injected. A matrix of wavelengths and time is obtained and by using PCA the matrix is deconvoluted into the effect's spectra and concentration profiles. This is founded on the well-known Lambert Beers law,  $A = \epsilon \cdot c \cdot l$ , where the extinction coefficient corresponds to spectra and concentration corresponds to the concentration profiles (see Figure 3).



**Figure 3.** Archimedes works that first a PSR is run and a matrix of wavelengths and time is obtained. By using PCA this matrix is deconvoluted into effect's concentration profile and spectra.

In a second step the sample is injected and the “analytical run” (AR) recorded. The concentration profiles for the background effects are determined by MLR. After subtraction of the background absorbance a residual is left which is related to the analytes in the sample. The square sum of this residual is calculated and integrated. At each peak apex in this residual square sum a spectrum is determined (see Figure 4). Using the spectrum from each analyte it is then possible to calculate the concentration profile for each analyte by MLR. It is also possible to calculate a background-free chromatogram at a certain wavelength.



**Figure 4.** After the PSR (Figure 3) an analytical run is run where the background profile is obtained by MLR. After subtraction of the background absorbance a residual is left. Any peaks remaining in the residual square sum corresponds to an analyte.

### ***3. 1. 1 Extracting effect vectors***

First of all the number of effect vectors that describes the background has to be calculated and this is done by entering the '3D-Dataprocessing' menu and then the 'Effect vectors' box (see Figure 5).

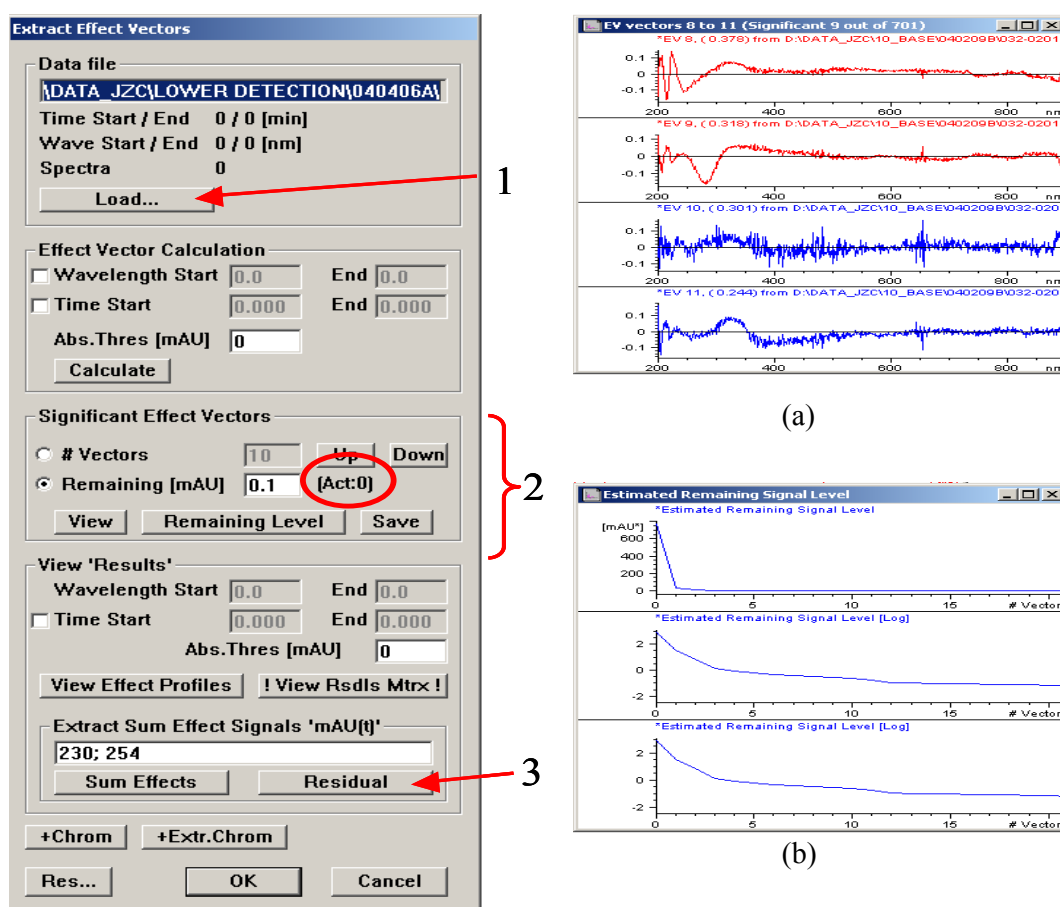
When this box is entered the next step is to load the PSR and choose the wavelength range, 200 nm and maximum of 150 nm plus the absorption range of the analytes. After this effect vector plots are calculated and the decision of how many are enough to describe the background is taken. There are three criteria of selecting the right number:

1. Looking at the effect vectors spectra plots and stopping when the first effect vector only shows noise.
2. Looking at the remaining level (it is the estimated remaining signal level against the number of effect vectors, i.e. it explains how much each effect vector describes the background) and stopping when the log-curve is planar.
3. Looking at the estimated remaining signal (which gives a rough estimation of the standard deviation of the remaining data that cannot be expressed by the effect vectors) and stopping when it has got a "low significance" (e.g. 0.2).

One should take enough effect vectors to describe the background, but as few as possible to avoid overfitting. Around 8-12 is normal.

When the right number of effect vectors has been selected the <save>-button is pressed and the effect vectors are saved under the filename “EFFVEC.REG” in the PSR directory.

The residual matrix can also be viewed and a residual curve (which display the square sum of the whole wavelength range). Also a specific wavelength and bandwidth can be extracted. The last two buttons at the end, <+Chrom> is for viewing all the loaded signals of the data file in the same window and <+Extr.Chrom>-button is the desired signals extracted form the original spectra matrix [2].



**Figure 5.** The user interface where the number of effect vectors from the blank is selected. 1) The blank is loaded, 2) The number of effect vectors are selected using three criteria, (a) effect vectors spectra, Remaining [mAU] which gives the rough standard deviation and (b) Remaining level which shows how much each vector is explaining the background 3) It is also possible to obtain residual plots and reconstructed chromatograms after a wavelength and bandwidth selection.



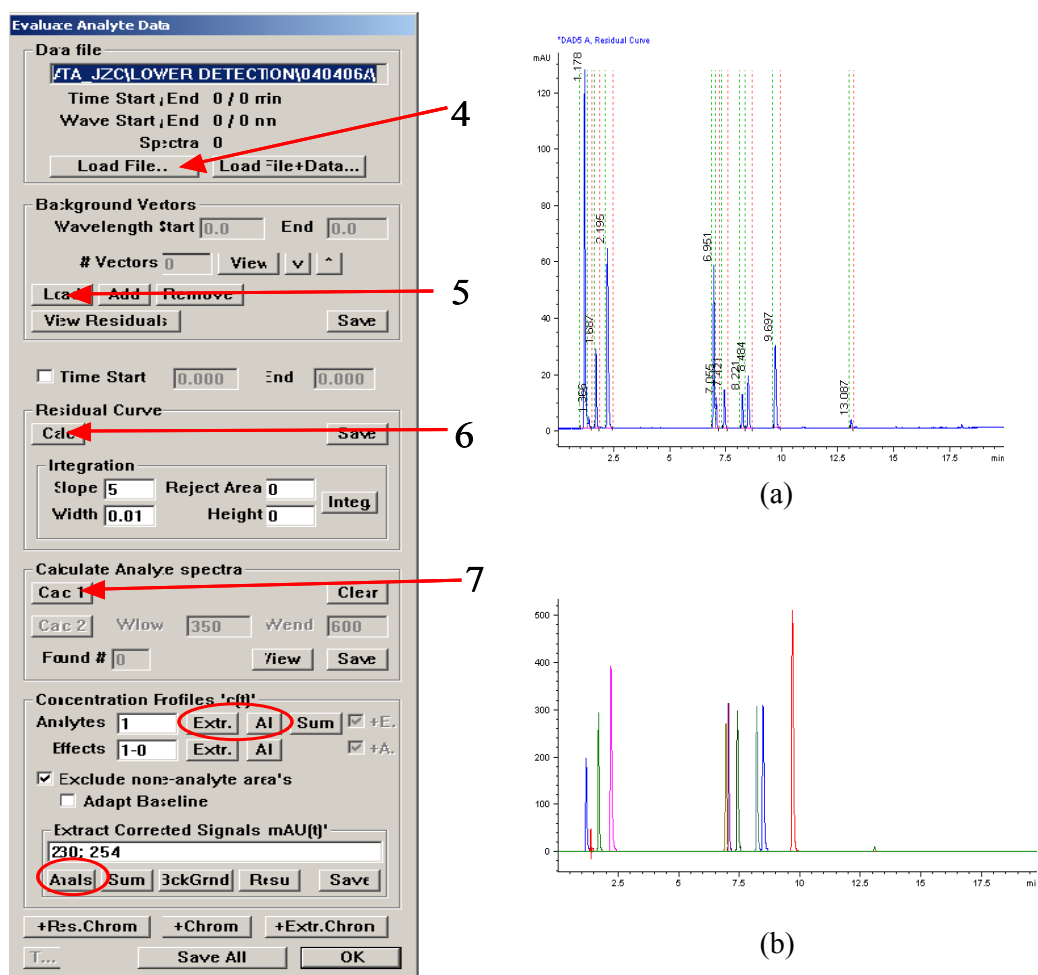
### ***3. 1. 2 Evaluating analyte data***

The next step is to get results, a wavelength independent concentration profile of the analytes, background-free chromatograms or spectra of the analytes, by entering the ‘Analytes...’ box (Figure 6). The analytical run is loaded, as well as the background vectors from the PSR. It is possible to combine several sets of background vectors by add/remove commands. Then the <save>-button is pressed and the vectors are saved for the analytical run under the file “BKGVEC. REG”.

To get the integrated residual curve for the data the <calculate>-button is pressed. If the <save>-button in the ‘Concentration profiles c(t)/Extracted corrected signals mAU(t)’ is pressed the peak areas are available after loading the residual curve as normal in ChemStation. Also the integration parameters can be changed to detect more analytes.

To get the spectra and then also later on the concentration profiles and the extracted corrected analyte signals (i.e. the background-free chromatogram at a selected wavelength and bandwidth) the analytes has to be detected in the residual curve to be able to get their spectra. The <Calc 1>-button is pressed to obtain the spectra and afterwards they are saved. The concentration profiles are shown by pressing the <All>-button or if just some specified analytes are selected the <Extr>-button is pressed. To get the extracted corrected analyte signals and the extracted residual curve a wavelength and a bandwidth (e.g. 210,4) is selected and the buttons <Anals> and <Residual> respectively are pressed to view the plots. For <+Chrom> and <+Extr. Chrom> see Section 3. 1. 1.

The next time the same analytical run is viewed the <load file+data>-button is pressed and all the saved calculations like effect vectors, residual curve, spectra, etc. are loaded [2].



**Figure 6.** User interface where the analytical run is evaluated after extracting the effect vectors (Figure 5), 4) Load the analytical run, 5) Load the effect vectors from the PSR, 6) Calculate (a) the residual curve, i.e. all that is left after subtracting the background, 7) Calculate the analyte's spectra and then (b) the concentration profiles and the extracted corrected analyte signals.

### 3. 1. 3 Suggested improvements of the user interface

Some improvements of Archimedes would be to be able to check one analyte at a time from a list of all detected analytes. As it is today it is difficult to know which analyte is, for example, number 3. Another alternative would be to use the retention time instead of the peak number. A scrolling function between the chromatograms would also be helpful. It would also be desirable to be able to handle the plots like they are handled in ChemStation, e.g. to be able to get overlay of spectra with spectra from ChemStation.

### 3. 2 Experiments made to investigate the capability of Archimedes

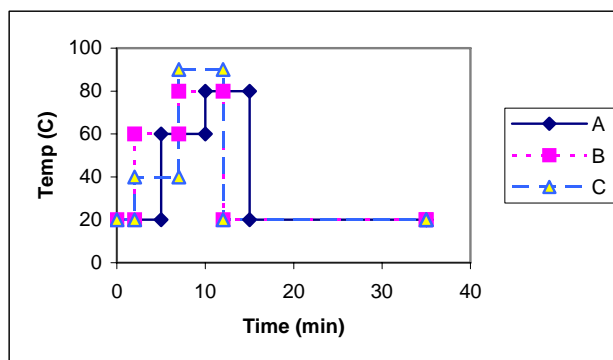
In order to investigate what types of disturbances Archimedes can model a series of different types of baseline disturbances was generated by inducing 1) temperature and 2) flow rate changes, 3) performing gradient runs with

different types of UV absorbing buffers, 4) employing columns with different type and extent of bleeding and, finally 5) generating ghost peaks. A more general detection, the signal-to-noise ratio, influence of number of effect vectors and wavelength range, the residual curve's integration parameters and spectra improvement were also investigated.

