

Calvia, Alessandro; Ferrari, Giorgio

**Working Paper**

## Nonlinear filtering of partially observed systems arising in singular stochastic optimal control

Center for Mathematical Economics Working Papers, No. 651

**Provided in Cooperation with:**

Center for Mathematical Economics (IMW), Bielefeld University

*Suggested Citation:* Calvia, Alessandro; Ferrari, Giorgio (2021) : Nonlinear filtering of partially observed systems arising in singular stochastic optimal control, Center for Mathematical Economics Working Papers, No. 651, Bielefeld University, Center for Mathematical Economics (IMW), Bielefeld, <https://nbn-resolving.de/urn:nbn:de:0070-pub-29554928>

This Version is available at:

<https://hdl.handle.net/10419/238129>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

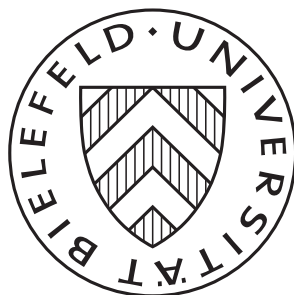


<https://creativecommons.org/licenses/by/4.0/>

# Nonlinear Filtering of Partially Observed Systems Arising in Singular Stochastic Optimal Control

---

Alessandro Calvia and Giorgio Ferrari



# NONLINEAR FILTERING OF PARTIALLY OBSERVED SYSTEMS ARISING IN SINGULAR STOCHASTIC OPTIMAL CONTROL

ALESSANDRO CALVIA AND GIORGIO FERRARI

**ABSTRACT.** This paper deals with a nonlinear filtering problem in which a multi-dimensional signal process is additively affected by a process  $\nu$  whose components have paths of bounded variation. The presence of the process  $\nu$  prevents from directly applying classical results and novel estimates need to be derived. By making use of the so-called reference probability measure approach, we derive the Zakai equation satisfied by the unnormalized filtering process, and then we deduce the corresponding Kushner-Stratonovich equation. Under the condition that the jump times of the process  $\nu$  do not accumulate over the considered time horizon, we show that the unnormalized filtering process is the unique solution to the Zakai equation, in the class of measure-valued processes having a square-integrable density. Our analysis paves the way to the study of stochastic control problems where a decision maker can exert singular controls in order to adjust the dynamics of an unobservable Itô-process.

**Keywords:** Stochastic filtering; singularly controlled systems; reference probability measure; Zakai equation; Kushner-Stratonovich equation.

**AMS 2020:** 93E11, 60G35, 60H15, 60J25, 60J76.

## 1. INTRODUCTION

This paper studies a stochastic filtering problem on a finite time horizon  $[0, T]$ ,  $T > 0$ , in which the dynamics of a multi-dimensional process  $X = (X_t)_{t \in [0, T]}$ , called *signal* or *unobserved process*, are additively affected by a process having components of bounded variation. The aim is to estimate the hidden state  $X_t$ , at each time  $t \in [0, T]$ , using the information provided by a further stochastic process  $Y = (Y_t)_{t \in [0, T]}$ , called *observed process*; said otherwise, we look for the conditional distribution of  $X_t$  given the available observation up to time  $t$ . This leads to derive an evolution equation for the filtering process, which is a probability measure-valued process satisfying, for any given bounded and measurable function  $\varphi: \mathbb{R}^m \rightarrow \mathbb{R}$ ,

$$\pi_t(\varphi) := \int_{\mathbb{R}^m} \varphi(x) \pi_t(dx) = \mathbb{E}[\varphi(X_t) \mid \mathcal{Y}_t], \quad t \in [0, T],$$

where  $(\mathcal{Y}_t)_{t \in [0, T]}$  is the natural filtration generated by  $Y$  and augmented by  $\mathbb{P}$ -null sets. The process  $\pi$  provides the best estimate (in the usual  $L^2$  sense) of the signal process  $X$ , given the available information obtained through the process  $Y$ .

Stochastic filtering is nowadays a well-established research topic. The literature on the subject is vast and many different applications have been studied: the reader may find a fairly detailed historical account in the book by Bain and Crisan [2]. Classic references are the books by Bensoussan [5], Kallianpur [26], Liptser and Shiryaev [32] (cf. also Brémaud [6, Chapter 4] for stochastic filtering with point process observation); more recent monographs are, e.g., the aforementioned book by Bain and Crisan [2], Crisan and Rozovskiĭ [15], and Xiong [37] (see also Cohen and Elliott [13, Chapter 22]). Recently, different cases where the signal and/or the observation processes can have discontinuous trajectories (as in the present work) have been studied and explicit filtering equations have been derived: see, for instance, Bandini et al. [4], Calvia [8], Ceci and Gerardi [11, 12], Ceci and Colaneri [9, 10], Confortola and Fuhrman [14], Grigelionis and Mikulevicius [23].

The main motivation of our analysis stems from the study of singular stochastic control problems under partial observation. Consider a continuous-time stochastic system whose position or level  $X_t$  at time  $t \in [0, T]$  is subject to random disturbances and can be adjusted instantaneously through (cumulative) actions that, as

---

A. Calvia

LUISS University, Department of Economics and Finance, Viale Romania 32, 00197 Rome (Italy).

E-mail: [acalvia@luiss.it](mailto:acalvia@luiss.it).

G. Ferrari

Bielefeld University, Center for Mathematical Economics (IMW), Universitätsstrasse 25, 33615, Bielefeld (Germany).

E-mail: [giorgio.ferrari@uni-bielefeld.de](mailto:giorgio.ferrari@uni-bielefeld.de).

This research was supported by the 2019 INdAM-GNMPA project *Problemi di controllo ottimo stocastico con osservazione parziale in dimensione infinita*, of which the first author was Principal Investigator. Financial support by the German Research Foundation (DFG) through the Collaborative Research Centre 1283 is also gratefully acknowledged by the second author.

functions of time, do not have to be absolutely continuous with respect to Lebesgue measure. In particular, they may present a Cantor-like component and/or a jump component. The use of such singular control policies is nowadays common in applications in Economics, Finance, Operations Research, as well as in Mathematical Biology. Typical examples are, amongst others, (ir)reversible investment choices (e.g., Riedel and Su [36]), dividends' payout (e.g., Reppen et al. [35]), inventory management problems (e.g., Harrison and Taksar [24]), as well as harvesting issues (e.g., Alvarez and Shepp [1]). Suppose also that the decision maker acting on the system is not able to observe the dynamics of the controlled process  $X$ , but she/he can only follow the evolution of a noisy process  $Y$ , whose drift is a function of the signal process. Mathematically, we assume that the pair  $(X, Y)$  is defined on a filtered complete probability space  $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$  and that its dynamics are given, for any  $t \in [0, T]$ , by the following system of SDEs:

$$\begin{cases} dX_t = b(t, X_t) dt + \sigma(t, X_t) dW_t + d\nu_t, & X_{0-} \sim \xi \in \mathcal{P}(\mathbb{R}^m), \\ dY_t = h(t, X_t) dt + \gamma(t) dB_t, & Y_0 = y \in \mathbb{R}^n. \end{cases} \quad (1.1)$$

Here:  $\xi$  is a given probability distribution on  $\mathbb{R}^m$ ;  $W$  and  $B$  are two independent  $\mathbb{F}$ -standard Brownian motions; coefficients  $b, \sigma, h, \gamma$  are suitable measurable functions;  $\nu$  is a càdlàg,  $\mathbb{R}^m$ -valued process with (components of) bounded variation, that is adapted to the previously introduced observation filtration  $(\mathcal{Y}_t)_{t \in [0, T]}$ .

Clearly, the decision maker might want to adjust the dynamics of  $X$  in order to optimize a given performance criterion. Since  $X$  is unobservable, this leads to a stochastic optimal control problem under partial observation, which can be tackled by deriving and studying the so-called *separated problem*, an equivalent problem under full information (see, e.g., Bensoussan [5]), where the signal  $X$  is formally replaced by its estimate provided by the filtering process  $\pi$ . However, to effectively solve the original optimization problem by means of the separated one, a first necessary step concerns the detailed study of the associated filtering problem.

To the best of our knowledge, the derivation of explicit filtering equations in the setting described above has not yet received attention in the literature. In this paper we provide a first contribution in this direction. Indeed, the recent literature treating singular stochastic control problems under partial observation assumes that the observed process, rather than the signal one, is additively controlled (cf. Callegaro et al. [7], De Angelis [16], Décamps and Villeneuve [17], and Federico et al. [22]). Clearly, such a modeling feature leads to a filtering analysis that is completely different from ours.

By making use of the so-called reference probability measure approach, we derive the *Zakai stochastic partial differential equation* (SPDE) satisfied by the so-called *unnormalized filtering process*, which is a measure-valued process, associated with the filtering process via a suitable change of probability measure. Then, we deduce the corresponding evolution equation for  $\pi$ , namely, the so-called *Kushner-Stratonovich equation* or *Fujisaki-Kallianpur-Kunita equation*. Furthermore, we show that the unnormalized filtering process is the unique solution to the Zakai equation, in the class of measure-valued processes having a square-integrable density. The latter result is proved under the technical requirement that the jump times of the process  $\nu$  affecting  $X$  in (1.1) do not accumulate over the considered time-horizon. Although such a condition clearly poses a restriction on the generality of the model, we also acknowledge that it is typically satisfied by optimal control processes arising in singular stochastic control problems. It is important to notice that establishing conditions under which the unnormalized filtering process possesses a density paves the way to recast the separated problem as a stochastic control problem in a Hilbert space, as we will briefly explain in the next section.

The rest of the introduction is now devoted to a discussion of our approach and results at a more technical level.

**1.1. Methodology and main results.** In this paper we are going to study the filtering problem described above through the so-called *reference probability approach*, that we briefly summarize here. To start, let us notice that the model introduced in (1.1) is somewhat ill-posed. In fact, the dynamics of the signal process  $X$  depend on the  $(\mathcal{Y}_t)_{t \in [0, T]}$ -adapted process  $\nu$  while, simultaneously, the dynamics of the observed process  $Y$  depend on  $X$ . Otherwise said, it is not clear how to define  $\nu$ , which has to be given *a priori*, and *circularity* arises if one attempts to introduce the partially observed system  $(X, Y)$  as in (1.1).

A possible way out of this impasse is to define  $Y$  as a given Gaussian process independent of  $X$  (see (2.2)). In this way, it makes sense to fix a  $(\mathcal{Y}_t)_{t \in [0, T]}$ -adapted process  $\nu$  and to define the dynamics of the signal process  $X$  as in the first SDE of (1.1) (see also (2.8)). Finally, under suitable assumptions, there exists a probability measure change (cf. (2.12)) that allows us to recover the dynamics of  $Y$  as in the second SDE of (1.1) (see also (2.13)). It is important to notice that the resulting probability depends on the initial law  $\xi$  of  $X_{0-}$  and on  $\nu$ .

To derive the associated Kushner-Stratonovich equation there are two main approaches in the literature: The *Innovations approach* and the aforementioned reference probability approach. Although it might be possible to derive the filtering dynamics in our context by using the former approach, we follow the latter method.

Our first main result is Theorem 3.4, where we deduce the Zakai equation verified by the unnormalized filtering process (see (3.3) for its definition). From this result, as a byproduct, we deduce in Theorem 3.6 the Kushner-Stratonovich equation satisfied by the filtering process. It is worth noticing that, given the presence of the bounded-variation process  $\nu$  in the dynamics of  $X$ , Theorem 3.4 cannot be obtained by invoking classical results, but novel estimates need to be derived (cf. Lemma A.1 and Proposition A.2). In particular, we employ a change of variable formula for Lebesgue-Stieltjes integrals.

It is clear that in applications, for instance to optimal control problems, establishing uniqueness of the solution to the Zakai equation or to the Kushner-Stratonovich equation is essential. In the literature there are several approaches to tackle this problem, most notably the following four: The filtered martingale problem approach, originally proposed by Kurtz and Ocone [31], and later extended to singular martingale problems in [29] (see also [28]); the PDE approach, as in the book by Bensoussan [5] (see also [2, Section 4.1]); the functional analytic approach, introduced by Lucic and Heunis [33] (see also [2, Section 4.2]); the density approach, studied in Kurtz and Xiong [30] (see also [2, Section 7] and [37]).

The first three methods allow to prove uniqueness of the solution to the Zakai equation in a suitable class of measure-valued processes. However, they do not guarantee that the unique measure-valued process solution to the Zakai equation admits a density process, a fact that has an impact on the study of the separated problem. Indeed, without requiring or establishing conditions guaranteeing existence of such a density process, the separated problem must be formulated in an appropriate Banach space of measures and, as a consequence, the *Hamilton-Jacobi-Bellman* (HJB) equation associated to the separated problem must be formulated in such a general setting as well. As a matter of fact, only recently some techniques have been developed to treat this case, predominantly in the theory of mean-field games (an application to optimal control problems with partial observation is given in [3]).

A more common approach in the literature considers, instead, the density process as the state variable for the separated problem. If it is possible to show that such a density process is the unique solution of a suitable SPDE in  $L^2(\mathbb{R}^m)$ , the so-called *Duncan-Mortensen-Zakai* equation, then this  $L^2(\mathbb{R}^m)$ -valued process can be equivalently used as state variable in the separated problem. This is particularly convenient, since for optimal control problems in Hilbert spaces a well-developed theory is available, at least in the regular case (see, e.g., the monograph by Fabbri et al. [21]). Therefore, in view of possible future applications to singular optimal control problems under partial observation, we adopted the density approach to prove that, under suitable assumptions, the unnormalized filtering process is the unique solution to the Zakai equation in the class of measure-valued processes admitting a density with respect to Lebesgue measure.

We show this result, first, in the case where  $\nu$  is a continuous process (cf. Theorem 4.6) and, then, in the case where the jump times of  $\nu$  do not accumulate in the time interval  $[0, T]$  (see Theorem 4.7). As we already observed, although this assumption prevents to achieve full generality, it has a clear interpretation and it is usually satisfied by the examples considered in the literature. From a technical side, it seems that a direct approach using the method proposed by [30] is not feasible to treat the case of accumulating jumps, due to difficulties in estimating crucial quantities in the arguments used, that are related to the jump component of filtering process. A possible workaround might consist in approximating the process  $\nu$  by cutting away jumps of size smaller than some  $\delta > 0$  and then, provided that a suitable tightness property holds, pass to the limit, as  $\delta \rightarrow 0$ , in the relevant equations. However, this is a delicate and lengthy reasoning, which is left for future research.

The rest of this paper is organized as follows. Section 1.2 provides notation used throughout this work. Section 2 introduces the filtering problem. The Zakai and Kushner-Stratonovich equations are then derived in Section 3, while the uniqueness of the solution to the Zakai equation is proved in Section 4. Finally, Appendix A collects the proof of technical results.

**1.2. Notation.** In this section we collect the main notation used in this work. Throughout the paper the set  $\mathbb{N}$  denotes the set of natural integers  $\mathbb{N} = \{1, 2, \dots\}$ ,  $\mathbb{N}_0 = \{0, 1, \dots\}$ , and  $\mathbb{R}$  is the set of real numbers.

For any  $m \times n$  matrix  $A = (a_{ij})$ , the symbol  $A^*$  denotes its transpose and  $\|A\|$  is its Frobenius norm; i.e.,  $\|A\| = (\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2)^{1/2}$ . For any  $x, y \in \mathbb{R}^d$ ,  $\|x\|$  denotes the Euclidean norm of  $x$  and  $x \cdot y = x^*y$  indicates the inner product of  $x$  and  $y$ . For a fixed Hilbert space  $H$ , we denote its inner product by  $\langle \cdot, \cdot \rangle$  and by  $\|\cdot\|_H$  its norm.

The symbol  $\mathbf{1}_C$  denotes the indicator function of a set  $C$ , while  $\mathbf{1}$  is the constant function equal to 1. The symbol  $\int_a^b$  denotes  $\int_{[a,b]}$  for any  $-\infty < a \leq b < +\infty$ .

For any  $d \in \mathbb{N}$  and  $T > 0$ , we denote by  $C_b^{1,2}([0, T] \times \mathbb{R}^d)$  the set of real-valued bounded measurable functions on  $[0, T] \times \mathbb{R}^d$ , that are continuously differentiable once with respect to the first variable and twice with respect to the second, with bounded derivatives. For any such function, the symbol  $\partial_t$  denotes the derivative with respect to the first variable, while  $D_x = (\partial_1, \dots, \partial_d)$  and  $D_x^2 = (\partial_{ij}^2)_{i,j=1}^d$  denote, respectively, the gradient and the Hessian matrix with respect to the second variable. Furthermore, we simply write  $C_b^2(\mathbb{R}^d)$ , when we are considering a real-valued bounded function on  $\mathbb{R}^d$  that is twice continuously differentiable with bounded derivatives.

For any  $d \in \mathbb{N}$  we indicate by  $L^2(\mathbb{R}^d)$  the set of all square-integrable functions with respect to Lebesgue measure and for all  $k \in \mathbb{N}$  we denote by  $W_k^2(\mathbb{R}^d)$  the Sobolev space of all functions  $f \in L^2(\mathbb{R}^d)$  such that the partial derivatives  $\partial^\alpha$  exist in the weak sense and are in  $L^2(\mathbb{R}^d)$ , whenever the multi-index  $\alpha = (\alpha_1, \dots, \alpha_d)$  is such that  $\alpha_1 + \dots + \alpha_d \leq k$ .

For a fixed metric space  $E$ , endowed with the Borel  $\sigma$ -algebra, we denote by  $\mathcal{P}(E)$ ,  $\mathcal{M}_+(E)$ , and  $\mathcal{M}(E)$  the sets of probability, finite positive, and finite signed measures on  $E$ , respectively. If  $\mu \in \mathcal{M}(E)$ , then  $|\mu| \in \mathcal{M}_+(E)$  is the total variation of  $\mu$ .

For any given càdlàg stochastic process  $Z = (Z_t)_{t \geq 0}$  defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , we denote by  $(Z_t^-)_{t \geq 0}$  the left-continuous version of  $Z$  (i.e.,  $Z_t^- = \lim_{s \rightarrow t^-} Z_s$ ,  $\mathbb{P}$ -a.s., for any  $t \geq 0$ ), and by  $\Delta Z_t := Z_t - Z_t^-$  the jump of  $Z$  at time  $t \geq 0$ . If  $Z$  has finite variation over  $[0, t]$ , for all  $t \geq 0$ ,  $|Z|$  (resp.  $Z^+$ ,  $Z^-$ ) is the *variation process* (resp. the positive part process, the negative part process) of  $Z$ , i.e., the process such that, for each  $t \in [0, T]$  and  $\omega \in \Omega$ ,  $|Z|_t(\omega)$  (resp.  $Z_t^+(\omega)$ ,  $Z_t^-(\omega)$ ) is the total variation (resp. the positive part, the negative part) of the function  $s \mapsto Z_s(\omega)$  on  $[0, t]$ . It is useful to remember that  $Z = Z^+ - Z^-$ ,  $|Z| = Z^+ + Z^-$ , and that  $Z^+$ ,  $Z^-$  are non-decreasing processes.

Finally, with the word *measurable* we refer to *Borel-measurable*, unless otherwise specified.

## 2. MODEL FORMULATION

Let  $T > 0$  be a given fixed time horizon and  $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$  be a complete filtered probability space, with  $\mathbb{F}$  satisfying the usual assumptions.

