

Marquette University

e-Publications@Marquette

Electrical and Computer Engineering Faculty
Research and Publications

Electrical and Computer Engineering,
Department of

5-2019

Quantitative Detection of Complex Mixtures using a Single Chemical Sensor: Analysis of Response Transients using Multi-Stage Estimation

Karthick Sothivelr
Marquette University

Florian Bender
Marquette University, florain.bender@marquette.edu

Fabien Josse
Marquette University, fabien.josse@marquette.edu

Edwin E. Yaz
Marquette University, edwin.yaz@marquette.edu

Antonio J. Ricco
Stanford University

Follow this and additional works at: https://epublications.marquette.edu/electric_fac



Part of the [Computer Engineering Commons](#), and the [Electrical and Computer Engineering Commons](#)

Recommended Citation

Sothivelr, Karthick; Bender, Florian; Josse, Fabien; Yaz, Edwin E.; and Ricco, Antonio J., "Quantitative Detection of Complex Mixtures using a Single Chemical Sensor: Analysis of Response Transients using Multi-Stage Estimation" (2019). *Electrical and Computer Engineering Faculty Research and Publications*. 745.

https://epublications.marquette.edu/electric_fac/745

Marquette University

e-Publications@Marquette

Faculty Research and Publications/College of

This paper is NOT THE PUBLISHED VERSION.

Access the published version via the link in the citation below.

ACS Sensors, Vol. 4, No. 6 (May 2019): 1682-1690. [DOI](#). This article is © American Chemical Society and permission has been granted for this version to appear in [e-Publications@Marquette](#). American Chemical Society does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from American Chemical Society.

Quantitative Detection of Complex Mixtures using a Single Chemical Sensor: Analysis of Response Transients using Multi-Stage Estimation

Karthik Sothivelr

Department of Electrical and Computer Engineering, Marquette University, Milwaukee, Wisconsin

Florian Bender

Department of Electrical and Computer Engineering, Marquette University, Milwaukee, Wisconsin

Fabien Josse

Department of Electrical and Computer Engineering, Marquette University, Milwaukee, Wisconsin

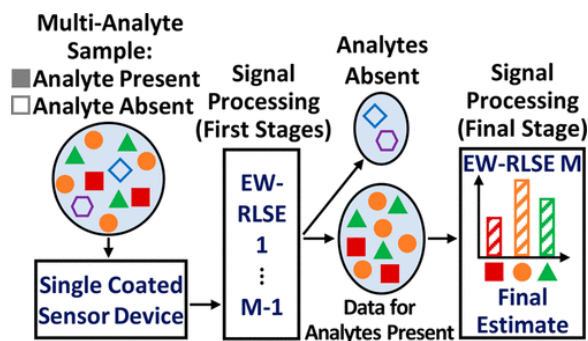
Edwin E. Yaz

Department of Electrical and Computer Engineering, Marquette University, Milwaukee, Wisconsin

Antonio J. Ricco

Department of Electrical Engineering, Center for Integrated Systems, Stanford University, Stanford, California

Abstract



Most chemical sensors are only partially selective to any specific target analyte(s), making identification and quantification of analyte mixtures challenging, a problem often addressed using arrays of partially selective sensors. This work presents and experimentally verifies a signal-processing technique based on estimation theory for online identification and quantification of multiple analytes using only the response data collected from a single polymer-coated sensor device. The demonstrated technique, based on multiple stages of exponentially weighted recursive least-squares estimation (EW-RLSE), first determines which of the analytes included in the sensor response model are absent from the mixture being analyzed; these are then eliminated from the model prior to executing the final stage of EW-RLSE, in which the sample's constituent analytes are more accurately quantified. The overall method is based on a sensor response model with specific parameters describing each coating-analyte pair and requires no initial assumptions regarding the concentrations of the analytes in a given sample. The technique was tested using the measured responses of polymer-coated shear-horizontal surface acoustic wave devices to multi-analyte mixtures of benzene, toluene, ethylbenzene, xylenes, and 1,2,4-trimethylbenzene in water. The results demonstrate how this method accurately identifies and quantifies the analytes present in a sample using the measured response of just a single sensor device. This effective, simple, lower-cost alternative to sensor arrays needs no arduous training protocol, just measurement of the response characteristics of each individual target analyte and the likely interferences and/or classes thereof.

KEYWORDS:

sensor signal processing, exponentially weighted recursive least-squares estimation, multi-analyte samples, liquid-phase detection, chemical sensor selectivity, false positives

Introduction

To enable real-world applications of chemical sensors, the sensors should exhibit adequate limits of detection (which are determined by sensitivity and levels of noise and drift) and selectivity to target analytes in the environment of interest. Both sensitivity and selectivity are highly dependent on the chemical interactions between the chemically sensitive layer and the target analytes.

While molecular specificity is feasible for large, complex target species, particularly biomolecules—by using antibodies, for example—perfect specificity is very challenging to obtain for small organic molecules in combination with rapidly reversible binding. Thus, chemically sensitive layers for “real-time-reversible” sensing often utilize a combination physical and weak chemical interactions: van der

Waals bonds, polar/nonpolar interactions, polarizability-related association being in the former category and hydrogen bonding in the latter. For this approach, the sensitivity, and sometimes the selectivity, of the sensor can be enhanced by selecting or synthesizing sensor coatings that match or mimic the chemical features of the target analyte(s),^{1,2} using the general principle of “like dissolves like.” However, the resulting sensor coatings are usually not highly selective to any specific analyte, though preference for a chemical class of analytes can result in the most successful cases. For instance, it is difficult to develop a rapidly reversible chemical sensor coating that interacts exclusively with benzene. The benzene sensor coating is likely to respond to chemically similar molecules as well, such as toluene, ethylbenzene, and, if aromatic character is not a factor, cyclohexane and other alkanes, making it difficult to accurately determine the concentration of benzene. The lack of sufficient selectivity limits the use of many chemical sensors in real-world applications. This is particularly true if identification and quantification of analytes in a mixture of compounds is required.

