Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008



# Observation Noise-Gain Detection For Markov Chains Observed Through Scaled Brownian Motion

W. P. Malcolm and Alain Bensoussan,

***Abstract*— In this preliminary article we consider the problem of estimating an unknown noise-gain for a Markov chain observed through a scaled Brownian motion. It is assumed that the unknown noise-gain is time invariant.**

**Two objectives are addressed in this work, 1) compute an estimation scheme that is fast, and 2) compute an estimation scheme without recourse to stochastic integration. To address the first objective we avoid the Expectation Maximization (EM) algorithm, instead we develop an estimation scheme for a finite number of candidate model hypotheses. To address the second objective we develop a version of the Gauge-Transformation technique introduced by J. M. C. Clark.**

***Index Terms*— Wonham Filter, Martingales, Reference Probability, Filtering, Detection.**

## I. INTRODUCTION

The celebrated Wonham Filter, (see [10]), and it variants, have received considerable attention in the literature. A common problem with the Wonham Filter, (and indeed all filters), is model calibration, that is, estimating the model parameters. Traditionally the parameters considered in calibrating Wonham Filters are the so-called drift-coefficients in the observation-process model and the rate-matrix for the hidden Markov process. It is well known that these parameters can be estimated offline, see for example [6]. It is also possible to estimate these parameters online, see for example [4]. Curiously, by contrast, the literature shows little emphasis upon the estimation of an unknown noise-gain term.

In this work we propose an online detection scheme for noise-gain estimation. Our main result is a closed form, finite-dimensional discrete-time recursion for identifying the best, (most likely), noise-gain from a candidate set of possible noise-gains. It is shown that the usual stochastic-integration in detection, (in our context), is not easily, if at all, eliminated form the scalar-valued classical detector. However, by considering an augmented state-space model including the hidden Markov chain, one can develop estimation dynamics amenable to the Clark Transformation and thereby eliminate all stochastic integrations.

This article is organised as follows. In §I-A we describe the class of dynamics we wish to consider and briefly summarize reference probability for the task being considered. In §II we derive the classical stochastic integral detector for a unknown noise-gain. Subsequently we compute the corresponding matrix-valued detector on an augmented state-space model. Finally in §II-C we compute our main result which is a pathwise-deterministic detection scheme.

W. P Malcolm is with the Mathematical Sciences Institute of the Australian National University, Building 57, 0200 Canberra Australia, Email: **malcolm@maths.anu.edu.au**

Alain Bensoussan is the Director at the International Center for Decision and Risk Analysis in the School of Management at the University of Texas at Dallas USA. PO Box 830688, Richardson Texas 75083-0688, Email **alain.bensoussan@utdallas.edu**

### A. *State Process Dynamics*

All stochastic dynamics are defined (initially), on a fixed probability space, that is, $\big(\Omega, \mathcal{F}, P\big)$.

Our partially observed dynamical system of interest consists of a continuous-time Markov chain observed through a linear mapping corrupted by additive noise, here a scaled standard Brownian motion with a constant, but unknown gain. To model the uncertainty in the noise gain we propose a model taking $M \in \mathbb{N}$ possible values, according to the state of a simple random variable.

Suppose a process $X = \{X_t; 0 \leq t\}$, is a continuous-time Markov process taking values in a discrete state space $\mathcal{L}$. The state space $\mathcal{L}$ is defined by the set $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ where $\boldsymbol{e}_i$ are vectors in $\mathbb{R}^n$. Each $\boldsymbol{e}_i$ is a column vectors with unity in the $i^{th}$ position and zero elsewhere. This convenient canonical representation for a Markov chain has been widely used in filtering and other areas and leads to the dynamics,

$$X_t = X_0 + \int_0^t A\, X_u\, du + M_t. \tag{1.1}$$

Here $M_t$ is a vector-valued $(P, \sigma\big\{X_u, u \leq t\big\})$-martingale and the parameter $A$ is an $n \times n$ matrix whose elements are the infinitesimal generators of the Markov chain.