Define on  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  two independent  $\mathbb{F}$ -adapted standard Brownian motions  $W$  and  $\bar{B}$ , taking values in  $\mathbb{R}^d$  and  $\mathbb{R}^n$ , respectively, with  $d, n \in \mathbb{N}$ . Let then  $\gamma: [0, T] \rightarrow \mathbb{R}^{n \times n}$  be a measurable function such that, for each  $t \in [0, T]$ ,  $\gamma(t)$  is symmetric, with  $\gamma_{ij}(t) \in L^2([0, T])$ , for all  $i, j = 1, \dots, n$ , and uniformly positive definite; that is, there exists  $\delta > 0$  such that for all  $t \in [0, T]$  and all  $x \in \mathbb{R}^n$

$$\gamma(t)x \cdot x \geq \delta \|x\|^2. \quad (2.1)$$

These requirements guarantee in particular that the *observed process*  $Y = (Y_t)_{t \in [0, T]}$ , defined as

$$Y_t = y + \int_0^t \gamma(t) d\bar{B}_t, \quad t \in [0, T], \quad y \in \mathbb{R}^n, \quad (2.2)$$

is an  $\mathbb{R}^n$ -valued  $\mathbb{F}$ -adapted martingale, of which we take a continuous version. Clearly, it holds

$$dY_t = \gamma(t) d\bar{B}_t, \quad t \in [0, T], \quad Y_0 = y \in \mathbb{R}^n. \quad (2.3)$$

*Remark 2.1.* It is not restrictive to require that  $\gamma$  is symmetric (and uniformly positive definite). Indeed, suppose that  $\bar{B}$  is an  $\mathbb{R}^k$ -valued  $\mathbb{F}$ -adapted standard Brownian motion and that  $\gamma: [0, T] \rightarrow \mathbb{R}^{n \times k}$  is such that  $\gamma\gamma^*(t) := \gamma(t)\gamma^*(t)$  is uniformly positive definite. Then, we can obtain an equivalent model defining the  $\mathbb{R}^n$ -valued  $\mathbb{F}$ -adapted standard Brownian motion  $\tilde{B} = (\tilde{B}_t)_{t \in [0, T]}$  through:

$$d\tilde{B}_t := (\gamma\gamma^*(t))^{-1/2} \gamma(t) d\bar{B}_t, \quad t \in [0, T].$$

In fact, in this case (2.3) becomes:

$$dY_t = (\gamma\gamma^*(t))^{1/2} d\tilde{B}_t, \quad t \in [0, T], \quad Y_0 = y \in \mathbb{R}^n,$$

and clearly  $(\gamma\gamma^*(t))^{1/2}$  is symmetric (and uniformly positive definite).

We indicate with the symbol  $\mathbb{Y}$  the completed natural filtration generated by  $Y$ , i.e.,  $\mathbb{Y} := (\mathcal{Y}_t)_{t \in [0, T]}$ , with  $\mathcal{Y}_t := \{Y_s: 0 \leq s \leq t\} \vee \mathcal{N}$ , where  $\mathcal{N}$  is the collection of all  $\mathbb{P}$ -null sets.

*Remark 2.2.* Notice that since  $\gamma$  is invertible,  $\mathbb{Y}$  coincides with the completed natural filtration generated by  $\bar{B}$  and is, therefore, right-continuous. These facts will be useful in the sequel.

Next, we consider a probability distribution  $\xi$  on  $\mathbb{R}^m$ ; measurable functions  $b: [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  and  $\sigma: [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}$ , with  $m \in \mathbb{N}$ ; a  $\mathbb{Y}$ -adapted, càdlàg,  $\mathbb{R}^m$ -valued process  $\nu$  whose components have paths of finite variation. We introduce the following requirements, that will be in force throughout the paper.

**Assumption 2.1.**

- (i) There exist constants  $C_b$  and  $L_b$  such that for all  $t \in [0, T]$

$$\|b(t, x) - b(t, x')\| \leq L_b \|x - x'\| \quad \text{and} \quad \|b(t, 0)\| \leq C_b, \quad \forall x, x' \in \mathbb{R}^m. \quad (2.4)$$

- (ii) There exist constants  $C_\sigma$  and  $L_\sigma$  such that for all  $t \in [0, T]$

$$\|\sigma(t, x) - \sigma(t, x')\| \leq L_\sigma \|x - x'\| \quad \text{and} \quad \|\sigma(t, 0)\| \leq C_\sigma, \quad \forall x, x' \in \mathbb{R}^m. \quad (2.5)$$

- (iii) The probability law  $\xi \in \mathcal{P}(\mathbb{R}^m)$  satisfies

$$\int_{\mathbb{R}^m} \|x\|^2 \xi(dx) < +\infty. \quad (2.6)$$

- (iv) The  $\mathbb{R}^m$ -valued process  $\nu$  is  $\mathbb{Y}$ -adapted, càdlàg, with  $\nu_{0-} = 0$ . Its components have paths of finite variation, which in particular satisfy

$$|\nu^i|_T \leq K, \quad \forall i = 1, \dots, m, \quad (2.7)$$

for some constant  $K > 0$ .

Under Assumption 2.1, for any such  $\nu$ , the following SDE for the signal process  $X = (X_t)_{t \in [0, T]}$  admits a unique strong solution:

$$dX_t = b(t, X_t) dt + \sigma(t, X_t) dW_t + d\nu_t, \quad t \in [0, T], \quad X_{0-} \sim \xi \in \mathcal{P}(\mathbb{R}^m). \quad (2.8)$$

It is important to bear in mind, especially in applications to optimal control problems, that the solution to (2.8) and all the quantities that are related to it depend on the probability distribution  $\xi$  and on  $\nu$ . However, for the ease of exposition, we will not stress this dependence in the sequel.

*Remark 2.3.* Conditions (2.4) and (2.5) ensure that SDE (2.8) admits a unique strong solution for any  $\nu$ . If we assume, in addition, that (2.6) and (2.7) hold, then we have that, for some constant  $\kappa$  depending on  $T$ ,  $b$ ,  $\sigma$ , and  $\nu$ ,

$$\mathbb{E}[\sup_{t \in [0, T]} \|X_t\|^2] \leq \kappa(1 + \mathbb{E}[\|X_{0-}\|^2]) < +\infty, \quad (2.9)$$

since  $\mathbb{E}[\|X_{0-}\|^2] = \int_{\mathbb{R}^m} \|x\|^2 \xi(dx)$ . Proofs of these statements are standard and can be found, for instance, in [13, 34].

We finally arrive to the model we intend to analyze *via* a change of measure. Let  $h: [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^n$  be a measurable function satisfying the following condition, that will stand from now on.

**Assumption 2.2.** There exists a constant  $C_h$  such that for all  $t \in [0, T]$

$$\|h(t, x)\| \leq C_h(1 + \|x\|), \quad \forall x \in \mathbb{R}^m. \quad (2.10)$$

For all  $t \in [0, T]$  define then:

$$\eta_t := \exp \left\{ \int_0^t \gamma^{-1}(s) h(s, X_s) d\bar{B}_s - \frac{1}{2} \int_0^t \|\gamma^{-1}(s) h(s, X_s)\|^2 ds \right\}. \quad (2.11)$$

By Proposition A.2,  $\eta$  is a  $(\mathbb{P}, \mathbb{F})$ -martingale, under Assumptions 2.1 and 2.2. Therefore, we can introduce the probability measure  $\tilde{\mathbb{P}}$  on  $(\Omega, \mathcal{F}_T)$  satisfying

$$\left. \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} \right|_{\mathcal{F}_T} = \eta_T. \quad (2.12)$$

By Girsanov's Theorem, the process  $B = (B_t)_{t \in [0, T]}$  given by  $B_t := \bar{B}_t - \int_0^t \gamma^{-1}(s) h(s, X_s) ds$ ,  $t \in [0, T]$ , is a  $(\tilde{\mathbb{P}}, \mathbb{F})$ -Brownian motion, and under  $\tilde{\mathbb{P}}$  the dynamics of the observed process are provided by the SDE:

$$dY_t = h(t, X_t) dt + \gamma(t) dB_t, \quad t \in [0, T], \quad Y_0 = y \in \mathbb{R}^n. \quad (2.13)$$

We see that equations (2.8) and (2.13) are formally equivalent to model (1.1). Observe, however, that the Brownian motion driving (2.13) is not a source of noise given *a priori*, but it is obtained through a probability measure change; moreover, our construction implies that it depends on the initial law  $\xi$  and on process  $\nu$ . This formulation is typical in optimal control problems under partial observation (see, e.g., [5, Chapter 8]) and has the advantage of avoiding the circularity problem discussed in the Introduction.

*Remark 2.4.* If the partially observed system defined by (2.8) and (2.13) describes the state variables of a singular optimal control problem, where  $\nu$  is the control process, then condition (2.7) implies that the singular control is of *finite fuel* type (see El Karoui and Karatzas [19], Karatzas et al. [27] for early contributions).

*Remark 2.5.* It is worth noticing that all the results in this paper remain valid if we allow  $b$  to depend also on  $\omega$ , as long as the map  $(\omega, t) \mapsto b(\omega, t, x)$  is  $\mathbb{Y}$ -adapted and càdlàg, for each  $x \in \mathbb{R}^m$ , and condition (2.4) holds uniformly with respect to  $\omega$  (i.e.,  $L_b$  and  $C_b$  do not depend on  $\omega$ ). To extend our subsequent results to this case, it suffices to apply the so-called *freezing lemma* whenever necessary.

This modeling flexibility is important when it comes to treating controlled dynamics where  $b$  is a deterministic function, depending on an additional parameter representing the action of a regular control  $\alpha = (\alpha_t)_{t \in [0, T]}$ . Clearly, this control must be càdlàg and  $\mathbb{Y}$ -adapted, i.e., based on the available information. The measurability requirement above ensures that the map  $(\omega, t) \mapsto b(t, x, \alpha_t(\omega))$  is  $\mathbb{Y}$ -adapted.

### 3. THE ZAKAI AND KUSHNER-STRATONOVICH EQUATIONS

In this section we will deduce the Zakai equation satisfied by the unnormalized filtering process, defined in (3.3). As a byproduct, we will deduce the Kushner-Stratonovich equation satisfied by the filtering process (see (3.1) for its definition). As anticipated in the Introduction, we will use the reference probability approach to achieve these results. The reference probability will be precisely  $\mathbb{P}$ , under which the observed process is Gaussian and satisfies (2.2). However, the probability measure that matters from a modelization point of view is  $\tilde{\mathbb{P}}$ , which defined in (2.12). Indeed, we will define the filtering process under this measure. It is important to bear in mind that  $\tilde{\mathbb{P}}$  and  $\mathbb{P}$  are equivalent probability measures. Hence, any result holding  $\mathbb{P}$ -a.s., holds also  $\tilde{\mathbb{P}}$ -a.s., and we will write only the first of these two wordings.

The following technical lemma is needed. Its proof is a consequence of the facts highlighted in Remark 2.2 and it is omitted (the reader may refer, for instance, to [2, Prop. 3.15]). In what follows we will denote  $\mathcal{Y} := \mathcal{Y}_T$ .

**Lemma 3.1.** *Let  $Z$  be an  $\mathcal{F}_t$ -measurable,  $\mathbb{P}$ -integrable random variable,  $t \in [0, T]$ . Then*

$$\mathbb{E}[Z \mid \mathcal{Y}_t] = \mathbb{E}[Z \mid \mathcal{Y}].$$

As previously anticipated, the *filtering process*  $\pi = (\pi_t)_{t \in [0, T]}$  is a  $\mathcal{P}(\mathbb{R}^m)$ -valued process providing the conditional law of the signal  $X$  at each time  $t \in [0, T]$ , given the available observation up to time  $t$ . It is defined for any bounded and measurable  $\varphi: [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}$  as:

$$\pi_t(\varphi_t) := \tilde{\mathbb{E}}[\varphi(t, X_t) \mid \mathcal{Y}_t], \quad t \in [0, T], \quad (3.1)$$

where  $\varphi_t(x) := \varphi(t, x)$ , for any  $(t, x) \in [0, T] \times \mathbb{R}^m$ . Since  $\mathbb{R}^m$  is a complete and separable metric space,  $\pi$  is a well-defined,  $\mathcal{P}(\mathbb{R}^m)$ -valued and  $\mathbb{Y}$ -adapted process.<sup>1</sup> Moreover,  $\pi$  admits a càdlàg modification, since  $X$  is càdlàg (see, e.g. [2, Cor. 2.26]). Hence, in the sequel we shall consider  $\pi$  as a  $\mathbb{Y}$ -progressively measurable process.

We recall the useful Kallianpur-Striebel formula, which holds thanks to Proposition A.2 for any bounded and measurable  $\varphi: [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}$  and for any fixed  $t \in [0, T]$  (for a proof see, e.g., [2, Prop. 3.16])

$$\pi_t(\varphi_t) = \frac{\mathbb{E}[\eta_t \varphi(t, X_t) \mid \mathcal{Y}]}{\mathbb{E}[\eta_t \mid \mathcal{Y}]}, \quad \mathbb{P}\text{-a.s.} \quad (3.2)$$

This formula allows us to define the measure-valued process  $\rho = (\rho_t)_{t \in [0, T]}$ , called *unnormalized conditional distribution* of  $X$ , or *unnormalized filtering process*, defined, for any bounded and measurable  $\varphi: [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}$ , as:

$$\rho_t(\varphi_t) := \mathbb{E}[\eta_t \varphi(t, X_t) \mid \mathcal{Y}_t], \quad t \in [0, T]. \quad (3.3)$$

Given the properties of  $\pi$  and of  $\eta$  it is possible to show (see, e.g., [2, Lemma 3.18]) that  $\rho$  is càdlàg and  $\mathbb{Y}$ -adapted, hence  $\mathbb{Y}$ -progressively measurable. Moreover, the Kallianpur-Striebel formula implies that for any bounded and measurable  $\varphi: [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}$  and for any fixed  $t \in [0, T]$ :

$$\pi_t(\varphi_t) = \frac{\rho_t(\varphi_t)}{\rho_t(1)}, \quad \mathbb{P}\text{-a.s.}, \quad (3.4)$$

where  $1: \mathbb{R}^m \rightarrow \mathbb{R}$  is the constant function equal to 1.

To describe the local dynamics of the signal process  $X$ , let us introduce the operator  $\mathcal{A}$ , defined for any  $\varphi \in C_b^{1,2}([0, T] \times \mathbb{R}^m)$  as:

$$\mathcal{A}\varphi(t, x) := D_x \varphi(t, x) \cdot b(t, x) + \frac{1}{2} \text{tr}(D_x^2 \varphi(t, x) \sigma \sigma^*(t, x)), \quad (t, x) \in [0, T] \times \mathbb{R}^m. \quad (3.5)$$

We can also define the family of operators  $\mathcal{A}_t$ ,  $t \in [0, T]$ , given by:

$$\mathcal{A}_t \varphi(x) = D_x \varphi(x) \cdot b(t, x) + \frac{1}{2} \text{tr}(D_x^2 \varphi(x) \sigma \sigma^*(t, x)), \quad x \in \mathbb{R}^m, \varphi \in C_b^2(\mathbb{R}^m).$$

<sup>1</sup>Without any particular assumptions on  $\mathbb{Y}$ , the filtering process is adapted with respect to the right-continuous enlargement of  $\mathbb{Y}$ . However, as previously observed, in our model  $\mathbb{Y}$  is already right-continuous.



To obtain the Zakai equation we need, first, to write the semimartingale decomposition of the process  $(\varphi(t, X_t))_{t \in [0, T]}$ . For any  $\varphi \in C_b^{1,2}([0, T] \times \mathbb{R}^m)$  we have, applying Itô's formula:

$$\begin{aligned} \varphi(t, X_t) &= \varphi(0, X_{0-}) + \int_0^t [\partial_s + \mathcal{A}] \varphi(s, X_s) ds + \int_0^t D_x \varphi(s, X_{s-}) d\nu_s \\ &\quad + \sum_{0 \leq s \leq t} \left[ \varphi(s, X_s) - \varphi(s, X_{s-}) - D_x \varphi(s, X_{s-}) \cdot \Delta \nu_s \right] + M_t^\varphi, \quad t \in [0, T]. \end{aligned} \quad (3.6)$$

Here,  $M_t^\varphi := \int_0^t D_x \varphi(s, X_s) \sigma(s, X_s) dW_s$ ,  $t \in [0, T]$ , is a square-integrable  $(\mathbb{P}, \mathbb{F})$ -martingale, thanks to conditions (2.4) and (2.5) (see also Remark 2.3).

We need the following two technical Lemmata. Up to minor modifications, their proofs follow that of [2, Lemma 3.21].

**Lemma 3.2.** *Let  $\Psi = (\Psi_t)_{t \in [0, T]}$  be a real-valued  $(\mathbb{P}, \mathbb{F})$ -progressively measurable process such that*

$$\mathbb{E} \left[ \int_0^T \Psi_s^2 ds \right] < +\infty.$$

*Then, for any  $j = 1, \dots, k$  we have*

$$\mathbb{E} \left[ \int_0^t \Psi_s d\bar{B}_s^j \mid \mathcal{Y} \right] = \int_0^t \mathbb{E}[\Psi_s \mid \mathcal{Y}] d\bar{B}_s^j, \quad t \in [0, T].$$

**Lemma 3.3.** *Let  $\Psi = (\Psi_t)_{t \in [0, T]}$  be a real-valued  $(\mathbb{P}, \mathbb{F})$ -progressively measurable process satisfying<sup>2</sup>*

$$\mathbb{E} \left[ \int_0^T \Psi_s^2 d\langle M^\varphi \rangle_s \right] < +\infty.$$

*Then,*

$$\mathbb{E} \left[ \int_0^t \Psi_s dM_s^\varphi \mid \mathcal{Y} \right] = 0, \quad t \in [0, T].$$

We are now ready to state the main result of this section, namely, to provide the Zakai equation.

**Theorem 3.4.** *Suppose that Assumptions 2.1 and 2.2 are satisfied and, moreover, that*

$$\int_{\mathbb{R}^m} \|x\|^3 \xi(dx) < +\infty. \quad (3.7)$$

*Then, for any  $\varphi \in C_b^{1,2}([0, T] \times \mathbb{R}^m)$ , the unnormalized conditional distribution  $\rho$  satisfies the Zakai equation:*

$$\begin{aligned} \rho_t(\varphi_t) &= \xi(\varphi_0) + \int_0^t \rho_s([\partial_s + \mathcal{A}_s] \varphi_s) ds + \int_0^t \rho_{s-}(D_x \varphi_s) d\nu_s + \int_0^t \gamma^{-1}(s) \rho_s(\varphi_s h_s) d\bar{B}_s \\ &\quad + \sum_{0 \leq s \leq t} \left[ \rho_{s-}(\varphi_s(\cdot + \Delta \nu_s)) - \varphi_s - D_x \varphi_s \cdot \Delta \nu_s \right], \quad \mathbb{P}\text{-a.s.}, \quad t \in [0, T], \end{aligned} \quad (3.8)$$

where  $\xi(\varphi_0) := \int_{\mathbb{R}^m} \varphi(0, x) \xi(dx)$  and, for all  $t \in [0, T]$ ,  $h_t(\cdot) := h(t, \cdot)$ ,

$$\begin{aligned} \int_0^t \rho_{s-}(D_x \varphi_s) d\nu_s &:= \sum_{i=1}^m \int_0^t \rho_{s-}(\partial_i \varphi_s) d\nu_s^i, \\ \int_0^t \gamma^{-1}(s) \rho_s(\varphi_s h_s) d\bar{B}_s &:= \sum_{i=1}^n \sum_{j=1}^n \int_0^t \gamma_{ij}^{-1}(s) \rho_s(\varphi_s h_s^j) d\bar{B}_s^i. \end{aligned}$$

*Proof.* Fix  $t \in [0, T]$  and  $\varphi \in C_b^{1,2}([0, T] \times \mathbb{R}^m)$ . Let us introduce the constants

$$C_\varphi := \sup_{t,x} |\varphi(t, x)|, \quad C'_\varphi := \sup_{t,x} \|D_x \varphi(t, x)\|, \quad C''_\varphi := \sup_{t,x} \|D_x^2 \varphi(t, x)\|,$$

where the suprema are taken over  $[0, T] \times \mathbb{R}^m$ . The proof is organized in several steps.

**Step 1. (Approximation)** For any fixed  $\varepsilon > 0$ , define the bounded process  $\eta^\varepsilon = (\eta_t^\varepsilon)_{t \in [0, T]}$ :

$$\eta_t^\varepsilon := \frac{\eta_t}{1 + \varepsilon \eta_t}, \quad t \in [0, T], \quad (3.9)$$

where  $\eta$  is defined in (2.11). Both  $\eta$  and  $\eta^\varepsilon$  have continuous trajectories and this fact will be used in what follows without further mention.

<sup>2</sup>If  $M$  is any  $(\mathbb{P}, \mathbb{F})$ -square integrable martingale,  $\langle M \rangle$  denotes its  $(\mathbb{P}, \mathbb{F})$ -predictable quadratic variation.