The most common approach to enhance the selectivity of a chemical sensor is to use an array of sensors, each with a different partially selective sensor coating.³⁻⁶ The sensors in the array respond to most if not all analytes in the sensing environment, but the pattern of responses from the sensors provides a unique fingerprint for each analyte or group of analytes.⁵ The sensor-array approach to detect and quantify target analytes is often facilitated by dimensionality reduction and pattern-recognition techniques⁷⁻¹¹ such as linear-discriminant analysis, principal-component analysis, and cluster analysis. However, using a sensor array for chemical sensing has potential drawbacks, among them the inability to obtain accurate results for multi-analyte mixtures and the cost and time required to define, develop, qualify, manufacture, and calibrate several different sensor-coating materials.

The sensor-array approach often works well only for single analytes (identification accuracy 95–98%¹³) and binary mixtures (identification accuracy ~89%^{13,14}). As the complexity of the mixture increases, the approach often will fail or produce inaccurate results for detection and quantification of target analytes. For instance, it has been shown in ref¹⁴ that the identification accuracy of ternary mixtures using sensor arrays is only ~3%. Other drawbacks of using the sensor-array approach include potential misclassification and longer time to detection, if steady-state features are used to identify and quantify the analytes. Misclassification errors are particularly likely if the chemical diversity (“chemical orthogonality”⁵) and partial selectivity of the sensor coatings in the array are insufficient. Moreover, the use of only one sensing parameter per sensor for classification, as is often the case, further increases the chances of misclassification.⁴

In many sensor array approaches, steady-state features are key to defining the response of each sensor in the array and to identify and quantify the analytes.⁶ In such cases, analytes cannot be identified and quantified, until the responses of all the sensors in the array reach steady state. The time to reach steady state is sometimes long, especially in liquid-phase detection, where diffusion coefficients in the sampled medium are orders of magnitude smaller than in the gas phase; response times can be on the order of several minutes to hours, an undesirable limitation for many real-world applications. Given all the drawbacks of the sensor-array approach, there is a need to develop a novel method to achieve selectivity without the use of sensor arrays for multi-analyte identification and quantification in liquid environments.

In this paper, a novel signal-processing technique based on estimation theory is investigated for multi-analyte identification and quantification in liquid environments using only the response from a single-sensor device. Specifically, this multistage exponentially weighted recursive least-squares (EW-RLSE) technique utilizes two sensing parameters, namely, the response time constant and the response sensitivity (i.e., signal magnitude vs analyte concentration) for each target or interfering species, to identify and quantify accurately the analytes in the liquid environment of interest. In this technique, EW-RLSE is implemented in autonomous multiple stages (in series), where the initial stages of EW-RLSE are used to determine and eliminate the analytes that are included in the model but happen to be absent from the particular sample under analysis. The final stage of EW-RLSE then uses a sensor-response model that includes only those analytes already determined to be present, providing more accurate analyte quantification. Note that the entire process is fully autonomous.

This newly demonstrated approach is based on a sensor-response model with specific parameters to characterize each coating-analyte pair; it does not require any initial assumptions about the concentrations of the analytes in the sample. One major advantage of this approach is that it only requires measurement data (frequency-shift transients) from a single-sensor device for the detection and quantification of multiple analytes present in the environment of interest. Moreover, this approach can be implemented in real time, as the measurement data are being collected; it only requires measurement of the first few minutes of the sensor response, not the entire response transient, to accurately detect and quantify the analytes present. Such real-time detection is important for many real-world applications, from industrial process monitoring to tracking the movement of environmental pollutants. In this paper, the signal-processing technique is demonstrated for the detection and quantification of multi-analyte mixtures of benzene, toluene, ethylbenzene, and xylenes (BTEX compounds) and 1,2,4-trimethylbenzene (TMB) using the frequency-shift response of a single polymer-coated shear-horizontal surface acoustic wave (SH-SAW) sensor operating in direct contact with the aqueous-phase samples.

Multi-Analyte Sensor Response Model

In general, the use of estimation-theory-based techniques such as EW-RLSE (used here) requires an accurate analytical model that describes the response of the sensor to the analytes in the sample; this is not unusual, as nearly all methods of analyzing sensor response data require thorough characterization of the sensor responses in some manner as a starting point. On the basis of the experimental data for various polymer-coated SH-SAW sensor responses to both single- and multi-analyte mixtures, three assumptions were made to formulate the general analytical model for the sensor response to a mixture of n analytes:

- (1) For concentrations of analytes from zero to ~ 50 ppm (the precise upper limit depending on the analyte and the polymer), sorption of a multi-analyte mixture by the polymer coating is assumed to obey Henry's law.¹⁵⁻¹⁷ Hence, for a dilute mixture of multiple soluble species, the sorption of one analyte does not affect the sorption of the other analytes in any way: the responses are additive. Such behavior requires free partitioning of analytes between polymer and contacting aqueous phase: the sorption process is fully reversible at room temperature (i.e., only physisorption and/or weak chemisorption occurs).

- (2) The polymer film is assumed to be sufficiently thin and the process of absorption sufficiently rapid that sorption follows first-order kinetics with respect to analyte concentration¹⁸ for the range of concentrations defined in assumption (1). Together, the first two assumptions imply that the concentration of the mixture in the coating at any time t is the sum of the concentrations of each individual i th analyte, $C_i(t)$, as it would be measured in a single-analyte response; that is, $C_{\text{mixture}}(t) = C_1(t) + C_2(t) + \dots + C_n(t)$.
- (3) It is assumed that the equilibrium frequency shifts resulting from sorption of each species are also mutually independent; that is, the frequency shift due to the mixture at any time t is the sum of the frequency shifts due to each analyte in the mixture at that time.

