The filtering equations revisited

Thomas Cass* Martin Clark† Dan Crisan‡

Abstract

The problem of nonlinear filtering has engendered a surprising number of mathematical techniques for its treatment. A notable example is the change-of–probability-measure method originally introduced by Kallianpur and Striebel to derive the filtering equations and the Bayes-like formula that bears their names. More recent work, however, has generally preferred other methods. In this paper, we reconsider the change-of-measure approach to the derivation of the filtering equations and show that many of the technical conditions present in previous work can be relaxed. The filtering equations are established for general Markov signal processes that can be described by a martingale-problem formulation. Two specific applications are treated. As part of the analysis, a criterion, of some independent interest, is given for an exponential local martingale to be a martingale, that is more general than the criteria of Kazamaki and Novikov.

Keywords: Measure Valued Processes; Non-Linear Filtering, Kallianpur-Striebel Formula, Change of Probability Measure Method, Kazamaki criterion.

1 Introduction

The aim of nonlinear filtering is to estimate an evolving dynamical system, customarily modelled by a stochastic process and called the signal process. The signal process cannot be measured directly, but only via a related process, termed the observation process. The filtering problem consists in computing the conditional distribution of the signal at the current time given the observation data accumulated up to that time. In order to describe the contribution of the paper, we start with a few historical comments on the subject.

The development of the modern theory of nonlinear filtering started in the sixties with the publications of Stratonovich [25, 26], Kushner [11, 12] and Shiryaev [24] for diffusions and Wonham for pure-jump Markov processes [27]; these introduced the basic form of the class of stochastic differential equations for the conditional distributions of partially observed Markov processes, which are now known generically as the filtering equation. This class of equations has inspired authors to introduce a rich variety of mathematical techniques to justify their structure, together with that of their unnormalized form, the Zakai (or Duncan-Mortensen-Zakai) equation, [7, 20, 30], and to establish the existence, uniqueness and regularity of their solutions. A description of much of the work on this equation and its generalizations can be found in [10] for papers before 1980, in [13, 14] for papers before 2000 and in [2, 5, 28] for more recent work.

For instance, Fujisaki, Kallianpur and Kunita [8] exploited a stochastic-integral representation theorem in order to enable them to express conditional distributions as functionals of an "innovations" martingale (a concept introduced in the Gaussian case by Kailath [17]). Krylov, Rozovsky, Pardoux [15, 16, 21], Chapter 6 in [5] and other authors developed a general theory of stochastic partial differential equations that led to a direct 'PDE' approach to the filtering equations, but there are

*Department of Mathematics, Imperial College London.
†Department of Electrical and Electronic Engineering, Imperial College London.
‡Department of Mathematics, Imperial College London.

many other approaches For example, see the work of Grigelionis and Mikulevicius on filtering for signal and observation processes with jumps [4, Chapter 4] and that of Kurtz and Nappo on the filtered martingale problem [4 Chapter 5]

A probabilistic approach, initially considered formally by Bucy [3], but developed in detail by Kallianpur and Striebel [18, 19], made use of a functional form of Bayes formula for processes, now known as the Kallianpur-Striebel formula. This technique, which is based on a change of probability measure that makes, at each time, the future observation process independent of past processes, is effective for filtering problems in which the observation process is of the "signal plus white noise" variety, where the signal is independent of the noise process, but less so for the "correlated case"; that is, for problems in which observed and unobserved components are coupled via a common noise process. For this reason, among probabilistic methods, the "innovations" approach is often preferred to the "change of measure" method. The awkwardness in its application results from the fact that an exponential local martingale, constructed via Girsanov's theorem as a process of potential densities, has to be verified as a true martingale, and this is generally requires ad hoc techniques peculiar to the particular filtering problem being considered.

In this paper we re-visit the change-of-measure method and show that it can be used to derive the filtering equations for a broad class of Markov processes with coupled observed and unobserved components. This class includes diffusions with jumps obeying only mild linear growth conditions on their characteristic coefficients. Propositions are also presented that serve to test whether the filtering equations are derivable by the change-of-measure method for a particular filtering problem.

As part of the analysis we give a criterion, of some independent interest, under which a continuous exponential local martingale is a genuine martingale. This is more general than the Kazamaki and Novikov criteria for such exponential local martingales, in the sense that it is implied by them and can be applied in cases where they are inapplicable.

2 The Filtering Framework

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space together with a filtration $(\mathcal{F}_t)_{t\geq 0}$ which satisfies the usual conditions[1]. On $(\Omega, \mathcal{F}, \mathbb{P})$ we consider an $\mathcal{F}_t$-adapted process $\bar{X}$ with càdlàg paths. The process $\bar{X}$ consists in a pair of processes X and Y, $\bar{X} = (X, Y)$. The process X is called the *signal* process and is assumed to take values in a complete separable metric space $\mathbb{S}$ (the state space). The process Y is assumed to take values in $\mathbb{R}^m$ and is called the *observation* process.

Let $\mathcal{B}(\mathbb{S} \times \mathbb{R}^m)$ be the associated product Borel σ-algebra on $\mathbb{S} \times \mathbb{R}^m$ and $b\mathcal{B}(\mathbb{S} \times \mathbb{R}^m)$ be the space of bounded $\mathcal{B}(\mathbb{S} \times \mathbb{R}^m)$-measurable functions. Let $A\colon b\mathcal{B}(\mathbb{S} \times \mathbb{R}^m) \to b\mathcal{B}(\mathbb{S} \times \mathbb{R}^m)$ and write $\mathcal{D}(A) \subseteq b\mathcal{B}(\mathbb{S} \times \mathbb{R}^m)$ for the domain of A. We assume that $\mathbf{1} \in \mathcal{D}(A)$ and $A\mathbf{1} = 0$. In the following we will assume that the distribution of X_0 is $\pi_0 \in \mathcal{P}(\mathbb{S})$ and that $Y_0 = 0$. Since $Y_0 = 0$, the initial distribution of X, is identical with the conditional distribution of X_0 given $\mathcal{Y}_0$ and we use the same notation for both. Further we will assume that $\bar{X}$ is a solution of the martingale problem for $(A, \pi_0 \times \delta_0)$. In other words, we assume that the process $M^\varphi = \{M_t^\varphi,\ t \geq 0\}$ defined as

$$M_t^\varphi = \varphi(\bar{X}_t) - \varphi(\bar{X}_0) - \int_0^t A\varphi(\bar{X}_s)\mathrm{d}s, \qquad t \geq 0, \tag{1}$$

is an $\mathcal{F}_t$-adapted martingale for any $\varphi \in \mathcal{D}(A)$. In addition, let $h = (h_i)_{i=1}^m : \mathbb{S} \to \mathbb{R}^m$ be a measurable function such that

$$P\left(\int_0^t \left|h^i(\bar{X}_s)\right|^2 ds < \infty\right) = 1. \tag{2}$$

[1]The probability space $(\Omega, \mathcal{F}, \mathbb{P})$ together with the filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfies the usual conditions provided: a. $\mathcal{F}$ is complete i.e. $A \subset B$, $B \in \mathcal{F}$ and $\mathbb{P}(B) = 0$ implies that $A \in \mathcal{F}$ and $\mathbb{P}(A) = 0$, b. The filtration $\mathcal{F}_t$ is right continuous i.e. $\mathcal{F}_t = \mathcal{F}_{t+}$. c. $\mathcal{F}_0$ (and consequently all $\mathcal{F}_t$ for $t \geq 0$) contains all the $\mathbb{P}$-null sets.

for all $t \geq 0$. Let W be a standard $\mathcal{F}_t$-adapted m-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P})$. We will assume that Y satisfies the following evolution equation

$$Y_t = \int_0^t h(\bar{X}_s)\,\mathrm{d}s + W_t. \tag{3}$$

To complete the description we need to identify the covariation process between $M^\varphi = \{M_t^\varphi,\ t \geq 0\}$ and W. For this we introduce m operators $B^i\colon b\mathcal{B}(\mathbb{S}\times\mathbb{R}^m) \to b\mathcal{B}(\mathbb{S}\times\mathbb{R}^m)$, $i = 1, ..., m$ with $\mathcal{D}(A) \subseteq \mathcal{D}(B^i) \subseteq b\mathcal{B}(\mathbb{S}\times\mathbb{R}^m)$. We assume that $\mathbf{1} \in \mathcal{D}(A)$ and $A\mathbf{1} = 0$. We will assume that,

$$\left\langle M^\varphi, W^i \right\rangle_t = \int_0^t B^i\varphi\left(\bar{X}_s\right)ds + \int_0^t \frac{\partial\varphi}{\partial y_i}\left(\bar{X}_s\right)ds, \tag{4}$$

for any $t \geq 0$ and for test functions φ both in the domain of A and with bounded partial derivatives in the y direction. In particular, for functions that are constant in the second component, then we have

$$\left\langle M^\varphi, W \right\rangle_t = \int_0^t B^i\varphi\left(X_s, Y_s\right)ds. \tag{5}$$

Let $\{\mathcal{Y}_t,\ t \geq 0\}$ be the usual augmentation of the filtration associated with the process Y, viz

$$\mathcal{Y}_t = \bigcap_{\varepsilon>0}\sigma(Y_s,\ s\in[0, t+\varepsilon]) \vee \mathcal{N}, \quad \mathcal{Y} = \bigvee_{t\in\mathbb{R}_+}\mathcal{Y}_t. \tag{6}$$

where $\mathcal{N}$ is that class of all $\mathbb{P}$-null sets. Note that Y_t is $\mathcal{F}_t$-adapted, hence $\mathcal{Y}_t \subset \mathcal{F}_t$. In the following we will assume that $\mathcal{Y}_t$ is a right continuous filtration.

Definition 1 *The filtering problem consists in determining the conditional distribution π_t of the signal X at time t given the information accumulated from observing Y in the interval $[0, t]$; that is, for $\varphi \in b\mathcal{B}(\mathbb{S})$, computing*

$$\pi_t(\varphi) = \mathbb{E}[\varphi(X_t) \mid \mathcal{Y}_t]. \tag{7}$$

There exists a suitable regularisation of the process $\pi = \{\pi_t,\ t \geq 0\}$, so that π_t is an optional $\mathcal{Y}_t$-adapted probability measure-valued process for which (7) holds almost surely[2]. In addition, since $\mathcal{Y}_t$ is right-continuous, it follows that π has a cadlag version (see Corollary 2.26 in [2]). In the following, we take π to be this version.

In the following we deduce the evolution equation for π. A new measure is constructed under which Y becomes a Brownian motion and π has a representation in terms of an associated unnormalised version ρ. This ρ is then shown to satisfy a linear evolution equation which leads to the evolution equation for π by an application of Itô's formula.

2.1 Preliminary Results

Definition 2 *We define $H^1(\mathbb{P})$ to be the set of càdlàg real-valued $\mathcal{F}_t$-martingales $M = \{M_t,\ t \geq 0\}$ such that the associated process $M^* = \{M_t^*,\ t \geq 0\}$ defined as $M_t^* := \sup_{0\leq s\leq t}|M_s|$ for $t \geq 0$ is a submartingale. In particular, $\mathbb{E}[M_t] < \infty$. for any $t \geq 0$.*

Remark 3 *$H^1(\mathbb{P})$ together with the distance function*

$$d(M, N) := \sum_{n=1}^\infty \frac{1}{2^n}\min\left(\mathbb{E}\left[(M-N)_n^*\right], 1\right)$$

[2] See Theorem 2.1 in [2].

is a Fréchet space with translation invariant metric. Suppose $(W_t)_{t\geq 0}$ is an $\mathbb{R}^d-$valued Brownian motion and $H=(H^i)_{i=1}^d$ is an $\mathcal{F}_t$-adapted measurable $\mathbb{R}^d$-valued process such that

$$P\left(\int_0^t |H_s|^2\, ds<\infty\right)=1. \tag{8}$$

Define $Z=(Z_t)_{t\geq 0}$ to be the exponential local martingale[3]

$$Z_t=\exp\left(\int_0^t H_s^\top dW_s-\frac{1}{2}\int_0^t |H_s|^2\, ds\right),$$

where $\int_0^t H_s^\top dW_s:=\sum_{i=1}^d\int_0^t H_s^i dW_s^i$.

