Anal. Chem. **2002,** *74,* 2293−2305

Statistical-Overlap Theory of Column Switching in Gas Chromatography: Applications to Flavor and Fragrance Compounds

Clint Samuel[†] and Joe M. Davis*

Department of Chemistry and Biochemistry, Southern Illinois University at Carbondale, Carbondale, Illinois 62901-4409

First-column gas chromatograms (GCs) of hundreds of flavor and fragrance compounds, and second-column GCs of specific regions of these GCs, are predicted using thermodynamic databases in commercial software. A statistical-overlap theory of column switching with cryogenic focusing then is developed by mimicking the predicted GCs by two kinds of Monte Carlo simulations. In the first kind, a probability distribution is calculated for the number of compounds in a region of the first-column GC, based on the number of observed peaks in the region, the number of observed peaks in the second-column GC, and the retention-time distributions and breadths of single-component peaks in both GCs. In the second kind, criteria are established for the theory's application. The theory is applied to 12 regions of first-column GCs. The theory predicts the number of compounds in all of them and shows that separation rarely is complete in second-column GCs, when 10 or more compounds are transferred between columns. The theory also rationalizes the tedious search required to find good separation conditions by showing that column-switching gas chromatography with cryogenic focusing is inherently statistical. The number of peaks in the second-column GC can be greater than, less than, or equal to the number of peaks in the relevant region of the first-column GC, and the good conditions sought by researchers to substantially improve separation correspond to favorable "rolls of the dice" found only by trial and error.

Multidimensional separations entail the subsequent alteration or enhancement of a separation by one or more columns of different selectivity.[1] One type of multidimensional separation is column switching,[1,2] in which effluent spanning part of the first separation is transferred to a second column. In gas chromatography, the transfer usually is made with a pressure-balancing Deans switch[3] or fast multiport rotary valve. Analysis is enhanced by augmenting the separation with hyphenated detection and computer-based identification.[4,5]

Column switching typically is applied to complex separations, in which some single-component peaks (i.e., the Gaussian-like profiles of pure compounds) overlap. The amount of overlap was quantified previously by statistical-overlap theory (SOT).[6−10] In SOT, the interval between successive single-component peaks (SCPs) is modeled by a random variable, and the breadth of SCPs usually is expressed by deterministic attributes. Overlap is gauged by calculating the probability that the interval between successive SCPs is less than the average SCP breadth.

The hypothesis investigated here is that the second separation in column-switching gas chromatography with cryogenic focusing rarely is complete, if more than 10 compounds are transferred, because a nonnegligible probability of overlap is realized. To test the hypothesis, we use SOT to calculate the conditional probability $P(M_1 \mid p_1, p_2)$ that M_1 compounds are in the transferred region of the first separation, given that p_1 peaks are observed in the region without column switching and p_2 peaks are observed in the second separation.

This group has published an analytical theory of $P(M_1 \mid p_1, p_2)$.[11] However, the theory was too simple to characterize separations because the number density and breadths of SCPs were considered to be constant. In reality, more SCPs are found in specific regions of separations than in others, and SCP breadths vary throughout separations. The analytical modifications of theory required to address these concerns appear to be difficult.

Recently this group has confirmed predictions of SOT based on a Poisson interval distribution by simulating complex gas chromatograms (GCs) with GC software. The predictions usually were excellent on nonpolar stationary phases[12] and were acceptable on polar ones.[13] The software also assisted us in justifying

* Corresponding author. E-mail: davis@chem.siu.edu. Phone: 618 453 6464. Fax: 618 453 6408.

[†] Current address: Department of Computer Science, Southern Illinois University at Carbondale, Carbondale, IL 62901.

(1) Bertsch, W. In *Multidimensional Chromatography: Techniques and Applications,* Chromatographic Science Series 50; Cortes, H. J., Ed.; Marcel Dekker: New York, 1990; pp 74−144.
(2) Schomburg, G. *J. Chromatogr.* **1995**, *703*, 309−325.
(3) Deans, D. R. *Chromatographia* **1968**, *1*, 18−22.
(4) Wilkins, C. L. *Anal. Chem.* **1994**, *66*, 295A−301A.
(5) Ragunathan, N.; Krock, K. A.; Klawun, C.; Sasaki, T. A.; Wilkins, C. L. *J. Chromatogr., A* **1995**, *703*, 335−382.
(6) Davis, J. M. In *Advances in Chromatography*; Brown, P., Grushka, E., Eds.; Marcel Dekker: New York, 1994; Vol. 34, pp 109−176.
(7) Felinger, A. In *Advances in Chromatography*; Brown, P., Grushka, E., Eds.; Marcel Dekker: New York, 1998; Vol. 39, pp 201−238.
(8) Dondi, F.; Bassi, A.; Cavazzini, A.; Pietrogrande, M. C. *Anal. Chem.* **1998**, *70*, 766−773.
(9) Felinger, A. *Data Analysis and Signal Processing in Chromatography*; Elsevier: Amsterdam, 1998; pp 331−408.
(10) Felinger, A.; Pietrogrande, M. C. *Anal. Chem.* **2001**, *73*, 619A−626A.
(11) Peters, M.; Davis, J. M. *Am. Lab.* **1998**, *30*, 18, 32C−38C.
(12) Samuel, C.; Davis, J. M. *J. Chromatogr., A* **1999**, *842*, 65−77.
(13) Samuel, C.; Davis, J. M. *J. Microcolumn Sep.* **2000**, *12*, 211−225.

10.1021/ac011235b CCC: $22.00
Published on Web 04/17/2002

the thermodynamic basis of SOT in programmed temperature gas chromatography.[14]

The successful application of SOT to predicted GCs stimulated us to use such GCs as templates, from which to develop a realistic SOT of column switching. We report here results based on distributions $P(M_1 \mid p_1, p_2)$ calculated from predicted GCs of flavors and fragrances, which have been studied by column-switching,[4,15–23] coupled-column,[24–27] and two-dimensional[28–31] gas chromatography. In the templates, we easily varied the number density and breadth of SCPs and avoided the simplifications noted above.

THEORY

The theory underlying calculation of $P(M_1 \mid p_1, p_2)$ is explained by example. Figure 1a is a predicted GC of $m = 94$ flavors and fragrances on the nonpolar stationary phase, 100% methylsilicone. The retention times and standard deviations (i.e., breadths) of the 94 Gaussian SCPs were predicted by GC software; their amplitudes were assigned arbitrarily by us. Only $p = 64$ peaks are observed because of overlap. Region 3, which spans 10−12 min, contains $m_1 = 24$ SCPs and $p_1 = 14$ peaks. We envisioned the transfer of this region, with prior focusing of all SCPs, to a second GC column having a polar poly(ethylene glycol) stationary phase. Figure 1b is the resultant GC predicted by the GC software. The SCP amplitudes are the same as in region 3, as would be the approximate case if matched detectors were used. The separation is improved markedly, as indicated by $p_2 = 22$ peaks, but still is incomplete ($m_1 = 24$).

To develop a SOT, these and other GCs are mimicked using two kinds of Monte Carlo simulations. In the first—which we call Monte Carlo simulations of the first kind—SCP breadths are represented by line segments, and the Monte Carlo equivalents of m_1, p_1, and p_2 are counted to evaluate $P(M_1 \mid p_1, p_2)$. In the second—which we call Monte Carlo simulations of the second kind—SCP breadths are represented by Gaussian profiles, and the numbers of peak maximums are counted.

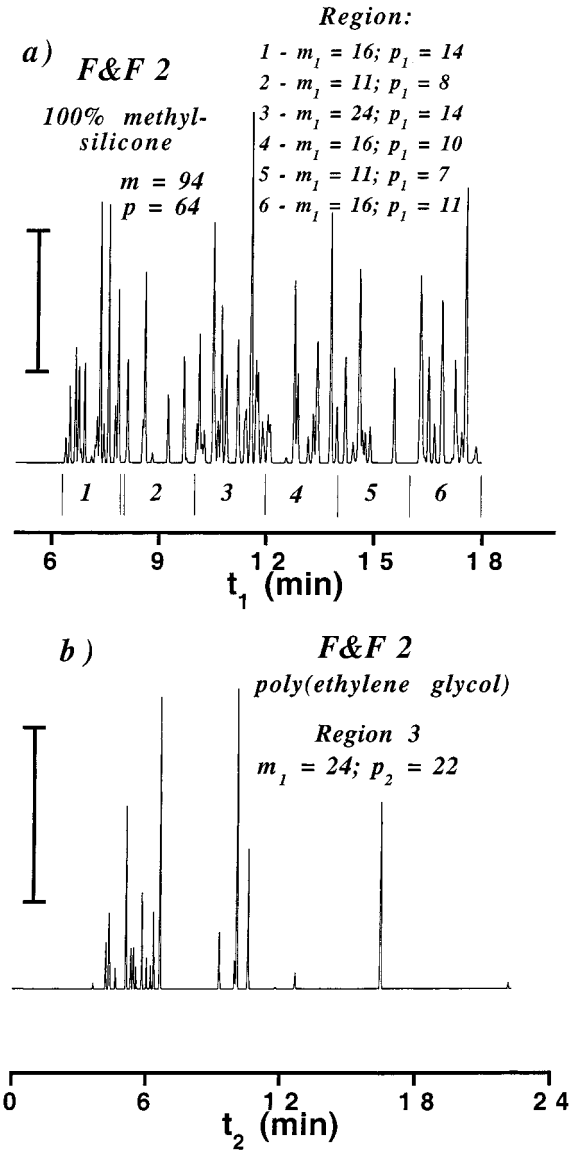

Figure 1. (a) First-column GC of flavors and fragrances, F&F 2, on 100% methylsilicone. GC is partitioned into six column-switched regions. (b) Second-column GC of region 3 on poly(ethylene glycol). Vertical bars indicate same response intensity.