From these assumptions, the sensor response to a mixture of n analytes can be modeled as the sum of the individual responses of each analyte in the mixture and is given by

$$\Delta f(t) = \sum_{i=1}^n aK_{p-w,i}C_{aq,i} \left[1 - e^{-\frac{t}{\tau_i}} \right] u_s(t)$$

(1)

where subscript $i = 1, 2, \dots, n$ refers to each analyte in the mixture, $\Delta f(t)$ is the frequency shift as a function of time t , a represents the sensitivity of the sensor platform (which also depends on coating thickness), $K_{p-w,i}$ is the polymer/water partition coefficient for a given analyte/coating pair, $C_{aq,i}$ is the ambient bulk aqueous-phase concentration of each analyte, τ_i is the response time constant of each analyte for a given coating, and $u_s(t)$ is the unit step function (where $u_s(t) = 1$ for $t \geq 0$ and $u_s(t) = 0$ for $t < 0$). Note that the term $aK_{p-w,i}$ represents the overall sensitivity of the sensor for a particular analyte. Alternatively, (1) can be written in terms of the concentration of each analyte in the coating at time t , $C_{\text{poly},i}(t)$, as

$$\Delta f(t) = \sum_{i=1}^n aC_{\text{poly},i}(t)$$

(2a)

where $C_{\text{poly},i}(t)$ is the solution to the following first-order differential equation:

$$\dot{C}_{\text{poly},i}(t) = -\frac{1}{\tau_i} C_{\text{poly},i}(t) + \frac{K_{p-w,i}}{\tau_i} C_{\text{amb},i}$$

(2b)

Equations (1) or (2) represent the general analytical model for the sensor response to any number of analytes (or nominal interferences) in the sample, provided that each analyte in the sample has been separately characterized for the selected coating in the concentration range of interest. Note also that, in some cases, multiple analytes or multiple interferences may have sufficiently similar values of τ_i that

their responses become indistinguishable; this can be either a flaw or a feature depending on whether the application demands individual quantitation of every species versus measuring a chemical class of analytes or interferences as a group.

In this paper, the multi-analyte sensor response model of (2) was normalized, discretized, and converted into state-space form before applying estimation-theory-based technique to perform analyte recognition and quantification. The model was normalized by dividing (2) by $K_{p-w,i}C_{aq,i}$. By defining variables as

$$m_i(t) = \frac{C_i(t)}{K_{p-w,i}C_{aq,i}}$$

(3a)

$$u_s(t) = \frac{C_{amb,i}}{C_{aq,i}}$$

(3b)

and

$$f_{\infty,i} = aK_{p-w,i}C_{aq,i}$$

(3c)

the following normalized differential equation and output equation are obtained as

$$\dot{m}_i(t) = -\frac{1}{\tau_i}m_i(t) + \frac{1}{\tau_i}u_s(t)$$

(4a)

and

$$\Delta f(t) = \sum_{i=1}^n f_{\infty,i}m_i(t)$$

(4b)

where, for analyte i , $m_i(t)$ represents the normalized concentration absorbed at time t , and $f_{\infty,i}$ is the equilibrium frequency shift.

Since the frequency-shift data are collected at discrete-time instants, $t = kT_s$, where T_s is the sampling period, and k is a non-negative integer, the normalized model of (4) was converted into a discrete-time model using Euler's first-order forward method:

$$\dot{m}_i(t) = \frac{m_{i,k+1} - m_{i,k}}{T_s}$$

(5)

which yields the following discrete-time equations.

$$m_{i,k+1} = (1 - S_i)m_{i,k} + S_i u_{s,k}$$

(6a)

$$\Delta f_k = \sum_{i=1}^n f_{\infty,i} m_{i,k} + w_k$$

(6b)

In (6), S_i is defined as

$$S_i = \frac{T_s}{\tau_i}$$

(7)

and is commonly referred to as the sorption rate constant. Note that in (6), the term w_k is added to represent the measurement noise with variance σ_w^2 , which is likely to be present during data collection; it can be either negative or positive. It is assumed that the measurement noise is white noise (uncorrelated in time). From (6), the state-space form of the multi-analyte sensor response model can be obtained by assigning state variables to the normalized concentrations $m_{i,k}$, absorbed at time instant k :

$$\begin{bmatrix} x_{k+1}^{(1)} \\ x_{k+1}^{(2)} \\ \vdots \\ x_{k+1}^{(n)} \end{bmatrix} = \begin{bmatrix} 1 - S_1 & 0 & & 0 \\ 0 & 1 - S_2 & & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & & 1 - S_n \end{bmatrix} \begin{bmatrix} x_k^{(1)} \\ x_k^{(2)} \\ \vdots \\ x_k^{(n)} \end{bmatrix} + \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{bmatrix} u_{s,k} = Ax_k + Bu_{s,k}$$

(8a)

and

$$y_k = \Delta f_k = \sum_{i=1}^n f_{\infty,i} x_k^{(i)} + w_k = C^T x_k + w_k$$

(8b)

where

$$x_k = [x_k^{(1)} x_k^{(2)} \dots x_k^{(n)}]^T$$

$$\mathbf{A} = \begin{bmatrix} 1 - S_1 & 0 & 0 \\ 0 & 1 - S_2 & 0 \\ 0 & 0 & \ddots \\ 0 & 0 & 0 & 1 - S_n \end{bmatrix}$$

$$\mathbf{B} = [S_1 S_2 \dots S_n]^T$$

and

$$\mathbf{C}^T = [f_{\infty,1} f_{\infty,2} f_{\infty,n}]$$

In (8), $x_k^{(1)}$ represents the normalized concentration of absorbed analyte at time instant k , and \mathbf{A} , \mathbf{B} , and \mathbf{C} represent the system matrices. Note that, for real-time n -analyte quantification, the unknown parameters that need to be estimated in the model defined by (8) are the equilibrium frequency shifts (i.e., $f_{\infty,i}$, $i = 1, 2, \dots, n$). With the estimated equilibrium frequency shift of each analyte $f_{\infty,i}$ in the mixture and known sensitivity value $aK_{p-w,i}$ for a given analyte/coating pair, the ambient aqueous concentration for each analyte $C_{aq,i}$ can be computed by dividing $f_{\infty,i}$ by $aK_{p-w,i}$.

$$C_{aq,i} = \frac{f_{\infty,i}}{aK_{p-w,i}}$$

(9)

Sensor Signal Processing

Multistage EW-RLSE was selected for this work, because it offers various advantages. General advantages of estimation-theory-based technique include the ability to accurately estimate the unknown parameters even in the presence of measurement noise, real-time data processing, and the compatibility with field implementation using a microcontroller. In particular, EW-RLSE offers additional advantages: it overcomes some practical limitations of other adaptive techniques such as least mean squares by providing a faster rate of convergence, while its performance is insensitive to variations in the eigenvalue spread of the correlation matrix of the known signal.¹⁹ Moreover, the implementation of EW-RLSE is less computationally intensive compared to techniques such as Kalman filter.¹⁹ Also note that, in EW-RLSE, for accurately estimating equilibrium frequency shifts, the measurement data are weighted exponentially to ensure that more credibility is given to the recent measurement data in an effort to obtain a more accurate result.^{20,21} Moreover, parameters (i.e., equilibrium frequency shift of each analyte, $f_{\infty,i}$) can be estimated even in the presence of slow time variations in their values because of this emphasis on recent measurements.