Lemma 4 (The $Z\log Z$ lemma) *For any $t\geq 0$ we have*

$$\sup_{\tau\in\mathcal{T}_t}\mathbb{E}\left[Z_\tau\log Z_\tau\right]=\frac{1}{2}\mathbb{E}\left[\int_0^t Z_s\left|H_s\right|^2 ds\right]\in[0,\infty], \tag{9}$$

where $\mathcal{T}_t$ is the set of $(\mathcal{F}_t)$-stopping times bounded by t. If furthermore the terms in (9) are finite, then they are both equal to $\mathbb{E}\left[Z_t\log Z_t\right]$. We also have

$$\mathbb{E}\left[Z_t^*\right]\leq\frac{e+1}{e-1}+\frac{e}{2\,(e-1)}\mathbb{E}\left[\int_0^t Z_s\left|H_s\right|^2 ds\right]\in[0,\infty]. \tag{10}$$

As an immediate consequence of this lemma we have

Corollary 5 *If the terms in (9) are finite,then $(Z_t)_{t\geq 0}$ is a genuine martingale, uniformly integrable over any finite interval $[0,t]$, that belongs to $H^1(\mathbb{P})$.*

Proof. Let $L_t:=Z_t\log Z_t$ for $t\geq 0$. If we assume that $\sup_{\tau\in\mathcal{T}_t}\mathbb{E}\left[L_\tau\right]$ is finite, then for all $K\geq e$

$$\sup_{\tau\in\mathcal{T}_t}\mathbb{E}\left[|Z_\tau|\,1_{\{|Z_\tau|\geq K\}}\right]\leq\sup_{\tau\in\mathcal{T}_t}\frac{\mathbb{E}\left[|Z_\tau|\log Z_\tau 1_{\{|Z_\tau|\geq K\}}\right]}{\log K}\leq\frac{1}{\log K}\left(\sup_{\tau\in\mathcal{T}_t}\mathbb{E}\left[L_\tau\right]+e^{-1}\right)$$

the right hand side of which tends to zero as $K\to\infty$. Hence the family random variables

$$\{Z_\tau:\tau\in\mathcal{T}_t\}$$

is uniformly integrable. Z is thus a martingale over $[0,t]$ and L, by Jensen's inequality, is a submartingale. Using $P\left(0<Z_t<\infty,\text{ for all } t<\infty\right)=1$ we have from Itô's formula that

$$L_t=\underbrace{\int_0^t\left(1+\log Z_s\right)Z_sH_s^\top dW_s}_{:=M_t}+\underbrace{\frac{1}{2}\int_0^t Z_s\left|H_s\right|^2 ds}_{:=A_t},$$

M is a local martingale, hence the stopped process $M_\cdot^{\sigma_n}:=M_{\cdot\wedge\sigma_n}$ is a martingale for some localising sequence $0\leq\sigma_n\leq\sigma_{n+1}\uparrow\infty$ as $n\to\infty$. For any $\tau\in\mathcal{T}_t$ we obtain

$$\mathbb{E}\left[L_\tau^{\sigma_n}\right]=\mathbb{E}\left[A_\tau^{\sigma_n}\right]\leq\mathbb{E}\left[L_\tau\right]\leq\mathbb{E}\left[L_t\right].$$

[3]Here and later if $a=(a_i)_{i=1}^d\in\mathbb{R}^d$, then $|a|^2=\sum_{i=1}^d a_i^2$. Hence, for example, in the expression for Z from $\int_0^t|H_s|^2\,ds=\sum_{i=1}^d\int_0^t\left(H_s^i\right)^2 ds$

Then, using Fatou's lemma[4] and the monotone convergence theorem, we have

$$\mathbb{E}\left[L_\tau\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[L_\tau^{\sigma_n}\right] = \liminf_{n\to\infty} \mathbb{E}\left[A_\tau^{\sigma_n}\right] = \mathbb{E}\left[A_\tau\right] \leq \mathbb{E}\left[L_\tau\right] \leq \mathbb{E}\left[L_t\right].$$

Finally taking the supremum over $\tau \in \mathcal{T}_t$ yields

$$\mathbb{E}\left[L_t\right] \leq \sup_{\tau\in\mathcal{T}_t} \mathbb{E}\left[L_\tau\right] \leq \sup_{\tau\in\mathcal{T}_t} \mathbb{E}\left[A_\tau\right] \leq \mathbb{E}\left[A_t\right] \leq \sup_{\tau\in\mathcal{T}_t} \mathbb{E}\left[L_\tau\right] \leq \mathbb{E}\left[L_t\right].$$

and the equality (9) holds in this case. If instead we know that $\mathbb{E}\left[A_t\right] < \infty$, then by defining the sequence of stopping times $(\tau_n)_{n=1}^{\infty}$, $0 \leq \tau_n \leq \tau_{n+1}$ by

$$\tau_n = \inf\left\{t \geq 0 : |Z_t| = \frac{1}{n} \text{ or } |Z_t| = n\right\}$$

we have

$$\mathbb{E}\left[M_{t\wedge\tau_n}^2\right] = \mathbb{E}\left[\int_0^{t\wedge\tau_n} \left(1+\log Z_s\right)^2 Z_s^2 \left|H_s\right|^2 ds\right] \leq 2n^2\left(1+\log n\right)^2 \mathbb{E}\left[A_t\right] < \infty.$$

From this we deduce that the stopped process $M_\cdot^{\tau_n} := M_{\cdot\wedge\tau_n}$ is a square-integrable martingale over $[0,t]$. Combining this with the fact that $A_{t\wedge\tau_n} \leq A_t$ yields

$$\mathbb{E}\left[L_{\tau\wedge\tau_n}\right] = \mathbb{E}\left[A_{\tau\wedge\tau_n}\right] \leq \mathbb{E}\left[A_t\right]$$

for any $\tau \in \mathcal{T}_t$. We notice that $\tau_n \uparrow \infty$, and hence $Z_{t\wedge\tau_n} \to Z_t$ a.s. as $n \to \infty$. Then applying Fatou's lemma and taking the supremum over all $\tau \in \mathcal{T}_t$ then gives that $\sup_{\tau\in\mathcal{T}_t} \mathbb{E}\left[L_\tau\right] \leq \mathbb{E}\left[A_t\right] < \infty$. The equality

$$\mathbb{E}\left[L_t\right] = \mathbb{E}\left[A_t\right] = \sup_{\tau\in\mathcal{T}_t} \mathbb{E}\left[L_\tau\right] \in [0,\infty)$$

then follows from the first part of the proof. It is clear from the argument that A_t is not integrable if and only if $\sup_{\tau\in\mathcal{T}_t} \mathbb{E}\left[L_\tau\right] = \infty$.

Turning attention to (10), we observe that the stopped process L^{τ_n} is a bounded submartingale, with a bounded martingale part given by M^{τ_n}. Hence, by a modification of a standard maximal inequality (see pg 52 in [22]), we deduce that

$$\begin{aligned}
\mathbb{E}\left[\left(Z^{\tau_n}\right)_t^*\right] &\leq \frac{e+1}{e-1} + \frac{e}{e-1}\mathbb{E}\left[L_{t\wedge\tau_n}\right] \\
&\leq \frac{e+1}{e-1} + \frac{e}{e-1}\mathbb{E}\left[A_{t\wedge\tau_n}\right].
\end{aligned}$$

The proof is finished by an application of the monotone convergence theorem. ■

Remark 6 (Kazamaki's criterion) *The criterion of locally finite transformed average energy:*

$$\mathbb{E}\left[\int_0^t Z_s \left|H_s\right|^2 ds\right] < \infty, \tag{11}$$

required in the lemma, turns out to be a criterion for the martingale nature of Z that is more generally applicable than Kazamaki's criterion: $\exp\left(\int_0^\cdot H_s^T dW_s\right)$ *is a submartingale*[5]*. To show that it is more*

[4] Which we may do since L is bounded from below by $-e^{-1}$.

[5] In turn, Kazamaki's criterion is more generally applicable that Novikov's criterion.

general we need to show (11) holds whenever Kazamaki's criterion holds. Assuming the latter therefore, we let $(\tau_n)_{n=1}^{\infty}$ *denote the localising sequence of stopping times defined*

$$\tau_n = \inf\left\{t \geq 0 : \int_0^t |H_s|^2\, ds \geq n\right\}$$

whence $Z_{\cdot\wedge\tau_n}$ *is a martingale and*

$$\begin{aligned}
& \mathbb{E}\left[\int_0^{t\wedge\tau_n} Z_s\, |H_s|^2\, ds\right] && (12)\\
= \ & \mathbb{E}\left[Z_{t\wedge\tau_n}\int_0^{t\wedge\tau_n} |H_s|^2\, ds\right] \\
= \ & \mathbb{E}\left[\exp\left(\int_0^{t\wedge\tau_n} H_s^T dW_s - \frac{1}{2}\int_0^{t\wedge\tau_n} |H_s|^2\, ds\right)\int_0^{t\wedge\tau_n} |H_s|^2\, ds\right] \\
\leq \ & \frac{2}{e}\mathbb{E}\left[\exp\left(\int_0^t H_s^T dW_s\right)\right] < \infty,
\end{aligned}$$

where the inequalities follow from the fact that $xe^{-x/2}$ *is bounded on* $(0,\infty)$ *by* $2e^{-1}$*, and Kazamaki's criterion. By applying the monotone convergence theorem to (12), we obtain*

$$\mathbb{E}\left[\int_0^t Z_s\, |H_s|^2\, ds\right] < \infty$$

for each $t < \infty$*.*

To show that condition (11) is strictly more general, in the sense that it can sometimes be applied when Kazamaki is unavailable, we can make use of a simple example introduced in Revuz and Yor [22] (page 366, Exercise ix.2.10.40) in which Kazamaki's criterion fails. Let W *be a scalar Brownian motion with* $W_0 = 0$ *and set* $H_t = \alpha W_t$ *for some* $\alpha > 0$*. As Revuz and Yor point out,* $Z_\cdot = \exp\left(\alpha\int_0^\cdot W_s dW_s - \frac{\alpha^2}{2}\int_0^\cdot W_s^2 ds\right)$ *is a true martingale on* $[0,\infty)$ *for all* α*, but* $\exp\left(\alpha\int_0^t W_s dW_s\right)$ *ceases to be a submartingale for* $t \geq \alpha^{-1}$*. However, under the transformed probability measure* $\tilde{\mathbb{P}}$*, defined on the* σ*-ring* $\cup_{t\geq 0}\mathcal{F}_t$ *by*

$$\left.\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = Z_t,$$

W *is turned into a Gaussian semimartingale satisfying*

$$W_t = \int_0^t \alpha W_s ds + B_t$$

for some $\left(\{\mathcal{F}_t\}_{t\geq 0}, \tilde{\mathbb{P}}\right)$ *Brownian motion* B*. But* W *can also be expressed as*

$$W_t = \int_0^t e^{\alpha(t-s)} dB_s$$

and then it is straightforward to show that for all $t \geq 0$

$$\mathbb{E}\left[\int_0^t Z_s H_s^2 ds\right] = \tilde{\mathbb{E}}\left[\alpha^2\int_0^t W_s^2 ds\right] = \frac{1}{4}\left(e^{2\alpha t} - 2\alpha t - 1\right).$$

Hence the transformed average energy condition is applicable in this case.