Applying Itô's formula we obtain

$$\eta_t^\varepsilon = \frac{1}{1+\varepsilon} - \int_0^t \frac{\varepsilon \eta_s^2}{(1+\varepsilon \eta_s)^3} \|\gamma^{-1}(s)h(s, X_s)\|^2 ds + \int_0^t \frac{\eta_s}{(1+\varepsilon \eta_s)^2} \gamma^{-1}(s)h(s, X_s) d\bar{B}_s.$$

Denoting by  $[\cdot, \cdot]$  the optional quadratic covariation operator, thanks to the integration by parts rule and recalling (3.6) we get

$$\begin{aligned} \eta_t^\varepsilon \varphi(t, X_t) &= \frac{\varphi(0, X_{0-})}{1+\varepsilon} + \int_0^t \eta_{s-}^\varepsilon d\varphi(s, X_s) + \int_0^t \varphi(s, X_{s-}) d\eta_s^\varepsilon + \int_0^t d[\eta^\varepsilon, \varphi(\cdot, X)]_s \\ &= \frac{\varphi(0, X_{0-})}{1+\varepsilon} + \int_0^t \eta_{s-}^\varepsilon [\partial_s + \mathcal{A}] \varphi(s, X_s) ds + \int_0^t \eta_{s-}^\varepsilon D_x \varphi(s, X_{s-}) d\nu_s \\ &\quad + \sum_{0 \leq s \leq t} \eta_{s-}^\varepsilon [\varphi(s, X_s) - \varphi(s, X_{s-}) - D_x \varphi(s, X_{s-}) \Delta \nu_s] + \int_0^t \eta_{s-}^\varepsilon dM_s^\varphi \\ &\quad - \int_0^t \frac{\varepsilon \eta_s^2 \varphi(s, X_{s-})}{(1+\varepsilon \eta_s)^3} \|\gamma^{-1}(s)h(s, X_s)\|^2 ds + \int_0^t \frac{\eta_s \varphi(s, X_{s-})}{(1+\varepsilon \eta_s)^2} \gamma^{-1}(s)h(s, X_s) d\bar{B}_s. \end{aligned} \quad (3.10)$$

**Step 2. (Projection onto  $\mathbb{Y}$ )** Notice that  $X_t = X_{t-} + \Delta \nu_t$ ,  $\mathbb{P}$ -a.s.,  $t \in [0, T]$ , and that, since  $\mathcal{Y}_{0-} = \mathcal{Y}_0 = \{\emptyset, \Omega\}$ , we have

$$\mathbb{E}[\varphi(0, X_{0-}) \mid \mathcal{Y}_{0-}] = \int_{\mathbb{R}^m} \varphi(0, x) \xi(dx) = \xi(\varphi_0).$$

Therefore, taking conditional expectation with respect to  $\mathcal{Y}$ , we have (rearranging some terms)

$$\begin{aligned} \mathbb{E}[\eta_t^\varepsilon \varphi(t, X_t) \mid \mathcal{Y}] &= \frac{\xi(\varphi_0)}{1+\varepsilon} + \mathbb{E}\left[\int_0^t \eta_{s-}^\varepsilon [\partial_s + \mathcal{A}] \varphi(s, X_s) ds \mid \mathcal{Y}\right] \\ &\quad + \mathbb{E}\left[\int_0^t \eta_{s-}^\varepsilon D_x \varphi(s, X_{s-}) d\nu_s \mid \mathcal{Y}\right] + \mathbb{E}\left[\int_0^t \frac{\eta_s \varphi(s, X_{s-})}{(1+\varepsilon \eta_s)^2} \gamma^{-1}(s)h(s, X_s) d\bar{B}_s \mid \mathcal{Y}\right] \\ &\quad + \mathbb{E}\left[\sum_{0 \leq s \leq t} \eta_{s-}^\varepsilon [\varphi(s, X_{s-} + \Delta \nu_s) - \varphi(s, X_{s-}) - D_x \varphi(s, X_{s-}) \cdot \Delta \nu_s] \mid \mathcal{Y}\right] \\ &\quad + \mathbb{E}\left[\int_0^t \eta_{s-}^\varepsilon dM_s^\varphi \mid \mathcal{Y}\right] - \mathbb{E}\left[\int_0^t \frac{\varepsilon \eta_s^2 \varphi(s, X_{s-})}{(1+\varepsilon \eta_s)^3} \|\gamma^{-1}(s)h(s, X_s)\|^2 ds \mid \mathcal{Y}\right]. \end{aligned} \quad (3.11)$$

We analyze now each of the terms appearing in (3.11). For any bounded  $\mathcal{Y}$ -measurable  $Z$ , thanks to conditions (2.4) and (2.5), there exists a constant  $C_1$ , depending on  $Z$ ,  $\varepsilon$ ,  $\varphi$ ,  $b$ , and  $\sigma$  such that

$$|Z \eta_t^\varepsilon [\partial_t + \mathcal{A}] \varphi(t, X_t)| \leq C_1(1 + \|X_t\|^2), \quad t \in [0, T],$$

which implies, using the estimate given in (2.9),

$$\mathbb{E}\left[\int_0^t Z \eta_s^\varepsilon [\partial_s + \mathcal{A}] \varphi(s, X_s) ds\right] \leq C \mathbb{E}\left[\int_0^t (1 + \|X_s\|^2) ds\right] \leq C_1 T[1 + \kappa(1 + \mathbb{E}[\|X_{0-}\|^2])] < +\infty.$$

Therefore, applying the tower rule and Fubini-Tonelli's theorem,

$$\mathbb{E}\left[Z \mathbb{E}\left[\int_0^t \eta_s^\varepsilon [\partial_s + \mathcal{A}] \varphi(s, X_s) ds \mid \mathcal{Y}\right]\right] = \mathbb{E}\left[Z \int_0^t \mathbb{E}[\eta_s^\varepsilon [\partial_s + \mathcal{A}] \varphi(s, X_s) \mid \mathcal{Y}] ds\right],$$

whence

$$\mathbb{E}\left[\int_0^t \eta_s^\varepsilon [\partial_s + \mathcal{A}] \varphi(s, X_s) ds \mid \mathcal{Y}\right] = \int_0^t \mathbb{E}[\eta_s^\varepsilon [\partial_s + \mathcal{A}] \varphi(s, X_s) \mid \mathcal{Y}] ds. \quad (3.12)$$

Similarly, for any bounded  $\mathcal{Y}$ -measurable  $Z$  we have that

$$\|Z \eta_t^\varepsilon D_x \varphi(t, X_{t-})\| \leq \frac{|Z| C'_\varphi}{\varepsilon} < +\infty, \quad d\mathbb{P} \otimes dt\text{-a.e.}$$

This fact will allow to use Fubini-Tonelli's theorem in formula (3.13) below. We need to introduce the changes of time associated to the processes  $\nu^{i,+}$  and  $\nu^{i,-}$ ,  $i = 1, \dots, m$ , defined as

$$C_t^{i,+} := \inf\{s \geq 0: \nu_s^{i,+} \geq t\}, \quad C_t^{i,-} := \inf\{s \geq 0: \nu_s^{i,-} \geq t\}, \quad t \geq 0, \quad i = 1, \dots, m,$$

where  $\nu^{i,+}$  (resp.  $\nu^{i,-}$ ) denotes the positive part (resp. negative part) process of the  $i$ -th component of process  $\nu$  (see the list of notation in Section 1.2 for a more detailed definition).

For each  $t \geq 0$  and  $i = 1, \dots, m$ ,  $C_t^{i,+}$  and  $C_t^{i,-}$  are  $\mathbb{Y}$ -stopping times (see, e.g., [18, Chapter VI, Def. 56] or [25, Proposition I.1.28]). Hence, applying the change of time formula (see, e.g., [18, Chapter VI, Equation (55.1)] or [25, Equation (1), p. 29]) and Fubini-Tonelli's theorem, we get

$$\begin{aligned}
 & \mathbb{E} \left[ Z \mathbb{E} \left[ \int_0^t \eta_s^\varepsilon D_x \varphi(s, X_{s-}) d\nu_s \mid \mathcal{Y} \right] \right] = \mathbb{E} \left[ \int_0^{+\infty} \mathbf{1}_{s \leq t} Z \eta_s^\varepsilon D_x \varphi(s, X_{s-}) d\nu_s \right] \\
 &= \sum_{i=1}^m \mathbb{E} \left[ \int_0^{+\infty} \mathbf{1}_{s \leq t} Z \eta_s^\varepsilon \partial_i \varphi(s, X_{s-}) d\nu_s^{i,+} \right] - \sum_{i=1}^m \mathbb{E} \left[ \int_0^{+\infty} \mathbf{1}_{s \leq t} Z \eta_s^\varepsilon \partial_i \varphi(s, X_{s-}) d\nu_s^{i,-} \right] \\
 &= \sum_{i=1}^m \mathbb{E} \left[ \int_0^{+\infty} \mathbf{1}_{C_s^{i,+} \leq t} Z \eta_{C_s^{i,+}}^\varepsilon \partial_i \varphi(C_s^{i,+}, X_{(C_s^{i,+})-}) \mathbf{1}_{C_s^{i,+} < +\infty} ds \right] \\
 &\quad - \sum_{i=1}^m \mathbb{E} \left[ \int_0^{+\infty} \mathbf{1}_{C_s^{i,-} \leq t} Z \eta_{C_s^{i,-}}^\varepsilon \partial_i \varphi(C_s^{i,-}, X_{(C_s^{i,-})-}) \mathbf{1}_{C_s^{i,-} < +\infty} ds \right] \\
 &= \sum_{i=1}^m \int_0^{+\infty} \mathbb{E} \left[ \mathbf{1}_{C_s^{i,+} \leq t} Z \mathbb{E} [\eta_{C_s^{i,+}}^\varepsilon \partial_i \varphi(C_s^{i,+}, X_{(C_s^{i,+})-}) \mid \mathcal{Y}] \mathbf{1}_{C_s^{i,+} < +\infty} \right] ds \\
 &\quad - \sum_{i=1}^m \int_0^{+\infty} \mathbb{E} \left[ \mathbf{1}_{C_s^{i,-} \leq t} Z \mathbb{E} [\eta_{C_s^{i,-}}^\varepsilon \partial_i \varphi(C_s^{i,-}, X_{(C_s^{i,-})-}) \mid \mathcal{Y}] \mathbf{1}_{C_s^{i,-} < +\infty} \right] ds \\
 &= \sum_{i=1}^m \mathbb{E} \left[ \int_0^{+\infty} \mathbf{1}_{s \leq t} Z \mathbb{E} [\eta_s^\varepsilon \partial_i \varphi(s, X_{s-}) \mid \mathcal{Y}] d\nu_s^{i,+} \right] - \sum_{i=1}^m \mathbb{E} \left[ \int_0^{+\infty} \mathbf{1}_{s \leq t} Z \mathbb{E} [\eta_s^\varepsilon \partial_i \varphi(s, X_{s-}) \mid \mathcal{Y}] d\nu_s^{i,-} \right] \\
 &= \mathbb{E} \left[ \int_0^{+\infty} \mathbf{1}_{s \leq t} Z \mathbb{E} [\eta_s^\varepsilon D_x \varphi(s, X_{s-}) \mid \mathcal{Y}] d\nu_s \right] = \mathbb{E} \left[ Z \int_0^t \mathbb{E} [\eta_s^\varepsilon D_x \varphi(s, X_{s-}) \mid \mathcal{Y}] d\nu_s \right], \tag{3.13}
 \end{aligned}$$

whence

$$\mathbb{E} \left[ \int_0^t \eta_s^\varepsilon D_x \varphi(s, X_{s-}) d\nu_s \mid \mathcal{Y} \right] = \int_0^t \mathbb{E} [\eta_s^\varepsilon D_x \varphi(s, X_{s-}) \mid \mathcal{Y}] d\nu_s. \tag{3.14}$$

Next, using (A.4) we obtain

$$\mathbb{E} \left[ \int_0^t \left( \frac{\eta_s^\varepsilon}{1 + \varepsilon \eta_s} \varphi(s, X_{s-}) \|\gamma^{-1}(s) h(s, X_s)\| \right)^2 ds \right] \leq \frac{C_\varphi^2}{\varepsilon^2} \mathbb{E} \left[ \int_0^t \|\gamma^{-1}(s) h(s, X_s)\|^2 ds \right] < +\infty,$$

hence, by Lemma 3.2 we have:

$$\mathbb{E} \left[ \int_0^t \frac{\eta_s \varphi(s, X_{s-})}{(1 + \varepsilon \eta_s)^2} \gamma^{-1}(s) h(s, X_s) d\overline{B}_s \mid \mathcal{Y} \right] = \int_0^t \mathbb{E} \left[ \frac{\eta_s \varphi(s, X_{s-})}{(1 + \varepsilon \eta_s)^2} \gamma^{-1}(s) h(s, X_s) \mid \mathcal{Y} \right] d\overline{B}_s. \tag{3.15}$$

Recalling that  $|\nu|_T^i \leq K$ ,  $\mathbb{P}$ -a.s., and hence  $|\Delta \nu_t^i| \leq K$ , for all  $t \in [0, T]$  and all  $i = 1, \dots, m$ ,  $\mathbb{P}$ -a.s., for any bounded  $\mathcal{Y}$ -measurable  $Z$  we have that

$$\begin{aligned}
 & \sum_{0 \leq s \leq t} \mathbb{E} \left| Z \eta_s^\varepsilon \left[ \varphi(s, X_{s-} + \Delta \nu_s) - \varphi(s, X_{s-}) - D_x \varphi(s, X_{s-}) \Delta \nu_s \right] \right| \leq \frac{|Z|}{\varepsilon} \frac{C_\varphi''}{2} \sum_{0 \leq s \leq t} \mathbb{E} [\|\Delta \nu_s\|^2] \\
 &= \frac{|Z|}{\varepsilon} \frac{C_\varphi''}{2} \mathbb{E} \left[ \sum_{0 \leq s \leq t} \Delta \nu_s^* \Delta \nu_s \right] \leq \frac{|Z|}{\varepsilon} \frac{C_\varphi''}{2} \sum_{i=1}^m \mathbb{E} \left[ \int_0^t |\Delta \nu_s^i| d|\nu^i|_s \right] \leq \frac{|Z|}{\varepsilon} \frac{C_\varphi''}{2} m K^2 < +\infty.
 \end{aligned}$$

Therefore, using once more Fubini-Tonelli's theorem

$$\begin{aligned}
 & \mathbb{E} \left[ Z \mathbb{E} \left[ \sum_{0 \leq s \leq t} \eta_s^\varepsilon \left[ \varphi(s, X_{s-} + \Delta \nu_s) - \varphi(s, X_{s-}) - D_x \varphi(s, X_{s-}) \cdot \Delta \nu_s \right] \mid \mathcal{Y} \right] \right] \\
 &= \mathbb{E} \left[ Z \sum_{0 \leq s \leq t} \mathbb{E} \left[ \eta_s^\varepsilon \left[ \varphi(s, X_{s-} + \Delta \nu_s) - \varphi(s, X_{s-}) - D_x \varphi(s, X_{s-}) \cdot \Delta \nu_s \right] \mid \mathcal{Y} \right] \right],
 \end{aligned}$$

and hence

$$\begin{aligned}
 & \mathbb{E} \left[ \sum_{0 \leq s \leq t} \eta_s^\varepsilon \left[ \varphi(s, X_{s-} + \Delta \nu_s) - \varphi(s, X_{s-}) - D_x \varphi(s, X_{s-}) \cdot \Delta \nu_s \right] \mid \mathcal{Y} \right] \\
 &= \sum_{0 \leq s \leq t} \mathbb{E} \left[ \eta_s^\varepsilon \left[ \varphi(s, X_{s-} + \Delta \nu_s) - \varphi(s, X_{s-}) - D_x \varphi(s, X_{s-}) \cdot \Delta \nu_s \right] \mid \mathcal{Y} \right]. \tag{3.16}
 \end{aligned}$$

Finally, being  $\eta^\varepsilon$  bounded, Lemma 3.3 entails  $\mathbb{E}[\int_0^t \eta_s^\varepsilon dM_s^\varphi \mid \mathcal{Y}] = 0$ , and, using the same rationale of the previous evaluations,

$$\mathbb{E}\left[\int_0^t \frac{\varepsilon \eta_s^2 \varphi(s, X_{s-})}{(1 + \varepsilon \eta_s)^3} \|\gamma^{-1}(s)h(s, X_s)\|^2 ds \mid \mathcal{Y}\right] = \int_0^t \mathbb{E}\left[\frac{\varepsilon \eta_s^2 \varphi(s, X_{s-})}{(1 + \varepsilon \eta_s)^3} \|\gamma^{-1}(s)h(s, X_s)\|^2 \mid \mathcal{Y}\right] ds. \quad (3.17)$$

Taking into account (3.12), (3.14), (3.15), (3.16), and (3.17), Equation (3.11) becomes

$$\begin{aligned} \mathbb{E}[\eta_t^\varepsilon \varphi(t, X_t) \mid \mathcal{Y}] &= \frac{\xi(\varphi_0)}{1 + \varepsilon} + \int_0^t \mathbb{E}[\eta_s^\varepsilon [\partial_s + \mathcal{A}] \varphi(s, X_s) \mid \mathcal{Y}] ds \\ &\quad - \int_0^t \mathbb{E}\left[\frac{\varepsilon \eta_s^2 \varphi(s, X_s)}{(1 + \varepsilon \eta_s)^3} \|\gamma^{-1}(s)h(s, X_s)\|^2 \mid \mathcal{Y}\right] ds \\ &\quad + \int_0^t \mathbb{E}\left[\frac{\eta_s \varphi(s, X_s)}{(1 + \varepsilon \eta_s)^2} \gamma^{-1}(s)h(s, X_s) \mid \mathcal{Y}\right] d\overline{B}_s + \int_0^t \mathbb{E}[\eta_s^\varepsilon D_x \varphi(s, X_{s-}) \mid \mathcal{Y}] d\nu_s \\ &\quad + \sum_{0 \leq s \leq t} \mathbb{E}\left[\eta_s^\varepsilon [\varphi(s, X_{s-} + \Delta\nu_s) - \varphi(s, X_{s-}) - D_x \varphi(s, X_{s-}) \cdot \Delta\nu_s] \mid \mathcal{Y}\right]. \end{aligned} \quad (3.18)$$

**Step 3. (Taking limits)** It remains to show that all the terms appearing in (3.18) converge appropriately to give (3.8). As  $\varepsilon \rightarrow 0$ , we have that  $\eta_t^\varepsilon \rightarrow \eta_t$ ,  $\mathbb{E}[\eta_t^\varepsilon \varphi(t, X_t) \mid \mathcal{Y}] \rightarrow \rho_t(\varphi)$ , for all  $t \in [0, T]$ , and

$$\mathbb{E}[\eta_t^\varepsilon [\partial_t + \mathcal{A}] \varphi(t, X_t) \mid \mathcal{Y}] \rightarrow \rho_t([\partial_t + \mathcal{A}] \varphi_t), \quad d\mathbb{P} \otimes dt\text{-a.e.}$$

Using boundedness of  $\varphi$  and (2.4), (2.5), we get that

$$|\mathbb{E}[\eta_t^\varepsilon [\partial_t + \mathcal{A}] \varphi(t, X_t) \mid \mathcal{Y}]| \leq C_2 \mathbb{E}[\eta_t(1 + \|X_t\|^2) \mid \mathcal{Y}], \quad t \in [0, T],$$

for some constant  $C_2$ , depending on  $\varphi$ ,  $b$ , and  $\sigma$ . The r.h.s. of this inequality is  $d\mathbb{P} \otimes dt$  integrable on  $\Omega \times [0, t]$ , since (apply again the tower rule and Fubini-Tonelli's theorem)

$$\mathbb{E}\left[\int_0^t C_2 \mathbb{E}[\eta_s(1 + \|X_s\|^2) \mid \mathcal{Y}] ds\right] \leq C_2 T \{1 + \kappa(1 + \tilde{\mathbb{E}}[\|X_{0-}\|^2])\} < +\infty,$$

where we used (2.9) (which holds also under  $\tilde{\mathbb{P}}$  because the dynamics of  $X$  does not change under this measure), and the fact that  $\tilde{\mathbb{E}}[\|X_{0-}\|^2] = \mathbb{E}[\|X_{0-}\|^2] < +\infty$ .

