

Hybrid Filters Modulated by Markov Chains with Two-time Scales

G. Yin¹ and S. Dey²

Abstract

We consider a class of hybrid filtering problems in discrete-time. The main feature is that the system is modulated by a Markov chain. Our main effort is to reduce the complexity of the underlying problems. Consider the case that the Markov chain has a large state space. Then the solution of the problem relies on solving a large number of filtering equations. By using the hierarchical structure of the system, we show that a reduced system of filtering equations can be obtained by aggregating the states of each recurrent class into one state. Extensions to inclusion of transient states and nonstationary cases are also treated.

KEY WORDS. Markov chain, filtering, near complete decomposability, weak convergence.

1 Introduction

We consider hybrid filtering problems in discrete time. By hybrid filters, we mean such filters modulated by a discrete-time Markov chain. Many problems in target tracking, speech recognition, telecommunication, and manufacturing require solutions of filtering problems involving a hidden Markov chain. Taking this into consideration, we assume that the system under consideration is influenced by a hidden Markov chain with finite state space in addition to the usual random system disturbances and observation noise. Since we often have to face complex and to be large-scale systems, although the state space of the Markov chain is finite, it inevitably contains a large number of states. In this paper, our focus is on reduction of complexity of such filtering problems involving large-scale hidden Markov chains.

Recently, linear systems with coefficients driven by a hidden Markov chain was considered in [20]. Discrete-time systems were studied in [1, 6, 11, 28] among oth-

ers. In [34], Zhang studied hybrid filters in continuous time and treated problems involving non-Gaussian noise. Our study is motivated by these recent developments and stems from the needs of treating hybrid systems.

In [25], Simon and Ando pointed out that various large-scale systems have hierarchical structures with different rates of various. These states are also naturally decomposable into different layers or hierarchy. The inherent hierarchy allows one to take advantage and to organize and reorganize the systems accordingly. Based on the decomposition and aggregation, Courtois dealt with the so-called nearly completely decomposable Markov chain models [7]. Recently, Dey derived reduced-complexity filtering results for hidden Markov models, in which the underlying Markov chains are nearly completely decomposable [9]. Such Markov structures have numerous applications in queueing and computer systems [7], multiple time-scale heterogeneous traffic modelling (e.g., variable bit rate video traffic [26]), manufacturing systems, operations research and many other biological and physical systems where a multiple time-scale or hierarchical behaviour is involved. Taking the approaches of [7] and [9] as our point of departure, to reduce the complexity of the underlying problem, we introduce a small parameter $\varepsilon > 0$ into the system. Note that the small parameter is used to reflect the high contrast of the transition rates of the Markov chain. For the subsequent asymptotic analysis, to obtain desired results, it is necessary to send $\varepsilon \rightarrow 0$, which can serve as a guideline for various applications and for approximation and heuristics. In real applications, however, ε might be a fixed constant and only the relative order of magnitude of this parameter matters. In our setup, we also consider a nearly completely decomposable Markovian model, in which the hidden Markov chain has a large state space. The transition probability matrix is a sum of a completely decomposable transition matrix and a generator of a continuous-time Markov chain. Following our systematic studies on singularly perturbed Markov chains in both continuous time and discrete time [14, 30, 31, 32, 35], we investigate the asymptotic properties of the filtering problem by means of weak convergence methods. We will demonstrate that a limit filtering problem can be derived in which the underlying Markov chain is replaced

¹Department of Mathematics, Wayne State University, Detroit, MI 48202. Email: gyin@math.wayne.edu. The research of this author was supported in part by the National Science Foundation under grant DMS-9877090.

²Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Victoria 3010, Australia. Email: sdey@ee.mu.oz.au.

by an averaged chain and the system coefficients are averaged out with respect to the stationary measures of each ergodic class. The reduction of complexity is more transparent when the transition matrix of the Markov chain consists of only one ergodic class. In this case, the limit filtering problem becomes a standard Kalman filter free of Markovian jump process.

The rest of the paper is arranged as follows. Section 2 presents the formulation of the problem and a number of preliminary results that are to be used in our study. Section 3 presents the main theorems. Section 4 gives remarks and a few extensions. The verbatim proofs and numerical results can be found in [29].

