

# Identification of Noise Covariances for State Estimation of a Continuous Fermenter using Extended EM Algorithm

Vinay A. Bavdekar\* Sachin C. Patwardhan\*,<sup>1</sup>

\* *Department of Chemical Engineering, Indian Institute of Technology  
Bombay, Mumbai, India*

---

**Abstract:** Despite developments in sensor technology, monitoring a biological process using regular sensor measurements is often difficult. Development of Bayesian state observers, such as extended Kalman filter (EKF), is an attractive alternative for soft-sensing of such complex systems. The performance of EKF is dependent on the accurate characterisation of the uncertainties in the state dynamics and in the measurements. In this work, an extended expectation maximisation (EM) algorithm is developed for estimation of the state and measurement noise covariances for the EKF using irregularly sampled multi-rate measurements. The efficacy of the proposed approach is demonstrated on a benchmark continuous fermenter system. The simulation results reveal that the proposed approach generates fairly accurate estimates of the noise covariances.

Keywords: Expectation Maximisation algorithm, nonlinear state estimation, extended Kalman filter, covariance identification, multi-rate sampled data systems

---

## 1. INTRODUCTION

The stringent quality control requirements on various products of fermentation processes make monitoring and control a crucial activity. Monitoring is a critical activity from an economic point of view. It facilitates in detection of abnormal behaviour of the process and helps in preventing production of off-quality products. Despite developments in sensor technology, monitoring a biological process using regular sensor measurements is often difficult. Concentration sensors are expensive and it is difficult to measure the components that are dilute in the bio-reactor. In some situations, these measurements are available after considerable delay and/or at irregular intervals through lab assays. Only a few of the crucial measurements such as concentration of dissolved oxygen and pH are available online and at regular intervals (Dochain, 2003). However, from the viewpoint of monitoring and control it is important to generate estimates of unmeasured / irregularly measured variables on-line at the fast rate. Dynamic model based estimation of unmeasured states or state observer is an attractive alternative for developing soft sensors for such complex systems.

Unstructured mechanistic models appear to be ideal candidates for the development of state estimators for a bio-reactor. These models are valid over a wide range of operating conditions and also provide a better insight into the evolution of the plant. State observers can be designed by using either the deterministic approach (Dochain, 2003) or the Bayesian approach (Rawlings and Bakshi, 2006). Deterministic approaches of observer design often exploit special structures of the nonlinearity and are, therefore,

applicable to only a certain class of nonlinear systems. On the other hand, the Bayesian approaches do not depend on the structure of the systems, and, hence, can be applied to a wider class of nonlinear systems. Moreover, it is easier to adapt the Bayesian approaches to deal with irregularly sampled multi-rate measurement scenario.

A critical aspect in the development of Bayesian state estimators is to have correct characterisation of the unmeasured disturbances. An incorrect choice of noise characteristics leads to deterioration in the performance of the state estimator and in the worst case, the estimator may diverge. A commonly used assumption in the development of Bayesian state estimation is that the state and measurement disturbances are additive, zero mean and Gaussian. Thus, the problem of noise characterisation is reduced to estimation of the covariances of these multivariate Gaussian distributions.

The methods for estimating the state and measurement noise covariances for the Kalman filter (KF) can be classified into four categories— covariance matching, correlation techniques, Bayesian and maximum likelihood methods. Covariance matching techniques compute the covariances from the residuals of the system. Correlation techniques estimate the covariance matrices by making use of the correlation between the system output or innovations. Mehra (1970) developed a recursive linear least squares algorithm, to estimate the state and measurement noise covariances,  $Q$  and  $R$ , respectively. This approach is restrictive as the number of unknowns in  $Q$  should be less than  $n \times p$ , where,  $n$  is the number of states and  $p$  is the number of outputs. Odelson et al. (2006) present an autocovariance least-squares method to estimate  $Q$  and  $R$  for the KF. The algorithm defines an autocovariance function between the

<sup>1</sup> Corresponding Author; Email: sachinp@iitb.ac.in, Tel: +91-22-2576-7211, Fax: +91-22-2572-6895

outputs and develops a linear least squares formulation to estimate  $Q$  and  $R$ .

Bayesian methods and maximum likelihood methods usually use traditional gradient-based numerical optimisation schemes to estimate the covariance matrices. The requirement to compute Jacobian / Hessian at each iteration in these methods can prove to be computationally intensive task as the objective function (like the MLE function) is highly nonlinear with respect to the parameters to be identified. The expectation-maximisation (EM) algorithm, developed by Dempster et al. (1977), on the other hand, is a derivative free iterative method for computation of maximum likelihood estimates of the model parameters. In particular, this method can be applied even when measurements are available at irregular intervals. Shumway and Stoffer (2000) and Raghavan et al. (2006) provide a framework to identify state space models along with the state and measurement noise covariances for linear systems using the EM algorithm and irregularly sampled data.