To avoid confusion in the subsequent discussion, the letters "GC" will refer to elution profiles predicted from thermodynamic databases in GC software. All GCs are predicted; none is experimental. In such profiles, both SCP standard deviations and retention times are deterministic attributes, whereas SCP amplitudes are random attributes. Panels a and b in Figure 1 are examples of such GCs. In contrast, the word "simulation" will refer to the two Monte Carlo mimics of GCs described in the preceding paragraph. In both simulation types, standard deviations are deterministic attributes whereas retention times are random ones.

Monte Carlo Simulation of the First Kind: First-Column GC. The amount of overlap in the first-column GC depends in part on the distribution of compounds' retention times. Figure 2a is the histogram of the $m = 94$ retention times of compounds in the first-column GC in Figure 1a, as calculated with statistics

(14) Davis, J. M.; Pompe, M.; Samuel, C. *Anal. Chem.* **2000**, *72*, 5700−5713.
(15) Wright, D. W.; Mahler, K. O.; Ballard, L. B. *J. Chromatogr. Sci.* **1986**, *24*, 60−65.
(16) Krock, K. A.; Ragunathan, N.; Wilkins, C. L. *Anal. Chem.* **1994**, *66*, 425−430.
(17) Krock, K. A.; Ragunathan, N.; Klawun, C.; Sasaki, T.; Wilkins, C. L. *Analyst* **1994**, *119*, 483−489.
(18) Weber, B.; Maas, B.; Mosandl, A. *J. Agric. Food Chem.* **1995**, *43*, 2438−2441.
(19) Tomlinson, M. J.; Wilkins, C. L. *J. High Resolut. Chromatogr.* **1998**, *21*, 347−354.
(20) Mondello, L.; Catalfamo, M.; Dugo, P.; Dugo, G. *J. Microcolumn Sep.* **1998**, 10, 203−212.
(21) Mondello, L.; Verzera, A.; Previti, P.; Crispo, F.; Dugo, G. *J. Agric. Food Chem.* **1998**, *46*, 4275−4282.
(22) Mondello, L.; Catalfamo, M.; Cotroneo, A.; Dugo, G.; McNair, H. *J. High Resolut. Chromatogr.* **1999**, 22, 350−356.
(23) Bicchi, C.; D′Amato, A.; Rubiolo, P. *J. Chromatogr., A* **1999**, 843, 99−121.
(24) Cartoni, G. P.; Goretti, G.; Monticelli, B.; Russo, M. V. *J. Chromatogr.* **1986**, *370*, 93−101.
(25) Cartoni, G. P.; Goretti, G.; Russo, M. V. *Chromatographia* **1987**, 23, 790−795.
(26) Chanegriha, N.; Baaliouamer, A.; Rolando, C. *J. Chromatogr., A* **1998**, *819*, 61−65.
(27) Veriotti, T.; Sacks, R. *Anal. Chem.* **2001**, *73*, 4395−4402.
(28) Dimandja, J. M. D.; Stanfill, S. B.; Grainger, J.; Patterson, D. G. *J. High Resolut. Chromatogr.* **2000**, 23, 208−214.
(29) Shellie, R.; Marriott, P.; Cornwell, C. *J. High Resolut. Chromatogr.* **2000**, 23, 554−560.
(30) Marriott, P.; Shellie, R.; Fergeus, J.; Ong, R.; Morrison, P. *Flavor Fragrance J.* **2000**, 15, 225−239.
(31) Shellie, R.; Marriott, P.; Cornwell, C. *J. Sep. Sci.* **2001**, 24, 823−830.

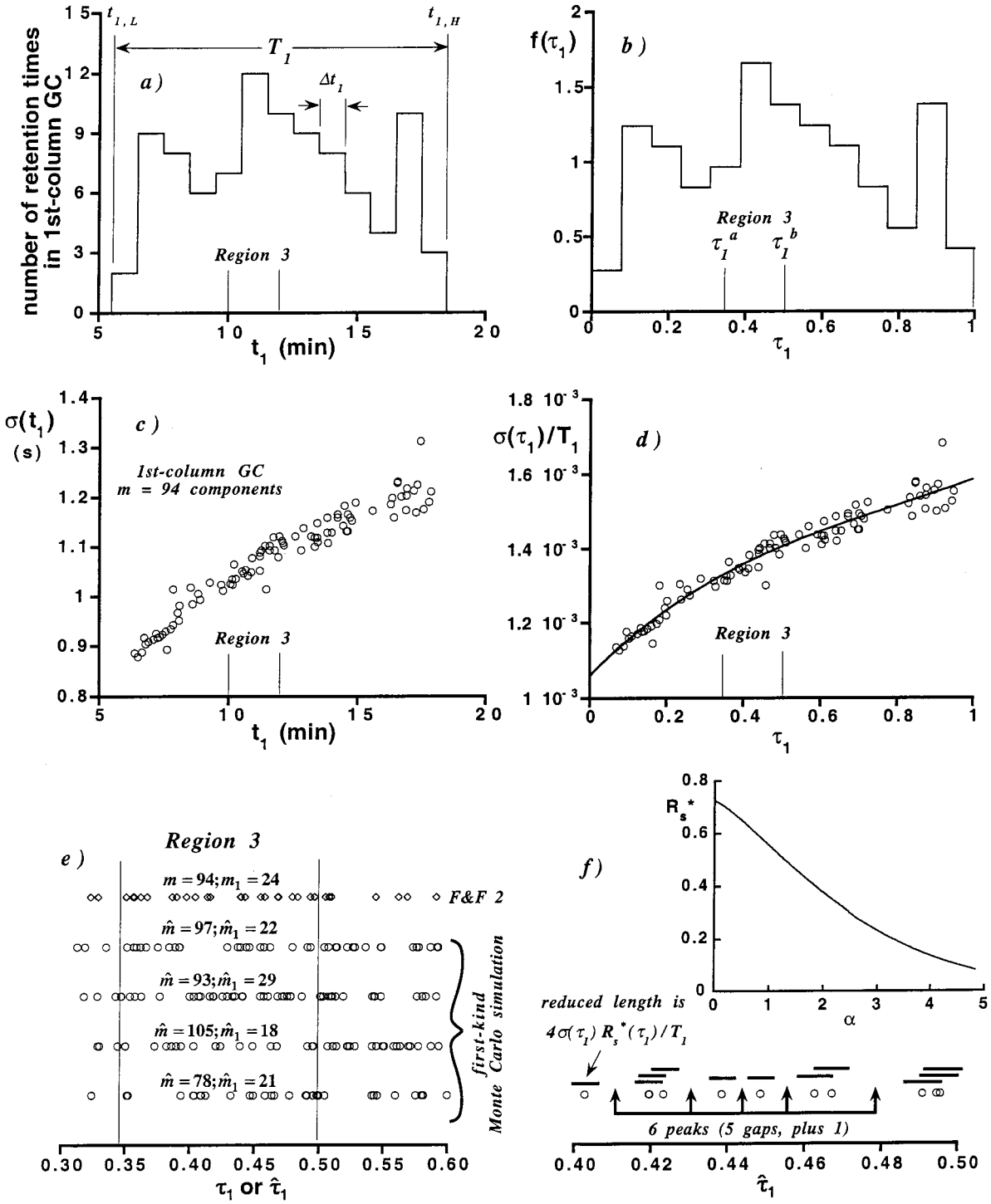

Figure 2. (a) Histogram of SCP retention times t_1 calculated from first-column GC in Figure 1a. (b) Frequency $f(\tau_1)$ vs reduced retention time τ_1 calculated from (a). (c) SCP standard deviation $\sigma(t_1)$ vs t_1 calculated from first-column GC in Figure 1a. (d) Reduced standard deviation $\sigma(\tau_1)/T_1$ vs τ_1 calculated from (c). Curve is least-squares fit of polynomial to data. (e) Reduced retention times τ_1 from first-column GC in Figure 1a and sequences of reduced retention times $\hat{\tau}_1$ from first-kind Monte Carlo simulations. (f) Line segments representing reduced SCP breadths and peaks formed by gaps between them. Inset: graph of resolution R_s^* vs saturation α.

software. The histogram's ordinate is the number of retention times between t_1 and $t_1 + \Delta t_1$, where t_1 is the first-column retention time and Δt_1 is the class-interval width shown in the figure. The lower and upper bounds of the histogram are $t_{1,L}$ and $t_{1,H}$ (the bounds lie beyond the first and last retention times because it is statistically improbable that the bounds and times coincide).

Figure 2b is the graph of frequency $f(\tau_1)$ versus τ_1 derived from the histogram. The abscissa τ_1 is the reduced retention time

$$\tau_1 = (t_1 - t_{1,L})/T_1 \quad (1)$$

where the temporal span T_1 of the separation equals $t_{1,H} - t_{1,L}$ as shown in Figure 2a. Consequently, the bounds of τ_1 are 0 and 1.

The ordinate of $f(\tau_1)$ is scaled such that the area under $f(\tau_1)$ is unity. Thus, $f(\tau_1)$ is a probability density function (pdf).

The amount of overlap in the first-column GC also depends on SCP breadths. Figure 2c is a graph of the standard deviations $\sigma(t_1)$ of the $m = 94$ Gaussian SCPs versus t_1, as predicted by GC software. Figure 2d is a related graph of reduced standard deviation $\sigma(\tau_1)/T_1$ versus τ_1. The curve is a least-squares polynomial fit, which will be used presently.

To simulate the first-column GC, $\hat{m}$ retention times are distributed between $\tau_1 = 0$ and $\tau_1 = 1$ using a Poisson distribution obeying frequency $f(\tau_1)$, as discussed elsewhere[14] (throughout the paper, the "hat" symbol indicates Monte Carlo random variables). Among different simulations, $\hat{m}$ varies randomly. Its average value is set equal to m, the number of compounds in the first-column GC. Since $\hat{m}$ is Poisson distributed, the number $\hat{m}_1$ of retention times in the column-switched region also is Poisson distributed. In Figure 2e, four random retention-time sequences so generated are compared to the deterministic retention times in the first-column GC.

