Stochastics and Partial Differential Equations: Analysis and Computations
https://doi.org/10.1007/s40072-025-00353-4

On partially observed jump diffusions I: the filtering equations

Fabian Germ[1,2] · **István Gyöngy**[1]

Received: 30 June 2023 / Revised: 27 January 2025 / Accepted: 4 February 2025

Abstract
This paper is the first part of a series of papers on filtering for partially observed jump diffusions satisfying a stochastic differential equation driven by Wiener processes and Poisson martingale measures. The coefficients of the equation only satisfy appropriate growth conditions. Some results in filtering theory of diffusion processes are extended to jump diffusions and equations for the time evolution of the conditional distribution and the unnormalised conditional distribution of the unobserved process at time t, given the observations until t, are presented.

Keywords Nonlinear filtering · Random measures · Lévy processes

Mathematics Subject Classification Primary 60G35 · 60H15; Secondary 60G57 · 60G51

1 Introduction

This is the first part of a series of papers on filtering of jump diffusions. We consider on a given complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ a $d + d'$-dimensional stochastic process $(Z_t)_{t\in[0,T]} = (X_t, Y_t)_{t\in[0,T]}$, satisfying the stochastic differential

✉ István Gyöngy
i.gyongy@ed.ac.uk

Fabian Germ
fgerm@ed.ac.uk

[1] School of Mathematics and Maxwell Institute, University of Edinburgh, Edinburgh, Scotland, UK

[2] Delft University of Technology, Delft, Netherlands

Published online: 01 April 2025

equation

$$
\begin{aligned}
dX_t &= b(t, Z_t)dt + \sigma(t, Z_t)dW_t + \rho(t, Z_t)dV_t \\
&\quad + \int_{\mathfrak{Z}_0} \eta(t, Z_{t-}, \mathfrak{z})\, \tilde{N}_0(d\mathfrak{z}, dt) + \int_{\mathfrak{Z}_1} \xi(t, Z_{t-}, \mathfrak{z})\, \tilde{N}_1(d\mathfrak{z}, dt), \\
dY_t &= B(t, Z_t)dt + dV_t + \int_{\mathfrak{Z}_1} \mathfrak{z}\, \tilde{N}_1(d\mathfrak{z}, dt),
\end{aligned}
\tag{1.1}
$$

on the interval $[0, T]$ for a given $\mathcal{F}_0$-measurable initial value $Z_0 = (X_0, Y_0)$, where $(W_t, V_t)_{t\geq 0}$ is $d_1 + d'$-dimensional $\mathcal{F}_t$-Wiener process, and $\tilde{N}_i(d\mathfrak{z}, dt) = N_i(d\mathfrak{z}, dt) - \nu_i(d\mathfrak{z})dt$, $i = 0, 1$, are independent $\mathcal{F}_t$-Poisson martingale measures on $\mathbb{R}_+ \times \mathfrak{Z}_i$ with σ-finite characteristic measures ν_0 and ν_1 on separable measurable spaces $(\mathfrak{Z}_0, \mathcal{Z}_0)$ and $(\mathfrak{Z}_1, \mathcal{Z}_1) = (\mathbb{R}^{d'} \setminus \{0\}, \mathcal{B}(\mathbb{R}^{d'} \setminus \{0\}))$, respectively. The mappings $b = (b^i)$, $B = (B^i)$, $\sigma = (\sigma^{ij})$ and $\rho = (\rho^{il})$ are Borel functions of $(t, z) = (t, x, y) \in \mathbb{R}_+ \times \mathbb{R}^{d+d'}$, with values in $\mathbb{R}^d$, $\mathbb{R}^{d'}$, $\mathbb{R}^{d\times d_1}$ and $\mathbb{R}^{d\times d'}$, respectively, and $\eta = (\eta^i)$ and $\xi = (\xi^i)$ are $\mathbb{R}^d$-valued $\mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^{d+d'}) \otimes \mathcal{Z}_0$-measurable and $\mathbb{R}^d$-valued $\mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^{d+d'}) \otimes \mathcal{Z}_1$-measurable functions of $(t, z, \mathfrak{z}_0) \in \mathbb{R}_+ \times \mathbb{R}^{d+d'} \times \mathfrak{Z}_0$ and $(t, z, \mathfrak{z}_1) \in \mathbb{R}_+ \times \mathbb{R}^{d+d'} \times \mathfrak{Z}_1$, respectively. Here, and later on, $\mathcal{B}(\mathcal{V})$ denotes the Borel σ-algebra on $\mathcal{V}$ for topological spaces $\mathcal{V}$.

We are concerned with the classic task of filtering theory: to calculate at each time t the mean square estimate of $f(X_t)$, a real-valued Borel function of the "unobservable component" X_t of Z_t, given the "observations" $\{Y_s : s \leq t\}$. Since, as it is well-known, this estimate is the conditional expectation

$$
\mathbb{E}(f(X_t)|Y_s : s \leq t) = \int_{\mathbb{R}^d} f(x)P_t(dx), \quad t \in [0, T],
$$

we are interested in equations for the evolution of $P_t(dx)$, the conditional distribution of X_t given $\{Y_s, s \leq t\}$. Their derivation under general assumptions on Eq. (1.1) is the aim of this paper. In the subsequent papers of this series we investigate the existence of the conditional density $\pi_t(x) = P_t(dx)/dx$ and its regularity properties.

There has been an immense interest in the development of filtering theory due to its wide applicability in various disciplines, be they of theoretical or applied nature. A vast amount of research has been done on filtering of partially observed processes governed by stochastic differential equations driven by Wiener processes, i.e., when $\eta = \xi = 0$ in (1.1), and a quite complete nonlinear filtering theory was built up, see for instance [7] for a historical account.
In this case it is well-known that $(P_t(dx))_{t\in[0,T]}$ satisfies a nonlinear stochastic PDE (SPDE), often called the Kushner-Shiryayev-Stratonovich equation in filtering theory. It is also well-known that this equation can be transformed into a linear SPDE, called Zakai equation, or Duncan-Mortensen-Zakai equation for $\mu_t(dx) = \lambda_t P_t(dx)$, the unnormalised conditional distribution, where $(\lambda_t)_{t\in[0,T]}$ is a positive normalizing stochastic process.

There exist several known methods of deriving the filtering equations for partially observed diffusion processes, three prominent of which are the "innovation method", the "reference measure method" and a "direct approach". The innovation method is based on "innovation process" representations, (see [23] and [11]), and the direct approach is based on suitable existence and uniqueness theorems for stochastic PDEs (see [18]). The reference probability method is employed in this paper, where we make use of the fact that by Girsanov's theorem one can introduce a new measure under which the observation σ-algebra, $\sigma(Y_s : s \leq t)$, is the product σ-algebra of three independent σ-algebras: the σ-algebra generated by the initial observation Y_0 and the σ-algebras generated by the Wiener process with the stochastic differential $B(t, Z_t)dt + dV_t$, and the Poisson random measure $N_1(d\mathfrak{z}, dt)$ on $[0, t] \times \mathfrak{Z}_1$, respectively. This structure of the observation σ-algebra makes it possible to calculate conditional expectations of functions of the process Z given the observations. (See, e.g., [3] for descriptions of various methods used in filtering theory.)

Recently, also filtering for jump diffusion systems have been intensively studied, which are most often modelled as SDEs driven by Wiener processes as well as random jump measures, a classical case of which are Poisson random measures. In an early article thereon, [24], the filtering equations were derived for uncorrelated continuous observations, as well as an observation process driven only by a jump process that has no common jumps with the signal. A similar system with continuous uncorrelated observations has also been considered in [25]. A more general nonlinear system with jumps in the observation process was considered in [2]. In [1] the filtering equations for a large class of uncorrelated linear systems with jumps are derived. In [13] a very general model is considered and a representation for optional projection of the signal process is derived. However, due to the generality a number of additional assumptions are imposed on their model and equations for the filtering measures are not obtained. In [5] and [6] the authors deal with a one-dimensional jump diffusion where observation and signal may have common jumps, by introducing a new random measure, nonzero only for observable jumps, relying on a construction in [4]. However, they impose a finiteness condition on the support of the integrand in front of the jump term, which translates to observing only finitely many jumps almost surely. In such a case, the jump measure and the associated compensator, also referred to as dual predictable projection, allow for a specific decomposition, see for instance [15, Sec. XI.4]. The filtering equations have been derived for a class of jump diffusion systems [26], later generalised to include correlated Wiener process noises in [27], however, it seems to us that certain important results needed for this derivation, including Lemma 3.2 in [26], also used in [27], do not hold, for instance if one considers the case of vanishing coefficients. A model where a correlation structure between the Lévy process noises in signal and observation is described using copulas is used in [10] to derive the Zakai equation.

In this paper we obtain the filtering equations for a jump diffusion system driven by correlated Wiener process, as well as correlated Poisson martingale measure noises. We impose common linear growth conditions. We do not assume any non-degeneracy conditions and allow for the number of jumps in any component of $(Z_t)_{t\geq 0}$ to be infinite over finite intervals. In order to obtain the equations, we generalise some results from filtering theory and in particular prove a "projection theorem" for a wide class of functions.

To the best of the author's knowledge, this is to date the most general jump-diffusion model for which the the filtering equations have been derived. Though a model with similar structure has been considered in [27], they do not include feedback from the observation in the signal and the jump noises of signal and observation are uncorrelated. Moreover, as mentioned already above, an important lemma needed for the projection theorem does not hold for vanishing coefficients.

We remark further that in our model (1.1) instead of dV_t in the observation process we may also consider $\Sigma(t, Y_t)dV_t$ with a matrix-valued function Σ of (t, y), such that it satisfies the linear growth condition and $\Sigma\Sigma^*$ is uniformly non-degenerate. However, by known methods this case can be treated in the same way. Indeed, adding Σ in our case would only introduce additional notation and hence we decided to consider Y of the form in (1.1).

It is also possible to consider a more involved coefficient in front of $\tilde{N}_1$, i.e., $\beta(t, Y_{t-}, \mathfrak{z})N_1(dt, d\mathfrak{z})$ in place of $\mathfrak{z}N_1(dt, d\mathfrak{z})$ in the equation for the observation process Y, where $\beta = \beta(t, y, \mathfrak{z})$ is a function of $t \in [0, T]$, $y \in \mathbb{R}^{d'}$ and $\mathfrak{z} \in \mathfrak{Z}_1 = \mathbb{R}^{d'}\setminus\{0\}$ into $\mathfrak{Z}_1$, and it is assumed that $Z = (X_t, Y_t)_{t\in[0,T]}$ is a solution to (1.1) with the coefficient β in front of $\tilde{N}_1$. To extend our main theorem, Theorem 2.1 to this case, as one can see from its proof, besides natural measurability and growth conditions, it is sufficient to assume a non-degeneracy condition on β ensuring that for any $t \in [0, T]$ the observations $\{Y_s : s \in [0, t]\}$ contain all information about the Poisson random measure N_1 until t, i.e., that the completed σ-algebra generated by $\{Y_s : s \le t\}$ contains the σ-algebra generated by $\{N_1((0, s], A) : s \le t, A \in \mathcal{Z}_1, \nu_1(A) < \infty\}$. It seems to us that such a non-degeneracy condition could be that β maps $[0, T] \times \mathbb{R}^{d'} \times \mathfrak{Z}_1$ into $\mathfrak{Z}_1$ such that $\beta(t, y, \cdot)$ is injective on the support of ν_1 for every $t \in [0, T]$ and $y \in \mathbb{R}^{d'}$.