Most of the work reported in literature is for estimating densities of noise in linear systems using the KF. However, since most biochemical processes are nonlinear, it is essential to use nonlinear state estimators for state estimation. There is very little work reported in the literature on the density estimation for nonlinear observers. In practice, nonlinear filters are tuned using heuristic approaches and by imposing diagonal parametrization of  $Q$  and  $R$ . Recently, Goodwin and Agüero (2005) have presented a modification of the EM algorithm for identification of nonlinear systems.

In this work it is proposed to use the extended Kalman filter (EKF), which is arguably the most widely used Bayesian estimator for nonlinear systems. EKF can be viewed as an approximate and computationally tractable sub-optimal solution to the sequential Bayesian estimation problem under the simplifying assumption that estimation error densities can be approximated as Gaussian. Valappil and Georgakis (2000) present a systematic method to estimate  $Q$  as arising from uncertainty in model parameters. Their approach assumes a structural uncertainty in the model that can be captured as variations in the parameters.  $Q$  is modelled as a function of the covariance of model parameter errors. The estimate of  $Q$  can either be obtained through successive linearisation of the system with respect to the model parameters or through Monte Carlo simulations of the parameters. It is assumed that the model parameter error covariance is known *a priori* and is constant. While this approach simplifies the tuning process, in practice the parameter error covariance matrix may not be known.

The EM algorithm has been successfully used for parameter identification of linear systems. It is possible to use the EM approach to estimate  $Q$  and  $R$  matrices associated with an EKF formulation, if some modifications are made in the expectation step of the EM algorithm. In this work, an extended EM algorithm is developed for estimation of  $Q$  and  $R$  for the EKF, when some of the data is irregularly sampled. The efficacy of the proposed approach is demonstrated by simulating a benchmark continuous fermenter system.

The organisation of the paper is as follows. In Section 2, the model and the assumptions involved for state estimation are described. The expressions for extended Kalman filter and extended Kalman smoother for irregularly sampled data are described in Section 3. The proposed extended EM-algorithm is described in Section 4. In Section 5, the results obtained by applying the extended EM-algorithm on a benchmark continuous fermenter are presented.

## 2. PRELIMINARIES

For continuous fermenters, the general unstructured mechanistic model can be written as

$$\begin{aligned}\dot{\mathbf{x}}(t) &= f(\mathbf{x}, \mathbf{u}, \mathbf{w}, t) \\ \mathbf{y}(t) &= h(\mathbf{x}, t)\end{aligned}\quad (1)$$

where,  $\mathbf{x} \in \mathbb{R}^n$  denotes the states,  $\mathbf{u} \in \mathbb{R}^m$  denotes the inputs,  $\mathbf{w} \in \mathbb{R}^l$  denotes the process noise and  $\mathbf{y} \in \mathbb{R}^p$  denotes the outputs.

For the purpose of developing a discrete state estimator, this model is discretised under the following assumptions: (a) the manipulated inputs are piecewise constant and (b) sampling time is assumed to be small enough so that the unmeasured disturbances can be treated to be piece-wise constant functions. Thus, the true dynamics of the process is assumed to evolve as follows

$$\begin{aligned}\mathbf{x}_{k+1} &= F(\mathbf{x}_k, \mathbf{u}_k) + \Gamma_d \mathbf{w}_k \\ \mathbf{y}_k &= h(\mathbf{x}_k) + \mathbf{v}_k\end{aligned}\quad (2)$$

where,  $F(\cdot)$  is the discrete time equivalent of  $f(\cdot)$ . The process noise,  $\mathbf{w}_k \sim \mathcal{N}(0, Q)$  is assumed to be an additive Gaussian white noise. If the state noise is modelled as entering the dynamics through some inputs, such as manipulated inputs, then,  $\Gamma_d = \Gamma_u = \left[\frac{\partial F}{\partial \mathbf{u}}\right]_{\mathbf{u}=\mathbf{u}_k}$ , else  $\Gamma_d = I_n$  and  $\mathbf{v}_k \sim \mathcal{N}(0, R)$  represents the measurement noise. It is also assumed that  $\mathbf{w}_k$  and  $\mathbf{v}_k$  are independent, identically distributed random variables and are mutually uncorrelated. In this work it is assumed that  $\Gamma_d = I_n$ .

## 3. THE EXTENDED KALMAN FILTER AND SMOOTHER

The EKF algorithm is made of two step namely, prediction and measurement update. The nonlinear equations of the plant model are linearised at every instant for update of estimation error covariance matrices. For the discrete time EKF, with initial conditions  $\hat{\mathbf{x}}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, P_0)$ , the prediction step is given by

$$\begin{aligned}\hat{\mathbf{x}}_{k+1|k} &= F(\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k) \\ P_{k+1|k} &= \Phi_k P_{k|k} \Phi_k^T + Q\end{aligned}\quad (3)$$

where,  $\Phi_k = \left[\frac{\partial F}{\partial \mathbf{x}}\right]_{\mathbf{x}=\hat{\mathbf{x}}_{k|k}}$ . The measurement update or the correction step is given by

$$\begin{aligned}\mathbf{e}_k &= \mathbf{y}_k - h(\hat{\mathbf{x}}_{k|k-1}) \\ K_k &= P_{k|k-1} C_k^T [C_k P_{k|k-1} C_k^T + R]^{-1} \\ \mathbf{x}_{k|k} &= \mathbf{x}_{k|k-1} + K_k \mathbf{e}_k \\ P_{k|k} &= [I - K_k C_k] P_{k|k-1}\end{aligned}\quad (4)$$

where,  $C_k = \left[\frac{\partial h}{\partial \mathbf{x}}\right]_{\mathbf{x}=\hat{\mathbf{x}}_{k|k-1}}$ ,  $K_k$  denotes the Kalman gain and  $\mathbf{y}_k$  is the measurement.

