Ann. I. H. Poincaré – PR 39, 6 (2003) 919–941

10.1016/S0246-0203(03)00022-0/FLA

ASYMPTOTIC STABILITY, ERGODICITY AND OTHER ASYMPTOTIC PROPERTIES OF THE NONLINEAR FILTER

A. BUDHIRAJA

Department of Statistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

Received 22 June 2002, accepted 27 December 2002

ABSTRACT. – In this work we study connections between various asymptotic properties of the nonlinear filter. It is assumed that the signal has a unique invariant probability measure. The key property of interest is expressed in terms of a relationship between the observation σ field and the tail σ field of the signal, in the stationary filtering problem. This property can be viewed as the permissibility of the interchange of the order of the operations of maximum and countable intersection for certain σ-fields. Under suitable conditions, it is shown that the above property is equivalent to various desirable properties of the filter such as

(a) uniqueness of invariant measure for the signal,
(b) uniqueness of invariant measure for the pair (signal, filter),
(c) a finite memory property of the filter,
(d) a property of finite time dependence between the signal and observation σ fields and
(e) asymptotic stability of the filter.

Previous works on the asymptotic stability of the filter for a variety of filtering models then identify a rich class of filtering problems for which the above equivalent properties hold.

Keywords: Nonlinear filtering; Invariant measures; Asymptotic stability; Measure valued processes

RÉSUMÉ. – Dans cet article, nous étudions les relations entre différentes propriétés asymptotiques du filtre nonlinéaire. On suppose que le signal à une unique mesure de probabilité invariante. La propriété principale nous donne une relation entre la tribu engendrée par les observations et la tribu asymptotique du signal dans le problème à du filtrage stationnaire. Cette propriété peut être vue comme la possibilité de changer l'ordre des opérations d'intersection dénombrable et de maximum pour certaines tribus.

Sous des conditions adéquates, on peut montrer que cette derniére propriété est équivalente à certaines propriétés recherchées pour les filtres, telles que

(a) l'unicité de la mesure invariante du signal.
(b) l'unicité de la mesure invariante du couple (signal, filtre).
(c) Une propriété de mémoire finie du filtre.

E-mail address: budhiraj@email.unc.edu (A. Budhiraja).

(d) Une relation entre la tribu du signal et la tribu engendrée par les observations.
(e) la stabilité asymptotique du filtre.

1. Introduction

In this work we will consider the classical model of nonlinear filtering. Namely, we have a pair of stochastic processes $(X_t, Y_t)_{t\geqslant 0}$ where (X_t) is called the signal process and (Y_t) the observation process. The signal is taken to be a Markov process with values in some Polish space E and the observations are given via the relation:

$$Y_t = \int_0^t h(X_s)\,ds + W_t, \tag{1.1}$$

where (W_t) is a standard d-dimensional Brownian motion independent of (X_t) and h, referred to as the observation function, is a map from $E \to \mathbb{R}^d$. Tho goal of nonlinear filtering is the study of the measure valued process (Π_t) which is the conditional distribution of X_t given $\sigma\{Y_s\colon\ 0 \leqslant s \leqslant t\}$. This measure valued process is called the nonlinear filter. In the current work we are primarily interested in the ergodicity and stability properties of the nonlinear filter. In recent years such a study has generated significant interest [22,29,23,30,19,28,3,15,2,11,24,8,1,16,26,7,9,27,12,17,4, 18,10,6,25,14].

The problem of invariant measures for filtering processes was first considered by Kunita [22]. In this classic paper Kunita showed, using the uniqueness of the solution of the Kushner–Stratonovich equation, that in the above filtering model if the signal is Feller–Markov with a compact, separable Hausdorff state space E then the optimal filter is also a Feller–Markov process with state space $\mathcal{P}(E)$, where $\mathcal{P}(E)$ is the space of all probability measures on E. Furthermore, [22] shows that if the signal in addition has a unique invariant measure μ for which (2.13) holds then the filter (Π_t) has a unique invariant measure. In subsequent papers Kunita [23] and Stettner [29] extended the above results to the case where the state space is a locally compact Polish space. In the above papers [22,23,29] the observation function h is assumed to be bounded. In a recent paper [4] the results of Kunita–Stettner were extended to the case of unbounded h and signals with state space an arbitrary Polish space. The proofs in [4] are of independent interest since unlike the arguments in [22,29] they do not rely on the uniqueness of the solution to Kushner–Stratonovich equation. Using the results of Kunita [22], Ocone and Pardoux [28], in a pioneering paper, studied the problem of asymptotic stability of filters. Roughly speaking, the property of asymptotic stability says that the distance between the optimal filter and an incorrectly initialized filter converges to 0 as time approaches ∞. More precisely, for $\nu \in \mathcal{P}(E)$ denote by Q_ν the measure induced by (Y_t) on $\mathcal{C} \doteq C([0,\infty)\colon\ \mathbb{R}^d)$ (the space of all continuous maps from $[0,\infty)$ to $\mathbb{R}^d$), when the Markov process (X_t) has the initial law ν. One can show that for every $\nu \in \mathcal{P}(E)$ there exists a family of measurable maps $\{\Lambda_t(\nu)\}_{t\geqslant 0}$ from $\mathcal{C}$ to $\mathcal{P}(E)$ such that if μ_1 is the law of $X(0)$, then $\Lambda_t(\mu_1)(Y_\cdot(\omega))$ is the optimal nonlinear filter whereas for any other $\mu_2 \in \mathcal{P}(E)$, $\Lambda_t(\mu_2)(Y_\cdot(\omega))$ is a suboptimal filter which is constructed under the erroneous assumption that the initial law of the signal is μ_2 instead of μ_1. We say that the filter is (μ_1, μ_2) asymptotically stable if for all $\phi \in C_b(E)$ (the space of real continuous and bounded functions on E)

$$\mathbb{E}_{Q_{\mu_1}}\left[\langle \Lambda_t(\mu_1), \phi\rangle - \langle \Lambda_t(\mu_2), \phi\rangle\right]^2 \tag{1.2}$$

converges to 0 as $t \to \infty$, where $\mathbb{E}_{Q_{\mu_1}}$ denotes the expectation with respect to the measure Q_{μ_1}. With a somewhat different goal in mind, Delyon and Zeitouni [19] (in an earlier work than [28]) had also studied the dependence of the optimal filter on the initial condition. In recent years various authors have considered the problem of asymptotic stability under different hypothesis [28,3,15,2,11,24,8,1,16,26,12,17,4,18,10,25].

Recently, it has been pointed out [13] that there is a gap in the proof of Lemma 3.5 of [22] which is the key step in the proof of the uniqueness of the invariant measure for the filter. The difficulty, as will be described below, lies in the statement made just below equation (3.21) of that paper. The gap is of serious concern since some of the results in [29,28,4,6] directly appeal to the argument of the above lemma. The basic problem can be described as follows. For the rest of this section we will assume that the signal process has a unique invariant measure μ, namely Assumption 2.2 holds. Consider the family of σ-fields $(\mathcal{G}_s^t, \mathcal{Z}_s^t)_{-\infty<s<t<\infty}$, defined in (2.17) and (2.16) respectively. Basically, the σ-fields are obtained via a "stationary filtering problem" on $(-\infty, \infty)$, with the signal and observation processes $(\tilde{\xi}_t)_{-\infty<t<\infty}$, $(\alpha_t)_{-\infty<t<\infty}$ defined on some probability space $(\Omega^{(1)}, \mathcal{B}(\Omega^{(1)}), R_\mu^{(1)})$ (see Section 2 for the precise definitions and construction). $\mathcal{Z}_s^t$ is the $R_\mu^{(1)}$ completion of the σ-field generated by the observations: $(\alpha_u - \alpha_s;\ s \leqslant u \leqslant t)$ and $\mathcal{G}_s^t$ is the $R_\mu^{(1)}$ completion of the σ-field generated by the signal $(\tilde{\xi}_u)_{s\leqslant u\leqslant t}$. We will extend the definition of $(\mathcal{G}_s^t, \mathcal{Z}_s^t)$ for $s, t = \infty, -\infty$ in an obvious manner. So, for example, $\mathcal{G}_{-\infty}^{-\infty}$ is defined to be $\bigcap_{t=-\infty}^{-\infty} \bigcup_{s=-\infty}^{t} \mathcal{G}_s^t$. It is well known (cf. [31]) that under Assumption 2.3, $\mathcal{G}_{-\infty}^{-\infty}$ is $R_\mu^{(1)}$ trivial. Now the key difficulty is the following. In the course of the proof of Theorem 3.3 of [22] (which assumes that Assumptions 2.2 and 2.3 hold), the following result is used

$$\mathcal{Z}^0_{-\infty} = \bigcap_{t=-\infty}^{\infty} \left(\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^t_{-\infty}\right). \tag{1.3}$$

Eq. (1.3) can be viewed as the permissibility of the interchange of the max operation $(\vee)$ and the intersection operation $(\bigcap_{t=-\infty}^{\infty})$ for the σ-fields $\mathcal{Z}^0_{-\infty}$ and $\{\mathcal{G}^t_{-\infty}\}_{t\in\mathbb{R}}$.

However, [22] does not provide a proof for (1.3). Moreover, the above statement, in general, may not hold. By this we mean that one can construct a probability space $(\Omega_0, \mathcal{F}_0, P_0)$ with sub-σ-fields $\mathcal{F}^*$, $(\mathcal{G}_n^*)_{n\geqslant 0}$, such that all the σ fields are P_0 complete and $\{\mathcal{G}_n^*\}_{n\geqslant 0}$ is a decreasing sequence such that $\bigcap_{n\geqslant 0} \mathcal{G}_n^*$ is P_0 trivial, however, $\bigcap_{n\geqslant 0}(\mathcal{F}^* \vee \mathcal{G}_n^*)$ strictly contains $\mathcal{F}^*$. We refer the reader to [32] and [33] (Exercise 4.12) for very instructive examples where the equality of the two σ fields fails to hold. We were pointed to these examples by Chigansky and Liptser [13] who also found an error in the example given by [32]. The problem is rather delicate in view of the following. For a probability measure P_1 on $(\Omega_0, \mathcal{F}_0)$ and sub-σ-fields $\mathcal{H}_i,\ i = 1, 2$, of $\mathcal{F}_0$, we will say that $\mathcal{H}_1 = \mathcal{H}_2$ (mod P_1) if the two σ fields are equal, modulo P_1 null sets. The paper

[32] shows that if $\mathcal{G}_n^*$ is separable for all n then the problem of equality of the σ-fields $\bigcap_{n\geqslant 0}(\mathcal{F}^* \vee \mathcal{G}_n^*)$ and $\mathcal{F}^*$ is equivalent to the statement that $\bigcap_{n\geqslant 0}\mathcal{G}_n^*$ is $P_0(\cdot \mid \mathcal{F}^*)$ trivial, a.s. P_0. Now note that the P_0 triviality of $\bigcap_{n\geqslant 0}\mathcal{G}_n^*$ implies that

$$P_0(A \mid \mathcal{F}^*) = P_0(A) = 0 \text{ or } 1, \text{ a.s. } [P_0], \quad \forall A \in \bigcap_{n\geqslant 0}\mathcal{G}_n^*. \tag{1.4}$$

Now suppose that there is a separable σ field $\mathcal{H}$ such that

$$\bigcap_{n\geqslant 0}\mathcal{G}_n^* = \mathcal{H}\ \big(\text{mod } P_0(\cdot \mid \mathcal{F}^*)\big), \text{ a.s. } P_0. \tag{1.5}$$

Then (1.4) implies that

$$P_0(A \mid \mathcal{F}^*) = P_0(A) = 0 \text{ or } 1 \quad \forall A \in \bigcap_{n\geqslant 0}\mathcal{G}_n^*, \text{ a.s. } [P_0], \tag{1.6}$$

thus showing that $\bigcap_{n\geqslant 0}\mathcal{G}_n^*$ is $P_0(\cdot \mid \mathcal{F}^*)$ trivial, a.s. P_0. Therefore the key difficulty in establishing the equality of the two σ-fields is proving the separability property of $\bigcap_{n\geqslant 0}\mathcal{G}_n^*$ stated in (1.5).