In Sect. 2, a fairly general condition for Girsanov's transformation and our main result are presented. In Sect. 3 a projection theorem covering a wide class of processes is proven, and thereby in the last section the filtering equations are derived.

Conditions and results on the existence and regularity of the filtering density are presented in the subsequent articles [8, 12] of this series.

We conclude with some notions and notations used throughout the paper. For an integer $n \ge 0$ the notation $C^n_b(\mathbb{R}^d)$ means the space of real-valued bounded continuous functions on $\mathbb{R}^d$, which have bounded and continuous derivatives up to order n. (If $n = 0$, then $C^0_b(\mathbb{R}^d) = C_b(\mathbb{R}^d)$ denotes the space of real-valued bounded continuous functions on $\mathbb{R}^d$). We denote by $\mathbb{M} = \mathbb{M}(\mathbb{R}^d)$ the set of finite Borel measures on $\mathbb{R}^d$. For $\mu \in \mathbb{M}$ we use the notation

$$\mu(\varphi) = \int_{\mathbb{R}^d} \varphi(x)\, \mu(dx)$$

for Borel functions φ on $\mathbb{R}^d$. We say that a function $\nu : \Omega \to \mathbb{M}$ is $\mathcal{G}$-measurable for a σ-algebra $\mathcal{G} \subset \mathcal{F}$, if $\nu(\varphi)$ is a $\mathcal{G}$-measurable random variable for every bounded Borel function φ on $\mathbb{R}^d$. An $\mathbb{M}$-valued stochastic process $\nu = (\nu_t)_{t\in[0,T]}$ is said to be weakly cadlag if almost surely $\nu_t(\varphi)$ is a cadlag function of t for all $\varphi \in C_b(\mathbb{R}^d)$. For such a process ν there is a set $\Omega' \subset \Omega$ of full probability and there is uniquely defined (up to indistinguishability) $\mathbb{M}$-valued processes $(\nu_{t-})_{t\in(0,T]}$ such that for every $\omega \in \Omega'$

$$\nu_{t-}(\varphi) = \lim_{s\uparrow t} \nu_s(\varphi) \quad \text{for all } \varphi \in C_b(\mathbb{R}^d) and t \in (0, T],$$

and for each $\omega \in \Omega'$ we have $\nu_{t-} = \nu_t$, for all but at most countably many $t \in (0, T]$. For processes $U = (U_t)_{t\in[0,T]}$ we use the notation $\mathcal{F}_t^U$ for the P-completion of the σ-algebra generated by $\{U_s : s \le t\}$. For a measure space $(\mathfrak{Z}, \mathcal{Z}, \nu)$ and $p \ge 1$ we use the notation $L_p(\mathfrak{Z})$ for the L_p-space of $\mathbb{R}^d$-valued $\mathcal{Z}$-measurable processes defined on $\mathfrak{Z}$. For σ-algebras $\mathcal{G}_i \subset \mathcal{F}$, $i = 1, 2$, the notation $\mathcal{G}_1 \vee \mathcal{G}_2$ means the P-completion of the smallest σ-algebra containing $\mathcal{G}_i$ for $i = 1, 2$. Finally, we always use without mention the summation convention, by which repeated integer valued indices imply a summation.

2 Formulation of the main results

We consider on a given finite interval $[0, T]$ a $d + d'$-dimensional stochastic process $Z = (Z_t)_{t\in[0,T]} = (X_t, Y_t)_{t\in[0,T]}$ carried by a complete probability space $(\Omega, \mathcal{F}, P)$, equipped with a filtration $(\mathcal{F}_t)_{t\ge0}$ such that $\mathcal{F}_0$ contains the P-null sets of $\mathcal{F}$. We assume that Z satisfies the stochastic differential equation (1.1) on the interval $[0, T]$, with an $\mathcal{F}_0$-measurable initial value $Z_0 = (X_0, Y_0)$.

Besides the natural measurability conditions on the coefficients b, σ, ρ, ξ, η and B, described in the Introduction, we assume the following conditions.

Assumption 2.1 (i) There are nonnegative constants K_0, K_1 and K_2 such that

$$|b(t,z)|^2 \le K_0 + K_1|z|^2, \quad |\sigma(t,z)|^2 + |\rho(t,z)|^2 + |B(t,z)|^2 \le K_0 + K_2|z|^2,$$
$$|\eta(t,z)|^2_{L_2(\mathfrak{Z}_0)} + |\xi(t,z)|^2_{L_2(\mathfrak{Z}_1)} \le K_0 + K_2|z|^2, \quad \int_{\mathfrak{Z}_1} |\mathfrak{z}|^2 \nu_1(d\mathfrak{z}) \le K_0$$

for all $z = (x, y) \in \mathbb{R}^{d+d'}$ and $t \in [0, T]$, and we have

(ii)

$$K_1\mathbb{E}|X_0| + K_2\mathbb{E}|X_0|^2 < \infty. \tag{2.1}$$

Note that in (2.1) we use the convention that $0 \times \infty = 0$, i.e., if $K_2 = 0$, then the finiteness of the second moment of $|X_0|$ is not required, and if $K_1 = K_2 = 0$ then Assumption 2.1 (ii) clearly holds.

The following moment estimate is known and can be easily proved by the help of well-known martingale inequalities.

Remark 2.1 If Assumption 2.1(i) holds, then for every $p \in [1, 2]$ and $A \in \mathcal{F}_0$ we have

$$\mathbb{E} \sup_{t\le T} \mathbf{1}_A |Z_t|^p \le N(1 + \mathbb{E}\mathbf{1}_A|Z_0|^p) \tag{2.2}$$

with a constant N depending only on T, K_0, K_1 and K_2.

Proof This estimate can be obtained in the same fashion as in the case of diffusion processes, see Section 2.5 in [17]. For the convenience of the reader we show below that (2.2) holds for any $\mathbb{R}^d$-valued $\mathcal{F}_t$-adapted cadlag process $(Z_t)_{t\in[0,T]}$ satisfying

$$dZ_t = f(t, Z_t)\,dt + g^k(t, Z_t)\,dw_t^k + \int_{\mathfrak{Z}} h^k(t, Z_{t-}, \mathfrak{z})\,\tilde{N}^k(dt, d\mathfrak{z})$$

on $[0, T]$, where f, g^k are Borel-measurable functions on $[0, T] \times \mathbb{R}^d$ with values in $\mathbb{R}^d$, and h^k is a $\mathcal{B}([0, T] \times \mathbb{R}^d \times \mathfrak{Z}_k)$-measurable function on $[0, T] \times \mathbb{R}^d \times \mathfrak{Z}_k$, with values in $\mathbb{R}^d$ for every integer $k \geq 1$, such that for a constant K,

$$|f(t,z)|^2 + \sum_k |g^k(t,z)|^2 + \sum_k |h^k(t,z,\cdot)|^2_{L_2(\mathfrak{Z}_k,\mathcal{Z}_k,\nu_k)} \leq K(1+|z|^2) \tag{2.3}$$

for $t \in [0, T]$ and $z \in \mathbb{R}^d$. Here $(w^k)_{k=1}^{\infty}$ is a sequence of independent $\mathcal{F}_t$-Wiener processes and $(\tilde{N}^k(dt, d\mathfrak{z}))_{k=1}^{\infty}$ is a sequence of independent $\mathcal{F}_t$-Poisson martingale measures on $[0, \infty) \times \mathfrak{Z}_k$ with σ-finite characteristic measures $\nu_k(d\mathfrak{z})$ on separable measurable spaces $(\mathfrak{Z}_k, \mathcal{Z}_k)$.

To prove (2.2) first note that for every stopping time $\tau \leq T$ and every $A \in \mathcal{F}_0$, $p \in [1, 2]$ we have

$$\mathbb{E} \sup_{s\leq\tau} \mathbf{1}_A |Z_s|^p \leq 4(I_0 + I_1 + I_2 + I_3),$$

where $I_0 := \mathbb{E}\mathbf{1}_A|Z_0|^p$,

$$I_1 := \mathbb{E}\sup_{s\leq\tau}\Big|\int_0^s \mathbf{1}_A f(r, Z_r)\,dr\Big|^p \leq N + N\mathbb{E}\left(\int_0^\tau \mathbf{1}_A|Z_r|^2\,dr\right)^{p/2},$$

$$I_2 := \mathbb{E}\sup_{s\leq\tau}\Big|\int_0^s \mathbf{1_A} g^k(r, Z_r)\,dw_r^k\Big|^p \leq N + N\mathbb{E}\left(\int_0^\tau \mathbf{1}_A|Z_r|^2\,dr\right)^{p/2}$$

$$I_3 := \mathbb{E}\sup_{s\leq\tau}\Big|\int_0^s \int_{\mathfrak{Z}} \mathbf{1_A} h^k(r, Z_r, \mathfrak{z})\,\tilde{N}^k(dr, d\mathfrak{z})\Big|^p \leq N + N\mathbb{E}\left(\int_0^\tau \mathbf{1}_A|Z_r|^2\,dr\right)^{p/2}$$

with a constant $N = N(T, K)$, due to (2.3), Jensen's inequality, and the Burkholder-Davis-Gundy inequality for (cadlag) local martingales with continuous Doob-Meyer processes. Define the stopping times $\tau_n := \inf\{s \in [0, T] : u(s) \geq n\} \wedge T$ for integers $n \geq 1$, where

$$u(s) := |Z_0|^2 + \int_0^s |Z_r|^2\,dr.$$

Then using the above estimate with $\tau_n \wedge t$ in place of τ, and with $A \cap \{\omega \in \Omega : \tau_n(\omega) > 0\} \in \mathcal{F}_0$ in place of A, we get

$$\mathbb{E}\sup_{s\le t} \mathbf{1_A}\mathbf{1}_{\tau_n>0}|Z_{s\wedge\tau_n}|^p \le 4\mathbb{E}\mathbf{1}_{\tau_n>0}\mathbf{1}_A|Z_0|^p + N + N\mathbb{E}\left(\int_0^{\tau_n\wedge t} \mathbf{1}_A|Z_r|^2\,dr\right)^{p/2} < \infty$$

with a constant $N = N(T, K)$. Thus, using

$$\mathbb{E}\left(\int_0^{\tau_n\wedge t} \mathbf{1}_A|Z_r|^2\,dr\right)^{p/2} \le \mathbb{E}\sup_{s\le t} \mathbf{1_A}\mathbf{1}_{\tau_n>0}|Z_{s\wedge\tau_n}|^{p/2}\left(\int_0^{\tau_n} \mathbf{1}_A|Z_r|\,dr\right)^{p/2}$$
$$\le \tfrac{1}{2}\mathbb{E}\sup_{s\le t} \mathbf{1_A}\mathbf{1}_{\tau_n>0}|Z_{s\wedge\tau_n}|^p + N\mathbb{E}\left(\int_0^t \mathbf{1}_A|Z_{s\wedge\tau_n}|\,ds\right)^p,$$

we obtain

$$\mathbb{E}\sup_{s\le t} \mathbf{1_A}\mathbf{1}_{\tau_n>0}|Z_{s\wedge\tau_n}|^p \le N + N\mathbb{E}\mathbf{1}_{\tau_n>0}\mathbf{1}_A|Z_0|^p + N\int_0^t \mathbb{E}\sup_{r\le s}\mathbf{1}_{\tau_n>0}\mathbf{1}_A|Z_{r\wedge\tau_n}|^p\,ds < \infty$$

with a constant $N = N(K, T)$, which by Gronwall's lemma yields

$$\mathbb{E}\sup_{s\le t} \mathbf{1_A}\mathbf{1}_{\tau_n>0}|Z_{t\wedge\tau_n}|^p \le N(1 + \mathbb{E}\mathbf{1}_A|Z_0|^p)$$

with a constant $N = N(K, T)$ for integers $n \ge 1$. Letting here $n \to \infty$ we finish the proof of the remark by using Fatou's lemma. □

We make also the following assumption.

