Control of Discrete-Time HMM Partially Observed Under Fractional Gaussian Noises

Robert J. Elliott [*] and Tak Kuen Siu [†]

Abstract

A discrete-time control problem of a finite-state hidden Markov chain partially observed in a fractional Gaussian process is discussed using filtering. The control problem is then recast as a separated problem with information variables given by the unnormalized conditional probabilities of the whole path of the hidden Markov chain. A dynamic programming result and a minimum principle are obtained.

Keywords: Discrete-Time Control; Hidden Markov Models; Fractional Gaussian Noises; Unnormalized Conditional Probability; Dynamic Programming; Minimum Principle.

[*]School of Mathematical Sciences, University of Adelaide, Adelaide, South Australia, Australia; Haskayne School of Business, University of Calgary, Calgary, Alberta, Canada; Email: relliott@ucalgary.ca.

[†]Department of Applied Finance and Actuarial Studies, Faculty of Business and Economics, Macquarie University, Sydney, Australia; Email: ktksiu2005@gmail.com.

1 Introduction

Hidden Markov Models (HMMs) are a powerful mathematical tool and have important applications in diverse fields, including information engineering, bio-informatics, bio-engineering, speech hearing, economics, finance, actuarial science, and many others. The monograph by Elliott et al. [3] provided a comprehensive discussion on HMMs, their stochastic calculus, filtering and control as well as some applications of HMMs. In Elliott et al. [3], a powerful technique based on a reference probability was adopted to discuss the filtering and control of HMMs. The central tenet of the reference probability approach is the use of the Bayes' rule for a measure change. It starts with a reference probability under which the model dynamics have simple forms. Then the real-world probability, under which the model dynamics have their original forms, is constructed from the reference probability via the measure change.

A typical HMM is that a hidden Markov chain is partially observed under Gaussian noise in discrete-time. In this model, the state process is the hidden Markov chain and the observation process is a Gaussian process with drift modulated by the chain. In a recent paper, Elliott and Deng [4] generalized this HMM to the case when the observation process is a fractional Gaussian process so that it incorporates long-term memory in the observation process. Indeed, long-term memory is an important feature of time series. Benoit Mandelbrot described the long-term memory as the "Joseph effect" and characterized this effect using fractal dimensions, (see Mandelbrot [8]). There are other models of long-term memory including fractional differentiation in a continuous-time setting and fractional differencing in discrete-time. Mandelbrot and Van Ness [7] proposed the use of fractional Brownian motion and fractional noise to model long-range dependence. Elliott and Deng [4] characterized long-term memory using fractional differencing and derived filters of the hidden Markov chain and related quantities. They also derived estimates of the model parameters using EM algorithm. Discrete-time stochastic optimal control problems are treated, for example, in Kumar and Varaiya [6] and Cairnes [1]. Elliott et al. [3] discussed the discrete-time, partially observed control problem using the reference probability approach. However, it seems that the discrete-time, partially observed control problem with long-term memory observations has, so far, not yet been treated in the existing literature.

In this paper, we discuss a discrete-time control problem for a finite-state hidden Markov chain partially observed in a fractional Gaussian process using filtering. The objective is to minimize a cost functional associated with the whole path of the hidden Markov chain. The case when the transition probability matrix of the chain depends on a control parameter is considered. We construct explicitly the reference probability and use the unnormalized conditional probabilities of the whole path of the hidden Markov chain, given the observations about the fractional Gaussian process, as information state variables. The control problem is then recast as a fully observed optimal control problem, where the unnormalized conditional probabilities play the role of information state variables. We give a dynamic programming result and a minimum principle.

This paper is organized as follows. The next section presents the model dynamics. In Section 3, we construct explicitly the reference probability. Section 4 derives a recursion for the unnormalized conditional probabilities of the whole path of the hidden Markov chain. In Section 5, we first present the control problem and its separated problem. We then discuss the separated problem using the dynamic programming and the minimum principles.

2 The Dynamics

We first introduce the concept of fractional differencing and then describe a Markov chain partially observed in a fractional Gaussian process.

2.1 Fractional Differencing

Let $\mathcal{Z}^+$ be the set of non-negative integers $\{0, 1, 2, \cdots\}$. Write $\mathcal{L}$ for the space of real-valued functions $f : \mathcal{Z}^+ \to \Re$. (We suppose that if $i < 0$, $f(i) = 0$.) Consequently, the function space $\mathcal{L}$ is isomorphic to the space of infinite sequences, say $f(0) = f_0$, $f(1) = f_1$, $\cdots$, $f(i) = f_i$, $\cdots$.

Definition 2.1. *For any $f, g \in \mathcal{L}$, the convolution product of f and g, denoted by $f \star g$,* *is defined by:*

$$(f * g)(n) := \sum_{i=0}^{\infty} f_i g_{n-i} = \sum_{i=0}^{n} f_i g_{n-i} \ .$$

Consider a function $I := (1, 0, 0, \cdots) \in \mathcal{L}$. Then for any function $f \in \mathcal{L}$,

$$(I * f)(n) = (f * I)(n) = f_n \ .$$

This is the identity operator for convolution multiplication.

Let $u^{*k} := u * u * \cdots u$, the k^{th} convolution power of u, for each $k = 0, 1, \cdots$. By convention, $u^{*0} := (1, 0, \cdots, 0) = I$. The following lemma gives an expression for u^{*k}. It follows directly from induction.

Lemma 2.1. *For each* $k = 1, 2, \cdots$,

$$u^{*k} := \left(1, \frac{k}{1!}, \frac{k(k+1)}{2!}, \frac{k(k+1)(k+2)}{3!}, \cdots\right) \ .$$

Indeed, Elliott and Miao [5] generalized this by defining u^{*k} for any $k \in \Re$. The following theorem is due to Elliott and Deng [4]. We state the result here without giving the proof.

