Identification of Noise Covariances for State Estimation of Autonomous Hybrid Systems

Vinay A. Bavdekar ∗ Sachin C. Patwardhan ∗.1

∗ Department of Chemical Engineering, Indian Institute of Technology
Bombay, Mumbai, India

Abstract: The occurrence of dynamic systems that involve both continuous and discrete state variables is becoming increasingly common in process industries, especially those that manufacture multiple products. To operate such processes efficiently, it is essential to monitor and tightly control the state variables associated with them. Recently, Prakash et al. (2010b) have proposed the use of derivative free state estimators, namely the unscented Kalman filter and ensemble Kalman filter, for estimating continuous and discrete states associated with autonomous hybrid systems. A critical aspect of developing such Bayesian state estimators is to have a reasonably accurate characterisation of such unmeasured disturbances and noise. In this work, it is proposed to identify noise covariances associated with autonomous hybrid systems from the operating data. The state estimators used for this purpose are UKF and EnKF. The problem of estimating the noise covariance matrices is formulated as a constrained optimisation problem, in which a suitable objective function of the innovation sequence is minimised. The efficacy of the proposed covariance estimation scheme is demonstrated by simulating the benchmark three-tank system from the literature.

Keywords: nonlinear state estimation, unscented Kalman filter, ensemble Kalman filter, autonomous hybrid systems, covariance identification

1. INTRODUCTION

The occurrence of dynamic systems that involve both continuous and discrete state variables is becoming increasingly common in process industries, especially those that manufacture multiple products. The continuous states variables relate to the states of the process, whereas the discrete variables are logic variables, which are influenced by the instantaneous values of the continuous states. Such systems are broadly classified as hybrid systems. To operate such processes efficiently, it is essential to monitor and tightly control the state variables associated with them. Since a first principles based mechanistic model is representative of the physical states of the process, it can be used to deduce the information of the trajectories of all internal unmeasured (or irregularly measured) states using the available measurements.

The problem of state estimation for autonomous hybrid systems has received very less attention in the literature. While the EKF is widely used for state estimation of nonlinear systems, its drawback is that it requires computation of the Jacobian of the nonlinear state and measurement equations at every sampling instant. Dynamic models for the hybrid systems have discontinuities that arise due to the discrete variables. The EKF, therefore, cannot be used for state estimation of such processes as the Jacobian cannot be computed for these discrete state variables. Ferrari-Trecate et al. (2000) proposed a moving horizon estimation (MHE) formulation for hybrid systems.

Recently, Prakash et al. (2010b) have proposed the use of derivative free state estimators, namely the unscented Kalman filter and ensemble Kalman filter, for state estimation of hybrid systems, that circumvent the problems caused by the EKF. These filters do not require linearisation of the system equations and the statistical properties of the state estimates are computed using sample statistics. Such filters can, therefore, be used for state estimation of autonomous hybrid systems. The unscented Kalman filter (UKF) (Julier and Uhlmann, 2004) computes and propagates the mean and covariances of the states and measurements using a deterministically chosen set of points. The ensemble Kalman filter (EnKF) (Evensen, 2007), belongs to a class of particle filters, in which the statistical properties of the estimates are generated using Monte Carlo sampling. Prakash et al. (2010b) demonstrate that these filters can be effectively used to construct estimates of the continuous as well as discrete states, under the assumption that the characteristic of the stochastic signals influencing the system dynamics is known accurately. In practice, however, the later assumption can prove to be a bottleneck in the implementation of these filters.