Remark 7 *We record four observations:*

1. *The proof does not require the a priori assumption that* $\mathbb{E}\left[\int_0^t |H_s|^2\, ds\right] < \infty$. *However observe that*
$$\mathbb{E}\left[\int_0^t Z_s\, |H_s|^2\, ds\right] = \mathbb{E}\left[\int_0^t \mathbb{E}\left[Z_t|\mathcal{F}_s\right] |H_s|^2\, ds\right] = \mathbb{E}\left[Z_t \int_0^t |H_s|^2\, ds\right].$$

2. *If the Brownian motion* W *is independent of* H *then using the sequence of stopping times* $(\tau_n)_{n=1}^{\infty}$, $0 \le \tau_n \le \tau_{n+1}$ *by*
$$\tau_n = \inf\{t \ge 0 : |H_t| \ge n\},$$
we get that
$$\begin{aligned}\mathbb{E}\left[Z_{t\wedge\tau_n}|H\right] &= \mathbb{E}\left[\exp\left(\int_0^{t\wedge\tau_n} H_s^\top dW_s - \frac{1}{2}\int_0^{t\wedge\tau_n} |H_s|^2\, ds\right)\middle| H\right] \\ &= \exp\left(-\frac{1}{2}\int_0^{t\wedge\tau_n} |H_s|^2\, ds\right) \mathbb{E}\left[\exp\left(\int_0^{t\wedge\tau_n} H_s^\top dW_s\right)\middle| H\right] = 1.\end{aligned}$$
In particular, the stopped process Z^{τ_n} *is a martingale. Moreover*
$$\mathbb{E}\left[\int_0^{t\wedge\tau_n} Z_s\, |H_s|^2\, ds\middle| H\right] = \int_0^t \mathbb{E}\left[\, Z_{s\wedge\tau_n}|\, H\right] |H_{s\wedge\tau_n}|^2\, ds = \int_0^{t\wedge\tau_n} |H_s|^2\, ds.$$
Hence, by an application of the monotone convergence theorem
$$\mathbb{E}\left[\int_0^t Z_s\, |H_s|^2\, ds\right] = \mathbb{E}\left[\int_0^t |H_s|^2\, ds\right].$$
By the same argument one can prove directly that Z *is a martingale under the weaker condition (8). This result is contained in Lemma 11.3.1 of [10].*

3. *Assume that* $\mathbb{E}\left[A_t\right] < \infty$ *for all* $t \ge 0$, *then* $(Z-1)$ *is a zero-mean martingale and* $\mathbb{E}\left[(Z-1)_t^*\right] < 1 + \mathbb{E}\left[Z_t^*\right] < \infty$. *Since* $\langle Z-1\rangle_t = \int_0^t Z_s^2\, |H_s|^2\, ds$ *the Burkholder-Davis-Gundy inequalities gives*
$$\mathbb{E}\left[\left(\int_0^t Z_s^2\, |H_s|^2\, ds\right)^{1/2}\right] < \infty$$
for all $t \ge 0$.

4. *The finiteness of the transformed average energy does not imply that the average energy itself is finite. The following example illustrates this. Let* $W = (W_t)_{0\le t\le 1}$ *be a one-dimensional* $(\mathcal{F}_t)_{0\le t\le 1}$ *-adapted Brownian motion with* $W_0 = 0$, *and suppose that* $\mathcal{F}_0$ *carries a uniform* $[0,1]$ *random variable which is independent of* W. *Then we will prove there exists an* $(\mathcal{F}_t)$*-optional process* $H = (H_t)_{0\le t\le 1}$ *such that the local martingale* Z *given by*
$$Z_t = \exp\left(\int_0^t H_s dW_s - \frac{1}{2}\int_0^t H_s^2 ds\right)$$
is a martingale on $[0,1]$ *for which*
$$\mathbb{E}\left[\int_0^1 Z_s H_s^2 ds\right] < \infty \text{ and } \mathbb{E}\left[\int_0^1 H_s^2 ds\right] = \infty.$$

To construct Z we will make use of the Gaussian martingale $B_t = \int_0^t \frac{1}{1-s} dW_s$ defined on $[0,1)$. We notice that $(1-t)\, B_t$ is a Brownian bridge on $[0,1)$ and the related process $V_t := \exp\left[B_t - \frac{t}{2(1-t)}\right]$ is just the martingale of densities on $(\mathcal{F}_t)_{0 \le t < 1}$ that turns W into a Brownian bridge, cf. [22]. But the property we exploit is the existence of a Brownian motion $\bar{B}$ on $[0,\infty)$ such that $B_t = \bar{B}_{\ \sigma(t)}$ wherein $\sigma(t) := t\,(1-t)^{-1}$. Let

$$X_t = \int_0^t \frac{V_s ds}{(1-s)^2}$$

be defined on $[0,1]$ and introduce the sequence of stopping times

$$T_n = \inf\left\{ t \ge 0 : X_t = \frac{n\,(1-t)}{n\,(1-t)+t} \right\}.$$

Since $X_{\cdot}$ is non-negative and increasing with $X_0 = 0$ and the function $t \mapsto \frac{n(1-t)}{n(1-t)+t}$ is strictly decreasing to 0, each T_n is strictly less than one. Furthermore the sequence $(T_n)_{n=1}^{\infty}$ increases to a limit $T_\infty \le 1$. We need to prove that $\mathbb{P}(T_\infty = 1) > 0$. Using the fact that

$$\lim_{n\to\infty} \frac{n\,(1-t)}{n\,(1-t)+t} = 1 \textit{ for all } t < 1,$$

it follows that $\mathbb{P}(T_\infty = 1) = \mathbb{P}(X_1 < 1)$. However,

$$\begin{aligned} X_1 &= \int_0^1 \frac{1}{(1-t)^2} \exp\left[B_t - \frac{t}{2\,(1-t)}\right] dt \\ &= \int_0^\infty \exp\left(\bar{B}_t - \frac{1}{2}s\right) ds \end{aligned}$$

and it is result of Dufresne [6] (see also Yor [29], page 15) that this latter integral is distributed as twice the inverse of a standard exponential random variable Y. In particular $\mathbb{P}(X_1 < 1) = \mathbb{P}(Y > 2) = e^{-2}$, from which it follows that $\mathbb{P}(T_\infty = 1) > 0$ and, therefore, $\mathbb{E}\left[\frac{T_\infty}{1-T_\infty}\right] = \infty$. The monotone convergence theorem implies that the sequence

$$m(n) := \mathbb{E}\left[\frac{T_n}{1-T_n}\right] \uparrow \infty \textit{ as } n \to \infty.$$

Let U be the uniform $[0,1]$ random variable on $\mathcal{F}_0$ referred to earlier. We can construct, as a measurable function of U, an integer random variable N satisfying

$$\mathbb{E}[m(N)] = \infty.$$

If T denotes the stopping time T_N then $T < 1$, but also

$$\mathbb{E}\left[\frac{T}{1-T}\right] = \mathbb{E}[m(N)] = \infty.$$

Finally we take $Z_t := M_{t\wedge T}$ on $[0,1]$ and define H to be the corresponding integrand

$$H_t = \begin{cases} (1-t)^{-1} & \textit{on } [0,T) \\ 0 & \textit{on } [T,1] \end{cases},$$

whereupon we have

$$\begin{aligned}\mathbb{E}\left[\int_0^1 Z_s H_s^2 ds\right] &= \mathbb{E}\left[X_T\right] = \mathbb{E}\left[\frac{N(1-T)}{N(1-T)+T}\right] < 1, \text{ but} \\ \mathbb{E}\left[\int_0^1 H_s^2 ds\right] &= \mathbb{E}\left[\int_0^T \frac{1}{(1-t)^2} dt\right] = \mathbb{E}\left[\frac{T}{1-T}\right] = \infty\end{aligned}$$

as required.

Remark 8 *For any $K > 0$, it is possible to decompose the local martingale M as*

$$M = M^{sq,K} + M^{d,K},$$

where $M^{sq,K}$ is a locally square-integrable martingale with jumps bounded by a constant K and $M^{d,K}$ is a purely discontinuous local martingale with locally integrable total variation, with jumps greater than K, in such a manner that the quadratic variation process $\left[M^{sq,K}, M^{d,K}\right]$ is identically equal to 0. In what follows we will discard the dependence on the constant K in the notation for $M^{sq,K}$ and $M^{d,K}$. The first part of the statement is essentially Proposition I.4.17 in [9] while the second part follows from Theorem I.4.18 of the same reference, i.e., from the classical decomposition of the local martingale M^{sq} into its continuous and purely discontinuous parts

$$M^{sq} = M^{sq,c} + M^{sq,d}.$$

We have that

$$\left[M^{sq}, M^d\right] = \left[M^{sq,c}, M^d\right] + \left[M^{sq,d}, M^d\right] = 0$$

as $\left[M^{sq,c}, M^d\right]$ is null since it is the quadratic variation between a continuous and a purely discontinuous martingale and since $\left[M^{sq,d}, M^d\right]$ since it is the quadratic variation of two purely discontinuous martingales with no jumps occurring at the same time.

For the following proposition, we introduce a positive $\mathcal{F}_t$-adapted cadlag semimartingale of the form

$$U_t = U_0 + \int_0^t a_s ds + M_t,$$

where a is a measurable $\mathcal{F}_t$-adapted process and M is a local $\mathcal{F}_t$-martingale null at zero[6]. We also assume that $E[U_0] < \infty$ and additionally that the quadratic variation processes $\left\langle W^i, M\right\rangle$ $i = 1, ..., m$ are absolutely continuous. In particular, there exists a measurable m-dimensional $\mathcal{F}_t$-adapted process $N = (N^i)_{i=1}^m$ such that

$$\left\langle W^i, M\right\rangle_t = \int_0^t N_s^i ds, \qquad t \geq 0, \qquad i = 1, ..., m.$$

Moreover we will assume that there exists a positive constant c such that

$$\max\left(|a_t|, |N_t|^2\right) \leq c \max\left(U_t, U_{t-}\right), \qquad t \geq 0. \tag{13}$$

[6] We will use the notation $[\cdot,\cdot]$ to denote the quadratic variation process of two local martingales. In addition, we will use the notation $\langle\cdot,\cdot\rangle$ to denote the predictable quadratic variation process of two locally square integrable martingales. The two processes coincide if one of the martingales is continuous. For further details see, for example, Chapter 4 of [23].

Proposition 9 *Assume that the $\mathcal{F}_t$-adapted measurable process $H = (H^i)_{i=1}^d$ satisfies the inequality*

$$|H_t|^2 \leq c \max(U_t, U_{t-}) \qquad t \geq 0. \tag{14}$$

Then the functions $t \to \mathbb{E}\left[Z_t |H_t|^2\right]$, $t \to \mathbb{E}\left[|H_t|^2\right]$ are locally bounded. In particular Lemma 4 allows us to deduce that the process Z is a $H^1(\mathbb{P})$ martingale.