Theorem 2.1. *For any* $r, s \in \Re$,

$$(u^r * u^s) = u^{*(r+s)} \ .$$

Corollary 2.1. *For any* $r \in \Re$,

$$u^r * u^{-r} = I \ .$$

This result follows directly from Theorem 2.1 and the convention that $u^{*0} = I$.

2.2 HMM Partially Observed in Fractional Gaussian Noise

Consider a discrete-time, N-state, hidden Markov chain $\mathbf{X} := \{\mathbf{X}_t | t \in \mathcal{Z}^+\}$ defined on a complete probability space $(\Omega, \mathcal{F}, P)$ with state space $\mathcal{S} := \{\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_N\} \subset \Re^N$.

Following the convention in Elliott et al. [3], without loss of generality, we identify the state space of the chain $\mathbf{X}$ to be a set of standard unit vectors $\mathcal{E} := \{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_N\} \in \Re^N$, where the j^{th} component of $\mathbf{e}_i$ is the Kronecker delta δ_{ij}.

To specify the probability law of the chain $\mathbf{X}$, we must define the transition probability matrix of the chain. Here we suppose that the transition probability matrix $\mathbf{A}(\cdot)$ of the Markov chain $\mathbf{X}$ depends on a control parameter θ taking values in some measurable space Θ. Let $Y := \{Y_t | t \in \mathcal{Z}^+\}$ be our observation process to be defined in the later part of this subsection. Write, for each $t \in \mathcal{Z}^+$, $\mathcal{F}_t^Y$ for P-completed σ-field generated by the values of the observation process Y up to and including time t. We suppose that for each $t \in \mathcal{Z}^+$, the control θ_t at time t is $\mathcal{F}_t^Y$-measurable. Write, for each $t \in \mathcal{Z}^+$, $\underline{\Theta}(t)$ for the space of such controls, and

$$\underline{\Theta}(k, k+l) := \underline{\Theta}(k) \cup \underline{\Theta}(k+1) \cup \cdots \cup \underline{\Theta}(k+l) \ .$$

Let T be a finite horizon. Then for each $\theta := (\theta_0, \theta_1, \cdots, \theta_{T-1}) \in \underline{\Theta}(0, T-1)$ with $\theta_t \in \underline{U}(t)$, $\mathbf{X}^\theta$ denotes the corresponding controlled Markov chain having the transition probability matrix $\mathbf{A}(\theta_t)$ at time t. Note that, for each $t = 0, 1, \cdots, T-1$, $\mathbf{A}(\theta_t) := [a_{ji}(\theta_t)]_{i,j=1,2,\cdots,N}$, where $a_{ji}(\theta_t)$ is the probability that the chain $\mathbf{X}$ transits from state $\mathbf{e}_i$ at time t to state $\mathbf{e}_j$ at time $t+1$, and this probability depends on the control parameter θ_t at time t. To simplify the notation, we suppress the superscript θ and write $\mathbf{X}$ for $\mathbf{X}^\theta$ unless otherwise stated.

For each $t \in \mathcal{Z}^+$, let $\mathcal{F}_t^{\mathbf{X}} := \sigma\{\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_t\} \vee \mathcal{N}$, the minimal σ-algebra generated by information about the values of the chain $\mathbf{X}$ up to and including time t and the collection $\mathcal{N}$ of P-null sets. Then with the canonical state space $\mathcal{E}$ of the chain $\mathbf{X}$, Elliott et al. [3] gave the following dynamics of the chain $\mathbf{X}$:

$$\mathbf{X}_{t+1} = \mathbf{A}(\theta_t)\mathbf{X}_t + \mathbf{M}_{t+1} \ . \tag{2.1}$$

Here $\mathbf{M} := \{\mathbf{M}_t | t \in \mathcal{Z}^+ \backslash \{0\}\}$ is an $\Re^N$-valued, martingale difference process.

We suppose that the chain $\mathbf{X}$ is not observed directly; rather, we observe a non-zero drift, fractional Gaussian, process Y to be defined in the sequel.

Consider a sequence of random variables $w := \{w_t | t \in \mathcal{Z}^+\}$ such that

1. $\{w_t | t \in \mathcal{Z}^+\}$ is a sequence of independent and identically distributed, (i.i.d.), random variables such that $w_t \sim N(0,1)$ and $w_0 = 0$, P-a.s;
2. w and $\mathbf{X}$ are stochastically independent under P.

Following Elliott and Deng [4], we define a fractional Gaussian noise $w^r := \{w_t^r | t \in \mathcal{Z}^+\}$ as:

$$w_t^r := (u^r * w)(t) = \sum_{k=0}^{\infty} u_k^r w_{t-k} = \sum_{k=0}^{t} u_k^r w_{t-k} \ ,$$

where $w_0^r = 0$, P-a.s.

Consequently, w^r is a sequence of Gaussian random variables which have long-memory and are correlated. We suppose that the observation process Y follows a fractional Gaussian process.

$$Y_t \quad = \quad \langle \mathbf{h}, \mathbf{X}_t \rangle + w_t^r \ , \quad t \in \mathcal{Z}^+ \ .$$

Here $\mathbf{h} := (h_1, h_2, \cdots, h_N)' \in \Re^N$ with $h_i \in \Re$ for each $i = 1, 2, \cdots, N$.

Using the fractional differencing and convolution product discussed in Section 2.1, we now define a Gaussian process $Z := \{Z_t | t \in \mathcal{T}\}$ associated with the observation process Y. The Gaussian process Z will be used in later parts of the paper.