The objective of this work is to show that the equality (1.3) is the central issue in the asymptotic study of the nonlinear filter. Our main result, Theorem 2.7, shows that, under suitable conditions (Assumptions 2.2, 2.3, 2.6), this equality is equivalent to various desirable properties of the nonlinear filter, such as: (a) uniqueness of invariant measure for the signal, (b) uniqueness of invariant measure for the pair (signal, filter), (c) finite memory property of the filter (see Definition 2.4), (d) finite time dependence of σ-fields $\{\mathcal{G}_{-\infty}^t\}_{t<0}$ with respect to the σ-field $\mathcal{Z}_{-\infty}^0$ (see Definition 2.5) and (e) asymptotic stability of the filter.

Of the above listed equivalent statements, the last statement, i.e., the asymptotic stability of the filters, has been the focus of most research. The first paper in this direction is [28] where asymptotic stability for Kalman filters was proved. In a sequence of papers Atar and Zeitouni [3,2] identify several important filtering problems, with a compact or countable state space for the signal, for which asymptotic stability holds. Other works on asymptotic stability for compact state space signals are [24,18]. The papers [11,12,25] study some signals in discrete time with non-compact state space for which asymptotic stability can be shown to hold. Atar [1] considers a continuous time filtering problem with non-compact state space and establishes asymptotic stability of the filter. Asymptotic stability for Benes filters is proved in [26]. In [8], asymptotic stability of the filter, for a compact state space signal model and point process observations, is proved.

The results of the above papers identify a rich family of filtering problems for which the equivalent conditions of Theorem 2.7 hold. However, we have been unable to prove the equality of the σ fields in (1.3) without making any additional assumptions than Assumptions 2.2, 2.3, 2.6. This remains a challenging open problem.

Finally we remark that the current work builds upon and borrows from several previous works. In particular, many arguments in this work are similar to those in [22, 23,29,28,4,6] and thus wherever possible we have referred the reader to previous papers.

2. Notation and the filtering model

Let E be a complete separable metric space and let $(\Omega, \mathcal{F}, P)$ be a probability space. Let (X_t) be a homogeneous Markov process with values in E with transition probability function $p(x, t, B)$, i.e. for $t, \tau > 0$, $x \in E$ and $B \in \mathcal{B}(E)$

$$P\big(X_{t+\tau} \in B \mid \sigma(X_u\colon u \leqslant \tau)\big) = p(X_\tau, t, B) \text{ a.s.,} \tag{2.1}$$

where for a Polish space S, $\mathcal{B}(S)$ denotes the Borel sigma field on S. Denote the distribution of X_0 by γ, i.e.

$$\gamma = P \circ (X_0)^{-1}. \tag{2.2}$$

Denote by $\mathcal{D} \doteq D([0,\infty), E)$, the Skorokhod space of E valued cadlag functions on $[0,\infty)$ and let $\xi_t(\cdot)$ be the coordinate process on $\mathcal{D}$, i.e. $\xi_t(\theta) \doteq \theta(t)$ for $\theta \in \mathcal{D}$.

We will assume that (X_t) admits a cadlag version, i.e. for all $(s, x) \in [0,\infty) \times E$ there exists a probability measure $P_{s,x}$ on $\mathcal{D}$ such that for $0 \leqslant s < t < \infty$, and $U \in \mathcal{B}(E)$,

$$P_{s,x}\big(\xi_t \in U \mid \sigma(\xi_u\colon u \leqslant s)\big) = p(\xi_s, t-s, U) \text{ a.s. } P_{s,x} \tag{2.3}$$

and

$$P_{s,x}(\xi_u = x,\ 0 \leqslant u \leqslant s) = 1. \tag{2.4}$$

For notational simplicity, $P_{0,x}$ will hereafter be denoted as P_x.

We will also assume that the Markov process is Feller, i.e. the map $x \to P_{s,x}$ is a continuous map from E to $\mathcal{P}(\mathcal{D})$, where for a Polish space S, $\mathcal{P}(S)$ denotes the space of probability measures on $(S, \mathcal{B}(S))$. Let (T_t) denote the semigroup corresponding to the Markov process (X_t), i.e. for $f \in BM(E)$ (for a Polish space S, $BM(S)$ denotes the space of bounded measurable functions on S),

$$(T_t f)(x) \doteq \int_{\mathcal{D}} f\big(\xi_t(\theta)\big)\, dP_x(\theta).$$

The Feller property of the Markov process gives that for $f \in C_b(E)$, $(T_t f) \in C_b(E)$.

The observation process is given as follows:

$$Y_t = \int_0^t h(X_u)\, du + W_t, \tag{2.5}$$

where $h : E \to \mathbb{R}^d$ is a continuous mapping and (W_t) is a $\mathbb{R}^d$-valued standard Wiener process, assumed to be independent of (X_t). Denote by Π_t the conditional distribution of X_t given past and current observations, i.e. for $A \in \mathcal{B}(E)$,

$$\Pi_t(A) \doteq P\big(X_t \in A \mid \sigma\{Y_u\colon 0 \leqslant u \leqslant t\}\big). \tag{2.6}$$

In order to study an incorrectly initialized filter we will introduce the following canonical setting, used in [4]. Let (β_t) be the canonical process on $\mathcal{C} \doteq C([0,\infty)\colon \mathbb{R}^d)$

(the space of continuous functions from $[0,\infty)$ to $\mathbb{R}^d$), i.e. $\beta_t(\eta) \doteq \eta(t)$ for $\eta \in \mathcal{C}$. Let Q be the standard Wiener measure on $(\mathcal{C}, \mathcal{B}(\mathcal{C}))$. Also set

$$(\widehat{\Omega}, \widehat{\mathcal{F}}) \doteq (\mathcal{D}, \mathcal{B}(\mathcal{D})) \otimes (\mathcal{C}, \mathcal{B}(\mathcal{C}))$$

and define for $\nu \in \mathcal{P}(E)$, $s > 0$

$$R_{s,\nu} \doteq P_{s,\nu} \otimes Q,$$

where $P_{s,\nu} \in \mathcal{P}(\mathcal{D})$ is defined as:

$$P_{s,\nu}(B) \doteq \int_E P_{s,x}(B)\, \nu(dx), \quad B \in \mathcal{B}(\mathcal{D}).$$

We will sometimes write $P_{0,\nu}$, $R_{0,\nu}$ as P_ν and R_ν respectively. Let $Z_t : \widehat{\Omega} \to \mathbb{R}$ be the stochastic process such that for all $0 \leqslant s \leqslant t$:

$$Z_t - Z_s = \int_s^t \langle h(\xi_u), d\beta_u \rangle, \text{ a.s. } R_{s,\nu}$$

for all $\nu \in \mathcal{P}(E)$, where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^d$. For the existence of such a common version see Theorem 3 in [21]. Next, for $0 \leqslant s \leqslant t$, let

$$q_{st} \doteq \exp\left(Z_t - Z_s - \frac{1}{2} \int_s^t \left\| h(\xi_u) \right\|^2 du \right).$$

For a Polish space S let $\mathcal{M}(S)$ denote the space of positive, finite measures on S. For $f \in BM(S)$ and $m \in \mathcal{M}(S)$ we will denote $\int_S f(x)\, dm(x)$ by $\langle m, f \rangle$ or $m(f)$.

For $\nu \in \mathcal{M}(E)$ and $0 \leqslant s \leqslant t < \infty$, define a $\mathcal{M}(E)$ valued process $\Gamma_{st}(\nu)$ on $\mathcal{C}$ as

$$\langle \Gamma_{st}(\nu)(\eta), f \rangle \doteq \int_E \int_{\mathcal{D}} f(\xi_t(\theta)) q_{st}(\theta, \eta)\, dP_{s,x}(\theta)\, d\nu(x), \ \eta\text{-a.s. } [Q]. \tag{2.7}$$

The measurability of the map $(s, t, \eta, \nu) \to \Gamma_{st}(\nu)(\eta)$ is a consequence of Theorem 3 in [21] which gives the measurability of the map $(t, \omega) \to Z_t(\omega)$.

Finally define for $0 \leqslant s \leqslant t$ and $\nu \in \mathcal{M}(E)$ a $\mathcal{P}(E)$ valued random variable $\Lambda_{st}(\nu)$ via the normalization of $\Gamma_{st}(\nu)$, i.e.

$$\Lambda_{st}(\nu) \doteq \frac{\Gamma_{st}(\nu)}{\langle \Gamma_{st}(\nu), 1 \rangle}.$$

Also, we define $\Gamma_t(\nu) \doteq \Gamma_{0t}(\nu)$ and $\Lambda_t(\nu) \doteq \Lambda_{0t}(\nu)$.

As a consequence of the Kallianpur–Striebel formula (see [20]) it follows that for $f \in BM(E)$

$$\langle \Pi_t(\omega), f \rangle = \langle \Lambda_{0t}(\gamma)(Y_\cdot(\omega)), f \rangle, \ \omega\text{-a.s. } [P], \ t \in (0, \infty). \tag{2.8}$$

By a filter initialized incorrectly at the probability measure γ_1 we mean the $\mathcal{P}(E)$ valued process, $\Pi_t^{\gamma_1}$ defined as:

$$\langle \Pi_t^{\gamma_1}(\omega), f\rangle \doteq \langle \Lambda_{0t}(\gamma_1)(Y_\cdot(\omega)), f\rangle, \quad f \in BM(E).$$

Let $\widetilde{\mathcal{F}}$ be the Q-completion of $\mathcal{B}(\mathcal{C})$ and $\widetilde{\mathcal{N}}$ be the class of Q-null sets in $\widetilde{\mathcal{F}}$. For $0 \leqslant s \leqslant t \leqslant \infty$, let $\mathcal{A}_s^t$ be the sub-σ-fields of $\widetilde{\mathcal{F}}$ defined by

$$\mathcal{A}_s^t = \sigma\big(\sigma(\beta_u - \beta_s\colon s \leqslant u \leqslant t) \cup \widetilde{\mathcal{N}}\big). \tag{2.9}$$

Next we introduce the probability measure on $\mathcal{C}$ under which the canonical process has the same law as the observation process. For an arbitrary $\nu \in \mathcal{P}(E)$ let $Q_\nu \in \mathcal{P}(\mathcal{C})$ be defined by

$$\frac{dQ_\nu}{dQ} = \Gamma_t(\nu)(E) \quad \text{on } \mathcal{A}_0^t,\ t \in [0,\infty). \tag{2.10}$$

It is easy to see that $P \circ Y^{-1} = Q_\gamma$.

