Analysis of Sensor Signals and Quantification of Analytes Based on Estimation Theory

Karthick Sothivelr
Marquette University

ANALYSIS OF SENSOR SIGNALS AND QUANTIFICATION OF ANALYTES BASED ON ESTIMATION THEORY

by

KARTHICK SOTHIVELR

A Thesis submitted to the Faculty of the Graduate School,
Marquette University,
in Partial Fulfillment of the Requirements for
the Degree of Masters of Science (Electrical and Computer Engineering).

Milwaukee, Wisconsin

August 2014

ABSTRACT

ANALYSIS OF SENSOR SIGNALS AND QUANTIFICATION OF ANALYTES BASED ON ESTIMATION THEORY

KARTHICK SOTHIVELR

MARQUETTE UNIVERSITY, 2014

Compact sensor systems for on-site monitoring of groundwater for trace organic compounds are currently under development. To permit near real-time analysis of samples containing multiple analytes, the present work investigates a sensor signal processing approach based on estimation theory, specifically using Kalman Filter and Extended Kalman Filter. As a first step towards the analysis of groundwater samples containing multiple compounds, the approach presented in this work permits estimation of analyte concentration(s) in binary mixtures and single analyte samples on-line, before the sensor response reaches steady-state. Sensor signals from binary mixtures and single analyte samples of BTEX compounds (benzene, toluene, ethylbenzene, and xylenes) were analyzed in this work because these compounds are good indicators of accidental release of fuel and oil into groundwater. Based on those previous experimental results, models for the sensor response to binary mixtures and single analyte samples are developed. These sensor response models were transformed into state-space representation so that estimation theory can be used to estimate the sensor parameters.

For the case of the single analyte system, one state-space form was developed and for the case of the two-analyte system, two different state-space forms were developed. These state-space forms were tested using the available measured data, and the results indicate that relatively accurate estimates of analyte concentration(s) could be obtained within a relatively short period of time (six minutes or less for the tested sensor system) well before the sensor response reaches steady-state. Also presented in this work are new techniques that enable correcting for linear baseline drift and outlier points in the measured data on-line. The linear baseline drift correction technique uses estimation theory (particularly Kalman Filter) to rapidly perform linear extrapolation and linear interpolation. The elimination of the outlier points in the sensor data was performed by using a combination of discrete low pass filter and Kalman Filter (or Extended Kalman Filter depending on the state-space form). These techniques were tested on measured data with linear baseline drift and outlier points and the results obtained indicate that these sensor signal pre-processing techniques are indeed capable of correcting for linear baseline drift and outlier points in real-time.

ACKNOWLEDGEMENTS

KARTHICK SOTHIVELR

I would like to thank my research advisors, Dr. Fabien Josse and Dr. Florian Bender, for their guidance and support throughout this work. This work would not have been possible without the countless hours of discussion on both technical aspects and writing style with Dr. Fabien Josse and Dr. Florian Bender. I would also like to thank Dr. Edwin Yaz for his discussions on estimation theory and various forms of Kalman Filter used throughout this work. Having such a well-versed group of faculty in the areas of sensors and estimation theory made it easy to check my ideas and arguments. Thanks are also due to my entire research committee (Dr. Fabien Josse, Dr. Florian Bender, Dr. Edwin Yaz, and Dr. Susan Schneider) for reading, correcting and making suggestions on how to improve this thesis. I also would like to acknowledge all the graduate students and faculty in the Microsensor Research Group at Marquette University for their help throughout this endeavor.

Special thanks to my family for their lifelong support and encouragement in all my endeavors. Also I would like to thank Arut Perum Johti (Supreme Grace-Light) for guiding me throughout this life.

Arut Perum Johti, Arut Perum Johti
Thaniperum Karunai, Arut Perum Johti

Vast Grace-Light, Vast Grace-Light
Supreme Compassion, Vast Grace-Light

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 General Background

Sensors are devices that allow the measurement of physical or chemical quantities and produce a signal that can be related to the quantity that the sensor is measuring. Typically, a sensor produces an electrical signal in response to the quantity that it is measuring. The electrical signal is then further processed by the signal processing unit, so that the electrical signal can be displayed in a way that is convenient to an observer. The signal processing unit can also be programmed to send a signal to a computer system or processor to take action if it senses any measurable change in the sensing environment. Basically, a sensor acts as a bridge between the real world and the electronic world. Sensors play a vital role in everyday life. Example of sensors used in everyday life include those used for monitoring the air quality in a room, monitoring the temperature in a room, checking the tire pressure, measuring the temperature of engine, and many more.

Sensors can be classified into two categories, physical and chemical. Physical sensors are used to measure any physical quantities such as temperature, force, velocity, acceleration, pressure, etc. Chemical sensors are devices that are used to detect or measure the concentration of chemical(s) in either the liquid or gas phase [1]. In chemical sensors, a chemical interaction will cause a measurable change in a property of the sensor. This chemical interaction is facilitated by a polymer or other compound that is placed on the sensor. Chemical sensors can further be classified into several groups

depending on the technology used in the design and operation of the sensor. These include optical devices, electrochemical devices, micro-electromechanical systems (MEMS) and acoustic wave devices. In the optical chemical sensors, changes in the particular optical parameters such as index of refraction, amount of absorbance, or intensity of photoluminescense are monitored [2]. The interaction between the target analyte and the polymer coating placed on the optical sensor, will result in the changes in a particular optical parameter and this change in the optical parameter can be related to the concentration of analyte [1, 2]. In the electrochemical devices, the interaction between the analyte and electrode are monitored. The interaction between the analyte and electrode will produce a measurable signal (e.g. change in conductivity) that can be related to the concentration of analyte [2]. Electrochemical devices include voltammetric sensors, potentiometric sensors and conductimetry sensors. In the MEMS devices, the properties of these micro-scaled devices are monitored and any changes observed in the mechanical or electrical properties of the device can be related to the concentration of analyte [3]. One of the most promising of the MEMS sensors is the microcantilever. A microcantilever is a diving-board-like structure usually only a few hundred microns in length. Microcantilever can be operated in the dynamic mode, where the resonant frequency and quality factor are monitored, or in the static mode, where the deflection of the microcantilever is monitored [4]. A microcantilever can be used as a chemical sensor when a polymer is placed on the surface of microcantilever. In dynamic mode, the analyte will interact with the polymer coating producing mass loading and stress effects which will change the resonant frequency and quality factor of the microcantilever. These changes can be related to the concentration of the analyte. On the other hand, in static

mode, the differential surface stress from analyte sorption causes the microcantilever to bend and the magnitude of the bending of the microcantilever can be related to the concentration of analyte. Acoustic wave devices use elastic waves at frequencies well above the audible range propagating in piezoelectric crystals. Typically, acoustic wave devices are operated between the frequencies of 1 MHz to slightly above 1000 MHz [5, 6]. The acoustic wave devices that are commonly used for sensing applications are quartz crystal microbalance (QCM) also known as thickness shear mode (TSM) resonator, surface acoustic wave (SAW) device, shear horizontal surface acoustic wave (SH-SAW) device, acoustic plate mode (APM) device and the flexural plate wave (FPW) device. Chemical interactions between the analyte and the polymer coating on the acoustic wave sensors cause a perturbation in the propagation characteristics of the wave. The changes in the frequency and attenuation of the wave are often used for detection of analyte.

As mentioned earlier, a chemical sensor is a sensor that is developed to detect the presence or measure the concentration of a chemical in either the liquid or gas phase. The detection of certain chemicals (or analytes) has become of great importance for human health. Leakages and spills from fuel and oil tanks, pipelines and other sources may contaminate groundwater, lakes, rivers and oceans posing a great threat to human health [7]. It is known that BTEX compounds (benzene, toluene, ethylbenzene and xylenes) are present in crude oil and its refined products [8] and in particular, benzene poses a great threat to human health. The United States Department of Health and Human Services (DHHS) classifies benzene as a human carcinogen [9]. Long-term exposure to significant levels of benzene could cause cancer, leukemia, and anemia. The current United States

Environmental Protection Agency (EPA) limit of benzene in drinking water requires benzene concentration to not exceed five parts per billion (ppb) [10]. There are various sources of benzene in the environment such as cigarette smoke and car exhaust but one of the worrying sources of benzene is due to the leakages from underground gasoline storage tanks. Gasoline contains an average of 0.62% of benzene (with a maximum of 1.3%) [11] and a leak can cause benzene to enter and contaminate soil and groundwater. Therefore, it is necessary to monitor the area around the gasoline storage tanks for any leakages to determine any presence of benzene and this can be done by using a chemical sensor. Extensive research is being conducted to develop an in-situ chemical sensor which can be used to monitor benzene concentration near the monitoring wells either continuously or frequently. In order to be used as an in-situ chemical sensor, the sensor must be able to respond rapidly to trace concentrations of benzene (that is on the order of five ppb as required by the EPA limit for drinking water). Such a chemical sensor can be made by using a shear horizontal surface acoustic wave (SH-SAW) device with a thin chemically selective coating [6]. The SH-SAW sensors made with the current sensing polymer layers respond quickly but have a limit of detection of 200 ppb [7]. Therefore, efforts are being made to increase the sensitivity of the SH-SAW sensor by investigating several polymers which could be used to detect smaller concentration of benzene than the current limit of detection of 200 ppb.

Most chemical sensors rely on the chemical interaction between the polymer and analyte. It is the polymer which dictates the sensitivity and selectivity of the chemical sensor. Therefore the selection of the polymer is very important for a particular

application. However, in many cases, the chemical interaction between the analyte and polymer is not highly selective. Typically, polymer coatings will absorb more than one analyte, thus most chemical sensors are only partially selective. The problem of partial selectivity of chemical sensors can be rectified by two methods. Method one is the most obvious solution, that is to use a single chemical sensor and select a polymer which only responds to one target analyte. This method is successful for certain types of chemicals [12, 13]. Method two is to use an array of many partially selective sensors (sensor array) to improve the selectivity of the sensor. Each sensor in a sensor array will respond to a wide variety of chemicals, however, the group of sensors, as a whole, will respond uniquely to different chemicals [14]. Therefore, using sensor arrays helps to detect individual components in a mixture [5]. However, both methods are always accompanied by signal processing to analyze the sensor response and also to quantify the target analyte. In a single chemical sensor, the sensor response is subject to noise and baseline fluctuations due to changes in the environmental conditions and the sensor response must be processed to eliminate the noise and baseline drift. In a sensor array, each and every sensor in the array is subject to noise and baseline fluctuation. Thus, each sensor in the array has to be processed separately to eliminate noise and baseline drift in the sensor response. Moreover, in a sensor array, further signal processing is required to process the sensor array as a whole to identify and also to quantify the target analyte. Therefore, signal processing is an important part of the chemical sensing process.

1.2 Review of Sensor Signal Processing Methods

From the previous section, it is obvious that sensor signal processing plays a major role in the chemical sensor system. The main purpose of signal processing is to analyze and to quantify the target analyte. This is true for single chemical sensor systems and also for chemical sensor arrays. In this section, a review of sensor signal processing methods will be given. Four types of sensor signal processing will be reviewed including baseline correction, time-to-detection (or steady-state information extraction), transient information extraction and sensor array processing (pattern recognition).

1.2.1 Baseline Correction

For most chemical sensor applications, a sample is collected from the environment and then transferred rapidly to a cell containing the sensors. When the flow of the sample to the sensor is sufficiently fast, concentration presented to the sensors, $C_{amb}(t)$, can be adequately represented by a step function as shown,

$$C_{amb}(t) = k\, C_{env}(t_s) u_s(t) \tag{1.1}$$

where $u_s(t)$ represents the unit step function, $C_{env}(t_s)$ is the environmental analyte concentration at the time the sample is collected and k represents the concentrating effect of the collector (k = 1 if the collector does not change the concentration) [15]. A typical

sensor response is shown in Fig. 1.1. As can be seen from Fig. 1.1, upon analyte exposure, the sensor will respond rapidly at first and then slowly as the transients decay and approaches steady-state (equilibrium). In order to identify and quantify the analyte(s) in a sample, steady-state features are often used. There are various steady-state features and the choice of steady-state feature to be used is largely dependent on the sensor platform [16]. Table 1.1 shows some proposed steady-state features and their formulae. Note, the steady-state feature shown in Fig. 1.1 is the steady-state difference measurement, i.e. the difference between the response at the steady-state, y_{ss} and the baseline response, y_b.

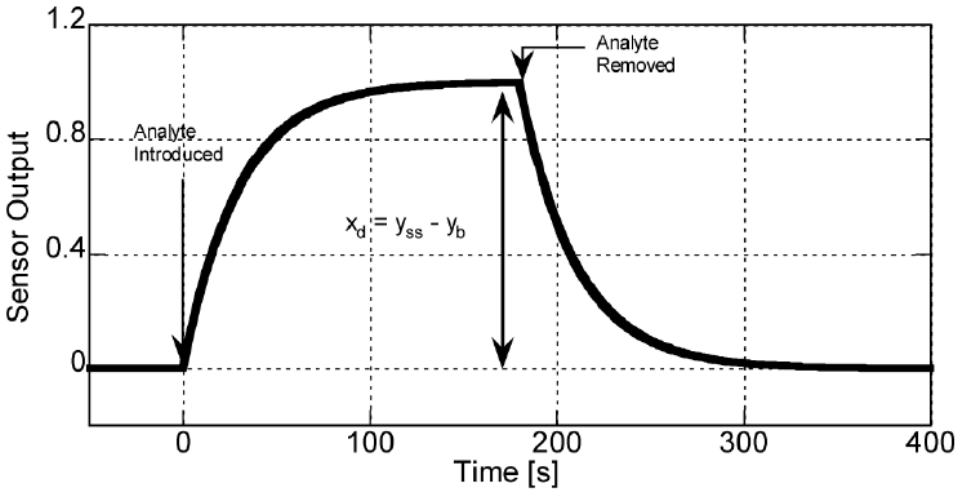

Figure 1.1: Typical chemical sensor response showing the difference measurement between the steady-state response and the baseline response [15].

Steady-State Feature	Formulae	Sensor Types
Difference	$x = y_{ss} - y_b$	Acoustic Wave Metal-Oxide Resistor
Relative	$x = y_{ss}/y_b$	Metal-Oxide Resistor Polymer Resistor
Fractional Change	$x = (y_{ss} - y_b)/y_b$	Metal-Oxide Resistor Polymer Resistor
Log Relative	$x = \ln(y_{ss}/y_b)$	Metal-Oxide Resistor

Table 1.1: List of various different steady-state features used for identification and quantification of chemical analytes [16].

There are two main issues with using the steady-state feature to identify and quantify the analyte(s) in a sample which includes baseline drift and time to detection. Baseline drift occurs due to fluctuations in environmental conditions such as changes in temperature and humidity. If the baseline measured just before the sensor is exposed to the analyte is used to calculate the steady-state feature, it will result in an incorrect measurement of the sensor response, as shown in Fig. 1.2.

There are several baseline correction techniques that allow one to estimate the true baseline during the exposure to the analyte such as linear extrapolation, linear interpolation, cubic interpolation and using estimation theory to estimate the baseline. Linear extrapolation and linear interpolation are often used for sensors with short response time. Using these two techniques it is implicitly assumed that the baseline remains constant during exposure. Linear extrapolation technique has the advantage of only requiring the data obtained before the sensor is exposed to the analyte. On the other

hand, linear interpolation requires data obtained both before the analyte is added and after it has been flushed from the sensor. However, linear interpolation is usually more accurate than linear extrapolation at the expense of additional time required for the analyte to be flushed from the sensor before baseline can be estimated. For sensors with longer response time, linear extrapolation and linear interpolation could lead to a poor estimate of the baseline. This is due to the possibility of baseline drift rate or direction changes during a response. In the case of baseline drift changing rate or even direction, cubic interpolation function and estimation theory can be used. In most cases, as the complexity of the baseline correction techniques increases, it will yield a better result as depicted in Fig. 1.2. As can be seen from Fig. 1.2, cubic interpolation is able to approximate the true sensor baseline better than linear extrapolation or linear interpolation. More details on using estimation theory to correct for baseline drift are given in [17]. By using, estimation theory, one can actually do the baseline drift correction in real time as the measurement is recorded, which could drastically shorten the time required to quantify the analyte from the sensor response.

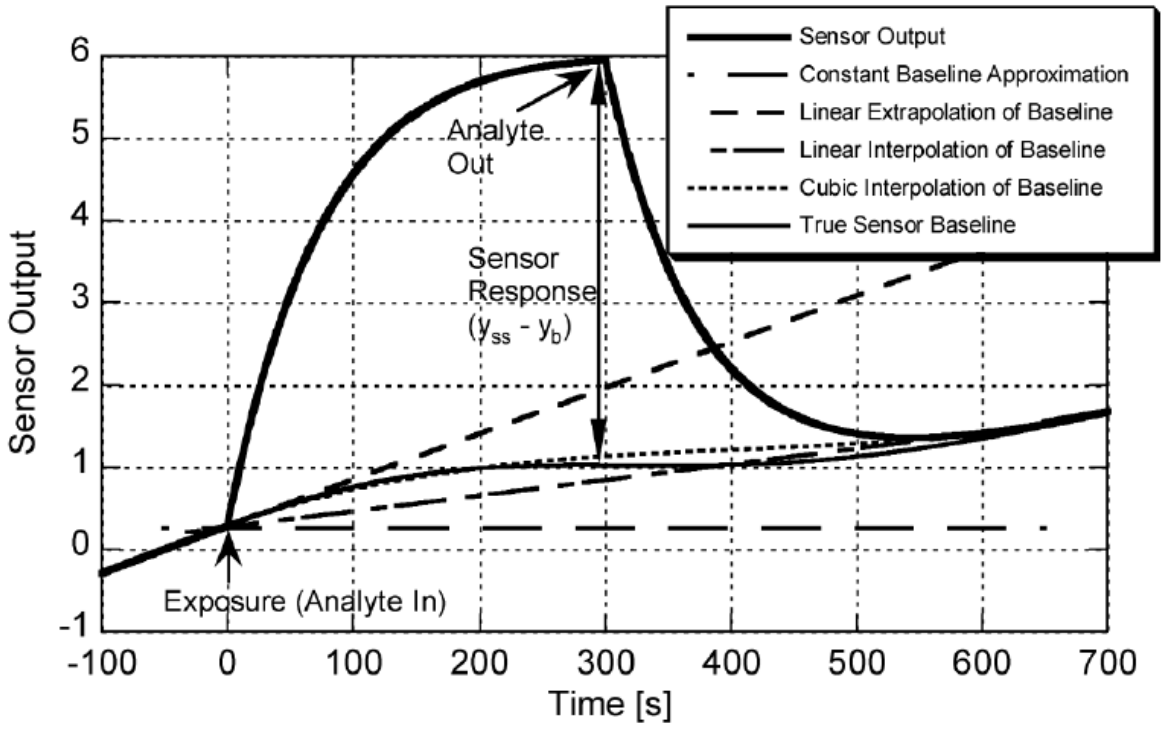

Figure 1.2: Illustration of several baseline correction techniques [15].

1.2.2 Time to Detection (Steady-State Information Extraction)

As mentioned earlier, the second issue associated with using steady-state features to identify and quantify the analyte(s) in the sample is time to detection. If the steady-state features are used identification and quantification of analyte(s) cannot be performed until the sensor response reaches its steady-state. In some cases, the sensor could take a fairly long time before it reaches steady-state. In these cases, if a dangerous chemical is present in the environment, the sensor would not be able to detect the dangerous chemical rapidly and the necessary remediation action could not be taken on time which could lead to undesirable outcome. Therefore, the time to detection is largely dependent on the

ability of the signal processing method used to rapidly extract the steady-state information from the sensor response.

Steady-state information can be extracted rapidly, if one uses estimation theory to estimate the steady-state of the sensor response well before the sensor response actually reaches steady-state. More details on estimation theory used to estimate the steady-state of a sensor response are given in the subsequent chapters.

Another approach to decrease the time required to quantify the analyte, is to use the initial derivative of sensor response instead of steady-state feature. By using initial derivative method, only the first few data points of the samples and an estimate of the initial derivative are required to quantify the analyte [15]. However, the initial derivative method is prone to flow effects (i.e. how quickly the sensor is exposed to the sample) and higher noise.

1.2.3 Transient Information Extraction

As mentioned earlier, the steady-state feature is commonly used to identify and to quantify the analyte. However, if one were to use the transient information together with the steady-state information, it will result in improved selectivity and increased recognition accuracy [18]. Therefore, transient information of a sensor response is vital in improving the identification of analyte.

One common approach to extract the transient information from the sensor response is by fitting the sensor response data with a single (or dual) exponential fit and determine the time constant from the exponential fit. This approach could take a long time because one has to wait until the response reaches steady-state before fitting the data to extract the transient information. In order to rapidly extract transient information, estimation theory can be used. By using estimation theory, one could actually estimate the transient information before the response reaches steady-state.

Other approach to extract the transient information includes transient integrals and dynamic slope that attempt to capture the transient information. More details on these two methods are given in [19]. Moreover, there is also some research conducted to evaluate the feasibility of applying Wavelets and Wavelet Transform methods to extract the transient information [20].

1.2.4 Sensor Array Processing

In a sensor array, further signal processing is required to identify and quantify the analyte. Analyte identification and quantification can be performed using a pattern recognition technique. Some common pattern recognition techniques include Bayesian analysis, nearest neighbor algorithm, linear discriminant functions and neural networks. The purpose of performing pattern recognition is to maximize the ability of a sensor array to identify and quantify the chemical or chemical mixtures. The pattern recognition process is divided into two steps, training and classification. The training step involves

teaching the pattern recognition algorithm how the sensor array will react to a known analyte [4]. The classification step is where the information learned in training is used to determine the analyte that is most likely to have caused the given response [4].

The data collected from a sensor array is typically very large. Therefore, a preprocessing technique is often used to reduce the dimensionality of the sensor array data without any loss of useful information before performing pattern recognition. Some of the common preprocessing technique includes principal component analysis (PCA) and linear discriminant analysis (LDA). Both PCA and LDA are used to find a smaller set of variables that are linear combinations of the original variables. The linear combinations are chosen in such a way that the pertinent information from the original data is retained in the lower dimension transformed space [4].

1.3 Problem Statement

As stated in section 1.1, the detection of certain chemicals has become of great importance for human health. This is particularly true in the detection of benzene in the groundwater sample. Long-term exposure to significant levels of benzene could cause cancer, leukemia, and anemia. Benzene is a volatile organic chemical and it is formed through natural or industrial processes [21]. Benzene is also a natural part of gasoline and typically, gasoline is stored in underground storage tanks (USTs). According to the United States Environmental Protection Agency (EPA), there is about 590 000 federally regulated USTs in the United States alone and about 6 000 leaks are reported annually

[10]. So, there is a high risk of benzene leaking into the soil and contaminating the groundwater. Therefore, regular monitoring of the area around the USTs is crucial to detect the leakage of benzene into the groundwater.

Currently, USTs are inspected at 2-3 year intervals and the groundwater samples collected at the monitoring wells have to be transported to a lab for analysis [10]. This current practice is time consuming and costly. Therefore, it is necessary to develop an in-situ chemical sensor and signal processing methods that is capable of rapidly analyzing and quantifying benzene. The process of developing an in-situ chemical sensor are currently being investigated and SH-SAW devices coated with certain type of polymers are showing promise in detecting benzene in trace amounts. However, the challenge in quantifying benzene is that, groundwater samples usually contain mixtures of multiple analytes which are chemically similar to benzene and it is difficult to extract the sensor response due to benzene alone.

In this work, efforts will be made to use estimation theory, in particular, Kalman Filter and Extended Kalman Filter (EKF) to analyze and to quantify benzene in binary mixtures of analyte. Although groundwater samples contain mixtures of multiple analytes, only binary mixtures of analytes will be considered in the present work. Furthermore, in this work only BTEX compounds (benzene, toluene, ethylbenzene and xylenes) and binary mixtures of BTEX compounds are considered. By using estimation theory (KF or EKF), it will be shown that the sensor parameters and concentrations of analyte(s) can be estimated in real-time well before the sensor response reaches steady-

state, thus saving time required to wait for the sensor response to reach steady-state before analyzing and quantifying the target analyte.

One of the challenges in using in-situ chemical sensors to monitor groundwater samples is the possibility of sensor baseline drift. Sensor baseline drift is a common problem that occurs in most chemical sensors; this is especially true for in-situ chemical sensors, where the sensor's environment is not controlled and temperature and humidity can fluctuate drastically. These influences of environmental parameters will not only cause the sensor baseline to drift but also might introduce some outlier points in the sensor measurement. The problem of sensor baseline drift and outlier points in the measurement is also addressed in this thesis. It will be shown that baseline drift can be corrected by using estimation theory, in particular, Kalman Filter (KF) and the outlier points can be corrected by using Low Pass Filter (LPF).

1.4 Organization of the thesis

The organization of the thesis is presented in this section. This thesis is presented in six chapters; Chapter 1 gives an introduction to the chemical sensors in general and specifically explains the importance of using a chemical sensor and signal processing methods to monitor the groundwater sample. Moreover, the importance of signal processing is emphasized and a brief review on signal processing methods is given. Chapter 1 also defines the problem that this thesis is addressing. In Chapter 2, a review of sensor signal processing using estimation theory, specifically Kalman Filter and

Extended Kalman Filter are discussed. Both Kalman Filter and Extended Kalman Filter are derived in this section and the algorithm on how to use Kalman Filter and Extended Kalman Filter to perform estimation is explained. In Chapter 3, modelling of the sensor responses to single and binary mixtures of analytes is shown. The discrete-time model and state-space model of both single analyte sensor response and two-analyte sensor response are given. In Chapter 4, the specifics of the Shear Horizontal Surface Acoustic Wave (SH-SAW) sensors that were used to collect the data analyzed in this thesis and the process of data acquisition using the SH-SAW sensor are discussed. Also in this chapter, the data processing including the baseline drift correction and the elimination of outlier points in the sensor data are explained. In Chapter 5, the estimation results using the data collected from the SH-SAW sensor are shown and the estimation results are discussed to highlight the advantages of using estimation theory to analyze and quantify both single and two-analyte system. Finally, Chapter 6 provides a summary of the work performed in this thesis, and also gives some suggestions regarding the possible extensions of this work for future research in the area.

2. SENSOR SIGNAL PROCESSING USING ESTIMATION THEORY: A REVIEW

2.1 Introduction

As mentioned in Chapter 1, sensor signal processing is at the core of chemical sensor systems and the purpose of sensor signal processing is to extract the performance criteria of the sensor. Some of the essential performance criteria for a chemical sensor include sensitivity, selectivity, response time and reproducibility. Sensitivity dictates the minimum amount of the target analyte that can be detected by the chemical sensor. In many applications, it is desired that small amounts of target analyte produce large changes in the measured signal. Therefore, low sensitivity might result in difficulties in the detection of trace amounts of the target analyte. Response time is the time required for a chemical sensor to respond when it is exposed to the target analyte. If the response time is long, it will result in difficulties in rapidly quantifying the target analyte. Moreover, poor selectivity and reproducibility as well as sensor aging and drift caused by environmental influences can all result in difficulties in the detection of target analyte [15]. All these issues, which limit the applicability of chemical sensors, could be improved by incorporating an effective signal processing method such as estimation theory in analyzing the sensor data.

Estimation theory is a branch of statistics and signal processing that deals with estimating the values of unknown parameters based on the measurement data [22, 23]. Basically, the estimation is done by using an estimator that attempts to approximate the

unknown parameters using the available measurements. There are three common problems in estimation, which consist of smoothing, filtering and prediction. Smoothing is a process where the past value of the unknown parameter is estimated using the available measurement data. Filtering is a process where the present value of the unknown parameter is estimated using the available measurement data and finally, prediction is a process where the future value of the unknown parameter is estimated using the available measurement data [15]. There are various forms of estimator and estimation methods which can be used to perform the estimation, and are Kalman Filter and its various derivatives, maximum likelihood estimators, Bayes Estimators, Cramer-Rao Bound, Wiener Filter, Particle Filter and Markov Chain Monte Carlo (MCMC). Estimation theory is used in numerous fields as shown in Table 2.1.

Area of application	Example application
Control Systems	Estimation of the position of a powerboat for correcting navigation in the presence of sensor and environmental noise.
Communications	Estimation of the carrier frequency of a signal for demodulation to the baseband in the presence of degradation noise.
Seismology	Estimation of the underground distance of an oil deposit in the presence of noisy sound reflections.
Biomedical	Estimation of the heart rate of a fetus in the presence of environmental noise.

Image Processing	Estimation of the position and orientation of an object from a camera image in the presence of lighting and background noise.
Radar Communications	Estimation of the delay in the received pulse echo in the presence of noise.
Speech Signal Processing	Estimation of the parameters of the speech model in the presence of speech variability and environmental noise.
Sensor Signal Processing	Estimation of the baseline drifts in the sensor response in the presence of noise.

Table 2.1: Applications of estimation theory [23].

As mentioned in chapter 1, in the present work, estimation theory and in particular Kalman Filter (KF) and one of its derivatives, Extended Kalman Filter (EKF) will be used extensively as a means toward chemical sensor signal processing. Therefore in this chapter, KF and EKF will be discussed rigorously. Both KF and EKF will be derived in this chapter and the algorithms on how to apply KF and EKF to perform estimation are explained. In section 2.4, some typical applications of various forms of Kalman Filters are given.

2.2 Kalman Filter

Kalman Filter (KF) is a set of mathematical equations that provides an efficient recursive means to estimate the state of a process in a way that minimizes the mean of the

squared error [24]. Basically, KF is capable of estimating the present value of an unknown state given the available measurement data. Moreover, KF can also be used to estimate the past and future value of the unknown state by appropriate modifications to the filter. Therefore, KF is a very powerful algorithm that can be used to estimate past, present and future states of a system and it can do so even when the precise nature of the modeled system is unknown [24].

Kalman Filter is named after R.E. Kalman, one of the primary developers of its theory. In 1960, R.E. Kalman first used Kalman Filter to obtain a recursive solution to the discrete-data linear filtering problem [24, 25]. Since then, there has been a tremendous research on Kalman Filter and today Kalman Filters are being used in many areas of applications, particularly in the area of assisted navigation.

2.2.1 Kalman Filter Derivation

Consider a general linear stochastic discrete-time system with internal states x_k, outputs y_k, inputs u_k, and time-varying system matrices A_k, B_k, C_k and D_k as following,

$$x_{k+1} = A_k x_k + B_k u_k + F_k v_k \tag{2.1a}$$

$$y_k = C_k x_k + D_k u_k + G_k w_k \tag{2.1b}$$

where, v_k is the process or state noise with covariance V_k, and w_k represent the measurement noise with covariance W_k. The cross covariance between the process and measurement noise is S_k. For the sensor systems considered in the present work, the internal states, x_k represent the parameters that need to be estimated such as time constant, steady-state value and concentration of analyte, the output, y_k represents the sensor measurement data and the input, u_k represents the unit step input. Assuming that the system in eq. 2.1 meets the detectability criteria (i.e. if all unstable modes of the system are observable) [26], then it is possible to estimate the unknown states, x_k of the system by using only the available measurement data, y_k [27].

The first step in the derivation of the Kalman Filter is to assume an estimator to estimate the new state of the system, $\hat{x}_{k+1}$. In order to form an estimator, one should look at the type of information that is available at any time, k. Typically, one will have access to three sources of information at time, k which include the present value of the state estimates, $\hat{x}_k$, the present value of the input, u_k and the present value of the measurement, y_k. By using this available information about the system, an estimator of the form given in eq. 2.2 can be formed [27].

$$\hat{x}_{k+1} = A_k\hat{x}_k + B_k u_k + K_k(y_k - \hat{y}_k) \tag{2.2}$$

where $\hat{y}_k$ is the estimate of the measurement given by eq. 2.3,

$$\hat{y}_k = C_k\hat{x}_k + D_k u_k$$

(2.3)

The goal is now to find the Kalman gain, K_k which minimizes the variances of the error given by,

$$e_{k+1} = x_{k+1} - \hat{x}_{k+1}$$

(2.4)

while the estimate of the unknown states remains unbiased (i.e. $E\{e_{k+1}\} = 0$). Before finding the Kalman gain, K_k, one has to find the error covariance, $P_{k+1} = E\{(e_{k+1})(e_{k+1})^T\}$ first. By substituting eq. 2.1(a), eq. 2.1(b), eq. 2.2 and eq. 2.3 into eq. 2.4, eq. 2.5 will be obtained after some manipulation.

$$e_{k+1} = (A_k - K_k C_k)e_k + F_k v_k - K_k G_k w_k$$

(2.5)

From eq. 2.5, one can find the error covariance,

$$P_{k+1} = E\{(e_{k+1})(e_{k+1})^T\}$$

$$P_{k+1} = A_k P_k A_k^T - A_k P_k C_k^T K_k^T - K_k C_k P_k A_k^T + K_k C_k P_k C_k^T K_k^T + F_k V_k F_k^T - K_k G_k S_k^T F_k^T - F_k S_k G_k^T K_k^T + K_k G_k W_k G_k^T K_k^T$$

(2.6)

Once the error covariance is found, the next step is to find the Kalman gain, K_k. Note that the error covariance matrix, P_{k+1} is a diagonal and symmetric matrix. Therefore, minimizing the error covariance matrix is equivalent to minimizing the trace of P_{k+1} (i.e. $Tr\{P_{k+1}\}$). Thus, the Kalman gain, K_k, can be found by taking the partial derivative of $Tr\{P_{k+1}\}$ with respect to K_k and solving it for K_k by setting the resulting equation to zero [28]. By taking the partial derivative of $Tr\{P_{k+1}\}$ with respect to K_k, eq. 2.7 will be obtained as,

$$\frac{\delta\, Tr\{P_{k+1}\}}{\delta\, K_k} = -2A_k P_k C_k^T - 2F_k S_k G_k^T + 2K_k(C_k P_k C_k^T + G_k W_k G_k^T) \tag{2.7}$$

Setting eq. 2.7 to zero and solving it for K_k, the following result will be obtained,

$$K_k = (A_k P_k C_k^T + F_k S_k G_k^T)(C_k P_k C_k^T + G_k W_k G_k^T)^{-1} \tag{2.8}$$

The Kalman gain obtained in eq. 2.8 is the value of the gain that would result in minimum error covariance at any given time k. The error covariance equation as given in eq. 2.6 can be further simplified using eq. 2.8,

$$P_{k+1} = A_k P_k A_k^T + F_k V_k F_k^T - (A_k P_k C_k^T + F_k S_k G_k^T)(C_k P_k C_k^T + G_k W_k G_k^T)^{-1} (C_k P_k A_k^T + G_k S_k^T F_k^T) \tag{2.9}$$

If the process and measurement noise are white noise which is usually the case for most systems and also for the sensor system considered in the present work, then their values will be uncorrelated with each other. Therefore, the cross-covariance, S_k, will be zero. If the cross-covariance is zero, eq. 2.8 and eq. 2.9 can be simplified to

$$K_k = A_k P_k C_k^T \left(C_k P_k C_k^T + G_k W_k G_k^T\right)^{-1} \tag{2.10}$$

$$P_{k+1} = A_k P_k A_k^T + F_k V_k F_k^T - A_k P_k C_k^T \left(C_k P_k C_k^T + G_k W_k G_k^T\right)^{-1} \left(C_k P_k A_k^T\right) \tag{2.11}$$

Finally, in the derivation of the Kalman Filter, the state update equation needs to be found. The state update equation can be found by substituting eq. 2.3 into eq. 2.2,

$$\hat{x}_{k+1} = A_k \hat{x}_k + B_k u_k + K_k \left(y_k - \left[C_k \hat{x}_k + D_k u_k\right]\right) \tag{2.12}$$

The last three equations, i.e. eq. 2.10, eq. 2.11, and eq. 2.12 lead to a recursive algorithm for updating the state estimate based on the measurement data. Since the estimation is performed in a recursive manner, only the current estimate of the states and the latest measurement value is required to update the state estimate. Therefore, by using Kalman Filter, the estimation can be performed in real time under strict memory requirements as one does not have to store all the measurement values.

2.2.2 Kalman Filter Algorithm

In the previous section, it has been established that eq. 2.10, eq. 2.11 and eq. 2.12 lead to a recursive algorithm that can be used to update and estimate the unknown states based on the measurement data. In this section, the algorithm on how to use Kalman Filter in order to perform estimation is explained.

Kalman Filter algorithm works in a two-step process consisting of the prediction step and correction step [24]. The prediction step is responsible for projecting the current state estimate and error covariance forward in time to obtain a new state estimate and the error covariance for the next time step. On the other hand, in the correction step, the new state estimate will be updated or corrected using the new measurement value with a weighted average, where more weight is assigned to estimates with higher certainty. Both steps are implemented simultaneously by eq. 2.10, eq. 2.11 and eq. 2.12.

Before applying the Kalman Filter algorithm, one has to assign a state variable to the unknown parameters that need to be estimated and determine the system matrices $A_k, B_k,\ C_k$ and D_k. One should also determine the measurement noise covariance, W_k and process noise covariance, V_k. In applying the Kalman Filter algorithm, the first step is to initialize the state estimate, $\hat{x}_0$ and error covariance, P_0 at time, $k = 0$. If the uncertainty about the system is high, one should set the initial error covariance, P_0 to be high as well [24]. For example, if the initial value of the state is completely unknown, one should make an educated guess and set a value for the initial

value of the state estimate, $\hat{x}_0$ and should also set the initial error covariance, P_0 to a large value. Once the initialization process is complete, the next step is to calculate the Kalman gain, K_k at time, $k = 0$. After that, by using the Kalman gain, K_k and the available measurement, y_k at time, $k = 0$, update the state estimate, $\hat{x}_{k+1}$ and error covariance, P_{k+1}. Note that, after the update process, one will obtain a new state estimate and a new error covariance for time, $k = 1$. Next, increment the time to $k = 1$ and by using the new state estimate and error covariance obtained at time, $k = 0$, repeat the process again starting with calculating the Kalman gain, K_k. Repeat the process until the error covariance becomes very small or until the new measurement is taken. Kalman Filter algorithm is summarized in Fig. 2.1.

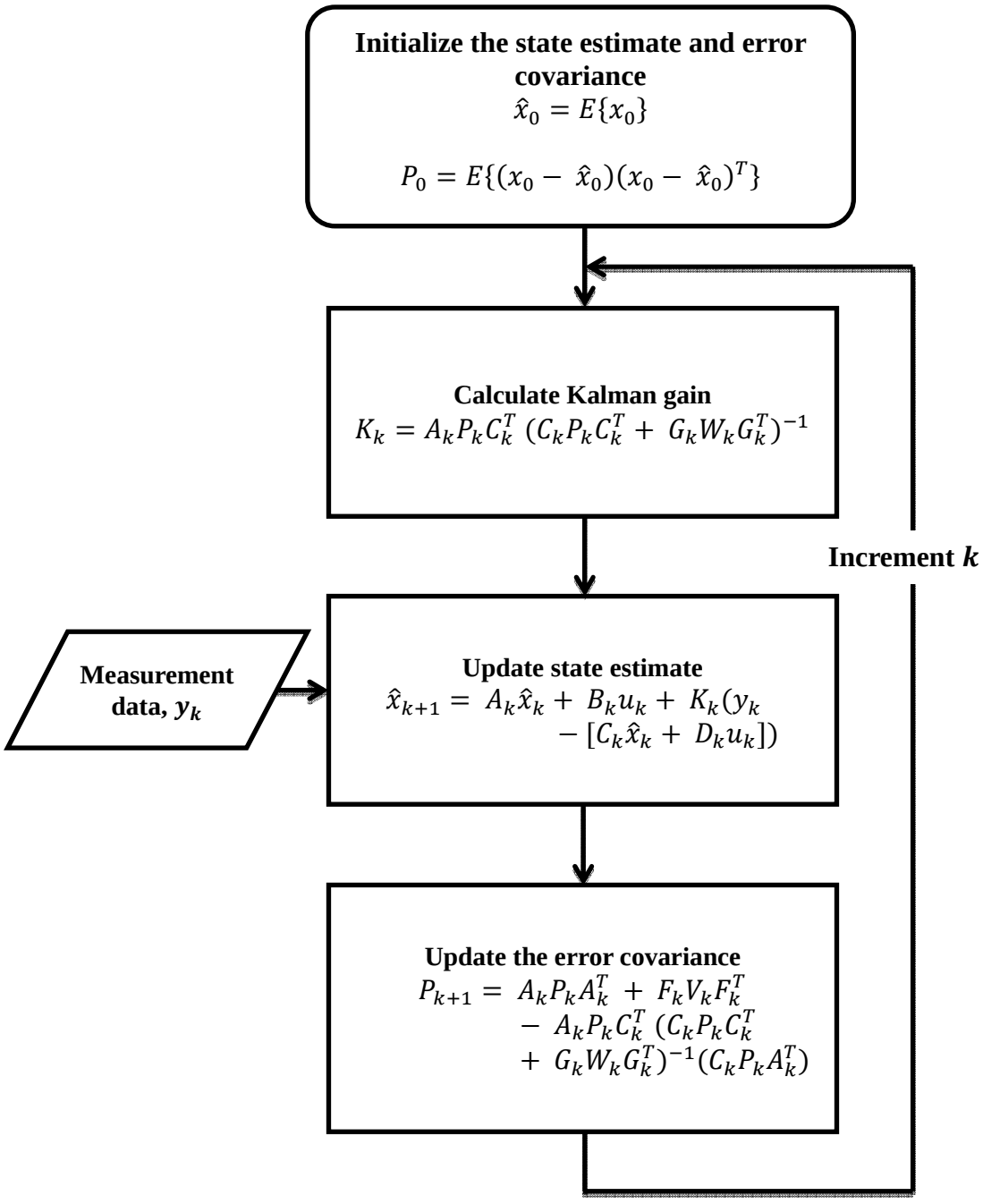

Figure 2.1: Flowchart of Kalman Filter algorithm.