Proof. Let $(T_n)_{n>0}$ be a localizing sequence of stopping times such that the stopped process $\left(M^{sq}_{T_n\wedge\cdot}\right)$ is a square integrable martingale and the process $\left(M^d_{T_n\wedge\cdot}\right)$ is a martingale with integrable total variation $Var\left(M^d\right)_{T_n\wedge\cdot}$. Now introduce the localizing sequence $(S_n)_{n>0}$ where

$$S_n = \inf\left\{t \geq 0 \left| \max\left\{Z_t, \int_0^t |a_s|\, ds, U_{t-}\right\} \geq n\right.\right\} \wedge T_n.$$

Note that the left continuity of the processes listed in the inner brackets implies that these processes, when stopped at S_n are bounded by n. Consider now the evolution equation for ZU, that is

$$Z_t U_t = U_0 + \int_0^t Z_s \left(a_s + H_s^\top N_s\right) ds + \int_0^t Z_s \left(H_s^\top dW_s + dM_s^{sq} + dM_s^d\right). \tag{15}$$

It follows that the expected value of $Z_t U_t$ is controlled by the sum of the expected values of the six terms on the right hand side of (15). The stochastic integral terms in (15), when stopped at S_n become genuine martingales. They can be controlled as follows:

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^{t\wedge S_n} Z_s U_{s-} H_s^\top dW_s\right)^2\right] &= \mathbb{E}\left[\int_0^{t\wedge S_n} Z_s^2 U_{s-}^2 |H_s|^2\, ds\right] \\
&\leq cn^4 \mathbb{E}\left[\int_0^{t\wedge S_n} \max(U_t, U_{t-})\, ds\right] \leq cn^5 t.
\end{aligned}$$

Here we have used the fact that, for all $t \geq 0$, $\int_0^t P(U_s \neq U_{s-})\, ds = 0$. We also have that

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^{t\wedge S_n} Z_s dM_s^{sq}\right)^2\right] &= \mathbb{E}\left[\int_0^{t\wedge S_n} Z_s^2 d\left\langle M^{sq}\right\rangle_s\right] \leq n^2 \mathbb{E}\left[\left\langle M^{sq}\right\rangle_{t\wedge S_n}\right] < \infty \\
\mathbb{E}\left[\left|\int_0^{t\wedge S_n} Z_s dM_s^d\right|\right] &\leq n\mathbb{E}\left[Var\left(M^d\right)_{S_n\wedge t}\right] < \infty.
\end{aligned}$$

By taking the expectation of both sides in (15) stopped at $t \wedge S_n$, we deduce that

$$\begin{aligned}
\mathbb{E}\left[Z_t U_t 1_{\{t\leq S_n\}}\right] &\leq \mathbb{E}\left[Z_{t\wedge S_n} U_{t\wedge S_n}\right] \\
&= \mathbb{E}[U_0] + \mathbb{E}\left[\int_0^{t\wedge S_n} Z_s\left(a_s + H_s^\top N_s\right) ds\right] \\
&\leq \mathbb{E}[U_0] + 2c\mathbb{E}\left[\int_0^t Z_s \max(U_s, U_{s-})\, 1_{\{s\leq S_n\}} ds\right] \\
&\leq \mathbb{E}[U_0] + 2c\int_0^t \mathbb{E}\left[Z_s U_s 1_{\{s\leq S_n\}}\right] ds \leq e^{2ct}\mathbb{E}[U_0] < \infty.
\end{aligned}$$

Note that the last inequality follows from Gronwall's lemma. Since $\lim_{n\to\infty} S_n = \infty$, we can then deduce by the monotone convergence theorem that, for all $t > 0$,

$$\sup_{s\in[0,t]} \mathbb{E}\left[Z_s U_s\right] \leq e^{2ct}\mathbb{E}\left[U_0\right]. \tag{16}$$

The local boundedness of $t \to \mathbb{E}\left[Z_t \left|H_t\right|^2\right]$ follows from (14) and (16). Similarly we show that for all $t > 0$,

$$\sup_{s\in[0,t]} \mathbb{E}\left[U_s\right] < \infty.$$

by using the above argument with $H = 0$ for all $t \geq 0$ (and therefore $Z_t = 1$). This in turn implies the local boundedness of the functions $t \to \mathbb{E}\left[\left|H_t\right|^2\right]$. ■

3 Two Particular Cases

3.1 The signal is a jump-diffusion process

We continue to assume that the observation process follows (3), and suppose that $X_t = (X_t^i)_{i=1}^d$, for all $t \geq 0$, is a cadlag solution of a d-dimensional stochastic differential equation. This is driven by a triplet (V,W,L) comprising a p-dimensional Brownian motion $V = (V^j)_{j=1}^p$, the m-dimensional Brownian motion $W = (W^j)_{j=1}^m$ driving the observation process Y, and an $\mathbb{R}^r$-valued Lévy process $L = (L^j)_{j=1}^r$ with no centred Gaussian component and with Lévy measure F such that $F(\{0\}) = 0$. viz.

$$X_t^i = X_0^i + \int_0^t f^i(X_{s-})\,\mathrm{d}s + \sum_{j=1}^p \int_0^t \sigma^{ij}(X_{s-})\,\mathrm{d}V_s^j + \sum_{k=1}^m \int_0^t \bar{\sigma}^{ik}(X_{s-})\,\mathrm{d}W_s^k + \sum_{l=1}^r \int_0^t \tilde{\sigma}^{il}(X_{s-})\,\mathrm{d}L_s^l, \tag{17}$$

for $i = 1,\ldots,d$. We write $f = (f^i)_{i=1}^d : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma = (\sigma^{ij})_{i=1,\ldots,d,j=1,\ldots,p} : \mathbb{R}^d \to \mathbb{R}^{d\times p}$, $\bar{\sigma} = (\bar{\sigma}^{ij})_{i=1,\ldots,d,j=1,\ldots,m} : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ and $\tilde{\sigma} = (\tilde{\sigma}^{ij})_{i=1,\ldots,d,j=1,\ldots,r} : \mathbb{R}^d \to \mathbb{R}^{d\times r}$.

We recall that a function $g : E \to F$ between two normed spaces $(E, ||\cdot||_E)$ and $(F, ||\cdot||_F)$ has *at most linear growth* if there exists $K < \infty$ such that

$$||g(e)||_F \leq K\left(1 + ||e||_E\right)$$

for all $e \in E$. We record the assumptions to be made on the coefficients in the equation (17).

Condition 10 *We assume f, σ, $\bar{\sigma}$ and $\tilde{\sigma}$ are Borel and have at most linear growth.*

We will use μ to denote the Poisson random measure associated with L, i.e. for every $t \geq 0$ and $A \in \mathcal{B}(\mathbb{R}^r \setminus \{0\})$ the random measure $\mu(t,\cdot)$ defined by

$$\mu(t, A) := \sum_{0\leq s\leq t} 1_A\left(\Delta L_s\right).$$

We let $\nu(t,\cdot) := F(\cdot)\,t = \mathbb{E}\left[\mu(1,\cdot)\right]t$, where $F(\cdot)$ is the Lévy measure of L, and denote the compensated measure by $\tilde{\mu}(t, A) = \mu(t, A) - \nu(t, A)$. L then has a Lévy-Ito decomposition of the form

$$L_t = at + \int_{0<|\rho|<1} \rho\tilde{\mu}(t, \mathrm{d}\rho) + \int_{|\rho|\geq 1} \rho\mu(t, \mathrm{d}\rho). \tag{18}$$

Condition 11 *Let $L=(L_t)_{t\geq 0}$ be a Lévy process with Lévy measure F. We assume the square integrability condition*

$$\int_{|\rho|\geq 1}\rho^2 F(\mathrm{d}\rho)<\infty.$$

Remark 12 *Whenever this condition is in force we have that*

$$\int_{|\rho|\geq 1}\rho F(\mathrm{d}\rho)<\infty \text{ for every } t\geq 0, \tag{19}$$

and hence the Lévy-Ito decomposition (18) may be rewritten as

$$L_t=bt+\int_{\mathbb{R}^r\setminus\{0\}}\rho\tilde{\mu}(t,\mathrm{d}\rho),$$

where $b:=a-\int_{|\rho|\geq 1}\rho F(\mathrm{d}\rho)$.

We continue to assume the dynamics for the observation process described in (3), and we now assume that (19) holds. We can restate this example in the language of Section 2.1 by noticing that the process $\bar{X}=(X,Y)$ is a solution to a martingale problem, with generator A now given by

$$\begin{aligned} A\phi(\bar{x}) &= A\phi(x,y) \\ &= \mathcal{L}\phi(x,y)+\int_{\mathbb{R}^r\setminus\{0\}}\left[\phi(x+\tilde{\sigma}(x)\eta,y)-\phi(x,y)-\sum_{i=1}^{d}\sum_{l=1}^{r}\frac{\partial\phi(x,y)}{\partial x_i}\tilde{\sigma}^{il}(x)\eta^l\right]F(d\eta) \end{aligned}$$

where

$$\mathcal{L}=\sum_{i=1}^{d}\tilde{f}^i(x)\frac{\partial}{\partial x_i}+\sum_{k=1}^{m}h^k(x,y)\frac{\partial}{\partial y_k}+\frac{1}{2}\sum_{i,j=1}^{d}\left(a^{ij}(x)+\bar{a}^{ij}(x)\right)\frac{\partial^2}{\partial x_i\partial x_j}+\frac{1}{2}\sum_{k=1}^{m}\frac{\partial^2}{\partial x_k^2},$$

with $\tilde{f}^i(x):=f^i(x)+b^i$, and $a=(a^{ij})_{i,j=1,\ldots,d}:\mathbb{R}^d\to\mathbb{R}^{d\times d}$, $\bar{a}=(\bar{a}^{ij})_{i,j=1,\ldots,d}:\mathbb{R}^d\to\mathbb{R}^{d\times d}$ are the matrix-valued function defined respectively as

$$a^{ij}=\frac{1}{2}\sum_{k=1}^{p}\sigma^{ik}\sigma^{jk}=\frac{1}{2}\left(\sigma\sigma^\top\right)^{ij} \text{ and } a^{ij}=\frac{1}{2}\sum_{k=1}^{m}\sigma^{ik}\sigma^{jk}=\frac{1}{2}\left(\bar{\sigma}\bar{\sigma}^\top\right)^{ij}$$

for all $i,j=1,\ldots,d$.

To ensure the filtering equations described in Section 6 can be applied to this example, we wish to establish that the functions $\mathbb{E}\left[Z_.\left|h(X_.)\right|^2\right]$ and $\mathbb{E}\left[\left|h(X_.)\right|^2\right]$ are locally bounded.

Corollary 13 *Assume the coefficients in (17) satisfy Conditions 10 and that $\bar{\sigma}$ is uniformly bounded. Let $X_t=(X_t^i)_{i=1}^d$ denote a $d-$dimensional jump-diffusion process which solves (17) for all $t\geq 0$. Suppose the driving Lévy process L has a Lévy measure F which satisfies $F(\{0\})=0$ and has no Gaussian part. Assume Condition 11 and further suppose that X_0,V,W and L are independent with $\mathbb{E}\left[|X_0|^2\right]<\infty$. Let $h:\mathbb{R}^d\to\mathbb{R}^m$ be any Borel measurable function for which there exists $K>0$ such that for all $x\in\mathbb{R}^d$*

$$|h(x)|\leq K(1+|x|),$$

and let $Z=(Z_t)_{t\geq 0}$ be the positive local martingale which solves $Z_t=1+\int_0^t Z_s h(X_s)^T dW_s$. Then $\mathbb{E}\left[Z_.\left|h(X_.)\right|^2\right]$ and $\mathbb{E}\left[\left|h(X_.)\right|^2\right]$ are locally bounded.