A word of the notation is in order. In what follows, we use K to denote a generic positive constant, which may take different values for different usage. For $z \in \mathbb{R}^{\ell_1 \times \ell_2}$ with some positive integers ℓ_1 and ℓ_2 , z' denotes its transpose. For a suitable function f , f_x and f_{xx} denote its first-order and second-order partial derivatives with respect to x .

2 Problem Formulation

Let $\varepsilon > 0$ be a small parameter and $\{\alpha_n^\varepsilon\}$ be a (time) homogeneous singularly perturbed Markov chain in discrete time with a finite state space \mathcal{M} having m elements and a transition matrix

$$P^\varepsilon = \tilde{P} + \varepsilon Q, \quad (1)$$

where \tilde{P} is an $m \times m$ transition matrix and $Q = (q_{i\ell})$ is a generator of a continuous-time homogeneous Markov chain, i.e., $q_{i\ell} \geq 0$ for $i \neq \ell$ and $\sum_{\ell} q_{i\ell} = 0$ for each i .

Suppose that for some $T > 0$ and $0 \leq n \leq \lfloor T/\varepsilon \rfloor$ (where $\lfloor z \rfloor$ denotes the largest integer part of z), $x_n^\varepsilon \in \mathbb{R}^r$ is the state to be estimated, y_n^ε is the corresponding observation, $A(i)$, $C(i)$, $\sigma_w(i)$, and $\sigma_v(i)$ are well defined for each $i \in \mathcal{M}$ (i.e., they are finite for each $i \in \mathcal{M}$). With initial data x_0 and y_0 , the hybrid filtering problem is concerned with the following linear system of equations:

$$\begin{aligned} x_{n+1}^\varepsilon &= x_n^\varepsilon + \varepsilon A(\alpha_n^\varepsilon)x_n^\varepsilon + \sqrt{\varepsilon}\sigma_w(\alpha_n^\varepsilon)w_n, \\ y_{n+1}^\varepsilon &= y_n^\varepsilon + \varepsilon C(\alpha_n^\varepsilon)x_n^\varepsilon + \sqrt{\varepsilon}\sigma_v(\alpha_n^\varepsilon)v_n, \end{aligned} \quad (2)$$

where $\{w_n\}$ and $\{v_n\}$ are the system disturbance and the observation noise, respectively. For ease of presentation, in what follows, we will suppress the floor-function notation $\lfloor \cdot \rfloor$ and write it as $0 \leq n \leq T/\varepsilon$ throughout. The use of the $\sqrt{\varepsilon}$ in the noise terms stems from the central limit scaling. Precise conditions on the noises will be provided later. In what follows, we will show that as $\varepsilon \rightarrow 0$, the above filtering problem has a limit. The limit filtering problem is

still modulated by a Markov chain. However, the total number of states of the limit Markov chain is equal to the number of recurrent groups or clusters l . As mentioned before, typically $l \ll m$ and by considering this limit filtering problem, substantial computational savings can be obtained. Although (2) is a discrete-time filtering problem, the limit under appropriate scaling is a continuous-time hybrid filtering problem. In the rest of the paper, our main effort is devoted to deriving the limit filtering problem. For solution of continuous-time hybrid filtering problems involving jump Markov processes, see [5, 8, 10, 11, 20]; see also [27] and the references therein for discrete time results.

In view of (1), the transition probabilities of α_n^ε are dominated by \tilde{P} . The structure of \tilde{P} is thus important. Since α_n^ε is a finite-state Markov chain, the Markov chain corresponding to the transition matrix \tilde{P} either consists of all recurrent states or it includes transient states in addition to recurrent states (see [13]). We first consider the case of inclusion of recurrent states only. Later we will discuss a generalization to the case where transient states are also included. Suppose that the matrix \tilde{P} is given by

$$\tilde{P} = \text{diag}(\tilde{P}^1, \dots, \tilde{P}^l) = \begin{pmatrix} \tilde{P}^1 & & \\ & \ddots & \\ & & \tilde{P}^l \end{pmatrix}, \quad (3)$$

where each $\tilde{P}^i \in \mathbb{R}^{m_i \times m_i}$ is itself a transition matrix and $\sum_{i=1}^l m_i = m$. Here and henceforth, by $\text{diag}(Z^1, \dots, Z^l)$, we mean a diagonal block matrix with matrix entries Z^1 through Z^l of appropriate dimensions. It is clear that for sufficiently small $\varepsilon > 0$, P^ε is close to \tilde{P} , so P^ε is a nearly completely decomposable transition matrix (see [7]). Note that typically for large scale Markovian systems, $l \ll m$ and therein lies the motivation for reducing computational complexity. Concerning the Markov chain, we assume the following condition.