2.3 Extended Kalman Filter

In this section, one of the most important derivatives of the Kalman Filter, the Extended Kalman Filter (EKF) will be discussed. As described in section 2.2, Kalman Filter addresses the problem of estimating the unknown states of a linear stochastic discrete-time system. If the system is non-linear, then some modifications need to be made, so that the estimation of non-linear state-space systems can be performed using Kalman Filter. This modification is done in the form of Taylor series expansion about the current state-estimate and neglecting the higher order terms (i.e. terms higher than first order) [29, 30]. Since this modification is just an extension to the original Kalman Filter, thus it is referred to as Extended Kalman Filter (EKF).

2.3.1 Extended Kalman Filter Derivation

Consider a general non-linear stochastic discrete-time system with internal states x_k, outputs y_k, inputs u_k, given by,

$$x_{k+1} = f(x_k, u_k, v_k) \tag{2.13}$$

$$y_k = h(x_k, u_k, w_k) \tag{2.14}$$

The non-linear system given in eq. 2.13 and eq. 2.14 can be linearized around the current state estimate by performing the Taylor series expansion and neglecting the higher order terms (i.e. terms higher than first order), which will lead to the following approximation,

$$x_{k+1} \cong f(\hat{x}_k, u_k, \bar{v}) + \left[\left(\frac{\partial f}{\partial x}\right)_{\substack{x=\hat{x}_k \\ u=u_k \\ v_k=\bar{v}}}\right] e_k + \left[\left(\frac{\partial f}{\partial v}\right)_{\substack{x=\hat{x}_k \\ u=u_k \\ v_k=\bar{v}}}\right] v_k \tag{2.15}$$

$$y_k \cong h(\hat{x}_k, u_k, \bar{w}) + \left[\left(\frac{\partial h}{\partial x}\right)_{\substack{x=\hat{x}_k \\ u=u_k \\ w_k=\bar{w}}}\right] e_k + \left[\left(\frac{\partial h}{\partial w}\right)_{\substack{x=\hat{x}_k \\ u=u_k \\ w_k=\bar{w}}}\right] w_k \tag{2.16}$$

Note that the partial derivatives are evaluated at the current state estimate, known input value and mean of noise. The partial derivatives are actually time-varying Jacobian matrices and can be redefined as shown,

$$A_k = \left(\frac{\partial f}{\partial x}\right)_{\substack{x=\hat{x}_k \\ u=u_k \\ v_k=\bar{v}}} \tag{2.17}$$

$$C_k = \left(\frac{\partial h}{\partial x}\right)_{\substack{x=\hat{x}_k \\ u=u_k \\ w_k=\bar{w}}} \tag{2.18}$$

$$F_k = \left(\frac{\partial f}{\partial v}\right)_{\substack{x=\hat{x}_k \\ u=u_k \\ v_k=\bar{v}}} \tag{2.19}$$

$$G_k = \left(\frac{\partial h}{\partial w}\right)_{\substack{x=\hat{x}_k \\ u=u_k \\ w_k=\bar{w}}} \tag{2.20}$$

For EKF, these Jacobian matrices will serve as the system matrices and can be used to perform the estimation in a similar fashion as Kalman Filter by using eq. 2.10, eq. 2.11 and eq. 2.12. However, some changes need to be made to eq. 2.10 to perform the estimation of nonlinear systems. These changes are as following,

$$\hat{x}_{k+1} = f(\hat{x}_k, u_k, \bar{v}) + K_k[y_k - h(\hat{x}_k, u_k, \bar{w})] \tag{2.21}$$

Therefore, for EKF, eq. 2.21, eq. 2.11 and eq. 2.12 will lead to a recursive algorithm for updating the state estimate based on the measurement data. It is important to note that EKF is not an optimal filter because Gaussianity of the probability distributions will not be preserved under a non-linear transformation [4]. However, the EKF does give useful estimates of the states and will demonstrate convergence for certain conditions [31]. The convergence of EKF is dependent on the initial value of the error covariance and the value of process and measurement noise. Therefore, to assure convergence in the sensor

system considered in the present work, care must be taken in setting the initial value of the error covariance, value of the process noise and value of the measurement noise.

2.3.2 Extended Kalman Filter Algorithm

In this section, the process on how to use Extended Kalman Filter (EKF) is explained. EKF works in a similar way as Kalman Filter. The main difference is, for EKF one extra step is required that is to linearize the non-linear system about the current state estimate. By linearizing the non-linear system, Jacobian matrices as given in eq. 2.17 to eq. 2.20 can be found and these matrices will serve as the system matrices for the estimation process.

Therefore, before applying the EKF algorithm, one has to first assign state variables to unknown parameters that need to be estimated and find the general form of Jacobian matrices. One should also determine the measurement noise covariance, W_k and mean, $\bar{w}$ and also process noise covariance, V_k and mean, $\bar{v}$. Similar to Kalman Filter algorithm, the first step in applying the EKF algorithm is to initialize the state estimate, $\hat{x}_0$ and error covariance, P_0 at time, $k = 0$. Once initialization is complete, the next step is to evaluate the Jacobian matrices at time $k = 0$. After that, the Kalman gain, K_k at time, $k = 0$ is calculated. By using the Kalman gain, K_k and the available measurement, y_k at time, $k = 0$, one can update the state estimate, $\hat{x}_{k+1}$ and error covariance, P_{k+1}. Once the update process is complete, one will obtain new state estimate and new error covariance for the next time step, i.e., for time, $k = 1$. Finally, the time is

incremented to $k = 1$ and by using the new state estimate and new error covariance obtained earlier, the process is repeated again starting with calculating the Kalman gain. The Extended Kalman Filter algorithm is summarized in Fig. 2.2.

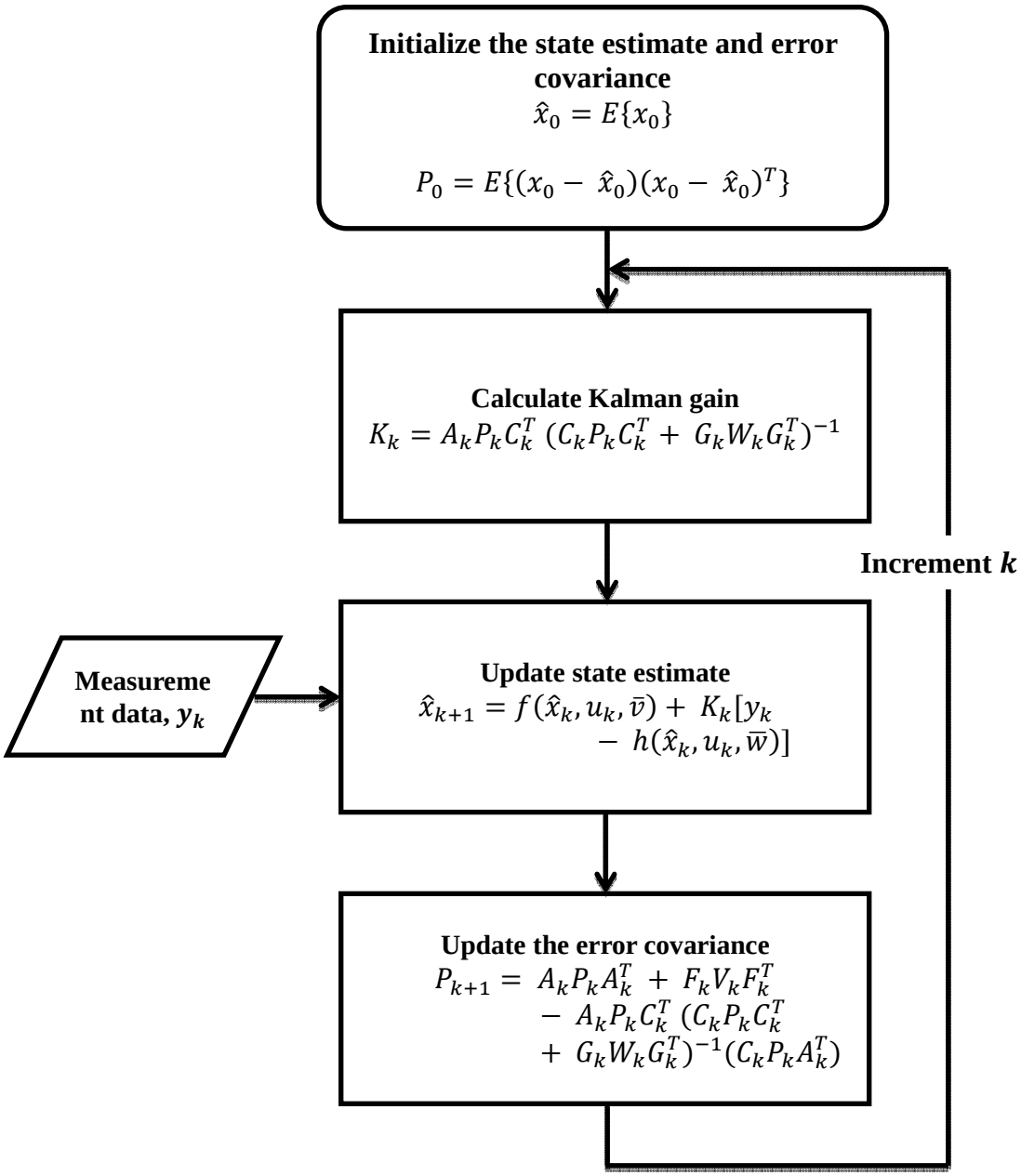

Figure 2.2: Flowchart of Extended Kalman Filter algorithm.

2.4 Applications of Kalman filter

In this section, some typical applications of various form of Kalman Filter will be reviewed. As mentioned earlier, since 1960, after R.E Kalman published his groundbreaking paper, Kalman Filter has been the focus of extensive research and applications [24]. Today, Kalman Filter and its derivatives have found many applications not only in the field of engineering and mathematics but also in the field of economics. Some typical application areas and example applications are listed in Table 2.2. In the present work, it will be shown that Kalman Filter and Extended Kalman Filter can also be used in sensor signal processing.

Area of application	Example applications
Navigation	To control and assist the navigation of automobiles, aircraft or spacecraft using the measured sensor data in a noisy environment [32].
Image processing	Using various forms of Kalman Filter to filter the noise out of the measured images.
Radar communications	Estimating the distance of the target object.
Control system	Active noise control in noisy systems [27].
Economics	Parameter estimation of linear and non-linear econometric models [33].

Speech signal processing	To estimate the parameters of the speech model and also to filter the noise out of the speech signal.
Forecasting	To estimate the parameters of the forecasting model using the measured data.
Sensor signal processing	To perform the baseline drift correction, to extract the transient information and to predict the steady-state information from the sensor response before the response reaches steady-state.

Table 2.2: Typical applications of various forms of Kalman Filter.

3. MODEL OF SENSOR RESPONSES TO SINGLE AND BINARY MIXTURES OF ANALYTES

3.1 Introduction

In this chapter, a model of the sensor response to a single analyte sample and a model of the sensor response to a mixture of two analytes are discussed. In order to model the sensor response for each case, several assumptions were made and are discussed in this chapter. Since the sensor data considered in the present work are collected at discrete-time instants, the discrete-time model of single analyte sensor response and two-analyte sensor response were found by using the Euler's continuous time approximation formula. The discrete-time model of the sensor response can also be used for sensor systems which are implemented using a microcontroller for which the sensors outputs are sampled at discrete-time instants. Moreover, the transformation of the discrete-time model of the sensor response to the state-space model of the sensor response is presented. Transforming the discrete-time model of the sensor response into the state-space model of the sensor response is important so that the unknown parameters from the sensor response can be estimated using discrete Kalman Filter (KF) or discrete Extended Kalman Filter (EKF) as discussed in the previous chapter. For the single analyte system, the unknown quantities that need to be estimated include normalized concentration of the analyte (ratio of the concentration of analyte in the coating at time, t to the product of the polymer-liquid partition coefficient and maximum ambient concentration), steady-state frequency shift and time constant. For the two-analyte system, two different models were developed, a nonlinear model and a linear model. The

main difference between the models lies in the formulation of their state-space form. For both models, it is assumed that the time constant for both analytes is known (from the single analyte experiments). For the nonlinear model, the unknown quantities that need to be estimated are the normalized concentration of each analyte and the steady-state frequency shift of each analyte. For the linear model, the unknown quantities that need to be estimated are the steady-state frequency shift of the analytes. This is because for the linear model, the normalized concentrations of each analyte were determined for every time instant by using the known time constant of the analytes and the sensor response model. The unknown quantities for both single analyte system and two-analyte system can be estimated using estimation theory (particularly KF or EKF). It should also be noted that the model of sensor responses presented in this chapter can be used for most sensor platforms.

3.2 Single Analyte System

In order to model the single analyte system, it has been assumed that the single analyte system obeys Henry's law (for concentrations below 50 ppm [7, 34, 35]) and it has been shown that this assumption is valid in [7, 34, 35]. Typically, when the sensor is exposed rapidly to an analyte of a given ambient concentration, the sensor will respond rapidly at first and then slowly as the system reaches equilibrium (i.e. one can assume a step profile in the concentration as a function of time). This is true for single analyte system and the process of analyte absorption can be assumed to be first order and is described by,

$$\dot{C}(t) = -\frac{1}{\tau}C(t) + \frac{\gamma_p}{\tau}\, C_{amb}(t) \tag{3.1}$$

where $C(t)$ is the concentration of analyte in the coating at time, t, $C_{amb}(t)$ is the ambient analyte concentration at time, t, τ is the response time constant, and γ_p is the polymer-liquid partition coefficient. For single analyte system, the measured frequency shift at time, t is given by,

$$\Delta f(t) = -aC(t) \tag{3.2}$$

where $\Delta f(t)$ is the frequency shift observed at time, t, and a is the steady-state frequency shift which is a function of the sensor platform, the sensor coating, and the analyte. Both eq. 3.1 and eq. 3.2 were normalized by dividing with $\gamma_p C_{max}$ (where C_{max} represent the maximum ambient concentration) and the following equations were obtained,

$$\frac{\dot{C}(t)}{\gamma_p C_{max}} = -\frac{1}{\tau}\frac{C(t)}{\gamma_p C_{max}} + \frac{1}{\tau}\frac{C_{amb}(t)}{C_{max}} \tag{3.3}$$

$$\Delta f(t) = -a\gamma_p C_{max} \frac{C(t)}{\gamma_p C_{max}} \tag{3.4}$$

By defining new variables as follows,

$$m(t) = \frac{C(t)}{\gamma_p \, C_{max}}$$

$$u_s(t) = \frac{C_{amb}(t)}{C_{max}}$$

$$\alpha = -a\gamma_p C_{max}$$

eq. 3.3 and eq. 3.4 were rewritten as,

$$\dot{m}(t) = -\frac{1}{\tau} m(t) + \frac{1}{\tau} u_s(t) \tag{3.5}$$

$$\Delta f(t) = -\alpha \, m(t) \tag{3.6}$$

where $m(t)$ represent the normalized concentration of absorbed analyte at time, t, α is the normalized steady-state frequency shift and $u_s(t)$ represents the unit step input (for

$t < 0$, $C_{amb} = 0$ and for $t > 0$, $C_{amb}(t) = C_{max}$). Eq. 3.5 and eq. 3.6 are the single analyte normalized equations which were used to represent the single analyte absorption.

3.2.1 Discrete-Time Model

The frequency shifts of the single analyte system are measured at discrete-time instants (i.e. $t = kT$, where T is the sampling period). Therefore, it is necessary to transform the continuous time model of the single analyte sensor response given in eq. 3.5 and eq. 3.6 into a discrete-time model. In order to discretize the continuous time model, Euler's continuous time approximation formula was used. Based on Euler's formula, the first derivative of the normalized concentration can be approximated by,

$$\dot{m}(t) = \frac{m_{k+1} - m_k}{T} \tag{3.7}$$

By applying Euler's approximation of the first derivative of the normalized concentration as given in eq. 3.7, the discrete-time model of the single analyte system is found to be,

$$m_{k+1} = \left(1 - \frac{T}{\tau}\right) m_k + \frac{T}{\tau} u_{s,k} + v_k \tag{3.8}$$

$$\Delta f_k = \alpha\, m_k + w_k \tag{3.9}$$

where v_k and w_k are added to represent the process and measurement noise that are present in the single analyte system (v_k and w_k are uncorrelated white noise with zero mean). By defining absorption rate constant as,

$$S = \frac{T}{\tau}$$

eq. 3.8 can be rewritten as,

$$m_{k+1} = (1 - S)m_k + S\,u_{s,k} + v_k \tag{3.10}$$

In conclusion, eq. 3.9 and eq. 3.10 represent the discrete-time model of the single analyte system.

3.2.2 State-Space Model

The discrete-time model of the single analyte system should be transformed into the state-space model, so that the unknown parameters from the single analyte response can be estimated using Extended Kalman Filter (EKF). As mentioned earlier, for the single analyte system, it is assumed that the normalized concentration of the analyte, the steady-state frequency shift and the time constant (absorption rate) are the unknown parameters. The normalized concentration is actually dependent on the absorption rate, thus by knowing the absorption rate, one could actually determine the normalized

concentration or vice versa but need to wait until the sensor response reaches steady-state. Therefore, to perform the estimation in real-time before the response reaches steady-state, one should assume both normalized concentration and absorption rate to be unknown, so that both unknowns can be estimated simultaneously. Note that if the steady-state frequency shift is assumed to be known, then only the other two unknowns, the normalized concentration and the absorption rate constant need to be estimated (this case is not considered in the present work).

In order to convert the discrete-time model into the state-space model, the first step is to assign state variables to the unknown parameters,

$$\begin{bmatrix} x_k^{(1)} \\ x_k^{(2)} \\ x_k^{(3)} \end{bmatrix} = \begin{bmatrix} m_k \\ S \\ \alpha \end{bmatrix}$$

and define the output as,

$$y_k = \Delta f_k$$

Eq. 3.9 and eq. 3.10 can be rewritten in terms of these state variables,

$$x_{k+1}^{(1)} = \left(1 - x_k^{(2)}\right) x_k^{(1)} + x_k^{(2)} u_{s,k} + v_k \tag{3.11}$$

$$y_k = x_k^{(3)} x_k^{(1)} + w_k$$

(3.12)

Since the absorption rate, S and the steady-state frequency shift, α are constants, the values of these parameters at time, $k+1$ are equal to their values at time, k.

$$x_{k+1}^{(2)} = x_k^{(2)}$$

(3.13)

$$x_{k+1}^{(3)} = x_k^{(3)}$$

(3.14)

Eq. 3.11, eq. 3.12, eq. 3.13 and eq. 3.14 form the state-space model of the single analyte system and can be rewritten as,

$$x_{k+1} = f(x_k, u_k, v_k) = \begin{bmatrix} x_{k+1}^{(1)} \\ x_{k+1}^{(2)} \\ x_{k+1}^{(3)} \end{bmatrix} = \begin{bmatrix} \left(1 - x_k^{(2)}\right) x_k^{(1)} + x_k^{(2)} u_{s,k} + v_k \\ x_k^{(2)} \\ x_k^{(3)} \end{bmatrix}$$

(3.15a)

$$y_k = h(x_k, u_k, w_k) = x_k^{(3)} x_k^{(1)} + w_k$$

(3.15b)

From eq. 3.15, it can be seen that the state-space model describing the single analyte sensor response is a nonlinear model. Therefore, to perform the estimation, EKF should

be used. In order to apply the EKF algorithm, the nonlinear system described by eq. 3.15 has to be linearized using Taylor series expansion. For the single analyte system, the result of linearization were obtained as described by eq. 3.16,

$$x_{k+1} = f(x_k, u_k, v_k) \approx f(\hat{x}_k, u_k, \bar{v}_k) + A_k(x_k - \hat{x}_k) + F_k v_k \tag{3.16a}$$

$$y_k = h(x_k, u_k, w_k) \approx h(\hat{x}_k, u_k, \bar{w}_k) + C_k(x_k - \hat{x}_k) + G_k w_k \tag{3.16b}$$

where,

$$f(\hat{x}_k, u_k, \bar{v}_k) = \begin{bmatrix} \left(1 - \hat{x}_k^{(2)}\right)\hat{x}_k^{(1)} + \hat{x}_k^{(2)} u_{s,k} \\ \hat{x}_k^{(2)} \\ \hat{x}_k^{(3)} \end{bmatrix}$$

$$h(\hat{x}_k, u_k, \bar{w}_k) = \hat{x}_k^{(3)}\hat{x}_k^{(1)}$$

$$A_k = \begin{bmatrix} \left(1 - \hat{x}_k^{(2)}\right) & -\hat{x}_k^{(1)} + u_{s,k} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$F_k = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$C_k = \begin{bmatrix} \hat{x}_k^{(3)} & 0 & \hat{x}_k^{(1)} \end{bmatrix}$$

$$G_k = [1]$$

By using the linearization result described by eq. 3.16, EKF algorithm can be applied to perform the estimation of the unknown parameters. The information on how to apply the EKF algorithm was discussed in chapter 2 and the EKF algorithm was summarized in Fig. 2.2.

3.3 Two-Analyte System

In order to model the two-analyte system, two main assumptions were made. The first assumption is that the mixture obeys Fick's law of absorption which states that when the mixture is extremely dilute, the sorption of one analyte into the polymer does not interfere with the sorption of the second analyte in any way. Free partitioning of the analyte between polymer and aqueous phase is assumed, implicating that the sorption process is reversible (i.e. only physisorption occurs). Fick's law of absorption is only valid for analyte concentrations below 50 ppm based on experimental observations [7, 34, 35]. From this assumption it follows that the concentration of the binary mixture in the coating at any time, t, is actually the sum of the concentration of each individual analyte in the mixture (i.e. $C_{mixture} = C_A(t) + C_B(t)$ for any time, t where subscript A and B represents two different analytes), and the process of analyte absorption can be assumed to be first order and is given by,

$$\dot{C}_A(t) = -\frac{1}{\tau_A} C_A(t) + \frac{\gamma_{p,A}}{\tau_A} C_{amb,A}(t) \tag{3.17a}$$

$$\dot{C}_B(t) = -\frac{1}{\tau_B} C_B(t) + \frac{\gamma_{p,B}}{\tau_B} C_{amb.B}(t) \tag{3.17b}$$

where $C_{amb,A}(t)$ and $C_{amb.B}(t)$ are the ambient analyte concentration of analyte A and analyte B at time, t, respectively, τ_A and τ_B are the response time constant of analyte A and analyte B respectively, and $\gamma_{p,A}$ and $\gamma_{p,B}$ are the polymer-liquid partition coefficient of analyte A and analyte B respectively.

For the two-analyte system, it is also assumed that the steady-state frequency shifts are also mutually unaffected, that is the frequency shift of the mixture at any time, t, is the sum of the frequency shifts due to each analyte in the mixture at any time, t. Therefore, the frequency shift for the binary mixture at any time, t, is given by,

$$\Delta f(t) = -a_A C_A(t) - a_B C_B(t) \tag{3.18}$$

where $\Delta f(t)$ is the frequency shift of the two-analyte system observed at time, t, a_A is the steady-state frequency shift of analyte A, a_B is the steady-state frequency shift of analyte B, $C_A(t)$ and $C_B(t)$ are the absorbed concentrations of the analyte A and analyte B in the

coating at time, t, respectively. From this assumption, it follows that the steady-state frequency shift for the mixture will also be the sum of the steady-state frequency shift of each analyte in the mixture and the response times of the analytes in the mixture will also be the same as those obtained from the single analyte measurements. Also note that this assumption is true for sufficiently low analyte concentrations in the range of parts per million (ppm) to parts per billion (ppb) because Henry's law can be applied for low concentrations of analytes (the investigation on the validity of Henry's law in the low concentration range of ppm to ppb is shown in [7, 34, 35]). Eq. 3.17a and eq. 3.17b were normalized by dividing with $\gamma_{p,A}C_{max,A}$ and $\gamma_{p,B}C_{max,B}$ respectively (where $C_{max,A}$ and $C_{max,B}$ represents the maximum ambient analyte concentrations of analyte A and analyte B correspondingly),

$$\frac{\dot{C}_A(t)}{\gamma_{p,A}\,C_{max,A}} = -\frac{1}{\tau_A}\frac{C_A(t)}{\gamma_{p,A}\,C_{max,A}} + \frac{1}{\tau_A}\,\frac{C_{amb,A}(t)}{C_{max,A}} \tag{3.19a}$$

$$\frac{\dot{C}_B(t)}{\gamma_{p,B}\,C_{max,B}} = -\frac{1}{\tau_B}\frac{C_B(t)}{\gamma_{p,B}\,C_{max,B}} + \frac{1}{\tau_B}\,\frac{C_{amb.B}(t)}{C_{max,B}} \tag{3.19b}$$

By defining new variables as follows,

$$m_A(t) = \frac{C_A(t)}{\gamma_{p,A}C_{max,A}}$$

$$m_B(t) = \frac{C_B(t)}{\gamma_{p,B} C_{max,B}}$$

$$u_A(t) = \frac{C_{amb,A}(t)}{C_{max,A}}$$

$$u_B(t) = \frac{C_{amb,B}(t)}{C_{max,B}}$$

$$\alpha_A = -a_A \gamma_{p,A} C_{max,A}$$

$$\alpha_B = -a_B \gamma_{p,B} C_{max,B}$$

eq. 3.18 and eq. 3.19 were rewritten as,

$$\dot{m}_A(t) = -\frac{1}{\tau_A} m_A(t) + \frac{1}{\tau_A} u_A(t) \tag{3.20a}$$

$$\dot{m}_B(t) = -\frac{1}{\tau_B} m_B(t) + \frac{1}{\tau_B} u_B(t) \tag{3.20b}$$

$$\Delta f(t) = \alpha_A m_A(t) + \alpha_B m_B(t) \tag{3.21}$$

where $m_A(t)$ and $m_B(t)$ represent the normalized concentrations of absorbed analyte A and analyte B respectively at time, t, α_A and α_B represent the normalized steady-state frequency shift of analyte A and analyte B respectively and $u_A(t)$ and $u_B(t)$ both represent the unit step input for analyte A and analyte B respectively. In conclusion, eq. 3.20 and eq. 3.21 are the two-analyte normalized equations that were used to represent the two-analyte absorption.

3.3.1 Discrete-Time Model

Similar to the single analyte system, the frequency shifts for the binary mixtures are also measured at discrete-time instants (i.e. $t = kT$, where T is the sampling period). Therefore, it is required to transform the continuous time model of the two-analyte sensor response given in eq. 3.20 and eq. 3.21 into a discrete-time model. Just as in the single analyte case, Euler's continuous time approximation formula was used and based on the Euler's formula, the first derivative of the normalized concentration of each analyte can be approximated by,

$$\dot{m}_A(t) = \frac{m_{A,k+1} - m_{A,k}}{T} \tag{3.22a}$$

$$\dot{m}_B(t) = \frac{m_{B,k+1} - m_{B,k}}{T} \tag{3.22b}$$

By using Euler's approximation as given in eq. 3.22, the discrete-time model of the two-analyte system is found to be as given,

$$m_{A,k+1} = (1 - S_A)m_{A,k} + S_A u_{A,k} + v_k \tag{3.23a}$$

$$m_{B,k+1} = (1 - S_B)m_{B,k} + S_B u_{B,k} + v_k \tag{3.23b}$$

$$\Delta f_k = \alpha_A m_{A,k} + \alpha_B m_{B,k} + w_k \tag{3.24}$$

where S_A and S_B are the absorption rate constant for analyte A and analyte B, respectively, and are defined as,

$$S_A = \frac{T}{\tau_A}$$

$$S_B = \frac{T}{\tau_B}$$

The terms v_k and w_k are added to represent the process and measurement noise that are present in the two-analyte system (v_k and w_k are uncorrelated white noise with zero mean). Eq. 3.23 and eq. 3.24 represent the discrete-time model of the two-analyte system.

3.3.2 Nonlinear Model

Two different state-space models for the two-analyte system were developed, one being a nonlinear model and the other a linear model. First, the formulation of the nonlinear model will be presented and discussed. As mentioned earlier, for the nonlinear model of the two-analyte system, the normalized concentration of each analyte and the steady-state frequency shift of each analyte are the unknown parameters that need to be estimated. Note that the absorption rate (or time constant) of each analyte does not have to be estimated because it is assumed to be known from the single analyte experiments. In order to obtain the state-space form of the nonlinear model, state variables are assigned to the unknown parameters that need to be estimated:

$$\begin{bmatrix} x_k^{(1)} \\ x_k^{(2)} \\ x_k^{(3)} \\ x_k^{(4)} \end{bmatrix} = \begin{bmatrix} m_{A,k} \\ m_{B,k} \\ \alpha_A \\ \alpha_B \end{bmatrix}$$

and define the output as,

$$y_k = \Delta f_k$$

By rewriting eq. 3.23 and eq. 3.24 in terms of these state variables, the following equations are obtained,

$$x_{k+1}^{(1)} = (1 - S_A)x_k^{(1)} + S_A u_{A,k} + v_k \tag{3.25a}$$

$$x_{k+1}^{(2)} = (1 - S_B)x_k^{(2)} + S_B u_{B,k} + v_k \tag{3.25b}$$

$$y_k = x_k^{(3)} x_k^{(1)} + x_k^{(4)} x_k^{(2)} + w_k \tag{3.26}$$

Since the steady-state frequency shift of analyte A, α_A, and steady-state frequency shift of analyte B, α_B, are a constant, the value of these parameters at time, $k + 1$ is equal to the value of the parameter at time, k.

$$x_{k+1}^{(3)} = x_k^{(3)} \tag{3.27}$$

$$x_{k+1}^{(4)} = x_k^{(4)} \tag{3.28}$$

By combining eq. 3.25, eq. 3.26, eq. 3.27, and eq. 3.28, the state-space form of the nonlinear model can be rewritten as,

$$x_{k+1} = f(x_k, u_k, v_k) = \begin{bmatrix} x_{k+1}^{(1)} \\ x_{k+1}^{(2)} \\ x_{k+1}^{(3)} \\ x_{k+1}^{(4)} \end{bmatrix} = \begin{bmatrix} (1 - S_A)x_k^{(1)} + S_A u_{A,k} + v_k \\ (1 - S_B)x_k^{(2)} + S_B u_{B,k} + v_k \\ x_k^{(3)} \\ x_k^{(4)} \end{bmatrix}$$

(3.29a)

$$y_k = h(x_k, u_k, w_k) = x_k^{(3)} x_k^{(1)} + x_k^{(4)} x_k^{(2)} + w_k$$

(3.29b)

From eq. 3.29, it can be seen that the state-space form is a nonlinear model; thus this formulation of the state-space form of the two-analyte system is known as the nonlinear model and EKF should be used to estimate the unknown parameters. In order to apply the EKF algorithm, the state-space form given in eq. 3.29 has to be linearized using Taylor series expansion. For the nonlinear model, the result of linearization is as follows,

$$x_{k+1} = f(x_k, u_k, v_k) \approx f(\hat{x}_k, u_k, \bar{v}_k) + A_k(x_k - \hat{x}_k) + F_k v_k$$

(3.30a)

$$y_k = h(x_k, u_k, w_k) \approx h(\hat{x}_k, u_k, \bar{w}_k) + C_k(x_k - \hat{x}_k) + G_k w_k$$

(3.30b)

where,

$$f(\hat{x}_k, u_k, \bar{v}_k) = \begin{bmatrix} (1 - S_A)\hat{x}_k^{(1)} + S_A u_{A,k} \\ (1 - S_B)\hat{x}_k^{(2)} + S_B u_{B,k} \\ \hat{x}_k^{(3)} \\ \hat{x}_k^{(4)} \end{bmatrix}$$

$$h(\hat{x}_k, u_k, \bar{w}_k) = \hat{x}_k^{(3)}\hat{x}_k^{(1)} + \hat{x}_k^{(4)}\hat{x}_k^{(2)}$$

$$A_k = \begin{bmatrix} (1-S_A) & 0 & 0 & 0 \\ 0 & (1-S_B) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$F_k = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$C_k = \begin{bmatrix} \hat{x}_k^{(3)} & \hat{x}_k^{(4)} & \hat{x}_k^{(1)} & \hat{x}_k^{(2)} \end{bmatrix}$$

$$G_k = [1]$$

By using the linearization result given in eq. 3.30, EKF algorithm can be applied to perform the estimation of the unknown parameters. The information on how to apply the EKF algorithm was discussed in chapter 2 and the EKF algorithm was summarized in Fig. 2.2.

3.3.3 Linear Model

In this section the formulation of the state-space form of the linear model of the two-analyte system will be presented. As mentioned earlier, for the linear model of the

two-analyte system, the unknown parameters that need to be estimated are only the steady-state frequency shifts of the analytes. The normalized concentration of each analyte does not have to be estimated because by using the known time constant of each analyte from the single analyte experiment and eq. 3.23, the normalized concentration for each analyte can be determined for every discrete-time instant. As a result, a simplified linear model can be obtained and the estimation of the unknown parameters can be performed using KF. In order to obtain the state-space form of the linear model, state variables are assigned to the unknown parameters that need to be estimated,

$$\begin{bmatrix} x_k^{(1)} \\ x_k^{(2)} \end{bmatrix} = \begin{bmatrix} \alpha_A \\ \alpha_B \end{bmatrix}$$

and define the output as,

$$y_k = \Delta f_k$$

Since the steady-state frequency shift of analyte A, α_A, and the steady-state frequency shift of analyte B, α_B, is a constant, the value of these parameters at time, $k+1$ is equal to the value of the parameter at time, k,

$$\begin{bmatrix} x_{k+1}^{(1)} \\ x_{k+2}^{(2)} \end{bmatrix} = \begin{bmatrix} x_k^{(1)} \\ x_k^{(2)} \end{bmatrix} \tag{3.31}$$

By rewriting eq. 3.24 and eq. 3.31 in terms of the state variables, the state-space form of the linear model can be obtained as follows,

$$\begin{bmatrix} x_{k+1}^{(1)} \\ x_{k+2}^{(2)} \end{bmatrix} = A \begin{bmatrix} x_k^{(1)} \\ x_k^{(2)} \end{bmatrix} \tag{3.32a}$$

$$y_k = C_k \begin{bmatrix} x_k^{(1)} \\ x_k^{(2)} \end{bmatrix} + Gw_k \tag{3.32b}$$

where,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$C_k = [m_{A,k} \quad m_{B,k}]$$

$$G = [1]$$

represent the system matrices.

Eq. 3.32 represents the state-space form of the linear model of the two-analyte system and from eq. 3.32, it can be seen that the state-space form is a linear model; thus

this formulation of the state-space form of the two-analyte system is known as the linear model. In this case, to estimate the unknown parameters, KF can be used. It should be noted that the time-varying system matrix, C_k is dependent on the values of the normalized concentration for each analyte (i.e. analyte A and analyte B) which can be determined for each discrete-time instant, k by using eq. 3.23. In conclusion, by using the system matrices and the state-space form of the linear model given in eq. 3.32, KF algorithm can be applied to perform the estimation of the unknown parameters. The information on how to apply the KF algorithm was discussed in chapter 2 and the KF algorithm was summarized in Fig. 2.1.

4. CHEMICAL SENSOR DATA ACQUISITION

4.1 Introduction

In this chapter, the specifics of the shear horizontal surface acoustic wave (SH-SAW) device that was used to collect the sensor data analyzed in the present work are discussed. The physics of the SH-SAW devices are discussed in detail. The SH-SAW sensor platform has been chosen to collect the data analyzed in the present work because it has been shown in [7, 34-36], that SH-SAW has the potential of being used as an in-situ chemical sensor to quantify BTEX compounds. All the preparations of the experimental setup and data collection were performed in the Microsensor Research Laboratory, Marquette University. The details of the experimental setup and data collection will also be discussed in this chapter. The information on the types of the polymers used in the present work to detect the target analytes is also given. The polymer films were coated on the sensor platform to enhance the sensitivity and to provide partial selectivity for the target analytes.

Moreover, data pre-processing techniques that are used in the present work are also discussed in this chapter. The data pre-processing discussed includes linear baseline drift correction and the elimination of outlier points in the sensor data. The linear baseline drift correction technique discussed in this chapter uses estimation theory (particularly Kalman Filter) to rapidly perform linear extrapolation and linear interpolation. The techniques discussed were tested on the actual experimental data and the results obtained

are shown. The elimination of the outlier points in the sensor data were performed by using a combination of discrete low pass filter and Kalman Filter (KF) (or Extended Kalman Filter (EKF)). The proposed technique was tested on the experimental data with outlier points to illustrate the effectiveness of the proposed technique and the results obtained are shown. Note that the data pre-processing had to be performed first before the data is used to perform the estimation of the analyte(s) in the sample.

4.2 Shear Horizontal Surface Acoustic Wave (SH-SAW) Devices

Shear horizontal surface acoustic wave (SH-SAW) devices are a specific type of sensor platform which can be used for biochemical detection in liquid environments. Compared to other types of acoustic wave devices which can be used for liquid-phase sensing applications (such as thickness shear mode, shear horizontal acoustic plate mode, and flexural plate wave), SH-SAW are preferred because surface waves are more sensitive to surface perturbations. Moreover, devices with high quality factor can be obtained. SH-SAW devices are also small, robust, and easy to incorporate into on-line low cost systems [36, 37]. SH-SAW devices are fabricated with a specific crystal structure that is rotated so that the wave in this crystal only supports the shear horizontal component of the surface wave [38]. Therefore, SH-SAW can be used effectively for liquid-phase sensing applications.

Compared to surface acoustic wave (SAW), also known as Rayleigh wave, the SH-SAW often propagates slightly deeper within the substrate [39, 40], hence preventing

the implementation of high-sensitivity detectors. However, the sensitivity of SH-SAW sensors can be increased by using a thin guiding layer on the device surface which will trap the acoustic energy near the sensing surface. Therefore, SH-SAW sensors are typically modeled as a multilayered structure [41]. The structure used in this investigation is a three-layer structure as shown in Fig. 4.1. This three-layer structure consists of the piezoelectric substrate with input and output interdigital transducers (IDTs) arranged in a delay line configuration, a polymer layer and a liquid layer. The purpose of the piezoelectric substrate is to convert the electrical signal into a mechanical signal (strain) (i.e. the acoustic wave) and also to serve as a support for the entire device. The polymer layer has a finite thickness, h and is assumed to have a lower shear wave velocity than the substrate (the latter is a precondition for the confinement of the SH-SAW to the surface). For the three-layer structure, the polymer layer serves as both a wave guiding layer and a chemically sensitive layer [42-44]. The liquid layer is assumed to be a Newtonian fluid because the solution(s) being tested are dilute aqueous solutions, and is used for transport of analyte molecules. Since the polymer layer is of thickness, h, the polymer layer is considered as a finite layer while the substrate and the liquid layer are considered as semi-infinite layers [42]. Note that, since the guided SH-SAW propagates in the x_1 direction and the particle displacement is parallel to the x_2 direction, the device is known as ‘shear-horizontal surface acoustic wave device’.

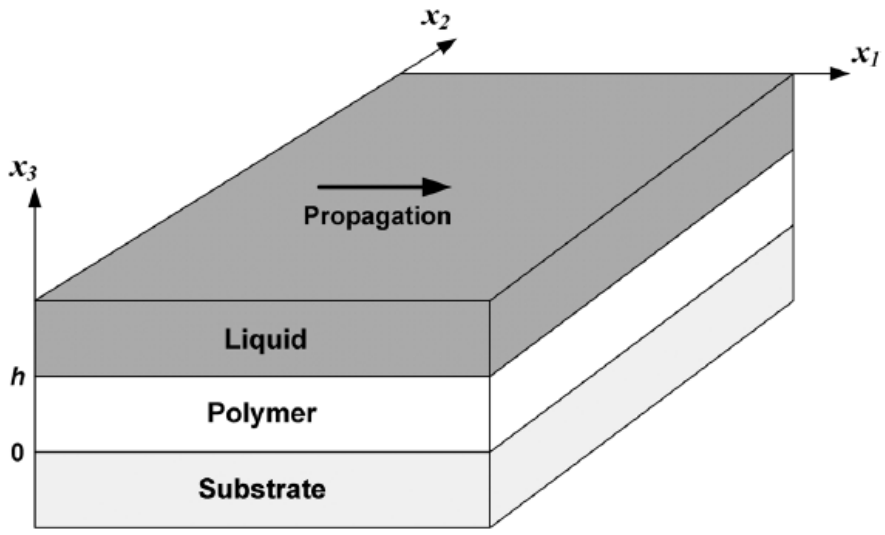

Figure 4.1: Three-layer structure and coordinate system. The guided SH-SAW will propagate in the $x_1$direction, x_2 is in the direction of the acoustic wave particle displacement, and x_3 is normal to the sensing surface [43].

In order to eliminate the common environmental interactions (such as temperature and pressure), a dual delay line configuration is used for the SH-SAW sensor design where one line serves as a sensing line and the other as a reference line. This design allows for the common environmental interactions producing the responses from both lines to be eliminated by subtraction (i.e. differential measurement) [37, 43]. A thin metal layer is also used between the two IDTs (input and output IDTs), to create an electrical short so that acoustoelectric interactions with the load can be eliminated [37, 43]. This is done in order to eliminate all electrical load interactions, so that only mechanical loading is present. Therefore, only sensing caused by mechanical loading (i.e. changes in the mechanical properties of the polymer coating) is considered in this work.