### B. *Observation Process Dynamics*

We suppose that the process $X$ is not observed directly, rather, we observe a scalar-valued process

$$y_t = \int_0^t \big\langle X_u, \boldsymbol{g}\big\rangle du + b\, W_t. \tag{1.2}$$

Here $W$ is a standard (scalar-valued) Wiener process on $(\Omega, \mathcal{F}, P)$ and $\boldsymbol{g} = \big(\langle \boldsymbol{g}, \boldsymbol{e}_1\rangle, \ldots, \langle \boldsymbol{g}, \boldsymbol{e}_n\rangle\big)' \in \mathbb{R}^n$, is a vector of the so called drift-coefficients, or levels for the Markov chain. The noise gain term $b \in \mathbb{R}$, is taken as unknown. Write

$$\mathcal{Y}_t = \sigma\big\{y_u, 0 \geq u \leq t\big\}. \tag{1.3}$$

To model the uncertainty in the noise gain term $b$, we suppose that $b$ can assume one of $M$ values according to the state of a simple random variable. We suppose that $\boldsymbol{\alpha}$ is a vector-valued simple random variable, whose state space is a collection of $M$ unit vectors $\{\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_M\}$. Here $\boldsymbol{f}_j$ has unity in the $jth$ position and zero elsewhere. Our

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

candidate values of the noise gain $b$ are written in vector form,

$$\boldsymbol{b} \triangleq (b_1, b_2, \ldots, b_M). \tag{1.4}$$

We now suppose that our observation process has dynamics

$$\begin{aligned} y_t &= \int_0^t \langle X_u, \boldsymbol{g} \rangle du + \langle \boldsymbol{b}, \boldsymbol{\alpha} \rangle W_t \\ &= \int_0^t \langle X_u, \boldsymbol{g} \rangle du + \Big( \sum_{j=1}^M \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle b_j \Big) W_t. \end{aligned} \tag{1.5}$$

The problem we wish to solve is, given a realisation $y = \{y_{t_1}, y_{t_2}, \ldots, y_{t_K}\}$, assumed to be a discrete-time sampling of the output generated by the dynamics at (1.1), estimate the probabilities

$$\widehat{p}^j \triangleq E[\mathbf{1}_{\{\omega | \boldsymbol{\alpha}(\omega) = b_j\}} \mid \mathcal{Y}_t] \;\; j = 1, 2, \ldots, M. \tag{1.6}$$

In what follows we will solve this problem via the change of probability measure techniques, moreover, our final solution will not involve stochastic integration.

### C. Reference Probability

As is now quite standard, we consider two probability measures, $P(\cdot)$ and $P^\dagger(\cdot)$. Under $P(\cdot)$, which we consider as the "real world" probability measure, our dynamical system evolves in time according to the dynamics given by (1.1). To facilitate the calculation of conditional expectations, we suppose that under the so-called reference measure $P^\dagger(\cdot)$, the observation process $y$ is a standard Brownian motion.

To begin our description of reference probability, we first suppose the observation process is generated by dynamics

$$y_t = \int_0^t \langle X_u, \boldsymbol{g} \rangle du + b_j W_t \tag{1.7}$$

Write

$$\Gamma_t^j \triangleq y_t / b_j = \int_0^t \langle X_u, \boldsymbol{g}/b_j \rangle du + W_t. \tag{1.8}$$

The combination of the scaled process dynamics for $\Gamma^j$ and the state process at (1.1), form a standard stochastic system, whose optimal filtering solution is the Wonham filter. We denote the real world probability measure for this system by $P^j(\cdot)$.

The Radon-Nikodym derivative for this system is, ¡

$$\begin{aligned} \Lambda_t^j &\triangleq \frac{dP^j}{dP^\dagger}\Big|_{\mathcal{G}_t^j} \\ &= \exp\Big( \int_0^t \langle X_u, \boldsymbol{g}/b_j \rangle d\Gamma_u^j - \\ &\qquad \tfrac{1}{2} \int_0^t \langle X_u, \boldsymbol{g}/b_j \rangle^2 du \Big) \\ &= 1 + \int_0^t \langle X_u, \boldsymbol{g}/b_j \rangle d\Gamma_u^j. \end{aligned} \tag{1.9}$$

Here $\mathcal{G}_t^j$ is the global sigma algebra generated by $\{y_u, X_u, \boldsymbol{\alpha}, 0 \le u \le t\}$, where here the process $y$ is assumed as generated by an observation model with noise gain $b_j$.

In what follows we wish to consider a collection of $M$ different stochastic systems, each with a unique noise gain term $b_j$. Recall that our individual systems are collectively written as a "single system", with dynamics given at (1.5). We now write the corresponding general Radon-Nikodym derivative as a convex combination of the individual noise-gain-specific Radon-Nikodym derivatives, that is,

$$\begin{aligned} \Lambda_t &\triangleq \sum_{j=1}^M \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \frac{dP^j}{dP^\dagger}\Big|_{\mathcal{G}_t^j} \\ &= \sum_{j=1}^M \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \Lambda_t^j \\ &= 1 + \sum_{j=1}^M \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \int_0^t \langle X_u, \boldsymbol{g}/b_j \rangle \Lambda_u^j d\Gamma_u^j \end{aligned} \tag{1.10}$$