- (A1) The transition probability matrix of the Markov chain α_n^ε is given by (1) with \tilde{P} specified in (3), and the state space of the Markov chain is

$$\begin{aligned} \mathcal{M} &= \mathcal{M}_1 \cup \mathcal{M}_2 \cap \dots \cup \mathcal{M}_l \\ &= \{s_{11}, \dots, s_{1m_1}\} \cup \dots \cup \{s_{l1}, \dots, s_{lm_l}\}. \end{aligned} \quad (4)$$

For each $i = 1, \dots, l$, $\mathcal{M}_i = \{s_{i1}, \dots, s_{im_i}\}$ is the state space corresponding to the transition matrix \tilde{P}^i and \tilde{P}^i is irreducible and aperiodic.

- (A2) $E|x_0|^2 < \infty$ and $E|y_0|^2 < \infty$. For each $i \in \mathcal{M}$, $A(i)$, $C(i)$, $\sigma_w(i)$, and $\sigma_v(i)$ are finite; $\sigma_w(i)\sigma_w'(i)$ and $\sigma_v(i)\sigma_v'(i)$ are positive definite matrices.
- (A3) The sequences $\{w_n\}$ and $\{v_n\}$ are independent of $\{\alpha_n^\varepsilon\}$ and are independent of each other. The

$\{w_n\}$ and $\{v_n\}$ are stationary martingale difference sequences (with zero mean) such that

$$\begin{aligned} Ew_n w_n' &= I, \quad E v_n v_n' = I, \\ E|w_n|^{2+\Delta} &< \infty, \quad \text{and} \\ E|v_n|^{2+\Delta} &< \infty \quad \text{for some } \Delta > 0. \end{aligned}$$

3 Main Results

3.1 Preliminary Results

Note that the probability vector

$$p_n^\varepsilon = (P(\alpha_n^\varepsilon = s_{ij})) \in \mathbb{R}^{1 \times m}$$

satisfies

$$p_{n+1}^\varepsilon = p_n^\varepsilon P^\varepsilon, \quad p_0^\varepsilon = p_0 \quad (5)$$

such that p_0 is the initial probability distribution. By (A1), the result in [31] yields the following lemma.

Lemma 3.1. *Suppose that condition (A1) hold. Then the following assertions hold:*

- (1) *Let ν^i the stationary distribution corresponding to the transition matrix \bar{P}_i for each $i = 1, \dots, l$. Then for some $0 < \lambda < 1$,*

$$p_n^\varepsilon = \theta(t) \text{diag}(\nu^1, \dots, \nu^l) + O(\varepsilon + \lambda^n), \quad (6)$$

where $\theta(t) = (\theta_1(t), \dots, \theta_l(t)) \in \mathbb{R}^{1 \times l}$ (with $t = \varepsilon n$) satisfies

$$\frac{d\theta(t)}{dt} = \theta(t) \bar{Q}, \quad \theta_i(0) = x_0^i \mathbb{1}_{m_i},$$

with

$$\begin{aligned} \bar{Q} &= \text{diag}(\nu^1, \dots, \nu^l) Q \tilde{\mathbb{I}}, \\ \tilde{\mathbb{I}} &= \text{diag}(\mathbb{1}_{m_1}, \dots, \mathbb{1}_{m_l}), \end{aligned} \quad (7)$$

where $\mathbb{1}_l$ denotes an l -dimensional column vector with all entries being 1.