Each simulated retention time is associated with an SCP, whose breadth is represented by a line segment centered about it as shown in Figure 2f. The reduced segment length at τ_1 is $4\sigma(\tau_1)R_s^*(\tau_1)/T_1$,[12] where $\sigma(\tau_1)/T_1$ is calculated from the polynomial in Figure 2d and R_s^* is the average minimum resolution required to separate successive SCPs.[32] The latter's value is predicted from SOT and depends on the interval distribution between SCPs, the amplitude distribution of SCPs, and the amount of overlap.[33] At τ_1, R_s^* satisfies the equation[12]

$$\alpha(\tau_1) = 4mf(\tau_1)\sigma(\tau_1)R_s^*(\tau_1)/T_1 \qquad (2)$$

such that the coordinate, $[\alpha(\tau_1), R_s^*(\tau_1)]$, lies on the graph of R_s^* versus saturation α shown in the inset of Figure 2f (this graph is specific to the Poisson distributions of intervals between SCPs and of SCP amplitudes,[33] the latter of which is observed often in natural-product mixtures[34,35]). By expressing eq 2 relative to the average number m of SCPs in simulations of the first-column GC, we have assumed that SOT applies to the entirety of the first-column GC. Saturation $\alpha(\tau_1)$ varies along the τ_1 axis.

Once reduced lengths are assigned, the number $\hat{p}_1$ of peaks in the column-switched region of the simulation is counted as the number of gaps between successive SCPs, whose associated segments do not overlap, plus 1. The overlap with segments outside the region is ignored. In Figure 2f, five such gaps exist, forming six peaks. Like $\hat{m}_1$, $\hat{p}_1$ is a random variable, whose value varies among different simulations.

Monte Carlo Simulation of the First Kind: Second-Column GC. We now simulate the second-column GC. Figure 3a is a histogram of the $m_1 = 24$ retention times in the second-column GC in Figure 1b, as calculated by statistics software. The reduced retention time τ_2 is

$$\tau_2 = (t_2 - t_{2,L})/T_2 \qquad (3)$$

where t_2 is the second-column retention time, $t_{2,L}$ and $t_{2,H}$ are the lower and upper histogram boundaries, and $T_2 = t_{2,H} - t_{2,L}$ is the temporal extent of the separation (see Figure 3a). Figure 3b is the frequency $f(\tau_2)$ versus τ_2 derived from the histogram; as before, $f(\tau_2)$ is a pdf. Figure 3c is a graph of the standard deviation $\sigma(t_2)$ of the $m_1 = 24$ Gaussian SCPs versus t_2; Figure 3d is the related graph of reduced standard deviation $\sigma(\tau_2)/T_2$ versus τ_2. The curve is a least-squares polynomial fit, which will be used presently.

Consider a simulation of the first-column GC having $\hat{m}_1$ Poisson-distributed SCPs in the column-switched region (see Figure 2e). To simulate the second-column GC, exactly $\hat{m}_1$ SCPs are distributed between $\tau_2 = 0$ and $\tau_2 = 1$ using uniformly random statistics obeying frequency $f(\tau_2)$, as discussed elsewhere.[14] As before, the $\hat{m}_1$ retention times are associated with SCP breadths represented by segments of reduced length $4\sigma(\tau_2)R_s^*(\tau_2)/T_2$, with $\sigma(\tau_2)/T_2$ predicted from the polynomial in Figure 3d and with $R_s^*(\tau_2)$ satisfying the equation

$$\alpha(\tau_2) = 4 <\hat{m}_1> f(\tau_2)\sigma(\tau_2)R_s^*(\tau_2)/T_2 \qquad (4a)$$

with $[\alpha(\tau_2), R_s^*(\tau_2)]$ lying on the graph of R_s^* versus α in Figure 2f. Quantity $<\hat{m}_1>$ is the average number of SCPs in simulations of the second-column GC and equals the average number of SCPs in the column-switched region of simulations of the first-column GC[36]

$$<\hat{m}_1> = m \int_{\tau_1^a}^{\tau_1^b} f(\tau_1)\, d\tau_1 \qquad (4b)$$

where the integration limits τ_1^a and τ_1^b are the reduced coordinates of the region's boundaries (see Figure 2b).

The number $\hat{p}_2$ of peaks in simulations of the second-column GC is counted exactly as is $\hat{p}_1$. Like $\hat{m}_1$ and $\hat{p}_1$, $\hat{p}_2$ is a random variable, whose value varies among simulations.

Our coupling of eq 4a and the relation between R_s^* and α for a Poisson interval distribution may seem wrong since the interval distribution in simulations of the second-column GC is uniformly random, not Poisson. However, for a Poisson counting process beginning at time $t_{2,L}$ and having the value $\hat{m}_1$ at time $t_{2,H}$, a stochastic theorem states that the arrival times of $\hat{m}_1$ Poisson events have a distribution identical to the order statistics corresponding to $\hat{m}_1$ independent random variables uniformly distributed on the interval, $(t_{2,L}, t_{2,H})$.[37] The converse of this theorem also is true: $\hat{m}_1$ independent random variables that are uniformly distributed on $(t_{2,L}, t_{2,H})$ have Poisson arrival times (and Poisson interval distributions) if $\hat{m}_1$ is Poisson distributed.

Assessment. SOT predicts that the average number of peaks in the column-switched region of simulations of the

(32) Felinger, A. *Anal. Chem.* **1997**, *69*, 2976−2979.
(33) Davis, J. M. *Anal. Chem.* **1997**, *69*, 3796−3805.
(34) Nagels, L. J.; Creten, W. L.; Vanpeperstraete, P. M. *Anal. Chem.* **1983**, *55*, 216−220.
(35) Dondi, F.; Kahie, Y. D.; Lodi, G.; Remelli, M.; Reschiglian, P.; Bighi, C. *Anal. Chim. Acta* **1986**, *191*, 261−273.
(36) Davis, J. M. *J. Microcolumn Sep.* **1995**, *7*, 3−15.
(37) Rowe, K.; Davis, J. M. *J. Chemom. Intell. Lab. Syst.* **1997**, *38*, 109−126.

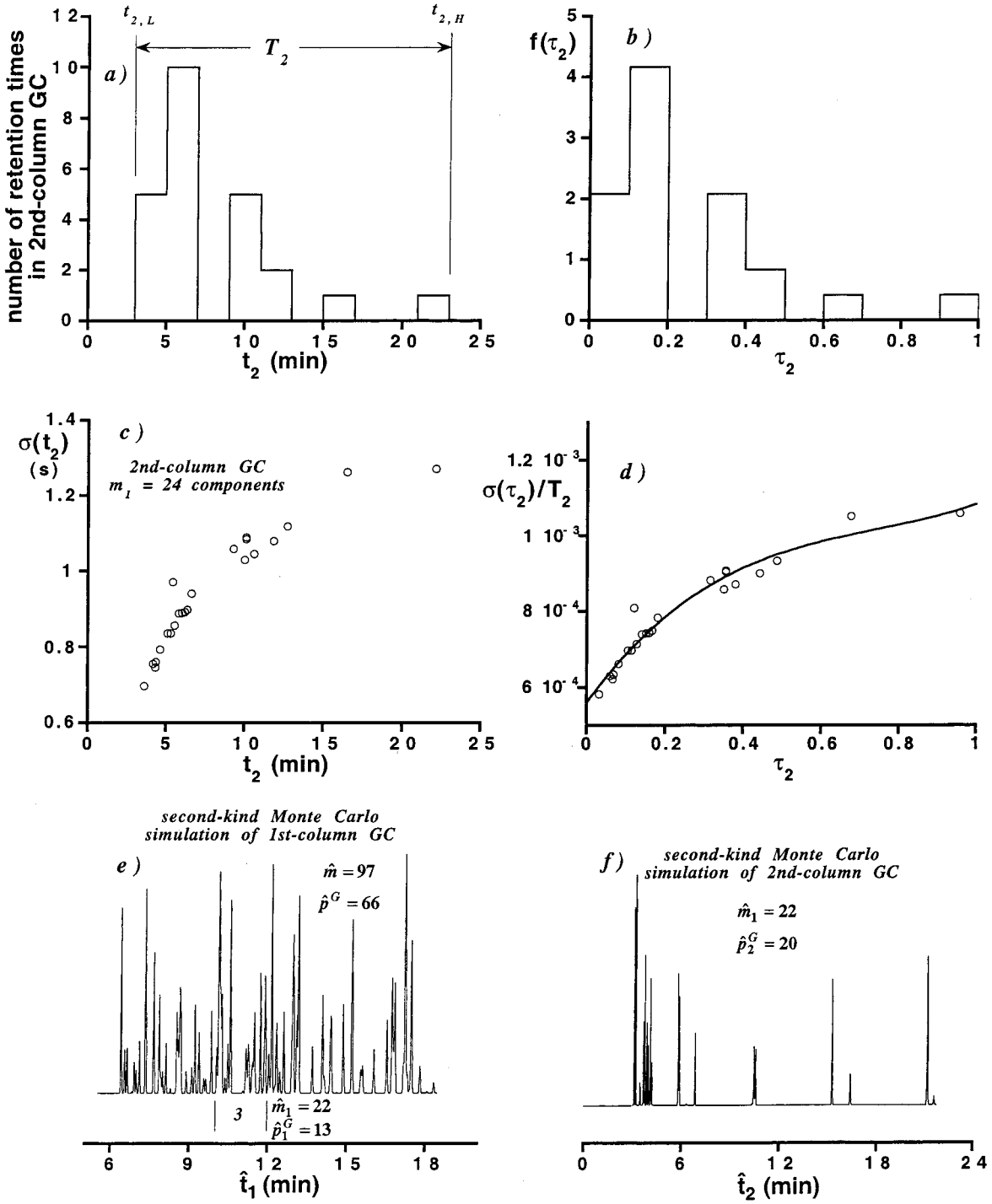

Figure 3. (a–d) As in Figure 2a–d, but for second-column GC in Figure 1b. (e) Second-kind Monte Carlo simulation of first-column GC in Figure 1a. (f) As in (e), but for second-column GC in Figure 1b.

first-column GC is[12,36]

$$m \int_{\tau_1^a}^{\tau_1^b} f(\tau_1) \exp[-\alpha(\tau_1)]\, d\tau_1 \qquad (5a)$$

and the variance of the peak number is[36,37]

$$m \int_{\tau_1^a}^{\tau_1^b} f(\tau_1) \exp[-\alpha(\tau_1)]\{1 - 2\alpha(\tau_1) \exp[-\alpha(\tau_1)]\}\, d\tau_1 \qquad (5b)$$

Thus, the average $<\hat{p}_1>$ and variance $\sigma^2_{\hat{p}_1}$ of random variable $\hat{p}_1$ should equal eqs 5a and 5b, respectively.