We now define our basic notion of asymptotic stability which was introduced by Ocone and Pardoux [28].

DEFINITION 2.1. – *Let $\mu_1, \mu_2 \in \mathcal{P}(E)$. We say that the filter is (μ_1,μ_2)-asymptotically stable if for all $\phi \in C_b(E)$*

$$\big|\langle \Lambda_t(\mu_1), \phi\rangle - \langle \Lambda_t(\mu_2), \phi\rangle\big|$$

converges to 0 *in Q_{μ_1}-probability as $t \to \infty$.*

We now proceed to describe the Markov properties of the filter. It was shown in [4] that $\{\pi_t\}_{t\geqslant 0}$ is a Feller–Markov process with associated semigroup:

$$(\mathcal{T}_t F)(\nu) \doteq \mathbb{E}_{Q_\nu}\big(F\big(\Lambda_t(\nu)\big)\big); \quad F \in BM\big(\mathcal{P}(E)\big);\ \nu \in \mathcal{P}(E).$$

In order to describe the Markov property of the (signal, filter) pair, we now introduce the following measure on $(\widehat{\Omega}, \widehat{\mathcal{F}})$ which corresponds to the law of the process $(X_t, Y_t)_{t\geqslant 0}$. For $\nu \in \mathcal{P}(E)$ define

$$\mathcal{K}_s^t(\nu) \doteq \sigma\big(\sigma\{\beta_u - \beta_s\colon s \leqslant u \leqslant t\} \cup \sigma\{\xi_u\colon s \leqslant u \leqslant t\} \cup \mathcal{N}\big), \tag{2.11}$$

where $\mathcal{N}$ is the class of all $R_{0,\nu}$ null sets. Now for fixed $\nu \in \mathcal{P}(E)$ define $\widehat{R}_{0,\nu}$ on $(\widehat{\Omega}, \widehat{\mathcal{F}})$ as follows:

$$\frac{d\widehat{R}_{0,\nu}}{dR_{0,\nu}}(\theta,\eta) \doteq q_{0t}(\theta,\eta) \quad \text{on } \mathcal{K}_0^t(\nu),\ t \geqslant 0. \tag{2.12}$$

Observing that

$$R_{0,\gamma} = P \circ (X_\cdot, Y_\cdot)^{-1},$$

it follows via an application of Girsanov's theorem (cf. [4], Section 4) that

$$\widehat{R}_{0,\gamma} = P \circ (X_\cdot, W_\cdot)^{-1}.$$

Now, for fixed $\nu \in \mathcal{P}(E)$ let Π_t^ν be the filter initialized at ν, defined as

$$\Pi_t^\nu(\omega) \doteq \Lambda_t(\nu)\big(Y_\cdot(\omega)\big).$$

It was shown in [4] that $((X_t, \Pi_t^\nu), \mathcal{F}_t)$ is a $E \times \mathcal{P}(E)$ valued Feller–Markov process on $(\Omega, \mathcal{F}, P)$ with associated semigroup $\{\mathcal{S}_t\}_{0 \leqslant t < \infty}$ defined as follows. For $F \in BM(E \times \mathcal{P}(E))$,

$$(\mathcal{S}_t F)(x, \lambda) \doteq \mathbb{E}_{\widehat{R}_{0,x}}\big[F(\Lambda_t(\lambda), \xi_t)\big]; \quad (x, \lambda) \in E \times \mathcal{P}(E).$$

The following two conditions will be assumed in many results of this paper.

Assumption 2.2. – There is a unique invariant probability measure, μ, for the semigroup (T_t).

Assumption 2.3. – For all $f \in C_b(E)$:

$$\limsup_{t \to \infty} \int_E \big|T_t f(x) - \langle \mu, f \rangle\big| \, \mu(dx) = 0. \tag{2.13}$$

The above assumption in the asymptotic study of the filter was introduced by Kunita [22], it is equivalent to the statement that $\mathcal{G}_{-\infty}^{-\infty}$ is $R_\mu^{(1)}$ – trivial (cf. [31]).

Following Stettner [29], denote for $\nu \in \mathcal{P}(E)$ and $A \in \mathcal{B}(\mathcal{P}(E))$,

$$m_t^\nu(A) \doteq (\mathcal{T}_t \mathcal{I}_A)(\nu) = E_{Q_\nu}\big(\mathcal{I}_A\big(\Lambda_t(\nu, \cdot)\big)\big) \tag{2.14}$$

and

$$M_t^\nu(A) \doteq \int_E (\mathcal{T}_t \mathcal{I}_A)(\delta_x) \, \nu(dx), \tag{2.15}$$

where $\mathcal{I}_A$ is the indicator function of the set A. We will now give representations for m_t^μ and M_t^μ as the laws of certain filtering processes in the stationary filtering problem introduced by Kunita [22]. Henceforth we will assume that Assumption 2.2 holds, i.e., there is a unique (T_t) invariant measure, μ. Let $\mathcal{D}_\mathbb{R} \equiv D((-\infty, \infty); E)$ denote the space of r.c.l.l. functions from $(-\infty, \infty)$ into E with Skorokhod topology and $\mathcal{C}_\mathbb{R} \equiv C((-\infty, \infty); \mathbb{R}^d)$ denote the space of continuous functions from $(-\infty, \infty)$ into $\mathbb{R}^d$ with topology of uniform convergence on compact subsets of $(-\infty, \infty)$. Let the coordinate processes on $\mathcal{D}_\mathbb{R}$ and $\mathcal{C}_\mathbb{R}$ be denoted by $(\tilde{\xi}_t(\cdot))$ and $(\tilde{\beta}_t(\cdot))$ respectively. Let $P_\mu^{(1)}$ be the unique measure on $(\mathcal{D}_\mathbb{R}, \mathcal{B}(\mathcal{D}_\mathbb{R}))$ which satisfies for $E_1, \ldots, E_n \in \mathcal{B}(\mathbb{R})$ and $-\infty < t_1 < t_2 < \cdots < t_n < \infty$,

$$P_\mu^{(1)}(\tilde{\xi}_{t_1} \in E_1, \ldots, \tilde{\xi}_{t_n} \in E_n) = \int_{E_1 \times \cdots \times E_n} \mu(dx_1) p(t_1, x_1, t_2, dx_2) \cdots p(t_{n-1}, x_{n-1}, t_n, dx_n).$$

Now let $Q^{(1)}$ be a probability measure on $(C_{\mathbb{R}}, \mathcal{B}(C_{\mathbb{R}}))$, such that for $-\infty < t_0 < t_1 < \cdots < t_n < \infty$,

$$\left(\frac{1}{\sqrt{t_1 - t_0}}(\tilde{\beta}_{t_1} - \tilde{\beta}_{t_0}), \ldots, \frac{1}{\sqrt{t_n - t_{n-1}}}(\tilde{\beta}_{t_n} - \tilde{\beta}_{t_{n-1}})\right)$$

are independent $N(0, I_{d\times d})$.

Let $\Omega^1 \doteq \mathcal{D}_{\mathbb{R}} \times C_{\mathbb{R}}$ and $R_\mu^{(1)} \doteq P_\mu^{(1)} \otimes Q^{(1)}$. Without loss of generality, we will consider the coordinate processes $(\tilde{\xi}_t)$, $(\tilde{\beta}_t)$ to be defined on the product space $(\Omega^1, \mathcal{B}(\Omega^1), R_\mu^{(1)})$. Let $\mathcal{F}^*$ be the completion of $\mathcal{B}(\Omega^1)$ under $R_\mu^{(1)}$. Define the observation process:

$$\alpha_t - \alpha_s \doteq \int_s^t h(\tilde{\xi}_u)\,du + \tilde{\beta}_t - \tilde{\beta}_s$$

and the sigma fields

$$\mathcal{Z}_s^t \doteq \sigma\left(\sigma(\alpha_v - \alpha_u; s \leqslant u \leqslant v \leqslant t) \cup \mathcal{N}^*\right), \tag{2.16}$$

$$\mathcal{G}_s^t = \sigma\left(\sigma(\tilde{\xi}_u\colon s \leqslant u \leqslant t) \cup \mathcal{N}^*\right), \tag{2.17}$$

where $-\infty \leqslant s < t \leqslant \infty$ and $\mathcal{N}^*$ is the class of $R_\mu^{(1)}$ null sets in $\mathcal{F}^*$. Further, let $\mathcal{G}_{-\infty}^{-\infty}$ be defined as

$$\mathcal{G}_{-\infty}^{-\infty} = \bigcap_{-\infty < t < \infty} \mathcal{G}_{-\infty}^t. \tag{2.18}$$

Now define for $-\infty < s < t < \infty$,

$$\bar{\pi}_{s,t}^{(0)} \doteq \Lambda_{t-s}(\mu)\left(\alpha^s\right),$$

where $\alpha^s : \Omega^1 \to C([0, \infty); \mathbb{R}^d)$ is defined as $\alpha_u^s(\omega) \doteq \alpha_{s+u}(\omega) - \alpha_s(\omega)$. Also define

$$\bar{\pi}_{s,t}^{(1)} \doteq \Lambda_{t-s}(\delta_{\tilde{\xi}_s})\left(\alpha^s\right).$$

Observe that for f in $BM(E)$

$$\bar{\pi}_{s,t}^{(0)}(f) = \mathbb{E}_{R_\mu^{(1)}}\left[f(\tilde{\xi}_t) \mid \mathcal{Z}_s^t\right] \quad \text{and} \quad \bar{\pi}_{s,t}^{(1)}(f) = \mathbb{E}_{R_\mu^{(1)}}\left[f(\tilde{\xi}_t) \mid \mathcal{Z}_s^t \vee \sigma(\tilde{\xi}_s)\right].$$

(For two sigma fields $\mathcal{L}_1$ and $\mathcal{L}_2$, $\mathcal{L}_1 \vee \mathcal{L}_2 \doteq \sigma(\mathcal{L}_1 \cup \mathcal{L}_2)$.) Also note that for F in $BM(\mathcal{P}(E))$

$$\begin{aligned}
\mathbb{E}_{R_\mu^{(1)}}\left[F\left(\bar{\pi}_{s,t}^{(1)}\right)\right] &= \mathbb{E}_{R_\mu^{(1)}}\left[F\left(\Lambda_{t-s}(\delta_{\tilde{\xi}_s})\left(\alpha^s\right)\right)\right] \\
&= \int_E \mathbb{E}_{Q_x} F\left(\Lambda_{t-s}(\delta_x)\right)\mu(dx) \\
&= \int_E (\mathcal{T}_{t-s}F)(\delta_x)\,\mu(dx) \qquad (2.19) \\
&= M_{t-s}^\mu(F). \qquad (2.20)
\end{aligned}$$

In a similar manner it is seen that

$$\mathbb{E}_{R_\mu^{(1)}}\big[F\big(\bar{\pi}_{s,t}^{(0)}\big)\big] = m_{t-s}^{\mu}(F). \tag{2.21}$$

A straightforward application of martingale convergence theorem shows that as $s \to -\infty$, almost surely the measure $\bar{\pi}_{s,t}^{(0)}$ converges (weakly) to the measure $\bar{\pi}_t^{(0)}$ defined as follows: For bounded and continuous function f on E

$$\bar{\pi}_t^{(0)}(f) \doteq \mathbb{E}_{R_\mu^{(1)}}\big[f(\tilde{\xi}_t) \mid \mathcal{Z}_{-\infty}^t\big]. \tag{2.22}$$

Furthermore we have that (cf. Lemma 3.3 of Kunita [22])

$$\bar{\pi}_{s,t}^{(1)}(f) = \mathbb{E}_{R_\mu^{(1)}}\big[f(\tilde{\xi}_t) \mid \mathcal{Z}_{-\infty}^t \vee \mathcal{G}_{-\infty}^s\big]$$

and thus by the reverse martingale convergence theorem we have that as $s \to -\infty$, $\bar{\pi}_{s,t}^{(1)}$ converges weakly to the measure $\bar{\pi}_t^{(1)}$ defined as

$$\bar{\pi}_t^{(1)}(f) \doteq \mathbb{E}_{R_\mu^{(1)}}\left[f(\tilde{\xi}_t) \mid \bigcap_{s=-\infty}^{\infty} \big(\mathcal{Z}_{-\infty}^t \vee \mathcal{G}_{-\infty}^s\big)\right]. \tag{2.23}$$

It is stated in [22] that under Assumptions 2.2 and 2.3, the above expectation is equal to $\mathbb{E}_{R_\mu^{(1)}}[f(\tilde{\xi}_t) \mid \mathcal{Z}_{-\infty}^t]$. However, we have been unable to prove that statement.

