ELSEVIER

Journal of Chromatography A, 868 (2000) 305–312

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Short communication

Towards a quantitative definition of perfection in chromatographic analyses

N. Dyson

The Croft, Kellfield Avenue, Low Fell, Tyne & Wear, NE9 5YP, UK

Received 13 April 1999; received in revised form 29 November 1999; accepted 29 November 1999

Abstract

The paper published by Ghaoui and Rothman [J. High Resolut. Chromatogr. 15 (1992) 36] and particularly its Fig. 5, is further considered here because it contains the germ of an idea of how to measure improvements to a chromatographic method, and how to define the goal of perfection in terms of ''zero defects'' as required by quality assurance schemes. From this, a new role emerges for signal averaging in capillary chromatography: a role to quantify and measure method improvements, and one which can be generally applied to measure improvements in instrument design too.

Keywords: Signal averaging; Zero defects; Method development; Detection, GC

1. Introduction

In 1992 Ghaoui and Rothman published a paper on improving detection limits in chromatography by signal averaging [1]. This work tested the improvements in design and performance of their capillary gas chromatograph–flame-ionisation detector by signal, or ensemble or boxcar, averaging repeated analyses of trace peaks in a sample to improve signal-to-noise ratio (S/N) and achieve a lowering of the detection limits of their experiments.

Their paper contains the germ of an idea of how to quantitatively assess a chromatographic experiment and judge whether it can be improved given the chromatographic equipment and methods involved.

Fig. 5 of their paper, reproduced below as Fig. 1 of this paper, shows area and height accumulation plots versus number of analysis repetitions (number of files [1]). The accumulated area, or ΣA, plot is linear, the sum of the measured peak areas increases uniformly as the number of peak measurements increase. In contrast, the accumulated height, or ΣH, plot is non linear and falls below the true sum of measured heights even though area and height are measured in the same experiment with similar precision.

The purpose of this paper is to suggest that the curvature of the ΣH plot is a measure of how well an analysis has been performed and can be used as a new diagnostic. It combines and demonstrates the imperfections of instrumental performance, method development and experiment control. The less the curvature, the better the experiment. In the limit, when the ΣH plot is linear, the experiment can be said to be perfect in terms of zero defects [2] as required by quality assurance schemes. In practice, this goal of perfect linearity is not achievable, because it requires the standard deviation of all analytical processes up to the detector to be zero. The value of measuring curvature is that real im-

