

# Dynamic Data Rectification Using Particle Filters

Tao Chen, Julian Morris and Elaine Martin\*

*School of Chemical Engineering and Advanced Materials,  
Newcastle University, Newcastle upon Tyne, NE1 7RU, U.K.*

\* E-mail: e.b.martin@ncl.ac.uk; Tel.: +44 191 222 6231; Fax: +44 191 222 5292.

## Abstract

The basis of dynamic data rectification is a dynamic process model. The successful application of the model requires the fulfilling of a number of objectives that are as wide-ranging as the estimation of the process states, process signal denoising and outlier detection and removal. Current approaches to dynamic data rectification include the conjunction of the Extended Kalman filter (EKF) and the expectation-maximization algorithm. However, this approach is limited due to the EKF being less applicable where the state and measurement functions are highly non-linear or where the posterior distribution of the states is non-Gaussian. This paper proposes an alternative approach whereby particle filters, based on the sequential Monte Carlo method, are utilized for dynamic data rectification. By formulating the rectification problem within a probabilistic framework, the particle filters generate Monte Carlo samples from the posterior distribution of the system states, and thus provide the basis for rectifying the process measurements. Furthermore, the proposed technique is capable of detecting changes in process operation and thus complements the task of process fault diagnosis. The appropriateness of particle filters for dynamic data rectification is demonstrated through their application to an illustrative non-linear dynamic system, and a benchmark pH neutralization process.

*Key words:* Dynamic data rectification; Filtering; Particle filters; Sequential Monte Carlo; State estimation.

## 1. Introduction

Mechanistic (phenomenological) models, based on mathematical representations of the physical and chemical behaviour of the process, have been widely applied for process optimization, process monitoring and process control. Whilst a set of differential equations are generally utilized in mechanistic modelling to describe the process of interest, in application studies they are typically discretized to give difference equations. These equations can be regarded as a general state space model (Jazwinski, 1970) with the difference variables defining the states. A typical state space model comprises both state and measurement functions,  $\mathbf{f}_k$  and  $\mathbf{h}_k$ , respectively :

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) + \mathbf{v}_{k-1} \quad (1)$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{n}_k \quad (2)$$

where  $k$  is the time index, and the vectors,  $\mathbf{x}$ ,  $\mathbf{u}$ , and  $\mathbf{z}$ , are the process states, the input variables and the process measurements, respectively.  $\mathbf{v}$  and  $\mathbf{n}$  denote independent and identically distributed noise for the process states and measurements, respectively. In practice the noise terms can take other forms, such as multiplicative and auto-correlated.

This study focuses on the model defined in Eqs. (1) and (2) due to its wide-spread applicability in process engineering (Chen et al., 2004; Romanenko and Castro, 2004; Terwiesch & Agarwal, 1995). However due to the presence of uncertainty both in the process states and the measurements (i.e.  $\mathbf{v}$  and  $\mathbf{n}$ ), the on-line estimation of the process states, and the removal of noise and outliers, i.e. random abrupt peaks, in the measurements, is critical to the application of a mechanistic model for process optimization, control and performance monitoring. This task is termed dynamic data rectification (or dynamic data reconciliation) (Darouach & Zasadzinski, 1991; Liebman, Edgar & Lasdon, 1992; Singhal & Seborg, 2000), and differs to the “steady” rectification techniques reported in the literature (Morad, Young & Svrcek, 2005; Tong & Crowe, 1997) which are designed around the steady state operation of processes and thus do not need to take into account the process dynamics. Dynamic data rectification is especially challenging when the state and/or measurement functions are highly non-linear, and the posterior distribution of the process states, is not Gaussian. These issues form the basis of this paper.

State estimation can itself be considered as an optimal filtering problem based on some pre-defined optimality criterion. By assuming that the probability distribution of the states is multivariate Gaussian, the Kalman filter (KF), and its non-linear extension, the extended Kalman filter (EKF) (Jazwinski, 1970) have been widely applied to estimate the states from noise-corrupted measurements, based on the system dynamics. Since the KF and EKF are not specifically designed to detect and remove outliers, Singhal and Seborg (2000) proposed a probabilistic formulation that combined the EKF and the expectation-maximization (EM) algorithm to attain the rectified measurements. The details are discussed in Section 2. One of the issues with this approach is that it is limited by the applicability of the EKF, which has been shown, in a number of applications, to give unsatisfactory state estimates, when the state and/or measurement equations are highly non-linear and the posterior distribution of the states is non-Gaussian (Chen et al., 2004; Chen, Morris & Martin, 2005; Terwiesch & Agarwal, 1995).

A number of alternative state estimation approaches have been proposed to address the fundamental problems associated with the EKF, i.e. the Gaussian assumption and the approximation of a non-linear system using a first-order Taylor series expansion. The unscented Kalman filter (UKF) (Julier, Uhlmann & Durrant-Whyte, 2000; Romanenko & Castro, 2004; Wan & van der Merwe, 2000) provides a better approximation to the non-linear behaviour by using the *unscented transformation*, in contrast to the Taylor series expansion, although it still assumes the states are Gaussian distributed. The point-mass filter, also termed the probability-grid filter (Bucy & Senne, 1971; Sorenson, 1988; Terwiesch & Agarwal, 1995) approximates the posterior distribution by discretizing the continuous state variables into grids. This approach materialises in an exponential increase in the computational cost with state dimension, thus limiting its widespread application.

More recently, particle filters (Arulampalam et al., 2002) have attained significant interest with respect to state estimation. The basic idea is to generate a large number of random samples (particles) using Monte Carlo methods to approximate the posterior distribution of the states. This approach materialises in the particles being adaptively concentrated in regions of high probability. This is in contrast to point-mass filters which adopt a pre-defined discretization approach to the state space, resulting in the particles being assumed to be uniformly distributed over the space. In addition, particle filters make use of the non-linear state and measurement functions, rather than a linear approximation as in the EKF, for state estimation. A number of applications of particle filters in process engineering have been reported in the literature (Chen et al., 2004; Chen, Morris &

Martin, 2005). These approaches specifically focus on state and/or parameter estimation but do not explicitly consider the rectification of measurements which may be affected by outliers.

This paper proposes the application of particle filters for dynamic data rectification with the goal of attaining satisfactory state estimates and detecting the presence of outliers, thereby recovering the “true” measurements from the corrupted values. The rectification problem will initially be presented within a probabilistic framework, prior to describing the specific application of particle filters for this task. In addition, this technique is shown to be applicable for the detection of process changes and thus complements the task of process fault diagnosis. The proposed approach for dynamic data rectification is illustrated by its application to a non-linear dynamic system, and a simulated benchmark pH neutralization reactor, and it is observed to exhibit improved performance over the rectification techniques based on the EKF and the EKF with the EM algorithm (EKF-EM) (Singhal & Seborg, 2000).

In summary this paper focuses on the formulation, and solution, of the dynamic data rectification problem utilising a probabilistic and statistical framework. In the literature considerable research has been reported to address this problem, most notably deterministic optimization algorithms using non-linear programming and moving horizon estimation (Abu-el-zeet, Becerra & Roberts, 2002; Liebman, Edgar & Lasdon, 1992). The execution of a comparative study between probabilistic filtering methods, such as the EKF and particle filter, is considered in this paper, but deterministic optimization based solutions are outwith the scope of this paper.

## 2. Dynamic Data Rectification

The key to data rectification is the definition of a noise model for the measurements. The measurement noise is typically considered to be described by a zero mean Gaussian distribution:  $p(\mathbf{n}_k) = G(\mathbf{n}_k; \mathbf{0}, \mathbf{\Sigma})$ , where  $\mathbf{\Sigma}$  is the covariance matrix. In the absence of prior knowledge about the correlation between the measured variables,  $\mathbf{\Sigma}$  is assumed to be a diagonal matrix:  $\mathbf{\Sigma} = \text{diag}(\sigma_1^2, \dots, \sigma_D^2)$ , where  $D$  is the dimension of the measured variables. To define the state space model, the process noise,  $\mathbf{v}_{k-1}$ , is also assumed to be Gaussian distributed with known covariance matrix:  $p(\mathbf{v}_{k-1}) = G(\mathbf{v}_{k-1}; \mathbf{0}, \mathbf{Q})$ . The Gaussian noise model for  $\mathbf{n}_k$  expresses the belief that there exists a certain level of “regular” noise (that is, natural noise inherent within the process measurements) with pre-specified covariance,  $\mathbf{\Sigma}$ . If there are outliers in the measurements, a

single Gaussian distribution will not account for the large variance of the ‘abrupt’ noise associated with the outliers. Thus to address the effect of outliers, a mixture model comprising two Gaussian distributions was proposed by Schick and Mitter (1994) that was adopted by Singhal and Seborg (2000):

$$\begin{aligned}
p(\mathbf{n}_k) &= (1 - \varepsilon) G(\mathbf{n}_k; \mathbf{0}, \Sigma) + \varepsilon G(\mathbf{n}_k; \mathbf{0}, b^2 \Sigma) \\
&= (1 - \varepsilon) \prod_{d=1}^D G(n_{k,d}; 0, \sigma_d^2) + \varepsilon \prod_{d=1}^D G(n_{k,d}; 0, b^2 \sigma_d^2)
\end{aligned} \tag{3}$$

where  $n_{k,d}$  is the noise for the  $d$ -th measured variable at time point  $k$ ,  $\varepsilon$  is the prior probability of outliers being present in the data, and  $b$  is a factor that accounts for the substantially larger standard deviation associated with the outliers.

