Investigations on the calculation of the third moments of elution peaks. I: Composite signals generated by adding up a mathematical function and experimental noise

Hong Gao [a,b], Paul G. Stevenson [b], Fabrice Gritti [b], Georges Guiochon [b,*]

[a] *Department of Physics, Anshan Normal University, Anshan, Liaoning 114005, PR China*
[b] *Department of Chemistry, University of Tennessee, Knoxville, TN 37996-1600, USA*

ABSTRACT

Elution peaks were generated by summing up a mathematical function and a previously recorded experimental noise. The first three statistical moments of these peaks were calculated using several data processing methods. The results obtained were analyzed. From this analysis, a method is developed for the calculation with a satisfying accuracy of the third moments of these combined signals. The method is applicable to real chromatographic peaks and makes it possible to determine accurate estimates of their third moments. This approach could be helpful for the investigation of the mass transfer processes in chromatographic columns because the third moment provides direct access to the external mass transfer coefficient.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

The detailed understanding of the different steps in the mass transfer process in chromatographic columns is a difficult and important problem that has attracted the attention of many researchers. The recent developments of the science and technology of chromatography [1–3] have improved our knowledge of the chromatographic mass transfer processes regarding the extra- and intra-particle mass transfer steps [4–17]. However, due to the complication of these processes and to the limited methods that can be used to directly explore them, our cognition of these processes is still limited, particularly for some of the steps involved for which the characteristic parameters are difficult to measure directly or still remain unknown for the lack of suitable models. A case in point is that of the external mass transfer coefficient which is usually determined according to empirical correlations [18–21].

Besides these correlations, two other methods are used to determine chromatographic mass transfer parameters. The first one consists in obtaining the values of these parameters through matching experimental peak profiles to those predicted by a theoretical model [22–28]. Another, more accurate method is provided by moment analysis. It is currently used only to determine the first two statistical moments, which characterize the retention time and the width of chromatographic peaks, respectively. The third moment characterizes peak symmetry, a property of peaks that expresses a different result of the mass transfer processes from that characterized by peak broadening. So, this moment provides a channel through which more information on the mass transfer processes can be obtained. Unfortunately, it remained ignored so far due to the poor accuracy of its measurement [29].

Although some investigations of the calculation of the statistical moments of chromatographic peaks have been published several decades ago [30–33], it was shown that they were not helpful for processing the experimental data acquired from the chromatographic systems available at the time. This was due in large part to the poor precision of the data produced by these instruments. More importantly, no practical method was available for the calculation of the higher statistical moments of chromatographic peaks with a satisfying degree of accuracy. Consideration of the third central moment was abandoned [1,34]. Recently, however, Gao and Lin showed that, using modern instruments, it was now possible to measure with a reasonable accuracy the third central moments of a band in a wide range of mobile phase flow rates and to derive from these measurements a reliable estimate of the external mass transfer coefficient [35–37]. However, only a limited range of experimental conditions was involved in that work. Therefore,

---

* Corresponding author. Tel.: +1 865 974 0733; fax: +1 865 974 2667.
  E-mail addresses: guiochon@utk.edu, guiochon@ion.chem.utk (G. Guiochon).

the requirements that must be followed to ensure the validity of this approach to measure the third moment should be established under a wider range of experimental conditions.

To validate the process of experimental measurements of third central moments, we developed a method for the calculation of realistically simulated data. Peak profiles are generated by combining a mathematical function and a segment of experimental baseline, previously recorded. This combination is processed as are real chromatographic peaks. Various methods of data acquisition and processing were adopted to calculate the third moments of these simulated signals. The performance of these methods can be estimated from the difference between the values of the third moments calculated with them and those provided by direct integration of the mathematical function. This is why we developed the method reported below, which permits an accurate calculation of the third moment of peak profiles, and its application to real chromatographic peaks.

2. Theory

2.1. Definition of the statistical moments of peak profiles

By definition, the nth central moment of an eluted band profile, μ_n, are given by the following equations [1]:

$$\mu_n = \frac{\int_0^\infty C(t)(t - \mu_1)^n \, dt}{\int_0^\infty C(t) \, dt} \tag{1}$$

$$\mu_1 = \frac{\int_0^\infty C(t)t \, dt}{\int_0^\infty C(t) \, dt} \tag{2}$$

where μ_1 is both the first moment of the peak and the average elution time of the band of solute, $C(t)$ is the concentration profile of the solute detected at the end of the column. The integrals in the denominators of these fractions ($\int_0^\infty C(t) \, dt$) or zero-order moment is the area of the peak; it is proportional to the amount of the sample injected in the column.

2.2. Fundamental importance of the third moment of chromatographic peaks

The set of mass balance equations of the general rate model of chromatography has no closed form algebraic solution, which makes difficult the interpretation of experimental results regarding the relationship between band broadening and column characteristics [1]. However, this equation can be solved in the Laplace domain, where the moments of the solution can be derived [38,39]. It is possible to transform these moments from the Laplace to the real domain. Algebraic relationships are thus provided between the coefficients of the general rate model and the different moments of the eluted band.

The equations of the general rate model are available in the literature(see [1] Chapter 6, Eqs. (6.58)–(6.64)). When the kinetics of adsorption–desorption on the stationary phase can be considered as fast enough for its contribution to band broadening to be neglected compared to the other contributions, the first three moments become:

• For the first moment:

$$\mu_1 = t_R = \frac{L}{u}(1 + k) + \frac{t_p}{2} \tag{3}$$

where t_R is the retention time of the band, k the retention factor of the compound considered ($k = F(\epsilon_p + (1 - \epsilon_p)K_a)$, F is the phase ratio, with $F = (1 - \epsilon_e)/\epsilon_e$, where ϵ_p and ϵ_e are the internal and external porosities of the column bed, respectively, K_a the adsorption equilibrium constant), L the column length, u the average interstitial velocity, and t_p the duration of injection of the sample.

• For the second moment:

$$\mu_2 = 2\frac{L}{u}\left[\frac{D_L}{u^2}(1 + k)^2 + \frac{k^2}{F}\left(\frac{d_p^2}{60 D_p} + \frac{d_p}{6k_f}\right)\right] + \frac{t_p^2}{12} \tag{4}$$

where D_L is the axial dispersion coefficient, combination of the dispersive effects of axial and eddy diffusion, D_p is the diffusion coefficient of the solute through the particles, d_p the average particle size, and k_f the external mass transfer coefficient.