0021-9673/00/$ – see front matter
PII: S0021-9673(99)01259-5

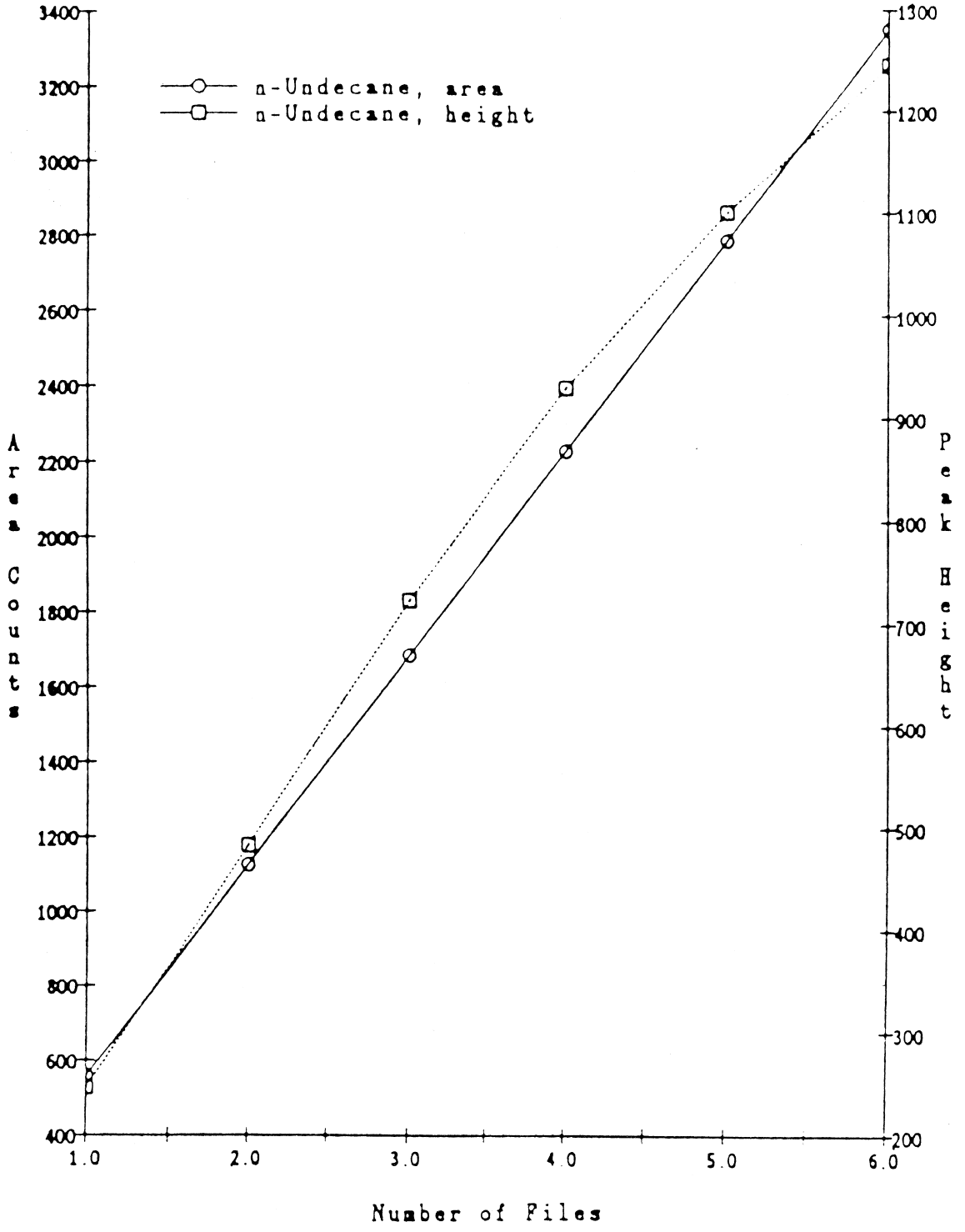

Fig. 1. Plot of area count and peak height for undecane, against number of injections or files summed. (Fig. 5 from [1]; reproduced with kind permission of Wiley–VCH).

provements in methods or in better instrument designs can be measured as reductions in curvature of the ΣH plot and residual curvature can be judged negligible for a given analysis. Curvature of the ΣH plot is an absolute measure of quality.

2. Features of the ΣH plot

2.1. Curvature

The curvature of the ΣH plot is a consequence of experimental variations in retention time. Consider the effect of such retention fluctuations on the summation of peak height and peak area data gathered from repeated analyses. The summation is made by the point by point addition of data from digitized chromatograms. If there is any movement in peak position, corresponding peak retention times measured at the maximum peak height ordinates will not be identical and when data ordinates measured at the average retention time are summed, the total will be less than the sum of the individual peak heights.

Retention times are sensitive indicators of a number of pre-detection factors including:

1. Inadequate synchronization between analysis start and start of data collection
2. Irreproducible injection technique
3. Varying column conditions — pressure, temperature, phase composition
4. Varying mobile phase — flow control, phase composition, leaks
5. Varying peak zone distribution due to varying injection or column conditions

If peak shape is constant, variations in retention time will have equivalent variations in peak start and end times. The net effect on a signal averaged peak, a peak created by the addition of correspondingly timed ''datums'' from each chromatogram, and divided by the number of chromatograms, will be an increase in peak width accompanied by a lowering of peak height compared to any of the component (real) peaks. The height of the signal averaged peak is always less than the numeric average of the component peak heights measured without regard to retention movements.

It will be noted that fluctuations in retention time, a random phenomenon, combine in the ΣH plot to produce a systematic effect on accumulated peak height. No matter how many peaks are added, the accumulated height, measured at the average retention time, will always be less than the sum of the individual heights. The greater the number of added peaks, the greater will be the shortfall in composite peak height. It is this feature which gives the ΣH plot a novel diagnostic value different to standard deviation.