Theoretically the above mixture model assumes *dependent* measurement noise, that is, the noise in one measured variable, is dependent on the noise in other measured variables. More formally:  $p(n_{k,1}, n_{k,2}, \dots, n_{k,D}) \neq \prod_{d=1}^D p(n_{k,d})$  (proof is given in Appendix 1). Thus the occurrence of an outlier in one measured variable is dependent on the occurrence of the outlier in other measured variables. However in practice, *independent* measurement noise may be appropriate. For example, if reactor temperature and feed flow rate are measured on a process, the measurement noise should be mutually independent, hence the outlier associated with temperature may be related to a temporary sensor malfunction and can be regarded as independent of flow rate. Based on the assumption of independence, this study adopts the noise model described in (Morad, Young & Svrcek, 2005):

$$p(\mathbf{n}_k) = \prod_{d=1}^D p(n_{k,d}) = \prod_{d=1}^D \left[ (1 - \varepsilon) G(n_{k,d}; 0, \sigma_d^2) + \varepsilon G(n_{k,d}; 0, b^2 \sigma_d^2) \right] \tag{4}$$

Let  $\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{u}_k)$  be the ‘‘noise-free’’ value of the process measurements. Based on this noise model, the likelihood,  $p(\mathbf{z}_k | \mathbf{x}_k)$ , is given by:

$$p(\mathbf{z}_k | \mathbf{x}_k) = \prod_{d=1}^D \left[ (1 - \varepsilon) G(y_{k,d} - z_{k,d}; 0, \sigma_d^2) + \varepsilon G(y_{k,d} - z_{k,d}; 0, b^2 \sigma_d^2) \right] \tag{5}$$

where  $y_{k,d}$  is the  $d$ -th variable of  $\mathbf{y}_k$ . According to Bayes' rule, the posterior distribution of the states given the measurements is proportional to the product of the prior and the likelihood:  $p(\mathbf{x}_k | \mathbf{z}_{1:k}) \propto p(\mathbf{x}_k | \mathbf{z}_{1:k-1})p(\mathbf{z}_k | \mathbf{x}_k)$ . Here  $\mathbf{z}_{1:k} = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k\}$  is the measurement sequence from time point 1 to  $k$ . Based on this posterior distribution, state estimation, outlier detection and removal can be achieved in a number of ways. For example Singhal and Seborg (2000) proposed the use of the EKF and the expectation-maximization (EM) algorithm to obtain the states and the rectified measurements, assuming the states are Gaussian distributed.

This paper proposes a more general formulation of the problem which does not assume any parametric probability distribution function for the states. The posterior distribution of the states,  $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ , is approximated by Monte Carlo samples using particle filters (see Section 3), which form the basis to obtain the rectified, "noise-free" process measurements,  $\mathbf{y}_k$ . By noting that  $\mathbf{y}_k$  is a function of  $\mathbf{x}_k$ , this paper defines the rectified measurements to be the expectation of  $\mathbf{y}_k$  with respect to the posterior distribution of  $\mathbf{x}_k$ :

$$\hat{\mathbf{y}}_k = \int \mathbf{y}_k p(\mathbf{x}_k | \mathbf{z}_{1:k}) d\mathbf{x}_k = \int \mathbf{h}_k(\mathbf{x}_k, \mathbf{u}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k}) d\mathbf{x}_k \quad (6)$$

Finally, discrete random variables,  $\{c_{k,d}, d = 1, \dots, D\}$ , are introduced to indicate whether the  $d$ -th measured variable,  $z_{k,d}$ , is an outlier:

$$c_{k,d} = \begin{cases} 0, & z_{k,d} \text{ is not an outlier} \\ 1, & z_{k,d} \text{ is an outlier} \end{cases} \quad (7)$$

Hence  $c_{k,d}$  is a Bernoulli random variable. The noise model in Eq. (4) states that the prior for these indicators is  $p(c_{k,d} = 1) = \varepsilon$ . Thus according to Bayes' rule, the probability that  $c_{k,d} = 1$ , given  $\mathbf{x}_k$  is:

$$\begin{aligned}
p(c_{k,d} = 1 | \mathbf{x}_k) &= \frac{p(c_{k,d} = 1)p(\mathbf{x}_k | c_{k,d} = 1)}{p(c_{k,d} = 0)p(\mathbf{x}_k | c_{k,d} = 0) + p(c_{k,d} = 1)p(\mathbf{x}_k | c_{k,d} = 1)} \\
&= \frac{\varepsilon G(y_{k,d} - z_{k,d}; 0, b^2 \sigma_d^2)}{(1 - \varepsilon) G(y_{k,d} - z_{k,d}; 0, \sigma_d^2) + \varepsilon G(y_{k,d} - z_{k,d}; 0, b^2 \sigma_d^2)} \tag{8}
\end{aligned}$$

Therefore the marginal posterior probability of  $c_{k,d} = 1$ , given the measurements  $\mathbf{z}_k$ , is the following integral as defined by the Chapman-Kolmogorov equation (Papoulis, 1984):

$$p(c_{k,d} = 1 | \mathbf{z}_k) = p(c_{k,d} = 1 | \mathbf{z}_{1:k}) = \int p(c_{k,d} = 1 | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k}) d\mathbf{x}_k \tag{9}$$

The value of  $p(c_{k,d} = 1 | \mathbf{z}_k)$  will be used to identify outliers in the process measurements. For a hard decision,  $p(c_{k,d} = 1 | \mathbf{z}_k) > 0.5$  would indicate the corresponding measured variable is an outlier. The integrals in Eq. (6) and (9) will be approximated using the same set of Monte Carlo samples that approximate the posterior of the states. The Monte Carlo method forms the basis of particle filters and is described in the next section.

### 3. Particle Filters for Dynamic Data Rectification

This section initially provides an overview of particle filters, based on the tutorial of Arulampalam et al. (2002), prior to describing the application of particle filters to the dynamic data rectification problem.

#### 3.1 Overview of Particle Filters

Particle filters are a class of statistical tools that sequentially estimate the system states from a state space model, that is, given the measurement sequence  $\mathbf{z}_{1:k}$ , the probability distribution function of the state  $\mathbf{x}_k$  can be inferred. The basic idea of particle filters is to approximate the distribution,  $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ , using a set of random samples (also called particles)  $\{\mathbf{x}_k^i, i = 1, \dots, N\}$  with associated weights  $\{w_k^i, i = 1, \dots, N\}$ :

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) \approx \sum_{i=1}^N w_k^i \delta(\mathbf{x}_k - \mathbf{x}_k^i) \quad (10)$$

where the weights sum to unity, and  $\delta(\mathbf{x})$  is an indicator function which is equal to unity if  $\mathbf{x} = \mathbf{0}$ ; otherwise it is equal to zero. Therefore the key step is to generate random samples from  $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ . However since  $p(\mathbf{x}_k | \mathbf{z}_{1:k})$  is not of the form of conventional probability distribution functions, direct sampling is not possible. Therefore importance sampling is applied to obtain the particles and associated weights (Robert & Casella, 1999, §3.3). The first step in importance sampling is to define an importance density  $q(\mathbf{x}_k | \mathbf{z}_{1:k})$  from which samples  $\mathbf{x}_k^i$  can be drawn (e.g. a standard Gaussian distribution function). Thus the weights are defined as:

$$w_k^i \propto \frac{p(\mathbf{x}_k^i | \mathbf{z}_{1:k})}{q(\mathbf{x}_k^i | \mathbf{z}_{1:k})} \quad (11)$$

For the sequential estimation problem, at time point  $k$ , the particles which approximate  $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})$  are passed through the state function which is updated with a new measurement,  $\mathbf{z}_k$  to approximate  $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ . It was shown in Arulampalam et al., (2002) that if the importance density is only dependent on the current measurement,  $\mathbf{z}_k$ , and the past state,  $\mathbf{x}_{k-1}$ , the weights can be updated as:

$$w_k^i \propto w_{k-1}^i \frac{p(\mathbf{z}_k | \mathbf{x}_k^i) p(\mathbf{x}_k^i | \mathbf{x}_{k-1}^i)}{q(\mathbf{x}_k^i | \mathbf{x}_{k-1}^i, \mathbf{z}_k)} \quad (12)$$

With these particles and associated weights, the mean of the states can be approximated by  $\hat{\mathbf{x}}_k = \sum_{i=1}^N w_k^i \mathbf{x}_k^i$ .

To implement particle filters, a number of issues need to be considered, including degeneracy, the selection of the importance density, and the number of particles required. These issues are now discussed briefly. Degeneracy is where, after a number of time points, only one particle has significant weight. Doucet (Doucet, 1998) showed that the variance of the importance weights increases over time, making degeneracy unavoidable. Thus considerable computational effort is



expended on updating particles whose contribution to the approximation of  $p(\mathbf{x}_k | \mathbf{z}_{1:k})$  is negligible. Re-sampling (Smith & Gelfand, 1992) has been used as a standard procedure to resolve this problem. In this case a new set of particles are generated by sampling with replacement from the original set  $\{\mathbf{x}_k^i, i = 1, \dots, N\}$  with probability  $\Pr(\mathbf{x}_k^j = \mathbf{x}_k^i) = w_k^i$ ,  $j$  denotes the particle index after re-sampling. The “updated relationship” is denoted, “parent( $j$ )= $i$ ”. The weights are re-set to  $1/N$  as the particles are independent and identically distributed and drawn from a discrete density function. By re-sampling, the particles with small weights will be eliminated.