The extended Kalman smoother calculates the smoothed estimates based on the entire set of measurements available. With initial conditions as  $\hat{\mathbf{x}}_{N|N}$  and  $P_{N|N}$ , the expressions for the smoother  $\forall k = N - 1, N - 2, \dots, 1$  are written as

$$\begin{aligned} J_{k-1} &= P_{k-1|k-1} \Phi_{k-1}^T (P_{k|k-1})^{-1} \\ \hat{\mathbf{x}}_{k-1|N} &= \hat{\mathbf{x}}_{k-1|k-1} + J_{k-1} (\hat{\mathbf{x}}_{k|N} - \hat{\mathbf{x}}_{k|k-1}) \\ P_{k-1|N} &= P_{k-1|k-1} \\ &\quad + J_{k-1} (P_{k|N} - P_{k|k-1}) J_{k-1}^T \end{aligned} \quad (5)$$

The matrices  $Q$  and  $R$  are the covariances of  $\mathbf{w}_k$  and  $\mathbf{v}_k$  respectively. The following assumptions about probability densities are implicit to the EKF algorithm

$$\begin{aligned} p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) &\approx \mathcal{N}(\hat{\mathbf{x}}_{k|k-1}, P_{k|k-1}) \\ p(\mathbf{x}_k | \mathbf{y}_{1:k}) &\approx \mathcal{N}(\hat{\mathbf{x}}_{k|k}, P_{k|k}) \end{aligned} \quad (6)$$

Similarly, the extended Kalman smoother (EKS) implicitly assumes

$$p(\mathbf{x}_k | \mathbf{y}_{1:N}) \approx \mathcal{N}(\hat{\mathbf{x}}_{k|N}, P_{k,N}) \quad (7)$$

### 3.1 Modified EKF for multi-rate data

The EKF expressions have to be suitably modified to handle multi-rate data. Let  $\mathbf{y}_{obs} \in R^{p_1}$  ( $p_1 \leq p$ ) be the set of measurements available at the  $k^{\text{th}}$  instant. Define a permutation matrix  $D_k$ , of dimensions  $p_1 \times p$ , that takes into consideration only the predictions for  $\mathbf{y}_{obs}$ . For example, if out of three measurements, the first and third are available at the  $k^{\text{th}}$  instant,

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

Let  $H_k = D_k C_k$ . The measurement update step in the EKF given by (4) is modified as follows

$$\begin{aligned} \mathbf{e}_k &= \mathbf{y}_{obs} - D_k h(\hat{\mathbf{x}}_{k|k-1}) \\ K_k &= P_{k|k-1} H_k^T [H_k P_{k|k-1} H_k^T + D_k R D_k^T]^{-1} \\ \mathbf{x}_{k|k} &= \mathbf{x}_{k|k-1} + K_k \mathbf{e}_k \\ P_{k|k} &= [I - K_k H_k] P_{k|k-1} \end{aligned} \quad (8)$$

If, at a particular sampling instant, none of the measurements are available, the correction step in (8) is not applied and only the predictions from the EKF are used.

The extended Kalman smoother is used with the predicted and filtered state estimates and covariances obtained from the EKF expressions modified for the multi-rate data case.

## 4. EM-BASED COVARIANCE ESTIMATION

To generate reliable estimates using EKF, it is important to know parameters  $(\boldsymbol{\mu}_0, P_0)$  and the covariances  $(Q, R)$  with a reasonable accuracy. However, in practice,  $Q$  is generally unknown and  $R$  may be known partially or fully. In such a scenario, these parameters can be estimated from the operating data. In the present work, it is assumed that the system under consideration is perturbed deliberately by introducing perturbations in the manipulated inputs in the control relevant frequency range. Let  $Z_N = \{Y_N, U_N\}$  denote the data collected during this exercise where  $Y_N = \{\mathbf{y}_1, \dots, \mathbf{y}_N\}$  and  $U_N = \{\mathbf{u}_1, \dots, \mathbf{u}_N\}$  represent the output and the input data sets, respectively, and  $N$  represents the number of data points. In this work, it is proposed

to use a modified version of the expectation maximisation algorithm to arrive at estimates of these unknown parameters using  $Z_N$ . To begin with, the conventional EM algorithm is briefly reviewed. The modifications necessary in the context of EKF are presented next.

