Towards Comprehensive Online Multidimensional Frequency Transform Separations (COMForTS)

Mark Trudgett

2014
University of Western Sydney

A thesis submitted to the University of Western Sydney in partial fulfilment of the requirements for admission to the degree of Doctor of Philosophy.

Declaration

I, Mark Jerome Eric Trudgett, declare that this thesis is my own work, and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished works of others has been acknowledged in the text, and a list of references is provided.

Signature of Candidate

Date

Dedication

For my parents, Kevin and Sylvia

For my wife, Teresa

For my son, Reuben

Preface and Acknowledgments

This thesis embodies the foundation research for a new analytical method that requires some specific instrumentation but is principally an approach to the processing of signals. The instrument built to test this method was a novel combination of existing technologies that may serve as a template for inexpensive, low-level instrumentation for schools and university undergraduates. Acquisition and processing of the instruments signals, according to the method's principles, was achieved through custom software. This software forms an integral part of this thesis and is included on the attached solid-state media, together with installation instructions.

Many people have helped and encouragement me throughout this work and they all deserve my deepest thanks. In particular, I wish to acknowledge the support of a University of Western Sydney Postgraduate Research Award and the support and contributions of my Academic Supervisors and mentors:

- Professor R. Andrew Shalliker (UWS) and
- Professor Georges Guiochon (UTK)

Those to whom special thanks are due:

- Dr Michael Berthon-Jones: For assistance with the coding of algorithms, sage advice and testing of software, ideas and my explanations.
- Associate Professor Glenn Stone: For helpful discussion and direction with the statistical and transformational treatment of COMForTS detectors signals.
- Jordan Trudgett: Writing of the Microsoft XNA 3D rendering routines.
- Kevin Trudgett: For tireless and patient testing of the COMForTS Software, proofreading and listening.
- Reuben Trudgett: Design and production of the COMForTS Creature (originally FFaTCat) Logo and Video. 10 years old then, 14 now. Thanks, son!

Abstract

Comprehensive multidimensional separations are today dominated by systems that are fundamentally limited to highly asymmetrical online separations sacrificing separation space, or to lengthy, time consuming offline separations. With the exception of pulse-modulated methods in gas chromatography, separations have thus been limited to two dimensions. In liquid chromatography, the highest efficiencies in terms of separation power are achieved through asymmetric comprehensive online two-dimensional separations where the capacity to separate analytes within a given time is severely restricted by the need to avoid under-sampling of the first dimension and by the subsequent limitation of peak capacity within the second dimension. These limitations have been overcome in this thesis by employing multi-dimensional detection whereby both the retention times and frequencies of analyte pulses are recorded. Within a given separation dimension, analyte pulses were related to their linear velocities. Time-dependent frequency analyses combined with knowledge of the physical dimensions of that separation dimension allowed the determination of both the times at which analytes entered that separation dimension and their retention times within that dimension. By this means, it is possible to reconstruct a virtual comprehensive multidimensional separation. This approach has been called Comprehensive Online Multidimensional Frequency Transform Separations (COMForTS).

Analyte pulses can be introduced either as physical pulses resulting from alternately switched valves or as virtual pulses produced from the combination of signals from multiple in-separation detectors. The principle of operation being the same in both cases, a semi-empirical computer model of a physically pulsed system was developed and its feasibility positively demonstrated in simulations of high-efficiency separations in two dimensions. In that model, time-dependent frequency spectra were obtained by short time Fourier transforms. Performance characteristics with respect to harmonics, overtones, pulse width, peak width and peak frequencies were identified. Separations of higher dimensionality were also shown to be possible.

Following this theoretical groundwork, a basic prototypical COMForTS instrument together with custom control software and signals processing system was designed and constructed based upon a less physically demanding on-column multipoint detector array. Practical online two-dimensional separations were performed in which varied and controlled degrees of peak wrap-around and physical overlap were generated, replicating the effects of first-dimension under sampling. It was simultaneously shown that while a conventional online two-dimensional separation failed, COMForTS was able to fully, and correctly, resolve all analytes. Improvements in peak production rates of between 26- and 41-fold were observed and quantitative results were obtained. Minimal interferences were observed when time-dependent frequency transforms were cross-correlated. Significant future improvements to both the efficacy and speed of signals processing were also identified.

COMForTS also places no restriction on the analysis time (or peak capacity) of second, or higher separation dimensions. Total analysis times increase only additively with dimensionality whilst increases in peak production rates are multiplicative – possibly approaching three orders of magnitude in three-dimensional separations.

Most significantly, COMForTS demonstrated an exceptionally high confidence in the purity of peaks, bypassing limitations imposed by current statistical peak overlap theory by applying conditional logic to the separation of analyte information.

COMForTS is limited in that it provides separation information rather than a physical separation and that, in its most efficient form, detection is limited to non-destructive, arrayed, in-separation methods. Nonetheless, the method exhibits great potential for applications that demand complex separations with rapid turnaround, such as in process control, comparative metabolomics, fingerprinting of natural products and rapid screening of complex samples. It is envisaged that these applications of the method would leverage emerging high-efficiency micro- and nanoscale separations technologies facilitating close to time-of-sampling analytical results.

Key Words: Comprehensive, online, two-dimensional, multidimensional, chromatography, frequency domain, Fourier, Radon, frequency transform, high-efficiency, separations, peak capacity, peak production, peak overlap

Table of Contents

COMForTS Digital Thesis Materials

Contents of the attached solid-state media:

1. A copy of this Thesis in Adobe Portable Document Format (.pdf)
2. COMForTS 3.0 Modelling Software and Source Code
3. COMForTS 4.0 Chromatography Suite Software and Source Code
4. COMForTS ECD Application and Firmware Source Code
5. COMForTS 3.0 & 4.0 Installation Guides & Software Documentation
6. Sample data for demonstration of COMForTS 3.0 & COMForTS 4.0

Ancillary Materials:

7. Microsoft .NET Framework 4 & XNA Framework / Game Studio
8. Adobe Reader 9.5
9. Microsoft Visual Studio Express 2012

Publications and Patents Derived from the Thesis

Publications

Trudgett, M. J. E., Guiochon, G., Shalliker, R. A. (2011). "Theoretical description of a new analytical technique: Comprehensive online multidimensional fast Fourier transform separations." Journal of Chromatography A **1218**(22): 3545-3554.

Patents

PCT/AU2012/000142
Method for analysing a sample comprising a plurality of analytes.
Priority from AU 2011905440, 23 December 2011
Applicant: The University of Western Sydney
Inventors: Mark Trudgett and Ross Andrew Shalliker

List of Figures

List of Tables

Symbols and Abbreviations

2DLC	Two-dimensional liquid chromatography
An	Anthracene
Ac	Acetone
ACN	Acetonitrile
AU	Absorbance units
b	Binary digit (bit)
B	Byte (8 bits)
C	Chrysene
CCD	Charge coupled device
COMForTS	Comprehensive Online Multidimensional Frequency Transform Separations
CPU	Central processing unit
C#	(Microsoft Visual) C# ("C-Sharp") programming language
$D_1 \ldots D_n$	Separation dimension in position n in sequence
DAD	Diode-array detector
Det.	Detector
DFT	Discrete Fourier transform
d_p	Particle size (stationary phase)
DRT	Discrete Radon transform
Et	Ethyl
(F)FT	(Fast) Fourier transform
FRT	Fast Radon transform
FTIR	Fourier transform infra-red
GC	Gas chromatography
HAc	Acetic acid
h	Hours
I.R.	Infra-red
inj.	Injection
k	Retention (capacity) factor
λ	Wavelength
LED	Light emitting diode
LC	Liquid chromatography
mAU	Milli-absorbance units (AU x 10^{-3})
mAU.s	Milli-absorbance unit × seconds

MCSCSD	Method for combining signals from closely-spaced detectors
Me	Methyl
MeOH	Methanol
MDLC	Multidimensional liquid chromatography
min	Minutes
MS	Mass Spectrometer
N	Naphthalene
ODS	Octadecyl silane
OH	Alcohol
P	Pyrene
PAH	Polynuclear aromatic hydrocarbon
PC	Personal computer
R	Resolution
R_c	Coefficient of linear correlation
RAM	Random access memory
R_f	Response factor
RMS	Root mean squared
RT	Radon transform
s	Seconds
SFC	Super-critical fluid chromatography
S/N or SNR	Signal to RMS noise ratio
SOT	Statistical overlap theory
T	Temperature
t	Time
UV	Ultraviolet (light)
VB	(Microsoft) Visual Basic (.NET) programming language
VIS	Visible (light)
VS	Microsoft Visual Studio software development environment
Vol	Volume

Chapter 1.
Introduction

This thesis represents the foundation research for a new analytical method named Comprehensive Online Multidimensional Frequency Transform Separations (COMForTS).

Forty years ago, Karger, Snyder, and Horvath [1] discussed multistep separation schemes for complex samples – a concept further developed and demonstrated a decade later in planar chromatography by Guiochon and Beaver [2-4]. Giddings [5] also appreciated the potential of multidimensional separations to yield vast improvements in resolving power over one-dimensional separations. Even in these pivotal years, Giddings and Guiochon recognised that the practical success of these separations would rely upon increased automation, including improved and extended detection systems coupled with fast data processing. Schure [6] points out that the most common reason for using multidimensional separations is the need to increase peak capacity, but Schure also uses 'dimensionality' not only in the sense of physical separation but in a sense that includes multidimensional detection. Diode-array detectors, for example, record absorbance data in the two dimensions of time and wavelength, and mass-selective detectors record over the domains of particle mass relative to charge and time. Over the past decade or so, there has been a growing realisation, particularly in the fields of metabolomics and proteomics [7-10], that the power of one-dimensional separations is woefully inadequate when applied to complex samples. In those fields, it is not surprising then that two-dimensional HPLC combined with mass spectrometry has become the standard analytical tool [11]. The most significant disadvantage of multidimensional separations is their typically, very long analysis times.

COMForTS was invented by the author as a response to the need to separate very complex samples within a reasonable time. Over recent years, improvements in physical separations systems have been mostly incremental

and it is beginning to appear as though the ultimate physical limits of those systems are not far away. In order then to achieve improvements of one or more orders of magnitude, a different approach was required: maximisation of the extraction of useful information contained within chromatographic signals.

From the nomenclature, it is evident that COMForTS is a multidimensional separation technique. The separation power of the technique arises from its use of multiple physical separation dimensions combined with multidimensional detection in a sense that is more fundamental than Schure's concept: COMForTS measures detection parameters not only in the time domain, but also in the frequency domain. It will be shown that this unique feature of COMForTS can greatly increase the ability of any chromatographic separation to obtain information about the chemical composition of a sample. In modern parlance, COMForTS is a 'separation turbocharger', that affords multiplicative increases in the separation information that can be obtained from any given system.

Throughout much of the present work, the focus is on separations in liquid chromatography (LC). However, with some limitations with regard to detection, the COMForTS technique can be applied equally to the improved performance of multidimensional gas chromatography (GC), capillary electrophoresis (CE) or indeed to multidimensional mixed-mode separations such as LC-CE.

To understand how this was done, and to appreciate its implications within the field of separation science, will require some exploration of the general characteristics, advantages and limitations of multidimensional separations, digital signals processing in the frequency domain and of software and systems design. This project spans a breadth of concepts not often encountered in a single work. As such, discussions will be kept necessarily and hopefully, pleasingly brief - limited to an overview sufficient to understand why this research was done, what it accomplished and how.

1.1. Multidimensional Separations

> *"The principal rationale behind multidimensional separations is that they offer a more effective as well as efficient way to generate high peak capacity and thus permit more complete resolution of complex mixtures. I suspect, however, that there is another motivation that attracts people to multidimensional separations: the resulting two dimensional chromatograms make fascinating pictures. Two-dimensional separation patterns are somehow more satisfying than a series of peaks in a one-dimensional chromatogram. The human mind is highly adept at dealing with complex information presented in the form of images and, despite the complexity, is able to quickly spot differences among such patterns."*
>
> *James W. Jorgenson in the Forward to "Multidimensional Liquid Chromatography" [6].*

Analytical chemists have long been aware of the limited number of analytes - the number of peaks - that may be resolved in a single chromatographic separation. At the same time, there is an ever-increasing need to separate progressively larger numbers of analytes in very complex samples.

Multidimensional chromatography has arisen as an increasingly important technique in fields requiring the analysis of highly complex mixtures, usually of natural origin. The fields of proteomics, metabolomics, and natural drug discovery, for example, deal with samples containing hundreds or even thousands of compounds that cannot be adequately or practically separated in one dimension (D_1). In such cases, further separation may be achieved by serial application of separations in further dimensions ($D_2 ... D_n$), which can be carried out in either time (S_t) or space (S_s) or combinations of both (such as $S_t \times S_s$).

1.1.1. Separations in time and space

In practice, it has been found that multiple time-domain dimensions ($S_t \times S_t$) provide the simplest and most efficacious separations [12]. The analytical requirements, constraints and limitations placed (by choice or necessity) on multidimensional chromatography may be summarised as follows:

a. The separation should be "comprehensive" in that all components pass through all dimensions.
b. No separation achieved in the series of previous dimensions ($D_1 ... D_n$) should be lost in any of the subsequent separation dimensions ($D_2 ... D_{n+1}$)
c. In the most common case of multiple time-dimension separations ($S_t \times S_t$), practical application can be achieved in a variety of modes:
d. Online:

   N-fractions per peak in D_n are collected, transferred to the next separation level, and directly subjected to analysis in D_{n+1} in real-time. In this case, the analysis time (T_{n+1}) in D_{n+1}, must be less than the time between subsequent fractions transferred to the second dimension, ideally less than one standard deviation of the peak distribution eluting from the first dimension [13]. Otherwise, chaotic band displacement can result. In HPLC, this may require D_{n+1} run times of only a few per cent of the D_n run time.
e. Offline:

   N-fractions in D_n are stored temporarily until subjected to analysis in D_{n+1} such that the run rime of D_{n+1} is no longer subject to the same constraint as in (d), but analyses are time consuming.

f. Stop-Start:
After an eluate plug has been delivered directly to D_{n+1} from D_n, the flow in D_n is stopped and the separation is then run in D_{n+1}. Upon completion of the D_{n+1} separation, the flow is resumed in D_n to deliver another eluate plug, and so on. The need to collect, store and manage large numbers of fractions, free of contamination, is thereby obviated. There is, however, no reduction in the overall analysis time compared with (e).

In offline analysis, relaxation of the constraint on D_n run times comes at the cost of a very large increase of the total analysis time, which is given by the series:

$$T_A = T_1 + N_1T_2 + N_2T_3 + \cdots \quad \text{Eq. (1-1)}$$

Where T_A is the total analysis time, T_1, T_2 ...T_n represent the analysis time in each respective dimension and N is the number of fractions transferred to each dimension. This drawback has been addressed by subjecting only specific fractions of interest from D_n to analysis in D_{n+1}: 'Heart-cutting' [14] – but this method can no longer be termed a 'comprehensive' separation. For comprehensive separations, the online method (d) is preferred due to its simplicity and speed of analysis. Unfortunately, the constraint placed on the analysis time in the second dimension is severe, and necessitates highly optimised, asymmetrical analyses that sacrifice much of the separation space that could otherwise be available in a multidimensional system [12]. Parallelisation of online second-dimensions has been employed as a means of achieving rapid symmetrical comprehensive separations [15], but remains limited in its D_1 retention time resolution and is a very expensive approach that necessitates significant optimisations.

Now, the crux of these difficulties – and the reason for the development of the above techniques – lies in the problem of being able to identify in D_n, the results of separations that have taken place in previous dimensions $D_{<n}$, such that requirement (b) may be met. In time-domain separations (such as in liquid chromatography and electrophoresis), components are identified by their retention time. If there were another way to persistently 'label' the components previously separated and thus to identify and maintain the

achievements of the $D_{<n}$ time-domain separations, the analysis-time constraint for online D_{n+1} could be removed and there would be no need for time-consuming offline separations.

1.1.2. Peak capacity

The number of peaks that may be separated by a given separation dimension is known as its peak capacity. First introduced by Giddings [16] in 1967 and elaborated upon by Grushka [17], peak capacity, n, is defined as:

$$dn = (4\sigma)^{-1} dt \qquad \text{Eq. (1-2)}$$

Under the reasonable assumption that the number of theoretical plates of a column (N) varies slowly, if at all, with increasing retention, Eq. (1-2) becomes Eq. (1-3) [18], using the retention times of the first (t_0) and last (t_r) peaks:

$$n = 1 + \frac{\sqrt{N}}{4} . ln\left(\frac{t_r}{t_0}\right) \qquad \text{Eq. (1-3)}$$

When measuring, rather than predicting peak capacity, single dimension peak capacity, assuming a resolution factor of one, is often estimated by:

$$n = \frac{t_r - t_0}{4\bar{\sigma}} \qquad \text{Eq. (1-4)}$$

using the mean variance ($\bar{\sigma}$) in peak width and difference between the retention times of the first (t_0) and last (t_r) peaks.

Under ideal conditions, the total peak capacity of a multidimensional separation ($n_{c,MD}$) is often described as the simple product of the individual peak capacities $n_{c,Dk}$.of each dimension, k (the 'product rule'):

$$n_{c,MD} = n_{c,D1} \times n_{c,D2} \times \ldots \times n_{c,Dk} \qquad \text{Eq. (1-5)}$$

When two-dimensional (2D) online separations are considered, under sampling of the first dimension can lead to a significant loss in peak capacity. Considerable work in quantifying – or predicting – this loss has been carried

out by Carr's group [19-23]. Li *et al* [22] obtained the following equation for the effective peak capacity ($n'_{c,2D}$) in online two-dimensional liquid chromatography (2DLC):

$$n'_{c,2D} = \frac{{}^1n_c \, {}^2n_c}{\sqrt{1 + 3.35\left(\frac{{}^2t_c \, {}^1n_c}{{}^1t_g}\right)^2}} \qquad \text{Eq. (1-6)}$$

where ${}^1n_c$ and ${}^2n_c$ are the first and second dimension peak capacities, ${}^1t_g$ is the first dimension gradient (or run) time, ${}^2t_c$ is the second dimension cycle time which is equal to the sum of the second dimension gradient time ${}^2t_g$ and the re-equilibration time (${}^2t_{reeq}$) with all times measured in minutes. In this equation, the denominator is always greater than unity and the total peak capacity of online 2DLC must always be less than the product of the dimensional peak capacities. To avoid under sampling and thereby maximise $n'_{c,2D}$, it can be seen from Eq. (1-6) that the second dimension cycle time (${}^2t_c$) must be very small compared to the first dimension gradient time (${}^1t_g$). If, however, under sampling is severe, the peak capacity becomes virtually independent of the first dimension peak capacity (${}^1n_c$) and $n'_{c,2D}$ can be approximated by [22]:

$$n'_{c,2D} \cong \frac{{}^1t_g \, {}^2n_c}{1.83 \, {}^2t_c} \qquad \text{Eq. (1-7)}$$

1.1.3. Statistical Peak Overlap

The above discussion of peak capacity assumes that all components of a sample elute at regular time intervals that correspond precisely to the required minimum resolution (usually R = 1). In reality, this is far from the case and Davis and Giddings [24] asserted that the retention of components in complex mixtures is random; a behaviour supported by the experimental work of Martin, Herman and Guiochon [25]. Consequently, we can only view the above definitions of peak capacity as expressions of the "ideal", or maximum, number of resolvable components. A most apposite description of its effect on chromatography and the magnitude of the over-estimation of peak capacity has already been made:

> *"…one has no real hope of resolving 100 components on a column with n = 100… One would intuitively expect that the random distribution of retention… would increase the space needed by each peak by a factor of approximately 1.5-2 and thus reduce the resolvable components to perhaps 50-70. … [However,] if one attempted to completely resolve a 50 component mixture on a column with n = 100, for which there is twice as much space as theoretically needed for each component peak, disappointment would ensue. Random positioning of retention volumes would lead to many components occupying the same space, and unoccupied and unused gaps would appear in embarrassing abundance."*
>
> *Joe Davis and Calvin Giddings, Anal. Chem. 55 (1983) 418 [24].*

Naturally, statistical overlap theory (SOT) has since received considerable attention both with respect to one dimensional separations and two dimensional separations [26-32]. With regard to the inherent nature of COMForTS (which will be dealt with in Chapter 2, Chapter 4 and Section 7.1), we are particularly interested in being able to determine the probability that a given peak in a chromatogram represents a single compound (component). Davis and Giddings showed that, the number of peaks p in a chromatogram is related to the peak capacity n_c and total number of components m by Eq. (1-8) [24].

$$\ln p = \ln m - m/n_c \qquad \text{Eq. (1-8)}$$

It is possible, using Eq. (1-8), to determine m by performing multiple separations at different peak capacities and plotting ln p against $1/n_c$. Knowing m, we can then calculate the probability that each peak equates to a single component. Martin *et al* [25] also considered the differing probabilities of peak overlap occurring near the extremities of retention and formulated a more useful relationship (Eq. (1-9)) between the probability P_x that all m components of a sample will be resolved in a separation of known peak capacity n_c. This equation can be re-arranged to give Eq. (1-10).

$$P_x = \left(1 - \frac{m-1}{n_c - 1}\right)^{m-2} \quad \text{Eq. (1-9)}$$

$$n_{c,min} = 1 + (m-1)/(1 - P_x^{1/(m-2)}) \quad \text{Eq. (1-10)}$$

where $n_{c,min}$ is the minimum peak capacity required to separate m components with a probability P_x that any given peak is pure.

Figure 1-1 shows, using Eq. (1-10), the relationship between the number of components m and the peak capacities n required for their separation at varying levels of certainty. It becomes immediately apparent that the number of components that may be truly resolved for a given peak capacity, is indeed embarrassingly small.

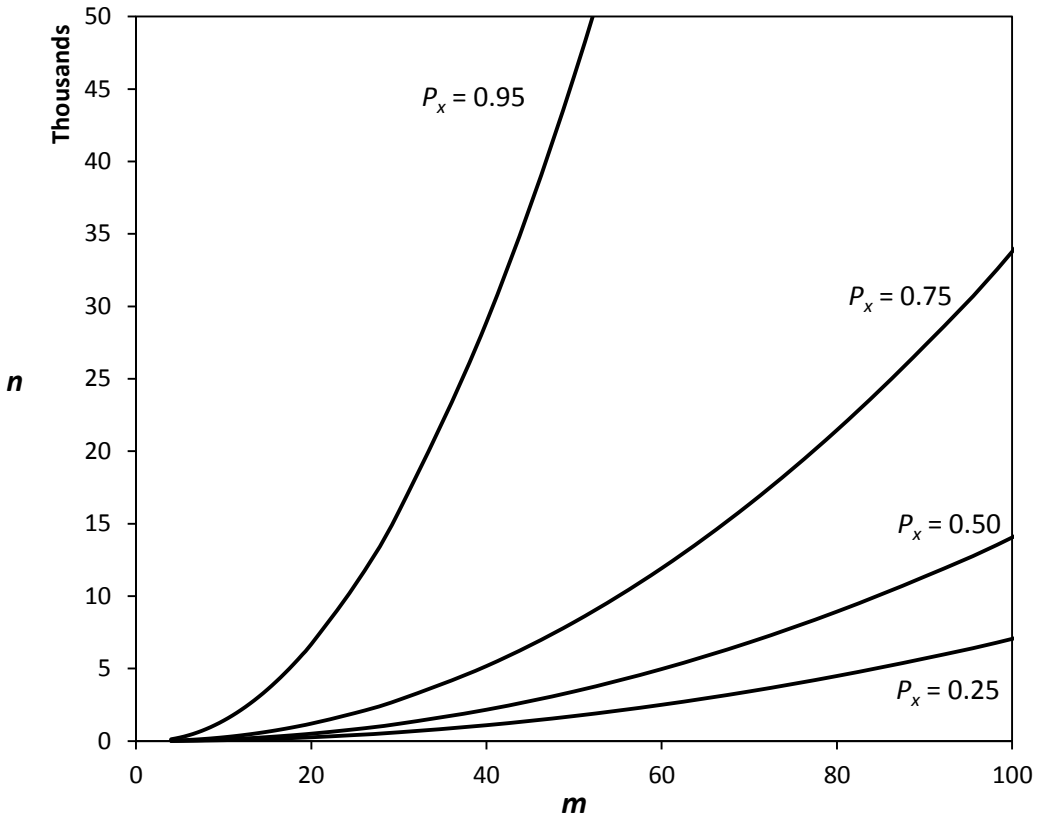


Figure 1-1: Probabilities of component overlap for different peak capacities.
The plot shows the peak capacities, n, required to ensure that all m components of a sample are separated with a probability of P_x from Eq. (1-10).

In a multidimensional separation system, we refer to the dimensions as being 'orthogonal' when the retention behaviour in one dimension is perfectly

uncorrelated to retention in any other dimension. In orthogonal systems, then, the retention of a set of m components is random in each of the dimensions k, and Eq. (1-9) can be extended, by the 'product rule' to apply to k-dimensional separations as shown in Eq. (1-11) [25].

$$Pb(k) = 1 - \left[1 - \left(1 - \frac{m-1}{n-1}\right)^{m-2}\right]^{k} \qquad \text{Eq. (1-11)}$$

In 2007, the maximum peak capacity (as calculated using Eq. (1-5) of state-of-the-art comprehensive 2DLC separations was of the order of 500–2000 and required analysis times between 30 and 60 minutes [33]. It is chastening to observe (using Eq. (1-11)) that such systems could not separate more than 23 components with a probability greater than 95% that each peak represented a single component. Certainly, multidimensional detection can afford a further multiplicative increase in "effective" peak capacity [6,31,33]. Stoll *et al* [33] noted that some 2DLC-MS separations (using Fourier transform ion cyclotron resonance (FT-ICR)) can effectively achieve total peak capacities as high as 25,000 (for 2DLC coupled with 12 T FT-ICR MS [34]), where all of 80 components may be resolved with the same purity criterion – provided, of course, that an exact-mass database containing those compound masses is available.

1.1.4. Peak production

When dealing with complex samples, our goal is not only to separate possibly thousands of compounds but also to do this quickly. Peak capacity gives us some idea of how many peaks may be resolved and SOT expresses a relationship between the number of resolved peaks and the probable number of resolved compounds. Peak production affords a measure of how quickly this can be accomplished.

Peak production is frequently referred to by the slightly more descriptive term "peak production rate", and is defined as the peak capacity per unit of analysis time. In the multidimensional separation literature, the peak capacity most often used to determine peak production is the simple, maximum peak capacity given by the product rule (Eq. (1-5)). The reasons for this are manifold and not confined to the fact that this method produces the largest,

most impressive figures. In online 2DLC, for example, if the second dimension is not under sampled (i.e. the second dimension cycle time is a small percentage of the first dimension's gradient time), then accounting for under sampling makes little difference to the result. Primarily, though, peak production permits an easy comparison of the relative speeds of different separations.

Offline 2DLC (Section 1.1.1) affords the greatest peak capacities but with deplorable peak production, where, in some instances, analysis times are measured in days. Significantly greater peak production is realised in online 2DLC with the trade-off being a loss in peak capacity. When the alarming increases in peak capacities required for true separation of greater numbers of analytes are considered, this deficiency in peak capacity appears almost insignificant compared to the resulting generous discount in analysis time. With high peak capacities and peak production in online 2DLC of the order of one peak per second [33,35,36], combined with the ability of liquid chromatography to separate a much wider range of analytes than other separation methods [37], one is not unexpectedly startled by the research focus on online 2DLC.

1.1.5. Effective peak capacity and production

An attempt has been made to limit the discussion thus far to physical peak capacity and hence physical peak production where the importance lays in the relative capabilities of varied separations methods. When we incorporate the capacity to separate analyte or peak information by methods such as multidimensional detection, which was touched upon in the preamble to this chapter, we now have another, non-physical dimension of separation. Such informational dimensions are not usually correlated to retention times in physical dimensions and the product rule can be used to estimate the *effective* peak capacity and *effective* peak production.

The example of 2DLC coupled with 12 T FT-ICR MS, used above to illustrate the devastating effect of SOT, had an *effective* system peak capacity of 25,000 and *effective* peak production of ~66 peaks per second. In contrast, SCX(5)-RP-HPLC with the same FT-ICR detection achieved an effective peak capacity of only 3500 but effective peak production of ~81 peaks / s [34]. Fast two

dimensional gas chromatography with time-of-flight MS detection (GC × GC-TOF-MS) has accomplished 17 peaks per second with an uncorrected n_c = ~4400 [38]. In each of these instances, the separation is not entirely physical and the peak capacity of MS detection is profoundly dependent upon the completeness of mass spectral databases.

Let us assume for the time being that 12 T FT-ICR MS mass-spectral libraries are perfectly complete, i.e. they contain data for every compound in existence. Since this would enable virtually unequivocal identification of physically overlapped analytes (that would appear in a standard chromatogram as a single peak), SOT becomes redundant and the effective peak capacity of such a system would better be described as the "analyte capacity" and is encompassed by Schure's concept of "dimensionality" [39]. This is a confounding area where MS detection has sometimes been found to be orthogonal and sometimes not [39]. The consequence is that the ability of MS detection to overcome the limitations on analyte resolution defined by SOT remains equivocal.

1.2. Frequency Domain Methods in Separations

To date, the frequency domain has been applied in chromatography essentially as a simple translation of retention time to migration speed through one dimension, providing improvements in signal to noise ratios. In multidimensional separations, it would be possible for two components that would otherwise be separated in the time domain, to present the same frequency at the detector. Similarly, components migrating at different speeds can elute at the same time in later dimensions. Rather than using the frequency domain or time domain alone, if we were to separate components in both the time and frequency domains, the separation of coincident components in subsequent dimensions becomes a problem of the separation of the frequencies that are coincident in time. This can be treated as a purely mathematical manipulation of the resultant signal by time-dependent frequency analysis.

Various mathematical methods have been applied to chromatography and analytical signal processing for many years, and of the transform methods, the Fourier transform is the most important and finds considerable application [32], albeit in various guises and differing purposes.

1.2.1. Fourier transform

The resolution of time-coincident frequencies can be treated as a purely mathematical manipulation of a pulse-containing signal by Fourier analysis, Eq. (1-12) [40]:

$$G(f) = \sum_{k=1}^{N} g(t).cos\left(\frac{2\pi k}{N}\right) - j.\sum_{k=1}^{N} g(t).sin\left(\frac{2\pi k}{N}\right) \quad \text{Eq. (1-12)}$$

where $G(f)$ is a frequency dependent function that is the discrete Fourier transform of a real, time dependent signal ($g(t)$): N is the number of data points and $k = 1, 2, \ldots, N$ [40]. Solution of the discrete Fourier transform is highly computationally intensive and, in practice, is achieved by using Fast Fourier Transform (FFT) algorithms that exploit the transform's symmetries and Euler's theorem [41] (Eq. (1-13)) (which gives rise to the integral form of the transform, Eq. (1-14). These are known as radix-2 algorithms and the Cooley-Tukey FFT algorithm (Eq. (1-15)) is probably the best known and most widely used example [42,43].

$$e^{i\theta} = cos\,\theta + i\,sin\,\theta \quad \text{Eq. (1-13)}$$

$$G(f) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} g(t).e^{-2\pi if}\,dt \quad \text{Eq. (1-14)}$$

$$X_k = \sum_{m=0}^{N/2-1} x_{2m} e^{-\frac{2\pi i}{N/2}mk} + e^{-\frac{2\pi i}{N}k}\sum_{m=0}^{N/2-1} x_{2m+1} e^{-\frac{2\pi i}{N/2}mk} \quad \text{Eq. (1-15)}$$

where N is the number of data points and must be an integer power of 2, i.e. $N = 2^x$ where x is a positive integer. The Fourier transform is symmetrical about the Nyquist limit, N_f:

$$N_f = {}^{N}/_{2} + 1 \qquad \text{Eq. (1-16)}$$

Where N_f is the number of frequencies that can be represented by a Fourier transform using a given number ($N = 2^x$) of data points and the '1' represents 0 Hz [40]. The FFT (Eq. (1-15)) thus calculates only the first half of the transform. Furthermore, Euler's relationship (Eq. (1-13)) allows a further division of the transform into its odd (m+1) and even (m) terms. Eq. (1-15) can be applied recursively on its own output until $N = 1$. The result of this 'divide and conquer' approach is a massive decrease in the total number of calculations, and hence the computational time required to perform the transform, allowing today's personal computers to perform FFT's on most analytical signals in real time.

It is important to note here that the Fourier transform is a one-dimensional transform in frequency ($G(f)$) of a one-dimensional time-domain signal ($g(t)$) (Eq. (1-14)). If, however, multiple Fourier transforms are performed over short (even overlapping) time intervals (i.e. segments or 'windows' of the time domain signal), then frequency-domain data become indexed in the time domain – and frequencies can be located in time. We can view this as a series of Fourier transforms on a window sliding over the data, or as a set of band-pass filters covering the frequency domain in fixed size bands [44]. This procedure, known as the short-time Fourier transform, then produces information in both the time and frequency domains. A general property of such systems, however, is that there is always a resolution 'trade-off' between the domains. Higher resolution in the frequency domain entails reduced resolution in the time domain and *vice versa*.

In chromatography, Allegri *et al* [45] used Fourier analysis as a means of deconvolution to resolve partially overlapped peaks in time-dependent (detector) signals. Fourier transforms have also found use in improvements to signal-to-noise ratios, peak resolution and the speed of in-process chromatography by either analysing the results of multiple overlapped sample injections (correlation chromatography [46,47]) or measuring analyte column migration speeds by multiple-point detection such as in Shah Convolution Fourier Transform (SCOFT) detection [48] and combinations of such methods [49]. Over time, these methods have been refined to include the application to

correlation chromatography of Hadamard transforms (a discrete Fourier transform in two variables) by employing complicated sample-injection protocols of pseudo-random binary sequences [50,51]. Further enhancements include the use of multiple parallel sample injection as in Fourier Transform Capillary Electrophoresis (FTCE) [52] to achieve the same result as that generated by SCOFT, but without the need for multiple-point detection: translation of the migration speed of an analyte into a frequency domain signal. Each of these applications has achieved notable improvements in signal to noise ratios and, in some cases, improved chromatographic resolution [52,53]. Each one replaces time-domain detection with frequency-domain detection. Most of these methods are used in process control analyses, in industrial environments and are unknown by most laboratory chromatographers.

1.2.2. Radon transform

Whilst the Radon transform had found no real use until the 1960's, its origin can be traced to Johann Radon's 1917 work "*On the determination of functions from integrals along certain manifolds*." Radon demonstrated how to construct a function of two variables from its integrals over all straight lines in the plane [54]. This has since had significance in fields in which absorbed or emitted radiation, travelling in straight lines, is recorded as a planar image. Because each pixel in the image is the sum (integral) of absorbance or emission functions, such images are, in effect, the Radon transforms of arrays of absorbance or emission functions and these functions can be recovered by performing an inverse Radon transform. In 1979, after exploring Radon's work, x-ray radiographer Allan Cormack and biomedical engineer Godfrey Hounsfield were awarded the Nobel Prize in Physiology and Medicine for their invention of computer-assisted tomography. The Radon transform is now found in astronomy, crystallography, electron microscopy, geophysics, material science, and optics and enjoys continued heavy use in computer assisted tomography (CAT) [54]. Radon transforms have also been put to use in image processing, pattern recognition and in chemistry for reconstruction of 3D chemical structures by 3D electron microscopy [55,56], yet no use has been reported in chromatography.

Mathematically, the classical Radon Transform (RT) projects a continuous 2D function f over straight lines as defined in its generalised form:

$$R(\rho,\theta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\delta(\rho - x.\cos\theta - y.\sin\theta)\,dxdy \qquad \text{Eq. (1-17)}$$

where δ(n) is the Kronecker delta function and ρ is the shortest distance from the origin of the coordinate system to the line. θ is an angle corresponding to the angular orientation of the line.

The result of the Radon transform is a (ρ, θ) space R, where a peak at coordinates (ρ_i, θ_i) denotes the existence in the original image, of a straight line of polar coordinates (ρ_i, θ_i) [57] as shown in Figure 1-2.

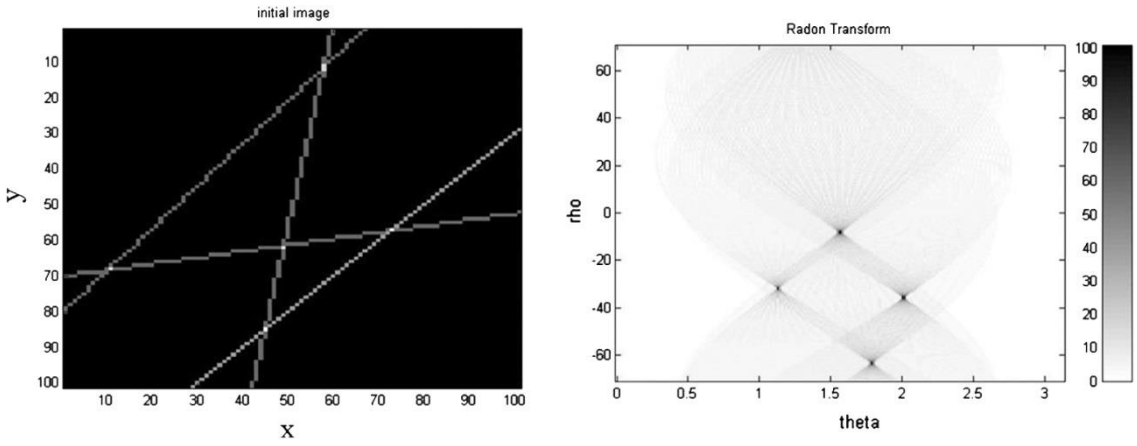


Figure 1-2: Effect of the Radon transform on images with straight lines.
When the classical RT is applied to the initial image (left), the resulting (ρ, θ) space (right) shows peaks corresponding to the polar parameters of the original straight lines. Reproduced from [57] with permission.

With respect to the Fourier transform, it can be shown that the two-dimensional Fourier transform ($\tilde{G}$, Eq. (1-18)) of a given two-dimensional function *f(x, y)*, is equivalent to the one-dimensional Fourier transform of the Radon transform of that function, Rf (Eq. (1-19)) [54].

$$\tilde{G}(\boldsymbol{k}) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y).e^{-i(\boldsymbol{k}.\boldsymbol{x})}\,dxdy \qquad \text{Eq. (1-18)}$$

where $\mathbf{k}$ = *(k, l)* and $\mathbf{x}$ = *(x, y).*

$$\tilde{G}(k) = G(Rf) \qquad \text{Eq. (1-19)}$$

One algorithm to compute the discrete Radon transform of a two-dimensional image consists of determining the points of a line and then summing up all pixels placed on that line. Having N slopes and between N and $2N$-1 displacement per slope this would require $O(N^3)$ additions where O is the image size. However, in order to produce high quality results, methods based on this theorem require interpolation from a rectangular grid into a circular grid, which increases the computation time considerably [58].

In 1996, Götz and Druckmüller [58] used the relationship between the Fourier and Radon transforms (Eq. (1-19)) to devise an algorithm for a fast Radon transform. As with the fast Fourier transform, the algorithm works on the same principle of decimation and greatly reduces the number of calculations required to evaluate the transform.

Even though chromatography has not seen any use of the Radon transform, there are distinct possibilities for its application to two-dimensional arrays of detector signals. If one considers that a convoluted system of objects with uniformly changing properties (such as chromatographic peaks and their position over time) may be represented as a system of overlapping straight lines, then the Radon transform could be used to characterise those straight lines.

1.2.3. Wavelet transforms

Time-dependent signals may be seen, geometrically, as the sum of an infinite series of sinusoidal waveforms, and the Fourier transform the frequency and phase coefficients of those waves. Wavelet transforms remove the sinusoidal constraint and allow fitting to non-sinusoidal but zero-average (wave) functions in time (or 'wavelets', ψ) (Eq. (1-20)) [59]. This fitting is performed when ψ is first normalised to one and centred about the time origin t = 0. ψ may then be shifted in time u and scaled by some factor s (Eq. (1-21) where, in the wavelet transform (Eq. (1-22)), relatively large scale-factors indicate the presence and magnitude of a matching wavelet with a time displacement of u within the original signal [59].

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \qquad \text{Eq. (1-20)}$$

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right) \qquad \text{Eq. (1-21)}$$

$$Wf(u,s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right)dt \qquad \text{Eq. (1-22)}$$

In practice, wavelet transforms can be quickly calculated by treating them as circular convolutions that are in turn calculated by the fast Four transform (FFT) [59]. Additionally, the choice of wavelet function has a significant effect on the resulting transform, and wavelets should be chosen to match as closely as possible the feature(s) of interest in the time-dependent signal [60].

Wavelet transforms have seen considerable application in separation science and there is a large body of literature to which Shao in particular has made valuable contributions, including, with Cai, an excellent review [61] of the use of wavelets in chemistry (see also [44,62]). Within chromatography, the principle applications have been to the analysis of multicomponent chromatograms [63-66], the resolution of partially resolved peaks [67-74] and the reduction of noise in chromatographic signals [75-78].

Of particular interest here was the use of wavelets by Eijkel, Kwok and Manz [60] to measure velocities of single particles flowing past an array of equally spaced detectors. This was a variation on the established use of the Fourier transform in Shah Convolution Fourier Transform detection (described in Section 1.3.1). Whilst Eijkel *et al* note that the wavelet has the advantage over Fourier transformation in that it reports both the frequency (hence velocity of particles) and time during which the particles were present in the system, they did not appear to have made any use of that information.

Like the short time Fourier transform, high resolution in the frequency and time domains is mutually exclusive. Interferences from harmonics and overtones are common and compounded by "oscillating interferences" that are centred in time and frequency between two fundamental frequency/time

peaks. Whilst such interferences are evident in the results of Eijkel *et al* [60], such was the remarkable improvement in signal to noise that they were nonetheless able to resolve two peaks from the original noise-dominated signal.

1.3. Introducing Frequencies into Chromatography.

Frequency domain signals can be introduced as pulses into chromatographic signals in two fundamentally different ways:

1. Signals from multiple physically displaced detectors can be combined. As the detectors are physically displaced, their signals are displaced in time and their summation results in a virtually pulsed signal.
2. Physical pulses can be introduced through a switching valve or through pulsed injection. The resulting physically pulsed signal can then be recorded with a single detector.

1.3.1. Virtual pulse modulation & arrayed on-column detection

The concept of virtual pulse modulation is typified by one-dimensional separations using SCOFT detection, a method pioneered by Kwok and Manz, in the period 1999 to 2003 [48,49,60,79-87]. This use of arrayed on-column detection to produce frequency-domain signals from physically summed [48,49,83] and digitally summed [88,89] time-domain signals has been well studied and shown to be effective for, in effect, measuring analyte velocities within a single separation dimension. The act of summing signals containing time-displaced analyte peaks creates one signal comprising a series of virtual pulses that can then be subjected to frequency analysis (Figure 1-3). Whilst the demonstrated goal of that research was to increase sensitivity by reducing signal to noise ratios, the approach has the singular disadvantage that it requires arrays of non-destructive, on-column detectors.

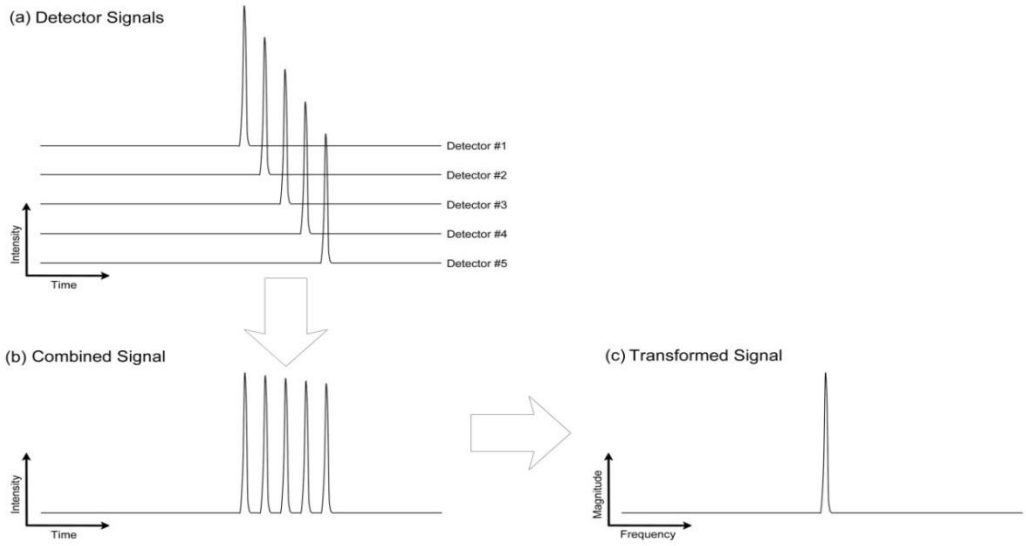


Figure 1-3: Frequency analysis of virtually pulsed signals.
Signals from multiple equally spaced detectors (a) are summed to produce a sequence of virtual pulses (b) the frequencies of which can be determined by a frequency transform (c).

1.3.2. Physical pulse modulation in multidimensional chromatography

Online comprehensive separations have been demonstrated in multidimensional gas chromatography via the use of inter-dimensional pulse modulation [90,91] and statistical analyses of the resulting signal. The identity and quantity of individual components in each dimension has been determined by the solution of multiple linear differential equations [92] and/or parallel factor analysis (PARAFAC) [90,93,94]. Such methods however, utilise very low frequency pulses (0.15 to 3 Hz [90]) whereby detection and deconvolution of overlapped peaks are achieved by statistical methods confined solely to the time domain. These chemometric algorithms are suitable only for target analytes and when multiple runs are combined for a single analysis all require alignment and pre-processing to enhance co-linearity of separations data. In combining multiple runs (though not in single runs), PARAFAC also requires an identical shape for a particular analyte peak in each run if it is not to output artificially generated peaks [10]. They also cannot resolve components that are not physically resolved in at least one dimension – unless there is a marked variation in peak shape between dimensions [90].

1.4. COMForTS Project Outline

The advantages of physically- or virtually-pulsed signals and time-dependent frequency domain analysis have not hitherto been combined to achieve comprehensive online multidimensional separations. Such a combination, by circumventing some of the physical constraints on peak capacity and peak production, has the potential to achieve performance gains that would be otherwise impossible.

Given the potential pitfalls and difficulties of such an approach as well as the time and the expense required to build some sort of prototypical test-bed, the following general procedure was adopted:

1. Develop a logically consistent theoretical basis and test that theory by computational modelling. (Chapter 2)
2. Perform a practical test of the principles of operation of COMForTS to ascertain whether locations-in-time of frequencies within detector signals adhere to the developed theory (Chapter 3)
3. Perform a practical test of the ability of COMForTS to overcome peak wrap-around and thereby increase peak production. (Chapter 4)

Of course, it is not possible to do the above without first designing and building an instrument and establishing digital control and data collection protocols (Chapter 5) in addition to developing a systematic and adaptable approach to the analysis of the resulting signals (Chapter 6).

Finally, Chapter 7 explores the nature and future of COMForTS: its application in assorted separations systems, requirements of and improvements to digital signals processing, applications in qualitative and quantitative analyses and the expected performance limitations of COMForTS itself.

Chapter 2.
Theory and Modelling of COMForTS

2.1. Introduction

Having described in Section 1.2 some means of separating overlapping frequencies and locating those frequencies in time, we must find some means of introducing these frequencies into the chromatograms without affecting the relative amplitudes of the component signals.

The imposition of a frequency domain signal would be most readily achieved by the application of pulses at a defined frequency to the eluate flow between separation dimensions. Because of longitudinal diffusion, such physical pulses will have a limited lifetime within a chromatographic system. This lifetime will depend upon the pulse frequency (the physical distance between pulse maxima) as well as the rate of longitudinal diffusion (band broadening) within the system. If such a system could be developed, the need for multiple (expensive and large) detectors would be obviated and only a single detector would be required with no limitations on the type of detection.

For both of these reasons, and that the resulting (virtually pulsed) signals from multiple detectors are comparable to physically pulsed signals, the generation of virtual pulses from detector arrays was not specifically modelled.

2.1.1. Effect of detector configuration on detected frequencies

If pulse modulation combined with detection in the frequency domain is to be employed to overcome peak wrap-around, the effects of the relative velocities of the frequency source and the velocity of pulses in the medium relative to the detector must be considered: the Doppler Effect.

Chromatographic detectors may be used in two distinctly different physical configurations: either 'on-column' where analyte signals are recorded whilst the analyte is *in* the separation process or, more commonly, 'post-column' where the separation is complete and the analyte signal is recorded as it passes through the detector with the same velocity as the mobile phase. The technique proposed here has been called Comprehensive Online Multidimensional Frequency Transform Separations (COMForTS). The present discussion will focus on online inter-dimensional pulsed injections followed by time-based sequential Fourier analyses i.e. short time Fourier transform. Because these two detector configurations and their effects on the detected frequencies have fundamental differences, COMForTS must be operated in two correspondingly different modes: COMForTS-*is* for in-separation detection and COMForTS-*ps* for post-separation detection (Figure 2-1).

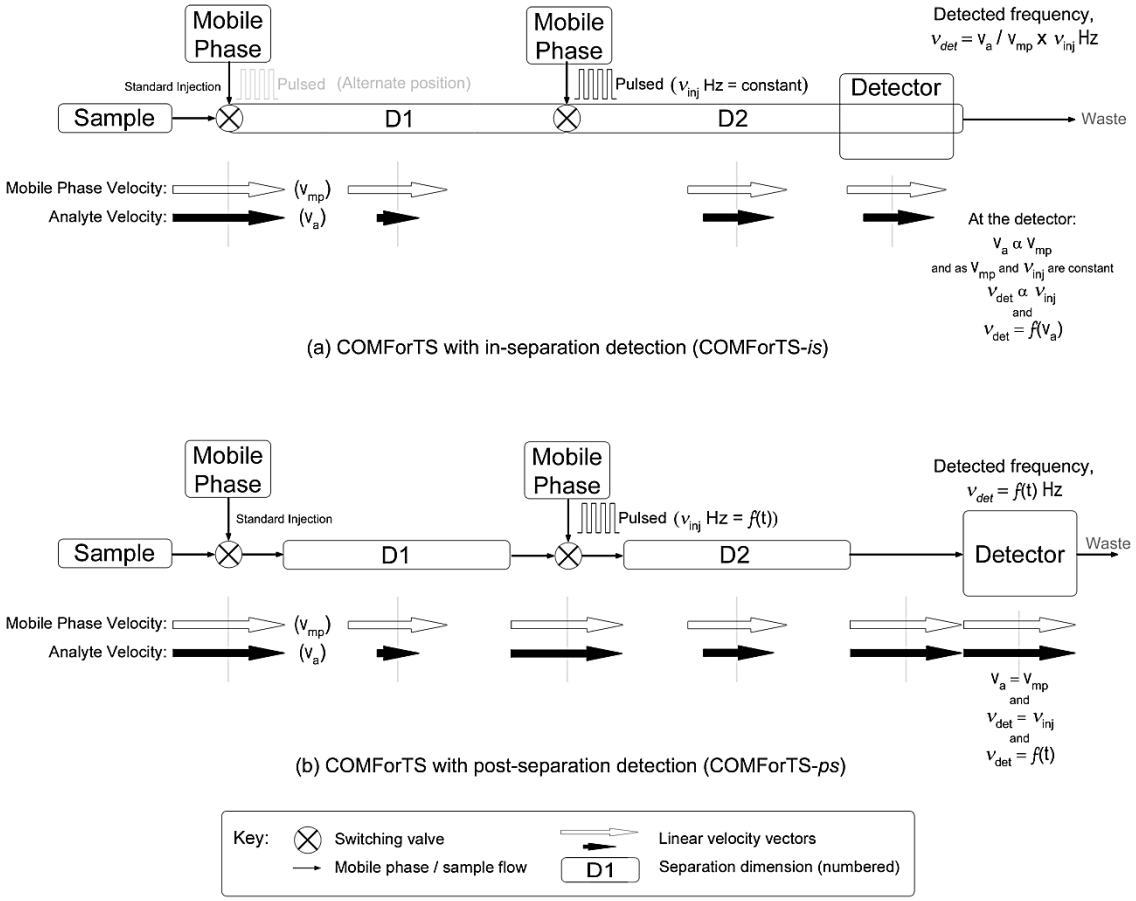


Figure 2-1: Detector configurations and measured pulse frequencies.
Configurations (a) and (b) describe the modelled instruments and their relationships between applied pulses, detector location and measured frequencies.

