Fault Detection in Process Systems using Hidden Markov Disturbance Models

Wee Chin Wong [*] **Jay H. Lee** [*]

[*] *School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: {weechin.wong, jay.lee}@chbe.gatech.edu).*

Abstract: Fault detection and diagnosis is critical for maintaining the health of process systems. Common fault signals include process and disturbance parameter changes, as well as sensor and actuator malfunctions (such as persistent drifts and biases). These may be characterized by the existence of latent 'fault' states. This work examines the effectiveness of a Hidden Markov Model framework for modeling such fault regimes. The proposed methodology may be interpreted as a generalization of a commonly-employed Mixture-of-Gaussians (Kesavan and Lee (1997)) approach and is demonstrated through a shell-and-tube heat exchanger problem. Furthermore, the flexibility of the method is shown in the context of detecting valve stiction. This is a significant problem in process industries where a valve's output suffers from excessive friction and is unable to track its input leading to degradation in closed-loop performance.

Keywords: Fault detection, Markov model, Sensor failure, Control valves.

1. INTRODUCTION

Tracking the closed-loop performance and health of process systems, although intuitively important, is oftentimes overlooked during the design of control solutions. Maintenance, required to mitigate the effects of system faults, typically necessitates expert personnel not found within normal plant situations (Kesavan and Lee (1997)). For this reason, multiple process monitoring algorithms have been developed so that such faults may be automatically detected, diagnosed and eventually removed.

Process monitoring methods may be further classified as i) data-driven ii) analytical and/ or iii) knowledge-based (Chiang et al. (2001)). The first involves statistical treatment of large quantities of process data and are typified by data-mining and machine learning techniques (such as principal and independent component analysis), statistical control charts and so on. Knowledge-based methods employ qualitative reasoning and are oftentimes rules-based with a strong logic underpinning. A thorough overview of all three classes is presented by Chiang et al. (2001) and the references therein.

This work, relying on dynamical models of the process for fault detection, is a particular type of analytical approach. Consequently, a necessary standing assumption is the availability of a mathematical model derived from first principles or otherwise. Given the wide-spread popularity of model-based control (such as Model Predictive Control), the controller's model can be readily ported over for the purpose of fault-detection. A model structure, such as in (1), (2), is therefore relevant in subsequent developments.

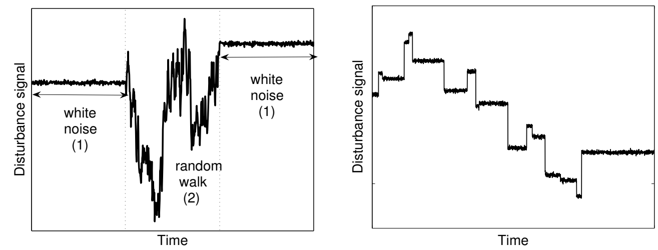
$$\begin{aligned} x_t &= f(x_{t-1}, \theta_{t-1}, u_{t-1}, \omega_t) \\ y_t &= g(x_t, \theta_t, v_t) \end{aligned} \tag{1}$$

$$\begin{aligned} \gamma_t &= \mathbb{A}_t \gamma_{t-1} + \mathbb{B}_t \varphi_t \\ \theta_t &= \mathbb{C}_t \gamma_t + e_t \end{aligned} \tag{2}$$

Here, $x_t \in \mathbb{R}^{n_x}$ represents the state at discrete time index t, $u_t \in \mathbb{R}^{n_u}$, the control input, and $y_t \in \mathbb{R}^{n_y}$, a noise-corrupted measurement signal. $\theta_t \in \mathbb{R}^{n_\theta}$ represents a fault vector with potentially time-varying dynamics governed by matrices $(\mathbb{A}_t, \mathbb{B}_t, \mathbb{C}_t)$ and noise vectors (φ_t, e_t) [1]. ω_t and v_t are process and measurement noise signals respectively. $f(\cdot)$, which may represent an integration of the continuous-time model over a unit sample-time, is the state transition map. Similarly, $g(\cdot)$ represents the state-to-output map.

