Sensors and Actuators B 202 (2014) 167–176

Contents lists available at ScienceDirect

Sensors and Actuators B: Chemical

journal homepage: www.elsevier.com/locate/snb

Multivariate curve resolution of co-eluting vapors from a gas chromatograph with microsensor array detector

Jonathan Bryant-Genevier[a,d], Kee Scholten[c,d], Sun Kyu Kim[a,d], Edward T. Zellers[a,b,c,d,*]

[a] Dept. of Environmental Health Sciences, University of Michigan, Ann Arbor, MI 48109-2029, United States
[b] Dept. of Chemistry, University of Michigan, Ann Arbor, MI 48109-1055, United States
[c] Applied Physics Program, University of Michigan, Ann Arbor, MI 48109-1040, United States
[d] Center for Wireless Integrated MicroSensing and Systems, University of Michigan, Ann Arbor, MI 48109-2122, United States

ARTICLE INFO

Article history:
Received 20 December 2013
Received in revised form 29 March 2014
Accepted 12 May 2014
Available online 17 May 2014

Keywords:
Microsensor array
Gas chromatography
Multivariate curve resolution
EFA
ALS
Chemometrics

ABSTRACT

The application of a hybrid multivariate curve resolution method, which combines evolving factor analysis (EFA) with alternating least squares (ALS), to the analysis of partially overlapping peaks from vapors measured by a microsensor-array gas chromatograph detector is described. The detector comprised an array of four chemiresistors coated with different sorptive thiolate-monolayer-protected gold nanoparticle (MPN) films. Three pairs of vapors, the members of which had array response pattern correlation coefficients, ρ, ranging from −0.57 to 0.85, were tested at different values of chromatographic resolution, R_s, and relative response ratio, *RRR*. Composite responses were equivalent to the sums of the responses to the individual components, but differences in peak asymmetry among the sensors in the array led to pattern distortions across the spans of all peaks. With data pre-processing to account for the latter, EFA correctly determined the chemical ranks of the binary composite peaks in 57 of the 63 cases (90%), with most errors observed for the most highly correlated pair. By using calibrated response patterns as inputs for the ALS refinements of EFA-extracted responses, the fidelity of recovered response patterns and elution profiles was sufficiently high to differentiate the composite peak components in 124 of 126 cases (98%) and to quantify them to within ±30% of actual values in 95 of 126 (75%) cases. Without such inputs, the corresponding rates were 112 of 126 (89%) and 68 of 126 (54%), respectively. In general, the *RRR* value was a more important determinant of performance than was the R_s value. The methodology and performance of EFA-ALS in this application are critically assessed.

1. Introduction

The use of microsensor arrays with sorptive interface layers for the quantitative analysis of mixtures of volatile organic compounds (VOC) generally requires upstream gas chromatographic (GC) separation because the diversity of responses achievable with stand-alone arrays of this type is severely limited; the composite pattern of responses from a mixture of three or more vapors typically cannot be reliably differentiated from those of the individual components or their lower-order mixtures [1–4]. The temporal separation provided by GC can alleviate this problem, but as GC systems continue to be miniaturized [5] and analysis times are reduced [6], chromatographic resolution invariably decreases, which raises the likelihood of overlapping peaks among the mixture components.

Although several reports have appeared on the use of hyphenated (μ)GC-microsensor array instrumentation for VOC mixture analysis [6–13] the chemometric analysis of data derived from such systems has not been studied extensively. In earlier explorations of this topic, we applied classification models derived from extended disjoint principal component regression (EDPCR) to assign identities to chromatographically resolved peaks corresponding to individual vapors and to composite peaks corresponding to unresolved or partially resolved binary mixtures, on the basis of the array response patterns derived from the corresponding peak areas [1–3,8,12]. If two peaks overlapped at all, then they were treated as an unresolved binary mixture, ignoring the potential value of partial chromatographic resolution.

Multivariate curve resolution (MCR) methods, which are commonly used to analyze chromatograms of complex mixtures with overlapping peaks obtained from gas and liquid chromatographic

* Corresponding author at: Dept. of Environmental Health Sciences, University of Michigan, Ann Arbor, MI 48109-2029, United States. Tel.: +1 734 936 0766; fax: +1 734 763 8095.
E-mail addresses: genevier@umich.edu (J. Bryant-Genevier), keesc@umich.edu (K. Scholten), airbuff@gmail.com (S.K. Kim), ezellers@umich.edu (E.T. Zellers).

http://dx.doi.org/10.1016/j.snb.2014.05.049
0925-4005/

analyzers equipped with spectrometric detection [14–28], may offer a more refined alternative to the analysis of such GC-microsensor array data. MCR methods can be used to determine the number of components in overlapped chromatographic peak composites and then to extract the spectrum and elution profile of each component without prior knowledge of the mixture composition. Extracted spectra can then be matched to those in a library to determine the identities of the individual analytes, and the recovered profiles can yield the mass or concentration by reference to calibration data. Examples of MCR algorithms applied to data from hyphenated chromatographic-spectrometric systems include evolving factor analysis (EFA) [14–19], SIMPLISMA [20], PARAFAC [21], alternating least squares (ALS) [15,19,24–26], and derivatives or combinations thereof [27,28]. The topic has been the subject of a recent review [29].

In EFA, singular value decompositions are performed progressively over sequential time segments of a composite peak profile, in both forward and reverse directions. Assuming that sufficient spectrometric discrimination of the partially resolved peak components is provided by the detector and that the responses from the components add linearly, then the rank of the data matrix should equal the number of analytes in the composite peak. By following EFA with ALS, iterative refinements are made to the initially extracted elution profiles and spectra to improve accuracy and efficiency [24]. It is also common to incorporate a priori problem-specific information in ALS to constrain the calculation and thereby further reduce the error in the final solution [17].

Since the response pattern produced by a microsensor array exposed to a vapor phase analyte is analogous to a spectrum produced by a spectrometer, such MCR methods should be readily applicable to microsensor arrays used as GC detectors. The primary difference is the amount of independent information about the analytes that can be derived from the detector output, which is expected to be much less for a microsensor array than for a spectrometer.

Mathematically, the raw chromatographic data matrix $\boldsymbol{Y}$, which consists of the collection of responses from each sensor (columns) at each value of time (rows), can be decomposed as follows:

$$\mathbf{Y} = \mathbf{CS} + \mathbf{E} \tag{1}$$

where $\boldsymbol{C}$ is the concentration matrix, the rows of which are the points in time and the columns of which are the concentrations of each component, and $\boldsymbol{S}$ is the sensor sensitivity matrix, the rows of which are the sensitivity pattern of the eluting components and the columns of which are the sensors in the array. To remove the influence of concentration, sensitivity patterns can be normalized (e.g., by constraining the range of sensitivities from 0 to 1). Residual errors are relegated to the error matrix $\boldsymbol{E}$.

Initial estimates of $\boldsymbol{C}$ and $\boldsymbol{S}$ are generated by EFA, which assumes that there are selective elution regions of each component, such that the first peak to start eluting is also the first peak to finish eluting. These estimates are subsequently refined using ALS, which employs the following iterative calculations, given $\boldsymbol{Y}$:

$$\mathbf{C} = \mathbf{YS}'(\mathbf{SS}')^{-1} \tag{2}$$

$$\mathbf{S}' = (\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'\mathbf{Y} \tag{3}$$

Eq. (2) is the least squares solution of Eq. (1) with respect to $\boldsymbol{C}$, assuming that $\boldsymbol{S}$ is known, and Eq. (3) is the least squares solution of the same equation assuming $\boldsymbol{C}$ is known. In practice, neither $\boldsymbol{C}$ nor $\boldsymbol{S}$ is known a priori, but by iterating between Eqs. (2) and (3), starting from the initial estimates of $\boldsymbol{C}$ and $\boldsymbol{S}$, one approaches an optimized solution for $\boldsymbol{Y}$. Optimization can be facilitated by applying constraints, such as non-negativity and unimodality, or by substituting calibration responses for EFA estimates, in situations where it is justified [17]. The ALS algorithm stops iteration at a preset convergence in the residual error matrix $\boldsymbol{E}$.

In a previous study from our laboratory, the application of EFA-ALS to *simulated data* from an array of microsensors used as the detector in a GC was explored [30]. Using sensitivities from an array of 4 chemiresistors (CR) coated with thiolate-monolayer-protected Au nanoparticles (MPNs) having different thiolate functionalities, Monte Carlo simulations were used to investigate whether EFA-ALS could detect and then recover the response patterns of the peak components from binary composites with high fidelity. Although experimental sensitivity values were employed, the overlaps and relative responses were simulated assuming idealized Gaussian peaks, and the (5%) random error superimposed on the sensor responses was also assumed to be Gaussian for the Monte Carlo simulations. Overall, the performance of the method and the array was quite good for the synthetic cases considered. Quantification accuracy was not explored.