In the case of in-separation detection (Figure 2-1a), the differing relative velocities of the analyte and the mobile phase with respect to the detector entail that the frequency presented to the detector by any particular analyte (ν_A) is proportional to its linear migration velocity (V_a) and related to the injection pulse frequency (ν_{inj}) and the mobile phase velocity (V_{mp}) in the Doppler relationship of classical physics [95]:

$$\nu_a = \nu_{inj} \times \frac{V_a}{V_{mp}} \qquad \text{Eq. (2-1)}$$

This expression can be re-written in terms of the retention factor of the analyte (k_A):

$$\nu_a = \nu_{inj} \times \frac{1}{k_A + 1} \qquad \text{Eq. (2-2)}$$

From Eq. (2-2), it is clear that the pulse frequency chosen for the injection (ν_{inj}) must be sufficiently high that late-eluting peaks (with high values of k_A) present to the detector frequencies that remain above - and are distinguishable from - low and zero-frequency components (from 0 Hz to about 5 Hz) that are often strong in the Fourier transform of such data [32].

It should be noted that the previously published configuration for COMForTS-*is* (Figure 2-2) [96] did not correctly represent the modelled system and would not exhibit the frequency / velocity relationships described.

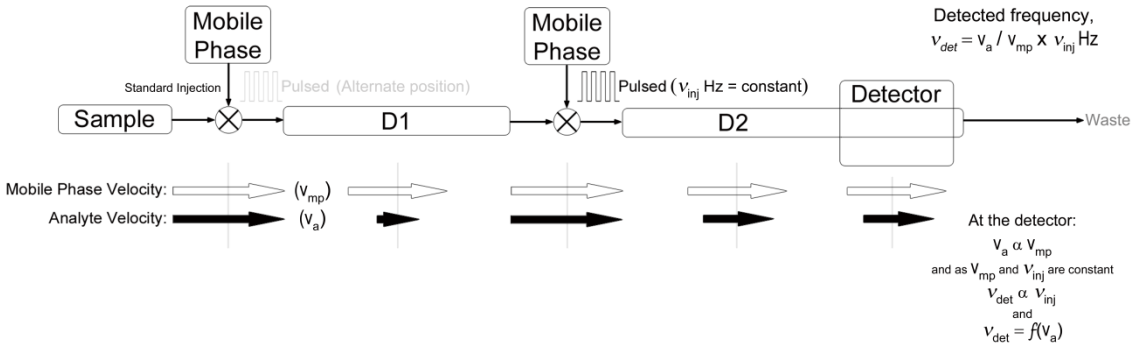


Figure 2-2: Previously reported COMForTS-*is* schematic representation.
The modelled system is accurately described in Figure 2-1a.

When post-separation detection is employed (Figure 2-1b), analytes experience the same linear velocity as the mobile phase – which is also the velocity at which they were travelling when the pulses were applied. There is therefore no Doppler shift in the detected frequencies and the detected frequency is the applied pulse frequency. If analytes are to carry different frequency 'labels' that are related to their retention time in the previous dimension, pulses must be applied at different frequencies as some function of time ($f(t)$):

$$v_t = f(t) = v_0 \pm f(n).t \qquad \text{Eq. (2-3)}$$

where ν_t is the pulse frequency at any given time, ν_0 is the initial pulse frequency, t is time and $f(n)$ is either constant or a function to optimise the distribution of applied frequencies over time.

2.1.2. Frequency domain processing of physically pulsed signals

Pulsed (on/off) injection would be expected to produce a waveform that approximates a square wave and it is known [97] that such waves cannot be represented by a single sinusoidal function but must rather be represented by the sum of a number of such functions as shown in Eq. (2-4). As each term in the FFT takes the form of $ze^{-i\theta}$ (see Eq. (1-15)) with no zero offset, the FFT of a square wave of frequency f will thus contain a signal at the fundamental frequency f as well as signals with harmonic frequencies of $3f$, $5f$, $7f$, … $n_o f$, where n_o is a positive, odd integer [97].

$$y = \frac{4}{\pi}.\left(sin(x) + \frac{sin(3x)}{3} + \frac{sin(5x)}{5} + \frac{sin(7x)}{7} \ \ldots\right) \qquad \text{Eq. (2-4)}$$

Square wave harmonics contain significant real data at predictable frequencies ($n_o f$) and with predictable magnitudes (M_n): The calculation of the magnitude (M_n) of the nth harmonic in the FFT of a square wave with fundamental magnitude M_f, pulse length ($t_2 - t_1$) and a repetition period of T [97] is given by:

$$M_n = \left|\frac{2M_f}{\pi n}.sin\left(\frac{\pi n(t_2 - t_1)}{T}\right)\right| \quad \text{Eq. (2-5)}$$

Any pulse-modulated separation process that is to be analysed in the frequency domain should therefore also make appropriate compensation for the presence of square wave harmonics.

2.1.3. Separations utilising more than two dimensions

When higher dimensionalities are considered, Eq. (2-2) and Eq. (2-3) can be applied to assign to each dimension a unique range of frequency labels. Alternatively, it is also possible to introduce multiple or multiplexed detectors or time-specific binary coded sequences [98,99] embedded in dimension-specific 'carrier' pulse frequencies; however, the complexities of the data analysis are beyond the present discussion.

2.1.4. Testing the principles of the proposed method

For the present, by the expedience of the introduction of a single frequency component, and the use of a single detector, very rapid comprehensive online two-dimensional separations can be achieved in real time without the lengthy analysis time concomitant with symmetric offline comprehensive separations. Moreover, components that are not resolved in the time domain in any dimension can be resolved in the frequency domain if the overlapping components are different in at least one dimension. The only constraint on the type of time-based separations to which COMForTS may be applied is the facility to apply pulses to the eluate flow. In other words, COMForTS would be equally applicable to, amongst others, liquid and gas chromatography and capillary electrophoresis.

The COMForTS technique combines the essentially digital physical process of 'on-off' pulsing and the entirely mathematical process of the Fourier transforms of the resulting signal. Because these processes are complicated and the lifetime of pulses within a separation will be limited by band broadening, it was expedient to test the above conceptualisation of COMForTS by computational modelling and simulation before attempting physical tests.

2.2. Experimental Simulation

The feasibility of COMForTS was assessed by construction of a software package (COMForTS 3.0) that involved the semi-empirical modelling of band broadening within a separation system and the simulation of time-domain separations based on defined retention data and the band-broadening model. Separations in the second and higher dimensions were treated by repeating the above process using the output of the previous dimensions' separation as the input for the next dimension. The package allowed programmed pulse modulation of the signal before 'injection' into any given dimension.

The liquid chromatographic separations of entirely arbitrary mixtures of components were simulated in both COMForTS-*is* and COMForTS-*ps* modes and the resulting simulated signals processed by a series of Fourier transforms of small time-slices of the simulated signal. The intent was to provide only a relatively uncomplicated demonstration of the main features, principles and possibilities of the proposed method and as such was not intended to represent a "best-case" separation at the limits of possibility. Nonetheless, some exploration was made of the conditions under which the performance of COMForTS may be maximised.

2.3. Materials and Methods

All data processing and calculations were performed by the COMForTS 3.0 modelling package, written in Microsoft Visual Basic .NET 2010 (Microsoft Corp., Redmond WA) as a proprietary executable program. Plotting functions were performed by COMForTS 3.0, Microsoft Excel 2007 and Mathematica 7 (Wolfram Research Inc., Champaign, IL).

COMForTS 3.0 runs natively on Microsoft Windows 32- or 64-bit operating systems from Windows XP on, and requires the Microsoft .NET Client Framework 4.0 or above. The work presented here was conducted with a 64-bit version of COMForTS 3.0 running under Microsoft Windows 7 64-bit on a

personal computer equipped with an AMD Phenom II X6 (six-core) CPU @ 3.4 GHz (Advanced Micro Devices, Sunnyvale, CA) and 8 GB RAM.

A copy of COMForTS 3.0 software is provided in the Digital Thesis Materials. The COMForTS 3.0 configuration files that define these simulations are contained within the folder named "…\Sample Data\COMForTS3" (or, if installed on your computer "My Documents\COMForTS\Samples\Modelling Examples"), for each COMForTS mode:

- Example_Six_Components_COMForTS-is.COMForTS
- Example_Six_Components_COMForTS-ps.COMForTS

2.3.1. Generation of simulated detector signals

Band broadening was modelled using a constant area, point-by-point Gaussian distribution biased for peak tailing and fronting. At each point in time, the standard deviation of the distribution was empirically defined as a function of the empirical rate at which an injected plug broadened in time within a given separation system. The retention times and tailing factors of each of the components were defined arbitrarily (or empirically) and elution was assumed to be isocratic. The efficacy of the band-broadening model and separation simulation was assessed by using empirical band-broadening rates, retention times and tailing factors to reproduce the empirical chromatogram from which these data were derived (Figure 2-3). The simulated chromatogram produced peaks at the correct retention times with peak shapes that differed only slightly on the tails of peaks below the level of about 5% of the peak height. The behaviour of the model with respect to changes in flow rate and column dimensions was as expected by chromatographic theory.

Samples of chromatographic noise were taken from a conventional HPLC-DAD system and were added to the final simulated signal, along with a linear baseline drift. The behaviour of the 'solvent peak' was also simulated (as an additional, unretained 'component' of the original injection) in order to facilitate visualisation of column void times and the relationships of retention time and detected frequencies to the pulse program. The resulting synthetic chromatograms thus contained real noise and close to real chromatographic distortions.

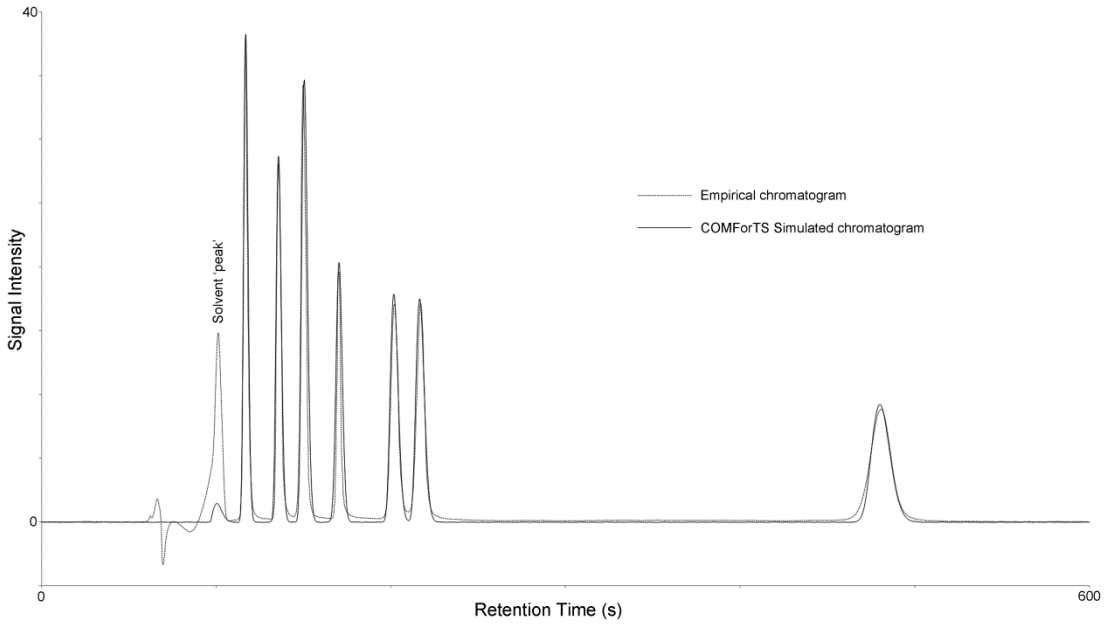


Figure 2-3: Comparison of simulated and empirical chromatograms.
The COMForTS simulation (solid line) used the empirical retention times, peak tailing (symmetry) factors and the band-broadening rate measured from the empirical data (dotted line).

The retention characteristics of a simulated six component 'sample' were construed, for the purposes of demonstration, such that components 'A', 'B' and 'C' exhibited low retention and low selectivity in the first dimension (Figure 2-4a) and *vice versa* in the second (Figure 2-4b). The inverse was held to be the case for components 'D', 'E' and 'F'. The result of such behaviour was that, of the six components, only component 'A' was fully resolved by a conventional symmetrical online separation (Figure 2-4b).

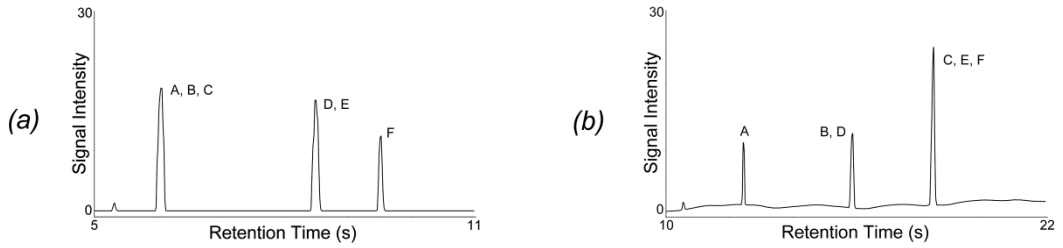


Figure 2-4: Simulated conventional online two-dimensional separation.
Six arbitrarily defined components were modelled. The first dimension separation result (a) and final two-dimensional result (b) show that components B, C, D and E are not resolved in either dimension.

This elution pattern also served to illustrate that even if an additional detector were used immediately after the first dimension, the expected result would be the resolution of no more than two of the six components. Peak tailing factors

(see [100]) were varied between 1.2 and 1.7 such that convoluted peaks consisted of components with both similar and dissimilar peak shapes.

Thus, the simulation conditions represent a set of circumstances under which no existing chromatographic method could achieve resolution of each of the six components within a comparable analysis time.

The COMForTS simulations were performed within the following liquid chromatography conditions and parameters:

- Column Size: 25 × 3 mm (in 1st and 2nd Dimensions)
- Overall rate of band broadening within the separation system: 0.005 s/s
- Injection Volume: 10 μL
- Eluent Flow Rate: 2.0 mL/min
- Run Times: 11 s in the first dimension and 11 s in the second.
- Total analysis time: 22 s
- Detector sampling rate: 256 Hz
- Signal generation resolution: 1024 Hz (1st Dimension), 512 Hz (2nd Dimension)
- COMForTS-*ps* pulse function: 14.0 Hz + 0.3 Hz / 0.05 s, t_0 = 5.3 s
- COMForTS-*is* pulse frequency: 30 Hz
- Total instrument noise was set to 1% RMS (using a sample of noise recorded from an HPLC-DAD instrument)
- A 10% sloping baseline was also incorporated into the COMForTS-*ps* detector signal.

Note that while the rate of band broadening is actually unit-less, we have specified units of "s/s" to emphasise that it describes the increase in band width (in seconds) per second of retention.

2.3.2. Frequency and time domain processing of the detector signal

Within the range of frequencies used in the simulated pulsed injection regimes, the lowest square wave harmonic frequencies would fall partially within the range of the analyses. If the magnitudes of the harmonic components (which are legitimate signals) were to be summed with the magnitude of the fundamental frequency component, it would be possible to

realise a significant improvement in the ratio of the detector signal to the background noise. For the sake of simplicity, however, COMForTS 3.0 was programmed with the facility to only identify (using Eq. (2-4)) and remove these harmonics at the lesser of either their measured levels or the mathematically predicted level (Eq. (2-5)).

COMForTS 3.0 was also equipped with some common apodization routines [40,41] to minimise the effects of performing Fourier transforms (which assume a continuous signal) on discrete data sets. Detector signals produced by separations in time also contain large tracts of data, such as the baseline, that contain virtually no frequency component. A sampling function [41] was used to mitigate the otherwise overwhelming near-zero frequencies contained in their transforms.

2.4. Results and Discussion

2.4.1. Separation of physically unresolved components

Comprehensive online two-dimensional frequency transform separations were simulated with both in-separation and post-separation detection, with frequency and time-domain information acquired by application of the short time Fourier transform to physically pulsed signals. The total simulated analysis time in both cases was 22 s and the actual signal processing was completed in less than two seconds - sufficiently quickly for real-time processing. When detection employed both the time and frequency domains, all six components were completely resolved in both COMForTS modes (Figure 2-5) even though components B, C, D and E were not physically resolved in either dimension.

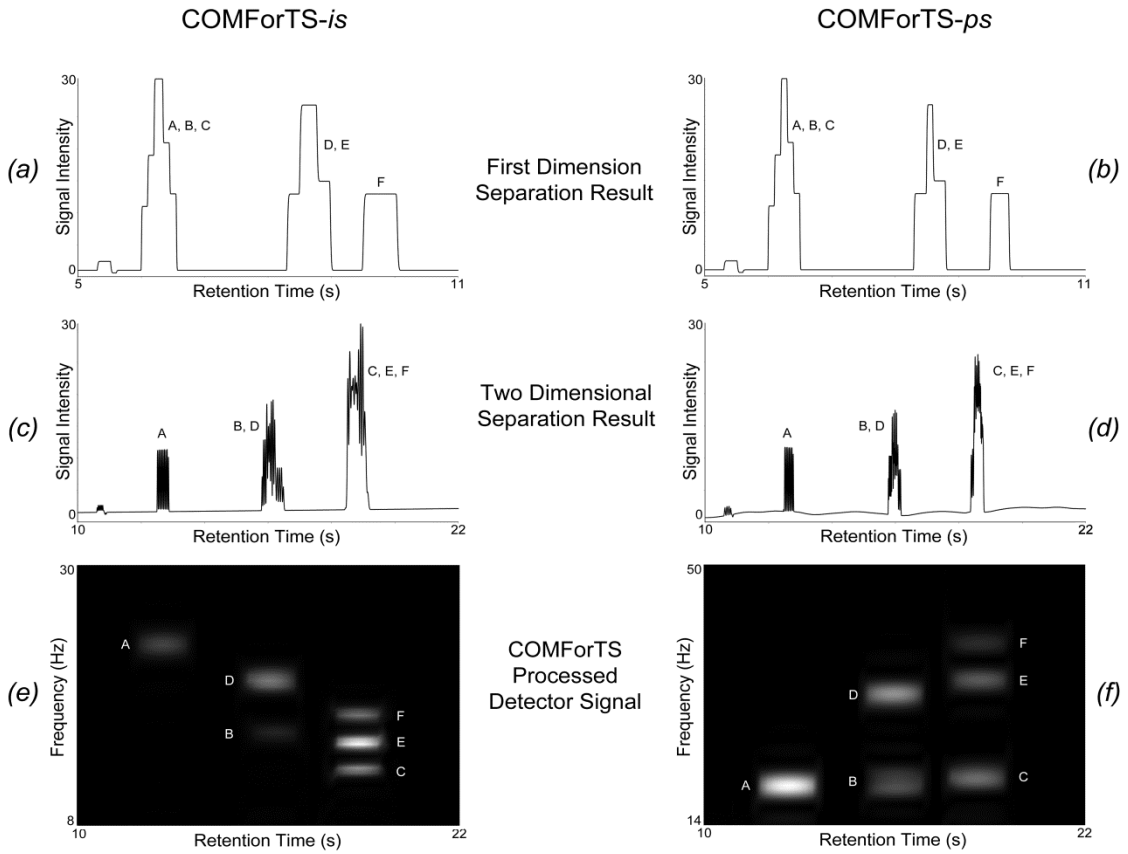


Figure 2-5: Simulated COMForTS separation of the six-component example. The eluate flow from the first dimension (a, b) was pulsed, according to each COMForTS mode, into the second dimension resulting in the final detector signal (c, d). 2D density plots of the COMForTS processed detector signals are shown in (e) and (f).

The roughly rectangular peaks observed in the first dimension (Figure 2-5 (a) and (b)) are the natural result of employing a relatively large (10 μL) injection plug. The effects of band broadening in a high-efficiency system do not become immediately apparent until these peaks are divided into smaller pulses as they pass into the second dimension (Figure 2-5 (c) and (d)).

2.4.2. Frequency labels and the reconstruction of separations in individual dimensions

Whilst there is absolutely no difference between the physical separations achieved in either COMForTS mode, as simulated analytes passed through an in-separation detector at a velocity that was lower than that of the mobile phase, it was expected in this mode (COMForTS-*is*), that apparent peak widths (Figure 2-5, (a) and (c)) would be greater than those observed with post-

column detection (COMForTS-*ps*, Figure 2-5, (b) and (d)). Differing patterns of detected frequency with respect to overall retention time are clearly evident in Figure 4 (e) and (f), which is also a direct result of the differing mechanisms of frequency labelling, described by Eq. (2-2) and Eq. (2-3), applicable to each of the COMForTS modes.

Let us first consider the frequency-retention pattern of the COMForTS-*is* separation (Figure 2-5e) in which a constant pulse-modulation frequency of 30 Hz was applied between the first and second dimensions. Because the detector is situated in the second dimension of the separation, the observed frequencies are proportional to the analyte velocity (Eq. (2-1), and hence retention time (Eq. (2-2) in the second dimension, where 30.0 Hz corresponds to the velocity of the mobile phase (i.e. a retention time equal to the void time (5.3 s)). Component D, for example, with a measured pulse frequency of 20.9 Hz has a retention time of 7.6 s in the second dimension (30.0 / 20.9 × 5.3 s). With an overall retention of 15.8 s (Figure 2-5 (c) and (e)), the first dimension retention time of Component D must have been 8.2 s (15.8 - 7.6 s) (Figure 2-5a). These relationships illustrate how the separation in any dimension can be determined and reconstructed by a series of such calculations based only on the signal from a single detector.

In the case of separations run with post-separation detection (COMForTS-*ps*), the relationship between the detected frequency and retention time is simpler because there is no difference between the linear velocities of the analytes and the mobile phase, and there is therefore no Doppler shift involved in the detected frequencies. The pulse frequencies are arbitrarily defined and related to the retention time in the dimension immediately prior to the application of pulses, by the pulsing regime defined by Eq. (2-3). If we examine the behaviour of Component C, we observe an overall retention time of 18.4 s (Figure 2-5, (d) and (f)) with an observed pulse frequency of 20.4 Hz (Figure 2-5f and Figure 2-6). Because the pulses were applied with a defined regime as a function of time (14.0 Hz + 0.3 Hz / 0.05 s, t_0 = 5.3 s), a detected frequency of 20.4 Hz corresponds directly to a retention time in the first dimension ((20.4 - 14.0) / 0.3 × 0.05) + 5.3 s = 6.4 s (Figure 2-5b). As the overall retention time was 18.4 s, the retention time in the second dimension alone must have been 12.0 s and we again demonstrate that the time-space

separations in individual dimensions can be reconstructed from the COMForTS detector signal.

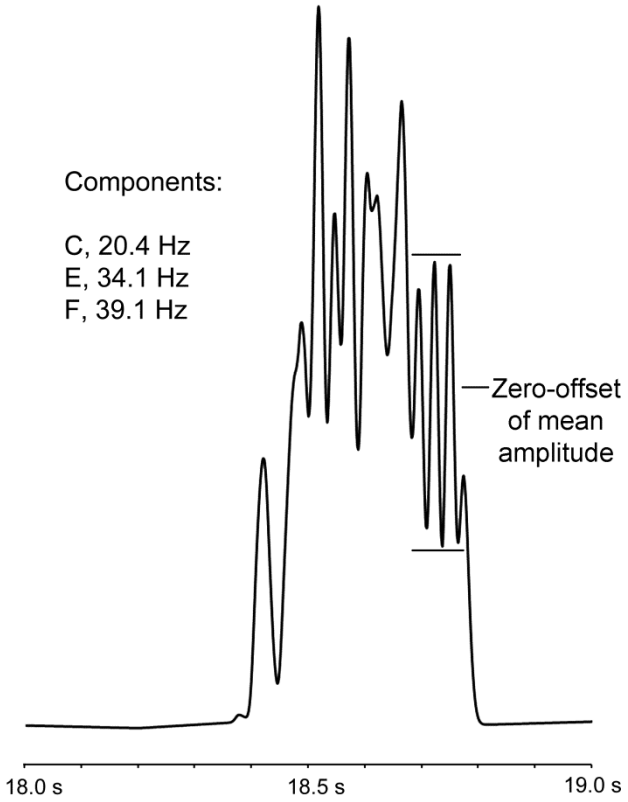


Figure 2-6: Detail of the simulated COMForTS-*ps* detector signal.
This view is of the latest eluting peak consisting of the co-eluting components C, E and F. The zero-offset of the mean pulse amplitude is also shown.

2.4.3. Detection in single and multiple domains

Figure 2-5 (c) and (d) describe the expected overlap of component peaks if detection were to be made in the time domain alone, producing only one pure peak and two convoluted peaks with overlapping pulses of varying frequency (Figure 2-6). Transforming the time domain signal to the frequency domain alone (Figure 2-7, main image) produced equally inadequate results with, for example, only three main frequency bands detected in the COMForTS-*ps* signal (Figure 2-5f and Figure 2-7, inset). In terms of resolution, however, maximum temporal resolution is found in the pure time-domain signal and maximum frequency resolution (Figure 2-7, inset) is attained in the Fourier transform of that entire signal. Using a single detection domain contributes minimal, but high resolution, information about the signal.

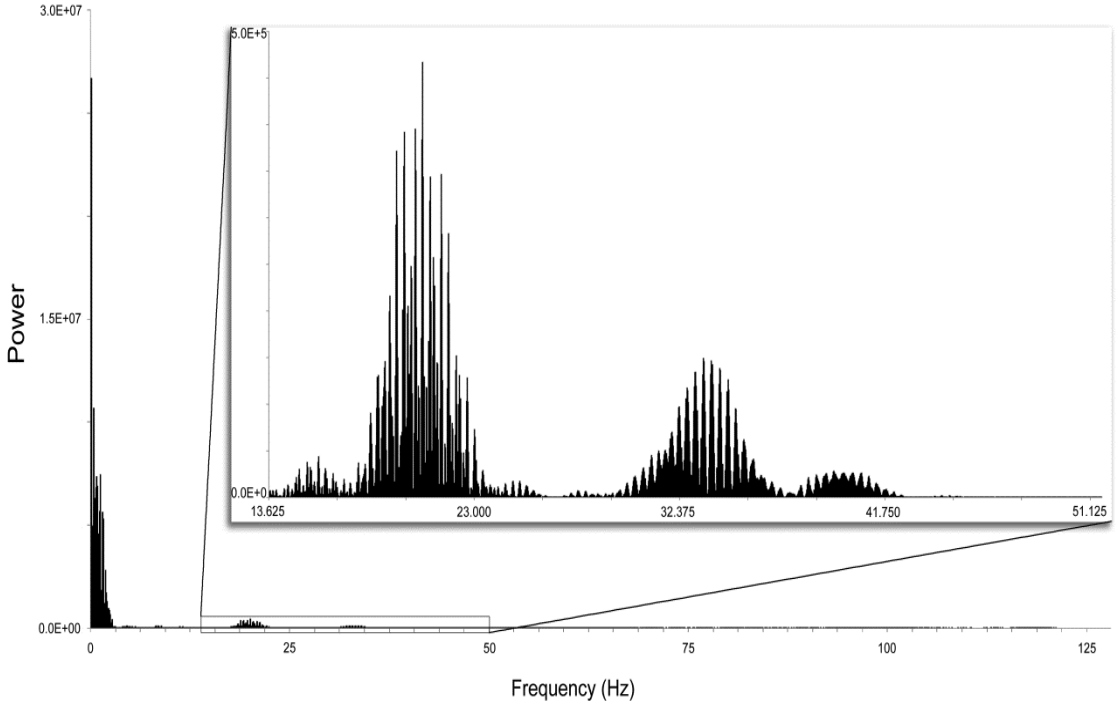


Figure 2-7: COMForTS-*ps* detector signal, high-resolution power spectrum. The range of frequencies shown are from zero Hertz to the Nyquist limit and (inset) the frequency range of applied pulses, 14 Hz to about 44 Hz. The detector sampling-rate was 256 Hz and the sampling time was 22 s.

2.4.4. COMForTS and digital signal processing

Because the distinguishing factor in COMForTS processing of detector signals is the combination of information from both domains by measurement of the frequencies present in small time-segments, it is necessary that some resolution is lost in both domains. The frequency resolution (df) of the Fourier transform is given by [97]:

$$df = \frac{1}{N} \times R \qquad \text{Eq. (2-6)}$$

where N is the number of data points used in the transform and R is the data sampling rate, from which it is evident that increasing N simultaneously reduces the resolution in time whilst increasing resolution in frequency (i.e. df is small). Whilst Fourier transforms were performed at 200 ms intervals, 256 data points (corresponding to a one-second sample of the detector signal) were needed to deliver adequate resolution in the frequency domain. Because the length of data used for transforms exceeded the length of data contained in a single transform interval, there is some overlap of transforms and an

associated increase in apparent peak width. This effect is most evident in the time width of COMForTS-processed peaks when compared to the widths of peaks in the raw detector signal as shown in Figure 2-5.

Evident also, in Figure 2-7, is some broadening of the component peaks along the frequency axis that, in this high-frequency-resolution plot, are resolved into discrete frequency components. This is undoubtedly the result of a lack of optimisation of the conditions chosen for the Discrete Fourier transform (DFT) and is a phenomenon known as 'leakage': A discrete Fourier transform uses a discrete number of data points (extracted from what is assumed to be a continuous data set) to return a discrete set of discrete frequencies. If any particular frequency actually present in the signal does not exactly match any of the discrete frequencies represented by the transform, the signal will be reported as the sum of two or more of the frequencies that *are* represented by the transform. This results in a 'spread' in the range of reported frequencies. These aspects of the DFT are well known and their effects are readily mitigated by appropriate processing techniques [101].

Whilst the function of our square-wave filter had been confirmed with pure rectangular pulses, little difference was observed in the filtered and unfiltered COMForTS results. In practice, the essentially square form of the injection pulses was distorted to a significant extent by diffusion, resulting in a signal that more closely resembled a sinusoidal waveform (see Figure 2-6) thereby greatly reducing the contributions of square-wave harmonics. This observation would be supported by Crabtree *et al* [48] who, while not then able offer an explanation, did note that observed harmonic contributions, when using SCOFT, were of both a lower order and a lower magnitude than those predicted in their own simulations. Nonetheless, the rudimentary processing methods applied in this experiment revealed low-level spurious frequencies (Figure 2-8) that were not anticipated by the author.

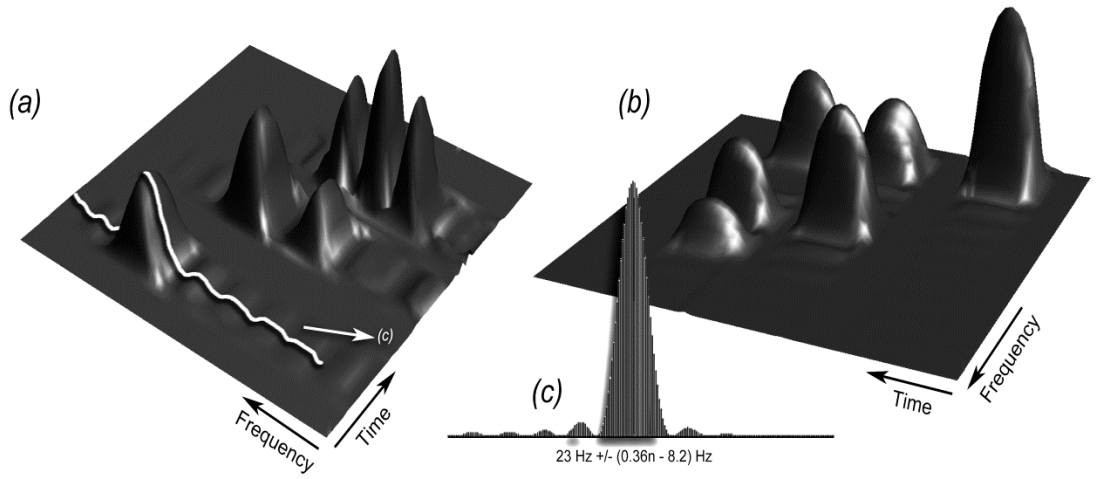


Figure 2-8: Surface plots of the COMForTS processed signals.
The signals result from COMForTS-*is* (a) and COMForTS-*ps* (b) simulations of the same six-component separation. A single frequency spectrum (c) obtained from the Fourier transform of the detector signal centred at the indicated retention time, displays a uniform series of overtones.

These processing artefacts were found to be of a uniform nature (Figure 2-8c) and of greater amplitude in the COMForTS-*is* chromatogram (Figure 2-8a) compared to COMForTS-*ps* chromatogram (Figure 2-8b). The reasons for this are yet to be established and it is unknown if the relationship would hold in all cases. However, in both COMForTS modes, the arithmetic mean of the applied pulses was not zero, and as the pulses decay, this mean value increases, as is evident in Figure 2-6. At the same time, the 'baseline' of a given signal segment does not necessarily have a slope of zero. These factors represent a constant or near constant contribution to the pure frequency component of the signal, creating an offset from the expected mean amplitude of zero. As for a square wave, this constant offset is represented by the Fourier transform as the sum of a series of frequency components as previously described in Section 2.1.2. Furthermore, standard apodization windows assume that while the data is finite, the signal is continuous, which is certainly not the case in chromatographic signals. That most of the interfering frequencies appear to take the form of overtones (Figure 2-8c) similar to those observed in un-windowed samples [41,97], and that the standard apodization windows employed appeared to have little effect, it is assumed that these interferences are due largely to inadequate apodization of the signal segments with smaller contributions from the zero-offset.

Even though the application of frequency analysis to discontinuous pulsed separations signals is entirely new, Felinger [32] describes a method that, by mirroring a signal sample, the time-broadening of COMForTS peaks would be halved and the effect of a sloping baseline is negated. Further reduction of this time broadening may also be achieved by cross-correlation with the pure time-domain signal and / or the pure frequency domain signal. The effect of a constant zero offset is easily dealt with by adjusting the mean of the detector signal sample to zero (removing the zero offset), which has no effect on the absolute amplitude of the frequency components. It would be reasonable to expect that these improvements, together with the application of state-of-the-art digital signal processing techniques could result in a doubling of the resolution in both the time and frequency domains as well as significant gains in signal to noise ratios afforded by filtering and / or compensation for the observed harmonics.

2.4.5. Sensitivity and noise

By imposing full pulse modulation, the pulse amplitude of component peaks is not appreciably altered and as the Fourier transform reports this same amplitude, there is no real sacrifice in terms of detector sensitivity. And though the digital signal processing methods applied in this simulation were of a very basic nature, the ratio of the signal to RMS noise in the COMForTS-*ps* chromatogram was about 2500:1 whereas the same ratio in the simulated detector signal was set to about 100:1; more than an order of magnitude improvement. This is not altogether surprising, given that other chromatographic applications of the Fast Fourier Transform (for example [75,83,102,103]) have been designed for this sole purpose. Three dimensional fully modulated GC separations processed using PARAFAC also improve signal to noise (by a factor of about ten) [90]. In contrast, methods employing partial pulse modulation have shown both a decrease in sensitivity and an increase in noise [92].

The performance of current COMForTS processing on signals with and without added noise was not found to be significantly different; indicating that the 'noise' in the COMForTS chromatograms consisted almost entirely of the uncompensated harmonics (Figure 2-8) already discussed. Removal of these harmonics with appropriate signal processing methods would yield a further

significant improvement in the already very high signal to noise of COMForTS separations.

2.4.6. Quantitative aspects of COMForTS

Within the context of the current examination of the principles of a proposed separations method based on the Fourier transform, it is not possible or justified to provide a detailed description or evaluation of the quantitative capabilities of COMForTS in practice. Nonetheless, it is possible to infer that because the Fourier transform is a linear operation and the amplitude of a sine wave is directly related to its Fourier magnitude [41], the amplitude (height) of pulsed signals will also be directly related to the Fourier magnitude. This sustains some direct relationship, whether linear or otherwise, to raw signal intensity and a quantitative measure of analyte concentration. Other aspects of the signal processing, including apodization and sampling functions may also affect the efficacy of quantitative analysis. In each of these cases, the mathematical processes are not only reversible and linear but are applied uniformly and with perfect reproducibility to the raw detector signal. There is little doubt that a stable and uniform relationship between analyte concentration and the height or area of COMForTS peaks will be established in practice.

2.4.7. COMForTS in three and more dimensions

The introduction into the analysis of frequency domain information by pulsed injection is by no means limited to two dimensions. It is possible, for example, to pulse injections between higher dimensions, utilising specific frequency ranges for each dimension or even to apply 'carrier' signals for binary coded dimension and time specific data that would then be carried through the separation by individual components. Such a scheme may be likened to the coding of digital FM radio signals and would afford the distinct possibility of rapid separations with much higher dimensionality. The Hadamard transform detection developed by Kaneta *et al.* [98] and used in capillary electrophoresis, to greatly increase sensitivity, is a cogent example of the practical application of a similar rationale that has seen a considerable range of applications extending into liquid and gas chromatography and beyond (for more recent developments and applications, see also [104-106]).

With a single detector, however, COMForTS is limited in the total separation time available due to the degradation of pulses over time. Nonetheless, as the retention time in a given dimension is related to the detected pulse frequency, the use of multiple inter (or intra-) dimensional detectors is no longer a moot point as it is with conventional online separations. Furthermore, the dimensionality of a system is only limited by the number of available detectors and the total analysis time only by the sum of the lifetime of pulses within each individual dimension. Even though the COMForTS simulation package was programmed with the capacity to simulate separations in higher dimensions, models of three-dimensional separations produced detector signals with measureable frequencies but of a complexity that was beyond the processing methods employed for this test of basic concepts. Such complexities may well provide an opportunity for the application of PARAFAC to frequency-domain data and exemplifies the growing need for collaboration between chemists and signal-processing professionals [107,108].

2.4.8. Analysis time, resolution and peak capacity

COMForTS processing was performed by calculating the Fourier transform of segments of the detector signal at prescribed intervals, which correspond to fractions collected for offline comprehensive analysis. With an effective sampling rate of five samples per second, a conventional offline comprehensive two-dimensional separation would require, by comparison, an analysis time nearly thirty times longer (see Eq. (1-1)) – if indeed it were possible to collect injectable fractions at 200 millisecond intervals. Using this offline analysis as a basis for comparison, we can begin to examine the more important aspect of specific peak capacities (peak capacity per unit of analysis time, also known as peak capacity production) of some multidimensional separations methods. Specific peak capacities were calculated relative to an equivalent comprehensive separation assuming that each method utilises separation dimensions with the same efficiency and peak capacity as the current example (Figure 2-4 and Section 2.4.1). The comparative results are listed in Table 2-1.

Separation Method	Relative Specific Peak Capacity* Two Dimensional	Three Dimensional
Comprehensive Offline[a]	1	0.52
Comprehensive Online Pulse Modulated with PARAFAC[b]	3.8	14
COMForTS with the current rudimentary processing[c]	3.7	26
COMForTS with optimised processing[d]	8.2	87

* Separation efficiencies are assumed to be the same as the simulated example
a. Based on 5.1 peaks per second in each dimension and 5 'cuts' per second
b. Based on 4 peaks per second per asymmetrical [90,92] dimension
c. Based on 1.8 peaks per second per symmetrical dimension
d. Based on 2.8 peaks per second (50% of the expected improvement)

Table 2-1: Estimated relative specific peak capacities.

Both the method presented here and other pulse modulation methods [1, 21] sacrifice some peak capacity due to the need to make unusually long injections such that a minimum number of pulse cycles are applied to each peak. These losses, which are greater for COMForTS, are more than compensated by the enhanced capacity afforded by higher dimensionality and the time-advantage gained by online separations. The COMForTS examples presented here have a specific peak capacity that is only slightly less, in two dimensions, than would be expected of a typical asymmetric pulse modulated separation with PARAFAC. If, however, only modest (50% of the expected) improvements are achieved by optimisation of COMForTS signal processing, the specific peak capacity of COMForTS has the potential to reach extraordinarily high levels. Furthermore, because COMForTS is able to resolve completely overlapped peaks as well as peaks that are not physically resolved in any dimension (provided only that no two specific peaks co-elute in all dimensions), a significantly greater proportion of this theoretical peak capacity is available to COMForTS systems. Given that statistical overlap theory predicts that a reasonably high percentage of components in multidimensional separations will elute as doublets or triplets [26,109], the facility of COMForTS to resolve these components is a marked advantage.

The primary disadvantage of the proposed method is that total peak capacity is, except for the first dimension, limited by the need for short, high efficiency separations. The alternative is to implement either multi-point detection or high total peak capacity (asymmetric) first dimension separations combined with shorter, symmetrical separations in succeeding dimensions.

2.4.9. Signal acquisition and processing requirements

COMForTS necessitates detectors that exhibit extremely fast (ms) response and allow sampling rates (possibly) in the kHz range (e.g. MS-TOF and perhaps FTMS [110]). Some detectors may require miniaturised flows cells (e.g. UV and IR) in order to produce sufficiently rapid response times. Upgraded analogue to digital conversion capabilities will be needed in order to meet the required signal-sampling rates. It is also evident that, while not often used with such short columns, the effects of gradient elution must be considered. By varying the migration speed of the analytes, solvent gradients will, when using COMForTS-*is* (but not COMForTS-*ps*), have a significant impact on the range of frequencies recorded over the width of individual analyte peaks. Separations systems such as GC in which eluent flow rates vary along the length of the column, will exhibit similar uniform changes in absolute linear velocities (in both COMForTS modes). Such behaviour is compatible with the demonstrated theory of COMForTS. In the case of COMForTS-*is* separations the result is a distortion of peak shapes along a time-frequency vector dictated by the solvent or flow-rate gradient and does not represent a material problem. COMForTS-*ps* separations are still achieved in variable flow-rate systems but the relationship between detected frequencies and first dimension retention times (Eq. (2-3) and Section 2.4.2) is complicated but not unresolvable.

Given the range and extent of processing requirements, COMForTS may appear to be computationally onerous; but not so much so that it might strain the capabilities of current personal computers. For example, the Fourier analyses performed in this experiment were executed in less than one-tenth real time. There will be an increased workload concomitant with higher retention and frequency resolution, harmonics-compensation and detection and decoding of binary sequences. Nonetheless, the workload is not anticipated to increase by more than a factor of four and most current multi-core CPU's would be capable of real-time signal processing and generation of two-dimensional COMForTS chromatograms.

2.4.10. Exploration of separation conditions.

Following from the above simulations, some exploration was made into conditions under which the COMForTS method could perhaps operate more efficiently, especially when very large numbers of analytes are involved.

To that end, a simulated separation of fifty analytes was constructed with arbitrary retention times defined much as one would expect from the first dimension separation of a series of five oligomers followed by the second dimension separation of their ten diastereomers (Table 2-2).

Component #	Retention Time (s)			Δt	Component #	Retention Time (s)			Δt
	D_1	D_2	Overall			D_1	D_2	Overall	
1	100.0	6.0	106.0		26	120.0	13.3	133.3	1.0
2	100.0	7.0	107.0	1.0	27	120.0	14.3	134.3	1.0
3	100.0	8.0	108.0	1.0	28	120.0	15.4	135.4	1.1
4	100.0	9.0	109.0	1.0	31	130.0	6.0	136.0	0.6
5	100.0	10.0	110.0	1.0	29	120.0	16.6	136.6	0.6
6	100.0	11.0	111.0	1.0	32	130.0	7.0	137.0	0.4
7	100.0	12.0	112.0	1.0	30	120.0	17.4	137.4	0.4
8	100.0	13.0	113.0	1.0	33	130.0	8.0	138.0	0.6
9	100.0	15.5	115.5	2.5	34	130.0	9.0	139.0	1.0
11	110.0	5.5	115.5	0.0	35	130.0	10.0	140.0	1.0
10	100.0	16.5	116.5	1.0	36	130.0	11.0	141.0	1.0
12	110.0	6.5	116.5	0.0	37	130.0	12.0	142.0	1.0
13	110.0	7.5	117.5	1.0	38	130.0	13.0	143.0	1.0
14	110.0	8.5	118.5	1.0	39	130.0	14.0	144.0	1.0
15	110.0	9.5	119.5	1.0	40	130.0	15.0	145.0	1.0
16	110.0	10.5	120.5	1.0	41	140.0	5.5	145.5	0.5
17	110.0	11.5	121.5	1.0	42	140.0	6.3	146.3	0.8
18	110.0	13.9	123.9	2.4	43	140.0	7.8	147.8	1.5
19	110.0	16.7	126.7	2.8	44	140.0	9.8	149.8	2.0
21	120.0	6.8	126.8	0.1	45	140.0	11.3	151.3	1.4
20	110.0	17.7	127.7	0.9	46	140.0	11.8	151.8	0.5
22	120.0	7.8	127.8	0.1	47	140.0	12.5	152.5	0.8
23	120.0	10.2	130.2	2.4	48	140.0	14.5	154.5	2.0
24	120.0	11.2	131.2	1.0	49	140.0	16.3	156.3	1.8
25	120.0	12.3	132.3	1.1	50	140.0	17.0	157.0	0.7

Table 2-2: Peak table of fifty overlapping D_1 and D_2 retention times.
Overall retention times and the differences in overall retention between adjacent peaks are also indicated for each of the 50 components.

Similar to the six-component simulation, a COMForTS simulation was performed with the following liquid chromatography conditions and parameters:

- First dimension column: 150 × 4.6 mm
- First dimension band broadening rate: 0.08 s/s (moderately efficient)
- Second dimension column: 25 × 3 mm

- Second dimension band broadening rate: 0.01 s/s (extremely efficient)
- Injection Volume: 10 μL
- Eluent Flow Rate: 2.0 mL/min
- Run Times: 150 s in the first dimension and 20 s in the second.
- Total analysis time: 170 s
- Detector sampling rate: 64 Hz
- Signal generation resolution: 128 Hz (1st Dimension), 64 Hz (2nd Dimension)
- COMForTS-*is* pulse frequency: 30 Hz
- Total instrument noise was set to 1% RMS (using a sample of noise recorded from an HPLC-DAD instrument)
- A 10% sloping baseline was also incorporated into the detector signal.
- One iteration of triangular apodization was applied to time samples prior to the Fourier transform

The COMForTS 3.0 configuration files that define this simulation are contained on the digital media attached to this thesis within the folder named "…\Sample Data\COMForTS3" (or, if installed on your computer "My Documents\COMForTS\Samples\Modelling Examples"), for each COMForTS mode:

- Asymmetric 50 - 150mmD1_is.COMForTS
- Asymmetric 50 - 150mmD1_ps.COMForTS

Whilst modelling was performed in both COMForTS-*is* and COMForTS-*ps* modes, the COMForTS-*ps* results were less remarkable and not presented here.

The generated first and second dimension detector signals are shown in Figure 2-9.

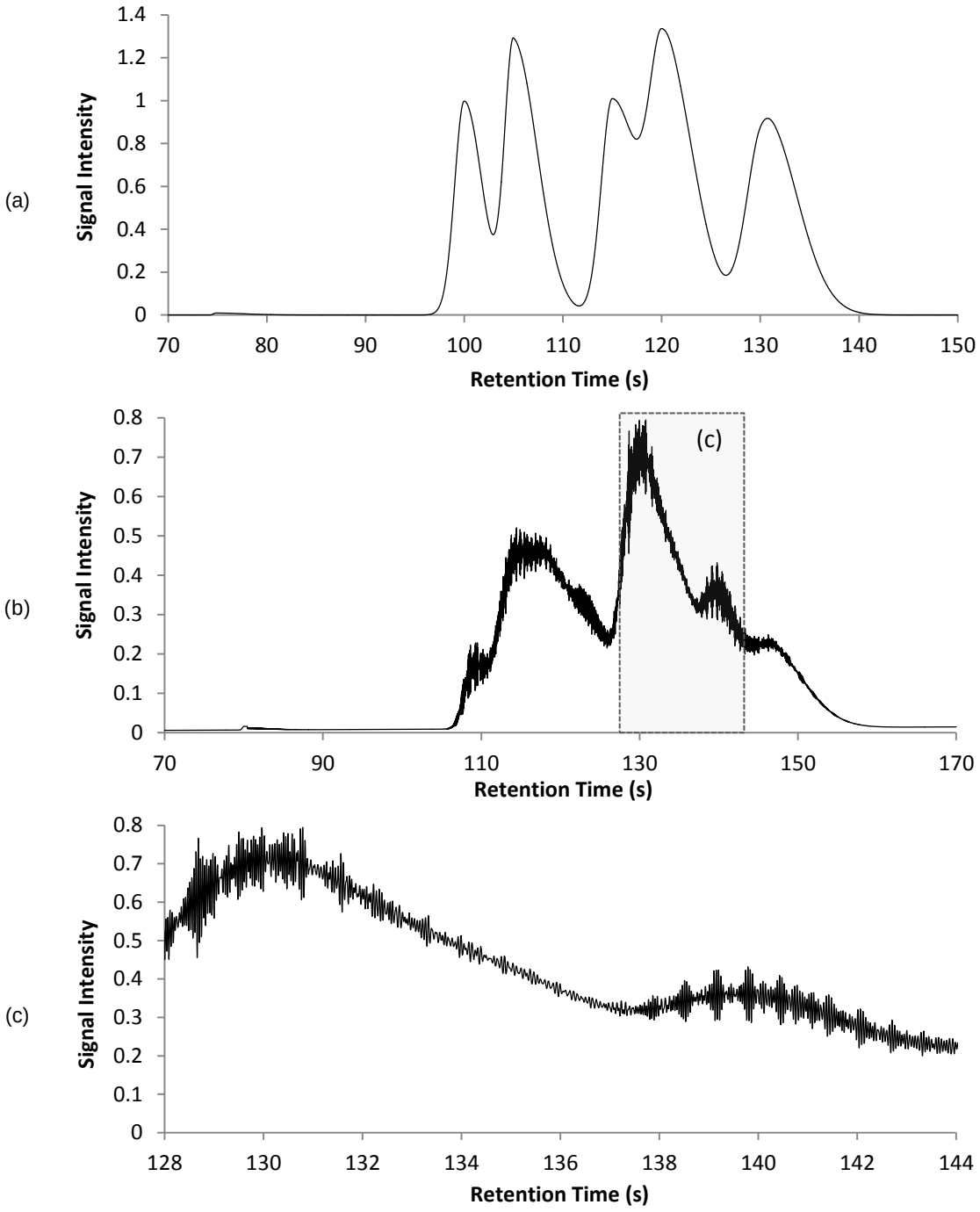


Figure 2-9: Asymmetric COMForTS-*is* separation of fifty components.
(a) Separation result at the end of the first dimension (b) Separation result at the end of the second dimension and (c) Enlarged shaded section of (b) showing strong low frequency (<1 Hz) harmonics and low pulse amplitude relative to signal strength.

Five broad, partially resolved peaks are evident in the first dimension signal (Figure 2-9a) and represent each of the five oligomers. When pulse-injected into a very short and extremely efficient column, the breadth of the first dimension peaks creates a longer time-domain signal that allows increased

resolution in the frequency domain. Even though the amplitude of the pulses is relatively small (compare Figure 2-9b with Figure 2-5c and d), the frequencies present remained readily resolvable, which is evident in the frequency/time density plots shown in Figure 2-10.

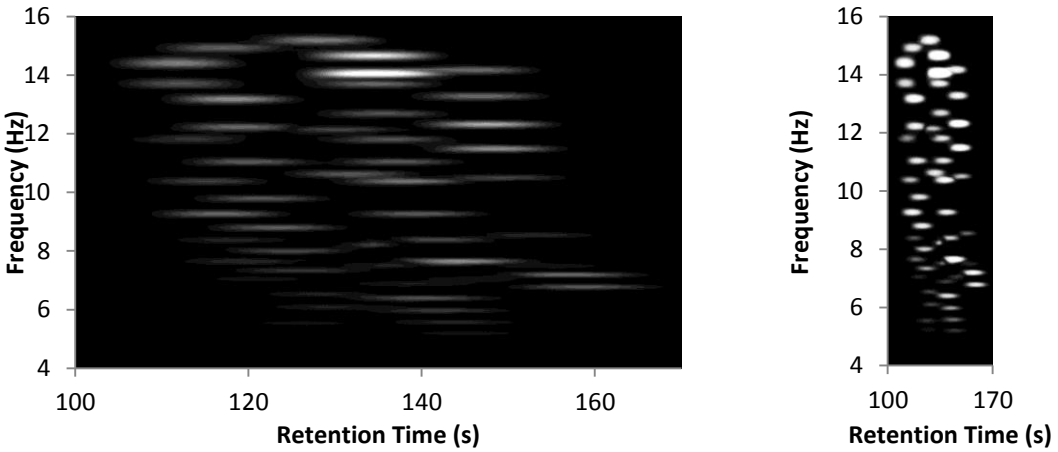


Figure 2-10: COMForTS-*is* frequency/time chromatogram.
An asymmetric COMForTS-*is* separation of fifty arbitrary components. (a) Density plot of raw result (b) Scale- and contrast-enhanced density plot, facilitating visual peak identification.