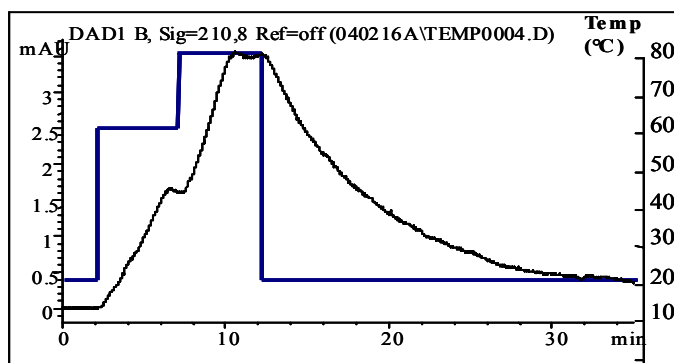
### 3. 2. 1 Temperature

The purpose was to investigate if Archimedes could compensate for temperature induced baseline disturbances. Baseline disturbances were generated by placing a 2 m x 0.12 mm i.d. capillary in a column oven, placed between the autosampler and the DAD. By programming different temperature programs (Figure 7) it is thereby possible to investigate if shifts in time (dotted line with squares in Figure 7) as well as in temperature (dotted line with triangles in Figure 7) can be compensated.

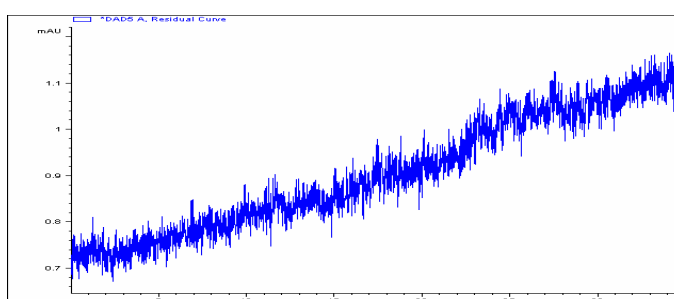
The raw signal is seen, with the temperature gradient, in Figure 8. Figure 9 shows the residual curve for AR 1 after calculating with Archimedes. The residual curve displays everything that is left after subtracting the background. Archimedes is able to compensate for temperature induced baseline disturbances because the residual curve in Figure 9 only shows noise. The drift in the residual curve is acceptable. It should not be stressed that the temperature gradient in this experiment is more than an order of magnitude larger than what can be expected during a conventional LC-analysis.



**Figure 7.** Temperature programs A (line with diamonds which is the calibration run), B (dotted line with squares which is the AR 1) and C (dotted line with triangles which is the AR 2).



**Figure 8.** Analytical run (AR) 1 at 210 nm with the temperature gradient B (for conditions see Section 2. 5).



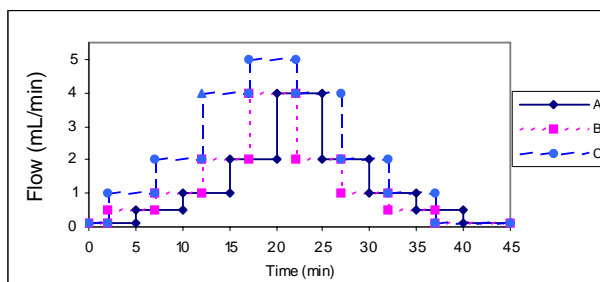
**Figure 9.** Residual curve with 9 effect vectors displaying only noise and an acceptable drift. It can therefore be concluded that Archimedes can compensate for temperature induced effects which are much larger than what can be expected during an LC run.

### 3. 2. 2 Flow/Pressure

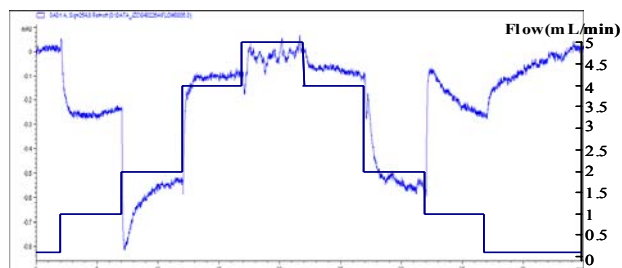
The purpose was to investigate if Archimedes could compensate for RI-and pressure effects caused by flow rate changes. The baseline disturbances were generated by running a flow program with a 2 m x 0.12 mm i.d. capillary placed between the autosampler and the DAD. By programming three different flow programs (Figure 10) it is thereby possible to investigate if shifts in time (dotted line with squares in Figure 10) as well as in flow (dotted line with triangles in Figure 10) can be compensated.

The raw signal, with the flow gradient, is seen in Figure 11 and the residual curve for AR 1 after calculating with Archimedes is seen in Figure 12. The residual curve shows that Archimedes can eliminate RI-and pressure effects because it is just showing noise. Once again it should be stressed that the flow rate changes in this experiment are much larger than what can be expected in a conventional LC-analysis. This result suggests that Archimedes can be used to facilitate flow rate programming which today is hampered by baseline disturbances. Flow rate programming in combination with monolithic columns

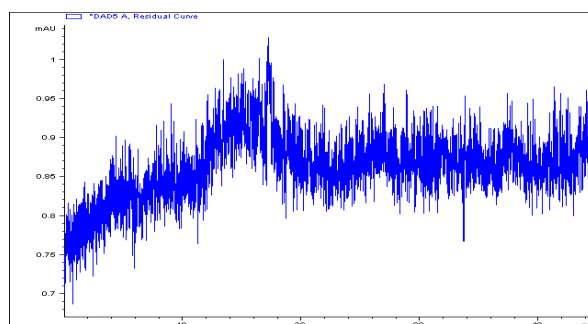
with low pressure drop can be used as an alternative to gradient chromatography. Since no re-equilibration is needed between runs, cycle time can be reduced.



**Figure 10.** Flow programs A (line with diamonds which is the calibration run), B (dotted line with squares which is the AR 1) and C (dotted line with triangles which is the AR 2).



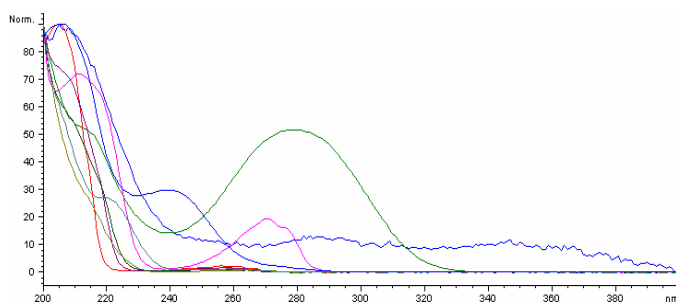
**Figure 11.** AR 2 at 254 nm with the flow gradient C (for conditions see Section 2. 5).



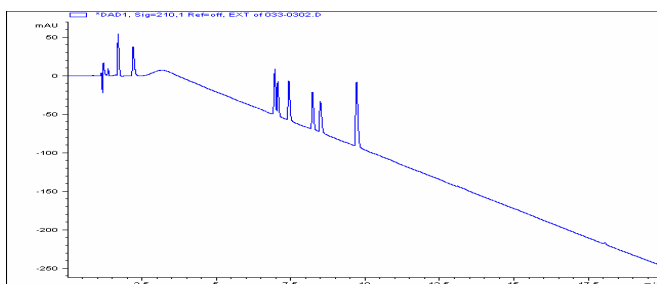
**Figure 12.** Residual curve with 8 effect vectors. No trends can be seen and thus it can be concluded that Archimedes can compensate for RI-and pressure induced effects.

### 3. 2. 3 A more general detection-“concentration” chromatograms

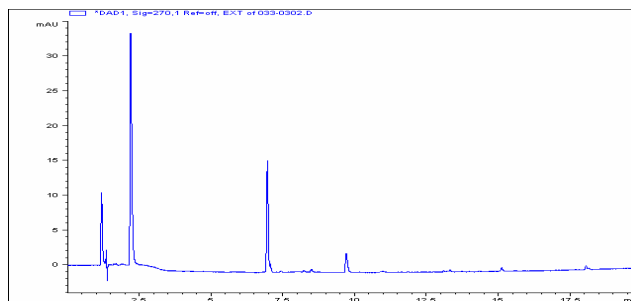
The purpose was to investigate if Archimedes could give a more general detection than conventional UV detection. According to the theory behind Archimedes the concentration profile should provide a wavelength independent signal. In this study 9 bases (nicotine, benzylamine, procainamide, ARD12495, diphenhydramine, remacemide, nortriptyline, bambuterol and phenol) with large differences in spectra were used as model substances (Figure 13). These spectra and the extracted chromatograms clearly shows that the response of the analytes vary significantly for different wavelengths (Figures 14a-b). Figure 15 shows the superimposed concentration profiles for all analytes. A comparison with Figures 14a-b and 15 clearly illustrates that the concentration profile provides a more general and wavelength independent detection. It should be stressed, however, that it is not a true concentration profile; it is a profile that is proportional to the concentration of one analyte. For example, two analytes present at the same molar concentration will not display the same area unless they have identical spectra.



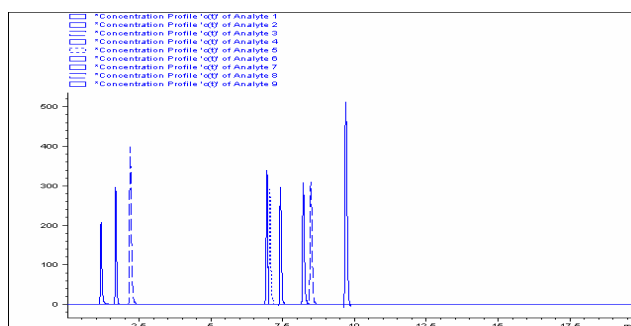
**Figure 13.** Spectra for 9 bases (nicotine, benzylamine, procainamide, ARD12495, diphenhydramine, remacemide, nortriptyline, bambuterol and phenol) where each analyte has its own wavelength maximum.



**Figure 14a.** Extracted chromatogram at 210 nm from original DAD matrix, where the analytes have different responses compared to another wavelength (see Figure 14b).



**Figure 14b.** Extracted chromatogram at 270 nm from original DAD matrix, where the analytes have different responses compared to another wavelength (see Figure 14a).



**Figure 15.** Concentration profile from Archimedes where all analytes can be seen if they can be detected at any wavelength. This concentration profile shows that Archimedes can reduce the wavelength independence and give a more general detection.

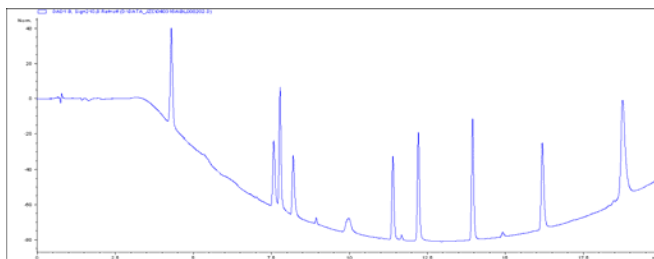
### 3. 2. 4 UV-absorbing buffers

In order to achieve a good performance with LC-MS it is necessary to use volatile buffers. Unfortunately such buffers absorb UV-light and consequently it is difficult to obtain good performance for UV detection when doing gradient LC-DAD-MS. The purpose, with this experiment, was to investigate if Archimedes could compensate for the different types of sloping baselines that UV-absorbing buffers give rise to when doing gradient chromatography. Three frequently used volatile buffers were investigated, 10 mM formic acid (pH 2.9), 10 mM trifluoroacetic acid (TFA, pH 2.1) and 20 mM ammonia (pH 10.8). The sample used for the analytical runs consisted of a series of pharmaceutical bases (see Section 3. 2. 3).

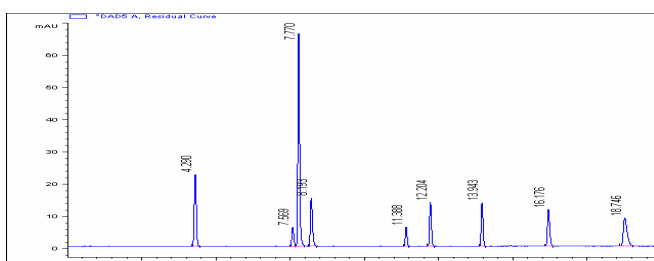
#### 3. 2. 4. 1 Ammonia

The raw signal is seen in Figure 16 and the residual curve after calculating with Archimedes is seen in Figures 17a-b. The plot is just showing noise,

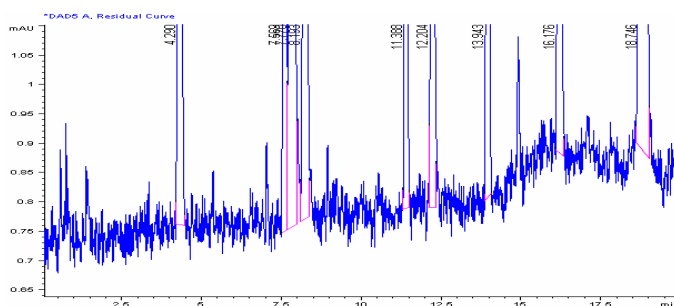
which shows that Archimedes is able to compensate for the background generated by ammonia.



**Figure 16.** AR 1 at 210 nm of the base mixture (see Section 3. 2. 3) with ammonia (for conditions see Section 2. 5).



**Figure 17a.** Residual curve which shows that Archimedes can compensate for the UV-absorbing buffer ammonia.

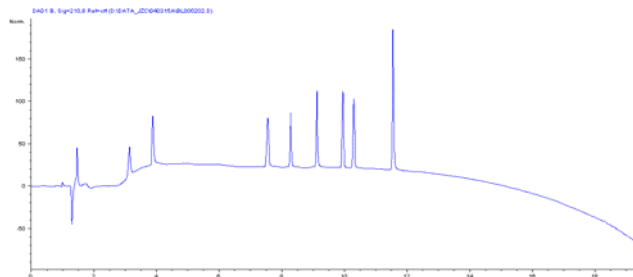


**Figure 17b.** Baseline of the residual curve in Figure 17a, which is just showing random noise, i.e. what is left after the subtraction of the background from the AR.