- (2) *With $n \leq T/\varepsilon$, the n -step transition probability matrix $(P^\varepsilon)^n$ satisfies*

$$(P^\varepsilon)^n = \Phi(t) + O(\varepsilon + \lambda^n), \quad (8)$$

where

$$\begin{aligned} \Phi(t) &= \tilde{\mathbb{I}} \Theta(t) \text{diag}(\nu^1, \dots, \nu^l) \\ \frac{d\Theta(t)}{dt} &= \Theta(t) \bar{Q}, \quad \Theta(0) = I. \end{aligned} \quad (9)$$

- (3) *As $\varepsilon \rightarrow 0$, $\bar{\alpha}^\varepsilon(\cdot)$ converges weakly to $\bar{\alpha}(\cdot)$, which is a continuous-time Markov chain with state space*

$\bar{\mathcal{M}} = \{1, \dots, l\}$ and generator \bar{Q} given by (7). Moreover, for the occupation measures defined by

$$\begin{aligned} o_{n,ij}^\varepsilon &= \varepsilon \sum_{k=0}^n [I_{\{\alpha_k^\varepsilon = s_{ij}\}} - \nu_j^i I_{\{\alpha_k^\varepsilon \in \mathcal{M}_i\}}], \\ &\text{for } i = 1, \dots, l, \quad j = 1, \dots, m_i, \end{aligned}$$

the following mean square estimates hold

$$\sup_{0 \leq n \leq T/\varepsilon} E|o_{n,ij}^\varepsilon|^2 = O(\varepsilon). \quad (10)$$

For $0 \leq n \leq T/\varepsilon$, define the interpolations $x^\varepsilon(\cdot)$ and $y^\varepsilon(\cdot)$ as

$$x^\varepsilon(t) = x_n^\varepsilon, \quad y^\varepsilon(t) = y_n^\varepsilon, \quad t \in [n\varepsilon, (n+1)\varepsilon]. \quad (11)$$

where x_n^ε and y_n^ε are given in (2). Then $x^\varepsilon(\cdot)$ and $y^\varepsilon(\cdot) \in D^r[0, T]$, which is the space of \mathbb{R}^r -valued functions that are right continuous, have left limits, endowed with the Skorohod topology [12, p. 122]. Using weak convergence methods, we will show that the interpolated processes converge weakly to $x(\cdot)$ and $y(\cdot)$, which satisfy continuous-time hybrid Kalman filtering equations. Following the approach of weak convergence methods [12, 15], we first show that the sequences of interests are tight, and then we characterize the limit processes by using martingale averaging techniques.

Owing to the assumption on the system and observation noise and $\sqrt{\varepsilon}$ scaling, the following lemma, known as the functional central limit theorem or Donsker's invariance theorem, holds. Its proof is standard; see for example, [12, Theorem 3.1, p. 351].

Lemma 3.2. *Set*

$$w^\varepsilon(t) = \sqrt{\varepsilon} \sum_{j=0}^{t/\varepsilon-1} w_j \quad \text{and} \quad v^\varepsilon(t) = \sqrt{\varepsilon} \sum_{j=0}^{t/\varepsilon-1} v_j. \quad (12)$$

Under (A3), $w^\varepsilon(\cdot)$ and $v^\varepsilon(\cdot)$ converge weakly to standard r -dimensional Brownian motions $w(\cdot)$ and $v(\cdot)$, respectively.

In the analysis to follow, we need the *a priori* bounds on $\{x_n^\varepsilon\}$ and $\{y_n^\varepsilon\}$, which are presented in the form of the following Lemma. The proof is provided in the appendix.

Lemma 3.3. *Suppose that conditions (A1)–(A3) hold. For $\{x_n^\varepsilon\}$ and $\{y_n^\varepsilon\}$ defined in (2), we have the following bounds:*

$$\sup_{0 \leq n \leq T/\varepsilon} E|x_n^\varepsilon|^2 < \infty, \quad \text{and} \quad \sup_{0 \leq n \leq T/\varepsilon} E|y_n^\varepsilon|^2 < \infty. \quad (13)$$

3.2 Weak Convergence

To proceed, let \mathcal{F}_n be the σ -algebra generated by $\{\alpha_j^\varepsilon, w_j, v_j : j \leq n\}$ and E_n be the conditional expectation w.r.t. \mathcal{F}_n ; let $\mathcal{F}_t^\varepsilon$ be the σ -algebra generated by $\{\alpha^\varepsilon(s), w^\varepsilon(s), v^\varepsilon(s) : s \leq t\}$ and E_t^ε be the conditional expectation w.r.t. $\mathcal{F}_t^\varepsilon$. We are to derive the tightness of $\{x^\varepsilon(\cdot)\}$ and $\{y^\varepsilon(\cdot)\}$. This is a compactness result, which is established by verifying a tightness criterion; the proof is in the appendix.