Whilst triangular apodization was found to be extremely effective in the reduction of the particular harmonics of this separation, Figure 2-9c clearly illustrates extensive low frequency harmonics. These are probably the result of the summation of fundamental 'analyte frequencies'. Whilst these visually evident harmonics fall outside the range of frequencies that could be physically attributed to an analyte (and therefore do not appear within the bounds of Figure 2-10), it remains unclear if these harmonics have overtones that do fall within the 'analyte range'. It is equally possible that less evident harmonics will masquerade as analyte peaks. However, that the expected fifty peaks are resolved is evidence in favour of an interference-free separation.

This simulation had a peak production rate of little more than 1 peak per second – a rate that was achieved with a symmetrical separation – but with significantly improved peak capacity.

2.5. Conclusions

These simulations have identified a number of the advantages and constraints that the practical application of physically pulsed COMForTS separations will enjoy or should deal with. At present, the main obstacles to the assembly and construction of an instrument are the need to produce programmatically defined pulses of relatively high frequency and to be able to detect and measure those pulses in high efficiency separations. That is not to say that this requisite technology does not already exist as does the technology for high-speed, high-data rate analogue to digital signal conversion but that it has simply not yet been applied in this field. Our work shows that these practical limitations can be overcome and that it is possible to achieve, with the aid of advanced signal processing methods, extremely rapid, sensitive and powerful separations.

The method of Comprehensive Online Multidimensional Frequency Transform Separations has been demonstrated by semi-empirical modelling and simulation to be theoretically sound and probably practically feasible within the very near future. Calculations show that it meets all the self-imposed requirements of comprehensive multidimensional separations. Due to its ability to numerically resolve components that are not physically resolved in any dimension, the method exhibits the potential for very high peak capacities per unit of analysis time together with high sensitivity, low noise and the facility to correlate peaks detected with multiple detectors. Conditions have been identified under which the total peak capacity can be significantly enhanced by increasing the number (or duration) of pulses recorded and used in the frequency transform. These unique advantages of the COMForTS method define its potential as an attractive possibility for high-speed multidimensional separations, where physical separation of the analytes is not required. Further explorations of this potential as well as its inherent possibilities for increased total peak capacity are undoubtedly warranted.

Chapter 3.
Proof of Concept I:
Principles of Operation

3.1. Overview

Whether analyte 'pulses' are produced via pulsed injection or arrayed-detection, the central precept of COMForTS is that analytes may be resolved in both the time and frequency domains. In either case, it is first necessary to demonstrate that the detected frequencies are uniformly related to analyte velocities (or injection frequency) and that the detected retention times correlate to conventionally measured retention times. Only once these conditions are met, is it then possible to investigate the efficacy of these measures with respect to the negation of the effects of peak wrap-around.

3.1.1. Pulsed injection

Observations of the COMForTS 3.0 modelled conditions in the previous chapter showed that higher frequencies gave improved resolution as did the number of pulse cycles recorded. These observed characteristics are supported by the literature [44]. At the same time, in chromatographic systems, higher frequency pulses diffused more quickly thus requiring separation dimensions of the highest efficiencies.

The use of switching valves to generate pulses presents particular practical difficulties: The valves must be able to operate at high frequencies and under high pressure and, in the case of COMForTS-*is*, the stationary phases must extend to the switching surfaces of the valves (Figure 2-1a). This approach requires extremely efficient separations at least in the second dimension.

Because the configuration of stationary phases required for COMForTS-*is* is not commercially available and not easily manufactured, the following experiments with pulsed injection sought to demonstrate that the measured pulse frequencies are a function of applied pulse frequency according to the principles established for COMForTS-*ps*.

3.1.2. Multipoint detection

If findings of the modelling experiments are extended to the virtually pulsed method (of multi-point detection), the requirement for high frequency translates to very closely spaced detectors. Similarly, the requirement for large numbers of pulses translates to the need for a large number of detectors.

It was concluded that some form of optical detection would be the simplest way to implement on-column detection and that a charge coupled device (CCD) could act as a large array of detectors by treating each pixel as an individual detector. In fact, this basic idea had already been demonstrated by McReynolds and Shippy [88,111] with respect to SCOFT and Hadamard transform fluorescence detection in capillary electrophoresis.

Absorbance detection in the UV-Vis range was chosen over fluorescence for both its greater potential for optical simplicity and that a greater variety of analytes would be detectable. A light source with a very narrow wavelength band of emission, combined with placement of a silica capillary column directly onto the face of the CCD would eliminate the need for filters and lenses with only a moderate loss in sensitivity.

There remained the problem, however that the physical length of chromatographic peaks, even in capillary columns is much greater than the length of peaks in micro-channel electrophoresis. Crabtree [48] had used a slit spacing (equivalent to detector spacing) of 700 μm with 300 μm slits while Kwok and Manz [86] used 40 μm slits, 70 μm apart. To realise adequate sensitivity in the frequency domain, all SCOFT detection had been carried out using 50-55 detectors. Manz also concluded that, when measuring the velocities of microspheres of 1 μm diameter, the optimum distance between detectors would be about 50 μm. Whilst not specifically stated by any of these researchers, it is understood that it is the physical widths of peaks that limits our ability to sum their signals. If the detectors are too close, individual

detector signals would become 'muddied' by signals from adjacent detectors. In a capillary liquid chromatography separation, physical peak widths within the column may be a millimetre to several centimetres in length and the required length of the detector array would be between 10 and 400 cm.

A method for combining signals from very closely spaced detectors (MCSCSD) was developed (see Appendix D) so that 'clear' frequency spectra could be obtained via the short time Fourier transform (1.2.1). Unfortunately, MCSCSD must lose substantial resolution in the frequency domain. MCSCSD was abandoned when an alternate, superior approach was devised. Instead of summing (or otherwise combining) the individual time-domain chromatograms from each detector, we can insert those individual chromatograms into a two-dimensional array of multiple, independent detector signals (detector number, on the y-axis) over time (seconds, on the x-axis). Analyte peaks moving past the individual detectors over time will then appear as straight lines where the slopes of these lines (detector count / s) gives us the analyte frequencies and the x-intercepts of the lines yield overall retention times. The human mind has little capacity to distinguish numerically, as does the Fourier transform, between multiple frequencies in a time domain signal. On the other hand, we do possess a natural facility to recognise patterns in images and are remarkably adept at distinguishing straight-line features - such as the lines produced by analyte peaks in arrayed detector signals. Ironically, it is somewhat less easy to obtain a rigorous mathematical description of these 'peak lines'.

In the course of this project, several algorithms were developed to meet this challenge. Each of these algorithms achieved varying degrees of success, depending upon the complexity of the straight-line systems within the arrayed chromatograms. The number of peak lines, their spacing and degree of overlap had a significant influence on their detectability. A simple line-search algorithm worked very well when peak lines did not overlap. An enhancement of this algorithm incorporated a 'genetic qualification' of the discovered lines: only the straightest lines between peak maxima were allowed to 'survive'. 'Unfit' lines were discarded and their assigned peak maxima returned to the 'genetic pool' from which new lines could be evolved. This 'genetic algorithm' worked surprisingly well, but like others of its ilk, could be notoriously slow. The algorithm used here was embodied in the "LINEAR2"

module of COMForTS 4.0 and is dealt with in more detail in Section 6.4.3. LINEAR2 is the genetic version of the simple line search algorithm that was implemented as "LINEAR". LINEAR was used for the initial processing of signals in these experiments and produced virtually identical results. LINEAR2 worked well with chromatograms of moderate complexity but either failed when confronted with higher complexity or ignored the limitations on human lifetimes.

A stubborn conviction that these straight lines were obvious, and therefore amenable to mathematical description, lead to an altogether different approach. The array of chromatograms may be viewed in three dimensions: time (x), detector number (y) and absorbance (z), where peak lines become ridges, or linear 'hills'. Berthon-Jones [private communication, 2013] suggested a more 'holistic', or 'whole-data' approach based on the concept of rotating or skewing the image whilst looking at it 'edge-on': Where no peak line is present, the 'horizon' would be flat, but as a peak line is rotated into one's line-of-sight, a large mound (appearing as a 'normal' chromatographic peak) would appear. One can imagine though, that in the region where two peak lines overlap, the view would not be a 'hill' but a more-or-less flat-topped or sloping 'mesa'. If, however, one were to take some point on the x-axis and *look* in different directions at the *sum* of all the absorbance values in each direction, that sum would be zero when no peak is present but would increase dramatically and reach a maximum when the 'looking' direction corresponds to the line's direction. Summing along the looking direction would thus produce a maximum regardless of whether one or more lines had contributed (by overlap) to that sum. This is the basic implementation of the algorithm later identified as the discrete Radon transform as described in Section 1.2.2. Its implementation in the "RADON" module of COMForTS 4.0 is discussed in Section 6.4.2.

3.2. Materials and Methods

3.2.1. Pulsed injections

Mobile phase compositions, analyte retention times and column performance were assessed with a Shimadzu HPLC system consisting of two mobile phase pumps (LC-10AD), a solvent mixer (FCV-10AL), auto-injector (SIL-10AD) and UV/Vis diode array detector (SPD-10A) operating at 257 nm and 275 nm.

For testing pulsed injections, two Shimadzu mobile phase pumps (LC-10AD) were connected directly to opposing ports of a VICI Valco Two Position valve (six port valve equipped with an actuator control module) with the valve's outlet connected to a 25 × 4.6 mm C18 Monolithic column. Post-column detection was conducted with the same Shimadzu UV/Vis diode array detector (SPD-10A) operating at 257 nm and 275 nm. Simultaneous on-column detection was performed with the COMForTS Electrical Conductivity Detector (ECD).

The COMForTS ECD was a custom-built Wheatstone bridge based detector with digital output. The design is described in the schematics in Appendix A and the completed arrangement shown in Appendix B. Plain text files of the detector's firmware (assembly source code (.asm) and machine code (.hex)) can be found in the Digital Thesis Materials folder named "…\Software\Source Code\COMForTS ECD\COMForTS_ECD_Firmware". Gold-plated electrodes were inserted into the stationary phase through the column wall and secured with epoxy resin. The electrodes were positioned at the longitudinal midpoint of the column (approximately 12.5 mm from the column inlet). Data acquisition was via the COMForTS ECD Control software running on a laptop computer equipped with a dual core 1GHz Intel Celeron CPU, 1 GB RAM, running Microsoft Windows XP (Service Pack 3) (Microsoft Corp., Redmond WA).

The COMForTS ECD Control software was written in Microsoft Visual Basic .NET 2010 as a proprietary executable program. Plotting functions were performed in Microsoft Excel 2010. COMForTS ECD Control runs natively on Microsoft Windows 32- or 64-bit operating systems from Windows XP on, and

requires the Microsoft .NET Client Framework 4.0. Microsoft Visual Studio 2010 Visual Basic source code and executables for COMForTS ECD Control can be found in the Digital Thesis Materials folder named "…\Software\Source Code\COMForTS ECD\COMForTS_ECD_Control".

Mobile phases were made with HPLC grade solvents and 80 MΩ Millipore water. All reagents were analytical grade.

3.2.1.1. Testing pulsed injections

Two mobile phase pumps (A and B) were connected to opposing ports on a six-port valve such that mobile phase from only one pump at a time was passed through the column to the detector. A mobile phase (A) was prepared and divided between the solvent reservoirs for each pump, at which point, an 'analyte' (salicylic acid) was added to one portion (B) of the mobile phase. Mobile phase 'B' thus serves as a first dimension 'peak' that can then be 'pulse-injected' into the second dimension column. The 'width' of this peak, and hence the number of pulses introduced into the system was controllable by limiting the overall time during which pulses were applied between solvents A and B. Both on-column (ECD) and post-column (UV) detection were used in an instrument configuration equivalent to a combination of those described in Figure 2-1 (b) and Figure 2-2. Both configurations correspond to COMForTS-*ps* mode of operation in that we expect in both cases that the measured frequencies will be equal to the applied frequencies.

The Shimadzu software was programmed, using its external events function, to toggle the six-port valve between solvents A and B at regular intervals (beginning with A). The maximum frequency achievable was restricted by the software to 1 / (0.01 min × 2) = 0.83 Hz. The pulse frequencies employed were 0.83 Hz and 0.42 Hz.

The average rate of band broadening (BBR) of the C18 monolith was measured for the homologous series of ethyl, methyl and propyl para-hydroxybenzoic acid (parabens) with methanol / water mobile phases for retention times up to three minutes. Injections (1 μL) of ~20 mg/L parabens yielded an average BBR of 0.1 with post-separation UV absorbance detection @ 275 nm. This measurement was made before modification of the column by the insertion of the ECD electrodes.

The retention times of salicylic acid and acetyl salicylic acid were determined by conventional injection to be 0.44 and 0.62 minutes respectively, with a mobile phase of 20% methanol in water buffered with acetate to pH ~ 4. This mobile phase was then used for the pulsed injections as mobile phase 'A'.

3.2.2. Multipoint detection

3.2.2.1. Instrumentation and Data Processing

A prototypical instrument and detector were constructed based on a capillary liquid chromatographic separation with CCD-based UV absorbance detection (Figure 3-1). The design, development and qualification process and means of overcoming the above difficulties, in particular the physical detector spacing requirement, are elaborated upon in Chapter 5 COMForTS Instrumentation, and in Chapter 6 Software for COMForTS Chromatographic Analysis. Details of the components of the COMForTS instrument itself and the sources of its components are provided in Appendix C.

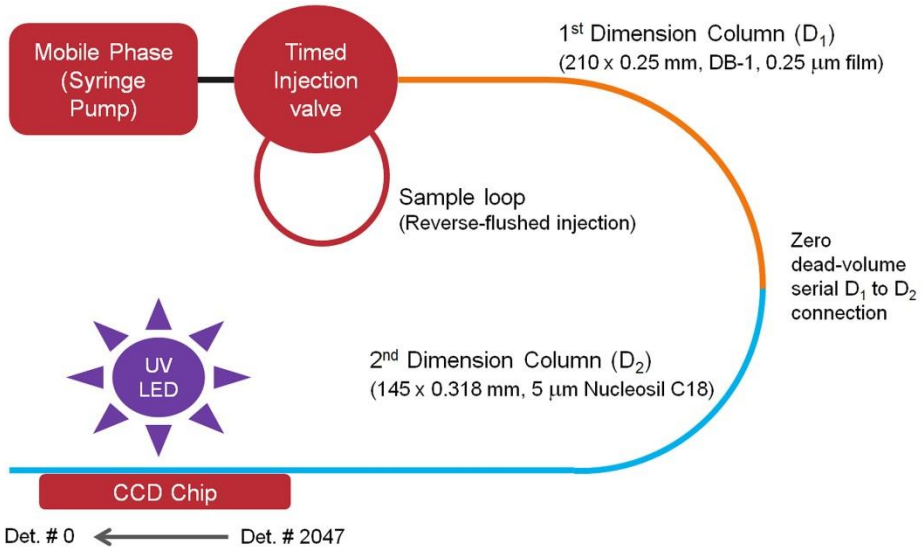


Figure 3-1: Schematic representation of the COMForTS Instrument.
D_1 directly abutted the injection valve outlet and was serially connected to D_2 via a zero dead volume connector. D_2 was placed directly against the CCD chip on its long axis, which consists of 2048 pixel lines. Each pixel line was treated as an individual detector.

Instrument control, data collection and processing were performed by the in-house developed COMForTS 4.0 Chromatography Suite (using the 'Radon' processing path), which is discussed in detail in Chapter 6. The COMForTS 4.0 Chromatography Suite software (source and executable code) is included in

the digital portion of this thesis. COMForTS 4.0 can be installed on the reader's computer by running the program "COMForTS_Installler.exe" from the root folder of the digital media. Representative raw data and finished chromatograms of interest are also provided in the "Sample Data" folder.

Data acquisition was carried out (using only the CICADA module of COMForTS 4.0) on a laptop computer equipped with a dual core 1GHz Intel® Celeron CPU (Intel Corporation, Santa Clara, California, USA) and 1 GB RAM running Microsoft Windows XP (Service Pack 3) (Microsoft Corp., Redmond WA).

This separate computer was used only because the Windows driver software for the CCD required Microsoft Windows XP. However, it did not have sufficient processing power to generate or process chromatograms from the raw data. All data processing and calculations were performed on another computer using the COMForTS 4.0 Chromatography Suite, written in Microsoft Visual C# .NET 2010 as a proprietary executable program. Plotting functions were performed by COMForTS 4.0 and Microsoft Excel 2010.

COMForTS 4.0 runs natively on Microsoft Windows 32- or 64-bit operating systems from Windows XP on, and requires the Microsoft .NET Client Framework 4.0 or above and Microsoft XNA Framework 4.0. The signals processing work presented here was conducted using a 64-bit version of COMForTS 4.0 running under Microsoft Windows 7 64-bit on a personal computer equipped with an Intel® i7-990X 3.4 GHz CPU (6 cores, 12 logical processors) and 12 GB DDR3 RAM.

The following detector conditions were employed, unless otherwise stated:

- Physical detector array: 2048 × 96, 14 μm square detectors (pixels)
- Absorbance wavelength: 255 nm ± 10 nm
- 255 nm UV LED current: 2 LEDs, 10 – 20 mA each (maximum output at 30 mA)
- Absorbance integration time: 150 ms
- CCD Amplifier, gain: 10 (maximum); mode: low noise (mode = 0)
- CCD Line-binning mode: binned (for 2048 detectors of 96 pixels each).
- CCD Binned Line count: 1 (i.e. no row binning)
- Sampling rate: 4 Hz

Basic performance qualification of the COMForTS instrument was established with respect to (unloaded) flow rate precision and accuracy, injection volume (injection timings) and fundamental operation of the detector (see Section 5.4).

3.2.2.2. Chromatography

As shown in Figure 3-1, two-dimensional liquid chromatographic separations were performed with serially connected capillary columns (D_1 and D_2) with on-column detection performed on D_2. The first dimension column (D_1) was a 210 × 0.250 mm open tubular capillary with 0.25 μm film thickness DB-1 stationary phase (cut from a GC column supplied by J&W Scientific, Folsom, California). The second dimension column (D_2) was a 145 × 0.318 mm fused silica capillary packed with 5μm Nucleosil C18 particles (see Section 5.3.4)

Both capillary columns were prepared in-house and fitted with standard HPLC 1/18" stainless steel connectors and ferrules with the aid of sleeves over the capillary ends, cut from 508 μm ID (orange) PEEK tubing (see Figure 5-3 and Figure 5-4).

All solvents were HPLC grade and all reagents were analytical grade. 80 MΩ Millipore water was used in all mobile phases. The following chromatographic conditions were employed unless otherwise stated:

- Mobile phase: 100% Acetonitrile
- Mobile phase flow rate: 5 μL/min
- Injection volume: 150 nL (injected over 1867 ms)

A solution of four polynuclear aromatic hydrocarbons (PAHs) in acetone was prepared as described in Table 3-1. This solution was known as "Stock X".

PAH	Concentration (g/L)
Naphthalene	1.3
Anthracene	0.48
Pyrene	0.48
Chrysene	0.47

Table 3-1: Concentrations of four PAHs in acetone, solution "Stock X".

3.2.2.3. Testing the ability of the COMForTS method to measure analyte frequencies and retention times.

The fundamental precept of COMForTS is that wrap-around can be overcome by simultaneous measurement in both the frequency and time domains (Chapter 1). To assess the ability of the method to measure both analyte pulse frequencies and analyte retention times, we again employed the characteristics of an unretained analyte. In this instance, the conditions imposed were:

1. A constant, known frequency was maintained using an unretained analyte at a constant mobile phase flow rate. By setting the frequency, we set the second dimension velocity and hence second-dimension retention time.
2. The apparent 'retention times' of the analyte (acetone) were controlled by making multiple injections of the analyte at controlled time intervals. As all 'analytes' are the same unretained acetone, the interval between injections may be interpreted as the first dimension retention time.

Measured frequencies in the second dimension stationary phase should thus correspond to a single frequency that also corresponds to a single second-dimension retention time. The first dimension retention times should match to the intervals between injections. Using the *intervals* between retention times automatically corrects for any error that would otherwise by introduced in the measurement of void times and injection delay times.

3.2.2.4. Testing the ability of the COMForTS method to resolve multiple analytes of differing retention over two separation dimensions to generate time/time chromatograms.

In the previous experiment, the 'analyte' frequencies were held constant and we were concerned only with being able to locate 'hot spots' of that frequency in time. Naturally, COMForTS depends upon the ability to resolve multiple different frequencies that are present in the chromatogram at different times.

This aspect of COMForTS was tested by injection of a solution of four PAHs in acetone (solution "Stock X"), using a mobile phase of 70% acetonitrile in water. These analytes (including acetone) are unresolved by the D_1 column and co-elute into the second dimension in which the analytes are resolved. We also

expect to be able to use the measured frequencies to determine the differing analyte velocities in the second dimension and hence the second dimension contribution to the total retention time. We can subsequently deduce the first dimension retention time. In this way, a conventional time/time chromatogram can be generated and compared to the known conditions.

3.3. Results and Discussion

3.3.1. Pulsed injection

This experiment was designed to replicate the pulsed injection of a first dimension analyte peak into the second dimension column, comparing applied pulse frequencies to measured pulse frequencies with both on-column and post-column detection.

This real-world experiment was modelled using COMForTS 3.0 in post-separation detection mode. The pulse rate was set to be (effectively) constant for the duration of the separation and the first dimension column parameters set such that the modelled first dimension had no effect on the separation. The COMForTS 3.0 configuration file that defines this simulation is contained on the digital media attached to this thesis within the folder named "…\Sample Data\COMForTS3" (or, if installed on your computer "My Documents\COMForTS\Samples\Modelling Examples"):

"Salicylic_&_acetyl_salicylic_acids_C18 monolith_COMForTS-ps.COMForTS"

Modelling was performed at a fixed frequency determined by the smallest programmable interval between valve-switching events: 0.01 minutes, corresponding to a maximum nominal frequency of 0.83 Hz. Whilst the modelling results were positive, indicating that measureable 0.8 Hz signals would be obtainable, the control of the switching of the valve itself was noticeably irregular. Nonetheless, the applied physical pulses were observed to approximate a 1 Hz signal.

Unfortunately, the background conductivity of the buffered mobile phase, combined with the low conductivity of salicylic acid necessitated that mobile

phase B contain 1 g/L salicylic acid. Acetyl salicylic acid at this (almost saturated) concentration produced no observable response in the ECD. In-separation and post-separation results were therefore only obtained for salicylic acid. At this concentration however, even small injections (pulses) of mobile phase B (1 g/L salicylic acid) appeared to result in significant column overload.

The results obtained when a pulse rate of 0.42 Hz was applied between mobile phases A and B over a period of 120 s are shown in Figure 3-2 (post-separation detection) and Figure 3-3 (in-separation detection).

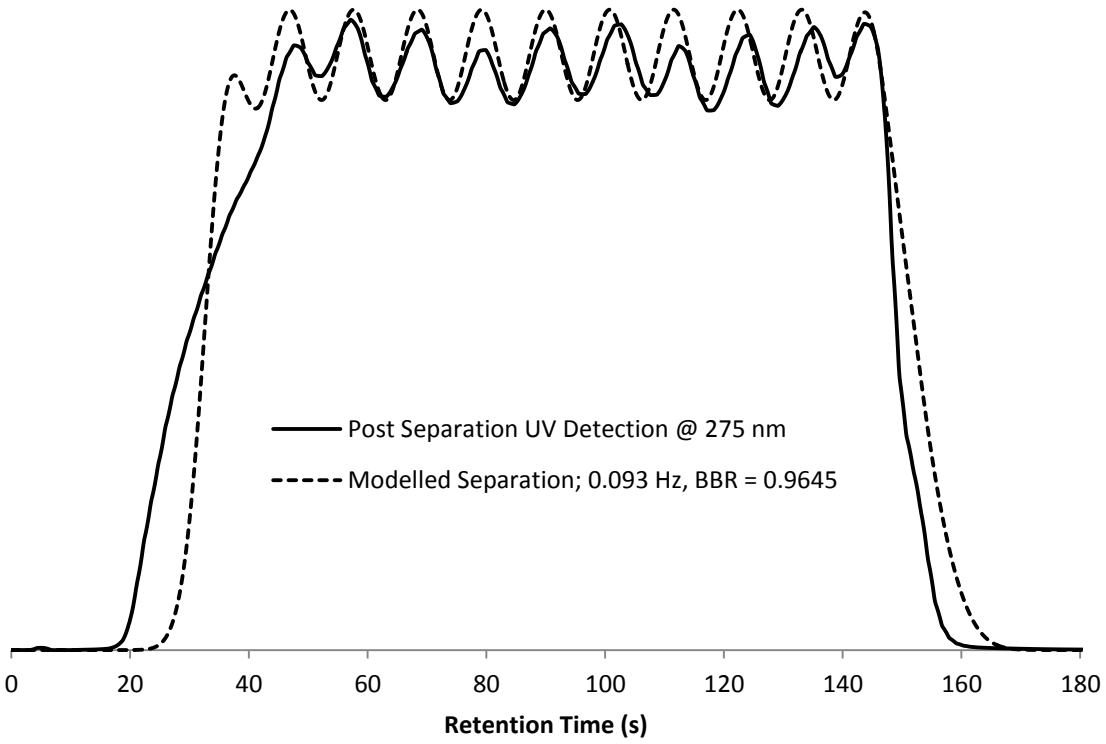


Figure 3-2: Pulsed injections with post-column detection.
Experimental and modelled results for pulsed salicylic acid with post-separation UV absorbance detection. Instrument configuration as described in Figure 2-1(c). The applied pulse rate was 0.42 Hz.

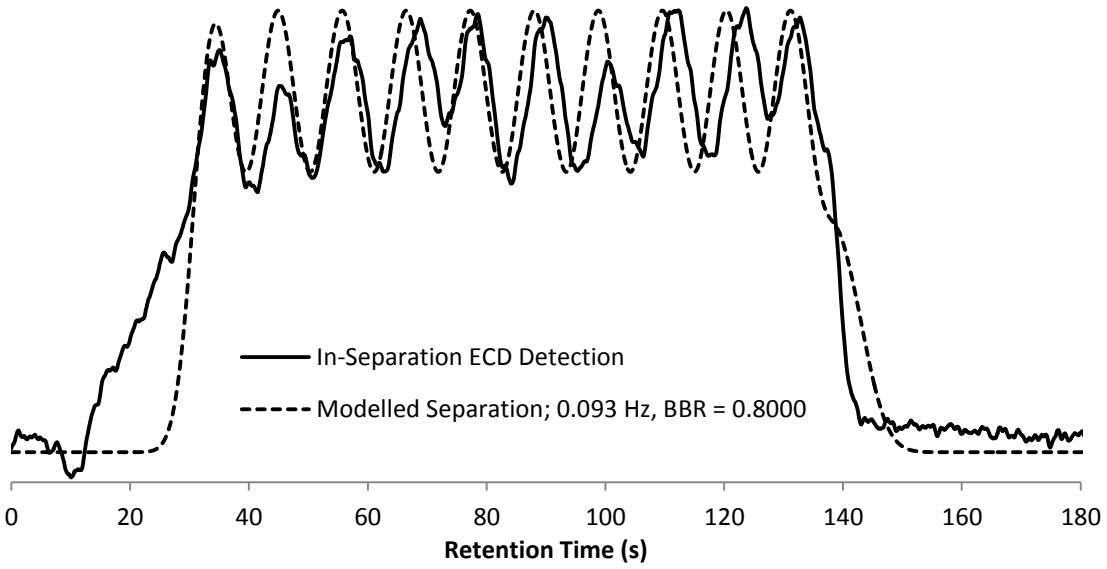


Figure 3-3: Pulsed injections with on-column detection.
Experimental and modelled results for pulsed salicylic acid with in-separation electrical conductivity detection (ECD). Instrument configuration as described in Figure 2-2. The applied pulse rate was 0.42 Hz.

Most curiously, the measured frequencies (0.093 Hz in both cases) did not correspond to the applied 0.42 Hz pulse rate. Similarly, an applied pulse rate of 0.83 Hz resulted in discernible frequencies of 0.19 Hz. Doubling the pulse frequency had clearly doubled the detected frequency as expected – and these results are summarised in Table 3-2.

Programmed Valve Switch			Nominal Pulse Frequency (Hz)	Measured Pulse Frequencies (Hz)		Nominal / Measured Frequency
Switch Interval		Period		Post-Separation	In-Separation	
min	sec	sec				
0.02	1.2	2.4	0.4167	0.093 +/- 0.002	0.092 +/- 0.002	4.5
0.01	0.6	1.2	0.8333	0.194 +/- 0.004	0.188 +/- 0.006	4.4

Table 3-2: Relationship between applied and measured pulse frequencies.

The results achieved with on-column and post-column detection were identical within the limits of error and followed the predicted and modelled patterns – except that the measured frequencies did not equal the nominally applied frequencies. Nevertheless, the *ratios* of nominal to measured frequency were found to be effectively constant (Table 3-2).

In order to arrive at some sort of understanding of these otherwise perplexing results, an attempt was made to replicate the experimental result by modelling. As shown in Figure 3-3 for the case of on-column detection (at the

longitudinal centre of the column), an applied pulse rate of 0.83 Hz produced similar modelled results in terms of diffused pulse amplitude, only if the BBR was greatly increased (from 0.1 to 0.8). Post column detection, which involved twice the residence time within the column, similarly produced the expected frequency and pulse amplitude only if the BBR were further increased to 0.96 (Figure 3-2). Both the ECD and UV detector signals also included a marked pre-cursive increase in response prior to the onset of a stable pulse. It would be reasonable to conclude then that the physical pulses were *in fact* at the measured frequencies – but that the severely overloaded peaks had physically 'combined' into a single, broadened, low frequency harmonic. That this 'harmonic' frequency bore a consistent direct proportionality to the nominally applied frequency supports this hypothesis.

These confounded results sustained the theory of COMForTS-*ps* only in that the detected frequency was indeed *proportional* (but not equal) to the applied frequency. The difficulties of even single-point on-column detection (peak distortion and loss of efficiency) were great; both sensitivity and analyte response were highly problematic. Furthermore, our modelling (discussed in Section 2.4.10) had shown that the peak capacity of COMForTS-*ps* would be far inferior to that of COMForTS-*is* and probably not worth pursuing.

With no immediately foreseeable way to manufacture columns or switching valves where the stationary phase of D_1 and D_2 extend to the 'switching surfaces' of the inter-dimensional valve (Figure 2-1), the research focus turned to virtual pulse generation by multipoint detection.

Whilst the COMForTS ECD was built with the (small) facility to piggy-back an array of additional, independent conductivity sensors (see Appendix A and Appendix B), this mode of on-column detection was deemed to have limited applicability as well as being difficult to achieve without detrimentally affecting column performance.

3.3.2. Multi-point detection

The Radon transform, when correlated to the results of LINEAR2 was able to identify all lines and unequivocally reject harmonics and overtones, even in complex chromatograms (which will be covered in detail in Chapter 4 and

Chapter 6). For these reasons and for consistency within the thesis, the results presented here are described in terms of the RADON results.

For each of the experiments, absorbance chromatograms were calculated for each of the 2048 detectors, and 50 equally spaced groups of 32 chromatograms were binned (using a Gaussian weighting profile) into single chromatograms to reduce noise. The data reported here were calculated on the basis of 50 'virtual' detectors spaced at trunc(2048 / 50) × 14 μm = 560 μm. Analyte pulse frequencies are, nonetheless, always reported with respect to the physical detector spacing of 14 μm.

3.3.2.1. Simultaneous measurements in frequency and time

Multiple injections of acetone were made at defined intervals into the system is outlined in Figure 3-1. After the initial injection, two further injections were made at 30 s intervals followed by three injections at 60 s intervals In a conventional single detector absorbance chromatogram, we would therefore expect to see five peaks of roughly uniform height (depending on resolution), spaced apart at retention time intervals corresponding to the intervals between injections. Using a single detector from the array, this expected result was indeed obtained and is shown in Figure 3-4.

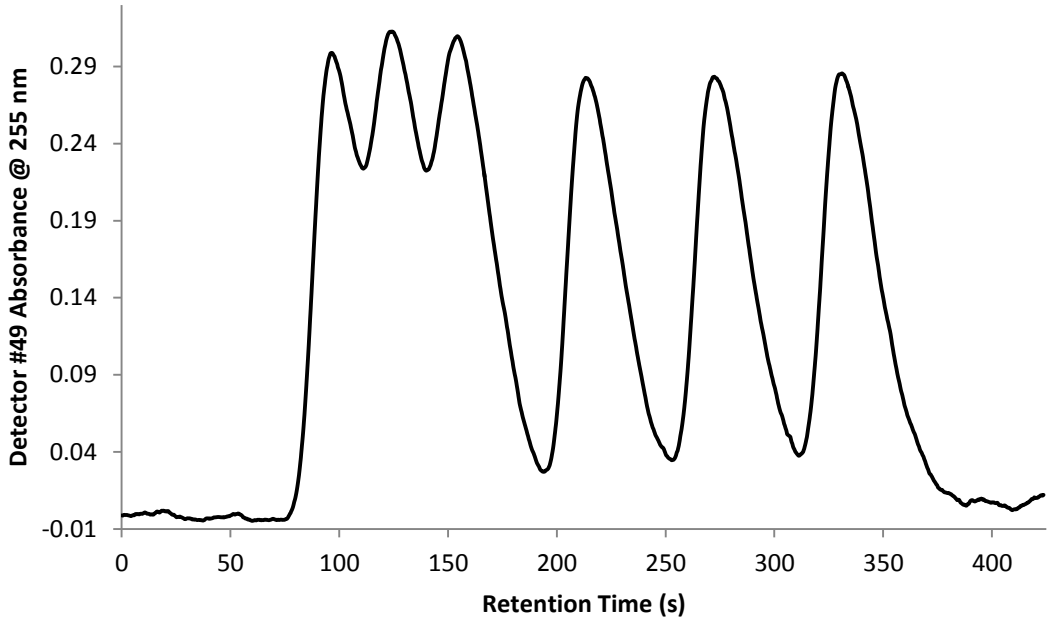


Figure 3-4: Chromatogram of five injections of an unretained analyte.
This is a conventional single-detector absorbance chromatogram that is the result of six injections of acetone. After initial injection, two further injections were made at 30 s intervals followed by three injections at 60 s intervals.

However, in contrast to conventional HPLC systems, the COMForTS system was equipped with an array of 2048 equally spaced detectors. Absorbance chromatograms were calculated for each detector and binned as defined above to create 50 chromatograms from 50 equally spaced 'virtual' detectors. Instead of using the SCOFT approach of summing these individual chromatograms to produce a single, one-dimensional (time domain) pulsed signal as in Figure 1-3, we can simply place these individual signals side-by-side in a two-dimensional array of detector signals against time. In other words, we are leaving the original signals much as depicted in the perspective view shown in Figure 1-3a.

The fifty individual time-domain chromatograms obtained in this experiment are shown in a density plot in Figure 3-5, using the colour scale to indicate absorbance. As an analyte peak moves over the first to last detector in the array (virtual detector # 0 to # 50) over time, we expect, if the analyte velocity is constant, to see a straight peak line that has a slope (detectors per second) that is directly proportional to the analyte velocity. The x-intercept of this line is the time at which the analyte peak passed the first detector in the array. This time is used as the *total* retention time of the analyte within the system.

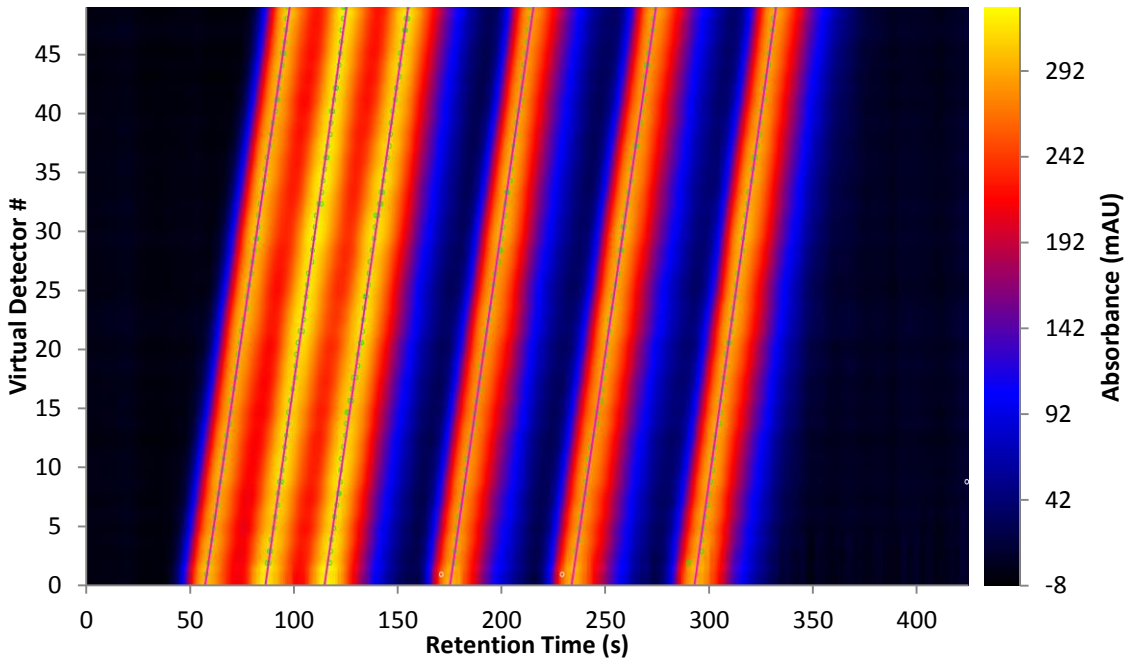


Figure 3-5: Density plot of arrayed multi-detector chromatograms.
Analyte peaks, moving across the detector array in time, appear as lines; the slopes and intercepts of which are derived from the Radon transform.

The visual interpretation of Figure 3-5 is straightforward:

- There are five parallel peak lines – therefore, there are five analytes travelling over the (second dimension, D_2) detector array at the same velocity
- Each peak line, however, has a different x-intercept, indicating different total retention times. The peak lines are obviously resolved.
- If these analytes are moving at the same velocity in D_2, they must have the same retention time in D_2 and the difference in retention time is entirely due to differences in the retention time in D_1.

…in this instance we *know* that acetone is not retained in D_1 and that the retention difference is therefore due entirely to the difference in injection time. In effect, we have replicated the result of a two-dimensional separation of five different analytes that are resolved in D_1 with no additional resolution provided by D_2.

The numerical evaluation of the slopes and x-intercepts of these lines can be achieved via a Radon transform (1.2.2), where peaks in the transform correspond to the polar coordinates of the lines within the two-dimensional array. The COMForTS 4.0 RADON module (see Section 6.4.2) uses a modified partial discrete Radon transform (DRT) where the coordinates of peaks within the transform correspond to the line's x-intercept (in time) and slope (in degrees). Application of this transform to the arrayed chromatogram image of Figure 3-5 yields the result shown in Figure 3-6. Because the DRT is a slow algorithm, the range of angles for which the transform was calculated was limited to the physically possible range of angles that *may* represent peak lines. For example, there is some minimum time required for a peak to traverse all detectors, which, in this case, corresponds to a line angle of 87 degrees.

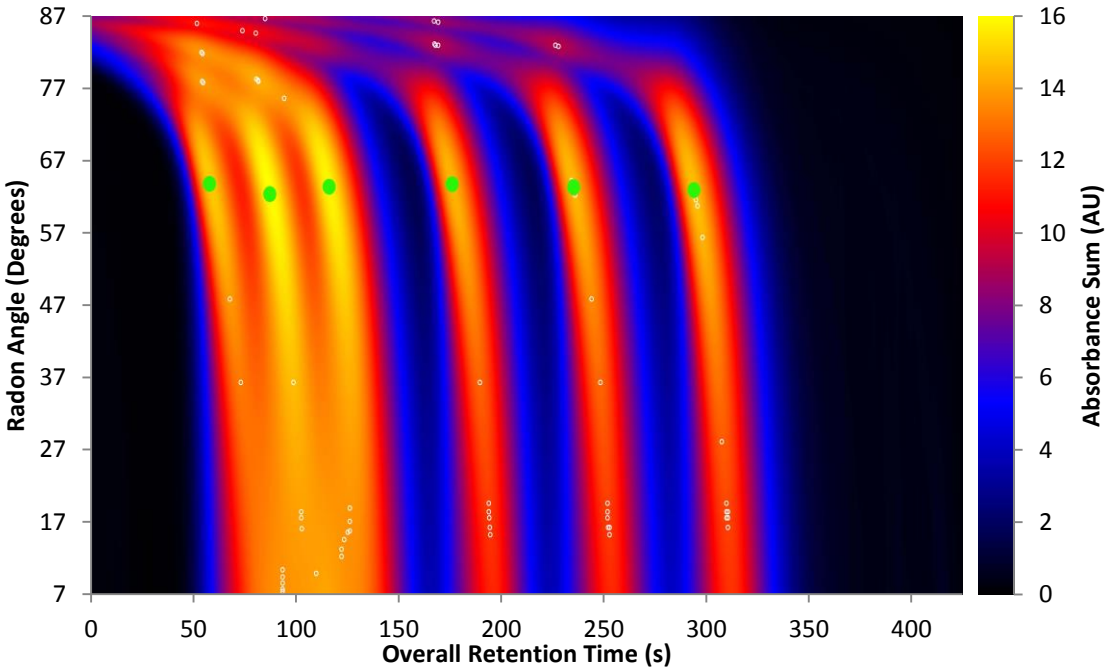


Figure 3-6: Frequency/time transform of six acetone peaks.
Discrete Radon transform (DRT) of arrayed detector chromatograms showing six fundamental peaks (marked in green) at the same Radon angle (frequency) but at different retention times. Detected false and harmonic peaks are marked in white.

A simple two-dimensional peak detection algorithm (assuming conical peaks and matching one-dimensional maxima in each axis) located the maxima in the DRT that are shown as small white circles. Clearly, some 'false peaks' have been detected along with harmonic peaks in the region between 77 and 87 degrees. These maxima, indicating the polar coordinates of lines, are easily translated into Cartesian line equations that can then be plotted over the chromatogram array and correlated with lines connecting one-dimensional (time domain) peak maxima in the arrayed detector chromatograms. The correlated lines are superimposed on the chromatogram intensity plot (Figure 3-5) and the corresponding peaks in the DRT (Figure 3-6) are marked in green. In this way, the location and slopes of lines are determined by the DRT and harmonics and 'false peaks' are easily identified as lines that do not, or cannot, exist between peak maxima within the chromatogram array.

3.3.2.2. Rejection of harmonics and overtones

This capability to correlate transform results to the raw chromatograms allows unequivocal rejection of harmonics and rejection of false peaks returned by less than perfect peak detection. This cannot be achieved when the absorbance from multiple detection points is acquired as a single, physically summed,

time domain signal as in SCOFT [48,83,85]. In fact, even when physically separate detectors were used [88], the digital signal sum was processed by Fourier or wavelet transform and exhibited significant interference from harmonics and overtones. As discussed earlier (in Sections 2.3.2 and 2.4.4), appropriate processing techniques may reduce, but probably not entirely eliminate, the effects of these interferences. If however, we retain the un-combined signals (individual detector chromatograms) in a two-dimensional array, it would be similarly possible to reject most interferences (resulting from short time FFTs or wavelet transforms) that do not correlate with the raw chromatograms.

3.3.2.3. Quantitation

If one takes a frequency transform, the magnitude measured at each frequency is proportional to the amplitude of the signal pulses. In this way, we can relate transform magnitude to peak (pulse) height as discussed and shown in Chapter 2. Unhappily, this relationship can break down when there are changes in the *relative* amplitudes of overlapping analyte pulses. This is just as true of the RT as it is of the FT and wavelet transform.

Using the transform results to identify peak lines within the original chromatogram array has the additional benefit of quantitation with far less potential for error. Following one line along the chromatogram array, the line follows the 'crest' that has an intensity equivalent to the peak height. Without significant band broadening during its passage over the detector, the peak height should remain more-or-less constant across the array. If two lines should cross each other (peak overlap), the intensity of both lines increases around the point of cross-over. We can see then that the *minimum* intensity along each line would remain a true representation of the peak height – and that this relationship would hold until a peak line is overlapped at *every* point along its length. At that point, the obscured line would remain detectable in the transform result but quantification on minimum height would be erroneous. In such cases, the obscured height may nonetheless be estimated by solving the set of simultaneous equations for the sum of the set of minimum peak heights of all lines that overlap at a given point along the overlapped line.

3.3.2.4. COMForTS peak areas

So far, we have within our chromatograms, detected peak lines in three dimensions: slope, intercept and height. These *lines*, by definition, do not have an area and COMForTS peaks do not have an area (or volume) in a conventional physical sense.

Consider an ordinary chromatographic peak as, principally, a random distribution of analyte molecules in time (as the result of random longitudinal diffusion within the column). The time of the peak maximum therefore represents the most *likely* time at which any given molecule would elute and the peak width may be seen as a roughly Gaussian distribution describing the range of possible retention times. Similarly, there is some uncertainty in measuring the slopes and intercepts of peak lines. COMForTS peaks should therefore have a 'width' that indicates the uncertainty in the measurement and a shape (in three dimensions) that indicates the probability that the identified peak line is located at the calculated position.

Returning to the arrayed detector chromatograms with conventional peak maxima identified in the time domain. Each of these peak maxima is therefore a point that should lie on the Radon-derived line. From a least squares linear regression on these points, we can extract the standard errors in the estimates of slope and intercept and therefore the expected standard deviations of the estimates of overall retention and analyte pulse frequency.

In a two-dimensional COMForTS chromatogram (as treated in this thesis), peaks have these characteristics:

- A maximum at the measured coordinates within the time and frequency domains.
- An overall height that corresponds to the analyte peak height
- A height that decreases in Gaussian proportion to the increasing distance from the coordinates of the peak maximum. This indicates the decreasing probability that the analyte would have the retention coordinates of that point.
- The *overall* width in each axis (frequency and time) is defined as eight times the standard deviation of the estimates (8σ, allowing for under-estimated errors) so that

- There is a >99% probability that the peak maximum lies somewhere within the area of the peak's base.

Detector signal noise, by way of its effect on conventional peak detection, is the largest contributor to the error in the determination of peak frequency and retention time.

COMForTS peak areas (over the baseline) therefore represent a location probability-field and not a physical peak width.

3.3.2.5. Conversion of frequency/time to time/time coordinates

At present, our retention coordinates are frequency and overall retention time. In the case of the Radon transform applied to an array of equally spaced detector signals, the frequency of the pulses over the detectors is directly proportional to velocity. For consistency, the COMForTS software reports the frequency with respect to physical detectors and not the virtual detectors as the number and spacing of these virtual detectors may vary.

A pulse frequency of, say 100 Hz corresponds, then, to an analyte peak crossing 100 detectors per second. As these physical detectors are 14 μm apart, the distance travelled by the analyte in one second is 100 × 14 μm = 1400 μm and its linear velocity *in the second dimension* is 1400 μm/s. From this, we can calculate the retention time T_2 of the analyte in D_2. Because we also know the total retention time T_1, the retention time in the first dimension, T_1 must be equal to $T_T - T_2$. The errors in the estimate of frequency and overall retention are similarly converted into D_1 and D_2 retention times. Because of its dependence on both T_2 (determined by frequency) and T_1, the error in the estimate in T_1 includes both the errors in frequency and in T_1.

When we apply the above conversion to the experimental data, we obtain the conventional representation of a two-dimensional time/time separation that is shown in Figure 3-7. This plot accords with our replication of a 2D separation in which all the separation takes place in D_1: we have five peaks with very close to the same retention time in D_2 but different retention times in D_1.

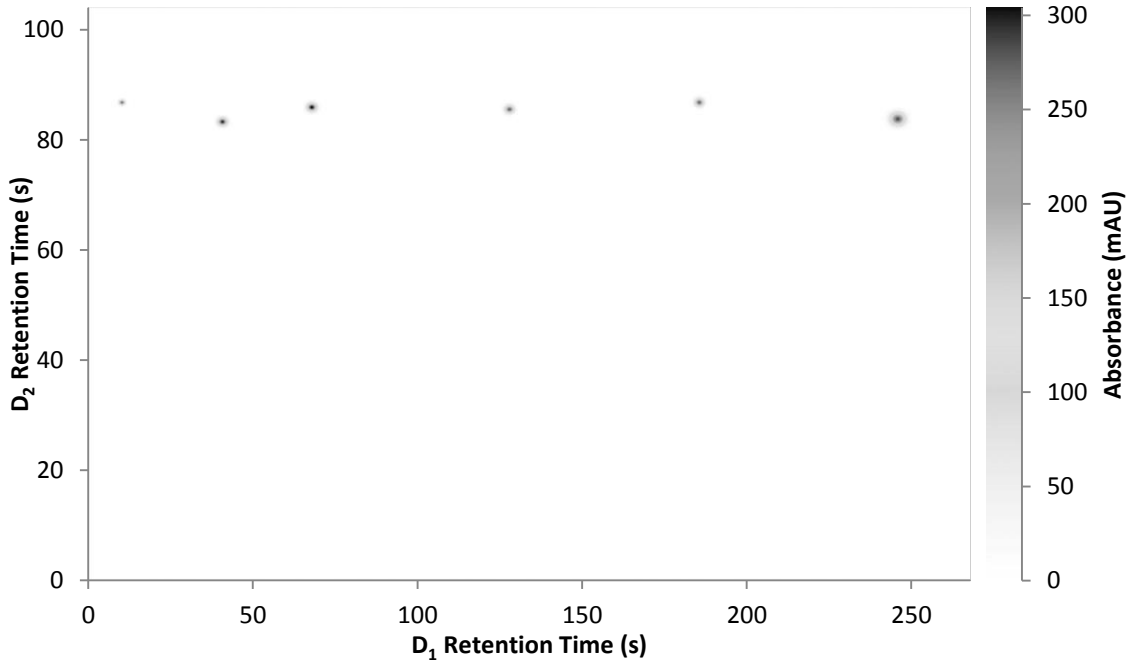


Figure 3-7: COMForTS 2D pseudo time/time chromatogram.
The analytical column (D_2) was assumed to be 60 mm in length with a zero-length first dimension (D_1) column. The differences in D_1 retention time are entirely the result of the interval between repeated injections.

The D_1 time intervals between these peaks correlate very closely to the intervals between the acetone injections. The overall retention time of acetone was known to be about 96 s (the first peak in Figure 3-4) which is in good agreement with the COMForTS result of 97 s and peak heights were also in good agreement (Table 3-3).

3.3.2.6. Comparative resolution and quantitation

When we compare the conventionally measured time domain signal (Figure 3-4) with the COMForTS result (Figure 3-7), COMForTS demonstrates a far superior ability to resolve peak information in the time domain (Figure 3-8). In effect, COMForTS has made, in this instance, 50 measurements of the overall retention time and the increased precision is the statistically expected result.