In contrast, a signal averaged peak will have an area equal to the numeric average of the component peak areas ($\pm$experimental error) and will be unaffected by small retention time movements or changes in peak shape though it will have a slightly broader base.

This is what Ghaoui and Rothman found. When they plotted accumulated peak area against the number of repetitions, the ΣA plot was linear but the ΣH plot was non linear and lagged the true sum of component peak heights, see Fig. 1.

2.2. Variations in peak shape

Ghaoui and Rothman attributed all of the non linearity in the ΣH plot to variations in retention times. However, variations in peak shape introduce subtle but widespread effects into measurement accuracy which have an effect on height but not area: variable peak asymmetry contributes to the non linearity of the ΣH plot but does not affect the ΣA plot. In fact, although they were working with trace peaks, Fig. 7 of their paper suggests that peak shape might have been a contributor.

Peak height is usually the largest peak dimension, and so is the most sensitive indicator to variations in peak shape — and the narrower the peak the greater the sensitivity and the greater will be the effect on the curvature of a ΣH plot. As peaks become narrower with better technology, the diagnostic value of the ΣH and ΣA plots increases. Averaged peak width would provide an equivalent diagnostic, but it is a smaller quantity and harder to measure with the same precision.

2.3. Flow sensitive and mass sensitive detectors

The flame ionisation detection used by Ghaoui and Rothman is a mass-sensitive detector. Injector and

column conditions which produce variations in peak shape will affect height but not area, and small variations in retention will have no effect on individual measured areas or heights. Flow sensitive detectors, as used in the various types of capillary liquid chromatography, will generate peak areas which are subject to an additional error from small random variations in analyte passage times through the detector cell. When this happens, peak areas will vary correspondingly but heights and average retention time are preserved. Most importantly, the random fluctuations in area do not combine in a systematic manner, and this distinguishes it from the effect of random variations in retention on averaged height. Compared to the flame ionisation detection area plot in Fig. 1 the ΣA plot derived from a flow sensitive detector will still be linear but the standard deviation of this plot will contain an additional error contribution from flow-rate.

3. Discussion

3.1. Signal or ensemble averaging

Signal averaging [3–5] involves the overlay and addition of data from repeated analyses so that systematic signals (peaks) add and always reinforce each other while random background signals (baseline noise) add but sometimes oppose and cancel each other. It is typically applied to trace analyses. Trace peaks on a noisy baseline are enhanced and easier to measure, noise is suppressed and interferes less with measurements. A spurious event which occurs in one experiment but no other is ''diluted'' by addition to event free repetitions. The accumulation of systematic signal is therefore larger than the accumulation of background noise, and one-off events fade, so the S/N ratio improves. The improvement is equal to $\sqrt{M}$ [3–6] where M is the number of repetitions, i.e. the number of chromatograms added together.

Signal averaging is only meant to deal with signal or y-axis fluctuations, but in chromatography it must also contend with retention time or x-axis fluctuations. These small random variations in retention limit the amount of signal conditioning which can be achieved. Ghaoui and Rothman managed a 3–4 fold improvement in S/N ratio but also became ''impatient'' with that other problem of signal averaging: the time required to gather good data from chromatograms with a low S/N ratio.

It is precisely the y-axis or retention time movements which create the non linearity of the ΣH plot, and they are used to advantage to provide this new application for signal averaging, or more accurately, signal summation. For this test peaks need not be trace and the S/N ratio can be large. For method inspection and validation, generating the ΣH plot is an occasional task so the time penalty is reduced.

3.2. The perfect chromatogram

If an analysis can be brought under such control that detector performance is maximized (or optimized) and a plot of ΣH versus number of chromatograms is linear, then it is proposed here that this is as good a working definition of the perfect chromatogram as is possible. It does not mean perfect in the sense of complete peak resolution or infinite S/N ratio, it means perfect in the quality control sense of zero defects and this definition would be valid under any laboratory regulation scheme. Moreover the definition is general and, apart from it being best applied to capillary peaks, it is independent of the type of chromatography.

