Proc. of the 43rd IEEE Conference on Decision and Control (CDC 2004)
Atlantis, Paradise Island, Bahamas, December 2004

Fault Detection in a Class of Stochastic Hybrid Systems

Eugenio Cinquemani, Mario Micheli and Giorgio Picci

***Abstract*— We consider the problem of *fault detection* for a Stochastic Hybrid System described by a state-space model with the following properties: the state equation evolves continuously in time according to a linear stochastic differential equation; noisy measurements of the continuous state are made available at discrete deterministic times, by a static linear equation; the parameters of both the state evolution and the measurement equation depend on a discrete state described as a continuous-time Markov chain. This model is particularly suitable for those applications where the frequency of measurements is comparable to the "rate of change" of the Markov chain, and allows the discrete state to switch in between measurements. We solve the problem of estimating both the continuous and the discrete state, given the measurements up to a certain time, in an on-line manner. We also provide the results of significant numerical simulations.**

I. Introduction

The problem of fault detection for Stochastic Hybrid Systems has often been formulated in terms of state estimation of discrete-time Jump Markov Linear Systems. There is already a rich literature on this topic; for example, see Tugnait [23][24], Bar-Shalom [2], Elliott *et al.* [9], Murphy [21], Logothetis & Krishnamurthy [19], Chen & Liu [3], Lerner *et al.* [18], Koutsoukos *et al.* [16][17], Doucet *et al.* [7][8], Hofbaur & Williams [12], Costa *et al.* [5][6], and Germani *et al.* [10]. Others have worked on the continuous-time version of the problem (both in the state dynamics *and* the measurement process). To name a few, see for instance Miller & Runggaldier [20], Hibey & Charalambous [11], Hu *et al.* [13], and Zhang [25].

In the present paper we study the problem of state estimation for a model where the continuous state x evolves in time described by a linear stochastic *differential equation*, and noisy measurements are acquired periodically at fixed deterministic *time instants*. The parameters of both the state equation and the measurement equation depend on a discrete state q which evolves in time as a *continuous-time Markov chain*. The goal is estimating the pair (x, q), given the available measurements. Note that the discrete state may switch (in principle, even more than once) *between* two different measurements. In this work we will restrict our attention to a typical *fault detection* setting, where all but one discrete states are absorbing (see, e.g., Nikiforov [22]). Such model is particularly well-suited for industrial applications where measurements are taken at a frequency that is comparable to the rate of possible abrupt changes (e.g. ruptures) in the plant.

The paper is organized as follows. In section II we introduce the general *continuous*-time dynamics, *discrete*-time measurement *stochastic hybrid model* and formalize the state estimation problems of our concern. We then focus on the special case of fault detection and introduce the switching time $t^\star$ and the final discrete state $q^\star$ as an alternative characterization of the $q(t)$ trajectory. In section III we derive a statistically equivalent model based on system discretization along fixed trajectories of q. Based on this result, the interpretation of the estimation of x as averaging of *conditional Kalman filters* – i.e. ordinary Kalman filters conditioned on $t^\star$ and $q^\star$ – is discussed in section IV. Section V presents methods for the efficient computation of the *a posteriori* joint probability distribution of $t^\star$ and $q^\star$, and shows its connection with the computation of the *a posteriori* density of x for *arbitrary values* of $t^\star$ and $q^\star$. The latter problem is studied in section VI, and is solved by way of an optimal algorithm for the update of the conditional Kalman recursion as a function of $t^\star$ and $q^\star$. Numerical results are shown and briefly discussed in section VII, whereas final comments and perspectives of our work are reported in Section VIII.

For reasons of space, all proofs will be omitted; we refer the reader to [4] for details.

II. Problem Formulation

Let $\mathcal{T} = \{t_k\}_{k\in\mathbb{N}_0}$ be the *deterministic* time sequence $t_k = kT$, where $T > 0$ is a given constant and $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$. Consider a finite *state space* $\mathcal{Q} = \{0, 1, 2, \ldots, N-1\}$ and let q denote its generic element. Assume that we are given matrix functions: $F : \mathcal{Q} \to \mathbb{R}^{n\times n}$, $G : \mathcal{Q} \to \mathbb{R}^{n\times m}$, $H : \mathcal{Q} \to \mathbb{R}^{p\times n}$, and $K : \mathcal{Q} \to \mathbb{R}^{p\times r}$, which assign to each value $q \in \mathcal{Q}$ a 4-tuple of matrices (F_q, G_q, H_q, K_q).

Consider the following dynamical model:[1]

$$\left\{ \begin{array}{rcl} \dot{x}(t) &=& F_{q(t)}x(t) + G_{q(t)}u(t) \\ y_k &=& H_{q(t_k)}x(t_k) + K_{q(t_k)}v_k \end{array} \right. , \quad t \in \mathbb{R},\ t_k \in \mathcal{T}, \tag{1}$$

where $u(t)$ is continuous-time white Gaussian noise, v_k is discrete-time white Gaussian noise and $x(t_0) \sim \mathcal{N}\big(\mu_0, \Sigma_0\big)$. Furthermore, $q(t)$, $t \in \mathbb{R}$ is a continuous-time, homogeneous *Markov chain* with assigned transition probabilities $T_{ij}(\Delta) \triangleq \mathbb{P}[q(t+\Delta) = j \,|\, q(t) = i]$ and initial probabilities $p_i \triangleq \mathbb{P}[q(t_0) = i]$, $i \in \mathcal{Q}$. We shall assume u, v, x_0 and q to be mutually independent.

Process $q(t)$ switches in time between different states in $\mathcal{Q}$, thus changing the parameters of both the state

This work was supported in part by the European Community through the project RECSYS of the V Framework Program. We wish to thank Professor Augusto Ferrante for his precious advice.

Eugenio Cinquemani and Giorgio Picci are with the Department of Information Engineering, University of Padova, I-35131 Padova, Italy eugenio.cinquemani@unipd.it giorgio.picci@unipd.it

Mario Micheli is with the Division of Applied Mathematics, Brown University, Providence RI 02912, USA mariom@dam.brown.edu

[1] We will refer to the first equation in (1) as *state equation*, and the second one as *measurement equation*.

equation and the measurement equation. Our problem is the following: given measurements up to time t_k, that is $y^k \triangleq \{y_0, \dots, y_k\}$ we wish to compute the "best" estimate for the joint state (x, q). More precisely, for $j, k \in \mathbb{N}_0$ we wish to compute the least squares estimate of the *continuous* state $x(t_j)$:

$$\hat{x}_{j|k} \triangleq \arg\min_{z \in \mathbb{R}^n} \mathbb{E}\left[||z - x(t_j)||^2 \,\middle|\, y^k \right] = \mathbb{E}\big[x(t_j) \,\big|\, y^k\big], \tag{2}$$

and at the same time the *a posteriori* probability distribution of the *discrete* state:

$$p_{j|k}(q) \triangleq \mathbb{P}\big[q(t_j) = q \,\big|\, y^k\big] \,. \tag{3}$$

We will mostly restrict our attention to the cases $j = k$ (filtering) and $j = k + 1$ (prediction).