The purpose of the current study was to evaluate the performance of such a hybrid EFA-ALS curve resolution algorithm with *experimental data* from an MPN-CR array used as a GC detector. Using the framework of our previous study [30], the capability to determine the number of components (i.e., the chemical rank) in each composite peak and to extract the pattern and concentration of each component were evaluated over a range of chromatographic resolution, R_s, and relative response ratio, *RRR*, values for pairs of vapors with different (Pearson) correlation coefficients, ρ.

After describing our methodology and various aspects of the individual and collective array responses for the three pairs of vapors tested, we present an assessment of the additivity of responses from the sensors in the binary composite peaks. The preprocessing steps required to address differences in retention time, t_R, and peak asymmetry among the sensors are then described. This is followed by an evaluation of rank determinations by EFA. The fidelity of the response patterns recovered by EFA-ALS to those in the calibration library, and the accuracy of quantification, are examined as a function of R_s, *RRR*, ρ, using two different "modes" (i.e., blind mode and informed mode) that differ in terms of the amount of a priori information used in the ALS step. Results are compared to those obtained using EDPCR, which ignores the partial chromatographic resolution of the components in the composite peaks.

2. Experimental and chemometric methods

2.1. Materials

All test compounds were purchased from Sigma–Aldrich/Fluka (Milwaukee, WI) or Acros/Fisher (Pittsburgh, PA) in >98% purity and were used as received. MPNs having monolayers derived from the following thiols were taken from existing supplies synthesized by a previously reported method [31]: n-octanethiol (C8), 6-phenoxyhexane-1-thiol (OPH), 4-(phenylethynyl)-benzenethiol (DPA), and methyl-6-mercaptohexanoate (HME). Average Au core diameters were ~4 nm. The thiolate monolayers on these MPNs span a sufficiently wide range of structures and vapor affinities to afford reasonably high selectivity among different VOCs [10,30], from the non-polar C8 to the more strongly dipolar HME.

2.2. Microsensor array

The CR array chip (2.0 cm × 1.2 cm) is the same as that used in several of our previous studies [9,10,12,32]. It contains four pairs of Au/Cr interdigital electrodes (IDEs) on a thermal-SiO_x/Si substrate. Each IDE has 24 electrode pairs with 5 μm widths/spaces, 450 μm length, and 410 μm overlap. MPN films were deposited by drop casting from solution to create multilayer films with baseline

resistance values, R_b, of 0.48, 1.0, 4.4, and 7.8 MΩ, for C8, DPA, OPH, HME, respectively (note: only one of each pair of similarly coated sensors was used). Film thicknesses were not measured, but previous work suggests that this method produces non-uniform films with average thicknesses in the range of 100–400 nm [32,33].

A Macor lid was sealed to the substrate at the periphery of the array using a self-adhesive VHB-tape gasket (3 M, St. Paul, MN) and thin film of silicone putty (Duraseal, 1531 Cotronics, Brooklyn, NY), creating a detector cell volume of 1.6 μL. Deactivated fused-silica capillaries (0.25 mm i.d., 0.32 mm o.d., Restek Corp., Belafonte, PA) were inserted into drilled inlet/outlet holes in the lid and sealed with epoxy (Hysol Epoxy Patch 1C, Henkel Corp., Rocky Hill, CT). Header pins were soldered to the contact pads, bent to a 90° angle, and inserted into a socket mounted on a printed circuit board (PCB).

Each CR was connected to a parallel matching reference resistor to create a voltage divider circuit, and the voltage across the CR sensor was amplified. A constant potential of 3 V from a CR2032 battery was applied to each CR circuit. The output voltage was collected at 20 Hz from each sensor, before and after amplification, and passed via a DAQ card to a laptop computer where the data were logged as a function of time. Changes in MPN film resistance, normalized to the baseline resistance (i.e., $\Delta R/R_b$) were calculated from the corresponding voltages by the data management software.

2.3. Testing configuration

The test set-up is shown in Fig. S1 of the Supplementary Data. The CR array was connected to the distal end of a conventional capillary column (6-m long, 250-μm i.d., 0.25-μm thick PDMS phase, Supelco, Belafonte, PA). The column was located inside the oven of a bench-scale GC (Model 7890, Agilent Technologies, Palo Alto, CA) and the PCB-mounted CR array was placed on the bench beside the GC. Fluidic connections to the capillary column and FID inside the GC were made with press-tight fittings using ~10-cm segments of deactivated fused-silica capillary fed through the oven-door seal. The stability of the sensor temperatures during temperature programming of the column in the oven was inferred by the lack of any baseline drift in the sensor outputs.

Samples from test atmospheres generated in N_2 in Supel Inert® gas sampling bags (Supelco) were drawn through a 25-μL gas sampling loop by a commercial mini-pump, and injected into the column via a 6-port valve under a flow of N_2 carrier gas. All chromatograms were run isothermally. Injected masses of each test compound were confirmed by an FID downstream from the CR array, which had been previously calibrated with solutions of the test compounds in CS_2. The carrier gas flow rate was 1.0 mL/min for all trials, as measured by a downstream bubble flow meter. Individual chromatograms were analyzed using GRAMS 32 (Thermo Fisher Scientific, Waltham, MA) and any shift in pre- and post-peak baselines were corrected by linear interpolation prior to further analyses of the peaks.

2.4. Screening tests and selection of vapor pairs

A set of 25 compounds, spanning a range of retention times from 40 to 210 s when run isothermally at 30 °C, was screened to find appropriate pairs for evaluating the EFA-ALS method. Two criteria were applied: the members of a pair had to have retention times similar enough to allow R_s to be adjusted over the desired range (0.1–1.0) with moderate changes in column temperature; and the value of ρ between the responses to the members of a pair had to be such that among the three pairs chosen there was a wide enough range of ρ values to permit the effect of pattern similarity on the performance of the EFA-ALS analyses to be assessed (see below for metrics). For assessing pattern similarity, the peak-maxima sensor sensitivities ($\Delta R/R_b$/ng) for each vapor were normalized by dividing by that of the most sensitive sensor in the array (MSS) for that vapor, and the normalized sensitivities from each of the four sensors for one member of a pair were regressed onto those for the other member (Fig. S2, Supplementary Data). The resulting values of ρ were calculated. Similar analyses were performed on the peak-area sensitivities.

Table 1
List of vapor pairs, measures of pattern diversity, and LODs.

Vapor pair	ρ[a]	ED[b]	LOD[c] (ng)
MIBK +	0.85	0.33	3.2
TOL			3.2
OCN +	0.59	0.79	4.1
BAC			2.7
NET +	−0.57	1.28	0.8
CHX			4.2

[a] Pair-wise Pearson correlation coefficient for normalized responses of pair members (see Fig. S2).
[b] Euclidean distance between normalized array response vectors of pair members.
[c] Based on the least sensitive sensor in the array (i.e., DPA for all vapors except NET, for which C8 was the least sensitive).

The following three pairs were selected: methyl isobutyl ketone + toluene (MIBK + TOL; $\rho = 0.85$), 1-octene + n-butyl acetate (OCN + BAC; $\rho = 0.59$), and nitroethane + cyclohexane (NET + CHX; $\rho = -0.57$) (see Table 1). The corresponding Euclidean distances (*ED*) between the vectors for each vapor in the pair, also calculated from the normalized sensitivities, were 0.33, 0.79, and 1.28 (Table 1). For reference, the maximum possible *ED* value for the normalized data is 2.0, as there are 4 sensors, each of which has a sensitivity of 0–1; thus, the maximum *ED* separating two vapors is $4^{1/2} = 2$.

2.5. Calibration, test conditions, preprocessing

Calibrations were performed with the individual test compounds by injecting discrete samples from bag test-atmospheres at 6 or 7 concentrations through the 6-port valve and sampling loop. The calibration spanned ≥10-fold range of masses between 20 and 600 ng for any compound. The mass range exceeded the range used for subsequent tests of binary mixtures. Plotting the $\Delta R/R_b$ values derived from the baseline-corrected peak maxima vs. injected mass with forced zero *y*-intercept yielded a linear calibration curve ($r^2 \geq 0.98$) the slope of which comprised the sensitivity for a given sensor-vapor combination (Fig. S3, Supplementary Data).