The second issue is how to select the importance density. One approach is to use  $q(\mathbf{x}_k | \mathbf{x}_{k-1}^i, \mathbf{z}_k) = p(\mathbf{x}_k | \mathbf{x}_{k-1}^i)$ , which will yield a simple form for updating the weights according to Eq. (12):  $w_k^i \propto w_{k-1}^i p(\mathbf{z}_k | \mathbf{x}_k^i)$ . However, as noted by Pitt and Shephard (1999), this importance density may be sensitive to the presence of outliers, and can be improved if it depends on the current measurement. This idea was further developed by Pitt and Shephard (1999) who proposed that  $\boldsymbol{\mu}_k^i$ , the mean of  $p(\mathbf{x}_k | \mathbf{x}_{k-1}^i)$ , is first calculated and then the importance density is redefined as:

$$q(\mathbf{x}_k | \mathbf{z}_k) \propto p(\mathbf{z}_k | \boldsymbol{\mu}_k^i) p(\mathbf{x}_k | \mathbf{x}_{k-1}^i) w_{k-1}^i \quad (13)$$

By utilising  $\boldsymbol{\mu}_k^i$ , new particles are generated from particles at the previous time point, conditional on the current measurement  $\mathbf{z}_k$ , which will be closer to the true states. The calculation of weights (Pitt and Shephard, 1999) is thus given by:

$$w_k^i \propto \frac{p(\mathbf{z}_k | \mathbf{x}_k^i)}{p(\mathbf{z}_k | \boldsymbol{\mu}_k^i)} \quad (14)$$

Particle filters with this importance density and re-sampling step are termed Auxiliary Sampling Importance Re-sampling (ASIR) filters. ASIR filters are used in this study to address the dynamic data rectification problem.

Finally, the number of particles required is not only dependent on the state dimension, but also on the specific distribution of the states. If the states are mutually independent, the number of required

particles increases exponentially with the state dimension, as per point-mass filters. However, in practical problems, the states tend to be correlated and thus the joint distribution will be concentrated along some “curve”, implying lower dimensionality than that defined for independent states. Therefore fewer particles are required to approximate the joint distribution. In practice, the number of particles is typically decided empirically by undertaking a series of initial experiments.

### 3.2 Particle Filters for Dynamic Data Rectification

The previous subsection has shown how Monte Carlo samples (particles) can be used to approximate the posterior distribution of the states. With these particles and associated weights, the calculation of the rectified measurements and the probability of outliers being present, necessitates the calculation of integrals that can be approximated by Monte Carlo integration (Robert & Casella, 1999):

$$\hat{\mathbf{y}}_k = \int \mathbf{h}_k(\mathbf{x}_k, \mathbf{u}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k}) d\mathbf{x}_k \approx \sum_{i=1}^N w_k^i \mathbf{h}_k(\mathbf{x}_k^i, \mathbf{u}_k) = \sum_{i=1}^N w_k^i \mathbf{y}_k^i \quad (15)$$

$$\begin{aligned} p(c_{k,d} = 1 | \mathbf{z}_k) &= \int p(c_{k,d} = 1 | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k}) d\mathbf{x} \\ &\approx \sum_{i=1}^N w_k^i \frac{\varepsilon G(y_{k,d}^i - z_{k,d}; 0, b^2 \sigma_d^2)}{(1 - \varepsilon) G(y_{k,d}^i - z_{k,d}; 0, \sigma_d^2) + \varepsilon G(y_{k,d}^i - z_{k,d}; 0, b^2 \sigma_d^2)} \end{aligned} \quad (16)$$

where  $\mathbf{y}_k^i = \mathbf{h}_k(\mathbf{x}_k^i, \mathbf{u}_k)$ , and  $y_{k,d}^i$  is the  $d$ -th variable of vector  $\mathbf{y}_k^i$ . In summary, the particle filtering algorithm for the current time point  $k$  is as follows:

(i) For  $i = 1 : N$

Calculate  $\boldsymbol{\mu}_k^i = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}^i, \mathbf{u}_{k-1})$

Calculate  $w_k^i \propto p(\mathbf{z}_k | \boldsymbol{\mu}_k^i) w_{k-1}^i$

End

Normalize  $w_k^i$  such that they sum to 1.

(ii) Re-sample. The particles are now indexed by  $j$ .

(iii) For  $j = 1 : N$

Draw  $\mathbf{x}_k^j$  from  $p(\mathbf{x} | \mathbf{x}_{k-1}^{\text{parent}(j)})$ .

Assign the weights to be:  $w_k^i \propto p(\mathbf{z}_k | \mathbf{x}_k^j) / p(\mathbf{z}_k | \boldsymbol{\mu}_k^{\text{parent}(j)})$ .

End

Normalize  $w_k^j$  such that they sum to 1.

- (iv) Calculate the mean of the states, the rectified measurements (Eq. (15)), and the probability of a measurement being an outlier (Eq. (16)).

#### 4. Simulation Study

In this section, the particle filtering technique is evaluated through its application to an illustrative non-linear dynamic system, and a simulated benchmark pH neutralization process. The results of the conventional EKF are shown only for illustration, as it is not specifically designed for detecting and removing outliers. The results of the EKF with the EM algorithm (EKF-EM) (Singhal & Seborg, 2000) are presented for comparison. Based on related research reported in the literature (Arulampalam et al., 2002; Chen et al., 2004; Chen, Morris & Martin, 2005) and preliminary experiments in the present study, two hundred particles were used for the particle filters in all examples. The root mean squared error (RMSE) is quoted as the quantitative measure of performance, with the RMSE between the “noise-free” and noise-contaminated measurements also being calculated as the baseline result for data rectification. An effective rectification algorithm is expected to give a lower RMSE than that for the noise-contaminated measurements.

In the simulation study, the process model is assumed to be known to the filtering algorithms. In practice the model will not be perfect, and thus additional approaches will require to be utilized to address the mismatch between the model and the real process. For example, if the mechanistic model comprises a number of parameters that cannot be accurately estimated from experimental data, these parameters can be augmented into the process states for joint state and parameter estimation (Terwiesch & Agarwal, 1995; Chen, Morris & Martin, 2005). Furthermore, the initial values of the process states can also be estimated if they are not known exactly (Chen, Morris & Martin, 2004). An in-depth investigation of these issues is outwith the scope of this paper.

##### 4.1 A Non-linear Dynamic System

This illustrative example considers the following set of state and measurement equations:

$$x_k = \frac{x_{k-1}}{2} + \frac{25x_{k-1}}{1+x_{k-1}^2} + 8\cos(1.2k) + v_{k-1} \quad (17)$$

$$z_k = \frac{x_k^2}{20} + n_k \quad (18)$$

where  $v_{k-1}$  and  $n_k$  are zero mean Gaussian random variables, both of which have a standard deviation of 5. The initial state for the simulation is  $x_0 = 0$ . This example has been analyzed in a number of publications (Arulampalam et al., 2002; Carlin, Polson & Stoffer, 1992; Chen et al., 2004) as a benchmark for the dynamic state estimation task. It is assumed *a priori* that 10% of the measurements are possible outliers, that is  $\varepsilon = 0.1$ ; and the standard deviation of the outliers is 10 times larger than that for the regular noise ( $b=10$ ). Unless otherwise stated, these parameters are used to simulate the dynamic system and the pH neutralization process, described in the next subsection. In practice these values are typically estimated from prior knowledge about the sensors, such as manufacturing specifications, and historical process data.

Figure 1 presents the results for the rectification error,  $\Delta z_k = y_k - \hat{y}_k$ , and the detection of outliers, which have a Gaussian distribution and that match the mixture noise model defined in Section 2. The 99% noise limits (the two dashed lines in the upper plot of Figure 1) are established based on the noise standard deviation used for the simulation:  $\pm 2.58\sigma_d$ . It can be seen that the rectification errors for the particle filter approach lie within the 99% noise limits in most cases and are smaller than those for the EKF-EM (results for the EKF are not shown due to their larger errors). As given in Table 1, the RMSEs of the rectified measurements are 4.96, 10.52 and 52.10, for the particle filter approach, the EKF-EM and the EKF respectively, compared with 12.22 for the original measurements. The lower part of Figure 1 is a plot of the probability of an observation being an outlier versus the occurrence of outliers. Table 1 also shows that the particle filter approach successfully detects 6 of the 8 true outliers, whereas the EKF-EM detects 5; however the EKF-EM incorrectly denotes several measurements with regular noise as outliers, that is, false alarms materialise.

(Figure 1 and Table 1 about here)

The unsatisfactory performance of the EKF based rectification techniques is mainly due to the non-Gaussian properties of the states, hence the Gaussian assumption in the EKF is a poor approximation. The analyses in Arulampalam et al. (2002) and Chen et al. (2004) clearly showed that over time, the system states exhibited skewed and multi-modal distributions, even if the initial states were Gaussian distributed. Therefore the EKF based techniques are unable to provide accurate state estimation (the RMSEs of state estimation were 8.65, 16.33 and 17.29 for the particle filter, EKF-EM and EKF respectively, given in Table 1), and hence are observed to be unable to rectify the measurements effectively. In general good performance of data rectification and outlier detection cannot be achieved in isolation of accurate state estimation.