Assumption 2.2 We have $\mathbb{E}\gamma_T = 1$, where

$$\gamma_t = \exp\left(-\int_0^t B(s, X_s, Y_s)\,dV_s - \tfrac{1}{2}\int_0^t |B(s, X_s, Y_s)|^2\,ds\right), \quad t \in [0, T]. \tag{2.4}$$

This assumption implies that the measure Q, defined by $dQ = \gamma_T dP$ on $\mathcal{F}$, is a probability measure equivalent to P, and hence by Girsanov's theorem under Q the process

$$\tilde{V}_t = \int_0^t B(s, X_s, Y_s)\,ds + V_t, \quad t \in [0, T], \tag{2.5}$$

is an $\mathcal{F}_t$-Wiener process.

To describe the evolution of the conditional distribution $P_t(dx) = P(X_t \in dx|Y_s, s \le t)$ for $t \in [0, T]$, we introduce the random differential operators

$$\mathcal{L}_t = a_t^{ij}(x)D_{ij} + b_t^i(x)D_i, \quad \mathcal{M}_t^k = \rho_t^{ik}(x)D_i + B_t^k(x), \quad k = 1, 2, \ldots, d',$$

where

$$a_t^{ij}(x) := \tfrac{1}{2}\sum_{k=1}^{d_1}(\sigma_t^{ik}\sigma_t^{jk})(x) + \tfrac{1}{2}\sum_{l=1}^{d'}(\rho_t^{il}\rho_t^{jl})(x), \quad \sigma_t^{ik}(x) := \sigma^{ik}(t,x,Y_t),$$
$$\rho_t^{il}(x) := \rho^{il}(t,x,Y_t),$$
$$b_t^i(x) := b^i(t,x,Y_t),\ B_t^k(x) := B^k(t,x,Y_t)$$

for $\omega \in \Omega$, $t \in [0,T]$, $x = (x^1,\dots,x^d) \in \mathbb{R}^d$, and $D_i = \partial/\partial x^i$, $D_{ij} = \partial^2/(\partial x^i \partial x^j)$ for $i, j = 1, 2\dots, d$. Moreover for every $t \in [0,T]$ and $\mathfrak{z} \in \mathfrak{Z}_1$ we introduce the random operators I_t^ξ and J_t^ξ defined by

$$I_t^\xi \varphi(x,\mathfrak{z}) = \varphi(x + \xi_t(x,\mathfrak{z}),\mathfrak{z}) - \varphi(x,\mathfrak{z}),$$
$$J_t^\xi \phi(x,\mathfrak{z}) = I_t^\xi \phi(x,\mathfrak{z}) - \sum_{i=1}^{d} \xi_t^i(x,\mathfrak{z}) D_i\phi(x,\mathfrak{z}) \tag{2.6}$$

for functions $\varphi = \varphi(x,\mathfrak{z})$ and $\phi = \phi(x,\mathfrak{z})$ of $x \in \mathbb{R}^d$ and $\mathfrak{z} \in \mathfrak{Z}_1$, and furthermore the random operators I_t^η and J_t^η, defined as I_t^ξ and J_t^ξ, respectively, with $\eta_t(x,\mathfrak{z})$ in place of $\xi_t(x,\mathfrak{z})$, where

$$\xi_t(x,\mathfrak{z}_1) := \xi(t,x,Y_{t-},\mathfrak{z}_1), \quad \eta_t(x,\mathfrak{z}_0) := \eta(t,x,Y_{t-},\mathfrak{z}_0)$$

for $\omega \in \Omega$, $t \in [0,T]$, $x \in \mathbb{R}^d$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$ for $i = 0, 1$ ($Y_{0-} := Y_0$).

Now we are in the position to formulate our main result. Recall that we denote by $(\mathcal{F}_t^Y)_{t\in[0,T]}$ the completed filtration generated by $(Y_t)_{t\in[0,T]}$.

Theorem 2.1 *Let Assumptions 2.1 and 2.2 hold. Then there exist measure-valued $\mathcal{F}_t^Y$-adapted weakly cadlag processes $(P_t)_{t\in[0,T]}$ and $(\mu_t)_{t\in[0,T]}$ such that*

$$P_t(\varphi) = \mu_t(\varphi)/\mu_t(\mathbf{1}), \quad \textit{for } \omega \in \Omega,\ t \in [0,T],$$
$$P_t(\varphi) = \mathbb{E}(\varphi(X_t)|\mathcal{F}_t^Y), \quad \mu_t(\varphi) = \mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t)|\mathcal{F}_t^Y)$$
$$\textit{(a.s.) for each } t \in [0,T],$$

for bounded Borel functions φ on $\mathbb{R}^d$, and for every $\varphi \in C_b^2(\mathbb{R}^d)$ almost surely

$$\mu_t(\varphi) = \mu_0(\varphi) + \int_0^t \mu_s(\mathcal{L}_s\varphi)\,ds + \int_0^t \mu_s(\mathcal{M}_s^k\varphi)\,d\tilde{V}_s^k + \int_0^t\int_{\mathfrak{Z}_0} \mu_s(J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds$$
$$+ \int_0^t\int_{\mathfrak{Z}_1} \mu_s(J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds + \int_0^t\int_{\mathfrak{Z}_1} \mu_{s-}(I_s^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z},ds), \tag{2.7}$$

and

$$P_t(\varphi) = P_0(\varphi) + \int_0^t P_s(\mathcal{L}_s\varphi)\,ds + \int_0^t \Big(P_s(\mathcal{M}_s^k\varphi) - P_s(\varphi)P_s(B_s^k)\Big)\,d\bar{V}_s^k$$
$$+ \int_0^t \int_{\mathfrak{Z}_0} P_s(J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds + \int_0^t \int_{\mathfrak{Z}_1} P_s(J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds \tag{2.8}$$
$$+ \int_0^t \int_{\mathfrak{Z}_1} P_{s-}(I_s^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z}, ds)$$

for all $t \in [0, T]$, *where* $(\tilde{V}_t)_{t\in[0,T]}$ *is given in* (2.5), *and the process* $(\bar{V}_t)_{t\in[0,T]}$ *is defined by*

$$d\bar{V}_t = d\tilde{V}_t - P_t(B_t)\,dt = dV_t + (B_t(X_t) - P_t(B_t))\,dt, \quad \bar{V}_0 = 0.$$

Remark 2.2 Clearly, $\bar{V} = (\bar{V}_t)_{t\in[0,T]}$ is a continuous process, starting from zero, and by the help of Lemma 4.2 below it is easy to see that it is $\mathcal{F}_t^Y$-adapted. Moreover, it is not difficult to see that $\bar{V}$ is a martingale (under P) with respect to $(\mathcal{F}_t)_{t\geq 0}$, with quadratic variation process $[\bar{V}]_t = t$, $t \in [0, T]$. Hence by Lévy's theorem, $\bar{V}$ is an $\mathcal{F}_t^Y$-Wiener process. It is called the *innovation process* in the case when the observation process does not have a stochastic integral component with respect to Poisson measures, i.e., when $\nu_1 = 0$. In this case it was conjectured that $(\bar{V}_s)_{s\in[0,t]}$ together with Y_0 carry the same information as the observation $(Y_s)_{s\in[0,t]}$, i.e., that the σ-algebra generated by $(\bar{V}_s)_{s\in[0,t]}$ and Y_0 coincides with the σ-algebra generated by $(Y_s)_{s\in[0,t]}$ for every t. An affirmative result concerning this conjecture, under quite general conditions on the filtering models (but without jump components) was proved in [19] and [14]. For our filtering model we conjecture that $(\bar{V}_s)_{s\in[0,t]}$, together with Y_0 and $\{\tilde{N}((0, s] \times \Gamma) : s \in [0, t], \Gamma \in \mathcal{Z}_1\}$ carry the same information as the observation $(Y_s)_{s\in[0,t]}$, if Assumption 2.1 holds and the coefficients of (1.1) satisfy an appropriate Lipschitz condition.

Remark 2.3 One can show, see the proof of Lemma 4.2, that $\tilde{N}^Y$, the compensated measure of jumps of the observation process Y, is almost surely the same as $\tilde{N}_1$,

$$Y_t^d := \int_0^t \int_{\mathfrak{Z}_1} \mathfrak{z}\,\tilde{N}^Y(ds, d\mathfrak{z}) = \int_0^t \int_{\mathfrak{Z}_1} \mathfrak{z}\,\tilde{N}_1(ds, d\mathfrak{z}), \quad t \in [0, T]$$

is the purely discontinuous component of Y, and

$$Y_t^c := Y_t - Y_t^d = Y_0 + \tilde{V}_t, \quad t \in [0, T]$$

is the continuous component of Y. Thus Eq. (2.7) can also be written as

$$\mu_t(\varphi) = \mu_0(\varphi) + \int_0^t \mu_s(\mathcal{L}_s\varphi)\,ds + \int_0^t \mu_s(\mathcal{M}_s^k\varphi)\,dY_s^{c\,k} + \int_0^t \int_{\mathfrak{Z}_0} \mu_s(J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds$$
$$+ \int_0^t \int_{\mathfrak{Z}_1} \mu_s(J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds + \int_0^t \int_{\mathfrak{Z}_1} \mu_{s-}(I_s^\xi\varphi)\,\tilde{N}^Y(d\mathfrak{z}, ds), \quad t \in [0, T],$$

where Y_t^{ck} denotes the k-th coordinate of Y^c. Notice that in contrast with the filtering equation of partially observed (continuous) diffusion processes, here not the whole observation process, but its continuous and purely discontinuous components, separately, drive the above equation. Thus it is problematic to approximate this equation only on the basis of discrete time observations, since in general one cannot effectively approximate Y^c and Y^d separately from discrete time observations.