According to our model the discrete state can switch between different values in $\mathcal{Q}$ *between* two successive measurements – in principle, even more than once: this makes the exact computation of the above estimates a formidable task. In the present paper, however, we shall limit ourselves to a *fault detection* setting. That is, we will assume the states $\mathcal{Q} \setminus \{0\}$ to be *absorbing*: for given positive parameters $\{\lambda_1, \dots, \lambda_{N-1}\}$, the *generator matrix* $\mathbf{G} = \left[\frac{\mathrm{d}T}{\mathrm{d}\Delta}\right]_{\Delta=0}$ of process $q(t)$ shall be [15]:

$$\mathbf{G} = \left[\begin{array}{c|ccc} -(\lambda_1 + \ldots + \lambda_{N-1}) & \cdots & \lambda_j & \cdots \\ \hline 0 & & & \\ \vdots & & 0 & \\ 0 & & & \end{array}\right] \tag{4}$$

that corresponds to a transition probability matrix $[T_{ij}(\Delta)]$ of the form:

$$T(\Delta) = \left[\begin{array}{c|ccc} e^{-\Lambda\Delta} & \cdots & \frac{\lambda_j}{\Lambda}\left(1 - e^{-\Lambda\Delta}\right) & \cdots \\ \hline 0 & & & \\ \vdots & & I_{N-1} & \\ 0 & & & \end{array}\right]$$

where $\Lambda \triangleq \sum_{i=1}^{N-1} \lambda_i$. The graphical representation of such Markov process is shown in Figure 1.

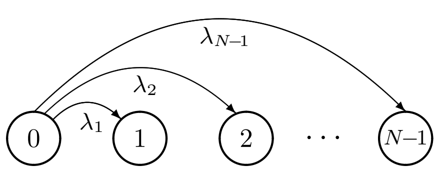

Fig. 1. Graphical representation of Markov Process (4).

In this setting, there can be *at most* one switching of the discrete state from state 0 to any of the absorbing states $\{1, \dots, N-1\}$. Hence, trajectory $q(t)$ is *characterized* by the *switching time* $t^\star$ (i.e. the time at which the event takes place) and the *new discrete state* $q^\star$ (i.e. the value of $q(t)$ for $t > t^\star$). The probability distribution function of $t^\star$ is

$$F_{t^\star}(t) = (1 - e^{-\Lambda t})p_0 + (1 - p_0)$$

for $t \geq 0$, and is undefined for $t < 0$. In particular, when $p_0 = 1$ we have that $t^\star \sim \mathcal{E}(\Lambda)$. We shall compute the *a posteriori* probabilities

$$\mathbb{P}\big[\, t^\star \leq t,\, q^\star = q \,\big|\, y^k \,\big] \tag{5}$$

from which probabilities (3) follow immediately:

$$p_{j|k}(q) = \begin{cases} \mathbb{P}[t^\star > t_j \,|\, y^k] & \text{for } q = 0 \\ \mathbb{P}[t^\star \leq t_j, q^\star = q \,|\, y^k] & \text{for } q \neq 0 \end{cases} .$$

In practice, this will be done by computing and integrating

$$f_{t^\star, q^\star}(t, q|y^k) \triangleq \frac{\mathrm{d}}{\mathrm{d}t}\, \mathbb{P}\big[\, t^\star \leq t,\, q^\star = q \,\big|\, y^k \,\big]; \tag{6}$$

with an abuse of notation we shall indicate the latter function by $f(t^\star, q^\star|y^k)$. Observe that

$$f(t^\star, q^\star|y^k) \propto f(y^k|t^\star, q^\star) \cdot f(t^\star, q^\star)$$

where

$$f(t^\star, q^\star = q) = f(t^\star)\, \mathbb{P}[q^\star = q]$$

(with $f(t^\star) = \Lambda e^{-\Lambda t^\star}$ for $t^\star > 0$ and $\mathbb{P}[q^\star = q] = \lambda_q/\Lambda$) can be easily computed from the system's parameters. Part of the paper, as we shall see, will be devoted to the iterative computation of factor $f(y^k|t^\star, q^\star)$. The *a posteriori* probability density of the switch time also follows:

$$f(t^\star|y^k) = \sum_{q=1}^{N-1} f(t^\star, q^\star = q|y^k).$$

For the sake of clarity, the assumption $p_0 = 1$ will be maintained throughout the rest of the paper.

III. The Conditioned System

Despite the stochastic nature of switching, one may fix the values of $t^\star$ and $q^\star$, i.e. the whole trajectory $q(t)$, and study the system associated to it. What is obtained is a standard linear time-varying Gaussian system with all the parameters determined by $q(t)$. Such a system can be *discretized*, i.e. one may define the *sampled state* $x_k \triangleq x(t_k)$ and the *discrete-time*, time-varying linear Gaussian system,

$$\begin{cases} x_{k+1} = A_k(t^\star, q^\star)\, x_k + u_k \\ y_k = C_k(t^\star, q^\star)\, x_k + D_k(t^\star, q^\star)\, v_k \end{cases} , \tag{7}$$

with $u_k \sim \mathcal{N}(0, Q_k(t^\star, q^\star))$ being a white process independent of $\{v_k\}$ and x_0, so that the joint statistical description of x_k and y_k is identical to that of the original system *given* $t^\star$ *and* $q^\star$. The equivalence is guaranteed by a suitable choice of the parameters $A_k(t^\star, q^\star)$, $Q_k(t^\star, q^\star)$, $C_k(t^\star, q^\star)$, $D_k(t^\star, q^\star)$, which will be discussed below. Clearly, (7) is a state-space representation of the random variables x_k and y_k *conditioned on* $t^\star$ and $q^\star$. Moreover, for changing values of $t^\star$ and $q^\star$, it describes a family of models corresponding to the different possible realizations of $q(t)$. In the sequel we will refer to (7) as the *conditioned system*.

A. Computation of the Conditioned System parameters

In this section we will assume that $q^\star$ is fixed and $t^\star$ takes value in a certain interval (t_h, t_{h+1}), with $t_h, t_{h+1} \in \mathcal{T}$. The interval is assumed to be open without loss of generality.