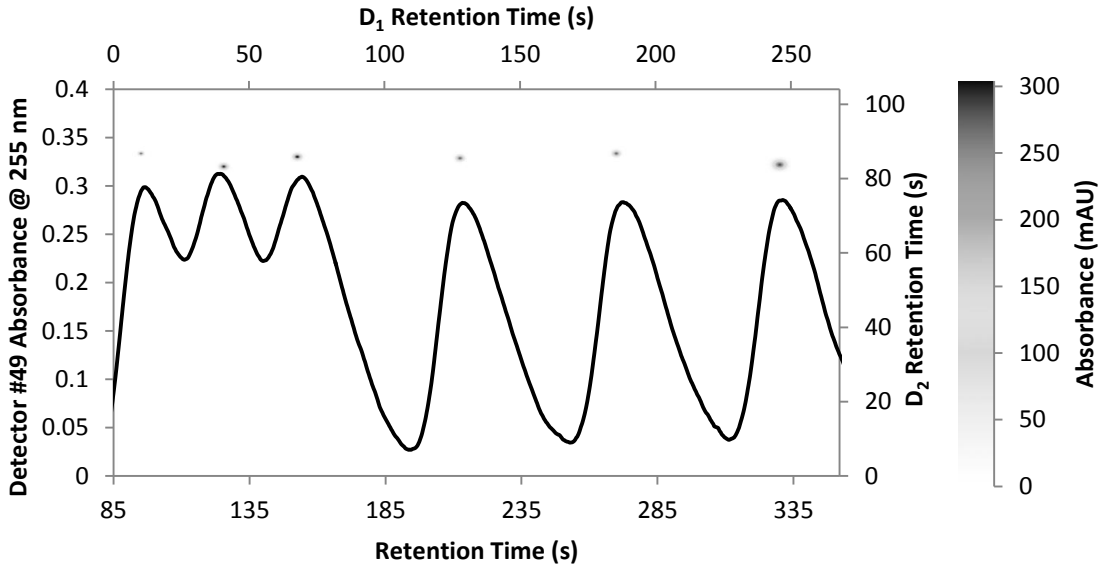


Figure 3-8: Comparative time-domain resolution.
A conventional chromatogram is shown (on the primary axes) against the COMForTS evaluation of the same separation (shown as time/time density plot on secondary axes).

Overall Retention Time (s)		
Conventional	COMForTS	% Rel Error
96	97	1.042
107	108	0.935
115	116	0.870
124	125	0.806
136	136	0.000
143	144	0.699
154	154	0.000
159	158	-0.629
173	173	0.000
184	183	-0.543
213	214	0.469
226	226	0.000
234	234	0.000
241	240	-0.415
272	272	0.000
295	295	0.000
298	298	0.000
309	309	0.000
330	330	0.000
370	369	-0.270
	Mean % Relative Error	0.148
	Std Dev in % Rel. Error	0.491

Peak Height (mAU)		
Conventional	COMForTS	% Rel Error
299	286	-4.348
217	215	-0.922
206	206	0.000
313	304	-2.875
220	214	-2.727
227	220	-3.084
309	304	-1.618
248	243	-2.016
221	220	-0.452
255	248	-2.745
283	276	-2.473
198	198	0.000
206	205	-0.485
244	236	-3.279
283	278	-1.767
209	204	-2.392
254	241	-5.118
196	196	0.000
286	284	-0.699
196	196	0
	Mean % Relative Error	-1.850
	Std Dev in % Rel. Error	1.514

Overall Retention Time Correlation		
	Conventional	*COMForTS*
Conventional	1	
COMForTS	0.999971541	1

Peak Height Correlation		
	Conventional	*COMForTS*
Conventional	1	
COMForTS	0.996394261	1

Table 3-3: Relative peak height and retention metrics.
The accuracy and precision of COMForTS retention times and peak heights are assessed relative to the equivalent conventionally measured metrics.

From the conventional chromatograms (of four replicates of this experiment), it is possible to measure the retention times and peak heights. By way of direct comparison, the same COMForTS metrics for the same separations show an almost perfect correlation to the conventionally measured results (Table 3-3).

If we assume that the COMForTS metrics are in error, we find that error to be ± 1.5% in overall retention time and ± 4.5% in peak height. Average COMForTS results differed from conventional results by less than 2%. However, the conventional metrics themselves showed similar variability, from which we can conclude that the COMForTS results were not significantly different to the conventional results provided by this instrument.

D_1 Retention Time Differences and Injection Intervals

Conventional		COMForTS		
Peak Times (s)	Inj. Interval (s)	Peak Times (s)	Inj. Interval (s)	% Rel Error
96		11		
124	28	41	31	9.429
154	30	68	27	-10.367
214	60	128	60	0.367
272	58	186	57	-0.914
331	59	246	60	2.288
136		58		
159	23	81	23	-2.000
184	25	107	26	4.200
241	57	162	56	-2.193
299	58	219	56	-2.948
115		38		
144	29	67	28	-0.211
173	30	98	31	4.305
234	61	159	61	-0.033
295	61	214	55	-9.738
40		15		
70	30	45	30	1.467
98	28	73	28	0.393
159	61	134	61	-0.590
219	60	194	60	-0.317
279	60	253	59	-1.033
			Mean % Relative Error	-0.458
			Std Dev in % Rel. Error	3.449

Correlation Between D_1 Retention Times and Injection Interval

	Conventional	*COMForTS*
Conventional	1	
COMForTS	0.993751449	1

Table 3-4: Accuracy and precision of COMForTS D_1 retention times.
D_1 retention times are the (intercept-derived) overall retention time less the (frequency-derived) D_2 retention time. With no physical D_1, the calculated D_1 retention time differences correspond to the time interval between successive injections.

The accuracy and precision of COMForTS-derived D_1 retention times depends, as noted above, on the measured frequency as well as overall retention. A comparison between the COMForTS and conventional injection intervals (D_1 retention differences) affords some measure of the accuracy and precision of frequency measurements (but *includes* contributions of the error in total retention). These comparative data are presented in Table 3-4.

Again, the correlation between COMForTS and conventional measurements is nearly perfect, though the 'error range' now extends to ± 17%. At the same time, the COMForTS result remained accurate, differing from the conventional result by less than 0.5% on average.

3.3.2.7. Generation of multicomponent time/time chromatograms from frequency/time data

We have now established that the COMForTS approach can accurately resolve separation information in the frequency domain and in the time domain. COMForTS time/time separations were also shown to correlate very strongly with conventional two-dimensional time/time separations, but in circumstances where only one frequency was present within the chromatogram. In other words, it has been shown that COMForTS can deliver excellent and true resolution in the first dimension but we have not demonstrated any resolution in the second dimension.

As in the previous experiment, the first dimension column (D_1) was a 210 × 0.250 mm DB-1 and the mobile phase was 70% acetonitrile in water. The test solution consisted of four PAHs (naphthalene, anthracene, pyrene and chrysene) dissolved in acetone. In this case, we know that D_1 gives no resolution, because the same analytes were not resolved with this mobile phase on a longer (450 mm) column with the same stationary phase and weaker mobile phase (50:50 ACN:Water) (Figure 3-9).

When this solution was injected into the system in which a 145 × 0.318 mm capillary packed with 5 μm C18 particles is connected serially as D_2, a separation was obtained. The separation so achieved and recorded conventionally with a single detector is shown in Figure 3-10.

Clearly, this separation is entirely achieved in D_2 and the components must therefore have differing linear velocities within the D_2 stationary phase.

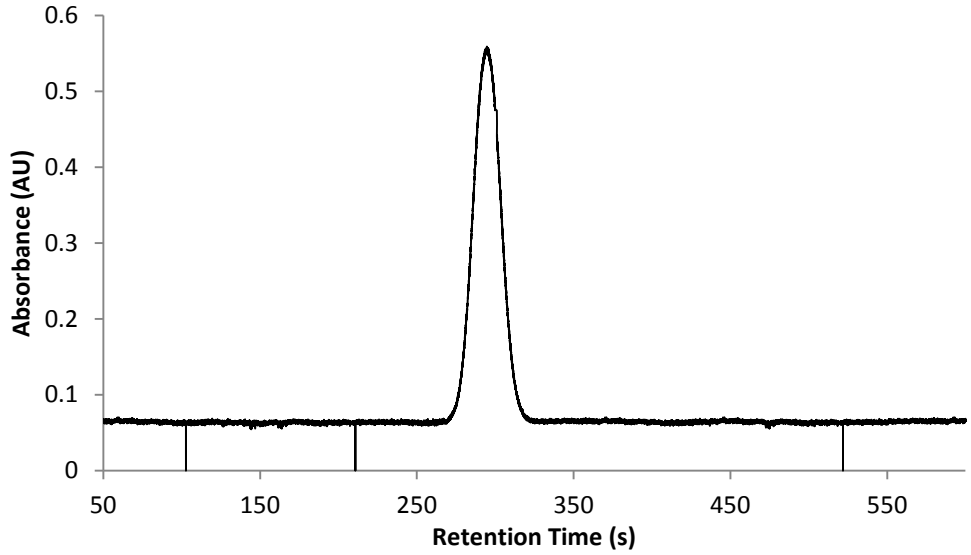


Figure 3-9: Conventional one-dimensional separation PAHs.
Column: 450 × 0.250 mm DB-1. The retention time of the single observed peak was just beyond the void time (265 s) and comprises unresolved signal contributions from of acetone and four PAHs. The mobile phase was 50:50 ACN:Water at a flow rate of 5 μL/min.

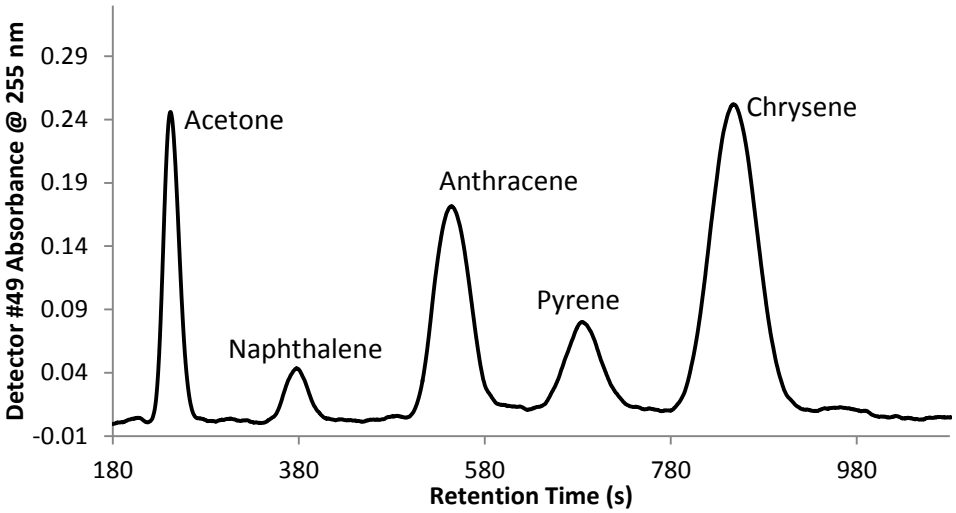


Figure 3-10: Conventional two-dimensional separation of PAHs.
The separation acetone and four PAHs was conducted in the online mode and recorded with a single on-column detector. The mobile phase was 70:30 ACN:Water at a flow rate of 5 μL/min.

With analytes of differing linear velocities in D_2, we should then see in the arrayed detector chromatograms a peak line for each component and each line should have a different slope. This was indeed the result obtained (Figure 3-11) and the Radon transform (Figure 3-12) displayed peaks corresponding to the coordinates of lines of different overall retention and at different angles.

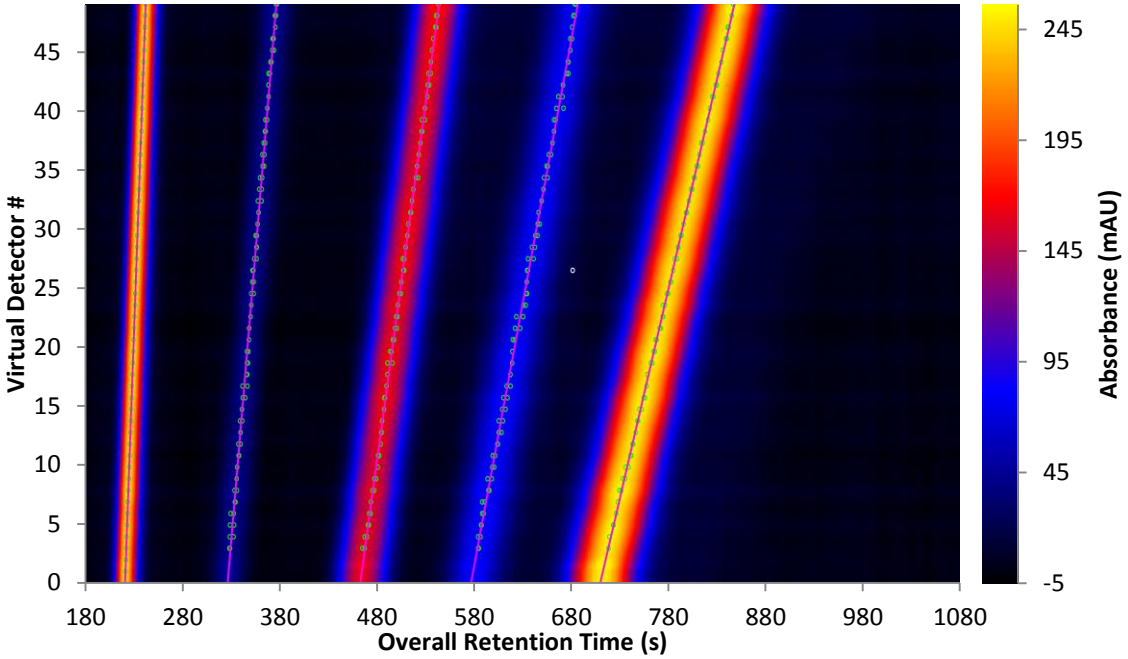


Figure 3-11: Density plot of arrayed multi-detector chromatograms.
This is an online two-dimensional separation of acetone and four PAHs. Slopes and intercepts of peak lines (shown in magenta) were derived from the Radon transform (Figure 3-12).

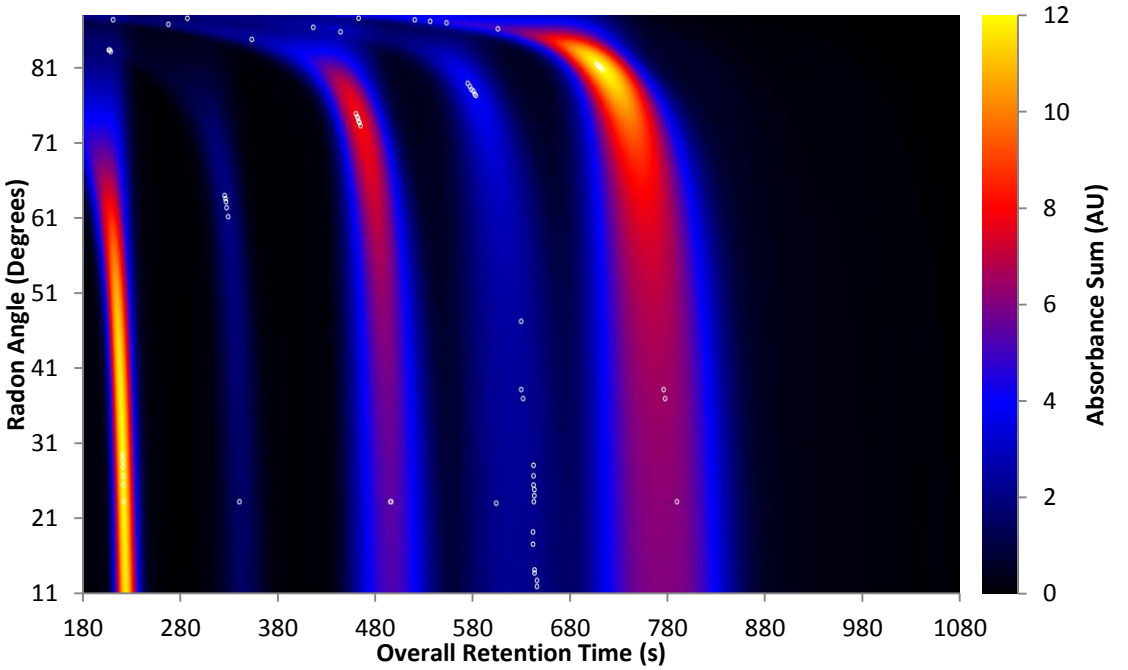


Figure 3-12: Radon transform with resolution in both time and frequency.
This is the Radon transform of Figure 3-11 showing five fundamental peaks at varying Radon angles and varying retention times.

The time and frequency domain data obtained for each analyte over seven replications of the experiment are summarised in Figure 3-13, with error bars for each metric indicating 95% confidence limits.

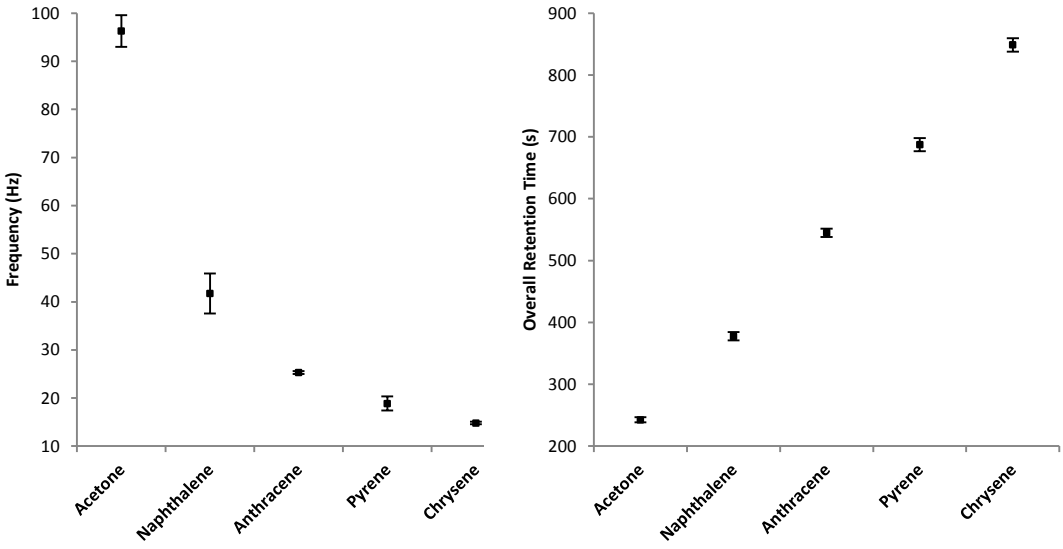


Figure 3-13: COMForTS time- and frequency-domain reproducibility.
Data for acetone and four PAHs are presented with error bars indicating 95% confidence limits (n = 7).

The measurement precision of the frequency of the naphthalene signal was notably worse than for any other peak. The ratio of signal/noise for naphthalene is approximately half that of the pyrene signal. This level of noise clearly created difficulties in the detection of the conventional peak maxima in the individual time-domain chromatograms. Because these peak maxima are used as a determinant of the 'goodness of fit' of the line equations it is unsurprising then that this is a ramification of signal noise within a COMForTS analysis.

As with the previous experiment, overall retention times were a good fit to the conventionally measured total retention times. Because there is no retention in D_1, the analyte pulse frequencies in D_2 should at least bear a linear relationship with overall retention. The left plot of Figure 3-15 indicates that this is not the case. When we plot these frequencies against the conventionally measured D_2 retention times (Figure 3-15, right), an almost perfectly linear relationship is revealed. It can only be concluded that their retention times in D_1 are, in fact different. Even though the error in frequency is generally larger than the error in overall retention and though D_1 retention times include *both* errors,

COMForTS was able to resolve each of these analytes in both D_1 and in D_2. These features are evident in the COMForTS chromatograms shown in the density plot of Figure 3-15 and in the surface plot of Figure 3-16.

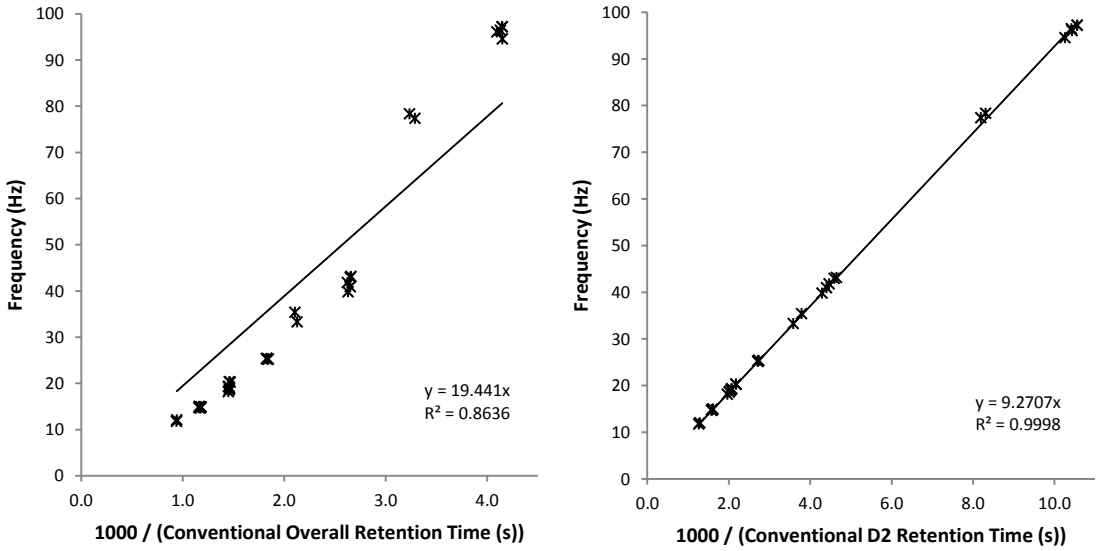


Figure 3-14: Correlation of analyte frequencies to velocity.
Examination of the correlation of peak frequency to average relative peak velocity and second dimension relative peak velocity. The discrepancy between the results indicates that the first dimension exhibited significant selectivity.

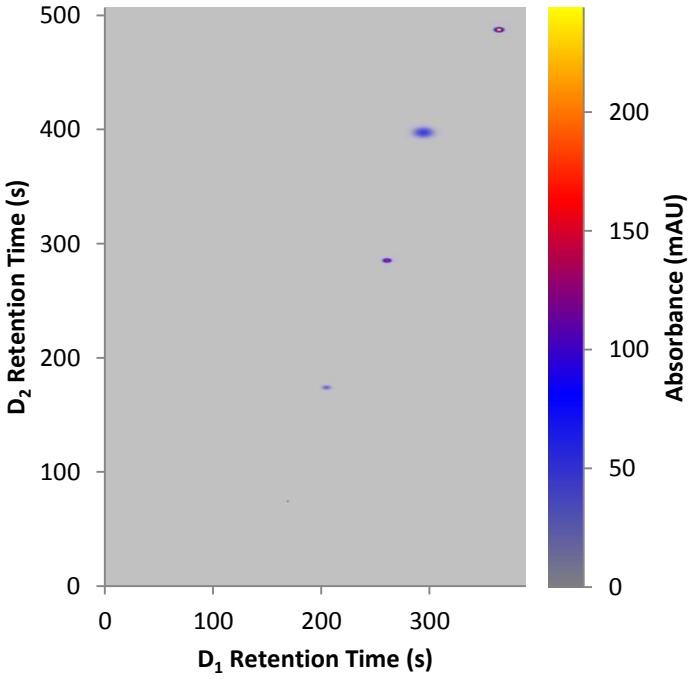


Figure 3-15: COMForTS 2D time/time chromatogram, density plot.
2D separation of acetone and four PAHs. D_1 = 210 × 0.15 mm DB-1; D_2 = 145 × 0.318 mm 5 µm C18. The acetone peak is more visible in Figure 3-16.

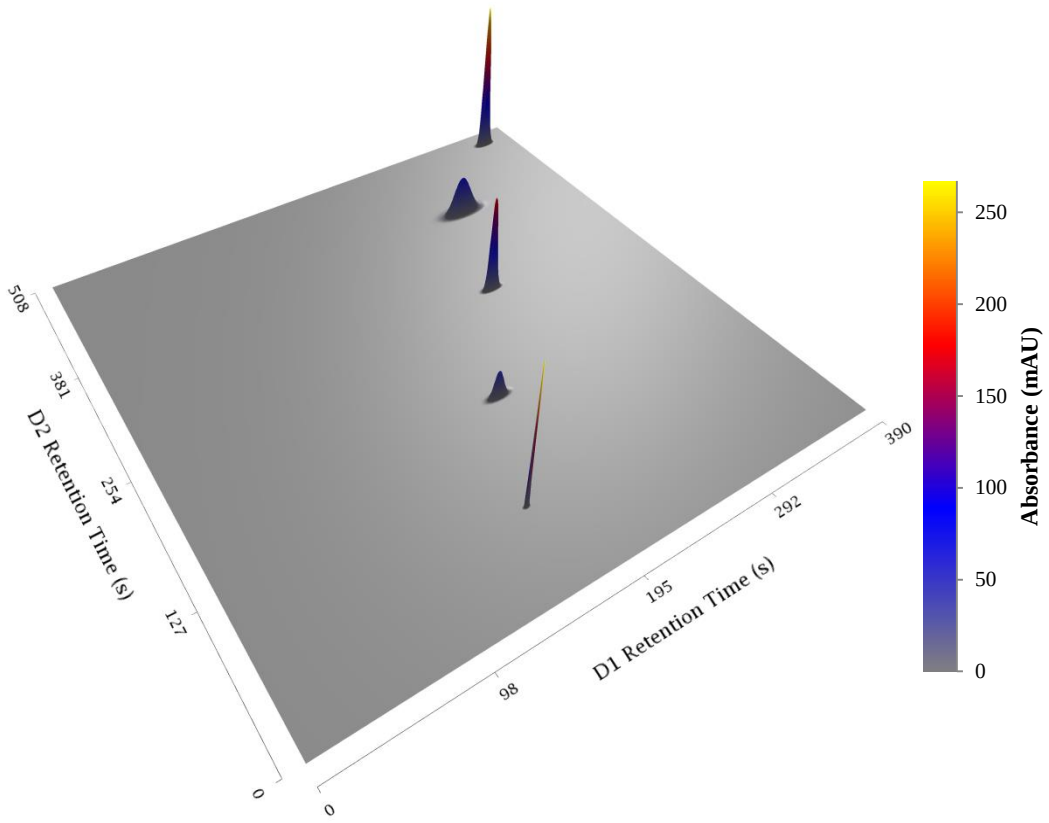


Figure 3-16: COMForTS 2D time/time chromatogram, surface plot.
2D separation of acetone and four PAHs. D_1 = 210 × 0.15 mm DB-1; D_2 = 145 × 0.318 mm 5 µm C18.

The acetone peak (at T_1 = 170 s, T_2 = 75 s) is so well defined as to be a very sharp 'pin-prick' in the two-dimensional density plot of Figure 3-15, and therefore hard to see. A far better representation is the 3D surface of Figure 3-16, where all five peaks are very clearly visible. Mathematically the two plots are interchangeable.

Signal noise in a COMForTS analysis, as mentioned earlier, is manifested in the uncertainty of peak locations. The COMForTS chromatograms (Figure 3-15 and Figure 3-16), with their perfectly flat baselines, only *appear* to be noise-free when compared to conventional 2D chromatograms. Low signal / noise ratios result in increased uncertainty in the measurement of analyte frequency and overall retention such as in the case of the two smallest peaks (naphthalene and pyrene) where the probability fields of the peak location cover significantly larger ranges of retention coordinates.

When the time axes (for retention in D_1 and D_2) are viewed at the same scale (Figure 3-15), we note that peak locations are more uncertain (wider) in D_1

than in D_2. This is because the determination of D_1 is dependent upon the errors in both frequency and overall retention.

An example of this COMForTS separation is provided in the Digital Thesis Materials. The chromatogram file is named "Stock X 70% ACN 150 nL 5 uLpm.COMForTS" and can be found on the attached digital media in the folder:

"[COMForTS]:\Sample Data\COMForTS4\Chromatograms\"

…or, if installed on the reader's computer, in the folder:

"C:\Users\[YourUsername]\Documents\COMForTS\Samples\COMForTS Chromatograms\"

The full processing history with tabulated results and plotted intermediate results (at each stage of analysis) can be inspected using the "COMForTS Chromatogram Viewer" module of the COMForTS 4.0 Chromatography Suite.

3.4. Conclusion

Physical pulse modulation in liquid chromatography was demonstrated and measured pulse frequencies were proportional to applied pulse frequencies when the switching valve was not in contact with the stationary phases of both separation dimensions. This was a partial confirmation of the theory of COMForTS-*ps* as outlined in Chapter 2. Nonetheless, the multifarious physical demands and limitations of this approach were highlighted. It was not possible to test the most promising mode, COMForTS-*is*, due to the even greater technical difficulties of the valve configuration, not to mention accurate control of switching rates.

Virtual pulse modulation through multipoint on-column detection was demonstrated using a CCD chip as a detector array. This provided 2048 physical detectors, spaced 14 μm apart. Because of the instrument's low signal/noise ratio, absorbance chromatograms were binned, creating 50 virtual detectors spaced at 560 μm. With physical peak widths of between 15 and 30

mm (more than 40 times the detector spacing), the SCOFT approach of signal summation across the detectors to produce a single pulsed time-domain signal (for Fourier or wavelet analysis) was impossible.

A different scheme for extracting frequency and time domain data was developed. Detector chromatograms were placed in a two dimensional array in which analyte peaks appear as lines running across the detector array in time. The Radon transform was then used to find the equations of these lines, yielding analyte pulse frequencies and overall retention.

It was conclusively demonstrated that the COMForTS approach was able to resolve multiple components simultaneously in both the time and frequency domains with significantly higher precision than was conventionally obtainable. The accuracies of the measurements of retention times and peak heights were at least equal to those of conventional measurements of the same separations.

In all, the principles of the method were demonstrated to operate in accordance with the developed theory and qualitative and quantitative measurements were no less accurate than conventionally obtained results. Furthermore, two-dimensional time/time COMForTS chromatograms reproduced contrived separation patterns with a precision that yielded separations of analyte information that were not conventionally obtainable in methods that rely entirely upon physical separation.

Chapter 4.
Proof of Concept II:
Enhanced Peak Production

4.1. Introduction

Simultaneous operation of the COMForTS instrument in a conventional single detector 2DLC mode and in 2D COMForTS mode was demonstrated in the previous chapter. Those experiments established the validity of the fundamental theory of the COMForTS method when applied to arrayed in-separation detection in that both frequency and time domain detection applied to the second dimension can be used to generate virtual time/time separations over two separation dimensions. However, the overriding purpose of COMForTS - the facility to increase peak capacity and hence peak production, by reducing or eliminating peak wrap-around – has not been shown.

To demonstrate the facility to overcome wrap-around, one could take, for example, a complex sample such as coffee or tobacco that has already been characterised by comprehensive offline 2DLC [112-115]. Repetition of the analysis in the COMForTS mode could perhaps allow us to say, "COMForTS found more peaks" or perhaps that "COMForTS found as many peaks but in much less time". One may perhaps even be able to conclude that the "COMForTS separation was comprehensive" – but in either case, it would not be possible to determine if peak wrap-around would have existed in online 2DLC experiments or indeed if there were any real correlation between analytes detected (as opposed to peaks) in either method. This 'replication of a known separation' approach has another significant practical problem: The prototypical COMForTS instrument, the columns and the software are clearly not 'state-of-the-art' equipment and could not be compared directly to achievements in high-speed online 2DLC (such as discussed in [36,116,117]).

The available equipment functions as a COMForTS instrument and can perform COMForTS separations and conventional online separations (where only one first-dimension 'sample' is passed through the second dimension).

The fact that the COMForTS instrument can perform simultaneously in COMForTS and conventional modes provides a direct comparison of relative rather than absolute performance. However, we can only measure relative performance in the face of peak wrap-around if we know the exact extent of the wrap-around and the total number of analytes and the retention characteristics of those analytes.

An analytical sample of known composition and known physical retention behaviour, combined with a controlled repeated-injection strategy can be used to generate complex, overlapped 'separations' in which we know where each individual 'analyte' is located within the chromatographic system at any given time.

This experiment therefore established a contrived complex separation such that the retention times of all 'analytes' were known and in which the degree of overlap of the analytes was controlled. Simultaneous conventional and COMForTS treatments again allowed a direct comparison of the relative performance of each approach under identical circumstances.

4.2. Materials and Methods

All analyses were performed using the COMForTS Instrument as previously described in Section 3.2.2.1 and detailed in Chapter 5 and Appendix C.

Instrument control, data acquisition and processing were performed by the COMForTS 4.0 Chromatography Suite (using the 'Radon' processing path), as described in Section 3.2.2.1 and detailed in Chapter 6. Representative raw data and finished chromatograms of interest are also provided in the "Sample Data" folder in the Digital Thesis Materials.

The experimental raw detector signals (with the file name extension ".COMForTS_Raw") for each of these analyses can be found in the Digital Thesis Materials in the folder:

"…\Sample Data\COMForTS4\RawData\2D Separation (DB-1 & C18) of PAHs\"

…or installed on the reader's computer in the folder:

"C:\Users\[YourUserName]\Documents\COMForTS\Samples\Raw Instrument Data\2D Separation (DB-1 & C18) of PAHs\"

For assistance with the data processing sequence, the reader should refer to the "COMForTS 4.0 Signals Processing – Walk-Through Guide" that can be accessed from the documentation index in the Digital Thesis Materials ("[COMForTS]:\Docs\Index.htm").

The corresponding completed COMForTS chromatograms of moderately and severely overlapped injection sequences are provided in the files named:

Moderate overlap: "Stock X 70% ACN 5xMultiInject 150 nL 5 uLpm.COMForTS"

Severe overlap: "Stock X 70% ACN 5xMultiInject 150 nL 5 uLpm_00.COMForTS"

COMForTS detector conditions and the data collection and processing hardware were as described in Section 3.2.2.1.

4.2.1. Chromatographic conditions

The instrument configuration was as shown in Figure 3-1, and the chromatographic conditions employed were identical to those described in Section 3.2.2.4, that, for convenience, are briefly summarised below.

Two-dimensional liquid chromatographic separations were performed with serially connected capillary columns (D_1 and D_2) with on-column detection performed on D_2. The first dimension column (D_1) was a 210 × 0.250 mm open tubular capillary with 0.25 μm film thickness DB-1 stationary phase (cut from a GC column supplied by J&W Scientific, Folsom, California). The second dimension column (D_2) was a 145 × 0.318 mm fused silica capillary packed with 5μm Nucleosil C18 particles (see Section 5.3.4)

Both capillary columns were prepared in-house and fitted with standard HPLC 1/18" stainless steel connectors and ferrules with the aid of sleeves over the capillary ends, cut from 508 μm ID (orange) PEEK tubing (see Figure 5-3 and Figure 5-4).

All solvents were HPLC grade and all reagents were analytical grade. 80 MΩ Millipore water was used in all mobile phases. The following chromatographic conditions were employed unless otherwise stated:

- Mobile phase: 70% Acetonitrile in water
- Mobile phase flow rate: 5 μL/min
- Injection volume: 150 nL (injected over 1867 ms)

A solution of four polynuclear aromatic hydrocarbons (PAHs) in acetone was prepared as described in Table 3-1. This solution was known as "Stock X". The two-dimensional retention behaviours of acetone and the four PAHs were characterised in Section 3.3.2, under the conditions used here.

4.2.2. Controlled generation of peak overlap

The COMForTS Instrument Control and Data Acquisition (CICADA) module of COMForTS 4.0 was programmed to allow manually initiated injections during the course of a normal chromatographic 'run'. This facility was used to perform sequences of five injections of the five-component solution ('Stock X') at defined intervals (measured with a stopwatch) that were known to be less than the time required to complete the separation of the previously injected sample. The degree of overlap between the separations resulting from subsequent injections was controlled by varying the injection interval.

Let us consider the PAHs in acetone as the five diastereomers of one particular oligomer. From Section 3.3.2, we know that they are not (conventionally) resolved in D_1, but we also know their overall retention times and that they *are* resolved on D_2. If we now make another injection after a short time delay, we can imagine that we have injected a different oligomer comprised, again, of five different diastereomers. If the assumption is made that the injections were made at the same time, it must then *appear* that the oligomers were resolved in D_1. This 5 × 5 injection strategy thus mimics one injection of a 25-

component sample – in which the retention times of all components are known in both dimensions.

Because the columns are serially connected, the conventional online 2DLC interpretation is that of one under sampled first-dimension sample where that under sampling is known to have resulted in peak wrap-around. An alternative interpretation is that the first dimension was well sampled but that the second dimension cycle time was too long. In this case, regardless of the mechanism of overlap, the total *number of analytes* in the sample is known and can be compared to the *number of peaks* detected.

Different injection sequences were used to test the efficacy of COMForTS in cases of both (a) 'moderate' and (b) 'severe' peak wrap-around (overlap):

(a) After the initial injection, three further injections were made at 360 s intervals with the final injection made after a 720 s interval.
(b) After the initial injection, three further injections were made at 240 s intervals with the final injection made after a 480 s interval.

A truly comprehensive separation based on this timed-injection strategy must yield 25 analyte peaks in five identical D_2 separations (of acetone and PAHs) spaced apart along the D_1 axis at the injection intervals.

4.3. Results and Discussion

4.3.1. Separation with moderate physical overlap

4.3.1.1. Conventional online 2DLC interpretation of the separation

A moderately overlapped separation was generated by a sequence of injections, (a) above, made at 360 s intervals with the final injection made after a 720 s interval. The timing of the final, fifth injection, was estimated to produce a separation very nearly free of overlap. The resulting conventional online 2DLC detector signal is shown in Figure 4-1, in which the 25 injected components have resulted in only 19 chromatographic peaks. There is little to no indication as to which peaks are 'pure' and which are not.

This chromatogram is the untreated absorbance signal from one detector (one pixel line on the CCD). The RMS noise was estimated to be ~4.3 mAU with peak-to-peak noise ~13.9 mAU. The ratio of signal to RMS noise was 64. Whilst the Radon transform itself is relatively immune to noise – that is effectively 'averaged-out' by summation of the absorbance values along lines of varying angles – the detection of chromatographic peaks in noisy one-dimensional signals is well known to be problematic. Because the location of these chromatographic peaks is correlated against the frequency transform, the quality of results will diminish with the quality of peak detection. With an array of 2048 detectors at our disposal, COMForTS 4.0 is equipped with a variety of noise-reduction methods that are discussed in Section 6.4.6.

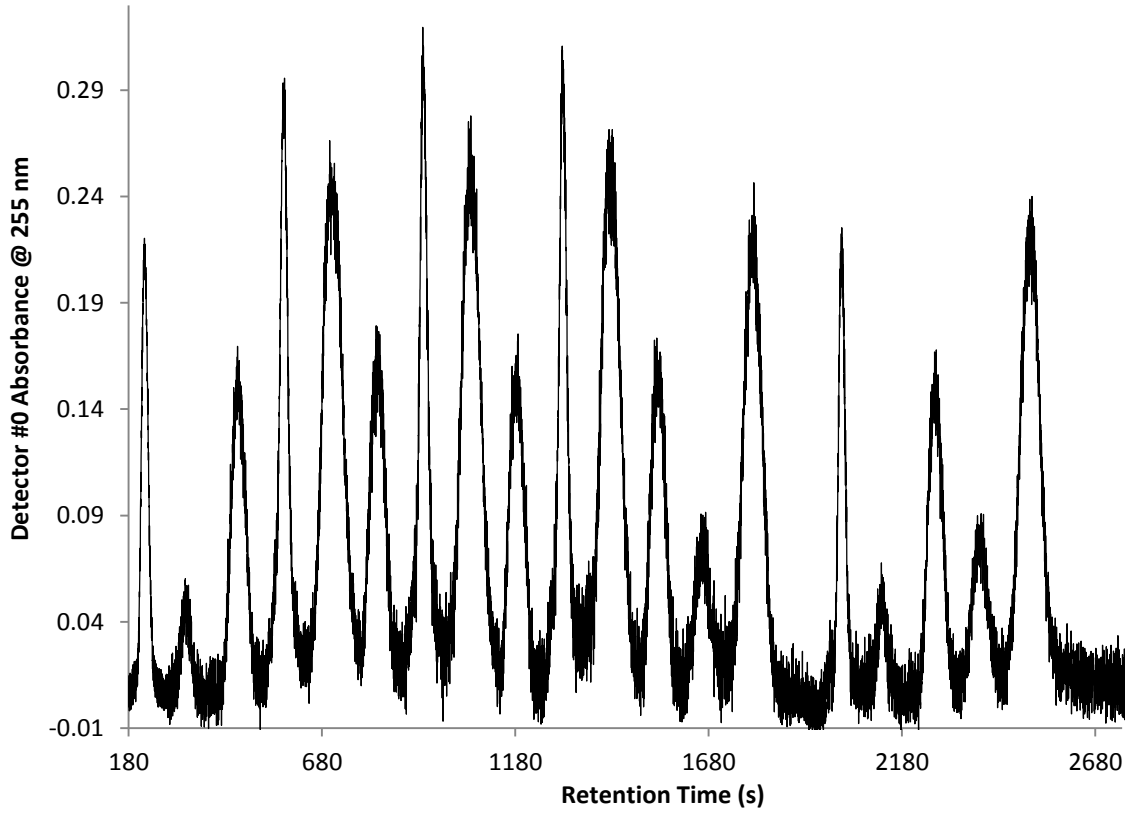


Figure 4-1: Conventional online 2DLC second-dimension detector signal.
The signal was the result of five successive injections of acetone, naphthalene, anthracene, pyrene and chrysene at intervals of 360 s and 720 s. Recorded as a conventional single-detector absorbance chromatogram, only 19 of the expected 25 peaks are apparent.

4.3.1.2. Treatment of arrayed detector chromatograms

To reduce both noise and processing time without losing the benefit in precision afforded by using a large number of detectors, chromatograms were

binned to yield 64 equally spaced virtual detectors each comprised of the Gaussian weighted chromatograms of 32 physical detectors. The array of chromatograms is shown in the density plot of Figure 4-2. It is relatively easy to distinguish 25 peak lines corresponding to 25 analytes with different overall retention and different D_2 linear velocities (line slopes). Several peak lines are seen to physically overlap at either the entry point to the detector array (at detector # 0) or at the exit point (detector # 64) – and this is the reason behind the "loss" of analyte peaks in the conventional single-detector signal. It is much less evident that, though not physically overlapped, there is a high degree of wrap-around: some peaks from one injection have eluted *before* peaks from the *previous* injection.

In Figure 4-2, it is quite difficult to discern the overlapped 5 × 5 retention pattern that has been imposed. When this image is vertically 'squashed' to make the slopes more apparent, the colour removed and contrast enhanced to yield distinct black lines, we can then overlay that image with lines having slopes and retentions that we know to be present (Figure 4-3).

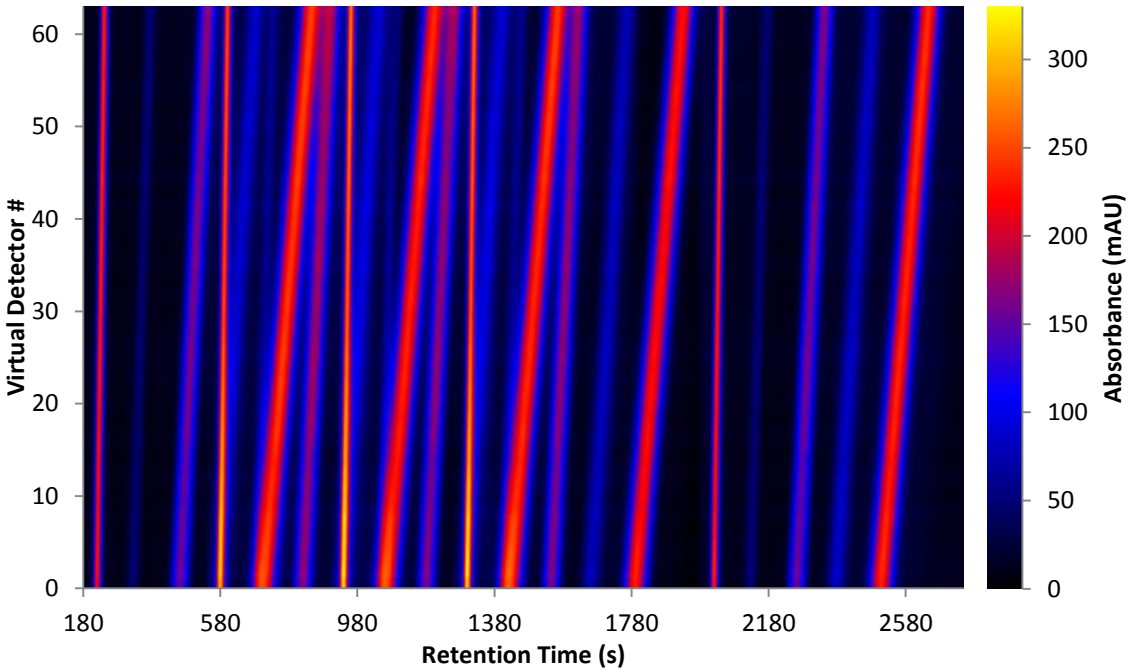


Figure 4-2: Multi-detector chromatograms arrayed in a density plot.
Noise has been reduced by binning 2048 physical detectors to 64 virtual detectors each comprised of 32 physical detectors. Twenty-five peak lines are clearly discernible in this moderately overlapped series.

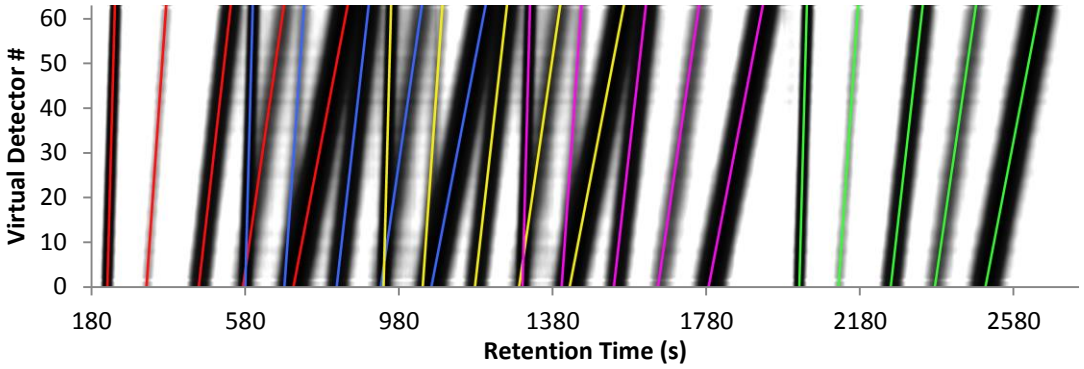


Figure 4-3: Expected and observed retention patterns of moderate overlap.
The expected retention patterns (coloured lines) are overlayed on the contrast-enhanced multi-detector chromatogram. A strong correspondence is evident between expected and observed overall retention and line slope (frequency). All peaks from a single injection are shown with the same colour, with the injection number coloured in order: red, blue, yellow, magenta and green.

All the components from one injection are coded (in Figure 4-3) with the one colour – the coloured lines representing the *expected* retention time and slope. Observationally, there is a very good fit between the expected separation and the observed separation. It is also worthwhile pointing out that, for example, the analyte peaks from the third injection (yellow) are intermingled with peaks from both the second injection (blue) and fourth injection (magenta). By a quirk of timing, some of these peaks are, nonetheless, physically resolved. By a quirk of fate, this demonstrates the uselessness of the conventional online 2DLC separation: not only has a full separation not been achieved, but also where a physical separation *has* been achieved, the first dimension retention times have become muddled and none of the 'peaks' in the (single sample of the) 2D chromatogram can have any meaning. This is indeed chaotic band displacement as discussed in Section 1.1.1.

4.3.1.3. Time dependent frequency domain transform

Let us now return to the COMForTS analysis of these data and the discovery of the slopes and intercepts of the lines in the arrayed detector chromatograms via their Radon transform. The partial discrete Radon transform of the arrayed detector signals (Figure 4-2) is shown in Figure 4-4. A 3D surface plot of a portion of the same transform is used to illustrate some features of the topography in Figure 4-5.

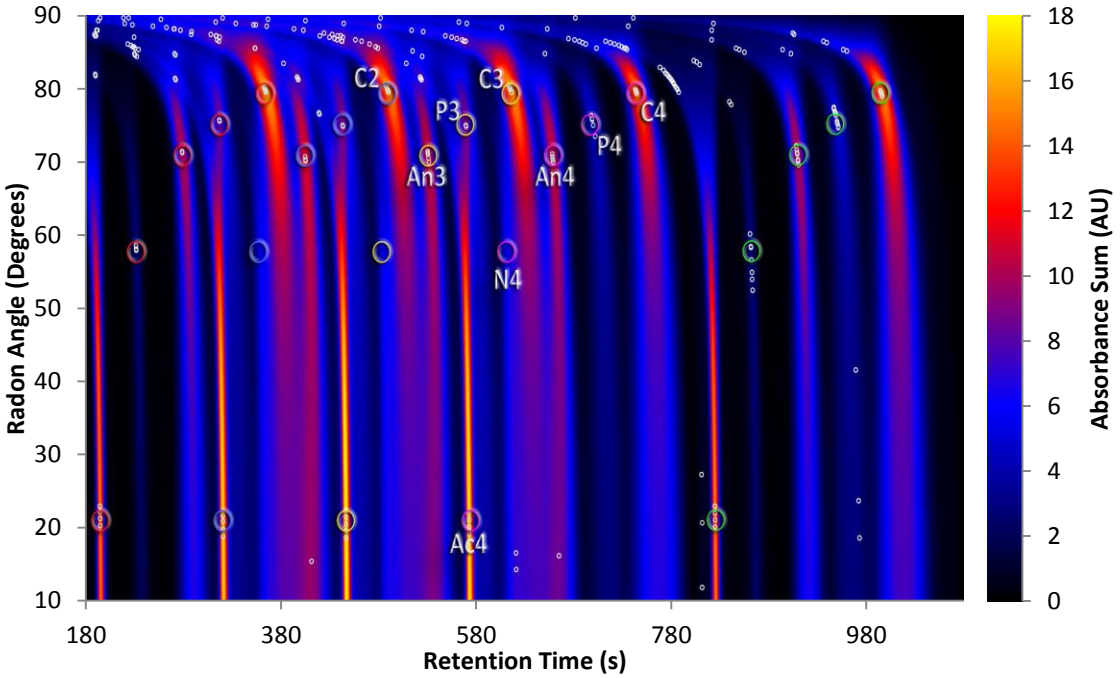


Figure 4-4: Radon transform of arrayed multi-detector chromatograms.
Peaks in the Radon transform correspond to the slope (frequency) and time-intercept (overall retention time) of peak lines. The expected peak locations for each injection are circled, with naphthalene (N) peaks not detected in injections 2, 3 and 4. Other peaks shown are acetone (Ac), anthracene (An), pyrene (P) and chrysene (C).

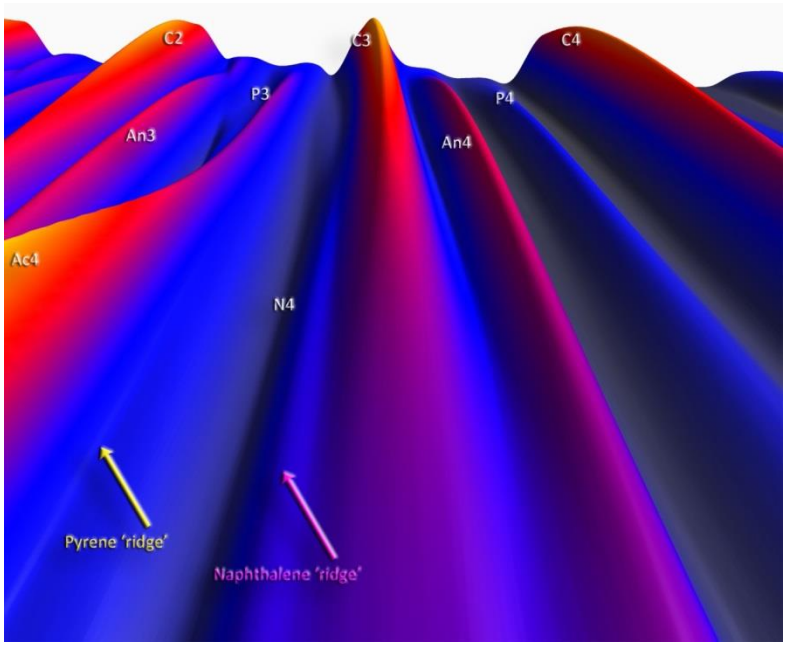


Figure 4-5: Surface plot of a section of the Radon transform.
The section corresponds to the labelled portion of Figure 4-4. Naphthalene typically produced very weak peak lines that remain evident as characteristic ridges in the Radon, even though no clear maximum was detected as was the case for P3.

Conventionally detected peaks within the Radon transform are marked with small white circles in Figure 4-4. As a peak in the Radon, at a particular angle and retention time, corresponds to a line in Figure 4-2, the expected Radon coordinates have been circled and again colour-coded according to the injection sequence. Apart from many harmonic and false peaks, the reader will note that no peaks corresponding to the second, third and fourth injections of naphthalene (marked 'N', at an angle of ~58°) were detected in the Radon.