We will prove Theorem 2.1 by deducing Eq. (2.8) from Eq. (2.7), which we obtain by taking, under Q, the conditional expectation of the terms in the equation for $\gamma_t^{-1}\varphi(X_t)$, given the observation $\{Y_s : s \le t\}$.

There are several known conditions ensuring that Assumption 2.2 is satisfied. For a simple proof for the well-known Novikov condition and Kazamaki condition, and their generalizations we refer to Exercise 6.8.VI in [20–22]. These conditions, are clearly satisfied if $|B|$ is bounded, but it does not seem to be easy to reformulate them in terms of the coefficients of the system of Eq. (1.1), if $|B|$ is unbounded. Here we give a condition, which together with Assumption 2.1(i) ensures that Assumption 2.2 holds.

Assumption 2.3 There is a constant K such that

$$-x^i\rho^{ik}(t,z)B^k(t,z) \le K(1+|z|^2) \quad \text{for all } t\in[0,T],\, z=(x,y)\in\mathbb{R}^{d+d'}.$$

Remark 2.4 Define the $\mathbb{R}^{(d+d')\times d'}$-valued function $\hat\rho$ by $\hat\rho^{jk} := \rho^{jk}$ for $j = 1,2,\dots,d, k=1,2,\dots,d'$ and $\hat\rho^{jk} := 0$ for $j=d+1,\dots,d+d', k=1,2,\dots,d'$. Then Assumption 2.3 means that the "one-sided linear growth" condition

$$zf(t,z) \le K(1+|z|^2), \quad t\in[0,T],\ z\in\mathbb{R}^{d+d'},$$

holds for the $\mathbb{R}^{d+d'}$-valued function $f=-\hat\rho B$, where zf denotes the standard inner product of the vectors $z, f\in\mathbb{R}^{d+d'}$. Clearly, this condition is essentially weaker than the linear growth condition on f (in $z\in\mathbb{R}^{d+d'}$), which obviously holds if one of the functions ρ and B is bounded in magnitude and the other satisfies the linear growth condition in Assumption 2.1 (i).

Theorem 2.2 *Let Assumptions 2.1(i) and 2.3 hold. Then $\mathbb{E}\gamma_T = 1$, i.e., Assumption 2.2 holds.*

Proof We want to prove $\mathbb{E}(\gamma_T\mathbf{1}_{|Z_0|\le R}) = P(|Z_0|\le R)$ for every constant $R>0$, since by monotone convergence it implies

$$\mathbb{E}\gamma_T = \lim_{R\to\infty}\mathbb{E}(\gamma_T\mathbf{1}_{|Z_0|\le R}) = \lim_{R\to\infty}P(|Z_0|\le R) = 1.$$

To this end we fix a constant $R>0$ and set $\bar\gamma_t := \gamma_t\mathbf{1}_{|Z_0|\le R}$. By Itô's formula

$$d\bar\gamma_t = -\bar\gamma_t B(t,Z_t)\,dV_t,$$

that shows that $\bar{\gamma}$ is a local $\mathcal{F}_t$-martingale. Thus $\mathbb{E}\bar{\gamma}_{T\wedge\tau_n} = P(|Z_0| \leq R)$ for an increasing sequence $(\tau_n)_{n=1}^{\infty}$ of stopping times τ_n such that τ_n converges to ∞ as $n \to \infty$, and $(\bar{\gamma}_{t\wedge\tau_n})_{t\in[0,T]}$ is a martingale for every n. Consequently, if we can show $\mathbb{E}\sup_{t\leq T}\bar{\gamma}_t < \infty$, then we can use Lebesgue's theorem on dominated convergence to get $\mathbb{E}\bar{\gamma}_T = P(|Z_0| \leq R)$. Define the stopping times

$$\tau_n = \inf\{t \in [0,T] : [\bar{\gamma}]_t \geq n\} \quad \text{for integers } n \geq 1,$$

where

$$[\bar{\gamma}]_t = \int_0^t \bar{\gamma}_s^2 |B(s, Z_s)|^2\, ds.$$

Note that $(\bar{\gamma}_t - \bar{\gamma}_0)_{t\leq T}$ is a continuous local martingale starting from 0, with quadratic variation process $[\bar{\gamma}]$. Clearly,

$$\sup_{t\leq T}\bar{\gamma}_{t\wedge\tau_n} \leq \bar{\gamma}_0 + \sup_{t\leq T}|\bar{\gamma}_{t\wedge\tau_n} - \bar{\gamma}_0| \leq 1 + \sup_{t\leq T}|\bar{\gamma}_{t\wedge\tau_n} - \bar{\gamma}_0|.$$

Thus, by standard estimates, using Davis' and Young's inequalities, we have

$$\begin{aligned}\mathbb{E}\sup_{t\leq T}\bar{\gamma}_{t\wedge\tau_n} \leq 1 + 3\mathbb{E}[\bar{\gamma}]_{T\wedge\tau_n}^{1/2} &\leq 1 + 3\mathbb{E}\sup_{t\leq T}\bar{\gamma}_{t\wedge\tau_n}^{1/2}\left(\int_0^{T\wedge\tau_n}\bar{\gamma}_t|B(t,Z_t)|^2\,dt\right)^{1/2}\\ &\leq 1 + \tfrac{1}{2}\mathbb{E}\sup_{t\leq T}\bar{\gamma}_{t\wedge\tau_n} + 5\mathbb{E}\int_0^T\bar{\gamma}_t|B(t,Z_t)|^2\,dt,\end{aligned}$$

which, after we subtract $\frac{1}{2}E\sup_{t\leq T}\bar{\gamma}_{t\wedge\tau_n}$ and let $n \to \infty$, by Fatou's lemma gives

$$\tfrac{1}{2}\mathbb{E}\sup_{t\leq T}\bar{\gamma}_t \leq 1 + 5\mathbb{E}\int_0^T\bar{\gamma}_t|B(t,Z_t)|^2\,dt \leq 1 + 5\mathbb{E}\int_0^T\bar{\gamma}_t(K_0 + K_2|Z_t|^2)\,dt.$$

Since $\mathbb{E}\bar{\gamma}_t \leq 1$, to show that the right-hand side of the last inequality is finite we need only prove that if $K_2 \neq 0$ then

$$\sup_{t\leq T}\mathbb{E}\bar{\gamma}_t|Z_t|^2 < \infty. \tag{2.9}$$

To this end we apply Itô's formula to $U_t := \bar{\gamma}_t|Z_t|^2$ and use Assumptions 2.1 (ii) and 2.3 to get

$$\begin{aligned}dU_t &= \bar{\gamma}_t(2X_t b(t,Z_t) + 2Y_t B(t,Z_t) + |\sigma(t,Z_t)|^2 + |\rho(t,Z_t)|^2 + 1)\,dt\\ &\quad - 2\bar{\gamma}_t(X_t\rho(t,Z_t)B(t,Z_t) + Y_t B_t(t,Z_t))\,dt + \bar{\gamma}_t\int_{\mathfrak{Z}_0}|\eta(t,Z_t,\mathfrak{z})|^2\,\nu_0(d\mathfrak{z})dt\\ &\quad + \bar{\gamma}_t\int_{\mathfrak{Z}_1}|\xi(t,Z_t,\mathfrak{z})|^2\,\nu_1(d\mathfrak{z})dt + \bar{\gamma}_t\int_{\mathfrak{Z}_1}|\mathfrak{z}|^2\,\nu_1(d\mathfrak{z})dt + dm_t\\ &\leq N\bar{\gamma}_t\,dt + NU_t\,dt + dm_t\end{aligned} \tag{2.10}$$

with a constant N and a cadlag local martingale m starting from zero. Hence by a standard stopping time argument and Gronwall's inequality we get a constant N such that

$$\sup_{t\leq T}\mathbb{E}U_{t\wedge\tau_n}\leq N(1+\mathbb{E}(\mathbf{1}_{|Z_0|\leq R}|Z_0|^2))<\infty$$

for an increasing sequence of stopping times $\tau_n\uparrow\infty$. Letting here $n\to\infty$ by Fatou's lemma we get (2.9), which finishes the proof of the theorem. □

3 Preliminaries

The following lemma is our main tool for calculating conditional expectations of Lebesgue and Itô stochastic integrals of simple processes under Q given $\mathcal{F}_t^Y$.

Lemma 3.1 *Let X and Y be random variables such that $\mathbb{E}|X|<\infty$, $\mathbb{E}|Y|<\infty$ and $\mathbb{E}|XY|<\infty$. Let $\mathcal{G}^1$, $\mathcal{G}^2$ and $\mathcal{G}$ be σ-algebras of events such that $\mathcal{G}^1\subset\mathcal{G}$, $\mathcal{G}^2$ is independent of $\mathcal{G}$, X is $\mathcal{G}$-measurable and Y is independent of $\mathcal{G}\vee\mathcal{G}^2$. Then almost surely*

$$\mathbb{E}(XY|\mathcal{G}^1\vee\mathcal{G}^2)=\mathbb{E}(X|\mathcal{G}^1)\mathbb{E}Y.$$

Proof The right-hand side of the above equation is a $\mathcal{G}^1$-measurable random variable, hence it is obviously $\mathcal{G}^1\vee\mathcal{G}^2$-measurable. Let $\mathcal{H}$ denote the family of $G\in\mathcal{G}^1\vee\mathcal{G}^2$ such that

$$\mathbb{E}Y\mathbb{E}(\mathbb{E}(X|\mathcal{G}^1)\mathbf{1}_G)=\mathbb{E}(XY\mathbf{1}_G).$$

Then $\mathcal{H}$ is a λ-system, and for $G=G_1\cap G_2$, $G_i\in\mathcal{G}^i$ we have

$$\begin{aligned}\mathbb{E}Y\mathbb{E}(\mathbb{E}(X|\mathcal{G}^1)\mathbf{1}_G)&=\mathbb{E}Y\mathbb{E}(\mathbb{E}(\mathbf{1}_{G_1}X|\mathcal{G}^1)\mathbf{1}_{G_2}))=\mathbb{E}Y\mathbb{E}(\mathbb{E}(\mathbf{1}_{G_1}X|\mathcal{G}^1))\mathbb{E}\mathbf{1}_{G_2}\\&=\mathbb{E}Y\mathbb{E}(\mathbf{1}_{G_1}X)\mathbb{E}\mathbf{1}_{G_2}=\mathbb{E}(XY\mathbf{1}_G),\end{aligned}$$

that shows that $\mathcal{H}$ contains the π-system $\{G_1\cap G_2: G_i\in\mathcal{G}^i, i=1,2\}$. Hence, by Dynkin's monotone class lemma $\mathcal{H}=\mathcal{G}^1\vee\mathcal{G}^2$, which completes the proof. □