Lemma 1: Let $q \in \mathcal{Q}$. If F_q and $-F_q$ have disjoint spectra, the *Lyapunov equation* $F_q J_q + J_q F_q^T = -G_q G_q^T$ admits a unique solution in J_q (see [1], pp. 203-204). $\square$

Based on this technical result, the parameters of (7) can easily be computed.

Proposition 1: Assume that F_q and $-F_q$ have disjoint spectra, $q \in \mathcal{Q}$. Then:

For $k \neq h$ (i.e. $t^\star \notin (t_k, t_{k+1})$),

$$\begin{aligned} A_k(t^\star, q^\star) &= A_q & Q_k(t^\star, q^\star) &= J_q - A_q J_q A_q^T \\ C_k(t^\star, q^\star) &= H_q & D_k(t^\star, q^\star) &= K_q \end{aligned}$$

where $A_q \triangleq e^{F_q T}$, and $q = 0$ if $k < h$, $q = q^\star$ if $k > h$;

For $k = h$ (i.e. $t^\star \in (t_k, t_{k+1})$),

$$\begin{aligned} A_k(t^\star, q^\star) &= \widetilde{A}_k(t^\star, q^\star)\widetilde{A}_k(t^\star, 0) \\ Q_k(t^\star, q^\star) &= S_k(t^\star, q^\star) - A_k(t^\star, q^\star) S_k(t^\star, 0) A_k^T(t^\star, q^\star) \\ C_k(t^\star, q^\star) &= H_0 \\ D_k(t^\star, q^\star) &= K_0 \end{aligned}$$

where $\widetilde{A}_k(t^\star, q^\star) \triangleq e^{F_{q^\star}(t_{k+1} - t^\star)}$, $\widetilde{A}_k(t^\star, 0) \triangleq e^{F_0(t^\star - t_k)}$ and, for every $q \in \mathcal{Q}$,

$$S_k(t^\star, q) = J_q - \widetilde{A}_k^{-1}(t^\star, q^\star) J_q \widetilde{A}_k^{-T}(t^\star, q^\star). \qquad \square$$

Remark 1. If the assumption on the spectrum of F_q fails to hold, Lemma 1 cannot be applied. For arbitrary matrices F_q, less convenient expressions for Q_k and S_k (where matrices J_q appear) can still be found.

Remark 2. For $k \neq h$, $A_k(t^\star, q^\star)$ and $Q_k(t^\star, q^\star)$ do not depend on the *specific* value of $t^\star$. In fact, they depend on $t^\star$ *only through* h. The same holds for $C_k(t^\star, q^\star)$, $D_k(t^\star, q^\star)$, for *any* k.

Remark 3. At this stage, all parameters of the conditioned system are expressed as explicit functions of $t^\star$, $q^\star$. Notice that J_q, $q \in \mathcal{Q}$ may be computed *offline* with arbitrary precision using standard numerical techniques. Parameters A_k, Q_k, $k \neq h$ can be computed offline as well.

IV. The Filtering Problem as Averaging of Kalman Filters

We shall now take a deeper look at the estimation problems we stated in section II. For any index j, consider the computation of $\hat{x}_{j|k}$. Application of the Total Probability Law to the probability density of x_j given y^k yields

$$f(x_j|y^k) = \sum_{q^\star=1}^{N-1} \int_0^{+\infty} f(x_j|t^\star, q^\star, y^k) f(t^\star, q^\star|y^k)\, \mathrm{d}t^\star. \quad (8)$$

We recognize $f(x_j|t^\star, q^\star, y^k)$ to be the *a posteriori* density of the state x_j given y^k of the conditioned system (7).

In the light of the discussion of section III, *for any fixed value of* $t^\star$ *and every* $q^\star$ it must hold that

$$f(x_j|t^\star, q^\star, y^k) \sim \mathcal{N}\big(\hat{x}_{j|k}(t^\star, q^\star), P_{j|k}(t^\star, q^\star)\big), \qquad (9)$$

where mean and variance may be interpreted as the minimum error variance estimate of x_j given y^k and the associated error covariance matrix for the conditioned system (see for instance [14]). In particular,

$$\hat{x}_{k|k}(t^\star, q^\star) \qquad (10)$$
$$\hat{x}_{k+1|k}(t^\star, q^\star) \qquad (11)$$

are the *conditional Kalman filter* and the *conditional Kalman predictor* for the corresponding conditioned system, and

$$P_{k|k}(t^\star, q^\star) \qquad (12)$$
$$P_{k+1|k}(t^\star, q^\star) \qquad (13)$$

are the respective error covariance matrices. These may be computed by a *conditional Kalman recursion*:

Measurement update:

$$\begin{aligned} \hat{x}_{k|k} &= \hat{x}_{k|k-1} + L_k(t^\star, q^\star)[y_k - C_k(t^\star, q^\star)\hat{x}_{k|k-1}] \\ P_{k|k} &= P_{k|k-1} - L_k(t^\star, q^\star) C_k(t^\star, q^\star) P_{k|k-1} \end{aligned} \qquad (14)$$

Time update:

$$\begin{aligned} \hat{x}_{k+1|k} &= A_k(t^\star, q^\star)\hat{x}_{k|k} \\ P_{k+1|k} &= A_k(t^\star, q^\star) P_{k|k} A_k^T(t^\star, q^\star) + Q_k(t^\star, q^\star) \end{aligned} \qquad (15)$$

where the (conditional) Kalman gain $L_k(t^\star, q^\star)$ depends on $C_k(t^\star, q^\star)$, $D_k(t^\star, q^\star)$ and $P_{k|k-1}$. In general, as all the system parameters that appear in the recursion depend on $t^\star$ and $q^\star$, so do (10)÷(13). By equation (8), estimate (2) is therefore equal to the conditional average

$$\hat{x}_{j|k} = \sum_{q^\star=1}^{N-1} \int_0^{+\infty} \hat{x}_{j|k}(t^\star, q^\star) f(t^\star, q^\star|y^k) \mathrm{d}t^\star. \qquad (16)$$

Hence, for $j = k$ (or $k+1$), we have a natural interpretation of $\hat{x}_{j|k}$ as *averaging of Kalman filters (or predictors)*. Note that (8) is an average of Gaussian densities – parameterized by $t^\star$ and $q^\star$ – weighted by $f(t^\star, q^\star|y^k)$. What is obtained in general is not at all Gaussian in x_j, hence there is no hope to compute $\hat{x}_{j|k}$ in a linear recursive manner [14].

It is now evident that the probability densities (9) and $f(t^\star, q^\star|y^k)$ play a fundamental role in the estimation process. Hence, with the computation of integral (16) in mind, the attention shifts to deriving *explicit* expressions for (9) (section VI) and $f(t^\star, q^\star|y^k)$ (sections V and VI).