4.3 Data Acquisition

The sensor data analyzed in this work was collected using the 36° YX-$LiTaO_3$ guided SH-SAW device as the sensing platform [34-36]. This device was fabricated with 10/80-nm-thick Cr/Au split finger pairs IDTs having the periodicity of 40µm, which will produce an operating frequency of 103MHz for the uncoated device [34-36]. As mentioned earlier, dual delay line configuration was used with a metalized path between the IDTs to eliminate the acoustoelectric interaction with the load. The sensing line was coated with the sorbent polymer coatings. The types of polymer coatings which are used to interact with the analytes of interest and to collect the sensor data include poly(ethyl acrylate) (PEA), poly(epichlorohydrin) (PECH), and poly(isobutylene) (PIB), all purchased from Sigma-Aldrich. The polymers were deposited on the sensing line from a solution by spin coating and baking for 15 minutes at 60°C which results in thicknesses, 1.0 µm for PEA, 0.6 µm for PECH, and 0.8 µm for PIB. The reference lines were coated with poly(methyl methacrylate) and baked for 120 minutes at 180°C, which will result in a glassy, non-sorbent coating so that the reference line will not absorb any analyte (i.e. it is chemically insensitive). All the BTEX analytes used in the experiment were purchased from Sigma-Aldrich and had purities of at least 98.5%.

The experimental set-up used to collect the sensor data consisted of a network analyzer (Agilent 8753ES) and a switch/control system (Agilent 3499A) to switch between the two SH-SAW delay lines. Note that for some measurements, an Agilent E5061B network analyzer and an Agilent 34980A switch/control system were used. In

order to perform the experiment, the SH-SAW sensor was placed inside a flow cell (which was designed at Marquette University Microsensors Laboratory) and a pump (Eppendorf EVA, in later experiments: Ismatec Reglo Digital MS) was used to pump the solutions into the flow cell. The solutions were pumped at sample flow rate of 0.4 ml/min. Before pumping the analyte solution into the flow cell for detection by the sensor, a reference solution (DI water) was pumped first. The reference solution was pumped until the output signal was stable. When the output was stable, an analyte sample was pumped into the flow cell for detection by the sensor. After the sensor response reaches the steady-state (or equilibrium), the reference solution was pumped again into the system to flush the flow cell and cause the analyte to desorb from the polymer coating on the sensor. This process was repeated periodically for different analyte samples and concentrations. The above procedures are well described in the literature [34-36]. The experiment was performed in an environment in which the temperature was held constant at 22.0 ± 0.1°C. The experiments were performed using the samples containing single analyte and also using the samples containing binary mixtures of analytes. The measurements were performed on the single analyte samples to determine the sensitivities, σ (in Hz/ppm) and the response time constant, τ (in s) for each coating/analyte combination considered in this work. By using the values of the sensitivity, σ, of each analyte from the single analyte measurements, the analyte concentration(s) can readily be extracted by dividing steady-state frequency shift, α, by the average value of sensitivity, σ [35]. The values of response time constant, τ, determined from the single analyte experiments can be used in the estimation process of the two-analyte system. Multiple single analyte measurements were performed and both

the values for the sensitivity, σ, and time constant, τ, are determined and are listed as average values for various coating/analyte combinations in tables 4.1 and 4.2, respectively.

Polymer	$\sigma_{benzene}$	$\sigma_{toluene}$	$\sigma_{ethylbenzene}$
1.0 µm PEA	244 ($\pm$27)	690 ($\pm$160)	2240 ($\pm$460)
0.6 µm PECH	109 ($\pm$9)	435 ($\pm$25)	1450 ($\pm$240)
0.8 µm PIB	63 ($\pm$5)	344 ($\pm$43)	1670 ($\pm$10)

Table 4.1: Measured average sensitivities, σ (in Hz/ppm) from multiple single analyte experiment for three different polymer coatings to various BTEX analytes. The standard errors (68% confidence interval) are given in parentheses [35].

Polymer	$\tau_{benzene}$	$\tau_{toluene}$	$\tau_{ethylbenzene}$
1.0 µm PEA	36.1 ($\pm$10.0)	76.7 ($\pm$6.0)	204 ($\pm$4.5)
0.6 µm PECH	26.5 ($\pm$8.4)	77.6 ($\pm$2.8)	175 ($\pm$13)
0.8 µm PIB	29.3 ($\pm$7.8)	84.2 ($\pm$6.5)	245 ($\pm$14)

Table 4.2: Measured average response times, τ (in s) from multiple single analyte experiment for three different polymer coatings to various BTEX analytes. The standard errors (68% confidence interval) are given in parentheses [35].

All the sensor data collected were recorded and the data collected exhibit linear baseline drift during the response. Therefore, pre-processing had to be done first in order to correct the data for baseline drift and to eliminate any outlier points in the data before using the data to perform the estimation process to quantify the analytes in the binary mixture (or single analyte sample).

4.4 Data Processing

As mentioned in the previous section, the raw data collected from the experiment exhibits a linear baseline drift. Therefore, the raw data has to be corrected for baseline drift before the data can be used to perform the estimation process to quantify the analytes in the binary mixture. The general approach to perform baseline drift correction is given in [17]. The baseline drift correction approach presented in [17] is based on Extended Kalman Filter (EKF) and works for both linear and nonlinear baseline drift. Since the sensor responds rapidly to the target analyte and the baseline drift observed in the data collected for the analysis in the present work exhibit a linear baseline drift, a simplified model of baseline drift correction method is proposed in this section. The linear baseline drift correction technique discussed in this section uses Kalman Filter (KF) to perform linear extrapolation and linear interpolation rapidly and can be viewed as a special case of the baseline drift correction technique presented in [17]. One should note that, the baseline correction technique presented in this section will only work for linear baseline drift and for sensors which responds rapidly to the target analyte. The proposed techniques were tested on the experimental data with linear drift and the results obtained are shown.

Occasionally, outlier points will be observed in the measurement data and these outlier points have to be eliminated or corrected in order to get an accurate result when one performs the estimation process to quantify the analyte(s) present in the sample. Outlier points will be recorded if the measurement noise is very high, sometimes during

the start of a new measurement, or if any changes in the boundary conditions occur, e.g. if the pump is briefly stopped when switching to a new sample. In this section, a new technique which is based on a combination of discrete low pass filter and EKF will be presented in order to eliminate the outlier points in the measurement data rapidly (in real-time). One should note that, the correction of outlier points can only be performed after the data point has been corrected for baseline drift. The proposed technique was tested on the experimental data with outlier points to illustrate the effectiveness of the proposed technique and the results obtained are shown.

4.4.1 Baseline Drift Correction

As mentioned earlier, linear baseline drift correction technique discussed in this section uses Kalman Filter (KF) to perform linear extrapolation or linear interpolation rapidly. Techniques to perform linear extrapolation using KF are explained first.

In order to perform linear extrapolation using KF, only the data obtained before the sensor is exposed to the analyte are used. Since the baseline drifts linearly in the present case, the baseline drift can be modeled as a first-order curve given by,

$$y\_baseline_k = a + bk + w_k \tag{4.1}$$

where a represent the y-intercept, b is the slope of the baseline, w_k represent the measurement noise and k represent the discrete time instant at which the baseline is measured. By using the measurement data recorded before the analyte is introduced to the sensor, one could actually estimate the constants a and b. Then by using the estimated value of the constants a and b, one could extrapolate the baseline during the sensor response (i.e. after the analyte has been introduced to the sensor) and perform the baseline drift correction by subtracting the baseline drift from the sensor response while measurements are taken. In order to estimate the constants a and b in real-time by using KF, the baseline drift model given in eq. 4.1 was transformed into the state-space model by assigning a state-variable to the parameters a and b,

$$\begin{bmatrix} x_k^{(1)} \\ x_k^{(2)} \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \tag{4.2}$$

Since the parameters that need to be estimated are constants, the values of the parameters cannot be changing in time. This can be represented by,

$$\begin{bmatrix} x_{k+1}^{(1)} \\ x_{k+2}^{(2)} \end{bmatrix} = \begin{bmatrix} x_k^{(1)} \\ x_k^{(2)} \end{bmatrix} \tag{4.3}$$

Then, the linear baseline drift model in the state-space form can be written as follows:

$$\begin{bmatrix} x_{k+1}^{(1)} \\ x_{k+2}^{(2)} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_k^{(1)} \\ x_k^{(2)} \end{bmatrix}$$

(4.4a)

$$y_k = \begin{bmatrix} 1 & k \end{bmatrix} \begin{bmatrix} x_k^{(1)} \\ x_k^{(2)} \end{bmatrix} + w_k$$

(4.4b)

By using the state-space model of the baseline drift and the measurement data before analyte exposure, KF algorithm as presented in chapter 2 can be used to estimate the constants a and b. Once the values of constants a and b are estimated, the baseline can then be extrapolated during the sensor response (i.e. after the analyte is introduced) in order to correct the measurement data for baseline drift while measurements are taken. The baseline corrected measurement can be found by subtracting the baseline value at a particular instant in time from the recorded measurement data at that same instant in time,

$$y\_corrected_k = y\_measurement_k - y\_baseline_k$$

(4.5)

Since the linear extrapolation only requires the data obtained before the analyte exposure to estimate the baseline, linear extrapolation can be performed in real-time.

If several samples are measured consecutively in the course of one experiment, linear interpolation using KF can be used to estimate the baseline to obtain a more accurate estimate. In order to perform linear interpolation using KF, data obtained both

before the analyte is added and after it has been flushed from the sensor are used. Both these data are used to determine the constant b which represents the slope of the baseline. The constant, a, which represents the y-intercept of the baseline is determined by using the data obtained before the analyte is added. Basically, one can use the same state-space model given in eq. 4.4 to perform linear interpolation. However, for linear interpolation, the linear baseline drift needs to be estimated twice, one by using the data obtained before the analyte exposure and another baseline estimation using the data obtained after the analyte has been flushed from the sensor. Then, the slope of the baseline, b, is determined by taking the average between the slopes of the two baselines and as for the y-intercept, a, one can assign the same value obtained for the y-intercept of the baseline estimated using the data obtained before analyte exposure. Once the constants a and b are determined, the baseline during the sensor response can be determined and subtracted from the measurement data to obtained the corrected measurement data, as illustrated by eq. 4.5. Since linear interpolation requires data obtained both before the analyte is added and after it has been flushed from the sensor to estimate the baseline, linear interpolation cannot be performed in real-time as the measurements are taken. This is because one has to wait until the data after the analyte has been flushed from the sensor are collected before estimating the baseline. However, linear interpolation can yield a more accurate result compared to linear extrapolation [17].

In order to prove the validity of the proposed linear baseline drift correction techniques explained in this section, the techniques were tested on the actual experimental data collected in the Microsensor Research Laboratory at Marquette

University. Here, the baseline drift correction results obtained by using data from two different experiments will be presented and discussed. First, the results obtained by using the raw experimental data collected by using PEA polymer coatings with thickness 1.0μm as shown in Fig. 4.2 are presented. As can be seen in Fig. 4.2, there is a baseline drift in the experimental data and the baseline drifts linearly for each individual sample. Therefore, the proposed baseline drift corrections techniques can be used. The result obtained by performing linear extrapolation is shown in Fig 4.3 and the result obtained by performing linear interpolation is shown in Fig 4.4.

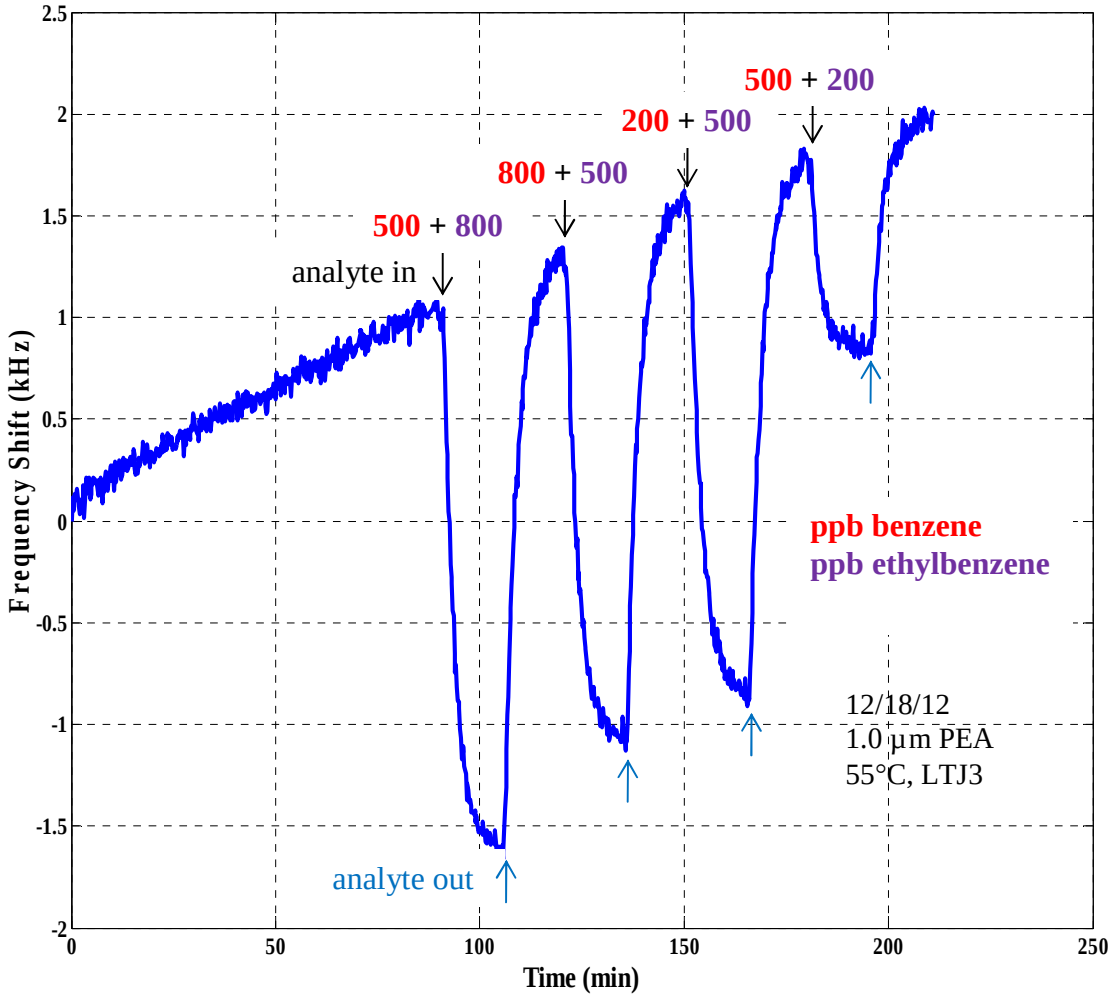

Figure 4.2: Raw experimental data with linear baseline drift. The experimental data shows the sensor response of a series of four different samples which are mixtures of benzene and ethylbenzene at different concentrations as specified in the figure above. In the figure, ppb stands for parts per billion (µg/L).

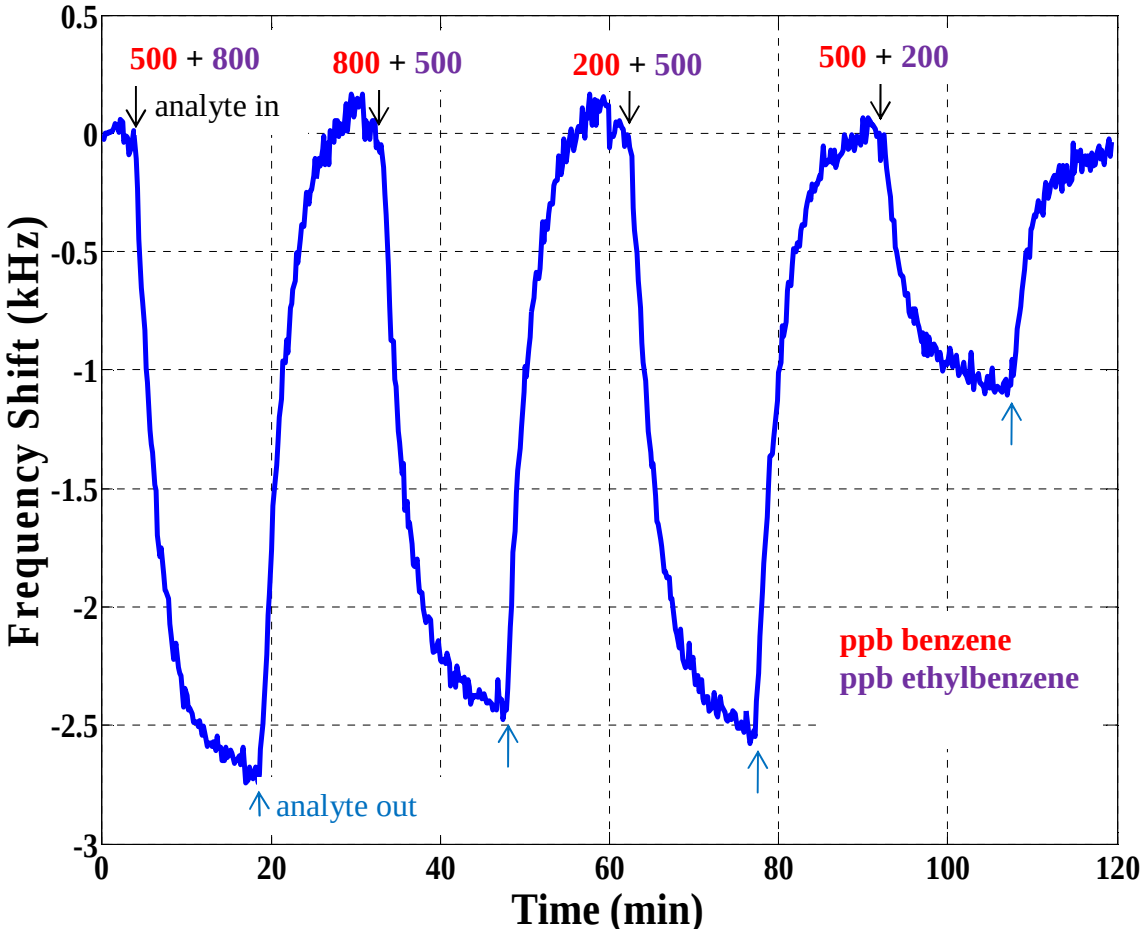

Figure 4.3: Baseline corrected result obtained by performing linear extrapolation using Kalman Filter on the raw experimental data shown in Fig. 4.2.

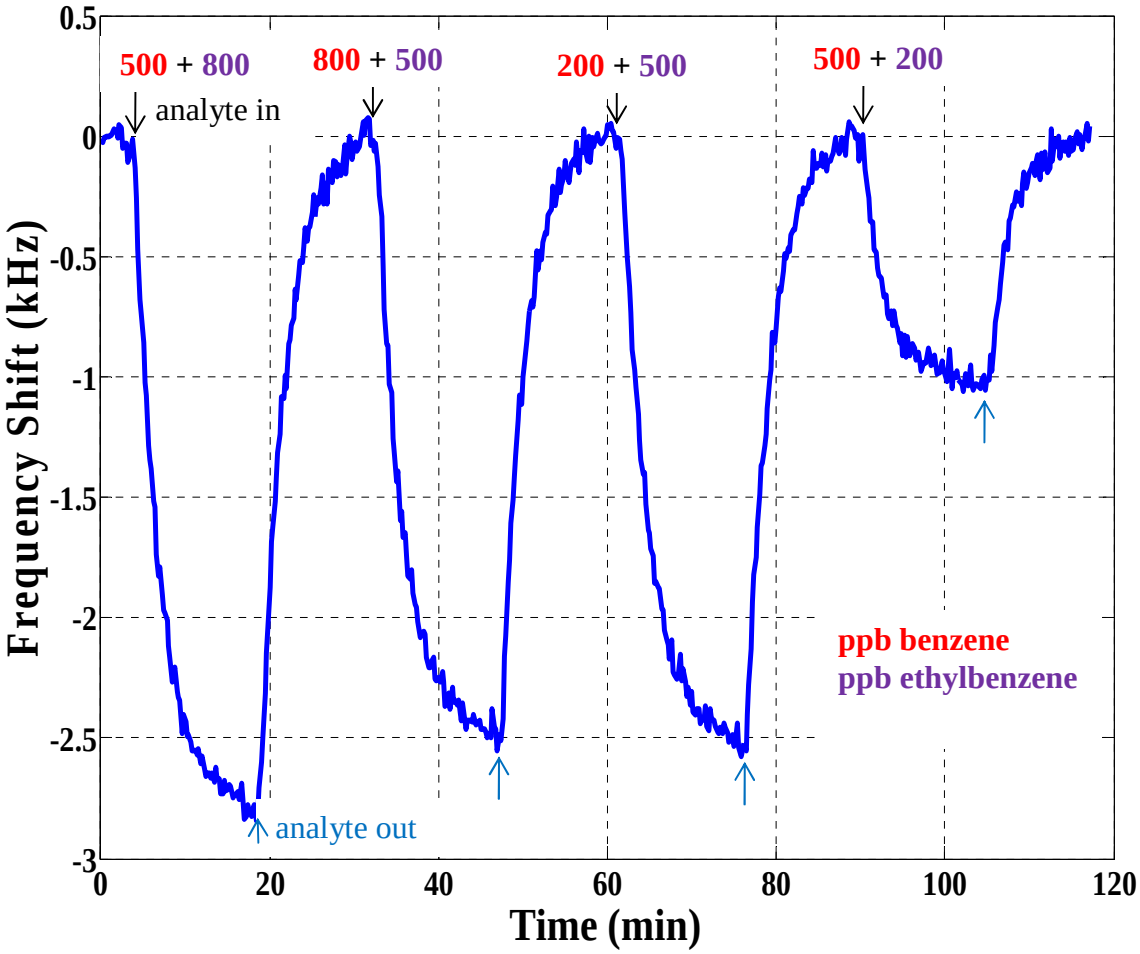

Figure 4.4: Baseline corrected result obtained by performing linear interpolation using Kalman Filter on the raw experimental data shown in Fig. 4.2.

As can be seen in Fig. 4.3 and Fig. 4.4, the raw experimental data from Fig. 4.2 has been corrected for the baseline drift by using the baseline drift correction techniques presented in this section. It should be noted that, as expected, the result obtained by

performing linear interpolation as shown in Fig. 4.4 is more accurate compared to the result obtained by performing linear extrapolation as shown in Fig. 4.3.

Fig. 4.5 shows another example of raw experimental data which were collected by using PECH polymer coatings with thickness 0.6µm. As can be seen in Fig. 4.5, the experiment consists of a series of four samples, where the first three samples consist of binary mixtures of benzene and toluene and the fourth sample just consists of benzene, which are measured consecutively in the course of one experiment. All the samples have different concentrations as indicated in Fig. 4.5. From Fig. 4.5, it can be noticed that there is a baseline drift in the experimental data and the baseline drifts linearly for each individual sample. Therefore, the proposed baseline drift correction techniques can be used. The result obtained by performing linear extrapolation is shown in Fig 4.6 and the result obtained by performing linear interpolation is shown in Fig 4.7.

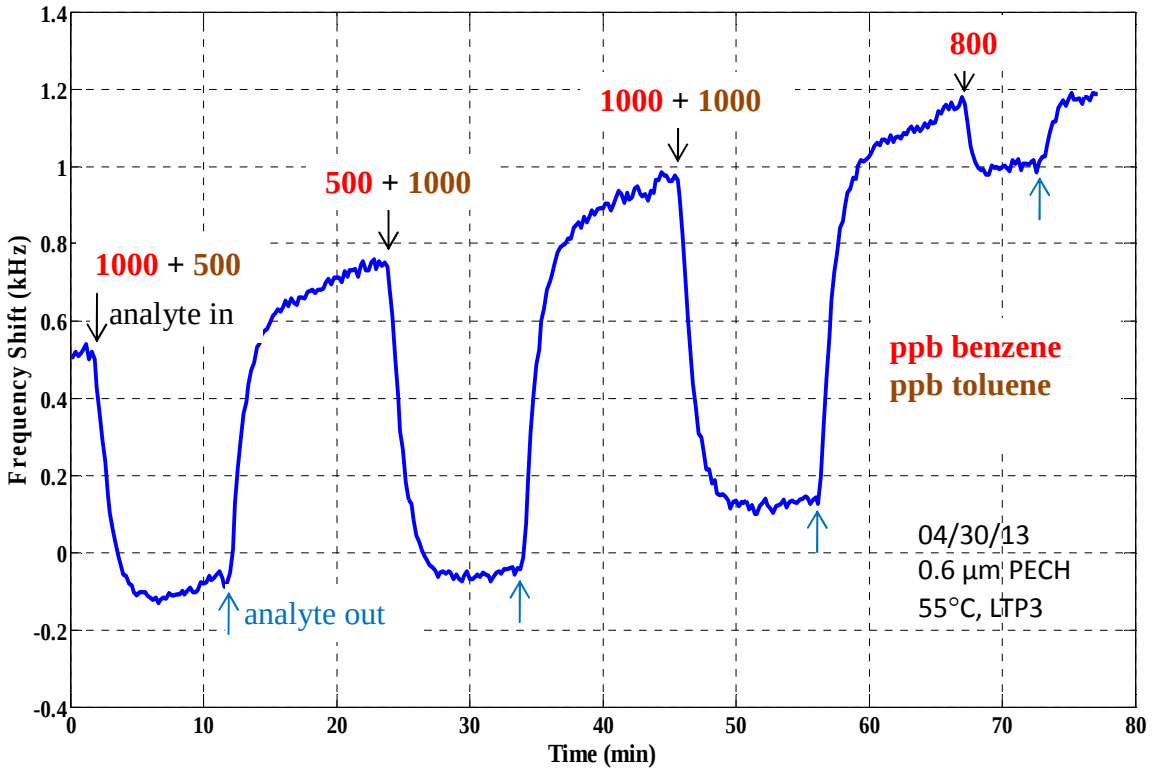

Figure 4.5: Raw experimental data with linear baseline drift. The experimental data shows the sensor response of a series of four different samples (the first three samples are binary mixtures of benzene and toluene and the fourth sample is a single analyte sample of benzene) at different concentrations as specified in the figure above.

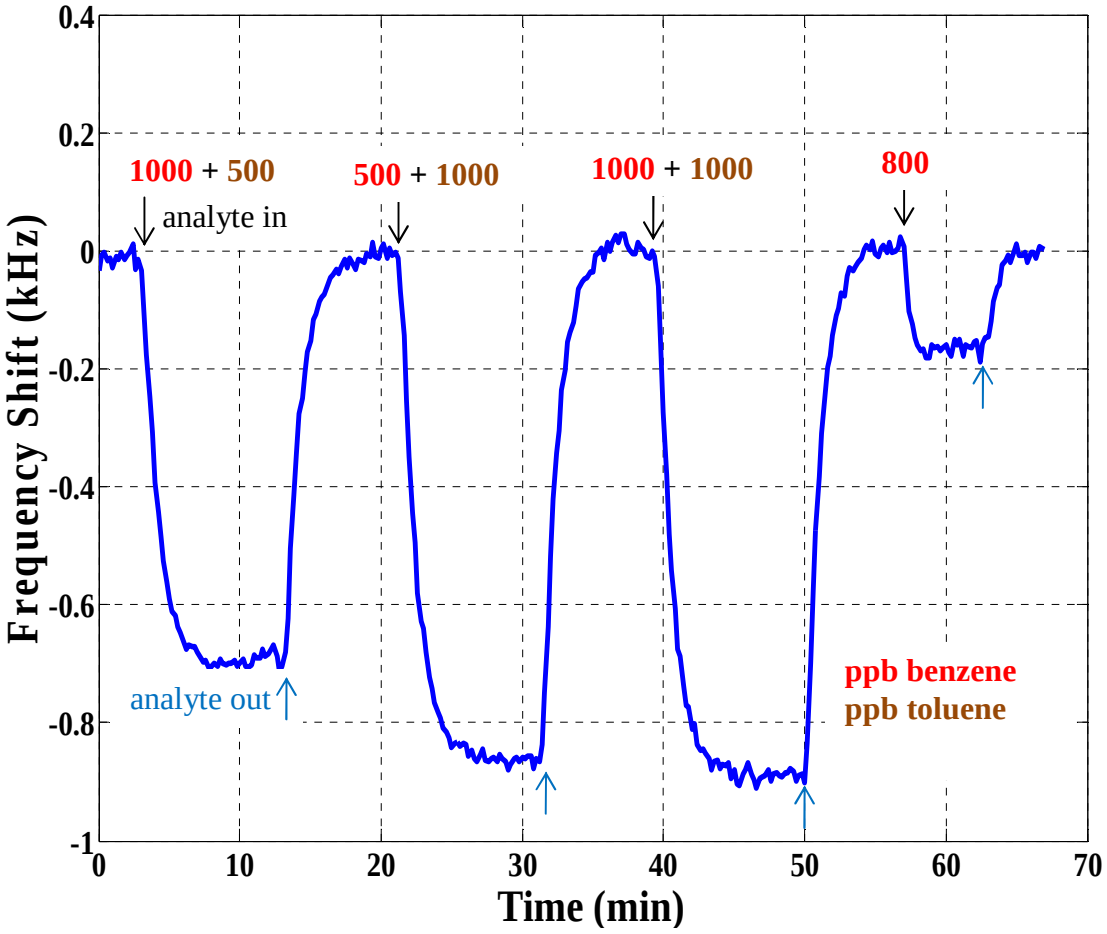

Figure 4.6: Baseline corrected result obtained by performing linear extrapolation using Kalman Filter on the raw experimental data shown in Fig. 4.5.

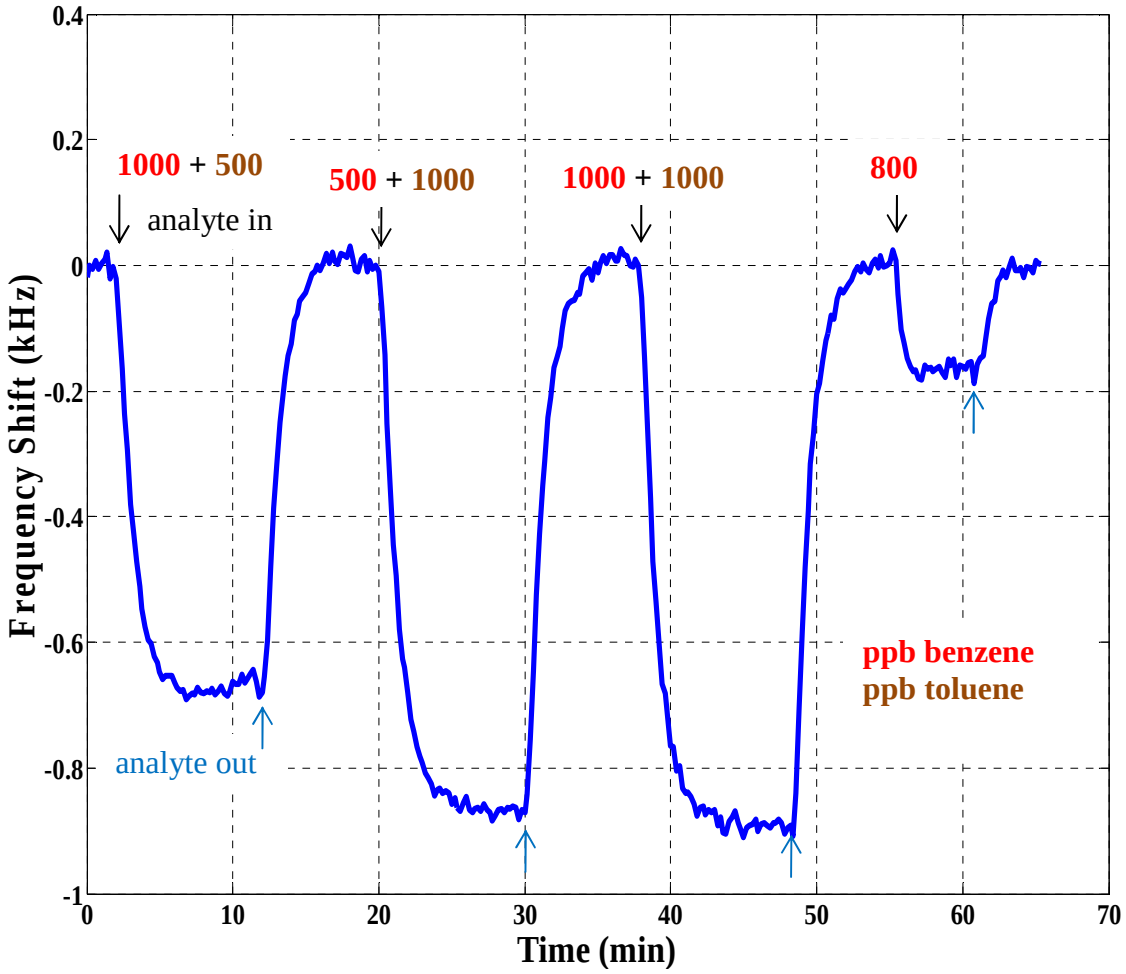

Figure 4.7: Baseline corrected result obtained by performing linear interpolation using Kalman Filter on the raw experimental data shown in Fig. 4.5.

As can be seen in Fig. 4.6 and Fig. 4.7, the raw experimental data from Fig. 4.5 has been corrected for the baseline drift. Based on the baseline corrections results obtained by using the techniques presented in this section, it can be concluded that the proposed baseline correction techniques are indeed capable of correcting for linear

baseline drift. Also, based on the results obtained it can be clearly seen that the result obtained by performing linear interpolation is more accurate compared to the result obtained by performing linear extrapolation. However, for use in the real-world application, linear extrapolation using KF can be used to correct for the baseline drift, so that the drift correction can be performed in real-time. Note that real groundwater samples might contain many compounds, some of which might show very long response times; in that case, it might not be practical to wait for the signal to return to baseline. Also in real-world applications, the sensor system may be installed in a fixed location, thus in the course of one experiment, only one sample will be tested. In that case, linear extrapolation can be used and can yield a sufficiently accurate result. Therefore, depending on the application, linear interpolation or linear extrapolation using KF can be used to correct for the baseline drift. In conclusion, the proposed baseline correction techniques presented in this section can be used to correct for linear baseline drift of the measurement data.

4.4.2 Correction of Outlier Points in Sensor Data

As stated earlier, outlier points in the experimental data are observed due to high measurement noise, sometimes when the sensor is exposed to the analyte(s) and also due to irregular changes in the boundary conditions. The outlier points can be eliminated or corrected in real-time by using a combination of a simple first-order discrete low-pass filter and KF (or EKF).

In order to design a simple discrete first-order low pass filter, the transfer function of a low-pass filter has to be discretized [45]. Consider the transfer function of a first-order low-pass filter as given in eq. 4.6,

$$H(s) = \frac{Y(s)}{U(s)} = \frac{(1/\tau)}{s + (1/\tau)} \tag{4.6}$$

where, τ represent the time constant of the filter, therefore, $1/\tau$ is the cut-off frequency of the filter, $Y(s)$ represent the output in s-domain and $U(s)$ represent the input in s-domain. The transfer function of the low-pass filter (given in eq. 4.6), can be rewritten as

$$\tau s Y(s) + Y(s) = U(s) \tag{4.7}$$

By taking the inverse Laplace transform of both sides of eq. 4.7, the following differential equation can be obtained:

$$\tau \dot{y}(t) + y(t) = u(t) \tag{4.8}$$

where $y(t)$ represent the output in time-domain and $u(t)$ represent the input in time-domain. In order to discretize eq. 4.8, Euler's Backward Differentiation method was

used. Based on the Euler's Backward Differentiation formula, the first derivative of the output, $y(t)$ can be approximated by,

$$\dot{y}(t) \approx \frac{y_k - y_{k-1}}{T} \tag{4.9}$$

where T is the sampling period. By applying the Euler's approximation as given in eq. 4.9, the discrete-time model of the low-pass filter is found to be

$$y_k = \left(\frac{T}{\tau + T}\right) u_k + \left(\frac{\tau}{\tau + T}\right) y_{k-1} \tag{4.10}$$

By defining a new variable, α, as

$$\alpha = \frac{T}{\tau + T} \tag{4.11}$$

eq. 4.10 can be rewritten as,

$$y_k = \alpha u_k + (1 - \alpha) y_{k-1} \tag{4.12}$$

where the variable, α is known as the filter parameter because of its dependence on the time constant, τ. Eq. 4.12 denotes a recursive relation between the output, y_k and the input, u_k. Therefore, by knowing the previous output value, y_{k-1} and the current input value, u_k, the current output value, y_k can be calculated. It should be noted that the input, u_k represent the data point that need to be filtered and the time constant of the filter, τ should be set to be equal to the time constant of the system that need to be filtered. Thus, for the single analyte system, time constant of the filter, τ should be set to be equal to the value of the time constant of the analyte response. For the two-analyte system, there will be two time constants which correspond to each analyte and one of the analyte will have a higher time constant than the other. Since the inverse of time constant can be related to the frequency of the analyte, one of the analyte will have a smaller frequency than the other. Specifically, the analyte with the higher time constant will have the smaller frequency. It is known that the cut-off frequency of the two-analyte system should be smaller than the smallest frequency (i.e. highest time constant) of the analyte responses. Therefore, as an approximation, for the two-analyte system, time constant of the filter, τ can be set to be equal to the time constant of the analyte with the highest time constant value. Note that, the process of finding the time constant of the two-analyte system is cumbersome; thus, the approximation that the time constant of the two-analyte system is equal to the time constant of the analyte with the highest time constant value can be used.

In order to implement the correction of outlier points in real-time, the discrete low-pass filter should be implemented together with the KF or EKF algorithm using the state-space model presented in chapter 3. The choice of using KF or EKF is dependent on

the state-space model used as discussed in chapter 3. KF (or EKF) will be used as a one-step ahead predictor to predict the next measurement data point. If the difference between the predicted measurement value and the actual measurement value is above a certain threshold set by the user (e.g. 0.01), the actual measurement point will be selected to be filtered by using the discrete low-pass filter. If the difference between the predicted measurement value and the actual measurement value is within the threshold set by the user, the actual measurement point will not be filtered using the discrete low-pass filter. By using this method only the actual outlier points in the measurement will be filtered. Therefore, outlier points can be eliminated and a better estimate of the sensor parameters can be rapidly obtained in real-time.

The proposed outlier point's correction technique was tested on the experimental data with outlier points to prove the feasibility of the proposed technique to correct the outlier points in the measurement data. The experimental data used to test the proposed technique to correct the outlier points were collected in the Microsensor Research Laboratory, at Marquette University. These data do not contain any outlier points. Therefore, in order to test the proposed technique, outlier points were introduced manually in the measurement data. It should be noted that the proposed technique was tested on multiple measurement data; however, in this section only a sample result will be presented. Additional outlier points correction results are presented in the Appendix A. Fig. 4.8 shows the result obtained after the correction of the outlier points co-plotted together with the measurement data with outlier points. Also shown in the figure is the fitted curve and fitted data expression for both data with outlier points and filtered data.

Note that both unfiltered data and filtered data were fitted by using dual-exponential fit and the fitting expression for both unfiltered data and filtered data is shown explicitly in Fig. 4.8. The measurement data used to test the proposed technique were obtained for a binary mixture sample which contains benzene and ethylbenzene with the concentration of 500 ppb and 800 ppb respectively. The data were collected for experiments performed using PEA polymer coatings with thickness 1.0μm and the raw data obtained had already been corrected for the baseline drift.

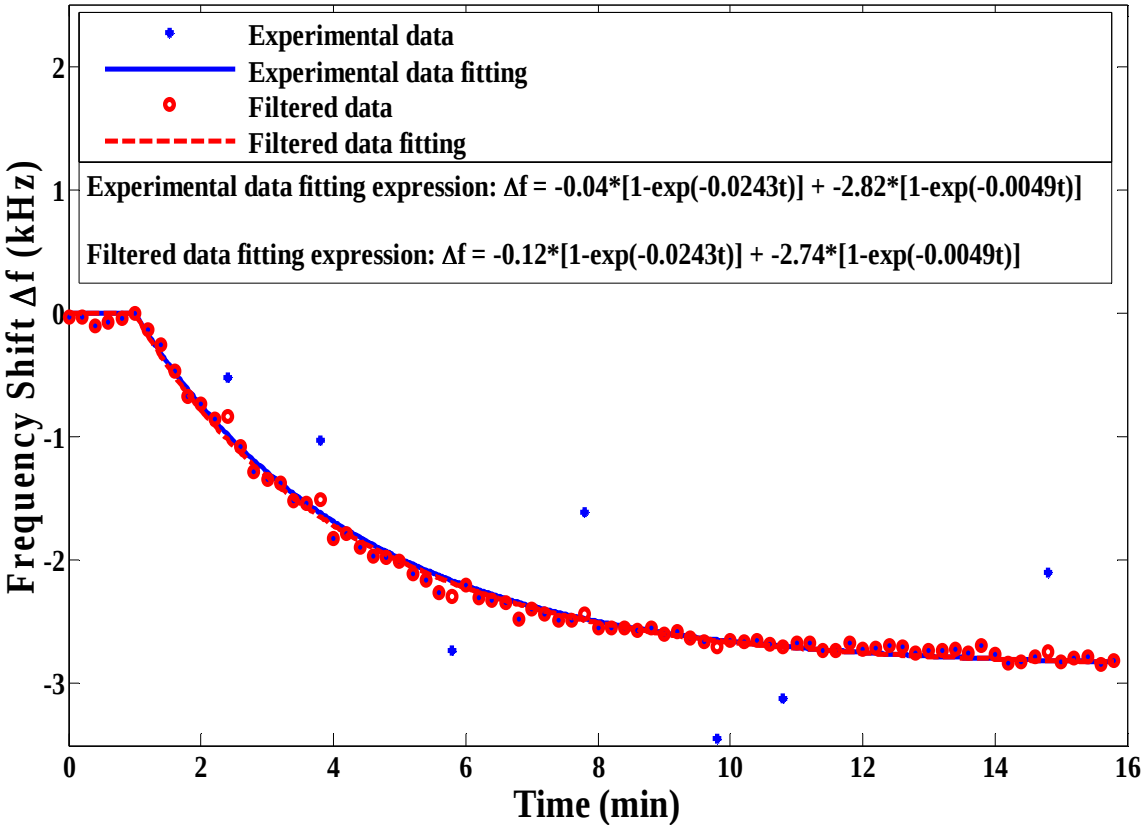

Figure 4.8: Outlier points corrected data co-plotted together with the measurement data with outlier points. The data are shown in two different colors where blue represents the measurement data with outlier points (unfiltered data) and red represents the outlier points corrected measurement data (filtered data). Both the data points and the curve fit for the data points are shown in the figure above. Also shown in the figure is the fitted data expression for both data with outlier points and filtered data.