To formulate a theorem on conditional expectations of Lebesgue and Itô integrals we consider a complete filtered probability space $(\Omega,\mathcal{F},P,\mathcal{F}_t)$ carrying independent $\mathcal{F}_t$-Wiener processes $W^i=(W_t^i)_{t\geq0}$ and independent $\mathcal{F}_t$-Poisson random measures $N_i=N_i(d\mathfrak{z},dt)$ with σ-finite characteristic measures ν_i on separable measurable spaces $(\mathfrak{Z}_i,\mathcal{Z}_i)$ for $i=0,1$, respectively. We denote by $\mathcal{G}_t$ the P-completion of the σ-algebra generated by the events of a σ-algebra $\mathcal{Y}_0\subset\mathcal{F}_0$ together with the random variables W_s^1 and $N_1((0,s]\times\Gamma)$ for $s\leq t$ and $\Gamma\in\mathcal{Z}_1$ such that $\nu_1(\Gamma)<\infty$. The predictable σ-algebras on $\Omega\times[0,T]$, relative to $(\mathcal{F}_t)_{t\in[0,T]}$ and $(\mathcal{G}_t)_{t\in[0,T]}$ are denoted by $\mathcal{P}_{\mathcal{F}}$ and $\mathcal{P}_{\mathcal{G}}$, respectively. The optional σ-algebras relative to $(\mathcal{F}_t)_{t\in[0,T]}$ and $(\mathcal{G}_t)_{t\in[0,T]}$ are denoted by $\mathcal{O}_{\mathcal{F}}$ and $\mathcal{O}_{\mathcal{G}}$, respectively.

The following definition will be frequently used.

Definition 3.1 Given a probability space $(\Omega,\mathcal{F},P)$ and a sub-σ-algebra $\mathcal{G}\subset\mathcal{F}$, we say that a random variable f is σ-integrable (with respect to P) relative to $\mathcal{G}$, if there exists an increasing sequence $(\Omega_n)_{n=1}^\infty$ such that $\bigcup_n\Omega_n=\Omega$, $\Omega_n\in\mathcal{G}$and $\mathbb{E}|f\mathbf{1}_{\Omega_n}|<\infty$ for all n.

One knows that for nonnegative random variables f and σ-algebras $\mathcal{G} \subset \mathcal{F}$ the conditional expectation $\mathbb{E}(f|\mathcal{G})$ is well-defined, and that for general random variables f we define the extended conditional expectation $\mathbb{E}(f|\mathcal{G})$ as $\mathbb{E}(f^+|\mathcal{G}) - \mathbb{E}(f^-|\mathcal{G})$ if $\mathbb{E}(|f||\mathcal{G}) < \infty$ (a.s.). It is not difficult to show that $\mathbb{E}(|f||\mathcal{G}) < \infty$ (a.s.) if and only if f is σ-integrable (with respect to P) relative to $\mathcal{G}$.

We consider real-valued $\mathcal{F} \otimes \mathcal{B}([0, T])$-measurable $\mathcal{F}_t$-adapted processes $f = (f_t)_{t\in[0,T]}$ and $g = (g_t)_{t\in[0,T]}$ on $\Omega \times [0, T]$, real-valued $\mathcal{F} \otimes \mathcal{B}([0, T]) \otimes \mathcal{Z}_i$-measurable functions $h^{(i)} = h_t^{(i)}(\omega, \mathfrak{z})$ of $(\omega, t, \mathfrak{z}) \in \Omega \times [0, T] \times \mathfrak{Z}_i$ for $i = 0, 1$, and a real-valued $\mathcal{F} \otimes \mathcal{B}([0, T]) \otimes \mathcal{Z}$-measurable function $h = h_t(\omega, \mathfrak{z})$ of $(\omega, t, \mathfrak{z}) \in \Omega \times [0, T] \times \mathfrak{Z}$, such that for every $t \in [0, T]$ the functions $h_t^{(i)}$ and h_t are $\mathcal{F}_t \otimes \mathcal{Z}_i$-measurable and $\mathcal{F}_t \otimes \mathcal{Z}$-measurable, respectively, for $i = 0, 1$, where $(\mathfrak{Z}, \mathcal{Z})$ is a separable measurable space, equipped with a σ-finite measure ν. Assume that almost surely

$$F := \left(\int_0^T |f_s|^2\, ds\right)^{1/2} < \infty \quad H^{(i)} := \left(\int_0^T \int_{\mathfrak{Z}_i} |h_s^{(i)}(\mathfrak{z})|^2\, \nu_i(d\mathfrak{z})ds\right)^{1/2} < \infty \tag{3.1}$$

$$G := \int_0^T |g_s|\, ds < \infty, \quad H := \int_0^T \int_{\mathfrak{Z}} |h_s(\mathfrak{z})|\, \nu(d\mathfrak{z})ds < \infty \tag{3.2}$$

for $i = 0, 1$. Then the processes

$$\alpha_t := \int_0^t g_s\, ds, \quad \delta_t := \int_0^t \int_{\mathfrak{Z}} h_s(\mathfrak{z})\, \nu(d\mathfrak{z})ds, \quad t \in [0, T],$$

and

$$\beta_t^{(i)} = \int_0^t f_s\, dW_s^i, \quad \delta_t^{(i)} = \int_0^t \int_{\mathfrak{Z}_i} h_s^{(i)}(\mathfrak{z})\, \tilde{N}_i(d\mathfrak{z}, ds), \quad t \in [0, T], \tag{3.3}$$

are well-defined for $i = 0, 1$, and we have the following theorem.

Theorem 3.2 *Assume the random variables F^r, G, H and $|H^{(i)}|^2$ for $i = 0, 1$, for some $r > 1$ are σ-integrable (with respect to P) relative to $\mathcal{G}_0$. Then for $t \in [0, T]$ we have*

$$\mathbb{E}(\beta_t^{(1)}|\mathcal{G}_t) = \int_0^t \hat{f}_s\, dW_s^1, \quad \mathbb{E}(\beta_t^{(0)}|\mathcal{G}_t) = 0, \tag{3.4}$$

$$\mathbb{E}(\alpha_t|\mathcal{G}_t) = \int_0^t \hat{g}_s\, ds, \quad \mathbb{E}(\delta_t|\mathcal{G}_t) = \int_0^t \int_{\mathfrak{Z}} \hat{h}_s(\mathfrak{z})\, \nu(d\mathfrak{z})ds, \tag{3.5}$$

$$\mathbb{E}(\delta_t^{(1)}|\mathcal{G}_t) = \int_0^t \int_{\mathfrak{Z}_1} \hat{h}_s^{(1)}(\mathfrak{z})\, \tilde{N}_1(d\mathfrak{z}, ds), \quad \mathbb{E}(\delta_t^{(0)}|\mathcal{G}_t) = 0 \tag{3.6}$$

almost surely for some $\mathcal{P}_{\mathcal{G}}$-measurable functions $\hat{f}$ and $\hat{g}$ on $\Omega \times [0, T]$, a $\mathcal{P}_{\mathcal{G}} \otimes \mathcal{Z}_1$-measurable function $\hat{h}^{(1)}$ on $\Omega \times [0, T] \times \mathfrak{Z}_1$, and a $\mathcal{P}_{\mathcal{G}} \otimes \mathcal{Z}$-measurable function $\hat{h}$

on $\Omega \times [0, T] \times \mathfrak{Z}$ *such that*

$$\hat{f}_t = \mathbb{E}(f_t|\mathcal{G}_t), \quad \hat{g}_t = \mathbb{E}(g_t|\mathcal{G}_t) \quad (a.s.) \text{ for } dt - a.\ e.\ t \in [0, T], \tag{3.7}$$

$$\hat{h}_t^{(1)} = \mathbb{E}(h_t^{(1)}(\mathfrak{z})|\mathcal{G}_t) \quad (a.s.) \text{for } dt \otimes \nu_1 - a.e.\ (t, \mathfrak{z}) \in [0, T] \times \mathfrak{Z}_1, \tag{3.8}$$

$$\hat{h}_t = \mathbb{E}(h_t(\mathfrak{z})|\mathcal{G}_t) \quad (a.s.) \text{ for } dt \otimes \nu - a.e.\ (t, \mathfrak{z}) \in [0, T] \times \mathfrak{Z}. \tag{3.9}$$

Proof Since F^r is σ-integrable with respect to $\mathcal{G}_0$, there is an increasing sequence $\Omega_n \in \mathcal{G}_0$ such that $\bigcup_{n=1}^{\infty} \Omega_n = \Omega$ and $\mathbb{E}(\mathbf{1}_{\Omega_n} F^r) < \infty$ for every integer $n \geq 1$. By the definition and elementary properties of (extended) conditional expectations and stochastic integrals, we have

$$\mathbf{1}_{\Omega_n}\mathbb{E}\left(\int_0^t f_s\, dW_s^i \Big| \mathcal{G}_t\right) = \mathbb{E}\left(\mathbf{1}_{\Omega_n}\int_0^t f_s\, dW_s^i \Big| \mathcal{G}_t\right) = \mathbb{E}\left(\int_0^t \mathbf{1}_{\Omega_n} f_s\, dW_s^i \Big| \mathcal{G}_t\right),$$

$$\mathbf{1}_{\Omega_n}\mathbb{E}(f_t|\mathcal{G}_t) = \mathbb{E}(\mathbf{1}_{\Omega_n} f_t|\mathcal{G}_t), \quad t \in [0, T]$$

for $i = 0, 1$ and every $n \geq 1$. Thus, taking $\mathbf{1}_{\Omega_n} f$ in place of f, we may assume that $\mathbb{E}F^r < \infty$. Similarly, we may also assume that $\mathbb{E}G$, $\mathbb{E}H$ and $\mathbb{E}|H^{(i)}|^2$ are finite in what follows below. Assume first that f belongs to $\mathcal{H}_0$, the set of simple processes of the form

$$f_t = \sum_{i=0}^{k-1} \xi_i \mathbb{1}_{(t_i, t_{i+1}]}(t), \tag{3.10}$$

where $0 = t_0 \leq \cdots \leq t_k = T$ are deterministic time instants, and ξ_i is a bounded $\mathcal{F}_{t_i}$-measurable random variable for every $i = 0, 1, \ldots, k-1$ for an integer $k \geq 1$. Then we have

$$\mathbb{E}\left(\int_0^t f_s\, dW_s^1 \Big| \mathcal{G}_t\right) = \sum_i \mathbb{E}\left(\xi_i (W^1_{t_{i+1}\wedge t} - W^1_{t_i \wedge t}) \big| \mathcal{G}_t\right), \quad \text{for } t \in [0, T]. \tag{3.11}$$