A consequence of the stochastic theorem discussed above is that the average number of peaks in simulations of the second-column GC is[12]

$$<\hat{m}_1> \int_0^1 f(\tau_2) \exp[-\alpha(\tau_2)]\, d\tau_2 \qquad (5c)$$

and the variance of the peak number is[37]

$$<\hat{m}_1> \int_0^1 f(\tau_2) \exp[-\alpha(\tau_2)]\{1 - 2\alpha(\tau_2) \exp[-\alpha(\tau_2)]\}\, d\tau_2 \qquad (5d)$$

with $<\hat{m}_1>$ equal to eq 4b. Thus, the average $<\hat{p}_2>$ and variance $\sigma^2_{\hat{p}_2}$ of random variable $\hat{p}_2$ should equal eqs 5c and 5d, respectively.

Calculation of $P(M_1 \mid p_1, p_2)$. Equation 5 is useful in characterizing the accuracy of moments of first-kind Monte Carlo simulations but not in calculating $P(M_1 \mid p_1, p_2)$. However, the calculation is simple. For every simulation of the first- and second-column GCs, a 3-tuple $(\hat{m}_1, \hat{p}_1, \hat{p}_2)$ is generated. First, the number N of times the 3-tuple, $(\hat{m}_1, \hat{p}_1 = p_1, \hat{p}_2 = p_2)$, is generated is counted for all $\hat{m}_1$. Then, the distribution $P(M_1 \mid p_1, p_2)$ is calculated for all M_1 values by dividing by N the number of times the specific 3-tuple, $(\hat{m}_1 = M_1, \hat{p}_1 = p_1, \hat{p}_2 = p_2)$, is generated.

The introduction of variable M_1 may seem puzzling. Although $\hat{m}_1$ and M_1 are both random variables for the number of SCPs in the column-switched region, $\hat{m}_1$ is a random variable in a Monte Carlo simulation. In contrast, the probability distribution is independent of the means of its calculation. Consequently, we use M_1 to describe the distribution.

To simplify matters, we have assumed until now that peak numbers p_1 and p_2 in the first- and second-column GCs are fixed attributes. In reality, p_1 and p_2 are determined by a specific sequence of SCP amplitudes and may differ if the amplitudes are varied. To assess the dependence of $P(M_1 \mid p_1, p_2)$ on SCP amplitudes in the first- and second-column GCs, the latter are varied and different distributions $P(M_1 \mid p_1, p_2)$ are calculated.

Calculations of Moments. We propose that if SOT is valid, then the best estimate of the number m_1 of compounds in the column-switched region is the mean $<M_1>$ of $P(M_1 \mid p_1, p_2)$[11]

$$<M_1> = \sum_{\max(p_1, p_2)}^{\infty} M_1 P(M_1 \mid p_1, p_2) \qquad (6a)$$

where the smallest M_1 value is the larger of p_1 and p_2 (i.e., "max (p_1, p_2)"), since at least as many SCPs as peaks are present. Quantity $<M_1>$ differs from $<\hat{m}_1>$, eq 4b, since it depends on the conditional observations of p_1 and p_2 peaks.

The standard deviation σ_{M_1} of $P(M_1 \mid p_1, p_2)$ is the square root of the variance

$$\sigma^2_{M_1} = \sum_{\max(p_1, p_2)}^{\infty} (M_1 - <M_1>)^2 P(M_1 \mid p_1, p_2) \qquad (6b)$$

Monte Carlo Simulation of the Second Kind: Calculation of Gaussian Profiles. Here, we assign Gaussian profiles of random amplitude to the Monte Carlo retention-time sequences described above. Specifically, the first-column GC is simulated by calculating the sum of all $\hat{m}$ Gaussians having standard deviations predicted from the polynomial fit to $\sigma(\tau_1)/T_1$, and the second-column GC is simulated by calculating the sum of $\hat{m}_1$ Gaussians having standard deviations predicted from the polynomial fit to $\sigma(\tau_2)/T_2$. Unlike in the first- and second-column GCs, however, the SCP amplitudes in the two simulations are not correlated since any amplitude correlation would be lost in a large sequence of simulations having uncorrelated retention times. The number of peak maximums in second-kind simulations of the first-column GC, the column-switched region, and the second-column GC are $\hat{p}^G$, $\hat{p}^G_1$, and $\hat{p}^G_2$, respectively (the superscript "G" signifies "Gaussian").

Unlike in Monte Carlo simulations of the first kind, $\hat{p}^G_1$ *is* affected by SCPs immediately outside the boundaries of the column-switched region, since their profiles spill into it. For this reason, Monte Carlo simulations of the second kind are more realistic than ones in which SCP breadths are represented by segment lengths. However, they are computationally intensive and are impractical for evaluating $P(M_1 \mid p_1, p_2)$. A small number of them is generated to assess if $\hat{p}_1$ and $\hat{p}_2$ differ from $\hat{p}^G_1$ and $\hat{p}^G_2$. If differences exist, then the distribution $P(M_1 \mid p_1, p_2)$ calculated from $\hat{p}_1$ and $\hat{p}_2$ is biased against interpreting the overlap of Gaussian peaks.

Panels e and f in Figure 3 are second-kind Monte Carlo simulations of the first- and second-column GCs in Figure 1. The simulations preserve the SCP distributions in the first- and second-column GCs but with random, not deterministic, SCP retention times. One observes the similarities among the peak and component numbers $m = 94$, $p = 64$, $m_1 = 24$, $p_1 = 14$, and $p_2 = 22$ in Figure 1a and b and their Monte Carlo counterparts, $\hat{m} = 97$, $\hat{p}^G = 66$, $\hat{m}_1 = 22$, $\hat{p}^G_1 = 13$, and $\hat{p}^G_2 = 20$ in Figure 3e and f.

Limitations of Theory. To develop the SOT, we have assumed that standard deviations $\sigma(\tau_1)$ and $\sigma(\tau_2)$, frequencies $f(\tau_1)$ and $f(\tau_2)$, and the number m of SCPs in the first-column GC are known. In practice, they must be determined. The standard deviations are determined by measuring the breadths of pure SCPs spanning the region of interest[38,39] or calculating them by SOT.[40] Frequencies $f(\tau_1)$ and $f(\tau_2)$ can be approximated from the retention times of peak maximums when overlap is not severe.[41] Severe overlap compromises the determination of $f(\tau_1)$ and $f(\tau_2)$, however, and this compromise currently limits the application of our SOT to separations of modest overlap unless the frequency is constant. The number m of SCPs in the first-column GC can be determined by regression methods when $f(\tau_1)$ is known.[41]

PROCEDURES

GC prediction. The retention times, t_1 and t_2, and standard deviations, $\sigma(t_1)$ and $\sigma(t_2)$, of SCPs in the first- and second-column GCs were calculated using the software, Pro ezGC for Windows (Restek Corp., Bellefonte, PA). Retention times were calculated from enthalpy and entropy changes between mobile and stationary phases; standard deviations were calculated from the Golay–Giddings equation.

Pro ezGC contains libraries of thousands of compounds on various stationary phases. Our study required that two libraries contain a large number of the same compounds on stationary phases of different polarity. Only libraries of flavors and fragrances or PCBs were satisfactory. We chose 247 flavors and fragrances on nonpolar 100% methylsilicone (first-column phase) and polar poly(ethylene glycol) (second-column phase). All GCs were predicted for 30-m capillaries having internal diameters of 250 μm and stationary-phase thicknesses of 0.25 μm. In the predictions, the first column was temperature programmed from 40 to 250 °C at 4 °C/min and the second column from 60 to 250 °C at 4 °C/min. Predictions were made under constant-pressure conditions,

(38) Herman, D. P.; Gonnord, M.-F.; Guiochon, G. *Anal. Chem.* **1984**, *56*, 995–1003.
(39) Dondi, F.; Gianferrara, T.; Reschiglian, P.; Pietrogrande, M. C.; Ebert, C.; Linda, P. *J. Chromatogr.* **1990**, *485*, 631–645.
(40) Dondi, F.; Betti, A.; Pasti, L.; Pietrogrande, M. C.; Felinger, A. *Anal. Chem.* **1993**, *65*, 2209–2222.
(41) Davis, J. M. *Anal. Chem.* **1994**, *66*, 735–746.