In view of (2.20) and (2.21) we have that M_u^{μ} and m_u^{μ} converge weakly as $u \to \infty$ to the law of $\bar{\pi}_t^{(0)}$ and $\bar{\pi}_t^{(1)}$ respectively, which also shows that the laws of $\bar{\pi}_t^{(0)}, \bar{\pi}_t^{(1)}$ are independent of t. Denote these laws as m^{μ} and M^{μ} respectively. Thus we have that

$$m_u^{\mu} \to m^{\mu}; \quad M_u^{\mu} \to M^{\mu}, \quad \text{as } u \to \infty. \tag{2.24}$$

Also note that since $(\mathcal{T}_t)$ is a Feller semigroup, it follows from (2.14) and (2.15) that both m^{μ} and M^{μ} have to be $(\mathcal{T}_t)$ invariant. This shows that there is at least one $(\mathcal{T}_t)$ invariant probability measure. In a similar way it is shown (cf. [4]) that there is at least one (S_t) invariant probability measure.

We now introduce the property of "finite memory of the filter". This property says that for large times (t), the filter initialized at any point $x \in E$ can be well approximated by a sub-optimal filter which is constructed using only the observations from the past τ units of time, for sufficiently large τ. This property was introduced in the filter stability problem by Ocone and Pardoux [28].

DEFINITION 2.4. – *We say that the filter has the finite memory property if for all $\phi \in C_b(E)$*

$$\limsup_{\tau\to\infty} \limsup_{t\to\infty} \mathbb{E}_{Q_{\delta_x}} \big|\big\langle \Lambda_t(\delta_x), \phi\big\rangle - \big\langle \Lambda_{t-\tau,t}(\delta_x T_{t-\tau}), \phi\big\rangle\big| = 0;\ x\text{-a.s. } [\mu]. \tag{2.25}$$

Finally, we give the following definition regarding the dependence between the signal and the observation process.

DEFINITION 2.5. – *We say that the sequence of σ-fields $\{\mathcal{G}^t_{-\infty}\}_{t<0}$ has a finite time dependence with respect to the σ-field $\mathcal{Z}^0_{-\infty}$, if for all $k \in \mathbb{N}$, $-\infty < t_1 < t_2 < \cdots < t_k < \infty$, $\phi_1, \ldots, \phi_k \in C_b(E)$ and $\varepsilon > 0$, there exists $\tau_\varepsilon \in (-\infty, t_1)$ and $t_\varepsilon < \tau_\varepsilon$ such that $\forall t \leqslant t_\varepsilon$*

$$\mathbb{E}_{R^{(1)}_\mu}\left|\mathbb{E}_{R^{(1)}_\mu}\left[\prod_{i=1}^k \phi_i(\tilde{\xi}_{t_i}) \,\Big|\, \mathcal{Z}^0_{-\infty} \vee \mathcal{G}^t_{-\infty}\right] - \mathbb{E}_{R^{(1)}_\mu}\left[\prod_{i=1}^k \phi_i(\tilde{\xi}_{t_i}) \,\Big|\, \mathcal{Z}^0_{\tau_\varepsilon} \vee \mathcal{G}^t_{-\infty}\right]\right| \leqslant \varepsilon.$$

The following assumption will be made at some places in this work.

Assumption 2.6. – For all $\nu_1, \nu_2 \in \mathcal{P}(E)$ there exists $t \in [0, \infty)$ such that $\nu_1 T_t$ is absolutely continuous with respect to $\nu_2 T_t$.

We can now state the main result in this work.

THEOREM 2.7. – *Suppose that Assumptions* 2.2, 2.3 *and* 2.6 *hold. Then the following are equivalent.*

(i) $\bigcap_{s=-\infty}^{\infty}(\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^s_{-\infty}) = \mathcal{Z}^0_{-\infty} \vee \mathcal{G}^{-\infty}_{-\infty}$.
(ii) $m^\mu = M^\mu$.
(iii) *The filter has a unique invariant measure, i.e. there is a unique $\mathcal{T}_t$ invariant probability measure.*
(iv) *The signal-filter pair has a unique invariant measure, i.e. there is a unique S_t invariant probability measure.*
(v) *For all $\nu_1, \nu_2 \in \mathcal{P}(E)$, the filter is (ν_1, ν_2) asymptotically stable.*
(vi) *The filter is (δ_x, μ) asymptotically stable for μ-almost every $x \in E$.*
(vii) *The filter has the finite memory property* (*Definition* 2.4)*.*
(viii) *The σ-fields $\{\mathcal{G}^t_{-\infty}\}_{t<0}$ have finite time dependence with respect to the σ-field $\mathcal{Z}^0_{-\infty}$* (*Definition* 2.5)*.*

Remark 2.8. – We show, in fact, a stronger result. Namely under Assumption 2.2 (alone), (v) $\Rightarrow$ (vi); (viii) $\Rightarrow$ (i); and (v) $\Rightarrow$ (iv) $\Rightarrow$ (ii) $\Rightarrow$ (iii). Assuming 2.3 in addition, (i) $\Rightarrow$ (ii). Finally, assuming 2.6 in addition (iii) $\Rightarrow$ (v) and (vi) $\Rightarrow$ (vii) $\Rightarrow$ (viii).

Proof. – (i) $\Rightarrow$ (ii) (Kunita [22]). Observe that m^μ is the probability law of $\bar{\pi}^{(1)}_0$ and M^μ is the probability law of $\bar{\pi}^{(0)}_0$. Assumption 2.3 implies that $\mathcal{G}^{-\infty}_{-\infty}$ is $R^{(1)}_\mu$ trivial. This combined with (i) immediately gives, in view of (2.22) and (2.23) that $\bar{\pi}^{(1)}_0 = \bar{\pi}^{(0)}_0$ a.s. and thus $m^\mu = M^\mu$.

(ii) $\Rightarrow$ (iii) (Kunita [22]). If Φ is some other $(\mathcal{T}_t)$ invariant probability measure, it follows (Theorem 3.2 [22]) that $m^\mu(F) \leqslant \Phi(F) \leqslant M^\mu(F)$ for all convex bounded and continuous functions F on $\mathcal{M}(E)$. Now (ii) implies that $m^\mu(F) = \Phi(F) = M^\mu(F)$ for all such F. Since the collection of all such F is a measure determining class, we have that $m^\mu = \Phi = M^\mu$.

(iii) $\Rightarrow$ (v) (Ocone and Pardoux [28]). Corollary 3.2 gives that $\nu_i T_t \to \mu$, for $i = 1, 2$, as $t \to \infty$. Also Assumption 2.6 implies that Q_{ν_1} is mutually absolutely continuous with respect to Q_{ν_2} (cf. Corollary 3.4 of [6]). Now Theorem 3.2 of [28] yield the implication. We remark that [28] considers the case when E is locally compact and h is bounded, however the result holds in the current setting as was shown in Theorem 7.3 of [4].

(v) ⇒ (vi) This is immediate.
(vi) ⇒ (vii) See Theorem 3.3.
(vii) ⇒ (viii) See Theorem 3.6.
(viii) ⇒ (i) See Theorem 3.8.
(v) ⇒ (iv) [7,6]. This is a direct consequence of Theorem 3.6 of [6].
(iv) ⇒ (ii) It is shown in Theorem 6.4 of [4] that if we define

$$\overline{m}^{\mu} \doteq R_{\mu}^{(1)} o\big(\bar{\pi}_0^{(1)}, \xi_0\big)^{-1}, \qquad \overline{M}^{\mu} \doteq R_{\mu}^{(1)} o\big(\bar{\pi}_0^{(0)}, \xi_0\big)^{-1},$$

then both $\overline{m}^{\mu}$ and $\overline{M}^{\mu}$ are (S_t) invariant. Now (iv) implies that $\overline{m}^{\mu} = \overline{M}^{\mu}$ and thus $(\overline{m}^{\mu})_1 = (\overline{M}^{\mu})_1$, where for $\rho \in \mathcal{P}(\mathcal{P}(E) \times E)$, we denote by $(\rho)_1$ the marginal on $\mathcal{P}(E)$. Since $(\overline{m}^{\mu})_1 = m^{\mu}$ and $(\overline{M}^{\mu})_1 = M^{\mu}$, (ii) follows. □

3. Proofs

Throughout this section we will assume that Assumption 2.2 holds. The following lemma will be used in the proof of Theorem 3.3.

LEMMA 3.1. – *Suppose that Assumption* 2.3 *holds. Let* $\nu \in \mathcal{P}(E)$ *be such that* $\nu T_\varepsilon \ll \mu$ *for some* $\varepsilon > 0$. *Then* $\nu T_t \to \mu$ *as* $t \to \infty$.

Proof. – We begin by noticing that for $f \in C_b(E)$ and $t \geqslant \varepsilon$,

$$(\nu T_t)(f) = \int_E (T_{t-\varepsilon} f)(x) \left(\frac{d\nu T_\varepsilon}{d\mu}(x) \right) \mu(dx).$$

This implies that for all $K \in (0, \infty)$

$$\begin{aligned} \big|(\nu T_t)(f) - \mu(f)\big| &= \left| \int_E \big((T_{t-\varepsilon} f)(x) - \mu(f)\big) \left(\frac{d\nu T_\varepsilon}{d\mu}(x) \right) \mu(dx) \right| \\ &\leqslant K \int_E \big|(T_{t-\varepsilon} f)(x) - \mu(f)\big| \, \mu(dx) \\ &\quad + 2 \sup_{x \in E} \big| f(x)\big| \int_E \left(\frac{d\nu T_\varepsilon}{d\mu}(x) \right) I_{\frac{d\nu T_\varepsilon}{d\mu}(x) > K} \, \mu(dx). \end{aligned}$$

Taking limit as $t \to \infty$, we have from Assumption 2.3 that

$$\limsup_{t \to \infty} \big|(\nu T_t)(f) - \mu(f)\big| \leqslant 2 \sup_{x \in E} \big| f(x)\big| \int_E \left(\frac{d\nu T_\varepsilon}{d\mu}(x) \right) I_{\frac{d\nu T_\varepsilon}{d\mu}(x) > K} \, \mu(dx).$$

The result now follows on taking limit as $K \to \infty$ in the above display. □

As an immediate consequence of the above result we have the following corollary.