For $0 \leq r \leq s \leq T$ define the σ-algebra

$$\mathcal{G}_{r,s} = \sigma(W_v^1 - W_u^1, N_1(\Gamma \times (u, v]) : r \leq u \leq v \leq s, \Gamma \in \mathcal{Z}_1, \nu_1(\Gamma) < \infty).$$

Then σ-algebras $\mathcal{G}_r$ and $\mathcal{G}_{r,s}$ are independent and $\mathcal{G}_s = \mathcal{G}_r \vee \mathcal{G}_{r,s}$. Thus, using Lemma 3.1 with $X := \xi_i$, $Y := 1$, $\mathcal{G}^1 := \mathcal{G}_{t_i}$, $\mathcal{G} := \mathcal{F}_{t_i}$ and $\mathcal{G}^2 := \mathcal{G}_{t_i,s}$ for $t_i \leq s \leq T$, we have

$$\mathbb{E}(\xi_i|\mathcal{G}_s) = \mathbb{E}(\xi_i|\mathcal{G}_{t_i}) \quad \text{for } i = 0, 1, 2, \ldots, k-1. \tag{3.12}$$

Hence for $t_i \leq s \leq t_{i+1} \leq t \leq T$,

$$\mathbb{E}(\xi_i(W^1_{t_{i+1}} - W^1_{t_i})|\mathcal{G}_t) = \mathbb{E}(\xi_i|\mathcal{G}_t)(W^1_{t_{i+1}} - W^1_{t_i}) = \mathbb{E}(\xi_i|\mathcal{G}_s)(W^1_{t_{i+1}} - W^1_{t_i}) \tag{3.13}$$

and for $t_j \leq s \leq t \leq T$,

$$\mathbb{E}(\xi_j(W^1_t - W^1_{t_j})|\mathcal{G}_t) = \mathbb{E}(\xi_j|\mathcal{G}_t)(W^1_t - W^1_{t_j}) = \mathbb{E}(\xi_j|\mathcal{G}_s)(W^1_t - W^1_{t_j}). \tag{3.14}$$

Consequently, defining $\hat{f}_s = \mathbb{E}(\xi_i|\mathcal{G}_s) = \mathbb{E}(f_s|\mathcal{G}_s)$ for $s \in (t_i, t_{i+1}]$, $i = 0, 1, \dots, k-1$, the function $\hat{f}$ on $\Omega \times [0, T]$ is $\mathcal{P}_{\mathcal{G}}$-measurable, and using (3.11) we can see that the first equation in (3.4) holds. Assume now that f is $\mathcal{F} \otimes \mathcal{B}([0, T])$-measurable and $\mathcal{F}_t$-adapted such that $\mathbb{E}F^r < \infty$. Then there are sequences $(f^n)_{n=1}^{\infty}$ and $(\hat{f}^n)_{n=1}^{\infty}$ such that $f^n \in \mathcal{H}_0$, $\hat{f}^n$ is $\mathcal{P}_{\mathcal{G}}$-measurable,

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T |f_t - f^n_t|^2\, dt\right)^{r/2} = 0, \tag{3.15}$$

and almost surely

$$\mathbb{E}(I_t(f^n)|\mathcal{G}_t) := \mathbb{E}\left(\int_0^t f^n_s\, dW^1_s \Big| \mathcal{G}_t\right) = \int_0^t \hat{f}^n_s\, dW^1_s =: I_t(\hat{f}^n) \quad \text{for all } t \in [0, T], \tag{3.16}$$

$$\hat{f}^n_t = \mathbb{E}(f^n_t|\mathcal{G}_t) \quad \text{for } dt - \text{a.e. } t \in [0, T] \tag{3.17}$$

for all $n \geq 1$. Using the Davis inequality, Doob's inequality, Jensen's and Burkholder's inequalities for any $r > 1$ we have

$$\begin{aligned}
\mathbb{E}\left(\int_0^T |\hat{f}^n_t - \hat{f}^m_t|^2\, dt\right)^{1/2} &\leq 3\mathbb{E}\sup_{t\leq T} |I_t(\hat{f}^n - \hat{f}^m)| \\
&= 3\mathbb{E} \sup_{t\in[0,T]\cap\mathbb{Q}} |\mathbb{E}\left(I_t(f^n - f^m)\big|\mathcal{G}_t\right)| \leq 3\mathbb{E} \sup_{t\in[0,T]\cap\mathbb{Q}} (\mathbb{E}(\sup_{s\leq T} |I_s(f^n - f^m)||\mathcal{G}_t) \\
&\leq 3\frac{r}{r-1}\left(\mathbb{E}\sup_{t\leq T} |I_t(f^n - f^m)|^r\right)^{1/r} \leq N\left(\mathbb{E}\left(\int_0^T |f^n_t - f^m_t|^2\, dt\right)^{r/2}\right)^{1/r},
\end{aligned}$$

where $\mathbb{Q}$ is the set of rational numbers and $N = N(r)$ is a constant, which gives

$$\lim_{n,m\to\infty} \mathbb{E}\left(\int_0^T |\hat{f}^n_t - \hat{f}^m_t|^2\, dt\right)^{1/2} = 0.$$

Thus there exists a $\mathcal{P}_{\mathcal{G}}$-measurable function $\hat{f}$ on $\Omega \times [0, T]$, such that

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T |\hat{f}_t - \hat{f}_t^n|^2\,dt\right)^{1/2} = 0, \tag{3.18}$$

which implies

$$\lim_{n\to\infty} \mathbb{E} \sup_{t\in[0,T]} |I_t(\hat{f}) - I_t(\hat{f}^n)| = 0. \tag{3.19}$$

Using Jensen's and Davis' inequalities again we have

$$\begin{aligned} &\mathbb{E}|\mathbb{E}(I_t(f)|\mathcal{G}_t) - \mathbb{E}(I_t(f^n)|\mathcal{G}_t)| \le \mathbb{E}\mathbb{E}(|I_t(f - f^n)||\mathcal{G}_t) \\ &= \mathbb{E}|I_t(f - f^n)| \le 3\mathbb{E}\left(\int_0^T |f_t - f_t^n|^2\,dt\right)^{1/2} \quad \text{for every } t \in [0, T], \end{aligned}$$

i.e., for $n \to \infty$

$$\mathbb{E}(I_t(f^n)|\mathcal{G}_t) \to \mathbb{E}(I_t(f)|\mathcal{G}_t) \quad \text{in } L_1(\Omega) \text{ for every } t \in [0, T]. \tag{3.20}$$

Thus letting $n \to \infty$ in Eq. (3.16), by virtue of (3.19) and (3.20) we get the first equation in (3.4). Clearly, (3.15) and (3.18) imply

$$\lim_{n\to\infty} \int_0^T \mathbb{E}|f_t - f_t^n| + \mathbb{E}|\hat{f}_t - \hat{f}_t^n|\,dt = 0.$$

Hence there is a subsequence $n_l \to \infty$ and a set $S \in \mathcal{B}([0, T])$ of Lebesgue measure 0 such that for $n_l \to \infty$,

$$f_t^{n_l} \to f_t \text{and a} \hat{f}_t^{n_l} \to \hat{f}_t \text{ in } L_1(\Omega) \text{ for each } t \in [0, T]\setminus S =: S^c,$$

and taking into account (3.17), we can assume that S is a dt-zero set such that we also have $\hat{f}_t^{n_l} = \mathbb{E}(f_t^{n_l}|\mathcal{G}_t)$ (a.s.) for every $t \in S^c$. Thus for $n_l \to \infty$ we have $\mathbb{E}(f_t^{n_l}|\mathcal{G}_t) \to \mathbb{E}(f_t|\mathcal{G}_t)$ in $L_1(\Omega)$ for each $t \in S^c$, which gives

$$\hat{f}_t = \mathbb{E}(f_t|\mathcal{G}_t) \quad \text{almost surely for every } t \in S^c,$$

i.e., the first equation in (3.7) holds. To prove the second equation in (3.4) we note that for ξ_i from the expression (3.10) we have

$$\mathbb{E}(\xi_i(W^0_{t_{i+1}\wedge t} - W^0_{t_i})|\mathcal{G}_t) = \mathbb{E}(\xi_i|\mathcal{G}_{t_i})\mathbb{E}(W^0_{t_{i+1}\wedge t} - W^0_{t_i}) = 0 \quad \text{for } i = 1, 2, \ldots, N-1, \tag{3.21}$$

by using Lemma 3.1 with $X = \xi_i$, $Y = W^0_{t_{i+1}\wedge t} - W^0_{t_i\wedge t}$ $\mathcal{G}^1 := \mathcal{G}_{t_i} \subset \mathcal{F}_{t_i} =: \mathcal{G}$ and $\mathcal{G}^2 := \mathcal{G}_{t_i,t}$ for $t_i \le t$. Hence we get the second equation in (3.4) for f given in (3.10), and the general case follows by approximation as above. To prove the first equation in (3.5) assume that g is given by the right-hand side of (3.10). Then using (3.12) we can see that

$$\hat{g}_t := \sum_{i=0}^{k-1} \mathbb{E}(\xi_i|\mathcal{G}_{t_i})\mathbf{1}_{(t_i,t_{i+1}]}(t) = \mathbb{E}(g_t|\mathcal{G}_t), \quad t \in [0,T],$$

and that the first equation in (3.5) and the second equation in (3.7) hold. Assume now that g is an $\mathcal{F}\otimes\mathcal{B}([0,T])$-measurable F_t-adapted random process such that $\mathbb{E}G < \infty$. Then there are sequences $(g^n)_{n=1}^{\infty}$ and $(\hat{g}^n)_{n=1}^{\infty}$ such that $g^n \in \mathcal{H}_0$, $\hat{g}^n$ is $\mathcal{P}_{\mathcal{G}}$-measurable,

$$\lim_{n\to\infty} \mathbb{E}\int_0^T |g_t - g_t^n|\,dt = 0,$$

and almost surely

$$\mathbb{E}\left(\int_0^t g_s^n\,ds\Big|\mathcal{G}_t\right) = \int_0^t \hat{g}_s^n\,ds \quad \text{for all } t \in [0,T],$$
$$\hat{g}_t^n = \mathbb{E}(g_t^n|\mathcal{G}_t) \quad \text{for } dt-\text{ a.e. } t \in [0,T].$$