Using the conditional form of the dominated convergence theorem, we have that, for all  $t \in [0, T]$ ,

$$\mathbb{E}\left[\int_0^t \mathbb{E}[\eta_s^\varepsilon [\partial_s + \mathcal{A}] \varphi(s, X_s) \mid \mathcal{Y}] ds \mid \mathcal{Y}\right] \rightarrow \mathbb{E}\left[\int_0^t \rho_s([\partial_s + \mathcal{A}] \varphi_s) ds \mid \mathcal{Y}\right], \quad \mathbb{P}\text{-a.s.},$$

as  $\varepsilon \rightarrow 0$ , whence, noticing that the integrals are  $\mathcal{Y}$ -measurable random variables,

$$\int_0^t \mathbb{E}[\eta_s^\varepsilon [\partial_s + \mathcal{A}] \varphi(s, X_s) \mid \mathcal{Y}] ds \rightarrow \int_0^t \rho_s([\partial_s + \mathcal{A}] \varphi_s) ds, \quad \mathbb{P}\text{-a.s.}, \quad \forall t \in [0, T].$$

We consider, now, the term on the second line of (3.18). We have that, for all  $t \in [0, T]$ ,

$$\mathbb{E}\left[\frac{\varepsilon \eta_t^2 \varphi(t, X_t)}{(1 + \varepsilon \eta_t)^3} \|\gamma^{-1}(t)h(t, X_t)\|^2 \mid \mathcal{Y}\right] \rightarrow 0,$$

as  $\varepsilon \rightarrow 0$ , and that

$$\mathbb{E}\left[\frac{\varepsilon \eta_t^2 \varphi(t, X_t)}{(1 + \varepsilon \eta_t)^3} \|\gamma^{-1}(t)h(t, X_t)\|^2 \mid \mathcal{Y}\right] \leq C_\varphi \mathbb{E}[\eta_t \|\gamma^{-1}(t)h(t, X_t)\|^2 \mid \mathcal{Y}].$$

The r.h.s. of the last inequality is  $d\mathbb{P} \otimes dt$  integrable on  $\Omega \times [0, t]$ , since

$$\mathbb{E}\left[\int_0^t C_\varphi \mathbb{E}[\eta_s \|\gamma^{-1}(s)h(s, X_s)\|^2 \mid \mathcal{Y}] ds\right] \leq n C_\varphi C_h C_\gamma T [1 + \kappa(1 + \tilde{\mathbb{E}}[\|X_{0-}\|^2])] < +\infty,$$

where we used (A.4), that holds also under  $\tilde{\mathbb{P}}$  (again, because the dynamics of  $X$  does not change under this measure), and the fact that  $\tilde{\mathbb{E}}[\|X_{0-}\|^2] = \mathbb{E}[\|X_{0-}\|^2] < +\infty$ .

Hence, reasoning as above, after applying the conditional form of the dominated convergence theorem we obtain that, for all  $t \in [0, T]$ , as  $\varepsilon \rightarrow 0$ ,

$$\int_0^t \mathbb{E}\left[\frac{\varepsilon \eta_s^2 \varphi(s, X_s)}{(1 + \varepsilon \eta_s)^3} \|\gamma^{-1}(s)h(s, X_s)\|^2 \mid \mathcal{Y}\right] ds \rightarrow 0, \quad \mathbb{P}\text{-a.s.}$$

Looking at the third line of (3.18), the next step is to show that

$$\int_0^t \mathbb{E}\left[\frac{\eta_s \varphi(s, X_s)}{(1 + \varepsilon \eta_s)^2} \gamma^{-1}(s)h(s, X_s) \mid \mathcal{Y}\right] d\overline{B}_s \rightarrow \int_0^t \gamma^{-1}(s) \rho_s(\varphi_s h_s) d\overline{B}_s, \quad \mathbb{P}\text{-a.s.}$$

The proof of this fact is standard (see, e.g., [2, Theorem 3.24 and Exercise 3.25.i] or [5, Theorem 4.1.1]). It is important to notice that condition (3.7) intervenes here.

Next, we examine the other integral in the third line of (3.18). We have that

$$\mathbb{E}[\eta_t^\varepsilon D_x \varphi(t, X_{t-}) \mid \mathcal{Y}] \longrightarrow \rho_{t-}(D_x \varphi_t), \quad \mathbb{P}\text{-a.s.},$$

as  $\varepsilon \rightarrow 0$ , for all  $t \in [0, T]$ . Notice that, for any  $t \in [0, T]$ ,

$$\|\mathbb{E}[\eta_t^\varepsilon D_x \varphi(t, X_{t-}) \mid \mathcal{Y}]\| \leq C'_\varphi \mathbb{E}[\eta_t \mid \mathcal{Y}].$$

Since  $\eta$  is non-negative, a  $\mathbb{Y}$ -optional version of  $\{\mathbb{E}[\mathbf{1}_{t \leq T} \eta_t \mid \mathcal{Y}_t]\}_{t \geq 0}$  is given by the  $\mathbb{Y}$ -optional projection of  $\{\mathbf{1}_{t \leq T} \eta_t\}_{t \geq 0}$  (see, e.g., [13, Corollary 7.6.8]). Therefore, applying [18, Chapter VI, Theorem 57] and using Lemma 3.1 we get that for all  $t \in [0, T]$ , and all  $i = 1, \dots, m$ ,

$$\mathbb{E} \left[ \int_0^T C'_\varphi \mathbb{E}[\eta_t \mid \mathcal{Y}] d|\nu^i|_t \right] = C'_\varphi \mathbb{E} \left[ \int_0^{+\infty} \mathbb{E}[\mathbf{1}_{t \leq T} \eta_t \mid \mathcal{Y}_t] d|\nu^i|_t \right] = C'_\varphi \mathbb{E} \left[ \int_0^T \eta_t d|\nu^i|_t \right] < +\infty,$$

where finiteness of  $\mathbb{E}[\int_0^T \eta_t d|\nu^i|_t]$  can be established with a reasoning analogous to the proof of (A.10). Therefore, we can apply the conditional form of the dominated convergence theorem, to obtain that, for all  $t \in [0, T]$ , as  $\varepsilon \rightarrow 0$ ,

$$\mathbb{E} \left[ \int_0^t \mathbb{E}[\eta_s^\varepsilon D_x \varphi(s, X_{s-}) \mid \mathcal{Y}] d\nu_s \mid \mathcal{Y} \right] \longrightarrow \mathbb{E} \left[ \int_0^t \rho_{s-}(D_x \varphi_s) d\nu_s \mid \mathcal{Y} \right], \quad \mathbb{P}\text{-a.s.}$$

Since the integrals are  $\mathcal{Y}$ -measurable random variables, this implies that, for all  $t \in [0, T]$ , as  $\varepsilon \rightarrow 0$ ,

$$\int_0^t \mathbb{E}[\eta_s^\varepsilon D_x \varphi(s, X_{s-}) \mid \mathcal{Y}] d\nu_s \longrightarrow \int_0^t \rho_{s-}(D_x \varphi_s) d\nu_s, \quad \mathbb{P}\text{-a.s.}$$

Finally, looking at the fourth line of (3.18), we have that, for all  $t \in [0, T]$ , as  $\varepsilon \rightarrow 0$ ,

$$\begin{aligned} \Lambda_t^\varepsilon &:= \mathbb{E} \left[ \eta_t^\varepsilon \left[ \varphi(t, X_{t-} + \Delta \nu_t) - \varphi(t, X_{t-}) - D_x \varphi(t, X_{t-}) \cdot \Delta \nu_t \right] \mid \mathcal{Y} \right] \\ &\longrightarrow \rho_{t-}(\varphi_t(\cdot + \Delta \nu_t) - \varphi_t - D_x \varphi_t \cdot \Delta \nu_t), \quad \mathbb{P}\text{-a.s.} \end{aligned}$$

Observe that, for any  $t \in [0, T]$ ,  $\Lambda_t^\varepsilon$  is bounded by  $\frac{1}{2} C''_\varphi \mathbb{E}[\eta_t \|\Delta \nu_t\|^2 \mid \mathcal{Y}]$ , which is positive and integrable with respect to the product of measure  $\mathbb{P}$  and the jump measure associated to  $\nu$ , since:

$$\begin{aligned} \mathbb{E} \left[ \sum_{0 \leq s \leq t} \frac{1}{2} C''_\varphi \mathbb{E}[\eta_s \|\Delta \nu_s\|^2 \mid \mathcal{Y}] \right] &= \frac{1}{2} C''_\varphi \sum_{0 \leq s \leq t} \mathbb{E}[\eta_s \|\Delta \nu_s\|^2] = \frac{1}{2} C''_\varphi \mathbb{E} \left[ \sum_{0 \leq s \leq t} \eta_s \Delta \nu_s \cdot \Delta \nu_s \right] \\ &\leq \frac{1}{2} C''_\varphi \sum_{i=1}^m \mathbb{E} \left[ \int_0^t \eta_s |\Delta \nu_s^i| d|\nu^i|_s \right] \leq \frac{1}{2} C''_\varphi K \sum_{i=1}^m \mathbb{E} \left[ \int_0^t \eta_s d|\nu^i|_s \right] < +\infty. \end{aligned}$$

By the conditional form of the dominated convergence theorem, we have that, for all  $t \in [0, T]$ , as  $\varepsilon \rightarrow 0$ ,

$$\mathbb{E} \left[ \sum_{0 \leq s \leq t} \Lambda_s^\varepsilon \mid \mathcal{Y} \right] \longrightarrow \mathbb{E} \left[ \sum_{0 \leq s \leq t} \rho_{s-}(\varphi_s(\cdot + \Delta \nu_s) - \varphi_s - D_x \varphi_s \cdot \Delta \nu_s) \mid \mathcal{Y} \right], \quad \mathbb{P}\text{-a.s.}$$

and since the sums are  $\mathcal{Y}$ -measurable random variables, this implies that, for all  $t \in [0, T]$ , as  $\varepsilon \rightarrow 0$ ,

$$\sum_{0 \leq s \leq t} \Lambda_s^\varepsilon \longrightarrow \sum_{0 \leq s \leq t} \rho_{s-}(\varphi_s(\cdot + \Delta \nu_s) - \varphi_s - D_x \varphi_s \cdot \Delta \nu_s), \quad \mathbb{P}\text{-a.s.} \quad \square$$

*Remark 3.1.* If the jump times of the process  $\nu$  do not accumulate over  $[0, T]$ , then the Zakai equation can be split into successive linear SPDEs between the jumps of  $\nu$  (i.e., of  $X$ ). Set  $T_0 = 0$ , denote by  $(T_n)_{n \in \mathbb{N}}$  the sequence of jump times of  $\nu$  and indicate by  $\nu^c$  the continuous part of  $\nu$ . Then, for any  $\varphi \in C_b^{1,2}([0, T] \times \mathbb{R}^m)$  and any  $n \in \mathbb{N}_0$  we have  $\mathbb{P}$ -a.s.

$$\begin{cases} d\rho_t(\varphi_t) = \rho_t([\partial_t + \mathcal{A}_t]\varphi_t)dt + \rho_{t-}(D_x \varphi_t) d\nu_t^c + \gamma^{-1}(t) \rho_t(\varphi_t h_t) d\bar{B}_t, & t \in [T_n \wedge T, T_{n+1} \wedge T), \\ \rho_{0-}(\varphi_0) = \xi(\varphi_0), \\ \rho_{T_n}(\varphi) = \rho_{T_n-}(\varphi_{T_n}(\cdot + \Delta \nu_{T_n})). \end{cases} \quad (3.19)$$

We are now ready to deduce, from the Zakai equation, the *Kushner-Stratonovich* equation, i.e., the equation satisfied by the filtering process  $\pi$ , defined in (3.1). The proof of the following two results follows essentially the same steps of [2, Lemma 3.29 and Theorem 3.30], up to necessary modifications due to the present setting (see, also, [5, Lemma 4.3.1 and Theorem 4.3.1]).

**Lemma 3.5.** *Under the same assumptions of Theorem 3.4, the process  $(\rho_t(1))_{t \in [0, T]}$  satisfies for all  $t \in [0, T]$*

$$\rho_t(1) = \exp \left\{ \int_0^t \gamma^{-1}(s) \pi_s(h_s) d\bar{B}_s - \frac{1}{2} \int_0^t \|\gamma^{-1}(s) \pi_s(h_s)\|^2 ds \right\}, \quad \mathbb{P}\text{-a.s.}$$

**Theorem 3.6.** *Under the same assumptions of Theorem 3.4, the process  $(\pi_t(\varphi))_{t \in [0, T]}$  satisfies for all  $t \in [0, T]$  and all  $\varphi \in C_b^{1,2}([0, T] \times \mathbb{R}^m)$  the Kushner-Stratonovich equation*

$$\begin{aligned} \pi_t(\varphi_t) &= \pi_{0-}(\varphi_0) + \int_0^t \pi_s([\partial_s + \mathcal{A}_s]\varphi_s) ds + \int_0^t \pi_{s-}(D_x \varphi_s) d\nu_s \\ &\quad + \int_0^t \gamma^{-1}(s) \left\{ \pi_s(\varphi_s h_s) - \pi_s(\varphi_s) \pi_s(h_s) \right\} [d\bar{B}_s - \gamma^{-1}(s) \pi_s(h_s) ds] \\ &\quad + \sum_{0 \leq s \leq t} \left[ \pi_{s-}(\varphi_s(\cdot + \Delta\nu_s) - \varphi_s - D_x \varphi_s \cdot \Delta\nu_s) \right], \quad \mathbb{P}\text{-a.s.} \end{aligned} \quad (3.20)$$

*Remark 3.2.* It is not difficult to show (see, e.g., [2, Proposition 2.30] or [5, Theorem 4.3.4]), that

$$I_t := \bar{B}_t - \gamma^{-1}(t) \pi_t(h_t), \quad t \in [0, T],$$

is a  $(\tilde{\mathbb{P}}, \mathbb{Y})$ -Brownian motion, the so-called *innovation process*. This allows to rewrite the Kushner-Stratonovich equation in the (perhaps more familiar) form

$$\begin{aligned} \pi_t(\varphi_t) &= \xi(\varphi_0) + \int_0^t \pi_s([\partial_s + \mathcal{A}_s]\varphi_s) ds + \int_0^t \pi_{s-}(D_x \varphi_s) d\nu_s \\ &\quad + \int_0^t \gamma^{-1}(s) \left\{ \pi_s(\varphi_s h_s) - \pi_s(\varphi_s) \pi_s(h_s) \right\} dI_s \\ &\quad + \sum_{0 \leq s \leq t} \left[ \pi_{s-}(\varphi_s(\cdot + \Delta\nu_s) - \varphi_s - D_x \varphi_s \cdot \Delta\nu_s) \right], \quad \tilde{\mathbb{P}}\text{-a.s.,} \quad t \in [0, T]. \end{aligned}$$

Notice, however, that in this setting the innovation process is not a Brownian motion given *a priori*, because it depends (through the density process  $\eta$ , and hence through  $X$ ), on the initial law  $\xi$  of the signal process and on process  $\nu$ .

*Remark 3.3.* Similarly to what stated in Remark 3.1, if the jump times of the process  $\nu$  do not accumulate over  $[0, T]$ , then the Kushner-Stratonovich equation can be split into successive nonlinear SPDEs between the jumps of  $\nu$  (i.e., of  $X$ ). Using the same notation of the aforementioned Remark, for any  $\varphi \in C_b^{1,2}([0, T] \times \mathbb{R}^m)$  and any  $n \in \mathbb{N}_0$  we have  $\mathbb{P}$ -a.s.

$$\begin{cases} d\pi_t(\varphi_t) = \pi_t([\partial_t + \mathcal{A}_t]\varphi_t) dt + \pi_{t-}(D_x \varphi_t) d\nu_t^c \\ \quad + \gamma^{-1}(t) \left\{ \pi_t(\varphi_t h_t) - \pi_t(\varphi_t) \pi_t(h_t) \right\} [d\bar{B}_t - \gamma^{-1}(t) \pi_t(h_t) dt], \quad t \in [T_n \wedge T, T_{n+1} \wedge T), \\ \pi_{0-}(\varphi_0) = \xi(\varphi_0), \\ \pi_{T_n}(\varphi) = \pi_{T_n-}(\varphi_{T_n}(\cdot + \Delta\nu_{T_n})). \end{cases} \quad (3.21)$$

#### 4. UNIQUENESS OF THE SOLUTION TO THE ZAKAI EQUATION

In this section we will address the issue of uniqueness of the solution to the Zakai equation (3.8), under the requirement that the jump times of the process  $\nu$  do not accumulate over  $[0, T]$ . Proving uniqueness is essential to characterize completely the unnormalized filtering process  $\rho$ , defined in (3.3), and is crucial in applications, e.g., in optimal control. Indeed, having ensured that (3.8) (or, equivalently, (3.20)) uniquely characterizes the conditional distribution of the signal given the observation, the filtering process can be employed as a state variable to solve the related separated optimal control problem (cf. [5]).

We follow the approach in [30] (see, also, [2, Chapter 7] and [37, Chapter 6]). The idea is to recast the measure-valued Zakai equation into an SPDE in the Hilbert space  $H := L^2(\mathbb{R}^m)$  and, therefore, to look for a density of  $\rho$  in this space. To accomplish that, we will smooth solutions to (3.8) using the heat kernel, and we will then use estimates in  $L^2(\mathbb{R}^m)$  in order to deduce the desired result. An important role in the subsequent analysis is played by the following lemma, whose proof can be found, e.g., in [2, Solution to Exercise 7.2].

**Lemma 4.1.** *Let  $\{\varphi_k\}_{k \in \mathbb{N}}$  be an orthonormal basis of  $H$  such that  $\varphi_k \in C_b(\mathbb{R}^m)$  for any  $k \in \mathbb{N}$ , and let  $\mu \in \mathcal{M}(\mathbb{R}^m)$  be a finite measure. If*

$$\sum_{k \in \mathbb{N}} [\mu(\varphi_k)]^2 < +\infty,$$

then  $\mu$  is absolutely continuous with respect to Lebesgue measure on  $\mathbb{R}^m$  and its density is square-integrable.

Let  $\psi_\varepsilon$  be the heat kernel, i.e., the function defined for each  $\varepsilon > 0$  as

$$\psi_\varepsilon(x) := \frac{1}{(2\pi\varepsilon)^{m/2}} e^{-\frac{\|x\|^2}{2\varepsilon}}, \quad x \in \mathbb{R}^m,$$

and for any Borel-measurable and bounded  $f$  and  $\varepsilon > 0$  define the operator

$$T_\varepsilon f(x) := \int_{\mathbb{R}^m} \psi_\varepsilon(x-y) f(y) dy, \quad x \in \mathbb{R}^m.$$

We also define the operator  $T_\varepsilon: \mathcal{M}(\mathbb{R}^m) \rightarrow \mathcal{M}(\mathbb{R}^m)$  given by

$$T_\varepsilon \mu(f) := \mu(T_\varepsilon f) = \int_{\mathbb{R}^m} f(y) \underbrace{\int_{\mathbb{R}^m} \psi_\varepsilon(x-y) \mu(dx)}_{:= T_\varepsilon \mu(y)} dy = \int_{\mathbb{R}^m} f(y) T_\varepsilon \mu(y) dy.$$

The equalities above imply that for any  $\mu \in \mathcal{M}(\mathbb{R}^m)$  the measure  $T_\varepsilon \mu$  always possesses a density with respect to Lebesgue measure, that we will still denote by  $T_\varepsilon \mu$ .

*Remark 4.1.* It is important to notice that, by [2, Exercise 7.3, point ii.],  $T_\varepsilon \mu \in W_k^2(\mathbb{R}^m)$ , for any  $\mu \in \mathcal{M}(\mathbb{R}^m)$ ,  $\varepsilon > 0$ , and  $k \in \mathbb{N}$ .

Further properties of these operators that will be used in the sequel are listed in the following Lemma (for its proof see, e.g., [2, Solution to Exercise 7.3] and [37, Lemma 6.7, Lemma 6.8]).

**Lemma 4.2.** *For any  $\mu \in \mathcal{M}(\mathbb{R}^m)$ ,  $h \in H$ , and  $\varepsilon > 0$  we have that:*

- i.  $\|T_{2\varepsilon}|\mu|\|_H \leq \|T_\varepsilon|\mu|\|_H$ , where  $|\mu|$  denotes the total variation measure of  $\mu$ ;
- ii.  $\|T_\varepsilon h\|_H \leq \|h\|_H$ ;
- iii.  $\langle T_\varepsilon \mu, h \rangle = \mu(T_\varepsilon h)$ ;
- iv. If, in addition,  $\partial_i h \in H$ ,  $i = 1, \dots, m$ , then  $\partial_i T_\varepsilon h = T_\varepsilon \partial_i h$  (with the partial derivative understood in the weak sense).
- v. If  $\varphi \in C_b^1(\mathbb{R}^m)$ , then  $\partial_i T_\varepsilon \varphi = T_\varepsilon (\partial_i \varphi)$ .

In this section we will work under the following hypotheses, in addition to Assumptions 2.1 and 2.2, concerning coefficients  $b$ ,  $\sigma$  and  $h$  appearing in SDEs (2.8) and (2.13). In what follows we will use the shorter notation

$$a(t, x) := \frac{1}{2} \sigma \sigma^*(t, x), \quad t \in [0, T], x \in \mathbb{R}^m. \quad (4.1)$$

**Assumption 4.1.** There exist constants  $K_b, K_\sigma, K_h$ , such that, for all  $i, j = 1, \dots, m$ , all  $\ell = 1, \dots, n$ , all  $t \in [0, T]$ , and all  $x \in \mathbb{R}^m$ ,

$$|b_i(t, x)| \leq K_b, \quad |a_{ij}(t, x)| \leq K_\sigma, \quad |h_\ell(t, x)| \leq K_h.$$

In the next section, we obtain the uniqueness result for the solution to the Zakai equation when the process  $\nu$  has continuous paths. This will be then exploited in Section 4.2 in order to obtain the uniqueness claim when  $\nu$  has jump times that do not accumulate over  $[0, T]$ .