Faults are typically manifested (Kesavan and Lee (1997)) as i) process parameter changes, and/ or ii) disturbance parameter changes, as well as iii) actuator and sensor problems – all captured by θ. Depending on circumstances, these may be sudden jumps (*e.g.* due to an abrupt introduction of significant sensor bias), or slow drifts or random walk-type changes (*e.g.* as a result of catalyst fouling) or even a mixture of both (Fig. 1). Such failure modes, which cannot be directly observed, and need to be estimated, are conveniently incorporated into the fault model (2) by adding the notion of latent states (denoted by r), each of which modifies the fault model (see (2)) differently. This work explores the use of a Hidden Markov chain, used previously to model realistic disturbances in the context of process control (Wong and Lee (2007)), to describe the temporal, probabilistic transitions between the latent states. Furthermore, this work can be interpreted as a generalization of the popular approach of assuming statistical independence, from one time period to the next, between hidden states. For example, at each time step t, Willsky (1976) and Kesavan and Lee (1997) allowed the statistics

[1] In practice, the user would model θ according to disturbance scenarios of interest.



(a) Intermittent drifts: white noise probabilistically interspersed with integrated white noise.

(b) Abrupt jumps.

Fig. 1. Possible disturbance signals (θ).

of φ_t and e_t to be described by a Mixture-Of-Gaussians (MOG) [2]. This captures the situation where faults that do occur happen infrequently but with significantly larger magnitudes. Persistent faults like drifts, which are easily described by the proposed Hidden Markov Model (HMM) approach, are captured in the MOG context by introducing additional states or non-linearities in the model.

The main contribution is to show that the aforementioned faults (abrupt jumps/ biases and drifts) can be better modeled and detected by the proposed method. Another novel application is in the context of detecting valve stiction, where it is demonstrated that the output of the valve (which is not normally measured) can be effectively tracked using the same proposed framework.

Section 2 provides the details behind an HMM, its subsequent use for fault detection and relevance to prior work. Section 3 demonstrates the effectiveness of the proposed method in the context of a heat exchanger. Section 4 explores the valve stiction issue before concluding remarks regarding future research are presented in Section 5.

2. FAULT MODELING USING HIDDEN MARKOV MODELS

HMMs represent a useful class of statistical models where a latent state, taking values from an alphabet $\mathcal{J} \in \{1, 2, \ldots, J \in \mathbb{Z}^+\}$ of cardinality J, transitions probabilistically in a Markovian [3] fashion from one sampling time to the next. Mathematically, a finite-state Markov chain is a sequence of random variables $(r_0, r_1, ..., r_t, ...)$, where the transition probability matrix $\Pi = (\pi_{ij}) = (pr(r_t = j|r_{t-1} = i), i, j \in J) : \sum_{j=1}^{J} \pi_{ij} = 1, \forall i \in \mathcal{J}$, governs the probabilistic temporal transitions. The term 'Hidden' signifies that the actual regime label is usually not known with complete certainty and must be inferred from available noisy measurements of itself or other related states. In the simplest case, each latent state has a probability distribution over a finite set of possible output symbols. All Markov chains under consideration are ergodic. For simplicity, the Markov chain is assumed to be at steady state, satisfying $\pi = \Pi'\pi$, where π is a column vector containing the unconditional and initial probabilities of each regime. HMMs have found widespread applications in science and engineering - ranging from speech recognition (Rabiner (1989)) to bioinformatics and diverse fields such as econometrics.

HMMs and their generalizations have been used in fault detection, with significant differences to our proposed approach. Smyth (1994), for example, did not consider an explicit fault model (*i.e.,* (2)). Instead, the process parameters are continuously estimated (in batch mode) and treated as output of an underlying Markov chain. This necessitates linking the process parameter vector to fault modes, which is not always possible. A recursive maximum a posteriori filter is then used for fault-mode detection. Huang (2008) suggested a similar (see Section 2.1) HMM approach to sensor problem diagnosis but limited considerations to faults in the output channels and input signals taking values from a finite, discrete set. Almeida and Park (2008) learned an HMM corresponding to each operating condition and, unlike the approach proposed in this work, does not make use of the process model. There, fault detection is achieved by a classification scheme that chooses the HMM that maximizes the probability of a given sequence of observations.

2.1 Proposed Fault Model: Intermittent Drifts & Abrupt Jumps

Following the successes in other fields, a generalization of (2) is considered by allowing the statistics of (φ_t, e_t) (and potentially the fault model parameters $(\mathbb{A}, \mathbb{B}, \mathbb{C})$) to vary according to a hidden Markov chain.