EW-RLSE belongs to the class of adaptive filters that recursively estimate the unknown parameters by minimizing a weighted quadratic least-squares cost (or objective) function related to the known signals (i.e., unknown parameters are estimated by minimizing the error between the actual measured data

and the estimated data). The general framework of EW-RLSE is described in ref²². To utilize EW-RLSE to estimate the unknown parameters, (8b) must be rearranged into the following form

$$y_k = \Delta f_k = x_k^T C + w_k \quad (10)$$

where $x_k^T = [x_k^{(1)} x_k^{(2)} \dots x_k^{(n)}]$ is a vector of known signals (the normalized concentrations) that can be determined at each time step k , using (8a), and $C = [f_{\infty,1} f_{\infty,2} \dots f_{\infty,n}]^T$ is a vector of unknown parameters (the equilibrium frequency shifts) that need to be estimated. The estimated output can be expressed as

$$\hat{y}_k = x_k^T \hat{C}_k \quad (11)$$

In (11), \hat{y}_k is the estimate of the output, and \hat{C}_k is the estimate of the unknown parameters at discrete-time instant k . The (measurement) estimation error can be defined as

$$e_k = y_k - \hat{y}_k \quad (12)$$

The EW-RLSE algorithm recursively estimates the unknown parameters by minimizing an exponentially weighted least-squares cost function as follows^{20,23}

$$\hat{C}_k = \hat{C}_{k-1} + P_{k-1} x_k (x_k^T P_{k-1} x_k + \lambda^2 \sigma_w^2)^{-1} e_k \quad (13)$$

where P_k is the inverse correlation matrix of the input vector, which is updated as

$$P_k = \frac{1}{\lambda^2} [P_{k-1} - P_{k-1} x_k (x_k^T P_{k-1} x_k + \lambda^2 \sigma_w^2)^{-1} x_k^T P_{k-1}] \quad (14)$$

In (13) and (14), λ is the exponentially weighted “forgetting factor” required to give less weight to past errors; it is assigned a value between 0 (no memory) and 1 (infinite memory). The EW-RLSE algorithm using (13) and (14), outlined elsewhere²² and summarized in Table 1, can be used to estimate the unknown parameters, $f_{\infty,i}$, recursively.

Table 1. Summary of EW-RLSE Algorithm

<p>Step 1: Initialize the vector of unknown parameters, \hat{C}_k, to a set of predefined values; assign initial values to the inverse correlation matrix P_k. Set a value for the forgetting factor, $0 < \lambda \leq 1$.</p>
<p>Step 2: When y_k and x_k become available, update the unknown parameters vector \hat{C}_k and inverse correlation matrix P_k using (13) and (14), respectively.</p>

Step 3: Repeat Step 2 as long as the new measurement data are collected, or until the unknown parameters converge to the actual value.

In this paper, EW-RLSE is implemented in multiple stages (in series), where initial stages of EW-RLSE are used to determine and eliminate the analytes from the model that are initially considered to be in the mixture but actually are absent (this may also include analytes with concentrations close to zero). The final stage of EW-RLSE with the corresponding sensor response model representing only the analytes present in the mixture will be used to obtain a more accurate quantification result of these analytes.

The first stage of EW-RLSE is implemented by assuming that n analytes could be present in a mixture. These n analytes include those that are actually present in the mixture, numbering p , and those that are actually absent from the mixture, numbering q ($n = p + q$). As indicated earlier, EW-RLSE is used to estimate the equilibrium frequency shifts of the analytes, $f_{\infty i}$, which in turn are used to estimate the concentrations of the analytes using (9). The estimated concentration of each analyte will be compared to the detection limit of the analyte for the selected polymer-coated sensor platform. If the estimated concentration of a particular analyte is negative or lower than the detection limit for that analyte, the concentration of the analyte will be assumed to be zero and the estimation process repeated without this analyte in the model. Thus, at the end of the first stage, if it is determined that z analyte(s) are absent from the mixture, the estimation process will be repeated using a sensor response model with $n = p + [q - z]$, where $z = 1, 2, \dots, q$. Note that, if at the end of any stage, no analytes are determined to be absent from the mixture (i.e., $q = 0$), the estimation process will be terminated at the end of that stage. Otherwise, the estimation process will be repeated using a new stage of EW-RLSE and the appropriate sensor response model as long as analytes are determined to be absent from the present stage of EW-RLSE. This selection process is important, because it was observed that the estimated concentrations associated with analytes that are actually absent from the sample can greatly influence the quantification result obtained for the analytes that are actually present. Therefore, use of the above multistage EW-RLSE results in more accurate analyte quantification. The estimation process is repeated, until all absent analytes are identified (i.e., $z = q$); that is, the final stage of EW-RLSE will only include the analytes that are actually present in the sample (i.e., $n = p$).

A block diagram depicting this sensor signal-processing technique is presented in Figure 1. While the initial $(M - 1)$ EW-RLSE stages along with appropriate sensor response models are used to determine the analytes that are absent from the mixture, the M th EW-RLSE will calculate the concentrations of the analytes present in the mixture based on the estimated equilibrium frequency shifts of the analytes (the actual equilibrium frequency shifts need not be measured, saving analysis time).