Theorem 3.4. *Suppose that (A1)–(A3) hold. Then $\{x^\varepsilon(\cdot)\}$ and $\{y^\varepsilon(\cdot)\}$ are tight in $D^r[0, T]$, where $D^r[0, T]$ is the space of \mathbb{R}^r -valued functions that are right continuous and have left limits, endowed with the Skorohod topology.*

Theorem 3.5. *Assume the conditions of Theorem 3.4. Then $x^\varepsilon(\cdot)$ and $y^\varepsilon(\cdot)$ converge weakly to $x(\cdot)$ and $y(\cdot)$, respectively, such that $x(\cdot)$ and $y(\cdot)$ are solutions of the filtering equations*

$$\begin{aligned} dx &= \bar{A}(\bar{\alpha}(t))xdt + \bar{\sigma}_w(\bar{\alpha}(t))dw, \\ dy &= \bar{C}(\bar{\alpha}(t))xdt + \bar{\sigma}_v(\bar{\alpha}(t))dv, \end{aligned} \quad (14)$$

where $w(\cdot)$ and $v(\cdot)$ are the independent r -dimensional standard Brownian motions given by Lemma 3.2,

$$\bar{A}(i) = \sum_{j=1}^{m_i} \nu_j^i A(s_{ij}), \quad \bar{B}(i) = \sum_{j=1}^{m_i} \nu_j^i B(s_{ij}), \quad \text{for each } i \in \bar{\mathcal{M}}, \quad (15)$$

and for each $i \in \bar{\mathcal{M}}$, $\bar{\sigma}_w(i)$ and $\bar{\sigma}_v(i)$ satisfy

$$\begin{aligned} \bar{\sigma}_w(i)\bar{\sigma}_w'(i) &= \sum_{j=1}^{m_i} \nu_j^i \sigma_w(s_{ij})\sigma_w'(s_{ij}), \\ \bar{\sigma}_v(i)\bar{\sigma}_v'(i) &= \sum_{j=1}^{m_i} \nu_j^i \sigma_v(s_{ij})\sigma_v'(s_{ij}). \end{aligned} \quad (16)$$

The reduction of complexity is particularly pronounced if the transition matrix (3) consists of only one ergodic class (i.e., P in (3) consists of only one block). That is, $P^\varepsilon = P + \varepsilon Q$ such that P is irreducible and aperiodic. It is easily seen that for sufficiently small $\varepsilon > 0$, P^ε is also irreducible. Consider the filtering problem (2). Similar to the previous case define $x^\varepsilon(\cdot)$ and $y^\varepsilon(\cdot)$ as the piecewise constant interpolations of x_k^ε and y_k^ε , respectively. Replace $\bar{A}(\cdot)$ and $\bar{\sigma}_w(\cdot)$ by

$$\bar{A}^0 = \sum_{j=1}^m A(j)\nu_j, \quad \text{and} \quad \bar{\sigma}_w^0(\bar{\sigma}_w^0)' = \sum_{j=1}^m \nu_j \sigma_w(j)\sigma_w'(j), \quad (17)$$

with $\nu = (\nu_1, \dots, \nu_m)$ denoting the stationary distribution of P . Similarly replace $\bar{C}(\cdot)$ and $\bar{\sigma}_v(\cdot)$ by \bar{C}^0 and $\bar{\sigma}_v^0$, respectively. The weak convergence of $\{x^\varepsilon(\cdot)\}, \{y^\varepsilon(\cdot)\}$

will still be obtained. The proofs are similar to the previous case. In fact, it is readily seen that Lemma 3.3 and Theorem 3.4 continue to hold. Lemma 3.1 still holds with obvious modifications and (10) (in Lemma 3.1) is changed to

$$\sup_{0 \leq n \leq T/\varepsilon} E \left[\varepsilon \sum_{k=0}^n [I_{\{\alpha_k^\varepsilon=j\}} - \nu_j] \right]^2 = O(\varepsilon).$$

Using this mean square estimate and similar arguments as before, we can show that Theorem 3.5 continues to hold. It is interesting to note that the limit filtering problem becomes a standard Kalman filter, in which the jump process effect has been completely averaged out. We state this as the following result.