• For the third moment:

$$\mu_3 = L\frac{12\epsilon_e D_L \left[\epsilon_e + (1 - \epsilon_e)J\right](1 - \epsilon_e)}{u^3\epsilon_e^3}\left(\frac{d_p J^2}{6k_f} + \frac{d_p^2 J^2}{60 D_p}\right)$$
$$+ L\frac{1 - \epsilon_e}{u\epsilon_e}\left(\frac{d_p^4}{420 D_p^2} + \frac{d_p^3}{30 k_f D_p} + \frac{d_p^2}{6k_f^2}\right)J^3$$
$$+ L\frac{12\epsilon_e^2 D_L^2}{u^5\epsilon_e^5}[\epsilon_e + (1 - \epsilon_e)J]^3 \tag{5}$$

with:

$$J = \epsilon_p + (1 - \epsilon_p)K_a \tag{6}$$

2.3. Calculation of the statistical moments from experimental chromatograms

Two main points must be carefully considered to perform accurate calculations of these statistical moments from the recorded chromatograms, the selection of the boundaries of the numerical integrals, which cannot extend to infinity in practice, and the stability of the baseline, which may drift during the recording of the peaks.

2.4. On the determination of the integral interval

According to the definitions provided above, exact values of the statistical moments should be obtained through integration of the elution band profiles from zero (the injection time) to an infinite time, which is practically impossible. In practice, all these integrations can be accomplished only within a finite integral interval. The determination of the most suitable integral intervals, especially that of the upper limit of integration for experimental peaks, which are always affected by a certain degree of tailing, becomes a problem [40].

Normally, the integrations for the moment calculations of mathematical functions are started and stopped at the points where the function is tangent to the baseline. But for experimental peaks that are always somewhat noisy, and especially for tailing peaks, the integral interval decided in this way may not be wide enough to provide an accurate value of the moments. This can be explained by Fig. 1 where three curves are shown. They represent a tailing peak profile generated by an exponential modified Gaussian (EMG) function and the integrands in Eqs. (1) and (2) which are integrated in the calculation of the second and the third central moments of the EMG function, respectively. The figure shows that the integration area under the curve $C(t)$ within the finite integral interval ending in the point indicated by the arrow 1 in the figure represents a good approximation of the zeroth moment. But, for the calculation of the second moment, the upper limit for the integration under the curve $C(t)(t - \mu_1)^2$ should not be less than the value indicated by the arrow 2 while for the integration of $C(t)(t - \mu_1)^3$, which gives the third moment, it should not be less than the time when the point indicated by the arrow 3, which includes the whole integral
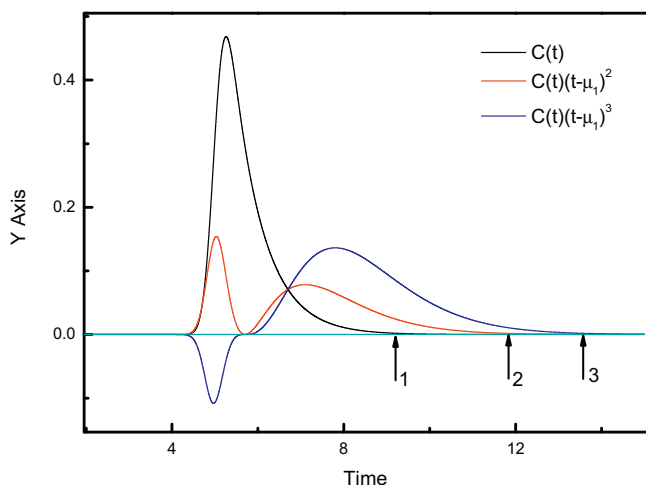
Fig. 1. Schematic of the determination of the integral interval for the integration of the second and third central moments of tailing peak profiles. The peak profile is generated by the exponentially modified Gaussian (EMG) function, where $\sigma = 0.2$, $t_g = 5.0$ s, $\tau = 0.7$, and $A = 0.54$.

area, is reached. Therefore, Fig. 1 shows that the integral intervals for the calculation of the moments of a tailing peak, and especially these intervals for higher moments, should be wide enough. Considering the curves of the integrands for the moment calculations, i.e. for $C(t)(t - \mu_1)^2$ and $C(t)(t - \mu_1)^3$ shown in Fig. 1, it would be most helpful to find an accurate procedure for the determination of the integral intervals.

It is important to note that, due to the formulas in Eqs. (1) and (2), the density plots of the contributions of the peak profile to its higher moments have two separate lobes with a null point between them. The practical consequence is that the central part of the elution profile, where the signal and the signal to noise ratio for experimental peaks are highest, contribute relatively little to the central moments while the wings of the profile have large contributions, which extend farther than the contributions of these wings to the zeroth and the first moments. This has considerable consequences for the handling of experimental results.

On the other hand, it is not true that the wider the integral interval, the more accurate the moments obtained. This is because any real chromatographic peak profile acquired through experiments is affected by a degree of baseline noise that cannot be avoided. Fig. 2 shows the effect of this noise on the integrands for the calculation of the first three moments of the peak profile. It is obvious that the baseline of the curve $t C(t)$ is clear and flat, which indicates that the noise has only a small effect on the result of the first moment integration. However, the effect of the noise for the curve $C(t)(t - \mu_1)^2$ can be seen near the ends of this curve, particularly on its tailing end. As for the curve $C(t)(t - \mu_1)^3$, the effect of the noise becomes important, even at locations that are not far remote from the center of the second lobe of this curve. Therefore, in order to reduce the influence of the baseline noise on the calculation results of higher moments, the integral interval for the integrand should not be excessively wide.