### 4.1 Review of EM Algorithm

Consider the problem of estimating parameters  $\Theta \equiv (\boldsymbol{\mu}_0, P_0, Q, R)$  from  $Z_N$  for a linear perturbation state space model of the form

$$\begin{aligned} \mathbf{x}_{k+1} &= A_k \mathbf{x}_k + B_k \mathbf{u}_k + \Gamma_d \mathbf{w}_k \\ \mathbf{y}_k &= C_k \mathbf{x}_k + v_k \end{aligned} \quad (9)$$

where the system matrices  $(A_k, B_k, \Gamma_d, C_k)$  are known *a priori*. It has been shown that the innovation sequence  $(e_k)$  generated by the Kalman filter developed for this system is a zero mean Gaussian white noise sequence with covariance

$$\Sigma_k = C_k P_{k|k-1} C_k^T + R \quad (10)$$

Thus, as suggested by Schweppe (1965), the parameter estimation problem can be formulated as an optimisation problem that maximises the likelihood function

$$L(\Theta | Z_N) = \prod_{k=1}^N \frac{1}{2\pi^{p/2} |\Sigma_k|^{1/2}} \exp \left[ -\frac{\mathbf{e}_k^T \Sigma_k^{-1} \mathbf{e}_k}{2} \right] \quad (11)$$

subject to constraint that  $(P_0, Q, R)$  are symmetric and positive definite. One possible way to solve this optimisation problem is Newton type method, which requires computations and gradients and Hessian of the log likelihood function,  $[\log L(\Theta | Z_N)]$ , with respect to  $\Theta$ . The requirement to estimate these derivatives makes this approach computationally complex.

Shumway and Stoffer (2000) have proposed an alternate derivative free iterative approach, called expectation maximisation (EM), to solve this problem. This approach starts with the objective of maximising the joint probability density function of the *complete data*  $(X_N, Z_N)$  set, which can be represented as follows

$$\begin{aligned} f_{\Theta}(X_N, Z_N) &= \frac{1}{2\pi^p |P_0|^{1/2}} \exp \left[ -\frac{(\mathbf{x}_0 - \boldsymbol{\mu}_0)^T P_0^{-1} (\mathbf{x}_0 - \boldsymbol{\mu}_0)}{2} \right] \\ &\times \prod_{k=1}^N \frac{1}{(2\pi)^{n/2} |Q|^{1/2}} \exp \left[ -\frac{\mathbf{w}_k^T Q^{-1} \mathbf{w}_k}{2} \right] \\ &\times \prod_{k=1}^N \frac{1}{(2\pi)^{p/2} |R|^{1/2}} \exp \left[ -\frac{\mathbf{v}_k^T R^{-1} \mathbf{v}_k}{2} \right] \end{aligned} \quad (12)$$

where,  $X_N = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_N\}$  represents the true states. In practice, however, the set  $X_N$  is not available. The problem of maximising  $f_{\Theta}(X_N, Z_N)$  with respect to  $\Theta$  is solved iteratively in the following step

- *Expectation (E-step)*: This step involves finding the expected value of the complete data log likelihood function, given the observed data set,  $Z_N$  and the previously estimated parameter vector,  $\Theta^{(k-1)}$ . This conditional expectation is obtained using estimates generated from the Kalman smoother.
- *Maximisation (M-step)*: This step involves maximising the log likelihood function,  $f_{\Theta}[\hat{X}_N(\Theta^{(k-1)}), Z_N, Z_N]$

with respect to the parameter vector  $\Theta$ , to generate  $\Theta^{(k)}$ . The parameters estimated in each iteration depend on the observed data and the smoothed state estimates and covariances obtained in the E-step.

The iterations are terminated when the condition that  $[\log L(\Theta^{(k)}|Z_N) - \log L(\Theta^{(k-1)}|Z_N)] \leq \varepsilon$  is satisfied, where  $\varepsilon$  is the specified tolerance and  $L(\Theta|Z_N)$  is defined by (11). The details of the derivation and proof for convergence can be found in the work by Raghavan et al. (2006). The algorithm guarantees increase in the likelihood function with successive EM steps and is guaranteed to converge to a local or global optimum depending on the shape of the objective function (Shumway and Stoffer, 2000).

#### 4.2 Extended EM Algorithm

If it is desired to estimate  $(\mu_0, P_0, Q, R)$  in the context of EKF, then it becomes necessary to modify the conventional EM algorithm. To begin with, it is assumed that the innovation sequence  $[\mathbf{e}_k = \mathbf{y}_k - h(\hat{\mathbf{x}}_{k|k-1})]$  generated by the Kalman filter is a Gaussian white noise with density function  $(\bar{0}, \Sigma_k)$ . For linear systems, it is possible to obtain closed form expressions for  $\log f_\Theta(X_N, Z_N)$  in the E-step. However, for nonlinear systems, it is not possible to obtain such closed form expressions. This difficulty arises from calculating the conditional expectations of the terms involving  $\mathbf{w}_k$  and  $\mathbf{v}_k$ . In the present work, it is proposed to use approximate expressions for  $\mathbf{w}_k$  and  $\mathbf{v}_k$  using Taylor series approximation in the neighbourhood of the smoothed estimates. Thus  $\mathbf{w}(k)$  and  $\mathbf{v}(k)$  are estimated as follows