Hence noting that by Tonelli's theorem and Jensen's inequality

$$\mathbb{E}\int_0^T |\hat{g}_t^n - \hat{g}_t^m|\,dt = \int_0^T \mathbb{E}|\mathbb{E}(g_t^n|\mathcal{G}_t) - \mathbb{E}(g_t^m|\mathcal{G}_t)|\,dt$$
$$\le \int_0^T \mathbb{E}\mathbb{E}(|g_t^n - g_t^m||\mathcal{G}_t)\,dt = \mathbb{E}\int_0^T |g_t^n - g_t^m|\,dt,$$

and repeating previous arguments we get a $\mathcal{P}_{\mathcal{F}}$-measurable $\hat{g}$ such that the first equation in (3.5) and the second equation in (3.7) hold. To prove the second equation in (3.5) we assume first that

$$h_t(\mathfrak{z}) = \sum_{i=0}^{k-1} \xi_i \mathbf{1}_{(t_i,t_{i+1}]\times\Gamma_i}(t,\mathfrak{z}), \tag{3.22}$$

for a partition $0 \le t_0 \le t_1 \le \cdots \le t_k = T$ of $[0,T]$, bounded $\mathcal{F}_{t_i}$-measurable random variables ξ_i and sets $\Gamma_i \in \mathcal{Z}$, $\nu(\Gamma_i) < \infty$ for $i = 0,\ldots,k-1$. Then

$$\mathbb{E}\left(\int_0^t \int_{\mathfrak{Z}} h_s(\mathfrak{z})\,\nu(d\mathfrak{z})\,ds\Big|\mathcal{G}_t\right) = \sum_{i=0}^{k-1} \mathbb{E}(\xi_i|\mathcal{G}_t)\nu(\Gamma_i)(t_{i+1}\wedge t - t_i \wedge t), \quad t \in [0,T].$$

Thus, since by virtue of (3.12) we have

$$\hat{h}_t(\mathfrak{z}) := \sum_{i=0}^{k-1} \mathbb{E}(\xi_i|\mathcal{G}_{t_i})\mathbf{1}_{(t_i,t_{i+1}]}(t)\mathbf{1}_{\Gamma_i}(\mathfrak{z}) = \mathbb{E}(h_t(\mathfrak{z})|\mathcal{G}_t), \quad t \in [0,T],\ \mathfrak{z} \in \mathfrak{Z},$$

for $\hat{h}$ the second equation in (3.5) and by definition (3.9) hold. Hence we can get these equations in the general case by a straightforward approximation procedure in the same way as the first equation in (3.5) and the second equation in (3.7) have been proved above.

Now we are going to prove (3.6). Assume first that $h^{(1)}$ is a simple function, given by the right-hand side of equation (3.22) with $\Gamma_i \in \mathcal{Z}_1$, $\nu_1(\Gamma_i) < \infty$, $i = 0, 1, \dots, k-1$. Then

$$\mathbb{E}\left(\int_0^t \int_{\mathfrak{Z}_1} h_s^{(1)}(\mathfrak{z})\, \tilde{N}_1(d\mathfrak{z}, ds)\Big|\mathcal{G}_t\right) = \sum_{i=0}^{k-1} \mathbb{E}\left(\xi_i \tilde{N}_1(\Gamma_i \times (t_i \wedge t, t_{i+1} \wedge t])\Big|\mathcal{G}_t\right).$$

In the same way as Eqs. (3.13) and (3.14) are obtained, by using (3.12) we get

$$\begin{aligned}&\mathbb{E}\left(\xi_i \tilde{N}_1(\Gamma_i \times (t_i, t_{i+1}])\Big|\mathcal{G}_t\right)\\ &\quad= \mathbb{E}(\xi_i|\mathcal{G}_{t_i})\tilde{N}_1(\Gamma_i \times (t_i, t_{i+1}]) = \mathbb{E}(\xi_i|\mathcal{G}_s)\tilde{N}_1(\Gamma_i \times (t_i, t_{i+1}])\end{aligned}$$

for $t_i \le s \le t_{i+1} \le t$, and

$$\mathbb{E}\left(\xi_j \tilde{N}_1(\Gamma_j \times (t_j, t])\Big|\mathcal{G}_t\right) = \mathbb{E}(\xi_j|\mathcal{G}_{t_j})\tilde{N}_1(\Gamma_j \times (t_j, t]) = \mathbb{E}(\xi_j|\mathcal{G}_s)\tilde{N}_1(\Gamma_j \times (t_j, t])$$

for $t_j \le s \le t \le t_{j+1}$. Thus for

$$\hat{h}_t^{(1)}(\mathfrak{z}) = \sum_{i=0}^{k-1} \mathbb{E}(\xi_i|\mathcal{G}_{t_i})\mathbf{1}_{(t_i,t_{i+1}]\times\Gamma_i}(t,\mathfrak{z}) = \mathbb{E}(h_t^{(1)}(\mathfrak{z})|\mathcal{G}_t),$$

equations in (3.6) and (3.8) hold. Assume now that $h^{(1)}$ is $\mathcal{F}\otimes\mathcal{B}([0,T])\otimes\mathcal{Z}$-measurable such that for every $t \in [0,T]$ the function $h_t^{(1)}$ is $\mathcal{F}_t \otimes \mathcal{Z}_1$-measurable and $\mathbb{E}|H^{(1)}|^2 < \infty$, where $H^{(1)}$ is defined in (3.1). Then there exist sequences $(h^n)_{n=1}^{\infty}$ and $(\hat{h}^n)_{n=1}^{\infty}$, such that h^n is a simple function of the form (3.22), $\hat{h}^n$ is a $\mathcal{P}_{\mathcal{G}} \otimes \mathcal{Z}_1$-measurable function,

$$\mathbb{E}(\tilde{I}_t(h^n)|\mathcal{G}_t) := \mathbb{E}\left(\int_0^t \int_{\mathfrak{Z}_1} h_s^n(\mathfrak{z})\, \tilde{N}_1(d\mathfrak{z}, ds)\Big|\mathcal{G}_t\right) = \int_0^t \int_{\mathfrak{Z}_1} \hat{h}_s^n(\mathfrak{z})\, \tilde{N}_1(d\mathfrak{z}, ds), \tag{3.23}$$

$$\hat{h}_t^n(\mathfrak{z}) = \mathbb{E}(h_t^n(\mathfrak{z})|\mathcal{G}_t), \quad \text{almost surely, for } \nu_1(d\mathfrak{z}) \otimes dt - \text{ a.e. } (\mathfrak{z}, t) \in \mathfrak{Z}_1 \times [0,T], \tag{3.24}$$

for every $n \geq 1$, and

$$\lim_{n\to\infty} \mathbb{E} \int_0^T \int_{\mathfrak{Z}_1} |h_t^{(1)}(\mathfrak{z}) - h_t^n(\mathfrak{z})|^2 \, \nu_1(d\mathfrak{z}) \, dt = 0. \tag{3.25}$$

Hence using Jensen's inequality we get

$$\lim_{n,m\to\infty} \mathbb{E} \int_0^T \int_{\mathfrak{Z}_1} |\hat{h}_t^n(\mathfrak{z}) - \hat{h}_t^m(\mathfrak{z})|^2 \, \nu_1(d\mathfrak{z})dt = 0,$$

which implies the existence of a $\mathcal{P}_{\mathcal{G}} \otimes \mathcal{Z}_1$-measurable function $\hat{h}^{(1)}$ such that

$$\lim_{n\to\infty} \mathbb{E} \int_0^T \int_{\mathfrak{Z}_1} |\hat{h}_t^{(1)}(\mathfrak{z}) - \hat{h}_t^n(\mathfrak{z})|^2 \, \nu_1(d\mathfrak{z}) \, dt = 0. \tag{3.26}$$

Thus letting $n \to \infty$ in (3.23) we obtain (3.6). By virtue of (3.25) and (3.26) there is a subsequence $n_l \to \infty$ and a set $A \in \mathcal{B}([0,T]) \otimes \mathcal{Z}_1$ such that $dt \otimes \nu_1(A) = 0$ and for $n_l \to \infty$

$$h_t^{n_l}(\mathfrak{z}) \to h_t^{(1)}(\mathfrak{z}) \quad \text{and} \quad \hat{h}_t^{n_l}(\mathfrak{z}) \to \hat{h}_t^{(1)}(\mathfrak{z}) \text{ in mean square}$$

for every $(t, \mathfrak{z}) \in A^c := [0,T] \times \mathfrak{Z}_1 \backslash A$. Consequently,

$$\mathbb{E}(h_t^{n_k}(\mathfrak{z})|\mathcal{G}_t) \to \mathbb{E}(h_t^{(1)}(\mathfrak{z})|\mathcal{G}_t) \quad \text{in mean square for every } (\mathfrak{z}, t) \in A^c,$$

and letting $n := n_l \to \infty$ in (3.24) we obtain $\mathbb{E}(h_t^{(1)}(\mathfrak{z})|\mathcal{G}_t) = \hat{h}_t^{(1)}(\mathfrak{z})$ for $(\mathfrak{z}, t) \in A^c$, which proves (3.8). To prove the second equation in (3.6) assume first that $h^{(0)}$ is a simple function of the form (3.22) with $\Gamma_i \in \mathcal{Z}_0$, $\nu_0(\Gamma_i) < \infty$ for $i = 0, 1, \dots k-1$. Just like (3.21) is obtained, by Lemma 3.1 we get

$$\mathbb{E}(\xi_i \tilde{N}_0(\Gamma_i \times (t_i \wedge t, t_{i+1} \wedge t])|\mathcal{G}_t) = \mathbb{E}(\xi_i|\mathcal{G}_{t_i})\mathbb{E}\tilde{N}_0(\Gamma_i \times (t_i \wedge t, t_{i+1} \wedge t]) = 0$$

for $i = 0, 1, \dots, k-1$ and $t \in [0,T]$, that implies the second equation in (3.6). Hence, we obtain the second equation in (3.6) for $\mathcal{O}_{\mathcal{F}}$-measurable functions satisfying (3.1) by approximation with simple functions. □

We can reformulate the above theorem by using the notion of optional projection of processes with respect to a given filtration. It is well-known (see for instance [15, Thm 5.1], [9, Thm 2.43]) that if $f = (f_t)_{t\in[0,T]}$ is a $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable process such that f_τ is σ-integrable (with respect to a probability measure P) relative to the σ-algebra $\mathcal{G}_\tau$ for every $\mathcal{G}_t$-stopping time $\tau \leq T$ (with respect to a P-complete filtration $(\mathcal{G}_t)_{t\in[0,T]}$), then there exists a unique (up to evanescence) $\mathcal{G}_t$-optional process ${}^of = ({}^of_t)_{t\in[0,T]}$ such that for every $\mathcal{G}_t$-stopping time $\tau \leq T$

$$\mathbb{E}(f_\tau|\mathcal{G}_\tau) = {}^of_\tau \quad \text{(a.s.)} .$$

The process ${}^{o}f$ is called the optional projection of f (under P with respect to $(\mathcal{G}_t)_{t\in[0,T]}$). If f is a cadlag process such that almost surely $\sup_{t\le T}|f_t|\le\eta$ for some σ-integrable random variable η with respect to P relative to $\mathcal{G}_0$, then almost surely the trajectories of ${}^{o}f$ have left and right limits at every $t\in(0,T]$ and $[0,T)$, respectively, and moreover, they are also almost surely right-continuous if $(\mathcal{G}_t)_{t\in[0,T]}$ is right-continuous. One can define the extended optional projection ${}^{o}f$ of every nonnegative $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable process $(f_t)_{t\in[0,T]}$ (under P with respect to $(\mathcal{G}_t)_{t\in[0,T]}$) by setting ${}^{o}f:=\lim_{n\to\infty}\mathbf{1}_S{}^{o}(f\wedge n)$, where S is the set of $(\omega,t)\in\Omega\times[0,T]$ such that $\lim_{n\to\infty}{}^{o}(f\wedge n)_t(\omega)$ exists, which may be infinite. For $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable stochastic processes $(f_t)_{t\in[0,T]}$ one defines the extended optional projection ${}^{o}f$ by ${}^{o}f={}^{o}(f^+)-{}^{o}(f^-)$ on the set $A=\{(\omega,t)\in\Omega\otimes[0,T]:{}^{o}(f^+)+{}^{o}(f^-)<\infty\}$ and ${}^{o}f=\infty$ on $\Omega\times[0,T]\backslash A$.