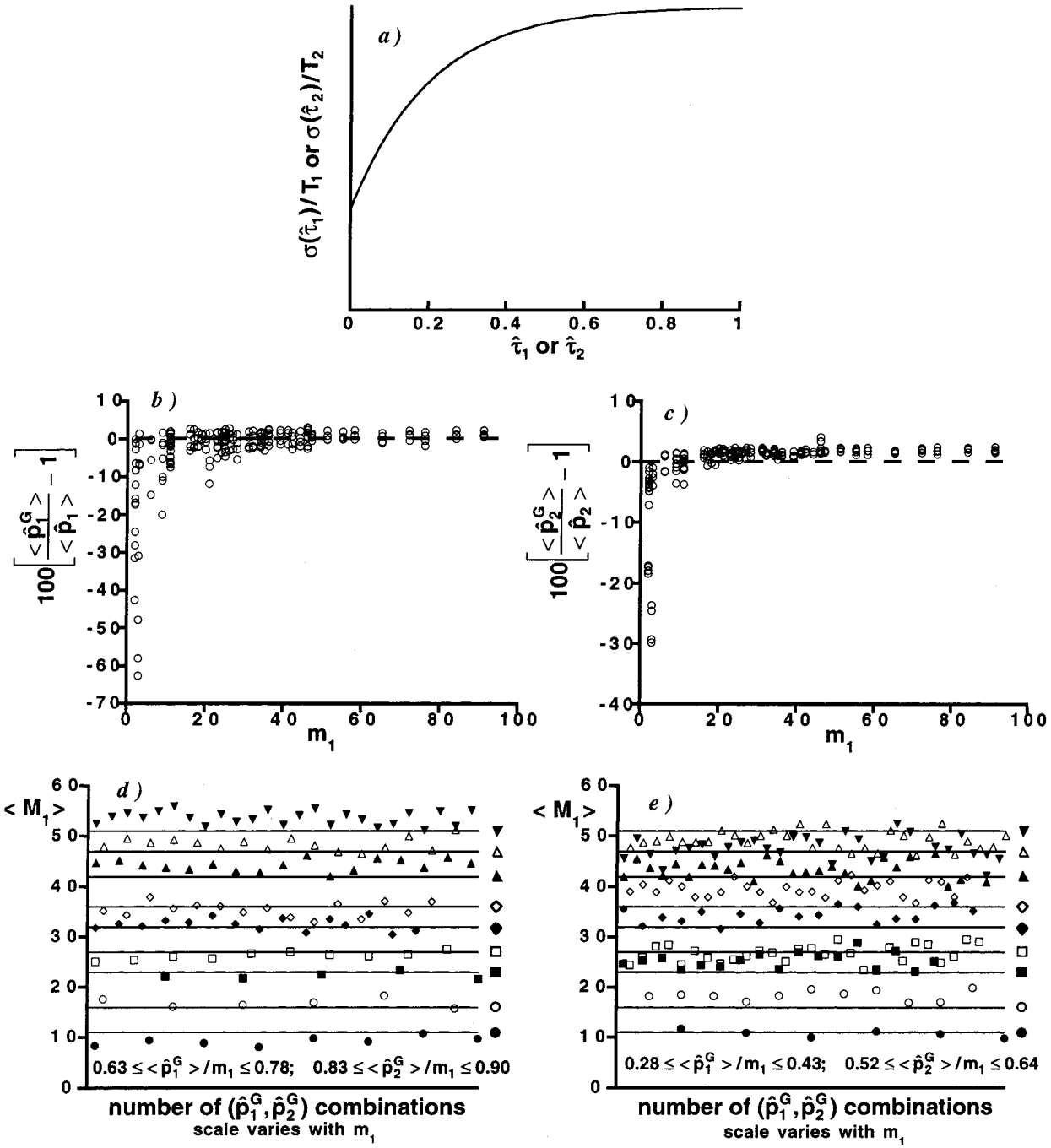

Figure 4. (a) Reduced standard deviation $\sigma(\hat{\tau}_1)/T_1$ or $\sigma(\hat{\tau}_2)/T_2$ vs $\hat{\tau}_1$ or $\hat{\tau}_2$, by which SOT was tested. (b) Percentage error between averages $<\hat{p}_1^G>$ and $<\hat{p}_1>$ vs m_1. (c) As in (b), but for $<\hat{p}_2^G>$ and $<\hat{p}_2>$. (e, f) Averages $<M_1>$ calculated from different $(\hat{p}_1^G, \hat{p}_2^G)$ combinations and distribution $P(M_1 \mid \hat{p}_1^G, \hat{p}_2^G)$.

with the linear velocity of the mobile phase, helium, equal to 30.0 and 28.8 cm/s in the first and second columns, respectively, at the lowest temperatures.

Two limitations of our use of Pro ezGC are noted. First, Pro ezGC does not predict column-switched GCs. To mimic column switching, we predicted two independent GCs—one on each stationary phase—and restricted the column-switched region to compounds in both libraries. This action mimicked developing the first-column GC, repeating the development and diverting the column-switched region to a cryofocuser, rapidly desorbing the cryofocused compounds into the second column, and developing the second-column GC. The mimicking is consistent with the actual column switching of flavors and fragrances, which often are cryotrapped to ensure timely volatilization and efficient loading.

Second, Pro ezGC computations are restricted to 100 or fewer compounds. The separation of 247 flavors and fragrances consequently was partitioned into three separate windows designated F&F 1, 2, and 3. The first-column GC of F&F 2 is a continuation of that of F&F 1, and the first-column GC of F&F 3 is a continuation of that of F&F 2. In other words, the first-column GCs of F&Fs 1−3 constitute a single chromatogram.

Frequency Generation. The histograms of the first- and second-column GC were calculated from Pro ezGC retention times

Table 1. Statistics of Linear Graphs of $<\hat{p}_1>$, $\sigma^2_{\hat{p}_1}$, $<\hat{p}_2>$, and $\sigma^2_{\hat{p}_2}$ vs Eqs 5a−d, Respectively.

relation	slope	intercept	correlation coefficient	abscissa range
eq 5a vs $<\hat{p}_1>$	1.0003	−0.2612	0.999 92	$1.0 \le <\hat{p}_1> \le 63.2$
eq 5b vs $\sigma^2_{\hat{p}_1}$	0.9936	0.0411	0.994 89	$0 \le \sigma^2_{\hat{p}_1} \le 31.6$
eq 5c vs $<\hat{p}_2>$	1.0044	−0.2974	0.999 93	$1.3 \le <\hat{p}_2> \le 78.6$
eq 5d vs $\sigma^2_{\hat{p}_2}$	0.9908	−0.3234	0.995 45	$0.2 \le \sigma^2_{\hat{p}_2} \le 58.2$

using the software, Minitab Release 12 for Windows (Minitab, Inc., State College, PA). A "nice numbers" algorithm determined the numbers of class intervals. The frequencies $f(\tau_1)$ and $f(\tau_2)$ were calculated using eqs 1 and 3 and the histogram boundaries.

Reduced Standard Deviations. The standard deviations $\sigma(t_1)$ and $\sigma(t_2)$ were converted to reduced standard deviations $\sigma(\tau_1)/T_1$ and $\sigma(\tau_2)/T_2$ using eqs 1 and 3 and the histogram boundaries. Polynomials were fit to the reduced standard deviations using KaleidaGraph 3.07 (Synergy Software, Reading, PA). The polynomial degree was chosen to obtain good fits while avoiding spurious oscillations.

Simulation. The algorithm for first- and second-kind Monte Carlo simulations was written in FORTRAN 77 and executed on a VArStation YMP 4665A Linux workstation (VA Research, Inc., Mountain View, CA). Random numbers were generated with function RAN3 in ref 42. The algorithm performed the following: (a) generated from Pro ezGC retention times and standard deviations 300 sequences of first- and second-column GCs containing Gaussian SCPs having exponentially distributed random amplitudes. In any sequence, the amplitude of a given compound was the same in both GCs. The numbers p_1 and p_2 of peak maximums were counted. Henceforth, we call the set of maximum numbers from one sequence a (p_1, p_2) combination; (b) generated 500 000 first-kind Monte Carlo simulations of the first- and second-column GCs. Segment lengths were determined with eqs 2 and 4, the polynomial fits to $\sigma(\tau_1)/T_1$ and $\sigma(\tau_2)/T_2$, and the graph of R^*_s versus α. In each simulation, the numbers $\hat{m}_1$, $\hat{p}_1$, and $\hat{p}_2$ were counted; (c) generated distributions $P(M_1 \mid p_1, p_2)$ for every (p_1, p_2) combination in the 300 sequences of first- and second-column GCs (see (a) above). The moments $<M_1>$ and $\sigma^2_{M_1}$ were calculated for each combination with eq 6; (d) generated 5000 second-kind Monte Carlo simulations of the first- and second-column GCs containing Gaussian SCPs having exponentially distributed random amplitudes. The amplitudes were not correlated between GC simulations. The SCP standard deviations were calculated from fits to $\sigma(\tau_1)/T_1$ and $\sigma(\tau_2)/T_2$. In each simulation, the numbers $\hat{p}^G_1$ and $\hat{p}^G_2$ were counted; (e) calculated averages and standard deviations of 5000 values of $\hat{m}_1$, $\hat{p}^G_1$, and $\hat{p}^G_2$; (f) evaluated eq 5 using Simpson's rule.

Testing of SOT. The GC simulation algorithm was tested using computer-generated inputs mimicking Pro ezGC data. Specifically, 28 sequences of (on average) $m = 100$ retention times were generated to mimic first-column GCs using a homogeneous Poisson distribution. In the 28 sequences, 50 column-switched regions were chosen randomly, with $2 \le m_1 \le 91$. Fifty sequences of m_1 retention times then were generated to mimic second-column GCs using a homogeneous uniformly random distribution. The frequencies $f(\tau_1)$ and $f(\tau_2)$ of all sequences were calculated as described above. The SCP standard deviations were modeled by Figure 4a, which is discussed below.

RESULTS AND DISCUSSION

Optimization of First-Column GCs. To justify the need of multidimensional separation, we optimized the first-column GCs of F&F 1−3 using the optimization algorithm of Pro ezGC. Accordingly, roughly 177 000 trial GCs of F&Fs 1−3 were calculated using a single temperature ramp, with the void-peak elution time, initial temperature, final temperature, and heating rate constrained to lie between 0 and 60 min, 0 and 250 °C, 250 and 310 °C, and 0 and 15 °C/min, respectively. We then considered only separations spanning 45 min or less between the first and last maximums. The optimal first-column GCs of F&Fs 1−3 contained (on average) 67, 74, and 43 maximums, respectively; the unoptimized first-column GCs contained (on average) 50, 66, and 37 maximums, respectively. In all cases, a smaller initial temperature and heating rate than we arbitrarily chose improved separation. However, since F&Fs 1−3 contain 87, 94, and 66 compounds, respectively, optimization left many compounds unresolved. Thus, the multidimensional separation of F&Fs 1−3 appears to be required, if one desires total resolution. We now examine how well unoptimized column switching provides total resolution.