Dynamical systems operate in the presence of various unmeasured disturbances. Also, the measurements available are corrupted with noise. A critical aspect of developing Bayesian state estimators is to have a reasonably accurate characterisation of such unmeasured disturbances and noise. An incorrect characterisation of the noise leads to deterioration in the performance of the state estimator and in the worst case, the estimator may diverge. There has been work reported in literature for estimating the covari-
ances (or densities) using the Kalman filter (KF) (Mehra, 1970; Odelson et al., 2006), for linear systems, and the EKF (Valappil and Georgakis, 2000; Bavdekar et al., 2011), for nonlinear systems. The later approach, as proposed, is suitable for covariance estimation for dynamic systems involving only the continuous states. In this work, it is proposed to extend this approach to estimation of noise covariances associated with autonomous hybrid systems from the operating data. The state estimators used for this purpose are UKF and EnKF. The problem of estimating the noise covariance matrices is formulated as a constrained optimisation problem in which a suitable objective function of the innovation sequence is minimised. The efficacy of the proposed covariance estimation scheme is demonstrated by simulating the benchmark three-tank system from the literature.

The paper is organised as follows. The process model, the assumptions involved and the UKF and EnKF algorithms for an autonomous hybrid system are discussed in Section 2. The formulation of the optimisation problem to identify the noise covariances is discussed in Section 3. In Section 4, the results obtained by applying the optimisation formulation on a benchmark three-tank hybrid system are discussed.

2. PRELIMINARIES

In this work, it is proposed to identify the covariances of the unmeasured disturbances and measurement noise, \( Q \) and \( R \) respectively, from operating input-output data. The details of the proposed approach are presented in this section.

2.1 Process and Model for State Estimation

The first principles model for an autonomous hybrid process can be described by the following equations

\[
\begin{align*}
\dot{x} &= f(x, \xi, u) \\
\dot{y}(t) &= h(x) + v(t)
\end{align*}
\]

\( x \in \mathbb{R}^n \) represents the continuous states of the process, \( \xi \in \mathbb{R}^{n_m} \) represents the discrete states of the system, \( u \in \mathbb{R}^m \) denotes the manipulated inputs, \( y \in \mathbb{R}^r \) represents the measured outputs and \( v \in \mathbb{R}^r \) denotes the measurement noise. The discrete states, \( \xi \), are a function of the continuous states, \( x \).

The following assumptions are made for simulating the process and modelling for state estimation:

1. The measurement noise can be modelled as a zero-mean white noise with Gaussian distribution, i.e. \( v(t) \sim \mathcal{N}(0, R) \).
2. The manipulated inputs are piece-wise constants over the sampling interval \( u(t) = u_k \) for \( t_k \leq t < t_k + T \)
3. The choice of the sampling time, \( T \), is small enough so that the variation in unmeasured disturbances can be treated as piece-wise constant functions. The unmeasured disturbances are assumed to evolve as a zero-mean white noise sequence with a Gaussian distribution, whose source is unknown. They are, therefore, modelled to enter the process as additive signals in the states.

Under Assumptions 1-3, a discrete-time representation of the true process can be obtained as follows

\[
x_{k+1} = x_k + \int_{kT}^{(k+1)T} f(x(\tau), \xi(\tau), u_k) \, d\tau + w_k
\]

\[
y_k = h(x_k) + v_k
\]

where, \( w_k \in \mathbb{R}^n \) denotes a white noise sequence, such that \( w_k \sim \mathcal{N}(0, Q) \). It is further assumed that \( w_k \) and \( v_k \) are mutually uncorrelated, independent and identically distributed random variables.

2.2 Unscented Kalman Filter

In the UKF algorithm (Julier and Uhlmann, 2004), the statistical properties of the states and measurements are approximated by using a deterministically chosen sample of points, known as ‘sigma points’. Each of the sigma point is transformed through the nonlinear equations and the statistical properties of the states and measurements are propagated through the measurement equations to generate the predictions for the next instant

\[
\hat{x}_{k+1|k,i} = F(\hat{x}_{k|k,i}, u_k)
\]