Corollary 3.6. *Let us consider problem (2) such that P is irreducible and aperiodic. Then $(x^\varepsilon(\cdot), y^\varepsilon(\cdot))$ converges weakly to $(x(\cdot), y(\cdot))$ that is the solution of the filtering problem*

$$\begin{aligned} dx(t) &= \bar{A}^0 x(t)dt + \bar{\sigma}_w^0 dw(t), \\ dy(t) &= \bar{C}^0 x(t)dt + \bar{\sigma}_v^0 dv(t). \end{aligned} \quad (18)$$

4 Further Remarks

Our main motivation stems from the effort of reduction of complexity. Regarding (2), note that the time horizon we are working with is $0 \leq n \leq \lfloor T/\varepsilon \rfloor$. As pointed out in [23], if we treat the discrete-time case directly, it can be reduced to an $m^{\lfloor T/\varepsilon \rfloor}$ -dimensional recursive system of equations, where m is the total number of states of the Markov chain. For us, m is a fairly large number. As a result, the amount of computation becomes practically un-trackable. One cannot complete the computation in polynomial time. By weak convergence methods, we have obtained a reduced or limit system of filtering equations. This limit system of equations allows us to find nearly optimal filtering, and the limit system has reduced complexity. In particular, if the transition matrix P given in (1) is irreducible, the limit becomes a Kalman filter (see Proposition 3.6).

For continuous-time Kalman filter problems with Markovian switching, it has been recognized (see [5, 10, 20]) that in general the problem is an infinite dimensional one just as in nonlinear filter case [19]. Nevertheless, Björk [5] proved that a finite-dimensional filter exists for a linear hybrid system if and only the observation is independent of the state variable. For the filtering problem considered in this paper, this requires the observation process in the limit problem being independent of state. Corresponding to such a requirement,

we can consider

$$\begin{aligned} x_{n+1}^\varepsilon &= x_n^\varepsilon + \varepsilon A(\alpha_n^\varepsilon) x_n^\varepsilon + \sqrt{\varepsilon} \sigma_w(\alpha_n^\varepsilon) w_n, \\ y_{n+1}^\varepsilon &= y_n^\varepsilon + \varepsilon C(\alpha_n^\varepsilon) x_n^\varepsilon + \sqrt{\varepsilon} \sigma_v v_n. \end{aligned} \quad (19)$$

Similar to the derivation of Theorem 3.5, we obtain the limit filtering equations

$$\begin{aligned} dx &= \bar{A}(\bar{\alpha}(t)) x dt + \bar{\sigma}_w(\bar{\alpha}(t)) dw, \\ dy &= \bar{C}(\bar{\alpha}(t)) x dt + \sigma_v dv. \end{aligned} \quad (20)$$

Note that the calculation of (19) leads to recursive filters of dimension $m^{lT/\varepsilon}$, whereas (20) yields a solution of finite-dimensional filtering problem.

In the previous sections, we have been mainly concerned with discrete-time filtering problems in which the Markov chain contains only recurrent states. Extensions to Markov chain including transient states can be considered. Let the transition probability be of the form (1). In lieu of (3), suppose the transition matrix \tilde{P} in (1) is given by

$$\tilde{P} = \begin{pmatrix} \tilde{P}^1 & & & \\ & \ddots & & \\ & & \tilde{P}^l & \\ \tilde{P}_*^1 & \dots & \tilde{P}_*^l & \tilde{P}_* \end{pmatrix}. \quad (21)$$

We can carry out the analysis similar to the recurrent-chain case.

Time-inhomogeneous Markov chains can be considered. In lieu of (1), assume that the transition probability matrix is nonstationary given by

$$P^\varepsilon(\varepsilon n) = \tilde{P}(\varepsilon n) + \varepsilon Q(\varepsilon n),$$

where $\tilde{P}(\varepsilon n)$ is the dominating part of the transition matrix. In this case, we can carry out the analysis as in the previous case although the details and notation are more involved.