A surface plot of the region in the Radon surrounding the expected location of the fourth naphthalene peak (marked 'N4') is shown in Figure 4-5. We see that 'N4' has generated a distinct ridge, much like the adjacent pyrene ('P3') ridge. Unlike the 'P3' ridge, the peak detection algorithm, by assuming a conventional conical peak shape, has failed to identify the weak maximum along the 'N4' ridge. Peak detection within the Radon was also hampered by applying the same slope-sensitivity and peak-width parameters as used for peak detection in the conventional chromatograms (and set through the software user interface) – when the topography of the Radon is decidedly different.

4.3.1.4. Correlation of frequency/time data to time domain data

Allowing the above software failure to pass and taking the successfully identified Radon-derived line equations, including false lines, we quickly find that some Radon lines correlate well with the time-domain peak maxima of the arrayed detector chromatograms (shown as small white circles in Figure 4-6). False lines show little or no correlation with these peak maxima, or lie at angles that are physically impossible or have minimum heights that are indistinguishable from the baseline noise – and therefore cannot be measureable analyte peaks.

Because not all peaks were detected in the Radon, some time-domain peaks will be present in the arrayed detector chromatograms that have not been identified as 'belonging' to a known line. A genetic line-search algorithm 'joins the dots' of these orphaned peaks to find the remaining 'missing' lines. The genetic algorithm is terribly slow with large numbers of points but works well to pick up the few lines (Figure 4-6) missed in the analysis of the Radon. Well-correlated time-domain peak maxima are shown in green in Figure 4-6.

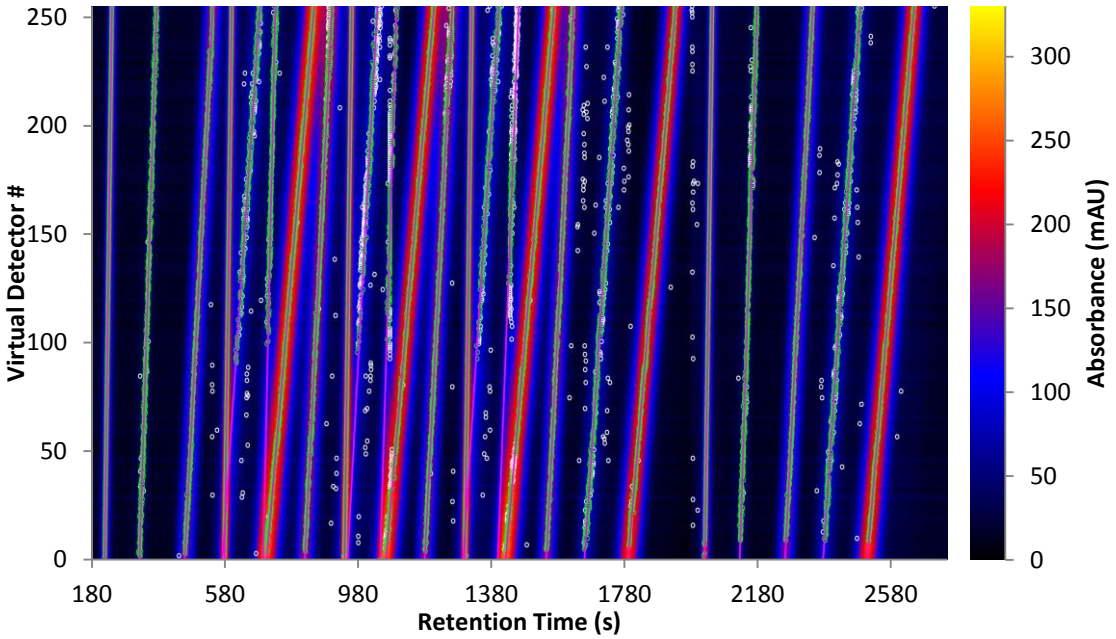


Figure 4-6: Overall frequency-transform result with calculated peak lines.
Peaks in the Radon transform are matched (green) to chromatographic peaks (white) in the arrayed chromatograms, allowing the rejection of transform harmonics. A genetic line-searching algorithm compensates for peaks poorly detected in the Radon. Confirmed peak lines are shown in magenta.

The overall transform result is shown in Figure 4-6, with calculated peak lines corresponding to the calculated slopes and intercepts, shown in magenta. These lines are both observationally consistent with the data and are consistent with their expected slopes and intercepts, within the instrument's limitations.

4.3.1.5. Conversion from the frequency/time domain to the time/time domain

With the slope giving the peak frequency in detectors per second, and the detectors being spaced at a uniform and known distance, we then know the analyte velocity in the second dimension. From the velocity and overall retention time, it is a trivial matter (as discussed in Section 3.3.2.5) to translate these data into first- and second-dimension retention times.

The time/time COMForTS chromatogram resulting from the frequency/time conversion is shown as a density plot in Figure 4-7 and as a 3D surface plot in Figure 4-8. The full resolution of all twenty-five analytes hardly requires mention – though some are identified with such precision that the small error in their location results in peaks that are almost impossible to see in the

density plot. Inspection of these figures also reveals an almost perfect correlation to the expected 5 × 5 retention pattern with peak heights within the instrument's normal range of variation.

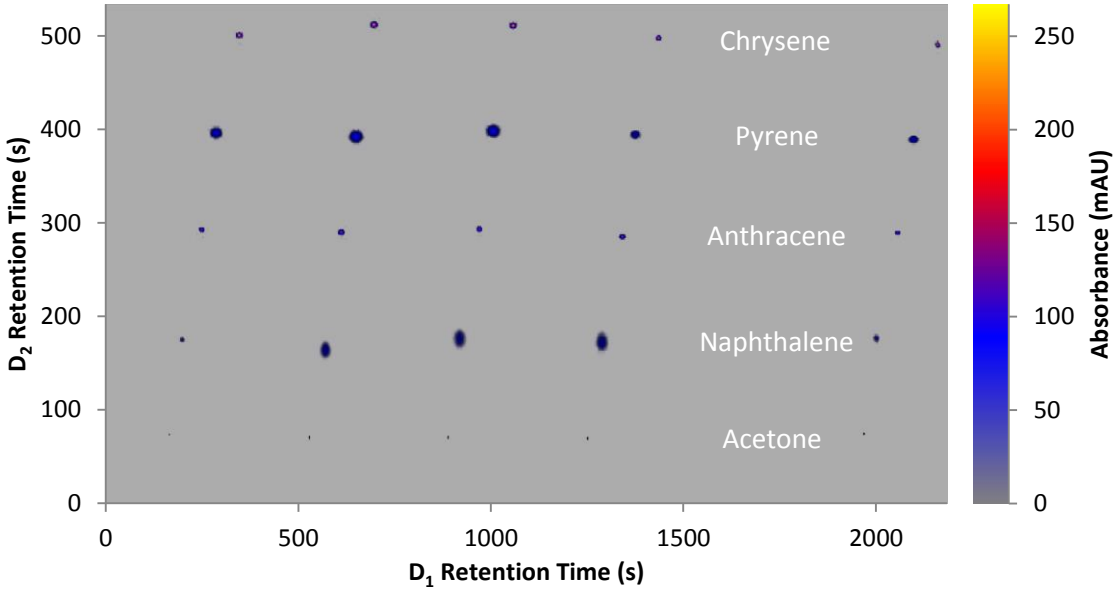


Figure 4-7: COMForTS chromatogram with moderate overlap – density plot.
Whilst absorbance values correspond directly to peak height in this density plot, peak areas represent a Gaussian retention-probability field rather than a physical time-domain peak area. Some exceptionally well-resolved peaks though present in the data, are not immediately apparent in density plots.

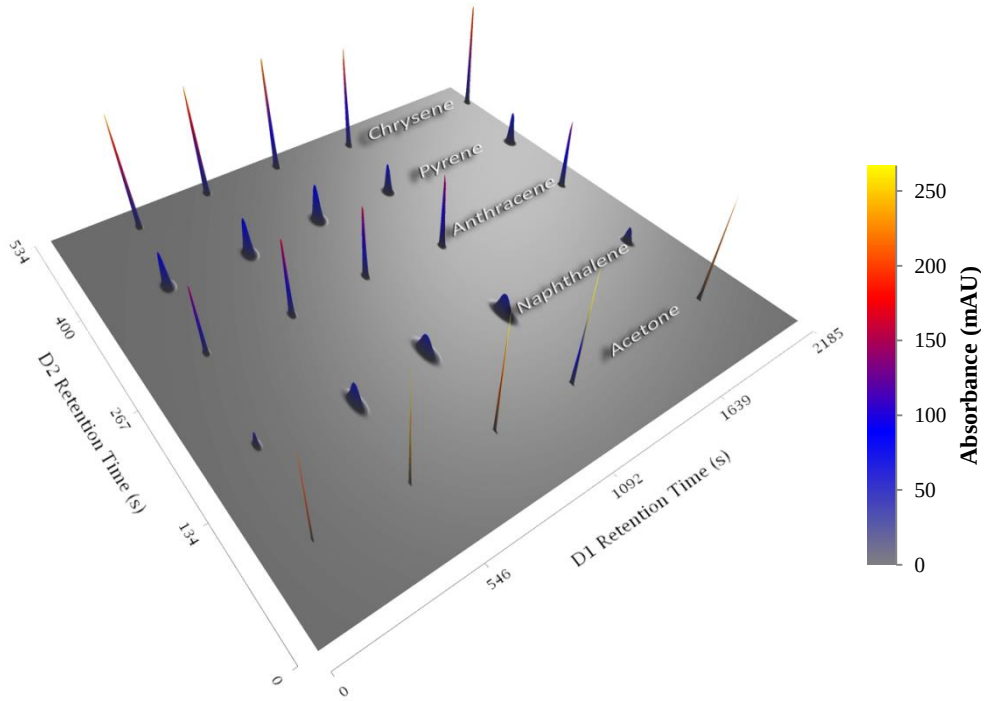


Figure 4-8: 3D surface plot of the 2D COMForTS chromatogram.
All 25 expected peaks are well resolved and conform very closely to the expected retention pattern resulting from five successive injections of the PAH solution.

This example demonstrated considerable peak wrap-around that did not eventuate in a high level of physical overlap but rather an 'interleaving' of D_1 'time slices'. COMForTS analysis, however, thoroughly elucidated the nature of the experiment, which was an artificially engineered sequence of five injections of the same five-component solution made at defined intervals.

4.3.1.6. Qualitative accuracy and precision

Furthermore, the present software demonstrated that there was no loss of accuracy at this level of complexity: the frequencies and retention times of each analyte (using 64 virtual detectors) were indistinguishable (Figure 4-9) from those measured when no overlap was present (using 50 virtual detectors). A substantial loss in the measurement precision of acetone frequencies was, however, unmistakable. The significant difference between acetone and the other analytes was that acetone was not retained in either dimension and its retention time and velocity are dependent entirely on the mobile phase flow rate. A low frequency variation (over a period of perhaps several minutes) in flow rate would result in more-or-less stable overall retention (the total run time was over 45 min) but measureable variation in velocity during the transit time of acetone over the detector array (~ 50 s). That the syringe pump gears were known to 'slip' on their axles when under high load (see Section 5.4.3) suggests that this may have been the cause.

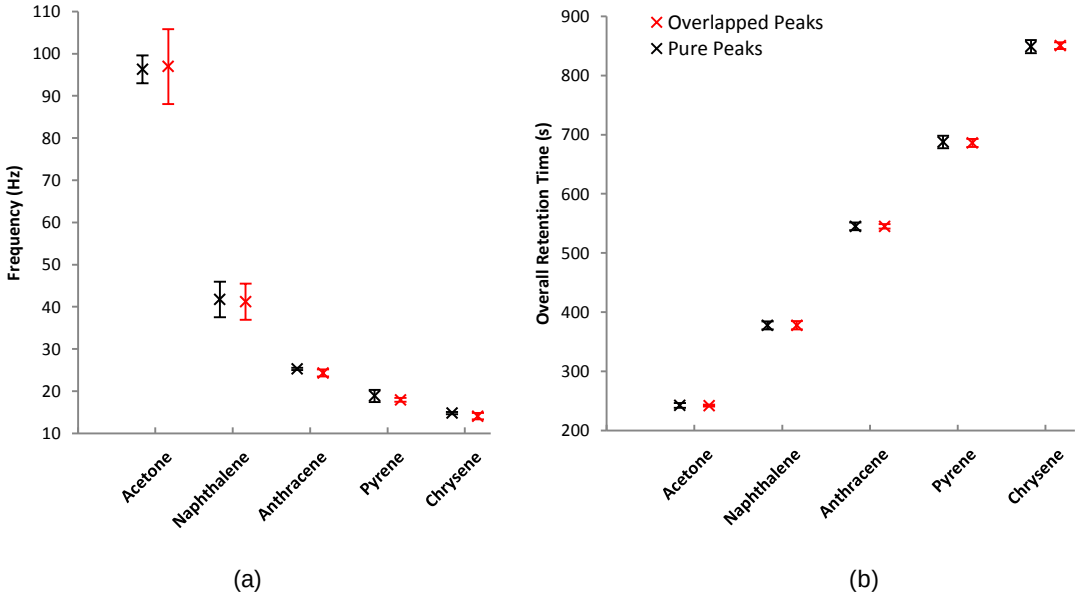


Figure 4-9: Accuracy and precision of COMForTS measurements.
The (a) peak frequency data and (b) overall retention time data from overlapped peaks (red) are compared to the same measurements made on pure analyte peaks (black). Error bars are at 95% confidence levels.

4.3.2. Separation with severe physical overlap

4.3.2.1. Conventional online 2DLC interpretation of the separation

In beginning an exploration of the capabilities and limitations of COMForTS, a severely overlapped separation was undertaken. As in the previous experiment, the separation was generated by a sequence of injections, as described in Section 4.2.2(b), but made at shorter 240 s intervals with the final injection made after a 480 s interval. The timings were estimated to produce a final injection that had little overlap. The resulting conventional online 2DLC detector signal is shown in Figure 4-10, in which the 25 injected components have resulted in only 15 chromatographic peaks. The experienced chromatographer would notice several distorted peaks, but otherwise there is little indication that extensive physical overlap and peak wrap-around has occurred.

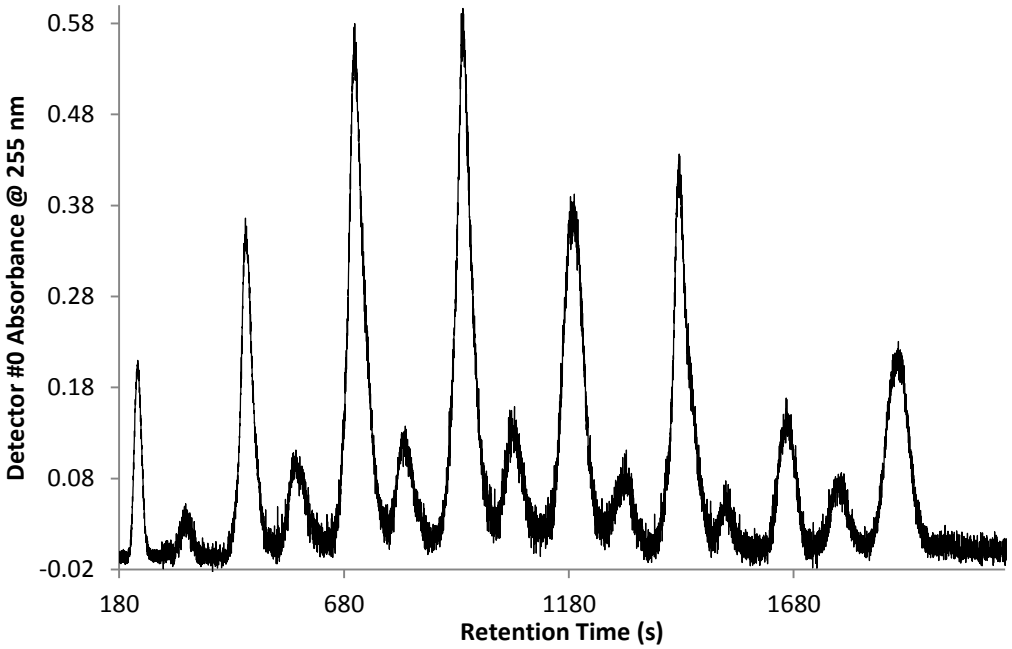


Figure 4-10: Online 2DLC detector signal with severe peak wrap-around.
This was produced by five successive injections of acetone, naphthalene, anthracene, pyrene and chrysene at intervals of 240 s and 480 s. Recorded as a conventional single-detector absorbance chromatogram, only 15 of the expected 25 peaks are apparent.

4.3.2.2. Treatment of arrayed detector chromatograms

The arrayed detector chromatograms (Figure 4-11a) show peak lines that conform well to the expected locations and line slopes (Figure 4-11b), but

peak detection within the Radon (Figure 4-12a) has failed in more cases and some line angle and retention coordinates differ considerably from their expected coordinates.

Signal noise was reduced by Gaussian weighted binning of 2048 physical detectors into 256 virtual detectors each comprising 8 physical detectors. The complexity of these data certainly warrants the use of a greater number of detectors than 256, but the COMForTS instrument's poor S/N requires as much data smoothing as possible in order to accurately identify time-domain peak maxima.

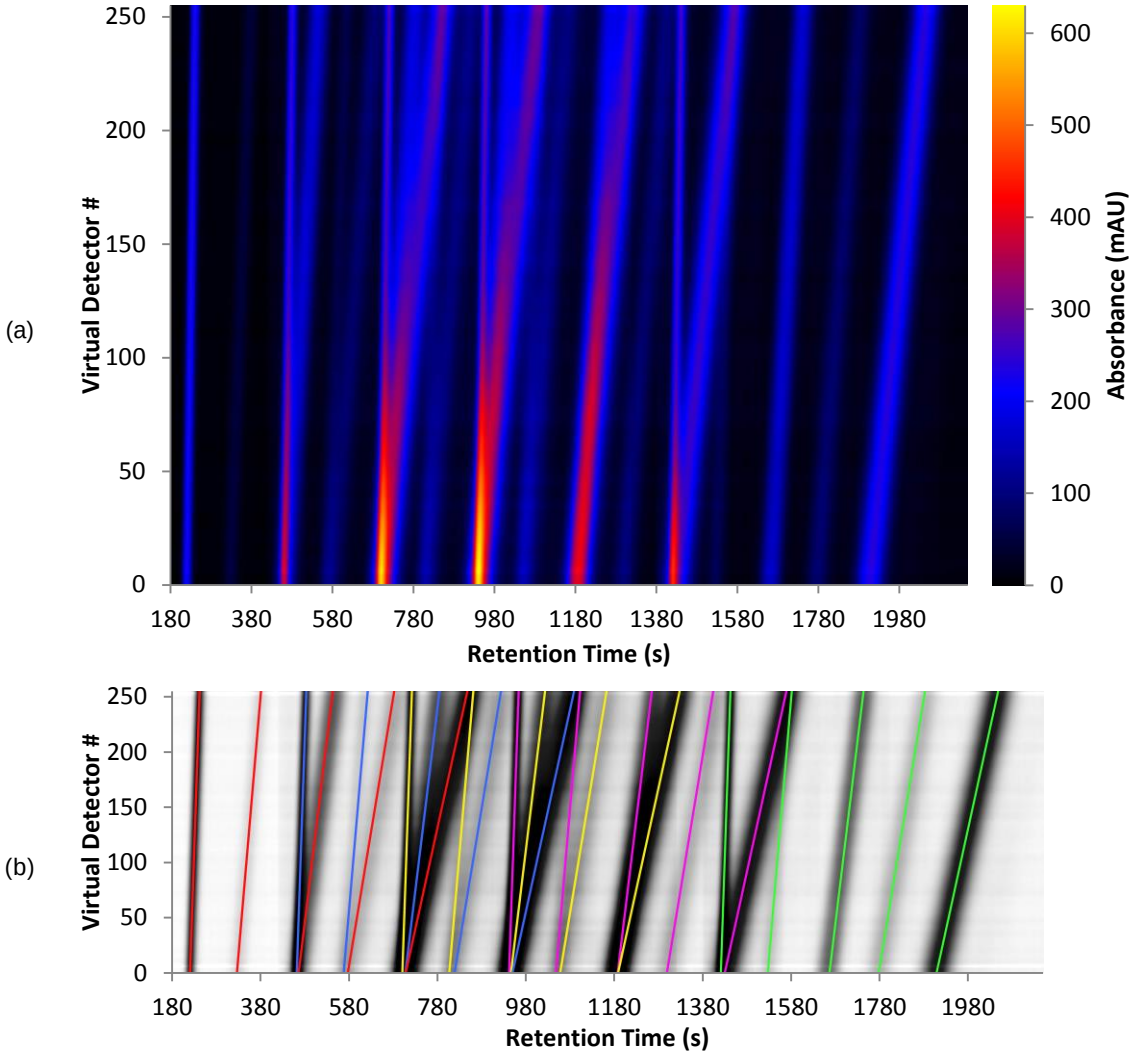
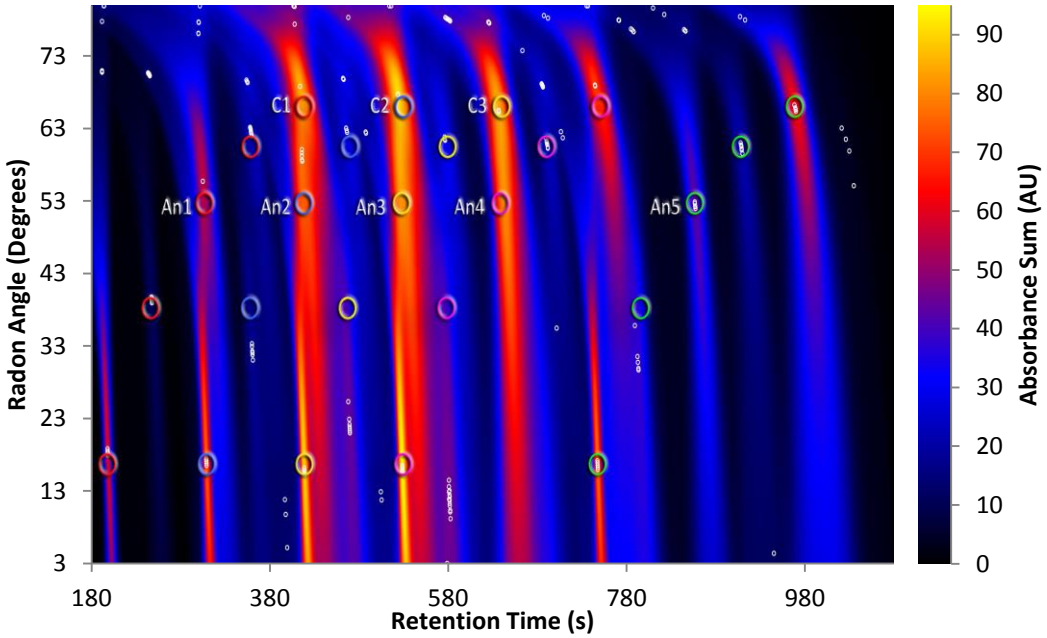


Figure 4-11: Multi-detector chromatograms of severely overlapped peaks.
The arrayed chromatograms are shown as (a) an absorbance-indicating density plot and (b) contrast-enhanced density plot overlaid with expected retention patterns. Twenty-five peak lines are discernible and a strong correspondence is evident between observed and expected retention patterns. Expected peak lines are coloured according to injection sequence: red, blue, yellow, magenta and green.

4.3.2.3. Time dependent frequency domain transform

The loss in precision (otherwise afforded by a higher number of detectors) combined with the inadequate conical peak detection within the Radon transform contributed to reduced efficacy of the algorithms. The Radon peaks shown in the surface plot of Figure 4-12b are clearly crescent-shaped and not conical and the anthracene peaks ('An' 1 to 4) were not detected despite their evidence (Figure 4-12b) in the Radon.

(a)



(b)

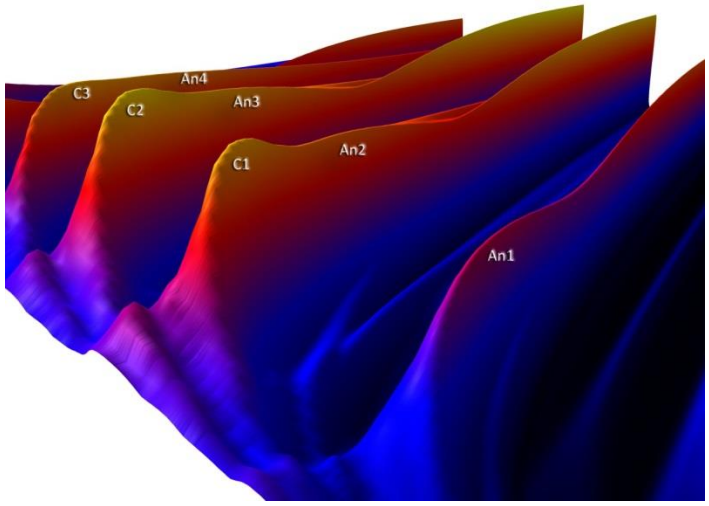


Figure 4-12: Radon transform of the arrayed detector signals.
This transform resulted from the severely overlapped separation. The intensity plot of the entire transform (a) shows coloured circles at the expected peak locations for each injection. Each analyte elutes at a characteristic frequency (indicated by the Radon angle) in increasing order: acetone (Ac), naphthalene (N), anthracene (An), pyrene (P) and chrysene (C). A detail section of the Radon is shown as a surface plot (b) where anthracene peaks are visible but were not identified by the peak-picking algorithm.

In COMForTS 4.0, detection and assessment of many peaks then fell entirely into the hands of the genetic line-search algorithm – simply because of peak detection that was not optimised for the Radon space. The result is that, even though this chromatogram was acquired over significantly less time than the moderately overlapped example and the volume of data proportionately less, the processing time is markedly longer. Still, the data acquired at 4 Hz over 35 minutes was processed through to completion in under 20 minutes. This was faster than real-time, but would benefit from the application of additional, relatively cheap, computing power.

4.3.2.4. Correlation of time / frequency data to time domain data

The final fully automated transform results, depending heavily on the linear search alone (with no cross-correlation), contained several lines that were clearly not peak lines, as well as a number of 'lines of best fit' that, to the human eye were not cleanly placed above the peak line evident in the arrayed chromatograms. The COMForTS 4.0 RADON module allows some user "direction" in that the user can 'show' the algorithm where the user believes a line should be by drawing that line on the screen plot. The genetic line search algorithm then employs the user-suggested line as a starting point in the search for a line of satisfactory fit to the data. The end-result of this process is shown in Figure 4-13.

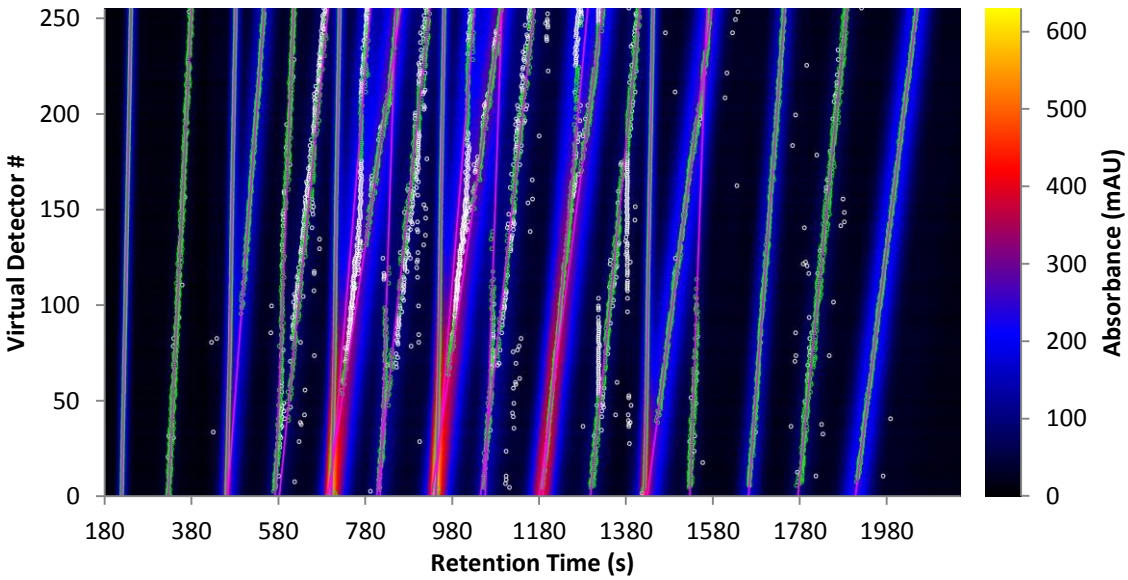


Figure 4-13: Overall frequency-transform result with severe overlap.
Peaks in the Radon transform are matched (green) to chromatographic peaks (white) in the arrayed chromatograms, allowing the rejection of transform harmonics. In the region from 580 to 1380 s, many chromatographic peaks were found that did not match calculated line data, thereby distorting the results.

There are many instances in Figure 4-13 where time-domain peaks (white) have somewhat suspect positions within the 'V' of strongly physically overlapped peaks. One such instance is in the region of about 1250 s between detectors 200 and 255. The 'better' time-domain maxima (further from the confluence of the 'V' were not used to construct the line because of the proliferation of nearby "poor" examples belonging to one – or both – of two peaks that were physically overlapped to a considerable extent during their transit times over the array.

4.3.2.5. Sources of error in the COMForTS analysis

Now, it is already apparent that the results of the signals processing (Figure 4-13) are not an entirely accurate reflection of the observed chromatographic behaviour (Figure 4-11b). We can see that the arrayed chromatograms at least very closely match the expected pattern dictated by the applied injection sequence and the known D_2 retention behaviour of the analytes. The problem, to put it anthropomorphically, is getting the computer to see this pattern.

In this severely overlapped experiment, peak detection within the Radon transform was inadequate and the COMForTS 4.0 analysis relied heavily on the results of the genetic line-search. The genetic line-search, for its part, relies entirely on the location of time-domain peak maxima in each of the detector chromatograms. Where the physical resolution is low and this is combined with low signal to noise, as it is in this example, even this task becomes challenging.

Unable to cross correlate the retention coordinates, and with only poorly measured data available, the software reports an increased uncertainty in the determination of those coordinates. The baseline area of COMForTS peaks reflects this uncertainty (as discussed in Section 3.3.2.4) and the consequent reduction in peak capacity.

This degree of physical overlap and signal noise can be dealt with by the human mind – through visual inspection of the arrayed signals - but the present algorithms of COMForTS 4.0 did not cope particularly well with extraction of the information that we can see to be present.

4.3.2.6. Conversion from the time/frequency domain to the time/time domain

Surprisingly respectable results were nonetheless achieved in the converted 2D time/time chromatogram (Figure 4-14).

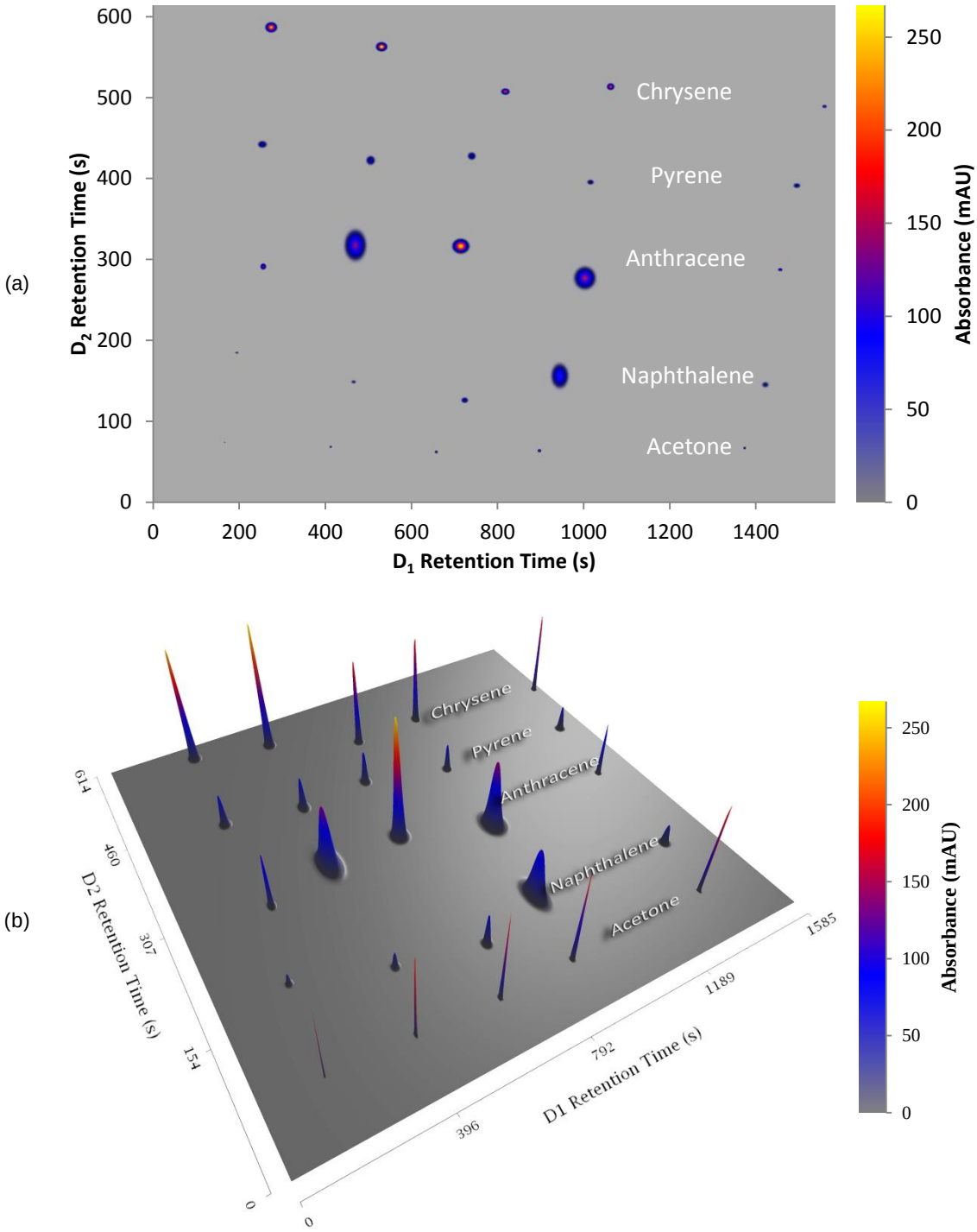


Figure 4-14: COMForTS chromatogram of a severely overlapped separation.
(a) Intensity and (b) surface plots of the final two-dimensional time/time chromatograms as determined by the COMForTS 4.0 software.

4.3.2.7. Qualitative accuracy and precision

As expected, the accuracy and precision of measured frequencies and retention times suffered (Figure 4-15) within this degree of complexity: at one stage, 80% of analyte peaks were *physically* overlapped and only five analytes were not ever physically overlapped. Most affected was naphthalene, with the lowest detector response. Bracketed by two peaks, acetone and anthracene, that were respectively 10× and 5× larger, it was statistically possible that, in the frequency domain alone, naphthalene might not be resolved from anthracene (Figure 4-15a).

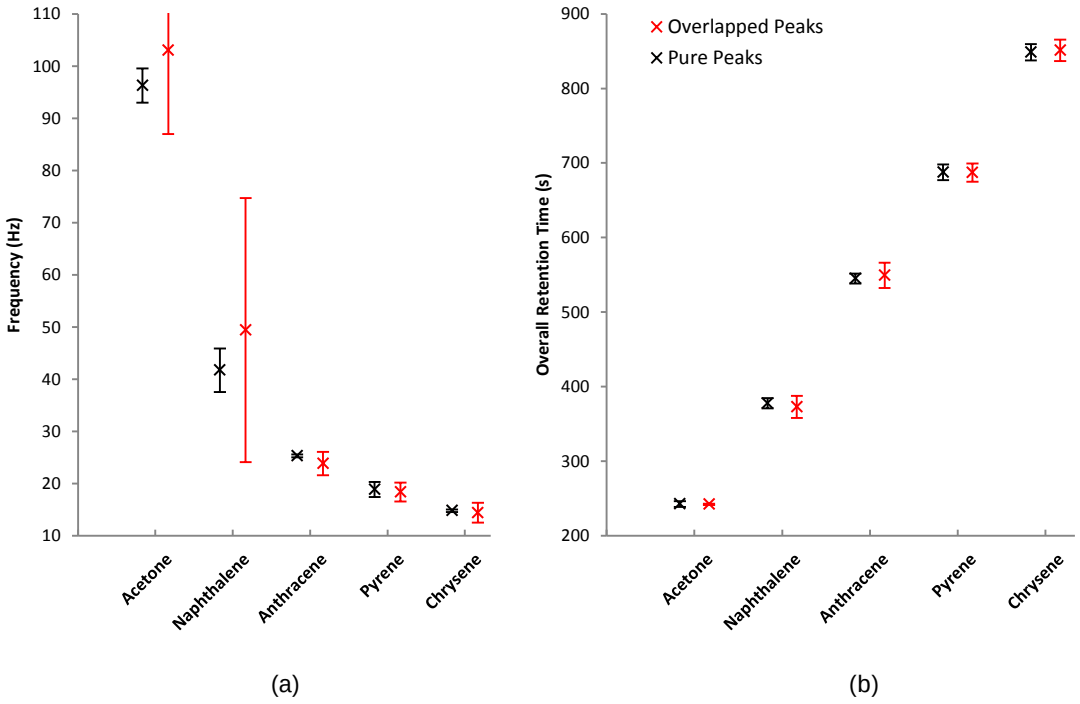


Figure 4-15: Accuracy and precision of COMForTS with severe overlap.
(a) Peak frequency and (b) Overall retention time under conditions of severe wrap-around (overlap). The data from overlapped peaks (red) are compared to the same measurements made on pure analyte peaks (black). Error bars are at 95% confidence levels.

Another factor affecting this particular analysis (but not COMForTS analyses in general) was small, irregular, instantaneous changes in the output of the UV LEDs. In Figure 4-13 in the region from 1280 to 1380 s, two vertical lines of marked "peaks" indicate two such changes in the baseline. Rapidly 'paging through' individual detector chromatograms (within COMForTS 4.0) reveals other 'stationary' distortions in the baseline that often adversely affect the shape of time-domain peak maxima. One may reasonably conclude that the increased errors in this result were a function more of the capabilities of the

instrument and its software than of the ability of the COMForTS method to resolve analytes within this extraordinary degree of wrap-around.

Whilst deformed by these errors, the COMForTS separation of the 25 severely overlapped analytes again correctly identified the multi-injection nature of the experiment.

4.3.3. Separation resolution

While the errors of the wrap-around generation experiments increase with the severity of the degree of overlap, these errors were measured relative to COMForTS' own evaluation of the same *2D* separation where there was no overlap.

If we go back to the conventional separation of the test solution on a longer DB-1 capillary, there was no observable separation whatsoever (see Figure 3-9). In contrast, a cross section of the moderately overlapped COMForTS separation, looking along the D_1 axis is shown in Figure 4-16. This reveals the consistent reversed-phase separation pattern (noted in Section 3.3.2.7) that one might expect from a DB-1 stationary phase.

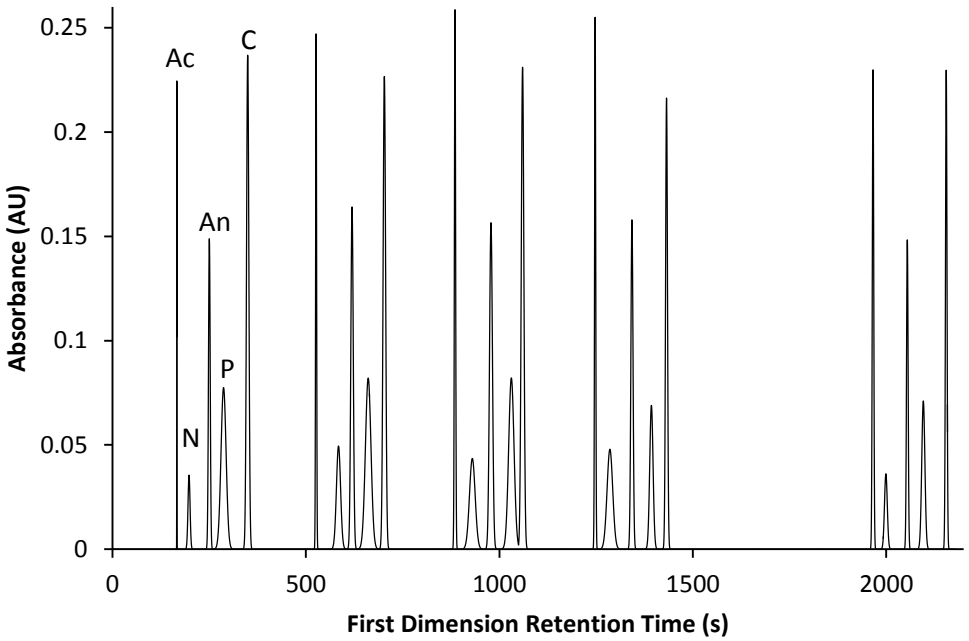


Figure 4-16: Cross section of the COMForTS first dimension separation.
From the separation of five successive, moderately overlapped injections of four PAHs in acetone. D_1 Column: 210 × 0.250 mm DB-1, 0.25 μm film thickness. Mobile phase was 70:30 ACN:Water at a flow rate of 5 μL/min. The components in order of elution within each injection group are acetone (Ac), naphthalene (N), anthracene (An), pyrene (P) and chrysene (C).

COMForTS time-domain resolution in the first dimension consistently far exceeded the resolution of the physical separation even when moderate physical overlap and severe wrap-around effects were present.

The resolution of COMForTS in the frequency domain was regularly inferior to the time domain resolution – and was largely limited by the detector binning required to reduce signal noise. Doubtless, this resolution will improve markedly in instruments with better S/N – or with more than 2048 physical detectors. All the same, the second dimension time-domain resolution of COMForTS, which is derived from the frequency domain, was at least equal to the physical resolution (Figure 4-17). The precision of retention times derived from overlapped frequency data was also roughly equal to that of conventionally measured retention times in non-overlapped separations.

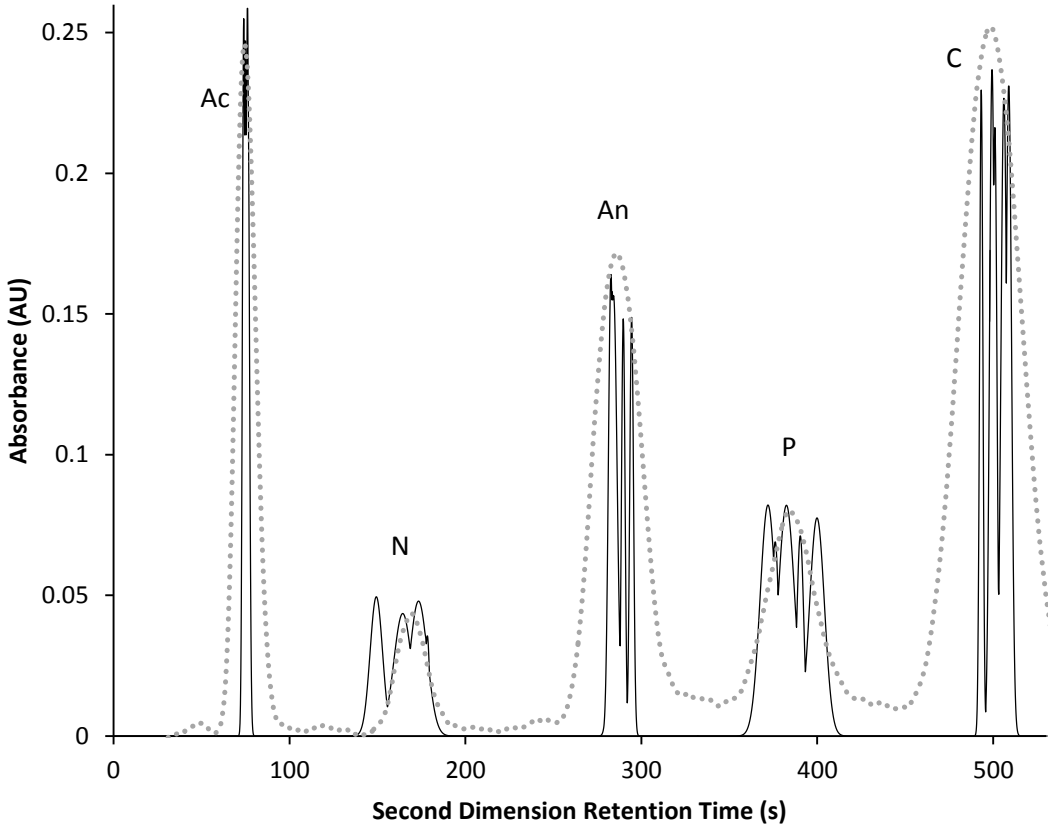


Figure 4-17: Cross section of the COMForTS second dimension separation.
From the separation of five successive, moderately overlapped injections of four PAHs in acetone. D_2 column: 145 × 0.318 mm 5μm Nucleosil C18. Mobile phase was 70:30 ACN:Water at a flow rate of 5 μL/min. The components in order of elution are acetone (Ac), naphthalene (N), anthracene (An), pyrene (P) and chrysene (C). The multiple peaks for each component correspond to each of the five injections. The dotted trace shows the relative, conventionally recorded physical resolution of the components.

4.3.4. Peak capacity and production

This thorough demonstration of the use of the time and frequency domains to overcome extensive peak wrap-around and physical peak overlap implies that COMForTS must also impart a large increase in effective peak capacity. Plainly, the peak capacity of the columns used was not up to today's standards and a direct comparison would be meaningless. What does matter, however, is that the separations were simultaneously treated as conventional online 2DLC separations. Such online 2DLC separations would come about if either the first dimension sampling time is too short or the second dimension cycle time is too long – or, indeed, some compromise between both.

The COMForTS separations and the conventional treatments were achieved in exactly the same total analysis time – the only difference being the number of peaks resolved in that time. If therefore, the chromatographic efficiency of the system were improved, the same separation could be achieved in less time – but still in the same total time for both methods. In other words, changing the time scales of the chromatograms in this experiment from seconds to milliseconds would make no difference to the relative numbers of peaks that would be resolved by either method.

The two-dimensional peak capacities $n_{c,2D}$ for both methods were calculated using Eq. (1-4) and the product rule, Eq. (1-5), though in the case of COMForTS, 8σ was used rather than 4 σ, where σ was the standard deviation of the estimates of peak location in the time/time domains. This was done to allow for underestimated or undetermined errors in the estimates. Peak production was calculated as $n_{c,2D}$ / (T_r – T_0). The results, together with the multiplicative increases afforded by COMForTS, are listed in Table 4-1.

		Conventional online 2DLC	COMForTS	COMForTS Improvement Factor
Moderate overlap	Peak Capacity, $n_{c,2D}$	20	832	41
	Peak production ($n_{c,2D}$/min)	0.5	19	41
Severe overlap	Peak Capacity	16	404	26
	Peak production ($n_{c,2D}$/min)	0.5	12	26

Table 4-1: Relative peak capacity and peak production COMForTS / 2DLC

4.3.5. COMForTS and statistical peak overlap

The basics of statistical peak overlap theory (SOT) discussed in Section 1.1.3 relate to the problem of not being able to resolve components that are physically overlapped. SOT describes the probability that such an overlap will occur in a system of a given peak capacity and known number of components.

In 2DLC, components may be physically overlapped in the first dimension but must not be physically overlapped in the second dimension if the number of observed peaks is to equal the number of components. This is clearly not the case for COMForTS. In each of the experimental separations, all components were physically overlapped in the first dimension (with four other peaks). In the second dimension, up to 80% of the 25 components were physically overlapped. Yet COMForTS fully resolved all 25 components. SOT states (using Eq. (1-11)) that it is essentially impossible to resolve 25 components into 25 peaks where the two dimensional physical peak capacity $n_{c,2D}$ is 20.

A unique feature of COMForTS is that it can demonstrably resolve physically overlapped peaks – with the only requirement being that no two analytes overlap each other in all separation dimensions. If we then consider a 2D COMForTS system, with $n_{c,2D}$ = 832 (as above) and that each dimension has the same peak capacity (i.e. $n_{c,D1}$ = $n_{c,D2}$ = 29), the probability that two peaks (out of 25) will co-elute in the first dimension is 25/29 = 0.86. However, the probability that these *same* two peaks *also* overlap in the second dimension is 0.86 × 1/29 = 0.03. The interpretation of this is that COMForTS has resolved 25 peaks with a 97% probability that each peak represents a *pure* component.

However, SOT, again using Eq. (1-11), demands that $n_{c,2D}$ must be greater than 3000 in order to have a 97% certainty that the peaks are pure.

At this stage, we understand well that COMForTS has not provided a physical separation, but provided increased 'dimensionality' by using the frequency domain in addition to the conventional time domain. The COMForTS $n_{c,2D}$ = 832 is therefore an *effective* peak capacity according to Stoll [33] and Schure [39]. Somewhat surprisingly, COMForTS has performed a separation that would *normally* require a peak capacity of 3000 (not 832). The COMForTS method must therefore make some 'unseen' contribution to dimensionality that produces an orthogonal peak capacity of 3000 / 832 = ~4.

4.4. Conclusion

The theoretical basis of COMForTS and the assertion that COMForTS could overcome peak wrap-around were experimentally demonstrated to be valid.

In separations constructed to contain extensive wrap-around and moderate to severe physical peak overlap, COMForTS resolved all components in all cases with up to 41-fold improvements in peak capacity and peak production over conventional 2DLC. The precision and accuracy of results translated to the time/time domain were at least equal to conventional separations performed with the same instrument.

Very good quantitative results, in terms of peak height, were obtained for moderately complex separations. Harmonics, overtones and false/missing peaks were eliminated when frequency transform data were correlated to conventional chromatograms arranged in a two dimensional array. In very complicated separations, deficiencies in the processing software did not always allow this correlation. As a consequence, the precision and accuracy of the results suffered, thereby broadening peaks and subsequently reducing the potential peak capacity. Under these difficult circumstances, lower but still meaningful improvements in peak capacity of the order of 26-fold were obtained. These deficiencies and their remedies will be discussed later in the thesis (Chapter 6). It is apparent there will be some loss in peak capacity with increasing separation complexity, but it would be expected to be much smaller than the extent observed here. Several areas of improvement in the software and algorithms have been identified from these experiments and are discussed further in Chapter 7.

Modelling of physically pulsed COMForTS separations had predicted that, with improved data treatment, COMForTS peak production could reach rates up to eight times greater than conventional online 2DLC (Table 2-1). It was anticipated that application of the method to virtually pulsed signals would yield similar results. It is usual that such predictions are eventually shown to be apocryphal but not so usual that they so greatly underestimate a benefit. Therefore, there must be some advantage imparted by the migration to virtual pulse detection that was not considered in the modelled separations.

At the cost of flexibility in detection methods and the concomitant increase in the cost of COMForTS detectors, virtual pulse generation through multi-point detection also removed the limitation imposed by the lifetime of physical pulses. It also ensured that, no matter what the pulse frequency, an equal and high number of pulses would be recorded for each analyte. The outcome is that D_2 separations are no longer limited in analysis time or peak capacity and that consistently large numbers of pulse measurements increase the resolution of information within that same physical peak capacity (as shown in Section 2.4.10).

Furthermore, there is a stark contrast between the D_1 resolution in COMForTS and in online 2DLC. By physically sampling D_1, online 2DLC loses resolution in D_1 and therefore presents less resolution than would be available in a D_1 detector signal [23]. Oppositely, these experiments showed that COMForTS had notably *greater* resolution of information in D_1 than was physically available.

Another somewhat startling result was found with respect to COMForTS and statistical peak overlap (discussed in Section 1.1.3), with SOT declaring that a separation of components as pure peaks with the level of confidence reported here cannot be achieved even with COMForTS' calculated effective peak capacity. In terms of SOT, if we call the capacity to separate components into individual peaks, the "analyte capacity", then facility to achieve that in *effect* (by an informational separation) may be called the "*effective* analyte capacity". Using this terminology, 2D COMForTS has demonstrated an "effective analyte capacity" some three orders of magnitude greater than conventional online 2DLC – and greater even than the increase ideally afforded by FT-ICR MS [33,34]. This subtle but important difference is not adequately described by Schure's concepts of dimensionality [31,39] and is clearly a matter worthy of further investigation.