Verification of SOT. Simulated separations containing SCPs having random retention times often are used to assess SOT.[6−9] In this section, the words "test GCs" refer to the mimics of Pro ezGC data used to test SOT.

Figure 4a is a graph of $\sigma(\hat{\tau}_1)/T_1$ or $\sigma(\hat{\tau}_2)/T_2$ versus $\hat{\tau}_1$ or $\hat{\tau}_2$, from which SCP standard deviations were calculated (the "hats" signify random retention times). The graph is empirical but similar to ones previously predicted by Pro ezGC for programmed temperature gas chromatography.[13] The ordinate was scaled such that the number of maximums equaled roughly 60−80% of m_1 in the best first-column test GCs (30−40% in the worst) and 80−90% of m_1 in the best second-column test GCs (50−60% in the worst).

Table 1 reports the slopes, intercepts, and correlation coefficients of linear graphs of $<\hat{p}_1>$, $\sigma^2_{\hat{p}_1}$, $<\hat{p}_2>$, and $\sigma^2_{\hat{p}_2}$ versus the predictions of eqs 5a−5d, respectively. The slopes, intercepts, and correlation coefficients approach 1, 0, and 1, respectively, and signify that the mean and variance of peak numbers in first-kind Monte Carlo simulations are indistinguishable from ones predicted by eq 5.

Figure 4b is a graph of the percentage error between $<\hat{p}_1>$ and $<\hat{p}^G_1>$, the average number of maximums in the column-switched region of second-kind Monte Carlo simulations of first-column test GCs, versus m_1. The error is almost 0 for large m_1, and its magnitude is less than 10% if $m_1 \ge 10$ or so. Figure 4c is

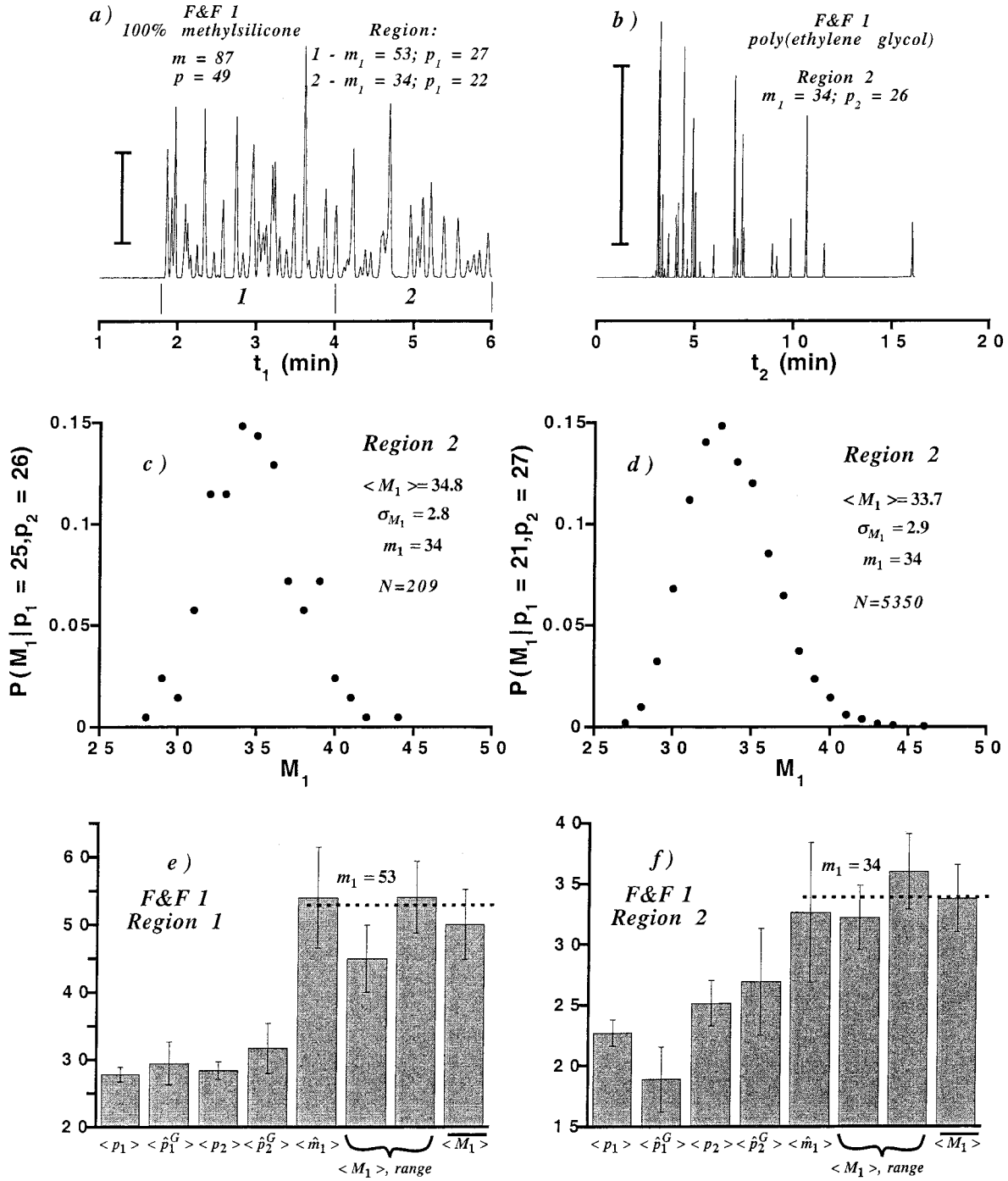

Figure 5. (a, b) As in Figure 1, but for volatile flavors and fragrances, F&F 1, eluting between 2 and 6 min in first-column GC. First-column GC is partitioned into two column-switched regions. (c, d) Distribution $P(M_1 \mid p_1, p_2)$ vs M_1 for different (p_1, p_2) combinations in region 2. (e) Averages and standard deviations of results determined from region 1 of first-column GC. (f) As in (e), but for region 2.

a similar graph but for $<\hat{p}_2>$ and $<\hat{p}_2^G>$, the average number of maximums in second-kind Monte Carlo simulations of second-column test GCs. The error is positive but its magnitude is small unless $m_1 \leq 3$. For small m_1, the error in $<\hat{p}_1^G>$ is larger than in $<\hat{p}_2^G>$, probably because SCP profiles immediately outside the region spill into it, increasing overlap and reducing $\hat{p}_1^G$ relative to $\hat{p}_1$. The error increases as m_1 decreases, because the region size decreases and approaches the SCP breadth.

Table 1 shows that $\hat{p}_1$ and $\hat{p}_2$ follow SOT, if retention times are random; panels b and c in Figure 4 show the errors between maximum numbers and $\hat{p}_1$ or $\hat{p}_2$ are less than 10% if $m_1 \geq 10$ or so. As explained earlier, the distribution $P(M_1 \mid p_1, p_2)$ is calculated by identifying the maximum numbers in first- and second-column GCs with $\hat{p}_1$ and $\hat{p}_2$, respectively. Thus, we expect the distribution to be accurate if retention times are random and $m_1 \geq 10$ or so.

Panels d and e in Figure 4 depict $<M_1>$ values calculated from $P(M_1 \mid \hat{p}_1^G, \hat{p}_2^G)$ for different $(\hat{p}_1^G, \hat{p}_2^G)$ combinations in test GCs, when $11 \leq m_1 \leq 51$. The lines represent different m_1 values and are paired with the appropriate symbol at the lines' right-hand edges (e.g., the bottom line represents $m_1 = 11$ and is paired with a filled circle; this line should be compared to filled circles). The $<M_1>$ values are not sensitive to different $(\hat{p}_1^G, \hat{p}_2^G)$ combinations