The mean of the predicted states and their error covariance matrix are obtained from the sigma points as follows

\[
\hat{x}_{k+1|k,i} = \sum_{i=0}^{2n} W_i \hat{x}_{k+1|k,i}
\]

\[
P_{k+1|k} = \sum_{i=0}^{2n} W_i [\hat{x}_{k+1|k,i} - \hat{x}_{k+1|k}] [\hat{x}_{k+1|k,i} - \hat{x}_{k+1|k}]^T + Q
\]

To obtain the output predictions, the sigma points are propagated through the measurement equations

\[
\hat{y}_{k+1|k,i} = h(\hat{x}_{k+1|k,i})
\]

Using the sigma points of output predictions, the innovations covariance matrix and the cross-covariance between the state estimation error and innovations are computed
\[
P_{ee,k+1} = \sum_{i=0}^{2n} W_i \left[ \hat{y}_{k+1|k,i} - \hat{y}_{k+1|k} \right] \left[ \hat{y}_{k+1|k,i} - \hat{y}_{k+1|k} \right]^T + R
\]

(8)

\[
P_{xe,k+1} = \sum_{i=0}^{2n} W_i \left[ x_{k+1|k,i} - \hat{x}_{k+1|k} \right] \left[ X_{k+1|k,i} - \hat{X}_{k+1|k} \right]^T
\]

(9)

where,

\[
\hat{y}_{k+1|k} = \sum_{i=0}^{2n} W_i Y_{k+1|k,i}
\]

(10)

The Kalman gain is computed using the above matrices

\[
K_{k+1} = P_{xe,k+1} \sum_{i=0}^{2n} W_i \left( Y_{k+1|k,i} - \hat{Y}_{k+1|k,i} \right)
\]

(11)

The filtered state estimates are then obtained as follows

\[
\hat{x}_{k+1|k} = \hat{x}_{k|k} + K_{k+1} (y_{k+1} - \hat{y}_{k+1|k})
\]

(12)

and the updated estimation error covariance is

\[
P_{k+1|k+1} = P_{k+1|k} - K_{k+1} P_{xe,k+1} K_{k+1}^T
\]

(13)

2.3 Ensemble Kalman Filter

The EnKF is initialised by sampling \( N \) particles, \( \tilde{x}_0^{(j)} \), from a suitable distribution, with mean \( \tilde{x}_0 \). At every sampling instant, random samples of the process noise, \( \beta \), and measurement noise, \( \nu_k \), are drawn from their distributions, \( \mathcal{N}(0,Q) \) and \( \mathcal{N}(0,R) \) respectively. The particles, \( \tilde{x}_k^{(j)} \), \( j = 1, 2, \ldots, N \), are propagated as follows:

\[
\tilde{x}_k^{(j)} = F \left( \tilde{x}_{k-1}^{(j)} - u_k, \tilde{\xi}_{k-1}^{(j)} \right) + \epsilon_k^{(j)}
\]

(14)

\[
\tilde{y}_k^{(j)} = h \left( \tilde{x}_k^{(j)} - \bar{y}_k \right) + v_k^{(j)}
\]

The sample mean and covariances of these particles are then estimated as follows

\[
\tilde{x}_{k|k} = \frac{1}{N} \sum_{j=1}^{N} \tilde{x}_k^{(j)}
\]

(15)

\[
\tilde{y}_{k|k} = \frac{1}{N} \sum_{j=1}^{N} \tilde{y}_k^{(j)}
\]

(16)

\[
P_{ee,k} = \frac{1}{N-1} \sum_{j=1}^{N} \left[ \tilde{y}_k^{(j)} - \tilde{y}_{k|k} \right] \left[ \tilde{y}_k^{(j)} - \tilde{y}_{k|k} \right]^T
\]

(17)

\[
P_{xe,k} = \frac{1}{N-1} \sum_{j=1}^{N} \left[ \tilde{x}_k^{(j)} - \tilde{x}_{k|k} \right] \left[ \tilde{y}_k^{(j)} - \tilde{y}_{k|k} \right]^T
\]

(18)

The Kalman gain is computed as

\[
K_k = P_{xe,k} P_{ee,k}^{-1}
\]