Here comes the conflict that we need to solve. The integral interval for moment calculation should be wide enough to reduce calculation error due to the loss of integration area, but it also should be sufficiently narrow to reduce or limit the contribution of the noise and the baseline drift. This conflict is most important for the calculation of the higher moment of tailing peaks. In part, the solution to this problem is in increasing the ratio of the signal to the noise of the peak profile while, at the same time, keeping the integral interval sufficiently wide to limit the truncation error.
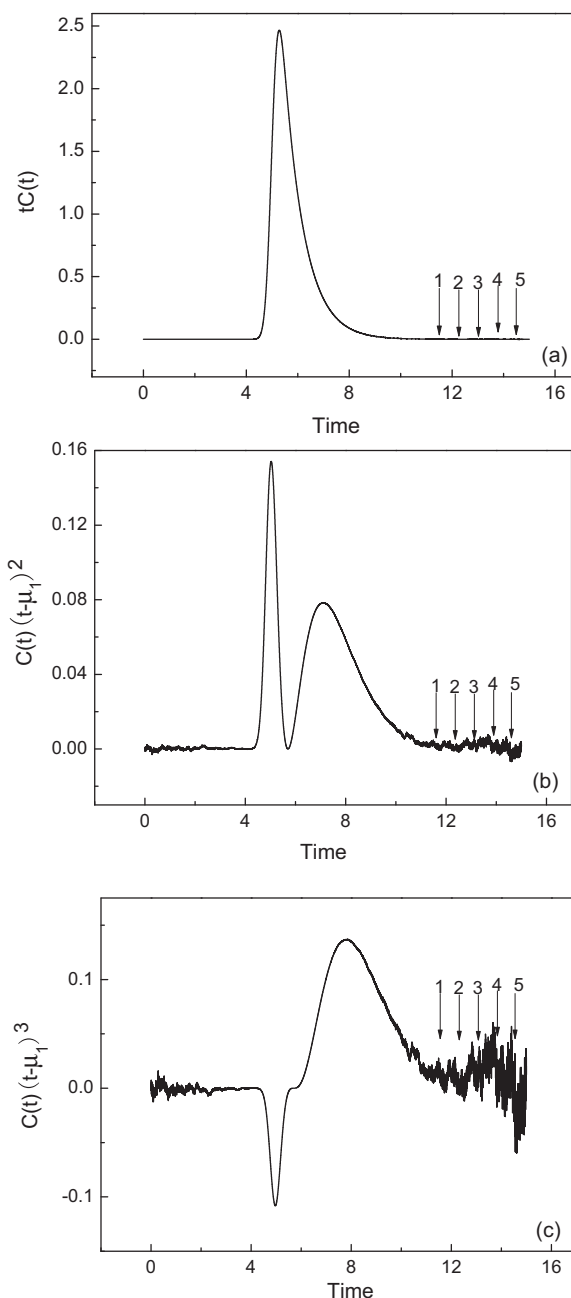

Fig. 2. Effect of the baseline noise on the integrands for calculations of the first three moments. The curves were generated from the same EMG function as those shown in Fig. 1 but a segment of noise was added (see details in Section 3). (a) Curve of integrand for calculating the first moment; (b) curve of integrand for calculating the second moment; and (c) curve of integrand for calculating the third moment.

Finally, besides reducing the signal noise, scientists should make sure that the signal is stable and that the baseline drift is limited. This renders important the methods for correcting the baseline drift. Different baseline levels will result in different integration area under the curve of the integrand and consequently will lead to different values of the moments. This specific problem is discussed in more detail in Section 4.

3. Calculations and experiments

The aim of this work is to estimate the accuracy with which the third moments can be calculated. For this purpose, we designed synthetic signals that combine peaks derived from mathematical

Table 1
Parameters used to generate EMG peaks and determine the integral intervals.

	A	t_g	σ	τ	w	n
Slight tailing peak	0.257	5.0	0.2	0.1	0.40	12
Strong tailing peak	0.540	5.0	0.2	0.7	0.15	42

functions and segments of noisy baseline recorded in conventional experiments. The moments of the mathematical functions are exactly known. The segments of baselines should not contribute to the moments but their noise and baseline drifts are realistic. The purpose of the exercise is to find a procedure to minimize their contributions to the calculated moments.

First, mathematical functions are selected to represent the chromatographic peak profile. Their statistical moments are accurately known; later in this work, these moment results are called the "real" moments of the (theoretical) profiles. Second, a segment of recorded baseline is added to the peak profiles produced by the mathematic function to simulate an experimental peak profile. Then, the statistical moments of the synthetic peaks are calculated; they are called the "simulated" moments. Finally, the accuracy of these simulated moments are calculated from the relative differences between them and the real moments. The precision is obtained from the relative standard deviations of a number of such calculations performed with the same mathematical function and different segments of the same, long baseline record.

3.1. Mathematical functions used to model peaks

The exponential modified Gaussian function was adopted in this work because it is the mathematical function most commonly used to represent chromatographic peaks [41,42]. It simulates rather well the influence of the sampling system, composed of empty tubes and small open vessels [1]. This function is written as follows:

$$C(t) = \frac{A}{2\tau} e^{\frac{\sigma^2}{2\tau^2} + \frac{t_g - t}{\tau}} \left[\operatorname{erf}\left(\frac{t - t_g}{\sigma\sqrt{2}} - \frac{\sigma}{\tau\sqrt{2}} \right) + 1 \right] \tag{7}$$

where A is the peak area, t_g is the retention time of the Gaussian component of the profile, σ its standard deviation and τ is the time constant of the exponential decay component of the real profile.

Two profiles with different extent of tailing but having the same height were used in this work. The values of the parameters used in the EMG function to generate these two profiles are listed in Table 1. The asymmetry factors at 10% peak height peaks are 1.09 and 3.08 which represents chromatographic peaks having a slight and a very strong degree of tailing, respectively.

3.2. Baseline and noise data acquisition

A very long section of baseline was recorded for 10 h using the UV-diode-array detector of a Waters Acquity UPLC System fitted with a Waters BEH C18 column (3.0 mm × 150 mm, average particle size 1.7 μm) flushed with the mobile phase at a flow rate of 0.8 mL/min. This mobile phase was composed of acetonitrile (HPLC Grade, Fisher Scientific, Fair Lawn, NJ, USA) and water (HPLC Grade, Fisher Scientific) in the volume ratio of 1:1. The inlet pressure of the instrument was slowly shifting between 866 and 875 bars while the temperature kept nearly constant close to 28 °C (27.992 °C to 28.007 °C) in these 10 h. The wavelength of the UV detector was set at 254 nm and the sampling rate was 40 points/s. The high flow rate and high inlet pressure adopted for the recording of the baseline signal were chosen to have a relatively high level of noise. Therefore, the baseline noise and drift were maximized. The recorded baseline noise is shown as Fig. 3.