Intermittent Drifts. In the case of one-dimensional intermittent drifts (Fig. (1a)), one has:

$$\begin{aligned} \gamma_{t+1} &= \gamma_t + \varphi_{r_{t+1}} \\ \theta_t &= \gamma_t + e_t \\ r_t &\in 1, 2 \\ \pi_{11} &\approx 1, \, \pi_{11} < 1 \\ \pi_{22} &\approx 1, \, \pi_{22} < 1 \end{aligned} \tag{3}$$

Here, φ_{r_t} and e_t are uncorrelated, zero-mean Gaussian signals with covariances (that may depend on r_t) of $Q^{\varphi}_{r_t}$ and Q^e_t. The abuse of notation on the subscript of φ emphasizes the dependence of the covariance of the noise signal on the underlying Markov chain. When $r_t = 1$ (*i.e.,*, the white-noise regime), $Q^{\varphi}_{r_t=1} \approx 0$. Random-walk type behavior occurs when the hidden state switches to $r_t = 2$, where $Q^{\varphi}_{r_t=2} >> 0$; Q^e_t is invariant to the hidden regime and of appropriate magnitude. Since it is common that there is low probability of switching once the system enters a particular regime, a diagonally-dominant Π is employed.

Abrupt Jumps. In the case of modeling abrupt jumps, (3) is adjusted such that $\pi_{11} = \pi_{12} = p \approx 1$, $p < 1$, so that $\Pi = [p, \, 1-p; p, \, 1-p]$. This ensures that the jump state (the second one, in this case) is infrequently accessed and when it is, a significant step-change occurs.

In this latter case, since it is assumed that the Markov chain is at steady state, this form of the transition matrix implies that the probability of entering a particular regime is independent of the current mode. It is thus clear that the HMM framework subsumes an MOG description.

[2] *i.e.,* at each time step, a member from a set of Gaussians, from which the noise signal is to be sampled, is selected with some time-invariant probability.

[3] transitions depend only upon the immediate past.

Fault detection and diagnosis is performed via state estimation (in particular to track θ) without the knowledge of the latent state trajectory. Hence, a brief mention of state estimation, based on a model resulting from the concatenation of (1), and (2) is necessary.

2.2 Fault Detection via State Estimation of Jump Markov Systems

Equations (1) and (2) can be merged to yield:

$$\begin{bmatrix} x_{t+1} \\ \gamma_{t+1} \end{bmatrix} = \mathcal{F}_{r_{t+1}}\Big(\begin{bmatrix} x_t \\ \gamma_t \end{bmatrix}, u_t, \xi_{r_{t+1}} \Big)$$
$$y_t = \mathcal{G}_{r_t}\Big(\begin{bmatrix} x_t \\ \gamma_t \end{bmatrix}, n_{r_t} \Big)$$
$$pr(r_t = j | r_{t-1} = i) = \pi_{ij} \tag{4}$$

Here, $\mathcal{F}$ is implicitly understood to include model structures and parameters from $\{f, \mathbb{A}, \mathbb{B}, \mathbb{C}\}$ and the hidden Markov chain. A similar remark is extended to $\mathcal{G}$. Besides $\mathcal{F}$ and $\mathcal{G}$, the statistics of the noise ξ (a concatenation of (ω, φ, e)) and n (a concatenation of (v, e)) can depend on r. The system represented by (4) is also termed a Markov jump system. Without knowledge of the sequence $(r_0, \ldots, r_t)$, the optimal filter involves averaging over an exponentially growing number of linear filters. The number of filters scales as J^t, where J is the cardinality of the set containing all possible realizations of r.

The following paragraphs outline the Generalized Pseudo Bayesian estimation algorithm of order 2 (GPB2), a popular sub-optimal method, developed by Bar-Shalom and Li (1993). The main idea to have trajectories whose last 2 terms differ be merged (via moment-matching) into a single Gaussian. Using the law of total probability and Bayes' Rule, it can be shown that:

$$x_{t+1|t+1} = \sum_{r_{t+1}} p(r_{t+1}|t+1) x_{t+1|(t+1,r_{t+1})}$$
$$x_{t+1|(t+1,r_{t+1})} \triangleq \sum_{r_t} x_{t+1|(t+1,r_{t+1},r_t)} p(r_t|r_{t+1}, t+1)$$
$$P_{t+1|t+1} = \sum_{r_{t+1}} \{(x_{t+1|t+1} - x_{t+1|(t+1,r_{t+1})})(\cdot)' + P_{t+1|t+1,r_{t+1}}\} p(r_{t+1}|t+1)$$
$$P_{t+1|t+1,r_{t+1}} = \sum_{r_t} \{(x_{t+1|t+1,r_{t+1}} - x_{t+1|(t+1,r_{t+1},r_t)})(\cdot)' + P_{t+1|t+1,r_{t+1},r_t}\} p(r_t|r_{t+1}, t+1)$$
$$p(r_t|r_{t+1}, t+1) = \frac{1}{c_1} p(y_{t+1}|t, r_{t+1}, r_t) p(r_t|r_t) p(r_t|t)$$
$$p(r_{t+1}|t+1) = \frac{1}{c_2} \sum_{r_t} p(y_{t+1}|t, r_{t+1}, r_t) p(r_{t+1}|r_t) p(r_t|t)$$