Proof. By exploiting Remark 12 we can rewrite the SDE governing X as

$$dX_t = \tilde{f}(X_{t-})\,\mathrm{d}t + \sigma(X_{t-})\,\mathrm{d}V_t + \bar{\sigma}(X_{t-})\,\mathrm{d}W_t + \int_{\mathbb{R}^r\setminus\{0\}} \tilde{\sigma}(X_{t-})\rho\,\tilde{\mu}\,(\mathrm{d}t, \mathrm{d}\rho)\,,$$

where $\tilde{f}(x) = f(x) + b$ (b is as given in Remark 12) is clearly still locally Lipschitz. In order to apply the local boundedness lemma we need to find a suitable process U and the component processes in its decomposition. To this end we let

$$U_t = 1 + |X_t|^2\,.$$

and use Itô's formula to obtain

$$U_t = 1 + |X_0|^2 + 2\int_0^t X_{s-}^T \mathrm{d}X_s + [X, X]_t\,,$$

where the quadratic variation $[X, X]$ may be computed as

$$\begin{aligned}
[X, X]_t &= \int_0^t \mathrm{tr}\left[\sigma\left(X_{s-}\right)^T \sigma(X_{s-}) + \bar{\sigma}\left(X_{s-}\right)^T \bar{\sigma}\left(X_{s-}\right)\right] \mathrm{d}s + \int_0^t \int_{\mathbb{R}^r\setminus\{0\}} \mathrm{tr}\left[\tilde{\sigma}(X_{s-})\rho\rho^T\tilde{\sigma}(X_{s-})^T\right] \mu\,(\mathrm{d}s, \mathrm{d}\rho) \\
&= \int_0^t \mathrm{tr}\left[\sigma\left(X_{s-}\right)^T \sigma(X_{s-}) + \bar{\sigma}\left(X_{s-}\right)^T \bar{\sigma}\left(X_{s-}\right)\right] \mathrm{d}s + \sum_{0\le s\le t} \mathrm{tr}\left[\tilde{\sigma}(X_{s-})\Delta L_s \Delta L_s^T \tilde{\sigma}(X_{s-})^T\right].
\end{aligned}$$

Hence we may write U as

$$U_t = U_0 + \int_0^t a_s \mathrm{d}s + M_t,$$

where

$$\begin{aligned}
U_0 &= 1 + |X_0|^2 \\
a_t &= 2X_{t-}^T \tilde{f}(X_{t-}) + \mathrm{tr}\left[\sigma\left(X_{t-}\right)^T \sigma(X_{t-}) + \bar{\sigma}\left(X_{t-}\right)^T \bar{\sigma}\left(X_{t-}\right)\right] \\
&\quad + \int_0^t \int_{\mathbb{R}^r\setminus\{0\}} \mathrm{tr}\left[\tilde{\sigma}(X_{s-})\rho\rho^T\tilde{\sigma}(X_{s-})^T\right] F\,(\mathrm{d}\rho)\,\mathrm{d}s
\end{aligned}$$

and M is the local martingale

$$M_t = \int_0^t 2X_{s-}^T \left[\sigma(X_{s-})\,\mathrm{d}V_s + \bar{\sigma}(X_{s-})\,\mathrm{d}W_s\right] + \int_0^t \int_{\mathbb{R}^r\setminus\{0\}} \mathrm{tr}\left[\tilde{\sigma}(X_{s-})\rho\rho^T\tilde{\sigma}(X_{s-})^T\right] \tilde{\mu}\,(\mathrm{d}s, \mathrm{d}\rho)\,.$$

Condition 10 on $\tilde{f}, \sigma, \bar{\sigma}$ and $\tilde{\sigma}$ ensures the existence of $C > 0$ such that

$$a_t \le C\left(U_{t-} \vee U_t\right),$$

moreover the boundedness of $\bar{\sigma}$ gives rise to the estimate

$$\left|\langle W, M\rangle_t'\right| = \left|\bar{\sigma}(X_{t-})X_{t-}\right| \le K\left|X_{t-}\right| \le K U_{t-}^{1/2}.$$

The result then follows from Proposition 9. ■

Remark 14 *We may adapt this example to the case where X be an $\{\mathcal{F}_t\}$-adapted Markov process with values in a finite state space I*

3.2 The *change-detection* filtering problem.

The following is a simple example with real-world applications which fits within the above framework. The effect we try to capture is a sudden change in the parameters of the model which describes the (stochastic) evolution of the observed process. The following illustrates how such an effect might be incorporated into the framework presented previously.

We assume that Y is the real-valued process with dynamics

$$Y_t = \int_0^t \left(b_0 + B1_{[T,\infty)}(s)\right) Y_s ds + W_t,$$

where $W = \{W_t,\ t \geq 0\}$ is a standard Brownian motion, b_0 a constant and B and T independent random variables, which are also independent of W. We also assume that $T \geq 0$ and that $\mathbb{E}\left[e^{\lambda B^2}\right] < \infty$ for all $\lambda \in \mathbb{R}$. The process $X_t = \left(X_t^1, X_t^2\right)$ is then defined by

$$X_t^1 = B \text{ and } X_t^2 = I_{[T,\infty)}(t), \qquad t \geq 0,$$

whereupon the process $\bar{X}_t = \left(X_t^1, X_t^2, Y_t\right)$ is adapted to the filtration

$$\{\mathcal{F}_t\}_{t\geq 0} := \left\{\sigma\left(B, I_{[T,\infty)}(s), W_s : s \leq t\right) \vee \mathcal{N}\right\}_{t\geq 0},$$

where $\mathcal{N}$ is the class of null sets of the completed σ-field $\mathcal{F}_\infty = \bar{\sigma}(B, T, W_s, s < \infty)$. We introduce the uniquely defined cadlag $(\mathcal{B}(\mathbb{R}) \times \mathcal{F}_t)$ −optional processes

$$\begin{aligned}
(t, b, \omega) &\mapsto H_t^b(\omega) = \left(b_0 + b1_{[T(\omega),\infty)}(t)\right) Y_t^b(\omega) \\
(t, b, \omega) &\mapsto Y_t^b(\omega) = \int_0^t H_s^b(\omega)\, ds + W_t(\omega),
\end{aligned}$$

and set $Z_t^b := \exp\left[-\int_0^t H_s^b dW_s - \frac{1}{2}\int_0^t (H_s)^2 ds\right]$. Notice that B is $\mathcal{F}_0$-measurable, and hence the continuous process $\left(Z_t^B\right)_{t\geq 0}$ is an $\{\mathcal{F}_t\}$-adapted exponential local martingale. Again, as in the previous example, we need to show that the functions $\mathbb{E}\left[Z_\cdot^B\left(H_\cdot^B\right)^2\right]$ and $\mathbb{E}\left[\left(H_\cdot^B\right)^2\right]$ are locally bounded. To do this, fix $b \in \mathbb{R}$ and take the terms U_t and c in Proposition 9 to be

$$U_t = U_t^b := 1 + \left(Y_t^b\right)^2 \text{ and } c = c(b) := 4 + (b_0 + b)^2.$$

Then we may verify that the conditions of Proposition 9 are satisfied. It is immediate from its proof that the conclusion of Proposition 9 can be strengthened to give the estimate

$$\max\left\{\mathbb{E}\left[Z_t^b\left(H_t^b\right)^2\right], \mathbb{E}\left[\left(H_t^b\right)^2\right]\right\} \leq e^{c(b)t}\mathbb{E}\left[U_0^b\right] = e^{c(b)t}.$$

Consequently

$$\mathbb{E}\left[Z_t^B\left(H_t^B\right)^2\right] = \mathbb{E}\left[\mathbb{E}\left[Z_t^b\left(H_t^b\right)^2\right]\Big|_{b=B}\right] \leq \mathbb{E}\left[e^{c(B)t}\right]$$

and similarly

$$\mathbb{E}\left[\left(H_t^B\right)^2\right] \leq \mathbb{E}\left[e^{c(B)t}\right].$$

These inequalities, together with the moment condition on B, give the required result.

4 The Change of Probability Measure Method

We now have all the ingredients required for introducing a probability measure with respect to which the process Y becomes a Brownian motion. We return to the set-up of Section 2. Define $Z = (Z_t)_{t\geq 0}$ to be the exponential local martingale

$$Z_t = \exp\left(-\int_0^t h\left(\bar{X}_s\right)^\top dW_s - \frac{1}{2}\int_0^t \left|h\left(\bar{X}_s\right)\right|^2 ds\right).$$

The change of probability measure method consists in modifying the probability measure on Ω by means of Girsanov's theorem. As we require Z to be a martingale in order to construct the change of measure, Lemma 4 suggests the following as a suitable condition to impose upon h,

$$\mathbb{E}\left[\int_0^t Z_s \left\|h(\bar{X}_s)\right\|^2 \,\mathrm{d}s\right] < \infty, \quad \forall t > 0. \tag{20}$$

Let us assume that (20) holds. Then, by Lemma 4, Z is a true martingale. Let $\tilde{\mathbb{P}}$ be the probability measure defined on the field $\bigcup_{0\leq t<\infty} \mathcal{F}_t$ that is specified by its Radon–Nikodym derivative Z_t on each $\mathcal{F}_t$ with respect to the corresponding trace of $\mathbb{P}$; that is, for each $t \geq 0$:

$$\left.\frac{\mathrm{d}\tilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = Z_t.$$

$\tilde{\mathbb{P}}$ restricted to each $\mathcal{F}_t$ is equivalent to $\mathbb{P}$ since Z_t is a positive random variable[7].

Let $\tilde{Z} = \{\tilde{Z}_t,\ t \geq 0\}$ be the process defined as $\tilde{Z}_t = Z_t^{-1}$ for $t \geq 0$. Under $\tilde{\mathbb{P}}$, $\tilde{Z}_t$ satisfies the following stochastic differential equation,

$$\mathrm{d}\tilde{Z}_t = \sum_{i=1}^m \tilde{Z}_t h^i(X_t)\,\mathrm{d}Y_t^i \tag{21}$$

and since $\tilde{Z}_0 = 1$,

$$\tilde{Z}_t = \exp\left(\sum_{i=1}^m \int_0^t h^i(X_s)\,\mathrm{d}Y_s^i - \frac{1}{2}\sum_{i=1}^m \int_0^t h^i(X_s)^2\,\mathrm{d}s\right), \tag{22}$$

then $\tilde{\mathbb{E}}[\tilde{Z}_t] = \mathbb{E}[\tilde{Z}_t Z_t] = 1$. So $\tilde{Z}$ is an $\mathcal{F}_t$-adapted martingale under $\tilde{\mathbb{P}}$ and

$$\left.\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} = \tilde{Z}_t \quad \text{for } t \geq 0.$$

$\mathbb{P}$ and $\tilde{\mathbb{P}}$ are therefore equivalent on each $\mathcal{F}_t$ for $t \geq 0$.

Proposition 15 *If condition (20) is satisfied, then under $\tilde{\mathbb{P}}$ the observation process Y is a Brownian motion. Let $\varphi \in \mathcal{D}(A)$ have bounded derivatives in the y-direction, and let $\tilde{M}^\varphi$ denote the semimartingale*

$$\tilde{M}_t^\varphi := M_t^\varphi + \int_0^t \sum_{i=1}^m \left(h^i B^i \varphi + \frac{\partial \varphi}{\partial y_i}\right)\left(\bar{X}_t\right)\mathrm{d}s.$$

Then the stochastic integral $\int_0^\cdot \tilde{Z}_s d\tilde{M}_s^\varphi$ is a zero-mean martingale under $\tilde{\mathbb{P}}$.

[7] Note that we have not defined $\tilde{\mathbb{P}}$ on $\mathcal{F}_\infty$, where $\mathcal{F}_\infty = \bigvee_{t=0}^\infty \mathcal{F}_t = \sigma\left(\bigcup_{0\leq t<\infty} \mathcal{F}_t\right)$.