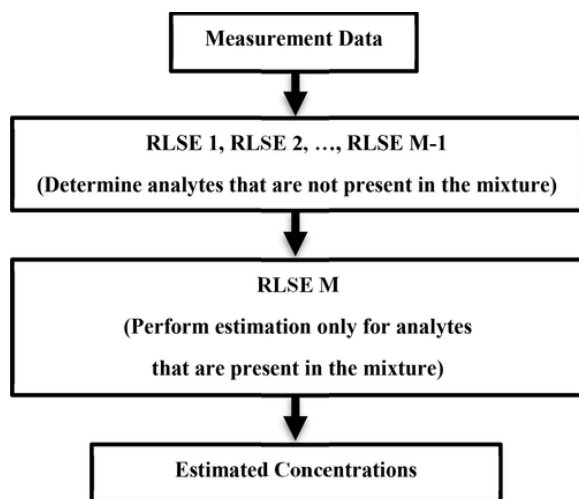


Figure 1. Block diagram of multistage EW-RLSE for online identification and quantification of analytes in a mixture.

Experimental Data Acquisition

The effectiveness of our approach was tested using the measured responses of polymer-coated SH-SAW sensors to various multi-analyte mixtures of BTEX and TMB, including single, binary, and ternary mixtures. The sensor response data analyzed in this work were collected using a 36° YX-LiTaO₃ guided SH-SAW device as the sensing platform.^{15,17,24} The fabrication details of the SH-SAW device used to collect the data can be found in our previous publications.^{15,17,24} A dual-delay-line configuration was used with one line serving as the sensing line and the other as the reference line. Sensing lines were either coated with poly(epichlorohydrin) (PECH) or poly(isobutylene) (PIB) (both purchased from Sigma-Aldrich). These polymers were deposited on the sensing line from solution in chloroform by spin coating and baking for 15 min at 55 °C. This resulted in thicknesses of 0.6 and 0.8 μm for PECH and PIB, respectively. The reference line was coated with poly(methyl methacrylate) (PMMA) (purchased from Scientific Polymer Products), then baked for 120 min at 180 °C, resulting in a glassy, nonsorbent coating, which has been found to be chemically insensitive in the concentration ranges of interest. The BTEX analytes and TMB used in the experiments were purchased from Sigma-Aldrich with purities of at least 98.5%.

Sensor data collection was performed as described in our previous publications.^{15,17,24} The measured signal is at least 3 times the root-mean-square (RMS) noise, with the average fluctuations in the measured signal due to RMS noise used as an indication of the error bars on the data. Moreover, a portable gas chromatograph-photoionization detector (GC-PID) system (Defiant Technologies FROG-4000) was used to obtain an independent measurement of BTEX concentrations. Note that the BTEX concentrations determined using GC-PID are subject to an average error of ±7%.²⁵

As mentioned earlier, our approach utilizes two sensing parameters, namely, response time constant and sensitivity for each analyte that might be present in the tested sample. Therefore, several single-analyte measurements were performed to determine the average values of response time constants and sensitivities for the BTEX compounds and TMB, as listed in Tables SI-1 (for 0.6 μm PECH) and SI-2 (for 0.8 μm PIB) in the Supporting Information accompanying this paper. The experimentally determined detection limits of the analytes of interest for the selected coatings are 100 ppb (parts per

billion by weight) for benzene, 50 ppb (for toluene), 30 ppb (for ethylbenzene), and 10 ppb (for TMB). These detection limits were used as reference values in our approach to determine whether the analyte is present or absent from the tested sample. It is also pertinent to note that identical values for response times and sensitivities are observed for ethylbenzene and the mixture of *m*-, *o*-, and *p*-xylenes. This is because those analytes are all chemical isomers (identical molecular formulas, the only differences being structural), and the associated responses therefore cannot be distinguished using the investigated sensor coatings.

Results and Discussion

The technique described in the previous sections makes it possible to detect and quantify the n analytes in a mixture that are most likely to have caused the measured sensor response, provided the characteristic response time constant and sensitivity of each single analyte/coating combination are known in advance. Experimental confirmation of this is described below for multi-analyte mixtures of BTEX compounds and TMB using the measured frequency-shift response obtained from a single polymer-coated SH-SAW sensor, which permits estimation of steady-state frequency shift for all mixture constituents. The approach was tested extensively using sensors coated with either 0.6 μm PECH or 0.8 μm PIB.

Note that, for the implementation of the reported technique, initially it is assumed that up to $n = 4$ analytes (benzene, toluene, a combination of the chemical isomers ethylbenzene and *m*-, *o*-, and *p*-xylenes, and 1,2,4-trimethylbenzene) could be present in the tested sample. The initial values of the unknown parameters (i.e., the steady-state frequency shifts $f_{\infty,i}$ associated with the analytes assumed to be in the sample) were set to zero for all the tested data for simplicity, because our technique is insensitive to these initial values. Once the unknown parameters have been initially estimated, the corresponding analyte concentrations are calculated using (9). The estimated concentrations of the analytes are then compared to the detection limit of the analytes (as listed in Tables SI-1 for 0.6 μm PECH and SI-2 for 0.8 μm PIB) to determine whether the analytes are actually present in the sample or not. After the various stages of EW-RLSE were gone through as described above, the final quantification result of the analytes was compared to the concentrations measured independently using GC-PID (and gas chromatography-mass spectrometry (GC-MS)).

Before more comprehensive experimental results are summarized at the end of this section, the results obtained from two individual tests are discussed to highlight the effectiveness of the technique as demonstrated. The first sample result, shown in Figure 2, was obtained using the sensor response data of a SH-SAW sensor coated with 0.8 μm PIB to a multi-analyte mixture of 590 ppb benzene (B), 420 ppb toluene (T), and 330 ppb of a combination of ethylbenzene and xylenes (EX). Figure 2 shows the measurement data (in blue) and the estimated sensor response curve (in red) along with the identification and quantification results. The measured and estimated sensor responses in Figure 2 are in good agreement, resulting in correct identification of the analytes in the sample and estimated analyte concentrations within $\pm 10\%$ of the values measured using GC-PID.