$$\begin{aligned} \mathbf{w}_k &= \mathbf{x}_k - F(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) \\ &\approx \mathbf{x}_k - \left[ F(\hat{\mathbf{x}}_{k-1|N}, \mathbf{u}_{k-1}) + \frac{\partial F}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1|N}} (\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|N}) \right] \end{aligned} \quad (13)$$

$$\begin{aligned} \mathbf{v}_k &= \mathbf{y}_k - h(\mathbf{x}_k) \\ &\approx \mathbf{y}_k - \left[ h(\hat{\mathbf{x}}_{k|N}) + \frac{\partial h}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k|N}} (\mathbf{x}_k - \hat{\mathbf{x}}_{k|N}) \right] \end{aligned} \quad (14)$$

With these modifications, the steps in the extended EM algorithm can be summarised as follows: At the  $j^{\text{th}}$  iteration ( $j = 1, 2, \dots$ ), taking the expectation of  $l(\Theta) = \log f_\Theta(X_N, Z_N)$  conditioned to  $Z_N$  and  $\Theta^{(j-1)}$ ,

$$\begin{aligned} E \left[ l(\Theta|Z_N, \Theta^{(j-1)}) \right] &= \log |P_0| + N \log |Q| + N \log |R| \\ &+ \text{tr} \left[ P_0^{-1} (P_{0|N} + (\mathbf{x}_{0|N} - \mu_0)(\mathbf{x}_{0|N} - \mu_0)^T) \right] \\ &+ \text{tr} \left[ Q^{-1} (\beta_1 - \beta_2 - \beta_2^T + \beta_3) \right] \\ &+ \text{tr} \left[ R^{-1} (\beta_4 - \beta_5 - \beta_5^T + \beta_6) \right] \end{aligned} \quad (15)$$

where  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  and  $\beta_6$  are functions of the observed data and the smoothed state and covariance estimates

$$\begin{aligned} \beta_1 &= \sum_{k=1}^N \left[ P_{k|N} + \hat{\mathbf{x}}_{k|N} \hat{\mathbf{x}}_{k|N}^T \right] \\ \beta_2 &= \sum_{k=1}^N \left[ P_{k,k-1|N} \Phi_{k-1}^T + \hat{\mathbf{x}}_{k|N} F^T(\hat{\mathbf{x}}_{k-1|N}, \mathbf{u}_{k-1}) \right] \\ \beta_3 &= \sum_{k=1}^N F(\hat{\mathbf{x}}_{k-1|N}, \mathbf{u}_{k-1}) F^T(\hat{\mathbf{x}}_{k-1|N}, \mathbf{u}_{k-1}) \\ &+ \Phi_{k-1|N} P_{k-1|N} \Phi_{k-1|N}^T \end{aligned} \quad (16)$$

$$\begin{aligned} \beta_4 &= \sum_{k=1}^N [\mathbf{y}_k \mathbf{y}_k^T] & \beta_5 &= \sum_{k=1}^N [\mathbf{y}_k h^T(\hat{\mathbf{x}}_{k|N})] \\ \beta_6 &= \sum_{k=1}^N C_k \left[ P_{k|N} + \hat{\mathbf{x}}_{k|N} \hat{\mathbf{x}}_{k|N}^T \right] C_k^T + h(\hat{\mathbf{x}}_{k|N}) h^T(\hat{\mathbf{x}}_{k|N}) \end{aligned} \quad (17)$$

The term  $P_{k,k-1|N}$  is called the lag-one covariance smoother and is given by

$$\begin{aligned} P_{N,N-1|N} &= \{I - K_N C_N\} \Phi_{N-1} P_{N-1|N-1} \\ &\quad \forall k = N \\ P_{k,k-1|N} &= P_{k|k} J_{k-1}^T \\ &\quad + J_k \{P_{k+1,k|N} - \Phi_k P_{k|k}\} J_{k-1}^T \\ &\quad \forall k = (N-1, N-2, \dots, 1) \end{aligned} \quad (18)$$

The expressions for  $\beta_i$ ,  $i = 1, 2, 3$  are obtained from the expression for  $\mathbf{w}_k$  in (13) and those for  $\beta_j$ ,  $j = 4, 5, 6$  from the expression for  $\mathbf{v}_k$  in (14).