The essential feature of the ΣH plot is its curvature which asymptotes towards linearity as complete instrumental control is achieved. Absence of curvature means that retention times and peak shapes are constant for the experiment; but even when precision is maximized, i.e. the standard deviation of retention time is minimized, there will still be a residual standard deviation and a residual curvature which cannot be reduced further without improving column or instrument specification. Plotting curvature makes these improvements easier to visualize and quantify, and the plot can distinguish between those modifications which bring benefit and those which do not.

In practice, the perfect chromatogram is unachievable because retention times and peak shape can never become absolutely stable, there will always be experimental error though the error (and imperfection) can become negligible. As instruments and techniques improve it will require more and more

replicate analyses to demonstrate non linearity in a ΣH plot. This trend becomes another measure of method quality.

3.3. Quantifying perfection and imperfections (non zero defects)

This working definition of perfection as a zero defect experiment and implying mainly capillary analyses, immediately lends itself to quantification in terms of plot curvature:

(a) A simple graphical measure is the shortfall in accumulated peak height expressed as a percentage of the real total of peak height as in Fig. 2. This measure must note the number of analyses involved, as a shortfall generated by many analyses is better than the same shortfall generated by only a few.

% Shortfall per added chromatogram

$= 1/M[h_1/(h_1 + h_2) \cdot 100\%]$ from Fig. 2

$= 1/M[\Sigma_M(y(t)_{avtr})/\Sigma_M H_{tr} \cdot 100\%]$

in which $\Sigma_M(y(t)_{avtr})$ = sum of M peak ordinates at average t_R, $\Sigma_M H_{tr}$ = sum of M separate peak heights, H_{tr}, and M = number of chromatograms.

The curvature of the graph can be seen as well as measured and this is a test where the analyst's judgement of linearity (or best achievable linearity) might be as useful as the numeric report. For speed and convenience, this graph and report would be a standard output from the data processor.

(b) More rigorously, it is possible by regression [7,8] to fit a polynomial to the accumulated data, i.e.

$$1/M[\Sigma_{i=1,2,..M}(y(t)_{avtr})] = 1/M[p_0 + p_1 Q_1 - \Sigma_{i=2,3,..M}(p_i Q^i)]$$

The intercept, p_0 should pass through the origin; if

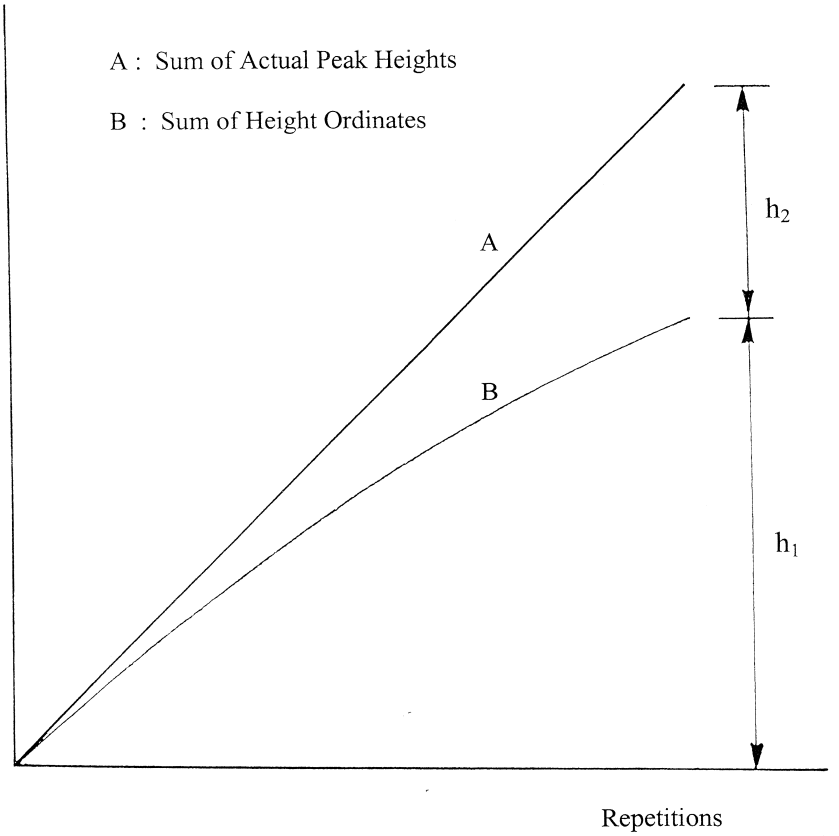

Fig. 2. ΣH curvature measured as a shortfall in accumulated peak height.