Proof. Lemma 4, together with condition 20, ensures that Z is a martingale (under $\mathbb{P}$) and that $\tilde{\mathbb{P}}$ is a probability measure on each $\mathcal{F}_t$.That Y becomes a Brownian motion under $\tilde{\mathbb{P}}$ is an immediate consequence of Girsanov's theorem. For brevity, let β denote the process defined by

$$\beta_t := \sum_{i=1}^{m} \left(h^i B^i \varphi + \frac{\partial \varphi}{\partial y_i} \right) (\bar{X}_t) ;$$

then $\tilde{M}_t^{\varphi}$ can be expressed as $M_t^{\varphi} + \int_0^t \beta_s \mathrm{d}s$. It also follows from (4) and the definition of $\tilde{Z}$ that $\left\langle M^{\varphi}, \tilde{Z} \right\rangle_t = \int_0^t \tilde{Z}_s \beta_s \mathrm{d}s$. But by Itô's integration-by-parts formula

$$\begin{aligned} \tilde{Z}_t M_t^{\varphi} &= \int_0^t M_s^{\varphi} \mathrm{d}\tilde{Z}_s + \int_0^t \tilde{Z}_s \mathrm{d}M_s^{\varphi} + \left\langle M^{\varphi}, \tilde{Z} \right\rangle_t \\ &= \int_0^t M_s^{\varphi} \mathrm{d}\tilde{Z}_s + \int_0^t \tilde{Z}_s \left(\mathrm{d}M_s^{\varphi} + \beta_s \mathrm{d}s \right) \\ &= \int_0^t M_s^{\varphi} \mathrm{d}\tilde{Z}_s + \int_0^t \tilde{Z}_s \mathrm{d}\tilde{M}_s^{\varphi}. \end{aligned} \tag{23}$$

However M^{φ} being a martingale under $\tilde{\mathbb{P}}$ implies that $\tilde{Z}M^{\varphi}$ is a martingale under $\tilde{\mathbb{P}}$, and the first integral on the right-hand side is a martingale under $\tilde{\mathbb{P}}$ because M^{φ} is bounded on finite intervals and $\tilde{Z}$ itself is a martingale. The conclusion of the proposition follows. ■

Remark 16 *Since $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are absolutely continuous with respect to each other, they have the same class of null sets $\mathcal{N}$ and therefore the (augmented) observation filtration is the same both under $\mathbb{P}$ and $\tilde{\mathbb{P}}$. Since Y is a Brownian motion under $\tilde{\mathbb{P}}$ it follows that the filtration $\{\mathcal{Y}_t,\ t \geq 0\}$ is right-continuous both under $\mathbb{P}$ and $\tilde{\mathbb{P}}$. To put it differently, $\{\mathcal{Y}_t,\ t \geq 0\}$ satisfies the usual conditions both under $\mathbb{P}$ and under $\tilde{\mathbb{P}}$.*

The following proposition is a consequence of the Brownian motion property of the process Y under $\tilde{\mathbb{P}}$.

Proposition 17 *Let U be an integrable $\mathcal{F}_t$-measurable random variable. Then we have*

$$\tilde{\mathbb{E}}[U \mid \mathcal{Y}_t] = \tilde{\mathbb{E}}[U \mid \mathcal{Y}]. \tag{24}$$

Proof. Let us denote by

$$\mathcal{Y}_t' = \sigma(Y_{t+u} - Y_t;\ u \geq 0);$$

then $\mathcal{Y} = \sigma(\mathcal{Y}_t, \mathcal{Y}_t')$. Under the probability measure $\tilde{\mathbb{P}}$ the σ-algebra $\mathcal{Y}_t' \subset \mathcal{Y}$ is independent of $\mathcal{F}_t$ because Y is an $\mathcal{F}_t$-adapted Brownian motion. Hence since U is $\mathcal{F}_t$-adapted using the property (f) of conditional expectation

$$\tilde{\mathbb{E}}[U \mid \mathcal{Y}_t] = \tilde{\mathbb{E}}[U \mid \sigma(\mathcal{Y}_t, \mathcal{Y}_t')] = \tilde{\mathbb{E}}[U \mid \mathcal{Y}].$$

■

5 Unnormalised Conditional Distribution

In this section we first prove the Kallianpur–Striebel formula and use this to define the unnormalized conditional distribution process. The notation $\tilde{\mathbb{P}}(\mathbb{P})$-a.s. below means that the result holds both $\tilde{\mathbb{P}}$-a.s. and $\mathbb{P}$-a.s. We only need to show that it holds true in the first sense since $\tilde{\mathbb{P}}$ and $\mathbb{P}$ are equivalent probability measures.

Proposition 18 (Kallianpur–Striebel) *Assume that condition (20) holds. For every $\varphi \in b\mathcal{B}(\mathbb{S})$, for fixed $t \in [0, \infty)$,*

$$\pi_t(\varphi) = \frac{\tilde{\mathbb{E}}[\tilde{Z}_t \varphi(X_t) \mid \mathcal{Y}]}{\tilde{\mathbb{E}}[\tilde{Z}_t \mid \mathcal{Y}]} \qquad \tilde{\mathbb{P}}(\mathbb{P})\text{-a.s.} \tag{25}$$

Proof. It is clear from the definition that $\tilde{Z}_t > 0$ $\tilde{\mathbb{P}}(\mathbb{P})$-a.s. as a consequence of which $\tilde{\mathbb{E}}[\tilde{Z}_t \mid \mathcal{Y}] > 0$ $\mathbb{P}$-a.s. and the right-hand side of (25) is well defined. It suffices to show that

$$\pi_t(\varphi)\tilde{\mathbb{E}}[\tilde{Z}_t \mid \mathcal{Y}_t] = \tilde{\mathbb{E}}[\tilde{Z}_t \varphi(X_t) \mid \mathcal{Y}_t] \qquad \tilde{\mathbb{P}}\text{-a.s.}$$

As both the left- and right-hand sides of this equation are $\mathcal{Y}_t$-measurable, this is equivalent to showing that for any bounded $\mathcal{Y}_t$-measurable random variable b,

$$\tilde{\mathbb{E}}[\pi_t(\varphi)\tilde{\mathbb{E}}[\tilde{Z}_t \mid \mathcal{Y}_t]b] = \tilde{\mathbb{E}}[\tilde{\mathbb{E}}[\tilde{Z}_t \varphi(X_t) \mid \mathcal{Y}_t]b].$$

A consequence of the definition of the process π_t is that $\pi_t\varphi = \mathbb{E}[\varphi(X_t) \mid \mathcal{Y}_t]$ $\tilde{\mathbb{P}}$-a.s., so from the definition of Kolmogorov conditional expectation

$$\mathbb{E}\left[\pi_t(\varphi)b\right] = \mathbb{E}\left[\varphi(X_t)b\right].$$

Writing this under the measure $\tilde{\mathbb{P}}$,

$$\tilde{\mathbb{E}}\left[\pi_t(\varphi)b\tilde{Z}_t\right] = \tilde{\mathbb{E}}\left[\varphi(X_t)b\tilde{Z}_t\right].$$

Since the function b is $\mathcal{Y}_t$-measurable, by the tower property of the conditional expectation,

$$\tilde{\mathbb{E}}\left[\pi_t(\varphi)\tilde{\mathbb{E}}[\tilde{Z}_t \mid \mathcal{Y}_t]b\right] = \tilde{\mathbb{E}}\left[\tilde{\mathbb{E}}[\varphi(X_t)\tilde{Z}_t \mid \mathcal{Y}_t]b\right]$$

which proves that the result holds $\tilde{\mathbb{P}}$-a.s. ■

Let $\zeta = \{\zeta_t,\ t \geq 0\}$ be the process defined by

$$\zeta_t = \tilde{\mathbb{E}}[\tilde{Z}_t \mid \mathcal{Y}_t], \tag{26}$$

then as $\tilde{Z}_t$ is an $\mathcal{F}_t$-martingale under $\tilde{\mathbb{P}}$ and $\mathcal{Y}_s \subseteq \mathcal{F}_s$, it follows that for $0 \leq s < t$,

$$\tilde{\mathbb{E}}[\zeta_t \mid \mathcal{Y}_s] = \tilde{\mathbb{E}}[\tilde{Z}_t|\mathcal{Y}_s] = \tilde{\mathbb{E}}\left[\tilde{\mathbb{E}}[\tilde{Z}_t \mid \mathcal{F}_s] \mid \mathcal{Y}_s\right] = \tilde{\mathbb{E}}[\tilde{Z}_s \mid \mathcal{Y}_s] = \zeta_s.$$

Therefore by Doob's regularization theorem (see Rogers and Williams, [23, Theorem II.67.7]) since the filtration $\mathcal{Y}_t$ satisfies the usual conditions we can choose a càdlàg version of ζ_t which is a $\mathcal{Y}_t$-martingale. In what follows, assume that $\{\zeta_t, t \geq 0\}$ has been chosen to be such a version. Given such a ζ, Proposition 18 suggests the following definition.

Definition 19 *Define the* unnormalised conditional distribution *of X to be the measure-valued process $\rho = \{\rho_t,\ t \geq 0\}$ given by $\rho_t = \zeta_t \pi_t$ for any $t \geq 0$.*

Lemma 20 *The process $\{\rho_t,\ t \geq 0\}$ is càdlàg and $\mathcal{Y}_t$-adapted. Furthermore, for any $t \geq 0$,*

$$\rho_t(\varphi) = \tilde{\mathbb{E}}\left[\tilde{Z}_t \varphi(X_t) \mid \mathcal{Y}_t\right] \qquad \tilde{\mathbb{P}}(\mathbb{P})\text{-a.s.} \tag{27}$$

Proof. Both $\pi_t(\varphi)$ and ζ_t are $\mathcal{Y}_t$-adapted. By construction $\{\zeta_t,\ t \geq 0\}$ is also càdlàg. We know that $\{\pi_t,\ t \geq 0\}$ is càdlàg and $\mathcal{Y}_t$-adapted; therefore the process $\{\rho_t, t \geq 0\}$ is also càdlàg and $\mathcal{Y}_t$-adapted.

For the second part, from Proposition 17 and Proposition 18 it follows that

$$\pi_t(\varphi)\tilde{\mathbb{E}}[\tilde{Z}_t \mid \mathcal{Y}_t] = \tilde{\mathbb{E}}[\tilde{Z}_t\varphi(X_t) \mid \mathcal{Y}_t] \qquad \tilde{\mathbb{P}}\text{-a.s.},$$

From (26), $\tilde{\mathbb{E}}[\tilde{Z}_t \mid \mathcal{Y}_t] = \zeta_t$ a.s. from which the result follows. ∎

Corollary 21 *Assume that condition (20) holds. For every $\varphi \in B(\mathbb{S})$,*

$$\pi_t(\varphi) = \frac{\rho_t(\varphi)}{\rho_t(\mathbf{1})} \qquad \forall t \in [0, \infty) \qquad \tilde{\mathbb{P}}(\mathbb{P})\text{-}a.s. \tag{28}$$

Proof. It is clear from Definition 19 that $\zeta_t = \rho_t(\mathbf{1})$. The result then follows immediately. ∎

The Kallianpur–Striebel formula explains the usage of the term *unnormalised* in the definition of ρ_t as the denominator $\rho_t(\mathbf{1})$ can be viewed as the normalising factor.