Suppose, for each $t \in \mathcal{Z}^+$,

$$Z_t := (u^{-r} * Y)(t) \ .$$

Note that

$$\begin{aligned} u^{-r} \quad &:= \quad \left(1, \frac{-r}{1!}, \frac{-r(-r+1)}{2!}, \frac{-r(-r+1)(-r+2)}{3!}, \cdots\right) \ , \\ \langle \mathbf{h}, \mathbf{X} \rangle \quad &:= \quad (\langle \mathbf{h}, \mathbf{X}_0 \rangle, \langle \mathbf{h}, \mathbf{X}_1 \rangle, \langle \mathbf{h}, \mathbf{X}_2 \rangle, \cdots) \ . \end{aligned}$$

Then by Theorem 2.1, it is not difficult to see that for each $t \in \mathcal{Z}^+$,

$$Z_t \quad = \quad (u^{-r} * \langle \mathbf{h}, \mathbf{X} \rangle)(t) + w_t \ .$$

Write, for each $t \in \mathcal{Z}^+$,

$$\gamma_t(\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_t) := (u^{-r} * \langle \mathbf{h}, \mathbf{X} \rangle)(t) \ .$$

Consequently,

$$Z_t = \gamma_t(\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_t) + w_t \ . \tag{2.2}$$

3 A Measure Change

We start with a reference probability measure $\bar{P}$ on $(\Omega, \mathcal{F})$ under which both the observation process Y and the Markov chain $\mathbf{X}$ have simple dynamics. That is, under $\bar{P}$,

1. $\{Z_t | t \in \mathcal{Z}^+\}$ is a sequence of i.i.d. random variables with common distribution $N(0, 1)$ and
2. $\{\mathbf{X}_t | t \in \mathcal{Z}^+\}$ is a sequence of i.i.d. random variables uniformly distributed over the set of unit standard vectors $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_N\}$.

In what follows, we construct the probability measure $P^{\gamma, \mathbf{A}}$ from $\bar{P}$ such that under $P^{\gamma, \mathbf{A}}$, the processes Z and $\mathbf{X}$ are governed by the dynamics (2.2) and (2.1), respectively.

Firstly, we specify the structure of information flow. Define $F^Z := \{\mathcal{F}_t^Z | t \in \mathcal{Z}^+\}$ and $G := \{\mathcal{G}_t | t \in \mathcal{Z}^+\}$ by:

$$\begin{aligned} \mathcal{F}_t^Z &:= \sigma\{Z_0, Z_1, \cdots, Z_t\} \vee \mathcal{N} \ , \\ \mathcal{G}_t &:= \sigma\{Z_0, Z_1, \cdots, Z_t, \mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_t\} \vee \mathcal{N} \ , \end{aligned}$$

and $\mathcal{F}_0^Z := \sigma\{\emptyset, \Omega\} \vee \mathcal{N}$ and $\mathcal{G}_0 := \sigma\{\mathbf{X}_0\} \vee \mathcal{N}$.

Note that the filtration generated by the process Z is equivalent to that generated by the observation process Y.

Let $\phi(z)$ is the density function of $N(0, 1)$. That is,

$$\phi(z) := \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \ .$$

Write, for each $t = 1, 2, \cdots, T-1$,

$$\mathbf{a}_{t+1}(\theta_t) := \mathbf{A}(\theta_t)\mathbf{X}_t \ ,$$

and, for each $i = 1, 2, \cdots, N$,

$$a^i_{t+1}(\theta_t) := \langle \mathbf{a}_{t+1}(\theta_t), \mathbf{e}_i \rangle = \langle \mathbf{A}(\theta_t)\mathbf{X}_t, \mathbf{e}_i \rangle \ ,$$

so that

$$\sum_{i=1}^{N} a^i_{t+1}(\theta_t) = 1 \ .$$

Define two G-adapted processes $\lambda^\gamma := \{\lambda^\gamma_t | t = 1, 2, \cdots, T\}$ and $\lambda^{\mathbf{A}} := \{\lambda^{\mathbf{A}}_t | t = 1, 2, \cdots, T\}$ by:

$$\begin{aligned} \lambda^\gamma_t &:= \frac{\phi(Z_t - \gamma_t(\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_t))}{\phi(Z_t)} \ , \\ \lambda^{\mathbf{A}}_t &:= \prod_{i=1}^{N} (N a^i_t(\theta_{t-1}))^{\langle \mathbf{X}_t, \mathbf{e}_i \rangle} \ . \end{aligned}$$

Set

$$\lambda^{\gamma,\mathbf{A}}_t := \lambda^\gamma_t \cdot \lambda^{\mathbf{A}}_t \ , \quad t = 1, 2, \cdots, T \ .$$

Consider another G-adapted process $\Lambda^{\gamma,\mathbf{A}} := \{\Lambda^{\gamma,\mathbf{A}}_t | t = 0, 1, \cdots, T\}$ defined by:

$$\begin{aligned} \Lambda^{\gamma,\mathbf{A}}_t &:= \prod_{k=1}^{t} \lambda^{\gamma,\mathbf{A}}_k \ , \quad t = 1, 2, \cdots, T \ , \\ \Lambda^{\gamma,\mathbf{A}}_0 &:= 1 \ . \end{aligned}$$

Then we have the following lemma.