The whole 10 h-baseline record was separated into ten sections lasting 15 min each. The locations of these sections are shown in
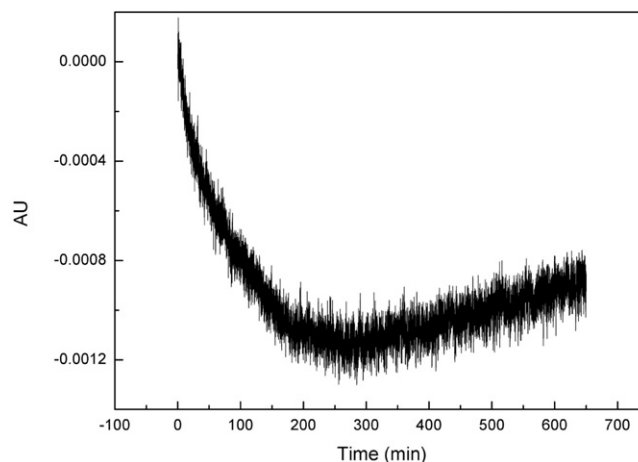

Fig. 3. Baseline recorded for 10 h. Details on the experimental conditions for this acquisition are provided in Section 3.2.

Table 2. Each section of noise was added to the peak profile generated by the EMG function, generating ten different synthetic peaks which simulate as many different experimental chromatograms.

The higher level of the signal to noise (S/N) ratios for these synthetic peaks considered in this work are provided in Table 2. The accurate measurement of the third central moment for practical chromatographic peaks is difficult, even for those having a high S/N ratio. The S/N ratios listed in Table 2 are defined as the peak height divided by the range of the noise (taken as four times the standard deviation of the noise [43] when the noise is measured for the baseline section considered, after its drift has been corrected). Because the heights of the two peak profiles are similar, with 0.46752 for the slightly tailing peak and 0.46825 for the strongly tailing one, similar S/N ratios are obtained after the same noise section is added to the theoretical functions.

3.3. Procedure for calculating the moments of synthetic peaks

Special methods were adopted for determining the integral intervals and different baseline processing methods are used in this work.

3.3.1. Determination of the integral intervals

The dilemma found when attempting to determine the integral interval, more specially the upper limit of the integration for calculating the higher moments was already stated in Section 2.4. This can be simply stated by writing that, in order to obtain reliable numerical values of the moments, we should choose an integral interval that is wide enough to provide a small calculation error due to the loss of integration area but which is, in the time, sufficiently narrow to limit the contribution of the noise and the baseline drift.

Therefore, the plot of $C(t)(t - \mu_1)^2 \sim t$ is a useful tool to determine the upper limit of the integrations for calculating the second and third moments. The integrations should end just before the moment when the whole area under the curve of $C(t)(t - \mu_1)^2 \sim t$ is involved, just as the position indicated by the second arrow in Fig. 1. Obviously, this rule cannot provide an exact value of the upper integration limit. This time fluctuates within a brief period, according to the subjective judgement of the operator. But it might result in only a small calculation error on the moments because the area that contributes mostly to the integrals was already included and it is much larger than the area variation due to a shift of the upper limit of integration.

This is clearly explained by the following moment calculation results. Let consider the situation shown in Fig. 2b. According to the

Table 2
Location of each section in the 10-h noise record and signal to noise ratios of synthetic peaks.

Number of noise section	Location of noise section (min)	Range of noise (AU)	S/N ratio (for slight tailing peaks)	S/N ratio (for strong tailing peaks)
1	0–15	4.54426E−4	2933	2937
2	60–75	3.03268E−4	3036	3041
3	120–135	2.82287E−4	3460	3465
4	180–195	3.04938E−4	2565	2569
5	240–255	3.10897E−4	2582	2586
6	300–315	3.32832E−4	2587	2591
7	360–375	3.33294E−4	2443	2447
8	420–435	2.78711E−4	2884	2889
9	480–495	2.75135E−4	2755	2760
10	540–555	3.23057E−4	2517	2521

previous analysis, the integrations to calculate the second and third moments can end later than but near the position indicated by the first arrow in 2b. To make clear the effect of a small shift of the upper limit of integration of the moments, the first three moments were re-calculated by ending their integrations at the positions indicated by arrows 1–5, respectively. During this calculation process, we used the third method for baseline correction which is discussed later in Section 4. The results are provided in Table 3. The relative decreases of these moments (second row in Table 3) are always less than 0.02, 2 and 10%, respectively. The moments calculated with different upper limits of integrations do not differ significantly from each other and all are within the range of acceptance. Similar arrows are marked in Fig. 2c, showing that the contribution of noise to the third moment increases when the integration ends later. So, although the upper limit of the integration may shift slightly, this shift is limited and causes only a small calculation error for the third moment.

In the following section, the effect of different methods of processing the baseline on the moments is discussed. The integral intervals for the moment calculations were determined after to the integrand plots of the EMG functions, before addition of the baseline segments, to avoid any effect of the noise. This means that the EMG functional peak itself and not the synthetic peak was considered for the determination of the integration interval. This permits a separation of the effects of the different baseline processing methods and those of peak integration, so the discussion assumes that the integral intervals for moment calculations were already determined properly.

For each group of moment calculations, for the ten synthetic peaks produced by adding ten sections of experimental noise to an EMG function peak, the same integral interval was used. The lower and the upper limits of integration are $t_g - nw\sigma$ and $t_g + n(1 - w)\sigma$, respectively, where n is an empirical coefficient [40] and w is a coefficient that corrects for peak asymmetry and for the shift between the peak apex and the center of the elution peak (w would be equal to 0.5 for a symmetrical peak). Different combinations of n and w were tried and a set of values was adopted so that the curve of the integrand (see Fig. 1) is placed near the center of the integral interval and the front and the rear wings of this curve are close to

the baseline. The values of the first moment used for calculating the curve $C(t)(t - \mu_1)^2 \sim t$ to determine the integral limits are the real first moments of the EMG functions. The values of the parameters w and n determined in this fashion are given in Table 1. The integral intervals for the slightly and the strongly tailing peaks are (4.04–6.44 min) and (3.74–12.14 min), respectively.

Although we are mostly concerned in this work with improving the calculating accuracy of the third moment, the integral interval for the moment calculation was determined mainly according to the integrand of the second moment. This is because the S/N ratio of the integrand of the second moment is higher than that of the third moment, which can be understood by comparing the plots in Fig. 2b and c. This approach is validated by the satisfactory results of the calculations of the third moments which are obtained in this way, as discussed later.

3.3.2. Baseline processing methods

In this work, we used the three following methods to handle the correction for the baseline drift in the calculations of the central moments of the synthetic peaks:

Method 1: There is no baseline correction in this method. After the integral intervals have been determined as described above, the moments are calculated directly according to Eqs. (1) and (2).