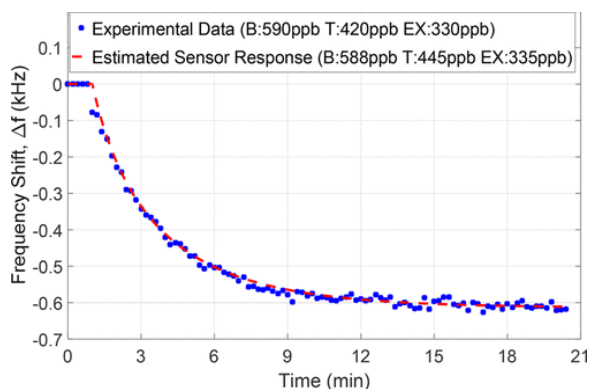


Figure 2. Measured response of an SH-SAW sensor coated with 0.8 μm PIB to multi-analyte mixtures of 590 ppb benzene (B), 420 ppb toluene (T), and 330 ppb ethylbenzene plus xylenes (EX). Also shown (red line) is the sensor response estimated using the multistage EW-RLSE. (inset) The actual and estimated analyte concentrations.

Table 2 compares the results obtained using just one stage of EW-RLSE with those from multistage EW-RLSE. It is evident that multistage EW-RLSE leads to more accurate results. Note that, for the result obtained using single-stage EW-RLSE, the estimated concentration associated (which is actually absent from the tested sample) greatly influenced the estimated concentrations of other analytes that are actually present. However, for multistage EW-RLSE, the same issue does not arise, as the estimation process was repeated after determining that 1,2,4-trimethylbenzene was not present (the first stage estimated a negative concentration of 1,2,4-trimethylbenzene).

Table 2. Comparison of Results Obtained using Single-Stage EW-RLSE and Multi-Stage EW-RLSE

analytes ^a	concentrations (ppb)		
	measured using GC-PID	single-stage (error)	multi-stage (error)
benzene	590	826 (40%)	588 (-0.3%)
toluene	420	331 (-21%)	445 (6.0%)
ethylbenzene & xylenes	330	337 (2.1%)	335 (1.5%)
1,2,4-trimethylbenzene	0	-8 (N/A)	0 (N/A)

^a Note: The tested sensor response data were collected using an SH-SAW device coated with 0.8 μm PIB.

Table 3 shows the identification and quantification results obtained using the data collected in the first 14, 16, 18, and 20 min after the multi-analyte sample came into contact with the sensor.

As Table 3 shows, for this multi-analyte sample, the data collected in the first 14 min are sufficient to identify and quantify the analytes with less than $\pm 20\%$ difference from the analyte concentrations measured by GC-PID. Using just the data collected for less than 14 min, however, resulted either in misidentification or inaccurate quantification of the analytes. Compared to previous results on binary analyte mixtures,²⁶⁾ a longer time is needed to accurately identify and quantify the analytes present in the mixture. This indicates that, as the number of analytes in the mixture increases, the minimum time to accurately detect and quantify the analytes in the sample may increase; to ascertain the consistency and extent of this dependence will require testing a wider range of samples and sensitive films.

Table 3. Identification and Quantification Results Obtained using the Measured Response Data^a Collected in the First 14, 16, 18, and 20 min after the Multi-Analyte Sample Was Introduced to the Sensor

identification and quantification result	identified analytes and their estimated concentrations in ppb (% difference to analyte concentration measured using GC-PID)
after 14 min	benzene: 714 (20%) toluene: 392 (-7%) ethylbenzene & xylenes: 347 (5%)
after 16 min	benzene: 527 (-11%) toluene: 469 (12%) ethylbenzene & xylenes: 330 (0%)
after 18 min	benzene: 534 (-10%) toluene: 466 (11%) ethylbenzene & xylenes: 330 (0%)
after 20 min	benzene: 588 (-1%) toluene: 445 (6%) ethylbenzene & xylenes: 335 (2%)

^a Of an SH-SAW sensor coated with 0.8 μm PIB to 590 ppb benzene, 420 ppb toluene, and 330 ppb ethylbenzene and xylenes.

Next, the results obtained for a 0.8 μm PIB-coated SH-SAW sensor exposed to 600 ppb TMB are presented in Figure 3 and Table 4. On the basis of these results, it can be concluded that the demonstrated technique accurately identifies and quantifies the analyte in the sample as TMB with less than $\pm 15\%$ difference from the TMB concentration measured using GC-PID. In fact, the analyte is correctly identified and quantified within $\pm 20\%$ of the value measured by GC-PID using only the measured data collected for the first 5 min, a small fraction of the total response time of TMB. In this case, valuable time (~ 20 min) could be saved by utilizing the proposed technique. Note that estimation of BTEX concentrations in the presence of various interferences, including TMB, has also been demonstrated in a recent publication by the authors.²⁷

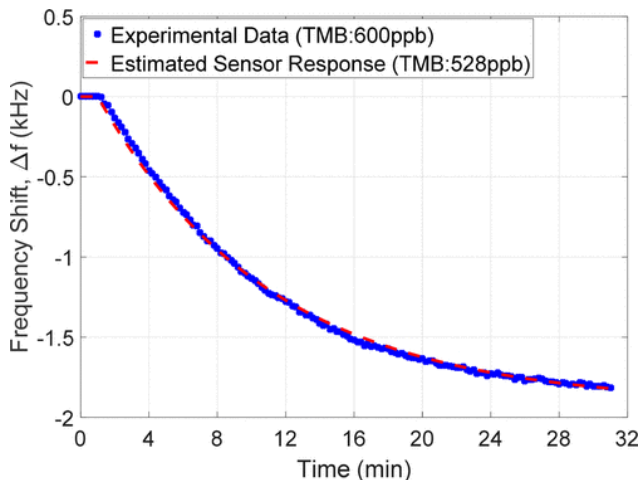


Figure 3. Measured response of an SH-SAW sensor coated with 0.8 μm PIB to 600 ppb 1,2,4-trimethylbenzene (TMB). Also shown (red line) is the sensor response estimated using the multistage EW-RLSE. (inset) The actual and estimated analyte concentrations.

Table 4. Identification and Quantification Results Obtained using the Measured Response Data^a Collected for the First 5, 8, 10, and 14 min after the Single-Analyte Sample Was Introduced to the Sensor

identification and quantification result	identified analytes and their estimated concentrations in ppb (% difference to analyte concentration measured using GC-PID)
after 5 min	1,2,4-trimethylbenzene: 484 (19%)
after 8 min	1,2,4-trimethylbenzene: 509 (15%)
after 10 min	1,2,4-trimethylbenzene: 518 (14%)
after 14 min	1,2,4-trimethylbenzene: 527 (12%)

^a Of an SH-SAW sensor coated with 0.8 μm PIB to 600 ppb 1,2,4-trimethylbenzene.