Lemma 3.1. $\Lambda^{\gamma,\mathbf{A}}$ *is a* $(G, \bar{P})$*-martingale.*

Proof. Write $\bar{\mathrm{E}}$ for expectation under $\bar{P}$. For each $t = 0, 1, \cdots, T-1$,

$$
\begin{aligned}
& \bar{\mathrm{E}}\Big[\frac{\Lambda_{t+1}^{\gamma,\mathbf{A}}}{\Lambda_t^{\gamma,\mathbf{A}}}|\mathcal{G}_t\Big] \\
= \;& \bar{\mathrm{E}}[\lambda_{t+1}^{\gamma,\mathbf{A}}|\mathcal{G}_t] \\
= \;& \bar{\mathrm{E}}\Big\{\prod_{i=1}^{N}(Na_{t+1}^{i}(\theta_t))^{\langle \mathbf{X}_{t+1},\mathbf{e}_i\rangle}\bar{\mathrm{E}}\Big[\frac{\phi(Z_{t+1}-\gamma_{t+1}(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_{t+1}))}{\phi(Z_{t+1})}|\mathcal{G}_t\vee\sigma\{\mathbf{X}_{t+1}\}\Big]|\mathcal{G}_t\Big\} \\
= \;& \bar{\mathrm{E}}\Big\{\prod_{i=1}^{N}(Na_{t+1}^{i}(\theta_t))^{\langle \mathbf{X}_{t+1},\mathbf{e}_i\rangle}\int_{\Re}\frac{\phi(z-\gamma_t(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_{t+1}))}{\phi(z)}\phi(z)dz|\mathcal{G}_t\Big\} \\
= \;& \bar{\mathrm{E}}\Big\{\prod_{i=1}^{N}(Na_{t+1}^{i}(\theta_t))^{\langle \mathbf{X}_{t+1},\mathbf{e}_i\rangle}|\mathcal{G}_t\Big\} \\
= \;& \sum_{i=1}^{N}Na_{t+1}^{i}(\theta_t)\bar{P}(\mathbf{X}_{t+1}=\mathbf{e}_i) \\
= \;& \sum_{i=1}^{N}Na_{t+1}^{i}(\theta_t)\frac{1}{N}=1\ , \quad \bar{P}\text{-a.s.}
\end{aligned}
$$

Hence the result follows. $\square$

We now define $P^{\gamma,\mathbf{A}}$ by putting:

$$\frac{dP^{\gamma,\mathbf{A}}}{d\bar{P}}\Big|_{\mathcal{G}_t} := \Lambda_t^{\gamma,\mathbf{A}}\ .$$

Then the following theorem gives the probability laws of the observation process Z and the chain $\mathbf{X}$ under $P^{\gamma,\mathbf{A}}$.

Theorem 3.1. *Under $P^{\gamma,\mathbf{A}}$,*

$$w_t := Z_t - \gamma_t(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_t)\ , \quad t = 1, 2, \cdots, T\ ,$$

is a sequence of $N(0,1)$, i.i.d., random variables. Further, for each $t = 0, 1, \cdots, T-1$,

$$E^{\gamma,\mathbf{A}}[\mathbf{X}_{t+1}|\mathcal{G}_t] = \mathbf{A}(\theta_t)\mathbf{X}_t\ .$$

Here $E^{\gamma,\mathbf{A}}$ is expectation under $P^{\gamma,\mathbf{A}}$.

Proof. Let $f : \Re \to \Re$ be a measurable test function. Then by a version of the Bayes' rule,

$$
\begin{aligned}
& \mathrm{E}^{\gamma,\mathbf{A}}[f(w_{t+1})|\mathcal{G}_t] \\
= & \frac{\bar{\mathrm{E}}[\Lambda^{\gamma,\mathbf{A}}_{t+1} f(w_{t+1})|\mathcal{G}_t]}{\bar{\mathrm{E}}[\Lambda^{\gamma,\mathbf{A}}_{t+1}|\mathcal{G}_t]} \\
= & \frac{\bar{\mathrm{E}}[\lambda^{\gamma,\mathbf{A}}_{t+1} f(w_{t+1})|\mathcal{G}_t]}{\bar{\mathrm{E}}[\lambda^{\gamma,\mathbf{A}}_{t+1}|\mathcal{G}_t]} \\
= & \bar{\mathrm{E}}[\lambda^{\gamma,\mathbf{A}}_{t+1} f(w_{t+1})|\mathcal{G}_t] \\
= & \bar{\mathrm{E}}\Big\{ \prod_{i=1}^{N} (N a^i_{t+1}(\theta_t))^{\langle \mathbf{X}_{t+1}, \mathbf{e}_i \rangle} \bar{\mathrm{E}}\Big[f(w_{t+1}) \frac{\phi(w_{t+1})}{\phi(Z_{t+1})} |\mathcal{G}_t \vee \sigma\{\mathbf{X}_{t+1}\}\Big] |\mathcal{G}_t \Big\} \\
= & \bar{\mathrm{E}}\Big[\prod_{i=1}^{N} (N a^i_{t+1}(\theta_t))^{\langle \mathbf{X}_{t+1}, \mathbf{e}_i \rangle} \int_{\Re} f(w)\phi(w) dw |\mathcal{G}_t \Big] \\
= & \Big(\int_{\Re} f(w)\phi(w) dw \Big) \times \bar{\mathrm{E}}\Big[\prod_{i=1}^{N} (N a^i_{t+1}(\theta_t))^{\langle \mathbf{X}_{t+1}, \mathbf{e}_i \rangle} |\mathcal{G}_t \Big] \\
= & \int_{\Re} f(w)\phi(w) dw \ .
\end{aligned}
$$

Since f is an arbitrary measurable function, the first statement of the theorem follows.

Similarly, we prove the second statement of the theorem. Again by a version of the Bayes' rule,

$$
\begin{aligned}
\mathrm{E}^{\gamma,\mathbf{A}}[\mathbf{X}_{t+1}|\mathcal{G}_t] &= \bar{\mathrm{E}}[\lambda^{\gamma,\mathbf{A}}_{t+1} \mathbf{X}_{t+1}|\mathcal{G}_t] \\
&= \bar{\mathrm{E}}\Big\{ \mathbf{X}_{t+1} \prod_{i=1}^{N} (N a^i_{t+1}(\theta_t))^{\langle \mathbf{X}_{t+1}, \mathbf{e}_i \rangle} \bar{\mathrm{E}}\Big[\frac{\phi(w_{t+1})}{\phi(Z_{t+1})} |\mathcal{G}_t \vee \sigma\{\mathbf{X}_{t+1}\}\Big] |\mathcal{G}_t \Big\} \\
&= \bar{\mathrm{E}}\Big[\mathbf{X}_{t+1} \prod_{i=1}^{N} (N a^i_{t+1}(\theta_t))^{\langle \mathbf{X}_{t+1}, \mathbf{e}_i \rangle} |\mathcal{G}_t \Big] \\
&= \sum_{i=1}^{N} a^i_{t+1}(\theta_t) \mathbf{e}_i \\
&= \sum_{i=1}^{N} \langle \mathbf{A}(\theta_t)\mathbf{X}_t, \mathbf{e}_i \rangle \, \mathbf{e}_i = \mathbf{A}(\theta_t)\mathbf{X}_t \ .
\end{aligned}
$$