COROLLARY 3.2. – *Suppose that Assumptions* 2.3, 2.6 *hold. Then for all* $\nu \in \mathcal{P}(E)$, $\nu T_t \to \mu$, *as* $t \to \infty$.

We now give the proof of the statement in Theorem 2.7 that (vi) implies (vii).

THEOREM 3.3. – *Suppose that Assumption* 2.3 *and* 2.6 *hold. Further suppose that the filter is* (δ_x, μ) *asymptotically stable for* μ*-almost every* $x \in E$. *Then the filter has the finite memory property in the sense of Definition* 2.4.

Proof. – Fix $\phi \in C_b(E)$. In view of Definition 2.4 it suffices to show (2.25). We begin by observing that from Definition 2.1 it follows that for μ almost every $x \in E$

$$\lim_{t\to\infty} \mathbb{E}_{Q_{\delta_x}} \left|\langle \Lambda_t(\delta_x), \phi\rangle - \langle \Lambda_t(\mu), \phi\rangle\right| = 0. \tag{3.1}$$

Next, following [28], we have that

$$\begin{aligned}
&\mathbb{E}_{Q_\mu}\left|\langle \Lambda_t(\mu), \phi\rangle - \langle \Lambda_{t-\tau,t}(\mu T_{t-\tau}), \phi\rangle\right|^2 \\
&\quad = \mathbb{E}_{Q_\mu}\left(\langle \Lambda_t(\mu), \phi\rangle\right)^2 + \mathbb{E}_{Q_\mu}\left(\langle \Lambda_{t-\tau,t}(\mu T_{t-\tau}), \phi\rangle\right)^2 \\
&\qquad - 2\mathbb{E}_{Q_\mu}\left(\langle \Lambda_t(\mu), \phi\rangle\langle \Lambda_{t-\tau,t}(\mu T_{t-\tau}), \phi\rangle\right) \\
&\quad = \mathbb{E}_{Q_\mu}\left(\langle \Lambda_t(\mu), \phi\rangle\right)^2 - \mathbb{E}_{Q_\mu}\left(\langle \Lambda_{t-\tau,t}(\mu T_{t-\tau}), \phi\rangle\right)^2 \\
&\quad = \mathbb{E}_{Q_\mu}\left(\langle \Lambda_t(\mu), \phi\rangle\right)^2 - \mathbb{E}_{Q_\mu}\left(\langle \Lambda_\tau(\mu), \phi\rangle\right)^2 \\
&\quad = m_t^\mu(F_\phi) - m_\tau^\mu(F_\phi),
\end{aligned}$$

where the last step follows from (2.14) and $F_\phi \in \mathcal{C}_b(\mathcal{P}(E))$ is defined as $F_\phi(\nu) \doteq \langle \nu, \phi\rangle^2$, $\nu \in \mathcal{P}(E)$. Taking limit as $t \to \infty$ and then $\tau \to \infty$ in the above display, we have from (2.24) that

$$\limsup_{\tau\to\infty} \limsup_{t\to\infty} \mathbb{E}_{Q_\mu}\left|\langle \Lambda_t(\mu), \phi\rangle - \langle \Lambda_{t-\tau,t}(\mu T_{t-\tau}), \phi\rangle\right| = 0. \tag{3.2}$$

Next, from Assumption 2.6 it follows that $Q_{\delta_x} \ll Q_\mu$. For a proof of this statement we refer the reader to Corollary 3.4 of [6]. Thus observing that

$$\left|\langle \Lambda_t(\mu), \phi\rangle - \langle \Lambda_{t-\tau,t}(\mu T_{t-\tau}), \phi\rangle\right| \leqslant 2\sup_{x\in E}\left|\phi(x)\right|,$$

we have from (3.2) that $\forall x \in E$,

$$\limsup_{\tau\to\infty} \limsup_{t\to\infty} \mathbb{E}_{Q_{\delta_x}}\left|\langle \Lambda_t(\mu), \phi\rangle - \langle \Lambda_{t-\tau,t}(\mu T_{t-\tau}), \phi\rangle\right| = 0. \tag{3.3}$$

We now note that

$$\begin{aligned}
&\mathbb{E}_{Q_{\delta_x}}\left|\langle \Lambda_t(\delta_x), \phi\rangle - \langle \Lambda_{t-\tau,t}(\delta_x T_{t-\tau}), \phi\rangle\right| \\
&\quad \leqslant \mathbb{E}_{Q_{\delta_x}}\left|\langle \Lambda_t(\delta_x), \phi\rangle - \langle \Lambda_t(\mu), \phi\rangle\right| \\
&\qquad + \mathbb{E}_{Q_{\delta_x}}\left|\langle \Lambda_t(\mu), \phi\rangle - \langle \Lambda_{t-\tau,t}(\mu T_{t-\tau}), \phi\rangle\right| \\
&\qquad + \mathbb{E}_{Q_{\delta_x}}\left|\langle \Lambda_{t-\tau,t}(\mu T_{t-\tau}), \phi\rangle - \langle \Lambda_{t-\tau,t}(\delta_x T_{t-\tau}), \phi\rangle\right|
\end{aligned} \tag{3.4}$$

From (3.1), (3.3) and (3.4) it follows that, in order to prove (2.25) it suffices to show that for μ almost every x and every fixed τ

$$\mathbb{E}_{Q_{\delta_x}}\left|\langle \Lambda_{t-\tau,t}(\mu T_{t-\tau}), \phi\rangle - \langle \Lambda_{t-\tau,t}(\delta_x T_{t-\tau}), \phi\rangle\right| \tag{3.5}$$

converges to 0 as $t \to \infty$. However (3.5) can be rewritten as

$$\mathbb{E}_{Q_{\delta_x T_{t-\tau}}} \left| \langle \Lambda_\tau(\mu T_{t-\tau}), \phi \rangle - \langle \Lambda_\tau(\delta_x T_{t-\tau}), \phi \rangle \right|. \tag{3.6}$$

Next, Corollary 3.2 gives that, $\delta_x T_{t-\tau}$ converges to μ as $t \to \infty$. Also, $\mu T_{t-\tau}$, being equal to μ, trivially converges to μ as $t \to \infty$. It now follows from Theorem 7.2 of [4] that the expression in (3.6), and therefore the expression in (3.5), converges to 0, as $t \to \infty$. This proves the theorem. □

The following lemma and the proposition following it will be used in the proof of Theorem 3.6.

LEMMA 3.4. – *Let $\{\nu_n\}$ be a sequence in $\mathcal{P}(E)$ such that $\nu_n \to \nu$ for some $\nu \in \mathcal{P}(E)$. Then for all $a \in (0, \infty)$*

$$\limsup_{K\to\infty} \limsup_{n\to\infty} \int_{\widehat{\Omega}} q_{0,a}(\theta, \eta) 1_{q_{0,a}(\theta,\eta)>K} \, dR_{\nu_n}(\theta, \eta) = 0.$$

Proof. – The proof of this result is contained in Theorem 5.1 of [4] and Theorem 3.2 of [5], however we sketch the argument for the sake of completeness.

Since $\nu_n \to \nu$ weakly as $n \to \infty$, the Feller property of (T_t) implies that $P_{\nu_n} \to P_\nu$ weakly as $n \to \infty$. Now let $(\widetilde{X}^n_t)$ and $(\widetilde{X}_t)$ be processes with values in $\mathcal{D}$ defined on some probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$ such that $\overline{P} \circ (\widetilde{X}^n_\cdot)^{-1} = P_{\nu_n}$, $\overline{P} \circ (\widetilde{X}_\cdot)^{-1} = P_\nu$ and $\widetilde{X}^n_\cdot \to \widetilde{X}_\cdot$ a.s. $\overline{P}$. Define

$$(\Omega_0, \mathcal{F}_0, R) \doteq \left(\overline{\Omega} \times \mathcal{C}, \overline{\mathcal{F}} \otimes \mathcal{B}(\mathcal{C}), \overline{P} \otimes Q \right)$$

and the processes $Z^n_\cdot$, $Z_\cdot$ on this space as

$$Z^n_a(\overline{\omega}, \eta) \doteq q_{0a}\left(\widetilde{X}^n(\overline{\omega}), \eta\right), \qquad Z_a(\overline{\omega}, \eta) \doteq q_{0a}\left(\widetilde{X}(\overline{\omega}), \eta\right).$$

Then it follows from the continuity of h that (cf. [5]) $Z^n_a \to Z_a$ in $L^1(R)$. This immediately yields that

$$\begin{aligned} &\limsup_{K\to\infty} \limsup_{n\to\infty} \int_{\widehat{\Omega}} q_{0,a}(\theta, \eta) 1_{q_{0,a}(\theta,\eta)>K} \, dR_{\nu_n}(\theta, \eta) \\ &= \limsup_{K\to\infty} \limsup_{n\to\infty} \int_{\Omega_0} Z^n_a 1_{Z^n_a > K} \, dR \\ &= 0. \end{aligned}$$

This proves the lemma. □

PROPOSITION 3.5. – *Let ϕ_i, $i = 1, \ldots, k$, be in $C_b(E)$. Let $C \in (1, \infty)$ be such that*

$$\prod_{i=1}^k \left| \phi_i(x_i) \right| \leqslant C, \quad \forall x_i \in E;\ i = 1, 2, \ldots, k. \tag{3.7}$$

Let

$$-\infty < t_1 < t_2 < \cdots < t_k = 0.$$

Also let $t, \tau_0 \in (-\infty, 0)$ *be such that* $t < \tau_0 < t_1$. *Then*

$$\mathbb{E}_{R_\mu^{(1)}}\left|\mathbb{E}_{R_\mu^{(1)}}\left[\prod_{i=1}^{k}\phi_i(\tilde{\xi}_{t_i}) \mid \mathcal{Z}_{-\infty}^{0} \vee \mathcal{G}_{-\infty}^{t}\right] - \mathbb{E}_{R_\mu^{(1)}}\left[\prod_{i=1}^{k}\phi_i(\tilde{\xi}_{t_i}) \mid \mathcal{Z}_{\tau_0}^{0} \vee \mathcal{G}_{-\infty}^{t}\right]\right| \tag{3.8}$$

is bounded above by $C\int_E (U_1(x) + U_2(x))\,\mu(dx)$, *where* $U_i(x)$ *for* $i = 1, 2$ *is defined as*

$$\int_{\mathcal{C}} \mathbb{E}_Q\big(\Gamma_{t^*}(\delta_x)(E)\big|\Lambda_{t^*}(\delta_x)\big(\Psi_i(\cdot, \eta')\big) - \Lambda_{t^*-\tau^*, t^*}(\delta_x T_{t^*-\tau^*})\big(\Psi_i(\cdot, \eta')\big)\big|\big)\, Q(d\eta'), \tag{3.9}$$

$t^* \doteq t_1 - t$, $\tau^* \doteq t_1 - \tau_0$ *and* $\Psi_i : E \times \mathcal{C}$ *is defined as follows*:

$$\Psi_1(x, \eta) = \phi_1(x)\int_{\mathcal{D}}\prod_{i=2}^{k}\phi_i(\xi_{t_i - t_{i-1}})q_{0,-t_1}(\theta, \eta)\, P_x(d\theta), \tag{3.10}$$

$$\Psi_2(x, \eta) \doteq \Gamma_{0,-t_1}(\delta_x)(E)(\eta) = \int_{\mathcal{D}} q_{0,-t_1}(\theta, \eta)\, P_x(d\theta). \tag{3.11}$$