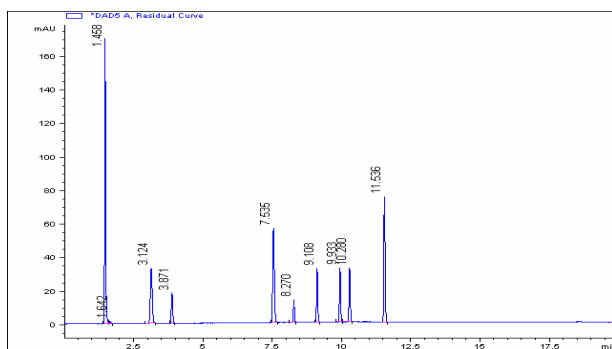
#### 3. 2. 4. 2 Trifluoroacetic acid

The raw signal is seen in Figure 18 and the residual curve after calculating with Archimedes in Figures 19a-b. As can be seen from these plots Archimedes is able to compensate also for TFA. A residual background can be observed but it is not large enough to cause problems.

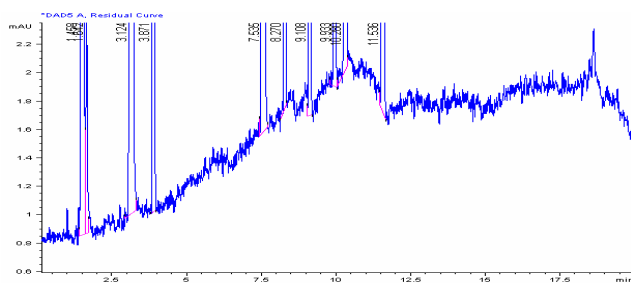




**Figure 18.** AR 1 at 210 nm with TFA (for conditions see section 2. 5).



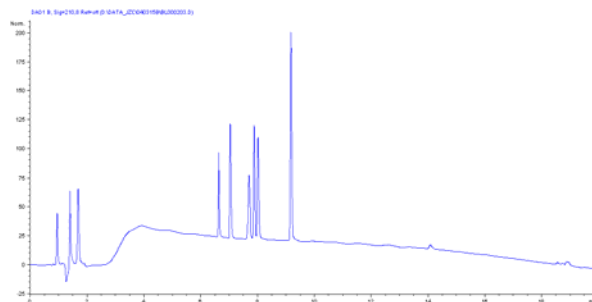
**Figure 19a.** Residual curve which shows that Archimedes is able to compensate for TFA.



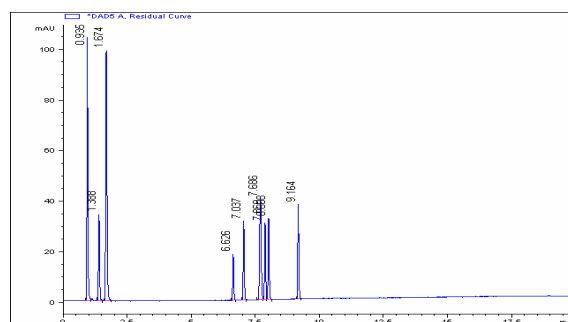
**Figure 19b.** Baseline of the residual curve in Figure 19a.

### 3. 2. 4. 3 Formic acid

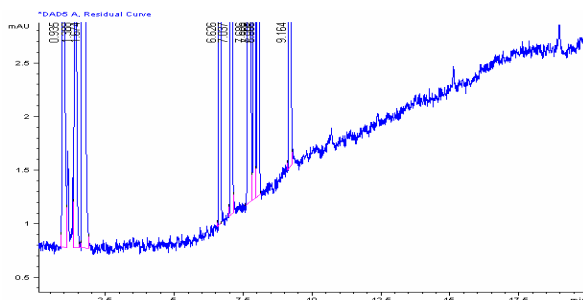
The raw signal is seen in Figure 20 and the residual curve after calculating with Archimedes is seen in Figures 21a-b, which shows that Archimedes can compensate also for formic acid.



**Figure 20.** AR 1 at 210 nm with formic acid (for conditions see Section 2. 5).



**Figure 21a.** Residual curve which shows that Archimedes can compensate for formic acid.



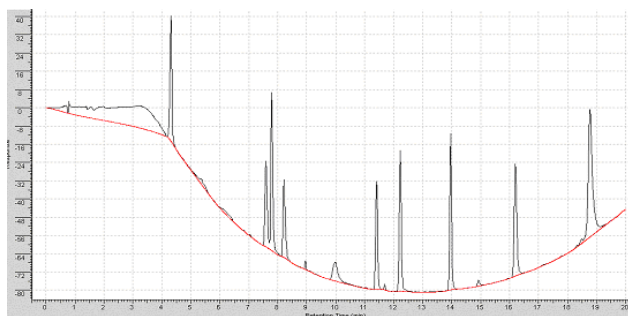
**Figure 21b.** Baseline of the residual curve in Figure 21a.

### ***3. 2. 5 A comparison with other methods for background subtraction***

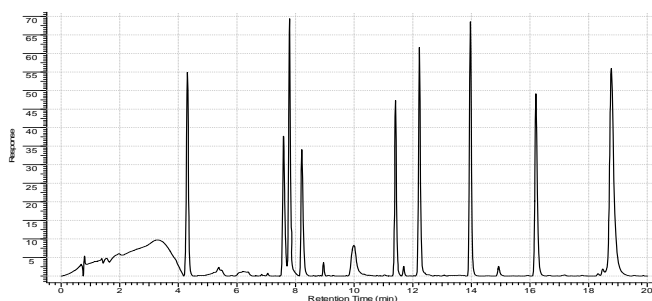
Another software that can be used to subtract background, ACD, was compared with Archimedes by using the UV-absorbing buffers data. ACD does not work like Archimedes, it is based on that a polynomial is fitted to the regions of the chromatogram that is considered to represent the baseline. An automatic fitting did not work very well as can be seen in Figures 22a-22b, which shows the results for ammonia (similar results or worse were obtained also for TFA and formic acid). However, there is another function, manual adjustment of the nodes, where the nodes of the baseline are moved to show

the software where the baseline is. This manual function works much better than the automatic one and it gives a nice straight baseline (Figure 23). Unfortunately, this manual function takes a lot of time and it is hard to use because you do not know if a peak is an analyte or not. Comparing the two Figures 23 and 17a it is showed in Figure 23, result from ACD, that there is an extra peak (encircled), which is a peak present in the blank. But in Figure 17a, result from Archimedes, this peak is not present because Archimedes compensates for all signals present in the blank run. Then why just not use ChemStation's own function 'Subtract blank run' (a function present in all chromatographic softwares)? Figure 24a shows one chromatogram before and 24b shows the result after subtraction of the corresponding blank run. By looking at these figures gives the answer to the question, the function is not used because it does not give a straight baseline, it is very uneven.

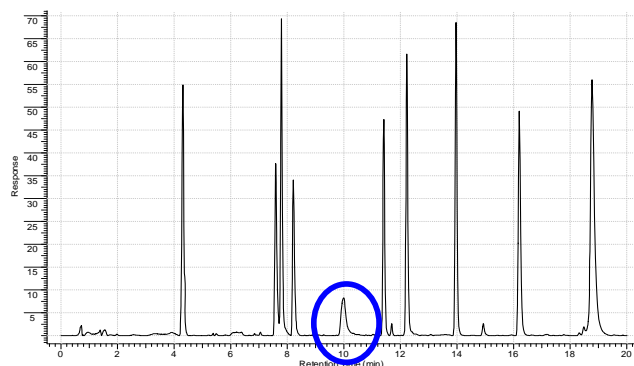
Another advantage with Archimedes is that the calculation goes automatically and a more general detection can be obtained due to the wavelength independent concentration profile. Also, as mentioned above, Archimedes takes the blank run into consideration and eliminates the ghost peaks.



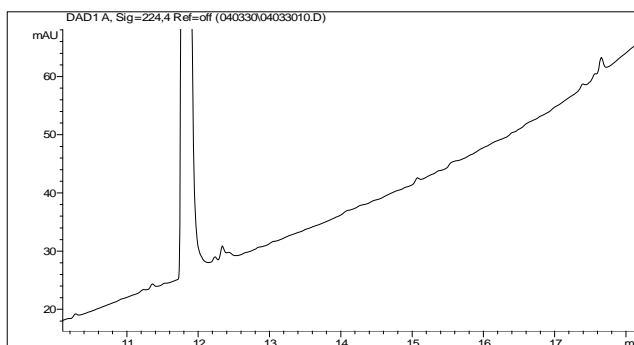
**Figure 22a.** Shows how ACD's, a software that gives straight baselines, automatic mathematical polynomial function fits the baseline.



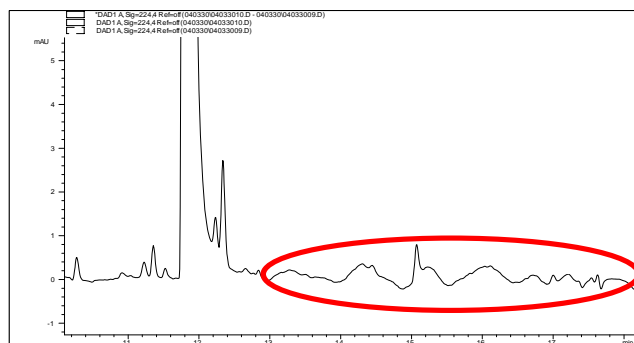
**Figure 22b.** The result after the automatic fit from figure 22a. It can be seen that it does not work very well.



**Figure 23.** Shows the result after the manual adjustment of the nodes with ACD. Here the results are much better compared to the automatic function (Figure 22). This function, however, takes a lot of time and it is hard to predict if a peak is an analyte or not. An example of this is the en-circled peak, which is from the blank.



**Figure 24a.** Original data for the sample X (for conditions see section 3. 2. 6), where the baseline is drifting.



**Figure 24b.** The result after using ChemStation's function Subtract blank run which does not work very well. The baseline is uneven and not straight at all.

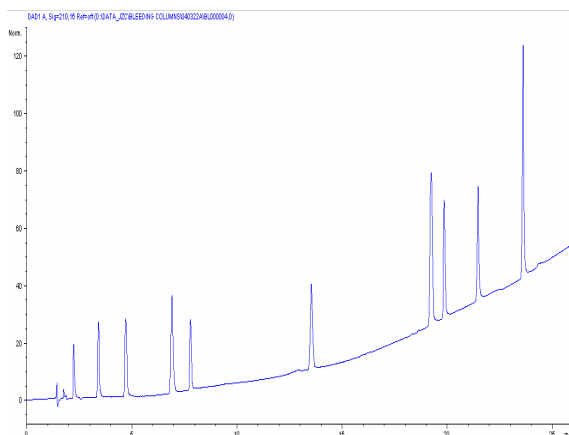
### 3. 2. 6 Bleeding columns

The purpose was to evaluate if Archimedes could compensate for bleeding columns with different functionality. The baseline disturbances were generated by running the columns Symmetry C18, Discovery HS F5 (fluorated phenyl) and Polaris Amide C18 (polar embedded i.e. amide-C18) with a base mixture containing nicotine, benzylamine, procainamide, ARD12495, diphenhydramine, remacemide, nortriptyline, terbutaline, salbutamol, ARC68397 and phenol.

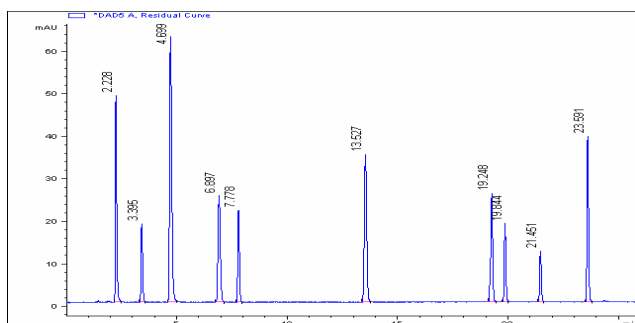
Generally C18 columns display a low background both with UV and MS detection. So called polar embedded columns, i.e. columns with amide, carbamide or urea between the silica and the alkyl chain, display a somewhat higher background with UV detection. With MS detection based on electrospray ionisation and in positive mode these columns does, however, display a significant background due to the polar moiety. Phenyl columns probably display the most severe bleeding that can be observed with UV-detection for any RPLC-column. This is due both to a low stability and the good UV-absorbance of the phenyl moiety.

#### 3. 2. 6. 1 Symmetry C18

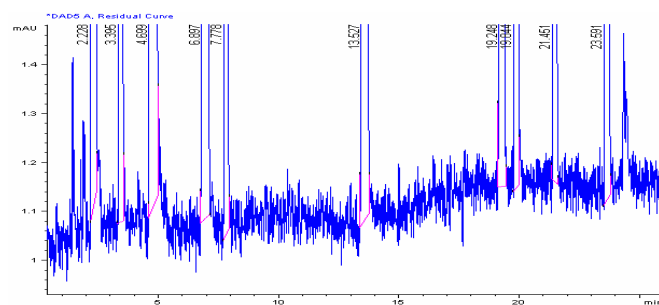
The raw signal is seen in Figure 25 where it can be seen that the column bleeds around 50 mAU, which is normal for a C18-column. The residual curve after calculation with Archimedes is seen in Figures 26a-b, which indicates that Archimedes is able to compensate for the bleeding generated by Symmetry C18.