For testing, seven static test-atmospheres of each binary mixture were generated and analyzed. Concentrations were adjusted to give average *RRR*s of approximately 1:10, 1:5, 1:2.5, 1:1, 2.5:1, 5:1, and 10:1, respectively, for the members of a given pair. Actual *RRR*s may have differed by 10–20% from the target values. For NET + CHX, *RRR* values of 1:20 and 20:1 were used instead of 2.5:1 and 1:2.5 to examine the performance of EFA-ALS for the least correlated pair under more challenging conditions. Fig. 1 provides some examples of Gaussian peaks from a hypothetical single-channel detector for selected combinations of R_s and *RRR*, which illustrate the nature of the composite peaks addressed here.

The minimum injected mass was greater than 5 times the limit of detection (LOD) for the least sensitive sensor in the array for all but one of the binary mixture cases (the exception was >3 × LOD). A single test was run at each value of R_s and *RRR*. All tests with a given vapor pair were conducted within a three day period, during which sensitivities remained quite stable. At the lowest oven temperature, the peaks for the members of any pair were fully resolved (i.e., $R_s > 4$). The areas of the fully resolved peaks were used as reference values for assessing the additivity of responses for the overlapped cases, and the peak maxima were used for assessing the accuracy of quantification of the profiles estimated by EFA-ALS.

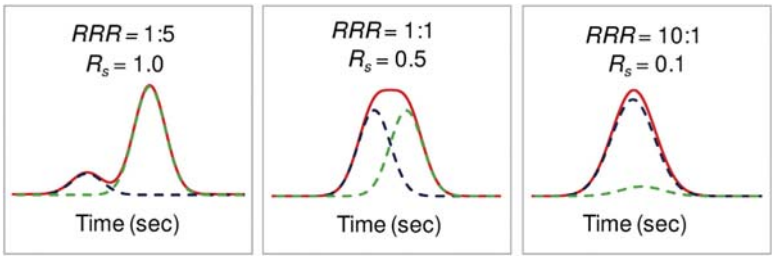

Fig. 1. Idealized (Gaussian) binary composite chromatograms with values of R_s and RRR selected arbitrarily to illustrate the nature and extent of peak overlap considered in this study. Since actual peaks from the sensors (and FID) are asymmetric to varying degrees, the overlap for a given nominal value of R_s will be greater than shown above.

Values of R_s were determined on the basis of the downstream FID traces using the following equation: $R_s = 0.59\Delta t_r/W_{1/2}$, where $W_{1/2}$ is the average full-width-at-half-maximum of the two peaks and Δt_r is the difference in retention time [34]. The net change in oven (column) temperature required to span R_s values of 0.1, 0.5, and 1.0 was roughly 35 °C in all cases, and the specific temperature required to achieve each R_s value was determined using the test atmosphere corresponding to an (average) *RRR* = 1:1. Since the peaks measured with the CR sensors were generally wider and less symmetric than those from the downstream FID, and the degree of asymmetry differed among the sensors (see below), the actual (i.e., sensor-wise) R_s values were always less than the stated (FID-wise) values. Prior to EFA-ALS analyses, it was necessary to correct for slight differences in retention times among the sensors and for differences in peak tailing among the sensors for a given vapor (described further below). Note that the *RRR* values are averaged over all sensors for each member of a pair/mixture. The range of individual *RRR* values spans a much larger range due to individual differences in sensitivity. For example, an *RRR* of 10:1 for the NET + CHX pair corresponds to a range of sensor-specific *RRR* values between 80:1 and 2:1, respectively.

2.6. EFA and ALS

The EFA algorithm developed by the Tauler group and made available on their website was used to determine the rank of each composite peak (after pre-processing; see below) using MatLab [35]. EFA was performed on the submatrices $\boldsymbol{Y}_i$ (i.e., each row in the matrix **Y** consisting of the set of sensor responses) to calculate the eigenvalues of the response vector for each retention time, increasing the window by 0.05 s for each evolution. This was done initially in the forward direction (i.e., increasing elution time). Retention times where sharp increases in the magnitude of the eigenvalues occur reveal when a compound begins to elute. A single noise threshold, established by analyzing single component chromatograms of each vapor, was applied to samples of all three mixtures to determine rank. EFA performed in the reverse direction, with respect to time, revealed the end of the eluting peak, and the elution range of each compound was determined by merging the results of the forward EFA and backward EFA [24].

Code written in-house in MatLab was used for generating initial estimates of concentration profiles and response patterns by EFA [30]. ALS was performed using the algorithm developed by the Lin group [36,37], with the constraint that all values in both the concentration and pattern matrices had to be positive.

EFA-ALS was performed in two modes. In "blind" mode, the estimates of normalized response patterns and concentration profiles used for ALS were taken directly from the output of EFA. In "informed" mode, the normalized library patterns for the pair members were used for ALS instead of the EFA-generated estimates. The latter was rationalized on the basis of the fact that the identities of the binary composite peaks would be known a priori. EFA-generated estimates of the concentration profiles were used for both blind-mode and informed-mode analyses. The convergence of the ALS algorithm was denoted by the point at which further iteration changed the residual error matrix $\boldsymbol{E}$ by $\leq 10^{-6}$.

2.7. Extracted pattern fidelity, vapor recognition, and quantification

Extracted response pattern for each mixture component was compared to the library patterns for the two components comprising the composite peak. Analyte masses were estimated by comparing the peak maxima of the reconstructed chromatogram from the most sensitive sensor in the array (vapor specific) to the appropriate calibration curve.

Pattern fidelity testing was performed in two stages using the *ED* as the metric of pattern fidelity. First, if the *ED* between the extracted and correct library vectors (i.e., patterns) was smaller than that between the extracted and alternative-vapor library patterns, then the vapor was considered to have been correctly recognized. If this "*ED*-proximity" criterion was met, then a more rigorous test of fidelity was applied by placing a limit on the value of the *ED* between the extracted and library patterns. An *ED* limit of 0.20 was established by calculating the 99% confidence interval around the largest standard deviation of replicate normalized response patterns created from the maxima of the fully resolved (reference) chromatograms. We refer to this in the discussion below as the "*ED*-threshold" criterion.

3. Results and discussion

3.1. General features of the data set

Calibration curves, derived from (baseline corrected) peak maxima or peak areas, were linear for all vapors and sensors, with forced-zero regression r^2 values > 0.98 in all cases. The former (Fig. S3, Supplementary Data) were used for quantification of the EFA-ALS extracted response profiles. The slopes (i.e., sensitivities) for a given vapor among the sensors in the array differed by as much as 11-fold (for CHX) and by as little as 3.3-fold (for NET).

Although most of the relative sensitivities are in accordance with what would be expected on the basis of vapor-MPN affinities, several exceptions were noted. For example, although the C8-coated sensor had the expected higher sensitivity toward the non-polar alkane and alkene vapors, it also had higher sensitivity toward the polar MIBK and BAC vapors than did the (polar) HME-coated sensor. In addition, the moderately polar OPH-coated sensor had higher peak-maxima sensitivity toward NET than did the HME sensor. In contrast, for the peak area data, the relative sensitivities are in accordance with what would be expected on the basis of vapor-MPN affinities; the C8 sensor retains its higher sensitivity for the non-polar vapors, and the HME sensor shows the highest sensitivity for the more polar MIBK, BAC, and NET vapors.

The differences in relative sensitivities between peak maxima and peak areas arise from the dependence of the sensor responses

Table 2
Asymmetry factors (AF) from each sensor for the 6 test vapors.[a]

Vapor	AF values [b]				
	C8	DPA	OPH	HME	FID
MIBK	1.4	3.0	1.6	4.1	1.2
TOL	1.3	2.1	1.6	2.5	1.1
OCN	1.1	1.5	1.4	2.0	0.9
BAC	1.3	2.0	1.7	2.9	1.2
NET	1.5	2.0	1.8	3.1	1.2
CHX	1.1	1.6	1.4	2.1	1.0

[a] Calculated from calibration samples. Injected masses ranged from 200 to 300 ng.
[b] Ratio of peak widths on either side of the peak maximum (back:front) measured at 10% of peak maxima.

on the rates of sorption or desorption of the vapors into or out of the MPN films, particularly for the polar vapors in the polar HME film [32]. Differences in the extent of tailing in the peaks among the sensors (discussed below) are also ascribable to this factor, and may have been enhanced by differences in film thickness, which were not well controlled during deposition. Sorption/desorption rates for vapors in C8 were apparently more rapid, resulting in less tailing than with sensors coated with other MPNs. The bar charts presented in Fig. S4 (Supplementary Data) reveal some differences in the pair-specific correlations of the normalized response patterns from peak-maxima and peak-area sensitivities. Normalized peak-maxima sensitivities were used to construct a PC score plot for qualitatively assessing the diversity of the array responses among these six vapor pairs. This is presented in Fig. S5 (Supplementary Data). As shown, the projected vectors of all three of the paired vapors considered are well separated, and the inter-vapor distances are consistent with the values of ρ and *ED* given in Table 1 and Fig. S4.