□

Consequently, from Theorem 3.1, under $P^{\gamma,\mathbf{A}}$,

$$\begin{aligned} Z_t &= \gamma_t(\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_t) + w_t \ , \\ \mathbf{X}_{t+1} &= \mathbf{A}(\theta_t)\mathbf{X}_t + \mathbf{M}_{t+1} \ . \end{aligned}$$

They are the real-world dynamics of the Z process and the chain $\mathbf{X}$ in Section 2.

4 Unnormalized Conditional Probabilities

In this section, we derive a recursive formula for the unnormalized conditional probabilities of the whole path of the hidden Markov chain $\mathbf{X}$. Firstly, by a version of the Bayes's rule,

$$\begin{aligned} & P^{\gamma,\mathbf{A}}(\mathbf{X}_0 = \mathbf{e}_{i_0}, \mathbf{X}_1 = \mathbf{e}_{i_1}, \cdots, \mathbf{X}_t = \mathbf{e}_{i_t} | \mathcal{F}_t^Z) \\ = \ & \mathrm{E}^{\gamma,\mathbf{A}}[\langle \mathbf{X}_0, \mathbf{e}_{i_0} \rangle \langle \mathbf{X}_1, \mathbf{e}_{i_1} \rangle \cdots \langle \mathbf{X}_t, \mathbf{e}_{i_t} \rangle | \mathcal{F}_t^Z] \\ = \ & \frac{\bar{\mathrm{E}}[\Lambda_t^{\gamma,\mathbf{A}} \langle \mathbf{X}_0, \mathbf{e}_{i_0} \rangle \langle \mathbf{X}_1, \mathbf{e}_{i_1} \rangle \cdots \langle \mathbf{X}_t, \mathbf{e}_{i_t} \rangle | \mathcal{F}_t^Z]}{\bar{\mathrm{E}}[\Lambda_t^{\gamma,\mathbf{A}} | \mathcal{F}_t^Z]} \end{aligned}$$

Define, for each $t = 0, 1, \cdots, T$ and any admissible control process θ,

$$q_t^\theta(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}) := \bar{\mathrm{E}}[\Lambda_t^{\gamma,\mathbf{A}} \langle \mathbf{X}_0, \mathbf{e}_{i_0} \rangle \langle \mathbf{X}_1, \mathbf{e}_{i_1} \rangle \cdots \langle \mathbf{X}_t, \mathbf{e}_{i_t} \rangle | \mathcal{F}_t^Z] \ ,$$

so $q_t^\theta(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t})$ is the unnormalized conditional probability that $\mathbf{X}_0 = \mathbf{e}_{i_0}$, $\mathbf{X}_1 = \mathbf{e}_{i_1}$, $\cdots$, $\mathbf{X}_t = \mathbf{e}_{i_t}$ given $\mathcal{F}_t^Z$ associated with the control process θ. Indeed, q_t^θ is a positive, not necessarily normalized, measure on the product space $\mathcal{E}^{\otimes(t+1)}$, the $(t+1)$-fold product of the canonical state space $\mathcal{E} := \{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_N\}$ of the chain $\mathbf{X}$. The following theorem gives a recursion for q_t^θ.

Theorem 4.1. *For each $t = 1, 2, \cdots, T$,*

$$\begin{aligned} & q_t^\theta(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}) \\ = \ & \frac{\phi(Z_t - \gamma_t(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t})) a_t^{i_t}(\theta_{t-1})}{\phi(Z_t)} q_{t-1}^\theta(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{t-1}}) \ . \end{aligned}$$

Proof.