V. Switching Time Estimation

In this section we shall present a technique for the computation of the *a posteriori* probability density $f(t^\star, q^\star|y^k)$ of the switching time and the final discrete state.

The *a posteriori* density $f(t^\star|y^k)$ is simply given by:

$$f(t^\star|y^k) = \sum_{q^\star=1}^{N-1} f(t^\star, q^\star|y^k);$$

analogously, a suitable integration of $f(t^\star, q^\star|y^k)$ yields the *a posteriori* distribution (3). On the other hand, as discussed in the previous section, density $f(t^\star, q^\star|y^k)$ is involved in the computation of the estimate of $x(t_j)$ as well.

$$\widetilde{\Sigma}_k(t^\star,q^\star) \triangleq \mathrm{diag}\big\{Q_{k-1}(t^\star,q^\star),Q_{k-2}(t^\star,q^\star),\ldots,Q_0(t^\star,q^\star),\Sigma_0\big\} \in \mathbb{R}^{(k+1)n\times(k+1)n} \qquad \text{(block diagonal matrix)} \quad (17)$$

$$\widetilde{\mu}_k \triangleq \begin{bmatrix} 0\\ 0\\ \vdots\\ 0\\ \mu_0 \end{bmatrix} \in \mathbb{R}^{(k+1)n}, \quad \Theta_k(t^\star,q^\star) \triangleq \begin{bmatrix} I & A_{k-1} & A_{k-1}A_{k-2} & \cdots & A_{k-1}A_{k-2}\ldots A_2A_1 & A_{k-1}A_{k-2}\ldots A_1A_0\\ 0 & I & A_{k-2} & \cdots & A_{k-2}A_{k-3}\ldots A_2A_1 & A_{k-2}A_{k-3}\ldots A_1A_0\\ 0 & 0 & I & \cdots & A_{k-3}A_{k-4}\ldots A_2A_1 & A_{k-3}A_{k-4}\ldots A_1A_0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & A_1 & A_1A_0\\ 0 & 0 & 0 & \cdots & I & A_0\\ 0 & 0 & 0 & \cdots & 0 & I \end{bmatrix} \quad (18)$$

We will obtain $f(t^\star,q^\star|y^k)$ by first computing $f(y^k|t^\star,q^\star)$ and then applying Bayes'rule. Two different methods for the computation of $f(y^k|t^\star,q^\star)$ are presented, both making use of the results of section III. For the time being, fix particular values for r.v.'s $t^\star$ and $q^\star$; keeping this in mind, sometimes we shall drop $t^\star$ and $q^\star$ from our notation.

A. Direct computation of $f(y^k|t^\star,q^\star)$.

Define vectors and matrices (17) and (18) (note that all the A_i's are functions of $t^\star$ and $q^\star$); also define:

$$\Xi_k(t^\star,q^\star) \triangleq \big[A_{k-1}\big|A_{k-1}A_{k-2}\big|\cdots\big|A_{k-1}A_{k-2}\ldots A_0\big],$$
$$\Upsilon_k(t^\star,q^\star) \triangleq \mathrm{diag}\{C_k(t^\star,q^\star),\ldots,C_0(t^\star,q^\star)\}$$
$$\Lambda_k(t^\star,q^\star) \triangleq \mathrm{diag}\{D_k(t^\star,q^\star),\ldots,D_0(t^\star,q^\star)\}.$$

The following proposition holds:

Proposition 2: With an abuse of notation, let

$$y^k \triangleq [y_k^T,y_{k-1}^T,\ldots,y_0^T]^T \in \mathbb{R}^{(k+1)\times p}$$

be the vector of all measurements up to time t_k. Then y^k conditioned on $t^\star$ and $q^\star$ has the following density:

$$f(y^k|\,t^\star,q^\star) \sim \mathcal{N}\left(\mu_{y^k}(t^\star,q^\star),\Sigma_{y^k}(t^\star,q^\star)\right),$$

where

$$\mu_{y^k}(t^\star,q^\star) \,=\, \Upsilon_k\Theta_k\,\widetilde{\mu}_k = \Upsilon_k \begin{bmatrix} A_{k-1}\cdots A_0\\ A_{k-2}\cdots A_0\\ \vdots\\ A_0\\ I \end{bmatrix}\mu_0,$$
$$\Sigma_{y^k}(t^\star,q^\star) \,=\, \Upsilon_k\Theta_k\,\widetilde{\Sigma}_k\Theta_k^T\Upsilon_k^T + \Lambda_k\Lambda_k^T\,.$$

The above covariance matrix may be computed by the following iteration on k:

- $\Sigma_{y^k}(t^\star,q^\star)$ is obtained by adding n rows and columns to $\Sigma_{y^{k-1}}(t^\star,q^\star)$ as follows:

$$\Sigma_{y^k} \,=\, \begin{bmatrix} \Phi_k & \Psi_k\\ \Psi_k^T & \Sigma_{y^{k-1}} \end{bmatrix}$$

where matrices Φ_k and Ψ_k are given by:

$$\Phi_k \,=\, C_k(\Xi_k\widetilde{\Sigma}_{k-1}\Xi_k^T + Q_{k-1})C_k^T + D_kD_k^T,$$
$$\Psi_k \,=\, C_k\Xi_k\widetilde{\Sigma}_{k-1}\Theta_{k-1}^T\Upsilon_{k-1}^T,$$

with initialization: $\Sigma_{y^0} = C_0\Sigma_0C_0^T + D_0D_0^T$. □

B. Iterative formulation for the computation of $f(y^k|t^\star,q^\star)$.

The above computation may look somewhat cumbersome: however it only requires the computation of the conditioned system's parameters (see section III).

Note that we may write density $f(y^k|t^\star,q^\star)$ simply as follows:

$$f(y^k|\,t^\star,q^\star) = f(y_k|\,t^\star,q^\star,y^{k-1})\,f(y^{k-1}|\,t^\star,q^\star); \quad (19)$$

this formula provides an iterative method for computing $f(y^k|t^\star,q^\star)$. Since y^k is a *given* vector of data, for fixed values of $t^\star$ and $q^\star$ we have that $f(y^{k-1}|t^\star,q^\star)$ is just a number that we carry on from the previous computation. Such number has to be multiplied by $f(y_k|\,t^\star,q^\star,y^{k-1})$, whose value is easily obtainable from $f(x_k|\,t^\star,q^\star,y^{k-1})$. The latter quantity plays a fundamental role in the estimation of continuous state x, as we saw in section IV; in section VI we will show a precise technique for computing it. However, if one is just interested in the joint posterior distribution of $t^\star$ and $q^\star$ and not in the estimation of the continuous state, the formulation given by Proposition 2 may be sufficient. Otherwise, once $f(x_k|\,t^\star,q^\star,y^{k-1})$ is known, the application of (19) is more appropriate.