Within each vapor pair, the elution order remained the same under all conditions: MIBK, OCN and NET were always the first to elute. All sensors exhibited tailing peak profiles for all vapors. To quantify the extent of tailing, the asymmetry factors (AF) presented in Table 2 were calculated for each vapor-sensor combination. The HME sensor, with AF values ranging from 2.1 to 4.1, consistently exhibited the most tailing, and the C8 sensor, with AF values from 1.1 to 1.5, consistently exhibited the least tailing. For reference, the AF values for the FID (also shown) ranged from 0.9 to only 1.2, despite being downstream from the array. Tailing was the greatest for the polar vapors MIBK and NET for all sensors. Note that reversing the position of the sensors in the flow path had no effect on tailing, which indicates that it was not due to any mixing anomalies in the detector cell. The larger AF values for the HME sensor may be due, in part, to its having a thicker coating film, but it appears that the affinity for polar vapors also contributes.

Of greater importance than the extent of tailing, however, was the *difference* in the extent of tailing among the sensors for a given vapor; ratios of AF values were as high as 2.9 and as small as 1.8 among the sensors for a given vapor. This, in turn, results in a change in the response pattern across the span of a peak, as discussed in more detail further below.

One additional factor is embedding of one peak entirely within the profile of another, which would violate the assumption in EFA that the first compound to elute is also the first to finish eluting. For Gaussian peaks, embedding should not occur for $R_S = 1.0$ or 0.5 under any of the *RRR* conditions considered here, whereas for $R_S = 0.1$, embedding would be expected at an $RRR \leq 1{:}5$ for the first eluting component and $RRR \geq 5{:}1$ for the second eluting compound. Clearly, embedding would be more prevalent at a given nominal R_S value in cases where the first peak tails significantly. Thus, embedding of TOL (in MIBK) and CHX (in NET) would occur when these were the minority components of a composite peak at $R_S = 0.1$ and $RRR > 2.5{:}1$, and embedding of BAC (in OCN) would occur at $R_S = 0.1$ and $RRR > 5{:}1$.

3.2. Additivity of responses for composite peaks

In all MCR methods it is assumed that the response to the two vapors in a composite peak is equivalent to what the sum of the responses to each vapor would have been if measured individually. To test for response additivity, for each test atmosphere containing a binary mixture, the peak areas obtained from the mixture components at the lowest elution temperature, at which they were completely resolved (i.e., $R_S > 4$), were summed and used as reference values. The composite peak areas from subsequent analyses of the same test atmosphere in which the peaks overlapped to varying extents (i.e., $R_S = 0.1$, 0.5, 1.0) were then compared to these reference values for each sensor. A focus was placed on the binary mixtures at $RRR = 1{:}1$ because samples at this value of *RRR* would be the most sensitive to any fractional deviations from additivity. To account for variations in the injected mass of replicate samples taken from a given test atmosphere, each

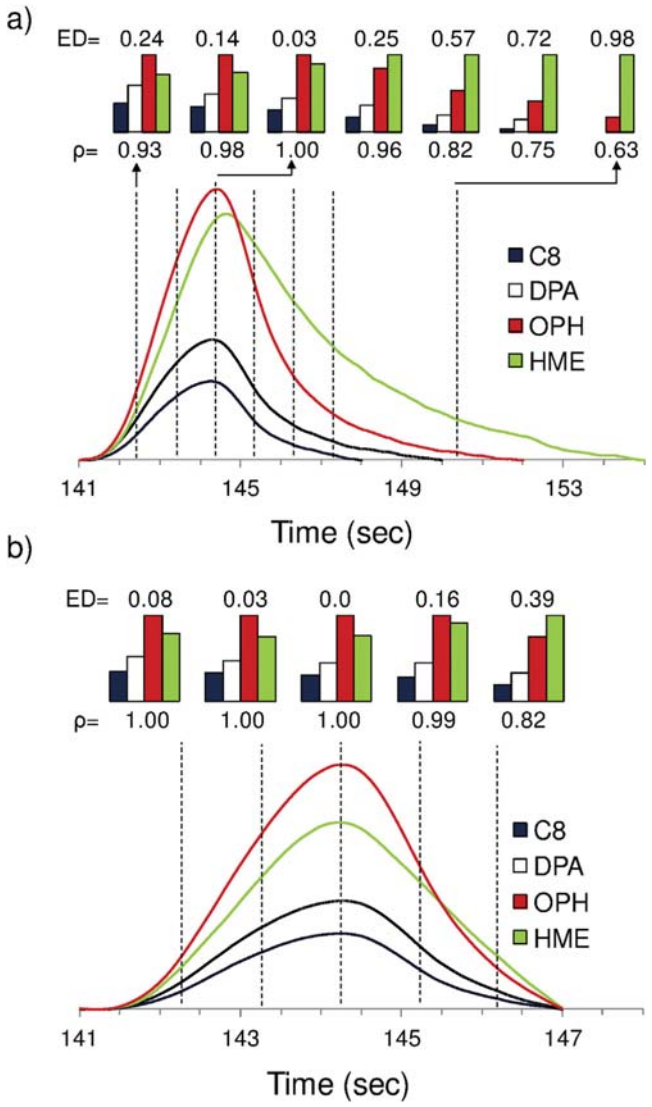

Fig. 2. Chromatograms for NET from all four sensors in the array (a) before and (b) after correcting for differential retention times and peak tailing. Normalized patterns at one second intervals across the peak are compared to that at the peak maximum by ρ and ED values.

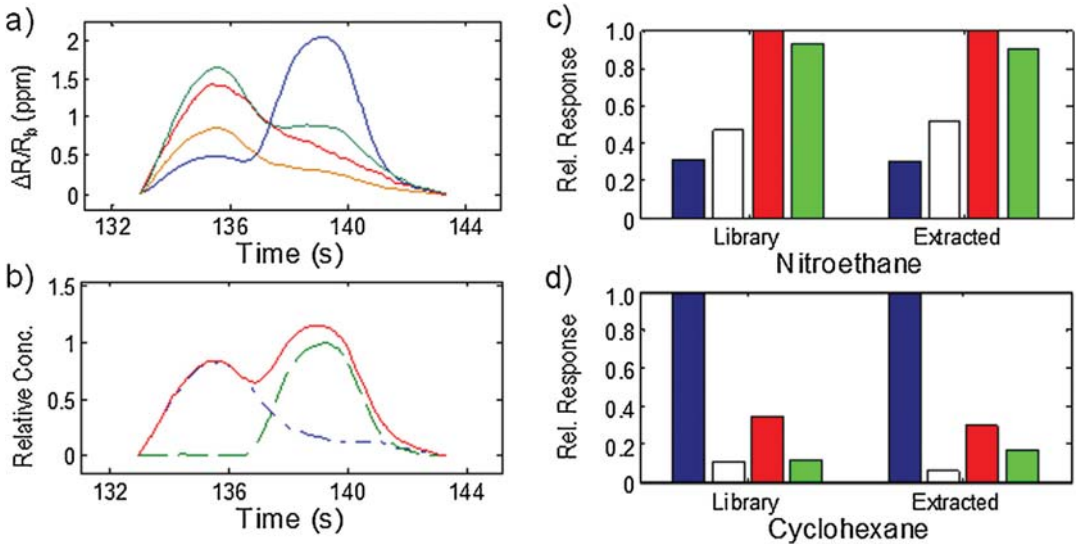

Fig. 3. Example of the output from the EFA-ALS analysis for NET + CHX at RRR = 1:1 and R_s = 0.5; (a) fully preprocessed binary-mixture sensor traces (baseline corrected, retention time adjusted, and windowed); (b) extracted unitless concentration profiles (C matrix) in informed mode; and the normalized library and extracted patterns (in order from l to *r*: C8, DPA, HME, OPH) for (c) NET and (d) CHX, also collected in informed mode.

sensor peak area value was divided by the corresponding FID peak area value for that sample and then multiplied by the average FID peak area for all samples from a given bag. Results are summarized in Table S1 (Supplementary Data). The average difference in peak areas among all sensors was 0.8% and the range of individual differences spanned from only −4.1 to 4.6% for the 36 cases considered. This confirms that the presence of a second vapor does not affect the response to another vapor over the ranges of sorbed masses under consideration. To the best of our knowledge, this is the first systematic testing of response additivity for this class of sensors.