This unprecedented level of certainty in peak purity has important implications for real-world applications of the COMForTS method, some of which will be touched upon in Section 7.8, Applications of COMForTS.

Chapter 5.
COMForTS Instrumentation

5.1. Introduction

In Chapter 2, a theoretical basis of COMForTS was developed with a particular emphasis on physical pulse modulation because of its associated advantages of flexibility in detector choice. The study in Chapter 3 also showed that physical pulse modulation was experimentally difficult (Section 3.3.1) and that a basic proof-of-concept would be more readily realisable in a virtually pulsed system, fundamentally operating in the 'COMForTS-*is*' mode, i.e. with in-separation detection.

Taking the instrument configuration for COMForTS-*is* (Figure 2-1a) and replacing the switching valve with a second-dimension array of equally spaced in-separation detectors forms a generalised physical configuration for COMForTS that is shown in Figure 5-1.

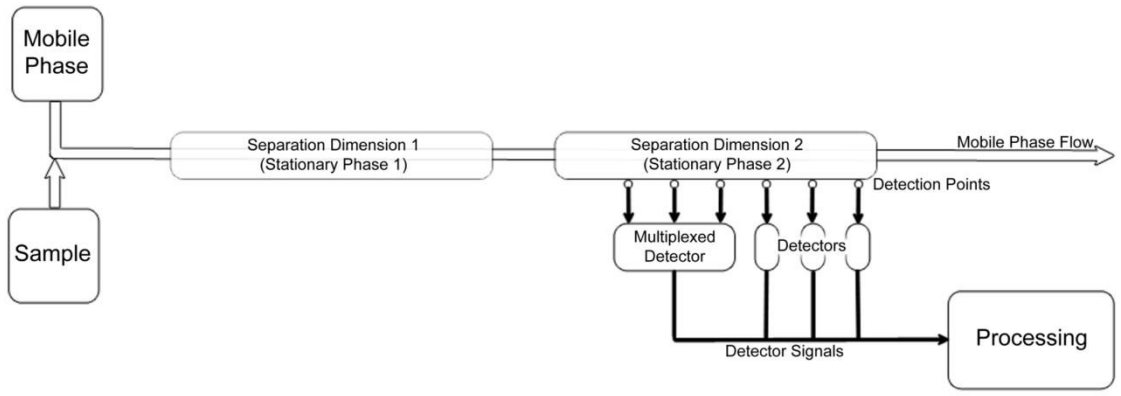


Figure 5-1: Generalised 2D COMForTS system with multipoint detection.
Multipoint detection may be achieved by using multiple detectors at equally spaced detection points or by using a single detector multiplexed to a number of detection points.

The only requirements of the separation dimensions are that analytes are separated based on differing velocities through differing stationary phases and that some form of arrayed in-separation detection may be implemented.

This chapter is intended to provide an understanding of the reasons for the particular choices, compromises and adaptations involved with the reconciliation of the theory of COMForTS with the construction of a practical instrument.

5.2. A Prototypical COMForTS Instrument

The chosen physical specification of the prototypical COMForTS instrument was that of a serially coupled two-dimensional capillary liquid chromatograph driven by a syringe pump, using timed injection and on-column CCD-based UV-absorbance detection.

In Figure 5-1, a section of tubing is shown between the separation dimensions. In the prototype, the two columns were joined by a zero dead volume connector – and this is the ideal. With zero dead volume between the dimensions, the time of exit from the first stationary phase D_1 is equal to the time of entry into the second D_2 stationary phase and no further calculations are required to determine D_1 retention times. On the other hand, a short section of tubing (such as a T-piece) may be necessary to allow for fluid connections where a change in mobile phase composition is desired (see Sections 7.4 and 7.5).

The arrangement of the major system components is shown in the photograph of the COMForTS prototype instrument in Figure 5-2. Many of the major components were available 'off-the-shelf', but some (such as the driver circuitry for control of the LED light sources (see Appendix A)) were built in-house.

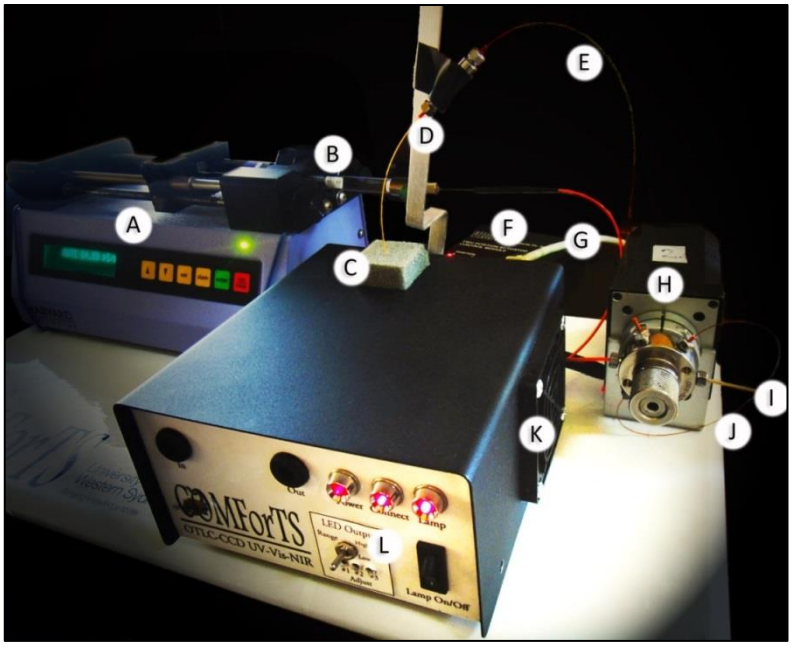


Figure 5-2: COMForTS Instrument Prototype.

Key to Figure 5-2:

A. Harvard Apparatus 11 Plus syringe pump
B. 1000 μL Pressure-Lok® syringe.
C. External light baffle
D. Second dimension column (packed capillary shown) connected to the first dimension column via a standard zero-dead-volume connector.
E. First dimension column (open-tube capillary shown) connected directly to injection valve (H).
F. VICI Valco two position micro electric valve actuator; Control Module
G. USB 2.0 communications for detector data acquisition and control
H. VICI Valco 6-port, 2-position valve with motor and gearbox assembly.
I. Sample introduction port
J. Capillary injection loop. The loop is reversed flushed directly onto the head of the first dimension column.
K. Detector ventilation fan.
L. LED High / Low power range selection and fine power-level adjustment (for setting optimum zero absorbance output level)

5.2.1. Solvent delivery

Isocratic solvent delivery was achieved by means of a syringe pump (Model 11 Plus, Harvard Apparatus, Holliston, Massachusetts) equipped with a 1000 μL Pressure-Lok® syringe. The syringe pump could be set to deliver different flow rates (in microlitres per minute) once the user had set the internal diameter of the syringe (in millimetres).

5.2.2. Sample injection

Our laboratory was not equipped with an injector capable of the low (sub-μL) injections required for capillary LC. Instead, injections were made using a sample loop (consisting of a short section of capillary tubing) using timed switching of a six-port two-position switching valve that was fitted with an electric actuator and microcontroller (Valco Instrument Company Inc., Houston, Texas) (see Figure 5-2).

Desired injection volumes were achieved by reverse flushing the filled sample loop for a period that corresponded to the required volume at a particular mobile phase flow rate.

5.2.3. Detector array

A large detection array is required in order to generate a sufficiently large number of pulses for the frequency transform. Most SCOFT experiments, for example, used 50 to 55 detectors [48,60,83]. Large light sensitive diode arrays are very expensive and not particularly sensitive. On the other hand, charge coupled devices (CCDs) are available in much larger arrays and offer greater sensitivity (up to 60% quantum efficiency at UV wavelengths and 80% in the visible region [118]) approaching that of photomultipliers. Furthermore, CCD's are readily available in two-dimensional arrays and each element (pixel) of a CCD can be sampled virtually simultaneously – as in digital photography. Each pixel can be treated as a separate detector and the entire array of detectors sampled at up to 1000 frames per second – more than capable of handling the relatively high data rates predicted by modelling (Sections 2.4.8 and 2.4.9).

Detection was by the COMForTS CCD-based, single beam, UV-absorbance detector (built in-house, as described in Appendix C) based on a Hamamatsu Photonics CCD chip that was sensitive to wavelengths from ultraviolet to near infrared.

The choice of a CCD as a detector array presented difficulties with respect to the spacing of the detectors, which is usually equal to the pixel width (ordinarily between 7 and 20 μm). For adequate pulse generation by summation of the individual detector signals (see Figure 1-3), the distance between the detectors should be greater than half the average peak-width.

Because peak widths of the order of a few millimetres to a few centimetres were expected, the spacing of the pixels on the CCD chip was far too small to allow simple summation of detector signals as had been previously [48,83].

As already discussed (see 3.1.2, 3.3.2 and 3.4), the problem of detector spacing, with respect to peak width, was resolved by the formulation of alternative treatments of the individual detector signals. The Radon transform was the most valuable of these treatments and its use is discussed throughout the thesis. Also developed, for use with the short time Fourier and wavelet transforms, was a method for combining signals from very closely spaced detectors (MCSCSD, discussed in Appendix D) and a linear transform by genetic line-searching (Section 6.4.3). Improvements to these methods beyond the present work are identified and discussed in Section 7.6.

These methods for detector signal treatment (Radon and MCSCSD) made it possible to design a compact instrument in which the detectors were very closely spaced. The instrument design that was finally implemented was thus a two-dimensional capillary liquid chromatograph where on-column detection was achieved by laying the second dimension column directly against the CCD's quartz coverslip with the whole illuminated by a UV LED as in Figure 3-1.

5.2.4. Open tubular D_1 capillary columns

Open tubular capillary columns (for the first dimension) were cut from a 25 m DB-1 GC column with 250 μm ID and 0.25 μm film (J&W Scientific, Folsom, California) and equipped with standard HPLC fittings by means of PEEK tubing sleeves. The entry end was connected directly to the injection port of the six-port valve and the exit end connected with a zero-dead-volume connector to the packed D_2 column.

5.2.5. Packed D_2 capillary columns

5.2.5.1. Column preparation

Packed capillary columns (for the second dimension) were manufactured in-house using 318 μm ID fused silica, polyimide-coated capillaries (TSP320450, Polymicro Technologies, Phoenix, Arizona) packed with 5 μm Nucleosil C18 particles (Sigma Aldrich, Australia).

A typical packed column used in this work is shown in Figure 5-3, with the final 60 mm of the polyimide coating removed to create a detection 'window'.

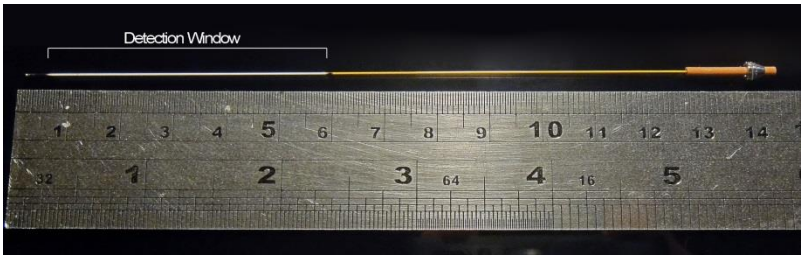


Figure 5-3: Packed capillary co1umn.
This column is a 145 × 0.318 mm 5μm Nucleosil® C18. PEEK sleeve and ferrule allow connection with standard HPLC fittings. The polyimide coating over the final 60 mm was ablated, prior to packing, to allow for optical on-column detection.

Polyimide coatings were removed by careful 'slicing' of the polyimide with a sharp craft knife or razor. This procedure was performed after packing of the column with stationary phase particles and, compared to removal by ablation, resulted in noticeably more robust windowed sections. Frits were made only in the outlet ends of columns by tamping silica particles (from Silica Sep-Pak Cartridges, Waters Corp., Milford, Massachusetts) into the terminal 3-5 mm of the capillary and sintering with a portable blowtorch.

Column connectors were standard stainless steel HPLC fittings, fitted over sleeves of 508 μm ID PEEK tubing (Figure 5-4). Whilst the capillary was initially placed flush to the end surface of the PEEK sleeve, compression of the ferrule to the extent required for a leak-free fit, often resulted in some expansion of the PEEK tubing slightly beyond the end of the capillary (Figure 5-4b). The dead volume thus created was estimated, by microscopy, to be ~ 20 nL, which was typically ~ 13% of the injection volume (most commonly, 150 nL).

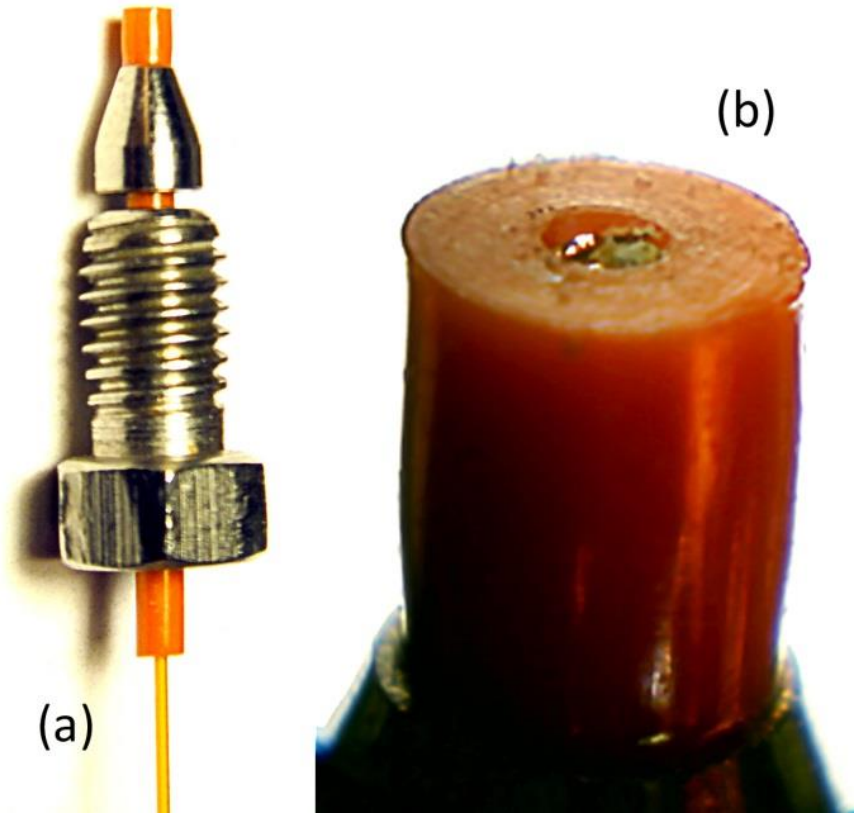


Figure 5-4: Capillary column connections.
Using standard HPLC fittings and PEEK tubing sleeves. (a) A 433 μm OD packed capillary (318 μm ID) is shown fitted to a 508 μm ID PEEK sleeve (b), with a dead volume of ~20 nL. Images are shown at magnifications of ~4× (a) and ~40× (b).

5.2.5.2. Column packing

Capillary columns were packed using an in-house developed upward slurry packing procedure. A thin slurry of particles in acetone was held within a reservoir (~1.6 mL) consisting of a standard stainless steel HPLC column blank. The capillary was located above the slurry reservoir. A syringe pump was used to deliver particles from the reservoir into the capillary. Magnified views of a typical packed capillary are shown in Figure 5-5.

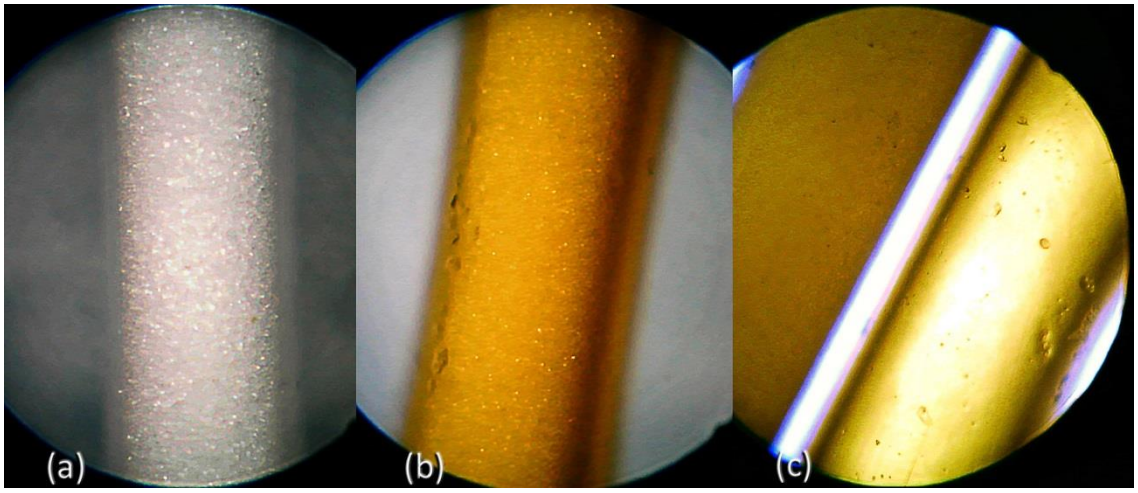


Figure 5-5: 318 μm ID capillary column packed with 5μm C18 particles.
(a) Windowed section and (b) polyimide-coated section with dry packing material and viewed by reflected light. (c) Light transmission of wetted packing material (left) is compared to that of an open-tube capillary (right).

Even though the columns were packed at relatively low pressure, no obvious voids or irregularities were observed over the length of the column when examined under an optical microscope (Figure 5-5 (a) and (b)). When the stationary phase was wetted with acetonitrile, the columns were translucent to a significant extent (Figure 5-5c).

5.2.6. Mounting of D_2 columns within the COMForTS detector

The CCD chip was mounted through a rectangular aperture in an aluminium plate to which the CCD driver circuit was attached (see Figure C-2). The front surface of the CCD was slightly recessed into the plate such that a small groove in the plate, above and below the CCD would locate a capillary column longitudinally over the sensor area as shown in Figure 5-6.

Because the width of the sensor area of the CCD (0.896 mm) was greater than the outer diameter of the capillary (0.433 mm), the capillary was fixed tightly within a masking slit of layered aluminium foil and polyethylene that was taped into position, firmly abutting the capillary. The majority of the light striking the CCD passed through the capillary. However, microscopic examination revealed irregularities in the slit edges of the order of 10 to 20 μm. These irregularities allowed what is effectively 'stray light' from the LEDs to impinge upon the sensor, slightly reducing the detector's relative response to absorbed light.

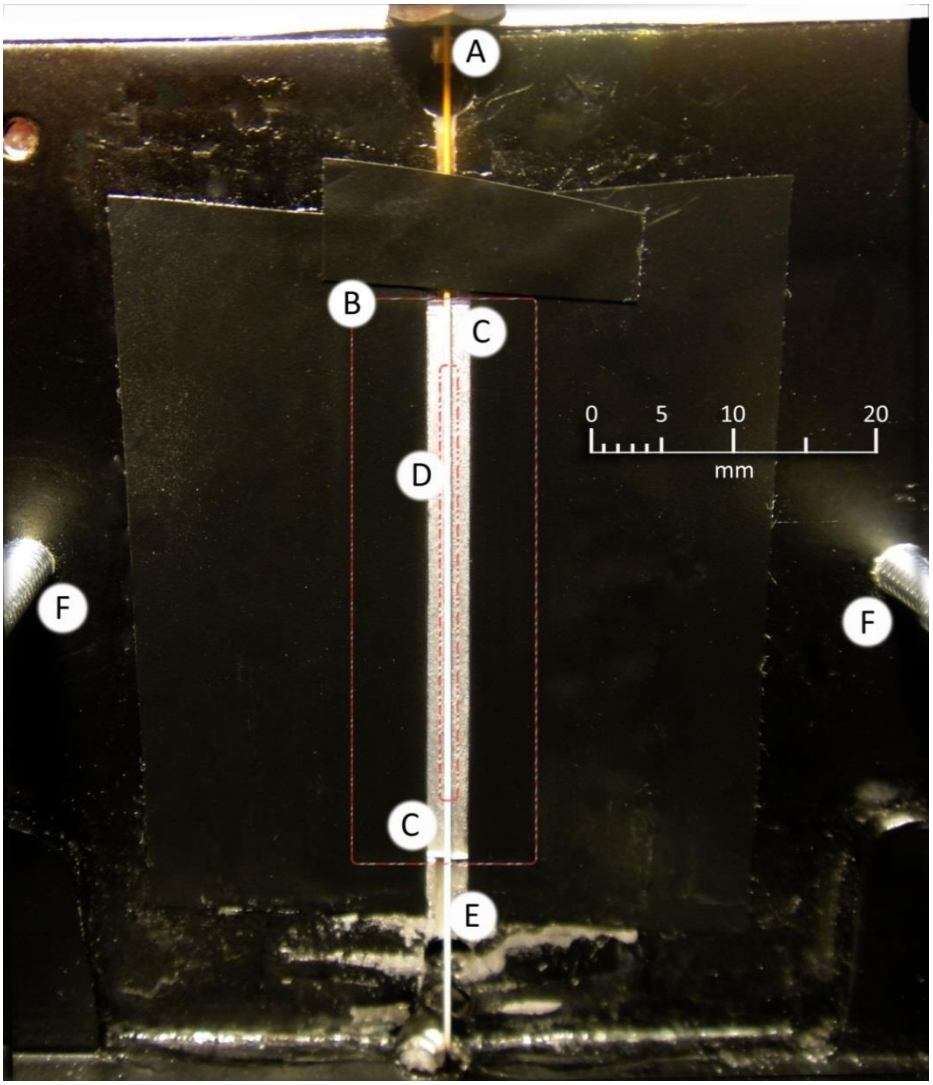


Figure 5-6: Packed capillary column mounted over CCD.
The D_2 packed capillary column is shown mounted in a masking slit longitudinally over the CCD sensor.

Key to Figure 5-6:

A. Second dimension packed capillary column and 3/16" stainless steel nut (coated section). Mobile phase flow is from top to bottom.

B. Location of the CCD chip (marked in red) mounted through an aperture in the mounting plate.

C. Left and right optically opaque masks (layered polyethylene and aluminium) mounted over the CCD sensor area (D) to form a slit between 320 and 400 µm in width.

D. CCD sensor area. The sensor area (indicated with the smaller red rectangle) is 0.896 mm wide and 28.672 mm high.

E. The windowed section of the second dimension capillary column extends over and beyond the CCD sensor area.

F. Mounting springs for the optical bench.

Environmental stray light was reduced as much as possible by the use of matt black coatings on fixed surfaces, light-absorbing 'flock' between removable parts and black PVC electrical insulating tape over joints.

5.3. Software Specification

The COMForTS method is both a new technique and reasonably computationally intensive. As such, there existed no single software package that could control a COMForTS instrument or process the data acquired from that instrument.

COMForTS is largely a data processing method. It does require multi-point on-column detection and it places some general requirements on separation characteristics, but it essentially revolves around the unique treatment of multi-point detection signals. Therefore, this project relies upon the software in order to execute and evaluate the experiments. If valid experiments are to be performed, there must be some means of ensuring the validity of the results and the repeatability of the experiments - and this should be done within a widely accepted framework.

The COMForTS instrument control and data acquisition (CICADA) software was written by the author (see Chapter 6) to meet the specifications described below. Both executable and source code for CICADA are included within the COMForTS 4.0 Chromatography Suite in the Digital Thesis Materials. Although the executable will not function (meaningfully) without a physical connection to the instrument, the functionality of the code can be traced through its source using Microsoft Visual Studio 2010 (or Visual Studio Express 2012, which accompanies the Digital Thesis Materials).

5.3.1. Instrument control

Ideally, one software application should set and record all operating conditions of the instrument. Whilst this was possible for the CCD driver board (via a USB 2.0 interface and Windows driver software supplied by Hamamatsu Photonics, Japan) and the six-port switching valve (via the valve controller's RS232 interface), the syringe pump was not equipped with a digital control interface.

The Microsoft .NET System.TimeSpan constructor was used to control and measure data sampling rates and injection timings with sub-millisecond accuracy. System.TimeSpan measures elapsed time in "system ticks" where

one system tick is 100 ns. The CICADA software was also equipped with a sampling rate self-tuning function: If the recorded times of data samples did not match the required sampling interval, the interval between sampling events was adjusted until the sampling rate was held constant at the nominal rate. Such measures are necessary in a multitasking operating system where other drains on CPU time and communications overheads can cause variability.

Whilst there exist methods to handle irregular sampling for the Fourier transform [119], it is difficult and adds considerably to computing overheads. It is therefore not generally desirable to allow irregularly sampled data.

5.3.2. Data collection

It is also necessary that the instrument control software record raw data together with the parameters used to acquire that data.

In CICADA, the pump flow rate must be entered manually, as well as the identification of the LEDs used as light sources in the detector. All other parameters can be set by the software. All instrument configuration data are recorded within a single data structure, along with the date and time of data collection. File names are automatically incremented to avoid over-writing existing data.

Unsigned 14-bit integer CCD frame data were recorded via the Hamamatsu driver interface in a one-dimensional array of 16-bit unsigned integers. These data were collected with the configuration data into the COMForTS "RawData" structure, which is saved to disk in a ('standard') serialized XML format.

Data files produced by CICADA (that contain raw COMForTS instrument data and instrument configuration and acquisition parameters) have the file name extension ".COMForTS _Raw".

5.3.3. Data processing

CICADA was programmed with the ability to collect, display and save raw data. It performs no modification or processing of the data whatsoever.

5.3.4. Commonality and adaptability

As much as possible, the software is 'modularised': Small independent, easily tested collections of code are treated as "objects". The "detector" object for example refers to the Hamamatsu C11287 driver interface. If a new type of CCD or interface is installed in the instrument, only the *contents* of the "detector" module need to be changed. The new detector module would have the same *external* appearance and functionality (such as having a "gain" control and the ability to capture a "frame" of data) and would therefore require no changes to any other part of the code.

This "object oriented" approach is an accepted programming practice that not only makes code more manageable and robust but also greatly simplifies the testing of code and the location and identification of errors – or 'bugs' – when they do occur.

Whilst COMForTS versions 1.0 to 3.0 were written in Microsoft Visual Basic .NET, it was decided to write COMForTS 4.0 in C#. The reasons for this lie in commonality and adaptability where the adaptability arises from the object-orientation of C# and the commonality arises from its similarity to another programming language: C++. Because C# is grammatically similar to Visual Basic, it was less difficult for the author to learn C# to meet the present purpose than to learn C++.

Amongst professional software developers, C++ is an almost universal language and C#, as a member of the 'C' family of languages, is far more easily understood by a larger number of programmers than is Visual Basic. Additionally, C# enjoys execution speeds of compiled-code that typically equal and often exceed the execution speeds of machine code compiled from C++. With a much wider user-base than Visual Basic, code written in C# (without Microsoft's .NET extensions) can be directly compiled for a large number of hardware and software platforms including those using Linux, UNIX, OS X and Android as well as Windows operating systems. Translation of C# to C++ for extended application in (typical) commercial environments is also relatively easy.

5.3.5. User interface

The CICADA user interface (Figure 5-7) presents the user with the means of setting instrument parameters (but only recording the pump flow rate, column type and LED type) and viewing the plotted raw detector signal.

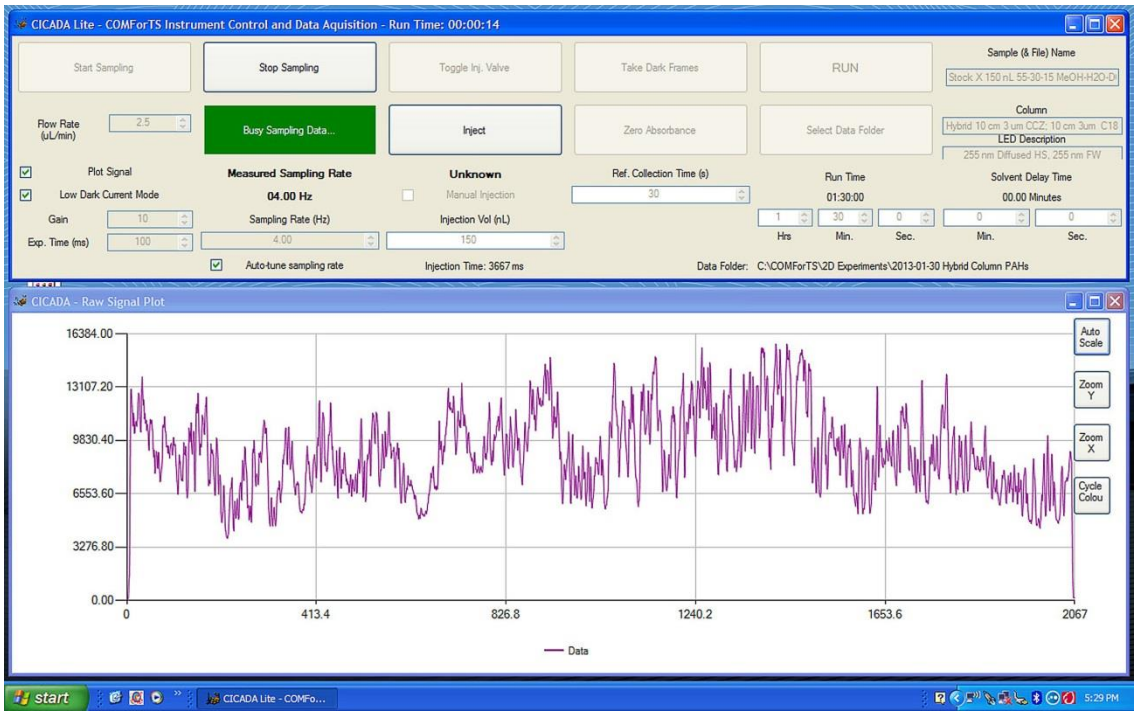


Figure 5-7: COMForTS Instrument Control and Data Acquisition (CICADA). CICADA user interface and typical CCD frame data. The CCD has 2048 active pixel lines out of 2068 total pixel lines. Solvent flow across the CCD is from right to left.

Being a single-beam instrument designed for absorbance detection, two sets of reference data are required: A 'dark' frame F_D and a 'zero-absorbance' frame F_Z, where a 'frame' is one 'exposure' across the entire CCD. Even though CICADA does not perform this calculation, these reference frames are required to calculate the absorbance A_S of a signal frame F_S (i.e. simultaneously across all detectors):

$$A_S = log_{10}\left(\frac{F_Z - F_D}{F_S - F_D}\right) \qquad \text{Eq. (5-1)}$$

To minimise the effects of noise, the reference frames are recorded over a user-definable period. The averaged values of the reference frames over that period were then used in Eq. (5-1).

The transmittance signal shown in Figure 5-7 shows the typical variation in pixel-to-pixel light transmission levels. When collecting data, analyte peaks appeared as localised drops in transmittance that usually extended about half the length of the CCD and travelled along the CCD detector array from right to left.

5.4. COMForTS Instrument Qualification

5.4.1. Detector response

When saturated, the Hamamatsu CCD exhibited erratic responses that often involved an increased output level when light levels were actually decreased. Consequently, once installed in the instrument and illuminated with two UV LEDs, it was first confirmed that the LEDs did not saturate the detector and that detector response was stable and directly proportional to light level (and never inversely proportional).

The two LEDs were mounted in the optical bench in such a way that the overall illumination across the CCD was as even as possible. With no column present and beginning with no power to the LEDs, the displayed detector response was expectedly low. Stepwise increases in the power of the upper LED, produced similar stepwise increments in the corresponding portion of the displayed CCD frame data (right hand side in Figure 5-7). The same consistent behaviour was noted for both LEDs and for stepwise decrements. The maximum current applied to the LEDs was less than 20% of their rated maximum.

This simple test confirmed that the CCD was not saturated and that it responded as expected to changes in light and that the raw data recorded by CICADA correctly matched 14-bit unsigned integer frame data. The very low-level signal that was recorded when the LEDs were not illuminated, did not vary with environmental light levels and indicated that virtually no environmental stray light was present.

5.4.2. Detector sampling rate

Accurate sampling rates could be set at 4, 8, 12 and 16 Hz. At other frequencies, the instrument achieved sampling rates accurate to within 2.5% of the nominal rate. In all cases, sampling intervals were found to be precise to within 0.01%, though an average error rate of 1 / 2000 data points was observed in which two to three data points were 'skipped'. This was not deemed a consequential error and subsequently all raw signals were assumed to have been sampled regularly. The actual time at which data frames were sampled are nonetheless recorded together with the signal at that time.

5.4.3. Flow rate precision and accuracy

In early testing, a very clear and consistent linear relationship was found to exist between the velocity of the mobile phase and the measured frequency of acetone injected at varying flow rates. However, the syringe pump was not able pump more than a nominal 4 μL/min through the two packed columns. It was for this reason that the experiments of Chapters 3 and 4 used only one packed column in combination with one open tubular capillary.

5.4.3.1. Flow rate calibration

Under conditions of low load (pumping through a 30 cm length of 0.5 mm ID PEEK tubing), the pump was set to deliver 5 μL/min. The actual flow rate (with the given syringe internal diameter) was determined by measuring the time required to pump 100 mg of water. A corrected syringe diameter was calculated and entered into the pump's configuration such that a nominal flow rate of 5.0 μL/min produced a measured flow rate of 5.0 μL/min. After calibration, there was less than 0.5% variation observed in unloaded flow rates measured over 20 min. All experiments were performed at the calibrated flow rate of 5 μL/min.

5.4.4. Injection volume calibration

In a six-port two-position ('A' and 'B') valve design, there is a finite time in which no solvent flows from the switching valve during the transition from position 'A' to position 'B', or vice-versa. Because the instrument's injection volume depends on an accurate flow rate and accurate timing of the flow time

through the sample loop, it was necessary to determine this solvent-flow 'dead' time.

An injection cycle requires two switching operations: From position A to B and then back to A. To calibrate the injection timing, switching commands were sent to the injector at varied and accurately controlled intervals and the resulting peak heights were measured. The height of a peak should reduce linearly with injection cycle time and reach a value of zero only when the cycle time is zero. From the results shown in Figure 5-8 the 'dead' time of one injection cycle was determined to be 66.6 ms. All injection times were hereafter calculated according to the nominal flow rate and incremented by 66.6 ms to compensate for the 'dead' time during the switching operation.

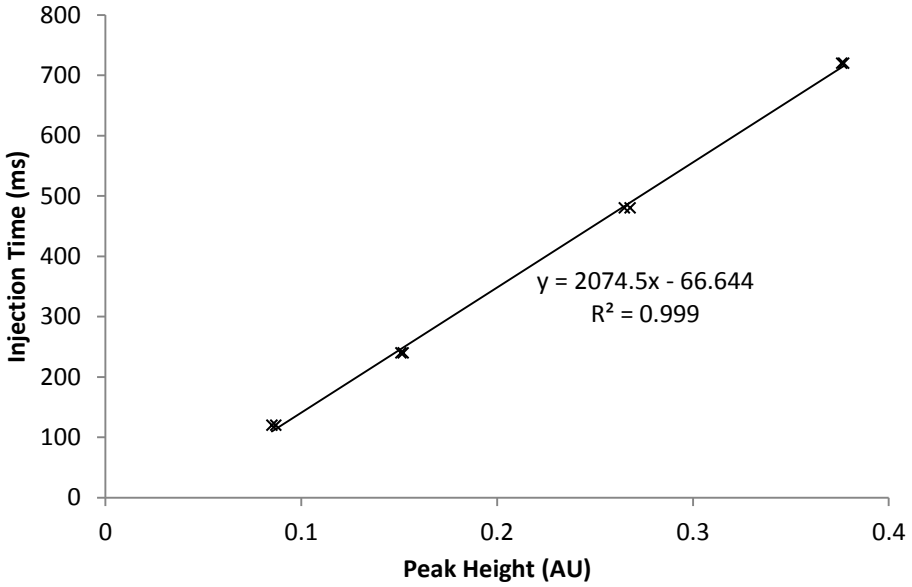


Figure 5-8: Calibration of injection timing.
The non-zero time intercept (-66.6 ms) indicates that during one injection cycle, there is no flow for 66.6 ms.

5.4.5. Chromatographic performance

The following data (Table 5-1) were recorded for an unretained analyte (10 repetitions) using a 210 × 0.250 mm DB-1 capillary connected serially to a 170 × 0.318 mm, 5μm d_p C18 capillary. Absorbance signals were measured at λ = 255 nm.

Peak retention time precision	0.2 % RSD
Peak height precision	3 % RSD
Peak Height	235 mAU
Peak tailing (symmetry) factor	1.2 ± 0.02
RMS Noise	3.2 mAU
Peak to Peak Noise	10.3 mAU
S/N (RMS)	74
Plate count, N (Tangential)	370
Plate Height, H (l = 380 mm)	1.0 mm

Table 5-1: COMForTS Instrument performance data.

The D_1 column was connected directly to an outlet port of the six-port valve. Significant band broadening beyond the expected band width was observed immediately following sample injection. It was surmised that this was due largely, if not entirely, to the (~ 20 nL) voids generated by the PEEK sleeves (Figure 5-4b). A further contribution would be expected from the two such voids associated with the capillary connections within the inter-dimensional zero dead volume connector. The resulting plate height H = 1 mm for the system was consequently close to two orders of magnitude greater than would be normally expected but was nonetheless sufficient for the full resolution of simple mixtures.

Most importantly, in terms of the primary concern of this project (i.e. the qualitative separation of overlapped peaks), the reproducibility of retention times was excellent and the overall peak shape (symmetry) was good. Given the low ratio of signal to noise, peak height reproducibility was also acceptable.

In the separation of five PAHs (described in Chapters 4 and 5), the variation of retention factors k between 0.56 and 2.51 did not exceed 0.76 % RSD and the variation in peak height did not exceed 4.3 % RSD for S/N values greater than 50. Such levels of variability are not unknown in standard HPLC systems working at similar noise levels [120]. Overall, the performance of the system was adequate for the present purposes whilst also presenting significant signals processing challenges that allowed the COMForTS method to be tested under conditions similar to those that would be expected of complex analyses.

5.5. Alternative instrument formats

There are many separation formats that would have equally met the separation and detection requirements that were discussed in Section 5.1 and shown in Figure 5-1. Such formats include gel electrophoresis, capillary electrophoresis (CE), HPLC, ultra-high performance liquid chromatography (UHPLC), supercritical fluid chromatography (SFC), capillary and microfluidic LC and gas chromatography (GC) and μGC. Indeed, hyphenated combinations of these methods could be employed.

In each of these separation formats, in-separation detection (or on-column detection, where the separation uses a 'column') has been demonstrated. On-column conductivity detection has been demonstrated in this project and has been well used elsewhere, while contactless conductivity detection is also relatively common [121-123]. Other demonstrated in-separation detection methods include electrochemical detection [124,125] (which has also been demonstrated, in principle, in parallel arrays [126]), fluorescence [52,127,128] and single-point and multipoint laser-induced fluorescence [52,129-131], optical (UV/Vis) absorbance [132-134] and Fabry-Pérot sensors [135,136] in GC [137] and μGC [136,138]. Therefore, it would be possible, given the appropriate resources, to build a COMForTS instrument in any of the above formats. Ultimately, this freedom in the choice of format could be exploited in a broad range of applications.

5.6. Summary

As the result of many design compromises and assembled at a cost of less than AU$5000, the prototypical COMForTS instrument demonstrated adequate performance that permitted the assessment of COMForTS theory under challenging conditions.

A particular benefit of the design was that, whilst equipped with an array of 2048 independent on-column detectors, conventional chromatograms were

easily obtained by the simple expedient of employing only one of those detectors. The unimpressive level of instrument performance was largely negated, in the experiments described in Chapters 3 and 4, by not relying on absolute determinations but rather by comparing the relative performance of the COMForTS method to simultaneously acquired, conventionally measured results.

Injection volumes and flow rates were confirmed to be accurate and reproducible within the range used in the experiments and sufficient for the present purposes. The noise level was high, largely due to the low current used to drive the UV-LEDs. Hemispherically lensed LEDs (UVTOPnnnTO39**HS**) were found, naturally to have much greater flux density, with their output focussed over a smaller area. Multiple LEDs (one with a 'masked' hemispherical lens and one with a flat window) were required to illuminate the CCD evenly. The use of two hemispherically lensed LEDs, would have greatly increased S/N due to their higher output and reduced scattering.

Improved high S/N, multi-wavelength, multipoint detection would be desirable and is certainly possible (whilst also affording some extra degree of dimensionality [33,39]) but at the time of design, the wholly theoretical understanding of COMForTS hardly justified such expense.

Peak heights showed slightly more variation than would be desirable for direct quantitative determinations. Nevertheless, the variation remained within a range that allowed meaningful comparative evaluations of simultaneously conducted conventional and COMForTS separations and their relative quantitative capabilities.

At this very early stage in COMForTS research, the constructed instrument served its designated purpose of the demonstration of a proof of principle, while simultaneously providing an opportunity to assess the difficulties that the method will face in real-world applications.

Chapter 6.
Software for COMForTS Chromatographic Analysis

"Good friends who come to read this book,
Strip yourselves first of affectation;
Do not assume a pained, shocked look,
For it contains no foul infection,
Yet teaches you no great perfection,
But lessons in the mirthful art,
The only subject of my heart.
While grief would consume and rot,
Mirth's my theme and tears are not,
For laughter is man's proper lot."

François Rabelais in the prologue to "Gargantua and Pantagruel"

6.1. Introduction

The COMForTS Chromatography Suite consists of more than 2,000 individual files based on 46,000 lines of Visual Basic and C# code – or 2,200 printed single-spaced pages. Whilst the software represents a substantial body of work, reading all of it is neither necessary nor of any real benefit. Significant portions are dedicated to enabling control over the flow of the operations and providing a means of inspecting intermediate and final results. Nonetheless, these portions are crucial to the correct application and operation of the code that is directly related to the analysis of the signals produced by COMForTS instruments. Furthermore, it is much easier to gain an appreciation of the function of the software by *using* the software to perform analyses rather than

by reading a description of the process. To that end, the COMForTS software, as well as samples of raw instrument data, are provided within the Digital Thesis Materials.

A guide to the software installation procedure will be found in the Digital Thesis Materials in the file named "[COMForTS]:\Docs\Installation Guide Files\COMForTS Package Installation Guide.htm". A guide to the full COMForTS documentation will be found in "[COMForTS]:\Docs\Index.htm".

The purpose of this chapter is therefore not to inflict the grief that is the detail, but to provide an overview of the COMForTS software, which is integral to the research project. It is expected that the chapter would be read in conjunction with the installed software and source code, although the following is primarily concerned with the reasons behind the software design and function and the evidence that it performed correctly.

6.2. Software Design

As briefly discussed with respect to the COMForTS Instrument Control and Data Acquisition (CICADA) module, the COMForTS 4.0 Chromatography suite was written in C# for two main reasons:

1) It is a widely understood language, easily translated into C++ for commercial and cross platform migration (i.e. to other operating systems apart from Windows.)
2) C# is a powerful, object oriented programming (OOP) language. This allows (reasonably) rapid development of robust, re-useable code that also executes quickly (see [139])

The benefits of using C# in scientific research include that the digital methods can be made transparent and open to scrutiny. In addition, the robustness of a dependency chain of modules furnishes a degree of confidence in the results in that "all components work as intended, therefore the overall result is as intended".

6.2.1. Software architecture for experiments

OOP is particularly appropriate for experimental software, as is the case here. In COMForTS, for example, there is the need to characterise the straight lines in arrayed chromatograms such as those in Figure 4-2. Two methods were used to do this: The Radon transform and a genetic line search algorithm. These methods were implemented in separate modules (RADON and LINEAR2 in the COMForTS 4.0 Suite) that can be used together or independently – or even replaced with another module that performs the task even better. A modular, object-oriented software architecture is shown in Figure 6-1. In COMForTS 4.0, the "System XYZ Core" is the "COMForTS Class Library". In the COMForTS processing chain, CICADA would be "Module 1", CCC (the chromatogram compiler), "Module 2" and "Module 3" may be RADON, LINEAR2, MCSCSD, or COMForTS_FFT (as is shown in the main COMForTS 4.0 interface).

A new module (or routine being tested in a temporary module) that fails miserably or is otherwise faulty, remains independent of the rest of the code and can be easily excised without detriment to other modules.

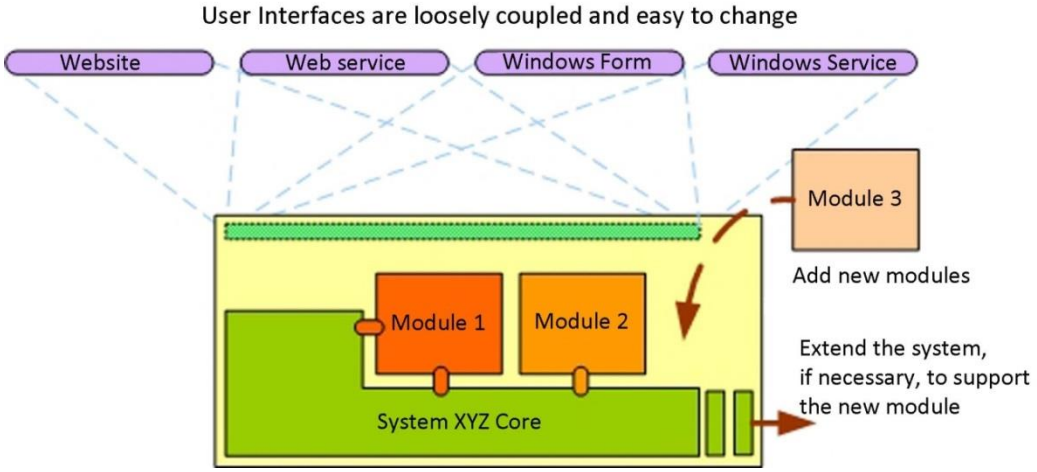


Figure 6-1: Object oriented software architecture (adapted from [139]).

LINEAR is an example of such a module that was replaced in this project with LINEAR2 – and both LINEAR2 and RADON are modules that can occupy the same 'slot' within a processing sequence.

COMForTS 4.0 obviously includes user interfaces but it also interfaces with Windows Services in order to, for example, communicate with the instrument

via USB and RS232 hardware interfaces, identify the number of CPUs and amount of RAM available for program execution across multiple threads, and to use graphics processing units (GPUs) to accelerate the rendering of 3D images. Whilst no web interfaces are used by COMForTS, such interfaces are easily "plugged-in" if one should wish to perform remote or distributed processing or upload and compare results to a database and so forth.

6.2.2. Navigating the COMForTS software

The COMForTS 4.0 (C#) and COMForTS 3.0 (Visual Basic) source code are provided in the Digital Thesis Materials in Microsoft Visual Studio 2010 format. When installed, the source code is located in the 'documents' folder "[Documents]\COMForTS\Source Code\".

The source itself is in plain text. Visual Basic code is found in files with the extension ".vb", and C# code in files with the extension ".cs". Both can be opened, read and edited in any text editor (which may require changing the filename extension to ".txt"). In practice, though, the plain text alone is virtually impossible to navigate and the structure of the software is completely obscured. It is preferable to view the source within the Visual Studio Integrated Development Environment (IDE). The freely distributable Microsoft Visual Studio Express 2012 is provided on the accompanying digital media.

In Visual Studio, the highest level of organisation is a "Solution" (".sln" files). A Solution contains at least one "Project" (".vbproj" or ".csproj" files) – and may contain many and these Projects may be written in different languages. The COMForTS 4.0 Solution contains 15 Projects, most of which are written in C#, but some are in Visual Basic. Visual Studio provides an accessible means of navigating the Solution structure: Its text editor 'understands' the programming language and code structure, using different colours to designate language key words, class names and in-line documentation ('comments' within the code, that have no effect on the program). This facility allows the user to quickly trace the program logic by 'jumping' to object definitions (whether contained in the same source code text file or not), following the sequence the code's execution.

The COMForTS 4.0 Visual Studio Solution ("COMForTS4.sln") contains all the modules described in this chapter. With the Digital Thesis Materials installed, the reader will find this file in the folder: "[Documents]\COMForTS\Source Code\COMForTS 4.0 Chromatography Suite\COMForTS4".

6.2.2.1. A brief tour of Microsoft Visual Studio

The reader familiar with Visual Studio and C# may freely ignore this section. It is intended only as a very brief primer to allow the unfamiliar reader the facility to navigate and understand the basic function of the COMForTS software.

In Visual Studio, the reader can open the COMForTS 4.0 solution by clicking File → Open → Project/Solution → [and selecting] "COMForTS4.sln" in the file browser dialogue.

The central area of the IDE will contain several 'tabbed' pages labelled with the names of the source code files (such as "COMForTSFlowForm.cs"), displayed in design mode (indicated by "[Design]" in the tab's title), which displays the *user interface* generated by the code. In the upper right part of the screen, the "Solution Explorer" window should be displayed (if not, click "View → Solution Explorer"). The Solution Explorer displays a structured list of all the Projects within the COMForTS4 solution and we will continue to refer to these projects (such as "CCC" and "RADON") as "modules".

The easiest way to understand how the software works is to follow the *events* that occur when the user performs an action in the user interface. For example, if we go to the "CCC.cs[Design]" tab, the user interface for the COMForTS Chromatogram Compiler is displayed. This is the "design view". To view the code that executes when the user clicks on the button labelled "Create Dark Frame", double-click on that button in the design view. The text editor will then open the "code view" at the appropriate section, which is an "*event handler*" that is invoked in the *event* that the user clicks on the button:

```
private void DarkFrame_button_Click(object sender, EventArgs e)
{
    DarkFileInfo = FileFunc.GetFileInfo(Constants.File.Type.DarkFrame);

    if (DarkFileInfo.Info != null)
    {
        // Read the Dark Frame data from the instrument's Raw Data:
        DarkFrameData = FileFunc.ReadRawDataFile(DarkFileInfo);
        // Convert the Uint16[] CCD Frames to frames (of type double)
```

```
            // of binned active pixels only:
            DarkFrames = ExtractActiveDetectors(DarkFrameData.Signal);

            // Get the dark frame that is the AVERAGE of all these dark frames:
            DarkFrame = AverageFrame(DarkFrames, DetectorCount);
            // now we are ready to subtract this dark frame from all raw data

            // Tell the user that it's done:
            Status_label.Text = "Dark Frame Calculated; Ready for Zero Absorbance";
            PlotReferenceFrame("Dark");
            ZeroAbsorbance_button.Enabled = true;
            DarkFrame_button.Enabled = false;
            SetPixelLineBinningControlsState(false); // Don't allow further changes
                                                     // to pixel-binning settings
            // as all files must be processed with the same settings.
        }
    }
```

The text within the code is coloured by the editor with the default settings of:

`ClassObject`, `MethodOrFunctionObject`, `DataObject`, `LanguageKeyword`, `//Non-Code Comments`, `"LiteralText"`

There are no spaces allowed in the names of objects, which renders them easily identifiable as a *single* object. The author has endeavoured to make the names of objects as descriptive as possible, reflecting the purpose, function or identity of the object.

In the IDE, placing the mouse cursor over class, object or method names will reveal information about the object definition. One would see, for example, that "`FileFunc`" is a "class COMForTS.FileFunc" i.e. `FileFunc` is a class-type object that is defined *within* the `COMForTS` class. In the same way, it can be seen that "`DarkFrame`" is a field array `[]` of type `double` that belongs to the `CCC` class object (the main class for this file).

In examining the function of the code, it is evident that `DarkFrame` is assigned a value that is returned by the function `AverageFrame()` that takes two arguments:

```
DarkFrame = AverageFrame(DarkFrames, DetectorCount);
```

The 'mouse over', of `AverageFrame` gives us the *signature* of that function, describing its return data type and the data types of its arguments. If we wish to see the *definition* of that function (i.e. *how* that function calculates the result that is being assigned to the `DarkFrame`), we can right-click on the function name and select "Go to Definition" from the pop-up menu, to find:

```
private double[] AverageFrame(double[] FullFrameVector, int FrameLength)
{
    double[] AvgFrame = new double[FrameLength];
```

```
        double[] SumFrame = new double[FrameLength];

        int FrameCount = FullFrameVector.Length / FrameLength;
        int FullFrameIndex;

        for (int j = 0; j < FrameLength; j++)
        {
            for (int i = 0; i < FullFrameVector.Length; i += FrameLength)
            {
                FullFrameIndex = i + j;
                SumFrame[j] += FullFrameVector[FullFrameIndex];
            }

        }

        for (int i = 0; i < AvgFrame.Length; i++)
        {
            AvgFrame[i] = SumFrame[i] / FrameCount;
        }

        return AvgFrame;
    }
```

The above function will `return` an array (`AvgFrame`) that is the average of a series of equally sized "framed" data. In this case, the `DarkFrames` are a series of CCD exposures (frames) taken over some period of time – and the purpose of this function is to find the one `DarkFrame` that exemplifies the CCD output when there is no light (i.e. dark) and little noise (by averaging over a large number of frames).