Proof. – We begin by observing that from the Markov property of the signal and the independence between observation noise and the signal, one can replace the two conditioning σ fields in (3.8) with $\mathcal{Z}_t^0 \vee \sigma\{\xi_t\}$ and $\mathcal{Z}_{\tau_0}^0 \vee \sigma\{\xi_t\}$, respectively. Thus, the expression in (3.8) can be rewritten as an expectation on $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{R}_{\delta_x})$ as,

$$\int_E \mathbb{E}_{\widehat{R}_{\delta_x}}\left|\mathbb{E}_{\widehat{R}_{\delta_x}}\left[\prod_{i=1}^{k}\phi_i(\xi_{t_i - t}) \mid \mathcal{A}_0^{-t}\right] - \mathbb{E}_{\widehat{R}_{\delta_x}}\left[\prod_{i=1}^{k}\phi_i(\xi_{t_i - t}) \mid \mathcal{A}_{\tau_0 - t}^{-t}\right]\right|\mu(dx). \tag{3.12}$$

Now an application of Bayes formula and a further conditioning yields that

$$\begin{aligned}
\mathbb{E}_{\widehat{R}_{\delta_x}}\left[\prod_{i=1}^{k}\phi_i(\xi_{t_i - t}) \mid \mathcal{A}_0^{-t}\right] &= \frac{\int_{\mathcal{D}}\prod_{i=1}^{k}\phi_i(\xi_{t_i - t})q_{0,-t}(\theta, \eta)P_x(d\theta)}{\int_{\mathcal{D}} q_{0,-t}(\theta, \eta)P_x(d\theta)} \\
&= \frac{\int_{\mathcal{D}}\Psi_1(\xi_{t_1 - t}, \gamma_{t_1 - t}(\eta))q_{0,t_1 - t}(\theta, \eta)P_x(d\theta)}{\int_{\mathcal{D}}\Psi_2(\xi_{t_1 - t}, \gamma_{t_1 - t}(\eta))q_{0,t_1 - t}(\theta, \eta)P_x(d\theta)} \\
&= \frac{\Gamma_{t_1 - t}(\delta_x)(\Psi_1(\cdot, \gamma_{t_1 - t}(\eta)))}{\Gamma_{t_1 - t}(\delta_x)(\Psi_2(\cdot, \gamma_{t_1 - t}(\eta)))} \\
&= \frac{\Lambda_{t_1 - t}(\delta_x)(\Psi_1(\cdot, \gamma_{t_1 - t}(\eta)))}{\Lambda_{t_1 - t}(\delta_x)(\Psi_2(\cdot, \gamma_{t_1 - t}(\eta)))}
\end{aligned} \tag{3.13}$$

where for $s > 0$, $\gamma_s : \mathcal{C} \to \mathcal{C}$ is defined as $\gamma_s(\eta)(u) \doteq \eta(u + s) - \eta(s)$ and $\Psi_i,\ i = 1, 2$, are as defined in the statement of the proposition. In exactly the same manner it is shown that

$$\mathbb{E}_{\widehat{R}_{\delta_x}}\left[\prod_{i=1}^{k}\phi_i(\xi_{t_i-t}) \mid \mathcal{A}_{\tau_0-t}^{-t}\right] = \frac{\int_{\mathcal{D}} \Psi_1(\xi_{t_1-t}, \gamma_{t_1-t}(\eta)) q_{\tau_0-t,t_1-t}(\theta,\eta) P_x(d\theta)}{\int_{\mathcal{D}} \Psi_2(\xi_{t_1-t}, \gamma_{t_1-t}(\eta)) q_{\tau_0-t,t_1-t}(\theta,\eta) P_x(d\theta)}$$
$$= \frac{\Lambda_{\tau_0-t,t_1-t}(T_{\tau_0-t}\delta_x)(\Psi_1(\cdot\,, \gamma_{t_1-t}(\eta)))}{\Lambda_{\tau_0-t,t_1-t}(T_{\tau_0-t}\delta_x)(\Psi_2(\cdot\,, \gamma_{t_1-t}(\eta)))}. \tag{3.14}$$

The above representations show that the term inside the integral in (3.12) is same as

$$\mathbb{E}_Q\left[\Gamma_{0,-t}(\delta_x)(E)\left|\frac{\Lambda_{t_1-t}(\delta_x)(\Psi_1(\cdot\,, \gamma_{t_1-t}(\eta)))}{\Lambda_{t_1-t}(\delta_x)(\Psi_2(\cdot\,, \gamma_{t_1-t}(\eta)))} - \frac{\Lambda_{\tau_0-t,t_1-t}(T_{\tau_0-t}\delta_x)(\Psi_1(\cdot\,, \gamma_{t_1-t}(\eta)))}{\Lambda_{\tau_0-t,t_1-t}(T_{\tau_0-t}\delta_x)(\Psi_2(\cdot\,, \gamma_{t_1-t}(\eta)))}\right|\right]. \tag{3.15}$$

Next observe that if U, V, U', V' are real numbers such that $|U'/V'| \leqslant D$ for some $D \geqslant 1$ then

$$\left|\frac{U}{V} - \frac{U'}{V'}\right| \leqslant D\left\{\left|\frac{U-U'}{V}\right| + \left|\frac{V-V'}{V}\right|\right\}.$$

Using this inequality, we have that the term in (3.15) is bounded by

$$C\sum_{i=1}^{2}\mathbb{E}_Q\Bigg[\Gamma_{0,-t}(\delta_x)(E) \left|\frac{\Lambda_{t_1-t}(\delta_x)(\Psi_i(\cdot\,, \gamma_{t_1-t}(\eta))) - \Lambda_{\tau_0-t,t_1-t}(T_{\tau_0-t}\delta_x)(\Psi_i(\cdot\,, \gamma_{t_1-t}(\eta)))}{\Lambda_{t_1-t}(\delta_x)(\Psi_2(\cdot\,, \gamma_{t_1-t}(\eta)))}\right|\Bigg]. \tag{3.16}$$

Next note that

$$\Lambda_{t_1-t}(\delta_x)\big(\Psi_2\big(\cdot\,, \gamma_{t_1-t}(\eta)\big)\big) = \frac{\Gamma_{t_1-t}(\delta_x)(\Psi_2(\cdot, \gamma_{t_1-t}(\eta)))}{\Gamma_{t_1-t}(\delta_x)(E)} = \frac{\Gamma_{0,-t}(\delta_x)(E)}{\Gamma_{t_1-t}(\delta_x)(E)}.$$

Using this equality we have that the term in (3.16) equals

$$C\sum_{i=1}^{2}\mathbb{E}_Q\Big[\Gamma_{t_1-t}(\delta_x)(E)\big|\Lambda_{t_1-t}(\delta_x)\big(\Psi_i\big(\cdot\,, \gamma_{t_1-t}(\eta)\big)\big) - \Lambda_{\tau_0-t,t_1-t}(T_{\tau_0-t}\delta_x)\big(\Psi_1\big(\cdot\,, \gamma_{t_1-t}(\eta)\big)\big)\big|\Big].$$

Replacing, in the above display $t_1 - t$ by t^* and $t_1 - \tau_0$ by τ^*, and observing that under Q, γ_{t^*} is independent of $\mathcal{A}_0^{t^*}$, we have that the above display equals $C(U_1(x) + U_2(x))$, where $U_i(x)$, $i = 1, 2$, are as defined in (3.9). Combining this observation with (3.12) we have the result. □

The following theorem shows that under Assumptions 2.3 and 2.6, (vii) implies (viii), where (vii) and (viii) are as in Theorem 2.7.

THEOREM 3.6. – *Suppose that Assumptions* 2.3 *and* 2.6 *hold. Further suppose that the filter has the finite memory property as defined in Definition* 2.4*. Then the* σ*-fields* $\{\mathcal{G}_{-\infty}^t\}_{t<0}$ *have finite time dependence* (*as defined in Definition* 2.5) *with respect to the* σ*-field* $\mathcal{Z}_{-\infty}^0$.

Proof. – Let $\phi_i, t_i,\ i = 1, \dots, k$, be as in Proposition 3.5. From Definition 2.5 we have that, it suffices to show that for μ-a.e. x and $i = 1, 2$,

$$\limsup_{\tau^*\to\infty} \limsup_{t^*\to\infty} U_i(x) = 0, \tag{3.17}$$

where $U_i(x)$ are defined in (3.9). For $K \in (0, \infty)$, we write

$$\begin{aligned} q_{0,-t}(\theta,\eta) &= q_{0,-t}(\theta,\eta) \wedge K + \left[q_{0,-t}(\theta,\eta) - K\right] I_{q_{0,-t}(\theta,\eta)\geqslant K} \\ &\equiv q^{(1)}(\theta,\eta) + q^{(2)}(\theta,\eta). \end{aligned}$$

Now for $i = 1, 2$, define $\Psi_i^{(1)}$ and $\Psi_i^{(2)}$ by replacing $q_{0,-t}$ in the definition of Ψ_i (see (3.10), (3.11)) by $q^{(1)}$ and $q^{(2)}$ respectively. Clearly, for $i = 1, 2$, $\Psi_i = \Psi_i^{(1)} + \Psi_i^{(2)}$ Observe that for $i = 1, 2$ and Q-a.e. η', $\Psi_i^{(1)}(\cdot, \eta') \in C_b(E)$. This implies that for μ almost every x,

$$\begin{aligned} &\limsup_{\tau^*\to\infty} \limsup_{t^*\to\infty} \int_{\mathcal{C}} \big(\mathbb{E}_Q \Gamma_{t^*}(\delta_x)(E) \big| \Lambda_{t^*}(\delta_x)\big(\Psi_i^{(1)}(\cdot\,,\eta')\big) \\ &\qquad - \Lambda_{t^*-\tau^*,t^*}(\delta_x T_{t^*-\tau^*})\big(\Psi_i^{(1)}(\cdot\,,\eta')\big)\big|\big) Q(d\eta') \\ &\quad\leqslant \int_{\mathcal{C}} \limsup_{\tau^*\to\infty} \limsup_{t^*\to\infty} \big(\mathbb{E}_Q \Gamma_{t^*}(\delta_x)(E) \big| \Lambda_{t^*}(\delta_x)\big(\Psi_i^{(1)}(\cdot\,,\eta')\big) \\ &\qquad - \Lambda_{t^*-\tau^*,t^*}(\delta_x T_{t^*-\tau^*})\big(\Psi_i^{(1)}(\cdot\,,\eta')\big)\big|\big) Q(d\eta') \\ &\quad= \int_{\mathcal{C}} \limsup_{\tau^*\to\infty} \limsup_{t^*\to\infty} \big(\mathbb{E}_{Q_{\delta_x}} \big| \Lambda_{t^*}(\delta_x)\big(\Psi_i^{(1)}(\cdot\,,\eta')\big) \\ &\qquad - \Lambda_{t^*-\tau^*,t^*}(\delta_x T_{t^*-\tau^*})\big(\Psi_i^{(1)}(\cdot\,,\eta')\big)\big|\big) Q(d\eta') \\ &\quad= 0, \end{aligned} \tag{3.18}$$

where the first inequality above follows on observing that the integrand in the first line of the display is uniformly bounded in (t^*, τ^*). and the last equality is a consequence of the finite memory property of the filter.