## II. $M$-ARY DETECTION SCHEMES

In $M$-ary detection, in our context, we are interested to estimate un-normalised conditional probabilities,

$$\widehat{q}_t^j \triangleq E^\dagger \big[ \Lambda_t \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \mid \mathcal{Y}_t \big] \in \mathbb{R}_+^n. \tag{2.11}$$

### A. Classical Results

THEOREM 1 *The recursion for the estimated un-normalised probability $\widehat{q}$ has dynamics*

$$\widehat{q}_t^j = \widehat{q}_0^j + \int_0^t \langle E^{P^j} \big[ X_u \mid \mathcal{Y}_t^j \big], \boldsymbol{g}/b_j \rangle \widehat{q}_u^j d\Gamma_u^j \tag{2.12}$$

Proof of Theorem 1.

$$\begin{aligned} \widehat{q}_t^j &\triangleq E^{P^\dagger} \big[ \Lambda_t \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \mid \mathcal{Y}_t \big] \\ &= E^{P^\dagger} \Big[ \Big( 1 + \sum_{\ell=1}^M \langle \boldsymbol{\alpha}, \boldsymbol{f}_\ell \rangle \int_0^t \langle X_u, \boldsymbol{g}/b_\ell \rangle \Lambda_u^\ell d\Gamma_u^\ell \Big) \times \\ &\qquad \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \mid \mathcal{Y}_t \Big] \\ &= E^{P^\dagger} \big[ \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \mid \mathcal{Y}_t \big] \\ &\quad + E^{P^\dagger} \Big[ \int_0^t \langle X_u, \boldsymbol{g}/b_j \rangle \Lambda_u^j d\Gamma_u^j \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \mid \mathcal{Y}_t \Big] \\ &= \widehat{q}_0^j + \int_0^t E^{P^\dagger} \Big[ \Lambda_u^j \langle X_u, \boldsymbol{g}/b_j \rangle \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \mid \mathcal{Y}_u \Big] d\Gamma_u^j \end{aligned} \tag{2.13}$$

To simplify the expectation in the last line above, we note the following factorization,

$$\begin{aligned} &E^{P^\dagger} \Big[ \Lambda_t^j \langle X_t, \boldsymbol{g}/b_j \rangle \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \mid \mathcal{Y}_t \Big] \\ &\quad = E^{P^j} \big[ \langle X_t, \boldsymbol{g}/b_j \rangle \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \mid \mathcal{Y}_t \big] \times E^{P^\dagger} \big[ \Lambda_t^j \mid \mathcal{Y}_t \big] \\ &= E^{P^j} \big[ \langle X_t, \boldsymbol{g}/b_j \rangle \mid \mathcal{Y}_t \,\&\, \boldsymbol{\alpha} = \boldsymbol{f}_j \big] \times E^{P^j} \big[ \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \mid \mathcal{Y}_t \big] \times \\ &\qquad E^{P^\dagger} \big[ \Lambda_t \mid \mathcal{Y}_t \big] \\ &\quad = \big\langle E^{P^j} [X_t \mid \mathcal{Y}_t], \boldsymbol{g}/b_j \big\rangle \, \widehat{q}_t^j \end{aligned} \tag{2.14}$$

□.

The recursive scheme in Theorem 1 estimates the probabilities $\{q_t^1, \ldots, q_t^M\}$, with which one may determine the best noise coefficient from the candidate set $\{b_1, \ldots, b_M\}$, either by a hard or soft decision, respectively, a MAP estimate, or an MLE.

The term in the integrand of equation (2.12), namely $E^{P^j}\big[X_t \mid \mathcal{Y}_t^j\big]$, can easily be evaluated without stochastic integration, however, the stochastic integral in (2.12) against the process $\Gamma^j$ must still be approximated. What we would like to do is eliminate this particular integration by identifying a pathwise-deterministic version of our $M$-ary detection scheme.

### B. Augmented State-Space Detection

In order to eliminate the stochastic integration in the dynamics at (2.12), we consider an estimation scheme on an augmented state space, defined by the cartesian product of $\boldsymbol{\alpha}$ and $X$. Our estimator on this larger state-space will be a joint, state-estimator, (filter), and $M$-ary detector, that is, an estimator computing probabilities for events of the form $\{\boldsymbol{\alpha} = \boldsymbol{f}_j \,\&\, X_t = \boldsymbol{e}_i\}$. It will be shown that there exists a version of this joint estimator which is independent of stochastic integration and, from which the $M$-ary detector can be directly recovered by marginalising out the hidden state process $X$. To begin, we define a new process,

$$Z_t \triangleq \boldsymbol{\alpha} X_t'. \tag{2.15}$$

The process $Z$ takes values on a finite matrix-valued basis $\mathcal{C}$, where, $Z \in \mathcal{C} = \big\{\boldsymbol{H}^{(j,i)}\big\}_{\substack{1\le j\le M\\ 1\le i\le n}}$, that is,