To investigate the robustness of the particle filters to the assumptions in the mixture noise model, two sets of additional simulations are considered. Firstly in practical situations, outliers may exhibit a non-Gaussian distribution. To denote this situation, the process is simulated where the outliers are sampled from an exponential distribution,  $p(n; \lambda) = \lambda e^{-\lambda n}$  with parameter  $\lambda^{-1} = b\sigma_d$ , this is equivalent to a standard deviation of  $b\sigma_d$ . The results are shown in Figure 2 and Table 2. The number of missed detections for both the particle filter and the EKF-EM is similar; however the EKF-EM results in 15 false alarms, in contrast to one false alarm when the particle filter is applied. The RMSE of the rectified measurements for the particle filter is 5.90, significantly smaller than that for the EKF-EM (RMSE=12.00). Consequently it can be concluded that the particle filter is capable of achieving reasonable rectification performance in this example when the distribution of outliers is non-Gaussian.

(Figure 2 and Table 2 about here)

The second study involves the modification of the probability of having Gaussian distributed outliers in the simulation (denoted by  $p_s$ ), while still retaining the prior of outliers being present within the rectification algorithms, i.e.  $\varepsilon = 0.1$ . The results are summarized in Tables 3 and 4, where for each case of  $p_s$ , the RMSE is averaged over 50 runs of the simulation. Table 3 summarises the RMSE for measurement rectification. For comparison, the RMSE for the measurements against their “noise-free” counterparts is included as the baseline case. It can be seen that the EKF fails to reduce the measurement error in all cases; the EKF-EM rectifies the data to some extent; and the particle filter achieves the lowest RMSE in all situations. The inferior performance of the EKF and the EKF-EM is again due to the non-Gaussian distributed system

states, as discussed previously. Even under the most extreme conditions, where approximately 90% of the measurements are defined to be outliers (this should not happen for normal processes in practice), the particle filter still achieves satisfactory performance, reducing the measurement error from 47.8 to 9.3, in terms of the RMSE. Table 4 summarises the corresponding RMSEs for state estimation. These results complement the conclusions drawn from Table 3, that is particle filters exhibit consistently superior performance compared to the EKF based data rectification techniques in this example, even if there exists considerable mismatch between the noise model assumptions ( $\varepsilon = 0.1$ ), and the real occurrence of outliers in the process measurements ( $p_s$  varies from 0.1 to 0.9).

(Table 3 and Table 4 about here)

#### 4.2 pH Neutralization Process

The pH neutralization process (McAvoy, Hsu & Lowenthal, 1972) takes place in a continuously stirred tank reactor shown in Figure 3. There are two input streams to the reactor, the acetic acid with concentration  $C_1$  at flow rate  $F_1$ , and the sodium hydroxide with concentration  $C_2$  at flow rate  $F_2$ . By assuming that the tank level is perfectly controlled, the mechanistic model of this process can be described as follows (McAvoy, Hsu & Lowenthal, 1972):

(Figure 3 about here)

$$\frac{d\xi}{dt} = \frac{F_1 C_1}{V} - \frac{(F_1 + F_2)\xi}{V} \quad (19)$$

$$\frac{d\zeta}{dt} = \frac{F_2 C_2}{V} - \frac{(F_1 + F_2)\zeta}{V} \quad (20)$$

$$\text{pH} = -\log_{10}[H^+] \quad (21)$$

The system states are given by  $\xi = [HAC] + [AC^-]$ , and  $\zeta = [Na^+]$ . Table 5 gives the definition and the values of the parameters used in the simulation. The only measured variable, pH, is indirectly related to the system states through the following equation:

(Table 5 about here)

$$[H^+]^3 + (K_a + \zeta)[H^+]^2 + (K_a(\zeta - \xi) - K_w)[H^+] - K_a K_w = 0 \quad (22)$$

In the simulation, the process noise is Gaussian with zero mean and diagonal covariance matrix:  $\text{Diag}(1e-6, 1e-6)$ . The standard deviation for the regular measurement noise is 0.2 and the initial states are  $(\xi, \zeta)_0 = (0.0432 \text{ mol/l}, 0.0432 \text{ mol/l})$ .

This neutralization process has a highly non-linear measurement function. Figure 4 shows a plot of the pH value as a function of the system states. It exhibits three distinct regions determined by the difference in the two states, that is  $|\xi - \zeta|$ . If  $|\xi - \zeta|$  is relatively large, the gradient of the pH with respect to the states is relatively small. On the other hand, if the two states are similar in magnitude, the gradient increases significantly. When the process moves from one region to another, for example due to a process disturbance, there will be a significant change in the gradients. This is equivalent to a large magnitude for the second-order derivatives of the pH with respect to the states. As a simplified example<sup>1</sup>, Figure 5 illustrates the pH value, its first-order and second-order derivatives with respect to  $\zeta$ , by fixing  $\xi = 0.04 \text{ mol/l}$ . It can be seen that the magnitude of the second-order derivative is large when  $\zeta$  is close to (but not equal to)  $0.04 \text{ mol/l}$ . At the extreme, that is when the second-order derivative attains its maximal value of  $1.88 \times 10^9$ , the corresponding first-order derivative is  $3.76 \times 10^4$ . Therefore, in this example, by only using the first-order derivative to approximate the non-linear functions, the EKF is expected to give low accuracy.

(Figure 4 and Figure 5 about here)

Figure 6 presents the results for the rectification error ( $\Delta\text{pH}$ ) and the detection of the Gaussian distributed outliers. Due to the poor approximation of the non-linearity by the EKF, the EKF-EM algorithm gives large rectification errors, and incorrectly identifies many measurements as outliers. The RMSE of the rectified data using the EKF-EM is 1.34, which is larger than that for the original measurements (RMSE = 0.57). On the other hand, the majority of the rectification errors for the particle filter are within the 99% noise limits, and result in a RMSE of 0.27. Figure 6 and Table 6

---

<sup>1</sup> This discussion serves as an illustrative example. A rigorous evaluation of the truncation error in the EKF requires the consideration of the Jacobian and Hessian matrices (even higher-order terms), as well as the state estimation error. See Chapter 9.6 in (Jazwinski, 1970) and references therein for a more comprehensive discussion.

show that the particle filter detects 8 of the 11 outliers, and gives no false alarms, as opposed to the 5 missed detections and 22 false alarms when applying the EKF-EM.

(Figure 6 and Table 6 about here)

The rectification error and detection of exponentially distributed outliers is illustrated in Figure 7 and summarised in Table 7. Similar to the previously analyzed non-linear dynamic system, the particle filter is observed to be reasonably robust to the mismatch of outlier distributions, and produces satisfactory rectified measurements. The particle filter successfully reduces the RMSE in the measurements from 0.93 to 0.35, whereas the EKF-EM can only achieve a RMSE of 1.60. The particle filter gives a similar missed detection value to that attained for the EKF-EM, but it results in only one false alarm, compared to 21 false alarms when using the EKF-EM.

(Figure 7 and Table 7 about here)

The final simulation considers the detection of a process change, which can be used as a surrogate for process fault detection and diagnosis. Unlike random outliers, an abnormal change tends to result in a systematic mismatch between the normal process model and the real process. This systematic process change can be identified by using a property of the Bernoulli distribution. Previously the Bernoulli random variables,  $c_{k,d}$ , were used to indicate that at time point  $k$ , the  $d$ -th measured variable was an outlier, with prior definition  $p(c_{k,d} = 1) = \varepsilon$ . The value of  $\varepsilon$  can be estimated from historical process data, and hence it is assumed to closely reflect reality. It is known (Papoulis, 1984) that the number of outliers,  $m$ , in a window of  $n$  time points, is a Binomial random variable with the following probability density function:

$$p(m; \varepsilon, n) = \binom{n}{m} \varepsilon^m (1 - \varepsilon)^{n-m} \quad (23)$$

Therefore a  $100(1 - \alpha)\%$  confidence limit can be established for  $m$ , that denotes the maximum number of outliers allowed,  $m_{100(1-\alpha)\%}$ , in a window of size  $n$ :



$$m_{100(1-\alpha)\%} = \max m, \quad \text{such that: } \sum_{g=0}^m \binom{n}{g} \varepsilon^g (1-\varepsilon)^{n-g} < 1-\alpha \quad (24)$$

The above confidence limit can be extended to consider multiple measured variables, by assuming independence between different variables with respect to the occurrence of outliers. This approach was discussed in Abraham & Chuang (1993) and Singhal & Seborg (2000). For  $\varepsilon = 0.1$  and  $n = 10$ , the values used in the present study, the maximum number of outliers allowed at a 99% confidence level is  $m_{0.99} = 3$ .

To simulate a process change in the pH neutralization reactor, a ramp,  $r(t)$ , was added to the first system state,  $\xi$ , as follows:

$$r(t) = \begin{cases} 0 & 0 \leq t < 50 \text{ min.} \\ 0.001(t - 50) & t \geq 50 \text{ min.} \end{cases} \quad (25)$$

Figure 8 denotes the detection of this ramp disturbance. The number of outliers is given by the number of observations that satisfy  $p(c_{k,d} = 1 | \mathbf{z}_k) > 0.5$ . The lower plot of Figure 8 shows that by using a particle filter, the ‘required’ number of outliers exceeds the 99% limit at time point 60 min., that is, the detection delay is 9 min. This result is acceptable since at the beginning of the process change, the magnitude of the ramp disturbance was relatively small. The upper plot also shows that the rectification error for the particle filter is still within the 99% noise limits until 57 min. On the other hand, due to its poor approximation of the highly non-linear measurement function, the EKF-EM results in a large number of false alarms of outliers (middle plot of Figure 8) well before 50 min. Hence the EKF-EM mistakenly identifies the process as exhibiting non-conforming behaviour at time point 18 min., when it is known that the process was running under normal conditions.