$$
\begin{aligned}
& q_t^\theta(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}) \\
= \ & \bar{\mathrm{E}}[\Lambda_t^{\gamma,\mathbf{A}} \langle \mathbf{X}_0, \mathbf{e}_{i_0}\rangle \langle \mathbf{X}_1, \mathbf{e}_{i_1}\rangle \cdots \langle \mathbf{X}_t, \mathbf{e}_{i_t}\rangle \,|\mathcal{F}_t^Z] \\
= \ & \bar{\mathrm{E}}\Bigg[\Lambda_{t-1}^{\gamma,\mathbf{A}} \frac{\phi(Z_t - \gamma_t(\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_t))}{\phi(Z_t)} \prod_{i=1}^{N} (N a^i(\theta_{t-1}))^{\langle \mathbf{X}_t, \mathbf{e}_i\rangle} \\
& \times \langle \mathbf{X}_0, \mathbf{e}_{i_0}\rangle \langle \mathbf{X}_1, \mathbf{e}_{i_1}\rangle \cdots \langle \mathbf{X}_t, \mathbf{e}_{i_t}\rangle \,|\mathcal{F}_t^Z\Bigg] \\
= \ & \bar{\mathrm{E}}\Bigg[\Lambda_{t-1}^{\gamma,\mathbf{A}} \langle \mathbf{X}_0, \mathbf{e}_{i_0}\rangle \langle \mathbf{X}_1, \mathbf{e}_{i_1}\rangle \cdots \langle \mathbf{X}_t, \mathbf{e}_{i_t}\rangle \,|\mathcal{F}_t^Z\Bigg] \frac{\phi(Z_t - \gamma_t(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}))}{\phi(Z_t)} \\
& \times N a^{i_t}(\theta_{t-1}) \\
= \ & \bar{\mathrm{E}}\Bigg\{\Lambda_{t-1}^{\gamma,\mathbf{A}} \langle \mathbf{X}_0, \mathbf{e}_{i_0}\rangle \langle \mathbf{X}_1, \mathbf{e}_{i_1}\rangle \cdots \big\langle \mathbf{X}_{t-1}, \mathbf{e}_{i_{t-1}}\big\rangle \,\bar{\mathrm{E}}\Bigg[\langle \mathbf{X}_t, \mathbf{e}_{i_t}\rangle \,|\mathcal{F}_t^Z \vee \mathcal{F}_{t-1}^{\mathbf{X}}\Bigg] |\mathcal{F}_t^Z\Bigg\} \\
& \times \frac{\phi(Z_t - \gamma_t(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}))}{\phi(Z_t)} N a^{i_t}(\theta_{t-1}) \\
= \ & \bar{\mathrm{E}}\Bigg[\Lambda_{t-1}^{\gamma,\mathbf{A}} \langle \mathbf{X}_0, \mathbf{e}_{i_0}\rangle \langle \mathbf{X}_1, \mathbf{e}_{i_1}\rangle \cdots \big\langle \mathbf{X}_{t-1}, \mathbf{e}_{i_{t-1}}\big\rangle \left(\frac{1}{N}\right) |\mathcal{F}_t^Z\Bigg] \\
& \times \frac{\phi(Z_t - \gamma_t(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}))}{\phi(Z_t)} N a^{i_t}(\theta_{t-1}) \\
= \ & \frac{\phi(Z_t - \gamma_t(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t})) a_t^{i_t}(\theta_{t-1})}{\phi(Z_t)} q_{t-1}^\theta(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{t-1}}) \ .
\end{aligned}
$$

□

Note that the future value q_{t+1}^θ only depends on the current value q_t^θ, but not the past values of q^θ. Consequently, we shall use q^θ as the information state variables in the separated form to be defined in the next section.

5 The Control Problem

Consider, for each $t = 0, 1, 2, \cdots, T$, measurable, bounded functions $L_t := L_t(\mathbf{X}_t, \theta_t)$. For any admissible control θ, we consider the following expected cost functional of θ:

$$J(\theta) := \mathrm{E}^{\gamma,\mathbf{A}}\left[\sum_{t=0}^{T} L_t\right]$$

For any two functions $f : \mathcal{E}^{\otimes(t+1)} \to \Re^+$ and $q : \mathcal{E}^{\otimes(t+1)} \to \Re^+$, write

$$\begin{aligned} & (f(\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_t) \odot g(\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_t)) \\ := & \sum_{i_0=1}^{N} \sum_{i_1=1}^{N} \cdots \sum_{i_t=1}^{N} f(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}) q(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}) \ . \end{aligned}$$

Then,

$$\begin{aligned} J(\theta) &:= \mathrm{E}^{\gamma,\mathbf{A}}\left[\sum_{t=0}^{T} L_t\right] \\ &= \bar{\mathrm{E}}\left[\Lambda_T^{\gamma,\mathbf{A}} \sum_{t=0}^{T} L_t\right] \\ &= \bar{\mathrm{E}}\left[\sum_{t=0}^{T} \Lambda_t^{\gamma,\mathbf{A}} L_t\right] \\ &= \sum_{t=0}^{T} \bar{\mathrm{E}}\{\bar{\mathrm{E}}[\Lambda_t^{\gamma,\mathbf{A}} L_t | \mathcal{F}_t^Z]\} \end{aligned}$$

Note that

$$
\begin{aligned}
& \bar{\mathrm{E}}[\Lambda_t^{\gamma,\mathbf{A}} L_t | \mathcal{F}_t^Z] \\
= & \sum_{i_0=1}^{N}\sum_{i_1=1}^{N}\cdots\sum_{i_t=1}^{N} \bar{\mathrm{E}}[\Lambda_t^{\gamma,\mathbf{A}} L_t \langle \mathbf{X}_0, \mathbf{e}_{i_0}\rangle \langle \mathbf{X}_1, \mathbf{e}_{i_1}\rangle \cdots \langle \mathbf{X}_t, \mathbf{e}_{i_t}\rangle | \mathcal{F}_t^Z] \\
= & \sum_{i_0=1}^{N}\sum_{i_1=1}^{N}\cdots\sum_{i_t=1}^{N} \bar{\mathrm{E}}[\Lambda_t^{\gamma,\mathbf{A}} \langle \mathbf{X}_0, \mathbf{e}_{i_0}\rangle \langle \mathbf{X}_1, \mathbf{e}_{i_1}\rangle \cdots \langle \mathbf{X}_t, \mathbf{e}_{i_t}\rangle | \mathcal{F}_t^Z] L_t(\mathbf{e}_{i_t}, \theta_t) \\
= & \sum_{i_0=1}^{N}\sum_{i_1=1}^{N}\cdots\sum_{i_t=1}^{N} q_t^{\theta}(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}) L_t(\mathbf{e}_{i_t}, \theta_t) \\
= & \sum_{i_0=1}^{N}\sum_{i_1=1}^{N}\cdots\sum_{i_t=1}^{N} q_t^{\theta}(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}) L_t(\mathbf{e}_{i_t}, \theta_t) \\
= & (q_t^{\theta}(\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_t) \odot L_t(\mathbf{X}_t, \theta_t)) \ .
\end{aligned}
$$

Consequently,

$$
J(\theta) = \sum_{t=0}^{T} \bar{\mathrm{E}}[(q_t^{\theta}(\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_t) \odot L_t(\mathbf{X}_t, \theta_t))]
$$

From Theorem 4.1,

$$
\begin{aligned}
& q_t^{\theta}(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}) \\
= & \frac{\phi(Z_t - \gamma_t(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t})) a_t^{i_t}(\theta_{t-1})}{\phi(Z_t)} q_{t-1}^{\theta}(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{t-1}}) \ .
\end{aligned} \tag{5.1}
$$

This is taken as the new information state variable with dynamics given by (5.1) so that the control problem is now represented in a separated form.