$$\mathcal{C} \triangleq \left\{ \begin{bmatrix} 1 & 0 & \ldots & 0\\ 0 & \ldots & \ldots & 0\\ \vdots & & \ddots & \vdots\\ 0 & \ldots & \ldots & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & \ldots & \ldots & 0\\ \vdots & \ldots & \ldots & \vdots\\ 0 & & \ddots & \vdots\\ 1 & 0 & \ldots & 0 \end{bmatrix}, \right.$$
$$\vdots$$
$$\left. \begin{bmatrix} 0 & \ldots & 0 & 1\\ \vdots & \ldots & \ldots & 0\\ \vdots & & \ddots & \vdots\\ 0 & \ldots & \ldots & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & \ldots & \ldots & 0\\ \vdots & \ldots & \ldots & \vdots\\ \vdots & & \ldots & 0\\ 0 & \ldots & 0 & 1 \end{bmatrix} \right\}. \tag{2.16}$$

We also note the following useful representation,

$$\begin{aligned} Z_t &= \sum_{j=1}^{M}\sum_{i=1}^{n}\big(\boldsymbol{f}_j' Z_t \boldsymbol{e}_i\big)\boldsymbol{H}^{(j,i)}\\ &= \sum_{j=1}^{M}\sum_{i=1}^{n}\big\langle \boldsymbol{\alpha}, \boldsymbol{f}_j\big\rangle\big\langle X_t, \boldsymbol{e}_i\big\rangle \boldsymbol{H}^{(j,i)}. \end{aligned} \tag{2.17}$$

The calculation below shows how the $M$-ary detector may be directly computed by marginalising out the hidden state process $X$. First, write

$$\mathbf{1}_n \triangleq \big(1, 1, \ldots, 1\big)'. \tag{2.18}$$

Now, suppose one has obtained matrix-valued dynamics for the expectation $E^{P^\dagger}\big[\Lambda_t \boldsymbol{\alpha} X_t' \mid \mathcal{Y}_t\big]$, we then note that,

$$\begin{aligned} E^{P^\dagger}\big[\Lambda_t \boldsymbol{\alpha} X_t' \mid \mathcal{Y}_t\big]\mathbf{1}_n &= E^{P^\dagger}\big[\Lambda_t \boldsymbol{\alpha}\big(X_t'\mathbf{1}_n\big) \mid \mathcal{Y}_t\big]\\ &= E^{P^\dagger}\Big[\Lambda_t \boldsymbol{\alpha} \times \\ &\qquad \Big(\sum_{\ell=1}^{n}\big\langle X_t, \boldsymbol{e}_\ell\big\rangle \boldsymbol{e}_\ell' \mathbf{1}_n\Big) \mid \mathcal{Y}_t\Big]\\ &= E^{P^\dagger}\Big[\Lambda_t \boldsymbol{\alpha}\big(\sum_{\ell=1}^{n}\big\langle X_t, \boldsymbol{e}_\ell\big\rangle\big) \mid \mathcal{Y}_t\Big]\\ &= E^{P^\dagger}\big[\Lambda_t \boldsymbol{\alpha} \mid \mathcal{Y}_t\big] \ \in \mathbb{R}^{M\times 1}. \end{aligned} \tag{2.19}$$

This calculation shows that the detector of interest may be directly recovered from the corresponding detector on an augmented state-space.

Write

$$\widetilde{q}_t \triangleq E^{P^\dagger}\big[\Lambda_t \boldsymbol{\alpha} X_t' \mid \mathcal{Y}_t\big] \ \in \ \mathbb{R}^{M\times n}. \tag{2.20}$$

THEOREM 2 *The matrix-valued process $\widetilde{q}$, defined at (2.20), has dynamics*

$$\begin{aligned} \widetilde{q}_t = \widetilde{q}_0 &+ \int_0^t \widetilde{q}_u A' du\\ &+ \sum_{j=1}^{M}\sum_{i=1}^{n}\big\langle \boldsymbol{e}_i, \boldsymbol{g}/b_j\big\rangle \int_0^t \Big(\boldsymbol{f}_j' \widetilde{q}_u \boldsymbol{e}_i\Big) d\Gamma_u^j \boldsymbol{H}^{(j,i)} \end{aligned} \tag{2.21}$$