There is a continuous-time analogue of the hybrid filtering problems. In lieu of (2), for $t \in [0, T]$, consider

$$\begin{aligned} dx^\varepsilon(t) &= A(\alpha^\varepsilon(t)) x^\varepsilon(t) dt + \sigma_w(\alpha^\varepsilon(t)) dw, \\ dy^\varepsilon(t) &= C(\alpha^\varepsilon(t)) x^\varepsilon(t) dt + \sigma_v(\alpha^\varepsilon(t)) dv, \end{aligned} \quad (22)$$

where $w(\cdot)$ and $v(\cdot)$ are independent standard Brownian motions, and where $\alpha^\varepsilon(\cdot)$ is a continuous-time singularly perturbed Markov chain with finite state space \mathcal{M} and with generator

$$Q^\varepsilon(t) = \frac{\tilde{Q}(t)}{\varepsilon} + \hat{Q}(t), \quad (23)$$

where both $\tilde{Q}(t)$ and $\hat{Q}(t)$ are generators. Weak convergence analysis can be carried out to obtain limit

systems. Suppose that $\tilde{Q}(t)$ is given by

$$\tilde{Q}(t) = \begin{pmatrix} \tilde{Q}^1(t) & & & \\ & \ddots & & \\ & & \tilde{Q}^l(t) & \\ \tilde{Q}_*^1(t) & \dots & \tilde{Q}_*^l(t) & \tilde{Q}_*(t) \end{pmatrix}. \quad (24)$$

For each $i \in \{1, \dots, l\}$, let $\tilde{Q}^i(t) = B(t) \tilde{Q}_{*,c}^i$, $\tilde{Q}_*(t) = B(t) \tilde{Q}_{*,c}$, where $B(t)$ is an $\mathbb{R}^{m_* \times m_*}$ matrix-valued function, and $\tilde{Q}_{*,c}^i \in \mathbb{R}^{m_* \times m_i}$ and $\tilde{Q}_{*,c} \in \mathbb{R}^{m_* \times m_*}$ are constant matrices. It is readily seen that $B(t)$ is invertible for each $t \in [0, T]$ and for each i ,

$$a_i(t) \stackrel{\text{def}}{=} -\tilde{Q}_*^{-1}(t) \tilde{Q}_*^i(t) \mathbb{1}_{m_i} = -\tilde{Q}_{*,c}^{-1} \tilde{Q}_{*,c}^i \mathbb{1}_{m_i} = a_i \quad (25)$$

is a time-independent vector. We can then obtain the limit filtering problem.

The possible extensions mentioned above together with their developments can be found in [29]. We omit their details here.

In this paper, it has been demonstrated that the formulation of hybrid system is natural. It is conceivable, such hybrid systems will play more and more important roles in the future.

References

- [1] Y. Bar-Shalom and X.R. Li, *Estimation and Tracking: Principles, Techniques, and Software*, Artech House Publishers, Norwood, MA, 1996.
- [2] D. P. Bertsekas, *Dynamic Programming and Stochastic Control*, Academic Press, New York, 1976.
- [3] T. R. Bielecki and L. Stettner, Ergodic control of a singularly perturbed Markov process in discrete time with general state and compact action spaces, *Appl. Math. Optim.*, **38** (1998), 261-281.
- [4] G. Blankenship, Singularly perturbed difference equations in optimal control problems, *IEEE Trans. Automat. Control* **T-AC 26** (1981), 911-917.
- [5] T. Björk, Finite dimensional optimal filters for a class of Ito processes with jumping parameters, *Stochastics*, **4** (1980), 167-183.
- [6] O.L.V. Costa, Linear minimum mean square error estimation for discrete-time Markov jump linear systems, *IEEE Trans. Automatic Control* **39** (1994), 1685-1689.
- [7] P. J. Courtois, *Decomposability: Queueing and Computer System Applications*, Academic Press, New York, 1977.
- [8] D.P. de Farias, J.C. Geromel, J.B.R. do Val and O.L.V. Costa, Output Feedback Control of Markov Jump Linear Systems in Continuous-Time, *IEEE Trans. Automat. Control*, **45** (2000), 944-949.