**Figure 25.** AR 1 at 210 nm with the column Symmetry C18 (for conditions see Section 2. 5).



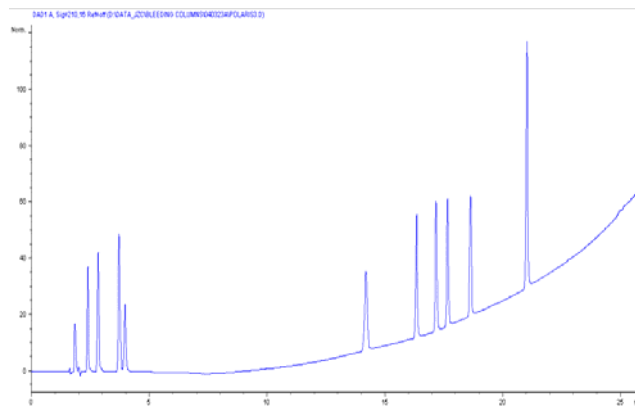
**Figure 26a.** Residual curve, which indicates that Archimedes can compensate for the bleeding column Symmetry C18.



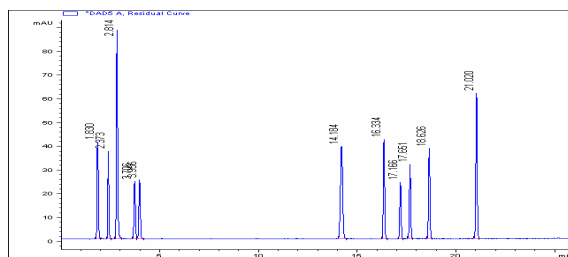
**Figure 26b.** Baseline of the residual curve in Figure 26a, which is showing random noise.

### 3. 2. 6. 2 *Polaris Amide C18*

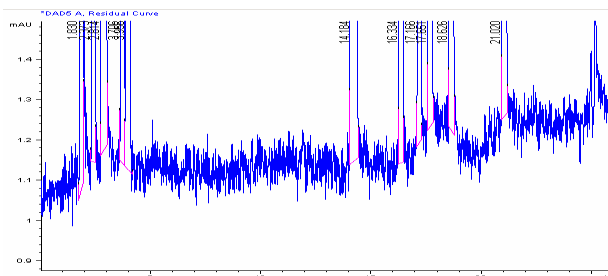
The raw signal is seen in Figure 27, where it also can be seen that the column bleeds around 70 mAU due to the amide. Residual curve after calculation with Archimedes is seen in Figures 28a-b, which indicates that Archimedes can compensate for Polaris Amide C18, because it is showing random noise.



**Figure 27.** AR 1 at 210 nm with the column Polaris Amide C18 (for conditions see section 2. 5).



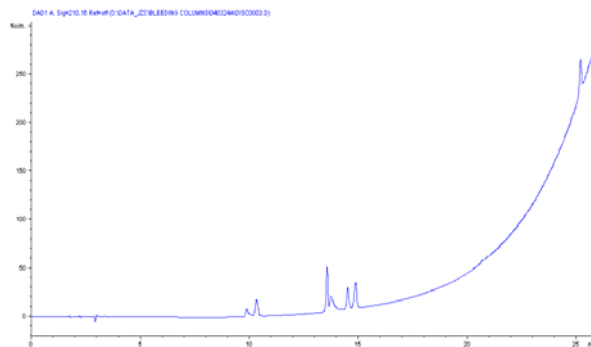
**Figure 28a.** Residual curve, which indicates that Archimedes can compensate for the bleeding column Polaris Amide C18.



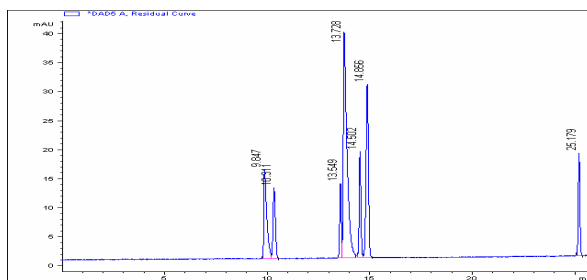
**Figure 28b.** Baseline of the residual curve in Figure 28a, which is showing random noise.

### 3. 2. 6. 3 Discovery HS F5

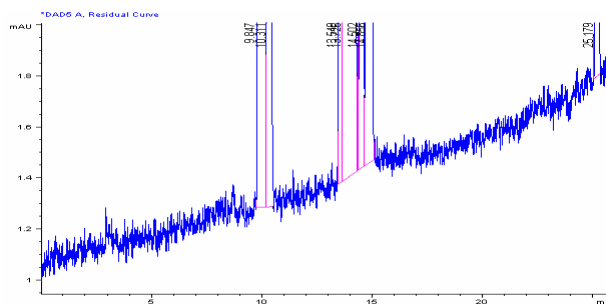
The raw signal is seen in Figure 29, where it can be seen that the column bleeds around 300 mAU due to the fluorated phenyl. Residual curve after calculation with Archimedes is seen in Figures 30a-b. This plot shows that Archimedes can compensate for the bleeding column Discovery HS F5. The residual background is relatively small and not considered to be a problem.



**Figure 29.** AR 1 at 210 nm with the column Discovery HS F5 (for conditions see section 2. 5).



**Figure 30a.** Residual curve which shows that Archimedes can compensate for the bleeding column Discovery HS F5.



**Figure 30b.** Baseline of the residual curve in Figure 30a.

### 3. 2. 7 Improved quality of UV spectra-“error-free spectra”

The purpose was to evaluate if Archimedes could give improved quality of spectra. This was investigated by taking spectra for the analyte at the retention time 12.9 min in a degraded sample X, which has been exposed to light, 1.2 MLuxh (corresponds to two months on a lab bench), and comparing the spectra without (manual, automatic and no reference) and with Archimedes (Figure 31). Spectra for the analyte at the retention time 12.9 min without Archimedes is seen in Figure 32 and the spectra with Archimedes is seen in Figure 33. As can be seen from the figures there is not any significant difference between the two spectra no matter what kind of reference is selected. One advantage with Archimedes is, however, that no reference spectrum has to be selected.

The second experiment was to evaluate the same data but now for an analyte with the retention time 13.4 min, that co-elutes with a component present in the blank (13.5 min), and comparing the spectra before (with automatic and manual reference) and after Archimedes (Figure 31).

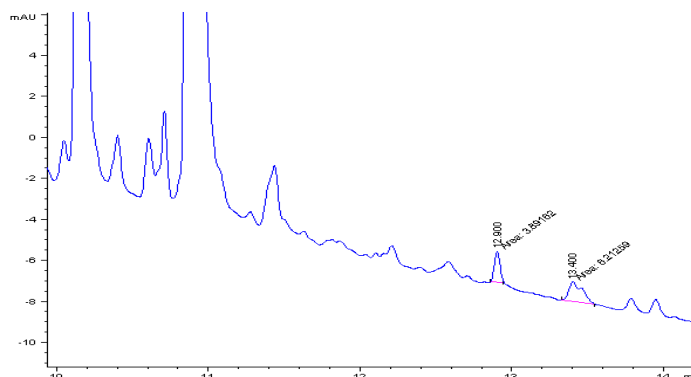
When investigated partially co-eluting peaks (Figure 34) the current automatic selection of reference spectra in ChemStation caused problems. A manual selection is necessary in order to get good results. Figure 36 gives the



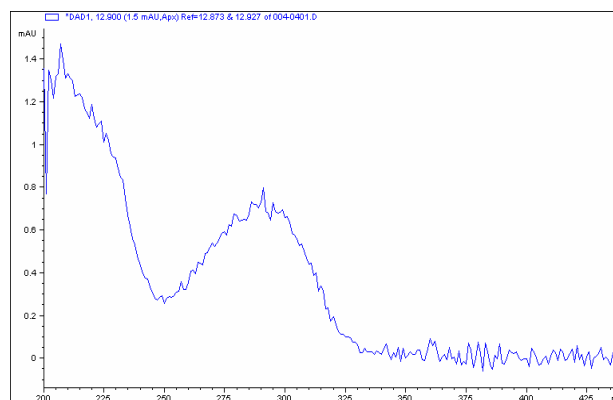
impression that Archimedes provides both a better signal-to-noise ratio (S/N) and spectral quality for overlapping peaks, in comparison with spectra without Archimedes (Figure 35).

In order to further investigate if this really is the case a more controlled experiment was conducted by running a PSR containing a “known ghost peak” (ARC68397) and then two analytical runs with the ghost peak separated from and co-eluting with an analyte (ARD12495) (Figures 41 and 43).

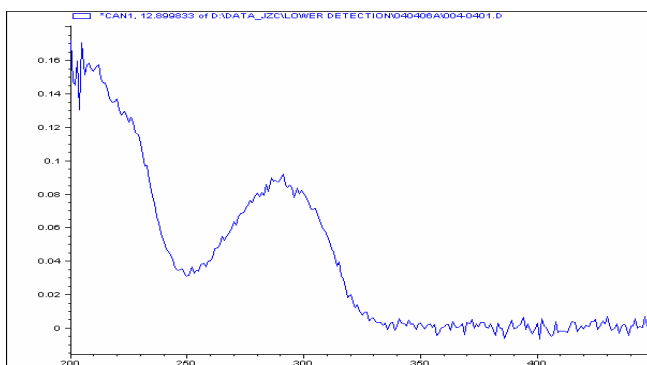
Figure 37 shows an overlay of the spectra for the two peaks when they are separated, overlapping and calculated from Archimedes. As can be seen from this figure any conclusions are hard to do because all three spectra differ. Archimedes is able to eliminate the ghost peak, however, the spectra for the analyte is affected for co-eluting peaks.



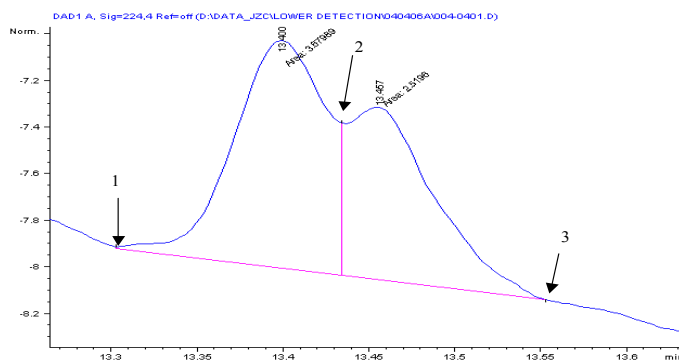
**Figure 31.** Original UV-signal where the co-eluting peak at 12.9 and 13.4 min is seen.



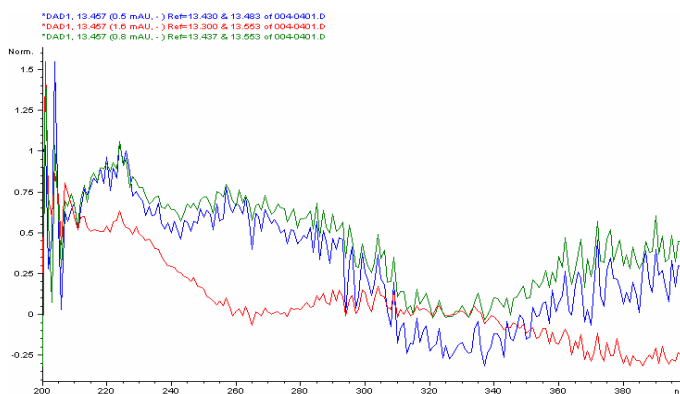
**Figure 32.** Spectra for the peak at  $t_r = 12.9$  min



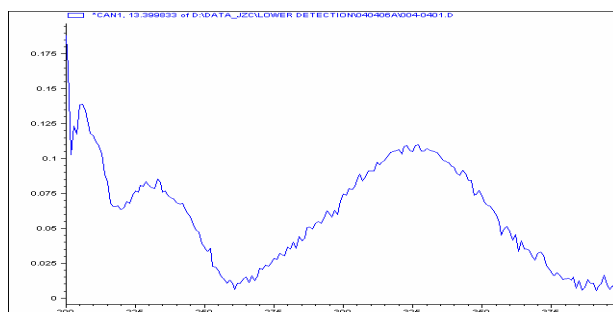
**Figure 33.** Spectra for the peak at  $t_r = 12.9$  min using Archimedes and comparing with Figure 32 and it is seen that there is not any difference between the spectra.



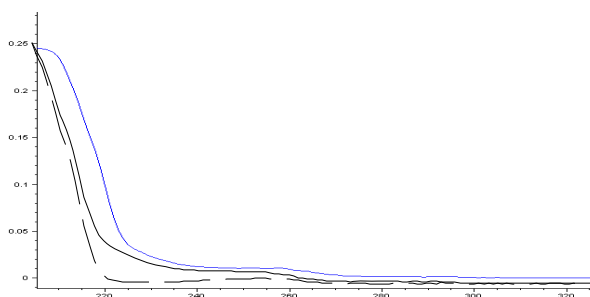
**Figure 34.** Chromatogram of overlapped peaks at  $t_r = 13.4$  min (ghost peak) and 13.5 min (analyte) detected at 224 nm (for conditions see section 2. 5).



**Figure 35.** Spectra with manual reference (green), is taken at arrows 2 and 3 and red is taken at arrows 1 and 3 in Figure 34 and automatic reference (blue) for the two co-eluting peaks at 13.4 and 13.5 min.