(19)

The filtered estimates of the particles are obtained as

\[
\hat{x}_k^{(j)} = \tilde{x}_{k|k} + K_k \left( y_k - \tilde{y}_k^{(j)} \right)
\]

(20)

The filtered estimate of the states is computed as a mean of the filtered particles.

\[
\hat{x}_{k|k} = \frac{1}{N} \sum_{j=1}^{N} \hat{x}_k^{(j)}
\]

(21)

2.4 Sampling of random variables for EnKF

To generate accurate sample statistics, it is necessary to have a good sample of particles. A good sample of particles avoids skewing of the statistics due to stray values of a few particles. To generate a good sample of particles for \( \hat{x}_0, \nu_k \) and \( \beta \), the following algorithm is used (Evensen, 2007): Sample a large ensemble of the state or noise vector, e.g. \( \beta \) times \( N \) and store the ensemble perturbations in a matrix \( \mathbf{A} \in \mathbb{R}^{N \times \beta N} \). For multivariable systems, it may be necessary to scale the ensembles in \( \mathbf{A} \). Then, the following steps are performed

i. Compute the SVD, \( \mathbf{USV}^T = \mathbf{A} \).

ii. Retain only the first \( N \) columns of \( \mathbf{S} \), referred to as \( \hat{\mathbf{S}} \) subsequently.

iii. Scale the non-zero singular values of \( \hat{\mathbf{S}} \) with \( \sqrt{\beta} \) to get the correct covariance in the new ensemble.

iv. Generate a new orthogonal matrix \( \mathbf{V} \in \mathbb{R}^{N \times N} \). This can be done by generating a \( \mathcal{N}(0, \mathbf{I}_N) \) matrix and taking its QR decomposition. The \( \mathbf{Q} \) matrix in the QR decomposition is an orthogonal matrix.

v. Generate the new ensemble as \( \hat{\mathbf{A}} = \hat{\mathbf{U}} \frac{1}{\sqrt{\beta}} \hat{\mathbf{S}} \mathbf{V}^T \), where \( \hat{\mathbf{U}} \) contains the first \( N \) singular vectors from \( \mathbf{U} \).

3. THE OPTIMISATION PROBLEM

To generate input-output data for the estimating the noise densities, the system under consideration is deliberately perturbed by introducing perturbations in the manipulated variables. Let, \( Z_{Nd} = \{ Y_{Nd}, U_{Nd} \} \) denote the data collected during the exercise, where \( Nd \) is the length of the data, \( Y_{Nd} = \{ y_1, y_2, \ldots, y_{Nd} \} \) and \( U_{Nd} = \{ u_1, u_2, \ldots, u_{Nd} \} \). The Maximum Likelihood Estimation (MLE) framework, maximises the following probability density function (pdf) of the observed data

\[
l(\Theta) = p(Z_{Nd} | \Theta)
\]

subject to \( \Theta \in \mathbb{S} \)

(22)

subject to \( \Theta \equiv (Q, R) \in \mathbb{R}^d \) represents the set of parameters to be estimated. \( \mathbb{S} \) represents a subset of \( \mathbb{R}^d \) and \( p(Z_{Nd} | \Theta) \) represents a suitable pdf in the context of the EnKF.

Assumption: The innovations sequence, \( \{ e_k \} \), is zero mean with a Gaussian distribution, with covariance computed using (17).