(Figure 8 about here)

## 5. Conclusions

This paper applied particle filters for on-line dynamic data rectification, which is aimed at addressing three tasks simultaneously: accurate state estimation, the detection of outliers or process changes, and the removal of measurement noise. In situations where the EKF is not applicable due

to highly non-linear behaviour and non-Gaussian distribution of states, particle filters were observed to achieve significantly enhanced results in terms of improved state estimation and measurement rectification. The proposed approach has been further shown to be able to detect process changes, which is potentially useful in model-based process fault detection.

The proposed methodology is potentially applicable for process fault detection and diagnosis, where the difference between the rectified and the actual process measurements,  $\hat{\mathbf{y}}_k - \mathbf{z}_k$ , could be used to facilitate fault diagnosis. Furthermore, in situations where historical process data is available for various fault scenarios, a pattern recognition system may be developed in conjunction with the proposed methodology to provide accurate identification and diagnosis of process faults. These issues will be investigated in future work.

### **Acknowledgments**

T. Chen would like to acknowledge the financial support from the UK EPSRC grant ‘‘Vertical Integration of Product Development and Manufacturing’’ (GR/R64407/01).and the UK ORS Award for his PhD study.

### **Appendix: Further Discussions on the Noise Model**

This appendix provides more detailed discussions on the dependence issue of the noise model in Eq. (3) and (4). The main conclusion is the following lemma:

*Lemma 1:* The random variables,  $(n_{k,1}, n_{k,2}, \dots, n_{k,D})$ , are independently distributed according to Eq. (3) if and only if:  $\varepsilon = 0$  or  $\varepsilon = 1$  or  $D=1$ .

*Proof:* Firstly, for the reason of clarity, the subscript for time index,  $k$ , is neglected in the proof. By definition,  $(n_1, n_2, \dots, n_D)$  are independently distributed if and only if  $p(n_1, n_2, \dots, n_D) = \prod_{d=1}^D p(n_d)$  holds for any possible values of  $(n_1, n_2, \dots, n_D)$ . To examine this condition for Eq. (3), the marginal distribution of  $n_d$  is calculated as follows:

$$\begin{aligned}
p(n_d) &= \int p(n_1, n_2, \dots, n_D) d(n_1, \dots, n_{d-1}, n_{d+1}, \dots, n_D) \\
&= (1-\varepsilon)G(n_d; 0, \sigma_d^2) \prod_{\substack{g=1 \\ g \neq d}}^D \int G(n_g; 0, \sigma_g^2) dn_g + \varepsilon G(n_d; 0, b^2 \sigma_d^2) \prod_{\substack{g=1 \\ g \neq d}}^D \int G(n_g; 0, b^2 \sigma_g^2) dn_g \\
&= (1-\varepsilon)G(n_d; 0, \sigma_d^2) + \varepsilon G(n_d; 0, b^2 \sigma_d^2)
\end{aligned} \tag{26}$$

The above derivation uses the fact that the integral of any probability distribution function is equal to one. The product of the marginal distribution is:

$$\prod_{d=1}^D p(n_d) = \prod_{d=1}^D [(1-\varepsilon)G(n_d; 0, \sigma_d^2) + \varepsilon G(n_d; 0, b^2 \sigma_d^2)] \tag{27}$$

Comparing the above equation with Eq. (3) shows that  $p(n_1, n_2, \dots, n_D) = \prod_{d=1}^D p(n_d)$  holds if and only if:  $\varepsilon = 0$  or  $\varepsilon = 1$  or  $D=1$ , and Lemma 1 is proved.  $\square$

The above procedure also shows that the marginal distributions of  $n_d$  are the same for Eq. (3) and (4), and the probability distribution in Eq. (4) clearly satisfies the condition of independence. In the situations where  $\varepsilon = 0$  or  $\varepsilon = 1$ , the mixture model in Eq. (3) degenerates to only one Gaussian distribution, and thus is not of interests in data rectification. Finally it should be noted that in the two simulation examples in this paper, the measurement is one-dimensional ( $D=1$ ), and thus the noise models in Eq. (3) and (4) are the same.

## References

- Abraham, B., Chuang, A. (1993). Expectation-Maximization Algorithm and the Estimation of Time Series Models in the Presence of Outliers, *Journal of Time Series Analysis*, 14, 221-234.
- Abu-el-zeet, Z. H., Becerra, V. M., Roberts, P. D. (2002). Combined Bias and Outlier Identification in Dynamic Data Reconciliation. *Computers and Chemical Engineering*, 26, 921-935.
- Arulampalam, M., Maskell, S., Gordon, N., Clapp, T. (2002). A Tutorial on Particle Filters for On-line Non-linear/Non-gaussian Bayesian Tracking. *IEEE Transactions on Signal Processing*, 50, 174-188.
- Bucy, R. S., Senne, K. D. (1971). Digital Synthesis of Non-linear Filters, *Automatica*, 7, 287-298.
- Carlin, B. P., Polson, N. G., Stoffer, D. S. (1992). A Monte Carlo approach to nonnormal and nonlinear state-space modelling. *Journal of American Statistical Association*, 87, 493-500.

- Chen, T., Morris, J., Martin, E. (2004) Particle filters for the estimation of a state space model. In *European Symposium on Computer-Aided Process Engineering 14. Computer-Aided Chemical Engineering*; Barbosa-Póvoa, A., Matos, H., Eds.; Elsevier BV: Lisbon, Portugal, 2004; Vol 18, pp 613-618.
- Chen, T., Morris, J., Martin, E. (2005) Particle Filters for State and Parameter Estimation in Batch Processes. *Journal of Process Control*, 15, 665-673.
- Chen, W.-S., Bakshi, B. R., Goel, P. K., Ungarala, S. (2004). Bayesian Estimation via Sequential Monte Carlo Sampling: Unconstrained Nonlinear Dynamic Systems. *Industrial and Engineering Chemistry Research*, 43, 4012-4025.
- Darouach, M., Zasadzinski, M. (1991). Data Reconciliation in Generalized Linear Dynamic Systems. *AIChE Journal*, 37, 193-201.
- Doucet, A. (1998) On Sequential Monte Carlo Methods for Bayesian Filtering. Technical Report, Department of Engineering, University of Cambridge, UK.
- Jazwinski, A. H. (1970) Stochastic Processes and Filtering Theory. Academic Press, New York.
- Julier, S. J., Uhlmann, J., Durrant-Whyte, H. F. (2000). A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators. *IEEE Transactions on Automatic Control*, 45, 477-482.
- Liebman, M. J., Edgar, T. F., Lasdon, L. S. (1992). Efficient Data Reconciliation and Estimation for Dynamic Processes Using Nonlinear Programming Techniques. *Computers and Chemical Engineering*, 16, 963-986.
- McAvoy, T. J., Hsu, E., Lowenthal, S. (1972). Dynamics of pH in Controlled Stirred Tank Reactor. *Industrial and Engineering Chemistry Process Design and Development*, 11, 68-70.
- Morad, K., Young, B. R., Svrcek, W. Y. (2005). Rectification of Plant Measurements Using a Statistical Framework. *Computers and Chemical Engineering*, 29, 919-940.
- Papoulis, A. (1984). Probability, Random Variables, and Stochastic Processes. 2nd Edition, New York: McGraw-Hill.
- Pitt, M., Shephard, N. (1999). Filtering via Simulation: Auxiliary Particle Filters. *Journal of American Statistical Association*, 94, 590-599.
- Robert, C., Casella, G. (1999). Monte Carlo Statistical Methods. Springer, New York.
- Romanenko, A., Castro, J. (2004). The Unscented Filter as an Alternative to the EKF for Nonlinear State Estimation: A Simulation Case Study. *Computers and Chemical Engineering*, 28, 347-355.
- Schick, I. C., Mitter, S. K. (1994). Robust Recursive Estimation in the Presence of Heavy-Tailed Observation Noise. *Annals of Statistics*, 22, 1045-1080.

- Singhal, A., Seborg, D. E. (2000). Dynamic Data Rectification Using the Expectation Maximization Algorithm. *AIChE Journal*, 46, 1556-1565.
- Smith, A. F. M., Gelfand, A. E. (1992). Bayesian Statistics without Tears: a Sampling-Resampling Perspective. *American Statistician*, 46, 84-88.
- Sorenson, H. W. (1988). Recursive Estimation for Non-linear Dynamic Systems. in Spall, J. C. (Eds.), *Bayesian Analysis of Time Series and Dynamic Models*, Marcel Dekker, New York, 127-165.
- Terwiesch, P., Agarwal, M. (1995). A Discretized Non-linear State Estimator for Batch Processes. *Computers and Chemical Engineering*, 19, 155-169.
- Tong, H., Crowe, C. M. (1997). Detecting Persistent Gross Errors by Sequential Analysis of Principal Components. *AIChE Journal*, 43, 1242-1249.
- Wan, E. A., van der Merwe, R. (2000). The Unscented Kalman Filter for Nonlinear Estimation. In *Proc. of Symposium on Adaptive Systems for Signal Processing, Communication, and Control*, Lake Louise, Canada.