**4.1. The case in which  $\nu$  has continuous paths.** We start our analysis with the following Lemma, which will play a fundamental role in the sequel. Its proof can be found in Appendix A.

**Lemma 4.3.** *Suppose that Assumption 4.1 holds. Let  $\zeta = (\zeta_t)_{t \in [0, T]}$  be a  $\mathbb{Y}$ -adapted, càdlàg,  $\mathcal{M}_+(\mathbb{R}^m)$ -valued solution of (3.8), with  $\zeta_{0-} = \xi \in \mathcal{P}(\mathbb{R}^m)$ . If  $\nu$  is continuous, then for any  $\varepsilon > 0$*

$$\mathbb{E}[\sup_{t \in [0, T]} \|T_\varepsilon \zeta_t - \zeta_t\|_H^2] < +\infty.$$

The next result is a useful estimate.

**Proposition 4.4.** *Suppose that Assumption 4.1 holds. Let  $\zeta = (\zeta_t)_{t \in [0, T]}$  be a  $\mathbb{Y}$ -adapted, càdlàg,  $\mathcal{M}(\mathbb{R}^m)$ -valued solution of (3.8), with  $\zeta_{0-} = \xi \in \mathcal{P}(\mathbb{R}^m)$ . Define the process*

$$A_t := t + \sum_{i=1}^m |\nu^i|_t, \quad t \in [0, T]. \quad (4.2)$$

*If  $\nu$  is continuous and if, for any  $\varepsilon > 0$ ,  $\mathbb{E}[\sup_{t \in [0, T]} \|T_\varepsilon |\zeta|_t - \zeta_t\|_H^2] < +\infty$ , then there exists a constant  $M > 0$  such that, for each  $\varepsilon > 0$  and all  $\mathbb{F}$ -stopping times  $\tau \leq t$ ,  $t \in [0, T]$ ,*

$$\mathbb{E}[\|T_\varepsilon \zeta_\tau - \zeta_\tau\|_H^2] \leq \|T_\varepsilon \zeta_{0-}\|_H^2 + M \int_0^{\tau^-} \mathbb{E}[\|T_\varepsilon |\zeta|_s - \zeta_s\|_H^2] dA_s. \quad (4.3)$$

*Proof.* To ease notations, for any  $\varepsilon > 0$  denote by  $Z_t^\varepsilon$  the process  $Z_t^\varepsilon := T_\varepsilon \zeta_t$ ,  $t \geq 0$ . Fix  $\varepsilon > 0$  and consider an orthonormal basis  $\{\varphi_k\}_{k \in \mathbb{N}}$  of  $H$  such that  $\varphi_k \in C_b^2(\mathbb{R}^m)$ , for any  $k \in \mathbb{N}$ . Writing the Zakai equation for the function  $T_\varepsilon \varphi_k$  (recall that  $\nu$  is continuous by assumption) we get:

$$\zeta_t(T_\varepsilon \varphi_k) = \xi(T_\varepsilon \varphi_k) + \int_0^t \zeta_s(\mathcal{A}_s T_\varepsilon \varphi_k) ds + \int_0^t \zeta_{s-}(\mathbf{D}_x T_\varepsilon \varphi_k) d\nu_s + \int_0^t \gamma^{-1}(s) \zeta_s(T_\varepsilon \varphi_k h_s) d\bar{B}_s, \quad (4.4)$$

for all  $t \in [0, T]$ . Notice that, for any  $\varphi \in C_b^2(\mathbb{R}^m)$  and any  $t \in [0, T]$ , we can write:

$$\mathcal{A}_t \varphi(x) = \sum_{i=1}^m b_i(t, x) \partial_i \varphi(x) + \sum_{i,j=1}^m a_{ij}(t, x) \partial_{ij} \varphi(x), \quad x \in \mathbb{R}^m,$$

where  $a$  is the function defined in (4.1). For any  $i, j = 1, \dots, m$ ,  $\ell = 1, \dots, n$ , and  $t \in [0, T]$ , we define the random measures on  $\mathbb{R}^m$ :

$$b_t^i \zeta_t(dx) := b_i(t, x) \zeta_t(dx), \quad a_t^{ij} \zeta_t(dx) := a_{ij}(t, x) \zeta_t(dx), \quad \gamma h_t^\ell \zeta_t(dx) := \sum_{p=1}^n \gamma_{\ell p}^{-1}(t) h_p(t, x) \zeta_t(dx).$$

These measures are  $\mathbb{P}$ -almost surely finite, for any  $t \in [0, T]$ , thanks to Assumption 4.1 and to (A.3) (see also (A.19) for the last measure).

Applying Lemma 4.2 and the integration by parts formula we get:

$$\begin{aligned} \zeta_t(\mathcal{A}_t T_\varepsilon \varphi_k) &= \sum_{i=1}^m \int_{\mathbb{R}^m} b_i(t, x) \partial_i T_\varepsilon \varphi_k(x) \zeta_t(dx) + \sum_{i,j=1}^m \int_{\mathbb{R}^m} a_{ij}(t, x) \partial_{ij} T_\varepsilon \varphi_k(x) \zeta_t(dx) \\ &= \sum_{i=1}^m \int_{\mathbb{R}^m} b_i(t, x) T_\varepsilon \partial_i \varphi_k(x) \zeta_t(dx) + \sum_{i,j=1}^m \int_{\mathbb{R}^m} a_{ij}(t, x) T_\varepsilon \partial_{ij} \varphi_k(x) \zeta_t(dx) \\ &= \sum_{i=1}^m b_t^i \zeta_t(T_\varepsilon \partial_i \varphi_k) + \sum_{i,j=1}^m a_t^{ij} \zeta_t(T_\varepsilon \partial_{ij} \varphi_k) \\ &= \sum_{i=1}^m \langle T_\varepsilon(b_t^i \zeta_t), \partial_i \varphi_k \rangle + \sum_{i,j=1}^m \langle T_\varepsilon(a_t^{ij} \zeta_t), \partial_{ij} \varphi_k \rangle = \sum_{i,j=1}^m \langle \varphi_k, \partial_{ij} T_\varepsilon(a_t^{ij} \zeta_t) \rangle - \sum_{i=1}^m \langle \varphi_k, \partial_i T_\varepsilon(b_t^i \zeta_t) \rangle. \end{aligned}$$

In a similar way, we obtain  $\zeta_t(\partial_i T_\varepsilon \varphi_k) = -\langle \varphi_k, \partial_i T_\varepsilon \zeta_t \rangle$ , and

$$\sum_{j=1}^n \gamma_{ij}^{-1}(t) \zeta_t(T_\varepsilon \varphi_k h_t^j) = \langle \varphi_k, T_\varepsilon(\gamma h_t^i \zeta_t) \rangle, \quad i = 1, \dots, n.$$

Putting together all these facts, we can rewrite (4.4) as

$$\begin{aligned} \langle \varphi_k, Z_t^\varepsilon \rangle &= \langle \varphi_k, Z_{0-}^\varepsilon \rangle + \sum_{i,j=1}^m \int_0^t \langle \varphi_k, \partial_{ij} T_\varepsilon(a_s^{ij} \zeta_s) \rangle ds - \sum_{i=1}^m \int_0^t \langle \varphi_k, \partial_i T_\varepsilon(b_s^i \zeta_s) \rangle ds \\ &\quad - \sum_{i=1}^m \int_0^t \langle \varphi_k, \partial_i T_\varepsilon \zeta_{s-} \rangle d\nu_s^i + \sum_{i=1}^n \int_0^t \langle \varphi_k, T_\varepsilon(\gamma h_s^i \zeta_s) \rangle d\bar{B}_s^i, \quad \mathbb{P}\text{-a.s.}, \quad t \in [0, T]. \end{aligned}$$

Applying Itô's formula we get that, for all  $t \in [0, T]$ ,  $\mathbb{P}$ -a.s.,

$$\begin{aligned} \langle \varphi_k, Z_t^\varepsilon \rangle^2 &= \langle \varphi_k, Z_{0-}^\varepsilon \rangle^2 + \sum_{i,j=1}^m \int_0^t 2 \langle \varphi_k, Z_s^\varepsilon \rangle \langle \varphi_k, \partial_{ij} T_\varepsilon(a_s^{ij} \zeta_s) \rangle ds \\ &\quad - \sum_{i=1}^m \int_0^t 2 \langle \varphi_k, Z_s^\varepsilon \rangle \langle \varphi_k, \partial_i T_\varepsilon(b_s^i \zeta_s) \rangle ds + \sum_{i=1}^n \int_0^t \langle \varphi_k, T_\varepsilon(\gamma h_s^i \zeta_s) \rangle^2 ds \\ &\quad - \sum_{i=1}^m \int_0^t 2 \langle \varphi_k, Z_{s-}^\varepsilon \rangle \langle \varphi_k, \partial_i T_\varepsilon \zeta_{s-} \rangle d\nu_s^i + \sum_{i=1}^n \int_0^t 2 \langle \varphi_k, Z_s^\varepsilon \rangle \langle \varphi_k, T_\varepsilon(\gamma h_s^i \zeta_s) \rangle d\bar{B}_s^i. \end{aligned}$$

Using Assumption 4.1 and (A.3), it is possible to show that the stochastic integral with respect to Brownian motion  $\bar{B}$  is a  $\mathbb{P}$ -martingale. By the optional sampling theorem, this stochastic integral has zero expectation even when evaluated at any bounded stopping time. Therefore, picking an  $\mathbb{F}$ -stopping time  $\tau \leq t$ , for arbitrary  $t \in [0, T]$ , summing over  $k$  up to  $N \in \mathbb{N}$ , and taking the expectation, by Fatou's lemma we have that

$$\mathbb{E} [\|Z_{\tau-}^\varepsilon\|_H^2] = \mathbb{E} \left[ \lim_{N \rightarrow \infty} \sum_{k=1}^N \langle \varphi_k, Z_{\tau-}^\varepsilon \rangle^2 \right] \leq \liminf_{N \rightarrow \infty} \mathbb{E} \left[ \sum_{k=1}^N \langle \varphi_k, Z_{\tau-}^\varepsilon \rangle^2 \right] \leq \|Z_{0-}^\varepsilon\|_H^2$$



$$\begin{aligned}
 & + \liminf_{N \rightarrow \infty} \left\{ \sum_{i,j=1}^m \mathbb{E} \left[ \int_0^{\tau^-} \sum_{k=1}^N 2 \langle \varphi_k, Z_s^\varepsilon \rangle \langle \varphi_k, \partial_{ij} T_\varepsilon(a_s^{ij} \zeta_s) \rangle ds \right] \right. \\
 & - \sum_{i=1}^m \mathbb{E} \left[ \int_0^{\tau^-} \sum_{k=1}^N 2 \langle \varphi_k, Z_s^\varepsilon \rangle \langle \varphi_k, \partial_i T_\varepsilon(b_s^i \zeta_s) \rangle ds \right] + \sum_{i=1}^n \mathbb{E} \left[ \int_0^{\tau^-} \sum_{k=1}^N \langle \varphi_k, T_\varepsilon(\gamma h_s^i \zeta_s) \rangle^2 ds \right] \\
 & \left. - \sum_{i=1}^m \mathbb{E} \left[ \int_0^{\tau^-} \sum_{k=1}^N 2 \langle \varphi_k, Z_{s-}^\varepsilon \rangle \langle \varphi_k, \partial_i T_\varepsilon \zeta_{s-} \rangle d\nu_s^i \right] \right\}, \tag{4.5}
 \end{aligned}$$

where we used the fact that, since  $Z_{0-}^\varepsilon \in H$ ,  $\lim_{N \rightarrow \infty} \sum_{k=1}^N \langle \varphi_k, Z_{0-}^\varepsilon \rangle^2 = \|Z_{0-}^\varepsilon\|_H^2$ . More generally, since  $Z_t^\varepsilon \in H$ , for all  $t \in [0, T]$ ,  $\mathbb{P}$ -a.s. (cf. Remark 4.1), we have that

$$\sum_{k=1}^N \langle \varphi_k, Z_t^\varepsilon \rangle^2 \leq \sum_{k=1}^\infty \langle \varphi_k, Z_t^\varepsilon \rangle^2 = \|Z_t^\varepsilon\|_H^2, \quad t \in [0, T]. \tag{4.6}$$

We want now to estimate the quantities appearing inside the limit inferior, in order to exchange the limit and the integrals in (4.5). First of all, let us notice that, thanks to Assumption 4.1, the following estimates hold  $\mathbb{P}$ -a.s., for all  $i, j = 1, \dots, m$ , all  $\ell = 1, \dots, n$ , and all  $t \in [0, T]$ :

$$\begin{aligned}
 \|\partial_{ij} T_\varepsilon(a_t^{ij} \zeta_t)\|_H^2 &\leq K_1 \|T_\varepsilon|\zeta|_t\|_H^2, & \|\partial_i T_\varepsilon(b_t^i \zeta_t)\|_H^2 &\leq K_2 \|T_\varepsilon|\zeta|_t\|_H^2, \\
 \|T_\varepsilon(\gamma h_t^\ell \zeta_t)\|_H^2 &\leq K_3 \|T_\varepsilon|\zeta|_t\|_H^2, & \|\partial_i T_\varepsilon \zeta_t\|_H^2 &\leq K_4 \|T_\varepsilon|\zeta|_t\|_H^2,
 \end{aligned}$$

where  $K_1 = K_1(\varepsilon, m, \sigma)$ ,  $K_2 = K_2(\varepsilon, m, b)$ ,  $K_3 = K_3(n, h, \gamma)$ ,  $K_4 = K_4(\varepsilon, m)$ . They can be proved following a reasoning analogous to that of [2, Lemma 7.5] (see also [37, Chapter 6]).

Recalling that  $2|ab| \leq a^2 + b^2$ , for all  $a, b \in \mathbb{R}$ , using the estimates provided above, Lemma 4.2, and (4.6), we get that, for all  $N \in \mathbb{N}$ , all  $i, j = 1, \dots, m$ , and all  $s \in [0, T]$ ,

$$\begin{aligned}
 \mathbf{1}_{s < \tau} \sum_{k=1}^N 2 \langle \varphi_k, Z_s^\varepsilon \rangle \langle \varphi_k, \partial_{ij} T_\varepsilon(a_s^{ij} \zeta_s) \rangle &\leq \sum_{k=1}^N \langle \varphi_k, Z_s^\varepsilon \rangle^2 + \sum_{k=1}^N \langle \varphi_k, \partial_{ij} T_\varepsilon(a_s^{ij} \zeta_s) \rangle^2 \\
 &\leq \|Z_s^\varepsilon\|_H^2 + \|\partial_{ij} T_\varepsilon(a_s^{ij} \zeta_s)\|_H^2 \leq (1 + K_1) \|T_{\varepsilon/2}|\zeta|_s\|_H^2.
 \end{aligned}$$

With analogous computations, we get, for all  $i = 1, \dots, m$ , all  $N \in \mathbb{N}$ , and all  $s \in [0, T]$ ,

$$\begin{aligned}
 \mathbf{1}_{s < \tau} \sum_{k=1}^N 2 \langle \varphi_k, Z_s^\varepsilon \rangle \langle \varphi_k, \partial_i T_\varepsilon(b_s^i \zeta_s) \rangle &\leq (1 + K_2) \|T_{\varepsilon/2}|\zeta|_s\|_H^2, \\
 \mathbf{1}_{s < \tau} \sum_{k=1}^N 2 \langle \varphi_k, Z_{s-}^\varepsilon \rangle \langle \varphi_k, \partial_i T_\varepsilon \zeta_{s-} \rangle &\leq (1 + K_4) \|T_{\varepsilon/2}|\zeta|_s\|_H^2,
 \end{aligned}$$

and, for all  $N \in \mathbb{N}$  and all  $s \in [0, T]$ ,

$$\sum_{i=1}^n \mathbf{1}_{s < \tau} \sum_{k=1}^N \langle \varphi_k, T_\varepsilon(\gamma h_s^i \zeta_s) \rangle^2 \leq n K_3 \|T_\varepsilon|\zeta|_s\|_H^2.$$

The terms appearing on the r.h.s. of these estimates are  $dt \otimes d\mathbb{P}$ - and  $d|\nu^i|_t \otimes d\mathbb{P}$ -integrable on  $[0, T] \times \Omega$ , for all  $i = 1, \dots, m$ , since, for any  $\varepsilon > 0$ ,

$$\begin{aligned}
 \mathbb{E} \left[ \int_0^T \|T_\varepsilon|\zeta|_s\|_H^2 ds \right] &\leq T \mathbb{E} \left[ \sup_{s \in [0, T]} \|T_\varepsilon|\zeta|_s\|_H^2 \right] < +\infty, \\
 \mathbb{E} \left[ \int_0^T \|T_\varepsilon|\zeta|_s\|_H^2 d|\nu^i|_s \right] &\leq K \mathbb{E} \left[ \sup_{s \in [0, T]} \|T_\varepsilon|\zeta|_s\|_H^2 \right] < +\infty.
 \end{aligned}$$

Therefore, by the dominated convergence theorem, we can pass to the limit in (4.5), as  $N \rightarrow \infty$ ,

$$\begin{aligned}
 \mathbb{E} [\|Z_{\tau-}^\varepsilon\|_H^2] &\leq \|Z_{0-}^\varepsilon\|_H^2 + \sum_{i,j=1}^m \mathbb{E} \left[ \int_0^{\tau^-} 2 \langle Z_s^\varepsilon, \partial_{ij} T_\varepsilon(a_s^{ij} \zeta_s) \rangle ds \right] - \sum_{i=1}^m \mathbb{E} \left[ \int_0^{\tau^-} 2 \langle Z_s^\varepsilon, \partial_i T_\varepsilon(b_s^i \zeta_s) \rangle ds \right] \\
 &+ \sum_{i=1}^n \mathbb{E} \left[ \int_0^{\tau^-} \|T_\varepsilon(\gamma h_s^i \zeta_s)\|_H^2 ds \right] - \sum_{i=1}^m \mathbb{E} \left[ \int_0^{\tau^-} \langle Z_{s-}^\varepsilon, \partial_i T_\varepsilon \zeta_{s-} \rangle d\nu_s^i \right], \tag{4.7}
 \end{aligned}$$

We finally get the claim, bounding the terms on the r.h.s. of (4.7) by using the following results: for the second one, apply [37, Lemma 6.11]; for the third and the last one, apply [37, Lemma 6.10]; for the fourth one, use the fact that the constant  $K_3$  above does not depend on  $\varepsilon$ .  $\square$

Proposition 4.4 allows to deduce that any  $\mathcal{M}_+(\mathbb{R}^m)$ -valued solution of the Zakai equation (3.8) admits a density with respect to Lebesgue measure.

**Proposition 4.5.** *Suppose that Assumption 4.1 holds. Let  $\zeta = (\zeta_t)_{t \in [0, T]}$  be a  $\mathbb{Y}$ -adapted, càdlàg,  $\mathcal{M}_+(\mathbb{R}^m)$ -valued solution of (3.8), with  $\zeta_{0-} = \xi \in \mathcal{P}(\mathbb{R}^m)$ . If  $\nu$  is continuous and if  $\xi$  admits a square-integrable density with respect to Lebesgue measure on  $\mathbb{R}^m$ , then there exists an  $H$ -valued process  $Z = (Z_t)_{t \in [0, T]}$  such that, for all  $t \in [0, T]$ ,*

$$\zeta_t(dx) = Z_t(x)dx, \quad \mathbb{P}\text{-a.s.}$$

Moreover,  $Z$  is  $\mathbb{Y}$ -adapted, continuous, and satisfies  $\mathbb{E}[\|Z_t\|_H^2] < +\infty$ , for all  $t \in [0, T]$ .