Method 2: There is a baseline correction in this method. The starting point of the integration is taken as the referential point, and the value of its y-coordinate is subtracted from the signal values of all the data points involved in the integration interval.

Method 3: This method also includes a baseline correction. Since the baseline segment used for the synthetic peak is known, it can be used to derive the baseline drift correction. First, two sections of this segment, those containing the initial and the final parts of the chromatograms and, hence, are outside of these integration limits, were fitted each one to a straight line, which was then subtracted from the part of the synthetic peak within the integral interval. This method eliminates the effect of the baseline shift on the results of the moment calculations.

Table 3
Relative difference of the fist three moments from their real values for the peak (shown in Fig. 2(a)) calculated with different upper limits of integration indicated by arrows 1–5, respectively.

	Relative difference of the fist three moments from their real values (%)				
	Integration ended in position of arrow 1	Integration ended in position of arrow 2	Integration ended in position of arrow 3	Integration ended in position of arrow 4	Integration ended in position of arrow 5
μ_1 (min)	−0.0235	−0.0158	−0.0107	−0.00316	0.00246
μ_2 (min^2)	−1.41	−0.960	−0.600	0.111	0.502
μ_3 (min^3)	−6.08	−3.98	−2.07	2.08	4.50

Table 4
Calculated first three moments for peak profiles generated by EMG functions.

	μ_1 (min)	μ_2 (min^2)	μ_3 (min^3)
Slight tailing peak	5.1	0.05	0.002
Strong tailing peak	5.7	0.53	0.686

4. Results and discussion

4.1. Values of moments calculated for the peak profiles generated by the EMG functions

The first three moments for the slightly and the strongly tailing peaks generated by the EMG functions were calculated according to the following equations [44]:

$$\mu_1 = t_g + \tau \tag{8}$$

$$\mu_2 = \sigma^2 + \tau^2 \tag{9}$$

$$\mu_3 = 2\tau^3 \tag{10}$$

The results are shown in Table 4 and illustrated by the heights of the white columns in Fig. 4.

4.2. Values of moments calculated without baseline correction in Method 1

The values calculated for the first three moments without baseline corrections (Method 1 of the baseline process) for the slightly and the strongly tailing synthetic peaks are shown in the first columns in Tables 5 and 6. They are also indicated visually by the deep grey columns in Fig. 4, left and right, respectively.

Obviously, the results obtained for the first moment are most accurate and precise. Although the results are slightly worse for the strongly than for the slightly tailing peaks, they are still quite satisfactory in this last case. This means that the baseline noise has little or nearly no influence on the result of calculations of the first moment when the (S/N) ratio satisfies the condition given in Table 2.

As for the second moment, the results of the calculations for the slightly tailing synthetic peaks are still fine but, in contrast, the results of these calculations for the strongly tailing peaks are much worse. The difference between the values of the real second moment of these peaks as generated by the EMG function and those of the simulated peaks is nearly 50%. Without baseline correction, the second moment of the strongly tailing peaks can certainly not be calculated accurately.

Finally, the values obtained for the third central moment without baseline correction are well beyond acceptance.

The values of the moments calculated above without any baseline correction are reported here to show the noticeable differences with their actual values, especially for the third moments. This illustrates the need for a baseline correction. Although the signal to noise ratios of the profiles discussed in this work are high, an accurate value of the moment cannot be obtained without including a baseline correction during the calculations. With the baseline correction methods used here, the accuracy of the moments is much improved when even a very simple processing method is adopted, as shown in the following discussions.

4.3. Values of moments calculated with the baseline correction in Method 2

The second method for baseline processing used in this work is only an expedient because it only corrects the position of the baseline level roughly. Yet, its effect is impressive, as illustrated in the second columns of Tables 5 and 6 and in Fig. 4 left and right.

Compared with the results obtained without baseline correction, the relative standard deviation (RSD) of the second moments of the synthetic peaks after baseline correction decreases nearly ten times for either slight or strong tailing peaks (see Tables 5 and 6). The lower relative differences (RD) between the second moments of the peak generated by the EMG function and the average of those calculated for the synthetic peaks also indicate the marked improvement of the calculation accuracy of the second moments of synthetic peaks obtained with this method.

Most importantly, reasonable result can now be obtained for the third moments. For the slightly tailing synthetic peaks, the RSD and the RD between the third moments average and the third moment of the EMG function are all lower than 10%, which is a considerable improvement. As for the strongly tailing synthetic peaks, due to the increase extent of their tailing, their calculated third moments are seriously scattered but the RD between their average and the "real" third moment of the EMG function is still less than 10%.

4.4. Values of moments calculated with the baseline correction in Method 3

The third method for processing the baseline of synthetic peaks is classical. The baseline level as well as the baseline shift (slope) is corrected with the method used. However, the drift is assumed to be linear, which remains approximate.

The first three moments calculated for the synthetic peaks with this method are given in the third columns of Tables 5 and 6 for the slightly and the strongly tailing peaks, respectively. As for the slightly tailing peaks, there is only a slight improvement of the results when Method 3 is used instead of Method 2, because the results provided by Method 2 were already satisfying.

For the strongly tailing peaks, in contrast, the results are markedly improved. The RSD of the second and third moments and also the Relative Difference between the averages of the higher central moments and those of the EMG function are almost ten times less than the corresponding results obtained with Method 2.

4.5. Development of a practical method for baseline correction

According to the analysis of the moment results obtained with the various baseline processing methods, it is clearly necessary to find a proper baseline correction method for accurately calculating the higher moments. Method 3 seems to be optimum, which is easily understood because with this method the effect of an uncorrected baseline drift can be removed almost totally. The integral intervals are properly determined, which decreases the influence of the noise. Then, satisfactory results can be calculated for the different central moments of the synthetic peaks because the two main sources of errors in the calculation of these moments are reduced as much as possible.

A proper baseline correction method applicable to actual experimental chromatographic peaks must now be developed, based on the results of this analysis. Although Method 3 is sufficiently efficient for synthetic peaks, it cannot be applied directly to real experimental peaks because the baseline noise cannot be extracted from a real chromatogram and no linear fitting result can be obtained. But consideration of this method is a useful reference basis for developing the new method needed.