C. Application of Bayes' rule.

The posterior density of $t^\star$ and $q^\star$ is given by:

$$f(t^\star,q^\star|y^k) = \frac{f(y^k|t^\star,q^\star)f(t^\star,q^\star)}{\sum_{q^\star=1}^{N-1}\int_0^\infty f(y^k|t^\star,q^\star)f(t^\star,q^\star)\,\mathrm{d}t^\star} \quad (20)$$

where $f(t^\star,q^\star=q) = \Lambda e^{-\Lambda t^\star}\mathbb{P}[q^\star=q]$ for $t^\star>0$. Since $f(y^k|t^\star,q^\star)$ is independent of the specific value assumed by $t^\star$ for $t^\star>t_k$, the denominator of (20) is given by

$$\sum_{q^\star=1}^{N-1}\Bigg[\int_{t_0}^{t_k} f(y^k|t^\star,q^\star)f(t^\star,q^\star)\ \mathrm{d}t^\star + \\ + f(y^k|t^\star>t_k,q^\star)\,\mathbb{P}[t^\star>t_k]\,\mathbb{P}[q^\star]\Bigg]. \quad (21)$$

In general, the above integral needs to be evaluated numerically. By exploiting the explicit expression of $f(y^k|t^\star,q^\star)$ (either computed by the direct method or by the iterative method, see section VI), this can be done in $\mathcal{O}(k^2)$ time by standard quadrature methods.

VI. CONDITIONAL KALMAN FILTERING

Following section IV, for any fixed values of $t^\star$ and $q^\star$ one may compute the a posteriori densities

$$f(x_k|t^\star, q^\star, y^k) \tag{22}$$
$$f(x_{k+1}|t^\star, q^\star, y^k) \tag{23}$$

at once by simply running the conditional Kalman recursion corresponding to $t^\star$, $q^\star$. In principle, the procedure allows to compute (22) and (23) for *any* value of $t^\star$ and every possible $q^\star$. In practice, however, it is not suited for the computation of integrals such as (8) and (21) (see also (19) and relevant comments), where (22) and (23) need to be known for each of the absorbing states $q^\star$ and either *for all* $t^\star$ or at least for a relatively large set of values of $t^\star$.

It turns out that the dependence on $t^\star$ and $q^\star$ can be singled out by a suitable restatement of (14) and (15). Fix $q^\star$ and an index $h \in \mathbb{N}_0$, and let $t^\star$ assume *any value* in the interval (t_h, t_{h+1}). We observe that:

(i) (10), (12) and (11), (13) are *independent of the specific* $t^\star$,$q^\star$ for $k \le h$ and $k < h$, respectively;

(ii) for $k \ge h+1$, (10)÷(13) depend on $t^\star$ *only through their new initialization* $\hat{x}_{h+1|h}(t^\star, q^\star)$, $P_{h+1|h}(t^\star, q^\star)$.

Indeed, the parameters of the conditioned system are constant w.r.t. $t^\star$ *and* $q^\star$ as long as $t_k < t_h$, since $q(t) \equiv 0$ for all $t < t_h$. Similarly, they are constant w.r.t. $t^\star$ for all $t_k > t_{h+1}$, since $q(t) \equiv q^\star$ for all $t > t_{h+1}$. Therefore, for *any* $t^\star \in (t_h, t_{h+1})$, (14) and (15) *evolve independently of* $t^\star$ before t_h and after t_{h+1}, whereas the role of $t^\star$ is concentrated in the time update at step $k = h$, with $q^\star$ selecting the new system parameters. Point (i) is easily formalized. Let $\hat{x}^0_{j|k}$ and $P^0_{j|k}$, with $j = k, k+1$, denote the Kalman estimates associated to the condition $q(t) \equiv 0$ (i.e. switch never occurring).

Proposition 3: It holds that

$$\hat{x}_{k|k}(t^\star, q^\star) = \hat{x}^0_{k|k}, \quad P_{k|k}(t^\star, q^\star) = P^0_{k|k}$$

for $k \le h$, and

$$\hat{x}_{k+1|k}(t^\star, q^\star) = \hat{x}^0_{k+1|k}, \quad P_{k+1|k}(t^\star, q^\star) = P^0_{k+1|k}$$

for $k < h$. □

The expressions of $\hat{x}_{h+1|h}(t^\star, q^\star)$, $P_{h+1|h}(t^\star, q^\star)$ follow.

Proposition 4: In the same hypotheses of Proposition 1,

$$\begin{aligned}
\hat{x}_{h+1|h}(t^\star, q^\star) &= A_h(t^\star, q^\star)\hat{x}^0_{h|h}, \\
P_{h+1|h}(t^\star, q^\star) &= J_{q^\star} + \widetilde{A}_h(t^\star, q^\star)(J_0 - J_{q^\star})\widetilde{A}^T_h(t^\star, q^\star) \\
&\quad + A_h(t^\star, q^\star)(P^0_{h|h} - J_0)A^T_h(t^\star, q^\star).
\end{aligned}$$ □

Recall that $A_h(t^\star, q^\star)$ and $\widetilde{A}_h(t^\star, q^\star)$ are known exponential matrices. Hence, new initial conditions for the recursion steps $k \ge h+1$ are given in terms of explicit functions of $t^\star$. It only remains to show how $\hat{x}_{h+1|h}(t^\star, q^\star)$, $P_{h+1|h}(t^\star, q^\star)$ affect (10)÷(13) for $k \ge h+1$. This constitutes the main result of the section.