Next note that for $i = 1, 2$,

$$\begin{aligned} &\int_{\mathcal{C}} \big(\mathbb{E}_Q \Gamma_{t^*}(\delta_x)(E) \big| \Lambda_{t^*}(\delta_x)\big(\Psi_i^{(2)}(\cdot\,,\eta')\big)\big|\big) Q(d\eta') \\ &\quad= \int_{\mathcal{C}} \big(\mathbb{E}_Q \big| \Gamma_{t^*}(\delta_x)\big(\Psi_i^{(2)}(\cdot\,,\eta')\big)\big|\big) Q(d\eta') \\ &\quad\leqslant C \int_{\widehat{\Omega}} q_{0,t^*}(\theta,\eta) q_{t^*,t^*-t_1}(\theta,\eta) 1_{q_{t^*,t^*-t_1}(\theta,\eta)>K}\, P_x(d\theta)\, Q(d\eta) \\ &\quad= C \int_{\widehat{\Omega}} q_{t^*,t^*-t_1}(\theta,\eta) 1_{q_{t^*,t^*-t_1}(\theta,\eta)>K}\, P_x(d\theta)\, Q(d\eta) \end{aligned}$$

$$= C \int_{\widehat{\Omega}} q_{0,-t_1}(\theta,\eta) 1_{q_{0,-t_1}(\theta,\eta)>K}\, P_{\delta_x T_{t^*}}(d\theta)\, Q(d\eta). \tag{3.19}$$

Corollary 3.2 gives that $\delta_x T_{t^*} \to \mu$ as $t^* \to \infty$. Using this observation in the above display, along with Lemma 3.4 we have that for $i = 1, 2$ and all $x \in E$,

$$\limsup_{K\to\infty} \limsup_{t^*\to\infty} \int_{\mathcal{C}} \big(\mathbb{E}_Q \Gamma_{t^*}(\delta_x)(E) \big| \Lambda_{t^*}(\delta_x)\big(\Psi_i^{(2)}(\cdot\,,\eta')\big)\big|\big)\, Q(d\eta') = 0. \tag{3.20}$$

Next consider

$$\begin{aligned}
&\int_{\mathcal{C}} \big(\mathbb{E}_Q \Gamma_{t^*}(\delta_x)(E) \big| \Lambda_{t^*-\tau^*,t^*}(\delta_x T_{t^*-\tau^*})\big(\Psi_i^{(2)}(\cdot\,,\eta')\big)\big|\big)\, Q(d\eta') \\
&\quad = \int_{\mathcal{C}} \big(\mathbb{E}_Q \Gamma_{t^*-\tau^*,t^*}(\delta_x T_{t^*-\tau^*})(E) \big| \Lambda_{t^*-\tau^*,t^*}(\delta_x T_{t^*-\tau^*})\big(\Psi_i^{(2)}(\cdot\,,\eta')\big)\big|\big)\, Q(d\eta') \\
&\quad = \int_{\mathcal{C}} \big(\mathbb{E}_Q \big| \Gamma_{t^*-\tau^*,t^*}(\delta_x T_{t^*-\tau^*})\big(\Psi_i^{(2)}(\cdot\,,\eta')\big)\big|\big)\, Q(d\eta') \\
&\quad \leqslant C \int_{\widehat{\Omega}} q_{t^*-\tau^*,t^*}(\theta,\eta) q_{t^*,t^*-t_1}(\theta,\eta) 1_{q_{t^*,t^*-t_1}(\theta,\eta)>K}\, P_x(d\theta)\, Q(d\eta) \\
&\quad \leqslant C \int_{\widehat{\Omega}} q_{0,-t_1}(\theta,\eta) 1_{q_{0,-t_1}(\theta,\eta)>K}\, P_{\delta_x T_{t^*}}(d\theta)\, Q(d\eta),
\end{aligned}$$

where the last step follows as in (3.19). Once more, in view of Corollary 3.2 and Lemma 3.4 we have that for $i = 1, 2$

$$\limsup_{K\to\infty} \limsup_{\tau^*\to\infty} \limsup_{t^*\to\infty} \int_{\mathcal{C}} \big(\mathbb{E}_Q \Gamma_{t^*}(\delta_x)(E) \big| \Lambda_{t^*-\tau^*,t^*}(\delta_x T_{t^*-\tau^*})\big(\Psi_i^{(2)}(\cdot\,,\eta')\big)\big|\big)\, Q(d\eta') = 0. \tag{3.21}$$

Finally, combining (3.18), (3.20) and (3.21), we have that for μ almost every x, (3.17) holds. This proves the result. □

We now proceed to the proof of the statement that (viii) implies (i). We will begin with the following lemma.

LEMMA 3.7. – *Let $a \in (0, \infty)$ be fixed. Then*

$$\bigcap_{t\leqslant 0} \big(\mathcal{Z}_{-a}^0 \vee \mathcal{G}_{-\infty}^t\big) = \mathcal{Z}_{-a}^0 \vee \mathcal{G}_{-\infty}^{-\infty}.$$

Proof. – Define $R^* \in \mathcal{P}(\Omega^1)$ by the relation

$$\frac{dR^*}{dR_\mu^{(1)}} \doteq \exp\left\{ -\int_{-a}^{0} h(\xi_u)\, d\alpha(u) + \frac{1}{2}\int_{-a}^{0} \big\|h(\xi_u)\big\|^2\, du \right\}.$$

Observe that,

$$\text{under } R^*,\ \mathcal{G}^0_{-\infty} \text{ is independent of } \mathcal{Z}^0_{-a}. \tag{3.22}$$

Though this is a standard fact, we sketch the proof of the statement in Appendix A. This immediately yields that

$$\bigcap_{t\leqslant 0}\left(\mathcal{Z}^0_{-a} \vee \mathcal{G}^t_{-\infty}\right) = \mathcal{Z}^0_{-a} \vee \mathcal{G}^{-\infty}_{-\infty} \ (\text{mod } R^*).$$

The lemma now follows on noting that $R^{(1)}_{\mu}$ and R^* are mutually absolutely continuous. □

THEOREM 3.8. – *Suppose that the σ-fields $\{\mathcal{G}^t_{-\infty}\}_{t<0}$ have finite time dependence with respect to the σ-field $\mathcal{Z}^0_{-\infty}$* (*Definition* 2.5). *Then*

$$\bigcap_{s=-\infty}^{\infty}\left(\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^s_{-\infty}\right) = \mathcal{Z}^0_{-\infty} \vee \mathcal{G}^{-\infty}_{-\infty}.$$

Proof. – Note that for every $s \in (-\infty, \infty)$

$$\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^s_{-\infty} \supseteq \mathcal{Z}^0_{-\infty} \vee \mathcal{G}^{-\infty}_{-\infty}.$$

Therefore, we have that

$$\bigcap_{s=-\infty}^{\infty}\left(\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^s_{-\infty}\right) \supseteq \mathcal{Z}^0_{-\infty} \vee \mathcal{G}^{-\infty}_{-\infty}.$$

It thus suffices to show that

$$\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^{-\infty}_{-\infty} \supseteq \bigcap_{s=-\infty}^{\infty}\left(\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^s_{-\infty}\right). \tag{3.23}$$

Next note that (3.23) will follow if we show that for all $F \in \bigcap_{s=-\infty}^{\infty}(\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^s_{-\infty})$

$$P\left(F \mid \mathcal{Z}^0_{-\infty} \vee \mathcal{G}^{-\infty}_{-\infty}\right) = 1_F, \text{ a.e.} \tag{3.24}$$

Since $\bigcap_{s=-\infty}^{\infty}(\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^s_{-\infty}) = \bigcap_{t\leqslant 0}(\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^t_{-\infty})$, the right side of (3.24) is $P(F \mid \bigcap_{t\leqslant 0}(\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^s_{-\infty}))$, a.e. and so we have that (3.23) will follow if we show that

$$P\left(F \mid \mathcal{Z}^0_{-\infty} \vee \mathcal{G}^{-\infty}_{-\infty}\right) = P\left(F \mid \bigcap_{t\leqslant 0}\left(\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^t_{-\infty}\right)\right), \text{ a.e.} \tag{3.25}$$

for all

$$F \in \mathcal{Z}^0_{-\infty} \vee \mathcal{G}^0_{-\infty} \supseteq \bigcap_{s=-\infty}^{\infty}\left(\mathcal{Z}^0_{-\infty} \vee \mathcal{G}^s_{-\infty}\right).$$

By the usual monotone class arguments it suffices to consider F of the form $F_1 \cap F_2$, where $F_1 \in \mathcal{Z}_{-\infty}^0$ and $F_2 \in \mathcal{G}_{-\infty}^0$. Thus (3.25) (and hence the result) will follow if we show that for all bounded random variables U_1 and U_2 such that U_1 is $\mathcal{G}_{-\infty}^0$ measurable and U_2 is $\mathcal{Z}_{-\infty}^0$ measurable

$$\mathbb{E}_{R_\mu^{(1)}}\left(U_1U_2 \mid \mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^{-\infty}\right) = \mathbb{E}_{R_\mu^{(1)}}\left(U_1U_2 \mid \bigcap_{t\leqslant 0}\left(\mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^t\right)\right). \tag{3.26}$$

Since U_2 is measurable with respect to both $\mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^{-\infty}$ and $\bigcap_{t\leqslant 0}(\mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^t)$, it follows that we can take (without loss of generality) $U_2 = 1$. Furthermore, via a monotone class argument it follows that, it is enough to prove that, given any $\phi_1, \ldots, \phi_k \in \mathcal{C}_b(E)$ and $t_1 < t_2 < \cdots < t_k = 0$,

$$\mathbb{E}_{R_\mu^{(1)}}\left(U \mid \mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^{-\infty}\right) = \mathbb{E}_{R_\mu^{(1)}}\left(U \mid \bigcap_{t\leqslant 0}\left(\mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^t\right)\right), \tag{3.27}$$

where

$$U \doteq \phi_1(\xi_{t_1}) \cdots \phi_k(\xi_{t_k}).$$

Let $\varepsilon > 0$ be arbitrary and let t_ε and τ_ε be as in Definition 2.5. Then for all $t \leqslant t_\varepsilon \leqslant \tau_\varepsilon < t_1$

$$\mathbb{E}_{R_\mu^{(1)}}\left|\mathbb{E}_{R_\mu^{(1)}}\left[U \mid \mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^t\right] - \mathbb{E}_{R_\mu^{(1)}}\left[U \mid \mathcal{Z}_{\tau_\varepsilon}^0 \vee \mathcal{G}_{-\infty}^t\right]\right| \leqslant \varepsilon.$$

Let C be as in (3.7), then $|U| \leqslant C$. This implies, on taking limit $t \to -\infty$ in the above display that

$$\mathbb{E}_{R_\mu^{(1)}}\left|\mathbb{E}_{R_\mu^{(1)}}\left[U \mid \bigcap_{t\leqslant 0}\left(\mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^t\right)\right] - \mathbb{E}_{R_\mu^{(1)}}\left[U \mid \bigcap_{t\leqslant 0}\left(\mathcal{Z}_{\tau_\varepsilon}^0 \vee \mathcal{G}_{-\infty}^t\right)\right]\right| \leqslant \varepsilon.$$