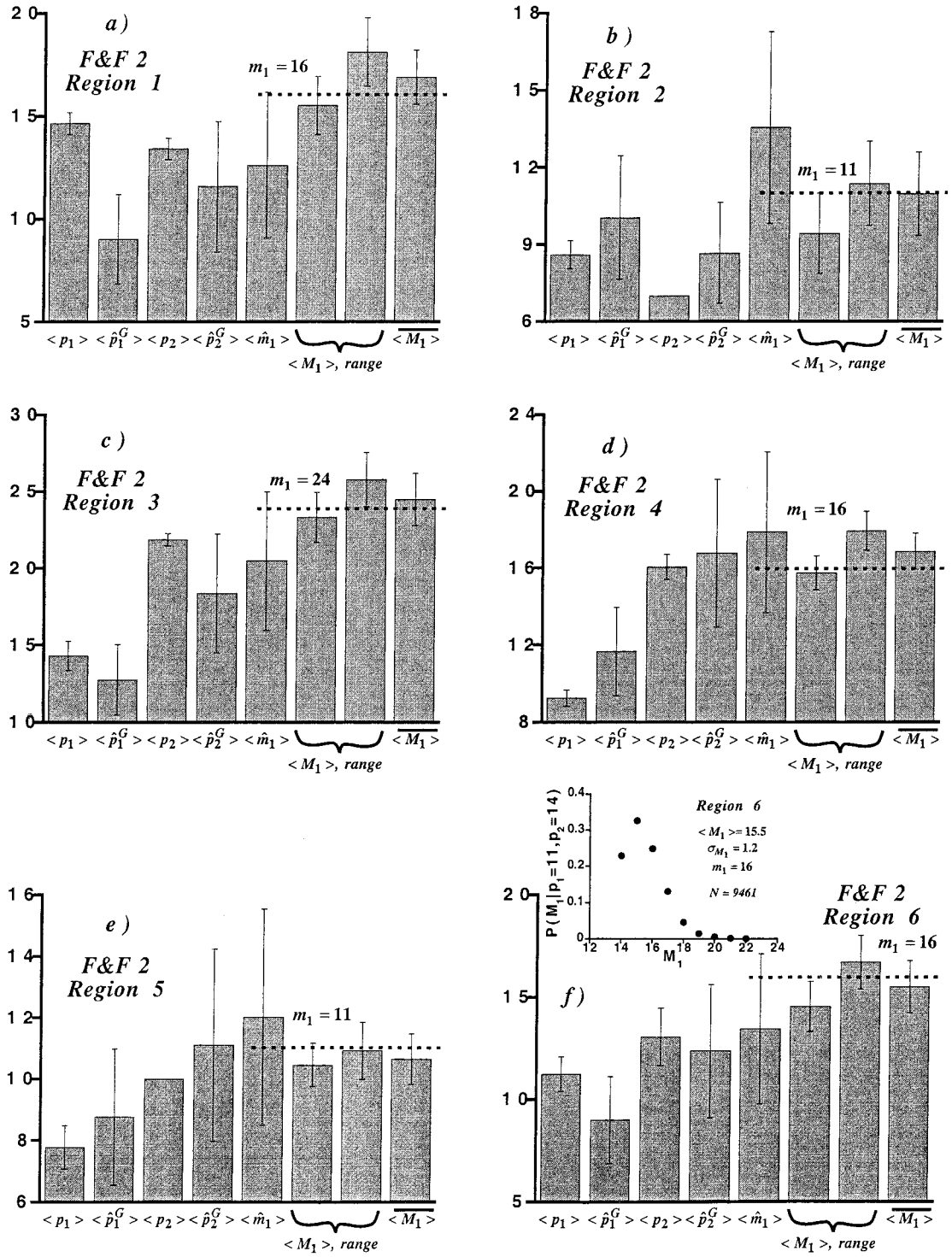

Figure 6. (a–f) As in Figure 5e, but for regions 1–6 of F&F 2 in Figure 1a. Inset in Figure 6f is distribution $P(M_1 \mid p_1, p_2)$ vs M_1 for specific (p_1, p_2) combination in region 6.

but are more accurate if predicted from test GCs having modest overlap (Figure 4d) instead of severe overlap (Figure 4e). This finding is consistent with previous studies by SOT.[6–9] However, the latter predictions are acceptable. On average, then, $<M_1>$ is a good estimate of m_1 if retention times are random and $m_1 \geq 10$.

Application of SOT to Flavors and Fragrances. Figure 5a is the predicted first-column GC of $m = 87$ highly volatile flavors and fragrances (F&F 1) on the nonpolar stationary phase, 100% methylsilicone. Only $p = 49$ peaks are observed because of overlap. The GC was partitioned into two contiguous regions; the numbers m_1 and p_1 of compounds and peaks in both are reported. The compounds in both regions then were "transferred" to the polar stationary phase, poly(ethylene glycol). Figure 5b is the predicted second-column GC of region 2. The former contains more peaks ($p_2 = 26$) than the latter ($p_1 = 22$), indicating an improvement of separation.

The improvement is insufficient, however. Panels c and d in Figure 5 are graphs of $P(M_1 \mid p_1, p_2)$ versus M_1 calculated from

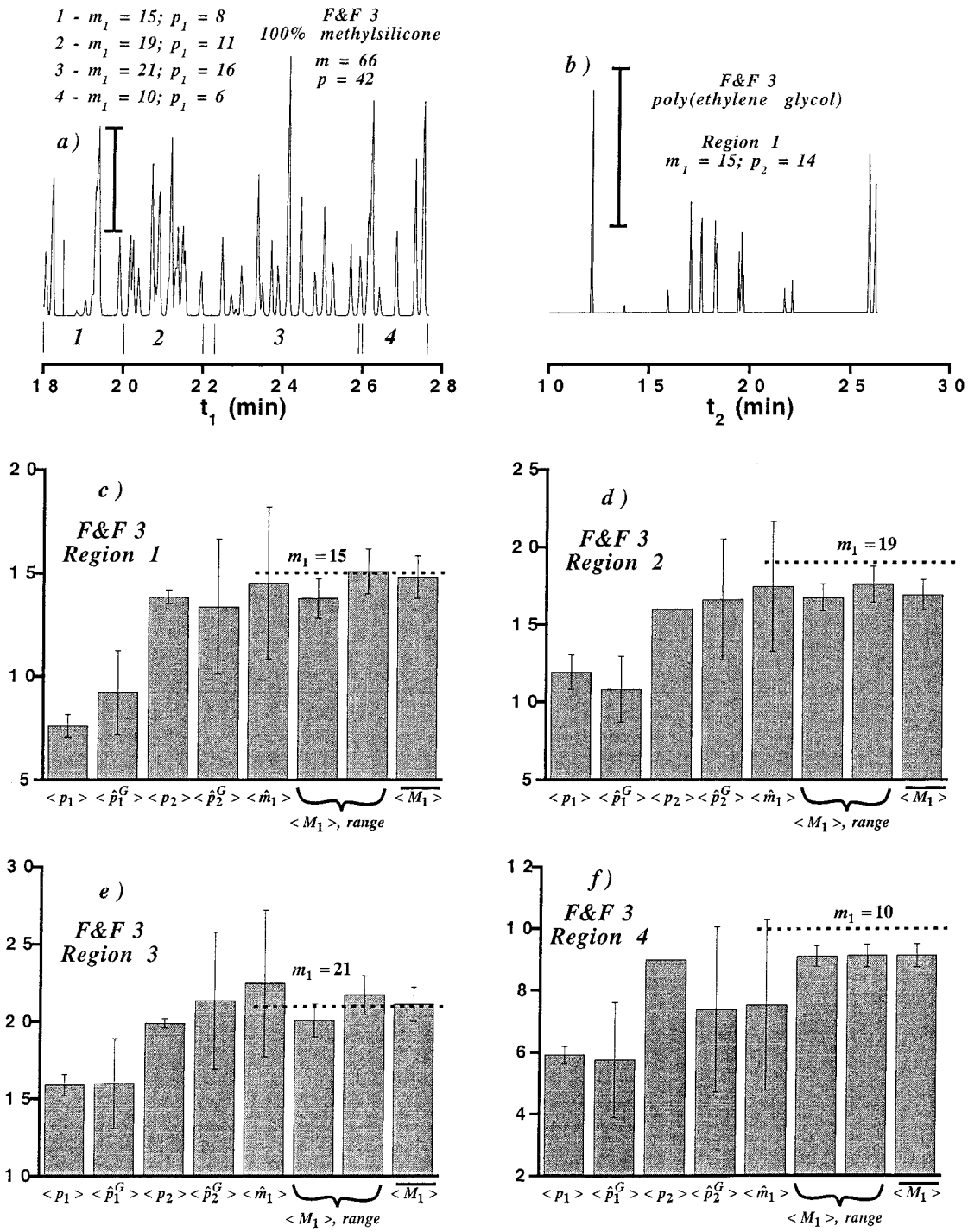

Figure 7. (a, b) As in Figure 1, but for semivolatile flavors and fragrances, F&F 3, eluting between 18 and 28 min in first-column GC. First-column GC is partitioned into four column-switched regions. (c−f) As in Figure 5e, but for regions 1−4 of F&F 3.

first-kind Monte Carlo simulations for two (p_1, p_2) combinations. The distribution in Figure 5d is the smoother of the two, because it was calculated from the larger number N of 3-tuples, ($\hat{m}_1$, $\hat{p}_1 = p_1$, $\hat{p}_2 = p_2$). Both distributions resemble skewed discretized Gaussians and scarcely differ, showing that $P(M_1 \mid p_1, p_2)$ is insensitive to different (p_1, p_2) combinations. The distributions show that the probability of m_1 equalling max (p_1, p_2) is very small. Figure 5c represents an outcome with no significant increase of separation ($p_1 = 25$, $p_2 = 26$), whereas Figure 5d represents one with a fairly large increase ($p_1 = 21$, $p_2 = 27$). However, in both cases, $<M_1>$ is about 34−35—much greater than the peak numbers—and in agreement with the number of compounds in region 2 ($m_1 = 34$).

Two conclusions can be drawn from the distributions. First, the finding that p_1 and p_2 are almost the same does not imply that separation is complete.[11] Rather, m_1 can exceed both if the separations have poor selectivity. Second, the favorable finding $p_2 > p_1$ does not mean that $p_2 \approx m_1$. Rather, m_1 can exceed p_2, even though the selectivity of the second separation is better than the first.

Panels e and f in Figure 5 are results obtained from the first- and second-column GCs and Monte Carlo first- and second-kind

simulations. The bars represent various averages, which are specified below the horizontal axis (e.g., the third bar from the left is the average number $<p_2>$ of peaks in second-column GCs). The second and third bars from the right are the largest and smallest $<M_1>$ values, respectively, calculated from extreme (p_1, p_2) combinations (together, they represent the $<M_1>$ range). The right-most bar is the global estimate $\overline{<M_1>}$ calculated by pooling, i.e., by weighting different $<M_1>$'s by the fraction of times a specific (p_1, p_2) combination occurred. The dotted line spanning the last four bars represents m_1 and should be compared to the bars' values. Standard deviations are represented by error bars; the absence of error bars signifies a standard deviation of zero. The standard deviation associated with $\overline{<M_1>}$ was calculated by pooling different variances $\sigma^2_{M_1}$.

In general, we can infer that SOT is valid if (A) $<p_1>$ and $<\hat{p}_1^G>$ agree within one to two standard deviations of $\hat{p}_1^G$, (B) $<p_2>$ and $<\hat{p}_2^G>$ agree within one to two standard deviations of $\hat{p}_2^G$, and (C) m_1 and the various $<M_1>$'s agree within one to two values of σ_{M_1}.