3.3. Intra-peak pattern fidelity

EFA entails a series of progressive (evolving) analyses across the span of each peak and implicitly assumes that response patterns remain constant. In light of the differential tailing discussed above, the extent of pattern distortion arising from differences in t_R and *AF* among the sensors for a given analyte peak was evaluated. A representative analysis is summarized in Fig. 2a, which shows the superimposed traces from all four sensors for a 200-ng injection of NET prior to any processing (other than baseline correction). There are small differences in t_R (0.05–0.5 s) and rather large differences in tailing (HME > DPA > OPH > C8; see Table 2). As shown by the bar chart patterns and accompanying values of ρ and *ED* generated at the indicated locations across the peak in Fig. 2a, the pattern is fairly stable over the first half of the peak. Beyond the maximum, however, the differential tailing leads to progressively greater pattern distortion. The C8 and DPA sensors return to baseline well before the OPH and HME sensors. Using the pattern from the (aligned) peak maxima as the reference, ρ decreases from 1 to 0.63 and *ED* increases from 0.03 to 0.98 among the instantaneous patterns in the right-hand portion of the peak.

The extent to which patterns were distorted among all six vapors can be summarized in terms of the average *ED* values across the entire span of a peak, which were 0.96, 0.85, 0.99, 0.83, 0.85, and 1.04 for MIBK, TOL, OCN, BAC, NET, and CHX, respectively, for representative injection masses of each. Note that these *ED* values approach or exceed the *ED* values separating the normalized patterns (derived from the aligned and preprocessed peak maxima) of the members of a pair (Table 1). Although most of the intra-peak distortion occurs in the latter part of the peak (∼40% of the area) and the pattern extracted by EFA is a weighted average of the instantaneous patterns across the entire span of the peak, this is clearly an unstable and undesirable situation.

To address this problem, t_R values were adjusted to align the four peak maxima and then a boundary was placed on the retention time interval at the point where the sensor giving the least tailing (i.e., C8) returned to baseline. For the other three sensors a new baseline was established by connecting the start of the peak with the point on the tail corresponding to this interval boundary. These initial and final time points were set to zero for each trace to establish a new baseline. The peak maxima and area were then recalculated on the basis of the new (common) baseline.

Fig. 2b shows the result of performing these pre-processing steps for NET. As shown, this greatly reduced, but did not eliminate, the extent of pattern distortion across the peak. Applying these preprocessing steps to all of the individual-vapor profiles, the *ED* between (aligned) peak-maxima sensitivity pattern and the average of the normalized response patterns determined across the span of the peak dropped from 0.96 to 0.24 for MIBK, from 0.85 to 0.35 for TOL, from 0.99 to 0.31 for OCN, from 0.83 to 0.27 for BAC, from 0.85 to 0.28 for NET, and from 1.04 to 0.35 for CHX. The reduction in pattern distortion is similar among all of the test vapors, but is a bit less for CHX, OCN and TOL because the influence of the tailing HME sensor remained relatively high for these vapors.

A dilemma arises for the composite binary peaks: it is not possible to specify the point at which the first peaks finish eluting in order to truncate the peaks as done in the individual-vapor cases discussed above, because it would depend on the degree of resolution and the relative contribution of the first peak to the overall pattern. An approximate solution was required. The approach taken was to adjust the baselines in the same manner as done with the individual-vapor peaks but for the composite peak as a whole. Thus, the end of the composite peak was truncated at the t_R value corresponding to the end of the most symmetric peak (i.e., C8) and a new baseline was drawn from the start of the peaks to this point on the trailing side of the composite. After this adjustment, the peaks all had the same starting and ending t_R values, and the degree of pattern distortion was minimized.

Table 3
Rank of binary composite peaks determined by EFA.

R_s	*RRR*[a]	Number of components		
		MIBK + TOL	OCN + BAC	NET + CHX
1.0	20:1	–	–	2
	10:1	2	2	2
	5:1	2	2	2
	2.5:1	2	2	–
	1:1	1	2	2
	1:2.5	2	2	–
	1:5	2	2	2
	1:10	2	2	2
	1:20	–	–	2
0.5	20:1	–	–	2
	10:1	2	2	2
	5:1	2	2	2
	2.5:1	2	2	–
	1:1	1	2	2
	1:2.5	2	2	–
	1:5	2	2	2
	1:10	2	2	2
	1:20	–	–	2
0.1	20:1	–	–	1
	10:1	2	2	1
	5:1	2	2	2
	2.5:1	2	2	–
	1:1	1	2	2
	1:2.5	2	2	–
	1:5	2	2	2
	1:10	1	2	2
	1:20	–	–	2

[a] *RRR* values are approximate. The first number in each *RRR* corresponds to the compound eluting first, i.e., MIBK, OCN, and NET, respectively. A common threshold value of log (eigenvalue) = 1.25 was applied to all cases.

3.4. Rank determinations

In most spectrometric detectors, the number of wavelengths or fragments used far exceeds the number of components being analyzed, and noise levels are minimized by averaging over many effective measurements. In contrast, with a 4-sensor array, the number of sensors barely exceeds the number of components being discriminated, and noise levels are averaged over a relatively small number of effective measurements. Differentiating the net signal from the noise with such a sensor array for the purpose of rank analysis is therefore more challenging; there are relatively few outputs upon which to establish an eigenvalue noise threshold (note also that, with a 4-sensor array, the maximum possible rank is mathematically limited to four). To address this, analyses of the individual-component (calibration) chromatograms were performed to empirically establish an eigenvalue noise threshold above which the change in log (eigenvalue) was attributable to a bonafide peak.

From the plots of log (eigenvalue) vs. t_R for all six vapors, a common noise threshold of 1.25 was established. By use of this threshold, the ranks of the binary composite peak profiles were determined, as shown in Table 3. At R_s = 0.1 the rank was correctly determined in 17 of the 21 (81%) composite peaks, and at R_s = 0.5 or 1.0 the correct rank was found for 40 of the 42 composite peaks (95%). Results at R_s = 0.5 and 1.0 were identical. All of the errors led to a rank of one, most (4 out of 6) occurred for MIBK + TOL, and most (4 out of 6) were for R_s = 0.1. For MIBK + TOL at *RRR* = 1:1 and R_s = 1.0 and 0.5, the second log (eigenvalue) traces fell only slightly below the threshold and visual inspection of the composite peak profile clearly indicated more than one chemical component. For the other rank errors, the determinations were unequivocal and were most likely due to pattern similarity for the remaining MIBK + TOL cases and to embedding of the minority component for the two NET + CHX cases.

3.5. Fidelity of extracted patterns

Notwithstanding the errors in rank determinations, initial EFA estimates of concentration profiles and response patterns assumed a rank of two. An example of the output from the EFA-ALS algorithm is shown in Fig. 3 for NET + CHX. Fig. 3a shows a set of preprocessed chromatograms of a binary mixture of NET and CHX. Fig. 3b shows the extracted relative concentration profile obtained by EFA-ALS (informed mode), and Fig. 3c and d presents the extracted response patterns of each pair component alongside the library reference pattern (derived from peak maxima). The ED values between extracted and library patterns were 0.11 for both vapors. For this particular case, the *ED* values met both the *ED*-proximity and *ED*-threshold criteria, and the quantification errors were −0.4% and −0.2%, for NET and CHX respectively.

Table 4 summarizes the results of EFA-ALS analyses of the entire data set in blind mode. The extracted-pattern fidelity was sufficient to yield correct assignments of identity in the vast majority of cases (112 out of 126; 89%) using the simple *ED*-proximity criterion. Even for MIBK + TOL, only four errors (10%) occurred over the entire range of R_s and *RRR* values. In all but two cases, the error in assigned identity occurred when the vapor was the minority component in the composite. Results for OCN + BAC were also quite good, with only two errors, both of which were for OCN at a low relative concentration. That OCN was mis-assigned at R_s = 1.0 is surprising given that it is the first eluting component of the pair. The errors obtained for CHX in the NET + CHX pair are also remarkable, particularly at the larger values of R_s. Examination of the extracted relative response patterns revealed that the contributions from HME to the CHX patterns were significantly higher than those in the library pattern, exceeding the contribution from C8 in several cases. Apparently, even at R_s values of 1.0 and 0.5, EFA cannot avoid distortion of the (partially embedded) CHX peak by the long tail of the HME peak for NET, despite the preprocessing steps performed to minimize this problem.