**Figure 36.** Spectra for the analyte at  $t_r = 13.4$  min using Archimedes. The result gives the impression that Archimedes provides both a better S/N and spectral quality for overlapping peaks.



**Figure 37.** Overlay of spectra for two separated peaks (lower dotted line), overlapped peaks (middle line) and spectra calculated from Archimedes (the upper line). These results indicate that Archimedes has some problems with overlapped peaks because all three spectra differ.

### 3. 2. 8 Elimination of ghost peaks

The purpose was to see if Archimedes could eliminate ghost peaks present in the PSR and separate the analyte from it during the analytical run, a very common problem when doing gradient chromatography. This was done by running sample X (Section 3. 2. 6), where the impurities are present at 0.1% level, and then evaluating the accuracy and repeatability for three peaks at the retention time 12.6, 12.9 and 13.4 min in 10 runs (see Figure 38).

Peak areas were calculated for the UV-signal and the extracted corrected analyte signals. A comparison between the peak areas of the UV-signal and the corresponding extracted corrected analyte signal gives equivalent results (Table 1). Also for the peak at 12.9 min, which is not co-eluting with any blank peaks, equivalent results are obtained. The co-eluting peaks at 12.6 min

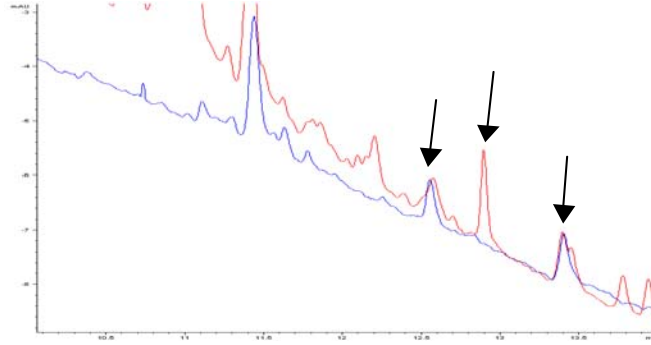
show a significantly lower peak area in the extracted corrected analyte signal. This indicates that Archimedes compensates for the blank peak. For the co-eluting peak at 13.4 min no significant difference can be observed and thus it seems like the elimination of the blank peak does not work.

An investigation of the difference between the UV traces of the average analytical run (blank peak and analyte peak integrated as one peak) and the average pure solvent run (Table 1) shows a good accuracy for the peak at 12.6 min whereas Archimedes overestimates the peak at 13.4 min. A closer look at the extracted corrected analyte signal (Figure 40) shows that the peak at 13.4 min have strange peak shape with perpendicular drop to the baseline at the start and end.

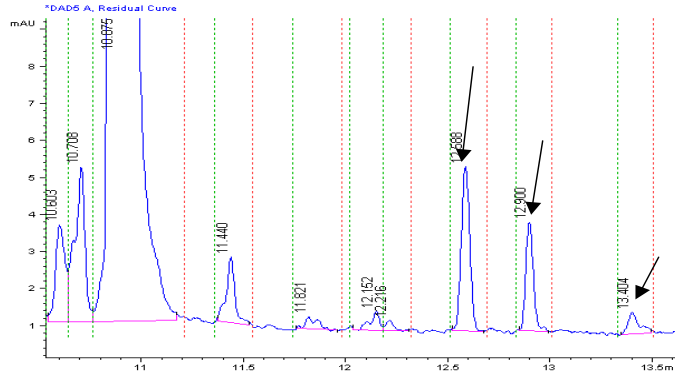
A comparison of the obtained relative standard deviation (Table 1) shows that the same RSD is obtained for the main peak in both UV-signal and the extracted corrected analyte signals. This RSD of 0.3% corresponds to the injection repeatability of the autosampler. The peak at 12.6 min which Archimedes manage to eliminate the blank peak for, displays a significant improvement, the RSD is reduced from 27% for the UV trace to 5% for the extracted corrected analyte signal. For the peak which Archimedes cannot eliminate the blank peak, the RSD increases from 5 to 11%. The peak, which does not co-elute at 12.9 min display the same RSD in both traces.

A more controlled experiment was conducted by running a PSR containing a “known ghost peak” (ARC68397) and then two analytical runs with the ghost peak separated from and co-eluting with an analyte (ARD12495) (Figures 41-44). Peak areas were calculated for the separated and the partially overlapped peaks with and without Archimedes.

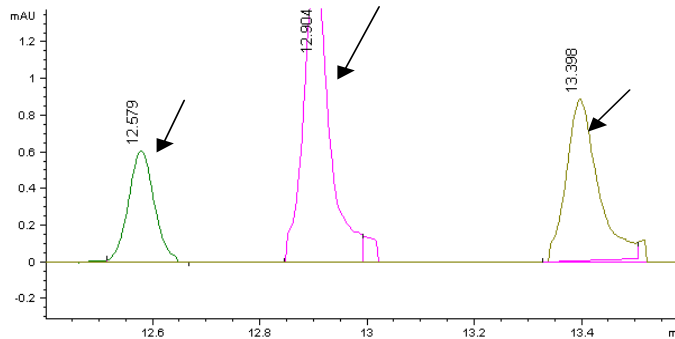
Figures 42 and 44 shows that Archimedes is able to eliminate ghost peaks and system peaks. But if one takes a closer look on the peak areas in Figures 42 and 44, it shows that Archimedes has some problems with the elimination. The total area for the partially overlapped peaks (Figure 43) is 311 mAU and it agrees well with the total area of the separated peaks (Figure 41), also 311 mAU. But the peak area for the extracted corrected analyte signals is 151 mAU, when the peak is separated (Figure 42) and 184 mAU, when it is overlapped (Figure 44). This big difference between the areas clearly shows that Archimedes has a serious problem with area and spectra for co-eluting peaks (see section 3. 2. 6).



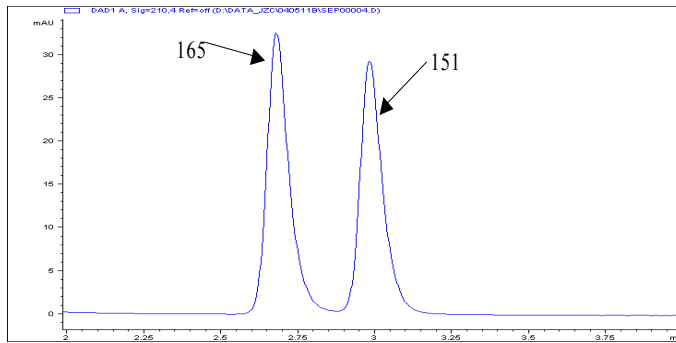
**Figure 38.** Overlay of PSR 2 (lower line) and AR 1 (upper line) detected at 224 nm (for conditions see section 2. 5) where the peaks at  $t_r = 12.6$  and 13.4 min are co-eluting with peaks present in the PSR.



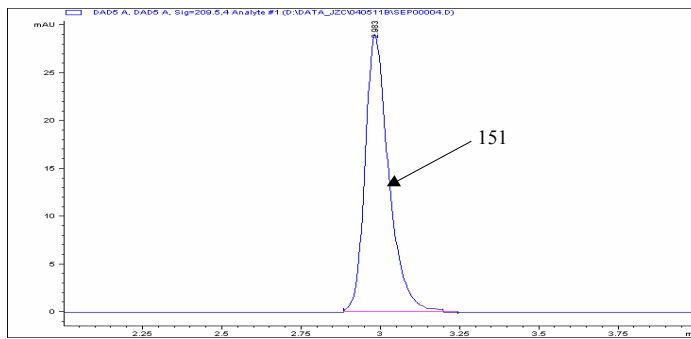
**Figure 39.** Residual curve from Archimedes where the ghost peaks from the PSR are eliminated (compare with Figure 38).



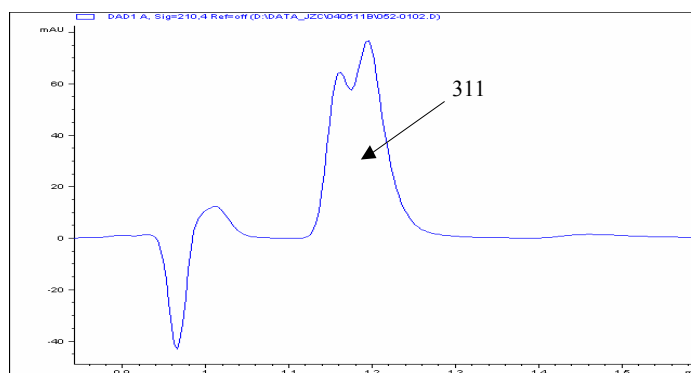
**Figure 40.** Extracted corrected analyte signals at 224 nm, where it is seen that the peak at 12.6 min is a “normal” peak. The peak at 13.4 min, however, has a strange peak shape and a baseline that is elevated resulting in an overestimated area.



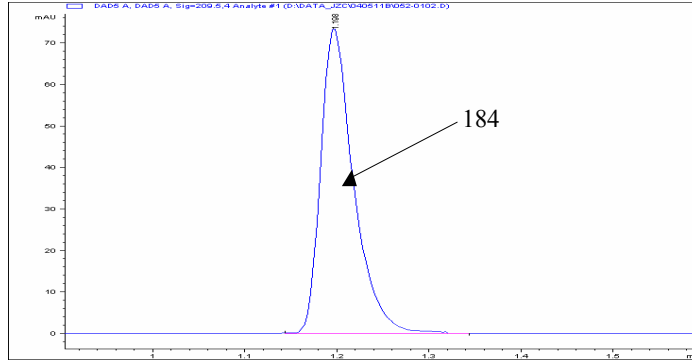
**Figure 41.** Original UV-signal at 210 nm where the first peak is the “known” ghost peak and the last peak is the analyte (for conditions see section 2. 5).



**Figure 42.** Extracted corrected analyte signal for Figure 41. Here it shows that the ghost peak is eliminated and that the peak area corresponds well to the original.



**Figure 43.** Original UV-signal at 210 nm where the analyte peak is co-eluting with a ghost peak present in the PSR (for conditions see section 2. 5).



**Figure 44.** Extracted corrected analyte signal for the co-eluting analytes in Figure 43. Also here it can be seen that Archimedes is able to eliminate ghost peaks and system peaks. The peak area do not, however, correspond to the original peak area. This indicates that Archimedes has a problem with areas for co-eluting peaks.

Table 1. Peak areas [mAU*s]									
	tr =10.8	tr =12.6	tr =12.9	tr =13.4		tr =10.8	tr =12.6	tr =12.9	tr =13.4
UV- signal at 224 nm	3282.8	4.8	5.1	3.5	Archimedes extr.correc.anals.signals	3295.1	1.7	5.4	2.8
	3260.6	4.0	5.1	3.7		3274.1	1.8	5	3.8
	3272.5	3.8	5.0	3.4		3283.6	1.7	5.1	3.7
	3272.7	3.1	4.8	3.5		3286.5	1.8	5.2	3.5
	3287.0	3.1	4.7	3.3		3300	1.7	5.4	3.7
	3272.1	2.9	5.0	3.4		3283.4	1.9	5.5	3.2
	3263.4	2.1	4.7	3.6		3275.9	1.8	5.1	2.9
	3283.8	2.4	5.2	3.1		3299	1.9	5.3	3.8
	3277.1	2.9	4.8	3.2		3291.7	1.7	5.1	3.7
	3263.6	2.3	5.0	3.5		3279.6	1.7	5.3	3.7
<b>Average</b>	3273.6	3.1	5.0	3.4	<b>Average</b>	3286.9	1.8	5.2	3.5
<b>Std dev</b>	9.2	0.8	0.2	0.2	<b>Std dev</b>	9.2	0.1	0.2	0.4
<b>RSD</b>	0.3	26.6	3.4	5.0	<b>RSD</b>	0.3	4.7	3.1	10.8

Table 1 continued								
Blank	tr =12.6	tr =13.4	AR (blank peak and analyte peak integrated as one peak)	tr =12.6	tr =13.4	Difference	tr =12.6	tr =13.4
UV- signal at 224 nm	1.0	3.9	UV- signal at 224 nm	4.0	6.4	<b>Average</b>	1.7	1.9
	2.9	4.1		3.7	6.1			
	1.1	4.0		3.7	5.7			
<b>Average</b>	1.6	4.0		3.6	6.1			
<b>Std dev</b>	1.1	0.1		3.3	6.0			
<b>RSD</b>	65.9	3.0		3.3	6.1			
				3.0	6.2			
				3.2	5.4			
				3.1	5.4			
				2.9	5.6			
			<b>Average</b>	3.4	5.9			
			<b>Std dev</b>	0.4	0.3			
			<b>RSD</b>	10.5	5.9			



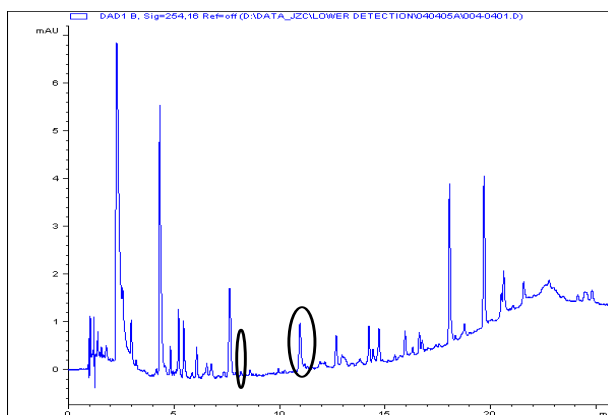
### 3. 2. 9 Improvement of signal-to-noise ratio

The purpose was to investigate if Archimedes could be used to improve the S/N and thereby give lower detection limits. Using sample X (see section 3. 2. 6) and by comparing the UV-signal and residual curve investigated this.