Minimising  $E[l(\Theta|Z_N, \Theta^{(j-1)})]$  with respect to the parameters, in this case  $Q$  and  $R$ , at every iteration is called the *maximisation step*. Since the mean and covariance of the initial state cannot be identified simultaneously, they are conventionally fixed as follows (Shumway and Stoffer, 2000)

$$\mu_0 = \hat{\mathbf{x}}_{0|N} \quad P_0 = P_{0|N} \quad (19)$$

Taking the partial derivative of  $E[l(\Theta|Z_N, \Theta^{(j-1)})]$  with respect to  $Q$  and  $R$ , and setting it to zero yields the following relation

$$Q = \frac{1}{N} (\beta_1 - \beta_2 - \beta_2^T + \beta_3) \quad (20)$$

$$R = \frac{1}{N} (\beta_4 - \beta_5 - \beta_5^T + \beta_6) \quad (21)$$

#### 4.3 Extended EM algorithm for multi-rate data

The EM-algorithm was essentially developed for efficient handling multi-rate data. Let  $N_0$  be the instances when none of the measurements are available,  $N_1$  the instances when atleast one measurement is available and  $N_a = N - (N_0 + N_1)$  instances when all measurements are available. In the M-step of the extended EM algorithm, the smoothed estimates of the states and covariances obtained from the modified EKF and EKS are used. The expressions for  $E(\mathbf{v}_k \mathbf{v}_k^T | Z_N)$  change as follows

$$\begin{aligned} \sum_{k=1}^N E(\mathbf{v}_k \mathbf{v}_k^T | Z_N) &= \sum_{k=1}^N E[(\mathbf{y}_k - h(\mathbf{x}_k))(\mathbf{y}_k - h(\mathbf{x}_k))^T | Z_N] \\ &= \sum_{N_a} [(\mathbf{y}_k - h(\hat{\mathbf{x}}_{k|N}))(\mathbf{y}_k - h(\hat{\mathbf{x}}_{k|N}))^T + C_k P_{k|N} C_k^T] \\ &\quad + \sum_{N_0} R^{(j-1)} + \sum_{N_1} R_1 \end{aligned} \quad (22)$$

where,  $R^{(j-1)}$  refers to the measurement noise covariance matrix obtained in the previous iteration (Raghavan et al., 2006).  $R_1$  is a matrix such that  $R_1(i, i) = (\mathbf{y}_{i,k} - h_i(\hat{\mathbf{x}}_{k|N}))(\mathbf{y}_{i,k} - h_i(\hat{\mathbf{x}}_{k|N}))^T + C_{i,k} P_{k|N} C_{i,k}^T$ , if the  $i^{\text{th}}$  value of the measurement vector is available and  $R_1(i, i) = R^{(j-1)}(i, i)$  if it is not available.

#### 4.4 Implementing the extended EM-algorithm

The extended EM-algorithm can be implemented as follows

- (1) Initialise the algorithm by selecting the values of  $\Theta^{(0)} = [\mu_0, Q_0, R_0]$  and fix  $P_0$ . Ensure that  $P_0$ ,  $Q_0$  and  $R_0$  are positive definite.
- (2) Compute the incomplete-data likelihood function  $-2 \log L(\Theta^{(j-1)}|Z_N)$  from (11).
- (3) Perform the E-step. Use (3)-(4), (5) and (18) to obtain the smoothed values of  $\hat{x}_{k|N}$ ,  $P_{k|N}$  and  $P_{k,k-1|N}$  for  $k = 1, 2, \dots, N$ , using parameters  $\Theta^{(j-1)}$ .
- (4) Perform the M-step, using (20) and (21), to obtain the new estimates of  $\Theta^{(j)} = [\mu_0, P_0, Q, R]$ .
- (5) For multirate data, use appropriate expressions of the extended EM-algorithm.
- (6) Repeat steps 2-4 till the objective function converges.

### 5. APPLICATION: CONTINUOUS FERMENTER

Production of ethanol by fermentation of glucose using *saccharomyces cerevisiae* yeast is a widely used fermentation process. A simplified version for such a continuous fermenter system can be described by the generalised model proposed by Henson and Seborg (1992). The model equations are described as follows

$$\begin{aligned}\dot{X} &= -DX + \mu X \\ \dot{S} &= D(S_f - S) - \frac{1}{Y_{X|S}}\mu X \\ \dot{P} &= -DP + (\alpha\mu + \beta)X\end{aligned}\quad (23)$$

where,  $X$  is the yeast(cell-mass) concentration,  $S$  is the glucose(substrate) concentration and  $P$  is the alcohol(product) concentration. The dilution rate,  $D$ , and feed substrate concentration,  $S_f$ , are available as manipulated inputs.  $\mu$  is the specific growth rate,  $Y_{X|S}$  is the cell-biomass yield,  $\alpha$  and  $\beta$  are yield parameters for the product. The specific growth rate model is assumed to exhibit both substrate and product inhibition:

$$\mu = \frac{\mu_m \left(1 - \frac{P}{P_m}\right) S}{K_m + S + \frac{S^2}{K_i}} \quad (24)$$

where,  $\mu_m$  is the maximum specific growth rate,  $P_m$  is the product saturation constant,  $K_m$  is the substrate saturation constant and  $K_i$  is the substrate inhibition constant. The substrate concentration,  $S$  and product concentration  $P$  are assumed to be available as measurements. The nominal parameters and operating conditions are given in Henson and Seborg (1992).