As shown in Table 4, however, many of the *ED* values are $\gg$0.20 and, if the more rigorous *ED*-threshold criterion for matching is applied, then only 33% (41 of 126) of the assignments were correct: 12% for MIBK + TOL, 45% for OCN + BAC, and 40% for NET + CHX. The majority of correct assignments (27/41, 66%) occurred for the predominant component of the mixture. Overall, *ED* values were generally larger for the second eluting compound, consistent with distortion arising from overlap by, or embedding within, the tail of the first peak.

Table 5 shows the results of EFA-ALS analysis in informed mode. The *ED* values between extracted and calibrated patterns (vectors) decreased by 0.20, on average, from those in blind mode, and the differences in *ED* values of the first and second components decreased. The correct assignment rate increased to 98% (124 of 126) on the basis of the simple *ED*-proximity criterion and to 54% (68 of 126) if the *ED*-threshold criterion is imposed. Notably, the large *ED* values for CHX at *RRR* = 5:1, 10:1, and 20:1 in blind mode were significantly reduced in informed mode. The 58 assignments that did not meet the *ED*-threshold criterion were distributed in accordance the pattern similarity of the pair members: 39 (67%) for MIBK + TOL, 14 (24%) for OCN + BAC, and 5 (9%) for NET + CHX. NET + CHX was the only pair for which the assignment errors showed a dependence on R_s, with the average *ED* reduced from 0.25 to 0.09 on going from R_s = 0.1 to R_s = 1.0. The improvement reflects the fact that the informed mode analyses circumvent pattern extraction errors by EFA.

3.6. Quantification

Tables 4 and 5 also present the errors in the component masses estimated from the extracted peak profiles on the basis of

Table 4
Pattern fidelity and quantification accuracy of profiles extracted by EFA-ALS in blind mode.

R_s	*RRR*	MIBK		TOL		OCN		BAC		NET		CHX	
		ED[a]	Error[b]	ED	Error	ED	Error	ED	Error	ED	Error	ED	Error
1.0	20:1[c]	–[d]	–	–	–	–	–	–	–	0.24	−**32**	**1.27**	28
	10:1	†0.08	−**77**[e]	0.73	**220**	†0.06	−0.1	0.21	**73**	0.29	−**34**	**1.29**	**65**
	5:1	†0.09	−**75**	0.74	**150**	†0.07	−0.2	†0.20	**45**	0.41	−12	**1.31**	**86**
	2.5:1	0.86	−**84**	0.70	**35**	†0.04	3.7	0.29	20	–	–	–	–
	1:1	0.46	−3.4	0.54	**35**	†0.07	7.9	0.24	16	0.23	1.1	†0.08	1.4
	1:2.5	0.91	**110**	**0.25**	13	†0.16	**58**	0.50	15	–	–	–	–
	1:5	0.28	−16	0.85	−2.8	0.30	**200**	0.66	−17	†0.10	−9.1	†0.12	2.0
	1:10	**0.29**[e]	2.2	0.82	3.5	**0.63**	**790**	0.80	−14	†0.19	**91**	†0.03	14
	1:20	–	–	–	–	–	–	–	–	†0.20	**79**	†0.02	13
0.5	20:1	–	–	–	–	–	–	–	–	0.93	−**31**	**1.35**	20
	10:1	0.66	−**62**	0.84	**200**	†0.07	−0.7	0.24	**59**	0.94	−25	**1.28**	**49**
	5:1	0.84	−**60**	†0.12	**130**	†0.06	1.8	0.21	**43**	0.68	−32	**1.35**	**44**
	2.5:1	0.82	−**67**	0.71	15	†0.08	5.4	†0.20	19	–	–	–	–
	1:1	0.83	4.6	0.85	22	†0.09	7.4	0.28	−8.6	†0.11	−0.5	†0.18	1.4
	1:2.5	0.59	**95**	†0.11	11	†0.14	**56**	0.42	7.0	–	–	–	–
	1:5	0.56	−**41**	0.80	−8.5	0.29	**180**	0.64	−1.9	0.32	−17	†0.05	6.1
	1:10	0.56	**200**	0.80	−0.4	0.51	**690**	0.67	−14	†0.17	**77**	†0.21	12
	1:20	–	–	–	–	–	–	–	–	0.32	**67**	†0.03	4.8
0.1	20:1	–	–	–	–	–	–	–	–	0.96	−**33**	**1.20**	**390**
	10:1	0.95	−**61**	†0.05	**180**	†0.07	10	0.23	**79**	0.98	−30	**1.16**	**230**
	5:1	0.53	−**62**	0.48	**130**	†0.11	−2.2	†0.15	**60**	0.54	−23	**0.97**	**130**
	2.5:1	0.65	11	0.80	−**86**	†0.13	11	†0.13	26	–	–	–	–
	1:1	0.73	**40**	0.82	21	†0.09	**60**	0.38	−19	†0.05	−17	0.22	3.9
	1:2.5	**0.60**	**55**	**1.14**	20	†0.17	**160**	0.44	−4.2	–	–	–	–
	1:5	0.58	**128**	0.66	−21	0.41	**250**	0.67	−3.6	0.26	−17	†0.08	6.9
	1:10	0.56	−**32**	0.61	2.2	**0.65**	**750**	0.80	−20	0.26	**73**	†0.02	13.8
	1:20	–	–	–	–	–	–	–	–	0.34	**85**	†0.06	10.6

[a] Euclidian distance between normalized extracted and library patterns.
[b] Quantification error (%) on the basis of peak maxima of MSS from extracted profiles.
[c] First number corresponds to compound eluting first, i.e., MIBK, OCN, and NET, respectively.
[d] Not tested.
[e] Bolded ED entries indicate errors in assigned identities on the basis of the ED value in 4-space and bolded error entries indicate values exceeding ±30% error.
† For these cases the ED value was within the more rigorous 'ED-threshold' criterion of 0.20 established for assigning correct identities.

Table 5
Pattern fidelity and quantification accuracy of profiles extracted by EFA-ALS in informed mode.

R_s	*RRR*	MIBK		TOL		OCN		BAC		NET		CHX	
		ED[a]	Error[b]	ED	Error	ED	Error	ED	Error	ED	Error	ED	Error
1.0	20:1[c]	–[d]	–	–	–	–	–	–	–	†0.05	−28	†0.00	−2.3
	10:1	0.48	−**48**[e]	0.23	**180**	†0.05	1.2	†0.06	6.4	†0.08	−27	†0.18	5.1
	5:1	0.46	−**28**	0.24	**120**	†0.05	0.8	†0.08	0.7	†0.07	3.5	†0.19	3.7
	2.5:1	†0.20	−28	0.22	17	†0.05	2.4	†0.14	7.7	–	–	–	–
	1:1	0.26	−8.1	0.26	27	†0.07	−1.3	†0.11	8.6	†0.13	7.2	†0.09	0.3
	1:2.5	0.71	−7.6	†0.12	0.4	†0.19	−22	0.37	−4.1	–	–	–	–
	1:5	0.68	**40**	0.56	−13	0.24	**175**	0.57	−17	†0.07	−7.5	†0.03	3.6
	1:10	0.70	**82**	0.56	−8.2	0.50	**749**	0.78	−14	†0.15	**97**	†0.04	15
	1:20	–	–	–	–	–	–	–	–	†0.11	**88**	†0.03	13
0.5	20:1	–	–	–	–	–	–	–	–	†0.03	−30	†0.00	−12
	10:1	0.49	−**39**	0.24	**92**	†0.07	−0.2	†	6.9	†0.07	−28	0.96	−21
	5:1	0.42	−25	0.21	**82**	†0.05	3.5	†0.05	26	†0.06	−2.7	†0.04	−12
	2.5:1	†0.20	−30	0.24	−26	†0.07	6.1	†0.08	13	–	–	–	–
	1:1	0.43	−18	0.34	−0.8	†0.09	−1.7	†0.13	0.4	†0.11	−0.4	†0.11	−0.2
	1:2.5	0.26	−15	0.28	−2.1	†0.16	−17	†0.14	−0.6	–	–	–	–
	1:5	0.47	14	0.56	−15	0.23	**140**	0.52	−1.9	†0.06	−8.6	†0.04	5.3
	1:10	0.29	**110**	0.56	−11	0.44	**670**	0.66	−18	†0.15	**79**	†0.04	15
	1:20	–	–	–	–	–	–	–	–	†0.09	**84**	†0.03	4.8
0.1	20:1	–	–	–	–	–	–	–	–	†0.02	−29	**1.20**	−16
	10:1	0.50	−**61**	0.29	**190**	†0.07	−8.1	†0.12	**46**	†0.02	−28	1.03	−22
	5:1	0.56	−**49**	0.40	**97**	†0.08	0.9	†0.11	20	†0.02	−10	0.88	−10
	2.5:1	0.66	−17	0.58	−**39**	†0.08	15	†0.10	20	–		–	–
	1:1	0.57	−13	0.28	9.9	†0.09	20	0.38	23	†0.05	−17	†0.09	−0.2
	1:2.5	0.59	**43**	**0.64**[e]	7.8	†0.19	28	†0.20	−11	–	–	–	–
	1:5	0.49	**64**	0.58	−13	0.22	**190**	0.47	−3.6	†0.11	−11	†0.03	6.1
	1:10	0.65	**110**	0.56	−11	0.43	**640**	0.68	−20	0.25	**74**	†0.01	14
	1:20	–	–	–	–	–	–	–	–	†0.11	**110**	†0.02	10