Combining the above observation with Lemma 3.7 we now have that

$$\mathbb{E}_{R_\mu^{(1)}}\left|\mathbb{E}_{R_\mu^{(1)}}\left[U \mid \bigcap_{t\leqslant 0}\left(\mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^t\right)\right] - \mathbb{E}_{R_\mu^{(1)}}\left[U \mid \mathcal{Z}_{\tau_\varepsilon}^0 \vee \mathcal{G}_{-\infty}^{-\infty}\right]\right| \leqslant \varepsilon.$$

Thus to every $\varepsilon > 0$, there exists a $\mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^{-\infty}$ measurable random variable V_ε; $|V_\varepsilon| \leqslant C$, such that

$$\mathbb{E}_{R_\mu^{(1)}}\left|\mathbb{E}_{R_\mu^{(1)}}\left[U \mid \bigcap_{t\leqslant 0}\left(\mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^t\right)\right] - V_\varepsilon\right| \leqslant \varepsilon.$$

This implies that $\mathbb{E}_{R_\mu^{(1)}}[U \mid \bigcap_{t\leqslant 0}(\mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^t)]$ is $\mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^{-\infty}$ measurable. Since $\bigcap_{t\leqslant 0}(\mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^t) \supseteq \mathcal{Z}_{-\infty}^0 \vee \mathcal{G}_{-\infty}^{-\infty}$, we have (3.27). □

Appendix A

Proof of (3.22). – We denote a typical element of Ω^1 by (ω_1, ω_2). For a fixed $\omega_1 \in \mathcal{D}_{\mathbb{R}}$, define $R^*_{\omega_1} \in \mathcal{P}(\mathcal{C}_{\mathbb{R}})$, via the relation

$$\frac{dR^*_{\omega_1}}{dQ^{(1)}} \doteq \exp\left\{-\int_{-a}^{0} h\big(\xi_u(\omega_1)\big)\, d\tilde{\beta}(u) - \frac{1}{2}\int_{-a}^{0} \big\|h\big(\xi_u(\omega_1)\big)\big\|^2 du\right\}.$$

Note that $R^*_{\omega_1}$ is indeed a probability measure on $\mathcal{C}_{\mathbb{R}}$ since

$$\int_{\mathcal{C}_{\mathbb{R}}} \exp\left\{-\int_{-a}^{0} h\big(\xi_u(\omega_1)\big)\, d\tilde{\beta}(u) - \frac{1}{2}\int_{-a}^{0} \big\|h\big(\xi_u(\omega_1)\big)\big\|^2 du\right\} dQ^{(1)} = 1.$$

Furthermore, by Girsanov's theorem $\{\alpha_u(\omega_1, \cdot) - \alpha_{-a}(\omega_1, \cdot)\}_{-a \leqslant u \leqslant 0}$ is a Wiener process under $R^*_{\omega_1}$. Now define processes $\{b_t\}_{-a \leqslant t \leqslant 0}$ and $\{\tilde{b}_t\}_{-a \leqslant t \leqslant 0}$ on $(\Omega^1, R^{(1)}_\mu)$ as follows:

$$b_u \doteq \tilde{\beta}_u - \tilde{\beta}_{-a}; \quad \tilde{b}_u \doteq \alpha_u - \alpha_{-a}; \quad -a \leqslant u \leqslant 0.$$

Then we have that $\forall\, \Psi \in C_b(C([-a, 0], \mathbb{R}^d))$

$$\begin{aligned}
&\mathbb{E}_{R^{(1)}_\mu}\left[\Psi(\tilde{b}) \exp\left\{-\int_{-a}^{0} h(\xi_u)\, d\tilde{\beta}(u) - \frac{1}{2}\int_{-a}^{0} \|h(\xi_u)\|^2 du\right\} \Big|\, \mathcal{G}^{\infty}_{-\infty}\right] \\
&\quad = \mathbb{E}_{R^{(1)}_\mu}(\Psi(b)),\ R^{(1)}_\mu \text{ a.s.}
\end{aligned}$$

Now let f be a $\mathcal{G}^{\infty}_{-\infty}$ measurable bounded random variable on $(\Omega^1, R^{(1)}_\mu)$. Then

$$\begin{aligned}
\mathbb{E}_{R^*}\big(f\Psi(\tilde{b})\big) &= \mathbb{E}_{R^{(1)}_\mu}\left[f\Psi(\tilde{b}) \exp\left\{-\int_{-a}^{0} h(\xi_u)\, d\tilde{\beta}(u) - \frac{1}{2}\int_{-a}^{0} \|h(\xi_u)\|^2 du\right\}\right] \\
&= \mathbb{E}_{R^{(1)}_\mu}\Bigg[f\mathbb{E}_{R^{(1)}_\mu}\Bigg(\Psi(\tilde{b}) \exp\bigg\{-\int_{-a}^{0} h(\xi_u)\, d\tilde{\beta}(u) \\
&\qquad - \frac{1}{2}\int_{-a}^{0} \|h(\xi_u)\|^2 du\bigg\} \mid \mathcal{G}^{\infty}_{-\infty}\Bigg)\Bigg] \\
&= \mathbb{E}_{R^{(1)}_\mu}\big(f\mathbb{E}_{R^{(1)}_\mu}\big(\Psi(b)\big)\big) \\
&= \mathbb{E}_{R^{(1)}_\mu}(f)\mathbb{E}_{R^*}\big(\Psi(\tilde{b})\big).
\end{aligned}$$

By taking $\Psi(\cdot) \equiv 1$ in the above display we see that $\mathbb{E}_{R^{(1)}_\mu}(f)$ on the right side of the above expression equals $\mathbb{E}_{R^*}(f)$. This proves (3.22). □

Acknowledgements

I am grateful to Professor R. Liptser and P. Chigansky for, pointing out the gap in the paper [22], showing me the examples where the equality of the two σ-fields fails and several other suggestions on this manuscript. I will also like to thank Professors Rami Atar and Ofer Zeitouni for their valuable comments.

REFERENCES

[1] R. Atar, Exponential stability for nonlinear filtering of diffusion processes in non-compact domain, Ann. Probab. 26 (1998) 1552–1574.
[2] R. Atar, O. Zeitouni, Exponential stability for nonlinear filtering, Annales de l'Institut H. Poincaré Probabilites et Statistique 33 (1997) 697–725.
[3] R. Atar, O. Zeitouni, Lyapunov exponents for finite state nonlinear filtering, SIAM J. Control Optim. 35 (1997) 36–55.
[4] A. Bhatt, A. Budhiraja, R. Karandikar, Markov property and ergodicity of the nonlinear filter, SIAM J. Control Optim. 39 (2000) 928–949.
[5] A. Bhatt, G. Kallianpur, R. Karandikar, Robustness of the nonlinear filter, Stochastic Process. Appl. 81 (1999) 247–254.
[6] A. Budhiraja, Ergodic properties of the nonlinear filter, Stochastic Process. Appl. 95 (2001) 1–24.
[7] A. Budhiraja, H.J. Kushner, Approximation and limit results for nonlinear filters over an infinite time interval, SIAM J. Control Optim. 37 (1997) 1946–1979.
[8] A. Budhiraja, H.J. Kushner, Robustness of nonlinear filters over the infinite time interval, SIAM J. Control Optim. 36 (1998) 1618–1637.
[9] A. Budhiraja, H.J. Kushner, Approximation and limit results for nonlinear filters over an infinite time interval: Part II, random sampling algorithms, SIAM J. Control Optim. 38 (2000) 1874–1908.
[10] A. Budhiraja, H.J. Kushner, Monte Carlo algorithms and asymptotic problems in nonlinear filtering, in: Stochastics in Finite/Infinite Dimensions, in: Trends Math., Birkhäuser, Boston, 2001, pp. 59–87.
[11] A. Budhiraja, D. Ocone, Exponential stability of discrete time filters without signal ergodicity, System Control Lett. 30 (1997) 185–193.
[12] A. Budhiraja, D. Ocone, Exponential stability in discrete time filtering for non-ergodic signals, Stochastic Process. Appl. 82 (1999) 245–257.
[13] P. Chigansky, R. Liptser, Private communication.
[14] P. Chigansky, P. Baxendale, R. Liptser, Asymptotic stability of the Wonham filter. Ergodic and nonergodic signals, Preprint in Arxiv.
[15] F. Cérou, Long time asymptotics for some dynamical noise free non linear filtering problems, Rapport de Recherche 2446, INRIA, December, 1994.
[16] J.M.C. Clark, D.L. Ocone, C. Coumarbatch, Relative entropy and error bounds for filtering of Markov processes, Math. Control Signals Syst. 12 (1999) 346–360.
[17] G. Da Prato, M. Fuhrman, P. Malliavin, Asymptotic ergodicity for the Zakai filtering equation, C. R. Acad. Sci. Paris Serie I 321 (1995) 613–616.
[18] P. Del Moral, A. Guionnet, On the stability of measure valued processes with applications to filtering, C. R. Acad. Sci. Paris Serie I 329 (1999) 429–434.
[19] B. Delyon, O. Zeitouni, Lyapunov exponents for filtering problems, in: Applied Stochastic Analysis (London, 1989), in: Stochastics Monogr., Vol. 5, Gordon and Breach, New York, 1991, pp. 511–521.

[20] G. Kallianpur, Stochastic Filtering Theory, Springer-Verlag, New York, 1980.
[21] R.L. Karandikar, On pathwise stochastic integration, Stochastic Process. Appl. 57 (1995) 11–18.
[22] H. Kunita, Asymptotic behavior of the nonlinear filtering errors of Markov processes, J. Multivariate Anal. 1 (1971) 365–393.
[23] H. Kunita, Ergodic properties of nonlinear filtering processes, in: K.C. Alexander, J.C. Watkins (Eds.), Spatial Stochastic Processes, 1991.
[24] F. Le Gland, L. Mevel, Exponential forgetting and geometric ergodicity in hidden Markov models, Math. Control Signals Syst. 13 (2000) 63–93.
[25] F. Le Gland, N. Oudjane, A robustification approach to stability and to uniform particle approximation of nonlinear filters: the example of pseudo-mixing signals, Preprint.
[26] D. Ocone, Asymptotic stability of Benes filters, Stochastic Anal. Appl. 17 (1999) 1053–1074.
[27] D. Ocone, Entropy inequalities and entropy dynamics in nonlinear filtering of diffusion processes, in: W. McEneaney, G. Yin, Q. Zhang (Eds.), Stochastic Analysis, Control, Optimization and Applications, 1999.
[28] D. Ocone, E. Pardoux, Asymptotic stability of the optimal filter with respect to its initial condition, SIAM J. Control Optim. 34 (1996) 226–243.
[29] L. Stettner, On invariant measures of filtering processes, in: K. Helmes, N. Christopeit, M. Kohlmann (Eds.), Stochastic Differential Systems, Proc. 4th Bad Honnef Conf., 1988, in: Lecture Notes in Control and Inform Sci., 1989, pp. 279–292.
[30] L. Stettner, Invariant measures of pair: State, approximate filtering process, Colloq. Math. LXII (1991) 347–352.
[31] H. Totoki, A class of special flows, Z. Wahr. Verw. Geb. 15 (1970) 157–167.
[32] H.V. Weizsäcker, Exchanging the order of taking suprema and countable intersections of σ algebras, Ann. Inst. Henri Poincaré B 19 (1) (1983) 91–100.
[33] D. Williams, Probability with Martingales, Cambridge University Press, 1991.