The baseline correction should adjust for the baseline level and for its slope. The following method was developed to accomplish this goal. First, two sections of the recorded peak profile are selected for the baseline correction. These sections should not be too narrow nor too wide, nor too far remote from the peak center, otherwise a satisfactory baseline correction cannot be obtained for the peak considered. We selected two sections in the following way: the distance between the initial point of the first baseline section selected
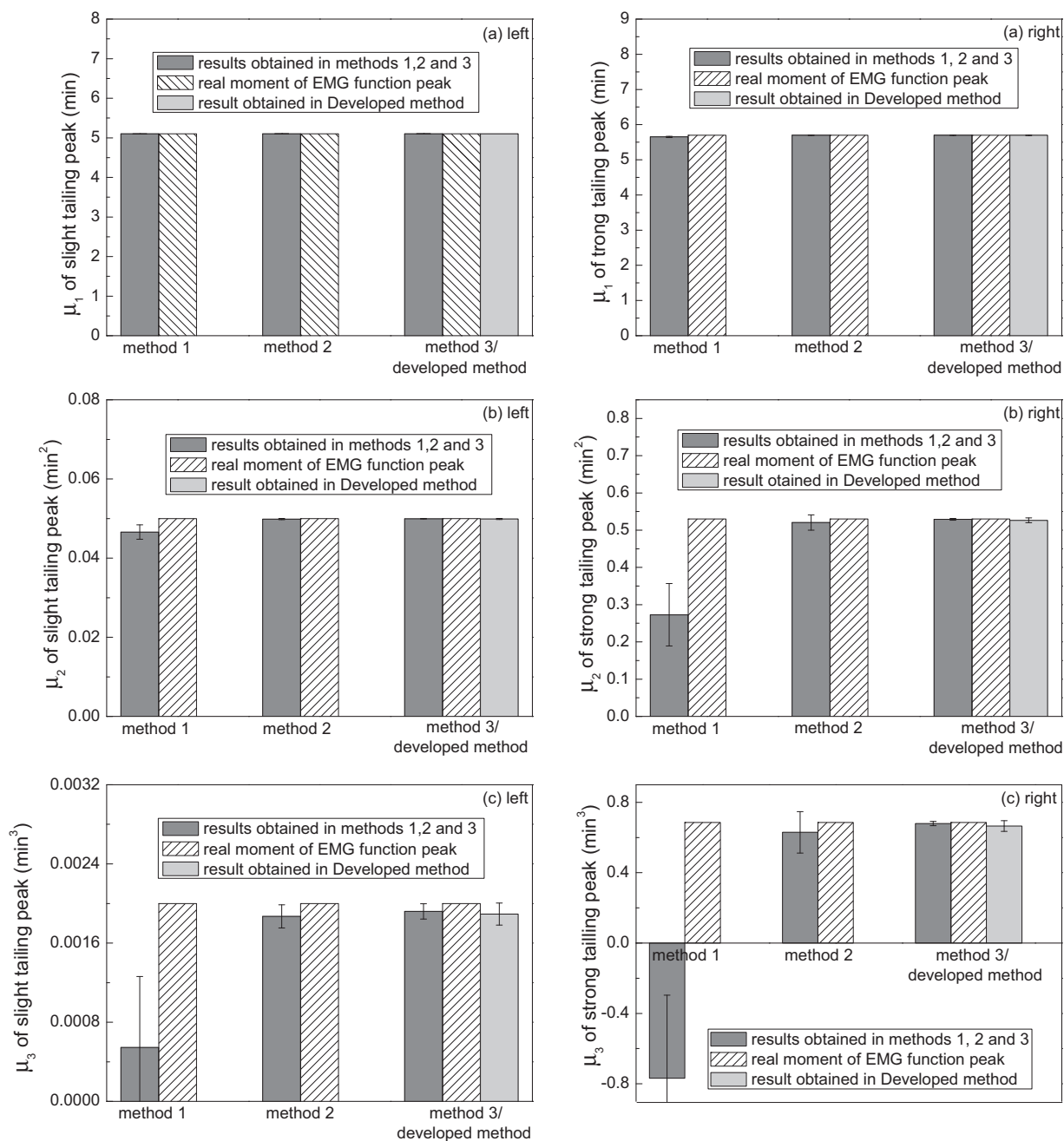
Fig. 4. (Left) Graphs comparing the moments of the weakly tailing synthetic peak and obtained using different baseline processing methods. (a) For the first moment; (b) for the second moment; and (c) for the third moment. (Right) Same for the strongly tailing peak.

and the lower limit of integration for the moment calculation is equal to the distance between the upper limit of integration and the end point of the second baseline section and the lengths of these two sections is equal to half the width of the integration interval. Then the data points recorded in these two sections are separately fitted to a straight line. The straight line joining the middle of these last two lines provides the baseline correction, which is finally subtracted from the peak profile.

To verify the validity of this method of baseline correction, the same method as just described was applied to the simulated peaks, i.e. to the baseline segment and the EMG function added to it and the simulated moments wer calculated.

Before these calculations, the proper integral interval is determined according to the method described in Section 3.3.1. However, this time, the synthetic peak, not the EMG function, was

the integrand for the second moment calculations (as shown in Fig. 2b). The integration interval was determined to ensure that the complete profile to be integrated was included within the integral interval while as much as possible of the noise part that might strongly influence the result was excluded. The value of the first moment used for determining the integral interval was calculated according to Eq. (2). The upper and lower limits of integration are determined easily as long as the complete area of the peak is involved, because, as shown above, it is easy to obtain accurate estimates of the first moment.

The results are shown in Table 7. They are satisfactory. They are also illustrated in Fig. 4 to compare them with the values of the moments obtained with the other baseline processing methods. It is clear that the accuracy of the moment calculations obtained with the developed method is similar to that achieved with Method 3.

Table 5
Moments' result for synthetic peaks of slight tailing with their baseline processed in first three methods.