In essence, criteria A and B are based on interpreting $<p_1>$ or $<p_2>$ as a single datum, averaged over different SCP amplitudes, of a specific retention-time distribution. One might envision comparing standard deviations; however, the standard deviations of p_1 and p_2 cannot be compared to those of $\hat{p}_1^G$ and $\hat{p}_2^G$ for reasons discussed elsewhere.[13]

It is clear these criteria are satisfied in Figure 5e and f. Separation in region 1 is not improved by column switching, whereas it is improved somewhat in region 2. In both cases, however, m_1 exceeds both p_1 and p_2 and the excess is predicted by SOT. The prediction of m_1 by $<M_1>$ shows that the deterministic retention times in the first- and second-column GCs are indistinguishable from ones simulated with random distributions.

Panels a−f in Figure 6 are similar to panels e anf f in Figure 5, except they correspond to regions 1−6 of the first-column GC in Figure 1a. These flavors and fragrances (F&F 2) are less volatile than those in Figure 5a and elute between 6 and 18 min. Because the regions have extents comparable to those in Figure 5a but contain fewer compounds, one might think that column switching should significantly improve separation.

Surprisingly, column switching actually decreases separation in the first two regions, while improving it in the last four. With one exception—the disagreement between $<p_1>$ and $<\hat{p}_1^G>$ in region 1—the criteria listed above are satisfied and SOT describes the results. Although column switching improves separation in four regions, separation is complete in only one of them (region 4). Interestingly, the improvement is largest in region 3, which contains more compounds than other regions. From this finding, we conclude that improvement of separation is only partly related to compound number and (unsurprisingly) also depends on selectivity.

The inset in Figure 6f is a graph of $P(M_1 \mid p_1, p_2)$ versus M_1 for region 6. Unlike the distributions in Figure 5c and d, it is not Gaussian-like; a fair probability exists that M_1 is equal to or only slightly greater than $\max(p_1, p_2)$. Such distributions were reported in our earlier study.[11] The distributions differ because the number of compounds in F&F 2, region 6 ($m_1 = 16$), is roughly half of that in F&F 1, region 2 ($m_1 = 34$). Consequently, the probability of separation is greater in the former than the latter.

The finding that separation is improved in some column-switched regions of F&F 2 but degraded in others is consistent with the statistical behavior of the separation. Some "rolls of the dice" are unfavorable, leading to a loss of separation, whereas others are favorable, leading to increased separation.

Figure 7a is the predicted first-column GC of $m = 66$ semivolatile flavors and fragrances (F&F 3) eluting between 18 and 28 min. Only $p = 42$ peaks are observed because of overlap. The GC was partitioned into four regions; the one compound between regions 3 and 4 was excluded from consideration, since it was absent from the poly(ethylene glycol) library of the GC software. Figure 7b is the predicted second-column GC of region 1. It contains many more peaks ($p_2 = 14$) than the column-switched region ($p_1 = 8$), but the separation still is incomplete ($m_1 = 15$).

Results are shown in Figure 7c−f. The same trends reported above are observed, with two exceptions. First, column switching improves separation in all four regions. Second, the predictions $<M_1>$ are somewhat low in regions 2 and 4 (this is addressed immediately below). However, our conclusion remains the same: separation still is incomplete in the second-column GC, despite improvement, and SOT predicts this outcome, although not always exactly.

In general, $\overline{<M_1>}$ is more representative than individual $<M_1>$ values, since it is determined from all (p_1, p_2) combinations. A prediction by SOT perhaps is questionable if m_1 lies outside two standard deviations of $\overline{<M_1>}$. In this study, m_1 lies within two standard deviations of $\overline{<M_1>}$ except in the two cases noted above. However, in both cases, the difference between m_1 and $\overline{<M_1>}$, plus two standard deviations, is only 0.1. Since m_1 must be integral, one can relax one's assessment and conclude these results are not questionable.

CONCLUSIONS

We draw four general conclusions. The first is that cryofocused column-switching gas chromatography without optimization is not as powerful as many researchers think, when 10 or more compounds are transferred. The second-column GC often contains overlapping SCPs.

This conclusion could have been reached without SOT by simply predicting first- and second-column GCs and interpreting results. The second conclusion—which does require SOT—is that the search for optimal separation conditions in cryofocused column-switching gas chromatography is a lengthy trial-and-error process *because* the separation is inherently statistical, i.e., because SOT describes the results. The same separation conditions can lead to improvement, no change, or loss of separation, depending on the nature of mixture constituents. The optimal conditions sought by researchers correspond to favorable "rolls of the dice", which rarely are obtained on initial efforts but only after tedious searches.

These conclusions may seem broad in light of our limited work. Among other things, we addressed only one mixture type, did not consider the enantioselective cyclodextrin phases that resolve flavors and fragrances,[18,20−23,43−45] and did not reverse the order of column polarity. Nevertheless, we think that if two GCs of a mixture each follow one-dimensional SOT, then a column-switched cryofocused separation based on the same two phases will follow column-switching SOT. Since one-dimensional SOT describes

overlap in predicted GCs of many mixtures on both polar and nonpolar phases,[12,13,46] our first two conclusions should apply to many column-switched separations.

The third conclusion is that SOT can be applied to multidimensional separations other than two-dimensional separations, which have been studied before.[47] The SOT describing the latter separations, in which spots or bi-Gaussians are displayed in a plane, is different from that developed here. The interpretation of multidimensional separations by different approaches supports the universality of SOT.

Our final conclusion is that Monte Carlo simulations permit a flexibility in SOT that is not possible with analytical theory. This flexibility includes the calculation of probability distributions for the number of mixture components (e.g., $P(M_1 \mid p_1, p_2)$), which has been done analytically only in the case of a one-dimensional separation containing uniformly distributed SCPs of constant density and breadth.[37,48] Such distributions contain all details of overlap statistics and are more useful than the moments of such distributions (e.g., eq 6). If one knew both the retention-time distribution(s) and statistics type(s) governing overlap—from either empiricism or theory—then such distributions could be calculated from simulations of one-dimensional separations, column-switched separations with and without preconcentration, two- and three-dimensional separations, and perhaps other types. In such simulations, SCP densities and breadths could be varied, different statistics types could be mixed, and even the periodicity of homologues could be introduced. Monte Carlo simulations consequently could broaden greatly the application of SOT.

Many column-switched gas chromatograms entail the transfer of fewer than 10 components as small heartcuts. This study does not address the modeling of overlap in such separations. Figure 4b does show that m_1 may be as small as 6, if the tolerable error between maximum numbers and $\hat{p}_1$ is relaxed from 10 to 20%. For $m_1 = 2$ or 3, however, the error can be much larger. The error probably results from asymptotic assumptions in the derivation of the relationship between α and R_s^* [33] (see Figure 2f). The modeling of overlap for very small m_1 requires a more sophisticated SOT than that used here. Perhaps Fourier analysis[7] or the pulse-point method[8,49] might be more useful.

(42) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in FORTRAN*; Cambridge University Press: Cambridge, 1992.
(43) Casabianca, H.; Graff, J. B.; Jame, P.; Perrucchietti, C.; Chastrette, M. *J. High Resolut. Chromatogr.* **1995**, *18*, 279−285.
(44) Mosandl, A. *Food Rev. Int.* **1995**, *11*, 597−664.
(45) Juchelka, D.; Mosandl, A. *Pharmazie* **1996**, *51*, 417−422.
(46) Davis, J. M.; Samuel, C. *J. High Resolut. Chromatogr.* **2000**, *23*, 235−244.
(47) Rowe, K.; Bowlin, D.; Zou, M.; Davis, J. M. *Anal. Chem.* **1995**, *67*, 2994−3003.
(48) Martin, M.; Herman, D. P.; Guiochon, G. *Anal. Chem.* **1986**, *58*, 2200−2207.
(49) Pietrogrande, M. C.; Coll, P.; Sternberg, R.; Szopa, C.; Navarro-Gonzalez, R.; Vidal-Majar, C.; Dondi, F. *J. Chromatogr., A* **2001**, 939, 69−77.

ACKNOWLEDGMENT

The authors thank Eric Good (Minitab, Inc.), John Garrett (Analytical Innovations, Inc.), and David Olive (Department of Mathematics, SIU-C) for helpful discussions. This work was funded by the National Science Foundation (Grant CHE-9714328).

SELECTIVE GLOSSARY

m	number of compounds in first-column GC
m_1	number of compounds in column-switched region of first-column GC
$\hat{m}$	number of retention times in Monte Carlo simulation of first-column GC
$\hat{m}_1$	number of retention times in column-switched region of Monte Carlo simulation of first-column GC
M_1	random variable for number of compounds in column-switched region of first-column GC
$<M_1>$	average of $P(M_1 \mid p_1, p_2)$
$\overline{<M_1>}$	global average of $<M_1>$
p	number of peak maximums in first-column GC
p_1	number of peak maximums in column-switched region of first-column GC
p_2	number of peak maximums in second-column GC
$\hat{p}_1$	number of peaks in column-switched region of first-kind Monte Carlo simulation of first-column GC
$\hat{p}_2$	number of peaks in first-kind Monte Carlo simulation of second-column GC
$\hat{p}^G$	number of peak maximums in second-kind Monte Carlo simulation of first-column GC
$\hat{p}_1^G$	number of peak maximums in column-switched region of second-kind Monte Carlo simulation of first-column GC
$\hat{p}_2^G$	number of peak maximums in second-kind Monte Carlo simulation of second-column GC
$P(M_1 \mid p_1, p_2)$	probability that M_1 compounds are in column-switched region of first-column GC, given p_1 peaks in the region and p_2 peaks in the second-column GC
α	saturation
σ_{M_1}	standard deviation of $P(M_1 \mid p_1, p_2)$

Received for review December 4, 2001. Accepted March 11, 2002.

AC011235B