In addition to the two sample results discussed above, further tests were performed to more fully characterize the detection and quantification of both multi-analyte and single-analyte samples using our method. Results obtained from these tests are summarized in Table 5, which includes data on SH-SAW sensors coated with 0.6 μm PECH and 0.8 μm PIB. Analysis of the results presented in Table 5 reveal an average (absolute) error of 6.3% for the estimated concentrations relative to the actual concentrations, with a standard deviation of $\pm 7.1\%$. Interestingly, if the sign of the error is considered, the average error is only -0.06% , indicating that over- and underestimates of concentration are similarly probable. Overall, these results demonstrate that this method accurately identifies and quantifies the analytes, within an average of $\pm 6.3\%$ of the values measured using GC-PID, in a mixture in near-real time using the measured response from just a single polymer-coated SH-SAW device.

Table 5. Summary of Identification and Quantification Results Obtained using Multi-Stage EW-RLSE and Frequency Shift Data of Polymer-Coated SH-SAW Sensors,^a Compared to Analyte Concentrations Measured using GC-PID

data	coating	actual concentrations ^b (in ppb)				estimated concentrations ^b (in ppb)			
		B	T	EX	TMB	B	T	EX	TMB
1	PIB	1920	0	0	0	1880	0	0	0
2		160	1010	350	0	220	970	350	0
3		0	0	370	0	0	0	370	0
4		1990	0	0	0	1940	0	0	0
5		590	420	330	0	590	450	340	0
6		0	0	0	2000	0	0	0	2010
7		0	0	0	600	0	0	0	530
8		170	450	360	0	140	470	350	0
9		0	930	0	0	0	920	0	0
10		0	0	0	1000	0	0	0	1030
11		0	640	0	0	0	630	0	0
12	PECH	0	0	0	500	0	0	0	430
13		390	810	70	0	450	720	80	0
14		1060	0	1410	0	880	0	1420	0
15		2260	740	0	0	2000	790	0	0
16		850	370	750	0	740	400	770	0
17		870	520	70	0	950	510	70	0
18		1950	0	1670	0	2110	0	1640	0
19		620	0	1260	0	660	0	1240	0
20		0	0	0	1000	0	0	0	920
21		0	0	0	2000	0	0	0	1930
22		980	340	0	0	940	330	0	0
23		0	0	1400	0	0	0	1440	0
24		1930	0	0	0	1950	0	0	0
25		0	360	0	0	0	360	0	0

^aCoated with either 0.6 μm PECH or 0.8 μm PIB as indicated.

^bB, T, EX, and TMB denote benzene, toluene, ethylbenzene and xylenes, and 1,2,4-trimethylbenzene, respectively.

Summary and Conclusions

A real-time sensor signal-processing technique based on multistage EW-RLSE for multi-analyte identification and quantification was demonstrated. The performance of the technique was experimentally verified for detection and quantification of multi-analyte mixtures of BTEX compounds and TMB using the frequency-shift response of a single polymer-coated SH-SAW sensor. This method can be used to detect and quantify n analytes in a mixture, provided the characteristic response time constants and sensitivities of the analyte/coating pairs are known. Importantly, however, no prior knowledge of the concentrations of the analytes in a given mixture is required to correctly identify and accurately quantify them.

The formulated sensor-response model achieves chemical selectivity by utilizing two sensor parameters, namely, the response time constants and the concentration sensitivities, resulting in reliable detection and accurate quantification in multi-analyte mixtures. The use of estimation theory permits times to detection to be significantly less and, in some cases, even less than the time required to reach equilibrium between the sample components and the sensing film, particularly for analytes with relatively long response times. Implementation of multiple stages of EW-RLSE eliminates analyte misidentification, even as it improves the accuracy of analyte quantification.

Measurement of 25 samples with up to three analytes resulted in 100% accurate identification of the analytes, in combination with an average quantification accuracy of $\pm 6.3\%$. Our method will obviate the need for sensor arrays in many applications involving detection and quantification of multi-analyte samples. Importantly, implementation of our approach does not require exact knowledge of all the analytes present in a given sample: it is sufficient to include those analytes in the model that could potentially be present in significant concentrations; the technique then can correctly identify and accurately quantify the analytes that are actually present in the tested sample.

Although the results and analyses presented here are for low concentrations of organic compounds directly detected in aqueous solution, the technique could also be used to detect and quantify analytes in gas-phase environments, provided sufficient sensor response data can be collected during the transient gas-sensor response and that the response times vary with the chemical characteristics of the analytes. Moreover, this technique is not restricted to the acoustic wave-based sensors reported here; it could be used with many other sensor platforms as well.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssensors.9b00564.

- Quantitative Detection of Complex Mixtures using a Single Chemical Sensor: Analysis of Response Transients using Multi-Stage Estimation. The average values of response time constants, sensitivities and detection limit for benzene, toluene, ethylbenzene, xylenes (BTEX) and 1,2,4-trimethylbenzene (TMB) used in the present approach (PDF)

Terms & Conditions

Most electronic Supporting Information files are available without a subscription to ACS Web Editions. Such files may be downloaded by article for research use (if there is a public use license linked to the

relevant article, that license may permit other uses). Permission may be obtained from ACS for other uses through requests via the RightsLink permission system: <http://pubs.acs.org/page/copyright/permissions.html>.

Notes

The authors declare no competing financial interest.