To simulate the process dynamics, the sampling time ( $T$ ) was chosen as 0.25 h. Input-output data was generated by subjecting the manipulated inputs to a pseudo-random binary signal(PRBS). The amplitude of perturbation in  $D$  is  $0.015 \text{ h}^{-1}$  and in  $S_f$  it is 2 g/l. White noise was added to the states and measurements to simulate the process and measurement noise respectively. The true values of the noise covariance matrices are shown in Table 1. It was assumed that the measurements of product concentration,  $P$ , are available at every sampling instant, while the measurements of substrate concentration,  $S$ , are available at irregular sampling intervals, which are integer multiples

of  $T$ . The maximum delay between any two successive samples of  $S$  is three. The case when the measurements of the substrate (glucose) concentration is available at every sampling instant is also investigated to serve as a benchmark.

The initial guess of the parameters used to start the EM algorithm are reported in Table 1. It may be noted that the initial estimates of  $Q$  are significantly different from the true values. The estimated optimum values of  $Q$  and  $R$  matrices, generated using the proposed extended EM algorithm on multi-rate and single rate measurements are compared in Table 2. The estimates of the initial state and its covariance are reported in Table 3. Fig. 1 reports the change in log likelihood function as a function of the iteration count in the E-step. This figure also plots the value of log likelihood function obtained when the true values of  $\Theta$  are used in EKF formulations in each case. It may be noted that the estimates  $[-\log L(\Theta|Z_N)]$  generated by EM monotonically converge to  $[-\log L(\Theta)]_{\text{true}}$  for the regularly sampled fast rate data. The performance of EM deteriorates for the multi-rate data case. A possible remedy to this problem is to increase data length. It may be noted that the measurement noise covariance,  $R$ , is estimated with fair accuracy in both the cases. This can be attributed to the fact that the measurements are linear function of states and the measurement noise is additive. The estimates of the state noise covariance,  $Q$ , are within an acceptable range. However, the error in the estimates is relatively higher. The observed discrepancy can be attributed to errors in approximating the estimation error densities through local linearisation. The performance of the EKF was also evaluated on a separate validation data set, using the sum of squared values of the estimation errors(SSE) as a metric for comparison. The SSE values of the single-rate and multi-rate EKF using the estimated  $Q$  and  $R$ , were compared with the corresponding values obtained using the true parameters of the EKF. These values are shown in Table 4. For the multi-rate case, a comparison of the estimation errors in  $X$  and  $S$ , for the estimates obtained using the initial guess of parameters,  $\Theta = (\hat{x}_0, P_0, Q, R)$ , and the estimated values of  $\Theta$ , are shown in Fig. 2. The SSE values and the graphical comparison presented in Fig. 2 indicate that the performance of the EKF implemented with estimated parameters, is close to that of the EKF implemented using true parameters. When the estimated optimum values of  $\Theta$  are used in the EKF, the state estimation optimum errors are significantly smaller, compared to the case when  $\Theta$  is set equal to the initial guess. The improvements obtained using the estimated optimum values are more pronounced in the case of the unmeasured state, i.e. biomass concentration ( $X$ ). A slight deterioration that is observed in the multi-rate case, can be attributed to the absence of regular measurements.

### 6. CONCLUSIONS

In this work, an extended EM-algorithm has been presented for the estimation of the state and measurement noise covariances for state estimation of nonlinear multi-rate sampled data systems using the EKF. The algorithm gives reasonably accurate estimates of the covariances, even in the case of multi-rate data. Simulation studies revealed that the algorithm converges even for signifi-

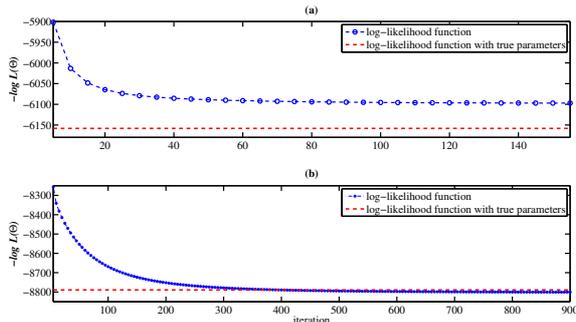


Fig. 1. Log-likelihood function vs. iteration. (a): multi-rate data. (b): single rate data

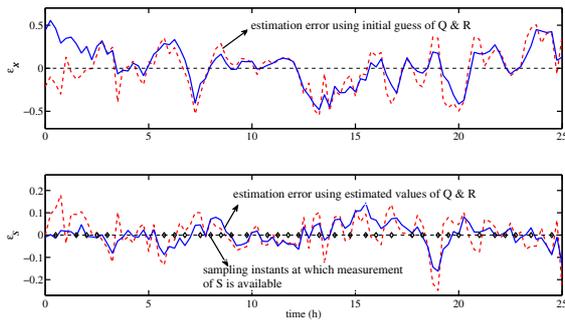


Fig. 2. Comparison of estimation errors obtained using the initial guess and the estimated parameters

cantly large errors in the initial estimates of the state noise covariance. The EKF approximates the transition densities,  $p(x_k|y_{1:k-1})$ , to be Gaussian and obtains them