*Proof.* As a consequence of Lemma 4.3, the assumptions of Proposition 4.4 hold and we have that for each  $\varepsilon > 0$  and all  $\mathbb{F}$ -stopping times  $\tau \leq t$ ,  $t \in [0, T]$ ,

$$\mathbb{E}[\|T_\varepsilon \zeta_{\tau-}\|_H^2] \leq \|T_\varepsilon \zeta_{0-}\|_H^2 + M \int_0^{\tau-} \mathbb{E}[\|T_\varepsilon \zeta_{s-}\|_H^2] dA_s.$$

Therefore, we can apply Lemma A.1 and get that, for all  $t \in [0, T]$ ,

$$\mathbb{E}[\|T_\varepsilon \zeta_{t-}\|_H^2] = \mathbb{E}[\|T_\varepsilon \zeta_t\|_H^2] \leq \|T_\varepsilon \zeta_{0-}\|_H^2 e^{M(T+mK)}, \quad (4.8)$$

where we used the fact that  $\zeta$  is continuous, since  $\nu$  is, and that  $A_t \leq A_T \leq T + mK$ , for all  $t \in [0, T]$ . Notice that, denoting by  $Z_{0-}$  the density of  $\xi$  with respect to Lebesgue measure on  $\mathbb{R}^m$ ,

$$T_\varepsilon \zeta_{0-}(y) = \int_{\mathbb{R}^m} \psi_\varepsilon(x - y) \xi(dx) = \int_{\mathbb{R}^m} \psi_\varepsilon(x - y) Z_{0-}(x) dx = T_\varepsilon Z_{0-}(y), \quad y \in \mathbb{R}^m.$$

By point ii. of Lemma 4.2 and since the constants appearing in (4.8) do not depend on  $\varepsilon$ , we get

$$\sup_{\varepsilon > 0} \mathbb{E}[\|T_\varepsilon \zeta_t\|_H^2] \leq \|Z_{0-}\|_H^2 e^{M(T+mK)}, \quad t \in [0, T].$$

Taking, as in the Proof of Proposition 4.4, an orthonormal basis  $\{\varphi_k\}_{k \in \mathbb{N}}$  of  $H$  such that  $\varphi_k \in C_b^2(\mathbb{R}^m)$ , for any  $k \in \mathbb{N}$ , the dominated convergence theorem entails that, for all  $k \in \mathbb{N}$ ,

$$\lim_{\varepsilon \rightarrow 0} \langle T_\varepsilon \zeta_t, \varphi_k \rangle = \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^m} \left\{ \int_{\mathbb{R}^m} \psi_\varepsilon(x - y) \varphi_k(y) dy \right\} \zeta_t(dx) = \int_{\mathbb{R}^m} \varphi_k(x) \zeta_t(dx) = \zeta_t(\varphi_k).$$

Applying Fatou's Lemma we get that, for all  $t \in [0, T]$ ,

$$\begin{aligned} \mathbb{E} \left[ \sum_{k=1}^{\infty} \zeta_t(\varphi_k)^2 \right] &= \mathbb{E} \left[ \sum_{k=1}^{\infty} \lim_{\varepsilon \rightarrow 0} \langle T_\varepsilon \zeta_t, \varphi_k \rangle^2 \right] \leq \liminf_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \sum_{k=1}^{\infty} \langle T_\varepsilon \zeta_t, \varphi_k \rangle^2 \right] \\ &\leq \sup_{\varepsilon > 0} \mathbb{E}[\|T_\varepsilon \zeta_t\|_H^2] \leq \|Z_{0-}\|_H^2 e^{M(T+mK)} < +\infty, \end{aligned} \quad (4.9)$$

and hence, from Lemma 4.1 we deduce that,  $\mathbb{P}$ -a.s.,  $\zeta_t$  is absolutely continuous with respect to Lebesgue measure on  $\mathbb{R}^m$ , for all  $t \in [0, T]$ . Moreover, its density process  $Z = (Z_t)_{t \in [0, T]}$  takes values in  $H$  and, by standard results, is  $\mathbb{Y}$ -adapted and continuous (because  $\nu$  is).

Finally, since  $\zeta_t(\varphi_k) = \int_{\mathbb{R}^m} \varphi_k(x) Z_t(x) dx = \langle \varphi_k, Z_t \rangle$ , for all  $k \in \mathbb{N}$ , and all  $t \in [0, T]$ , we get

$$\mathbb{E}[\|Z_t\|_H^2] = \mathbb{E} \left[ \sum_{k=1}^{\infty} \langle \varphi_k, Z_t \rangle^2 \right] = \mathbb{E} \left[ \sum_{k=1}^{\infty} \zeta_t(\varphi_k)^2 \right] < +\infty, \quad t \in [0, T]. \quad \square$$

We are now ready to state our first uniqueness result for the solution to the Zakai equation, in the case where  $\nu$  is continuous.

**Theorem 4.6.** *Suppose that Assumptions 2.1, 2.2, 4.1, and (3.7) hold. If  $\nu$  is continuous and if  $\xi \in \mathcal{P}(\mathbb{R}^m)$  admits a square-integrable density with respect to Lebesgue measure on  $\mathbb{R}^m$ , then the unnormalized filtering process  $\rho$ , defined in (3.3), is the unique  $\mathbb{Y}$ -adapted, continuous,  $\mathcal{M}_+(\mathbb{R}^m)$ -valued solution to the Zakai equation (3.8).*

Moreover, there exists a  $\mathbb{Y}$ -adapted, continuous,  $H$ -valued process  $p = (p_t)_{t \in [0, T]}$  satisfying, for all  $t \in [0, T]$ ,  $\mathbb{E}[\|p_t\|_H^2] < +\infty$  and  $\rho_t(dx) = p_t(x)dx$ ,  $\mathbb{P}$ -a.s.

*Proof.* Clearly, the unnormalized filtering process  $\rho$ , defined in (3.3), is a  $\mathbb{Y}$ -adapted, continuous (since  $\nu$  is),  $\mathcal{M}_+(\mathbb{R}^m)$ -valued solution to (3.8). Therefore, the second part of the statement follows directly from Proposition 4.5.

Uniqueness can be established as follows. Let  $\zeta^{(1)}, \zeta^{(2)}$  be two  $\mathbb{Y}$ -adapted, càdlàg,  $\mathcal{M}_+(\mathbb{R}^m)$ -valued solutions to (3.8). Define  $\zeta := \zeta^{(1)} - \zeta^{(2)} \in \mathcal{M}(\mathbb{R}^m)$  and let  $Z := Z^{(1)} - Z^{(2)} \in H$  be its density process, where  $Z^{(1)}$  and  $Z^{(2)}$  are the density processes of  $\zeta^{(1)}$  and  $\zeta^{(2)}$ , respectively, which exist thanks to Proposition 4.5.

Standard facts from measure theory show that, for all non-negative, bounded, measurable functions  $\varphi: \mathbb{R}^m \rightarrow \mathbb{R}$  and all  $t \in [0, T]$ ,  $|\zeta|_t(\varphi) \leq \zeta_t^{(1)}(\varphi) + \zeta_t^{(2)}(\varphi)$ . From this fact, applying Lemma 4.3 we deduce that

$$\mathbb{E}[\sup_{t \in [0, T]} \|T_\varepsilon |\zeta|_{t-}\|_H^2] \leq 2\mathbb{E}[\sup_{t \in [0, T]} \|T_\varepsilon \zeta_{t-}^{(1)}\|_H^2] + 2\mathbb{E}[\sup_{t \in [0, T]} \|T_\varepsilon \zeta_{t-}^{(2)}\|_H^2] < +\infty.$$

Therefore, from Proposition 4.4 we get that for all  $\varepsilon > 0$  and all  $\mathbb{F}$ -stopping times  $\tau \leq t$ ,  $t \in [0, T]$ ,

$$\mathbb{E}[\|T_\varepsilon \zeta_{\tau-}\|_H^2] \leq M \int_0^{\tau-} \mathbb{E}[\|T_\varepsilon |\zeta|_{s-}\|_H^2] dA_s,$$

where  $A$  is defined in (4.2). An application of the dominated convergence theorem shows that  $\|T_\varepsilon |\zeta|_{t-}\|_H^2 \rightarrow \|Z_t\|_H^2$ , as  $\varepsilon \rightarrow 0$ , for all  $t \in [0, T]$ , and hence, by Fatou's lemma

$$\begin{aligned} \mathbb{E}[\|Z_t\|_H^2] &= \mathbb{E}[\lim_{\varepsilon \rightarrow 0} \|T_\varepsilon \zeta_{\tau-}\|_H^2] \leq \liminf_{\varepsilon \rightarrow 0} \mathbb{E}[\|T_\varepsilon \zeta_{\tau-}\|_H^2] \\ &\leq \liminf_{\varepsilon \rightarrow 0} M \int_0^{\tau-} \mathbb{E}[\|T_\varepsilon |\zeta|_{s-}\|_H^2] dA_s = M \int_0^{\tau-} \mathbb{E}[\|Z_{s-}\|_H^2] dA_s. \end{aligned}$$

Finally, Proposition 4.5 ensures that

$$\mathbb{E}[\|Z_t\|_H^2] \leq 2\mathbb{E}[\|Z_t^{(1)}\|_H^2] + 2\mathbb{E}[\|Z_t^{(2)}\|_H^2] < +\infty, \quad \text{for all } t \in [0, T],$$

This allows us to use Lemma A.1 to get that, for all  $t \in [0, T]$ ,  $\mathbb{E}[\|Z_t\|_H^2] = \mathbb{E}[\|Z_t\|_H^2] = 0$ , whence we obtain  $\|Z_t\|_H^2 = 0$ ,  $\mathbb{P}$ -a.s., and therefore uniqueness of the solution to the Zakai equation.  $\square$

**4.2. The case in which the jump times of  $\nu$  do not accumulate.** Exploiting the recursive structure of (3.19), we can prove uniqueness of the solution to the Zakai equation (3.8), also in the case where the jump times of  $\nu$  do not accumulate.

**Theorem 4.7.** *Suppose that Assumptions 2.1, 2.2, 4.1, and (3.7) hold. If the jump times of  $\nu$  do not accumulate over  $[0, T]$  and if  $\xi \in \mathcal{P}(\mathbb{R}^m)$  admits a square-integrable density with respect to Lebesgue measure on  $\mathbb{R}^m$ , then the unnormalized filtering process  $\rho$ , defined in (3.3), is the unique  $\mathbb{Y}$ -adapted, càdlàg,  $\mathcal{M}_+(\mathbb{R}^m)$ -valued solution to the Zakai equation (3.8).*

*Moreover, there exists a  $\mathbb{Y}$ -adapted, càdlàg,  $H$ -valued process  $p = (p_t)_{t \in [0, T]}$  satisfying, for all  $t \in [0, T]$ ,  $\mathbb{E}[\|p_t\|_H^2] < +\infty$  and  $\rho_t(dx) = p_t(x)dx$ ,  $\mathbb{P}$ -a.s.*

*Proof.* Let us denote by  $\rho$  the unnormalized filtering process associated with the initial law  $\xi$  and process  $\nu$ , and by  $p_{0-}$  the density of  $\xi$  with respect to Lebesgue measure on  $\mathbb{R}^m$ . Let  $T_0 = 0$  and define the sequence of jump times of  $\nu$

$$T_n := \inf\{t > T_{n-1} : \Delta\nu_t \neq 0\}, \quad n \in \mathbb{N},$$

with the usual convention  $\inf \emptyset = +\infty$ . Recall that also  $T_0$  can be a jump time of  $\nu$ . Moreover, since the jump times of  $\nu$  do not accumulate over  $[0, T]$ , we have that  $T_n \leq T_{n+1}$ ,  $\mathbb{P}$ -a.s., and  $T_n < +\infty \implies T_n < T_{n+1}$ , for all  $n \in \mathbb{N}_0$ .

We start noticing that the formula  $\rho_{T_n}(\varphi) = \rho_{T_n-}(\varphi_{T_n}(\cdot + \Delta\nu_{T_n}))$ ,  $n \in \mathbb{N}_0$ , appearing in (3.19) holds for all  $\varphi \in C_b(\mathbb{R}^m)$ . Indeed, continuity of the observation filtration  $\mathbb{Y}$  implies that

$$\tilde{\mathbb{E}}[\varphi(X_{T_n-}) | \mathcal{Y}_{T_n}] = \tilde{\mathbb{E}}[\varphi(X_{T_n-}) | \mathcal{Y}_{T_n-}] = \pi_{T_n-}(\varphi).$$

Using continuity of process  $\eta$ , Kallianpur-Striebel formula (3.2), and the freezing lemma, we get

$$\begin{aligned} \rho_{T_n}(\varphi) &= \tilde{\mathbb{E}}[\varphi(X_{T_n}) | \mathcal{Y}_{T_n}] \mathbb{E}[\eta_{T_n} | \mathcal{Y}] = \tilde{\mathbb{E}}[\varphi(X_{T_n-} + \Delta\nu_{T_n}) | \mathcal{Y}_{T_n}] \mathbb{E}[\eta_{T_n-} | \mathcal{Y}] \\ &= \pi_{T_n-}(\varphi(\cdot + \Delta\nu_{T_n})) \mathbb{E}[\eta_{T_n-} | \mathcal{Y}] = \rho_{T_n-}(\varphi(\cdot + \Delta\nu_{T_n})), \end{aligned}$$

for all  $\varphi \in C_b(\mathbb{R}^m)$  and all  $n \in \mathbb{N}_0$ . This, in turn, entails that if  $\rho_{T_n-}$  admits a density  $p_{T_n-}$  with respect to Lebesgue measure, then

$$\int_{\mathbb{R}^m} \varphi(x) \rho_{T_n}(dx) = \rho_{T_n-}(\varphi(\cdot + \Delta\nu_{T_n})) = \int_{\mathbb{R}^m} \varphi(x + \Delta\nu_{T_n}) p_{T_n-}(x) dx = \int_{\mathbb{R}^m} \varphi(x) p_{T_n-}(x - \Delta\nu_{T_n}) dx.$$

Therefore, since  $C_b(\mathbb{R}^m)$  is a separating set (see, e.g., [20, Chapter 3, Section 4]), we have the equivalence of measures  $\rho_{T_n}(dx)$  and  $p_{T_n^-}(x - \Delta\nu_{T_n})dx$ , implying that  $\rho_{T_n}$  admits density with respect to Lebesgue measure on  $\mathbb{R}^m$ , given by  $p_{T_n^-}(\cdot - \Delta\nu_{T_n})$ .

We can now use the recursive structure of (3.19) to get the claim. Define the process

$$\nu_t^{(1)} := \nu_t \mathbf{1}_{t < T_1} + \nu_{T_1} \mathbf{1}_{t \geq T_1}, \quad t \in [0, T],$$

and the random measure  $\xi^{(1)}(dx) := p_{0^-}(x - \Delta\nu_0)dx$ , on  $\mathbb{R}^m$ . Consider, for all  $\varphi \in C_b^2(\mathbb{R}^m)$ , the Zakai equation

$$\begin{aligned} \rho_t^{(1)}(\varphi) &= \xi^{(1)}(\varphi) + \int_0^t \rho_s^{(1)}([\partial_s + \mathcal{A}_s]\varphi) ds \\ &\quad + \int_0^t \rho_{s^-}^{(1)}(D_x \varphi) d\nu_s^{(1)} + \int_0^t \gamma^{-1}(s) \rho_s^{(1)}(\varphi h_s) d\bar{B}_s, \quad \mathbb{P}\text{-a.s.}, \quad t \in [0, T]. \end{aligned} \quad (4.10)$$

Since  $\nu^{(1)}$  satisfies point (iv) of Assumption 2.1, we have that (4.10) is the Zakai equation for the filtering problem of the partially observed system (2.8)–(2.13), with initial law  $\xi^{(1)}$  and process  $\nu^{(1)}$ , which is continuous on  $[0, T]$ . Therefore, by Theorem 4.6,  $\rho^{(1)}$  is its unique solution and admits a density  $p^{(1)}$  with respect to Lebesgue measure on  $\mathbb{R}^m$ , with  $\mathbb{E}[\|p_t^{(1)}\|_H^2] < +\infty$ , for each  $t \in [0, T]$ . It is clear that, since  $\nu_t = \nu_t^{(1)}$  on  $\{t < T_1\}$ , we have that  $\rho_t = \rho_t^{(1)}$  on the same set, and hence  $\rho_t$  admits density  $p_t^{(1)}$  on  $\{t < T_1\}$ .

Next, let us define the process

$$\nu_t^{(2)} := \nu_{t+T_1} \mathbf{1}_{t < T_2-T_1} + \nu_{T_2} \mathbf{1}_{t \geq T_2-T_1}, \quad t \in [0, T],$$

and the random measure  $\xi^{(2)}(dx) = p_{T_1^-}(x - \Delta\nu_{T_1})dx$ , on  $\mathbb{R}^m$ . Consider, for all  $\varphi \in C_b^2(\mathbb{R}^m)$ , the Zakai equation

$$\begin{aligned} \rho_t^{(2)}(\varphi) &= \xi^{(2)}(\varphi) + \int_0^t \rho_s^{(2)}([\partial_s + \mathcal{A}_{s+T_1}]\varphi) ds \\ &\quad + \int_0^t \rho_{s^-}^{(2)}(D_x \varphi) d\nu_s^{(2)} + \int_0^t \gamma^{-1}(s+T_1) \rho_s^{(2)}(\varphi h_{s+T_1}) d\bar{B}_{s+T_1}, \quad \mathbb{P}\text{-a.s.}, \quad t \in [0, T]. \end{aligned} \quad (4.11)$$

Since  $\nu^{(2)}$  satisfies point (iv) of Assumption 2.1, we have that (4.11) is the Zakai equation for the filtering problem of the partially observed system (2.8)–(2.13), with initial law  $\xi^{(2)}$  and process  $\nu^{(2)}$ , which is continuous on  $[0, T]$ . Therefore, by Theorem 4.6,  $\rho^{(2)}$  is its unique solution and admits a density  $p^{(2)}$  with respect to Lebesgue measure on  $\mathbb{R}^m$ , with  $\mathbb{E}[\|p_t^{(2)}\|_H^2] < +\infty$ , for each  $t \in [0, T]$ . It is clear that, since  $\nu_t = \nu_{t-T_1}^{(2)}$  on  $\{T_1 \leq t < T_2\}$ , we have that  $\rho_t = \rho_{t-T_1}^{(2)}$  on the same set, and hence  $\rho_t$  admits density  $p_{t-T_1}^{(2)}$  on  $\{T_1 \leq t < T_2\}$ .

Continuing in this manner, we construct a sequence of solutions  $(\rho^{(n)})_{n \in \mathbb{N}}$  and corresponding density processes  $(p^{(n)})_{n \in \mathbb{N}}$ . We deduce that the unnormalized filtering process is represented by

$$\rho_t = \sum_{n=1}^{\infty} \rho_{t-T_n}^{(n)} \mathbf{1}_{T_{n-1} \leq t < T_n}, \quad t \in [0, T],$$

and hence is the unique  $\mathbb{Y}$ -adapted, càdlàg,  $\mathcal{M}_+(\mathbb{R}^m)$ -valued solution to the Zakai equation (3.8), admitting a  $\mathbb{Y}$ -adapted, càdlàg,  $H$ -valued density process  $p$ , given by

$$p_t = \sum_{n=1}^{\infty} p_{t-T_n}^{(n)} \mathbf{1}_{T_{n-1} \leq t < T_n}, \quad t \in [0, T].$$

The fact that  $\mathbb{E}[\|p_t\|_H^2] < +\infty$ , for all  $t \in [0, T]$ , follows from the analogous property for each of the processes  $p^{(n)}$ ,  $n \in \mathbb{N}$ .  $\square$

## APPENDIX A. TECHNICAL RESULTS

Let us recall that if  $A$  (defined on a given filtered complete probability space) is a càdlàg, adapted, non-negative process, with  $A_{0^-} = 0$ , and  $H$  is an optional process, satisfying  $\int_0^t |H_s| dA_s < +\infty$ , for all  $t \geq 0$ ,  $\mathbb{P}$ -a.s., then for any stopping time  $\tau$  we have that

$$\int_0^{\tau^-} H_s dA_s := \int_0^{+\infty} H_s \mathbf{1}_{s < \tau} dA_s.$$

**Lemma A.1.** *Let  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  be a given filtered complete probability space, fix  $T > 0$ , and let  $A$  and  $H$  be two càdlàg,  $\mathbb{F}$ -adapted real-valued processes. Suppose that  $A$  is non-decreasing, with  $A_{0-} = 0$  and  $A_T \leq K$ ,  $\mathbb{P}$ -a.s., for some constant  $K > 0$ , and that  $H$  satisfies one of the following:*

- a.  $\mathbb{E}[\sup_{t \in [0, T]} |H_{t-}|] < +\infty$ ;*
- b.  $H$  is non-negative and such that  $\mathbb{E}[H_{t-}] < +\infty$ , for all  $t \in [0, T]$ .*

*Assume, moreover, that for any  $\mathbb{F}$ -stopping time  $\tau \leq T$  we have*

$$\mathbb{E}[H_{\tau-}] \leq M + \mathbb{E} \left[ \int_0^{\tau-} H_{s-} dA_s \right], \quad (\text{A.1})$$

*for some constant  $M$ . Then  $\mathbb{E}[H_{T-}] \leq Me^K$ .*

*Proof.* The following reasoning is inspired by the proof of [25, Lemma IX.6.3]. Let us define

$$\tilde{A}_t := A_t \mathbf{1}_{t < T} + K \mathbf{1}_{t \geq T}, \quad t \geq 0.$$

$\tilde{A}$  is still a càdlàg,  $\mathbb{F}$ -adapted and non-decreasing process, with  $\tilde{A}_{0-} = 0$ . Moreover, for any stopping time  $\tau \leq T$ , random measures  $\mathbf{1}_{s < \tau} dA_s$  and  $\mathbf{1}_{s < \tau} d\tilde{A}_s$  agree, therefore (A.1) implies

$$\mathbb{E}[H_{\tau-}] \leq M + \mathbb{E} \left[ \int_0^{\tau-} H_{s-} d\tilde{A}_s \right]. \quad (\text{A.2})$$

Next, define  $C_t := \inf\{s \geq 0 : \tilde{A}_s \geq t\}$ ,  $t \geq 0$ , which (see, e.g., [18, Chapter VI, Def. 56] or [25, Proposition I.1.28]) is an  $\mathbb{F}$ -stopping time for all  $t \geq 0$ , satisfying  $C_t \leq T$ , thanks to the definition of  $\tilde{A}$ .