References

1. Adhikari, P.; Alderson, L.; Bender, F.; Ricco, A.; Josse, F. Investigation of Polymer–Plasticizer Blends as SH-SAW Sensor Coatings for Detection of Benzene in Water with High Sensitivity and Long-Term Stability. *ACS Sensors* **2017**, *2*, 157–164, DOI: 10.1021/acssensors.6b00659
2. Jones, Y.; Li, Z.; Johnson, M.; Josse, F.; Hossenlopp, J. ATR-FTIR spectroscopic analysis of sorption of aqueous analytes into polymer coatings used with guided SH-SAW sensors. *IEEE Sens. J.* **2005**, *5*, 1175–1184, DOI: 10.1109/JSEN.2005.859231
3. Zaromb, S.; Stetter, J. Theoretical basis for identification and measurement of air contaminants using an array of sensors having partly overlapping selectivities. *Sens. Actuators* **1984**, *6*, 225–243, DOI: 10.1016/0250-6874(84)85019-2
4. Albert, K.; Lewis, N.; Schauer, C.; Sotzing, G.; Stitzel, S.; Vaid, T.; Walt, D. Cross-Reactive Chemical Sensor Arrays. *Chem. Rev.* **2000**, *100*, 2595–2626, DOI: 10.1021/cr980102w
5. Ricco, A.; Crooks, R.; Osbourn, G. Surface Acoustic Wave Chemical Sensor Arrays: New Chemically Sensitive Interfaces Combined with Novel Cluster Analysis To Detect Volatile Organic Compounds and Mixtures. *Acc. Chem. Res.* **1998**, *31*, 289–296, DOI: 10.1021/ar9600749
6. Gardner, J.; Bartlett, P. *Electronic noses: Principles and Applications*; Oxford University Press: New York, 1999.
7. Duda, R.; Hart, P.; Stork, D. *Pattern classification*; Wiley: New York, 2001.
8. Webb, A.; Copsey, K. *Statistical pattern recognition*; Wiley: Hoboken, NJ, 2011.
9. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27, DOI: 10.1109/TIT.1967.1053964
10. Devroye, L. On the Inequality of Cover and Hart in Nearest Neighbor Discrimination. *IEEE Trans. Pattern Anal. Mach. Intell.* **1981**, *PAMI3*, 75–78, DOI: 10.1109/TPAMI.1981.4767052
11. Bishop, C. M. *Neural Networks for Pattern Recognition*; Oxford University Press: Oxford, UK, 1995.
12. Gardner, J. Detection of vapours and odours from a multisensor array using pattern recognition Part 1. Principal component and cluster analysis. *Sens. Actuators, B* **1991**, *4*, 109–115, DOI: 10.1016/0925-4005(91)80185-M
13. Josse, F.; Dahint, R.; Shah, S.; Houndegla, E. Metal ion solutions identification using acoustic plate mode sensors and principal component analysis. *Proc. SPIE* **1999**, *2*, DOI: 10.1117/12.370271
14. Hsieh, M.; Zellers, E. Limits of Recognition for Simple Vapor Mixtures Determined with a Microsensor Array. *Anal. Chem.* **2004**, *76*, 1885–1895, DOI: 10.1021/ac035294w
15. Bender, F.; Josse, F.; Ricco, A. J. Influence of Ambient Parameters on the Response of Polymer-Coated SH-Surface Acoustic Wave Sensors to Aromatic Analytes in Liquid-Phase Detection. *Joint Conf. IEEE IFCS and EFTF Proc.* **2011**, 422–427, DOI: 10.1109/FCS.2011.5977752

16. Bender, F.; Mohler, R.; Ricco, A. J.; Josse, F. Quantification of Benzene in Groundwater Using SH-Surface Acoustic Wave Sensors. *IMCS Proc.* **2012**, 473– 476
17. Bender, F.; Mohler, R.; Ricco, A.; Josse, F. Identification and Quantification of Aqueous Aromatic Hydrocarbons Using SH-Surface Acoustic Wave Sensors. *Anal. Chem.* **2014**, *86*, 1794– 1799, DOI: 10.1021/ac403724f
18. Sothivelr, K.; Bender, F.; Josse, F.; Ricco, A.; Yaz, E.; Mohler, R.; Kolhatkar, R. Detection and Quantification of Aromatic Hydrocarbon Compounds in Water Using SH-SAW Sensors and Estimation-Theory-Based Signal Processing. *ACS Sens.* **2016**, *1*, 63– 72, DOI: 10.1021/acssensors.5b00024
19. Li, T. On Exponentially Weighted Recursive Least Squares for Estimating Time-Varying Parameters and its Application to Computer Workload Forecasting. *J. Stat. Theory Practice* **2008**, *2*, 339– 354, DOI: 10.1080/15598608.2008.10411879
20. Lewis, F.; Xie, L.; Popa, D.; Lewis, F. *Optimal and robust estimation*, 2nd ed.; CRC Press: Boca Raton, FL, 2008.
21. Brogan, W. *Modern control theory*; Prentice-Hall: Englewood Cliffs, NJ, 1991.
22. Simon, D. *Optimal state estimation*; Wiley-Interscience: Hoboken, NJ, 2006.
23. Goodwin, G.; Sin, K. *Adaptive filtering prediction and control*; Dover Publications: Mineola, NY, 2009.
24. Bender, F.; Josse, F.; Mohler, R. E.; Ricco, A. J. Design of SH-Surface Acoustic Wave Sensors for Detection of ppb Concentrations of BTEX in Water. *Joint UFFC, EFTF and PFM Symp.* **2013**, 628– 631, DOI: 10.1109/EFTF-IFC.2013.6702067
25. *A Micro-GC Based Chemical Analysis System*; Defiant Technologies, Inc., 2018. [Online]. Available: <http://www.defiant-tech.com/pdfs/Pittcon%202014%20A%20Micro-GC%20Based%20Chemical%20Analysis%20System.pdf>. [Accessed Jan 8, 2018].
26. Sothivelr, K.; Bender, F.; Yaz, E. E.; Josse, F.; Mohler, R. E.; Ricco, A. Near-Real-Time Analysis of Binary Mixtures of Organic Compounds in Water using SH-SAW Sensors and Estimation Theory. *IEEE Sens. 2014 Proc.* **2014**, 578– 581, DOI: 10.1109/ICSENS.2014.6985064
27. Sothivelr, K.; Bender, F.; Josse, F.; Yaz, E. E.; Ricco, A. Obtaining Chemical Selectivity from a Single, Nonselective Sensing Film: Two-Stage Adaptive Estimation Scheme with Multiparameter Measurement to Quantify Mixture Components and Interferents. *ACS Sens.* **2018**, *3*, 1656– 1665, DOI: 10.1021/acssensors.8b00353