As can be seen in Fig. 4.8, the measurement data (in blue) contains some outlier points and in the outlier points corrected data (in red) the outlier points have been filtered. Based on the result shown in Fig. 4.8, it can be observed that the proposed technique in this section is capable of eliminating any outlier points in the measurement data. Moreover, the concentrations of the analytes determined by using the fitted parameters of the filtered data are closer to the actual concentrations of the analytes compared to the concentrations of the analytes determined by using the fitted parameters of the unfiltered data as shown in Table 4.3. Note that the concentrations of the analytes were extracted by using the parameters obtained by fitting the data with dual-exponential fit and average values of the sensitivities of the single analytes given in Table 4.1. Also shown in Table 4.3 are the percentage differences between the estimated and the actual concentrations of the analytes.

Analytes	**Nominal Concentration (ppb)**	**Estimated Concentration using Unfiltered Data (ppb)** **(% difference with nominal concentration)**	**Estimated Concentration using Filtered Data (ppb)** **(% difference with nominal concentration)**
Benzene	**500**	**164 (67.2 %)**	**492 (1.6 %)**
Ethylbenzene	**800**	**1259 (57.4 %)**	**1223 (53 %)**

Table 4.3: Nominal concentration of the analytes, estimated concentration of the analytes using unfiltered data and estimated concentration of the analytes using filtered data obtained by using the measurement data for a binary mixture of benzene and ethylbenzene. Also shown in the table are the percentage differences between the estimated concentrations and nominal concentrations of the analytes.

As can be seen from Table 4.3, the estimated concentrations of the analytes determined from the unfiltered data fit do not agree well with the actual concentrations of benzene and ethylbenzene with percentage difference of 67.2% for benzene and 57.4% for ethylbenzene. For the outlier points corrected measurement data (filtered data), the concentrations of benzene and ethylbenzene are found to be closer to the actual concentration of benzene and ethylbenzene with percentage difference of 1.6% for benzene and 53% for ethylbenzene. Therefore, the concentrations of the analytes determined from the filtered data fit are much closer to the actual values of the concentrations of benzene and ethylbenzene compared to those from the unfiltered data. It should be noted that the discrepancies between estimated and actual concentrations of the analytes are to be expected due to the error introduced by the inaccuracy of the manual sample mixing procedure [35]. Based on the result obtained, it can be concluded that the outlier points corrected data could yield a more accurate estimate of the analyte concentrations compared to the measurement data with outlier points. This emphasizes the importance of eliminating outlier points in the measurement data.

Since the corrections of outlier points were performed by using a combination of discrete low-pass filter and KF (or EKF), the correction process can be performed in real-time as the measurements are being recorded. In conclusion, the proposed outlier points correction techniques presented in this section can be used to eliminate any outlier points in the measurement data in real-time and consequently, could yield a more accurate estimation results.

5. ESTIMATION RESULTS AND DISCUSSION

5.1 Introduction

In this chapter, the estimation results obtained for both single and two-analyte systems by using the models developed in chapter 3 are presented. The estimation process was performed on the sensor data collected in the Microsensor Research Laboratory at Marquette University using the SH-SAW sensor. The process of data acquisition had already been discussed in chapter 4. Before the data collected were used to perform the estimation process by using Kalman Filter (KF) or Extended Kalman Filter (EKF) to estimate the unknown parameters, the data were first pre-processed to correct the data for any baseline drift and also to eliminate the outlier points in the data using the techniques presented in chapter 4. The signal-processing steps that the data has to undergo before the estimate of the unknown parameters could be obtained are summarized in the block diagram shown in Fig. 5.1. Note that, all the signal-processing steps shown in the block diagram of Fig. 5.1 can be performed simultaneously in real-time.

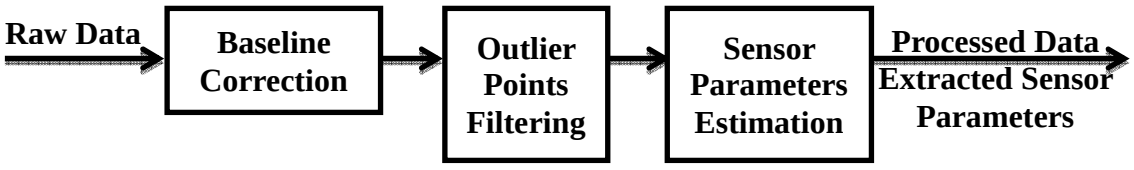

Figure 5.1: Block diagram showing the steps that the data has to undergo before the estimate of the unknown parameters can be obtained.

The estimation results presented in this chapter include plots showing the estimated sensor response co-plotted together with the measured sensor response. The plots also will contain pertinent information about the estimated parameters and the parameters obtained by fitting the measurement data. Note that for the single analyte system, the measurement data were fitted with single exponential fits and for the two-analyte system, the measurement data were fitted with dual-exponential fits. Both the estimated parameters and the fitted parameters are shown in the plot, so that the estimated parameters can be readily compared to the fitted parameters. By using the estimated steady-state frequency shift (or steady-state frequency shift determined by fitting the measurement data) and measured average sensitivities, σ (in Hz/ppm) values for various coating/analyte combinations as presented in Table 4.1, the analyte concentration(s) in the sample can be determined by using the following equations,

$$C_{fit} = \frac{\alpha_{fit}}{\sigma} \tag{5.1}$$

$$C_{est} = \frac{\alpha_{est}}{\sigma} \tag{5.2}$$

where, C_{fit} represents the concentration of the analyte(s) determined by using steady-state frequency shift obtained by fitting the measurement data, α_{fit}, and C_{est} represents the concentration of the analyte(s) determined by using the estimated steady-state frequency shift, α_{est}. C_{fit} and C_{est} are compared to each other to validate the

estimation results obtained. It should be noted that for most cases, the values of C_{est} and C_{fit} will be different than the actual ambient concentration, C_{amb} because of the error introduced during the sample preparations and also due to the volatility of the hydrocarbon analytes. Such discrepancies have been found from independent measurements of the prepared and tested samples to be as high as 13% [35]. Moreover, the chapter also includes a discussion on how rapidly the estimation results could be obtained. Several estimated sensor responses obtained after a specific amount of time are compared to determine the minimum time required to obtain a good estimate of the unknown parameters.

5.2 Single Analyte Estimation Results

In this section, the single analyte estimation results are presented. The single analyte state-space model developed in chapter 3 is used to perform the estimation by using EKF algorithm. For this system, there are three unknown parameters that were estimated which include normalized concentration of the analyte, adsorption rate (i.e. inverse of time constant) and steady-state frequency shift. Note that by using eq. 5.2, the estimated concentrations of the analyte, C_{est} are determined.

Single analyte estimations were tested on various single analyte measurement data, however, in this section only the estimation results obtained for three different measurement data are presented and discussed. For all cases, three types of estimation result figures will be presented and discussed. One of the figures will contain the

information about the estimated sensor response, estimated sensor parameters, measurement data, and measurement data fitting (with the fitting parameters). Note that the single analyte measurement data were fitted using single exponential fits. The second figure will contain the information about the estimated normalized concentration of the analyte along with the theoretical normalized concentration of the analyte. The theoretical normalized concentration of the analyte was determined by using the average time constants values of the analyte/coating pair given in Table 4.2 and the following equation,

$$m(t) = (1 - e^{-\frac{t}{\tau}}) \tag{5.3}$$

where $m(t)$ represents the normalized concentration of absorbed analyte at time, t and τ represents the time constant of the analyte. Note that eq. 5.3 was obtained by solving the normalized concentration differential equation as given by eq. 3.5. The final figure will show several estimated sensor responses obtained after a certain number of minutes plotted along with the measurement data and measurement data fitting. Following this figure is a table which shows the estimated parameters corresponding to the estimated sensor responses shown in the final figure along with the percentage difference between these estimated parameters and the parameters obtained from experimental data fitting. This is done to determine the minimum time required to obtain a good estimate of the sensor parameters.

First, the estimation results obtained by using the measurement data of the sensor response of an SH-SAW sensor coated with 1.0µm PEA to 1000 ppb ethylbenzene are discussed and Fig. 5.2 through Fig. 5.4 show the estimation results obtained by using this data. In Fig. 5.2, the measurement data (blue asterisk), measurement data fitting (blue curve) and the estimated sensor response (red curve) are shown. Also shown in Fig. 5.2 are the estimated sensor parameters along with the parameters determined by fitting the measurement data. As can be seen from Fig. 5.2, both the estimated steady-state frequency shift and adsorption rate are in close agreement (i.e. less than $\pm 1\%$ difference) with the steady-state frequency shift and adsorption rate determined by fitting the measurement data. By using eq. 5.1 and eq. 5.2, C_{fit} and C_{est} are found to be around 813 ppb and 808 ppb, respectively. The values for C_{est} and C_{fit} are also in good agreement (i.e. less than $\pm 1\%$ difference) with each other; this result should be expected because their steady-state frequency shifts are in good agreement as well. The estimation result obtained for the normalized concentration of ethylbenzene versus time (in red) is shown in Fig. 5.3. Also shown in Fig. 5.3 is the theoretical normalized concentration of ethylbenzene (in blue). Based on Fig. 5.3, it can be seen that the estimated normalized concentration of ethylbenzene is in very good agreement with the theoretical normalized concentration of ethylbenzene and it approaches a value of one as time increases. Note that normalized concentration is the ratio of concentration of the analyte in the coating at time, t to the maximum ambient concentration of the analyte and as time increases the ratio should approach a value of one (i.e. the analyte concentration in the coating and maximum ambient concentration times the partition coefficient are equal) as observed in Fig. 5.3. Finally in Fig. 5.4, three estimated sensor responses obtained using the

measurement data collected for the first 2, 3 and 4 minutes after the analyte has been introduced to the sensor are shown. Also shown in Fig. 5.4 are the measurement data, measurement data fitting and the estimated sensor response using all the data points. The estimated sensor parameters obtained using the measurement data collected for the first 2, 3 and 4 minutes after the analyte has been introduced to the sensor are given in Table 5.1. Based on Table 5.1 and Fig. 5.4, it can be seen that the estimated sensor response and sensor parameters obtained using the measurement data collected for the first 3 minutes (and above) agree well (i.e. less than $\pm 10\%$ difference) with the measurement data and measurement data fitting. Therefore, the data collected for the first 3 minutes are sufficient to obtain a good estimate of the unknown parameters using the EKF algorithm. This means that, by using single analyte estimation technique presented in this work, one could quantify the analyte well before the sensor response reaches steady-state.

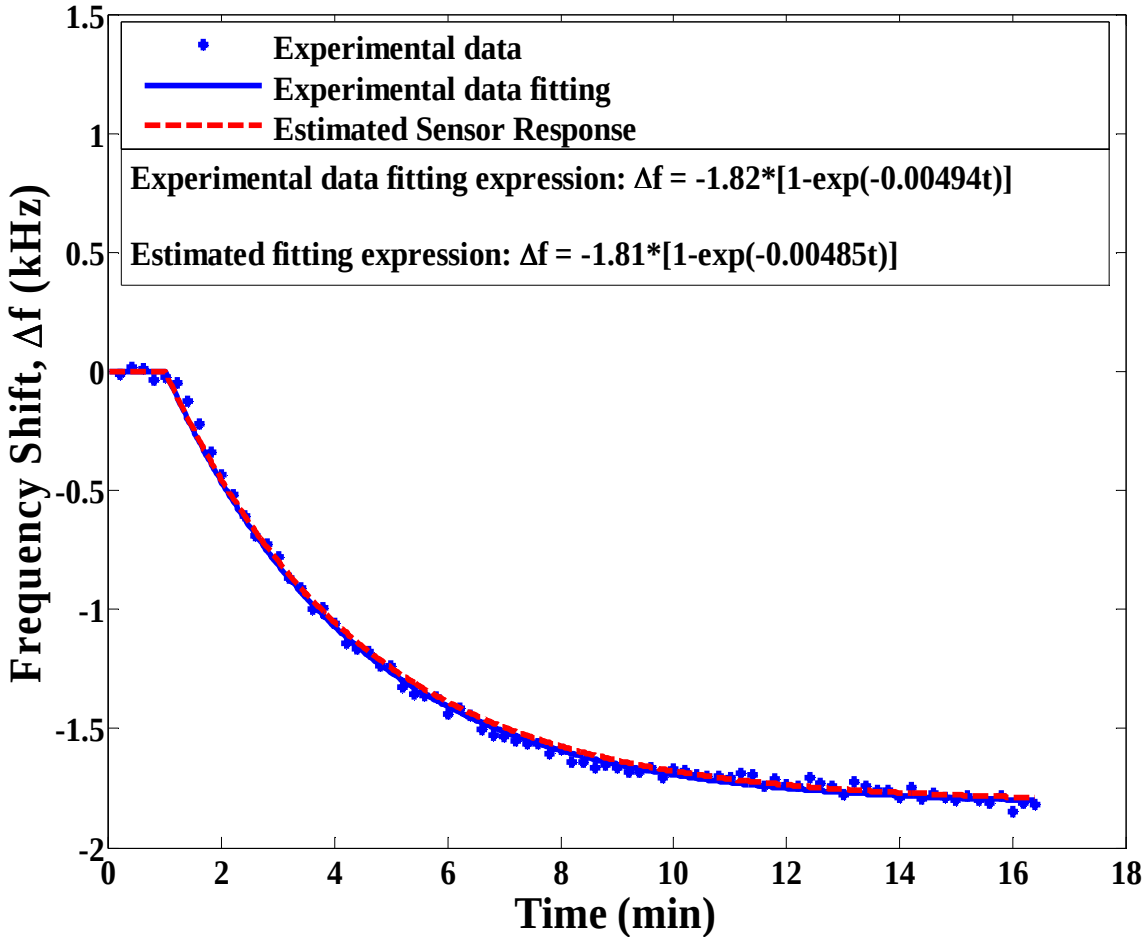

Figure 5.2: Response of a SH-SAW sensor coated with 1.0μm PEA to 1000 ppb ethylbenzene (blue curve) along with the estimated sensor response (red curve). Also shown in the figure are the estimated sensor parameters along with the parameters determined by fitting the measurement data.

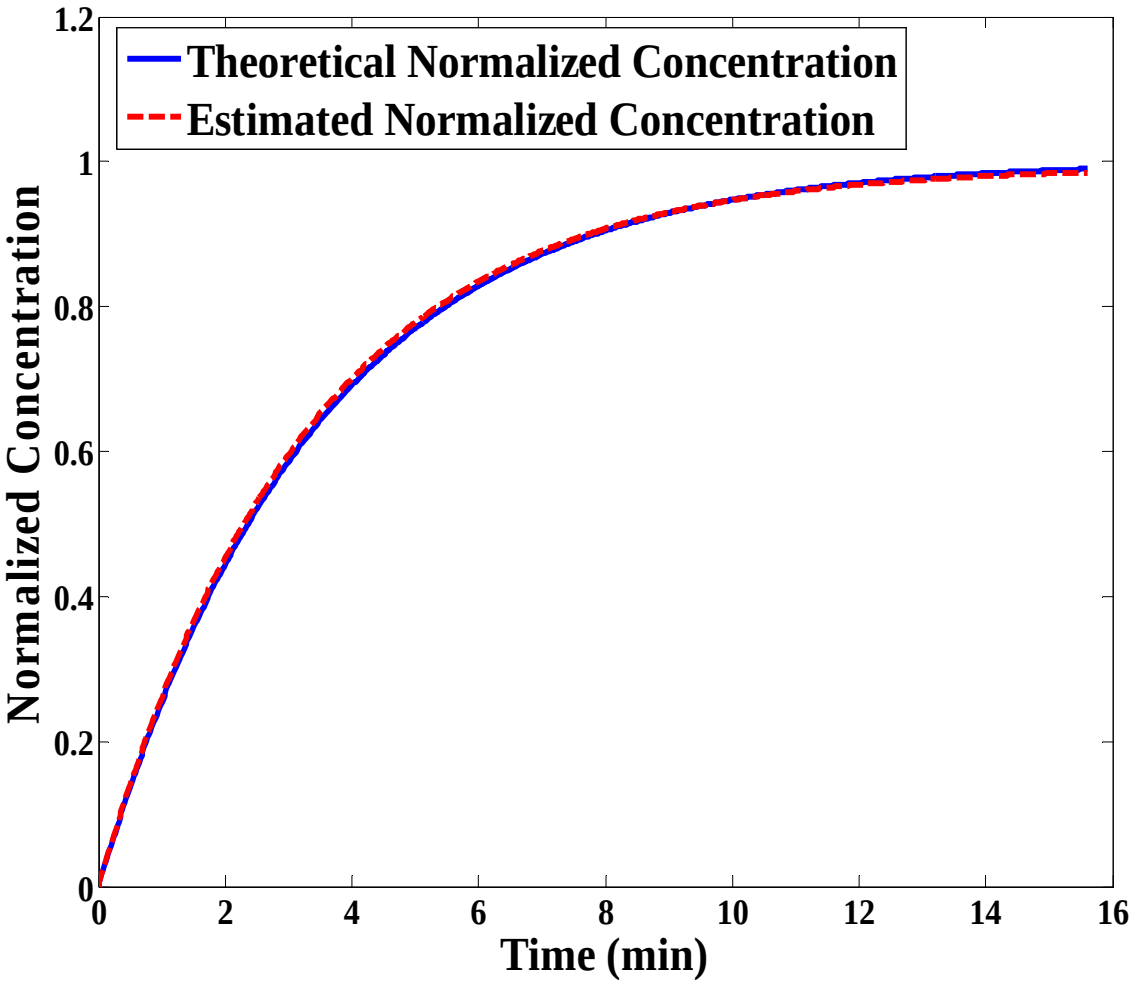

Figure 5.3: Estimated normalized concentration of ethylbenzene co-plotted with the theoretical normalized concentration of ethylbenzene for 1.0µm PEA.

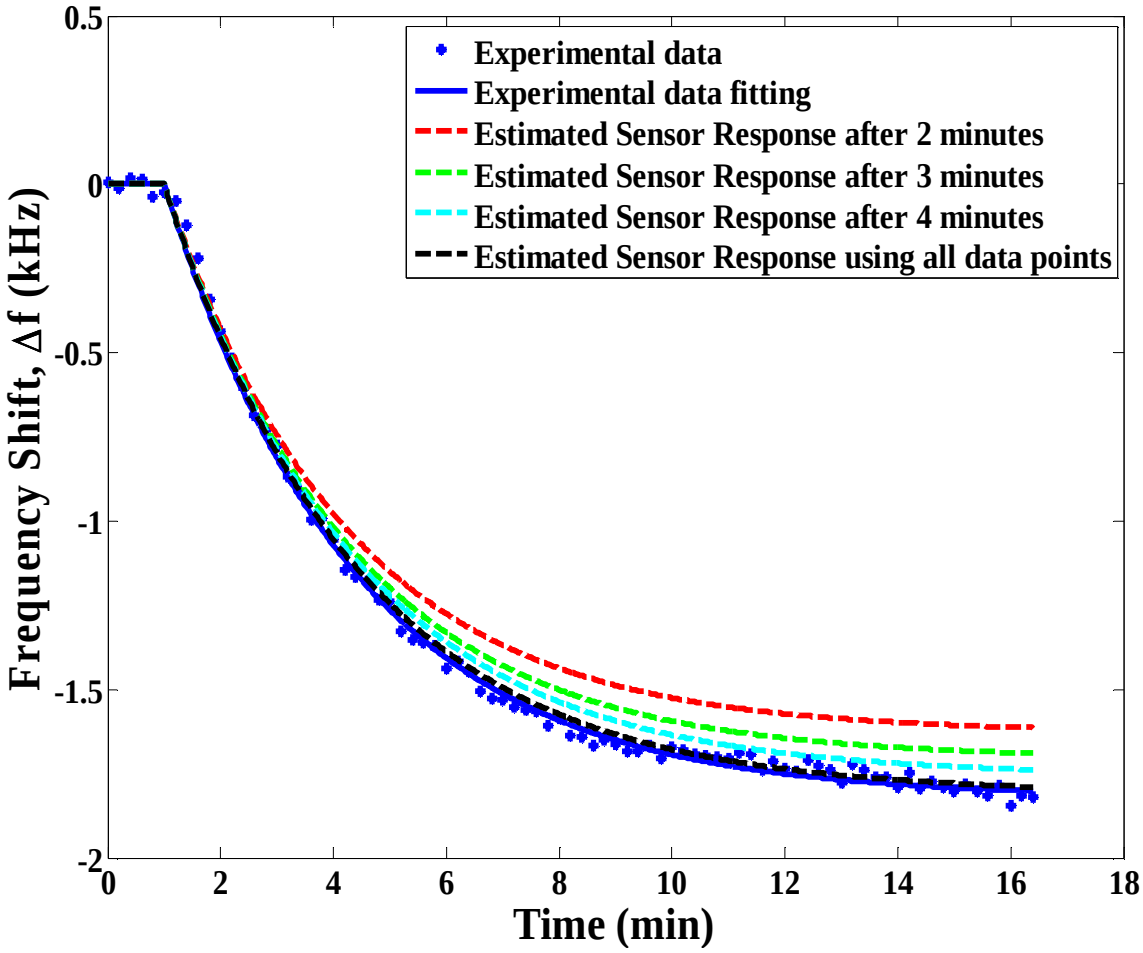

Figure 5.4: Estimated sensor response of 1000 ppb ethylbenzene (for 1.0µm PEA coating) obtained using the measurement data collected for the first 2, 3 and 4 minutes after the analyte has been introduced to the sensor co-plotted together with the measurement data, measurement data fitting and also the estimated sensor response using all the measurement data points.

Estimated Sensor Parameters (Measurement Data 1)	Steady-State Frequency Shift, α_{est} (kHz) (% difference with α_{fit})	Time Constant, τ_{est} (s) (% difference with τ_{fit})
After 2 minutes	**-1.63 (10.53 %)**	**196.08 (3.20 %)**
After 3 minutes	**-1.71 (6.31 %)**	**197.07 (2.36 %)**
After 4 minutes	**-1.76 (3.49 %)**	**202.02 (0.17 %)**
Using all data points	**-1.82 (0.55 %)**	**206.19 (1.86 %)**

Table 5.1: Estimated sensor parameters of 1000 ppb ethylbenzene (for 1.0µm PEA coating) obtained using the measurement data collected for the first 2, 3 and 4 minutes after the analyte has been introduced to the sensor along with the estimated sensor parameters obtained using all the data points. Also given in the table are the percentage differences between the estimated sensor parameters and sensor parameters determined by fitting the measurement data.

Next, the estimation results obtained by using the measurement data for the sensor response of a SH-SAW sensor coated with 0.6µm PECH to 1000 ppb benzene are discussed. The following figures, Fig. 5.5, Fig. 5.6 and Fig. 5.7 and Table 5.2 show the estimation results obtained for this case.

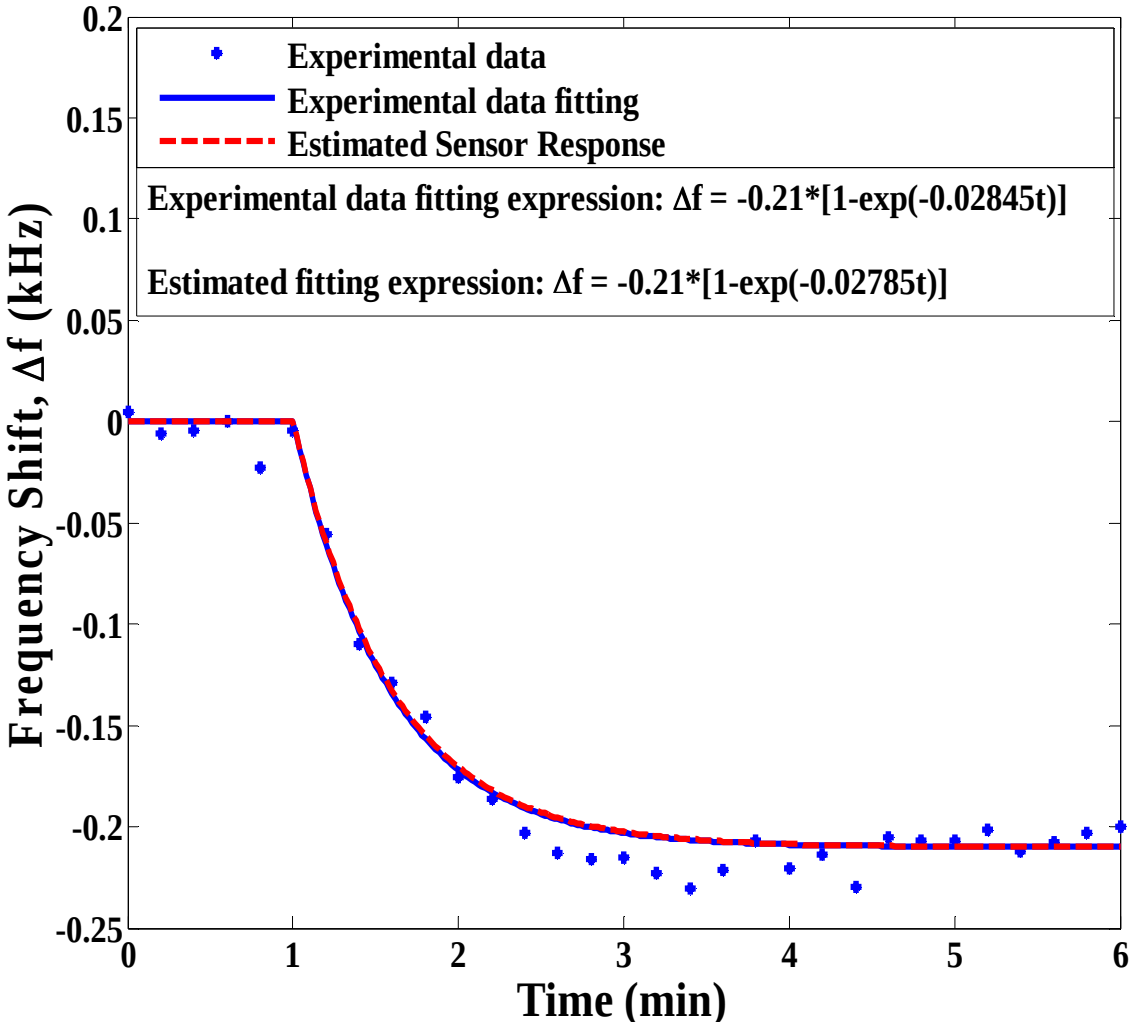

Figure 5.5: Response of a SH-SAW sensor coated with 0.6μm PECH to 1000 ppb benzene (blue curve) along with the estimated sensor response (red curve). Also shown in the figure are the estimated sensor parameters along with the parameters determined by fitting the m measurement data.

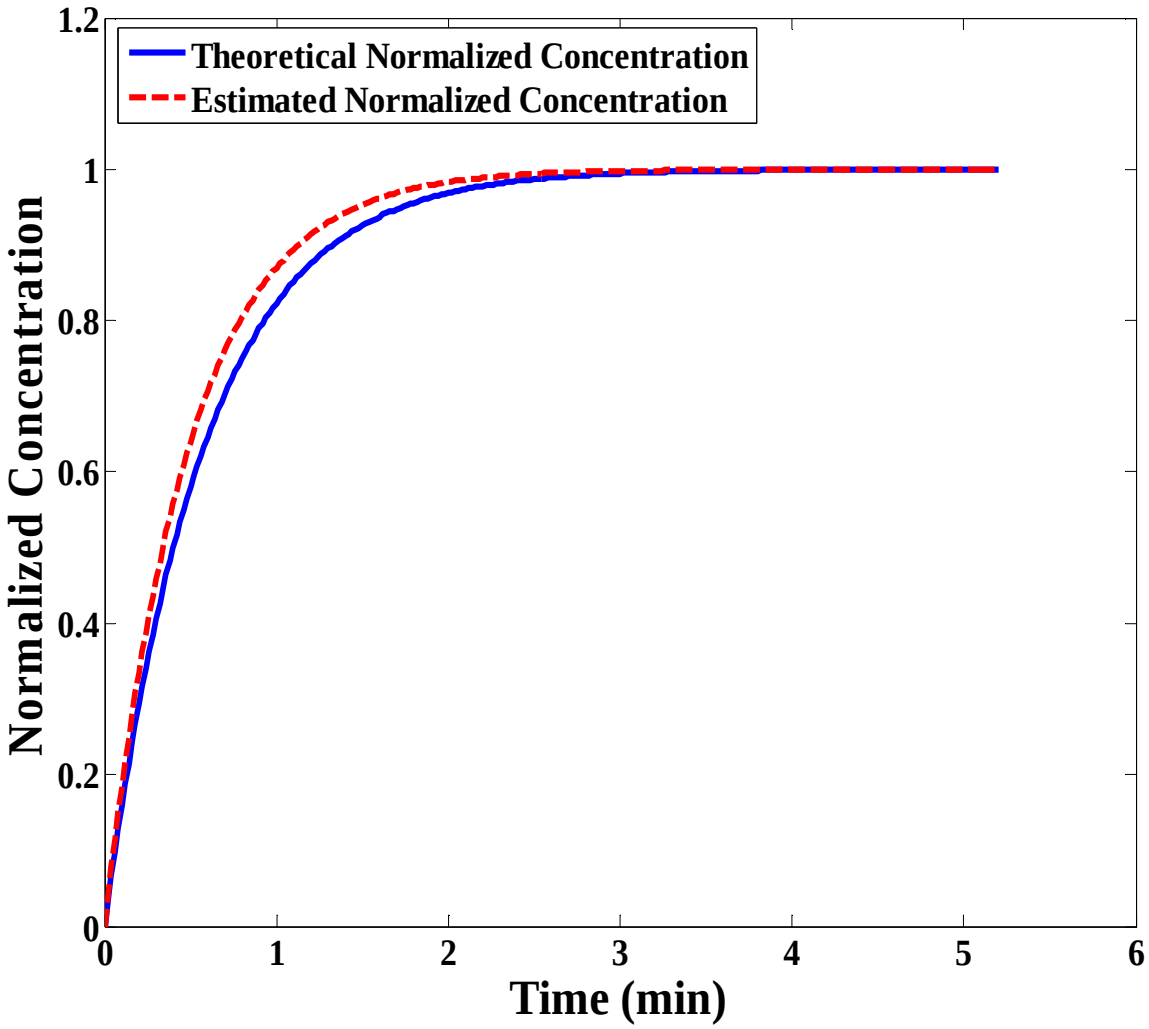

Figure 5.6: Estimated normalized concentration of benzene co-plotted with the theoretical normalized concentration of benzene for 0.6µm PECH.

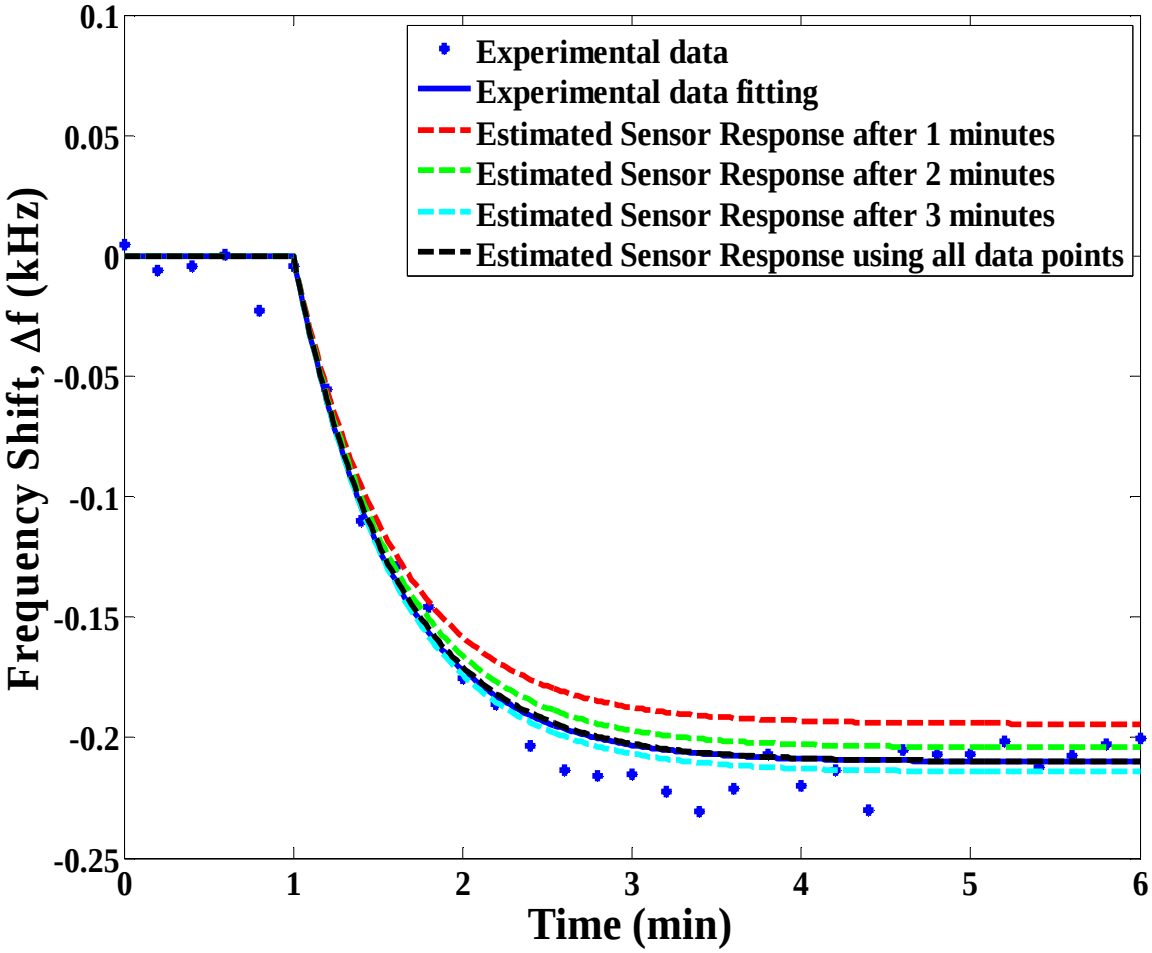

Figure 5.7: Estimated sensor response to 1000 ppb benzene (for 0.6μm PECH coating) obtained using the measurement data collected for the first 1, 2 and 3 minutes after the analyte has been introduced to the sensor co-plotted together with the measurement data, measurement data fitting and also the estimated sensor response using all the data points.

Estimated Sensor Parameters (Measurement Data 2)	**Steady-State Frequency Shift, α_{est} (kHz) (% difference with α_{fit})**	**Time Constant, τ_{est} (s) (% difference with τ_{fit})**
After 1 minutes	**-0.1943 (7.48 %)**	**35.74 (1.64 %)**
After 2 minutes	**-0.2042 (2.76 %)**	**35.82 (1.87 %)**
After 3 minutes	**-0.2143 (2.05 %)**	**35.94 (2.20 %)**
Using all data points	**-0.21 (0 %)**	**35.91 (2.11 %)**

Table 5.2: Estimated sensor parameters for 1000 ppb benzene (for 0.6µm PECH coating) obtained using the measurement data collected for the first 1, 2 and 3 minutes after the analyte has been introduced to the sensor along with the estimated sensor parameters obtained using all the data points. Also given in the table are the percentage differences between the estimated sensor parameters and sensor parameters determined by fitting the measurement data.

As can be seen from Fig. 5.5, both the estimated steady-state frequency shifts and adsorption rate for the second case are also in agreement (i.e. less than $\pm 2\%$ difference) with the steady-state frequency shift and absorption rate determined by fitting the measurement data. Both C_{fit} and C_{est} are found to be equal to 817 ppb. In Fig. 5.6, the estimation result obtained for the normalized concentration of benzene versus time is shown along with the theoretical normalized concentration of benzene for 0.6µm PECH coating. As depicted in Fig 5.6, the estimated normalized concentration and theoretical normalized concentration are in good agreement with each other. This further validates

the estimation results obtained for this second case. Also note that from Fig. 5.6, the normalized concentration of benzene reaches the value of one much faster than the normalized concentration of ethylbenzene shown in Fig.5.3. This is because benzene has a much faster response time (smaller time constant) compared to ethylbenzene. Fig. 5.7 shows the estimated sensor responses and Table 5.2 shows the estimated sensor parameters obtained using the measurement data collected for the first 1, 2 and 3 minutes after the analyte has been introduced to the sensor. Based on Table 5.2 and Fig. 5.7, it can be seen that the estimated sensor response and sensor parameters obtained using the measurement data collected for the first 1 minutes (and above) agree well (i.e. less than $\pm 10\%$ difference) with the measurement data and measurement data fitting. Therefore, the data collected for one minute are sufficient to obtain a good estimate of the unknown parameters using the EKF algorithm. These results indicate that the sensor parameters could be estimated long before the sensor response reaches steady-state. Therefore, once again it can be concluded that by using the state-space model of single analyte and EKF algorithm presented in the present work, the analyte could be quantified well before the sensor response reaches steady-state.

Finally for the third case, the estimation results obtained by using the measurement data for the sensor response of a SH-SAW sensor coated with 0.8μm PIB to 1000 ppb toluene are discussed. Fig. 5.8 through Fig 5.10 and Table 5.3 show the estimation results obtained for this third case.

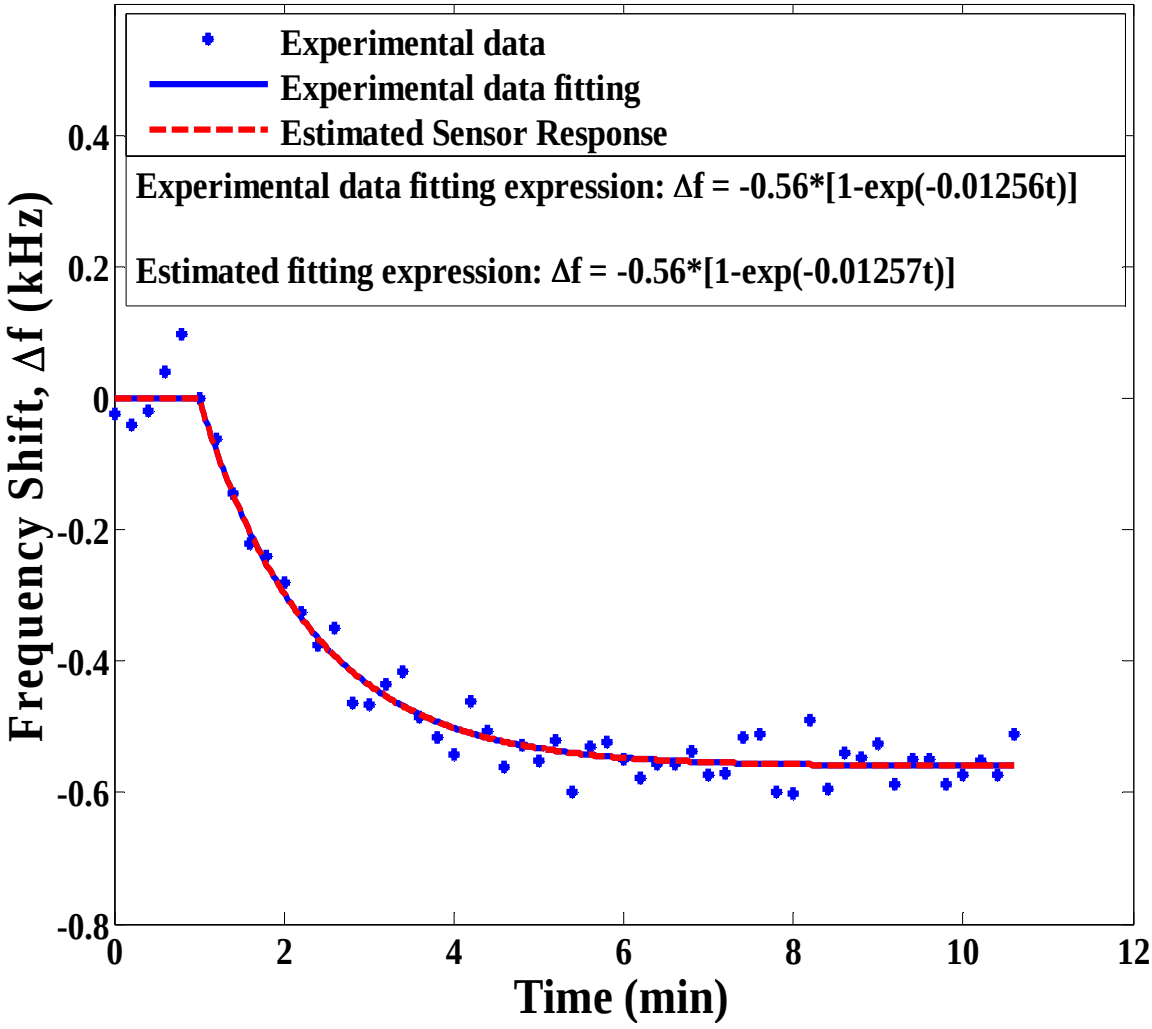

Figure 5.8: Response of a SH-SAW sensor coated with 0.8μm PIB to 1000 ppb toluene (blue curve) along with the estimated sensor response (red curve). Also shown in the figure are the estimated sensor parameters along with the parameters determined by fitting the measurement data.