- [9] S. Dey, Reduced-complexity filtering for partially observed nearly completely decomposable Markov chains, *IEEE Trans. Signal Processing*, **48**, (2000), 3334-3344.
- [10] F. Dufour and P. Bertrand, The filtering problem for continuous-time linear systems with Markovian switching coefficients, *Syst. Control Lett.*, **23** (1994), 453-461.
- [11] F. Dufour and R.J. Elliott, Adaptive control of linear systems with Markov perturbations, *IEEE Trans. Automatic Control* **43** (1997), 351-372.
- [12] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, J. Wiley, New York, 1986.
- [13] M. Iosifescu, *Finite Markov Processes and Their Applications*, Wiley, Chichester, 1980.
- [14] R. Z. Khasminskii, G. Yin and Q. Zhang, Asymptotic expansions of singularly perturbed systems involving rapidly fluctuating Markov chains, *SIAM J. Appl. Math.* **56** (1996), 277-293.
- [15] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes, with applications to Stochastic Systems Theory*, MIT Press, Cambridge, MA, 1984.
- [16] H. J. Kushner, *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, Birkhäuser, Boston, 1990.
- [17] H. J. Kushner and G. Yin, *Stochastic Approximation Algorithms and Applications*, Springer-Verlag, New York, 1997.
- [18] A. Il'in, R.Z. Khasminskii, and G. Yin, Singularly perturbed switching diffusions: Rapid switchings and fast diffusions, *J. Optim. Theory Appl.*, **102** (1999), 555-591.
- [19] R. S. Liptser and A. N. Shiriyayev, *Statistics of Random Processes I & II*, Springer-Verlag, New York, 2001.
- [20] B.M. Miller and W.J. Runggaldier, Kalman filtering for linear systems with coefficients driven by a hidden Markov jump process, *Syst. Control Lett.* **31** (1997), 93-102.
- [21] A. A. Pervozvanskii and V. G. Gaitsgori, *Theory of Suboptimal Decisions: Decomposition and Aggregation*, Kluwer, Dordrecht, 1988.
- [22] R. G. Phillips and P. V. Kokotovic, A singular perturbation approach to modelling and control of Markov chains, *IEEE Trans. Automat. Control* **26** (1981), 1087-1094.
- [23] W.J. Runggaldier and C. Visentin, Combined filtering and parameter estimation: Approximation and robustness, *Automatica*, **26** (1990), 401-404.
- [24] S. P. Sethi and Q. Zhang, *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Birkhäuser, Boston, 1994.
- [25] H. A. Simon and A. Ando, Aggregation of variables in dynamic systems, *Econometrica* **29** (1961), 111-138.
- [26] D. N. C. Tse, R. G. Gallager, and J. N. Tsitsiklis, Statistical multiplexing of multiple time-scale Markov streams, *IEEE J. Selected Areas Comm.* **13** (1995), 1028-1038.
- [27] A. Doucet, N.J. Gordon and V. Krishnamurthy, Particle Filtering for state estimation for jump Markov linear systems, *IEEE Trans. Signal Processing*, **49** (2001), 613-624.
- [28] C. Yang, Y. Bar-Shalom, and C.-F. Lin, Discrete-time point process filter for mode estimation, *IEEE Trans. Automatic Control*, **37** (1992), 1812-1816.
- [29] G. Yin and S. Dey, Weak convergence of hybrid filtering problems involving nearly completely decomposable hidden Markov chains, to appear in *SIAM J. Control Optim.*
- [30] G. Yin and Q. Zhang, *Continuous-time Markov Chains and Applications: A Singular Perturbation Approach*, Springer-Verlag, New York, 1998.
- [31] G. Yin and Q. Zhang, Singularly perturbed discrete-time Markov chains, *SIAM J. Appl. Math.* (2000) **61** (2000), 834-854.
- [32] G. Yin, Q. Zhang, and G. Badowski, Asymptotic properties of a singularly perturbed Markov chain with inclusion of transient states, *Ann. Appl. Probab.* **10** (2000), 549-572.
- [33] G. Yin, Q. Zhang, and G. Badowski, Decomposition and aggregation of large-dimensional Markov chains in discrete time, in *Proc. 40th Conf. Decision Control*, 1687-1692, 2001.
- [34] Q. Zhang, Hybrid filtering for linear systems with non-Gaussian disturbances, *IEEE Trans. Automatic Control*, **45** (2000), 50-61.
- [35] Q. Zhang and G. Yin, On nearly optimal controls of hybrid LQG problems, *IEEE Trans. Automat. Control*, **44** (1999), 2271-2282.