*The corresponding normalised detector probability is computed by normalising over all candidate model-hypotheses and hidden state process values, that is*

$$p(X_t = \boldsymbol{e}_i \,\&\, \boldsymbol{\alpha} = \boldsymbol{f}_j \mid \mathcal{Y}_t) = \frac{\boldsymbol{f}_j' \widetilde{q}_t \boldsymbol{e}_i}{\mathbf{1}_M' \widetilde{q}_t \mathbf{1}_n}. \tag{2.22}$$

**Proof of Theorem 2**

Theorem 2 can be established by using the Ito product rule and a stochastic calculus version of Fubini's Theorem developed in [11].

$$\begin{aligned} \Lambda_t \boldsymbol{\alpha} X_t' = \boldsymbol{\alpha} X_0' &+ \int_0^t \Lambda_u \boldsymbol{\alpha} X_u' A' du + \int_0^t \Lambda_u \boldsymbol{\alpha} dM_u'\\ &+ \int_0^t \sum_{\ell=1}^{M}\big\langle \boldsymbol{\alpha}, \boldsymbol{f}_\ell\big\rangle\big\langle X_u, \boldsymbol{g}/b_\ell\big\rangle \Lambda_u^\ell \boldsymbol{\alpha} X_u' d\Gamma_u^\ell \end{aligned} \tag{2.23}$$

Now conditioning the dynamics at (2.23), under the measure $P^\dagger(\cdot)$, and on the information $\mathcal{Y}_t$, we get,

$$\begin{aligned} \widetilde{q}_t = \widetilde{q}_0 &+ \int_0^t \widetilde{q}_u A' du\\ &+ \sum_{\ell=1}^{M}\int_0^t E^{P^\dagger}\big[\Lambda_u^\ell\big\langle \boldsymbol{\alpha}, \boldsymbol{f}_\ell\big\rangle\big\langle X_u, \boldsymbol{g}/b_\ell\big\rangle \boldsymbol{\alpha} X_u' \mid \mathcal{Y}_u\big] d\Gamma_u^\ell \end{aligned} \tag{2.24}$$

Recalling the representation at (2.17), we see that

$$\sum_{\ell=1}^{M}\int_0^t E^{P^\dagger}\left[\Lambda_u^\ell\langle\boldsymbol{\alpha},\boldsymbol{f}_\ell\rangle\langle X_u,\boldsymbol{g}/b_\ell\rangle\boldsymbol{\alpha}X_u' \mid \mathcal{Y}_u\right]d\Gamma_u^\ell$$
$$=\sum_{\ell=1}^{M}\int_0^t E^{P^\dagger}\Big[\Lambda_u^\ell\langle\boldsymbol{\alpha},\boldsymbol{f}_\ell\rangle\langle X_u,\boldsymbol{g}/b_\ell\rangle\times$$
$$\sum_{j=1}^{M}\sum_{i=1}^{n}\langle\boldsymbol{\alpha},\boldsymbol{f}_j\rangle\langle X_u,\boldsymbol{e}_i\rangle\boldsymbol{H}^{(j,i)} \mid \mathcal{Y}_u\Big]d\Gamma_u^\ell$$
$$=\sum_{j=1}^{M}\sum_{i=1}^{n}\int_0^t E^{P^\dagger}\left[\Lambda_u^j\langle\boldsymbol{e}_i,\boldsymbol{g}/b_j\rangle\times\right.$$
$$\left.\langle\boldsymbol{\alpha},\boldsymbol{f}_j\rangle\langle X_u,\boldsymbol{e}_i\rangle\boldsymbol{H}^{(j,i)} \mid \mathcal{Y}_u\right]d\Gamma_u^j$$
$$=\sum_{j=1}^{M}\sum_{i=1}^{n}\langle\boldsymbol{e}_i,\boldsymbol{g}/b_j\rangle\int_0^t\Big(\boldsymbol{f}'\widetilde{q}_u\boldsymbol{e}_i\Big)d\Gamma_u^j\boldsymbol{H}^{(j,i)} \quad (2.25)$$

□

To check the correspondence of the dynamics give at (2.21), with those given at (2.12), we project the joint estimator on (2.21) down onto the state space of the simple random variable $\boldsymbol{\alpha}$. That is, we consider product $\widetilde{q}_t\mathbf{1}_n$ which is the vector-valued form of the $M$-ary detector for the simple random variable $\boldsymbol{\alpha}$.

It is immediate that the desired correspondence holds for the two matrix-valued terms $\widetilde{q}_t$ and $\widetilde{q}_0$. Checking the bounded variation term in (2.21) we see that

$$\int_0^t\widetilde{q}_uA'du\mathbf{1}_n=\int_0^t\widetilde{q}_u\Big(A'\mathbf{1}_n\Big)du = \int_0^t\widetilde{q}_u\mathbf{0}du=0. \quad (2.26)$$

Recall that the rate matrix in our formulation has zero column-sums.