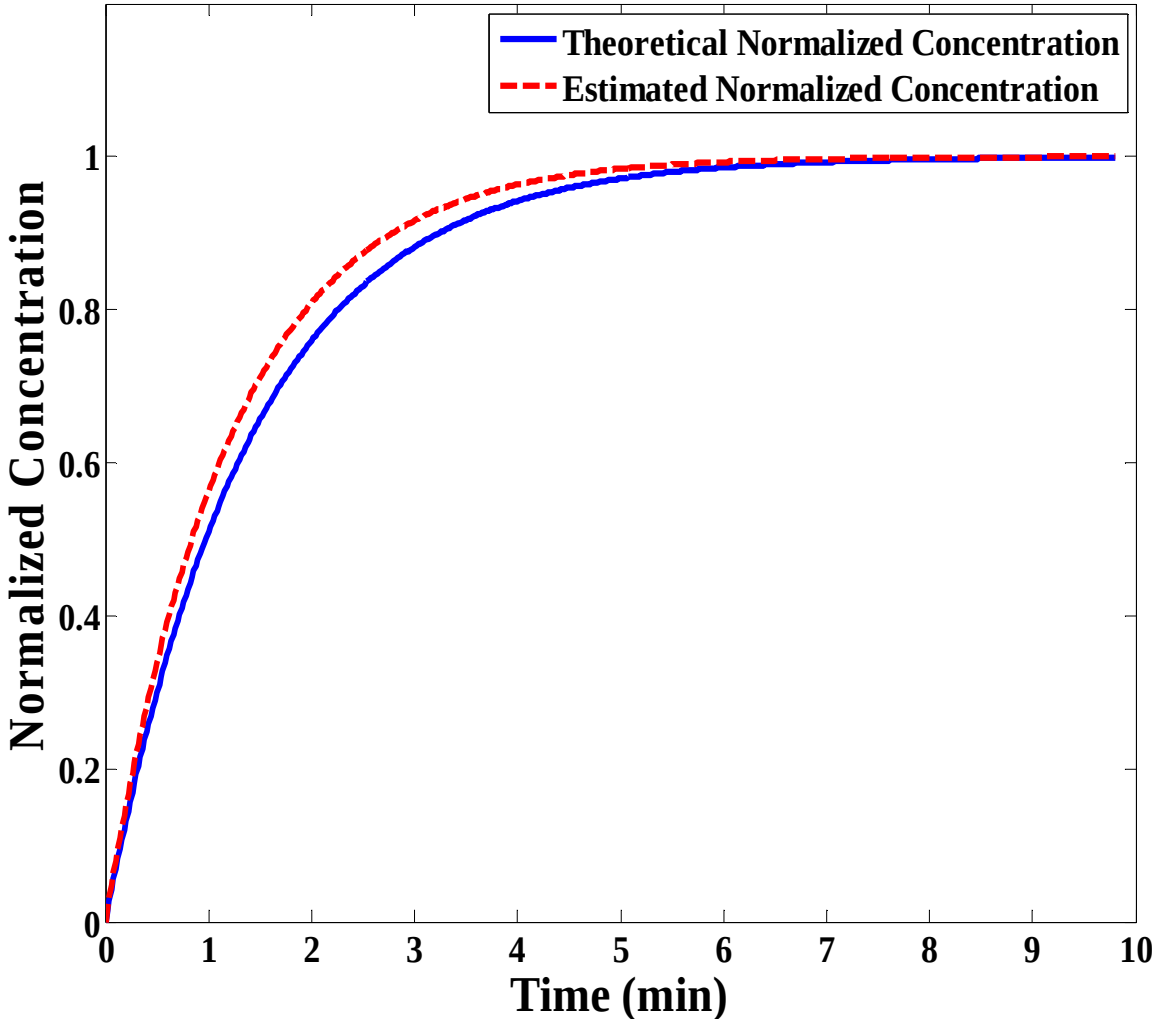

Figure 5.9: Estimated normalized concentration of toluene co-plotted with the theoretical normalized concentration of toluene for 0.8μm PIB.

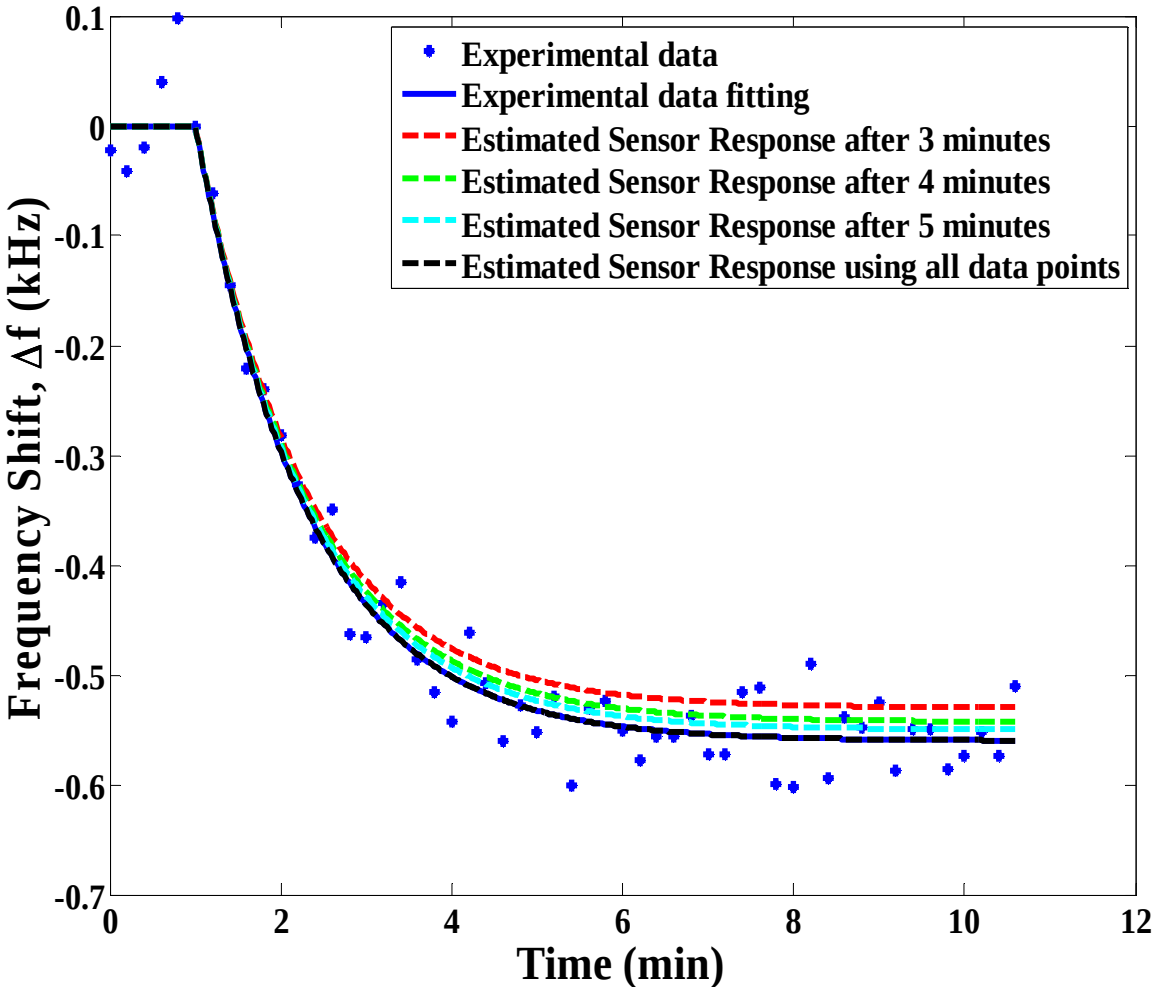

Figure 5.10: Estimated sensor response to 1000 ppb toluene (for 0.8µm PIB coating) obtained using the measurement data collected for the first 3, 4 and 5 minutes after the analyte has been introduced to the sensor co-plotted together with the measurement data, measurement data fitting and also the estimated sensor response using all the data points.

Estimated Sensor Parameters (Measurement Data 3)	**Steady-state Frequency Shift, α_{est} (kHz)** **(% difference with α_{fit})**	**Time Constant, τ_{est} (s)** **(% difference with τ_{fit})**
After 3 minutes	**-0.5299 (5.38 %)**	**78.74 (1.13 %)**
After 4 minutes	**-0.5428 (3.07 %)**	**79.00 (0.80 %)**
After 5 minutes	**-0.5501 (1.77 %)**	**79.26 (0.46 %)**
Using all data points	**-0.56 (0 %)**	**79.58 (0.07 %)**

Table 5.3: Estimated sensor parameters for 1000 ppb toluene (for 0.8µm PIB coating) obtained using the measurement data collected for the first 3, 4 and 5 minutes after the analyte has been introduced to the sensor along with the estimated sensor parameters obtained using all the data points. Also given in the table are the percentage differences between the estimated sensor parameters and sensor parameters determined by fitting the measurement data.

As indicated in Fig. 5.8, the estimated steady-state frequency shift and absorption rate for the third case are also in conformity (i.e. less than $\pm 1\%$ difference) with the steady-state frequency shift and absorption rate determined by fitting the measurement data. Since the estimated steady-state frequency shift and steady-state frequency shift determined by fitting the measurement data closely match each other, C_{fit} and C_{est} are also expected to be in close agreement. In fact, for this third case both the values of C_{fit} and C_{est} are approximately equal to 1447 ppb. Next in Fig. 5.9, the estimation result obtained for the normalized concentration of toluene versus time is depicted along with the theoretical

normalized concentration of toluene for 0.8µm PIB coating. Based on Fig. 5.9, it can be seen that the estimated normalized concentration and theoretical normalized concentration are in good agreement with each other and this further validates the estimation results obtain for the third case. Also note that in Fig. 5.9, the normalized concentration of toluene reaches the value of one much faster than the normalized concentration of ethylbenzene as shown in Fig.5.3 but slower than the normalized concentration of benzene as shown in Fig. 5.6. This should be expected because the response time of toluene is between those of benzene and ethylbenzene. Fig. 5.10 shows the estimated sensor responses and Table 5.3 shows the estimated sensor parameters obtained using the measurement data collected for the first 3, 4 and 5 minutes after the analyte has been introduced to the sensor. Based on Table 5.3 and Fig. 5.10, it can be seen that the estimated sensor response and sensor parameters obtained using the measurement data collected for the first 4 minutes (and above) agree well (i.e. less than $\pm$10% difference) with the measurement data and measurement data fitting. Therefore, the data collected for the first 4 minutes are sufficient to obtain a good estimate of the unknown parameters using the EKF algorithm. Again this third case also proves that by using the single analyte estimation process presented in the present work, the sensor parameters of single analyte could be estimated well before the sensor response reaches steady-state.

All the estimated sensor parameters along with the maximum percentage difference with the sensor parameters determined by fitting the measurement data for all the three cases are summarized in Table 5.4. It should be noted that the observed

discrepancies between C_{fit} and C_{est} with the actual concentration can be attributed to the error in the sample preparations as well as volatility of the hydrocarbon analytes. From Table 5.4, it can be seen that all the estimated sensor parameters agree well (i.e. less than $\pm 10\%$ difference) with the sensor parameters determined by fitting the measurement data. Therefore, one could obtain the same result either by fitting the measurement data or by using the estimation technique for the single analyte sample presented in this work. Although both methods produce approximately the same results, the advantage of the method presented in this thesis lies in the fact that only the measurement data collected for the first few minutes (i.e. around 1-4 minutes) are needed to quantify the analyte with sufficient accuracy. This means that the sensor parameters could be estimated in less than half the time required for the sensor response to reach steady-state. Therefore, one could obtain an accurate estimate of the sensor parameters of single analyte well before the sensor response reaches steady-state.

Data	Polymer	Steady-State Frequency Shift (kHz)		Time Constant (s)		Concentration (ppb)		Percentage Difference (with the fitting parameters)
		α_{fit}	α_{est}	τ_{fit}	τ_{est}	C_{fit}	C_{est}	
1: 1000 ppb ethylbenzene	**1.0µm PEA**	**-1.82**	**-1.81**	**202.36**	**204.08**	**813**	**808**	**0.85%**
2: 1000 ppb benzene	**0.6µm PECH**	**-0.21**	**-0.21**	**35.15**	**35.91**	**817**	**817**	**2.16%**
3: 1000 ppb toluene	**0.8µm PIB**	**-0.56**	**-0.56**	**79.63**	**84.21**	**1447**	**1447**	**5.75%**

Table 5.4: Estimated sensor parameters (steady-state frequency shift and time constant) and concentration along with sensor parameters and concentration determined from measurement data fit. Also given is the maximum percentage difference between the estimated sensor parameters and sensor parameters determined by fitting the measurement data with single exponential fit.

5.3 Two-Analyte Estimation Results

In this section, the two-analyte estimation results are presented. The two-analyte state-space models (i.e. both nonlinear and linear model) presented in chapter 3 are used to perform the estimation by using EKF or KF algorithm depending on the state-space model used. For the nonlinear model, EKF algorithm was used to perform the estimation and for the linear model, KF algorithm was used. Estimations of binary mixtures using both state-space models were performed on various measurement data; however, in this section only the estimation results obtained for two different measurement data are presented and discussed. Also it should be noted that, in order to compare the different

models (i.e. nonlinear and linear model), the same measurement data were used for both models to estimate the unknown parameters. After the results obtained for each model are discussed, comparisons between the two models are made.

5.3.1 Nonlinear Model

For the nonlinear model, there are four unknown quantities that were estimated which include the normalized concentrations of each analyte and the steady-state frequency shifts of each analyte. The time constants of each analyte are not estimated because it is assumed that the values of the time constants for each analyte are known from the single analyte measurement. For the present work, the average time constants for various coating/analyte pair given in Table 4.2 (in chapter 4) were used as the time constant of the corresponding analytes. As mentioned earlier, only the estimation results obtained from two different measurement data (i.e. two different cases) will be presented and discussed. For these two cases, four types of estimation result figures will be presented and discussed. The first figure will contain the information about the estimated sensor response, estimated sensor parameters, measurement data, and measurement data fitting (with the fitting parameters). Note that the two-analyte measurement data were fitted using dual-exponential fits. Next, the following two figures will show the result obtained for the estimated normalized concentration of each analyte in the binary mixture sample along with its theoretical normalized concentration determined by using eq. 5.3. Finally, the fourth figure will show several estimated sensor responses determined after a certain number of minutes plotted along with the measurement data and measurement

data fitting. Following this figure is a table which shows the estimated parameters corresponding to the estimated sensor responses shown in the fourth figure along with the percentage difference between these estimated parameters and the parameters obtained from experimental data fitting. This is done to determine the minimum time required to obtain a good estimate of the sensor parameters.

First, the estimation results obtained by using the measurement data of the sensor response of SH-SAW sensor coated with 1.0μm PEA to binary mixture of 500 ppb benzene and 200 ppb ethylbenzene are discussed. The following figures, Figs. 5.11 – 5.14 and Table 5.5, show the estimation results obtained by using this binary mixture data and the nonlinear model. Fig. 5.11 shows the measurement data (blue asterisk), measurement data fitting (blue curve) and the estimated sensor response (red curve). Also shown in Fig. 5.11 are the estimated steady-state frequency shifts along with the steady-state frequency shifts determined by fitting the measurement data. As can be seen from Fig. 5.11, the estimated steady-state frequency shift for each analyte is in agreement (i.e. less than ± 6 % difference) with the steady-state frequency shift of the analytes determined by fitting the measurement data. Note that, the measurement data were fitted by using dual-exponential fits to extract the steady-state frequency shift of each analyte in the binary mixture sample. Moreover, by using eq. 5.1 and eq. 5.2, C_{fit} and C_{est} for both analytes are found to be around 285 ppb and 303 ppb, respectively, for the concentration of benzene and 280 ppb and 283 ppb, respectively for the concentration of ethylbenzene. Therefore, C_{fit} and C_{est} found for both analytes are within ± 6 % difference. Fig. 5.12 and Fig. 5.13 show the estimation results obtained for the normalized concentration of

each analyte versus time along with its theoretical normalized concentration determined by using eq. 5.3. As can be seen from Fig. 5.12 and Fig. 5.13, both the estimated normalized concentration of each analyte (i.e. benzene and ethylbenzene) are in conformity with its theoretical normalized concentration. This further validates the estimation results obtained for this case. Fig. 5.14 shows the estimated sensor responses and Table 5.5 shows the estimated steady-state frequency shifts obtained using the measurement data collected for the first 2, 3 and 4 minutes after the binary mixture has been introduced to the sensor. Based on Table 5.5 and Fig. 5.14, it can be seen that the estimated sensor response and sensor parameters obtained using the measurement data collected for the first 3 minutes (and above) agree well (i.e. less than $\pm 20\%$ difference) with the measurement data and measurement data fitting. Therefore, the steady-state frequency shifts of the analytes could be estimated even before the sensor response reaches steady-state and thus, the analytes in the binary mixture can be quantified rapidly by using the nonlinear model of the two-analyte system.

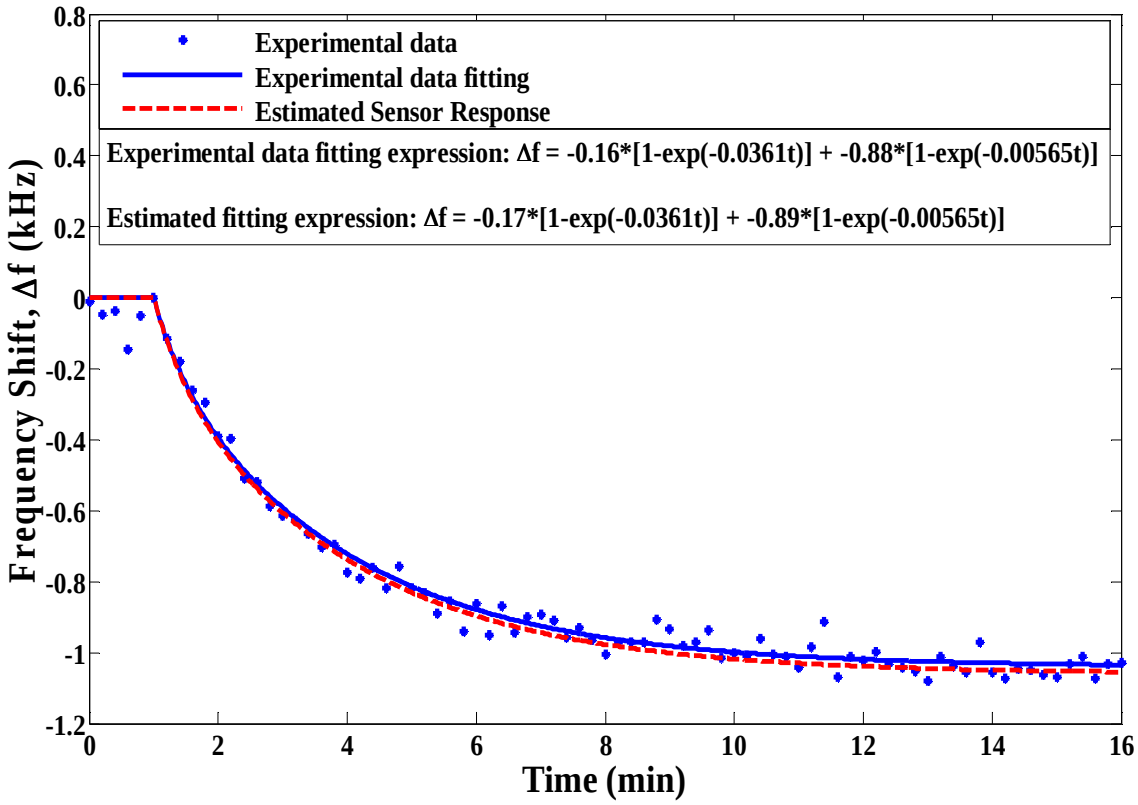

Figure 5.11: Response of a SH-SAW sensor coated with 1.0μm PEA to a binary mixture of 500 ppb benzene and 200 ppb ethylbenzene (blue curve) along with the estimated sensor response using the nonlinear model of the two-analyte system (red curve). Also shown in the figure are the estimated steady-state frequency shifts along with the steady-state frequency shifts determined by fitting the measurement data using dual-exponential fit.

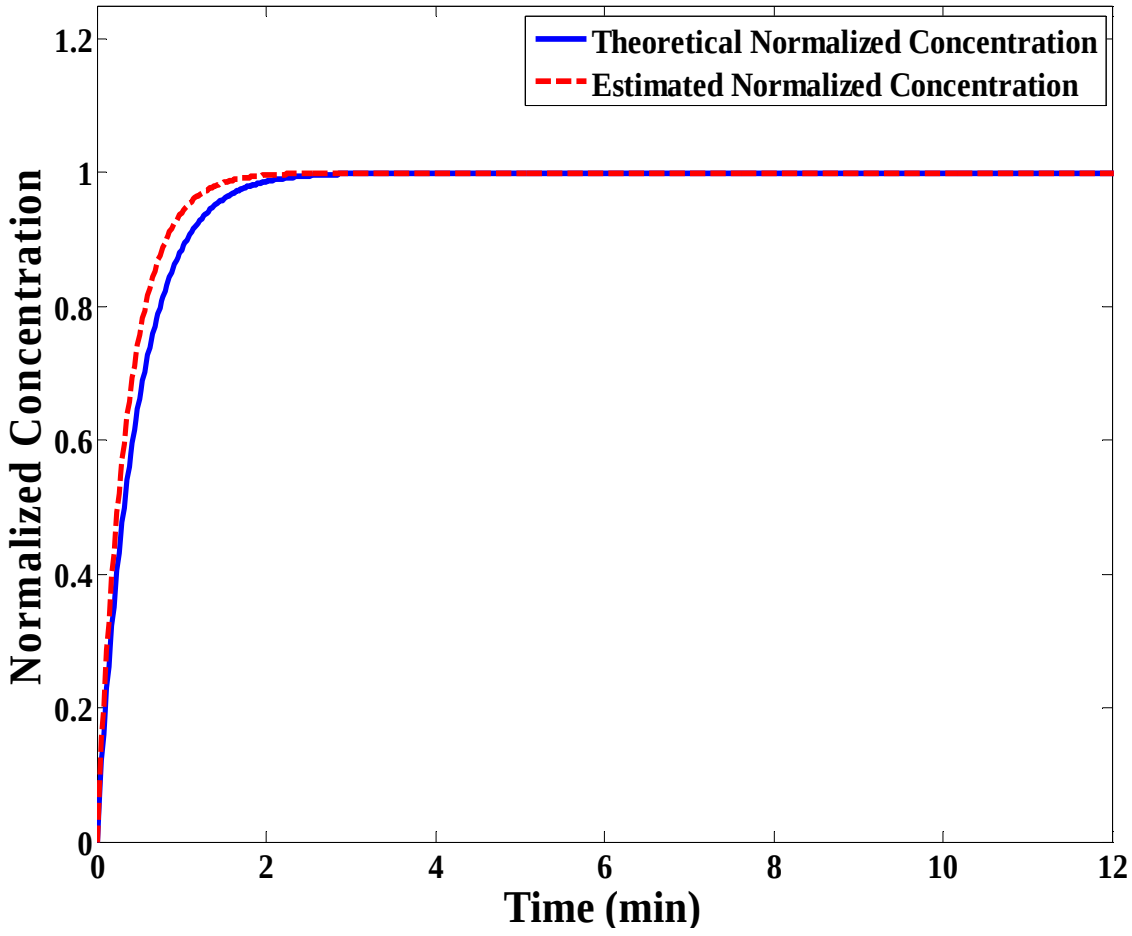

Figure 5.12: Estimated normalized concentration of benzene using the nonlinear model of two-analyte system co-plotted with the theoretical normalized concentration of benzene for 1.0μm PEA.

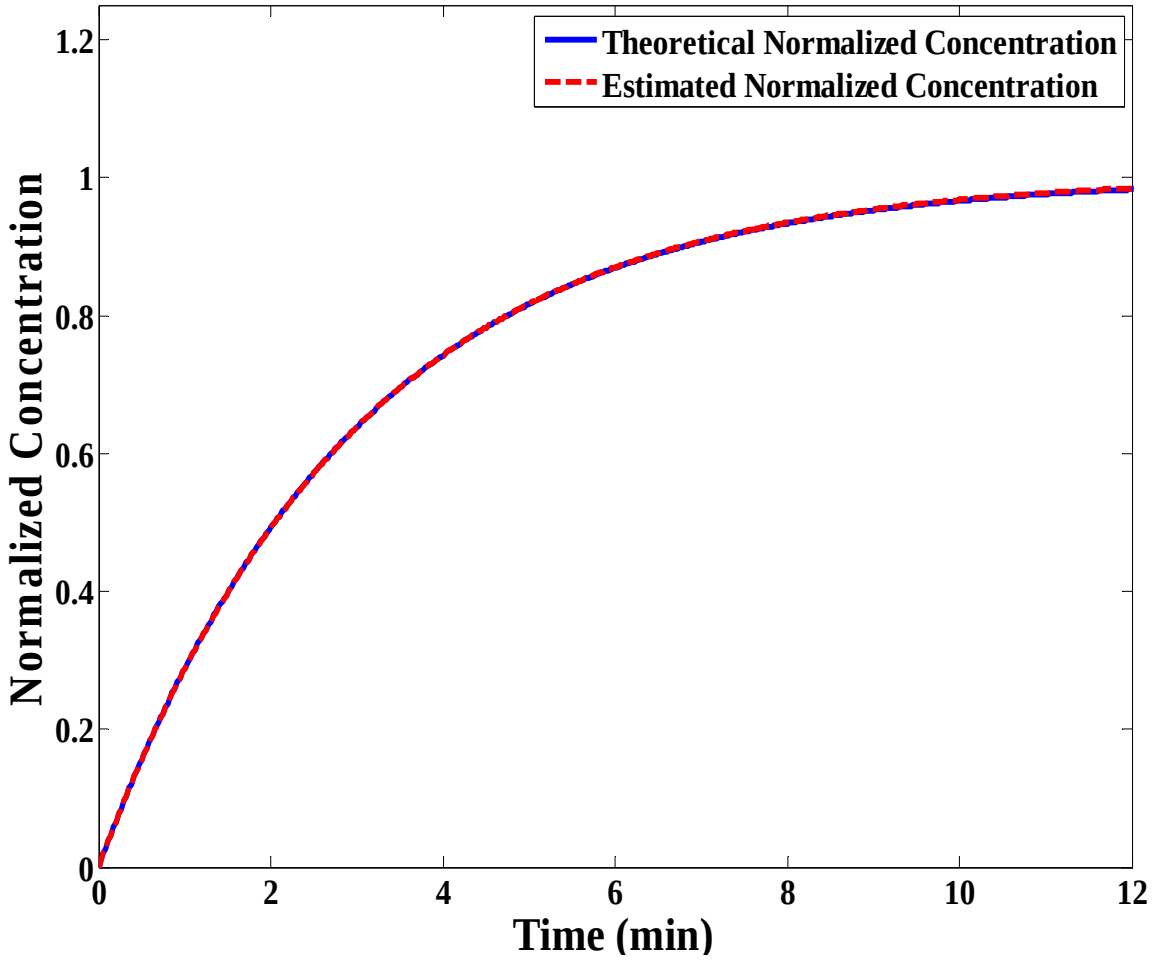

Figure 5.13: Estimated normalized concentration of ethylbenzene using the nonlinear model of two-analyte system co-plotted with the theoretical normalized concentration of ethylbenzene for 1.0µm PEA.

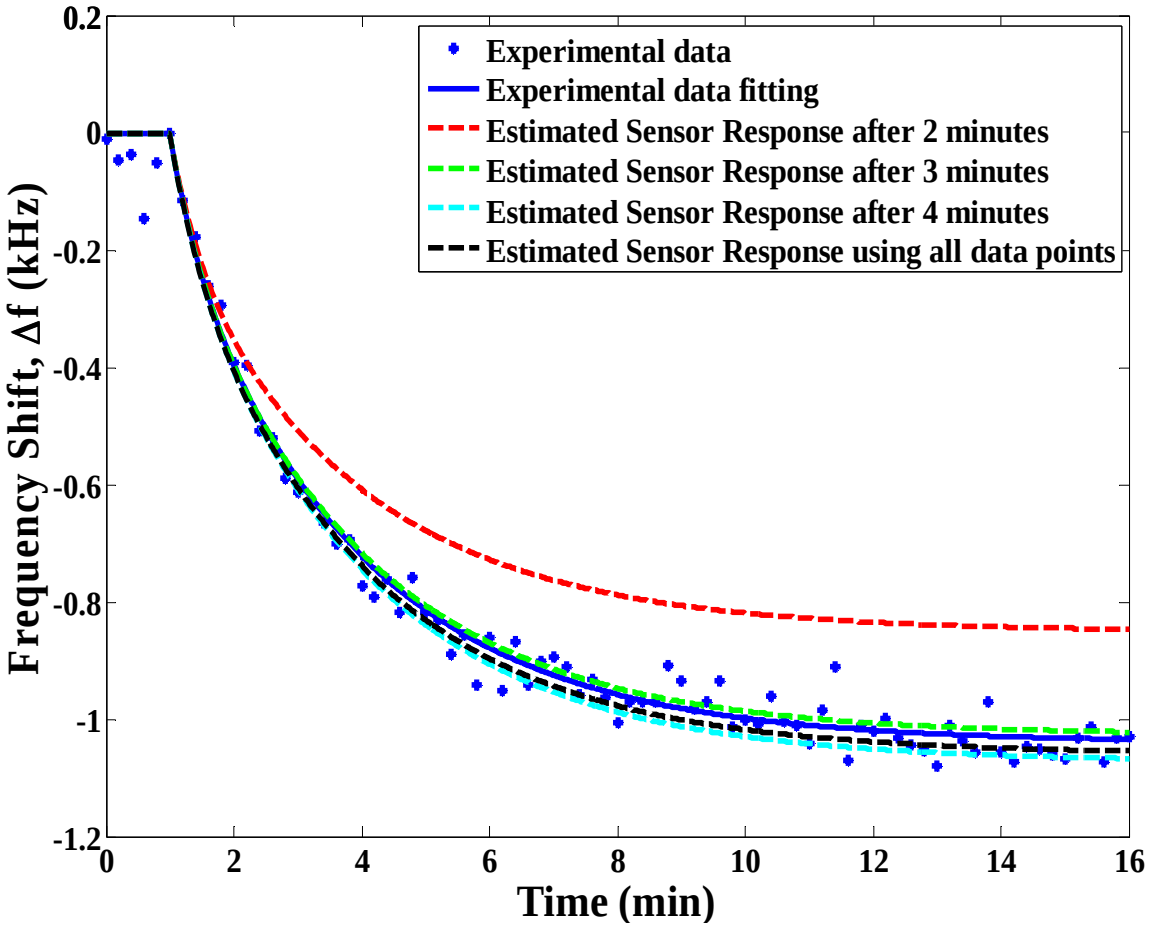

Figure 5.14: Estimated sensor response to a binary mixture of 500 ppb benzene and 200 ppb ethylbenzene (with 1.0µm PEA coating) using the nonlinear model of two-analyte system obtained using the measurement data collected for the first 2, 3 and 4 minutes after the binary mixture sample has been introduced to the sensor co-plotted together with the measurement data, measurement data fitting and also the estimated sensor response using all the data points.

Estimated Steady-State Frequency Shift (Measurement Data 1)	**Steady-State Frequency Shift of Benzene, α_{est} (kHz)** **(% difference with α_{fit})**	**Steady-State Frequency Shift of Ethylbenzene, α_{est} (kHz)** **(% difference with α_{fit})**
After 2 minutes	**-0.1809 (13.06 %)**	**-0.6699 (23.88 %)**
After 3 minutes	**-0.1678 (4.88 %)**	**-0.8594 (2.34 %)**
After 4 minutes	**-0.1647 (2.94 %)**	**-0.9080 (3.18 %)**
Using all data points	**-0.1652 (3.25 %)**	**-0.8896 (1.09 %)**

Table 5.5: Estimated steady-state frequency shifts for a binary mixture of 500 ppb benzene and 200 ppb ethylbenzene (with 1.0μm PEA coating) using the nonlinear model of two-analyte system obtained using the measurement data collected for the first 2, 3 and 4 minutes after the binary mixture sample has been introduced to the sensor along with the estimated steady-state frequency shifts obtained using all the data points. Also given in the table are the percentage differences between the estimated steady-state frequency shifts and steady-state frequency shifts determined by fitting the measurement data.

Next for the second case, the estimation results obtained by using the measurement data of the sensor response of SH-SAW sensor coated with 0.6μm PECH to binary mixture of 1000 ppb benzene and 500 ppb toluene are discussed. Figs. 5.15 - 5.18 and Table 5.6 show the estimation results obtained by using the nonlinear model for this second measurement data.

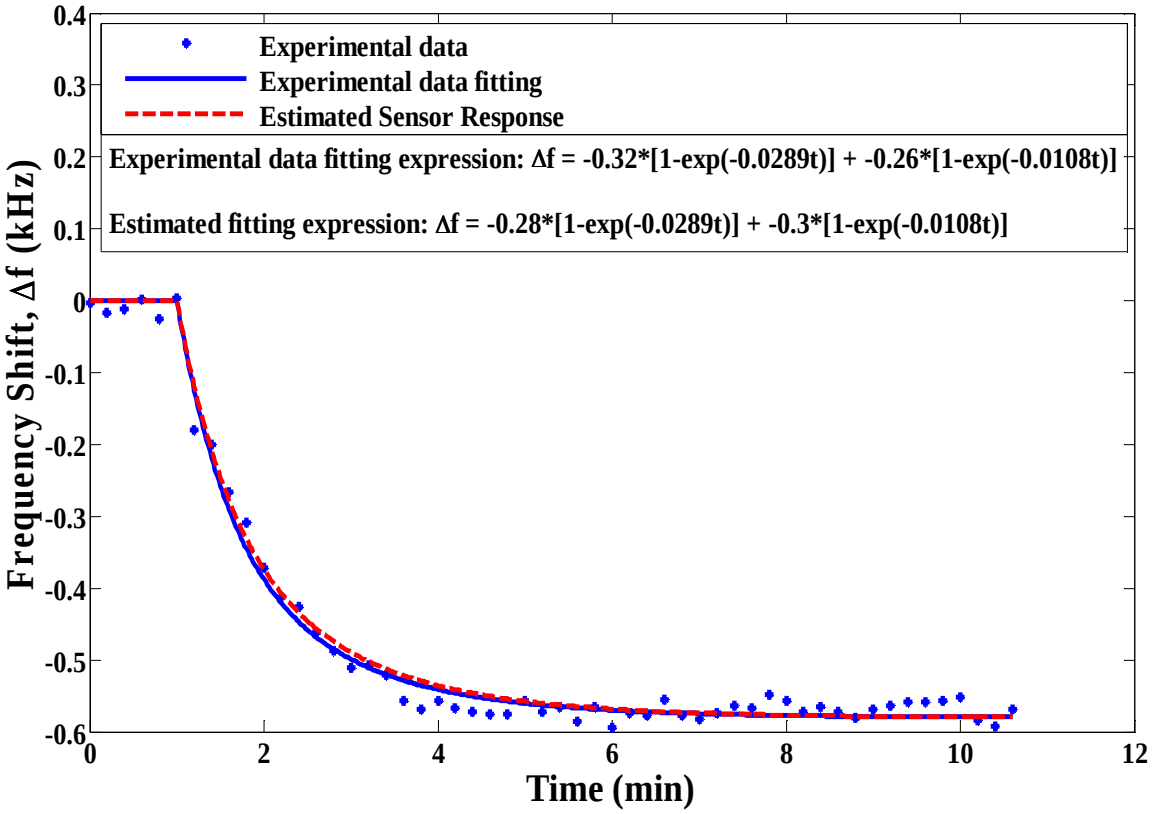

Figure 5.15: Response of a SH-SAW sensor coated with 0.6µm PECH to a binary mixture of 1000 ppb benzene and 500 ppb toluene (blue curve) along with the estimated sensor response using the nonlinear model of two-analyte system (red curve). Also shown in the figure are the estimated steady-state frequency shifts along with the steady-state frequency shifts determined by fitting the measurement data using dual-exponential fit.

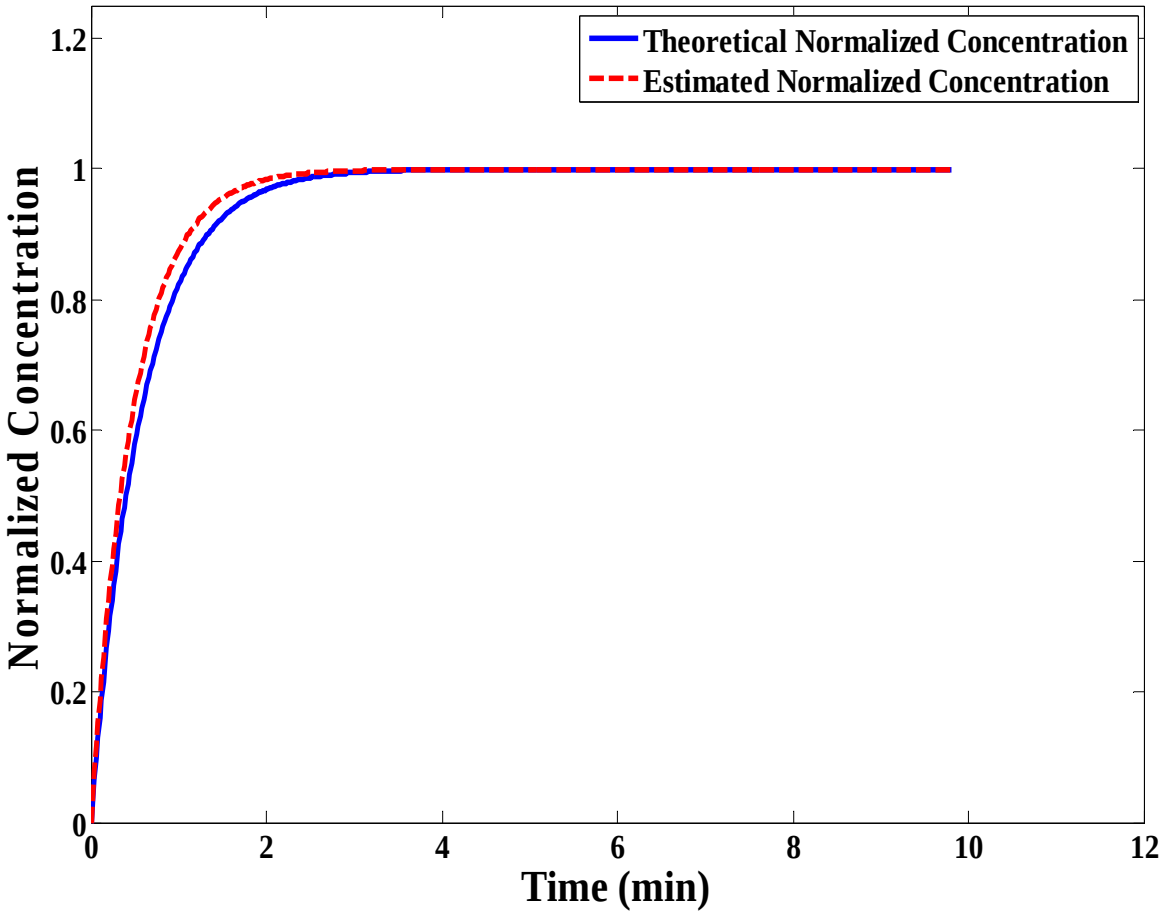

Figure 5.16: Estimated normalized concentration of benzene using the nonlinear model of two-analyte system co-plotted with the theoretical normalized concentration of benzene for 0.6µm PECH.

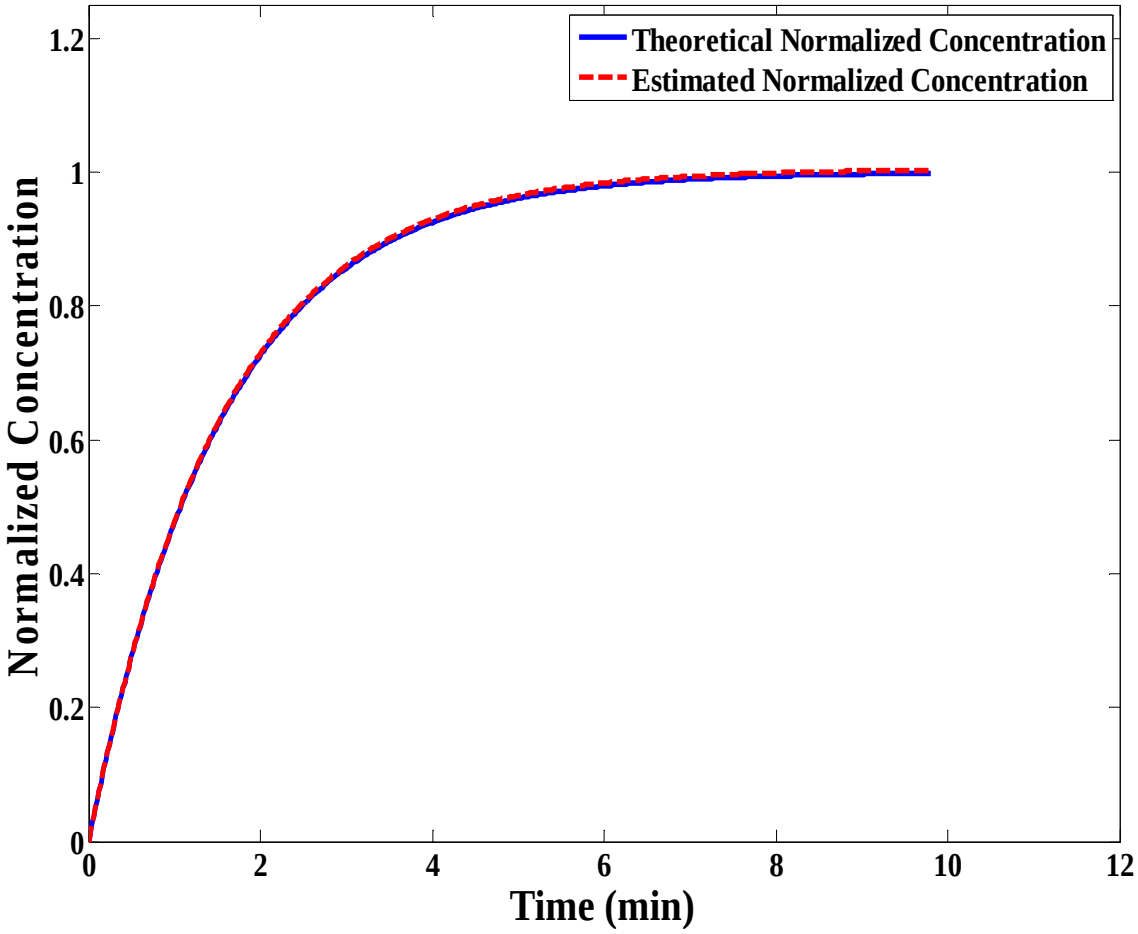

Figure 5.17: Estimated normalized concentration of toluene using the nonlinear model of two-analyte system co-plotted with the theoretical normalized concentration of toluene for 0.6µm PECH.