[a] Euclidian distance between normalized extracted and library patterns.
[b] Quantification error (%) on the basis of peak maxima of MSS from extracted profiles.
[c] First number corresponds to compound eluting first, i.e., MIBK, OCN, and NET, respectively.
[d] Not tested.
[e] Bolded ED entries indicate errors in assigned identities on the basis of the ED value in 4-space and bolded error entries indicate values exceeding ±30% error.
† For these cases the ED value was within the more rigorous 'ED-threshold' criterion of 0.20 established for assigning correct identities.

peak-maxima calibrations. As mentioned above, results were based on the MSS, rather than on the average or median of all sensors in the array. For assessing the results, an arbitrary benchmark of ±30% error in the estimated mass of each component, relative to that obtained (from the MSS) from the calibration library, was used. Estimates exceeding this benchmark are presented in bold typeface in Tables 4 and 5.

In blind mode (Table 4), only 54% (68/126) of the mass estimates had errors less than the benchmark error value, and the number of errant cases decreased in the order MIBK + TOL > OCN + BAC > NET + CHX. In only 3 of 21 samples were both of the components of the MIBK + TOL composite accurately quantified. The largest errors were observed for (first-eluting) OCN when it was the minority component in the OCN + BAC composite, regardless of the R_S value. Inspection of the extracted patterns revealed that a significant fraction of the C8 sensor response to BAC had been attributed to the OCN, and since C8 is the MSS for OCN, this led to large overestimates in the OCN concentration. Since embedding of the OCN peaks within the larger BAC peaks would only occur at $R_S = 0.1$, this factor does not explain the error.

In general, errors in quantification of the majority component were smaller and less frequent than the errors of the minority component and there was a trend toward underestimation of the former and overestimation of the latter. The quantification errors did not show a particularly strong dependence on R_S for any of the pairs (in either mode), but was significantly increased (in both modes) when the *RRR* exceeded 5:1 or 1:5. OCN + BAC was the only pair showing a discernible correlation between pattern fidelity and quantification accuracy. For this pair, across all *RRR* and R_S, a weak linear trend was apparent between *ED* and quantification error. The trend was positive for OCN and negative for BAC, with $r^2 > 0.69$.

In informed mode (Table 5) 75% (95/126) of the sample mass estimates fell within the benchmark of ±30% error. The increase, relative to blind mode, follows from the improvement in the accuracy of extracted patterns obtained in informed mode. The trends noted for blind mode persisted with informed mode, but there were fewer errors and the magnitudes of the errors were generally reduced. Notably, the errors in quantification for OCN remained extremely high when it was the minority component.

4. Conclusions

This is the first report on the application of MCR to experimental data from a microsensor array employed as the detector for a GC. In delineating the methodological details of the hybrid EFA-ALS algorithm used here, several important factors relevant to the implementation of multivariate data analysis for this application were characterized, including the pattern similarity, relative concentration, chromatographic resolution, and peak asymmetries of the components of the binary peak composites.

The general finding that the response to two vapors is additive confirms that this critical assumption of EFA can be met for MPN-coated CR sensors. The extent of response pattern distortion across the span of the peaks observed was unexpected, and our expedient approach to resolving this in the binary composite peaks was inherently approximate; the residual distortions undoubtedly contributed to decreases in the fidelity of extracted response patterns. This could pose a serious impediment to the general utility of EFA in this type of application. Adjusting MPN film thicknesses or operating conditions might help to minimize this phenomenon. Alternative EFA algorithms that are more resilient to pattern distortion might also be applied [29].

EFA was able to confirm the number (and elution order) of components in composite peaks in the vast majority of cases. Yet, EFA-ALS in blind mode faltered in quantitatively analyzing many of the binary composites under the conditions examined. Accordingly, we not recommended utilizing blind mode for quantification. This reflects the limitations of the CR array in generating data of sufficient diversity, as well as the difficulty in applying EFA-ALS to a low-dimensional data set with time-varying response patterns. Results suggest that vapor pairs having response patterns with ρ values ≥ 0.85 (perhaps lower) are too highly correlated to be effectively analyzed. Although performance was sensitive to R_S, it was much more sensitive to *RRR*, with *RRR* values <1:5 or >5:1 being problematic, especially for the minority component. As ρ deceases and *RRR* approaches 1:1, accurate quantification is possible in informed mode.

All the analyses considered here were performed with a constrained library comprising only the two members of a given pair. For EFA-ALS in blind mode using the *ED*-proximity criterion, this allows for extracted patterns that differ greatly from the calibration pattern to still yield correct assignments, because there was only one other competing vapor in the library. If there were more members in the library, the errors in assigning identities would undoubtedly increase. Using EFA-ALS in informed mode, which entailed inputting calibration response patterns in lieu of EFA-extracted patterns for the ALS step, led to significant improvements in performance in both extracted pattern fidelity and quantification accuracy. To the extent feasible, this mode is highly recommended. Ongoing work is aimed at further understanding, and attempting to resolve, some of the factors found to degrade performance, including the application of more sophisticated data processing and MCR algorithms, so that the full value of microsensor array detection in μGC systems can be realized.

Acknowledgments

The authors are grateful to Lindsay K. Wright for preparing the sensor arrays used in this study and to Nicolas Nuñovero for the design and implementation of the electronic software and hardware used for CR data acquisition. This work was supported in part by the Department of Defense, ESTCP Grant ER-200702 through a subcontract with Integrated Science and Technology, Inc.; by the Department of Homeland Security, Science and Technology Directorate Cooperative Agreement No. 06-G-024; by the National Science Foundation Grant ECCS 1128157; and by the National Institute for Occupational Safety and Health Grant 1-R01-OH-010297. These sponsoring organizations had no role in the design or conduct of this research or in the decision to pursue publication. Devices were fabricated in the Lurie Nanofabrication Facility, a member of the National Nanotechnology Infrastructure Network, which is supported by the National Science Foundation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.snb.2014.05.049.

References

[1] C. Jin, P. Kurzawski, A. Hierlemann, E.T. Zellers, Evaluation of multitransducer arrays for the determination of organic vapor mixtures, Anal. Chem. 80 (2008) 227–236.
[2] J. Park, W.A. Groves, E.T. Zellers, Vapor recognition with small arrays of polymer-coated microsensors: a comprehensive analysis, Anal. Chem. 71 (1999) 3877–3886.
[3] M.D. Hsieh, E.T. Zellers, Limits of recognition for simple vapor mixtures determined with a microsensor array, Anal. Chem. 76 (2004) 1885–1895.
[4] M.D. Hsieh, E.T. Zellers, Adaptation and evaluation of a personal electronic nose for selective multivapor analysis, J. Occup. Environ. Hyg. 1 (2004) 149–160.
[5] Y. Qin, Y.B. Gianchandani, A facile, standardized fabrication approach and scalable architecture for a micro gas chromatography system with integrated pump, in: Proc. 17th Intnl. Conf. Solid-State Sensors, Actuators, and Microsystems, Transducers '13, Barcelona, June 16–20, 2013, pp. 2755–2758.