Lemma 22 **i.** *Let $\{u_t,\ t \geq 0\}$ be an $\mathcal{F}_t$-progressively measurable process such that for all $t \geq 0$, we have*

$$\tilde{\mathbb{E}}\left[\left(\int_0^t u_s^2 \,\mathrm{d}s\right)^{1/2}\right] < \infty; \tag{29}$$

then, for all $t \geq 0$, and $j = 1, \ldots, m$, we have

$$\tilde{\mathbb{E}}\left[\int_0^t u_s \,\mathrm{d}Y_s^j \,\middle|\, \mathcal{Y}\right] = \int_0^t \tilde{\mathbb{E}}[u_s \mid \mathcal{Y}] \,\mathrm{d}Y_s^j. \tag{30}$$

ii. *Let $\tilde{M}^\varphi$ be as defined in Proposition 15. Then for all $t \geq 0$*

$$\tilde{\mathbb{E}}\left[\int_0^t \tilde{Z}_s \,\mathrm{d}\tilde{M}_s^\varphi \,\middle|\, \mathcal{Y}\right] = \sum_{j=1}^m \int_0^t \tilde{\mathbb{E}}\left[\left(B^j\varphi + \frac{\partial \varphi}{\partial y_j}\right)(\bar{X}_s)\,\tilde{Z}_s \,\middle|\, \mathcal{Y}\right] \mathrm{d}Y_s^j, \tag{31}$$

Proof.

i. To deduce the results we introduce the set of uniformly bounded test random variables

$$S_t = \left\{\varepsilon_t = \exp\!\left(i\int_0^t r_s^\top \,\mathrm{d}Y_s + \frac{1}{2}\int_0^t \|r_s\|^2 \,\mathrm{d}s\right) : r \in L^\infty\left([0,t], \mathbb{R}^m\right)\right\}. \tag{32}$$

Then S_t is a total set. That is, if $a \in L^1(\Omega, \mathcal{Y}_t, \tilde{\mathbb{P}})$ and $\tilde{\mathbb{E}}[a\varepsilon_t] = 0$, for all $\varepsilon_t \in S_t$, then $a = 0$ $\tilde{\mathbb{P}}$-a.s. For a proof of this result see, for example, Lemma B.39 page 355 in Bain and Crisan [2]. In addition, if $\varepsilon_t \in S_t$, then

$$\varepsilon_t = 1 + \int_0^t i\varepsilon_s r_s^\top \,\mathrm{d}Y_s.$$

From condition (29) it follows, by Burkholder-Davis-Gundy's inequalities that both processes $t \to \int_0^t u_s \,\mathrm{d}Y_s^j$ and $t \to \int_0^t \tilde{\mathbb{E}}\left[u_s \mid \mathcal{Y}\right] \mathrm{d}Y_s^j$ belong to $H^1(\tilde{\mathbb{P}})$. In particular they are zero-mean martingales. We observe the following sequence of identities

$$\begin{aligned}
\tilde{\mathbb{E}}\left[\varepsilon_t\tilde{\mathbb{E}}\left[\int_0^t u_s\,\mathrm{d}Y_s^j\,\middle|\,\mathcal{Y}\right]\right] &= \tilde{\mathbb{E}}\left[\varepsilon_t\int_0^t u_s\,\mathrm{d}Y_s^j\right]\\
&= \tilde{\mathbb{E}}\left[\int_0^t u_s\,\mathrm{d}Y_s^j\right] + \tilde{\mathbb{E}}\left[\int_0^t i\varepsilon_s r_s^j u_s\,\mathrm{d}s\right]\\
&= \tilde{\mathbb{E}}\left[\tilde{\mathbb{E}}\left[\int_0^t i\varepsilon_s r_s^j u_s\,\mathrm{d}s\,\middle|\,\mathcal{Y}\right]\right]\\
&= \tilde{\mathbb{E}}\left[\int_0^t i\varepsilon_s r_s^j\,\tilde{\mathbb{E}}[u_s\mid\mathcal{Y}]\,\mathrm{d}s\right]\\
&= \tilde{\mathbb{E}}\left[\varepsilon_t\int_0^t \tilde{\mathbb{E}}[u_s\mid\mathcal{Y}]\,\mathrm{d}Y_s^j\right],
\end{aligned}$$

which completes the proof of (30).

ii. From Proposition 15 we know that $\int_0^\cdot \tilde{Z}_s\,\mathrm{d}\tilde{M}_s^\varphi$ is a zero-mean martingale under $\tilde{\mathbb{P}}$. It is therefore integrable and its conditional expectation is well defined. Notice that

$$\left\langle \tilde{M}^\varphi, Y^j\right\rangle_t = \left\langle M^\varphi, W^j\right\rangle_t = \int_0^t \left(B^j\varphi + \frac{\partial\varphi}{\partial y_j}\right)\left(\bar{X}_s\right)\,\mathrm{d}s$$

The rest of the proof of (31) is similar to that of (30). Once again we choose ε_t from the set S_t and in this case we obtain the following sequence of identities.

$$\begin{aligned}
\tilde{\mathbb{E}}\left[\varepsilon_t\tilde{\mathbb{E}}\left[\int_0^t \tilde{Z}_s\,\mathrm{d}\tilde{M}_s^\varphi\,\middle|\,\mathcal{Y}\right]\right] &= \tilde{\mathbb{E}}\left[\varepsilon_t\int_0^t \tilde{Z}_s\,\mathrm{d}\tilde{M}_s^\varphi\right]\\
&= \tilde{\mathbb{E}}\left[\int_0^t \tilde{Z}_s\,\mathrm{d}\tilde{M}_s^\varphi\right] + \sum_{j=1}^m \tilde{\mathbb{E}}\left\langle\int_0^\cdot i\varepsilon_s r_s^j\,\mathrm{d}Y_s^j, \int_0^\cdot \tilde{Z}_s\,\mathrm{d}\tilde{M}_s^\varphi\right\rangle_t\\
&= \tilde{\mathbb{E}}\left[\int_0^t \tilde{Z}_s\,\mathrm{d}\tilde{M}_s^\varphi\right] + \sum_{j=1}^m \tilde{\mathbb{E}}\int_0^t i\varepsilon_s r_s^j\tilde{Z}_s\,\mathrm{d}\left\langle\tilde{M}^\varphi, Y^j\right\rangle_s\\
&= \sum_{j=1}^m \tilde{\mathbb{E}}\int_0^t i\varepsilon_s r_s^j\tilde{Z}_s\,\left(B^j\varphi + \frac{\partial\varphi}{\partial y_j}\right)\left(\bar{X}_s\right)\mathrm{d}s\\
&= \sum_{j=1}^m \tilde{\mathbb{E}}\left[\varepsilon_t\int_0^t \tilde{\mathbb{E}}\left[\left(B^j\varphi + \frac{\partial\varphi}{\partial y_j}\right)\left(\bar{X}_s\right)\tilde{Z}_s\,\middle|\,\mathcal{Y}\right]\mathrm{d}Y_s^j\right].
\end{aligned}$$

As the identities hold for an arbitrary choice of $\varepsilon_t \in S_t$, the proof of (31) is complete

■

6 The Filtering Equations

To simplify the analysis, we will impose onto $\tilde{Z}$ a similar condition to (20). More precisely, we will assume that,

$$\tilde{\mathbb{E}}\left[\int_0^t \tilde{Z}_s\left\|h(\bar{X}_s)\right\|^2\,\mathrm{d}s\right] < \infty, \quad \forall t > 0. \tag{33}$$

Reverting back to $\mathbb{P}$, condition (33) is equivalent to

$$\mathbb{E}\left[\int_0^t \left\|h(\bar{X}_s)\right\|^2 \,\mathrm{d}s\right] < \infty, \quad \forall t > 0. \tag{34}$$

From Corollary 5, it follows that $\tilde{Z}$ is an $H^1(\tilde{\mathbb{P}})$-martingale. Then $\left(\tilde{Z}_\cdot - 1\right)$ is a zero-mean martingale and $\mathbb{E}\left[\left(\tilde{Z}_\cdot - 1\right)^*_t\right] < 1+ \mathbb{E}\left[\tilde{Z}^*_t\right] < \infty$. Since $\left\langle \tilde{Z}_\cdot - 1\right\rangle_t = \int_0^t \tilde{Z}_s^2 \left|h(\bar{X}_s)\right|^2 ds$ the Burkholder-Davis-Gundy inequalities give

$$\mathbb{E}\left[\left(\int_0^t \tilde{Z}_s^2 \left|h(\bar{X}_s)\right|^2 ds\right)^{1/2}\right] < \infty \tag{35}$$

for all $t \geq 0$ and hence, for any $\varphi \in b\mathcal{B}(\mathbb{S}\times\mathbb{R}^m)$, the processes

$$\begin{aligned}
t &\to \int_0^t \varphi(\bar{X}_t)\tilde{Z}_t h(\bar{X}_s)^\top \mathrm{d}Y_s \\
t &\to \int_0^t \tilde{\mathbb{E}}[\varphi(\bar{X}_t)\tilde{Z}_t h(\bar{X}_s)^\top \mid \mathcal{Y}_t]\,\mathrm{d}Y_s
\end{aligned}$$

are zero-mean $H^1(\tilde{\mathbb{P}})$ martingales. In the following, for any function $\varphi \in b\mathcal{B}(\mathbb{S}\times\mathbb{R}^m)$ such that $\varphi \in \mathcal{D}(A)$ and that has bounded partial derivatives in the y direction we will denote by $D_i\varphi, \quad j = 1,\ldots,m$ the functions

$$D_j\varphi = h^j\left(\varphi + B^j\varphi + \frac{\partial\varphi}{\partial y_j}\right) \qquad j = 1,\ldots,m.$$

Theorem 23 *If conditions (20) and (33) are satisfied then,*

$$\tilde{\mathbb{E}}[\tilde{Z}_t\varphi(\bar{X}_t)\mid\mathcal{Y}] = \pi_0(\varphi) + \int_0^t \tilde{\mathbb{E}}[\tilde{Z}_s A\varphi(\bar{X}_s)\mid\mathcal{Y}]\,\mathrm{d}s + \sum_{j=1}^m \tilde{\mathbb{E}}[\tilde{Z}_s D_j\varphi(\bar{X}_s)\mid\mathcal{Y}]\mathrm{d}Y_s^j \tag{36}$$

for any $\varphi \in b\mathcal{B}(\mathbb{S}\times\mathbb{R}^m)$ be a function such that $\varphi, \varphi^2 \in \mathcal{D}(A)$ and that has bounded partial derivatives in the y direction. In particular the process ρ_t satisfies the following evolution equation

$$\rho_t(\varphi) = \rho_0(\varphi) + \int_0^t \rho_s\left(A\varphi\right)\,\mathrm{d}s + \int_0^t \rho_s((h^\top + B^\top)\varphi)\,\mathrm{d}Y_s, \qquad \tilde{\mathbb{P}}\text{-a.s. } \forall t \geq 0 \tag{37}$$

for any function $\varphi \in b\mathcal{B}(\mathbb{S})$ be a function such that $\varphi \in \mathcal{D}(A)$.