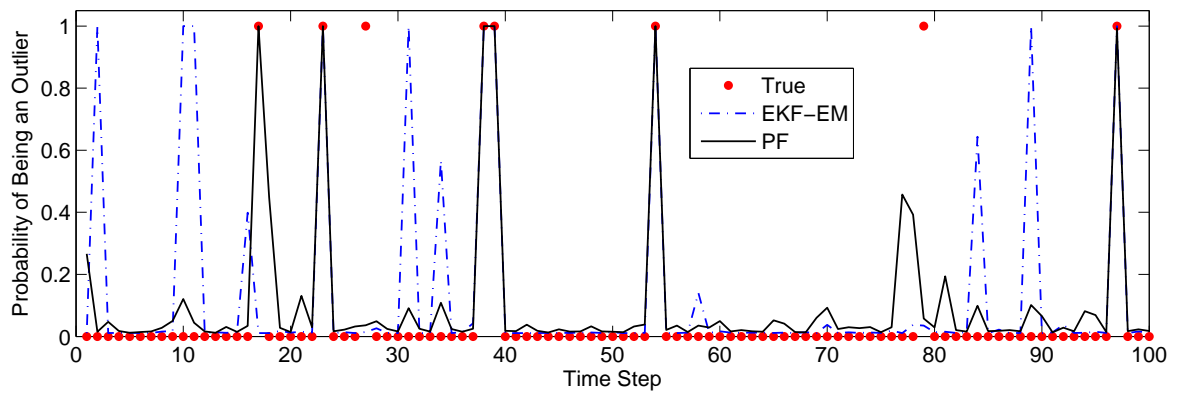
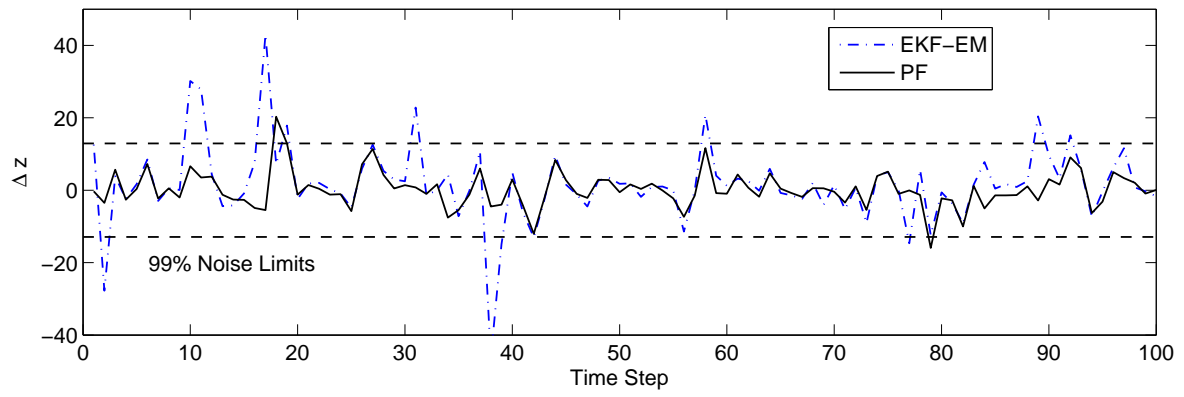


Figure 1: Rectification results for the non-linear dynamic system with outliers having a Gaussian distribution.

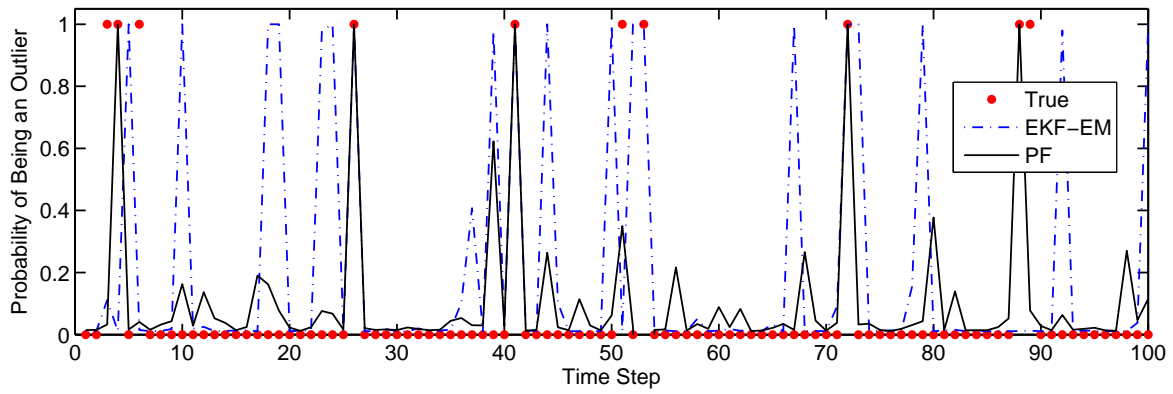
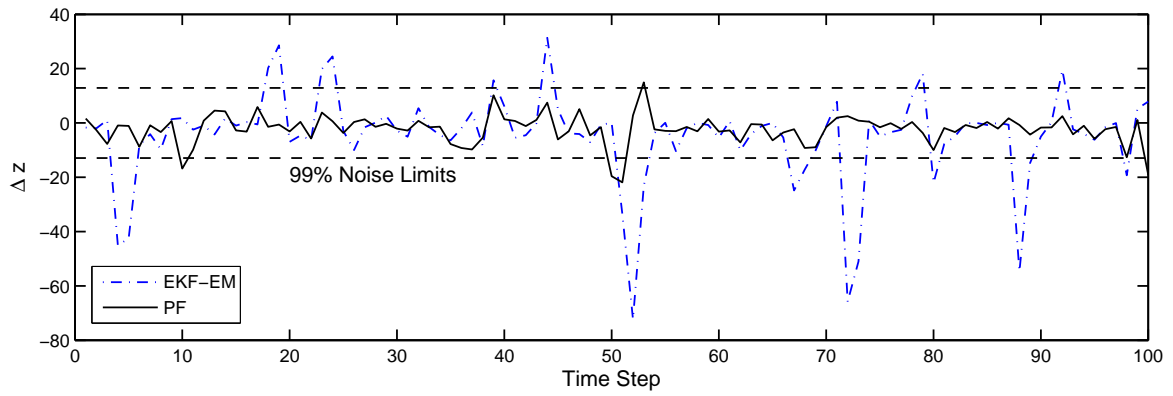


Figure 2: Rectification results for the non-linear dynamic system with outliers having an exponential distribution.

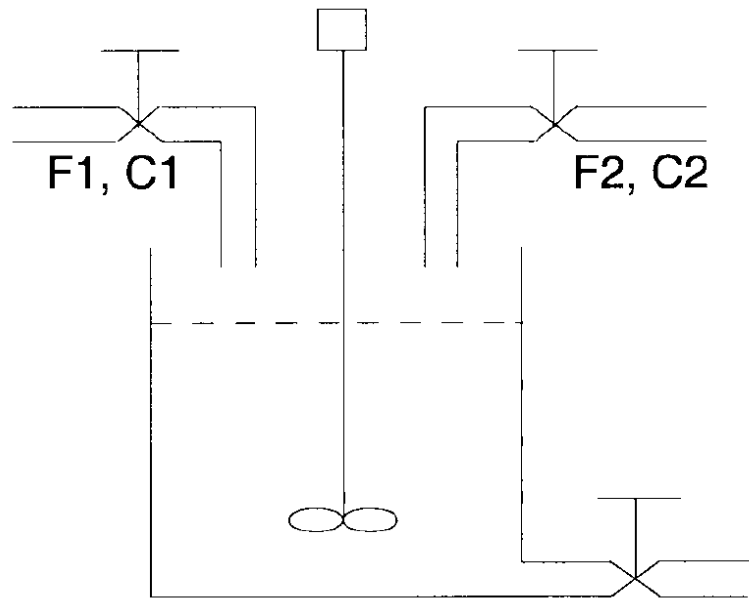


Figure 3: A continuous stirred tank reactor for pH neutralization.