To check the correspondence of the Martingale term in (2.21), we first note that

$$\boldsymbol{H}^{(j,i)}\mathbf{1}_n=\boldsymbol{f}_j. \quad (2.27)$$

Further, using the representations at (2.17), we see that

$$\int_0^t\Big(\boldsymbol{f}_j'\widetilde{q}_u\boldsymbol{e}_i\Big)d\Gamma_u^j\boldsymbol{f}_j$$
$$=\int_0^t E^{P^\dagger}\left[\Lambda_u\langle X_u,\boldsymbol{e}_i\rangle\langle\boldsymbol{\alpha},\boldsymbol{f}_j\rangle \mid \mathcal{Y}_u\right]d\Gamma_u^j\boldsymbol{f}_j$$
$$=\int_0^t E^{P^j}\left[\langle X_u,\boldsymbol{e}_i\rangle\langle\boldsymbol{\alpha},\boldsymbol{f}_j \mid \mathcal{Y}_u\right]E^{P^j}\left[\langle\boldsymbol{\alpha},\boldsymbol{f}_j\rangle \mid \mathcal{Y}_u\right]$$
$$\times E^{P^\dagger}\left[\Lambda_u \mid \mathcal{Y}_u\right]d\Gamma_u^j\boldsymbol{f}_j$$
$$=\int_0^t\langle E^{P^j}\left[X_u \mid \mathcal{Y}_u\right],\boldsymbol{e}_i\rangle q_u^jd\Gamma_u^j\boldsymbol{f}_j \quad (2.28)$$

Finally, we note that

$$\sum_{i=1}^{n}\langle\boldsymbol{e}_i,\boldsymbol{g}/b_j\rangle\int_0^t\langle E^{P^j}\left[X_u \mid \mathcal{Y}_u\right],\boldsymbol{e}_i\rangle q_u^jd\Gamma_u^j\boldsymbol{f}_j$$
$$=\int_0^t\langle E^{P^j}\left[X_u \mid \mathcal{Y}_u\right],\boldsymbol{g}/b_j\rangle q_u^jd\Gamma_u^j\boldsymbol{f}_j \quad (2.29)$$

This calculation establishes the correspondence between the dynamics at (2.12) and the dynamics at (2.21).

### *C. Pathwise-Determinsitic Detection Schemes*

In this section we extend a seminal idea of J. M. C. Clark, (see [2]), to identify a version of the dynamics at equation (2.21) independent of stochastic integration. The particular transformation used here is a Hadamard product transformation. The following definitions are needed for our calculations.

DEFINITION 1 *Define a matrix-valued process* $\Phi$

$$\Phi_t\triangleq\left[\phi_t^{(j,i)}\right]_{\substack{i=1,\dots,n\\ j=1,\dots,M}}\in\mathbb{R}^{M\times n}, \quad (2.30)$$

*where*

$$\phi_t^{(j,i)}\triangleq\exp\Big(\langle\boldsymbol{g}/b_j,\boldsymbol{e}_i\rangle\Gamma_t^j-\tfrac{1}{2}\langle\boldsymbol{g}/b_j,\boldsymbol{e}_i\rangle^2t\Big), \quad (2.31)$$

$\forall\ (j,i), j=1,\dots,M, i=1,\dots,n.$

DEFINITION 2 *Define a constant Matrix $K$, where*

$$K\triangleq\begin{bmatrix}\langle\boldsymbol{g}/b_1,\boldsymbol{e}_1\rangle & \langle\boldsymbol{g}/b_1,\boldsymbol{e}_2\rangle & \dots & \langle\boldsymbol{g}/b_1,\boldsymbol{e}_n\rangle\\ \langle\boldsymbol{g}/b_2,\boldsymbol{e}_1\rangle & \langle\boldsymbol{g}/b_2,\boldsymbol{e}_2\rangle & \dots & \langle\boldsymbol{g}/b_2,\boldsymbol{e}_n\rangle\\ \vdots & \vdots & \ddots & \vdots\\ \langle\boldsymbol{g}/b_M,\boldsymbol{e}_1\rangle & \langle\boldsymbol{g}/b_M,\boldsymbol{e}_2\rangle & \dots & \langle\boldsymbol{g}/b_M,\boldsymbol{e}_n\rangle\end{bmatrix}. \quad (2.32)$$

DEFINITION 3 *Define a new matrix-valued stochastic process $\Psi$, where*

$$\Psi_t\triangleq\Phi_t^{-1}\odot\widetilde{q}_t. \quad (2.33)$$

*Here the symbol $\odot$ denotes the Hadamard matrix product. Note that this product is commutative.*

What we would like to do, is compute dynamics for the process $\Psi$. To this end, we start by applying the Ito rule we obtain the dynamics for the process $\Phi^{-1}$, that is,

$$\Phi_t^{-1}=\Phi_0^{-1}+\int_0^t\Big(K\odot K\Big)\odot\Phi_u^{-1}du-$$
$$\sum_{j=1}^{M}\sum_{i=1}^{n}\int_0^t\Big(K\odot\Phi_u^{-1}\odot\boldsymbol{H}^{(j,i)}\Big)d\Gamma_u^j. \quad (2.34)$$

REMARK 1 The values taken by the process $\Phi_t^{-1}$ are not regular matrix inverses, rather Hadamard inverses.