Using this assumption, the likelihood function in (22) can be written in terms of the pdf of the innovation sequence as

\[
p(Z_{Nd} | \Theta) \approx \prod_{j=1}^{Nd} p(e_j | \Theta)
\]

(23)

The problem of estimating \( \Theta = (Q, R) \) can be formulated as an optimisation problem that minimises the negative of the log-likelihood function of the innovations
\[ \Theta_{ML} = \arg \min_{\Theta} \left[ L(\Theta | Z_{N_d}) \right] \]
\[ L(\Theta | Z_{N_d}) = \sum_{j=1}^{N_d} \left( \log \det (P_{ee, j}) + e_j^T P_{ee, j}^{-1} e_j \right) \]  
(24)

Subject to

EnKF equations (14) – (21),
\[ Q > 0, \ R > 0 \]

The objective function in (24) is nonlinear and non-convex, with respect to \( \Theta \). The resulting optimisation problem can be solved using any of the standard available numerical techniques for constrained optimisation, such as sequential quadratic programming.

**Note:** When the EnKF is used for state estimation, the problem of state estimates becoming inconsistent with the physical limits might arise when gradient-based algorithms are used to solve the optimisation problem. It is possible that an intermediate gradient step may result in a large value of \( Q \) and \( R \). This would result in large values of samples of \( w^{(j)} \) and \( v^{(j)} \) when compared to the states and measurements, thus resulting in the unconstrained estimated states violating the feasible region. This problem can be circumvented by using the constrained EnKF, which would help maintain the consistency of the state estimates, in accordance with their physical limits, during the subsequent search steps. The details of the C-EnKF are discussed in Appendix A.

### 3.1 Parametrisation of the covariance matrices

The parametrisation of the covariance matrices has a bearing on the number of variables to be identified. If \( Q \) and \( R \) are formulated as full matrices, the number of variables to be identified are \( \frac{n(n+1)}{2} + \frac{r(r+1)}{2} \). On the other hand, the covariance matrix \( Q \) can be parametrised as a diagonal matrix, which is a reasonable assumption when the noise is arising from independent physical variables. The measurement noise covariance matrix \( R \), on the other hand, can be parametrised as a diagonal matrix, as it is reasonable to assume that the sensors function independently of each other. This, therefore, reduces the number of variables to be estimated to \((n + r)\), thereby also reducing the number length of the operating data required.

**Note:** In many cases, the variances of the measurements (e.g., temperature, level, etc.) can be obtained independently by analysing the measurements of the variable under consideration, under steady state operating conditions. Thus, the measurement covariance matrix \( R \) can be obtained from independent experiments. In such a scenario, the optimisation problem defined by Eq. 24 is further reduced to identification of only the process noise covariance matrix, \( Q \).

### 4. APPLICATION: HYBRID THREE-TANK PROBLEM

The three-tank system (Blanke et al., 2007) is used as a case study to assess the efficacy of the proposed approach. The mechanistic model for the system is given by the following equations

\[ A \frac{dh_1}{dt} = q_{max} u_1 - q_2 - q_3 - q_5 \]
\[ A \frac{dh_2}{dt} = q_2 + q_3 - q_4 - q_7 - q_8 \]  
(25)
\[ A \frac{dh_3}{dt} = q_{max} u_5 + q_4 + q_7 \]

where, \( u_1 \) and \( u_5 \) are the manipulated variables and
\[ q_3 = k_3 \text{sgn} (h_1 - h_2) \sqrt{|h_1 - h_2| u_4} \]
\[ q_4 = k_4 \text{sgn} (h_2 - h_3) \sqrt{|h_2 - h_3| u_4} \]
\[ q_5 = k_5 \sqrt{h_1 + h_2 u_6} \]
\[ q_7 = k_6 \sqrt{h_2 + h_3 u_8} \]