it does not, the value provides more diagnostic data about sample degradation, the detector or data processor.

The term p_1Q_1 is equal to the peak height measured at average t_R, alternatively it is the true peak height minus the average shortfall per peak.

All of the non linearity is contained in the $\Sigma_{i=2,3,..M}$ (p_iQ^i) term which is negative because the ΣH plot is less than linear.

The $1/M[\Sigma_{i=2,3,..M}$ $(p_iQ^i)]$ term is a measure of the imperfection of the analysis. The purpose of method development is to make p_0 and this term less than a user defined threshold and negligible.

It follows that if the ΣH plot is as near linear as possible and no more curvature can be removed from it by refining the experiment, the experiment is under as much control as it can be.

This does not mean that the analysis cannot be improved by using another column to provide better resolution, or by using a more sensitive detector to give bigger peak areas for the same injected quantity. It only means that the experiment cannot be improved with existing equipment and methods. In quality terms, method development has reached a plateau. Further improvements will come from using better components or technique and the improvement can be measured as a reduction in $1/M[\Sigma_{i=2,3,..M}$ $(p_iQ^i)]$.

3.4. Number of chromatograms in a ΣH plot

Ghaoui and Rothman produced a measurable curvature with only six chromatograms. The number of chromatograms required to produce a measurable curvature depends on the individual analysis, some are inherently more difficult than others.

When curvature measured as $1/M$ $[\Sigma_{i=2,3,..M}$ $(p_iQ^i]$ becomes small, by improved experiment design, more curvature can be introduced into the ΣH plot by increasing the number of chromatograms included, but the quantity $1/M$ $[\Sigma_{i=2,3,..M}$ $(p_iQ^i)]$ which can be interpreted as "$\Sigma_{i=2,3,..M}$ (p_iQ^i) per added chromatogram" will become smaller as each improvement is made. This trend of general improvement can be monitored over long periods by recording M and $1/M$ $[\Sigma_{i=2,3,..M}$ $(p_iQ^i)]$ separately. The more chromatograms required to produce a significant curvature, the better the experiment is performed.

3.5. Test sample

A general purpose sample, analogous to the Grob mixture [9], might be designed to include analyte species selected to test different parts of the system for their contributions to the non linearity. For example, analytes might be selected to provide

1. A molecular mass range to test injector discrimination
2. A polarity range to provide a range of peak asymmetries
3. A concentration range to monitor detector linearity (and to stay in range)
4. Trace components to measure the S/N improvement as Ghaoui and Rothman did

How complex this sample becomes will reflect the individual needs of a laboratory and the ingenuity of the sample maker. Too many components should be avoided if this creates peak overlap or makes the analysis time too long. The linear dynamic range of the detector must not be exceeded though inspection of the ΣA plots of the largest peaks will show if it has. Where possible and for efficiency, these tests might be integrated with existing system suitability tests.

3.6. Limitations

To ensure accurate and not merely precise peak measurement, and to overcome the limits of commercial data processors [11], the measured peaks in a sample (which is not necessarily all of the peaks) must be fully resolved. Peak symmetry is preferable but not essential; analysts should note however that if peak shape and asymmetry are variable within the experiment, the data processor must have adequate sampling frequency as peak sampling requirements change with peak shape [10–12].

3.6.1. Data measurement

The data processor must be capable of sampling the narrowest and most asymmetric peaks sufficiently as under-sampling leads to peak shape distortion

[11,12]. Other signal conditioning and measurement algorithms are embedded in the software and are unknown to the user; it can only be assumed that they measure peaks with accuracy and without distortion.