Now, recalling the dynamics for the matrix-valued process $\widetilde{q}$, we see that,

$$\begin{aligned} d\Big(\Phi_t^{-1}\odot\widetilde{q}_t\Big) &= \Big(d\Phi_t^{-1}\Big)\odot\widetilde{q}_t + \Phi_t^{-1}\odot\Big(d\widetilde{q}_t\Big) + d\Big[\Phi_t^{-1}\odot\widetilde{q}_t\Big] \\ &= \Big(K\odot K\Big)\odot\Phi_t^{-1}\odot\widetilde{q}_t dt \\ &\quad -\sum_{j=1}^{M}\sum_{i=1}^{n} K\odot\Phi_t^{-1}\boldsymbol{H}^{(j,i)}\odot\widetilde{q}_t d\Gamma_t^j \\ &\quad +\Phi_t^{-1}\odot\widetilde{q}_t A' dt + \Phi_t^{-1}\odot \\ &\qquad \Big(\sum_{j=1}^{M}\sum_{i=1}^{n}\langle\boldsymbol{e}_i,\boldsymbol{g}/b_j\rangle\Big(\boldsymbol{f}_j'\widetilde{q}_t\boldsymbol{e}_i\Big)\boldsymbol{H}^{(j,i)}\Big)d\Gamma_t^j \\ &\quad -\Big(\sum_{j=1}^{M}\sum_{i=1}^{n} K\odot\Phi_t^{-1}\odot \\ &\quad \boldsymbol{H}^{(j,i)}\Big)\odot\Big(\sum_{j=1}^{M}\sum_{i=1}^{n}\langle\boldsymbol{e}_i,\boldsymbol{g}/b_j\rangle\Big(\boldsymbol{f}_j'\widetilde{q}_t\boldsymbol{e}_i\Big)\boldsymbol{H}^{(j,i)}\Big)dt. \end{aligned} \tag{2.35}$$

The only surviving term in the expansion above is the rate matrix term, so it follows that

$$\frac{d\Psi_t}{dt} = \Phi_t^{-1}\odot\Big(\widetilde{q}_t A'\Big)dt. \tag{2.36}$$

The dynamics at (2.36) are a pathwise deterministic linear ordinary differential equation (LODE). This LODE may be discretized in a variety of ways, for brevity we consider an Euler scheme on a regular time partition, whose epochs are labelled $t_k, t_{k+1}, \dots$ etc. The Euler scheme leads to the approximation,

$$\begin{aligned} \Psi_{t-k} &= \Psi_{t_{k-1}} + \int_{t_{k-1}}^{t_k}\Phi_u^{-1}\odot\Big(\widetilde{q}_t A'\Big)du \\ &\approx \Psi_{t_{k-1}} + \Phi_{t_{k-1}}^{-1}\odot\Big(\widetilde{q}_{t_{k-1}}A'\Big)\Delta_t \end{aligned} \tag{2.37}$$

Recalling the definition of $\Psi$, we take the Hadamard product $\Phi_{t_k}\odot\Psi_{t_k}$ and get

$$\begin{aligned} \widetilde{q}_{t_k} &= \Phi_{t_k}\odot\Phi_{t_{k-1}}^{-1}\odot\widetilde{q}_{t_{k-1}} + \\ &\qquad \Phi_{t_k}\odot\Phi_{t_{k-1}}^{-1}\odot\widetilde{q}_{t_{k-1}}A'\Delta_t \\ &= \Phi_{t_k}\odot\Phi_{t_{k-1}}^{-1}\odot\Big[\boldsymbol{I}_{m\times n} + \widetilde{q}_{t_{k-1}}A'\Delta_t\Big] \end{aligned} \tag{2.38}$$

Here $\boldsymbol{I}_{M\times n}$ is an $M\times n$ matrix with unity in each element.

REMARK 2 The discrete-time matrix-valued recursion at (2.38) provides a scheme to compute un-normalised probabilities corresponding to all of the joint events $\{\boldsymbol{\alpha} = \boldsymbol{f}_j \& X_{t_k} = \boldsymbol{e}_i\}$ without recourse to stochastic integration.