The flow \( q_2 \) can take four possible values, depending on the water levels \( h_1 \) and \( h_2 \), as follows
\[ \begin{align*}
q_2 &= z_1 k_2 \sqrt{|h_1 - h_2| u_2} \quad \{h_1, h_2 \leq h_T\} \\
&= z_2 k_2 \sqrt{(h_1 - h_T) u_2} \quad \{(h_1 > h_T) \text{ and } (h_2 \leq h_T)\} \\
&= z_3 k_2 \sqrt{(h_2 - h_T) u_2} \quad \{(h_1 \leq h_T) \text{ and } (h_2 > h_T)\} \\
&= z_4 k_2 \sqrt{|h_1 - h_2| u_2} \quad \{(h_1, h_2 > h_T)\}
\end{align*} \]  
(27)

where, \[ z_1 = \begin{cases}
0 & \{(h_1 \leq h_T) \text{ AND } (h_2 \leq h_T)\} \text{ OR } \{(h_1 = h_2)\} \\
+1 & \{(h_1 > h_T) \text{ AND } (h_2 \leq h_T)\} \text{ OR } \{(h_1 > h_T) \text{ AND } (h_2 > h_T)\} \text{ AND } \{(h_1 = h_2)\} \\
-1 & \{(h_1 \leq h_T) \text{ AND } (h_2 > h_T)\} \text{ OR } \{(h_1 > h_T) \text{ AND } (h_2 > h_T)\} \text{ AND } \{(h_1 = h_2)\}
\end{cases} \]

(28)

Similarly, the flow \( q_7 \), too can take four possible values, depending on the values of \( h_2 \) and \( h_3 \)
\[ \begin{align*}
q_7 &= z_1 k_7 \sqrt{|h_2 - h_3| u_7} \quad \{h_2, h_3 \leq h_T\} \\
&= z_2 k_7 \sqrt{(h_2 - h_T) u_7} \quad \{(h_2 > h_T) \text{ and } (h_3 \leq h_T)\} \\
&= z_3 k_7 \sqrt{(h_3 - h_T) u_7} \quad \{(h_2 \leq h_T) \text{ and } (h_3 > h_T)\} \\
&= z_4 k_7 \sqrt{|h_2 - h_3| u_7} \quad \{(h_2, h_3 > h_T)\}
\end{align*} \]  
(29)

where, \[ z_2 = \begin{cases}
0 & \{(h_2 \leq h_T) \text{ AND } (h_3 \leq h_T)\} \text{ OR } \{(h_2 = h_3)\} \\
+1 & \{(h_2 > h_T) \text{ AND } (h_3 \leq h_T)\} \text{ OR } \{(h_2 > h_T) \text{ AND } (h_3 > h_T)\} \text{ AND } \{(h_2 = h_3)\} \\
-1 & \{(h_2 \leq h_T) \text{ AND } (h_3 > h_T)\} \text{ OR } \{(h_2 > h_T) \text{ AND } (h_3 > h_T)\} \text{ AND } \{(h_2 = h_3)\}
\end{cases} \]

(30)

The range of inputs \( u_i, (i = 1, 2, \ldots, 7) \) is in the interval \([0, 1]\). It may be noted that the values of the flows \( q_2 \) and \( q_7 \) depend on the discrete variables \( z_1 \) and \( z_2 \) respectively. The discrete variables, \( z_1 \) and \( z_2 \), can take the values \{-1, 0, +1\} depending, solely, on the liquid levels in the tanks.

The levels in tanks 1 and 3, \( h_1 \) and \( h_3 \) respectively, are available as measurements. The objective is to estimate the level of water in each of the three tanks. The model parameters and the steady state operating conditions are given in Table 4.
Fig. 1. The three-tank hybrid system