Not all data processors provide the procedures to allow signal or ensemble averaging of chromatograms, but it is an easy task for manufacturers to add them to those processors which do not. The changes are in software and will only incur a fixed development cost. An increase in data storage capacity may become necessary if more chromatograms are archived.

3.6.2. Detector linearity and sensitivity

The curvature of the ΣH plot is independent of the detector provided the detector is fully functional and working within its linear range. Detector sensitivity need not be maximized to be judged as having ''zero defects'' but it must be stable and sufficient. If there is a loss of sensitivity, if detection becomes non-linear, or if there is in-line sample or column degradation, it may contribute to the non linearity of the ΣH plot, but such events will also degrade the ΣA plot which is therefore retained for comparison. It is assumed here that curvature of the ΣH plot is measured while the ΣA plot remains linear. This is not unreasonable: it is an assumption that installation and maintenance related problems are absent, detector sensitivity is optimized, and the analyst is at the stage of fine-tuning his/her methods for best performance.

Carr and co-workers [13,14] have pointed out that the linear dynamic range for height measurements is generally less that that of area, and different species can have different linear dynamic ranges on the same detector. Thus the assumption that because the ΣA plot is linear, the detector is working inside the linear dynamic range of height is not infallible.

4. Conclusion

Using signal averaging to monitor system performance and stability is a new application, and uses peak sizes much larger than trace. It depends on the conversion of random fluctuations in retention time into a systematic reduction of measured peak height. This new application provides a scheme for monitoring improvements to a chromatographic experiment when methods are changed or the equipment is upgraded.

It is a means to selectively measure improvements to specific parts of the experiment, the injector, flow controllers or column, and is another validation test for the transfer of methods to different sites. Round robin tests might choose to include a measure of $1/M[\Sigma_{i=2,3,..M}\ (p_i Q^i)]$ from ΣH plots as an initial comparison of the labs involved.

Instrument manufacturers can quantify instrument improvements (especially to injectors) by measuring and reporting reductions in non linearity of the ΣH plot brought about by design improvements in their new systems. Column manufacturers might find peak shape degradation (seen as peak height reduction) a more sensitive and a quantifiable indicator of column suitability or degradation.

As a system performance test, signal averaging of repeated analyses may be too sensitive for some chromatographic analyses and more demanding than the analysis warrants, but it will undoubtedly demonstrate the experimental skills of those analysts who can achieve near zero defects and it introduces the quantitative idea that a chromatograph can be judged to be performing at its best.

References

[1] L. Ghaoui, L.D. Rothman, J. High Resolut. Chromatogr. 15 (1992) 36.
[2] P.B. Crosby, Quality is Free, McGraw Hill, New York, 1979.
[3] D.A. Skoog, Principles of Instrumental Analysis, 3rd Ed., Saunders, New York, 1985.
[4] E.R. Davies, Electronics, Noise and Signal Recovery, Academic Press, London, 1993.
[5] A. Fellinger, Data Analysis and Signal Processing in Chromatography, Data Handling in Science and Technology, Vol. 21, Elsevier, Amsterdam, 1998.
[6] R.Q. Thompson, J. Chem. Educ. 62 (1986) 866.
[7] R.L. Burden, J.D. Faires, Numerical Analysis, 6th Ed., Brookes/Cole, Pacific Grove, CA, 1997.
[8] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, L. Kaufman, Chemometrics: A Textbook, Elsevier, Amsterdam, 1988.
[9] K. Grob, G. Grob, K. Grob Jr., J. Chromatogr. 156 (1978) 1.

[10] D.T. Rossi, J. Chromatographic Sci. 26 (1988) 101.
[11] N. Dyson, Chromatographic Integration Methods (RSC Chromatography Monographs), 2nd Edition, Royal Society of Chemistry, Cambridge, 1998.
[12] N. Dyson, J. Chromatogr. A 842 (1999) 320.
[13] P.W. Carr, Anal. Chem. 52 (1980) 1746.
[14] L.M. McDowell, W.E. Barber, P.W. Carr, Anal. Chem. 53 (1981) 1373.