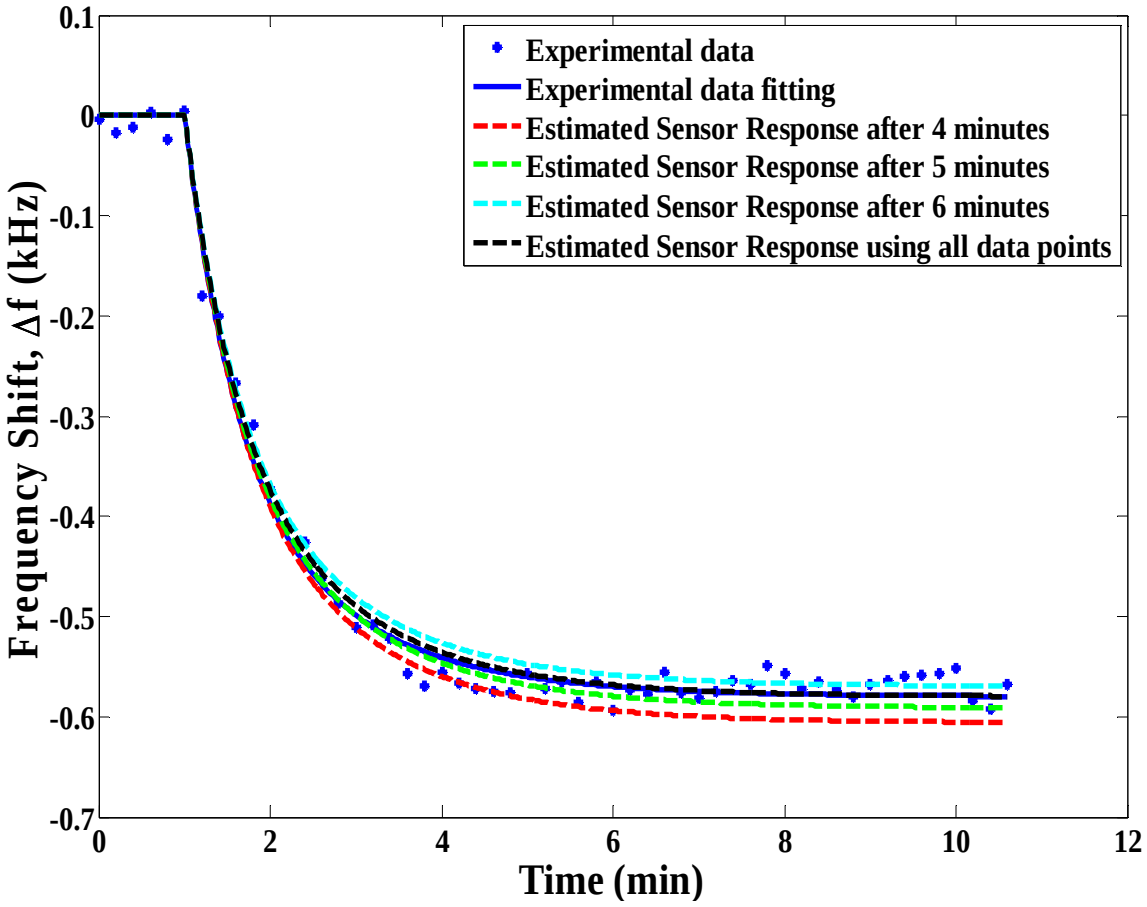

Figure 5.18: Estimated sensor response to a binary mixture of 1000 ppb benzene and 500 ppb toluene (with 0.6μm PECH coating) using the nonlinear model of two-analyte system obtained using the measurement data collected for the first 4, 5 and 6 minutes after the binary mixture sample has been introduced to the sensor co-plotted together with the measurement data, measurement data fitting and also the estimated sensor response using all the data points.

Estimated Steady-State Frequency Shift (Measurement Data 2)	Steady-State Frequency Shift of Benzene, α_{est} (kHz) (% difference with α_{fit})	Steady-State Frequency Shift of Toluene, α_{est} (kHz) (% difference with α_{fit})
After 4 minutes	**-0.2936 (8.25 %)**	**-0.3125 (20.19 %)**
After 5 minutes	**-0.2869 (10.34 %)**	**-0.3045 (17.11 %)**
After 6 minutes	**-0.2763 (13.66 %)**	**-0.2935 (12.88 %)**
Using all data points	**-0.2843 (11.16 %)**	**-0.3015 (15.96 %)**

Table 5.6: Estimated steady-state frequency shifts for a binary mixture of 1000 ppb benzene and 500 ppb toluene (with 0.6µm PECH coating) using the nonlinear model of two-analyte system obtained using the measurement data collected for the first 4, 5 and 6 minutes after the binary mixture sample has been introduced to the sensor along with the estimated steady-state frequency shifts obtained using all the data points. Also given in the table are the percentage differences between the estimated steady-state frequency shifts and steady-state frequency shifts determined by fitting the measurement data.

As depicted in Fig. 5.15, the estimated steady-state frequency shift for each analyte for this second case is also in agreement (i.e. less than ± 15 % difference) with the steady-state frequency shift of the analytes determined by fitting the measurement data. The values of C_{fit} and C_{est} for both analytes in the sample are found to be equal to 1245 ppb and 1090 ppb, respectively, for the concentration of benzene and 349 ppb and 402 ppb, respectively, for the concentration of toluene. Note that, C_{fit} and C_{est} found for both analytes are within ± 15 % difference. It should be noted that C_{est} are closer to the actual

concentration values. Fig. 5.16 and Fig. 5.17 show the estimation results obtained for the normalized concentration of each analyte versus time along with its theoretical normalized concentration determined by using eq. 5.3. As can be seen from Fig. 5.16 and Fig. 5.17, both the estimated normalized concentration of each analyte (i.e. benzene and toluene) are in conformity with its theoretical normalized concentration. This further validates the estimation results obtained for this second case. Fig. 5.18 shows the estimated sensor responses and Table 5.6 shows the estimated steady-state frequency shifts obtained using the measurement data collected for the first 4, 5 and 6 minutes after the binary mixture has been introduced to the sensor. Based on Table 5.6 and Fig. 5.18, it can be seen that the estimated sensor response and sensor parameters obtained using the measurement data collected for the first 5 minutes (and above) agree well (i.e. less than $\pm 20\%$ difference) with the measurement data and measurement data fitting. Therefore, again it has been shown that the steady-state frequency shifts of the analytes could be estimated even before the sensor response reaches steady-state using the nonlinear model of two-analyte system.

5.3.2 Linear Model

For the linear model, there are only two unknown quantities that were estimated which are the steady-state frequency shifts of each analyte. The time constants of each analyte are not estimated because it is assumed that they are known from the single analyte measurement. As stated earlier, for this present work the average time constants for various coating/analyte pairs given in Table 4.2 (in chapter 4) were used as the time

constant of the corresponding analyte. Similar to the previous section, only the estimation results obtained from two different measurement data (i.e. two different cases) will be presented and discussed. In fact, the same measurement data were used so that the efficiency of both models can be compared. For these two cases, two types of estimation result figures will be presented and discussed. The first figure will contain the information about the estimated sensor response, estimated sensor parameters, measurement data, and measurement data fitting (with the fitting parameters). Note that the two-analyte measurement data were fitted using dual-exponential fits. The second figure will show several estimated sensor responses determined after a certain number of minutes plotted along with the measurement data and measurement data fitting. Following this figure is a table which shows the estimated steady-state frequency shifts corresponding to the estimated sensor responses shown in the second figure along with the percentage difference between these estimated steady-state frequency shifts and the steady-state frequency shifts obtained from experimental data fitting. This is done to determine the minimum time required to obtain a good estimate of the sensor parameters.

First, the estimation results obtained by using the measurement data of the sensor response of SH-SAW sensor coated with 1.0μm PEA to binary mixture of 500 ppb benzene and 200 ppb ethylbenzene are discussed. Fig. 5.19, Fig. 5.20 and Table 5.7 show the estimation results obtained by using this binary mixture data and the linear model of the two-analyte system. As can be seen from Fig. 5.19, the estimated steady-state frequency shift for each analyte is in agreement (i.e. less than ± 6 % difference) with the steady-state frequency shift of the analytes determined by fitting the measurement data.

Moreover, by using eq. 5.1 and eq. 5.2, C_{fit} and C_{est} for both analytes are found to be around 285 ppb and 303 ppb, respectively for the concentration of benzene and 280 ppb and 277 ppb, respectively for the concentration of ethylbenzene. Therefore, C_{fit} and C_{est} found for both analytes are within ± 6 % difference. Fig. 5.20 shows the estimated sensor responses and Table 5.7 shows the estimated sensor parameters obtained using the measurement data collected for the first 4, 5 and 6 minutes after the binary mixture has been introduced to the sensor. Based on Table 5.7 and Fig. 5.20, it can be seen that the estimated sensor response and steady-state frequency shifts obtained using the measurement data collected for the first 6 minutes (and above) agree well (i.e. less than $\pm 20\%$ difference) with the measurement data and measurement data fitting. Therefore, the steady-state frequency shifts of the analytes could be estimated even before the sensor response reaches steady-state and thus, the analytes in the binary mixture can be quantified rapidly by using the linear model of the two-analyte system.

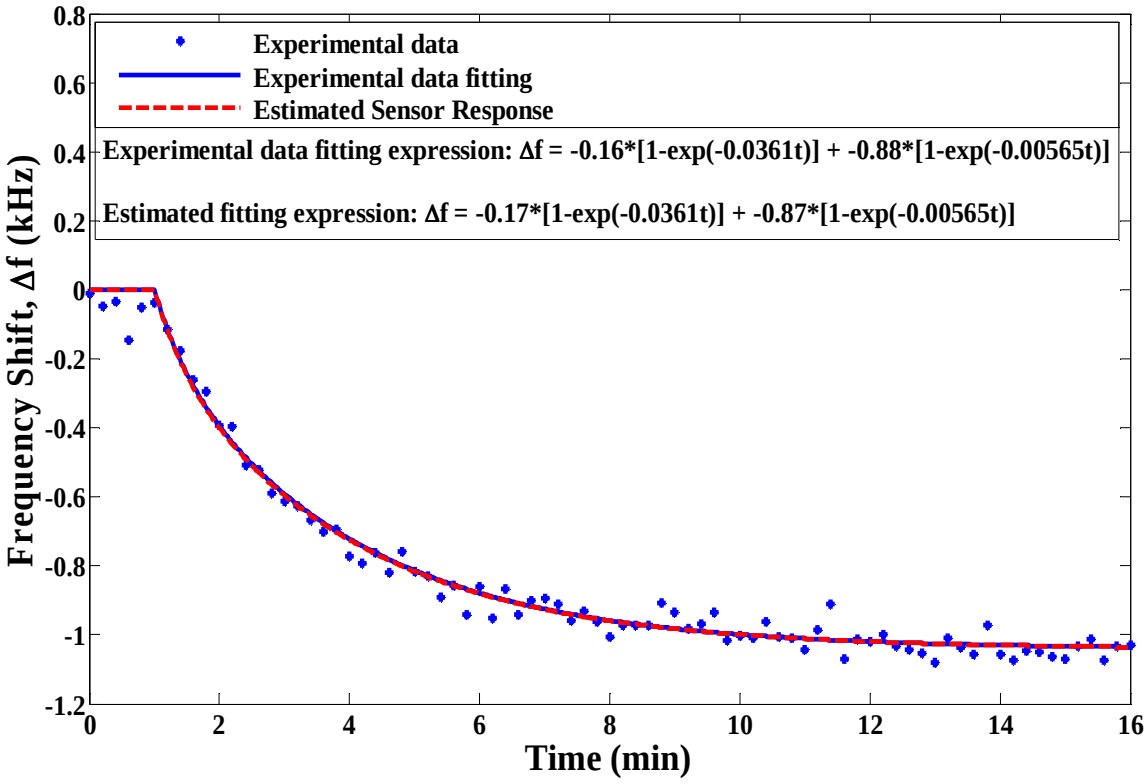

Figure 5.19: Response of a SH-SAW sensor coated with 1.0μm PEA to a binary mixture of 500 ppb benzene and 200 ppb ethylbenzene (blue curve) along with the estimated sensor response using the linear model of the two-analyte system (red curve). Also shown in the figure are the estimated steady-state frequency shifts along with the steady-state frequency shifts determined by fitting the measurement data using dual-exponential fit.

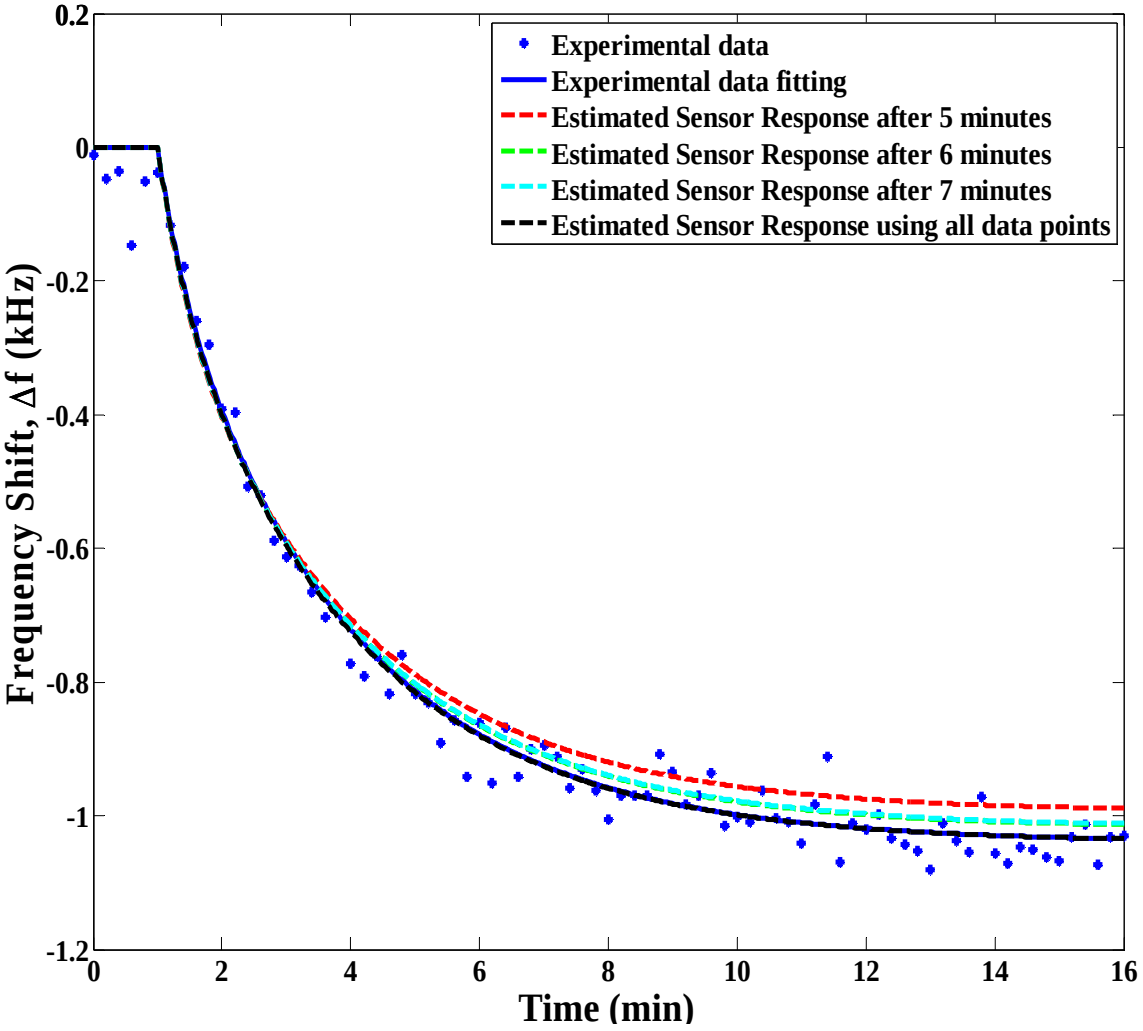

Figure 5.20: Estimated sensor response to a binary mixture of 500 ppb benzene and 200 ppb ethylbenzene (with 1.0µm PEA coating) using the linear model of two-analyte system obtained using the measurement data collected for the first 5, 6 and 7 minutes after the binary mixture sample has been introduced to the sensor co-plotted together with the measurement data, measurement data fitting and also the estimated sensor response using all the data points. Note that the green dashed line (estimated sensor response after 6 minutes) coincides with the light blue dashed line (estimated sensor response after 7 minutes).

Estimated Steady-State Frequency Shift (Measurement Data 1)	Steady-State Frequency Shift of Benzene, α_{est} (kHz) (% difference with α_{fit})	Steady-State Frequency Shift of Ethylbenzene, α_{est} (kHz) (% difference with α_{fit})
After 5 minutes	**-0.1973 (23.31 %)**	**-0.7968 (9.45 %)**
After 6 minutes	**-0.1817 (13.56 %)**	**-0.8372 (4.86 %)**
After 7 minutes	**-0.1829 (14.31 %)**	**-0.8343 (5.19 %)**
Using all data points	**-0.1652 (3.25 %)**	**-0.8688 (1.27 %)**

Table 5.7: Estimated steady-state frequency shifts for a mixture of 500 ppb benzene and 200 ppb ethylbenzene (with 1.0µm PEA coating) using the linear model of two-analyte system obtained using the measurement data collected for the first 5, 6 and 7 minutes after the binary mixture sample has been introduced to the sensor along with the estimated steady-state frequency shifts obtained using all the data points. Also given in the table are the percentage differences between the estimated steady-state frequency shifts and steady-state frequency shifts determined by fitting the measurement data.

For the second case, the estimation results obtained by using the measurement data of the sensor response of SH-SAW sensor coated with 0.6µm PECH to binary mixture of 1000 ppb benzene and 500 ppb toluene are discussed. Fig. 5.21, Fig. 5.22 and Table 5.8 show the estimation results obtained by using the linear model of the two-analyte system for this second measurement data.

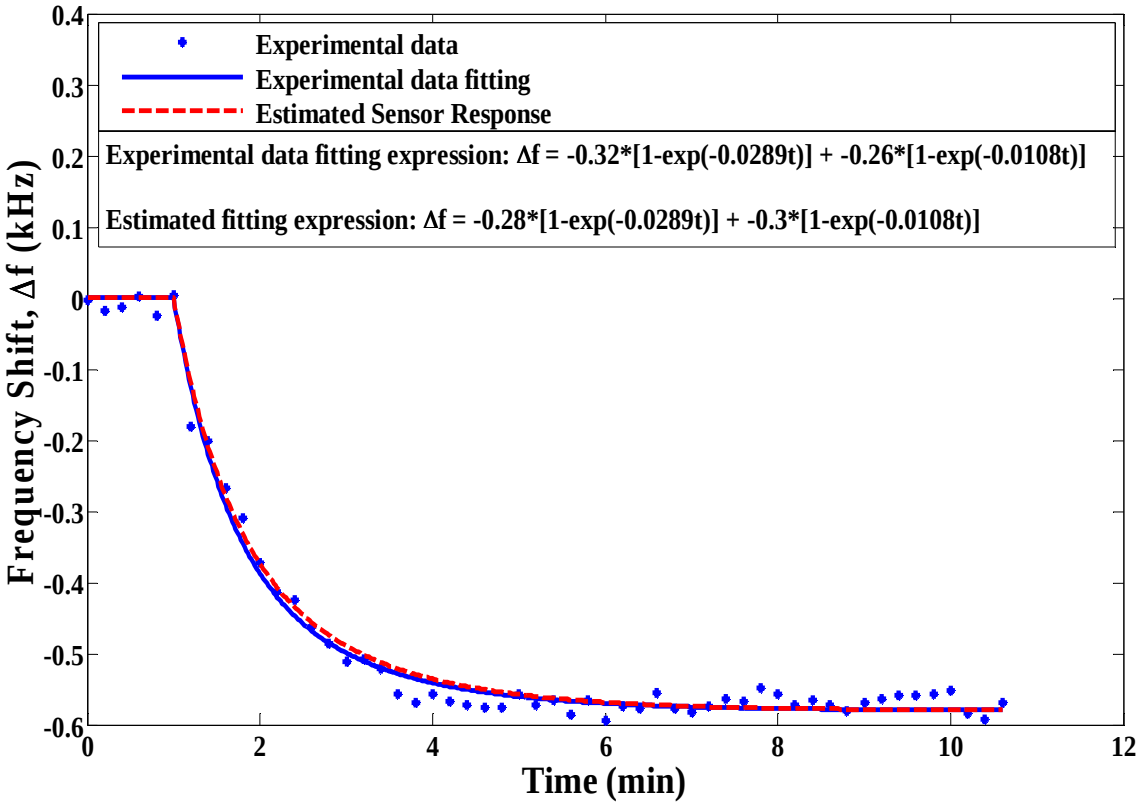

Figure 5.21: Response of a SH-SAW sensor coated with 0.6µm PECH to a binary mixture of 1000 ppb benzene and 500 ppb toluene (blue curve) along with the estimated sensor response using the linear model of the two-analyte system (red curve). Also shown in the figure are the estimated steady-state frequency shifts along with the steady-state frequency shifts determined by fitting the measurement data using dual-exponential fit.

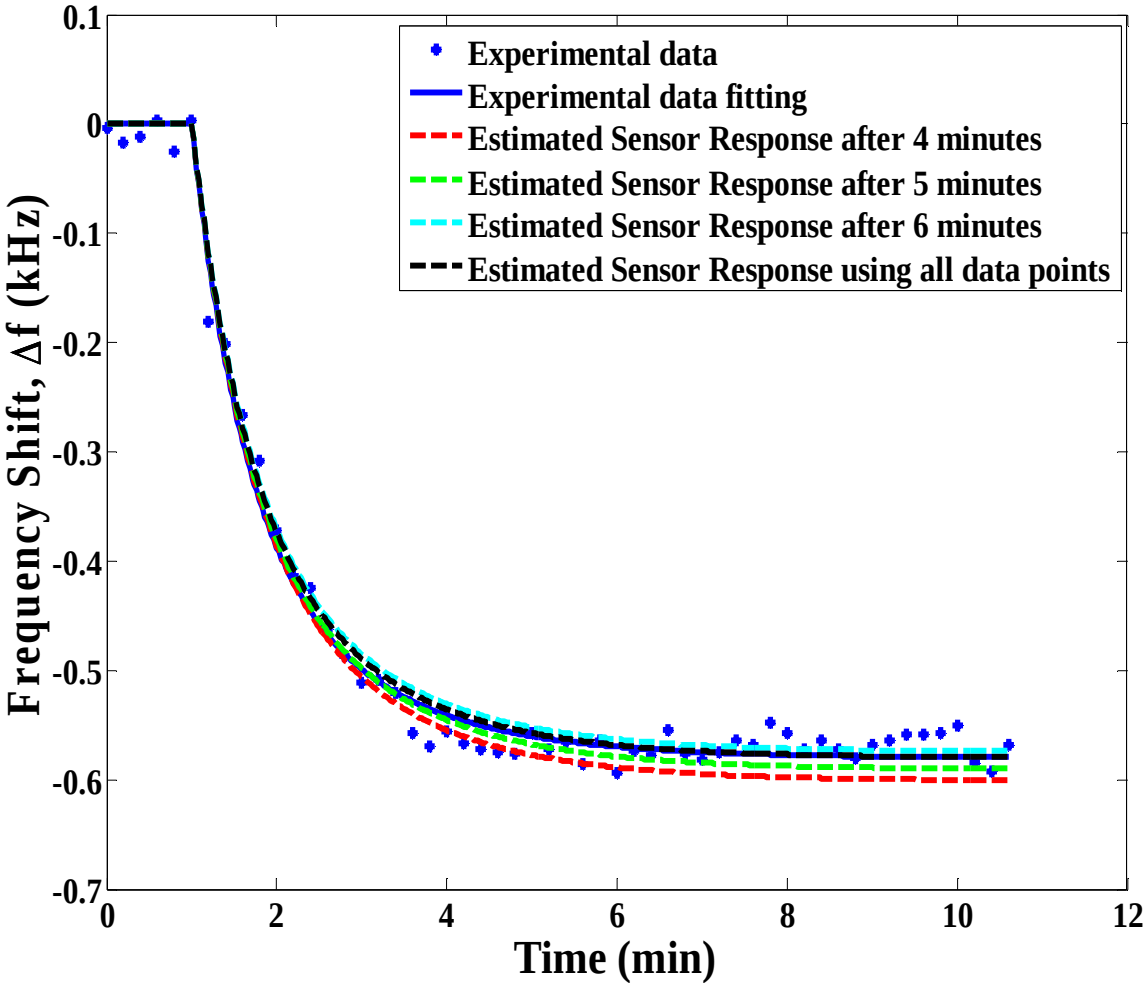

Figure 5.22: Estimated sensor response to a binary mixture of 1000 ppb benzene and 500 ppb toluene (with 0.6μm PECH coating) using the linear model of the two-analyte system obtained using the measurement data collected for the first 4, 5 and 6 minutes after the binary mixture sample has been introduced to the sensor co-plotted together with the measurement data, measurement data fitting and also the estimated sensor response using all the data points.

Estimated Steady-State Frequency Shift (Measurement Data 2)	Steady-State Frequency Shift of Benzene, α_{est} (kHz) (% difference with α_{fit})	Steady-State Frequency Shift of Ethylbenzene, α_{est} (kHz) (% difference with α_{fit})
After 4 minutes	**-0.2868 (10.38 %)**	**-0.3143 (20.88 %)**
After 5 minutes	**-0.2849 (10.97 %)**	**-0.3058 (17.62 %)**
After 6 minutes	**-0.2785 (12.97 %)**	**-0.2961 (13.88 %)**
Using all data points	**-0.2814 (12.06 %)**	**-0.2988 (14.92 %)**

Table 5.8: Estimated steady-state frequency shifts for a binary mixture of 1000 ppb benzene and 500 ppb toluene (with 0.6µm PECH coating) using the linear model of two-analyte system obtained using the measurement data collected for the first 4, 5 and 6 minutes after the binary mixture sample has been introduced to the sensor along with the estimated steady-state frequency shifts obtained using all the data points. Also given in the table are the percentage differences between the estimated steady-state frequency shifts and steady-state frequency shifts determined by fitting the measurement data.

As can be seen from Fig. 5.21, the estimated steady-state frequency shift for each analyte for this second case is also in agreement (i.e. less than ± 15 % difference) with the steady-state frequency shift of the analytes determined by fitting the measurement data. The values for C_{fit} and C_{est} for both analytes in the sample are found to be equal to 1245 ppb and 1090 ppb, respectively for the concentration of benzene and 349 ppb and 402 ppb, respectively for the concentration of toluene. Note that, C_{fit} and C_{est} found for both analytes are within ± 15 % difference. It should be noted that the values of C_{est} are closer

to the actual concentrations of the analytes. Fig. 5.22 shows the estimated sensor responses and Table 5.8 shows the estimated steady-state frequency shifts obtained using the measurement data collected for the first 4, 5 and 6 minutes after the binary mixture has been introduced to the sensor. Based on Table 5.8 and Fig. 5.22, it can be seen that the estimated sensor response and sensor parameters obtained using the measurement data collected for the first 5 minutes (and above) agree well (i.e. less than $\pm 20\%$ difference) with the measurement data and measurement data fitting. Therefore, again it has been shown that the steady-state frequency shifts of the analytes could be estimated even before the sensor response reaches steady-state using the linear model of the two-analyte system.

5.3.3 Summary on Two-Analyte Estimation Results

In this section, the summary of the two-analyte estimation results attained by using the nonlinear and linear model are presented. All estimation results obtained by using these two different models are summarized in Table 5.9 through Table 5.11. It should be noted that the observed discrepancies between C_{fit} and C_{est} with the actual concentrations can be attributed to the error in the sample preparations as well as volatility of the hydrocarbon analytes.

	Measurement Data 1 (500 ppb benzene + 200 ppb ethylbenzene) (1.0 μm PEA)				Measurement Data 2 (1000 ppb benzene + 500 ppb toluene) (0.6 μm PECH)			
	Benzene (kHz)		Ethylbenzene (kHz)		Benzene (kHz)		Toluene (kHz)	
	α_{fit}	α_{est}	α_{fit}	α_{est}	α_{fit}	α_{est}	α_{fit}	α_{est}
Nonlinear Model	**-0.16**	**-0.17 (6 %)**	**-0.88**	**-0.89 (1 %)**	**-0.32**	**-0.28 (13 %)**	**-0.26**	**-0.30 (15 %)**
Linear Model	**-0.16**	**-0.17 (6 %))**	**-0.88**	**-0.87 (1 %)**	**-0.32**	**-0.28 (13 %)**	**-0.26**	**-0.30 (15 %)**

Table 5.9: Estimated steady-state frequency shift, α_{est}, and steady-state frequency shift obtained by measurement data fit, α_{fit} for the two different measurement data. Note that the estimated steady-state frequency shift attained by using the two different models of the two-analyte system is presented. Given in parentheses are the percentage difference between the estimated steady-state frequency shifts and the steady-state frequency shifts obtained by measurement data fit.

	Measurement Data 1 (500 ppb benzene + 200 ppb ethylbenzene) (1.0 μm PEA)				Measurement Data 2 (1000 ppb benzene + 500 ppb toluene) (0.6 μm PECH)			
	Benzene (ppb)		Ethylbenzene (ppb)		Benzene (ppb)		Toluene (ppb)	
	C_{fit}	C_{est}	C_{fit}	C_{est}	C_{fit}	C_{est}	C_{fit}	C_{est}
Nonlinear Model	**285**	**303 (6 %)**	**280**	**283 (1 %)**	**1245**	**1090 (12 %)**	**349**	**402 (15 %)**
Linear Model	**285**	**303 (6 %)**	**280**	**277 (1 %)**	**1245**	**1090 (12 %)**	**349**	**402 (15 %)**

Table 5.10: Estimated concentration, C_{est}, and concentration determined from measurement data fit, C_{fit} for the two different measurement data. Note that the estimated concentration attained by using the two different models of the two-analyte system is given. Given in parentheses are the percentage difference between the estimated concentration and the concentration determined from measurement data fit.

	Measurement Data 1 (1000 ppb benzene + 1000 ppb ethylbenzene) (1.0 μm PEA)	**Measurement Data 2 (1000 ppb benzene + 500 ppb toluene) (0.6 μm PECH)**
	Minimum Estimation Time (minute)	**Minimum Estimation Time (minute)**
Nonlinear Model	**3**	**5**
Linear Model	**6**	**5**

Table 5.11: Minimum estimation time required to obtain a good estimate of sensor response (or parameters) for the two different models using the two different measurement data. Note that the results shown in the table are not absolute, and could be further improved by minimizing the measurement noise.

Based on Table 5.9 and Table 5.10, it can be seen that the estimated steady-state frequency shift and concentration of the analytes by using the nonlinear and linear model of the two-analyte system agree well (i.e. less than ± 15 % difference) with the steady-state frequency shift and concentration of analytes determined from fitting the measurement data. Therefore, it can be concluded that almost the same results can be attained either by using the estimation techniques for the binary mixture sample presented in this work or by fitting the measurement data. However, it should be noted that the advantage of using estimation theory to quantify the analytes lies in the fact that the analytes could be quantified well before the sensor-response reaches steady-state. From Table 5.11, it can be seen that by using the estimation techniques presented in this work,

the time required to quantify the analytes in the binary mixture is about 3-6 minutes which is less than half the time required for the sensor response to reach steady-state. As a result of using estimation theory, one could obtain an accurate estimate of the steady-state frequency shift of the analytes in the binary mixture even before the sensor response reaches steady-state. Note that from Tables 5.9 - 5.11, it can be seen that for the two-analyte system, the percentage difference and the minimum time required for estimation are larger than for the single analyte system; this is to be expected because quantifying the two analytes requires accurate evaluation of the deviations of the response to the analytes from a single exponential curve, and these deviations can be small if the two analytes have response time constants in the same order of magnitude.

Moreover, from the estimation results presented in Tables 5.9 - 5.11, it can be inferred that both nonlinear and linear model of the two-analyte system presented in this thesis perform equally well. Both the models are capable of producing estimates of the steady-state frequency shift well before the sensor response reaches steady-state. Although, both models performs equally well, the linear model is slightly better than the nonlinear model because, for the linear model, the estimation is performed by using KF which is an optimal filter (i.e. the convergence of the unknown quantities to their actual values are guaranteed if the system meets detectability criteria). For the linear model, there are only two unknown quantities that are being estimated. Another advantage of the linear model is that, its state space model can be readily extended to the case of multiple analytes (three or more analytes in the sample), so that the steady-state frequency shift of multiple analytes can be estimated.

6. SUMMARY, CONCLUSIONS AND FUTURE WORK

6.1 Summary

The objective of this work was to use estimation theory, in particular Kalman Filter (KF) and Extended Kalman Filter (EKF) to analyze and to quantify the analyte(s) in binary mixtures and single analyte samples of BTEX compounds (benzene, toluene, ethylbenzene and xylenes) in real-time. Also discussed in this work were the process of linear baseline drift correction using KF and the process of correcting for outlier points in the sensor data using a combination of discrete low pass filter and KF (or EKF depending on the model used). Note that both linear baseline drift correction and outlier points correction techniques presented in this thesis can be performed in real-time. Since KF and EKF are used extensively as a means toward signal processing in this thesis, the theory of KF and EKF were reviewed first. Under this review, the formulation of KF and EKF were presented. Also discussed were the algorithms on how to apply KF and EKF to estimate the unknown parameters.

Next, the models for the sensor responses to single and binary mixtures of analytes were discussed. For the case of the single analyte system, the sensor response model was developed by assuming that the single analyte system obeys Henry's law for concentrations of analyte below 50 ppm [7, 34, 35]. On the other hand, for the case of binary mixtures of analytes, the sensor response model was developed by first assuming Henry's law and by also assuming that the mixture obeys Fick's law of absorption which

states that when the mixture is extremely dilute, the sorption of one analyte into the polymer does not interfere with the sorption of the second analyte in any way. Free partitioning of the analyte between polymer and aqueous phase is assumed, implicating that the sorption process is reversible (i.e. only physisorption occurs). It is also assumed that the steady-state frequency shifts of each analyte in the binary mixture are additive. Moreover, since the sensor data considered in the present work are collected at discrete-time instants, the discrete-time model of the single analyte system and two-analyte system were found by using Euler's continuous time approximation formula. In order to apply estimation theory, the discrete-time version of the single analyte and two-analyte systems were transformed into the state-space form. For the case of the single analyte system, only one state-space model was developed and for the case of the two-analyte system, two different state-space models were developed where one is known as the nonlinear model because its state-space model is nonlinear and the other is known as linear model because its state-space model is linear. Note that the state-space models are dependent on the unknown parameters that need to be estimated. For the single analyte system, it is assumed that the normalized concentration of the analyte, steady-state frequency shift and time constant are unknown and based on these unknown quantities, the state-space model turns out to be nonlinear. Therefore, for the single analyte system, EKF algorithm has to be used to estimate the unknown parameters. For the nonlinear model of the two-analyte system, it is assumed that the normalized concentration of each analyte and the steady-state frequency shift of each analyte are unknown and must be estimated using EKF algorithm. For the linear model of the two-analyte system, it is assumed that the steady-state frequency shift of each analyte is unknown and must be

estimated using KF. Note that for the linear model of the two-analyte system, the normalized concentration of the analytes can be determined for each time instant by using the sensor response model and the known time constant of the analytes.

In order to show the validity of the estimation theory (in particular KF and EKF) to estimate the unknown parameters of the sensor response to single analyte samples and binary mixture samples, the proposed state-space models were tested on the measured data collected in the Microsensor Research Laboratory at Marquette University using a shear horizontal surface acoustic wave (SH-SAW) sensor coated with various chemically sensitive polymers. Before discussing the estimation results obtained for the single analyte system and the two-analyte system, the fundamentals of SH-SAW sensors and the process of data acquisition were first reviewed. Note that the types of polymers used to perform the experiments were also indicated. It has been noticed that the measured data collected in the lab exhibit a linear baseline drift and also in some cases, outlier points have been observed in the measured data (note that actual data collected in the field may also exhibit linear baseline drift and might produce some outlier points). Therefore, before using the measured data to perform the estimation process, the data have to be corrected for the linear baseline drift and the outlier points. In this work, it has been proposed to use a simplified technique of linear baseline drift correction using KF which is based on the baseline drift correction technique presented in [17]. For the correction of outlier points, a new technique has been proposed using a discrete low pass filter and KF (or EKF depending on the state-space model of the system). The data pre-processing techniques (i.e. linear baseline drift correction and the elimination of outlier points in the

measured data) were tested on the measured data and the results obtained are presented to prove the validity of the proposed technique. Also, it should be noted that since the proposed data pre-processing techniques uses estimation theory, the data pre-processing can be done in real time as the data are recorded.

Finally, the estimation results obtained for the single analyte system and for the two-analyte system using both nonlinear and linear model were presented and discussed. Also for the two-analyte system, the performance of the two different models was compared. For all models (i.e. one single analyte state-space model and two different state-space models of the two-analyte system), the estimated sensor response was co-plotted with the measured data and measured data fit so that the estimated response can be readily compared to the measured response. Based on the estimation results obtained, all the estimated sensor responses for all models showed a good agreement with the actual sensor responses. Furthermore, the estimated sensor parameters were also in conformity with the sensor parameters determined by fitting the measurement data. It has also been shown that the sensor parameters could be estimated in less than half the time required for the sensor response to reach steady-state. Therefore, by using the estimation technique presented in this thesis, the analyte(s) could be quantified rapidly. Moreover, based on the estimation results obtained for the two-analyte system using both nonlinear and linear models, it has been shown that both models perform equally well and are capable of estimating the unknown parameters rapidly.

6.2 Conclusions

One state-space model for the single analyte system and two different state-space models for the two-analyte system were developed which enables estimation theory to be applied to estimate the unknown parameters of the sensor response to BTEX compounds using a SH-SAW sensor. Based on the accurate estimation results obtained by using these state-space models, it can be concluded that the models developed were accurate mathematical representations of the systems (i.e. one analyte and two-analyte system).

Typically, the sensor response of analyte(s) (in particular BTEX compounds) may take several minutes in the liquid phase to reach steady-state. However, by using the estimation techniques demonstrated in this thesis, in particular KF (or EKF depending on the state-space model), the analyte(s) could be quantified rapidly. Therefore, using the estimation techniques, sensor parameters can be accurately estimated well before the sensor response reaches steady-state and this can significantly improve the time to quantification of the analyte(s). In this research, it has been shown that the time to quantification of the analyte(s) could be reduced to about less than half the time required for the sensor response to reach steady-state. This means that the concentration of the analyte(s) in the sample can be determined rapidly and based on this concentration level of the analyte(s) in the sample, mitigation plans could be carried out earlier. Shortening sensor exposure times may also improve accuracy, repeatability, and coating longevity.

This research also demonstrated new data pre-processing techniques which are capable of correcting for linear baseline drift and outlier points in the measurement data. The linear baseline drift correction technique is based on KF so that the baseline drift can be corrected in real-time. Normally, the raw measured data collected in the field for the sensor responses to BTEX compounds will exhibit linear baseline drift as well as a rapid response time. Therefore, the linear baseline drift correction technique demonstrated could be employed to correct the linear baseline drift in real-time. The outlier points correction technique is based on a combination of discrete low pass filter and KF (or EKF) so that the outlier points can be eliminated rapidly (in real-time) as soon as these outlier points are detected in the measured data. Note that the outlier points have to be removed from the measured data to obtain an accurate estimate of the sensor parameters. Outlier points will be recorded if the measurement noise is very high, sometimes during the start of a new measurement, or if any changes in the boundary conditions at the device surface occur. Therefore, the outlier points correction technique demonstrated could be employed to filter out the outlier points in measured data in real-time.

Since all the sensor signal processing techniques presented in this thesis could be performed in real-time, these techniques can be used in real world applications to rapidly quantify the analyte(s) in samples. All the signal processing techniques presented can be implemented in a single microcontroller or in a smart sensor system. Such a sensor system will have the capability of correcting the measured data for any linear baseline drift, correcting the measured data for any outlier points and, at the same time, use the corrected measured data to estimate the sensor parameters rapidly (in real-time).

Moreover, such a compact sensor signal processing system will reduce the cost of chemical sensing as the data could be processed immediately in the field with minimum manpower. There are various potential applications for a compact sensor signal processing system. One such possible application will be in monitoring of spill clean-ups. In this case, the level of remaining contamination should be regularly checked and by using a compact sensor signal processing system, the groundwater contamination level could be monitored remotely and the result could be transmitted to the company which performs the clean-up. Other applications could be the legally required periodic groundwater monitoring around underground storage tanks, or the monitoring of the plume in a sub-surface marine oil spill [46].