Table 1. Three tank system: model parameters and steady states

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( q_{\text{max}} )</td>
<td>( 2 \times 10^{-4} )</td>
<td>( k_2 )</td>
<td>( 3 \times 10^{-4} )</td>
</tr>
<tr>
<td>( k_4 )</td>
<td>( 6 \times 10^{-4} )</td>
<td>( k_6 )</td>
<td>( 2 \times 10^{-4} )</td>
</tr>
<tr>
<td>( k_7 )</td>
<td>( 3 \times 10^{-4} )</td>
<td>( h_2 )</td>
<td>0.3</td>
</tr>
<tr>
<td>( h_4 )</td>
<td>0.02</td>
<td>( A )</td>
<td>0.0108</td>
</tr>
<tr>
<td>Steady State Values</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( h_1 )</td>
<td>0.280</td>
<td>( h_2 )</td>
<td>0.267</td>
</tr>
<tr>
<td>( h_3 )</td>
<td>0.303</td>
<td>( u_2 )</td>
<td>0.700</td>
</tr>
<tr>
<td>( u_3 )</td>
<td>0.400</td>
<td>( u_4 )</td>
<td>0.600</td>
</tr>
<tr>
<td>( u_5 )</td>
<td>0.400</td>
<td>( u_6 )</td>
<td>0.300</td>
</tr>
<tr>
<td>( u_7 )</td>
<td>0.700</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To simulate the process dynamics, the sampling time \( T \) was chosen as 10 s. The input-output data was generated by subjecting the manipulated inputs to a pseudo-random binary signal (PRBS). The amplitude of perturbation in \( u_1 \) is 0.012 and in \( u_5 \) it is 0.06. To simulate the effect of process noise and measurement noise, a Gaussian white noise \( w_k \sim \mathcal{N}(0, Q) \) was added to the states and \( v_k \sim \mathcal{N}(0, R) \) was added to the measurements, using the values given in Table 4.

Table 2. Values of \( Q \)

<table>
<thead>
<tr>
<th></th>
<th>( 10^{-4} \times \text{diag} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Values</td>
<td>[1.232 0.960 1.082]</td>
</tr>
<tr>
<td>Initial Guess</td>
<td>[3.600 3.000 3.300]</td>
</tr>
<tr>
<td>Estimated (UKF)</td>
<td>[0.947 1.332 0.791]</td>
</tr>
<tr>
<td>Estimated (EnKF)</td>
<td>[1.335 1.154 1.217]</td>
</tr>
</tbody>
</table>

In this work, it is assumed that the measurement noise covariance, \( R \), is known. The task is, therefore, to identify the process noise covariance, \( Q \), from the input-output data. The problem of estimation of \( Q \) was solved for both, the UKF and EnKF. The initial guess of \( Q \), used for the optimisation problem in both cases, UKF and EnKF, is given in Table 4. The optimisation problem was solved using the \textit{fmincon} function in Matlab. For the UKF, the optimum value of \( Q \) obtained is shown in Table 4. The change in the value of the log-likelihood function (Eq. 24) with every iteration is shown in Fig. 2. This figure also shows the value of the log-likelihood function obtained when the true values of \( Q \) were used in the UKF, henceforth referred to as the ideal value of the log-likelihood function. From Fig. 2, it can be seen that the value of the log-likelihood function generated during the gradient based optimisation search, converges close to the ideal value of the log-likelihood function, albeit not monotonically.

For the EnKF, the number of ensemble particles chosen, for the states, process and measurement noise, were \( N = 50 \). The optimum value of the covariance obtained, is given in Table 4. The change in the value of the log-likelihood function (Eq. 24) with very iteration is shown in Fig. 3. It may be noted, from Fig. 3, that the value of the log-likelihood function generated during the gradient based optimisation search, converges close to the ideal value of the log-likelihood function.
parametrisation being \( Q = \alpha I_n \), where \( \alpha \) is a positive scalar. The performance of the UKF and EnKF is also compared for different values of \( \alpha \). During the simulations, it was found that both the filters diverge, if \( \alpha > 0.002 \). The SSE values can be seen in Table 4 for the UKF and the EnKF. As can be seen from the SSE values in Table 4, the performance of both the state estimators when the estimated values of \( Q \) are used is closer to that when the true values are used.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>UKF</th>
<th>EnKF</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = 0.001 )</td>
<td>0.0336</td>
<td>0.0490</td>
</tr>
<tr>
<td>( \alpha = 0.0005 )</td>
<td>0.0390</td>
<td>0.0428</td>
</tr>
<tr>
<td>Initial Guess</td>
<td>0.0336</td>
<td>0.0399</td>
</tr>
<tr>
<td>True</td>
<td>0.0282</td>
<td>0.0312</td>
</tr>
<tr>
<td>Estimated</td>
<td>0.0280</td>
<td>0.0319</td>
</tr>
</tbody>
</table>