	Method 1			Method 2			Method 3		
	μ_1 (min)	μ_2 (min^2)	μ_3 (min^3)	μ_1 (min)	μ_2 (min^2)	μ_3 (min^3)	μ_1 (min)	μ_2 (min^2)	μ_3 (min^3)
1 [a]	5.100	0.04967	0.001722	5.100	0.04960	0.001673	5.100	0.04995	0.001872
2 [a]	5.099	0.04735	8.086E−4	5.100	0.04974	0.001808	5.100	0.05003	0.001897
3 [a]	5.099	0.04628	4.947E−4	5.100	0.04996	0.001884	5.100	0.05003	0.002019
4 [a]	5.099	0.04532	1.228E−4	5.100	0.04989	0.001949	5.100	0.04992	0.001986
5 [a]	5.099	0.04510	−9.981E−5	5.100	0.04981	0.001841	5.100	0.04992	0.001859
6 [a]	5.100	0.04992	0.001905	5.100	0.04992	0.001905	5.100	0.04995	0.001924
7 [a]	5.098	0.04517	−1.084E−5	5.100	0.04978	0.001862	5.100	0.04981	0.001882
8 [a]	5.099	0.04553	2.320E−4	5.100	0.05013	0.002089	5.100	0.04999	0.002030
9 [a]	5.099	0.04564	1.626E−4	5.100	0.04999	0.001939	5.100	0.04996	0.001936
10 [a]	5.099	0.04570	1.023E−4	5.100	0.04970	0.001733	5.100	0.04984	0.001786
$\overline{\mu}_i$ [b]	5.099	0.04657	5.440E−4	5.100	0.04985	0.001868	5.100	0.04994	0.001919
RSD (%) [c]	0.01060	3.910	132.0	0.003400	0.3110	6.240	0.002370	0.1410	3.980
$(\mu_i - \mu_{i,real})/\mu_{i,real} \times 100\%$ [d]	−0.0229	−6.860	−72.80	−0.002350	−0.3000	−6.500	−0.001960	−0.1200	−4.000

[a] Number of noise section being added to generate synthetic peak.
[b] Average value of moment. $\mu_i (i = 1, 3)$ represents the first, second and third moment, respectively.
[c] Relative standard deviation of moments for synthetic peaks.
[d] Relative difference between the calculated moments of peaks generated by EMG functions ($\mu_{i,real}$) and those of synthetic peaks.

Table 6
Moments' result for synthetic peaks of strong tailing with their baseline processed in first three methods.

	Method 1			Method 2			Method 3		
	μ_1 (min)	μ_2 (min^2)	μ_3 (min^3)	μ_1 (min)	μ_2 (min^2)	μ_3 (min^3)	μ_1 (min)	μ_2 (min^2)	μ_3 (min^3)
1	5.691	0.4796	0.3926	5.691	0.4810	0.4004	5.700	0.5293	0.6813
2	5.666	0.3526	−0.3108	5.692	0.4894	0.4582	5.700	0.5296	0.6811
3	5.652	0.2754	−0.7554	5.696	0.5086	0.5607	5.700	0.5295	0.6796
4	5.642	0.2231	−1.052	5.701	0.5340	0.7047	5.699	0.5252	0.6554
5	5.639	0.2125	−1.103	5.701	0.5377	0.7351	5.700	0.5284	0.6838
6	5.639	0.2108	−1.123	5.700	0.5306	0.6859	5.701	0.5331	0.7005
7	5.642	0.2278	−1.024	5.701	0.5331	0.7009	5.700	0.5307	0.6869
8	5.645	0.2432	−0.9328	5.702	0.5394	0.7389	5.700	0.5307	0.6884
9	5.645	0.2422	−0.9412	5.699	0.5249	0.6551	5.700	0.5275	0.6698
10	5.649	0.2610	−0.8304	5.699	0.5249	0.6582	5.699	0.5271	0.6681
$\overline{\mu}_i$	5.651	0.2728	−0.7681	5.698	0.5204	0.6298	5.700	0.5291	0.6795
RSD (%)	0.2833	30.69	−61.41	0.06774	3.949	18.72	0.007108	0.4139	1.840
$(\mu_i - \mu_{i,real})/\mu_{i,real} \times 100\%$	−0.8586	−48.52	−212.0	−0.03123	−1.817	−8.194	−0.002281	−0.1679	−0.9490

Table 7
Re-calculated moments of the synthetic peaks which considered as the representation for real chromatographic peaks.

	Slight tailing synthetic peaks			Strong tailing synthetic peaks		
	μ_1 (min)	μ_2 (min^2)	μ_3 (min^3)	μ_1 (min)	μ_2 (min^2)	μ_3 (min^3)
1	5.100	0.04984	0.001760	5.698	0.5257	0.6633
2	5.100	0.05000	0.001820	5.701	0.5326	0.6950
3	5.100	0.04988	0.001980	5.698	0.5241	0.6529
4	5.100	0.04985	0.001920	5.701	0.5336	0.6955
5	5.100	0.04984	0.001930	5.695	0.5159	0.6177
6	5.100	0.04972	0.001870	5.697	0.5257	0.6639
7	5.100	0.04980	0.001840	5.703	0.5367	0.7147
8	5.100	0.05015	0.002140	5.698	0.5258	0.6685
9	5.100	0.04978	0.001770	5.699	0.5199	0.6297
10	5.100	0.04990	0.001900	5.695	0.5242	0.6528
$\overline{\mu}_i$	5.100	0.04988	0.001893	5.698	0.5264	0.6654
RSD (%)	0.003000	0.2440	5.910	0.04299	1.204	4.507
$(\mu_i - \mu_{i,real})/\mu_{i,real} \times 100\%$	−0.002157	−0.2400	−5.500	−0.02737	−0.6755	−3.001

This means that the developed method is valid since it properly corrects the baseline level and slope.

Finally, we note that all the values calculated for the moments of the synthetic peaks and listed in the Tables are less than the true, real moments of the EMG functional peaks. This is due to the limited integral intervals that must be adopted for the synthetic peaks. After the noise sections were added, the integral intervals must be shortened to eliminate the contributions of the parts of the integrands that are influenced most strongly by the noise, Consequently, parts of the profiles that should be involved are lost.

However, adopting the proper method of determination of the integration intervals provides low values of the Relative Difference of the moments. Even for the third moment of the strongly tailing peak, the relative error made is still below 10%.

5. Conclusion

The results of our work show that it might be possible to calculate estimates of the third central moments of chromatographic peaks recorded for the purpose of physical chemistry

investigations of mass transfer mechanisms. When proper attention and care are paid to achieve baseline stability and a reliable instrument with advanced computer control is used, the values obtained may be sufficiently accurate to permit the interpretation of the results using the equations developped by Kubin [38] and Kucera [39] in their works on the solution of the general rate model [1]. We generated peak profiles by adding up segments of realistic baseline signal recorded with noise and baseline drift to EMG functional peaks having different degrees of tailing. We were able to calculate the first moment and the second and third moments of these profiles with a precision of 0.002, 0.14, and 4% for the first second, and third moments respectively, and with systematic errors of 0.002, 0.12, and 4%, all three errors being negative.