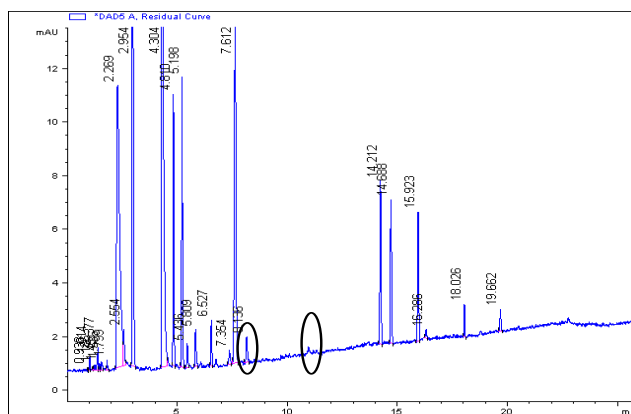
In order to generate an extracted corrected analyte signal (or concentration signal) with Archimedes it is necessary to detect the analyte in the residual curve and subsequently extract its spectrum. Thus the S/N of the residual curve defines the detection limit.

Archimedes has the advantage of a reduced wavelength dependence, i.e. a more general detection, which facilitates detection in comparison to detection at a single wavelength. On the other hand the residual plot displays the residual square sum of the whole wavelength range and this results in a relatively high noise level. One can therefore expect to find both better and worse detection limits when using Archimedes. This is illustrated in Figures 45-46 below.

If the spectrum of the analyte is known, e.g. in an assay, it is likely that Archimedes could be used in order to reduce the detection limit. This functionality is, however, not implemented in the current version even if it should be possible in theory.



**Figure 45.** UV-signal recorded at 224 nm (for conditions see Section 2. 5). The peak at  $t_r=11$  min is detected in the UV-signal as seen from the figure, which gives a better detection with UV. For the peak at  $t_r=8$  min, however, is detected with Archimedes and gives a worse detection with UV (comparison with Figure 46).



**Figure 46.** Residual curve for sample X, where the peak at  $t_r = 8$  min is detected with Archimedes which shows that Archimedes gives a better detection. For the peak at  $t_r = 11$  min is, however, not detected in the residual curve which gives a worse detection with Archimedes (comparison with Figure 45).

### 3. 2. 10 Influence of number of effect vectors and wavelength range

The purpose was to see if the number of effect vectors and the wavelength range had an influence on the resulting residual curve, e.g. if the noise or the analyte signals are affected. Calculations with increasing wavelength ranges 200-300, 200-400, 200-500, 200-700 and 200-900 nm were done. After finding the optimal range-the lowest wavelength, 200, was changed to 210 and 220 nm. Also here the optimal wavelength was found. Subsequently the number of effect vectors, in this case 12, 13, 14, 15 and 16 were evaluated.

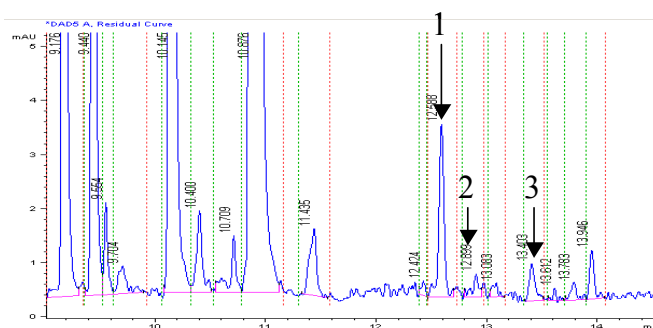
Starting with the influence of the wavelengths, a comparison was made between the residual curves which showed that the interval 200-300 nm does not show all the analytes as for e.g. the residual curve for 200-400 nm or 200-900 nm (see Figures 47-49). The explanation to this is that the analytes absorb above 300 nm (see Figure 52). The same behaviour with the loss of analyte signal is seen when the wavelength is changed from 200 to 210 and 220 nm. These results are what can be expected. It is however more difficult to explain why the analyte with the retention time 12.9 min which has max absorbance below 300 nm should be less affected than the analyte with the retention time 13.4 min which has max absorbance above 300 nm (Figures 47 and 52).

There are no significant differences between the residual curves for 200-400, 200-500, 200-700 and 200-900 nm. It appears, however, that the noise is somewhat lower when reducing the wavelength range. One important advantage with the narrower range is that a significantly shorter time is needed for processing.

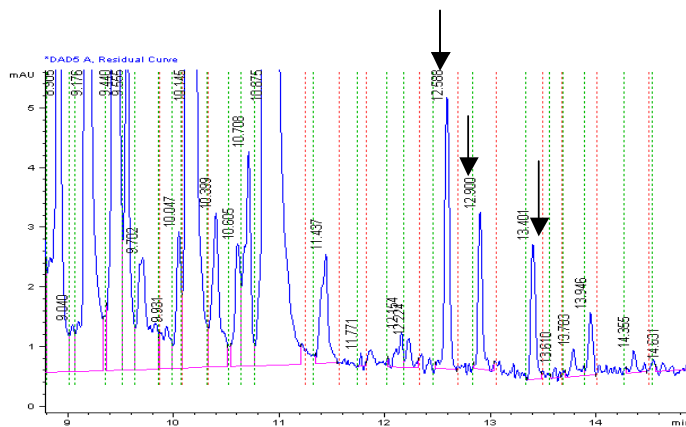
These findings are well in line with the recommendations given for Archimedes; one should choose a start wavelength of 200 nm and take approximately 150 nm broader range than the absorption range of analyte.

A comparison of the influence of number of effect vectors i.e. 12, 13, 14, 15 and 16 do not show any significant differences, except for the analyte with the retention time 11.4 min. At 12 effect vectors it is two overlapped peaks. But when adding more effect vectors the peak gets smaller and turns into one peak at 16 effect vectors (see Figures 50-51). An explanation to this is that 12 effect vectors is not enough to explain the background, it requires 16 effect vectors.

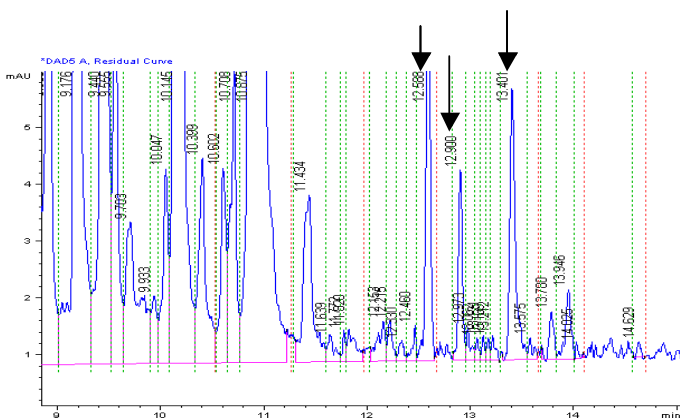
To select the right number of effect vectors manually is difficult. One chemometrical methodology, which could automatize this task, is so called cross validation. Cross validation is used to evaluate which mathematical function that best describes a relationship. It is also used within certain softwares for PCA in order to determine the right number of principal components. Thus it should also be possible to use in order to determine the right number of effect vectors. The principle for cross validation can be illustrated with the determination of which polynomial that best fits a set of experimental data. A higher grade of a polynomial will always give a better fit to the data points (Figure 54). At a certain point one, however, starts to modulate noise and thereby a more unstable and a less robust model is obtained, i.e. we get an overfitting (Figure 53). In cross validation one takes out a data point and calculates the residual between the data point and the fitted polynomial. After this calculation the data point is put back again. This is done with all the data points for all the different polynomials. Subsequently a comparison of the residual sum of squares obtained for the different polynomials will reveal the most appropriate model, i.e. the model with the smallest residual square sum (Figure 54).



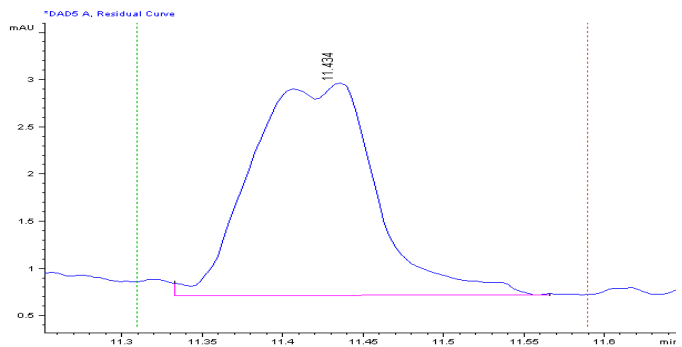
**Figure 47.** Residual curve with the wavelength interval 200–300 nm, where it can be seen, if comparing with Figures 48-49, that some analytes not are detected because it is a too narrow interval, i.e. the analytes absorbs above 300 nm (also see Figure 52).



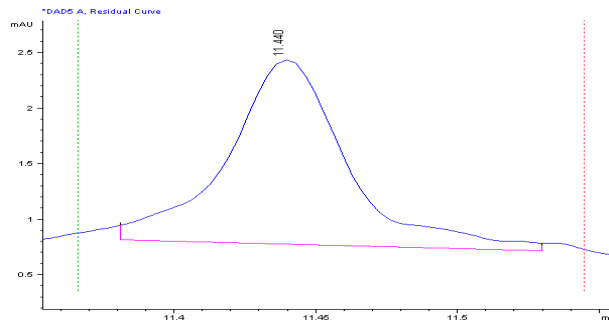
**Figure 48.** Residual curve with the wavelengths 200-400 nm, where all the analytes are detected.



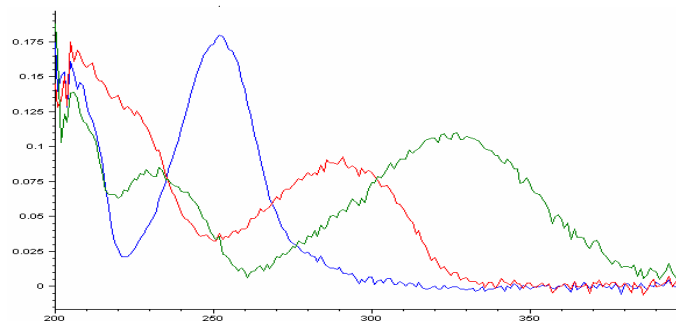
**Figure 49.** Residual curve with wavelengths 200-900nm, where all the analytes are detected.



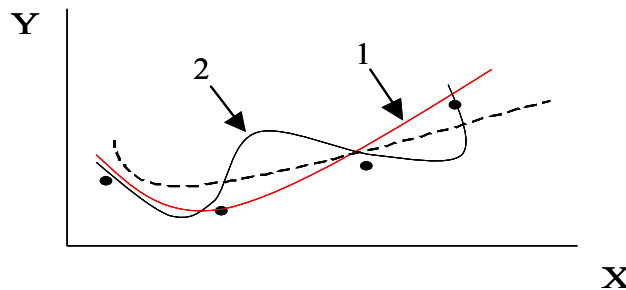
**Figure 50.** Analyte at  $t_r = 11.4$  min in a residual curve with 12 effect vectors. It can be seen that there are two overlapped peaks because the background cannot be totally described with only 12 effect vectors (compare with Figure 51).



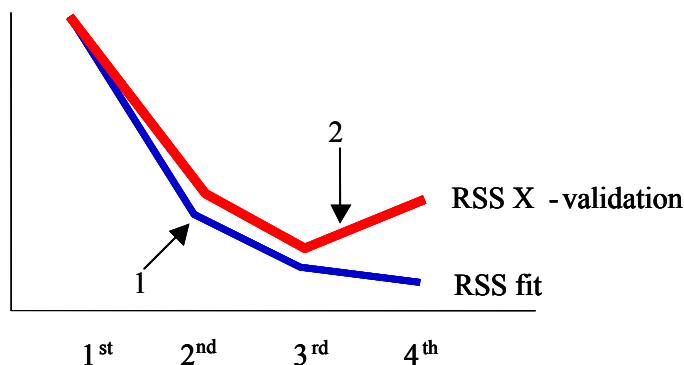
**Figure 51.** Analyte at  $t_r = 11.4$  min in a residual curve with 16 effect vectors, where it is seen that the same peak as in Figure 50 is one. The reason for this is that the whole background is described.



**Figure 52.** Spectra from Archimedes, for the analytes at  $t_r = 12.6$  min (blue line), 12.9 min (red line) and 13.4 min (green line), which absorbs above 300 nm and cannot be detected at a too narrow wavelength interval (see Figures 47-49).



**Figure 53.** An illustration of the problem with overfitting. The red line (1) shows the true function/relationship, the black dots represents the measured values which deviate from the true function due to the experimental noise, the dotted line is a fitted second grade polynomial and the black line (2) is a fourth grade polynomial. The fourth grade polynomial is an overfitting because noise is modulated and thus it deviates from the true model even if it gives a better fit to the experimental values.



**Figure 54.** An illustration of how the residual square sum (RSS) differs for fitting and cross validation. The lower curve (1) shows that a higher grade of polynomial gives a lower residual sum of squares when fitting a polynomial to the experimental data. However, when cross validation is used (upper curve 2), a method used to evaluate how well a model fits the true function, it is seen that an optimal residual sum of square is reached for a lower order polynomial. This lower order polynomial gives the most true and robust model.