Proposition 5: Assume that $C_{q^\star}$ is full row rank. Let $C^\dagger_{q^\star} = C^T_{q^\star}(C_{q^\star}C^T_{q^\star})^{-1}$, $\Delta_{q^\star} = C^T_{q^\star}(D_{q^\star}D^T_{q^\star})^{-1}C_{q^\star}$ and

$$Z_{q^\star} = \begin{bmatrix} A^{-T}_{q^\star} & A^{-T}_{q^\star}\Delta_{q^\star} \\ Q_{q^\star}A^{-T}_{q^\star} & A_{q^\star} + Q_{q^\star}A^{-T}_{q^\star}\Delta_{q^\star} \end{bmatrix}.$$

For $k \ge h$, define[2]

$$\begin{aligned}
\Pi_k &= (Z_{q^\star})^{k-h}, \\
\Gamma_{k+1} &= \mathrm{diag}(A^T_{q^\star}, A^T_{q^\star})Z_{q^\star}\Pi_k
\end{aligned}$$

and the recursions

$$\begin{aligned}
N_{k+1} &= [A^T_{q^\star}\Pi^{1,1}_{k+1} - \Pi^{1,1}_k]^T C^\dagger_{q^\star} y_{k+1} + N_k, \\
M_{k+1} &= [A^T_{q^\star}\Pi^{1,2}_{k+1} - \Pi^{1,2}_k]^T C^\dagger_{q^\star} y_{k+1} + M_k
\end{aligned}$$

initialized by $N_h = 0$, $M_h = 0$. Then, for $k \ge h+1$,

$$\begin{aligned}
\hat{x}_{k|k} &= [\Gamma^{1,1}_k + \Gamma^{1,2}_k P_{h+1|h}]^{-T} \cdot \\
&\quad \cdot [\hat{x}_{h+1|h} + N_k + P_{h+1|h}M_k], \\
P_{k|k} &= [\Gamma^{1,1}_k + \Gamma^{1,2}_k P_{h+1|h}]^{-T} \cdot [\Gamma^{2,1}_k + \Gamma^{2,2}_k P_{h+1|h}]^T \\
&\quad - A^{-1}_k Q_k A^{-T}_k, \\
\hat{x}_{k+1|k} &= [\Pi^{1,1}_k + \Pi^{1,2}_k P_{h+1|h}]^{-T} \cdot \\
&\quad \cdot [\hat{x}_{h+1|h} + N_k + P_{h+1|h}M_k], \\
P_{k+1|k} &= [\Pi^{1,1}_k + \Pi^{1,2}_k P_{h+1|h}]^{-T} \cdot [\Pi^{2,1}_k + \Pi^{2,2}_k P_{h+1|h}]^T;
\end{aligned}$$

superscript $^{(i,j)}$ indicates the (i,j)-th matrix block. □

Remark 1. Observe that Π_k, N_k, M_k and Γ_k do not depend on the specific value of $t^\star$. Hence, $\hat{x}_{k|k}$, $P_{k|k}$ and $\hat{x}_{k+1|k}$, $P_{k+1|k}$ depend on $t^\star$ only through $\hat{x}_{h+1|h}(t^\star, q^\star)$ and $P_{h+1|h}(t^\star, q^\star)$. Notice that the time update step of the above proposition also holds for $k = h$.

Remark 2. In practice, $\Pi_k(h, q^\star)$ and $\Gamma_k(h, q^\star)$ can be computed *offline* for every possible $q^\star$. Notice that $\Pi_k(h, q^\star) = \Pi_{k-h}(0, q^\star)$ and $\Gamma_k(h, q^\star) = \Gamma_{k-h}(0, q^\star)$, hence it suffices to carry out their computation for $h = 0$.

Propositions 3÷5 provide an algorithm for the recursive update of the conditional estimates (10)÷(13) for every possible $q^\star$ and every $t^\star$ *in the interval* (t_h, t_{h+1}). Extension to arbitrary values of $t^\star$ is immediate: at any time k, the (piecewise) expressions of (10)÷(13) are found by considering their restrictions (w.r.t. $t^\star$) to each of the $k+2$ intervals

$$(t_0, t_1), \ldots, (t_h, t_{h+1}), \ldots, (t_k, t_{k+1}), (t_{k+1}, +\infty).$$

Finally, the expressions of (22), (23) follow from Equation (9). The procedure requires the offline computation of $\Pi_k(0, q^\star)$, $\Gamma_k(0, q^\star)$ and the online computation of $N_k(h, q^\star)$, $M_k(h, q^\star)$ for each $h \le k$ (restrictions to (t_h, t_{h+1}), Propositions 4 and 5), plus the iteration of a standard Kalman recursion up to step k (restriction to $(t_{k+1}, +\infty)$, Proposition 3). In sums, we get the following recursive algorithm:

Offline: compute $\Pi_k(0, q^\star)$, $\Gamma_k(0, q^\star)$, $P^0_{k|k}$, $P^0_{k|k-1}$, $k \ge 0$;

[2] It is in fact $\Pi_k = \Pi_k(h, q^\star)$, $N_k = N_k(h, q^\star)$, $M_k = M_k(h, q^\star)$, $\Gamma_k = \Gamma_k(h, q^\star)$. For notational conciseness, their dependence on h and $q^\star$ and the dependence of $\hat{x}_{\cdot|k}$, $P_{\cdot|k}$ on $t^\star$ and $q^\star$ is not reported here.

Initialization: set $\hat{x}_{0|-1} = \mu_0$;
Iteration ($k \geq 0$): as measurement y_k arrives,

1) compute $\hat{x}^0_{k|k}$ from $\hat{x}^0_{k|k-1}$, $P^0_{k|k-1}$;
2) for $h = 0, \ldots, k-1$, compute $N_k(h, q^\star)$, $M_k(h, q^\star)$ from $N_{k-1}(h, q^\star)$, $M_{k-1}(h, q^\star)$; set $N_k(k, q^\star) = 0$, $M_k(k, q^\star) = 0$; compute $\hat{x}^0_{k+1|k}$ from $\hat{x}^0_{k|k}$.

Of course, the initialization step gives the parameters that are needed to represent $f(x_0|\cdot, y^{-1})$, whereas points 1 and 2 of the iteration step yield the parameters to represent $f(x_k|\cdot, y^k)$ and $f(x_{k+1}|\cdot, y^k)$, respectively. *With this scheme, a complete, explicit representation of all* (22) *and* (23) *up to index k is computed with complexity $\mathcal{O}(k^2)$.*

A. Application to switching time estimation

Based on expressions (19) and (20) of section V we get the following result.

Proposition 6: The *a posteriori* density $f(t^\star, q^\star|y^k)$ can be computed as follows:

$$f(t^\star, q^\star|y^k) = \frac{\prod_{j=0}^k f(y_j|t^\star, q^\star, y^{j-1}) f(t^\star, q^\star)}{(\ldots)},$$

where

$$\begin{aligned}(\ldots) = \sum_{q^\star=1}^{N-1} \Bigg\{ & \prod_{j=0}^k f(y_j|t^\star > t_k, q^\star, y^{j-1}) \mathbb{P}[t^\star > t_k] + \\ & + \sum_{h=0}^{k-1} \Bigg[\prod_{j=0}^h f(y_j|t^\star > t_h, q^\star, y^{j-1}) \cdot \\ & \cdot \int_{t_h}^{t_{h+1}} \prod_{j=h+1}^k f(y_j|t^\star, q^\star, y^{j-1}) f(t^\star, q^\star) \mathrm{d}t^\star \Bigg] \Bigg\}.\end{aligned}$$

In fact, all factors of the type $f(y_j|t^\star, q^\star, y^{j-1})$ may trivially be deduced from the corresponding densities $f(x_j|t^\star, q^\star, y^{j-1})$. By considering their restrictions to the relevant interval of integration, one may then apply the algorithm presented above and suitable numerical quadrature so to obtain an efficient evaluation of all integrals, i.e. of the normalization factor. Similarly, this representation of $f(t^\star, q^\star|y^k)$ is extremely well suited for a piecewise computation of integral (16) as well as of the probability function (3) and of the conditional expectation $\mathbb{E}[t^\star|y^k]$, the latter being of course the best estimate of $t^\star$ given y^k.