Based on these results, we developed a protocol for measuring these moments for actual peaks in true chromatograms, provided that they are well resolved from other components of the sample. This method involves:

- Determining the integral intervals for the moment calculation. The curve $C(t)(t - \mu_1)^2 \sim t$ is an important tool to determine the proper integral intervals for the moment calculation. The beginning and ending parts of the peak, which contribute significantly to the second and third moments and are greatly influenced by the baseline noise due to their low (S/N) ratio, can be identified more clearly in this curve than in the peak profile itself. The loss of part of the integral area for moment calculation can be reduced and the effect of the noise can be excluded
- Correcting for the baseline drift. The baseline level and its slope were corrected linearly in this work. This correction is easy to perform in practice and its effect in improving the moment accuracy is remarkable, as shown in the foregoing discussion.

This method was validated by applying it to the generated peaks. With this method, the calculated values of the third central moments of real chromatographic peaks which have a moderate degree of tailing and an (S/N) ratio greater than 2500 should be accurate within better than 10%. Even with strongly tailing peaks, a similar result is possible.

Since we now have a valid method to determine the third moment of elution bands with a satisfactory accuracy, we plan to systematically apply this method to measure the value of the coefficient of external mass transfer of modern HPLC columns and to study its variations with the experimental conditions, trying to shed some needed light on the nature and properties of this poorly known parameter, and following the pioneering work of Gao and Lin [35–37].

References

[1] G. Guiochon, A. Felinger, A. Katti, D. Shirazi, Fundamentals of Nonlinear and Preparative Chromatography, Elsevier, Amsterdam, The Netherlands, 2006.
[2] F. Gritti, G. Guiochon, J. Chromatogr. A 1218 (2011), doi:10.1016/j.chroma.2011.04.058.
[3] F. Gritti, G. Guiochon, J. Chromatogr. A 1218 (2011), doi:10.1016/j.chroma.2011.07.014.
[4] F. Gritti, G. Guiochon, AIChE J. 57 (2011) 346–358.
[5] F. Gritti, G. Guiochon, J. Chromatogr. A 1218 (2011) 896–906.
[6] F. Gritti, G. Guiochon, AIChE J. 57 (2011) 333–345.
[7] F. Gritti, G. Guiochon, J. Chromatogr. A 1218 (2011) 907–921.
[8] F. Gritti, G. Guiochon, J. Chromatogr. A 1217 (2010) 8167–8180.
[9] F. Gritti, G. Guiochon, J. Chromatogr. A 1217 (2010) 5069–5084.
[10] F. Gritti, G. Guiochon, J. Chromatogr. A 1217 (2010) 5137–5151.
[11] L. Asnin, K. Horváth, G. Guiochon, J. Chromatogr. A 1217 (2010) 1320–1331.
[12] F. Gritti, G. Guiochon, J. Chromatogr. A 1216 (2009) 4752–4767.
[13] F. Gritti, G. Guiochon, Anal. Chem. 81 (2009) 2723–2736.
[14] K. Kaczmarski, G. Guiochon, Anal. Chem. 79 (2007) 4648.
[15] K. Miyabe, Y. Matsumoto, G. Guiochon, Anal. Chem. 79 (2007) 1970–1982.
[16] A. Cavazzini, F. Gritti, K. Kaczmarski, N. Marchetti, G. Guiochon, Anal. Chem. 79 (2007) 5972–5979.
[17] F. Gritti, G. Guiochon, Anal. Chem. 78 (2006) 5329–5347.
[18] K. Miyabe, N. Ando, T. Nakamura, G. Guiochon, Chem. Eng. Sci. 65 (2010) 5950.
[19] K. Miyabe, M. Ando, N. Ando, G. Guiochon, J. Chromatogr. A 1210 (2008) 60–67.
[20] K. Miyabe, Y. Kawaguchi, G. Guiochon, J. Chromatogr. A 1217 (2010) 3053.
[21] F. Gritti, G. Guiochon, Chem. Eng. Sci., submitted for publication.
[22] L. Hong, A. Felinger, K. Kaczmarski, G. Guiochon, Chem. Eng. Sci. 59 (2004) 3399–3412.
[23] K.K.H. Kim, G. Guiochon, Chem. Eng. Sci. 61 (2006) 4249–4267.
[24] P. Persson, H. Kempe, G. Zacchi, B. Nilsson, Chem. Eng. Res. Des. 82 (2004) 517.
[25] P. Persson, H. Kempe, G. Zacchi, B. Nilsson, Process Biochem. 40 (2005) 1649.
[26] C. Chang, A.M. Lenhoff, J. Chromatogr. A 827 (1998) 281.
[27] M.A. Fernandez, G. Carta, J. Chromatogr. A 746 (1996) 169.
[28] M.A. Fernandez, W.S. Laughinghouse, G. Carta, J. Chromatogr. A 746 (1996) 185.
[29] C. Vidal-Madjar, G. Guiochon, J. Chromatogr. 142 (1977) 61.
[30] H.C. Smit, H.L. Walg, Chromatographia 9 (1976) 483.
[31] S.N. Chesler, S.P. Cram, Anal. Chem. 43 (1971) 1922.
[32] S.N. Chesler, S.P. Cram, Anal. Chem. 44 (1972) 2240.
[33] D.J. Anderson, R.R. Walters, J. Chromatogr. Sci. 22 (1984) 353.
[34] B. Lin, G. Guiochon, Modeling for Preparative Chromatography, Elsevier, Amsterdam, The Netherlands, 2003.
[35] H. Gao, X. Wu, L. Bingchang, J. Chromatogr. Sci. 48 (2010) 478.
[36] H. Gao, X. Wu, L. Bingchang, J. Chromatogr. Sci. 48 (2010) 742.
[37] H. Gao, B. Lin, Comp. Chem. Eng. 34 (2010) 277–285.
[38] M. Kubin, Coll. Czech Chem. Commun. 30 (1965) 2900.
[39] E. Kučera, J. Chromatogr. 19 (1965) 237.
[40] P.G. Stevenson, G. Guiochon, J. Chromatogr. A 1218 (2011) 8255–8263.
[41] V.B.D. Marco, G.G. Bombi, J. Chromatogr. A 931 (2001) 1.
[42] E. Grushka, Anal. Chem. 44 (1972) 1733.
[43] A. Felinger, Data Analysis and Signal Processing in Chromatography, Elsevier, Amsterdam, The Netherlands, 1998.
[44] W.W. Yau, Anal. Chem. 49 (1977) 395.