[6] M. Li, E.B. Myers, H.X. Tang, S.J. Aldridge, H.C. McCaig, J.J. Whiting, R.J. Simonson, N.S. Lewis, M.L. Roukes, Nanoelectromechanical resonator arrays for ultrafast, gas-phase chromatographic chemical analysis, Nano Lett. 10 (2010) 3899–3903.
[7] P.R. Lewis, R.P. Manginell, D.R. Adkins, R.J. Kottenstette, D.R. Wheeler, S.S. Sokolowski, D.E. Trudell, J.E. Byrnes, M. Okandan, J.M. Bauer, R.G. Manley, G.C. Frye-Mason, Recent advances in the gas-phase MicroChemLab, IEEE Sens. J. 6 (2006) 784–795.
[8] C.-J. Lu, W.H. Steinecker, W.-C. Tian, M. Agah, J.A. Potkay, M.C. Oborny, J. Nichols, H.K.L. Chan, J. Driscoll, R.D. Sacks, S.W. Pang, K.D. Wise, E.T. Zellers, First-generation hybrid MEMS gas chromatograph, Lab Chip 5 (2005) 1123–1131.
[9] Q. Zhong, W. Steinecker, E.T. Zellers, Characterization of a high-performance portable GC with a chemiresistor array detector, Analyst 134 (2009) 283–293.
[10] S.K. Kim, H. Chang, E.T. Zellers, Microfabricated gas chromatograph for selective determination of trichloroethylene at sub-parts-per-billion concentrations in complex mixtures, Anal. Chem. 83 (2011) 7198–7206.
[11] P.J. Chapman, F. Vogt, P. Dutta, P.G. Datskos, G.L. Devault, M.J. Sepaniak, Facile hyphenation of gas chromatography and a microcantilever array sensor for enhanced selectivity, Anal. Chem. 79 (2007) 364–370.
[12] W.R. Collin, G. Serrano, L. Wright, H. Chang, N. Nuñovero, E.T. Zellers, Microfabricated gas chromatograph for rapid, trace-level determinations of gas–phase explosive marker compounds, Anal. Chem. 86 (2014) 655–663.
[13] C.J. Lu, J. Whiting, R.D. Sacks, E.T. Zellers, Portable gas chromatograph with tunable retention and sensor array detection for determination of complex vapor mixtures, Anal. Chem. 75 (2003) 1400–1409.
[14] M. Maeder, Evolving factor analysis for the resolution of overlapping chromatographic peaks, Anal. Chem. 59 (1987) 527–530.
[15] J. Kuligowski, G. Quintás, R. Tauler, B. Lendl, M. Guardia, Background correction and multivariate curve resolution of online liquid chromatography with infrared spectrometric detection, Anal. Chem. 83 (2011) 4855–4862.
[16] H.R. Keller, D.L. Massart, Evolving factor analysis, Chemometr. Intell. Lab. 12 (1991) 209–224.
[17] A. de Juan, R. Tauler, Chemometrics applied to unravel multicomponent processes and mixtures, revisiting latest trends in multivariate resolution, Anal. Chim. Acta 500 (2003) 195–210.
[18] R. Tauler, Multivariate curve resolution applied to second order data, Chemometr. Intell. Lab. Sys. 30 (1995) 133–146.
[19] E. Pére-Trepat, S. Lacorte, R. Tauler, Solving liquid chromatography mass spectrometry co-elution problems in the analysis of environmental samples by multivariate curve resolution, J. Chromatogr. A 1096 (2005) 111–122.
[20] W. Windig, N.B. Gallagher, J.M. Shaver, B.M. Wise, A new approach for interactive self-modeling mixture analysis, Chemometr. Intell. Lab. Sys. 77 (2005) 85–96.
[21] J.C. Hoggard, R.E. Synovec, Parallel factor analysis (PARAFAC) of target analytes in GC × GC-TOFMS data: automated selection of a model with an appropriate number of factors, Anal. Chem. 79 (2007) 1611–1619.
[22] R. Manne, On the resolution problem in hyphenated chromatography, Chemometr. Intell. Lab. Sys. 27 (1995) 89–93.
[23] J.S. Nadeau, B.W. Wright, R.E. Synovec, Chemometric analysis of gas chromatography–mass spectrometry data using fast retention time alignment via a total ion current shift function, Talanta 81 (2010) 120–128.
[24] A. de Juan, R. Tauler, Factor analysis of hyphenated chromatographic data: exploration, resolution and quantification of multicomponent systems, J. Chromatogr. A 1158 (2007) 184–195.
[25] S. Mas, A. de Juan, R. Tauler, A.C. Olivieri, G.M. Escandar, Application of chemometric methods to environmental analysis of organic pollutants: a review, Talanta 80 (2010) 1052–1067.
[26] J. Jaumot, R. Gargallo, A. de Juan, R. Tauler, A graphical user-friendly interface for MCR-ALS: a new tool for multivariate curve resolution in MATLAB, Chemometr. Intell. Lab. Sys. 76 (2005) 101–110.
[27] H. Abdollahi, R. Tauler, Uniqueness and rotation ambiguities in multivariate curve resolution methods, Chemometr. Intell. Lab. Sys. 108 (2011) 100–111.
[28] R. Tauler, A.K. Smilde, B.J. Kowalski, Local rank, three-way data analysis and ambiguity in multivariate curve resolution, J. Chemometr. 9 (1995) 31–58.
[29] L.W. Hantao, H.G. Aleme, M.P. Pedroso, G.P. Sabin, R.J. Poppi, F. Augusto, Multivariate curve resolution combined with gas chromatography to enhance analytical separation in complex samples: a review, Analyt. Chim. Acta 731 (2012) 11–23.
[30] C. Jin, E.T. Zellers, Chemometric analysis of gas chromatographic peaks measured with a microsensor array: methodology and performance assessment, Sensor Actuat. B-Chem. 139 (2009) 548–556.
[31] M.P. Rowe, K.E. Plass, K. Kim, C. Kurdak, E.T. Zellers, A.J. Matzger, Single-phase synthesis of functionalized gold nanoparticles, Chem. Mater. 16 (2004) 3513–3517.
[32] L. Wright, E.T. Zellers, A nanoparticle-coated chemiresistor array as a microscale gas chromatograph detector for explosive marker compounds: flow rate and temperature effects, Analyst 138 (2013) 6860–6868.
[33] E. Covington, F.I. Bohrer, C. Xu, E.T. Zellers, Ç. Kurdak, Densely integrated array of chemiresistor vapor sensors with electron beam patterned monolayer-protected gold nanoparticle interface films, Lab Chip 10 (2010) 3058–3060.
[34] J.V. Hinshaw, Anatomy of a peak, LCGC Europe 17 (2004) 216–223.
[35] R. Tauler Centre d'Investigació i Desenvolupament, (2004), http://www.cid.csic.es/homes/rtaqam/, Accessed Feb., 2013.
[36] C.-J. Lin, Projected gradient methods for non-negative matrix factorization, Neural Comput. 19 (2007) 2756–2779.
[37] C.-J. Lin National Taiwan University, (2005), http://www.csie.ntu.edu.tw/~cjlin/index.html, accessed July, 2012.

Biographies

Jonathan Bryant-Genevier received a B.S. in chemistry from the University of Maryland Baltimore County in 2008, and is currently pursuing a Ph.D. in environmental health sciences at the University of Michigan, under the advisement of Professor Zellers. His research focuses on the development of micro-instrumentation for applications in exposure assessment, and on the use of chemometric algorithms in conjunction with such systems.

Kee Scholten earned a B.S. degree in applied physics from the California Institute of Technology, Pasadena, CA, USA, in 2009. He is currently pursuing a Ph.D. degree in applied physics at the University of Michigan, Ann Arbor, MI, USA, working with Professor Zellers. His current research interests include MEMS-based microfluidic systems employing micro-optical and micro-electronic sensors and sensor arrays.

Sun Kyu Kim earned B.S. and M.S. degrees in Environmental Engineering from Daejon University in Korea. He earned a Ph.D. degree in Environmental Health Sciences from the University of Michigan in 2012, working with Professor Zellers on the development and characterization of μGC instrumentation for the determination of VOC mixture components. He is currently employed as a senior researcher at Samsung Display, Co., Ltd.

Edward T. Zellers (B.A. Chemistry, Rutgers; MS, Ph.D., Environmental Health Sciences, University of California, Berkeley) joined the faculty at the University of Michigan in 1987 and is currently appointed as a Professor in the Department of Environmental Health Sciences (EHS) and in the Department of Chemistry. He is Director of the Occupational Health Program in EHS. Since 2000, he has been the Leader of the Environmental Sensors and Systems Thrust in the Michigan Center for Wireless Integrated MicroSensing and Systems (WIMS^2). He has been involved in research on chemical microsensors and micro-instrumentation for volatile organic compounds for over 27 years, with particular interests in the design and physicochemical characterization of devices and materials related to the capture, preconcentration, separation, and sensing of multi-vapor mixtures.