The term $p(y_{t+1}|t, r_{t+1}, r_t)$ refers to the probability density of the corresponding one-step ahead output prediction. $x_{t+1|(t+1,r_{t+1})}$ refers to the estimate of x_{t+1} given output measurements $\{y_0, \ldots, y_{t+1}\}$ and a certain realization of r_{t+1}; $P_{t+1|(t+1,r_{t+1})}$ denotes the corresponding error covariance matrix. The pair $(x_{t+1|(t+1,r_{t+1},r_t)}, P_{t+1|(t+1,r_{t+1},r_t)})$ are similarly defined. It is noted that starting from $(x_{t|(t,r_t)}, P_{t|(t,r_t)})$, a single application of the time and measurement update steps of the (extended) Kalman filter yields these latter quantities. c_1 and c_2 are normalizing constants such that the merging probabilities $p(r_t|r_{t+1}, t+1)$ and $p(r_{t+1}|, t+1)$ sum to unity.

2.3 A-posteriori Regime Estimation

If required, a prediction and/ or filtered estimate of the hidden regime can be obtained viz:

$$\hat{r}_{t+1|t} = \arg\max_{r_{t+1}} \Big\{ p(r_{t+1}|t) \triangleq \sum_{r_t} pr(r_{t+1}|r_t) \cdot pr(r_t|t) \Big\}$$
$$\hat{r}_{t|t} = \arg\max_{r_t} \{p(r_t|t)\} \tag{5}$$

3. EXAMPLE 1: FAULT TRACKING IN A SHELL & TUBE HEAT EXCHANGER

In this example, the usefulness of the proposed method in detecting faults is studied in the context of a shell and tube heat exchanger (6) also considered by Kesavan and Lee (1997). In particular, we contrast the proposed HMM approach against an MOG method (Kesavan and Lee (1997)) in modeling the latent states that govern the fault signals (see Section 3.1 for simulation details). The main difference is that the latter framework assumes that each latent state occurs with a (time-invariant) probability that is independent of the previous realization. The governing non-linear ordinary differential equations used for simulation but not estimator design, are:

$$\frac{dT_c}{dt} = \frac{q_c}{V_c}(T_{ci} - T_c) + \frac{\alpha_c}{V_c}(T_h - T_c)$$
$$\frac{dT_h}{dt} = \frac{q_h}{V_h}(T_{hi} - T_h) - \frac{\alpha_h}{V_h}(T_h - T_c)$$
$$y = \begin{pmatrix} T_c \\ T_h \end{pmatrix} + \mu_v + v \tag{6}$$

Here, the measured state variables are the temperatures of the hot and cold streams respectively: $[T_c; T_h]$. $[T_{ci}; T_{hi}]$ are the temperatures of the incoming cold and hot streams respectively. $[\alpha_c; \alpha_h]$ are system parameters reflecting the heat transfer coefficient, heat transfer area, density, specific heat capacity of the cold and hot streams respectively. Similarly, $[q_h; q_c]$ are the flow rates of the hot and cold streams and represent the degrees of freedom available to a controller. $[V_c; V_h]$ are the volumes of the cold and hot sides. Steady-state values are reported in Table 1. v refers to zero-mean measurement noise of covariance $R \triangleq \mathbb{E}[vv']$. μ_v is nominally a null vector but might be subject to changes due to disturbances.

3.1 Simulation Conditions

Although a variety of fault types may be considered (*e.g.* those affecting the various input and output channels and/ or changes in parameters (α_c, α_h), as discussed in Section 1), for clarity of exposition, only two different fault types are assumed. Furthermore, these affect only the cold side. Given initially quiescent conditions (see Table 1), one considers:

(1) An abrupt step that is normally distributed with zero mean [L/min] and variance q_u^{hi} [$L^2/\min^2$] affecting the input channel on the cold side (q_c) at some unknown time t_u. This may be thought of as a sudden bias developing in the input channel:

$$q_{ct} = q_{ct-1} + \varphi_t^u \cdot \delta(t, t_u),\ \varphi_t^u \sim \mathcal{N}(0, q_u^{hi}) \quad (7)$$

$\delta(\cdot,\cdot)$ is the Dirac delta function. q_u^{hi} has a value of 2 in the following experiments.