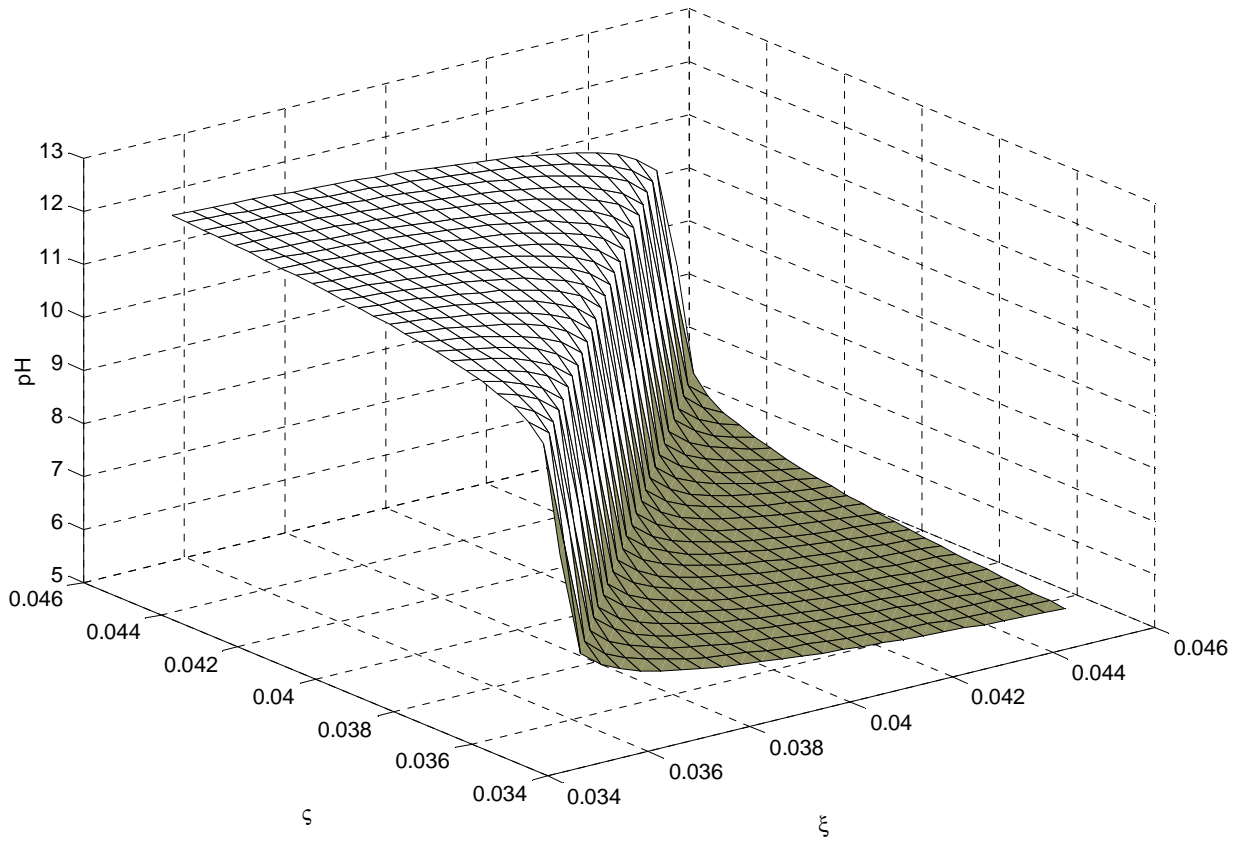


Figure 4: pH as a function of the states.

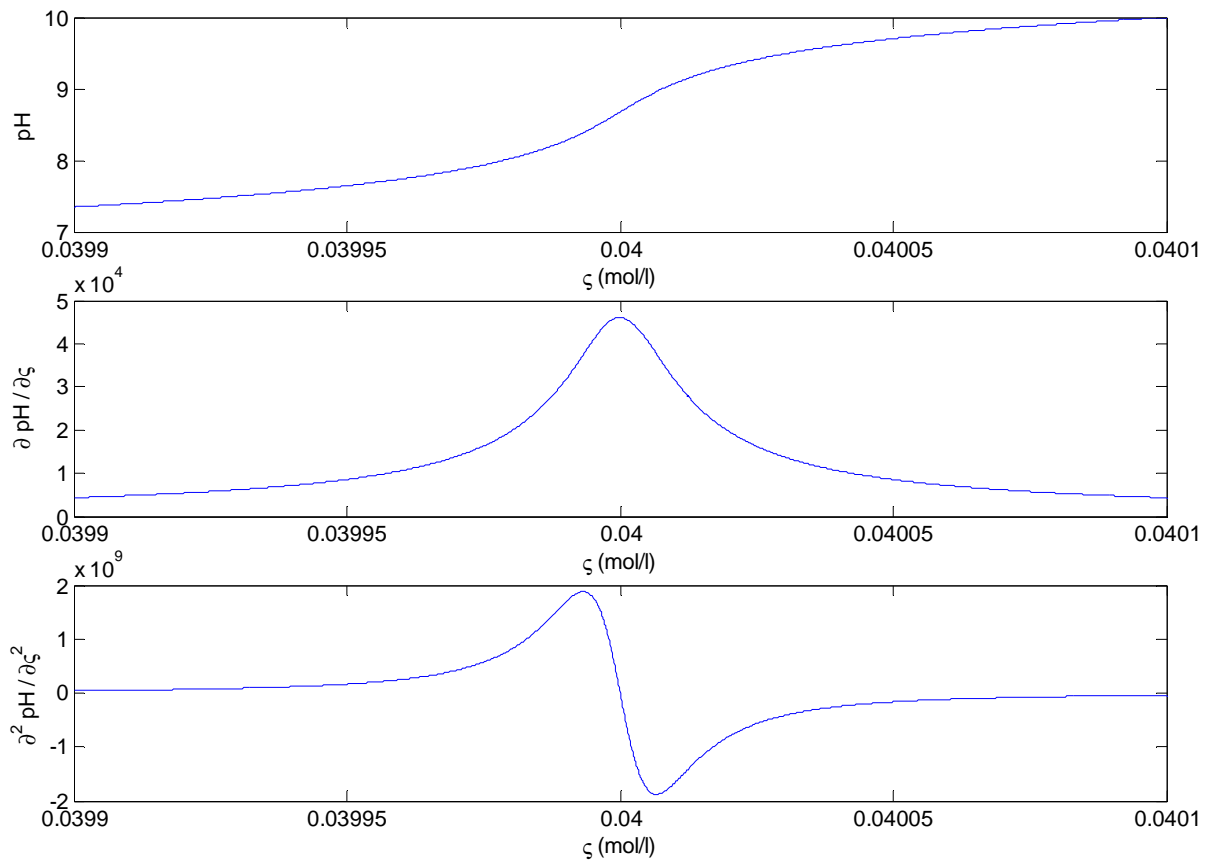


Figure 5: The value of pH, its first and second derivatives with respect to  $\zeta$  in the range from 0.0399 mol/l to 0.0401 mol/l, by fixing  $\xi = 0.04$  mol/l.

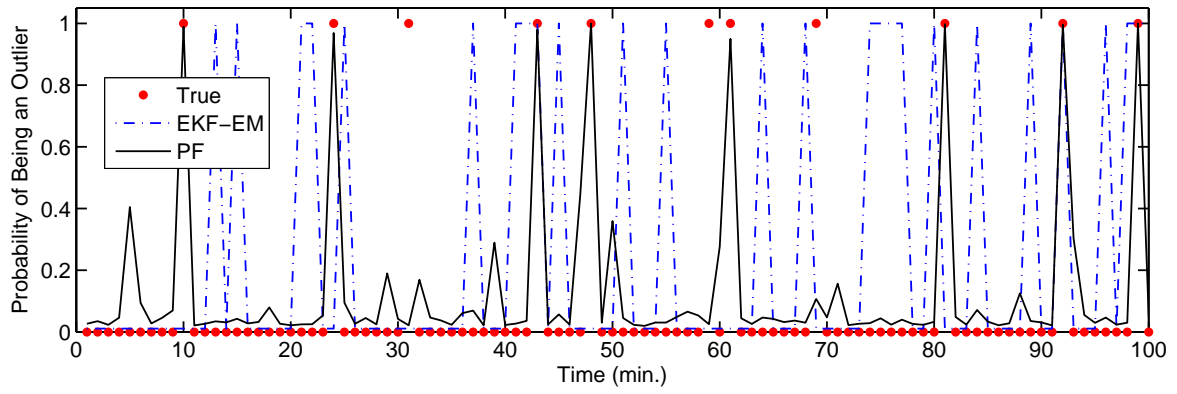
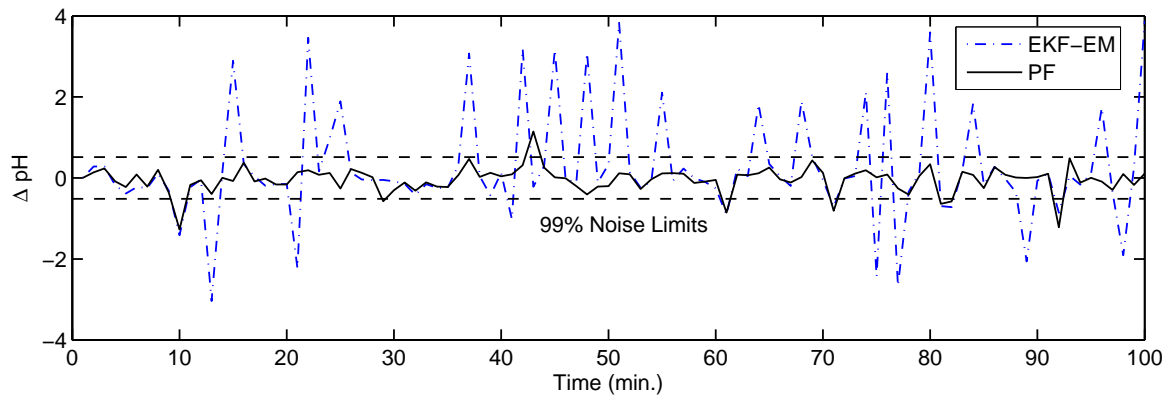


Figure 6: Rectification results for the pH neutralization process with outliers having a Gaussian distribution.