In what follows, we give a dynamic programming result and a minimum principle for the control problem.

The value function for the control problem is as follows:

$$
\begin{aligned}
V(t,q) \ &:= \ \bigwedge_{\theta \in \underline{\Theta}(t,T-1)} \bar{\mathrm{E}}\Big[\sum_{k=t}^{T} (q_k^{\theta}(\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_k) \odot L_k(\mathbf{X}_k, \theta_k)) | q_t^{\theta} = q\Big] \\
&:= \ \bigwedge_{\theta \in \underline{\Theta}(t,T-1)} V(t,q,\theta) \ ,
\end{aligned}
$$

so the value function is defined as the essential infimum under $\bar{P}$. We set

$$V(T, q) := \bar{\mathrm{E}}[(q \odot L_T(\mathbf{X}_T, \theta_T))] \ .$$

The following theorem gives the dynamic programming principle for the control problem.

Theorem 5.1. *For each $t = 1, 2, \cdots, T$, let*

$$\alpha_t := \frac{\phi(Z_t - \gamma_t(\mathbf{e}_{i_0}, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}))a_t^{i_t}(\theta_{t-1})}{\phi(Z_t)} \ .$$

Then the value functions $V(t, q)$, $t = 0, 1, \cdots, T-1$, satisfy the following backward recursion:

$$\begin{aligned} V(t, q) &= \bigwedge_{\theta \in \underline{\Theta}(t,t)} \bar{E}[(q_t^\theta(\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_t) \odot L_t(\mathbf{X}_t, \theta_t)) \\ &\quad + V(t+1, q_{t+1}^\theta(\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{X}_{t+1}))|q_t^\theta = q] \\ &= \bigwedge_{\theta \in \underline{\Theta}(t,t)} \bar{E}[(q \odot L_t(\mathbf{X}_t, \theta_t)) + V(t+1, q\alpha_{t+1})|q_t^\theta = q] \ , \end{aligned} \tag{5.2}$$

with terminal condition:

$$V(T, q) := \bar{E}[(q \odot L_T(\mathbf{X}_T, \theta_T))] \ .$$

Proof.

$$
\begin{aligned}
V(t,q) &:= \bigwedge_{\theta\in\underline{\Theta}(t,T-1)} V(t,q,\theta)\\
&= \bigwedge_{\theta\in\underline{\Theta}(t,t)} \bigwedge_{\theta\in\underline{\Theta}(t+1,T-1)} V(t,q,\theta)\\
&= \bigwedge_{\theta\in\underline{\Theta}(t,t)} \bigwedge_{\theta\in\underline{\Theta}(t+1,T-1)} \bar{\mathrm{E}}\Big[\sum_{k=t}^{T}(q_k^\theta(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_k)\odot L_k(\mathbf{X}_k,\theta_k))|q_t^\theta=q\Big]\\
&= \bigwedge_{\theta\in\underline{\Theta}(t,t)} \bigwedge_{\theta\in\underline{\Theta}(t+1,T-1)} \bar{\mathrm{E}}\Big\{\bar{\mathrm{E}}\Big[(q_t^\theta(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_t)\odot L_t(\mathbf{X}_t,\theta_t))\\
&\quad + \sum_{k=t+1}^{T}(q_k^\theta(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_k)\odot L_k(\mathbf{X}_k,\theta_k))|\mathcal{F}_{t+1}^Z\Big]|q_t^\theta=q\Big\}\\
&= \bigwedge_{\theta\in\underline{\Theta}(t,t)} \Big\{\bar{\mathrm{E}}\Big[(q_t^\theta(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_t)\odot L_t(\mathbf{X}_t,\theta_t))|q_t^\theta=q\Big]\\
&\quad + \bigwedge_{\theta\in\underline{\Theta}(t+1,T-1)} \bar{\mathrm{E}}\Big\{\bar{\mathrm{E}}\Big[\sum_{k=t+1}^{T}(q_k^\theta(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_k)\odot L_k(\mathbf{X}_k,\theta_k))|\mathcal{F}_{t+1}^Z\Big]|q_t^\theta=q\Big\}
\end{aligned}
$$

By the lattice property for the controls, (see, for example, Elliott et al. [2], Lemma 16.14 therein), the inner minimization and first expectation can be interchanged, so this is

$$
\begin{aligned}
&= \bigwedge_{\theta\in\underline{\Theta}(t,t)} \Big\{\bar{\mathrm{E}}\Big[(q_t^\theta(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_t)\odot L_t(\mathbf{X}_t,\theta_t))|q_t^\theta=q\Big]\\
&\quad +\bar{\mathrm{E}}\Big\{\bigwedge_{\theta\in\underline{\Theta}(t+1,T-1)} \bar{\mathrm{E}}\Big[\sum_{k=t+1}^{T}(q_k^\theta(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_k)\odot L_k(\mathbf{X}_k,\theta_k))|\mathcal{F}_{t+1}^Z\Big]|q_t^\theta=q\Big\}\\
&= \bigwedge_{\theta\in\underline{\Theta}(t,t)} \Big\{\bar{\mathrm{E}}\Big[(q_t^\theta(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_t)\odot L_t(\mathbf{X}_t,\theta_t))|q_t^\theta=q\Big]\\
&\quad +\bar{\mathrm{E}}\Big[V(t+1,q_{t+1}^\theta)|q_t^\theta=q\Big]\Big\}
\end{aligned}
$$