VII. Numerical Example

In this section we will show numerical results concerning a specific example of Stochastic Hybrid System. Due to limitations in space, we will only pursue a qualitative analysis, restricting our attention on the computation of $f(t^\star, q^\star|y^k)$ with the method outlined in Section VI-A ($\mathbb{E}[t^\star|y^k]$ and $p_{k|k}(\cdot)$ easily follow).

Let $\mathcal{Q} = \{0, 1, 2\}$. Consider system (1) with parameters $T = 0.5$, $\mu_0 = 0$, $\Sigma_0 = 0.1 \cdot I_{2\times 2}$. We chose all 4-tuples (F_q, G_q, H_q, K_q) to be

$$\left(\begin{bmatrix} -0.4 & 0.6 \\ c_q & -0.5 \end{bmatrix}, \begin{bmatrix} 0.02 \\ 0.02 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0.02 & 0 \\ 0 & 0.02 \end{bmatrix} \right)$$

where $c_0 = 0$, $c_1 = -2$, $c_2 = 0.5$. That is, only the state evolution matrix changes with q. This modifies the character of the system from stable to oscillatory or unstable, with the spectrum of F_q (for $q = 0, 1, 2$) given by $\sigma(F_0) = \{-0.4, -0.5\}$, $\sigma(F_1) = \{-0.45 \pm i1.0943\}$, $\sigma(F_2) = \{-1, 0.1\}$. The Markov chain underlying the evolution of $q(t)$ is set to start from $q(0) = 0$, i.e. $p_0 = 1$; it is also assumed to privilege jumps towards $q = 2$: $\lambda_1 = 0.06$, while $\lambda_2 = 0.08$. With this choice, $\mathbb{P}[q^\star = 1] \simeq 0.43$, $\mathbb{P}[q^\star = 2] \simeq 0.57$. Moreover, $\mathbb{E}[t^\star] \simeq 7.14$.

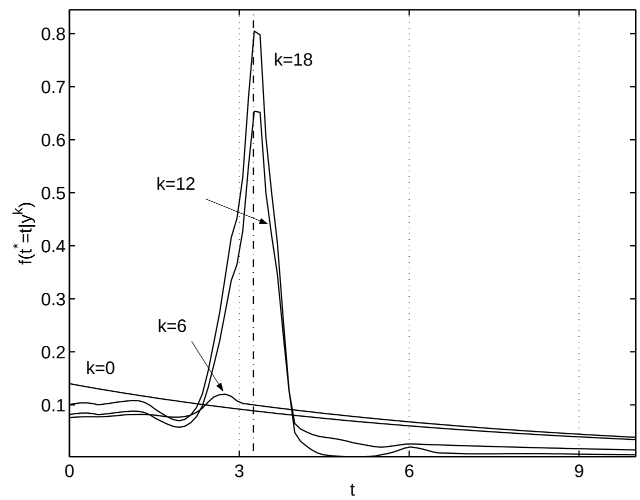

Fig. 2. Density function $f(t^\star|y^k)$ plotted for $k = 0, 6, 12$ and 18. The dash-dotted vertical line marks the *actual* switching time.

In the simulations, we started off the system from $x(0) = 0$. We then randomly generated x and y up to time $k_{\max} \cdot T$, with $k_{\max} = 20$, for a jump of $q(t)$ from 0 to $q^\star = 1$ occurring at $t^\star = 3.5$, i.e. considerably before the expected time. The values of $t^\star$ and $q^\star$ have been chosen manually by the programmer, i.e. they have not been simulated as random variables. This does not affect the application of the proposed method. Note the exiguity of measurements, in accordance with the motivations for model (1). The algorithm of Section VI is then applied to the data y_k, $k = 0, \ldots, k_{\max}$, along with Simpson's adaptive method for the numerical integrations.

Figure 2 shows for different values of k the *a posteriori* density of $t^\star$ given y^k computed by suitable application of Proposition 6. Its evolution from the exponential prior to a density roughly concentrated around the true switching instant may be observed. Also notice the exponential tail of $f(t^\star|y^k)$ for $t^\star > t_k$. The evolution of the conditional expectation of $t^\star$ and of the conditional probability function of $q(t_k)$ given y^k are reported in Figure 3. After the switch occurs and new measurements are collected, the $t^\star$ estimates become significant and tend to correctly detect the switching event. Of course, the quality of the estimates strongly relates to the nature of the systems (F_q, G_q, H_q, K_q), $q = 0, 1, 2$. In the present case, the effectiveness of the estimates is certainly favored by the change in the stability of the system determined by the switch.

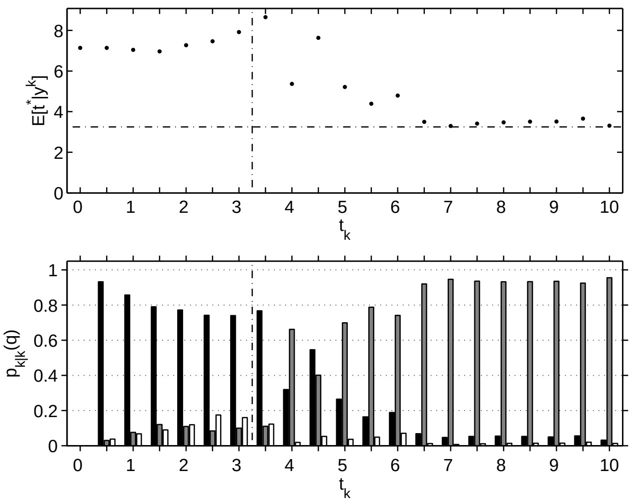

Fig. 3. Evolution of the expectation $\mathbb{E}[t^\star|y^k]$ (above) and of the probability function $p_{k|k}(q)$ (below; left bar: $p_{k|k}(0)$, center bar: $p_{k|k}(1)$; right bar: $p_{k|k}(2)$). The dash-dotted lines mark the *actual* switching time.