Finally, it should be pointed out that the sensor signal processing techniques presented in this thesis can be generalized to any type of chemical sensor platforms used to detect single analytes or binary mixtures of analytes, and are not specific to the SH-SAW sensor platform. The techniques should work equally well on sensor data collected using other sensor platforms such as microcantilever-based sensors, optical chemical sensors and other acoustic wave-based sensors.

6.3 Future Work

The work presented in thesis could be expanded in many ways. Further improvements in the sensor signal processing may be possible. In this section, a few possible future research proposals are listed.

(a)

In this thesis, the state-space models for the two-analyte system were developed by assuming that the time constant of each analyte in the binary mixture are known. However in some applications, the time constants of each analyte in the mixture will be unknown. Therefore, as a possible extension, this assumption could be relaxed in order to obtain a more general state-space model for the two-analyte system so that the state-space model obtained can be used in more applications. If this extension were made, one does not have to know the time constants of each analyte in the binary mixture beforehand as the generalized state-space model can be used to estimate the time constant of each analyte together with its steady-state frequency shift. This general state-space model also could be used to perform estimation on any unknown binary mixture sample.

(b)

In the present work, only the case of quantification of single analyte and two analytes in a sample using estimation theory were considered. However, many real world applications require the ability of a sensor system to quantify several analytes in a complex mixture. Therefore, this work could be naturally expanded to the case of quantification of multiple analytes in a complex mixture (i.e. more than two analytes). In particular, the linear model of the two-analyte system presented in this thesis could be easily modified so that the multiple analytes in a mixture could be quantified.

(c)

As mentioned in the problem statement (chapter 1) of this work, the ultimate goal of this research is to quantify benzene in groundwater samples which typically contain mixtures of multiple analytes. As a first step towards this goal, in this thesis it has been shown that benzene in a binary mixture sample could be quantified. The case considered could be expanded so that benzene in multiple analyte samples could also be quantified (i.e. benzene in mixtures of three or more analytes could be quantified). Therefore, future work could consist of extracting only benzene data from the sensor response of samples containing multiple analytes.

(d)

The sensor signal processing presented focused mainly on the quantification of the analyte(s) in a sample. Another aspect of sensor signal processing which is the identification of the analyte(s) in the sample could also be explored. Generally it is of interest to identify as well as quantify the analyte(s) that are present in a sample. Quantification aspect of the analyte(s) in the sample can be performed by using the techniques presented in this thesis. As for the identification of the analyte(s) in an unknown sample, a new approach using estimation theory could be investigated so that both identification and quantification of the analyte(s) could be performed rapidly. In order to perform identification, a sensor array with different coatings had to be used so that the target analyte could be identified more accurately. The use of estimation theory

in the identification process will enable the responses of these multiple sensors to be processed simultaneously and by using pattern recognition, the target analyte can be identified in real-time.

(e)

The sensor signal processing techniques presented in this thesis uses estimation theory in particular, Kalman Filter (KF) and Extended Kalman Filter (EKF) to quantify the analytes in real time. Besides KF and EKF, there are some other estimation theories that could have been used to perform the sensor signal processing. Therefore, as a future work in this area, one could investigate the feasibility of using other estimation theories to quantify the analytes. A comparison of the various approaches will then determine which of these estimation theories is best suited for the quantification of the analyte(s) in a given sample.

REFERENCES

[1] A.Hulanicki, S.Glab and F.Ingman, "*Chemical Sensors Definitions and Classification*," Pure & Appl. Chem., Vol. 63, No.9, pp. 1247-1250, 1991.

[2] A.Lobnik, M.Turel, and S.K. Urek, "*Optical Chemical Sensors: Design and Applications*," in Advances in Chemical Sensors, W.Wen, Ed. ISBN: 978-953-307-792-5, InTech, DOI: 10.5772/31534. [Online] 2012, http://www.intechopen.com/books/advances-in-chemical-sensors/optical-chemical-sensors-design-and-applications (Accessed: 14 September 2013).

[3] N.Lavrik, M.J.Sepaniak, and P.G. Datskos, "*Cantilever transducers as a platform for chemical and biological sensors*," Review of Scientific Instruments. Vol. 75, no.7, pp. 2229-2253, July 2004.

[4] M.J. Wenzel, "Modelling the Transient Response of Microcantilever Sensors and Analyte Classification Using Estimation Theory," M.S. Thesis, Marquette University, Milwaukee, WI, U.S.A, Nov 2006.

[5] S.Shah, "*Principal Component Analysis of APM Sensor Signals for Identification of Dilute Metal Ion solutions*," M.S. Thesis, Marquette University, Milwaukee, WI, U.S.A, May 1999.

[6] D.S. Ballantine, R.M. White, S.J. Martin, *et al*., *Acoustic Wave Sensors, Theory, Design, and Physico-Chemical Applications*, Academic Press, 1997.

[7] F.Josse, *et al*., "*Quantification of Benzene in Groundwater using SH-Surface Acoustic Wave sensors*," in The 14th International Meeting on Chemical Sensors, Nuremberg, Germany, pp. 473-476, 2012.

[8] J.S. Arey, P.M. Gschwend, Estimating Partition Coefficients for Fuel-Water Systems: Developing Linear Solvation Energy Relationships Using Linear Solvent Strength Theory To Handle Mixtures, *Environ. Sci. Technol.* 39, pp. 2702-2710.

[9] U.S. Department of Health and Human Services, "*Toxicological Profile for Benzene*," pp. 1-377, 2007.

[10] U.S. Environmental Protection Agency, FY 2011 Annual Report On The Underground Storage Tank Program (2012); EPA-510-R-12-001, pp.1-6, March 2012.

[11] American Cancer Society, “Benzene,” Internet: http://www.cancer.org/cancer/cancercauses/othercarcinogens/intheworkplace/benzene , [September 26, 2013].

[12] Xiong, Weizhen and Kale, Girish M., “*Novel high-selectivity NO2 sensor incorporating mixed-oxide electrode*” *Sensors and Actuators B,* vol. 114, pp. 101-108, 2006.

[13] Price, Gareth J. and Drake, Philip L., “*Potassium selective quartz crystal microbalance chemical sensors using functionalized copolymer coatings*” *Sensors and Actuators B,* vol. 114, pp. 466-472, 2006.

[14] Albert, Keith J., Lewis, Nathan S., Schauer, Caroline L., Sotzing, Gregory A., Stitzel, Shannon E., Vaid, Thomas P., and Walt, David R. “*Cross-Reactive Chemical Sensor Arrays*” *Chemical Reviews,* vol. 100, pp. 2595-2626, 2000.

[15] M.J. Wenzel, “*Polymer-Coated and Polymer-Based Microcantilever Chemical Sensors: Analysis and Sensor Signal Processing*,” Ph.D. Dissertation, Marquette University, Milwaukee, WI, U.S.A, 2009.

[16] J. W. Gardner and P. N. Bartlett, “*Electronic Noses: Principles and Applications*,” (Oxford University Press, New York, 1999).

[17] M.J. Wenzel, A.M. Brown, F.Josse, and E.Yaz, “*Online Drift Compensation for Chemical Sensors Using Estimation Theory*,” IEEE Sensors Journal, Vol.11, No. 1, pp.225-232, January 2011.

[18] A.Bermak, S.B.Belhouari, S.Minghua, D.Martinez, “*Pattern Recognition Techniques for Odor Discrimination in Gas Sensor Array*,” Encyclopedia of Sensors, Vol. X, pp.1-17, 2006.

[19] S.B.Belhouari, A.Bermak, W.Guangfen, P.C.H.Chan, “*On the Use of the Transient Information for Gas Identification using Microelectronic Gas Sensor*,” EEE Department, Hong Kong University of Science and Technology, Kowloon, Hong Kong, pp. 693-696, 2004.

[20] K.B.Jagler, “*Wavelet Signal Processing for Transient Feature Extraction*,” prepared for Air Force Office of Scientific Research, pp. 1-79, March 15, 1992.

[21] U.S. Environmental Protection Agency, “*Basic Information about Benzene in Drinking Water*,” 21 5 2012. [Online]. Available: http://water.epa.gov/drink/contaminants/basicinformation/benzene.cfm [Accessed 9/12/2012].

[22] S.M.Kay, “*Fundamentals of Statistical Signal Processing: Estimation Theory*,” Upper Saddle River, NJ: Prentice Hall, 1993.

[23] T.Roberto (2005, August 30), "*Estimation Theory for Engineers*," Available: http://www.ee.uwa.edu.au/~roberto/teach/Estimation_Theory.pdf [Accessed: September 20, 2013].

[24] W.Greg, B.Gary (2001), "*An Introduction to the Kalman Filter"*, University of North Carolina, Chapel Hill, NC Available: http://www.cs.unc.edu/~welch [September 21, 2012].

[25] Kalman, R. E., "*A New Approach to Linear Filtering and Prediction Problems*" *Transactions of the ASME - Journal of Basic Engineering*, vol. 82, pp. 35-45, 1960.

[26] E. Costa, J. do Val, and M.Fragoso, "A New Approach to Detectability of Discrete-Time Infinite Markov Jump Linear Systems," SIAM J. Control Optim., Vol. 43, No. 6, pp. 2132-2156, 2005.

[27] E. Yaz, EECE 6340, Class Lecture, Topic: "*Kalman Filter*," Faculty of Electrical and Computer Engineering, Marquette University, Milwaukee, WI, February 2013.

[28] R.G. Brown, P.Y.C. Hwang, "*Introduction to Random Signals and Applied Kalman Filtering*," 2nd Ed. New York: John Wiley and Sons, Inc, 1992.

[29] E. Yaz, EECE 6340, Class Lecture, Topic: "*Extended Kalman Filter*," Faculty of Electrical and Computer Engineering, Marquette University, Milwaukee, WI, April 2013.

[30] D. Simon, *Optimal State Estimation: Kalman, H∞, and Nonlinear Approaches*, (Wiley, New York, 2006).

[31] K. Reif, S. Günther, E. Yaz, and R. Unbehauen, "Stochastic stability of the discrete-time extended Kalman filter," *IEEE Transactions on Automatic Control*, vol. 44 (4), pp. 714-728, April 1999.

[32] T.Dursun, "*Kalman Filter and Its Applications in Navigation*," M.S. Thesis, California State University, Northridge, CA, December 2012.

[33] A.Michael, "*The Importance of Kalman Filtering Methods for Economic Systems*" in Annals of Economic and Social Measurement, Volume 3, number 1, 1974, pp.49-64 Available: http://www.nber.org/chapters/c9994.pdf [September 22, 2013].

[34] F.Bender, F.Josse, A.J.Ricco. "*Influence of Ambient Parameters on the Response of Polymer-Coated SH-Surface Acoustic Wave Sensors to Aromatic Analytes in Liquid-Phase Detection*," Joint Conference of the IEEE International Frequency Control & The European Frequency & Time Forum, pp. 1-6, 2011.

[35] F.Bender, R.E.Mohler, A.J.Ricco, and F.Josse. “Identification and Quantification of Aqueous Aromatic Hydrocarbons Using SH-Surface Acoustic Wave Sensors,” *Analytical Chem.*, vol. 86, pp. 1794-1799, January 2014.

[36] F.Bender, F.Josse, R.E.Mohler, and A.J.Ricco. “*Design of SH-Surface Acoustic Wave Sensors for Detection of ppb Concentrations of BTEX in Water*,” Joint UFFC, EFTF and PFM Symposium, pp. 628-631, 2013.

[37] F.Josse, F.Bender, and R.W.Cernosek. “Guided Shear Horizontal Surface Acoustic Wave Sensors for Chemical and Biochemical Detection in Liquids,” *Analytical Chem.*, vol. 73. No. 24, pp. 5937-5944, Dec 2001.

[38] R. Lenisa, “*Chemically Sensitive Polymer Coatings for SH-Surface Acoustic Wave Sensors for The Detection of Benzene in Water*,” M.S. Thesis, Marquette University, Milwaukee, WI, U.S.A, August 2013.

[39] D.L.Lee. “Analysis of energy trapping effects for SH-type waves on rotated Y-cut quartz,” *IEEE Trans. Son. Ultrason.*, SU-28, pp. 330-341, 1981.

[40] F.Josse and D.L.Lee. “Analysis of excitation, interaction and detection of surface and bulk acoustic waves on piezoelectric substrates,” *IEEE Trans. Son. And Ultrason.*, SU-29, pp. 262-273, 1982.

[41] T. Zhou, “*Theoretical Modeling of Acoustic Waves in Layered Structure Chemical Sonsors and Biosensors*,” M.S. Thesis, Marquette University, Milwaukee, WI, 1992.

[42] Z.H.Li, Y.Jones, J.Hossenlopp, R.Cernosek, and F.Josse, “Analysis of liquid-phase chemical detection using guided shear horizontal-surface acoustic wave sensors,” *Analytical Chem.*, vol. 77, No. 14, pp. 4595-4603, July 2005.

[43] A.K.Mensah-Brown, D.Mlambo, F.Josse, S.C.Schneider. “Analysis of the Detection of Organophosphate Pesticides in Aqueous Solutions Using Hydrogen-Bond Acidic Coating on SH-SAW Devices,” *IEEE Sensors Journal*, vol. 12, No. 5, pp. 893-903, May 2012.

[44] A.K.Mensah-Brown. “Analysis of the Detection of Organophosphate Pesticides in Aqueous Solutions Using Polymer-Coated SH-SAW Devices,” Ph.D. Dissertation, Marquette University, Milwaukee, WI, U.S.A, 2010.

[45] F.Haugen. “*Derivation of a Discrete-Time Lowpass Filter*,” TechTeach, Norwegian Institute of Technology, Trondheim, Norway, pp. 1-3, March 2008.

[46] R.N.Conmy, *et al.*, "*Submersible Optical Sensors Exposed to Chemically Dispersed Crude Oil: Wave Tank Simulations for Improved Oil Spill Monitoring*," Environ. Sci. Technol., pp. 1803-1810 February 2014.

APPENDIX A: ADDITIONAL OUTLIER POINTS CORRECTION RESULTS

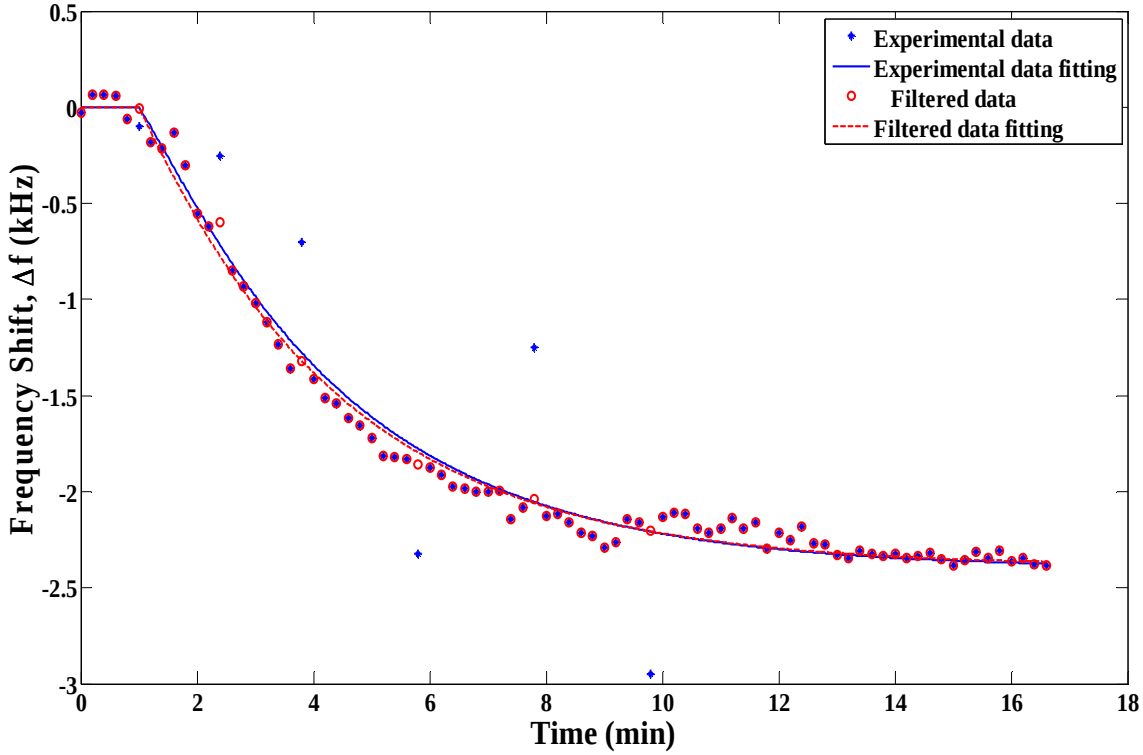

Figure A. 1: Outlier points corrected data co-plotted together with the measurement data with outlier points. The data are shown in two different colors where blue represents the measurement data with outlier points (unfiltered data) and red represents the outlier points corrected measurement data (filtered data). Both the data points and the curve fit for the data points are shown in the figure above.

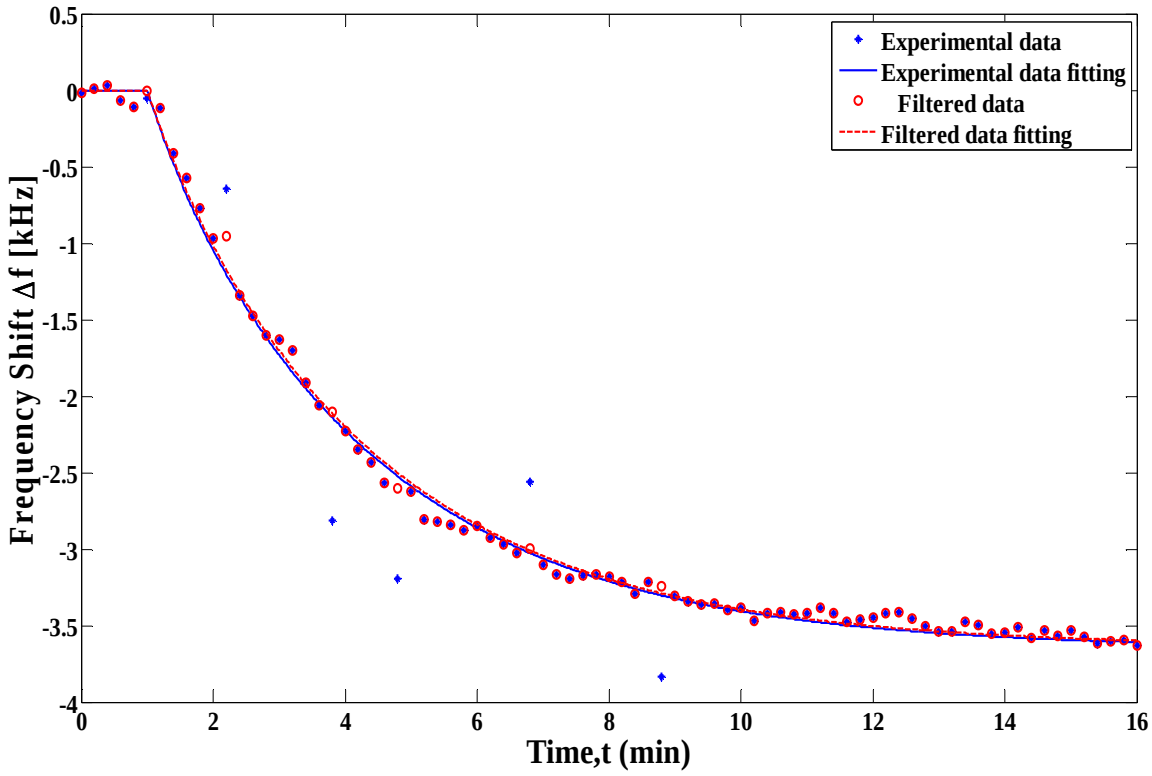

Figure A. 2: Outlier points corrected data co-plotted together with the measurement data with outlier points. The data are shown in two different colors where blue represents the measurement data with outlier points (unfiltered data) and red represents the outlier points corrected measurement data (filtered data). Both the data points and the curve fit for the data points are shown in the figure above.

APPENDIX B: MATLAB CODES

B.1 MATLAB Code for Outlier Points Filtering

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Author: KARTHICK SOTHIVELR
% File Name: outlier_correction_con.m
% Date (Created): 02-27-2013
% Date (Modified): 02-17-2014
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Description:
% Program to implement discrete first-order low pass filter to
eliminate
% the outlier points in the measurement data.
% Discrete Low-Pass Filter are implemented together with KF using the
% Linear Model of the Two-Analyte System.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Cleaning
clear all
close all
clc
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Open and read the measurement file
FID = fopen('op_thesis2.ini','r'); % 121211PEABEJ3binaryCor
data = textscan(FID,'%f %f');
fclose(FID);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Order the data (Based on Analyte)
y1 = data{2}(80:154); yb1 = data{2}(75:79);        % Benzene (800ppb)
y2 = data{2}(231:305); yb2 = data{2}(226:230);     % Benzene (200ppb)
% y3 = data{2}(375:456); yb3 = data{2}(370:374);     % Benzene (200ppb)
% y4 = data{2}(525:605); yb4 = data{2}(520:524);     % Benzene +
Ethylbenzene (500ppb + 1000ppb)

% Sampling Period:
T = 12; % in seconds

% Time Constants ( in sec) from Averaged Single Analyte Table:
tauA = 36.1+5; % Benzene
tauB = 204+1.5; % Ethylbenzene
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Select the case to be analyzed
y = y1; yb = yb1;
kmax = length(y); % Length of the measurement data points
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
ssA = 0.244; % Benzene
ssB = 2.24; % Toluene
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Introducing Outlier Points in the Data
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
y(8)=y(8)+ 0.5;
y(15)=y(15)+0.7;
y(25)=y(25)-0.5;
y(35)=y(35)+0.9;
y(45)=y(45)-0.8;
y(50)=y(50)-0.5;
y(70)=y(70)+0.7;
y_ori = y;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Discrete Low-Pass Filter Parameter
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
tau = tauB; % Select the time constant value (largest value)
alpha = T/(tau + T); diff=0;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Initialize Kalman Filter variables
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
yhat = zeros(1,kmax); x = zeros(2,kmax+1); P = zeros(2,2,kmax+1);
x(:,1) = [0; 0]; % Initial state vector (Initial state estimate)
P(:,:,1) = diag([1000, 1000]); % Initial error covariance Matrix
W = 10; V = 0; % Covariance of Measurement and process noise
G = [1]; % Matrix G (1 by 1 Matrix)
U = ones(kmax,1); % Step Input
m_A=0; m_B=0; % Initial value of the normalized Concentration
% System Matrices (A and F)
A = eye(2);
F = [1;1];
% Adsorption Rate Constant:
Sa=(T/tauA); % Benzene
Sb=(T/tauB); % Ethylbenzene

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%************** Correction of Outlier Points and KF Scheme
****************
% Loop to evaluate each instant in time
for i=1:kmax
    C = [m_A(i) m_B(i)];% C Matrix
    % Simulate the normalized concentration values:
    m_A = (1 - Sa)*m_A + Sa*U;
    m_B = (1 - Sb)*m_B + Sb*U;
    % Estimated measurement
```

```
    yhat(i)=C*x(:,i);


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    % Discrete Low-Pass Filter Scheme:
    if i==1
        y(i) = alpha*y(i);
    end
    if i>1
        diff = abs(y(i) - yhat(i));
    end
    if diff>=0.3
        y(i) = alpha*y(i) + (1-alpha)*y(i-1);
    end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

    K = (A*P(:,:,i)*C')/((C*P(:,:,i)*C' + G*W*G')) ;  % Kalman Gain
    x(:,i+1) = A*x(:,i) + K*(y(i)-C*x(:,i)); % State Update Equation
    % Error Covariance Update Equation:
    P(:,:,i+1) = (A-K*C)*P(:,:,i)*(A-K*C)' + K*G*W*G'*K' + F*V*F';
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Plot and Analysis
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
vk=0:kmax-1;
vt=(0:0.1:kmax-1)';

% Experimental Data Fitting
fun = ['a*(1-exp(' num2str(-(T/tauA)) '*x))+b*(1-exp(' num2str(-
(T/tauB)) '*x))'];
hfit = fit((0:kmax-1)',y_ori,fun,'StartPoint',[1 1]);
% Steady-State Sensitivity
a = round(hfit.a * 1e2)/1e2; % Analyte A
b = round(hfit.b * 1e2)/1e2; % Analyte B
vfit = a*(1-exp(-(T/tauA).*vt)) + b*(1-exp(-(T/tauB).*vt));

% Filtered Data Fitting
fun2 = ['fa*(1-exp(' num2str(-(T/tauA)) '*x))+fb*(1-exp(' num2str(-
(T/tauB)) '*x))'];
hfit2 = fit((0:kmax-1)',y,fun2,'StartPoint',[1 1]);
% Steady-State Sensitivity
fa = round(hfit2.fa * 1e2)/1e2; % Analyte A
fb = round(hfit2.fb * 1e2)/1e2; % Analyte B
vfil = fa*(1-exp(-(T/tauA).*vt)) + fb*(1-exp(-(T/tauB).*vt));

% Converting time step number to minutes and add baselines
y_ori = [yb(1:5);y_ori]; y=[yb(1:5);y];
vfit=[((0:0.1:4.9)'*0);vfit]; vfil=[((0:0.1:4.9)'*0);vfil];
vk=((0:length(y_ori)-1)')*(T/60); vt=(0:0.1:length(y)-1)'*(T/60);

% Plot
```

```
figure(1)
% Plot of Frequency Shift vs Time step
h=plot(vk, y_ori, '*b',vt, vfit, '-b', vk, y, 'or', vt, vfil, '--r');
%title ('Frequency Shift vs Time','FontSize',24)
LineWidth = [1];
xlabel('Time,t (min)','FontSize',14); ylabel('Frequency Shift \Deltaf
(kHz)','FontSize',14)
h_legend = legend('Experimental data','Experimental data fitting', ...
        'Filtered data','Filtered data fitting');
set(h_legend,'FontSize',15);
set(h,'LineWidth',3)
axis([0 16 -3.5 1.5])
annotation('textbox', [0.2,0.4,0.1,0.1],...
           'String', ['Experimental data fitting expression: \Deltaf =
', ...
        num2str(a) '*[1-exp(-' num2str(round((1/tauA)*1e4)/1e4) 't)] +
' ...
        num2str(b) '*[1-exp(-' num2str(round((1/tauB)*1e4)/1e4)
't)]']);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%% End of Program
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

B.2 MATLAB Code for Linear Baseline Drift Correction

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Author: KARTHICK SOTHIVELR
% File Name: baseline_corr_li.m
% Date (Created): 02-27-2013
% Date (Modified): 02-17-2014
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Description:
% Program to implement baseline drift correction using Kalman Filter
% Linear Interpolation Using KF are used to correct for baseline drift
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Cleaning
clear all
close all
clc
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Open file and read the data from the file
FID = fopen('newset.ini','r'); % 110620Pibbett2
data = textscan(FID,'%f %f');
fclose(FID);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%****************** Specify the Number of Responses
**********************
N_single = 1;       % Total number of single analyte response
N_mixture = 3;      % Total number of binary mixture response
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Measurement Information:
T = 12; % Sampling Period (in seconds)
W  = 100; % Measurement noise variance
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%************************* Single Analyte
********************************
tau_res_single = [29.78]; % Time constant for the single analyte
responses

analyte_in_single = [650 714]; % Point of analyte in or the end point
analyte_out_single = [681]; % Point of analyte out
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%***************** Binary Mixture
****************************************
% Time constant for the binary mixture
tau_res_mixture = [29.78 82.88 179]; % Benzene, Toluene, Ethylbenzene

analyte_in_mixture = [199 353 503]; % Point of analyte in or the end
point
```

```
analyte_out_mixture = [253 403 556]; % Point of analyte out
%**********************************************************************
***
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%************************ Linear Interpolation Info
***********************
% Number of points that need to be averaged on baseline
avg_points = 15;
% Number of points to check to determine the best fit(Check Starting
Point)
points = 5;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%***************************** Counters
***********************************
ff=0; hh=0; gg=1; ll=0; bestfit=0; kk=1;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% ********************* Sorting the analytes
******************************
analyte_in = sort([analyte_in_single analyte_in_mixture]);
% Single Analyte Response will be analyzed first, then two analyte
response
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%************************** Reset the Counters
****************************
ll=0; bestfit=0; kk=1; pp=1;% Counters
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%************************ Binary Mixture Section
**************************
while (kk<=N_mixture)

  while (pp<=4)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %******************* Linear Interpolation Section
*********************
    % Baseline before the response
  yfir = data{2}(analyte_in_mixture(kk)-
avg_points:analyte_in_mixture(kk));
```

```
% Determine the baseline after the response
for jj=1:length(analyte_in) % To determine the second baseline
    if analyte_in(jj)>analyte_in_mixture(kk)
        break
    end
end

% Baseline after the response
ysec = data{2}(analyte_in(jj)-avg_points:analyte_in(jj));

% Response
y1 = data{2}(analyte_in_mixture(kk):analyte_out_mixture(kk));

% Average of baseline before the response
avg1 = sum(yfir)/length(yfir);
% Average of baseline after the response
avg2 = sum(ysec)/length(ysec);

% Slope of the baseline drift
b = (avg2-avg1)/(analyte_in(jj) - analyte_in_mixture(kk));
% y-intercept of the baseline drift
a = (avg1);

% Added Part
if kk==1
    change = 15;
else
    change = 10;
end

% Response to be baseline corrected
yy = data{2}(analyte_in_mixture(kk)-change:analyte_in(jj)-70);

% Section to identify and select the time constant of the analytes
tau=tau_res_mixture; % Sorption time constant
% Case of Benzene + Toluene
if (pp==1)
    tauA=tau(1); tauB=tau(2);
end
% Case of Benzene + Ethylbenzene
if (pp==2)
    tauA=tau(1); tauB=tau(3);
end
% Case of Benzene + Toluene
if (pp==3)
    tauA=tau(1); tauB=tau(2);
end
% Select the Correct Case (Based on Best Curve Fit)
if (pp==4)
    [rr,place]=min(err_store);
    if (place==1)
        tauA=tau(1); tauB=tau(2);
    end
    if (place==2)
        tauA=tau(1); tauB=tau(3);
```

```
        end
        if (place==3)
            tauA=tau(2); tauB=tau(3);
        end

        vk2 = 0:length(yy)-1;

        % Interpolate the Linear Baseline drift
        yfit = a + b*vk2;

        % Linear Baseline Drift Correction
        yy2 = yy - yfit';

        %If the starting point produces the best curve fit store the
result
        if (bestfit==1)
            for j=1:length(yy2) % Keep track of all correction
                hh=hh+1;
                yall(hh)=yy2(j);
                yall2(hh)=yy(j);
            end
        end
    end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %**********************Select the best curve
fit**********************
    if (bestfit==1)
        pp = pp+1;
        bestfit=2;
        analyte_in_mixture(kk)=ori;
    end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %*****************************EKF
Section*****************************
    % EKF to estimate the concentration/s.s. sensitivity
    y = y1; ymeas=y1;
    kmax = length(y);

    % Adsorption Rate
    Sa=(T/tauA); Sb=(T/tauB); % Analyte A and B

    % Start EKF Scheme
    x = zeros(4,kmax+1); P = zeros(4,4,kmax+1); yhat = zeros(1,kmax);
    % Initial state vector
    x(:,1) = [0; 0; 0; 0];
    % Initial error covariance Matrix
    P(:,:,1) = diag([0.01, 0.01, 25000, 100000]);
```

```
    V = 3*eye(2); % Covariance matrix of process noise v
    G = [1]; % Matrix G (dh/dw)  (1 by 1 Matrix)
    U = ones(kmax,1); % Step Input

    % Loop to evaluate each instant in time
    for i=1:kmax
    C = [x(3,i) x(4,i) x(1,i)  x(2,i)]; % C Matrix (dh/dx)
    h = [x(3,i)*x(1,i) + x(4,i)*x(2,i)]; % h Matrix
    y(i) = y(i) - a - b*(i-1); % Baseline Corrected Measurement

    %*****************Measurement Update
Equations**********************
    x1 = x(:,i); P1 = P(:,:,i);
    x(:,i+1) = x1 + P1*C'*(inv((C*P1)*C' + G*W*G')) * [y(i) - h];
    P(:,:,i+1) = P1 - P1*C'*(inv(C*P1*C' + G*W*G'))*C*P1;
    yhat(1,i) = h;

    %**********************Time Update
Equation*************************
    % f matrix
f = [(1-Sa)*x(1,i+1)+Sa*U(i); (1-Sb)*x(2,i+1)+Sb*U(i); x(3,i+1);
x(4,i+1)];
    % A matrix (df/dx)
    A = [1-Sa 0 0 0; 0 1-Sb 0 0; 0 0 1 0; 0 0 0 1];
    % F matrix (df/dv)
    F = [1 0;0 1;0 0;0 0];

    x(:,i+1) = f; % Time update the state estimate

    P1 = P(:,:,i+1);
    P(:,:,i+1) = A*P1*A' + F*V*F'; % Time update the error covariance
    end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %*********************** Analysis and Plot
*************************
    % Time vector
    vk=(0:kmax-1)'; vt=(0:0.1:kmax-1)';

    % Experimental Data Curve fitting to find S and alpha
    fun = ['a*(1-exp(' num2str(-Sa) '*x))+b*(1-exp(' num2str(-Sb)
'*x))'];
    hfit = fit((0:kmax-1)',y,fun,'StartPoint',[1 1]);

    % Steady-state sensitivity
    x03 = round(hfit.a * 1e4)/1e4; % Analyte A
    x04 = round(hfit.b * 1e4)/1e4; % Analyte B

    % Experimental Data Fitting
    vfit = x03*(1-exp(-Sa.*vt)) + x04*(1-exp(-Sb.*vt));
```

```
    % Estiamted Steady-state sensitivity (Analyte A)
    Ea = round(x(3,kmax)*1e4)/1e4;
    % Estimated Steady-state sensitivity (Analyte B)
    Eb = round(x(4,kmax)*1e4)/1e4;

    % Estimated Frequency Shift
    vest_test = Ea*(1-exp(-Sa.*vk)) + Eb*(1-exp(-Sb.*vk));


%**************************************************************************
    %~~~~~~~~~~~~~~~~~~~~~Section to find the
bestfit~~~~~~~~~~~~~~~~~~~~
    error = abs(y - vest_test);

    if bestfit==0
        ll = ll + 1;
        if (ll==1)
            ori=analyte_in_mixture(kk);
        end
        mean_err(ll) = mean(error);
        analyte_in_mixture(kk)=analyte_in_mixture(kk) + 1;

        if (ll==points)
            [err,index]=min(mean_err);
            analyte_in_mixture(kk)=ori + (index-1);
            bestfit=1;
            err_store(pp) = err;
        end
    end

%~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

%**************************************************************************


%**************************************************************************
    %~~~~~~ Identify the Analyte in the Mixture and Plot the Result
~~~~~~~
    if (pp<=4)&&(bestfit==2)
        bestfit=0;
        ll=0;
    end

    if (bestfit==2)

        if (place==1)
            fprintf('\nResponse %d :',(kk))
            fprintf(' Benzene & Toluene\n\n');
        end
        if (place==2)
            fprintf('\nResponse %d :',(kk))
            fprintf(' Benzene & Ethylbenzene\n\n');
        end
        if (place==3)
            fprintf('\nResponse %d :',(kk))
```

```
            fprintf(' Toluene & Ethylbenzene\n\n');
        end

        % Estimated Frequency Shift
        vest = Ea*(1-exp(-Sa.*vt)) + Eb*(1-exp(-Sb.*vt));

        % Add the baseline (First 5 points)
        ynew = [yy2(1:5);y]; yhatnew=[yy2(1:5);yhat'];
        vfit=[((0:0.1:4.9)'*0);vfit]; vest=[((0:0.1:4.9)'*0);vest];

        % Convert time to minutes
        vvk=((0:length(ynew)-1)')*(T/60);
        vvt=(0:0.1:length(ynew)-1)'*(T/60);

        % Plot of Frequency Shift vs Time
        figure(ff+1)
        plot(vvk, ynew, '*b', vvt, vfit, '-b', vvk, yhatnew, 'or', vvt,
vest, '--r')
        title ('Frequency Shift vs Time')
        xlabel('Time (min)'); ylabel('Frequency Shift \Deltaf [kHz]')
        legend('Experimental data',    ['Experimental data Fitting:
\Deltaf=', ...
            num2str(x03) '*[1-exp(-' num2str(Sa) 'k)] + ' ...
            num2str(x04) '*[1-exp(-' num2str(Sb) 'k)]'], ...
           'Direct Estimation', ['Estimated Expression: \Deltaf=' ...
           num2str(Ea) '*[1-exp(-' num2str(Sa) 'k)] + ' ...
           num2str(Eb) '*[1-exp(-' num2str(Sb) 'k)]']);

        % Update the counters
        ff=ff+1;
        kk=kk+1;
        bestfit=0;
        ll=0;
    end

%*********************************************************************
  end
  pp=1;
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%************************* Reset the Counters
**************************
ll=0; bestfit=0; kk=1; pp=1;% Counters
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%********************** Single Analyte Response
*************************
while (kk<=N_single)
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %****************** Linear Interpolation Section
*********************
    % Baseline before the response
    yfir = data{2}(analyte_in_single(kk)-
avg_points:analyte_in_single(kk));
    % Determine the baseline after the response
    for jj=1:length(analyte_in)
        if analyte_in(jj)>analyte_in_single(kk)
            break
        end
    end

    % Baseline after the response
    ysec = data{2}(analyte_in(jj)-avg_points:analyte_in(jj));

    % Response (Data points)
    y1 = data{2}(analyte_in_single(kk):analyte_out_single(kk));

    % Average of the points in baseline before the response
    avg1 = sum(yfir)/length(yfir);
    % Average of the points in baseline after the response
    avg2 = sum(ysec)/length(ysec);

    % Slope of the baseline drift
    b = (avg2-avg1)/(analyte_in(jj) - analyte_in_single(kk));
    % y-intercept of the baseline drift
    a = (avg1);

    % Determine the Response (Data points) to be baseline corrected
    yy = data{2}(analyte_in_single(kk)-10:analyte_in(jj)-10);

    vk2 = 0:length(yy)-1; % Specify the discrete-time instant
    % Linear drift
    yfit = a + b*vk2; % Linear Interpolation

    % Baseline Drift correction
    yy2 = yy - yfit';

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    % If the starting point produces the best curve fit store the
result
    if bestfit==1
        for j=1:length(yy2) % Keep track of all correction
            hh=hh+1;
            yall(hh)=yy2(j);
            yall2(hh)=yy(j);
        end
        bestfit=2;
    end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %**************************** EKF Section
*****************************
    % EKF to estimate the concentration/s.s. sen sensitivity
    y = y1; ymeas=y1;
    tau=tau_res_single(kk); % Sorption time constant
    kmax = length(y);

    x02 = T/tau; % S (Initial Estimate of Adsorption Rate)
    x03 = -1; % alpha (Initial Estimate of Steady-State Sensitivity)

    % Start EKF Scheme
    x = zeros(3,kmax+1); P = zeros(3,3,kmax+1); yhat = zeros(1,kmax);
    x(:,1) = [0; x02+0.1; x03]; % Initial state vector
    P(:,:,1) = diag([0.01, 100, 10000]); % Initial error covariance
Matrix

    V = [1]; % Covariance matrix of process noise v (1 by 1 Matrix)
    G = [1]; % Matrix G (dh/dw)  (1 by 1 Matrix)
    U = ones(kmax,1); % Step Input

    % Loop to evaluate each instant in time
    for i=1:kmax
        C = [x(3,i) 0 x(1,i)]; % C Matrix (dh/dx)
        h = [x(3,i)*x(1,i)]; % h Matrix
        y(i) = y(i) - a - b*(i-1); % Baseline Corrected Measurement

        %*****************Measurement Update
Equations*********************
        x1 = x(:,i); P1 = P(:,:,i);
        x(:,i+1) = x1 + P1*C'*(inv((C*P1)*C' + G*W*G')) * [y(i) - h];
        P(:,:,i+1) = P1 - P1*C'*(inv(C*P1*C' + G*W*G'))*C*P1;

        yhat(1,i) = h;

        %**********************Time Update
Equation************************
        % f matrix
        f = [(1-x(2,i+1))*x(1,i+1)+x(2,i+1)*U(i); x(2,i+1); x(3,i+1)];
        % The A matrix (df/dx)
        A = [1-x(2,i+1) U(i)-x(1,i+1) 0; 0 1 0; 0 0 1];
        % F matrix (df/dv)
        F = [1; 0; 0];

        x(:,i+1) = f; % Time update the state estimate
        P1 = P(:,:,i+1);
        P(:,:,i+1) = A*P1*A' + F*V*F'; % Time update the error
covariance
    end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %************************* Analysis and Plot
************************
    % Time vector
    vk=(0:kmax-1)'; vt=(0:0.1:kmax-1)';

    % Experimental Data Curve fitting to find S and alpha
    hfit = fit((0:kmax-1)',y,'a*(1-exp(-S*x))','StartPoint',[1 1]);
    x02 = round(hfit.S * 1e4)/1e4; % S
    x03 = round(hfit.a * 1e4)/1e4; % alpha
    vfit = x03*(1-exp(-x02.*vt)); % Experimental Data Fitting


%**************************************************************************
    % ~~~~~~~~~~~~~~~~~~~~Section to find the
bestfit~~~~~~~~~~~~~~~~~~~~~~
    vfit_test = x03*(1-exp(-x02.*vk));
    error = abs (y - vfit_test); % Error in the fit
    % Checking the Starting Point
    if bestfit==0
        ll = ll + 1;
        if ll==1
            ori=analyte_in_single(kk);
        end
        mean_err(ll) = mean(error);
        analyte_in_single(kk)=analyte_in_single(kk) + 1;

        if ll==points
            [err,index]=min(mean_err);
            analyte_in_single(kk)=ori + (index-1);
            bestfit=1;
        end
    end
    %
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

%**************************************************************************


%**************************************************************************
    % If the starting point produces best curve fit Plot the Results
    if (bestfit==2)
        ES = round(x(2,kmax)*1e4)/1e4; % Estimated adsorption rate (S)
        % Estimated Steady-state sensitivity (alpha)
        Ea = round(x(3,kmax)*1e4)/1e4;
        % Estimated Frequency Shift
        vest = Ea*(1-exp(-ES.*vt));

        % Adding some baseline
        ynew = [yy2(1:5);y]; yhatnew=[yy2(1:5);yhat']; % The first 5
points
        vfit=[((0:0.1:4.9)'*0);vfit]; vest=[((0:0.1:4.9)'*0);vest];
```

```
        % Convert time to minutes
        vvk=((0:length(ynew)-1)')*(T/60); vvt=(0:0.1:length(ynew)-
1)'*(T/60);

        % Plot the result
        figure(ff+1)
        % Plot of Frequency Shift vs Time step
        plot(vvk, ynew, '*b', vvt, vfit, '-b', vvk, yhatnew, 'or', vvt,
vest, '--r')
        title ('Frequency Shift vs Time')
        xlabel('Time (min)'); ylabel('Frequency Shift \Deltaf [kHz]')
        legend('Experimental data',    ['Experimental data Fitting:
\Deltaf=', ...
            num2str(x03) '*[1-exp(-' num2str(x02) 't)]'], ...
           'Direct Estimation', ['Estimated Expression: \Deltaf=' ...
           num2str(Ea) '*[1-exp(-' num2str(ES) 't)]']);

        % Update the Counters
        ff=ff+1;
        kk=kk+1;
        bestfit=0;
        ll=0;
    end

%*********************************************************************

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%********************** Baseline Corrected Plot
*************************
% Time vector
vk = 0:length(yall)-1; vk = vk.*(12/60);

% Plot the Baseline Corrected Result
figure(ff+1)
h=plot(vk,yall);
set(h,'LineWidth',1.5)
%title('Frequency Shift','FontSize',16)
xlabel('Time (min)','FontSize',16)
ylabel('Frequency Shift (kHz)','FontSize',16)
grid on
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%



figure(ff+2)
h=plot((data{1}(1:713)),(data{2}(1:713)));
set(h,'LineWidth',1.5)
%title('Frequency Shift of RAW Data (With Baseline
Drift)','FontSize',16)
xlabel('Time (min)','FontSize',16)
```

```
ylabel('Frequency Shift (kHz)','FontSize',16)
grid on

vk = 0:length(yall2)-1; vk = vk.*(12/60);
figure(ff+3)
h=plot(vk,yall2);
set(h,'LineWidth',1.5)
%title('Frequency Shift of RAW Data (With Baseline
Drift)','FontSize',16)
xlabel('Time (min)','FontSize',16)
ylabel('Frequency Shift (kHz)','FontSize',16)
grid on
axis([0 80 -0.4 1.4])
```