Where the function of the code may not be immediately obvious, the author has included documentation in the form of `//comments` that explain the purpose or reasons behind the code or for elucidation of the overall procedure. In most cases, the (button-click) event handler will provide sufficient information about the *process* of the COMForTS software without having to go into programming details or "housekeeping" activities (such as managing the user interface to display "progress" etc.).

Each of the COMForTS modules can be navigated in a similar fashion by following the events that are invoked by 'button clicks' in the user interface.

6.2.3. Modularity and data integrity

6.2.3.1. Divide and conquer

By dividing tasks into smaller independent tasks, it is possible to reduce the complexity of code and thereby better elucidate the overall function of the program while at the time minimising the opportunity for error. The division

of code into smaller functional units is a process known as "factoring", since the result is a 'product' of the code's 'factors'.

As a basic example, let us look at the mildly complicated process of the calculation of absorbance values from arrays of signals that were acquired from arrays of single-beam absorbance detectors. As discussed in Section 5.3.5, CICADA produces a series of CCD image frames: A 'dark' frame F_D, a 'zero-absorbance' F_Z frame and a 'signal' frame F_S. The absorbance A_S of a signal frame F_S (i.e. a set of signals recorded simultaneously across all detectors) may be calculated using Eq. (5-1): $A_S = \log_{10}((F_Z - F_D)/(F_S - F_D))$. Even though Eq. (5-1) is a succinct representation of the process, the verbal description is already becoming complicated – and coded instructions quickly become even more verbose and obscure. At the same time, it is readily apparent that the equation is composed of simpler functions: Subtraction, division etc. Therefore, the code may be simplified by using a small collection of functions, noting that a method is a type of object and functions are methods that also *return* an object of some designated type.

One such function would serve, for example, to subtract all the values in one array from the corresponding values in another:

```
public static double[] Subtract(double[] Array1, double[] Array2,
                                int StartIndex = 0, int StopIndex = 0)
{
    double[] Result = new double[Array1.Length];

    if (StopIndex == 0) { StopIndex = Array1.GetUpperBound(0); }

    if ((Array1.Length == Array2.Length)
            && (StopIndex > StartIndex)
            && (StopIndex <= Array1.GetUpperBound(0)))
    {
        for (int i = StartIndex; i <= StopIndex; i++)
        {
            Result[i] = Array1[i] - Array2[i];
        }
    }

    return Result;
}
```

This function actually belongs to the "ArrayMaths" class object, which is part of the "ArrayFunctions" class, which is part of the COMForTS (core) class. Similarly, there is a function to *divide* all values in one array by all values in another and that function's definition looks much the same except that it is called "Divide", and the key line of code is "`Array1[i] /= Array2[i];`". Another similarly easily tested "Log10" function is provided, that returns the base-10

log of every element in an array. Now, it is an easy matter to show that *these* three functions work – and will always work. If the code to evaluate Eq. (5-1) were written out explicitly, we would require two error-free copies of the subtraction routine above, a division routine and a base-10 log routine. The resulting code is lengthy and disguises its overall purpose – and introduces further opportunities for error.

Instead, if the incidental processes of array subtraction and division are 'factored' into self-contained functions, the absorbance calculation may be written as:

```
double[] AbsorbanceFrame_As = CalculateAbsorbanceFrame(Fz, Fd, Fs);

private static double[] CalculateAbsorbanceFrame(double[] Fz,
                                                 double[] Fd,
                                                 double[] Fs)
{
    double[] Fz_minus_Fd = ArrayMaths.Subtract(ref Fz, Fd);
    // Fz_minus_Fd is now equal to (Fz - Fd)
    double[] Fs_minus_Fd = ArrayMaths.Subtract(ref Fs, Fd);
    // Fs_minus_Fd is now equal to (Fz - Fd)
    double[] As = ArrayMaths.Log10(ArrayMaths.Divide(Fz_minus_Fd, Fs_minus_Fd));

    return As;
}
```

…and an otherwise complicated task becomes not only quite easy, but the correct implementation of Eq. (5-1) is readily discernible. (N.B.: This is not *quite* the way that it was done because this particular process is complicated by the arrangement of many CCD frames within a single array that *also* includes inactive pixel data.) Because we can see that Eq. (5-1) is correctly implemented in the function `CalculateAbsorbanceFrame()`, and we know that its lower-level functions work correctly, it is valid to say that `CalculateAbsorbanceFrame()` must also work correctly. The COMForTS 4.0 code has been extensively 'factored' in this way such that there is a level of confidence that the easily discernible correct function of individual parts equates to the correct function of the whole.

C# is also a "strongly typed" language that brings another noteworthy advantage: C# will not allow the data *types* of the input or output to be anything other than that specified in the method signature. For example, one *cannot* accidentally place the resulting absorbance frame (an array of double-precision floating-point values) into an integer array (`int[]`) only to find at some later time that all the absorbance results are either 0, 1 or 2…

In short, the modular design of the software and the use of C# allowed the development of 'experimental' software with a high degree of adaptability and *confidence* that the software performed as intended. With such an approach, most errors and their causes become glaringly obvious very quickly.

6.2.3.2. Preserving data in the face or error

A prime consideration in the laboratory is the preservation of raw data and its integrity. When working with *experimental software*, there is a need to preserve the intermediate results – together with the details of how those results were obtained.

Within the COMForTS Suite, this data preservation and electronic 'audit trail' are accomplished in that the output of each module *includes* its input data in *addition* to its own results and operating parameters. Hard disk space is cheap. Re-processing 50 chromatograms from the instrument data because of a mistake in the *last* of four modules is a much more daunting, time-consuming task than is simply re-processing the last step.

The COMForTS software must also adequately handle high volumes of data and a complicated history of processing parameters, signal details and measurements. The programmer must also be able to find the appropriate data or parameter when required – and be able to *see* that the correct data is being used. Some level of organisation is required: Data structures.

COMForTS 4.0 data structures are defined within the `DataStruct` class within the `COMForTS` base class: `COMForTS.DataStruct` (and are found in the file "COMForTS_ClassLibrary.cs"). The data structures that are used in the main processing sequence are summarised in Table 6-1.

Data Structure	Data Stored	Included Input Data Structures	File name extension for data saved in this format
RawData	Instrument parameters and raw CCD signal	None	.COMForTS_Raw
DetData	Absorbance chromatograms for each individual detector and their generation parameters	RawData	.COMForTS_Det
RadonResult	The Radon transform and the frequency/time data derived from the transform and their generation parameters	RawData DetData	.COMForTS_Rad
COMForTSChromatogram	Time/time chromatograms derived from Frequency/Time data and their generation parameters	RawData DetData RadonResult	.COMForTS

Table 6-1: Principal COMForTS 4.0 data structures.

Note the accumulation of data throughout the processing sequence. When the COMForTS Chromatogram Viewer is used to examine completed chromatograms (in the COMForTSChromatogram data structure), the entire sequence and intermediate results are available for inspection. This feature makes it particularly easy to identify processing routes or user-defined parameters that either positively or adversely affect the ultimate result. Such examination can identify deficiencies of certain modules in handling their assigned tasks and thereby guide future development of both the software and the data treatment methods themselves.

6.3. Software validation

In commercial or industrial software the validation of the software often consists of little more than statements to the effect that the software was logically constructed from a logical series of unambiguous statements and therefore must produce the correct results [140]. The above discussion amounts to the same claim for the COMForTS 4.0 Chromatography Suite. It would be better to have some sort of external validation whereby a known data set is processed and the software's result compared to the known result.

Part of the problem, of course, in validating the COMForTS software was that the software was intended to be used to show that the COMForTS method works in practice. Because it was not *known*, or could not assumed, that COMForTS works as an analytical technique, it was not possible to generate precisely *known* detector signals against which the software's *true* performance could be tested [140]. Nevertheless, there are similar internal input/output relationships that demonstrate that the software performed the function that was intended.

The design of the proof-of-concept experiments described in Section 3.2.2 and throughout Chapter 4, generated a well-defined expectation of the *nature* of the results. It would seem like a circular argument to claim, "The software worked because we found the designed results" when it has been previously claimed that "the manifestation of designed results show that the technique works". In truth, they are co-dependent. Neither can be true without the other being true and neither can be false without the other also being false. In other words, for the experiments to yield the designed result, both the method *and* the software must function correctly.

Other portions of COMForTS are 'self-checking'. For example, it is well understood that peaks in the Radon transform *should* correspond to coordinates of lines in the image – and these were shown to do so by plotting the calculated lines over the original image (such as in Figure 4-6). Similarly, peaks detected within one- and two-dimensional data are plotted and compared to the parent data. In fact, in most cases, errors in the function of the software would result in immediately apparent nonsense (if the code worked at all).

Other functions within the code are also readily amenable to visual inspection but are not shown to the user in the normal course of operations. The efficacy of peak detection and data smoothing functions and strategies are examples of such and will be discussed later (see Section 6.4.6). To facilitate such visual - or numeric – inspection, tools to extract COMForTS data, as tab-delimited text, were created and are available from the 'Tools' menu of the COMForTS 4.0 main screen. These text-based data may then be inspected for validity and/or correctness using third-party packages such as Microsoft Excel and Wolfram Mathematica. In some instances, modules such as the "FFT_TestShell"

(accessed from the 'Tools' menu under "Explore Fast Fourier Transforms") were created specifically to quickly test the function of a *group* of functions.

6.4. Core Algorithms

6.4.1. The Fourier transform

The Fast Fourier transform described in Eq. (1-15) and its inverse transform are implemented in `COMForTS.Fourier.Transform` (located in the file "Fourier.cs"). Apodization (windowing) functions in the forward and reverse directions are defined in `COMForTS.Fourier.Window`. These groups of functions were tested in both directions against synthesised signals of known frequencies and amplitudes. The test module written for this purpose was described in Section 6.3.

Short time Fourier transforms are achieved by applying the Fast Fourier transform, sampling and windowing functions to 'moving boxcar' segments of a signal.

6.4.2. Partial discrete Radon transform

The basic procedure for performing the discrete Radon transform (DRT) is described in Section 1.2.2. Because this is a slow routine, it allows the user to perform a *partial* transform over only the range of frequencies (line angles) that may be encountered within the data. A multi-threaded implementation of this partial discrete Radon transform is defined in `COMForTS.Radon` (located in the file "Radon.cs").

Each thread is responsible for calculating a single column in the Radon. Each Radon column is a range of angles at a given retention time and each point is the sum of the absorbance values of all points that lie on the line at angle θ drawn through the x-axis (of the arrayed chromatograms) at that retention time.

A simple linear interpolation routine is used to minimise inaccuracies in the DRT brought about by the conversion from polar coordinates to integer-only

Cartesian coordinates. This routine increases the accuracy of the DRT but also increases processing time, and may be 'turned-off' in the user interface by selecting the "Integer-only Radon" option. Because the DRT is cross-correlated with the Linear transform (below), there was no apparent difference in *overall* transform results that relied upon interpolated input or input from integer Cartesian coordinates only.

6.4.3. Linear transform

Being an exploration of a method of identifying straight lines in the arrayed chromatograms, the Linear transform was not factored as a specific method and advanced to the 'status' of a class within the `COMForTS` class.

The Linear transform falls into a class of notoriously slow but effective "genetic" algorithms. The actual code will be found in the `FindPeakLines_Button_Click()` event handler located in the LINEAR_Form.cs file within the LINEAR2 module.

The function of this algorithm (and its "genetic" appellation) may best be described in words:

1) Start with a 'pool of genetic material': PeakPoints, which are the peak maxima in the arrayed one dimensional detector chromatograms.
2) Spawn a population of lines from a few PeakPoints that are nearest to each other.
3) 'Healthy' lines are straight lines and are allowed to grow.
4) 'Mutant' lines are not straight – they are 'unfit' and are not allowed to survive.
5) Some healthy lines may begin to mutate (bend) and are 'killed-off'.
6) Some healthy lines may be following the same evolutionary path from different points (i.e. two segments of the same straight line) – and these lines are allowed to coalesce.
7) Some otherwise healthy lines may not fit into the 'social structure' (they may have un-allowed slopes or intercepts) and are also 'killed-off'.
8) The genetic material (PeakPoints) that was used by dead 'mutant' or 'asocial' lines is returned to the 'genetic pool'.

9) Evolution of straight lines continues by returning to step (1) until there is no genetic material left or no more healthy lines can be spawned.

This approach achieved excellent results, in less time than the Radon transform, when applied to the moderately overlapped separation described in Chapter 4. There were usually only one or two lines (depending on the peak detection parameters used) that did not correlate well with the Radon-derived lines. However, with heavy physical overlap, many 'healthy' lines can degrade (Step 5, above) and the algorithm takes considerable time to evolve a stable population of healthy lines – and many lines remain unidentified, with the un-used genetic material (PeakPoints) being eventually deemed 'unable to support life'.

However, when peak detection in the Radon failed, this algorithm was able to pick-up the few missing healthy lines from the little genetic material left over from the Radon-derived line population. The measures of line fitness (linearity) that were used in this algorithm also allowed an estimate of the error in the determination of the line equations: their 'goodness' of fit to the arrayed detector chromatograms.

The reader can examine the function of this algorithm by watching line 'evolution' in progress when running the LINEAR2 module on a .COMForTS_Det file produced by the COMForTS Chromatogram Compiler. In the LINEAR2 result plot, growing 'adolescent' lines are shown with green peak points and mature, adult lines shown in magenta.

6.4.4. A method for combining signals from closely spaced detectors

When combining the signals from broad peaks, even when alternately inverted to allow small displacements, this method (MCSCSD) operated as intended but did not produce sufficient resolution in the frequency domain – and was not used in this thesis. It is nonetheless included in the COMForTS 4.0 package and discussed in detail in Appendix D. The reasons for its inclusion being that an essentially binary interpretation of lines identified in the Radon/Linear transform would have no width and MCSCSD would once again become relevant.

It was observed that, when summed, alternately inverted chromatographic peaks create a zero-average signal that quite closely resembles a sine wave. The obvious consequence being that the FFT of such a signal produces a very sharp peak at the fundamental frequency and only simple overtones that are easily filtered.

If chromatographic peaks, in the arrayed chromatograms, are replaced with alternately inverted unit-width Gaussian peaks that are constructed around each maximum, the short-time FFTs of the sums of *these* signals could produce very well defined frequency spectra. This has significant potential for further cross-correlation between the Radon and Linear transforms, improving the accuracy and precision of the overall transform result. The FFT is such a fast transform that the added computational demands would not be outrageous.

6.4.5. Frequency/time to time/time conversion

Time-dependent frequency transforms return frequencies present at a given time. The frequency f of peak lines provided by the Radon is given by the line slope, in detectors per second. Because the detectors are evenly spaced at a known distance d, this corresponds directly to a distance per second, or velocity v_{D2} (in the second dimension). We also know that the overall retention time T_t (which is equal to T_1 + T_2), is given by the x-intercept of those lines. The velocity in the second dimension is given by Eq. (6-1):

$$v_{D2} = fd \qquad \text{Eq. (6-1)}$$

and, knowing the physical length L_2 of D_2, the retention time in the second dimension T_2 may be calculated:

$$T_2 = v_{D2}/L_2 \qquad \text{Eq. (6-2)}$$

and:

$$T_1 = T_t - T_2 \qquad \text{Eq. (6-3)}$$

Now that retention *time* data for both dimensions are known, the errors in the estimates of frequency and total retention are similarly translated into errors in the estimates of T_1 and T_2. Three-dimensional Gaussian peaks are subsequently constructed with a height that corresponds to the peak height in the chromatogram but with total widths in the T_1 and T_2 axes that correspond to eight times the standard deviation of the error in each axis (allowing for the possibility of, as yet, undetermined contributions to line errors).

This frequency/time (line equation) to time/time 3D peak conversion is accomplished in the `ConvertLineToChrompoint` method within the `FT_TT_Form` class defined in the file "FT_TT_Form.cs".

6.4.6. Ancillary algorithms

Data smoothing was achieved either by averaging signals (by detector or chromatogram binning) or by a low-pass Fourier filter or by a windowed linear least squares smoothing algorithm. Except for binning, smoothing was limited such that the difference between the raw signal and the smoothed signal was never greater than the measured RMS noise. The smoothing methods were shown to have varying degrees of effectiveness in both terms of the degree of improvement to S/N and their faithfulness to a human interpretation of the raw signal (Figure 6-2 and Figure 6-3).

The linear least squares smoothing was similar to a normal boxcar smoothing, but instead of the simple average, the line of best fit was calculated and the central value of the window determined from the line equation. The window length was adapted to each signal such that the original and smoothed signal differed by no more than the RMS noise. The only disadvantage of this algorithm is its lack of speed, though it proved to follow the original signal accurately, with little apparent loss of detail (Figure 6-3). Similarly, the low-pass Fourier transform filter removed high frequencies from the transform until the recovered signal (the inverse transform) differed from the original signal at any point by no more than the RMS noise. Operation of this filter was very fast but often produced slightly "wavy" baselines that could interfere with the detection of low-intensity peaks.

Peak detection was by second derivative zero crossing, which appeared to be adequate. The implementation, though, produced a small *systematic* error in

the time-location of the maxima. Evidence of this may be seen in Figure 6-2 where peak maxima are marked at the correct height but ~2 s before of the actual maximum.

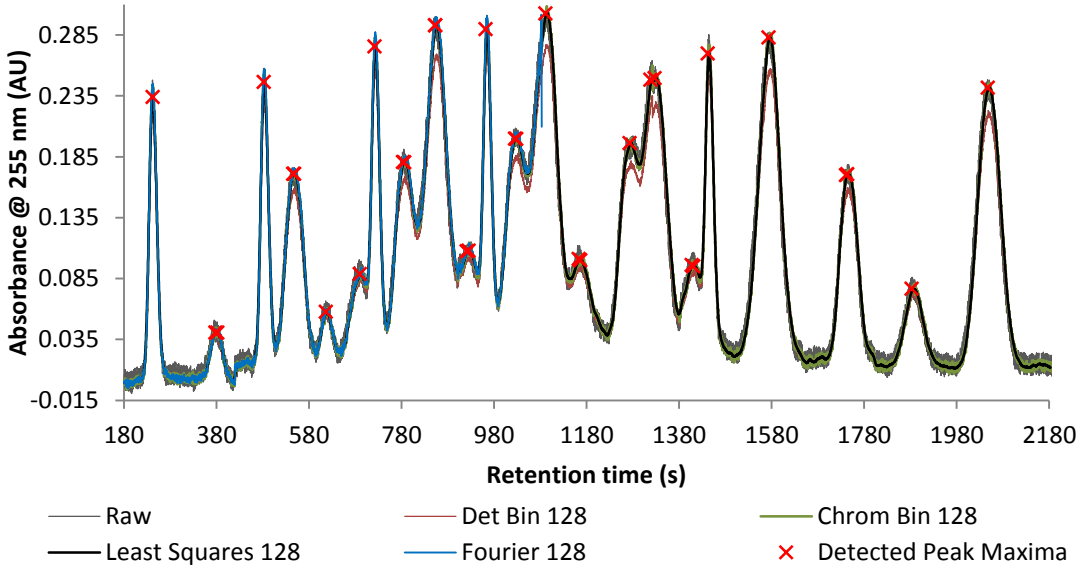


Figure 6-2: Overall effect of different smoothing functions.
Except for the raw signal, other signals were smoothed, binned signals from 128 × 16 virtual detectors. Adaptive linear least squares smoothing produced the best results but was also the slowest algorithm.

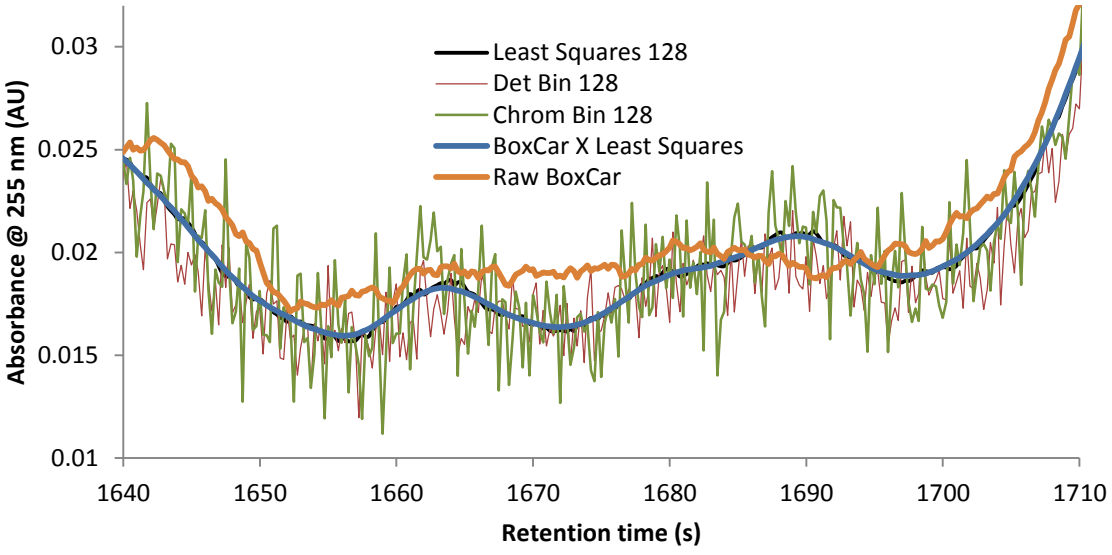


Figure 6-3: Smoothing for fine detail.
Detector binning often resulted in a loss of signal detail in spite of being more effective at reducing noise. Chromatogram binning, followed by adaptive least squares smoothing greatly reduced noise whilst maintaining signal detail. Boxcar smoothing of the least-squares signal (not implemented during the experiments), increased S/N by a factor of 1000 with no apparent loss of detail.

6.5. Conclusion

The COMForTS 4.0 Chromatography Suite was developed as a suite of independent, experimental modules that were each designed to complete discrete stages in the processing of COMForTS signals.

The choice of development language, environment and software architecture helped to ensure the production of reliable software that functioned as intended. Where possible, the overall function of the software was tested against expected results from known inputs. Whilst there are many instances in which the accuracy or precision of the software could be improved, the overall function of the software was found to be correct and some possibilities for enhanced or alternate approaches were identified.

Custom data structures were also employed to ensure that the integrity of raw data and results of experimental intermediate processing steps was maintained.

Future enhancements are easily incorporated into the software architecture by 'swapping-out' individual modules of code or method and function objects as required.

Chapter 7.
COMForTS

7.1. Fundamental Interpretation of the Nature of COMForTS

The fundamental types of separation are described in Section 1.1.1, separations in time S_t and separations in space S_s. We can see that COMForTS, with multiple equally *spaced* detectors, effectively transforms *one* separation dimension in time S_t into a separation in two dimensions:

$$C(S_t) = S_t \times S_s \qquad \text{Eq. (7-1)}$$

where C is the effective function of COMForTS in terms of dimensionality.

This simultaneous inclusion of the space domain allows us to say that if two analytes have different speeds (*s/t*) they cannot be in the same place at the same time unless they began their journey from different places. Similarly, two analytes in different places cannot have the same speed unless they started in different places. This reasoning of (space × time) mutual exclusivity effectively creates another *logical* dimension of separation.

If this separation due to logic "S_L", is included in the COMForTS 'transformation function' C, Eq. (7-1) becomes:

$$C(S_t) = S_t \times S_s \times S_L \qquad \text{Eq. (7-2)}$$

However, the distance and time properties of S_s and S_t allow us, if we know the physical length l_{St} of S_t, to define another separation in time S_{t0} that is the starting state of S_t.

We can see then that COMForTS, with knowledge of the separation distance l_{St}, transforms one separation in time S_t into a logical four-dimensional separation:

$$C(S_t, l_{St}) = S_{t0} \times (S_t \times S_s \times S_L) \qquad \text{Eq. (7-3)}$$

where the conventional 2D peak capacity $n_{c,2D}$ is provided by $S_{t0} \times S_t$ and the effective peak capacity n_{ce} is provided by $S_{t0} \times S_t \times S_s$. It may be then concluded that the difference between the measured effective peak capacity and the observed 'effective analyte capacity', as discussed in Section 4.4, is due to S_L. and that this effective analyte capacity is given by $S_{t1} \times S_{t2} \times S_s \times S_L$.

The logical separation dimension S_L cannot be perfectly orthogonal because logical conditions exist that are not exclusive, e.g. two analytes *can* be at the same place at the same time whilst beginning their journey from the same location. It is only when the logic is exclusive that it contributes to dimensionality.

7.2. Separations in Three or More Dimensions

It should also be evident from Eq. (7-3) that the starting state, S_{t0}, can be the *result* a multidimensional separation itself, and that the dimensionality of COMForTS therefore has no theoretical limit. Nevertheless, there are practical limits to the number of columns that can be serially connected. In addition, there are thermodynamic limitations to the degree of orthogonality that could be achieved by the addition of those columns. There is little doubt, however, that at least three physical separation dimensions would be advantageous as demonstrated in gas chromatography [90,91] and other, mixed mode separations [141].

In applying COMForTS to physical three-dimensional separations in time ($S_{t1} \times S_{t2} \times S_{t3}$), Eq. (7-3) indicates that COMForTS may be used only to link pairs of separations. Therefore, two COMForTS transforms, C_1 and C_2, are required:

$$C_1(S_{t2}, l_{St2}) = S_{t1} \times (S_{t2} \times S_{s1} \times S_L)$$ Eq. (7-4)

and

$$C_2(S_{t3}, l_{St3}) = S_{t2} \times (S_{t3} \times S_{s2} \times S_L)$$ Eq. (7-5)

where S_{t2} is the result of the first *and* second dimension separations. The analogous physical arrangement is shown in Figure 7-1, with arrayed on-column detection (S_{s1} and S_{s2}) applied to both S_{t2} and S_{t3}.

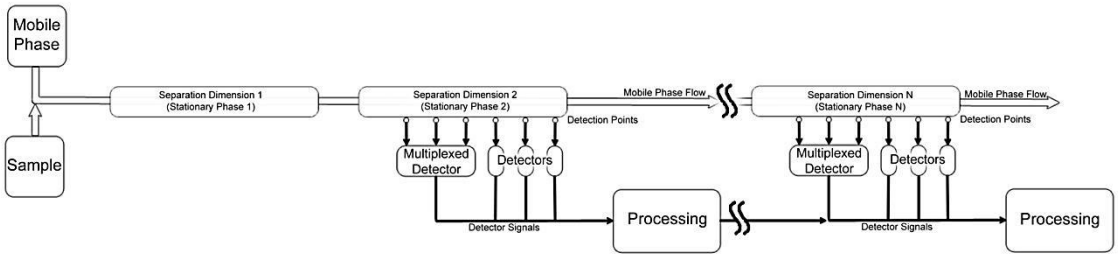


Figure 7-1: COMForTS system schematic with N separation dimensions.

The overall dimensionality (Eq. (7-6)) of the logical combination of C_1 and C_2 depends upon the mutually acceptable intersection of two logical dimensions S_{L1} and S_{L2} and the separation must therefore lose some benefit of this dimensionality.

$$C_1(S_{t2}, l_{St2}) \cup C_2(S_{t3}, l_{St3}) = S_{t1} \times S_{t2} \times S_{t3} \times S_{s1} \times S_{s2} \times (S_{L1} \cap S_{L2})$$ Eq. (7-6)

We can understand the loss caused by an increased logical requirement when we inspect the retention pattern described in Table 7-1:

Analyte	T_1	T_2	T_3
A	1	2	3
B	2	1	2

Table 7-1: Problematic three-dimensional retention patterns.
The retention times T_n are shown for each dimension D_n.

Note that A and B co-elute from D_2 but are resolved by both C_1 and C_2 as indicated in Figure 7-2. However, the COMForTS function C_2 cannot *identify* the peak at T_3 = 3 as belonging specifically to A or specifically to B but only to one *or* the other. In this example, the three-dimensional retention coordinates

of analyte A are clearly (1,2,3) – but COMForTS would not be able to correlate the T_3 and T_1 times if (T_1 + T_2) are equal for two or more analytes. It could only report, for both components, that the retention coordinates are *either* (1,2,3) *or* (2,1,2).

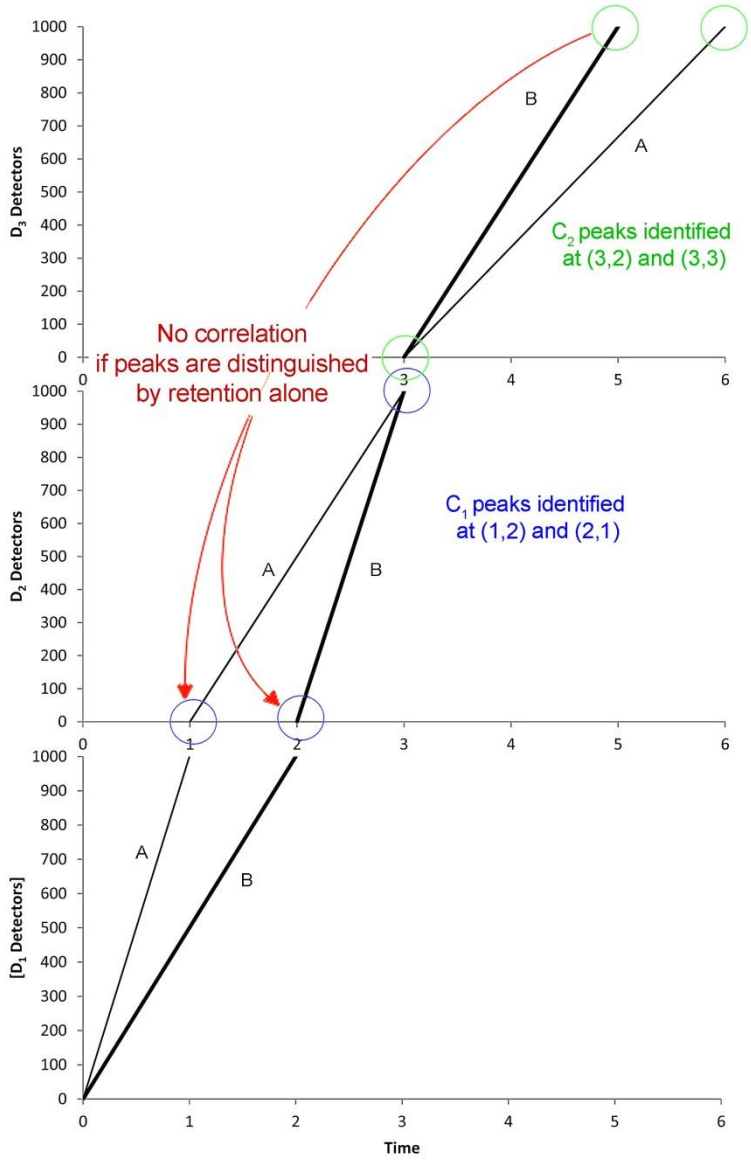


Figure 7-2: Problematic retention patterns as described in Table 7-1.
Shown are multi-detector chromatograms for each of three separation dimensions where each detector array extends the full length of the column. The correlation between D_3 retention times and D_1 retention times is lost if peaks co-elute from D_2. Relative peak heights (indicated by line thickness) or spectra can be used to restore the correlation. Retention coordinates identified in C_1 and C_2 are indicated in blue and green respectively.

Fortunately, there are other differences between the analyte peaks that can be exploited, the most obvious being that they can have different heights. We could therefore say that two *logical* peaks can co-elute only when, for

example, their retention times *and* peak heights are equal. In Figure 7-2, a quick examination of the height ratios of the two peaks in D_1 compared to their ratios in D_3, could conclusively identify the 'owners' of the peaks in D_3. Similarly, the dimensionality of the detector [39] could be augmented, adding the condition that the overall detector responses must be equal (or in some fixed proportion).

Intuitively, the loss in *analyte* capacity that would result from co-eluting peaks that *also* share the same peak height (or peak height proportion) *and* the same UV spectrum (for example) would be relatively small. Nonetheless, this is a condition that requires attention.

7.3. COMForTS Separations Formats

A range of possible separations formats for the COMForTS prototype, and the detection options available were discussed in Section 5.5. This research identified that a very large array of in-separation detectors is highly beneficial. In addition, the accurate physical length of separation dimensions and the distance between detectors must be known if a correlation is to be made between COMForTS separations and the individual physical separations achieved in each of the constituent one-dimensional separations.

The peak capacity, peak production and analyte capacity advantage of COMForTS is not absolute. COMForTS yields some factor of improvement over the capacity of the system to which it is applied – and, to achieve its potential, should be applied to the most efficient separations amenable to the basic conditions that were reiterated in the above discussion.

It is the author's opinion that these requirements would be most easily met in the first instance by chip-based μ-LC combined with CCD-based multi-wavelength UV absorbance detection. Chips containing two, three or even four separation dimensions can be manufactured with great precision and placed in a known position over detectors, again with great precision.

With peak capacities of 120 in 20-minute separations [142], an orthogonal two-dimensional μ-LC COMForTS system could achieve a not unimpressive peak capacity of 120 × 120 × 41 = 590000 over 40 minutes – or 246 peaks per second. The analyte capacity of such a system would be unprecedented.

7.4. Mobile Phase Changes

The introduction of a mobile phase change between the separation dimensions has no effect on the frequencies and retention times measured in the second dimension as long the separation remains isocratic. However, the introduction of a new mobile phase requires that there must be a non-zero volume between the dimensions (as in Figure 5-1). In this situation, the simple relationship between the frequency/time and time/time domains as described in Section 6.4.5 is no longer quite true: The time at which an analyte has entered the second dimension can be calculated from the velocity and overall retention, but if there is some delay between dimensions, the D_2 entry time is not equal to the D_1 retention time.

The remedy for the problem is not quite as simple as calculating the 'flow time' of the inter-dimensional void: It will take some time for analytes leaving D_1 to accelerate to the mobile phase velocity before then decelerating into the second dimension. One solution is to perform an empirical calibration of this void time or even add a detector array to the first dimension and use the accurate velocities measured there to determine the true time of exit from the first dimension and hence deduce the inter-dimensional flow time.

Changing mobile phases between dimensions is certainly possible – and certainly desirable in terms of the likely gains to be made in orthogonality – but it is not without complications if true time/time separations are desired.

7.5. Gradient Separations

Throughout this project, we have dealt only with isocratic separations. Analyte peaks therefore appear as (theoretically) perfectly straight lines in the arrayed chromatograms such as shown in Figure 3-5 and in Figure 4-2. This is because analytes are travelling over evenly spaced detectors at a uniform velocity. If gradient elution is employed, the velocity of analytes will change over time and straight lines will not be visible. Nonetheless, if the gradient is some linear function of time - that need not be a *straight* line function - there will still be some linear relationship between analyte velocity (and distance travelled over the detectors) and time.

In the case of a parabolic solvent gradient, analyte velocities will increase at some constant rate, i.e. acceleration is constant and the velocity v at any given time t is $u + at$, where u is the initial velocity. It is easy to show then that the distance d travelled over the detectors is given by the parabolic relationship:

$$d = ut + \frac{1}{2}at^2 \qquad \text{Eq. (7-7)}$$

To perform a Radon transform to identify such parabolic lines in arrayed chromatograms 'only' requires the calculation of sums along parabolic lines at varying angles *with* varying values of a. Fortunately, a generalised multi-directional discrete Radon transform has been developed that performs just such a task and has been specifically demonstrated to identify parabolas at varying degrees of rotation [57]. Of course, such a Radon transform has a three-(or higher)-dimensional space and makes proportionately greater demands on computing resources.

In short, as long as a solvent gradient produces a *linear* change in analyte velocity, a Radon transform can be constructed to identify such lines within an image. The relationship between time and distance does not need to be *rectilinear*.

7.6. Signals Processing

Although the processing of even single-detector chromatograms is computationally intensive, advances in digital technology have rendered delays almost unnoticeable. When, however, there are 2048 detector signals to be processed, substantially more computing power and attention to software design are needed in order to minimise delay. The COMForTS 4.0 Chromatography Suite was written with that in mind but also under the assumption that memory and hard disk space were unlimited. Nonetheless, execution of two key functions remains noticeably challenged: noise reduction and the time-dependent frequency transform.

There are certainly more appropriate noise-reduction methods than those used here and the COMForTS analysis would undoubtedly benefit from their use, particularly in terms of speed. Felinger and Guiochon [140] showed that even commercial chromatography software, that otherwise displayed great accuracy, was noticeably affected when S/N fell below about 100 with the quality of results decreasing dramatically for S/N <64. Most of the raw detector signals from the COMForTS prototype instrument displayed S/N in this troublesome region between about 68 and 75, and large numbers of detector chromatograms had to be binned in order to detect peaks accurately within that level of noise. Sacrificing detectors entails a loss in precision. One would therefore expect that better and faster hardware and software noise reduction would make a noticeable improvement in precision by removing the need to bin chromatograms.

Implementation of the fast Radon transform and more appropriate peak detection to suit the topography of the transform would be a remarkable boon. Combined with a binary treatment of the arrayed chromatograms (peaks in the arrayed chromatograms are either present or not), the fast Radon becomes a fast Hough transform, which is reported to be the more usual method of finding straight lines in images [58]. This binary treatment permits the application of the MCSCSD approach (Appendix D) to the short time fast Fourier transform – which is orders of magnitude faster than the author's Linear transform. The additional transform data provides the possibility of

notable gains in precision and an improved capability to cross correlate multiple transforms for harmonic and overtone rejection.

Unfortunately, with the addition of 'dimensional' detection (such as DAD UV/Vis), the volume of data and demands on system resources will increase significantly.

7.7. Performance Expectations and Limitations

The results produced by the present software revealed that, even with moderately physically overlapped samples, the calculated results were not an optimal fit to the observed data – and that the quality of that fit decreased with complexity. In both cases, the separation signals themselves followed the expected patterns (Figure 4-3 and Figure 4-11b). One could then justifiably say that the observed improvement in peak production by a factor of 26 (See Section 4.3.4) is the worst that could expected from COMForTS, but that a factor of 41-fold *may* be approaching the best that we could expect. In high S/N systems with 1000+ detectors and much smaller peak widths compared with detector spacing, it would be possible to halve the present errors and double the improvement in peak production to 80-fold. Such an improvement would not be unreasonable in light of the relative performance of the COMForTS prototype (see 5.4.5).

A limiting factor is the ratio of peak width to separation complexity. Separations that are more complex need narrower peaks widths – and this can be achieved through increasing chromatographic efficiency and / or reduction of the data to essentially binary information. Image sharpening algorithms in combination with the Hough transform have significant potential application in this respect alone.

There is also the consideration that highly crowded analytes may result in analyte-analyte interactions on the column, resulting in distorted or abnormally bent trajectories across the detector array. This would also be a matter for further investigation.

7.8. Applications of COMForTS

While noting that COMForTS is not intended to resolve peaks physically, and that it requires arrays of non-destructive, in-separation detectors, it nevertheless increases peak production in the region of 40-fold. At the same time, COMForTS provides an extraordinary increase in confidence in peak purity: the ability to identify individual analytes.

Applications of multidimensional separations to the fields of metabolomics and proteomics are rapidly expanding and COMForTS has conspicuous relevance in these areas. COMForTS' real strength, however, lies in providing complex separation information very quickly and with unparalleled confidence that the appearance of a peak in the chromatogram equates with the presence of a single pure compound.

This strength has particular relevance to 'bio-prospecting' and 'fingerprinting' applications. The identification of chemical differences between natural products known to have different or specific properties has become a valuable approach in the search for natural medicines and commercially valuable compounds such as those with insecticidal or germicidal properties. 'Fingerprinting' of natural products or complex samples of natural origin is a significant and growing area [143-146]. Such a strategy has been used to create "biological passports" that are used to identify differences in the metabolic profiles of élite athletes as a means of detecting the use of illicit or performance enhancing drugs [147-149].

Whilst COMForTS does not yield a physical separation, it does describe the conditions under which a physical separation can be obtained. If a compound of interest were to be found in a three dimensional COMForTS separation, its retention characteristics would be known in all three dimensions. In a *targeted* offline separation, only one 'cut' from each separation dimension would therefore be sufficient to yield a physical separation of a targeted compound – without the need to perform a *comprehensive* offline multidimensional separation. One such targeted offline separation would be achieved in the same total analysis time as the COMForTS separation. The time saved by

removing the need to perform a physical separation until it is *required* would be substantial.

7.9. Conclusion

"Je m'en vais chercher un grand peut-être." - François Rabelais

The present research began with a 'great perhaps', that was developed through the theory, computationally modelled and, in the laboratory, carefully tested against that theory and demonstrated to be not a 'perhaps', but an actuality.

Eschewing physically pulsed signals in favour of multipoint detection yielded two-dimensional peak production rates beyond expectation. Methods for time-dependent frequency transformation of signals from arrayed detectors that were spaced apart at small fractions of the chromatographic peak width were demonstrated for the first time. The COMForTS 4.0 Chromatography Suite was designed as experimental software, sufficient to demonstrate the function of COMForTS. Nevertheless, it produced results with such precision that it yielded time-domain resolution far greater than the physical resolution.

Importantly, the fundamental nature of COMForTS separations was elucidated. The reasons behind the observed 40-fold improvement (over online 2DLC) in peak production and massive increase in "analyte capacity" reside mainly in the use of logic to provide a separation of analyte information. COMForTS achieved in the prototypical instrument, with a physical peak capacity of 16, a separation that would theoretically require a peak capacity greater than 3000: a separation of 25 analytes with 97% confidence in peak purity. It was also shown that this extraordinary power of COMForTS could be multiplied by extension to separations in three or more dimensions.

The instrument and software built in this project were far from perfect but their function – the defeat of peak wrap-around – was demonstrated sufficiently well to justify the direction of further research effort towards Comprehensive Online Multidimensional Frequency Transform Separations.

For now, new applications may be COMForTS' next 'great perhaps'.

References

[1] B.L. Karger, L.R. Snyder, C. Horvath, An Introduction to Separation Science, Wiley & Sons, New York, 1973.

[2] G. Guiochon, M.F. Gonnord, M. Zakaria, L.A. Beaver, A.M. Siouffi, Chromatographia 17 (1983) 121.

[3] G. Guiochon, L.A. Beaver, M.F. Gonnord, A.M. Siouffi, M. Zakaria, J. Chromatogr. 255 (1983) 415.

[4] L.A. Beaver, G.A. Guiochon, in USPTO (Editor), Varex Corp., USA . United States of America, 1984.

[5] J.C. Giddings, Anal. Chem. 56 (1984) 1258A.

[6] S.A. Cohen, M.R. Schure, Multidimensional Liquid Chromatography: Theory and Applications in Industrial Chemistry and the Life Sciences, John Wiley & Sons, Inc., Hoboken, New Jersey, 2008.

[7] R.C. Allen, M.G. John, S.C. Rutan, M.R. Filgueira, P.W. Carr, J. Chromatogr. A 1254 (2012) 51.

[8] R.E. Mohler, K.M. Dombek, J.C. Hoggard, K.M. Pierce, E.T. Young, R.E. Synovec, Analyst 132 (2007) 756.

[9] S.E.G. Porter, D.R. Stoll, S.C. Rutan, P.W. Carr, J.D. Cohen, Anal. Chem. 78 (2006) 5559.

[10] K.M. Pierce, J.C. Hoggard, R.E. Mohler, R.E. Synovec, J. Chromatogr. A 1184 (2008) 341.

[11] D.R. Stoll, X. Wang, P.W. Carr, Anal. Chem. 80 (2008) 268.

[12] G. Guiochon, N. Marchetti, K. Mriziq, R.A. Shalliker, J. Chromatogr. A 1189 (2008) 109.

[13] R.E. Murphy, M.R. Schure, J.P. Foley, Anal. Chem. 70 (1998) 1585.

[14] J.V. Seeley, J. Chromatogr. A 962 (2002) 21.

[15] J.N. Fairchild, K. Horvath, G. Guiochon, J. Chromatogr. A 1216 (2009) 6210.

[16] J. Calvin Giddings, Anal. Chem. 39 (1967) 1027.

[17] E. Grushka, J. Chromatogr. A 316 (1984) 81.

[18] J. Krupčík, J. Garaj, P. Čellár, G. Guiochon, J. Chromatogr. A 312 (1984) 1.

[19] J.M. Davis, D.R. Stoll, P.W. Carr, Anal. Chem. 80 (2008) 8122.

[20] M.R. Filgueira, Y. Huang, K. Witt, C. Castells, P.W. Carr, Anal. Chem. 83 (2011) 9531.

[21] H. Gu, Y. Huang, P.W. Carr, J. Chromatogr. A 1218 (2011) 64.

[22] X. Li, D.R. Stoll, P.W. Carr, Anal. Chem. 81 (2009) 845.

[23] L.W. Potts, D.R. Stoll, X. Li, P.W. Carr, J. Chromatogr. A 1217 (2010) 5700.

[24] J.M. Davis, J.C. Giddings, Anal. Chem. 55 (1983) 418.

[25] M. Martin, D.P. Herman, G. Guiochon, Anal. Chem. 58 (1986) 2200.

[26] J.M. Davis, J. Chromatogr. A 831 (1999) 37.

[27] J.M. Davis, J. Chromatogr. A 449 (1988) 41.

[28] J.M. Davis, P.W. Carr, Anal. Chem. 81 (2009) 1198.

[29] A. Felinger, in Advances in Chromatography, Vol 39, 1998, p. 201.

[30] A. Felinger, Anal. Chem. 69 (1997) 2976.
[31] M.R. Schure, J.M. Davis, J. Chromatogr. A 1218 (2011) 9297.
[32] A. Felinger, Data Analysis and Signal Processing in Chromatography, Elsevier, Amsterdam, 1998.
[33] D.R. Stoll, X. Li, X. Wang, P.W. Carr, S.E.G. Porter, S.C. Rutan, J. Chromatogr. A 1168 (2007) 3.
[34] J.L. Frahm, B.E. Howard, S. Heber, D.C. Muddiman, J. Mass Spectrom. 41 (2006) 281.
[35] X. Li, P.W. Carr, J. Chromatogr. A 1218 (2011) 2214.
[36] D.R. Stoll, J.D. Cohen, P.W. Carr, J. Chromatogr. A 1122 (2006) 123.
[37] L.R. Snyder, J.J. Kirkland, Introduction to modern liquid chromatography, John Wiley & Sons, Inc., New York,2nd, 1978.
[38] B.D. Fitz, R.B. Wilson, B.A. Parsons, J.C. Hoggard, R.E. Synovec, J. Chromatogr. A 1266 (2012) 116.
[39] M.R. Schure, J. Chromatogr. A 1218 (2011) 293.
[40] R. De Levie, Advanced Excel for scientific data analysis, Oxford University Press, Oxford ; New York, 2004.
[41] C.L. Phillips, Signals, systems, and transforms, Prentice Hall, Upper Saddle River, NJ :, 2008.
[42] J.W. Cooley, J.W. Tukey, Math. Comput. 19 (1965) 297.
[43] D.N. Rockmore, Comput. Sci. Eng. 2 (2000) 60.
[44] B. Walczak (B. Walczak),B. Walczaks), Wavelets in chemistry, Elsevier Science B.V., Amsterdam, 2000.
[45] D. Allegri, G. Mori, R. Seeber, Analyst 121 (1996) 1359.
[46] D.C. Villalanti, M.F. Burke, J.B. Phillips, Anal. Chem. 51 (1979) 2222.
[47] H.C. Smit, Trac-Trends Anal. Chem. 2 (1983) 1.
[48] H.J. Crabtree, M.U. Kopp, A. Manz, Anal. Chem. 71 (1999) 2130.
[49] Y.C. Kwok, A. Manz, Electrophoresis 22 (2001) 222.
[50] J.G. Brayan, D.J. Malcolmelawes, C.D. Mew, S. Xie, J. Autom. Chem. 17 (1995) 77.
[51] M. Kaljurand, H.C. Smit, Chemom. Intell. Lab. Syst. 79 (2005) 65.
[52] P.B. Allen, B.R. Doepker, D.T. Chiu, Anal. Chem. 79 (2007) 6807.
[53] H.P. Chen, H.J. Liao, C.M. Huang, S.C. Wang, S.N. Yu, J. Chromatogr. A 1217 (2010) 2804.
[54] L. Debnath, D. Bhatta, Integral transforms and their applications, Chapman & Hall/CRC Boca Raton,2nd, 2007.
[55] M. Radermacher, T. Ruiz, T. Clason, S. Benjamin, U. Brandt, V. Zickermann, J. Struct. Biol. 154 (2006) 269.
[56] S. Benjamin, M. Radermacher, J. Kirchberger, T. Schöneberg, A. Edelmann, T. Ruiz, J. Struct. Biol. 168 (2009) 345.
[57] I. Elouedi, R. Fournier, A. Nait-Ali, A. Hamouda, Signal Processing 93 (2013) 345.
[58] W.A. Götz, H.J. Druckmüller, Pattern Recognition 29 (1996) 711.
[59] S.G. Mallat, A wavelet tour of signal processing, Academic Press, San Diego :, 1999.
[60] J.C.T. Eijkel, Y.C. Kwok, A. Manz, Lab Chip 1 (2001) 122.
[61] X.G. Shao, W.S. Cai, Rev. Anal. Chem 17 (1998) 235.
[62] X.G. Shao, A.K.M. Leung, F.T. Chau, Acc. Chem. Res. 36 (2003) 276.

[63] X.G. Shao, Z.C. Liu, W. Cai, Analyst 134 (2009) 2095.
[64] X.G. Shao, W.S. Cai, P.Y. Sun, Chemom. Intell. Lab. Syst. 43 (1998) 147.
[65] B.F. Liu, Y. Sera, N. Matsubara, K. Otsuka, S. Terabe, Electrophoresis 24 (2003) 3260.
[66] A. Kuczynska, B. Walczak, Acta Chromatographica 12 (2002) 11.
[67] X.G. Shao, C.X. Ma, Anal. Lett. 36 (2003) 1261.
[68] Y. Yu, L. Wang, R. Zhao, D. Shangguan, G.Q. Li, Chinese Journal of Analytical Chemistry 31 (2003) 1062.
[69] Z.M. Zhang, S. Chen, Y.Z. Liang, Analyst 135 (2010) 1138.
[70] X.G. Shao, G.Q. Wang, S.F. Wang, Q.D. Su, Anal. Chem. 76 (2004) 5143.
[71] X.G. Shao, L. Sun, Anal. Lett. 34 (2001) 267.
[72] J.D. Dimitrov, Anal. Bioanal. Chem. 379 (2004) 312.
[73] X.Q. Zhang, H. Liu, J.B. Zheng, H. Kao, Progress in Chemistry 14 (2002) 174.
[74] Y. Zhang, J. Mo, T. Xie, P. Cai, X. Zou, Anal. Chim. Acta 437 (2001) 151.
[75] J. Lasa, I. Sliwka, J. Rosiek, K. Wal, Chem. Anal. (Warsaw) 46 (2001) 529.
[76] W.D. Cao, X.Y. Chen, X.R. Yang, E.K. Wang, Electrophoresis 24 (2003) 3124.
[77] S. Cappadona, F. Levander, M. Jansson, P. James, S. Cerutti, L. Pattini, Anal. Chem. 80 (2008) 4960.
[78] A. Felinger, M. Kare, Chemom. Intell. Lab. Syst. 72 (2004) 225.
[79] Y. Cai, D. Janasek, J. West, J. Franzke, A. Manz, Lab Chip 8 (2008) 1784.
[80] J.C.T. Eijkel, A. van den Berg, A. Manz, Electrophoresis 25 (2004) 243.
[81] K.B. Mogensen, Y.C. Kwok, J.C.T. Eijkel, N.J. Petersen, A. Manz, J.P. Kutter, Anal. Chem. 75 (2003) 4931.
[82] X. Yang, G. Jenkins, J. Franzke, A. Manz, Lab Chip 5 (2005) 764.
[83] Y.C. Kwok, A. Manz, Analyst 126 (2001) 1640.
[84] Y.C. Kwok, A. Manz, J. Chromatogr. A 924 (2001) 177.
[85] Y.C. Kwok, A. Manz, in A. VandenBerg, P. Bergveld, W. Olthuis (Editors), Micro Total Analysis Systems 2000, Proceedings, 2000, p. 603.
[86] Y.C. Kwok, N.T. Jeffery, A. Manz, Anal. Chem. 73 (2001) 1748.
[87] A. Manz, Lab Chip 8 (2008) 13.
[88] J.A. McReynolds, P. Edirisinghe, S.A. Shippy, Anal. Chem. 74 (2002) 5063.
[89] S. Kitagawa, K. Mitsuya, C. Chaiyasut, T. Tsuda, J. Sep. Sci. 26 (2003) 1169.
[90] N.E. Watson, W.C. Siegler, J.C. Hoggard, R.E. Synovec, Anal. Chem. 79 (2007) 8270.
[91] W.C. Siegler, J.A. Crank, D.W. Armstrong, R.E. Synovec, J. Chromatogr. A 1217 (2010) 3144.
[92] H.M. Cai, S.D. Stearns, Anal. Chem. 76 (2004) 6064.
[93] R.E. Synovec, B.J. Prazen, K.J. Johnson, C.G. Fraga, C.A. Bruckner, Adv. Chromatogr. 42 (2003) 1.
[94] A.E. Sinha, C.G. Fraga, B.J. Prazen, R.E. Synovec, J. Chromatogr. A 1027 (2004) 269.
[95] D. Halliday, R. Resnick, Physics: Part 1, Wiley, New York,International, 1966.