VIII. Conclusions

In this paper we have presented an on-line method for estimating the joint state (x, q) of a class of Stochastic Hybrid Systems, characterized by a state-space model where the continuous state evolves according to a linear SDE, the discrete state is a Markov process, while noisy measurements of the continuous state are discrete and periodic in time. Such class is suitable for applications where discrete state jumps may occur at a rate that is comparable to the frequency at which measurements are taken. We focused on the problem of *fault detection*.

For a *given* trajectory of the discrete state $q(t)$ the problem is solvable by applying ordinary Kalman filtering. In order to solve our problem, however, we must average these Kalman filters against the *a posteriori* joint distribution of $(t^\star, q^\star)$. This averaging operation eliminates the Gaussian nature of the estimate, which cannot therefore be described in a parametric way. However, we managed to formulate an on-line algorithm that is exact *up to* the numerical averaging operation. Note, e.g., that any approximation (due to the numerical computation of integral (8)) that is introduced for the calculation of $f(x_k|y^k)$ does *not* influence the degree of approximation of $f(x_\ell|y^\ell)$ for $\ell > k$, since the latter density is not computed directly from the former. We also provided a numerical example that shows the effectiveness of our method.

A future direction for our research will be to extend our algorithm to models where the Markov chain, instead of having all (but one) absorbing states, is allowed multiple switches between two consecutive measurements. Also, studying the convergence of densities such as $f(t^\star|y^k)$ as $k \to \infty$ is of utmost interest; such rate would certainly depend on the *relative* dynamics (stability, modes of convergence or divergence, etc.) of the continuous-time systems that correspond to different values of the discrete state.

References

[1] R. Bhatia. *Matrix Analysis*. Springer-Verlag New York Inc., New York, USA, 1997.

[2] H. A. P. Blom and Y. Bar-Shalom. The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Transactions on Automatic Control*, 33(8):780–783, Aug. 1988.

[3] R. Chen and J. S. Liu. Mixture Kalman Filters. *Journal of the Royal Statistical Society – Series B*, 62:493–508, 2000.

[4] E. Cinquemani, M. Micheli, and G. Picci. State estimation for a class of stochastic hybrid systems. Paper in preparation.

[5] O. L. V. Costa. Linear minimum mean square error estimation for discrete-time Markovian jump linear systems. *IEEE Transactions on Automatic Control*, 39(8):1685–1689, Aug. 1994.

[6] O. L. V. Costa and S. Guerra. Stationary filter for linear minimum mean square error estimator of discrete-time Markovian jump systems. *IEEE Transactions on Automatic Control*, 47(8):1351–1356, Aug. 2003.

[7] A. Doucet and C. Andrieu. Iterative algorithms for state estimation of Jump Markov linear systems. *IEEE Transactions on Signal Processing*, 49(6):1216–1227, June 2001.

[8] A. Doucet, A. Logothetis, and V. Krishnamurthy. Stochastic sampling algorithms for state estimation of Jump Markov linear systems. *IEEE Transactions on Automatic Control*, 45(2):188–201, 2000.

[9] R. J. Elliott, F. Dufour, and D. D. Sworder. Exact hybrid filters in discrete time. *IEEE Transactions on Automatic Control*, 41(12):1807–1810, Dec. 1996.

[10] A. Germani, C. Manes, and P. Palumbo. Polynomial filtering for stochastic systems with Markovian switching coefficients. In *Proceedings of the 42nd IEEE Conference on Decision and Control (CDC 2003)*, pages 1392–1397, Maui, Hawaii, Dec. 2003.

[11] J. L. Hibey and C. D. Charalambous. Conditional densities for continuous-time nonlinear hybrid systems with application to fault detection. *IEEE Transactions on Automatic Control*, 44(11):2164–2169, Nov. 1999.

[12] M. W. Hofbaur and B. C. Williams. Mode estimation of probabilistic hybrid systems. In *Hybrid Systems: Computation and Control (HSCC 2002)*, Lecture Notes on Computer Sciences. Springer Verlag, 2002.

[13] J. Hu, J. Lygeros, and S. S. Sastry. Towards a Theory of Stochastic Hybrid Systems. In *Third International Workshop on Hybrid Systems: Computation and Control*, Pittsburgh, PA, 2000. Springer Verlag Lecture Notes on Computer Science, vol. 1790.

[14] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, London, 1970.

[15] M. Kijima. *Markov Processes for Stochastic Modeling*. CRC Press, London, UK, 1997.

[16] X. Koutsoukos, J. Kurien, and F. Zhao. Monitoring and diagnosis of hybrid systems using particle filtering methods. In *Proceedings of the Fifteenth International Symposium on the Mathematical Theory of Networks and Systems (MTNS '02)*, University of Notre Dame, South Bend, Indiana, Aug. 2002.

[17] X. Koutsoukos, J. Kurien, and F. Zhao. Estimation of distributed hybrid systems using particle filtering methods. In *Hybrid Systems: Computation and Control (HSCC 2003)*. Springer Verlag Lecture Notes on Computer Science, vol. 2623, Pittsburgh, PA, 2003.

[18] U. Lerner, R. Parr, D. Koller, and G. Biswas. Bayesian fault detection and diagnosis in dynamic systems. In *Proceedings of the 17th National Conference on Artificial Intelligence (AAAI)*, pages 531–537, Austin, Texas, July 2000.

[19] A. Logothetis and V. Krishnamurthy. Expectation maximization algorithms for map estimation of jump Markov linear systems. *IEEE Transactions on Signal Processing*, 47(8):2139–2156, Aug. 1999.

[20] B. M. Miller and W. J. Runggaldier. Kalman filtering for linear systems with coefficients driven by a hiddem Markov jump process. *Systems and Control Letters*, 31:93–102, 1997.

[21] K. P. Murphy. Switching Kalman filters. Report 98-10, Compaq Cambridge Research Laboratory, 1998.

[22] I. V. Nikiforov. Optimal sequential detection and isolation of changes in stochastic systems. Technical Report 2063, Institut National de Recherche en Informatique et en Automatique (INRIA), Sept. 1993.

[23] J. K. Tugnait. Adaptive estimation and identification for discrete systems with Markov jump parameters. *IEEE Transactions on Automatic Control*, 27(5):1054–1065, Oct. 1982.

[24] J. K. Tugnait. Detection and estimation for abruptly changing systems. *Automatica*, 18(5):607–615, 1982.

[25] Q. Zhang. Hybrid filtering for linear systems with non-Gaussian disturbances. *IEEE Transactions on Automatic Control*, 45(1):50–61, Jan. 2000.