Proof. Using Itô's formula and integration-by-parts, we find

$$\begin{aligned}
\mathrm{d}\left(\tilde{Z}_t\varphi(\bar{X}_t)\right) &= \tilde{Z}_t A\varphi(\bar{X}_t)\,\mathrm{d}t + \tilde{Z}_t\mathrm{d}M_t^\varphi + \varphi(\bar{X}_t)\tilde{Z}_t h^\top(\bar{X}_t)\,\mathrm{d}Y_t + \sum_{j=1}^m \tilde{Z}_t h^i(\bar{X}_t)\left\langle M^\varphi, Y^i\right\rangle_t \\
&= \tilde{Z}_t\left[A\varphi(\bar{X}_t) + \sum_{j=1}^m h^i(\bar{X}_t)\left(B^i\varphi\left(\bar{X}_t\right) + \frac{\partial\varphi}{\partial y_i}\left(\bar{X}_t\right)\right)\right]\mathrm{d}t + \tilde{Z}_t\mathrm{d}M_t^\varphi \\
&\quad +\varphi(\bar{X}_t)\tilde{Z}_t h^\top(\bar{X}_t)\,\mathrm{d}Y_t \\
&= \tilde{Z}_t A\varphi(\bar{X}_t)\mathrm{d}t + \tilde{Z}_t\mathrm{d}\tilde{M}_t^\varphi + \varphi(\bar{X}_t)\tilde{Z}_t h^\top(\bar{X}_t)\,\mathrm{d}Y_t.
\end{aligned} \tag{38}$$

We next take the conditional expectation with respect to $\mathcal{Y}$ and obtain

$$\tilde{\mathbb{E}}[\tilde{Z}_t\varphi(\bar{X}_t) \quad | \quad \mathcal{Y}] = \tilde{\mathbb{E}}[\tilde{Z}_0\varphi(\bar{X}_t) \mid \mathcal{Y}] + \int_0^t \tilde{\mathbb{E}}[\tilde{Z}_t A\varphi(\bar{X}_t) \mid \mathcal{Y}]\,\mathrm{d}s$$
$$+\tilde{\mathbb{E}}\left[\int_0^t \tilde{Z}_s \mathrm{d}\tilde{M}_s^{\varphi} \mid \mathcal{Y}\right] + \tilde{\mathbb{E}}\left[\int_0^t \varphi(\bar{X}_s)\tilde{Z}_s h^{\top}(\bar{X}_s)\,\mathrm{d}Y_s \mid \mathcal{Y}\right], \tag{39}$$

where we have used Fubini's theorem (the conditional version) to get the second term on the right hand side of (39). Observe that, since $\tilde{Z}$ is an $H^1(\tilde{\mathbb{P}})$-martingale, we have

$$\tilde{\mathbb{E}}\left[\left(\int_0^t \tilde{Z}_s^2\,\mathrm{d}s\right)^{1/2}\right] \le \sqrt{t}\tilde{\mathbb{E}}\left[\tilde{Z}_s^*\right] < \infty.$$

Also from (35) we get that

$$\tilde{\mathbb{E}}\left[\left(\int_0^t \left(\varphi(\bar{X}_s)\tilde{Z}_s h^j(\bar{X}_s)\right)^2 \mathrm{d}s\right)^{1/2}\right] \le ||\varphi||\,\mathbb{E}\left[\left(\int_0^t \tilde{Z}_s^2 \left|h(\bar{X}_s)\right|^2 ds\right)^{1/2}\right] < \infty.$$

In other words condition (29) is satisfied for $u = \varphi\tilde{Z}h^j$. The identity (36) then follows from (39) by applying (30) and (31). Identity (37) follows immediately after observing that the terms containing the partial derivatives in the y direction $\frac{\partial\varphi}{\partial y_i}$ are zero since the function no longer depends on y. ∎

Theorem 24 *If conditions (20) and (33) are satisfied then the conditional distribution of the signal π_t satisfies the following evolution equation*

$$\pi_t(\varphi) = \pi_0(\varphi) + \int_0^t \pi_s(A\varphi)\,\mathrm{d}s$$
$$+\int_0^t \left(\pi_s(\varphi h^{\top}) - \pi_s(h^{\top})\pi_s(\varphi) + \pi_t(B^{\top}\varphi)\right)(\mathrm{d}Y_s - \pi_s(h)\,\mathrm{d}s), \tag{40}$$

for any $\varphi \in \mathcal{D}(A)$.

Proof. Since $A\mathbf{1}$ =0, it follows from (1) that $M^{\mathbf{1}} \equiv 0$, which together with (4) implies that

$$\int_0^t B^i\mathbf{1}\left(\bar{X}_s\right) ds = 0,$$

for any $t \ge 0$ and $i = 1, ..., m$, so

$$\sum_{j=1}^m \int_0^t \rho_s\left(h^j B^j \mathbf{1}\right)\,\mathrm{d}s = 0.$$

Hence, from (37), one obtains that $\rho_t(\mathbf{1})$ satisfies the following equation

$$\rho_t(\mathbf{1}) = 1 + \int_0^t \rho_s(h^{\top})\,\mathrm{d}Y_s.$$

Let $(U_n)_{n>0}$ be the sequence of stopping times

$$U_n = \inf\left\{t \ge 0 \,\middle|\, \rho_t(\mathbf{1}) \le \frac{1}{n}\right\}.$$

Then

$$\rho_t^{U_n}(\mathbf{1}) = \rho_{t\wedge U_n}(\mathbf{1}) = 1 + \int_0^{t\wedge U_n} \rho_s(h^\top)\,\mathrm{d}Y_s,$$

We apply Itô's formula to the stopped process $t \to \rho_{t\wedge U_n}(\mathbf{1})$ and the function $x \mapsto \frac{1}{x}$ to obtain that

$$\frac{1}{\rho_t^{U_n}(\mathbf{1})} = 1 - \int_0^{t\wedge U_n} \frac{\rho_s(h^\top)}{\rho_s(\mathbf{1})^2}\mathrm{d}Y_s + \int_0^{t\wedge U_n} \frac{\rho_s(h^\top)\rho_s(h)}{\rho_s(\mathbf{1})^{\,3}}\,\mathrm{d}s \tag{41}$$

By using (stochastic) integration by parts, (41), the equation for $\rho_t(\varphi)$ and the Kallianpur–Striebel formula, we obtain

$$\begin{aligned} \frac{\rho_t^{U_n}(\varphi)}{\rho_t^{U_n}(\mathbf{1})} &= \pi_0(\varphi) + \int_0^{t\wedge U_n} \pi_s(A\varphi)\,\mathrm{d}s + \int_0^{t\wedge U_n} \pi_s((h^\top + B^\top)\varphi)\,\mathrm{d}Y_s - \int_0^{t\wedge U_n} \pi_s(\varphi)\pi_s(h^\top)\mathrm{d}Y_s \\ &\quad + \int_0^{t\wedge U_n} \pi_s(\varphi)\pi_s(h^\top)\pi_s(h)\,\mathrm{d}s - \int_0^{t\wedge U_n} \pi_s((h^\top + B^\top)\varphi)\,\pi_s(h)\mathrm{d}s \end{aligned}$$

As $\lim_{n\to\infty} U_n = \infty$ almost surely, we obtain the result by taking the limit as n tends to infinity. ∎

Remark 25 *The jump-diffusion example and the change detection model discussed in Section 3 both satisfy conditions (20) and (34). Therefore the two previous theorems can be applied to these two cases.*

References

[1] D. Applebaum, Lévy processes and stochastic calculus. Cambridge Studies in Advanced Mathematics, 93. Cambridge University Press, 2004

[2] A. Bain, D. Crisan, *Fundamentals of Stochastic Filtering*, Stochastic Modelling and Applied Probability, Vol 60, Springer Verlag, 2008.

[3] R. S. Bucy. Nonlinear filtering. IEEE Trans. Automatic Control, AC-10:198, 1965.

[4] J. M. C. Clark, D. Crisan, *On a robust version of the integral representation formula of nonlinear filtering*, Probab. Theory Related Fields 133, no. 1, 43-56, 2005.

[5] D. Crisan, B. Rozovsky, *The Oxford handbook of nonlinear filtering*, Oxford Univ. Press, Oxford, 2011.

[6] D. Dufresne, *The distribution of a perpetuity, with applications to risk theory and pension funding*, Scand. Actuar. J. 1990, no. 1-2, 39–79.

[7] T. E. Duncan, *Likelihood functions for stochastic signals in white noise,* Information and Control, 16 pages 303-310, 1970.

[8] M. Fujisaki, G. Kallianpur, and H. Kunita, *Stochastic differential equations for the non linear filtering problem*, Osaka J. Math., 9 pp 19-40, 1972.

[9] J. Jacod, A. N. Shiryaev, *Limit theorems for stochastic processes*, 288. Springer-Verlag, Berlin, 1987.

[10] G. Kallianpur, *Stochastic filtering theory*, Applications of Mathematics, 13. Springer-Verlag, New York-Berlin, 1980.

[11] H. Kushner, *On the differential equations satisfied by conditional densities of Markov processes, with applications.* SIAM J. Control, 2:106–119, 1964.

[12] H.J. Kushner, *Dynamical equations for optimal nonlinear filtering.* J. Differential Equations, 3:179–190, 1967.

[13] R. S. Liptser, A. N. Shiryaev, *Statistics of random processes. II. Applications,* Translated from the 1974 Russian original by A. B. Aries. Second, revised and expanded edition. Applications of Mathematics (New York), 6. Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin, 2001.

[14] R. S. Liptser, A. N. Shiryaev, *Statistics of random processes. I. General theory.* Translated from the 1974 Russian original by A. B. Aries. Second, revised and expanded edition. Applications of Mathematics (New York), 5. Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin, 2001.

[15] N. V. Krylov and B. L. Rozovsky. The Cauchy problem for linear stochastic partial differential equations. Izv. Akad. Nauk SSSR Ser. Mat., 41(6):1329–1347, 1448, 1977.

[16] N. V. Krylov and B. L. Rozovsky. Conditional distributions of diffusion processes. Izv. Akad. Nauk SSSR Ser. Mat., 42(2):356–378,470, 1978.

[17] T. Kailath. An innovations approach to least-squares estimation. I. Linear filtering in additive white noise. IEEE Trans. Autom. Control, AC-13:646–655, 1968.

[18] G. Kallianpur and C. Striebel (1968). Estimation of stochastic systems: arbitrary system process with additive noise observation errors. Ann. Math. Statist., 39, 785-801.

[19] G. Kallianpur and C. Striebel (1969). Stochastic differential equations occurring in the estimation of continuous parameter stochastic processes. Teor. Veroyatn. Primen., 14, no. 4, 597-622.

[20] R. E. Mortensen. *Stochastic optimal control with noisy observations*, Internat. J. Control, 1(4) pages 455-464, 1966.

[21] E. Pardoux. Equations aux dérivées partielles stochastiques non linéarires monotones. PhD thesis, Univ Paris XI, Orsay, 1975.

[22] D. Revuz, M. Yor, *Continuous martingales and Brownian motion.* Third edition, 293. Springer-Verlag, Berlin, 1999.

[23] L.C. G. Rogers, D. Williams, *Diffusions, Markov processes, and martingales.* Vol. 1. Reprint of the second (1994) edition. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000.

[24] A. N. Shiryaev, *On stochastic equations in the theory of conditional Markov process,* Teor. Verojatnost. i Primenen., Vol. 11, 200-206, 1966.

[25] R. L. Stratonovich. *Application of the theory of Markov processes for optimum filtration of signals.* Radio Eng. Electron. Phys, 1:119, 1960.

[26] R. L. Stratonovich. *Conditional Markov processes. Theory Probability Applications* , 5(2): pp 156-178, 1960.

[27] W. M. Wonham, *Some applications of stochastic differential equations to optimal nonlinear filtering.* J. Soc. Indust. Appl. Math. Ser. A Control 2 pp 347–369 (1965).

[28] J. Xiong, *An introduction to stochastic filtering theory. Oxford Graduate Texts in Mathematics*, 18. Oxford University Press, Oxford, 2008.

[29] M. Yor, *Exponential functionals of Brownian motion and related processes*, Springer (Berlin) , 2001.

[30] M. Zakai, *On the optimal filtering of diffusion processes,* Z. Wahrs.und Verw. Gebiete, 11 pp 230-243, 1969.