While the optimisation algorithm yields similar results of the covariances, for both the UKF and EnKF, there is a considerable difference in the amount of time taken for the algorithm to converge. On a computer equipped with 4 GB RAM and Intel Core i7, 2.67 GHz processor, the optimisation algorithm with the UKF as a state estimator takes around 9500 seconds to converge to a solution. When the EnKF is used, with \( N = 50 \), the optimisation algorithm takes around 52000 seconds to converge to a solution.

5. CONCLUSIONS

In this work, a constrained optimisation formulation has been presented to obtain the maximum likelihood estimates of the covariances of the process and measurement noise for state estimation of hybrid systems using derivative-free state estimators namely, the UKF and EnKF. Simulation studies on the benchmark three tank hybrid system reveal that the algorithm yields reasonably accurate estimates of the process noise covariance matrix. However, there is a considerable difference in the time taken by the optimisation algorithm to converge to a solution, depending on which state estimator is used. In case of the EnKF, it was also found that it is essential to have a good set of samples for initial state, process and measurement noise to avoid numerical ill-conditioning in the C-EnKF and the optimisation algorithm. The studies also indicate a significant improvement in the performance of the state estimators when the optimum values of the covariance matrices are used. The performance of the state estimators using the estimated values of the covariances is comparable to the performance when the true values of the covariances are used. It was observed that the convergence obtained using fmincon function of the Optimization Toolbox in Matlab is good only when the initial guess for the covariance matrix parameters is close to the true values. Efforts are being made to improve the optimisation scheme so as to achieve larger region of convergence.

REFERENCES


Appendix A. CONSTRAINED ENKF (C-ENKF)

To account for the bounds on the state estimates, Prakash et al. (2010a) have proposed a modification in the measurement update step. The time update step in the C-EnKF is similar to the EnKF. Predicted particles, \( \tilde{x}_{k|k-1}^{(i),c} \), lying outside the feasible limits are projected back onto the constrained space

\[
\tilde{x}_{k|k-1}^{(i),c} = \text{Pr} \left[ \tilde{x}_{k|k-1}^{(i)} \right]
\]

For the measurement update step, the following modifications are done. The sample covariance of the predicted state estimates is obtained as

\[
P_{k|k-1} = \frac{1}{N} \sum_{j=1}^{N} \left[ \tilde{x}_{k|k-1}^{(j),c} - \hat{x}_{k|k-1} \right] \left[ \tilde{x}_{k|k-1}^{(j),c} - \hat{x}_{k|k-1} \right]^T
\]

For the \( j \)th particle, the constrained optimisation problem is formulated as follows

\[
\begin{align*}
\min_{\tilde{x}_{k|k}^{(j)}} & \left[ \epsilon_{k|k}^{(j)} \right]^T P_{k|k-1}^{-1} \epsilon_{k|k}^{(j)} + \left[ \epsilon_{k|k}^{(j)} \right]^T R^{-1} \epsilon_{k|k}^{(j)} \\
\epsilon_{k|k}^{(j)} &= \tilde{x}_{k|k-1}^{(j),c} - \hat{x}_{k|k}^{(j)} \\
\epsilon_{k|k}^{(j)} &= y_k - \left( h \left( \tilde{x}_{k|k-1}^{(j),c} + v_{k}^{(j)} \right) \right) \\
\text{Subject to} & \\
x_{L} \leq \tilde{x}_{k|k}^{(j)} \leq x_{H}
\end{align*}
\]

The solution of the above optimisation problem yields the updated particles \( z_{k|k}^{(j)}(j = 1, 2, \ldots, N) \). The updated estimate of the states is computed using Eq. 21.