(2) A sudden drift (see Fig. 1a) affecting the sensor relaying T_c (*i.e.,* y_1) measurements between an unknown time span: $\mathcal{T} \triangleq [t_{y,1}, t_{y,2}]$. Namely, one has:

$$\mu_{v,1_t} = \mu_{v,1_{t-1}} + \varphi_t^y \quad (8)$$

where $\mathbb{E}[\varphi_t\varphi_t'] = q_y^{hi} = 0.5$ if $t \in \mathcal{T}$ and $\mathbb{E}[\varphi_t\varphi_t'] = q_y^{lo} = 10^{-10} \approx 0$, for other time periods. $\mu_{v,2}$ remains at the origin for all time.

The above non-linear model is not available for state estimation. Instead, a version linearized about the nominal operating conditions is available. With a sampling time of 0.5 min, $A = [0.91,\ 0.03; 0.03, 0.91]$, $B =$ [-0.12, 0.002; -0.002, 0.12], $C = diag([1,\ 1])$. Measurement covariance, R, is set to $diag([0.5,\ 0.5])$ and known. Since estimation is the focus of this example, the system is run in the absence of feedback control.

3.2 Proposed HMM Method to Handle Abrupt Jumps & Intermittent Drifts

The following Markov jump linear model, a specialization of (4), is employed:

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + b\theta_t^u + \omega_{t+1} \\ \theta_{t+1}^u &= \theta_t^u + \varphi_{r_{t+1}}^u \\ \theta_{t+1}^y &= \theta_t^y + \varphi_{r_{t+1}}^y \\ y_t &= Cx_t + \theta_t^y + v_t \end{aligned} \quad (9)$$

where x_t, the state variable at discrete time index t are deviations from $[T_c^*; T_h^*]$. Similarly, the vector $u_t \in \mathbb{R}^2$ represents deviations from $[q_c^*; q_h^*]$. b represents the first column of matrix B, consistent with the fact that disturbances enter the q_c channel. $[\theta^u;\ \theta^y]$ are input and output disturbance state variables respectively. Both θ^u and θ^y are modeled as integrators but distinguished by the effects of the hidden Markov regime on the second moments of φ^u and φ^y. Consistent with the assumption of an abrupt jump, the covariance of φ^u is assumed to be large with a small probability, and vice versa. θ^y is naturally modeled as an intermittent drift (see (3)). Details are given in the following paragraphs.

Table 1. Nominal steady state operating conditions

Variable	Value	Units
$q_c^* = q_h^*$	10	L/min
T_{ci}^*	25	^{o}C
T_{hi}^*	100	^{o}C
T_c^*	43.75	^{o}C
T_h^*	81.25	^{o}C
α_c^*	5	m^3/min
α_h^*	5	m^3/min
$V_c^* = V_h^*$	75	L

A four-regime Markov chain is considered. These regimes represent the following scenarios:

(1) No disturbance in input channel, No disturbance in output channel ('LO-LO')
(2) No disturbance in input channel, Drifting disturbance in output channel ('LO-HI')
(3) Abrupt disturbance in input channel, No disturbance in output channel ('HI-LO')
(4) Abrupt disturbance in input channel, Drifting disturbance in output channel ('HI-HI')

Accordingly, a simple method for determining the values of the transition probability matrix (Π) is proposed. Per the earlier discussion (Section 2.1), two (sub) transition probability matrices are appropriate for the input (Π^u) and output channels (Π^y) respectively, the first state being the 'normal' regime in both cases.

$$\Pi^u = \begin{pmatrix} 0.99 & 0.01 \\ 0.99 & 0.01 \end{pmatrix};\ \Pi^y = \begin{pmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{pmatrix} \quad (10)$$

An overall transition probability matrix (Π) accounting for the four scenarios can be obtained by assuming statistical independence between the input and output channels. For example in computing π_{23}, one has transitions between the 'normal' to 'abnormal' state for the input channel and the opposite transitions for the output channels so that

$$\pi_{23} = \pi_{21}^u \pi_{12}^y \quad (11)$$

The overall Π [4] is:

$$\begin{pmatrix} 0.98 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.98 & 0.01 & 0.01 \\ 0.98 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.98 & 0.01 & 0.01 \end{pmatrix}$$

In accordance to the noise statistics of the possible fault scenarios, the covariance of the overall noise vector $\xi_t \triangleq [\omega_t,\ \varphi_t^u,\ \varphi_t^y]$ for the 4 regimes are:

(1) 'LO-LO': $\mathbb{E}[\xi_t\xi_t'] = diag([10^{-10},\ 10^{-10},\ 10^{-10},\ q_y^{lo}])$
(2) 'LO-HI': $\mathbb{E}[\xi_t\xi_t'] = diag([10^{-10},\ 10^{-10},\ 10^{-10},\ q_y^{hi}]$
(3) 'HI-LO': $\mathbb{E}[\xi_t\xi_t'] = diag([10^{-10},\ 10^{-10},\ q_u^{hi},\ q_y^{lo}]$
(4) 'HI-HI': $\mathbb{E}[\xi_t\xi_t'] = diag([10^{-10},\ 10^{-10},\ q_u^{hi},\ q_y^{hi}]$

Process noise ω is negligible compared to θ^u and will be assumed to be absent for simplicity.

3.3 Alternative MOG Description

If one were to be restricted to an MOG description of the latent regime, then an additional state (θ^β) is required:

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + b\theta_t^u + \omega_{t+1} \\ \theta_{t+1}^u &= \theta_t^u + \varphi_{t+1}^u \\ \theta_{t+1}^\beta &= \theta_t^\beta + \varphi_{t+1}^\beta \\ \theta_{t+1}^y &= \theta_t^y + \theta_t^\beta \\ y_t &= Cx_t + \theta_t^y + v_t \end{aligned} \quad (12)$$

Similar to (9), θ^u refers to the input channel disturbance and is modeled as an abrupt jump. However, the output

[4] the rows do not sum to unity due to rounding errors

disturbance (θ^y) is now modeled as a double integrator (driven by θ^β). θ^β itself may be interpreted as a velocity term and is driven by φ^β which is set to have a small covariance (10^{-10}) with large probability and a large covariance (of q_y^{hi}) with small probability. This captures the (rare) event of a velocity change when the output disturbance transitions from the white-noise regime to the random-walk mode and vice versa (see Fig. 1(a)). In this case, the sub transition matrices for the input and output channels are:

$$\Pi^u = \Pi^y = \begin{pmatrix} 0.99 & 0.01 \\ 0.99 & 0.01 \end{pmatrix}$$

The overall transition matrix may be obtained as before, per (11). The covariance of the overall noise vector $\xi_t \triangleq [\omega_t,\, \varphi_t^u,\, \varphi^\beta,\, \varphi_t^y]$ for the 4 regimes are:

(1) $\mathbb{E}[\xi_t\xi_t'] = diag([10^{-10},\, 10^{-10},\, 10^{-10},\, q_y^{lo},\, 10^{-10}])$
(2) $\mathbb{E}[\xi_t\xi_t'] = diag([10^{-10},\, 10^{-10},\, 10^{-10},\, q_y^{hi},\, 10^{-10}]$
(3) $\mathbb{E}[\xi_t\xi_t'] = diag([10^{-10},\, 10^{-10},\, q_u^{hi},\, q_y^{lo},\, 10^{-10}]$
(4) $\mathbb{E}[\xi_t\xi_t'] = diag([10^{-10},\, 10^{-10},\, q_u^{hi},\, q_y^{hi},\, 10^{-10}]$

3.4 Example 1: Results

Table 2 presents a summary (average over 100 realizations) of the state-estimation error for both the input and output channel. A typical realization is depicted in Fig. 2.

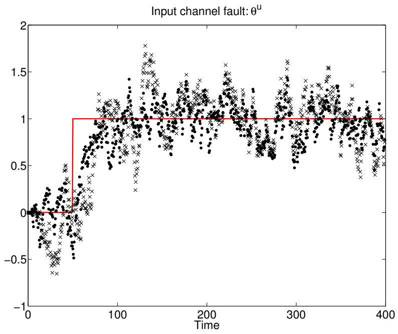
Table 2. 2-norm of state-estimation error (Average of 100 realizations)

Channel	Proposed see (9)	MOG approach see (12)
Input	11.4	12.9
Output	13.3	19.7

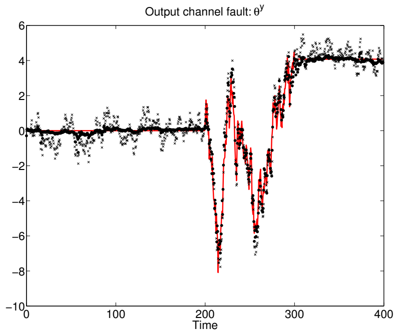
Due to the similarities in modeling the abrupt jump in the output channel, it can be seen from Fig. 2(a) and the first line of Table 2 that the performance of the state estimator corresponding to both approaches yield similar performances. However, the MOG approach fares significantly worse than the proposed HMM approach in tracking the fault signal (which is an intermittent drift) corresponding to the output channel (see Fig. 2(b) and the second row of Table 2).