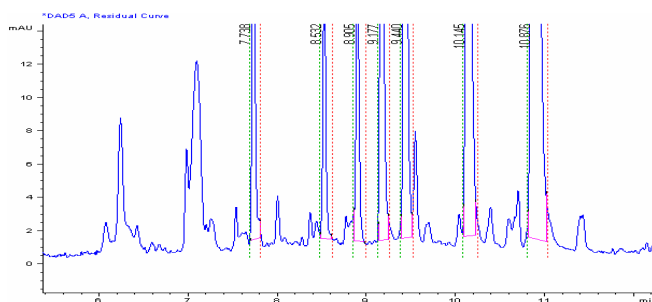
### 3. 2. 11 Does the integration parameters of the residual curve matter?

The purpose was to see if changing the integration parameters of the residual curve had an affect on the result i.e. extracted corrected analyte signals. The data set generated in Section 3. 2. 6 were used for this investigation (sample X).

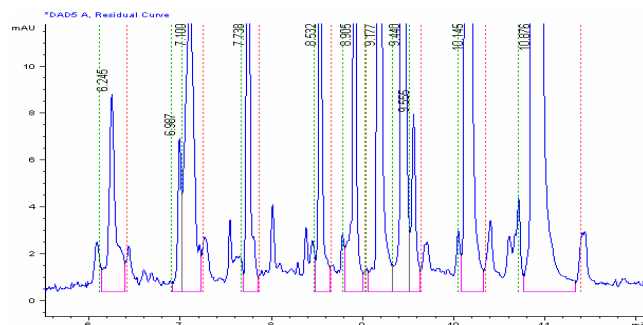
It was necessary to employ quite large differences in the integration parameters in order to obtain differences in the integration as can be seen in Figures 55-56. Thus it seems like the integration of the residual curve is quite robust. As can be seen in Figures 55-57 the start and stop of the peak in the residual curve is reflected in the extracted corrected analyte signal. The reason for this is probably that the start and end time from the integration of the residual curve is used to define the time range for which Archimedes calculates the extracted corrected analyte signal (or concentration profile). This limitation in time has probably been done in order to speed up the calculations. To overcome this problem the peak interval could be extended to +/- 6 standard deviations of the peak and thus also tailing/fronting peaks should be better defined.

The only information that really is needed from the residual curve is where to determine the spectrum of the analyte. Thus the integration does not have to be perfect as long as the peak apex is integrated. If the spectra of the analytes are known before hand it is not necessary at all to integrate the residual curve. The integration of the concentration profile or extracted corrected analyte signal for an analyte is truly automatic, it is not necessary to draw a baseline.

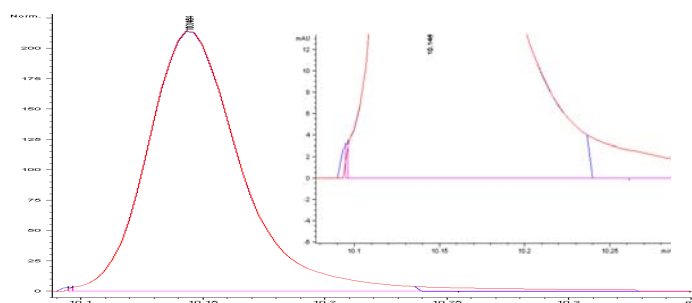
The integration is performed by simply summarizing the signal for each time point (outside the peak the signal is zero by definition).



**Figure 55.** Residual curve with the integration parameters: slope: 200, height: 5, reject area: 5 and width: 0.1, which gives a good integration (for conditions see Section 2. 5).



**Figure 56.** Residual curve with the integration parameters: slope: 1, height: 1, reject area: 1 and width: 0.1, which gives a bad integration (for conditions see Section 2. 5).



**Figure 57.** Overlay of the extracted corrected analyte signals for a good (upper curve) and bad (lower curve) integration of the residual curve (see also Figures 55 and 56). The peaks are identical except in the beginning and end, where the bad integration peak is cut off at the end. The reason is that the definition of the peak start and end is not well defined.

### 3.3 Potential application areas

Potential application areas for Archimedes are 1) LC in general, i.e. elimination of sloping baselines and ghost peaks, a reduced wavelength dependence and a truly automated integration. Because of the automated integration the evaluation time of the data is reduced and the working environment gets better. No more “mouse elbow” due to the automated integration. 2) LC-DAD-MS, i.e. to get good baselines with both UV and MS. Today this is a problem since MS compatible buffers display bothersome UV absorbance, 3) High throughput analysis (HTA), i.e. by co-eluting analytes the analysis time is minimized. The current version of the software cannot do this. In theory it should, however, be possible to do one slow run with well separated standards in order to determine the spectra of the analytes. Subsequently during the analysis the amount of modifier is increased and the analytes are partially co-eluted in order to reduce the cycle time. A data set has been generated (Section 3.2.12 in Appendix 1) but it cannot yet be evaluated. Archimedes also facilitates the use of flow programmed separations as an alternative to gradient separations that does not require re-equilibration. 4) Process Analytical Technology<sup>1</sup> (PAT), i.e. minimal cycle times through co-elution and flow programmed separations in combination with truly automatic integration (the latter of critical importance in a production environment).

### 3.4 “Bugs” and concerns

#### 3.4.1 Bugs

When calculating spectra, sometimes one of two possible error-messages appears: *DTP\_Eval\_1:Parameter 3: A value is outside the legal range* or *DTP\_Eval\_1: Parameter 2: The Table was not found*. This is concerning because if it is not able to calculate the spectra, neither can the extracted corrected analyte signals nor the concentration profiles be extracted.

#### 3.4.2 Concerns

When calculating the extracted corrected analyte signals for overlapped peaks the peak shapes are strange, the obtained spectrum is incorrect and the area does not agree with the original area (Section 3.2.6 and 3.2.7).

---

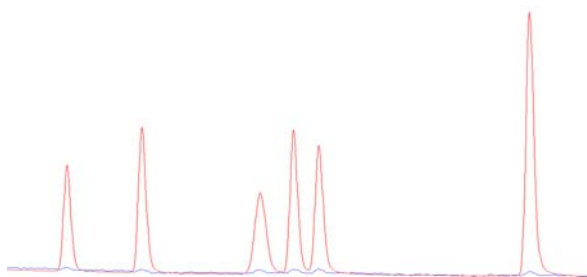
<sup>1</sup> PAT is a system for analysis and control of manufacturing processes based on timely measurements, i.e. during processing, of critical and quality parameters and performance aspects of raw and in-process materials with the goal of ensuring final product quality. PAT also involves optimal application of process analytical chemistry tools, feed back control strategies and information management tools and/or product/process optimization strategies to the manufacture of pharmaceuticals. There are three ways of measurements within PAT, at-line (“lab-wise” measurements conducted in the process hall on samples withdrawn from a manufacturing process), in-line (automated measurements on samples in the process vessel), off-line (traditional where a sample is drawn, and transferred to a distant laboratory) and on-line (automated measurements conducted in the process hall on samples withdrawn from a manufacturing process. Samples may be transferred back to the process after measurement) [22].



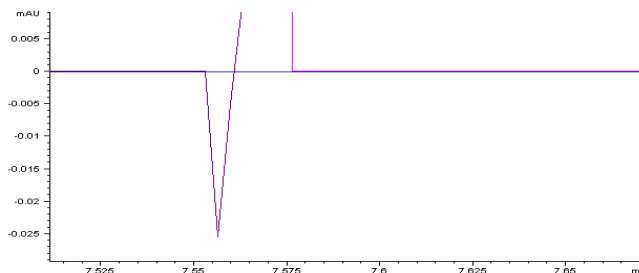
The extracted corrected analyte signals do not give Gaussian peaks for analytes at high retention time, i.e. the baseline seems to be elevated or the background signal not completely eliminated (Section 3. 2. 7).

Background peaks can be observed at the retention of analytes in the extracted corrected background (Figure 58). A comparison between the recorded DAD-signal and the extracted corrected background signal shows a significant difference. This difference seems to depend on the absolute intensity of the residual background signal (Data file: 040315b in Appendix 1).

There is also a problem with the definition of start and end point of the peaks in the extracted corrected analyte signals, i.e. you get a significant perpendicular drop to zero and for some analytes a negative signal at the start of the peak is observed (Figure 59) (Data file: 040315b in Appendix 1).



**Figure 58.** Illustrates the recorded DAD-signal (upper curve) and the extracted corrected background (lower curve) where background peaks can be observed at the retention of analytes in the extracted corrected background (Data file: 040315b in Appendix 1). This difference seems to depend on the absolute intensity of the residual background signal.



**Figure 59.** Illustrates a problem with the definition of start and end point of peaks in the extracted corrected analyte signals, i.e. a negative signal at the start of the peaks is observed (Data file: 040315b in Appendix 1).

## 4 Conclusions

Archimedes has been able to compensate for all the baseline disturbances investigated so far.

It has been shown that the subtract blank run function which is implemented in most chromatographic softwares today not result in an acceptable result (at least not for determination of related substances). It has also been shown that the algorithm implemented in ACD/ChromManager works well for a manual definition and subtraction of a baseline. It is, however, time consuming and cannot eliminate ghost peaks.

Archimedes gives a concentration profile, which makes the analyte response less wavelength dependent, and thus Archimedes provides a more general detection than conventional UV/DAD detection.

The selection of the relevant number of effect vectors is of critical importance. Unfortunately this selection is difficult with the current tools available. A possible solution for a more objective selection, which also can be automatized, would be cross validation.

As can be expected the wavelength range has a large importance. If the spectra of the analytes is known it is easy to select a suitable range. A wide range does not reduce the quality of the data. It does, however, increase the time needed for processing significantly.

The detection limit can both be better and worse when using Archimedes. If the spectrum of the analyte is unknown it is our impression that the S/N will be better in the UV trace. However, the detection limit will be lower in the concentration profile if the analyte has a higher absorbance at another wavelength. If the spectrum of the analyte is known it is likely that the detection limit will be lower with Archimedes.

For well separated peaks no improvement in spectral quality could be observed with Archimedes. For overlapping peaks it seems like Archimedes have advantages to offer. With the current algorithm the accuracy of both spectrum and peak area is, however, not acceptable for such peaks.

The integration parameters of the residual curve do affect the resulting extracted corrected analyte signal as well as concentration profile. This is not acceptable if Archimedes should be claimed to provide truly automated integration. We do, however, believe that this problem is relatively easy to overcome.

## **5 Acknowledgments**

I would like to thank my supervisor Patrik Petersson at AstraZeneca R&D Lund, for giving me the opportunity to work with this project at AstraZeneca, and for all his help through the work and for good supervision.

A big thanks to Aida Vajraca at AstraZeneca R&D Lund, for helping me finding this project.

Also, big thanks to Margareta Sandahl for all her help and answers.

## 6 References

- [1] R. Tauler, D. Barceló, *Anal. Chem.*, 12, (1993), 319-327
- [2] Papers from Agilent inkl. Archimedes Toolbox, (2004)
- [3] A. J. Owen, The diode-array advantage in UV/Visible spectroscopy, *Hewlett-Packard Company*, 8, (1988), 22-30
- [4] W. Lindberg, J. Öhman, S. Wold, *Anal. Chem.*, 58, (1986), 301
- [5] H. H. Willard, L. L. Merritt, Jr., J. A. Dean, F. A. Settle, Jr., *Instrumental methods of analysis 7<sup>th</sup> Ed.*, *Wadsworth publishing company*, 1988, 626-629, 35
- [6] R. G. Brereton, *Chemometrics: data analysis for the laboratory and chemical plant*, *John Wiley & Sons Ltd*, 2003, 184-209, 212-213, 219-223, 284-292, 315-317
- [7] L. Eriksson, E. Johansson, N. Kettaneh- Wold and S. Wold, *Multi- and Megavariate Data Analysis Principles and Applications*, *Umetrics*, 2001, 7-70
- [8] J. N. Miller and J. C. Miller, *Statistics and Chemometrics for Analytical Chemistry 4<sup>th</sup> Ed.*, *Prentice Hall*, 2000, 214-236
- [9] S. Wold, K Esbensen and P. Geladi, *Chemom. Intell. Lab. Syst.*, 2, (1987), 37-52
- [10] K. Wiberg, M. Andersson, A. Hagman, S. P. Jacobsson, *J. Chromatogr. A*, 1029, (2004), 13-20
- [11] W. Lindberg, J. Öhman, S. Wold, *Anal. Chem.*, 58, (1986), 299-303
- [12] M. Josefson and L. Tekenbergs- Hjelte, *J. of Pharmaceutical & Biomedical analysis*, 9, (1991), 643- 654
- [13] R. A. Johnson, D. W. Wichern, *Applied multivariate statistical analysis 5<sup>th</sup> Ed.*, *Prentice Hall*, 2002, 1-10, 354-414
- [14] <http://www.chemweb.com/alchem/articles/985883672445.html>  
(2004-03-15 at 3 p.m)

- [15] M. Maeder, *Anal. Chem.*, 59, (1987), 530-533
- [16] H. Shen, J. Wang, Y. Liang, K. Petersson, M. Josefsson, J. Gottfries, F. Lee, *Chemom. Intell. Lab. Syst.*, 37, (1997), 261-269
- [17] M. Maeder, A. D. Zuberbuehler, *Analytica Chimica Acta*, 181, (1986), 287-291
- [18] H. R. Keller, D. L. Massart, Y. Z. Linag and O. M. Kvalheim, *Analytica Chimica Acta*, 267, (1992), 63-71
- [19] M. Maeder and A. Zilian, *Chemom. Intell. Lab. Syst.*, 3, (1988), 205-213
- [20] C. Xu, Y. Liang, Y. Li and Y. Du, *Analyst*, 128, (2003), 75-81
- [21] D. W. Osten and B. R. Kowalski, *Anal. Chem.*, 56, (1984), 991-995
- [22] <http://www.fda.gov/cder/OPS/PAT.htm>  
(2004-06-10 at 2.20 p.m.)