[96] M.J.E. Trudgett, G. Guiochon, R.A. Shalliker, J. Chromatogr. A 1218 (2011) 3545.

[97] S. Goldman, Frequency analysis, modulation and noise, McGraw-Hill, New York, 1948.

[98] T. Kaneta, Y. Yamaguchi, T. Imasaka, Anal. Chem. 71 (1999) 5444.

[99] O. Trapp, Angew. Chem. Int. Ed. 46 (2007) 5609.

[100] L.R. Snyder, J.J. Kirkland, J.L. Glajch, Practical HPLC Method Development, John Wiley & Sons, Inc.,2, 1997.

[101] K.V. Rangarao, Digital signal processing a practitioner's approach, John Wiley, Chichester, England, 2005.

[102] A. Economou, P.R. Fielden, A.J. Packham, Analyst 121 (1996) 97.

[103] A. Felinger, T.L. Pap, J. Inczedy, Anal. Chim. Acta 248 (1991) 441.

[104] O. Trapp, J. Chromatogr. A 1217 (2010) 6640.

[105] C.H. Lin, T. Kaneta, H.M. Chen, W.X. Chen, H.W. Chang, J.T. Liu, Anal. Chem. 80 (2008) 5755.

[106] Z.Z. Fan, C.H. Lin, H.W. Chang, T. Kaneta, J. Chromatogr. A 1217 (2010) 755.

[107] A.G. Marshall, Fourier, Hadamard, and Hilbert Transforms in Chemistry, Plenum Press, New York, 1982.

[108] J.B. Phillips, Abstr. Pap. Am. Chem. Soc. 193 (1987) 113.

[109] J.M. Davis, J. Sep. Sci. 28 (2005) 347.

[110] R.C. Willoughby, E. Sheehan, S. Mitrovich, A global view of LC/MS : how to solve your most challenging analytical problems, Global View Publ., Pittsburgh, Pa.,1st, 1998.

[111] J.A. McReynolds, S.A. Shippy, Anal. Chem. 76 (2004) 3214.

[112] M. Mnatsakanyan, T.A. Goodie, X.A. Conlan, P.S. Francis, G.P. McDermott, N.W. Barnett, D. Shock, F. Gritti, G. Guiochon, R.A. Shalliker, Talanta 81 (2010) 837.

[113] M. Mnatsakanyan, P.G. Stevenson, X.A. Conlan, P.S. Francis, T.A. Goodie, G.P. McDermott, N.W. Barnett, R.A. Shalliker, Talanta 82 (2010) 1358.

[114] M. Mnatsakanyan, P.G. Stevenson, D. Shock, X.A. Conlan, T.A. Goodie, K.N. Spencer, N.W. Barnett, P.S. Francis, R.A. Shalliker, Talanta 82 (2010) 1349.

[115] S. Pravadali, D.N. Bassanese, X.A. Conlan, P.S. Francis, Z.M. Smith, J.M. Terry, R.A. Shalliker, Anal. Chim. Acta 803 (2013) 188.

[116] P.W. Carr, D.R. Stoll, X. Wang, Anal. Chem. 83 (2011) 1890.

[117] C. Paek, Y. Huang, M.R. Filgueira, A.V. McCormick, P.W. Carr, J. Chromatogr. A 1229 (2012) 129.

[118] Hamamatsu, CCD Image Sensors S11071/S10420-01 Series, Hamamatsu Photonics, Hamamatsu City, Japan, 2011.

[119] W.H. Press, Numerical recipes: the art of scientific computing, Cambridge University Press, New York,3rd, 2007.

[120] M.J.E. Trudgett, The Ultra-Trace Level Analysis of Xanthates by High Performance Liquid Chromatography, M.Sc.(Hons) Thesis, University of Western Sydney, Sydney, NSW, 2005.

[121] D. Connolly, L.P. Barron, E. Gillespie, B. Paull, Chromatographia 70 (2009) 915.

[122] H.F. Lau, N.M. Quek, W.S. Law, J.H. Zhao, P.C. Hauser, S.F.Y. Li, Electrophoresis 32 (2011) 1190.

[123] W. Li, Y. Pan, Y. Liu, X. Zhang, J. Ye, Q. Chu, Chromatographia (2013) 1.
[124] A.M. Krstulovic, H. Colin, G.A. Guiochon, Adv. Chromatogr. (N. Y.) 24 (1984) 83.
[125] M.A. Schwarz, P.C. Hauser, Lab Chip 1 (2001) 1.
[126] J.A. Abia, J. Putnam, K. Mriziq, G.A. Guiochon, J. Chromatogr. A 1217 (2010) 1695.
[127] K.L. Braun, S. Hapuarachchi, F.M. Fernandez, C.A. Aspinwall, Electrophoresis 28 (2007) 3115.
[128] M.B. Smalley, J.M. Shaver, L.B. McGown, Anal. Chem. 65 (1993) 3466.
[129] R. Garrido-Medina, J.C. Diez-Masa, M. De Frutos, Electrophoresis 34 (2013) 2295.
[130] A. Mazouchi, B.J. Dodgson, D.W. Wegman, S.N. Krylov, C.C. Gradinaru, Analyst 137 (2012) 5538.
[131] Y. Zhu, N. Chen, Q. Li, Q. Fang, Analyst 138 (2013) 4642.
[132] E. Vinner, M. Stievenart, L. Humbert, D. Mathieu, M. Lhermitte, Biomed. Chromatogr. 15 (2001) 342.
[133] K.B. Mogensen, O. Gustafsson, P.S. Nunes, J.P. Kutter, in, 2008.
[134] M.S. Galhiane, S.R. Rissato, B.M. Apon, J. Liq. Chromatogr. Related Technol. 28 (2005) 2349.
[135] Y. Guo, X. Fan, in, 2012.
[136] K. Reddy, J. Liu, M.K.K. Oo, X. Fan, Journal of Microelectromechanical Systems 22 (2013) 1174.
[137] J.H. Seo, J. Liu, X. Fan, K. Kurabayashi, Lab on a Chip - Miniaturisation for Chemistry and Biology 13 (2013) 851.
[138] J. Liu, J.H. Seo, Y. Li, D. Chen, K. Kurabayashi, X. Fan, Lab on a Chip - Miniaturisation for Chemistry and Biology 13 (2013) 818.
[139] Nirosh, Introduction to Object Oriented Programming Concepts (OOP) and More, Code Project (Online: http://www.codeproject.com/), 2013.
[140] A. Felinger, G. Guiochon, J. Chromatogr. A 913 (2001) 221.
[141] A.W. Moore, J.W. Jorgenson, Anal. Chem. 67 (1995) 3456.
[142] J. Mohr, R. Swart, M. Samonig, G. Böhm, C.G. Huber, Proteomics 10 (2010) 3598.
[143] D. Megson, R. Kalin, P.J. Worsfold, C. Gauchotte-Lindsay, D.G. Patterson, M.C. Lohan, S. Comber, T.A. Brown, G. O'Sullivan, J. Chromatogr. A 1318 (2013) 276.
[144] N. Ali, A. Thanga Thirupathi, K.S. Kesavanarayanan, H. Sumathy, V. Gayathri, International Journal of Pharmacy and Pharmaceutical Sciences 5 (2013) 340.
[145] H.J. de Geus, I. Aidos, J. de Boer, J.B. Luten, U.A.T. Brinkman, J. Chromatogr. A 910 (2001) 95.
[146] M. Schäffer, T. Gröger, M. Pütz, R. Zimmermann, Forensic Toxicology 31 (2013) 251.
[147] M. Thevis, A. Thomas, W. Schänzer, Anal. Bioanal. Chem. (2002) 1.
[148] P. Van Renterghem, P.E. Sottas, M. Saugy, P. Van Eenoo, Anal. Chim. Acta 768 (2013) 41.
[149] C. Reichel, Forensic Science International 213 (2011) 20.

Appendix A.
COMForTS ECD Design and Schematics

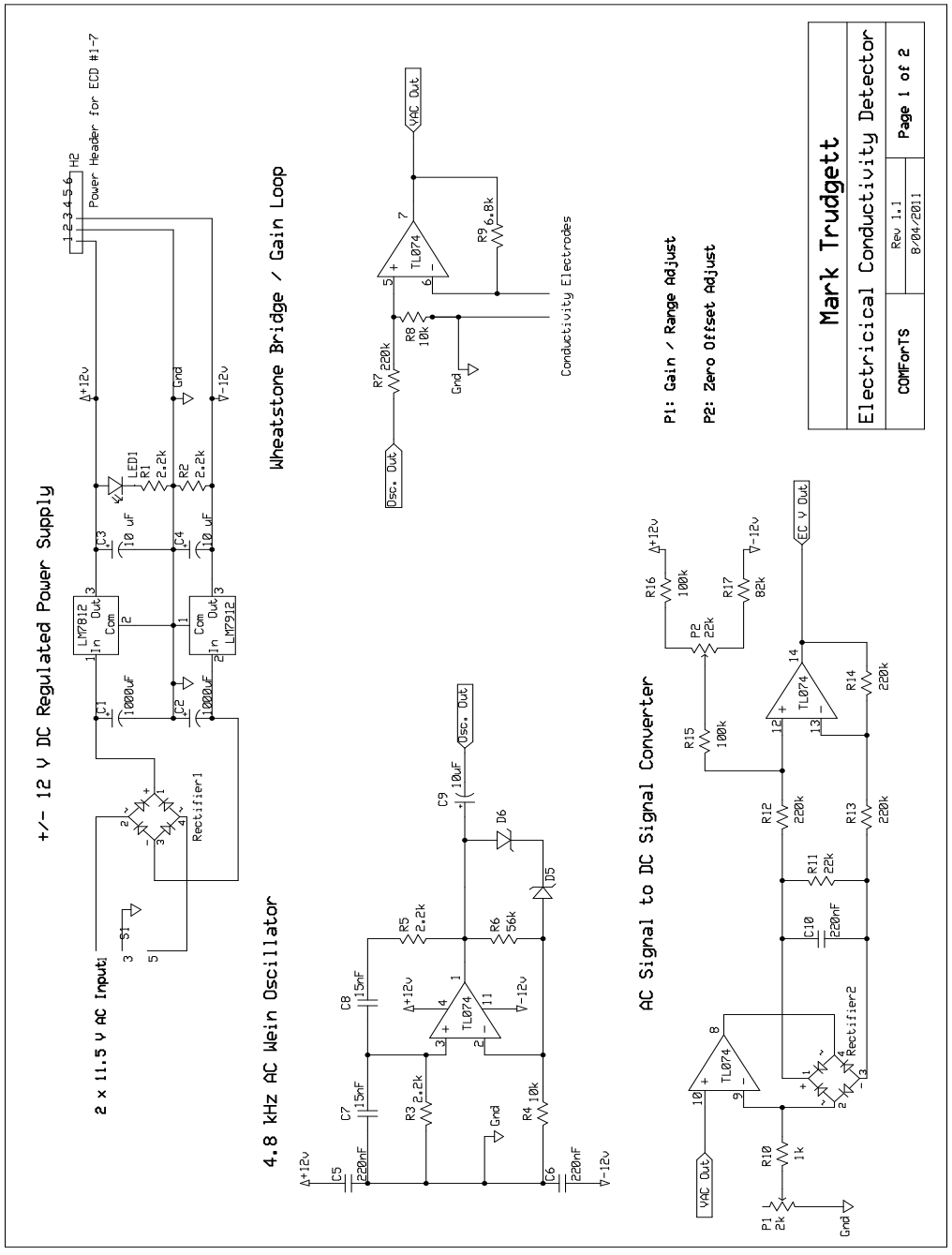


Figure A-1: COMForTS Digital Electrical Conductivity Detector Schematic 1. Schematic for power supplies, Wein oscillator, Wheatstone bridge and AC signal rectifier / amplifier.

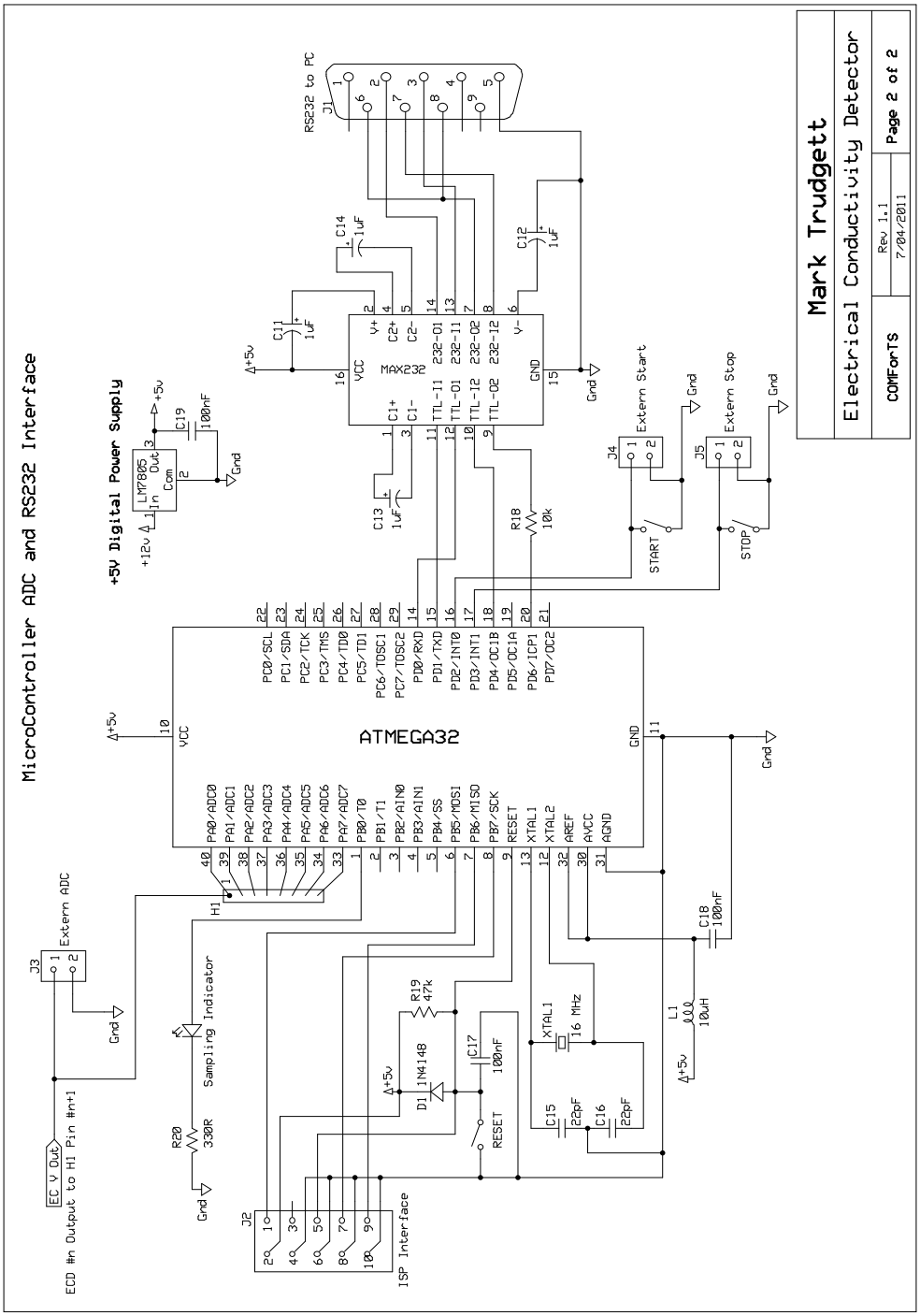


Figure A-2: COMForTS Digital Electrical Conductivity Detector Schematic 2. Schematic for microcontroller, programming interface, A/D conversion and serial digital (RS232) PC interface.

Firmware (for the Atmega32 microcontroller) and Windows software for this detector are to be found in the "[COMForTS]:\Source Code\COMForTS ECD" folder of the Digital Thesis Materials.

Appendix B.
Completed COMForTS ECD

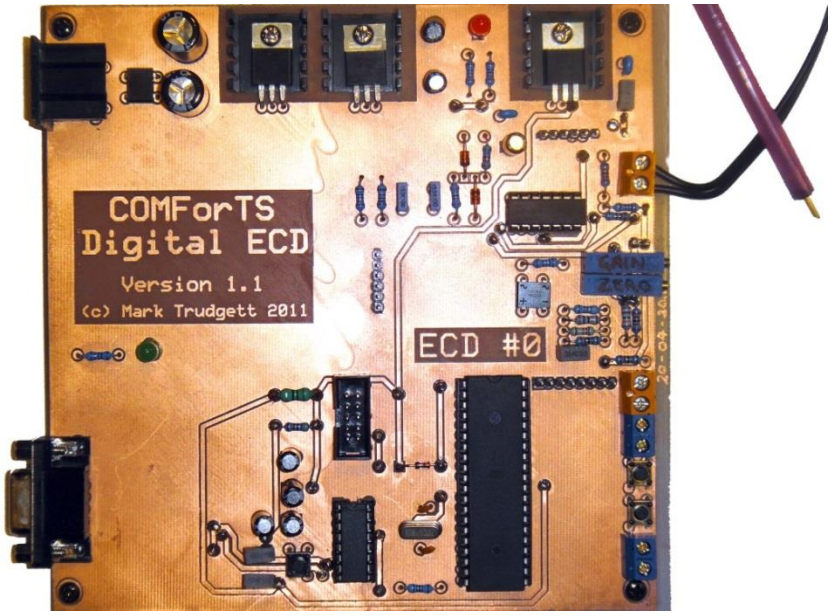


Figure B-1: COMForTS ECD circuit board, upper surface.
Test electrodes (comprising gold PCB pins) are attached. This board includes power supplies (top section) and the microcontroller interface (lower section) as well as a single ECD (ECD #0). Up to seven additional independent ECDs, for arrayed detection, can be connected and 'stacked' to the headers located above and below ECD #0

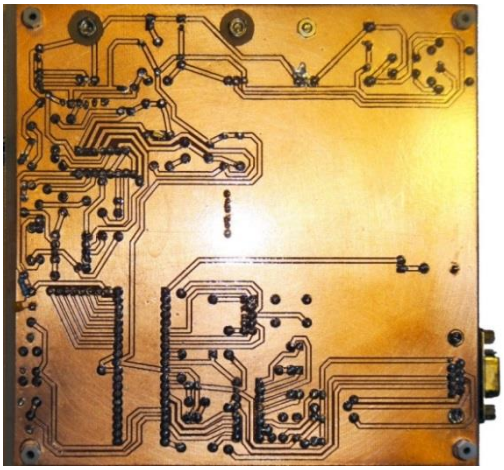


Figure B-2: COMForTS ECD circuit board, lower surface.
The RS232C communications port is located in the lower right. The detector was found to be noisy (due to the non-sinusoidal output of the Wein oscillator) but, nonetheless, sensitive to ~5 mg / L sodium chloride in water.

Appendix C.
CCD UV Absorbance Detector

The internal layout and major components of the detector are shown in Figure C-1.

Parts list

- UV CCD chip (S10420-1106-01, Hamamatsu Photonics, Japan). 2048 × 96 active 14 μm pixels (sensor area: 28.672 × 0.896 mm) (Figure C-2)
- CCD driver circuit (C11287, Hamamatsu Photonics, Japan) (Figure C-2)
- 255 nm UV LEDs (UVTOP255TO39HS and UVTOP255TO39FW, Sensor Electronic Technology, Inc., (SETi), South Carolina, USA)
- LED driver circuit. Designed (Figure C-3) and built in house (Figure C-4)
- Active and passive electronic components were obtained with 1% tolerances where applicable from Jaycar Electronics, Rydalmere, Australia.
- Miscellaneous hardware (connectors, screw, bolts, case etc.) were obtained from Jaycar Electronics, Rydalmere, Australia and general hardware suppliers.

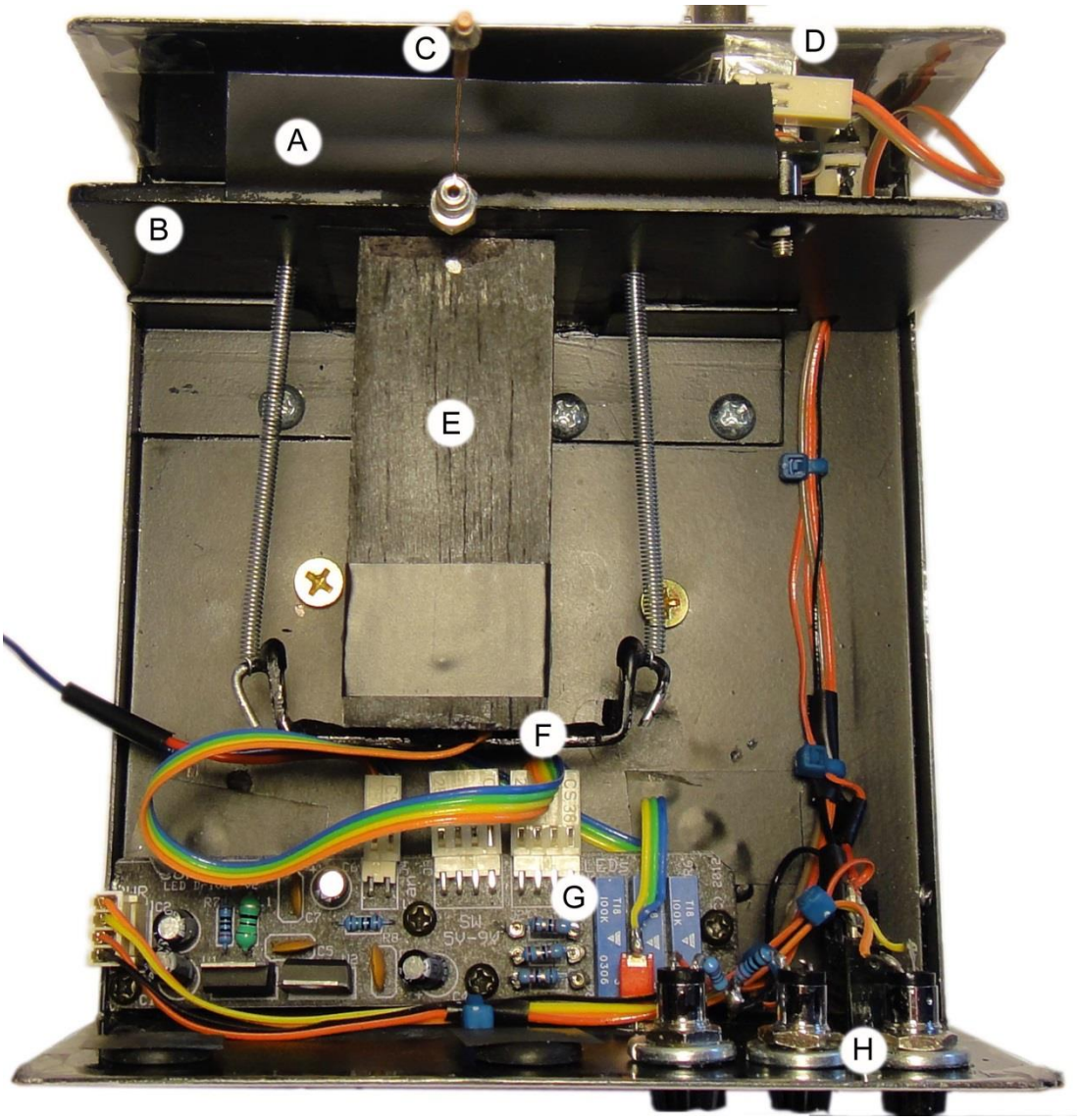


Figure C-1: Internal layout of the COMForTS CCD UV Absorbance Detector.
The image scale is approximately 1:1

Key to Figure C-1:

A. Light baffle over Hamamatsu C11287 CCD Driver Board
B. CCD Driver board mounting plate.
C. Capillary column (packed) passing over CCD sensor
D. USB 2.0 interface and common ground connector
E. Optical bench
F. LED light source circuit board (not visible) and clamp
G. LED driver circuit board
H. Front-panel status indicators.

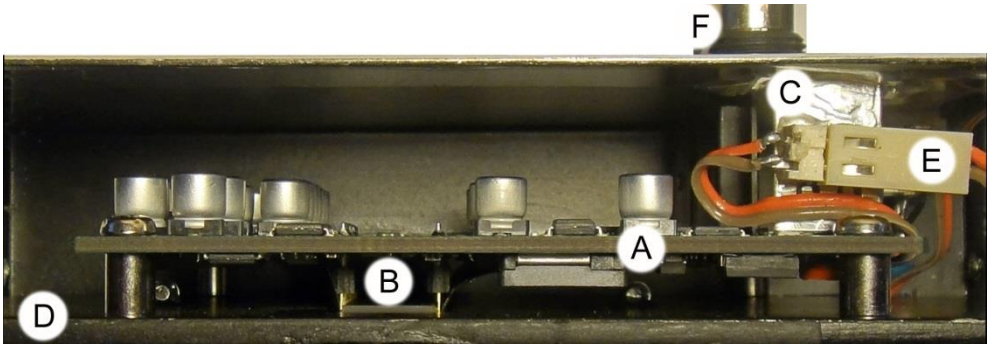


Figure C-2: Hamamatsu C11287 CCD Driver Board.
Shown *in situ* with an image scale of approximately 1.4:1

Key to Figure C-2:

A. Hamamatsu C11287 CCD Driver Board
B. Hamamatsu UV CCD chip S10420-1106-01. The fused silica window of the chip is flush with the lower (vertical) surface of mounting plate (D).
C. USB 2.0 interface. Windows PC control is via the driver UsbCamIF.sys and function library DCamUSB.dll.
D. CCD Driver board mounting plate.
E. +5V and 0V connector for CCD board status and common ground.

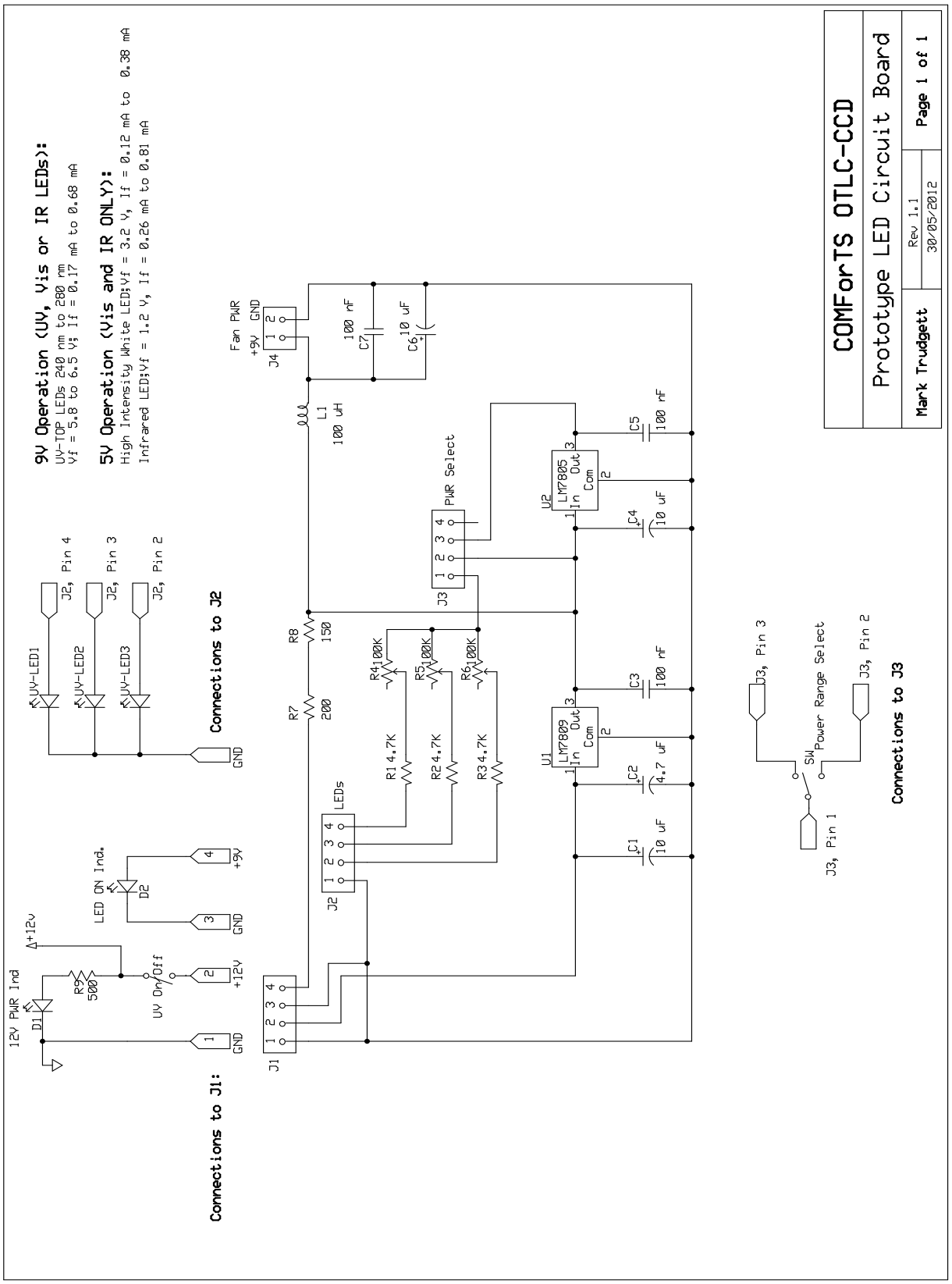


Figure C-3: LED Driver Circuit Schematic.
The driver circuit provides regulated voltage and noise-filtered power to three independently controlled LEDs.

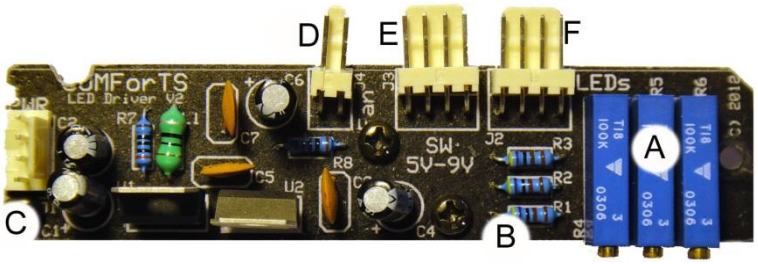


Figure C-4: LED Driver Board.
Drives up to three LEDs and supplies stable +5V and +9V with high and low range variable power to a variety of common LEDs. Image scale is approximately 1:1

Key to Figure C-4:

A. 10-turn current limiting potentiometers control the power supplied to each of three LEDs
B. High-range (maximum) current-limiting resistors. Shown permanently mounted here but later changed to a PCB socket mount to allow substitution (see Figure C-1 (G)).
C. +12V (independent) power supply connection and common ground connection.
D. Power supply connection (+9V) to case ventilation fan.
E. Connection to high / low range LED power selection switch
F. Connection to LED circuit board. The LED circuit board provides wired PCB sockets for mounting the LEDs in the optical bench.

Appendix D.
A Method for Combining Signals from Closely Spaced Detectors (MCSCSD)

D.1 Background

MCSCSD is a method for combining multiple time-domain signals derived from closely spaced (effectively independent) detectors for the purpose of transformation into a single frequency-domain signal. This method applies to chromatographic signals where "closely-spaced" is defined as a distance that is less than or equal to the physical peak width (i.e. the physical length of an analyte peak) within a given separation dimension (such as a chromatographic column).

This method is applicable to all multipoint detection methods that derive frequency domain data (such as SCOFT) or time-dependent frequency domain data (COMForTS) from the combination of signals from multiple detectors (or multiple detection points).

MCSCSD was devised by the author as a solution to the problem of separation dimensions being unnecessarily lengthy (for the required physical separation) in the COMForTS prototype. This column length created an undesirably bulky system with long analysis times.

D.2 The problem

In order to obtain a true frequency spectrum from multiple detector signals (in the COMForTS or SCOFT methods), the entirety of information embodied in individual detector signals must be preserved in the combined signal.

Generally, this would imply that the combined signal is explicitly calculated as the sum of each of the individual detector signals. When such signals are derived from "well-spaced" detectors (i.e. detectors that are spaced at a distance that is greater than the physical analyte peak width as in Figure 1-3a), the simple sum of signals produces a series of discrete pulses (Figure 1-3b). These pulses can then be easily transformed to produce a strong and clear frequency domain response (Figure 1-3c). However, as the spacing of detectors decreases below that of the analyte peak width (Figure D-1), the combined signal becomes 'muddied' (Figure D-2) due to the fact that an analyte peak is a broad signal spread out over time rather than a discrete digital pulse. These 'muddied' signals include only a small component that can be identified at the frequency of the analyte peaks. They also contain a large component that contains no frequency information that is thus interpreted by the Fourier transform as the sum of a set of sine and cosine signals that are not related to the frequency of the analyte peaks. Such signals thus produce transformed signals show a weak analyte frequency response with strong near-zero frequency components and strong harmonic frequencies, that respectively weaken and strengthen with decreased detector spacing (Figure D-3).

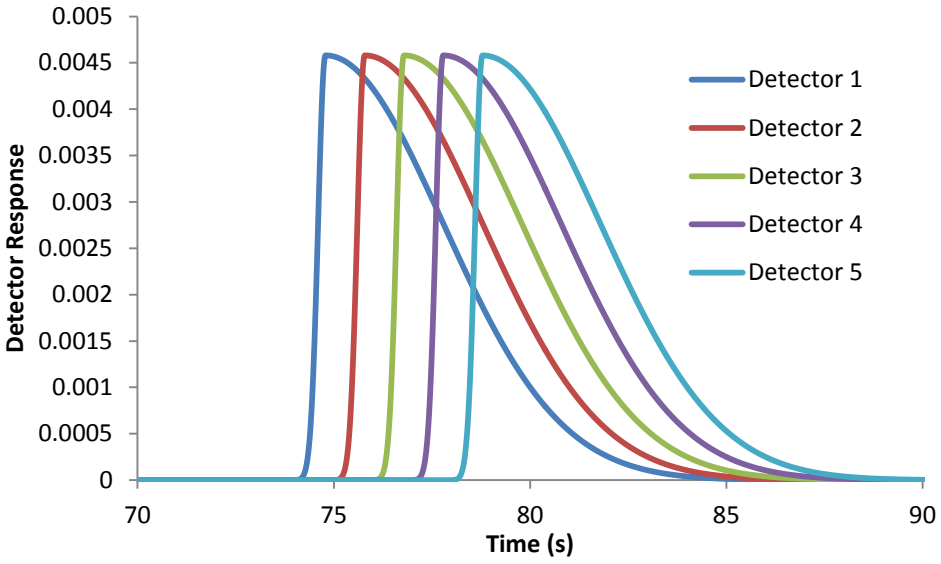


Figure D-1: Time domain signal from closely spaced detectors.
Signals are shown from five detectors where the distance between detectors is one twelfth of the physical width of the peaks within the column. Highly asymmetric peaks are shown because their sum (Figure D-2) is more clearly pulsed than that of symmetric peaks.

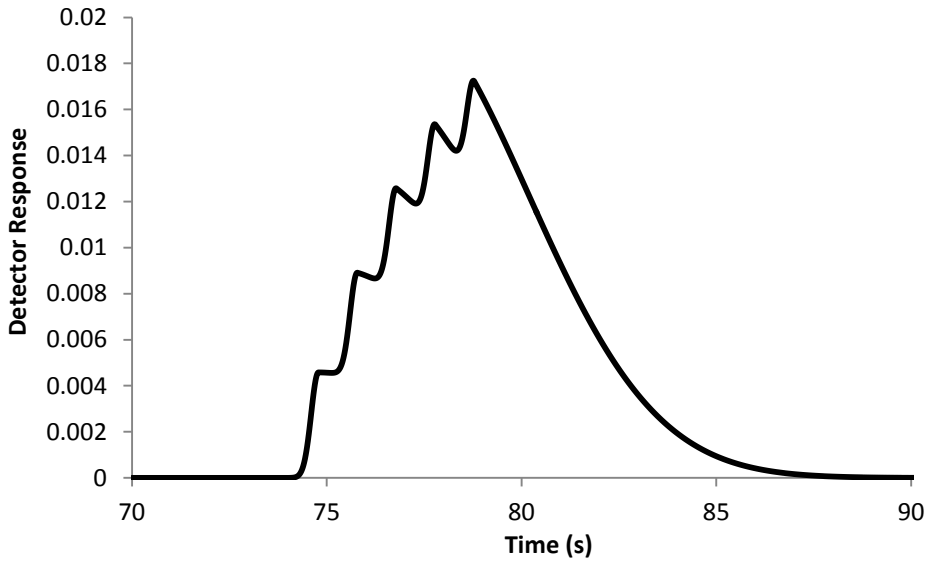


Figure D-2: Pulsed signals from closely spaced detectors.
This signal is the sum of the signals shown in Figure D-1 and shows a virtually pulsed 1 Hz signal.

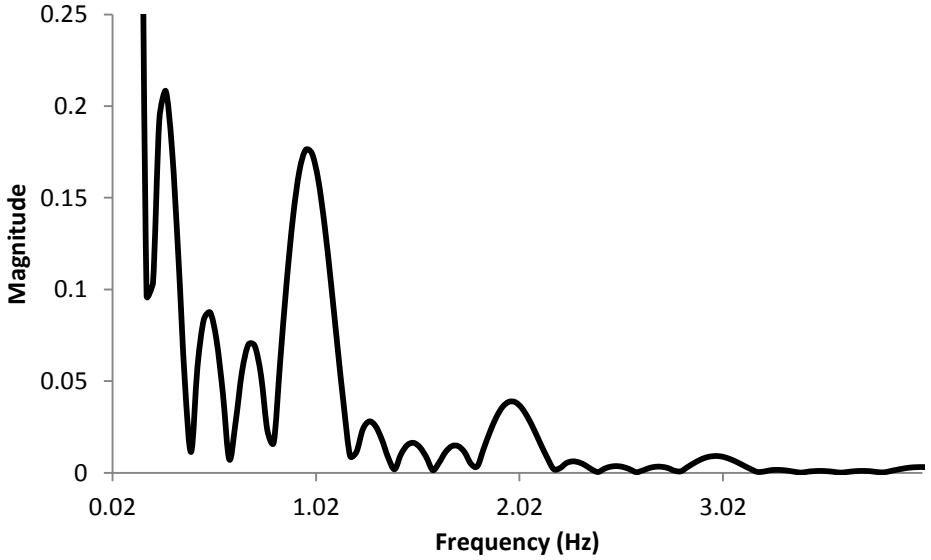


Figure D-3: Fourier transform of virtually pulsed 1 Hz signal.
This is the magnitude spectrum of the 1 Hz signal shown in Figure D-2. Whilst the fundamental frequency is evident, it is by no means the predominant frequency.

Note that the fundamental frequency (1 Hz) is not the major frequency component in the transformed signal. There are also several strong overtones (at 2×, 3× and 4× the fundamental frequency) and numerous strong harmonics, some with magnitudes greater than 120% of the magnitude of the fundamental.

This means that with a conventional data treatment, the size of instruments that employ multi-point detection (where greater sensitivity is provided by a larger number of detectors) must be large with respect to analyte peak widths. For example, at the HPLC/UHPLC scale, analyte peaks are approximately 8 cm in width and the (rigid) column would need to be at least 4 m long (as opposed to a common 10 cm length) in order to accommodate (a modest) 50 detection points. This not only makes instruments unusually bulky for the relevant chromatographic scale but also necessarily increases the time required for an analysis.

D.3 The Solution

The solution to this difficulty lies in the fact that detectors that are physically close produce time-domain signals that, when combined, have pulses that are closely spaced in the time domain. However, if before combining the signals, we artificially displace the time domains of each detector signal (by a fixed and known degree) we can obtain signal pulses that *are* well spaced *in time* (i.e. the peak width in time, P_w, is now less than the sampled pulse period P_s).

As frequency is defined as the reciprocal of period, in a combined signal the frequency F_s of an analyte peak signal is similarly related to its period (P_s):

$$F_s = 1/P_s \qquad \text{Eq. (D-1)}$$

and

$$P_s = 1/F_s \qquad \text{Eq. (D-2)}$$

If the detector signals are artificially displaced in time (Figure D-4), the period of the signals (P_s) may be extended by a known amount (d), such that, when summed (Figure D-5), the period of the displaced signal (P_d) is $P_s + d$, and the frequency of the displaced signal (F_d) is:

$$F_d = 1/(P_s + d)$$ Eq. (D-3)

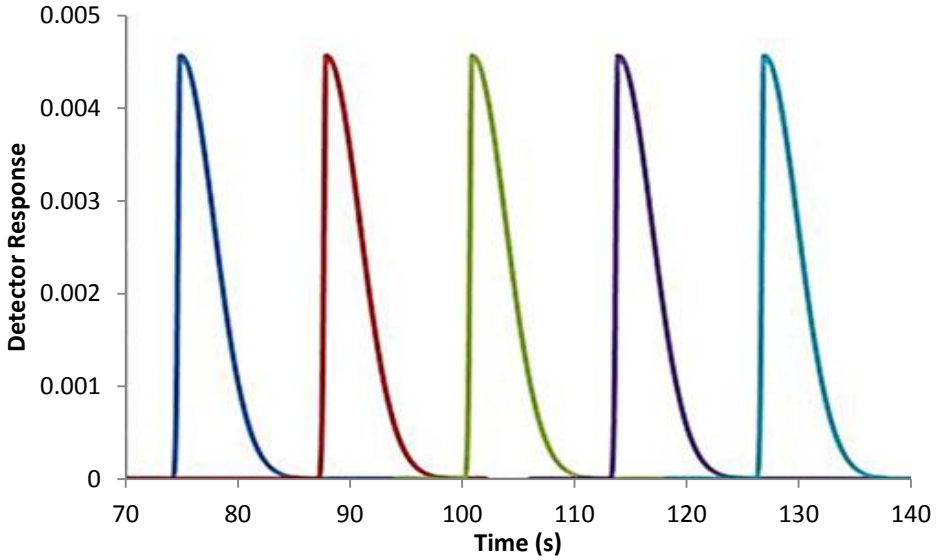


Figure D-4: Time-displaced signals, where d = 12 s
These displaced signals are derived from those shown in Figure D-1.

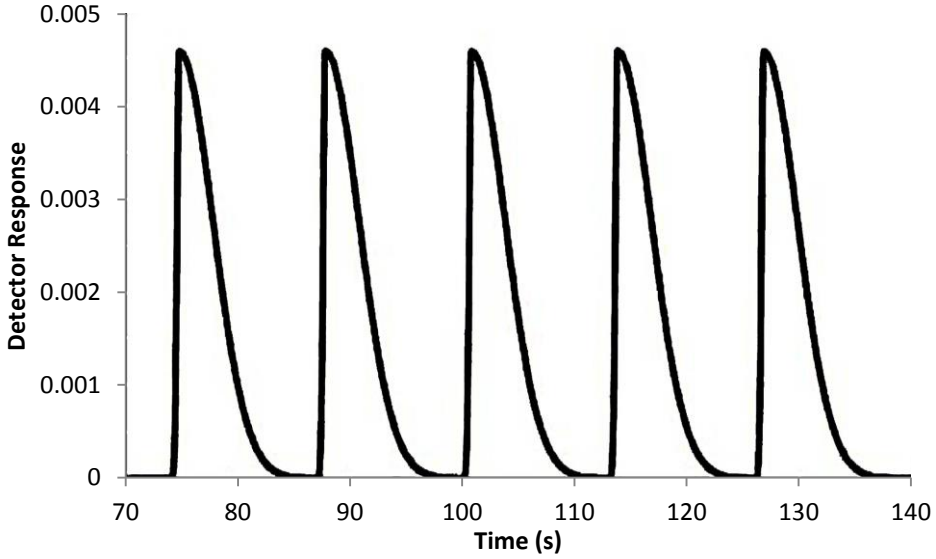


Figure D-5: The result of summation of time displaced signals.
The time displacement, d is equal to the peak width in time, 12 s.

The (Fourier) transform of a signal containing the frequency F_d will produce a well-defined and intense frequency component corresponding to F_d, (Figure D-6), if d is chosen such that (P_s + d) > P_w. This manipulation on its own is sufficient to obtain frequency spectra of analyte pulses – but the frequencies reported do not represent the actual frequency of the pulses.

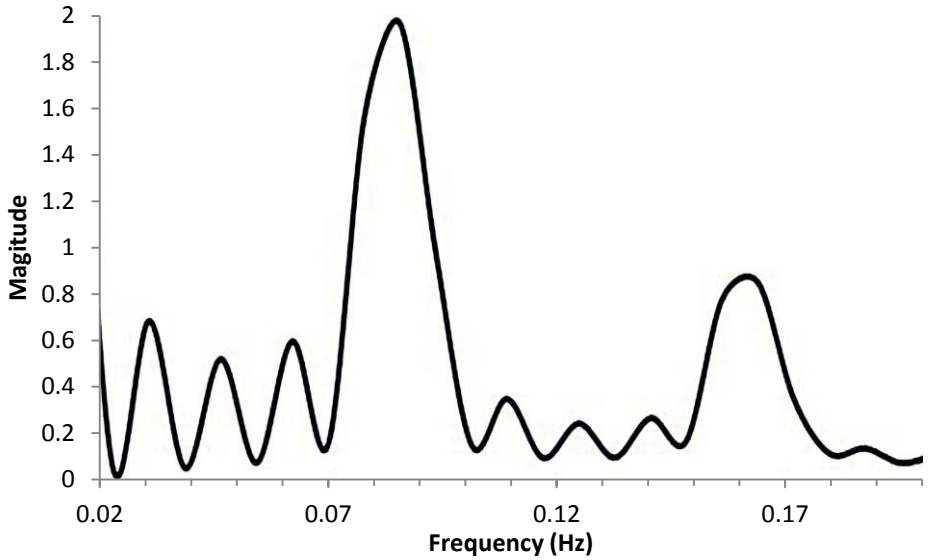


Figure D-6: Fourier transform of the combined displaced signals.

Note that the above transform data were produced using Microsoft Excel 2010, in which the implementation of the Fast Fourier transform (FFT) is approximate and not commutative. Note, however that the fundamental frequency (1 Hz in Figure D-3 and 0.08 Hz in Figure D-6) has a magnitude more than ten times greater than that of the undisplaced summed signal. The MCSCD method has produced one simple overtone at 2× the fundamental and relatively smaller harmonic contributions (<40% of the fundamental magnitude), using the same Excel FFT routine as in Figure D-3. The displaced frequency F_d is as expected (within the error of Excel's calculations) and higher resolution transforms would allow the following calculations.

Substituting P_s from Eq. (D-2) into Eq. (D-3), and solving for P_s yields:

$$P_s = \frac{1}{F_d} - d \qquad \text{Eq. (D-4)}$$

and as the true frequency of the analyte pulses F_s is given by Eq. (D-1):

$$F_s = 1/\left(\frac{1}{F_d} - d\right) \qquad \text{Eq. (D-5)}$$

Eq. (D-4) and Eq. (D-5) allow calculation of the true analyte peak velocity from the displaced frequency F_d reported by the transform of the displaced signals. Knowing the velocity of an analyte peak allows calculation of its retention time

(the time required to transverse a separation dimension of known length). This is important for the correlation of analyte retention times between dimensions in multidimensional separations.

Thus, the artificial time-displacement of detector signals from closely spaced detectors can provide strong, clear frequency spectra from which the true analyte frequency and velocities can be calculated.

To produce the combined signal S_c, the algorithm for combining individual time-dependent signals, S_0 to S_n, from very closely spaced detectors (numbered 0 to n) such that a constant displacement, d, is achieved is thus:

$$S_c = \sum_{n=0}^{n} S_{dn}(T_{Sn} + n \times d) \qquad \text{Eq. (D-6)}$$

where T_{Sn} is the vector of time values (time domain, only) of Sn and the function $S_{dn}()$ combines the resulting displaced-time vector with the original undisplaced intensity vector, resulting (mathematically) in a two dimensional signal in displaced time and undisplaced intensity, for each detector n.

The result of applying function $S_{dn}()$ to the raw detector signals (Figure D-1) is shown graphically Figure D-4.

The result of applying the entire algorithm (Eq. (D-6)) to these same raw detector signals is shown graphically in Figure D-5.

D.4 Software implementation

The algorithm (Eq. (D-6)) could be implemented in a variety of different ways depending on the nature of the detectors and the data that they produce. For example, some detectors produce time series signals that are composed of two-dimensional mass spectra. The following is an example of how this algorithm could be implemented in software (using Visual C#) (`Comments are in green, and begin with "//"`), for multiple identical detectors that simultaneously record only a single 'intensity' value at fixed time intervals:

```
using System;

namespace MCSCSD
{
    public class MCSCD_DemoCode
    {

        int NumberOfDetectors = 50;             // Assume that we have fifty detectors
        double d_Seconds;                       // The displacement, d, in time (seconds)
        double SamplingPeriod_Seconds = 0.1;    // All signal samples are taken at the same
time and at 0.1s intervals

        double[] SignalData = new double[2000];    // An array containing a single individual
signal intensity vector, sampled at a constant rate. i.e. the ArrayIndex x SamplingPeriod =
time. In this case, SignalData(20) gives the intensity at 20 x 0.1 s = 2 s
        double[][] AllRawSignals = new double[50][];    // An array containing all of the
individual two dimensional signal vectors
        double[] CombinedSignal;                        // A vector containing the combined
signals from all detectors


        public void PerformMCSCD()
        {
            for (int DetNum = 0; DetNum < NumberOfDetectors; DetNum++) // For each of the
detectors
            {
                SignalData = AllRawSignals[DetNum];                    // extract the current
detector's signal. Because data has been sampled at a regular rate and at the same time for
each detector, we can displace the signals in time by simply moving all the data to higher
index values within the array:
                int IndexShift = Convert.ToInt32((DetNum * d_Seconds) /
            SamplingPeriod_Seconds); // d_Seconds should be chosen to yield an integer result.
                // But the original array is already full, so we need to resize the array (i.e.
increase the number of indices available) whilst preserving its contents)
                // to accommodate the shift to higher indices:
                int NewArraySize = SignalData.Length + IndexShift;
                System.Array.Resize(ref SignalData, NewArraySize);
                // Similarly, the CombinedSignal array must be large enough to hold the shifted
signals:
                if (CombinedSignal.Length < NewArraySize)
                     { System.Array.Resize (ref CombinedSignal, NewArraySize); }

                for (int i = SignalData.GetUpperBound(1); i >=0 ; i--) //Starting from the END
of the signal
                {
                    SignalData[i] = SignalData[i - IndexShift];         // Shift the signal by
the required time (i.e. to a higher array index)
                    CombinedSignal[i] = CombinedSignal[i] + SignalData[i]; // ADD this shifted
signal to the combined signal.
                }
            }
        }
    }
}
```

The complexity of the algorithm's implementation will increase with the complexity of detector signals. Furthermore, the algorithm can be further enhanced by alternate inversion of chromatograms, producing a roughly sinusoidal, zero-average signal that admits well of the Fourier transform.