Table 1. True and Initial guess of parameters

	True	Initial Guess
$Q$	$10^{-3} \times \begin{bmatrix} 11.1 & 0 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0 & 139.2 \end{bmatrix}$	$\begin{bmatrix} 1.5 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 15 \end{bmatrix}$
$R$	$10^{-3} \times \begin{bmatrix} 5.6 & 0 \\ 0 & 15.6 \end{bmatrix}$	$\begin{bmatrix} 0.01 & 0 \\ 0 & 0.03 \end{bmatrix}$
$\mathbf{x}_0$	$\begin{bmatrix} 7.5306 \\ 2.5723 \\ 26.6109 \end{bmatrix}$	$\begin{bmatrix} 7.7417 \\ 2.6444 \\ 27.3570 \end{bmatrix}$
$P_0$	$10^{-3} \times \begin{bmatrix} 0.495 & 0 & 0 \\ 0 & 0.577 & 0 \\ 0 & 0 & 0.386 \end{bmatrix}$	$\begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}$

Table 2. Estimated values of  $Q$  and  $R$

	Single-rate	Multi-rate
$Q$	$10^{-3} \times \begin{bmatrix} 11.2 & -0.4 & 0.9 \\ -0.4 & 0.6 & -1.4 \\ 0.9 & -1.4 & 151.0 \end{bmatrix}$	$10^{-3} \times \begin{bmatrix} 10.7 & -0.2 & 0.6 \\ -0.2 & 0.7 & -1.2 \\ 0.6 & -1.2 & 150.6 \end{bmatrix}$
$R$	$10^{-3} \times \begin{bmatrix} 5.6 & 0 \\ 0 & 15.2 \end{bmatrix}$	$10^{-3} \times \begin{bmatrix} 5.5 & 0 \\ 0 & 15.4 \end{bmatrix}$

Table 3. Estimated initial conditions

Case	$\mathbf{x}_0 - \hat{\mathbf{x}}_0$	$\hat{P}_0$
Single Rate	$\begin{bmatrix} 0.43800 \\ 0.00380 \\ -0.0142 \end{bmatrix}$	$10^{-3} \times \begin{bmatrix} 0.1281 & 0.0128 & 0.0005 \\ 0.0128 & 0.0065 & -0.0008 \\ 0.0005 & -0.0008 & 0.0344 \end{bmatrix}$
Multi rate	$\begin{bmatrix} 0.44030 \\ -0.0100 \\ -0.0165 \end{bmatrix}$	$10^{-3} \times \begin{bmatrix} 0.1663 & 0.0145 & 0.0005 \\ 0.0145 & 0.0080 & -0.0009 \\ 0.0005 & -0.0009 & 0.0432 \end{bmatrix}$

Table 4. SSE values of the estimation errors

Parameters	$\varepsilon_X$	$\varepsilon_S$	$\varepsilon_P$
	<b>Multi</b>	<b>Rate Case</b>	
Initial Guess	37.3598	3.2571	7.3037
True	30.0700	1.7953	6.4479
Estimated	31.3182	1.7631	6.4546
	<b>Single</b>	<b>Rate Case</b>	
Initial Guess	47.6591	2.5509	7.3036
True	28.6048	1.2641	6.4485
Estimated	29.9715	1.2427	6.4593

using linearised system matrices. It may be possible to obtain improved estimates of the noise covariances if better approximations are obtained for the transition densities. This direction of research is currently under investigation.

## REFERENCES

- A. P. Dempster, N. M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc. B*, 3:1–38, 1977.
- D. Dochain. State and parameter estimation in chemical and biochemical processes: A tutorial. *J. Process Contr.*, 13:801–818, 2003.
- G. C. Goodwin and J. C. Agüero. Approximate EM algorithms for parameter and state estimation in nonlinear stochastic models. In *Proc. 44th IEEE Conference on Decision and Control*, pages 368–373, December 2005.
- M. A. Henson and D. E. Seborg. Nonlinear control strategies for continuous fermenters. *Chem. Eng. Sci.*, 47(4):821–835, 1992.
- R. K. Mehra. On the identification of variances and adaptive Kalman filtering. *IEEE Trans. Autom. Control*, 15(12):175–184, 1970.
- B. J. Odelson, M. R. Rajamani, and J. B. Rawlings. A new autocovariance least-squares method for estimating noise covariances. *Automatica*, 42:303–308, 2006.
- H. Raghavan, A. K. Tangirala, R. B. Gopaluni, and S. L. Shah. Identification of chemical processes with irregular output sampling. *Control Eng. Pract.*, 14:467–480, 2006.
- J.B. Rawlings and B.R. Bakshi. Particle filtering and moving horizon estimation. *Comput. Chem. Eng.*, 30: 1529–1541, 2006.
- F. Scheppe. Evaluation of likelihood functions for Gaussian signals. *IEEE Trans. Inf. Theory*, 11(1):61 – 70, 1965.
- R. H. Shumway and D. S. Stoffer. *Time Series Analysis and Its Applications*. Springer-Verlag New York, Inc., New York, 2000.
- J. Valappil and C. Georgakis. Systematic estimation of state noise statistics for extended Kalman filters. *AIChE J.*, 46(2):292–308, 2000.