We now fix  $t \in [0, K]$ . Using (A.2), we get

$$\mathbb{E}[H_{(C_t)-}] \leq M + \mathbb{E} \left[ \int_0^{+\infty} H_{s-} \mathbf{1}_{s < C_t} d\tilde{A}_s \right] = M + \mathbb{E} \left[ \int_0^{+\infty} H_{(C_u)-} \mathbf{1}_{C_u < C_t} du \right].$$

Since  $C$  is a non-decreasing process, we have that  $\{C_u < C_t\} \subset \{u < t\}$ , and hence  $\mathbf{1}_{C_u < C_t} \leq \mathbf{1}_{u < t}$ . Therefore

$$\mathbb{E}[H_{(C_t)-}] \leq M + \mathbb{E} \left[ \int_0^t H_{(C_u)-} du \right].$$

If  $H$  satisfies condition b. we can directly apply Fubini-Tonelli's theorem as below. If, instead, condition a. holds, since  $C_u \leq T$  and, for each fixed  $\omega \in \Omega$ , the image of the map  $u \mapsto C_u(\omega)$  is a subset of  $[0, T]$ , we have that  $\sup_{u \in [0, K]} |H_{(C_u)-}| \leq \sup_{s \in [0, T]} |H_{s-}|$ , so

$$\mathbb{E} \left[ \int_0^t |H_{(C_u)-}| du \right] \leq K \mathbb{E} \left[ \sup_{s \in [0, T]} |H_{s-}| \right] < +\infty.$$

Therefore, we can apply Fubini-Tonelli's theorem and get

$$\mathbb{E}[H_{(C_t)-}] \leq M + \int_0^t \mathbb{E}[H_{(C_u)-}] du,$$

whence we obtain, from the usual Gronwall's lemma,  $\mathbb{E}[H_{(C_t)-}] \leq Me^t$ . Thanks to the definition of  $\tilde{A}$ , we have that  $C_K = T$  and the claim follows letting  $t = K$  in the last inequality.  $\square$

**Proposition A.2.** *Under Assumptions 2.1 and 2.2, the process  $\eta$ , defined in (2.11), is a  $(\mathbb{P}, \mathbb{F})$ -martingale.*

*Proof.* Let us notice, first, a fact that will be useful in this proof. It can be easily shown that condition (2.1) implies, for some constant  $C_\gamma$ ,

$$\|\gamma^{-1}(t)\| \leq C_\gamma, \quad \forall t \in [0, T]. \quad (\text{A.3})$$

Let us define, for all  $t \in [0, T]$ ,

$$Z_t := \int_0^t \gamma^{-1}(s) h(s, X_s) d\bar{B}_s.$$

Thanks to condition (2.10) and using (A.3) and (2.9), we easily get

$$\begin{aligned} \mathbb{E} \left[ \int_0^T \|\gamma^{-1}(s) h(s, X_s)\|^2 ds \right] &\leq n \mathbb{E} \left[ \int_0^T \|\gamma^{-1}(s)\|^2 \|h(s, X_s)\|^2 ds \right] \\ &\leq n C_h C_\gamma \mathbb{E} \left[ \int_0^T (1 + \|X_s\|^2) ds \right] \leq n C_h C_\gamma T [1 + \kappa(1 + \mathbb{E}[\|X_{0-}\|^2])] < +\infty. \end{aligned} \quad (\text{A.4})$$

Therefore,  $Z$  is an  $(\mathbb{F}, \mathbb{P})$ -martingale, and hence  $\eta$ , which is the Doléans-Dade exponential of  $Z$ , is a non-negative local  $(\mathbb{F}, \mathbb{P})$ -martingale (see, e.g., [13, Lemma 15.3.2]). Thus, to prove the claim it is enough to show that  $\mathbb{E}[\eta_t] = 1$  for all  $t \in [0, T]$ .

We start proving, first, that  $\mathbb{E}[\eta_t \|X_t\|^2] \leq C$ , for all  $t \in [0, T]$ , where  $C$  is an appropriately chosen constant. For the sake of brevity, let us write  $b_s := b(s, X_s)$ ,  $\sigma_s := \sigma(s, X_s)$ , and  $h_s := h(s, X_s)$ . Applying Itô's formula we get

$$\begin{aligned} \|X_t\|^2 &= \|X_{0-}\|^2 + \int_0^t [2X_{s-}^* b_s + \|\sigma_s\|^2] ds + 2 \int_0^t X_{s-}^* \sigma_s dW_s \\ &\quad + 2 \int_0^t X_{s-} d\nu_s + \sum_{0 \leq s \leq t} \{\|X_s\|^2 - \|X_{s-}\|^2 - 2X_{s-} \cdot \Delta\nu_s\}, \end{aligned}$$

and using the integration by parts rule we have

$$\begin{aligned} \eta_t \|X_t\|^2 &= \|X_{0-}\|^2 + \int_0^t [2\eta_{s-} X_{s-}^* b_s + \eta_{s-} \|\sigma_s\|^2] ds + 2 \int_0^t \eta_{s-} X_{s-}^* \sigma_s dW_s \\ &\quad + \int_0^t \|X_{s-}\|^2 \eta_s \gamma^{-1}(s) h_s d\bar{B}_s + 2 \int_0^t \eta_{s-} X_{s-} d\nu_s \\ &\quad + \sum_{0 \leq s \leq t} \eta_{s-} \{\|X_s\|^2 - \|X_{s-}\|^2 - 2X_{s-} \cdot \Delta\nu_s\}. \end{aligned}$$

Therefore, for any fixed  $\varepsilon > 0$ , we obtain

$$\begin{aligned} \frac{\eta_t \|X_t\|^2}{1 + \varepsilon \eta_t \|X_t\|^2} &= \frac{\|X_{0-}\|^2}{1 + \varepsilon \|X_{0-}\|^2} + \int_0^t \frac{\eta_{s-}}{[1 + \varepsilon \eta_{s-} \|X_{s-}\|^2]^2} [2X_{s-}^* b_s + \|\sigma_s\|^2] ds \\ &\quad - \int_0^t \frac{\varepsilon \eta_{s-}^2}{[1 + \varepsilon \eta_{s-} \|X_{s-}\|^2]^3} [4\|X_{s-}^* \sigma_s\|^2 + \|X_{s-}\|^4 \|\gamma^{-1}(s) h_s\|^2] ds \\ &\quad + \int_0^t \frac{2\eta_{s-}}{[1 + \varepsilon \eta_{s-} \|X_{s-}\|^2]^2} X_{s-} d\nu_s^c + \int_0^t \frac{2\eta_{s-}}{[1 + \varepsilon \eta_{s-} \|X_{s-}\|^2]^2} X_{s-}^* \sigma_s dW_s \\ &\quad + \int_0^t \frac{\eta_{s-} \|X_{s-}\|^2}{[1 + \varepsilon \eta_{s-} \|X_{s-}\|^2]^2} \gamma^{-1}(s) h_s d\bar{B}_s + \sum_{0 \leq s \leq t} \left\{ \frac{\eta_s \|X_s\|^2}{1 + \varepsilon \eta_s \|X_s\|^2} - \frac{\eta_{s-} \|X_{s-}\|^2}{1 + \varepsilon \eta_{s-} \|X_{s-}\|^2} \right\}, \end{aligned} \quad (\text{A.5})$$

where  $\nu^c$  denotes the continuous part of the process  $\nu$ .

With standard estimates (see, e.g., [2, Solution to Exercise 3.11]) it is possible to show that the stochastic integrals with respect to Brownian motions  $W$  and  $\bar{B}$  are  $(\mathbb{F}, \mathbb{P})$ -martingales. This implies, thanks to the optional sampling theorem, that these stochastic integrals have zero expectation even when evaluated at any bounded stopping time. Fixing a  $\mathbb{F}$ -stopping time  $\tau \leq t$ , for arbitrary  $t \in [0, T]$ , taking the expectation and noticing that the third term in (A.5) is non-negative, we get

$$\begin{aligned} \mathbb{E} \left[ \frac{\eta_{\tau-} \|X_{\tau-}\|^2}{1 + \varepsilon \eta_{\tau-} \|X_{\tau-}\|^2} \right] &\leq \mathbb{E} \left[ \frac{\|X_{0-}\|^2}{1 + \varepsilon \|X_{0-}\|^2} \right] \\ &\quad + \mathbb{E} \left[ \int_0^{\tau-} \frac{\eta_{s-} [2X_{s-}^* b_s + \|\sigma_s\|^2]}{[1 + \varepsilon \eta_{s-} \|X_{s-}\|^2]^2} ds \right] + \mathbb{E} \left[ \int_0^{\tau-} \frac{2\eta_{s-} X_{s-}}{[1 + \varepsilon \eta_{s-} \|X_{s-}\|^2]^2} d\nu_s^c \right] \\ &\quad + \mathbb{E} \left[ \sum_{0 \leq s < \tau} \left\{ \frac{\eta_s \|X_s\|^2}{1 + \varepsilon \eta_s \|X_s\|^2} - \frac{\eta_{s-} \|X_{s-}\|^2}{1 + \varepsilon \eta_{s-} \|X_{s-}\|^2} \right\} \right]. \end{aligned} \quad (\text{A.6})$$

We proceed, now, to find suitable estimates for the terms appearing in (A.6).

Notice that, thanks to conditions (2.4) and (2.5), we have that for some constant  $C_1$

$$|2X_{s-}^* b_s + \|\sigma_s\|^2| \leq C_1(1 + \|X_{s-}\|^2), \quad \mathbb{P}\text{-a.s.}, \quad s \in [0, T],$$

Recalling that  $\eta$  is non-negative and that  $\mathbb{E}[\eta_t] \leq 1$ , for any  $t \in [0, T]$ , we get

$$\begin{aligned} \mathbb{E} \left[ \int_0^{\tau-} \frac{\eta_{s-} [2X_{s-}^* b_s + \|\sigma_s\|^2]}{[1 + \varepsilon \eta_{s-} \|X_{s-}\|^2]^2} ds \right] &\leq C_1 \mathbb{E} \left[ \int_0^{\tau-} \frac{\eta_{s-} (1 + \|X_{s-}\|^2)}{[1 + \varepsilon \eta_{s-} \|X_{s-}\|^2]^2} ds \right] \\ &\leq C_1 \mathbb{E} \left[ \int_0^{\tau-} \eta_{s-} ds \right] + C_1 \mathbb{E} \left[ \int_0^{\tau-} \frac{\eta_{s-} \|X_{s-}\|^2}{1 + \varepsilon \eta_{s-} \|X_{s-}\|^2} ds \right] \\ &\leq C_1 T + C_1 \mathbb{E} \left[ \int_0^{\tau-} \frac{\eta_{s-} \|X_{s-}\|^2}{1 + \varepsilon \eta_{s-} \|X_{s-}\|^2} ds \right]. \end{aligned} \quad (\text{A.7})$$

Next, we see that

$$\begin{aligned}
 & \mathbb{E} \left[ \int_0^{\tau^-} \frac{2\eta_{s-} X_{s-}}{[1 + \varepsilon\eta_{s-} \|X_{s-}\|^2]^2} d\nu_s^c \right] = \sum_{i=1}^m \mathbb{E} \left[ \int_0^{\tau^-} \frac{2\eta_{s-} X_{s-}^i}{[1 + \varepsilon\eta_{s-} \|X_{s-}\|^2]^2} d\nu_s^{i,c} \right] \\
 & \leq \sum_{i=1}^m \mathbb{E} \left[ \int_0^{\tau^-} \frac{2\eta_{s-} |X_{s-}^i|}{[1 + \varepsilon\eta_{s-} \|X_{s-}\|^2]^2} d|\nu^{i,c}|_s \right] \leq \sum_{i=1}^m \mathbb{E} \left[ \int_0^{\tau^-} \frac{\eta_{s-} (1 + |X_{s-}^i|^2)}{[1 + \varepsilon\eta_{s-} \|X_{s-}\|^2]^2} d|\nu^{i,c}|_s \right] \\
 & \leq \sum_{i=1}^m \mathbb{E} \left[ \int_0^{\tau^-} \eta_{s-} d|\nu^{i,c}|_s \right] + \sum_{i=1}^m \mathbb{E} \left[ \int_0^{\tau^-} \frac{\eta_{s-} \|X_{s-}\|^2}{1 + \varepsilon\eta_{s-} \|X_{s-}\|^2} d|\nu^{i,c}|_s \right]. \tag{A.8}
 \end{aligned}$$

Similarly to what we did in the proof of Lemma A.1, let us define

$$\tilde{\nu}_t^i := |\nu^i|_t \mathbf{1}_{t < T} + K \mathbf{1}_{t \geq T}, \quad t \geq 0, \quad i = 1, \dots, m.$$

For each  $i = 1, \dots, m$ ,  $\tilde{\nu}^i$  is a  $\mathbb{Y}$ -adapted, càdlàg, non-decreasing process, with  $\tilde{\nu}_{0-}^i = 0$ . Moreover, random measures  $\mathbf{1}_{s < \tau} d|\nu^i|_s$  and  $\mathbf{1}_{s < \tau} d\tilde{\nu}_s^i$  agree, therefore

$$\mathbb{E} \left[ \int_0^{\tau^-} \eta_{s-} d|\nu^i|_s \right] = \mathbb{E} \left[ \int_0^{\tau^-} \eta_{s-} d\tilde{\nu}_s^i \right], \quad i = 1, \dots, m,$$

and, in particular,

$$\mathbb{E} \left[ \int_0^{\tau^-} \eta_{s-} d|\nu^{i,c}|_s \right] = \mathbb{E} \left[ \int_0^{\tau^-} \eta_{s-} d\tilde{\nu}_s^{i,c} \right], \quad i = 1, \dots, m.$$

Let us define the changes of time  $C_t^i := \inf\{s \geq 0 : \tilde{\nu}_s^i \geq t\}$ , for all  $t \geq 0$  and all  $i = 1, \dots, m$ . Then, noticing that  $\{C_s^i \leq t\} = \{\tilde{\nu}_t^i \geq s\}$  and recalling that  $\eta$  is non-negative and  $\tilde{\nu}_T^i = K$ , we get

$$\begin{aligned}
 & \mathbb{E} \left[ \int_0^{\tau^-} \eta_{s-} d\tilde{\nu}_s^{i,c} \right] \leq \mathbb{E} \left[ \int_0^T \eta_{s-} d\tilde{\nu}_s^{i,c} \right] \leq \mathbb{E} \left[ \int_0^T \eta_{s-} d\tilde{\nu}_s^i \right] = \mathbb{E} \left[ \int_0^{+\infty} \eta_{(C_s^i)-} \mathbf{1}_{C_s^i \leq T} ds \right] \\
 & = \mathbb{E} \left[ \int_0^{+\infty} \eta_{(C_s^i)-} \mathbf{1}_{s \leq \tilde{\nu}_T^i} ds \right] = \mathbb{E} \left[ \int_0^K \eta_{(C_s^i)-} ds \right] = \int_0^K \mathbb{E}[\eta_{(C_s^i)-}] ds.
 \end{aligned}$$

Since  $\mathbb{E}[\eta_t] \leq 1$ , for any  $t \in [0, T]$ , and  $C_s^i \leq T$ , for all  $s \in [0, K]$ , we get that

$$\mathbb{E} \left[ \int_0^{\tau^-} \eta_{s-} d|\nu^{i,c}|_s \right] = \mathbb{E} \left[ \int_0^{\tau^-} \eta_{s-} d\tilde{\nu}_s^{i,c} \right] \leq K, \quad i = 1, \dots, m. \tag{A.9}$$

Similarly, we obtain also

$$\mathbb{E} \left[ \int_0^{\tau^-} \eta_{s-} d|\nu^i|_s \right] = \mathbb{E} \left[ \int_0^{\tau^-} \eta_{s-} d\tilde{\nu}_s^i \right] \leq K, \quad i = 1, \dots, m. \tag{A.10}$$

Therefore, putting together (A.8) and (A.9) we obtain

$$\mathbb{E} \left[ \int_0^{\tau^-} \frac{2\eta_{s-}}{[1 + \varepsilon\eta_{s-} \|X_{s-}\|^2]^2} X_{s-} d\nu_s^c \right] \leq mK + \sum_{i=1}^m \mathbb{E} \left[ \int_0^{\tau^-} \frac{\eta_{s-} \|X_{s-}\|^2}{1 + \varepsilon\eta_{s-} \|X_{s-}\|^2} d|\nu^{i,c}|_s \right] \tag{A.11}$$

We are left with estimating the last term of (A.6). We have:

$$\begin{aligned}
 & \mathbb{E} \left[ \sum_{0 \leq s < \tau} \left\{ \frac{\eta_s \|X_s\|^2}{1 + \varepsilon\eta_s \|X_s\|^2} - \frac{\eta_{s-} \|X_{s-}\|^2}{1 + \varepsilon\eta_{s-} \|X_{s-}\|^2} \right\} \right] \leq \mathbb{E} \left[ \sum_{0 \leq s < \tau} \left\{ \frac{\eta_{s-} (\|X_s\|^2 - \|X_{s-}\|^2)}{1 + \varepsilon\eta_{s-} \|X_{s-}\|^2} \right\} \right] \\
 & = \mathbb{E} \left[ \sum_{0 \leq s < \tau} \left\{ \frac{\eta_{s-} (\|\Delta\nu_s\|^2 + 2X_{s-} \cdot \Delta\nu_s)}{1 + \varepsilon\eta_{s-} \|X_{s-}\|^2} \right\} \right] \leq \mathbb{E} \left[ \sum_{0 \leq s < \tau} \sum_{i=1}^m \left\{ \frac{\eta_{s-} (|\Delta\nu_s^i| + 1 + |X_{s-}^i|^2)}{1 + \varepsilon\eta_{s-} \|X_{s-}\|^2} |\Delta\nu_s^i| \right\} \right],
 \end{aligned}$$

where we used the fact that  $\eta$  is continuous. Since all quantities in the last term are non negative and  $|\Delta\nu_s^i| \leq K$ , for all  $s \in [0, T]$  and all  $i = 1, \dots, m$ ,  $\mathbb{P}$ -a.s., we get that

$$\begin{aligned}
 & \mathbb{E} \left[ \sum_{0 \leq s < \tau} \sum_{i=1}^m \left\{ \frac{\eta_{s-} (|\Delta\nu_s^i| + 1 + |X_{s-}^i|^2)}{1 + \varepsilon\eta_{s-} \|X_{s-}\|^2} |\Delta\nu_s^i| \right\} \right] \\
 & \leq (1 + K) \sum_{i=1}^m \mathbb{E} \left[ \sum_{0 \leq s < \tau} \eta_{s-} |\Delta\nu_s^i| \right] + \sum_{i=1}^m \mathbb{E} \left[ \sum_{0 \leq s < \tau} \left\{ \frac{\eta_{s-} \|X_{s-}\|^2}{1 + \varepsilon\eta_{s-} \|X_{s-}\|^2} |\Delta\nu_s^i| \right\} \right] \\
 & \leq (1 + K) \sum_{i=1}^m \mathbb{E} \left[ \int_0^{\tau^-} \eta_{s-} d|\nu^i|_s \right] + \sum_{i=1}^m \mathbb{E} \left[ \sum_{0 \leq s < \tau} \left\{ \frac{\eta_{s-} \|X_{s-}\|^2}{1 + \varepsilon\eta_{s-} \|X_{s-}\|^2} \Delta|\nu^i|_s \right\} \right]
 \end{aligned}$$

$$\leq mK(1+K) + \sum_{i=1}^m \mathbb{E} \left[ \sum_{0 \leq s < \tau} \left\{ \frac{\eta_{s-} \|X_{s-}\|^2}{1 + \varepsilon \eta_{s-} \|X_{s-}\|^2} \Delta |\nu^i|_s \right\} \right], \quad (\text{A.12})$$

where we used (A.10) and the fact that  $|\Delta \nu^i| = \Delta |\nu^i|$ .