B.3 MATLAB Code for Single Analyte Estimation

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Author: KARTHICK SOTHIVELR
% File Name: Single_Analyte.m
% Date (Created): 02-27-2013
% Date (Modified): 02-17-2014
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Description:
% Program to estimate steady-state sensitivity, time constant and
% concentration of the analyte using the single analyte model.
Estimation
% were performed using Extended Kalman Filter.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Cleaning
clear all
close all
clc
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Open and read the measurement file
FID = fopen('BE_PEA_1st.ini','r'); % BE_PEA 1st (old one)
data = textscan(FID,'%f %f');
fclose(FID);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Order the data (Based on Analyte)
y1 = data{2}(80:155); yb1 = data{2}(75:79);        % Ethylbenzene
(1000ppb)
y2 = data{2}(229:306); yb2 = data{2}(224:228);     % Ethylbenzene
(1000ppb)
%y3 = data{2}(377:410); yb3 = data{2}(217:221);     % Benzene + Toluene
(500ppb + 500ppb)

% Sampling Period:
T = 12; % in seconds

% Time Constants ( in sec) from Averaged Single Analyte Table:
tauA = 204; % Ethylbenzene

% Steady-State Sensitivity from Averaged Single Analyte Table:
ssA = 2.24; % Ethylbenzene

% Adsorption Rate Constant:
Sa=(T/tauA); % Benzene
Sa = round(Sa*1e4)/1e4; % Analyte A

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Select the case to be analyzed
```

```
y = y2; yb = yb2;
kmax = length(y); % Length of the measurement data points
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Initialize Extended Kalman Filter variables
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
yhat = zeros(1,kmax); x = zeros(3,kmax+1); P = zeros(3,3,kmax+1);
x(:,1) = [0; Sa+0.001; -1]; % Initial state vector (Initial estimate)
P(:,:,1) = diag([1,1,100000]); % Initial error covariance Matrix
V = 0.00001; % Covariance matrix of process noise v
G = [1]; % Matrix G (dh/dw)  (1 by 1 Matrix)
U = ones(kmax,1); % Step Input
W = 100; % Covariance of Measurement noise

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%******************** Extended Kalman Filter Scheme
***********************
% Loop to evaluate each instant in time
for i=1:kmax
    % System Matrices
    C = [x(3,i) 0 x(1,i)]; % C Matrix (dh/dx)
    h = [x(3,i)*x(1,i)]; % h Matrix
    % f matrix
    f = [(1-x(2,i))*x(1,i)+x(2,i)*U(i); x(2,i); x(3,i)];
    % The A matrix (df/dx)
    A = [1-x(2,i) U(i)-x(1,i) 0; 0 1 0; 0 0 1];
    % F matrix (df/dv)
    F = [1; 0; 0];

    % Estimated measurement
    yhat(i)=h;

    % Setting the first measurement value to be 0
%     if i==1
%         y(i) = y(i)*0;
%     end

    K = (A*P(:,:,i)*C')/((C*P(:,:,i)*C' + G*W*G')) ;  % Kalman Gain
    x(:,i+1) = f + K*(y(i)-h); % State Update Equation
    % Error Covariance Update Equation:
    P(:,:,i+1) = (A-K*C)*P(:,:,i)*(A-K*C)' + K*G*W*G'*K' + F*V*F';
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Plot and Analysis
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Time vector
vk=(0:kmax-1)'; vt=(0:0.1:kmax-1)';
```

```
% Experimental Data Curve fitting to find S and alpha
hfit = fit((0:kmax-1)',y,'a*(1-exp(-S*x))','StartPoint',[1 1]);
S = round(hfit.S * 1e4)/1e4; % S
a = round(hfit.a * 1e2)/1e2; % alpha
vfit = a*(1-exp(-S.*vt)); % Experimental Data Fitting

% Estimated adsorption rate (S)
ES = round(x(2,kmax)*1e4)/1e4;
% Estimated Steady-state sensitivity (alpha)
Ea = round(x(3,kmax)*1e2)/1e2;
% Estimated Frequency Shift
vest = Ea*(1-exp(-ES.*vt));

% Adding some baseline
y = [yb(1:5);y]; yhat=[yb(1:5);yhat']; % The first 5 points
vfit=[((0:0.1:4.9)'*0);vfit]; vest=[((0:0.1:4.9)'*0);vest];

% Convert time to minutes
vk=((0:length(y)-1)')*(T/60); vt=(0:0.1:length(y)-1)'*(T/60);

% Plot of Frequency Shift
figure(1)
h=plot(vk, y, '*b', vt, vfit, '-w',vt, vfit, '-b', vt, vest, '--w', vt,
vest, '--r');
xlabel('Time, t (min)','FontSize',14); ylabel('Frequency Shift, \Deltaf
[kHz]','FontSize',14)
h_legend = legend('Experimental data', ['Experimental data fitting
expression: \Deltaf = ', ...
num2str(a) '*[1-exp(-' num2str(S/T) 't)]'], ...
'Experimental data fitting', ...
['Estimated fitting expression: \Deltaf = ' ...
num2str(Ea) '*[1-exp(-' num2str(ES/T) 't)]'], ...
'Estimated Sensor Response');
set(h_legend,'FontSize',15);
set(h,'LineWidth',3)

% Estimated Concentration
Con = abs(a)/(ssA);
Con_A = abs(Ea)/(ssA);


fprintf('The estimated concentration of Analyte A (in ppm) is \n')
disp(Con_A)
fprintf('The concentration of Analyte A (in ppm) (from fitting
parameter) is \n')
disp(Con)

figure(2)
h=plot(vk, y, '*b', vt, vfit, '-b', vt, vest, '--r');
xlabel('Time, t (min)','FontSize',14); ylabel('Frequency Shift, \Deltaf
(kHz)','FontSize',14)
h_legend = legend('Experimental data','Experimental data fitting', ...
'Estimated Sensor Response');
set(h_legend,'FontSize',15);
```

```
set(h,'LineWidth',3)
annotation('textbox', [0.2,0.4,0.1,0.1],...
    'String', ['Experimental data fitting expression: \Deltaf = ', ...
    num2str(a) '*[1-exp(-' num2str(round((S/T)*1e5)/1e5) 't)]'...
    'Estimated fitting expression: \Deltaf = ' ...
num2str(Ea) '*[1-exp(-' num2str(ES/T) 't)]']);

% ['Experimental data fitting expression: \Deltaf = ', ...
% num2str(a) '*[1-exp(-' num2str(S/T) 't)]'], ...

%['Estimated fitting expression: \Deltaf = ' ...
%num2str(Ea) '*[1-exp(-' num2str(ES/T) 't)]'], ...
% round((1/tauA)*1e4)/1e4

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Normalized Concentration
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
vk=(0:kmax)'; vt=(0:0.1:kmax)'; % Time

mA = (1 - exp(-Sa.*vt)); % Normalized Analyte A Concentration
mApoint = (1 - exp(-Sa.*vk));

% Estimated Normalized Concentration of Analyte A
EmA = x(1,:);
mAfit = fit((0:kmax)',EmA','EAa*(1-exp(-mAS*x))','StartPoint',[1 1]);
CA = mAfit.EAa;
mSA = mAfit.mAS;
EmmA = CA*(1 - exp(-mSA.*vt));

vk=vk*(T/60); vt=vt*(T/60); % Convert time to minute

% Plot of Normalized Concentration
figure(3)
h=plot(vt, mA, '-b', vt, EmmA, '--w', vt, EmmA, '--r');
xlabel('Time, t (min)','FontSize',14); ylabel('Normalized
Concentration','FontSize',14)
h_legend = legend(['Theoretical Normalized Concentration: m = '...
num2str(1) '*[1-exp(-' num2str(Sa/T) '*t)]'], ...
['Estimated Norm. Con. Expression: m = ' ...
num2str(CA) '*[1-exp(-' num2str(mSA/T) '*t)]'], ...
'Estimated Normalized Concentration');
set(h_legend,'FontSize',15);
set(h,'LineWidth',3)


figure(4)
h=plot(vt, mA, '-b', vt, EmmA, '--r');
xlabel('Time, t (min)','FontSize',14); ylabel('Normalized
Concentration','FontSize',14)
h_legend = legend('Theoretical Normalized Concentration',...
'Estimated Normalized Concentration');
set(h_legend,'FontSize',15);
set(h,'LineWidth',3)
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Time Analysis (Time to quantification)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

% Time vector
vk=(0:kmax-1)'; vt=(0:0.1:kmax-1)';

% Several Estimated adsorption rate (S)
ES1 = round(x(2,10)*1e4)/1e4 % after 2 minutes
ES2 = round(x(2,15)*1e4)/1e4 % after 3 minutes
ES3 = round(x(2,20)*1e4)/1e4 % after 4 minutes
ES4 = round(x(2,25)*1e4)/1e4; % after 5 minutes
ES5 = round(x(2,30)*1e4)/1e4; % after 6 minutes

% Several Estimated Steady-state sensitivity (alpha)
Ea1 = round(x(3,10)*1e4)/1e4 % after 4 minutes
Ea2 = round(x(3,15)*1e4)/1e4 % after 5 minutes
Ea3 = round(x(3,20)*1e4)/1e4 % after 6 minutes
Ea4 = round(x(3,25)*1e4)/1e4; % after 10 minutes
Ea5 = round(x(3,30)*1e4)/1e4; % after 12 minutes

% Several Estimated Frequency Shift
vest1 = Ea1*(1-exp(-ES1.*vt)); % after 4 minutes
vest2 = Ea2*(1-exp(-ES2.*vt)); % after 6 minutes
vest3 = Ea3*(1-exp(-ES3.*vt)); % after 8 minutes
vest4 = Ea4*(1-exp(-ES4.*vt)); % after 10 minutes
vest5 = Ea5*(1-exp(-ES5.*vt)); % after 12 minutes

vest1=[((0:0.1:4.9)'*0);vest1];
vest2=[((0:0.1:4.9)'*0);vest2];
vest3=[((0:0.1:4.9)'*0);vest3];
vest4=[((0:0.1:4.9)'*0);vest4];
vest5=[((0:0.1:4.9)'*0);vest5];

pd1 = ((ES1-S)/S)*100
pd2 = ((ES2-S)/S)*100
pd3 = ((ES3-S)/S)*100
pd4 = ((ES4-S)/S)*100
pd5 = ((ES5-S)/S)*100

pda1 = ((Ea1-a)/a)*100
pda2 = ((Ea2-a)/a)*100
pda3 = ((Ea3-a)/a)*100
pda4 = ((Ea4-a)/a)*100
pda5 = ((Ea5-a)/a)*100

pd = ((ES-S)/S)*100
pda = ((Ea-a)/a)*100

% Change time to minutes
```

```
vt=(0:0.1:length(y)-1)'*(T/60); vk=((0:length(y)-1)')*(T/60);

% Plot
figure(5)
h=plot(vk, y, '*b', vt, vfit, '-b', vt, vest1,'--r', vt, vest2,'--g',
vt, vest3,'--c', vt, vest,'--k');
xlabel('Time, t (min)','FontSize',14); ylabel('Frequency Shift, \Deltaf
(kHz)','FontSize',14)
h_legend = legend('Experimental data', ...
'Experimental data fitting', ...
'Estimated Sensor Response after 2 minutes', ...
'Estimated Sensor Response after 3 minutes', ...
'Estimated Sensor Response after 4 minutes', ...
'Estimated Sensor Response using all data points');

set(h_legend,'FontSize',15);
set(h,'LineWidth',3)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% %%%%%%%%%%%%%%%%%%%%%%%%%%%% End of Program
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

B.4 MATLAB Code for Two-Analyte Estimation (Nonlinear Model)

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Author: KARTHICK SOTHIVELR
% File Name: Two_Analyte.m
% Date (Created): 02-27-2013
% Date (Modified): 02-17-2014
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Description:
% Program to estimate the steady-state sensitvity and concentrations of
the
% binary mixture using the two analyte system model. Estimation
% were performed using Extended Kalman Filter.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Cleaning
clear all
close all
clc
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Open and read the measurement file
FID = fopen('thesis_BT_2nd.ini','r'); % 2nd new one
data = textscan(FID,'%f %f');
fclose(FID);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Order the data (Based on Analyte)
y1 = data{2}(53:78); yb1 = data{2}(48:52);         % Benzene (1000ppb)
y2 = data{2}(127:175); yb2 = data{2}(122:126);     % Benzene + Toluene
(1000ppb + 500ppb)
y3 = data{2}(222:271); yb3 = data{2}(217:221);     % Benzene + Toluene
(500ppb + 500ppb)
y4 = data{2}(367:417); yb4 = data{2}(362:366);     % Benzene + Toluene
(200ppb + 500ppb)

% Sampling Period:
T = 12; % in seconds

% Time Constants ( in sec) from Averaged Single Analyte Table:
tauA = 34.6; % Benzene
tauB = 92.6; % Toluene

% Steady-State Sensitivity from Averaged Single Analyte Table:
ssA = 0.257; % Benzene
ssB = 0.746; % Toluene

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Select the case to be analyzed
y = y2; yb = yb2;
```

```
kmax = length(y); % Length of the measurement data points
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Initialize Extended Kalman Filter variables
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
yhat = zeros(1,kmax); x = zeros(4,kmax+1); P = zeros(4,4,kmax+1);
x(:,1) = [0; 0; 0; 0]; % Initial state vector (Initial estimate)
P(:,:,1) = diag([0.1,0.1,500,1000]); % Initial error covariance Matrix
V = 10; % Covariance matrix of process noise v
G = [1]; % Matrix G (dh/dw)  (1 by 1 Matrix)
U = ones(kmax,1); % Step Input
W = 100; % Covariance of Measurement noise

% Adsorption Rate Constant:
Sa=(T/tauA); % Benzene
Sb=(T/tauB); % Ethylbenzene
Sa = round(Sa*1e4)/1e4; % Analyte A
Sb = round(Sb*1e4)/1e4; % Analyte B

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%******************** Extended Kalman Filter Scheme
**********************
% Loop to evaluate each instant in time
for i=1:kmax
    % System Matrices
    C = [x(3,i) x(4,i) x(1,i)  x(2,i)]; % C Matrix (dh/dx)
    h = [x(3,i)*x(1,i) + x(4,i)*x(2,i)]; % h Matrix
    % f matrix
    f = [(1-Sa)*x(1,i)+Sa*U(i); (1-Sb)*x(2,i)+Sb*U(i); x(3,i); x(4,i)];
    A = [1-Sa 0 0 0; 0 1-Sb 0 0; 0 0 1 0; 0 0 0 1]; % The A matrix
(df/dx)
    F = [1 1 0 0]; % The F matrix (df/dv)

    % Estimated measurement
    yhat(i)=h;

    % Setting the first measurement value to be 0
%     if i==1
%         y(i) = y(i)*0;
%     end

    K = (A*P(:,:,i)*C')/((C*P(:,:,i)*C' + G*W*G')) ;  % Kalman Gain
    x(:,i+1) = f + K*(y(i)-h); % State Update Equation
    % Error Covariance Update Equation:
    P(:,:,i+1) = (A-K*C)*P(:,:,i)*(A-K*C)' + K*G*W*G'*K' + F*V*F';
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Plot and Analysis
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
vk=0:kmax-1;
vt=(0:0.1:kmax-1)';

% Experimental Data Fitting
fun=['a*(1-exp(' ...
    num2str(-(T/tauA)) '*x))+b*(1-exp(' num2str(-(T/tauB)) '*x))'];
hfit = fit((0:kmax-1)',y,fun,'StartPoint',[1 1]);

% Steady-State Sensitivity
a = round(hfit.a * 1e2)/1e2; % Analyte A
b = round(hfit.b * 1e2)/1e2; % Analyte B
vfit = a*(1-exp(-(T/tauA).*vt)) + b*(1-exp(-(T/tauB).*vt));

% Estimated Steady-State Sensitivity
Ea = round(x(3,kmax)*1e4)/1e4; % Analyte A
Eb = round(x(4,kmax)*1e4)/1e4; % Analyte B
% Estimated Frequency Shift
vest = Ea*(1-exp(-Sa.*vt)) + Eb*(1-exp(-Sb.*vt));

% Converting time step number to minutes and add baselines
y = [yb(1:5);y]; yhat = [yb(1:5);yhat'];
vfit=[((0:0.1:4.9)'*0);vfit]; vest=[((0:0.1:4.9)'*0);vest];
vk=((0:length(y)-1)')*(T/60); vt=(0:0.1:length(y)-1)'*(T/60);

% Plot
figure(1)
plot(vk, y, '*b', vt, vfit, '-b', vt, vest, '--r')
title ('Frequency Shift vs Time')
xlabel('Time (min)'); ylabel('Frequency Shift \Deltaf [kHz]')
legend('Experimental data',    ['Experimental data Fitting: \Deltaf=',
...
num2str(a) '*[1-exp(-' num2str(Sa) 'k)] + ' ...
num2str(b) '*[1-exp(-' num2str(Sb) 'k)]'],['Estimated data
Fitting:\Deltaf=' ...
num2str(Ea) '*[1-exp(-' num2str(Sa) 'k)] + ' ...
num2str(Eb) '*[1-exp(-' num2str(Sb) 'k)]']);


figure(2)
% Plot of Frequency Shift vs Time step
h=plot(vk, y, '*b', vt, vfit, '-w',vt, vfit, '-b', vt, vest, '--w', vt,
vest, '--r');
%title ('Frequency Shift vs Time','FontSize',24)
%LineWidth = [3];
xlabel('Time, t (min)','FontSize',14); ylabel('Frequency Shift, \Deltaf
[kHz]','FontSize',14)
h_legend = legend('Experimental data', ['Experimental data fitting
expression: \Deltaf = ', ...
        num2str(a) '*[1-exp(-' num2str(round((Sa/T)*1e4)/1e4) 't)] + '
...
        num2str(b) '*[1-exp(-' num2str(round((Sb/T)*1e5)/1e5) 't)]'],
...
        'Experimental data fitting', ...
```

```
        ['Estimated fitting expression: \Deltaf = ' ...
        num2str(Ea) '*[1-exp(-' num2str(round((Sa/T)*1e4)/1e4) 't)] + '
...
        num2str(Eb) '*[1-exp(-' num2str(round((Sb/T)*1e5)/1e5) 't)]'],
...
        'Estimated Sensor Response');
set(h_legend,'FontSize',15);
set(h,'LineWidth',3)

figure(3)
% Plot of Frequency Shift vs Time step
h=plot(vk, y, '*b', vt, vfit, '-b', vt, vest, '--r');
%title ('Frequency Shift vs Time','FontSize',24)
%LineWidth = [3];
xlabel('Time (min)','FontSize',14); ylabel('Frequency Shift, \Deltaf
(kHz)','FontSize',14)
h_legend = legend('Experimental data', 'Experimental data fitting', ...
            'Estimated Sensor Response');
set(h_legend,'FontSize',15);
set(h,'LineWidth',3)
annotation('textbox', [0.2,0.4,0.1,0.1],...
    'String', ['Experimental data fitting expression: \Deltaf = ', ...
        num2str(a) '*[1-exp(-' num2str(round((Sa/T)*1e4)/1e4) 't)] + '
...
        num2str(b) '*[1-exp(-' num2str(round((Sb/T)*1e5)/1e5) 't)]'...
        'Estimated fitting expression: \Deltaf = ' ...
        num2str(Ea) '*[1-exp(-' num2str(round((Sa/T)*1e4)/1e4) 't)] + '
...
        num2str(Eb) '*[1-exp(-' num2str(round((Sb/T)*1e5)/1e5) 't)]']);


% Estimated Concentration
Con_A = abs(Ea)/(ssA);
Con_B = abs(Eb)/(ssB);

ConA = abs(a)/(ssA);
ConB = abs(b)/(ssB);


fprintf('The estimated concentration of Analyte A (in ppm) is \n')
disp(Con_A)

fprintf('The estimated concentration of Analyte B (in ppm) is \n')
disp(Con_B)

fprintf('The concentration of Analyte A (in ppm) (from fitting
parameter) is \n')
disp(ConA)

fprintf('The concentration of Analyte B (in ppm) (from fitting
parameter) is \n')
disp(ConB)


%Info
```

```
%plot(vk, y, '*b', vt, vfit, '-b', vk, yhat, 'or', vt, vest, '--r')

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Normalized Concentration
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
vk=(0:kmax)'; vt=(0:0.1:kmax)'; % Time

mA = (1 - exp(-Sa.*vt)); % Normalized Analyte A Concentration
mApoint = (1 - exp(-Sa.*vk));

mB = (1 - exp(-Sb.*vt)); % Normalized Analyte B Concentration
mBpoint = (1 - exp(-Sb.*vk));

% Estimated Normalized Concentration of Analyte A
EmA = x(1,:);
mAfit = fit((0:kmax)',EmA','EAa*(1-exp(-mAS*x))','StartPoint',[1 1]);
CA = mAfit.EAa;
mSA = mAfit.mAS;
EmmA = CA*(1 - exp(-mSA.*vt));

% Estimated Normalized Concentration of Analyte A
EmB = x(2,:);
mBfit = fit((0:kmax)',EmB','EBb*(1-exp(-mBS*x))','StartPoint',[1 1]);
CB = mBfit.EBb;
mSB = mBfit.mBS;
EmmB = CB*(1 - exp(-mSB.*vt));

vk=vk*(T/60); vt=vt*(T/60); % Convert time to minute

% Plot of Normalized Concentration
% Analyte A
% figure(4)
% h=plot(vt, mA, '-b', vt, EmmA, '--w', vt, EmmA, '--r');
% xlabel('Time, t (min)','FontSize',14); ylabel('Normalized
Concentration','FontSize',14)
% h_legend = legend(['Theoretical Normalized Concentration: m = '...
% num2str(1) '*[1-exp(-' num2str(Sa/T) '*t)]'], ...
% ['Estimated Norm. Con. Expression: m = ' ...
% num2str(CA) '*[1-exp(-' num2str(mSA/T) '*t)]'], ...
% 'Estimated Normalized Concentration');
% axis([0 12 0 1.25])
% set(h_legend,'FontSize',15);
% set(h,'LineWidth',3)

figure(4)
h=plot(vt, mA, '-b', vt, EmmA, '--r');
xlabel('Time (min)','FontSize',14); ylabel('Normalized
Concentration','FontSize',14)
h_legend = legend('Theoretical Normalized Concentration', ...
'Estimated Normalized Concentration');
axis([0 12 0 1.25])
set(h_legend,'FontSize',15);
set(h,'LineWidth',3)
```

```
% Analyte B
% figure(5)
% h=plot(vt, mB, '-b', vt, EmmB, '--w', vt, EmmB, '--r');
% xlabel('Time, t (min)','FontSize',14); ylabel('Normalized
Concentration','FontSize',14)
% h_legend = legend(['Theoretical Normalized Concentration: m = '...
% num2str(1) '*[1-exp(-' num2str(Sb/T) '*t)]'], ...
% ['Estimated Norm. Con. Expression: m = ' ...
% num2str(CB) '*[1-exp(-' num2str(mSB/T) '*t)]'], ...
% 'Estimated Normalized Concentration');
% axis([0 12 0 1.25])
% set(h_legend,'FontSize',15);
% set(h,'LineWidth',3)

figure(5)
h=plot(vt, mB, '-b', vt, EmmB, '--r');
xlabel('Time (min)','FontSize',14); ylabel('Normalized
Concentration','FontSize',14)
h_legend = legend('Theoretical Normalized Concentration',...
'Estimated Normalized Concentration');
axis([0 12 0 1.25])
set(h_legend,'FontSize',15);
set(h,'LineWidth',3)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Time Analysis (Time to quantification)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

% Time vector
vk=(0:kmax-1)'; vt=(0:0.1:kmax-1)';

% Several Estimated Steady-state sensitivity (Analyte A)
Ea1 = round(x(3,20)*1e4)/1e4; % after 2 minutes
Ea2 = round(x(3,25)*1e4)/1e4; % after 3 minutes
Ea3 = round(x(3,30)*1e4)/1e4; % after 4 minutes
Ea4 = round(x(3,35)*1e4)/1e4; % after 7 minutes
Ea5 = round(x(3,40)*1e4)/1e4; % after 8 minutes

% Several Estimated Steady-state sensitivity (Analyte B)
Eb1 = round(x(4,20)*1e4)/1e4; % after 4 minutes
Eb2 = round(x(4,25)*1e4)/1e4; % after 5 minutes
Eb3 = round(x(4,30)*1e4)/1e4; % after 6 minutes
Eb4 = round(x(4,35)*1e4)/1e4; % after 7 minutes
Eb5 = round(x(4,40)*1e4)/1e4; % after 8 minutes

% Several Estimated Frequency Shift
vest1 = Ea1*(1-exp(-Sa.*vt)) + Eb1*(1-exp(-Sb.*vt));
vest2 = Ea2*(1-exp(-Sa.*vt)) + Eb2*(1-exp(-Sb.*vt));
vest3 = Ea3*(1-exp(-Sa.*vt)) + Eb3*(1-exp(-Sb.*vt));
```

```
vest4 = Ea4*(1-exp(-Sa.*vt)) + Eb4*(1-exp(-Sb.*vt));
vest5 = Ea5*(1-exp(-Sa.*vt)) + Eb5*(1-exp(-Sb.*vt));

vest1=[((0:0.1:4.9)'*0);vest1];
vest2=[((0:0.1:4.9)'*0);vest2];
vest3=[((0:0.1:4.9)'*0);vest3];
vest4=[((0:0.1:4.9)'*0);vest4];
vest5=[((0:0.1:4.9)'*0);vest5];

pda1 = ((Ea1-a)/a)*100
pda2 = ((Ea2-a)/a)*100
pda3 = ((Ea3-a)/a)*100
pda4 = ((Ea4-a)/a)*100
pda5 = ((Ea5-a)/a)*100

pdb1 = ((Eb1-b)/b)*100
pdb2 = ((Eb2-b)/b)*100
pdb3 = ((Eb3-b)/b)*100
pdb4 = ((Eb4-b)/b)*100
pdb5 = ((Eb5-b)/b)*100

pda = ((Ea-a)/a)*100
pdb = ((Eb-b)/b)*100

% Change time to minutes
vt=(0:0.1:length(y)-1)'*(T/60); vk=((0:length(y)-1)')*(T/60);

% Plot
figure(6)
h=plot(vk, y, '*b', vt, vfit, '-b', vt, vest1,'--r', vt, vest2,'--g',
vt, vest3,'--c', vt, vest,'--k');
xlabel('Time (min)','FontSize',14); ylabel('Frequency Shift, \Deltaf
(kHz)','FontSize',14)
h_legend = legend('Experimental data', ...
'Experimental data fitting', ...
'Estimated Sensor Response after 4 minutes', ...
'Estimated Sensor Response after 5 minutes', ...
'Estimated Sensor Response after 6 minutes', ...
'Estimated Sensor Response using all data points');

set(h_legend,'FontSize',15);
set(h,'LineWidth',3)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%


%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%
% %%%%%%%%%%%%%%%%%%%%%%%%%% End of Program
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

B.5 MATLAB Code for Two-Analyte Estimation (Linear Model)

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Author: KARTHICK SOTHIVELR
% File Name: alternative_2Analyte.m
% Date (Created): 02-27-2013
% Date (Modified): 02-17-2014
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Description:
% Program to estimate the steady-state sensitvity and concentrations of
the
% binary mixture using the alternative two analyte system model.
Estimation
% were performed using Kalman Filter.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Cleaning
clear all
close all
clc
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Open and read the measurement file
FID = fopen('thesis_BT_2nd.ini','r'); % 121211PEABEJ3binaryCor
data = textscan(FID,'%f %f');
fclose(FID);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Order the data (Based on Analyte)
y1 = data{2}(53:78); yb1 = data{2}(48:52);         % Benzene (1000ppb)
y2 = data{2}(127:175); yb2 = data{2}(122:126);     % Benzene + Toluene
(1000ppb + 500ppb)
y3 = data{2}(222:271); yb3 = data{2}(217:221);     % Benzene + Toluene
(500ppb + 500ppb)
y4 = data{2}(367:417); yb4 = data{2}(362:366);     % Benzene + Toluene
(200ppb + 500ppb)

% Sampling Period:
T = 12; % in seconds

% Time Constants ( in sec) from Averaged Single Analyte Table:
tauA = 34.6; % Benzene
tauB = 92.6; % Toluene

% Steady-State Sensitivity from Averaged Single Analyte Table:
ssA = 0.257; % Benzene
ssB = 0.746; % Toluene

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Select the case to be analyzed
```

```
y = y2; yb = yb2;
kmax = length(y); % Length of the measurement data points
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Initialize Kalman Filter variables
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
yhat = zeros(1,kmax); x = zeros(2,kmax+1); P = zeros(2,2,kmax+1);
x(:,1) = [0; 0]; % Initial state vector (Initial state estimate)
P(:,:,1) = diag([1000, 500]); % Initial error covariance Matrix
W = 10; V = 1; % Covariance of Measurement and process noise
G = [1]; % Matrix G (1 by 1 Matrix)
U = ones(kmax,1); % Step Input
m_A=0; m_B=0; % Initial value of the normalized Concentration
% System Matrices (A and F)
A = eye(2);
F = [1;1];
% Adsorption Rate Constant:
Sa=(T/tauA); % Benzene
Sb=(T/tauB); % Ethylbenzene
Sa = round(Sa*1e4)/1e4; % Analyte A
Sb = round(Sb*1e4)/1e4; % Analyte B

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
%************************ Kalman Filter Scheme
**************************
% Loop to evaluate each instant in time
for i=1:kmax
    C = [m_A(i) m_B(i)];% C Matrix
    % Simulate the normalized concentration values:
    m_A = (1 - Sa)*m_A + Sa*U;
    m_B = (1 - Sb)*m_B + Sb*U;
    % Estimated measurement
    yhat(i)=C*x(:,i);

    K = (A*P(:,:,i)*C')/((C*P(:,:,i)*C' + G*W*G')) ;  % Kalman Gain
    x(:,i+1) = A*x(:,i) + K*(y(i)-C*x(:,i)); % State Update Equation
    % Error Covariance Update Equation:
    P(:,:,i+1) = (A-K*C)*P(:,:,i)*(A-K*C)' + K*G*W*G'*K' + F*V*F';
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Plot and Analysis
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
vk=0:kmax-1;
vt=(0:0.1:kmax-1)';

% Experimental Data Fitting
fun = ['a*(1-exp(' num2str(-(T/tauA)) ...
    '*x))+b*(1-exp(' num2str(-(T/tauB)) '*x))'];
```

```
hfit = fit((0:kmax-1)',y,fun,'StartPoint',[1 1]);
% Steady-State Sensitivity
a = round(hfit.a * 1e2)/1e2; % Analyte A
b = round(hfit.b * 1e2)/1e2; % Analyte B
vfit = a*(1-exp(-(T/tauA).*vt)) + b*(1-exp(-(T/tauB).*vt));

% Estimated Steady-State Sensitivity
Ea = round(x(1,kmax)*1e4)/1e4; % Analyte A
Eb = round(x(2,kmax)*1e4)/1e4; % Analyte B
% Estimated Frequency Shift
vest = Ea*(1-exp(-Sa.*vt)) + Eb*(1-exp(-Sb.*vt));

% Converting time step number to minutes and add baselines
y = [yb(1:5);y]; yhat = [yb(1:5);yhat'];
vfit=[((0:0.1:4.9)'*0);vfit]; vest=[((0:0.1:4.9)'*0);vest];
vk=((0:length(y)-1)')*(T/60); vt=(0:0.1:length(y)-1)'*(T/60);

% Plot
figure(1)
plot(vk, y, '*b', vt, vfit, '-b', vt, vest, '--r')
title ('Frequency Shift vs Time')
xlabel('Time (min)'); ylabel('Frequency Shift \Deltaf [kHz]')
legend('Experimental data',    ['Experimental data Fitting: \Deltaf=',
...
num2str(a) '*[1-exp(-' num2str(Sa) 'k)] + ' ...
num2str(b) '*[1-exp(-' num2str(Sb) 'k)]'],['Estimated data
Fitting:\Deltaf=' ...
num2str(Ea) '*[1-exp(-' num2str(Sa) 'k)] + ' ...
num2str(Eb) '*[1-exp(-' num2str(Sb) 'k)]']);

figure(2)
% Plot of Frequency Shift vs Time step
h=plot(vk, y, '*b', vt, vfit, '-w',vt, vfit, '-b', vt, vest, '--w', vt,
vest, '--r');
%title ('Frequency Shift vs Time','FontSize',24)
%LineWidth = [3];
xlabel('Time, t (min)','FontSize',14); ylabel('Frequency Shift, \Deltaf
[kHz]','FontSize',14)
h_legend = legend('Experimental data', ['Experimental data fitting
expression: \Deltaf = ', ...
        num2str(a) '*[1-exp(-' num2str(round((Sa/T)*1e4)/1e4) 't)] + '
...
        num2str(b) '*[1-exp(-' num2str(round((Sb/T)*1e5)/1e5) 't)]'],
...
        'Experimental data fitting', ...
        ['Estimated fitting expression: \Deltaf = ' ...
        num2str(Ea) '*[1-exp(-' num2str(round((Sa/T)*1e4)/1e4) 't)] + '
...
        num2str(Eb) '*[1-exp(-' num2str(round((Sb/T)*1e5)/1e5) 't)]'],
...
        'Estimated Sensor Response');
set(h_legend,'FontSize',15);
set(h,'LineWidth',3)
% annotation('textbox', [0.2,0.4,0.1,0.1],...
%            'String', 'Estimated Concentration: Benzene (1743 ppb) and
Toluene (897 ppb)');
```

```
figure(3)
% Plot of Frequency Shift vs Time step
h=plot(vk, y, '*b', vt, vfit, '-b', vt, vest, '--r');
%title ('Frequency Shift vs Time','FontSize',24)
%LineWidth = [3];
xlabel('Time (min)','FontSize',14); ylabel('Frequency Shift, \Deltaf
(kHz)','FontSize',14)
h_legend = legend('Experimental data', 'Experimental data fitting', ...
            'Estimated Sensor Response');
set(h_legend,'FontSize',15);
set(h,'LineWidth',3)
annotation('textbox', [0.2,0.4,0.1,0.1],...
    'String', ['Experimental data fitting expression: \Deltaf = ', ...
        num2str(a) '*[1-exp(-' num2str(round((Sa/T)*1e4)/1e4) 't)] + '
...
        num2str(b) '*[1-exp(-' num2str(round((Sb/T)*1e5)/1e5) 't)]'...
        'Estimated fitting expression: \Deltaf = ' ...
        num2str(Ea) '*[1-exp(-' num2str(round((Sa/T)*1e4)/1e4) 't)] + '
...
        num2str(Eb) '*[1-exp(-' num2str(round((Sb/T)*1e5)/1e5) 't)]']);


% Estimated Concentration
Con_A = abs(Ea)/(ssA);
Con_B = abs(Eb)/(ssB);

ConA = abs(a)/(ssA);
ConB = abs(b)/(ssB);


fprintf('The estimated concentration of Analyte A (in ppm) is \n')
disp(Con_A)

fprintf('The estimated concentration of Analyte B (in ppm) is \n')
disp(Con_B)

fprintf('The concentration of Analyte A (in ppm) (from fitting
parameter) is \n')
disp(ConA)

fprintf('The concentration of Analyte B (in ppm) (from fitting
parameter) is \n')
disp(ConB)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Time Analysis (Time to quantification)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%

% Time vector
vk=(0:kmax-1)'; vt=(0:0.1:kmax-1)';
```

```
% Several Estimated Steady-state sensitivity (Analyte A)
Ea1 = round(x(1,20)*1e4)/1e4; % after 4 minutes
Ea2 = round(x(1,25)*1e4)/1e4; % after 5 minutes
Ea3 = round(x(1,30)*1e4)/1e4; % after 6 minutes
Ea4 = round(x(1,35)*1e4)/1e4; % after 7 minutes
Ea5 = round(x(1,40)*1e4)/1e4; % after 8 minutes

% Several Estimated Steady-state sensitivity (Analyte B)
Eb1 = round(x(2,20)*1e4)/1e4; % after 4 minutes
Eb2 = round(x(2,25)*1e4)/1e4; % after 5 minutes
Eb3 = round(x(2,30)*1e4)/1e4; % after 6 minutes
Eb4 = round(x(2,35)*1e4)/1e4; % after 7 minutes
Eb5 = round(x(2,40)*1e4)/1e4; % after 8 minutes

% Several Estimated Frequency Shift
vest1 = Ea1*(1-exp(-Sa.*vt)) + Eb1*(1-exp(-Sb.*vt));
vest2 = Ea2*(1-exp(-Sa.*vt)) + Eb2*(1-exp(-Sb.*vt));
vest3 = Ea3*(1-exp(-Sa.*vt)) + Eb3*(1-exp(-Sb.*vt));
vest4 = Ea4*(1-exp(-Sa.*vt)) + Eb4*(1-exp(-Sb.*vt));
vest5 = Ea5*(1-exp(-Sa.*vt)) + Eb5*(1-exp(-Sb.*vt));

vest1=[((0:0.1:4.9)'*0);vest1];
vest2=[((0:0.1:4.9)'*0);vest2];
vest3=[((0:0.1:4.9)'*0);vest3];
vest4=[((0:0.1:4.9)'*0);vest4];
vest5=[((0:0.1:4.9)'*0);vest5];

pda1 = ((Ea1-a)/a)*100
pda2 = ((Ea2-a)/a)*100
pda3 = ((Ea3-a)/a)*100
pda4 = ((Ea4-a)/a)*100
pda5 = ((Ea5-a)/a)*100

pdb1 = ((Eb1-b)/b)*100
pdb2 = ((Eb2-b)/b)*100
pdb3 = ((Eb3-b)/b)*100
pdb4 = ((Eb4-b)/b)*100
pdb5 = ((Eb5-b)/b)*100

pda = ((Ea-a)/a)*100
pdb = ((Eb-b)/b)*100


% Change time to minutes
vt=(0:0.1:length(y)-1)'*(T/60); vk=((0:length(y)-1)')*(T/60);

% Plot
figure(5)
h=plot(vk, y, '*b', vt, vfit, '-b', vt, vest1,'--r', vt, vest2,'--g',
vt, vest3,'--c', vt, vest,'--k');
xlabel('Time, t (min)','FontSize',14); ylabel('Frequency Shift, \Deltaf
(kHz)','FontSize',14)
h_legend = legend('Experimental data', ...
'Experimental data fitting', ...
'Estimated Sensor Response after 4 minutes', ...
```

```
'Estimated Sensor Response after 5 minutes', ...
'Estimated Sensor Response after 6 minutes', ...
'Estimated Sensor Response using all data points');

set(h_legend,'FontSize',15);
set(h,'LineWidth',3)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%


%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%
% %%%%%%%%%%%%%%%%%%%%%%%%%% End of Program
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```