## Appendix 1: Data files

The parameter settings of the DAD were the same for all the experiments:

Peakwidth	0.05 min
Slit	2 nm
Spectrum store	All
Spectrum range	200- 950 nm
Spectrum step	1 nm
Spectrum threshold	N/A
Lamps	UV and VIS on
Time	Stop time as pump Post time off
Autobalance	Pre run on Post run off
Margin for negative absorbance	1000 mAU

### 3. 2. 1 *Temperature*

Mobile phases: Water and acetonitrile (50:50 v/v)

Flow: 0.40 mL/min

Signals stored at: 210,8 and 254,8 nm (reference off)

Sequence table:

	<b>Data file</b>
PSR 1	040216a\Temp0002
PSR 2	040216a\Temp0003
AR 1	040216a\Temp0004
AR 2	040216a\Temp0005
PSR 3	040216a\Temp0006

### 3. 2. 2 *Flow/Pressure*

Mobile phases: Water

Signals stored at: 210,8 and 254,8 nm (reference off)

Sequence table:

	<b>Data file</b>
PSR 1	040226a\Flow0002
PSR 2	040226a\Flow0003
AR 1	040226a\Flow0004
AR 2	040226a\Flow0005
PSR 3	040226a\Flow0006

### **3. 2. 3 A more general detection-“concentration” chromatograms**

Mobile phases: 10 mM formic acid and acetonitrile. The acetonitrile gradient were 5-95% for 20 min

Flow: 0.64 mL/min

Injection volume: 2 µL

Column temperature: 40 °C

Column: Zorbax SB C18 150 x 3 mm, 3 µm

Signals stored at: 254,8 and 210,8 (reference 300,100)

Sequence table:

	<b>Data file</b>
PSR 1	040209b\031-0101
PSR 2	040209b\032-0201
AR 1	040209b\033-0301
AR 2	040209b\033-0302
PSR 3	040209b\034-0401

### **3. 2. 4 UV-absorbing buffers**

The mobile phases: A) 10 mM formic acid in water and 10 mM formic acid in 95% acetonitrile

B) 10 mM TFA in water and 10 mM TFA in 95 % acetonitrile

C) 20 mM ammonia in water and 20 mM ammonia in 95% acetonitrile

All three gradients were 5-100% acetonitrile for 20 min, post time 10 min.

Flow: 0.64 mL/min

Injection volume: 2 µL

Column temperature: 20 °C

Column: XTerra MS C18 150 x 3.0 mm 3.5 µm

Signals stored at: 254 and 210 nm (reference off).

### 3. 2. 4. 1 Ammonia

Sequence table:

<b>Ammonia (040316a)</b>	<b>Data file</b>
PSR 1	040316a/B1000200
PSR 2	040316a/B1000201
AR 1	040316a/B1000202
PSR 3	040316a/B1000203

### 3. 2. 4. 2 TFA

Sequence table:

<b>TFA (040315a)</b>	<b>Data file</b>
PSR 1	040315a/B1000200
PSR 2	040315a/B1000201
AR 1	040315a/B1000202
PSR 3	040315a/B1000203

### 3. 2. 4. 3 Formic acid

Sequence table:

<b>Formic acid (040315b)</b>	<b>Data file</b>
PSR 1	040315b/B1000201
PSR 2	040315b/B1000202
AR 1	040315b/B1000203
PSR 3	040315b/B1000204

### 3. 2. 6 Bleeding columns

Mobile phases: 20 mM  $\text{KH}_2\text{PO}_4$  (pH 2.7) in water and 20 mM  $\text{KH}_2\text{PO}_4$  (pH 2.7) in methanol/water (65:35 v/v)

Flow: 0.21 mL/min.

The gradient was (with a post time of 15 min):

<b>Time (min)</b>	<b>% <math>\text{KH}_2\text{PO}_4</math> in MeOH/<math>\text{H}_2\text{O}</math></b>
0	5
20	100
25	100
26	5

Injection volume: 1  $\mu\text{L}$



Column temperature: 60 °C.

The three columns were:

- A) Symmetry C18 150 x 2 mm, 3 µm
- B) Discovery HS F5 150 x 2.1 mm, 3 µm
- C) Polaris Amide C18 150 x 2 mm, 3µm

Since these columns have an i.d. of 2 mm the detector was equipped with a semi micro cell to minimize band broadening.

Signals stored at: 210,16 (reference off) and 254,16 nm (reference 450,80 nm)

### 3. 2. 6. 1 Symmetry C18

Sequence table:

	<b>Data file</b>
PSR 1	040322a/BI000002
PSR 2	040322a/BI000003
AR 1	040322a/BI000004
PSR 3	040322a/BI000005

### 3. 2. 6. 2 Polaris Amide C18

Sequence table:

	<b>Data file</b>
PSR 1	040323a/Polaris1
PSR 2	040323a/Polaris2
AR 1	040323a/Polaris3
PSR 3	040323a/Polaris4

### 3. 2. 6. 3 Discovery HS F5

Sequence table:

	<b>Data file</b>
PSR 1	040324a/Disc0001
PSR 2	040324a/Disc0002
AR 1	040324a/Disc0003
PSR 3	040324a/Disc0004

### 3. 2. 7 Improved quality of UV spectra-”error-free spectra”

Mobile phases: 10 mM NH<sub>4</sub>Ac in water and 10 mM NH<sub>4</sub>Ac in acetonitrile/water (95:5 v/v)

Flow: 0.64 mL/min

Injection volume: 7 µL

Column temperature: 36 °C

Column: HyPurity C18 150 x 3 mm, 3  $\mu$ m  
Signals stored at: 224 nm (reference off).

The gradient was:

<b>Time (min)</b>	<b>% NH<sub>4</sub>Ac in acetonitrile/H<sub>2</sub>O</b>
0	5
2.5	5
17.5	100
25.5	100
25.6	5
33	5

Sequence table:

	<b>Data file</b>
PSR 1	040406a/002-0201
PSR 2	040406a/003-0301
AR 1	040406a/004-0401
AR 2	040406a/004-0501
AR 3	040406a/004-0601
AR 4	040406a/004-0701
AR 5	040406a/004-0801
AR 6	040406a/004-0901
AR 7	040406a/004-1001
AR 8	040406a/004-1101
AR 9	040406a/004-1201
AR 10	040406a/004-1301
PSR 3	040406a/005-1401

For the more controlled experiment see Section 3. 3. 8 below.

### ***3. 2. 8 Elimination of ghost peaks***

Mobile phases: 10 mM NH<sub>4</sub>Ac in water and 10 mM NH<sub>4</sub>Ac in acetonitrile/water (95:5 v/v)

Flow: 0.64 mL/min

Injection volume: 7  $\mu$ L

Column temperature: 36 °C

Column: HyPurity C18 150 x 3 mm, 3  $\mu$ m

Signals stored at: 224 nm (reference off).

The gradient was:

Time (min)	% NH <sub>4</sub> Ac in acetonitrile/H <sub>2</sub> O
0	5
2.5	5
17.5	100
25.5	100
25.6	5
33	5

Sequence table:

	Data file
PSR 1	040406a/002-0201
PSR 2	040406a/003-0301
AR1	040406a/004-0401
AR 2	040406a/004-0501
AR 3	040406a/004-0601
AR 4	040406a/004-0701
AR 5	040406a/004-0801
AR 6	040406a/004-0901
AR 7	040406a/004-1001
AR 8	040406a/004-1101
AR 9	040406a/004-1201
AR 10	040406a/004-1301
PSR 3	040406a/005-1401

For the more controlled experiment where a PSR containing a “known ghost peak” (ARC68397) and two analytical runs with the ghost peak separated from and co-eluting with an analyte (ARD12495), had the following conditions:

Mobile phases: For the separated analytes the mobile phases were 65 % of 20 mM KH<sub>2</sub>PO<sub>4</sub> in methanol/water (65:35 v/v) and 35% of 20 mM KH<sub>2</sub>PO<sub>4</sub> in water. But for the partially overlapped peaks the mobile phases were 90% of 20 mM KH<sub>2</sub>PO<sub>4</sub> in methanol/water (65:35 v/v) and 10% of methanol

Analysis time: 20 min

Flow: 0.64 mL/min

Injection volume: 2 µL

Column temperature: 60 °C

Column: Zorbax SB C18 150 x 3 mm, 3 µm

Signals stored at: 210 and 254 nm (reference off)

Sequence table for the separated analytes:

	<b>Datafile</b>
PSR 1	040511b\ Sep00001
PSR 2	040511b\ Sep00002
AR 1	040511b\ Sep00003
AR 2	040511b\ Sep00004
PSR 3	040511b\ Sep00005

Sequence table for the partially overlapped analytes:

	<b>Datafile</b>
PSR 1	040511b\ PSR00001
PSR 2	040511b\ PSR00002
AR 1	040511b\ 052-0101
AR 2	040511b\ 052-0201
PSR 3	040511b\ PSR00003

### **3. 2. 9 Improvement of signal-to-noise ratio**

Mobile phases: 10 mM NH<sub>4</sub>Ac in water and 10 mM NH<sub>4</sub>Ac in acetonitrile/water (95:5 v/v)

Flow: 0.64 mL/min

Injection volume: 7 µL

Column temperature: 36 °C

Column: HyPurity C18 150 x 3 mm, 3 µm

Signals stored at: 224 nm (reference off)

The gradient was:

<b>Time (min)</b>	<b>% NH<sub>4</sub>Ac in acetonitrile/H<sub>2</sub>O</b>
0	5
2.5	5
17.5	100
25.5	100
25.6	5
33	5

Sequence table:

	<b>Data file</b>
PSR 1	040406a/002-0201
PSR 2	040406a/003-0301
AR1	040406a/004-0401
AR 2	040406a/004-0501
AR 3	040406a/004-0601
AR 4	040406a/004-0701
AR 5	040406a/004-0801
AR 6	040406a/004-0901
AR 7	040406a/004-1001
AR 8	040406a/004-1101
AR 9	040406a/004-1201
AR 10	040406a/004-1301
PSR 3	040406a/005-1401

### ***3. 2. 10 Influence of number of effect vectors and wavelength range***

Mobile phases: 10 mM NH<sub>4</sub>Ac in water and 10 mM NH<sub>4</sub>Ac in acetonitrile/water (95:5 v/v).

Flow: 0.64 mL/min

Injection volume: 7 µL

Column temperature: 36 °C

Column: HyPurity C18 150 x 3 mm, 3 µm

Signals stored at: 224 nm (reference off)

The gradient was:

<b>Time (min)</b>	<b>% NH<sub>4</sub>Ac in acetonitrile/H<sub>2</sub>O</b>
0	5
2.5	5
17.5	100
25.5	100
25.6	5
33	5

Sequence table:

	<b>Data file</b>
PSR 1	040406a/002-0201
PSR 2	040406a/003-0301
AR 1	040406a/004-0401
AR 2	040406a/004-0501
AR 3	040406a/004-0601
AR 4	040406a/004-0701
AR 5	040406a/004-0801
AR 6	040406a/004-0901
AR 7	040406a/004-1001
AR 8	040406a/004-1101
AR 9	040406a/004-1201
AR 10	040406a/004-1301
PSR 3	040406a/005-1401

**3. 2. 11 Does the integration parameters of the residual curve matter?**

Mobile phases: 10 mM NH<sub>4</sub>Ac in water and 10 mM NH<sub>4</sub>Ac in acetonitrile/water (95:5 v/v).

Flow: 0.64 mL/min

Injection volume: 7 µL

Column temperature: 36 °C

Column: HyPurity C18 150 x 3 mm, 3 µm

Signals stored at: 224 nm (reference off)

The gradient was:

<b>Time (min)</b>	<b>% NH<sub>4</sub>Ac in acetonitrile/H<sub>2</sub>O</b>
0	5
2.5	5
17.5	100
25.5	100
25.6	5
33	5

Sequence table:

	<b>Data file</b>
PSR 1	040406a/002-0201
PSR 2	040406a/003-0301
AR 1	040406a/004-0401
AR 2	040406a/004-0501
AR 3	040406a/004-0601
AR 4	040406a/004-0701
AR 5	040406a/004-0801
AR 6	040406a/004-0901
AR 7	040406a/004-1001
AR 8	040406a/004-1101
AR 9	040406a/004-1201
AR 10	040406a/004-1301
PSR 3	040406a/005-1401

**3. 2. 12 Infinite chromatographic resolution: Deconvolution of partially co-eluting peaks**

Mobile phases: 20 mM KH<sub>2</sub>PO<sub>4</sub> in methanol/water (65:35 v/v) and 20 mM KH<sub>2</sub>PO<sub>4</sub> in water.

Flow: 0.64 mL/min

Injection volume: 2 µL

Column temperature: 60 °C

Column: Zorbax SB C18 150 x 3 mm, 3 µm

Signals stored at: 210 and 254 nm (reference off)

Sequence table:

<b>% of modifier</b>	<b>Datafile</b>
60	040414b
65	040414c
70	040414d
75	040415a
80	040415b
85	040415c
90	040415d