Finally, to recover our estimator of interest, that is, a vector of conditional probabilities for the state of the simple random variable $\boldsymbol{\alpha}$, we compute

$$\left\{\frac{\widetilde{q}_{t_k}}{\boldsymbol{1}_M'\widetilde{q}_{t_k}\boldsymbol{1}_n}\right\}\boldsymbol{1}_n = \begin{bmatrix}\widehat{P}(\boldsymbol{\alpha}=\boldsymbol{f}_1\mid\mathcal{Y}_{t_k})\\ \widehat{P}(\boldsymbol{\alpha}=\boldsymbol{f}_2\mid\mathcal{Y}_{t_k})\\ \vdots \\ \widehat{P}(\boldsymbol{\alpha}=\boldsymbol{f}_m\mid\mathcal{Y}_{t_k})\end{bmatrix}. \tag{2.39}$$

THEOREM 3 *For the process $\alpha X'$, the quantity*

$$\pi(\boldsymbol{\alpha}X_t') \triangleq \frac{\Phi_t\odot\Psi_t}{\boldsymbol{1}_M'\big(\Phi_t\odot\Psi_t\big)\boldsymbol{1}_n}. \tag{2.40}$$

*defines a locally Lipschitz continuous version of the expectation $E\big[\boldsymbol{\alpha}X_t'\mid\mathcal{Y}_t\big]$ in the space of observation sample paths.*

Theorem 3 is stated here without proof. Similar proofs are given in [2].

## III. CONCLUSION

In this article a detection scheme is developed to identify the best of a candidate set of noise-gain values for a Markov chain observed through a scaled Brownian motion. To eliminate stochastic integration in the resulting detector dynamics, a transformation based upon the ideas of J. M. C. Clark are applied through a Hadamard product transformation. This transformation is applied to an augmented state-space filter which is essential a special Wonham filter on a product state-space. A specific version of the detector dynamics is identified as a pathwise-deterministic linear ordinary differential equation. This LODE may be discretized in a variety of ways. It was shown that implementing the Euler approximation leads to a discrete-time recursion for direct computation of the estimated detector probabilities $\widehat{P}(\boldsymbol{\alpha}=\boldsymbol{f}_j\mid\mathcal{Y}_t)$.

## IV. ACKNOWLEDGMENT

W. P. Malcolm would like to sincerely thank Professor Alain Bensoussan and the School of Management at the University of Texas at Dallas, for hospitality during November 2007.

## REFERENCES

[1] A. Bensoussan, *Stochastic control of partially observable systems*. *Cambridge University Press* Cambridge (UK) 1992.

[2] J. M. C. Clark, The design of robust approximations to the stochastic differential equations of nonlinear filtering. In : J.Skwirzynski (ed.), *Communication systems and random process theory* Sijthoff and Noordhoof (1978).

[3] M. H. A. Davis, Pathwise Nonlinear Filtering, Stochastic Systems: The Mathematics of Filtering and Applications (Editors M. Hazewinkel and J. C. Willems), NATO Advanced Study Institute, Reidel Dordrecht, July 1980, pp. 505-528.

[4] R. J. Elliott and W. P. Malcolm, Robust $M$-ary Detection Filters and smoothers for continuous-time jump Markov systems, IEEE Transactions on Automatic Control, Volume 40, Issue 7, pp. 1046-1045, July 2004.

[5] R. J. Elliott, *Stochastic Calculus and its Applications*, Springer Verlag 1982.

[6] M. R. James, V. Krishnamurthy and F. Le Gland, Time Discretization for Continuous Time Filters and Smoothers for HMM Parameter Estimation, Transactions on Information Theory, Volume 42, Number 2, March 1998.

[7] W. P. Malcolm and R. J. Elliott, A General Smoothing Equation For Poisson Observations, 38th I.E.E.E Conference on Decision and Control, December 1999, Phoenix USA.

[8] W. P. Malcolm, Robust Filtering And Estimation With Poisson Observations, PhD Thesis, The Australian National University, 1999.

[9] W. P. Malcolm, R. J. Elliott and J. van der Hoek, On The Numerical Stability of Time-Discretised State Estimation *via* Clark Transformations, 42nd I.E.E.E Conference on Decision and Control, December 2003, Hawai USA.

[10] M. W. Wonham, Some Applications of Stochastic Differential Equations to Nonlinear Filtering, SIAM Journal Of Control and Optimisation, Volume 2, Number 3, pp. 357-369, 1965.

[11] E. Wong and B. Hajek, Stochastic Processes in Engineering Systems, Springer Verlag, New York, 1985.