Therefore, feeding (A.7), (A.11), and (A.12) back into (A.6), we obtain

$$\mathbb{E} \left[ \frac{\eta_{\tau-} \|X_{\tau-}\|^2}{1 + \varepsilon \eta_{\tau-} \|X_{\tau-}\|^2} \right] \leq M \left\{ 1 + \mathbb{E} \left[ \int_0^{\tau-} \frac{\eta_{s-} \|X_{s-}\|^2}{1 + \varepsilon \eta_{s-} \|X_{s-}\|^2} dA_s \right] \right\}, \quad (\text{A.13})$$

where  $M$  is a suitable constant, not depending on  $\varepsilon$ , and  $A$  is the process

$$A_t := t + \sum_{i=1}^m |\nu^i|_t, \quad t \in [0, T].$$

Clearly,  $A$  is a càdlàg,  $\mathbb{Y}$ - (and hence  $\mathbb{F}$ -) adapted, non-negative process, with  $A_{0-} = 0$  and  $A_T \leq T + mK$ . Moreover,  $\frac{\eta_{t-} \|X_{t-}\|^2}{1 + \varepsilon \eta_{t-} \|X_{t-}\|^2} \leq \frac{1}{\varepsilon}$ , for all  $t \in [0, T]$ ,  $\mathbb{P}$ -a.s. Therefore, we can apply Lemma A.1 and obtain

$$\mathbb{E} \left[ \frac{\eta_{t-} \|X_{t-}\|^2}{1 + \varepsilon \eta_{t-} \|X_{t-}\|^2} \right] \leq M e^{M(T+mK)}.$$

Recalling that  $\eta$  is continuous we get, applying Fatou's lemma,

$$\mathbb{E}[\eta_t \|X_{t-}\|^2] = \mathbb{E} \left[ \lim_{\varepsilon \rightarrow 0} \frac{\eta_t \|X_{t-}\|^2}{1 + \varepsilon \eta_t \|X_{t-}\|^2} \right] \leq \liminf_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \frac{\eta_t \|X_{t-}\|^2}{1 + \varepsilon \eta_t \|X_{t-}\|^2} \right] \leq M e^{M(T+mK)}. \quad (\text{A.14})$$

It is important to stress that (A.14) holds for any  $t \in [0, T]$ , since  $t$  was arbitrarily chosen.

Now we can finally obtain that  $\mathbb{E}[\eta_t] = 1$ , for all  $t \in [0, T]$ . By Itô's formula, for an arbitrarily fixed  $\varepsilon > 0$  and all  $t \in [0, T]$ ,

$$\frac{\eta_t}{1 + \varepsilon \eta_t} = \frac{1}{1 + \varepsilon} - \int_0^t \frac{\varepsilon \eta_s^2}{(1 + \varepsilon \eta_s)^3} \|\gamma^{-1}(s) h(s, X_s)\|^2 ds + \int_0^t \frac{\eta_s}{(1 + \varepsilon \eta_s)^2} \gamma^{-1}(s) h(s, X_s) d\bar{B}_s.$$

Thanks to conditions (2.10) and (2.1), standard computations show that the stochastic integral is a  $(\mathbb{P}, \mathbb{F})$ -martingale. Therefore, taking the expectation we get

$$\mathbb{E} \left[ \frac{\eta_t}{1 + \varepsilon \eta_t} \right] = \frac{1}{1 + \varepsilon} - \mathbb{E} \left[ \int_0^t \frac{\varepsilon \eta_s^2}{(1 + \varepsilon \eta_s)^3} \|\gamma^{-1}(s) h(s, X_s)\|^2 ds \right].$$

Notice that  $\frac{\varepsilon \eta_s^2}{(1 + \varepsilon \eta_s)^3} \|\gamma^{-1}(s) h(s, X_s)\|^2 \rightarrow 0$ , as  $\varepsilon \rightarrow 0$ ,  $d\mathbb{P} \otimes dt$ -a.s. Moreover,

$$\frac{\varepsilon \eta_s^2}{(1 + \varepsilon \eta_s)^3} \|\gamma^{-1}(s) h(s, X_s)\|^2 \leq \eta_s \|\gamma^{-1}(s) h(s, X_s)\|^2, \quad s \in [0, T],$$

that, using conditions (2.10) and (2.1), satisfies (see also (A.4))

$$\begin{aligned} \mathbb{E} \left[ \int_0^T \eta_s \|\gamma^{-1}(s) h(s, X_s)\|^2 ds \right] &\leq n C_h C_\gamma \mathbb{E} \left[ \int_0^T \eta_s (1 + \|X_s\|^2) ds \right] \\ &= n C_h C_\gamma \left\{ \int_0^T \mathbb{E}[\eta_s] ds + \int_0^T \mathbb{E}[\eta_s \|X_{s-}\|^2] ds \right\} \leq n C_h C_\gamma T [1 + M e^{M(T+mK)}], \end{aligned}$$

where we used the fact that  $\mathbb{E}[\eta_t] \leq 1$ , for all  $t \in [0, T]$ , and (A.14). Similarly,  $\frac{\eta_t}{1 + \varepsilon \eta_t} \rightarrow \eta_t$ , as  $\varepsilon \rightarrow 0$ ,  $d\mathbb{P} \otimes dt$ -a.s., and  $\mathbb{E}[\int_0^T \eta_s ds] \leq T$ . Therefore, by the dominated convergence theorem

$$\mathbb{E}[\eta_t] = \lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \frac{\eta_t}{1 + \varepsilon \eta_t} \right] = \lim_{\varepsilon \rightarrow 0} \left\{ \frac{1}{1 + \varepsilon} - \mathbb{E} \left[ \int_0^t \frac{\varepsilon \eta_s^2}{(1 + \varepsilon \eta_s)^3} \|\gamma^{-1}(s) h(s, X_s)\|^2 ds \right] \right\} = 1,$$

and this concludes the proof.  $\square$

*Proof of Lemma 4.3.* Fix  $\varepsilon > 0$ . To start, let us notice that continuity of process  $\nu$  implies that also  $\zeta$  is continuous and, therefore,  $\zeta_t = \zeta_{t-}$  and  $T_\varepsilon \zeta_t = T_\varepsilon \zeta_{t-}$ ,  $dt \otimes d\mathbb{P}$ -almost everywhere.

Since  $\psi_{2\varepsilon}$  is bounded by  $(4\pi\varepsilon)^{-\frac{m}{2}}$ , we get that for all  $t \in [0, T]$ ,

$$\|T_\varepsilon \zeta_t\|_H^2 = \int_{\mathbb{R}^m} \left[ \int_{\mathbb{R}^m} \psi_\varepsilon(x-y) \zeta_t(dx) \right]^2 dy \quad (\text{A.15})$$

$$= \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \psi_\varepsilon(x-y) \psi_\varepsilon(z-y) \zeta_t(dx) \zeta_t(dz) dy \quad (\text{A.16})$$

$$= \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \psi_{2\varepsilon}(x-z) \zeta_t(dx) \zeta_t(dz) \leq (4\pi\varepsilon)^{-\frac{m}{2}} \zeta_t(1)^2. \quad (\text{A.17})$$



Taking into account (3.8) and the fact that  $\nu$  is continuous, the process  $\zeta(1)$  satisfies

$$\zeta_t(1) = 1 + \int_0^t \gamma^{-1}(s) \zeta_s(h_s) d\bar{B}_s, \quad t \in [0, T],$$

where  $h_t(\cdot) := h(t, \cdot)$ ,  $t \in [0, T]$ . Thanks to Assumption 4.1,  $\zeta_t(h_t) < +\infty$ ,  $\mathbb{P}$ -a.s., for all  $t \in [0, T]$ . Therefore, since  $\zeta_t$  is  $\mathbb{P}$ -a.s. a finite (non-negative) measure, for any  $t \in [0, T]$ , we get that  $\zeta(1)$  is a non-negative  $(\mathbb{P}, \mathbb{Y})$ -local martingale, and hence a  $(\mathbb{P}, \mathbb{Y})$ -supermartingale.

The next step is to prove that  $\zeta(1)$  is a square-integrable<sup>3</sup>  $(\mathbb{P}, \mathbb{Y})$ -martingale. We follow, first, a reasoning analogous to that of [5, Lemma 4.3.1] (see also [2, Lemma 3.29]) to provide an explicit representation of  $\zeta(1)$ . By Itô's formula we obtain, for any  $\delta > 0$  and all  $t \in [0, T]$ ,

$$\begin{aligned} \log(\sqrt{\delta + \zeta_t(1)^2}) &= \log(\sqrt{1 + \delta}) + \int_0^t \frac{\zeta_s(1)}{\delta + \zeta_s(1)^2} \gamma^{-1}(s) \zeta_s(h_s) d\bar{B}_s \\ &+ \frac{1}{2} \int_0^t \frac{\delta - \zeta_s(1)^2}{[\delta + \zeta_s(1)^2]^2} \sum_{i=1}^n \left( \sum_{j=1}^n \gamma_{ij}^{-1}(s) \zeta_s(h_s^j) \right)^2 ds. \end{aligned} \quad (\text{A.18})$$

Since, thanks to Assumption 4.1 and (A.3),

$$\sum_{i=1}^n \left( \sum_{j=1}^n \gamma_{ij}^{-1}(t) \zeta_t(h_t^j) \right)^2 \leq (nC_\gamma K_h \zeta_t(1))^2, \quad \mathbb{P}\text{-a.s.}, \quad \forall t \in [0, T], \quad (\text{A.19})$$

and  $\frac{\delta - \zeta_t(1)^2}{[\delta + \zeta_t(1)^2]^2} \leq \frac{1}{\delta + \zeta_t(1)^2}$ ,  $\mathbb{P}$ -a.s., for all  $t \in [0, T]$ , we have

$$\frac{\zeta_s(1)^2}{[\delta + \zeta_s(1)^2]^2} \sum_{i=1}^n \left( \sum_{j=1}^n \gamma_{ij}^{-1}(t) \zeta_t(h_t^j) \right)^2 \leq \left( nC_\gamma K_h \frac{\zeta_s(1)^2}{\delta + \zeta_s(1)^2} \right)^2 \leq (nC_\gamma K_h)^2, \quad \forall t \in [0, T],$$

and

$$\frac{\delta - \zeta_s(1)^2}{[\delta + \zeta_s(1)^2]^2} \sum_{i=1}^n \left( \sum_{j=1}^n \gamma_{ij}^{-1}(s) \zeta_s(h_s^j) \right)^2 \leq \frac{\zeta_s(1)^2}{\delta + \zeta_s(1)^2} (nC_\gamma K_h)^2 \leq (nC_\gamma K_h)^2, \quad \forall t \in [0, T].$$

Both the r.h.s. of the last two inequalities are integrable on  $[0, T]$ , therefore we can pass to the limit, as  $\delta \rightarrow 0$ , in (A.18), getting that, for all  $t \in [0, T]$ ,

$$\log(\zeta_t(1)) = 1 + \int_0^t \gamma^{-1}(s) \zeta_s^1(h_s) d\bar{B}_s - \frac{1}{2} \int_0^t \sum_{i=1}^n \left( \sum_{j=1}^n \gamma_{ij}^{-1}(s) \zeta_s^1(h_s^j) \right)^2 ds, \quad (\text{A.20})$$

where  $\zeta_t^1(dx) := \frac{\zeta_t(dx)}{\zeta_t(1)}$ ,  $t \in [0, T]$ , is the normalized process associated to  $\zeta$ . From (A.20) we get the explicit representation for  $\zeta(1)$ , i.e., for all  $t \in [0, T]$ ,

$$\zeta_t(1) = \exp \left\{ \int_0^t \gamma^{-1}(s) \zeta_s^1(h_s) d\bar{B}_s - \frac{1}{2} \int_0^t \sum_{i=1}^n \left( \sum_{j=1}^n \gamma_{ij}^{-1}(s) \zeta_s^1(h_s^j) \right)^2 ds \right\}. \quad (\text{A.21})$$

This entails that  $\zeta(1)$  coincides with the Doléans-Dade exponential of the continuous  $(\mathbb{P}, \mathbb{Y})$ -local martingale  $\int_0^t \gamma^{-1}(s) \zeta_s^1(h_s) d\bar{B}_s$ ,  $t \in [0, T]$ . Using once more (A.19) we have that, for any  $k > 1$ ,

$$\mathbb{E} \left[ \exp \left\{ \frac{k}{2} \int_0^T \sum_{i=1}^n \left( \sum_{j=1}^n \gamma_{ij}^{-1}(t) \zeta_t^1(h_t^j) \right)^2 dt \right\} \right] \leq \exp \left\{ \frac{kT(nC_\gamma K_h)^2}{2} \right\}.$$

Applying [13, Theorem 15.4.6], we get that, for any  $p > 1$ ,  $\zeta(1)$  is a  $p$ -integrable (in particular, square-integrable)  $(\mathbb{P}, \mathbb{Y})$ -martingale. Therefore, from (A.15) we get

$$\mathbb{E} \left[ \sup_{t \in [0, T]} \|T_\varepsilon \zeta_t\|_H^2 \right] \leq (4\pi\varepsilon)^{-\frac{m}{2}} \mathbb{E} \left[ \sup_{t \in [0, T]} \zeta_t(1)^2 \right] < +\infty,$$

whence, recalling the remark at the beginning of the proof, the claim.  $\square$

<sup>3</sup>If  $M = (M_t)_{t \in [0, T]}$  is any martingale, we say that  $M$  is a  $p$ -integrable martingale, with  $p \geq 1$ , if  $\mathbb{E}[\sup_{t \in [0, T]} |M_t|^p]^{1/p} < +\infty$ .

## REFERENCES

- [1] L.H.R. Alvarez and L.A. Shepp. Optimal harvesting of stochastically fluctuating populations. *J. Math. Biol.*, 37:155–177, 1998.
- [2] A. Bain and D. Crisan. *Fundamentals of Stochastic Filtering*. Springer, New York, 2009.
- [3] E. Bandini, A. Cosso, M. Fuhrman, and H. Pham. Randomized filtering and Bellman equation in Wasserstein space for partial observation control problem. *Stochastic Process. Appl.*, 129(2):674–711, 2019.
- [4] E. Bandini, A. Calvia, and K. Colaneri. Stochastic filtering of a pure jump process with predictable jumps and path-dependent local characteristics. Preprint, arXiv:2004.12944, 2021.
- [5] A. Bensoussan. *Stochastic control of partially observable systems*. Cambridge University Press, Cambridge, 1992.
- [6] P. Brémaud. *Point Processes and Queues*. Springer Series in Statistics. Springer-Verlag, New York, 1981.
- [7] G. Callegaro, C. Ceci, and G. Ferrari. Optimal reduction of public debt under partial observation of the economic growth. *Fin. Stoch.*, 24(4):1083–1132, 2020.
- [8] A. Calvia. Stochastic filtering and optimal control of pure jump Markov processes with noise-free partial observation. *ESAIM: Control, Optimisation and Calculus of Variations*, 26:25, 2020. doi: 10.1051/cocv/2019020.
- [9] C. Ceci and K. Colaneri. Nonlinear filtering for jump diffusion observations. *Adv. in Appl. Probab.*, 44(3):678–701, 2012. doi: 10.1239/aap/1346955260.
- [10] C. Ceci and K. Colaneri. The Zakai equation of nonlinear filtering for jump-diffusion observations: existence and uniqueness. *Appl. Math. Optim.*, 69(1):47–82, 2014. doi: 10.1007/s00245-013-9217-1.
- [11] C. Ceci and A. Gerardi. Filtering of a markov jump process with counting observations. *Applied Mathematics and Optimization*, 42(1):1–18, 2000.
- [12] C. Ceci and A. Gerardi. Nonlinear filtering equation of a jump process with counting observations. *Acta Applicandae Mathematica*, 66(2):139–154, 2001.
- [13] S. N. Cohen and R. J. Elliott. *Stochastic calculus and applications*. Probability and its Applications. Springer, Cham, second edition, 2015.
- [14] F. Confortola and M. Fuhrman. Filtering of continuous-time Markov chains with noise-free observation and applications. *Stochastics An International Journal of Probability and Stochastic Processes*, 85(2): 216–251, 2013.
- [15] D. Crisan and B. Rozovskiĭ, editors. *The Oxford handbook of nonlinear filtering*. Oxford University Press, Oxford, 2011.
- [16] T. De Angelis. Optimal dividends with partial information and stopping of a degenerate reflecting diffusion. *Fin. Stoch.*, 24(1):71–123, 2020.
- [17] J.-P. Décamps and S. Villeneuve. Dynamics of cash holdings, learning about profitability, and access to the market. Preprint, TSE Working Paper, n. 19-1046, 2020.
- [18] C. Dellacherie and P.-A. Meyer. *Probabilities and potential*. B, volume 72 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1982. Theory of martingales, Translated from the French by J. P. Wilson.
- [19] N. El Karoui and I. Karatzas. Probabilistic aspects of finite-fuel, reflected follower problems. *Acta Appl. Math.*, 11:223–258, 1988.
- [20] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.
- [21] G. Fabbri, F. Gozzi, and A. Swiech. *Stochastic optimal control in infinite dimension*, volume 82 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2017. Dynamic programming and HJB equations, With a contribution by Marco Fuhrman and Gianmario Tessitore.
- [22] S. Federico, G. Ferrari, and N. Rodosthenous. Two-sided singular control of an inventory with unknown demand trend. Preprint, arXiv:2102.11555, 2021.
- [23] B. Grigelionis and R. Mikulevicius. Nonlinear filtering equations for processes with jumps. In D. Crisan and B. Rozovskiĭ, editors, *The Oxford Handbook of Nonlinear Filtering*. Oxford University Press, 2011.
- [24] J.M. Harrison and M.I. Taksar. Instantaneous control of brownian motion. *Math. Oper. Res.*, 8(3): 439–453, 1983.
- [25] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, second edition, 2003. doi: 10.1007/978-3-662-05265-5.

- [26] G. Kallianpur. *Stochastic filtering theory*, volume 13 of *Applications of Mathematics*. Springer-Verlag, New York-Berlin, 1980.
- [27] I. Karatzas, D. Ocone, H. Wang, and M. Zervos. Finite-fuel singular control with discretionary stopping. *Stoch. Stoch. Rep.*, 71(1-2):1–50, 2000.
- [28] T. G. Kurtz and G. Nappo. The filtered martingale problem. In *The Oxford handbook of nonlinear filtering*, pages 129–165. Oxford Univ. Press, Oxford, 2011.
- [29] T. G. Kurtz and R. H. Stockbridge. Stationary solutions and forward equations for controlled and singular martingale problems. *Electron. J. Probab.*, 6:no. 17, 52, 2001. doi: 10.1214/EJP.v6-90.
- [30] T. G. Kurtz and J. Xiong. Particle representations for a class of nonlinear SPDEs. *Stochastic Process. Appl.*, 83(1):103–126, 1999. doi: 10.1016/S0304-4149(99)00024-1.
- [31] T.G. Kurtz and D. Ocone. Unique characterization of condition distribution in nonlinear filtering. *Ann. Probab.*, 16:80–107, 1988.
- [32] R. S. Liptser and A. N. Shiryaev. *Statistics of random processes. I*, volume 5 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, expanded edition, 2001. General theory, Translated from the 1974 Russian original by A. B. Aries, Stochastic Modelling and Applied Probability.
- [33] V. M. Lucic and A. J. Heunis. On uniqueness of solutions for the stochastic differential equations of nonlinear filtering. *Ann. Appl. Probab.*, 11(1):182–209, 2001.
- [34] P. E. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2004. Stochastic Modelling and Applied Probability.
- [35] M. Reppen, J.-C. Rochet, and M. Soner. Optimal dividend policies with random profitability. *Math. Fin.*, 30(1):228–259, 2020.
- [36] F. Riedel and X. Su. On irreversible investment. *Fin. Stoch.*, 15(4):607–633, 2011.
- [37] J. Xiong. *An Introduction to Stochastic Filtering Theory*. Oxford University Press, New York, 2008.