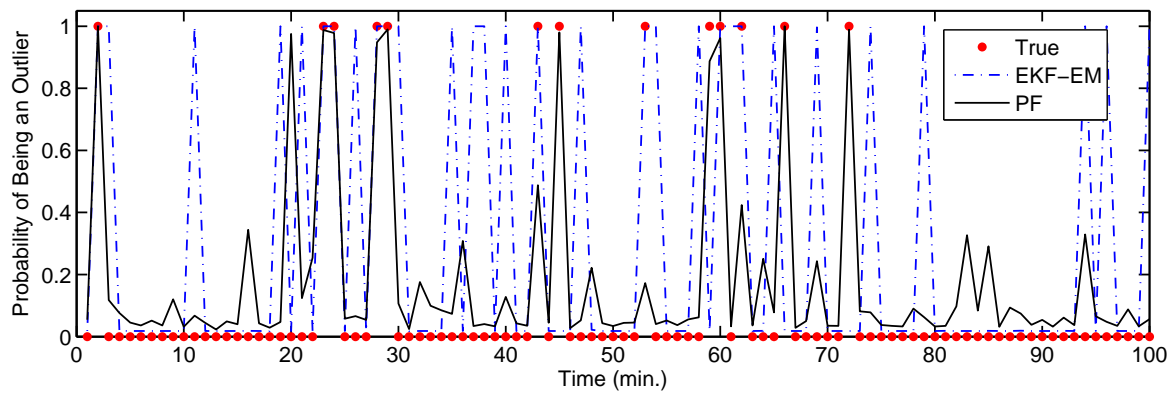
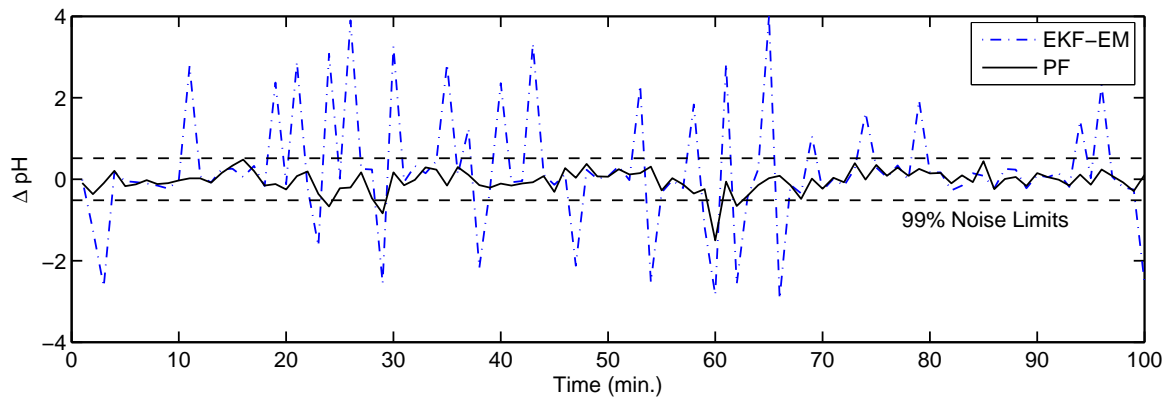


Figure 7: Rectification results for the pH neutralization process with outliers having an exponential distribution.

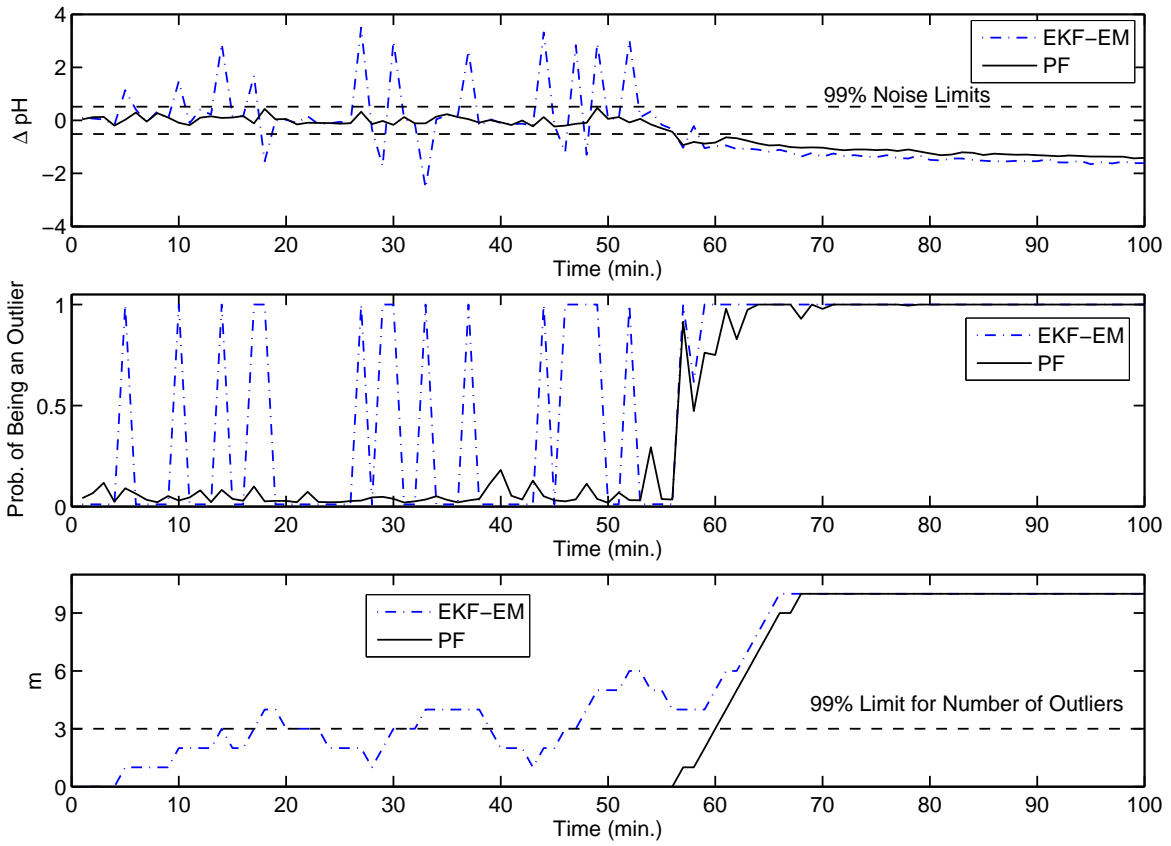


Figure 8: Detection of ramp disturbance for the pH neutralization process.

Table 1: Rectification results for the non-linear dynamic system with 8 outliers having a Gaussian distribution. RMSE-Rec: RMSE for rectified measurement; RMSE-State: RMSE for state estimation.

	RMSE-Rec	RMSE-State	False Alarm	Missed Detection
EKF	52.10	17.29	-	-
EKF-EM	10.52	16.33	7	3
PF	4.96	8.65	0	2

Table 2: Rectification results for the non-linear dynamic system with 10 outliers having an exponential distribution. RMSE-Rec: RMSE for rectified measurement; RMSE-State: RMSE for state estimation.

	RMSE-Rec	RMSE-State	False Alarm	Missed Detection
EKF	55.04	24.11	-	-
EKF-EM	12.00	19.18	15	4
PF	5.90	9.02	1	5

Table 3: Measurement rectification results (RMSE) for the non-linear dynamic system with different probabilities ( $p_s$ ) of outliers being present in the simulation.

$p_s$	0.1	0.3	0.5	0.7	0.9
Measured	15.3	26.8	35.6	41.6	47.8
EKF	290.9	366.5	416.8	488.2	424.5
EKF-EM	9.5	12.9	15.8	17.2	18.8
PF	4.7	6.0	7.1	8.4	9.3



Table 4: State estimation results (RMSE) for the non-linear dynamic system with different probabilities ( $p_s$ ) of outliers being present in the simulation.

$p_s$	0.1	0.3	0.5	0.7	0.9
EKF	24.0	29.5	35.6	38.7	40.3
EKF-EM	18.1	19.4	20.6	22.6	24.1
PF	8.0	8.7	9.5	10.3	11.1

Table 5: The physical parameters used in the simulation of the pH neutralization process.

Parameter	Definition	Value
$V$	Volume of the tank reactor	1 l
$F_1$	Flow rate of acid	0.081 l/min
$F_2$	Flow rate of base	0.512 l/min
$C_1$	Concentration of acetic acid in $F_1$	0.32 mol/l
$C_2$	Concentration of sodium hydroxide in $F_2$	0.0505 mol/l
$K_a$	Acid equilibrium constant	$1.8 \times 10^{-5}$
$K_w$	Water equilibrium constant	$1.0 \times 10^{-14}$

Table 6: Rectification results for the pH neutralization process with 11 outliers having a Gaussian distribution. RMSE-Rec: RMSE for rectified measurement; RMSE-State: average RMSE for the estimation of the two states.

	RMSE-Rec	RMSE-State ( $\times 10^{-3}$ )	False Alarm	Missed Detection
EKF	1.46	2.65	-	-
EKF-EM	1.34	0.93	22	5
PF	0.27	0.78	0	3

Table 7: Rectification results for the pH neutralization process with 13 outliers having an exponential distribution. RMSE-Rec: RMSE for rectified measurement; RMSE-State: average RMSE for the estimation of the two states.

	RMSE-Rec	RMSE-State ( $\times 10^{-3}$ )	False Alarm	Missed Detection
EKF	1.86	2.86	-	-
EKF-EM	1.60	1.03	21	4
PF	0.35	0.75	1	3