4. EXAMPLE 2: VALVE STICTION

Valve stiction is a common problem in control valves, the latter being widely used in process industries (Choudhury et al. (2005)). Due to the effects of friction, the output (u^x) of the control valve does not track its input (u^c) (*i.e.,* the control signal prescribed by the controller) instantaneously. Instead, u^x has been observed to demonstrate a delayed and sluggish response to u^c, where the valve 'sticks' to its current position if changes in the control signal (and/ or the absolute magnitude itself) are insufficiently large to overcome friction effects. This is usually to the detriment of closed-loop performance. It is assumed that the plant is linear and therefore parameterized by matrices (A, B, C), where A is the state-transition map, B, the input-to-state map and C, the state-to-output map. Technical definitions, first-principles and empirical models of stiction can be found in the articles by Choudhury et al. (2005, 2008) and the references therein. For simplicity,



(a) Tracking θ^u



(b) Tracking θ^y

Fig. 2. Tracking θ^u and θ^y. Comparing the proposed HMM vs. MOG approaches. Legend: solid line - actual fault signal; Dots ($\cdot$) - HMM; Crosses (x) - MOG

an efficient single-parameter model employed by Stenman et al. (2003) and Srinivasan and Rengaswamy (2005) for stiction detection is used for simulations in the sequel:

$$u_t^x = \begin{cases} u_{t-1}^x, \text{ if } |u_t^c - u_{t-1}^x| \leq d \\ u_t^c, \text{ otherwise} \end{cases} \tag{13}$$

where d represents the valve stiction band. The larger the value of d, the more severe the stiction problem.

The detection, diagnosis and compensation-for valve stiction has received much attention in academia and industry. Based on (13), Stenman et al. (2003) proposed a suitable model for detecting stiction:

$$u_t^x = \tilde{\delta}_t \cdot u_{t-1}^x + (1 - \tilde{\delta}_t) \cdot u_t^c$$

where $\tilde{\delta}_t$ is a binary (0/1) mode parameter occurring with a certain (i.i.d) probability.

For the same purpose of stiction detection and estimating the typically unmeasured u_t^x, we allow $\tilde{\delta}_t$ to have statistics governed by an underlying Markov chain so that observations reflecting persistent 'stickiness' can be more effectively modeled. Also, instead of identifying the segmentation sequence $\{\tilde{\delta}_1, \ldots, \tilde{\delta}_t\}$ that maximizes the posterior quantity $pr(\tilde{\delta}_1, \ldots, \tilde{\delta}_t | y_1, \ldots, y_t)$ through dynamic programming, we propose a novel Markov jump linear description that is consistent with (13) to be used by a GPB2 state-estimator:

$$\begin{pmatrix} x_t \\ u^x_{t-1} \end{pmatrix} = \begin{pmatrix} A & B^x_{r_{t-1}} \\ 0 & \tilde{\delta}_{r_{t-1}} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ u^x_{t-2} \end{pmatrix} + \begin{pmatrix} B^c_{r_{t-1}} \\ 1-\tilde{\delta}_{r_{t-1}} \end{pmatrix} u^c_{t-1}$$

$$y_t = (\, C \;\; 0\,) \begin{pmatrix} x_t \\ u^x_{t-1} \end{pmatrix} + v_t \tag{14}$$

When $r = 1$, stiction is absent, $\tilde{\delta} = 0$, $B^x = 0$, $B^c = B$. Conversely, when $r = 2$, stiction is present, $\tilde{\delta} = 1$, $B^x = B$, $B^c = 0$.

4.1 Simulation Studies: Mixing Tank

For simulation studies, we consider a simple isothermal mixing-tank (of cross-sectional area $\mathcal{A}$) with an outlet stream whose flow-rate is controlled by a valve (with resistance $\mathcal{R}$):

$$\frac{dm}{dt} = \frac{1}{\mathcal{A}}(q_1 + q_2 - \frac{m}{\mathcal{R}}) \tag{15}$$

The controlled (and also measured) variable is the liquid level (m). The flow-rate of the first stream, q_1, is a measured disturbance whereas that of the other stream (q_2) represents the manipulated variable. A PI controller (with gain K_c, and integral time constant τ_I) is given by:

$$u^c_t = u^c_{t-1} + K_c[e_t - e_{t-1} + \frac{h}{\tau_I} e_t],\ e_t \triangleq l - y_t$$