By Theorem 4.1,

$$q_{t+1}^\theta=\alpha_{t+1}q_t^\theta\ .$$

Consequently,

$$\begin{aligned} & V(t,q) \\ = & \bigwedge_{\theta\in\underline{\Theta}(t,t)} \left\{ \bar{\mathrm{E}}\left[(q_t^\theta(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_t)\odot L_t(\mathbf{X}_t,\theta_t)) | q_t^\theta = q \right] \right. \\ & \left. +\bar{\mathrm{E}}\left[V(t+1,\alpha_{t+1}q_t^\theta) | q_t^\theta = q \right] \right\} \\ = & \bigwedge_{\theta\in\underline{\Theta}(t,t)} \bar{\mathrm{E}}[(q\odot L_t(\mathbf{X}_t,\theta_t)) + V(t+1,q\alpha_{t+1})|q_t^\theta = q] \ . \end{aligned}$$

□

A control process $\theta\in\underline{\Theta}(0,T-1)$ is said to be separated if θ_t depends on $(Z_0,Z_1,\cdots,Z_t)$ only through the information state q_t^θ. Write $\underline{\Theta}_s(0,T-1)$ for the space of separated controls. Then we have the following theorem.

Lemma 5.1. *For each $t=0,1,\cdots,T-1$,*

$$V(t,q) = \bigwedge_{\theta\in\underline{\Theta}_s(0,T-1)} V(t,q,\theta) \ .$$

Proof. We prove the result by backward induction in t. Firstly, it is clear that

$$\begin{aligned} V(T,q) &= \bigwedge_{\theta\in\underline{\Theta}(T,T)} V(T,q,\theta) \\ &= \bigwedge_{\theta\in\underline{\Theta}(T,T)} \bar{\mathrm{E}}[(q\odot L_T(\mathbf{X}_T,\theta))] \\ &= \bigwedge_{\theta\in\underline{\Theta}_s(T,T)} \bar{\mathrm{E}}[(q\odot L_T(\mathbf{X}_T,\theta))] \ , \end{aligned}$$

so the result holds for $t=T$.

Then by Formula (5.2),

$$V(t,q) \ = \bigwedge_{\theta\in\underline{\Theta}(t,t)} \bar{\mathrm{E}}[(q\odot L_t(\mathbf{X}_t,\theta_t)) + V(t+1,q\alpha_{t+1})|q_t^\theta = q]$$

Clearly, a minimizing θ_k, (or a sequence of minimizing θ_k's), depends only on the information $q_t^\theta = q$. Consequently,

$$\begin{aligned} V(t,q) &= \bigwedge_{\theta\in\underline{\Theta}_s(t,t)} \bar{\mathrm{E}}\Big[(q \odot L_t(\mathbf{X}_t,\theta_t)) + \bigwedge_{\theta\in\underline{\Theta}_s(t+1,T-1)} V(t+1,q\alpha_{t+1})|q_t^\theta = q\Big] \\ &= \bigwedge_{\theta\in\underline{\Theta}_s(t,T-1)} V(t,q,\theta)\ . \end{aligned}$$

□

Then we have a minimum principle of the form presented in the following theorem.

Theorem 5.2. *Suppose θ^* is a control such that, for each measure $q_t^\theta(\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_t)$, θ^* achieves the minimum in (5.2). Then,*

$$V(t,q,\theta^*) = V(t,q)\ ,$$

and θ^* is an optimal control.

Proof. Again we use backward induction in t. It is clear that

$$V(T,q,\theta^*) = \bar{\mathrm{E}}[\langle q, L_T(\mathbf{X}_T,\theta^*)\rangle] = V(T,q)\ .$$

Suppose the statement is true for $k = t+1, t+2, \cdots, T$. We wish to prove that it is also true for $k = t$. Then

$$\begin{aligned} V(t,q,\theta_t^*) &= \bar{\mathrm{E}}[(q, L_t(\mathbf{X}_t,\theta_t^*)) + V(t+1, q_{t+1}^{\theta^*}, \theta^*)|q_t^\theta = q] \\ &= \bar{\mathrm{E}}[(q, L_t(\mathbf{X}_t,\theta_t^*)) + V(t+1, q_{t+1}^{\theta^*})|q_t^\theta = q] \\ &= V(t,q)\ . \end{aligned}$$

Now for any other $\theta \in \underline{\Theta}(0,T)$,

$$V(t,q,\theta^*) = V(t,q) \le V(t,q,\theta)\ .$$

In particular, this holds true when $t = 0$. Consequently,

$$V(0,q,\theta^*) \le V(0,q,\theta)\ ,$$

so the statement is true. □

Acknowledgement: The authors would like to thank Australian Research Council, (ARC), for their support.

References

[1] P.E. Caines, Linear Stochastic Systems, Wiley, New York, 1988.

[2] R.J. Elliott, Stochastic Calculus and Applications, Springer, Berlin-Heidelberg-New York, 1982.

[3] R.J. Elliott, L. Aggoun, J.B. Moore, Hidden Markov Models: Estimation and Control, Springer, Berlin-Heidelberg-New York, 1995.

[4] R.J. Elliott, J. Deng, A filter for a hidden Markov chain observed in fractional Gaussian noise, Haskayne School of Business, University of Calgary, Preprint, 2009.

[5] R.J. Elliott, M. Hong, Fractional differencing in discrete time, Haskayne School of Business, University of Calgary. Preprint, 2009.

[6] P.R. Kumar, P.P. Varaiya, Stochastic Systems, Prentice Hall, Englewood Cliffs, NJ, 1986.

[7] B.B. Mandelbrot, J.W. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM Review 10 (1968) 422437.

[8] B.B. Mandelbrot, Fractals: Form, Chance, and Dimensions, Free Press, New York, 1977.