Here l is the set-point, nominally calibrated to a value of 6. For ease, $\mathcal{A}$, $\mathcal{R}$, K_c, τ_I and the measured disturbance signal q_1, are all set to nominal values of 1. A relatively large value for the stiction band is employed: $d = 0.5$. A sampling time of $h = 0.05$ is employed, resulting in the following parametrization to be used by the state estimator: $A = 0.951$, $B = 0.0488$ and $C = 1$. Measurement noise is set to have a known covariance of $R \triangleq \mathbb{E}[v_t v_t'] = 10^{-4}$. To reflect the high degree of stiction, the transition probability matrix Π is:

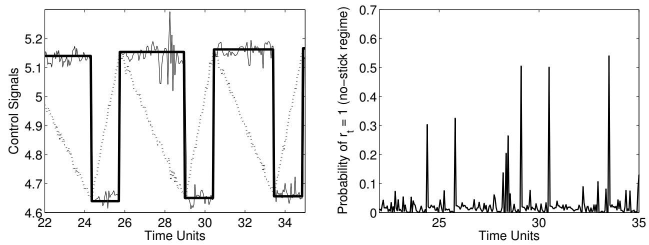
$$\begin{pmatrix} 0.01 & 0.99 \\ 0.01 & 0.99 \end{pmatrix} \tag{16}$$

4.2 Results: Estimating Valve Output & Detecting Stiction

Tracking results for a typical closed-loop realization are shown in Fig. 3. The existence of the cycles in u^c and u^x (Fig. 3(a)) is due to the presence of integral action as well as the valve stiction phenomenon. From Fig. 3(a), it can be seen that the proposed methodology is able to estimate u^x. Observing the (a-posteriori) probability (see (5) and Fig. 3(b)) of the first mode (or equivalently the second) via reveals the time instances where a switch occurs (by means of the probability peaks). Doing so represents an effective way for detecting stiction.

5. CONCLUSIONS & FUTURE WORK

The main contribution of this work is to show that the common faults (abrupt jumps/ biases and drifts) can be better modeled and detected by the proposed HMM-based method. Another novel application is in the context of detecting valve stiction, where it is demonstrated that the output of the valve (which is not normally measured) can be effectively estimated. Future work involves extending the problem to large scale systems (*e.g.* a network of unit operations) of industrial interest.



(a) Time series plot of valve input and output and estimate of valve output. Legend: dotted: u^c; bold line: u^x; line: $\hat{u}^x$

(b) Probability of mode 1

Fig. 3. Tracking unmeasured valve output in mixing-tank example.

REFERENCES

Almeida, G.M. and Park, S.W. (2008). Fault detection and diagnosis in the damadaics benchmark actuator systems - a hidden markov model approach. In *17th IFAC World Congress*, 12419. Seoul, Korea.

Bar-Shalom, Y. and Li, X.R. (1993). *Estimation and Tracking: Principles, Techniques, and Software*. Artech House.

Chiang, L., Russell, E., and Braatz, R. (2001). *Fault detection and diagnosis in industrial systems*. Advanced textbooks in control and signal processing. Springer.

Choudhury, M.S., Jain, M., and Shah, S. (2008). Stiction - definition, modeling, detection and quantification. *Journal of Process Control*, 18, 232–243.

Choudhury, M.S., Thornhill, N.F., and Shah, S. (2005). Modeling valve stiction. *Control Engineering Practice*, 13, 641–658.

Huang, B. (2008). Bayesian methods for control loop monitoring and diagnosis. *Journal of process control*, 18, 829–838.

Kesavan, P. and Lee, J.H. (1997). Diagnostic tools for multivariable model-based control. *Industrial & Engineering Chemistry Research*, 36, 2725–2738.

Rabiner, L.R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257–286.

Smyth, P. (1994). Hidden markov models for fault detection in dynamic systems. *Pattern recognition*, 27, 149–164.

Srinivasan, R. and Rengaswamy, R. (2005). Control loop performance assessment: 1. a qualitative approach for stiction diagnosis. *Industrial & Engineering Chemistry Research*, 44, 6708–6718.

Stenman, A., Gustafsson, G., and Forsman, K. (2003). A segementation-based method for detection of stiction in control valves. *International journal of adaptive control and signal processing*, 17, 625–634.

Willsky, A.S. (1976). A survey of design methods for failure detection in dynamic systems. *Automatica*, 12, 601–611.

Wong, W.C. and Lee, J.H. (2007). Disturbance modeling for process control via hidden markov models. In *8th International Symposium on Dynamics and Control of Process Systems*. Cancun, Mexico.