Notice that if for a $t\in[0,T]$ the extended conditional expectations $\mathbb{E}(f_t^+|\mathcal{G}_t)$ and $\mathbb{E}(f_t^-|\mathcal{G}_t)$ are almost surely finite, then they are almost surely equal to ${}^{o}(f_t^+)$ and ${}^{o}(f_t^-)$, respectively, meaning that we have ${}^{o}f_t=\mathbb{E}(f_t|\mathcal{G}_t)$ (a.s.). Let $h=(h_t(\mathfrak{z}))$ be an $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{Z}$-measurable function on $\Omega\times\mathbb{R}_+\times\mathfrak{Z}$. Then by the help of the Monotone Class Theorem it is not difficult to show the existence of an $\mathcal{O}_{\mathcal{G}}\otimes\mathcal{Z}$-measurable function, which for each fixed $\mathfrak{z}\in\mathfrak{Z}$ gives the (possibly extended) $\mathcal{O}_{\mathcal{G}}$-optional projection of $h(\mathfrak{z}):=(h_t(\mathfrak{z}))_{t\ge0}$. We denote this function by ${}^{o}h$, and call it the (extended) $\mathcal{O}_{\mathcal{G}}$-optional projection of h.

Corollary 3.3 *Assume the random variables F, $H^{(i)}$ and G, H, defined in (3.1) and (3.2), respectively, are σ-integrable relative to $\mathcal{G}_0$ for $i=0,1$. Assume moreover that almost surely*

$$\int_0^T|{}^{o}f_t|^2dt<\infty,\quad\int_0^T\int_{\mathfrak{Z}_i}|{}^{o}h_t^{(i)}(\mathfrak{z})|^2\,\nu_i(d\mathfrak{z})dt<\infty\quad\textit{for } i=0,1,\tag{3.27}$$

where ${}^{o}f$ and ${}^{o}h^{(i)}$ are the (extended) $\mathcal{O}_{\mathcal{G}}$-optional projections of f and $h^{(i)}$, respectively. Then for every $t\in[0,T]$ Eqs. (3.4), (3.5) and (3.6) hold almost surely with the $\mathcal{O}_{\mathcal{G}}$-optional projections ${}^{o}f$, ${}^{o}g$, ${}^{o}h^{(i)}$ and ${}^{o}h$ in place of $\hat{f}$, $\hat{g}$, $\hat{h}^{(i)}$ and $\hat{h}$, respectively, for $i=0,1$. Moreover, there is a dt-null set $T_0\subset[0,T]$, a $dt\otimes\nu_1$-null set $B_0\subset[0,T]\times\mathfrak{Z}_1$ and a $dt\otimes\nu$-null set $B\subset[0,T]\times\mathfrak{Z}$, such that

(i) *for each $t\in[0,T]\backslash T_0$ the random variable $|f_t|+|g_t|$ is σ-integrable relative to $\mathcal{G}_0$ and*

$$\mathbb{E}(f_t|\mathcal{G}_t)={}^{o}f_t\in\mathbb{R},\ \text{(a.s.)},\quad\mathbb{E}(g_t|\mathcal{G}_t)={}^{o}g_t\in\mathbb{R}\,\text{(a.s.)},\tag{3.28}$$

(ii) *for each $(t,\mathfrak{z})\in[0,T]\times\mathfrak{Z}_1\backslash B_0$ the random variable $|h_t^{(1)}(\mathfrak{z})|$ is σ-integrable relative to $\mathcal{G}_0$ and*

$$\mathbb{E}(h_t^{(1)}(\mathfrak{z})|\mathcal{G}_t)={}^{o}{h_t}^{(1)}(\mathfrak{z})\in\mathbb{R}\,\text{(a.s.)},\tag{3.29}$$

(iii) *for each* $(t, \mathfrak{z}) \in [0, T] \times \mathfrak{Z} \backslash B$ *the random variable* $|h_t(\mathfrak{z})|$ *is* σ*-integrable relative to* $\mathcal{G}_0$*, and*

$$\mathbb{E}(h_t(\mathfrak{z})|\mathcal{G}_t) = {}^{o}h_t(\mathfrak{z}) \in \mathbb{R} \text{ (a.s.)}. \tag{3.30}$$

Proof Just like in the proof of the previous lemma without loss of generality we may and will assume that F, G, H and $H^{(i)}$, $i = 0, 1$, have finite expectation. Thus by Minkowski's inequality and Tonelli's theorem we have

$$\left(\int_0^T (\mathbb{E}|f_t|)^2\,dt\right)^{1/2} \le \mathbb{E}\left(\int_0^T |f_t|^2\,dt\right)^{1/2} < \infty, \quad \int_0^T \mathbb{E}|g_t|\,dt = \mathbb{E}\int_0^T |g_t|\,dt < \infty,$$
$$\int_0^T \int_{\mathfrak{Z}} \mathbb{E}|h_t(\mathfrak{z})|\,\nu(d\mathfrak{z})\,dt = \mathbb{E}\int_0^T \int_{\mathfrak{Z}} |h_t(\mathfrak{z})|\,\nu(d\mathfrak{z})\,dt < \infty$$
$$\left(\int_0^T \int_{\mathfrak{Z}_1} (\mathbb{E}|h_t^{(1)}(\mathfrak{z})|)^2\,\nu_1(d\mathfrak{z})\,dt\right)^{1/2} \le \mathbb{E}\left(\int_0^T \int_{\mathfrak{Z}_1} |h_t^{(1)}(\mathfrak{z})|^2 \nu_1(d\mathfrak{z})\,dt\right)^{1/2} < \infty.$$

Therefore $\mathbb{E}|f_t| + \mathbb{E}|g_t| < \infty$ for dt-almost every $t \in [0, T]$, $\mathbb{E}|h_t(\mathfrak{z})| < \infty$ for $dt \otimes \nu$-a.e. $(t, \mathfrak{z}) \in [0, T] \times \mathfrak{Z}$, and $\mathbb{E}|h_t^{(1)}(\mathfrak{z})| < \infty$ for $dt \otimes \nu_1$-a.e. $(t, \mathfrak{z}) \in [0, T] \times \mathfrak{Z}_1$, i.e., we get (3.28), (3.29) and (3.30). Hence due to (3.7) and (3.9) we have (3.5) with ${}^{o}g$ and ${}^{o}h$ in place of $\hat{g}$ and $\hat{h}$, respectively. We also have (3.4) and (3.6) with ${}^{o}f$ and ${}^{o}h^{(1)}$ in place of $\hat{f}$ and $\hat{h}^{(1)}$, provided F^r and $|H^{(i)}|^2$ are σ-integrable relative to $\mathcal{G}_0$ for $i = 0, 1$ for some $r > 1$. Thus it remains to prove (3.4) and (3.6) with ${}^{o}f$ and ${}^{o}h^{(1)}$ in place of $\hat{f}$ and $\hat{h}^{(1)}$, respectively, under the condition that F and $H^{(i)}$ are σ-integrable relative to $\mathcal{G}_0$ for $i = 0, 1$, and (3.27) holds. We show only (3.6) under these conditions, because (3.4) can be proven similarly. To this end define $h^{(1)n} = \mathbf{1}_{\mathfrak{Z}^n}(-n \vee h^{(1)} \wedge n)$ for integers $n \ge 1$, where $(\mathfrak{Z}^n)_{n=1}^{\infty}$ is an increasing sequence of sets $\mathfrak{Z}^n \in \mathcal{Z}_1$ such that $\bigcup_{n=1}^{\infty} \mathfrak{Z}^n = \mathfrak{Z}_1$ and $\nu_1(\mathfrak{Z}^n) < \infty$ for every $n \ge 1$. Then for each $t \in [0, T]$

$$\mathbb{E}(\delta_t^{(1)n}|\mathcal{G}_t) = \int_0^t \int_{\mathfrak{Z}_1} {}^{o}h_s^{(1)n}(\mathfrak{z})\,\tilde{N}_1(d\mathfrak{z}, ds) \quad \text{(a.s.)}, \tag{3.31}$$

where $\delta^{(1)n}$ is defined as $\delta^{(1)}$ in (3.3), but with $h^{(1)n}$ in place of $h^{(1)}$. Note that

$$|{}^{o}h^{(1)n}| \le |{}^{o}h^{(1)}| \quad P \otimes dt \otimes \nu_1 \text{-almost every } (\omega, t, \mathfrak{z}) \in \Omega \times [0, T] \times \mathfrak{Z}_1,$$

and for $n \to$ we have ${}^{o}h_s^{(1)n}(\mathfrak{z}) \to {}^{o}h_s^{(1)}(\mathfrak{z})$ almost surely for every $(s, \mathfrak{z}) \in [0, T] \times \mathfrak{Z}_1$ such that ${}^{o}h_s^{(1)}(\mathfrak{z}) \neq \infty$. Hence due to condition (3.27), by Lebesgue's theorem on dominated convergence we have

$$\int_0^T \int_{\mathfrak{Z}_1} |{}^{o}h_s^{(1)n}(\mathfrak{z}) - {}^{o}h_s^{(1)}(\mathfrak{z})|^2\,\nu_1(d\mathfrak{z})dt \to 0 \text{ (a.s.)} \text{ as n} \to \infty,$$

which implies

$$\int_0^t \int_{\mathfrak{Z}_1} {}^o h_s^{(1)n}(\mathfrak{z})\, \tilde{N}_1(d\mathfrak{z}, ds) \to \int_0^t \int_{\mathfrak{Z}_1} {}^o h_s^{(1)}(\mathfrak{z})\, \tilde{N}_1(d\mathfrak{z}, ds) \tag{3.32}$$

in probability, uniformly in $t \in [0, T]$. Using obvious properties of conditional expectations, by Davis' inequality and Lebesgue's theorem on dominated convergence we get

$$\begin{aligned} &\lim_{n\to\infty} \mathbb{E}|\mathbb{E}(\delta_t^{(1)n}|\mathcal{G}_t) - \mathbb{E}(\delta_t^{(1)}|\mathcal{G}_t)| \\ &\quad\le \lim_{n\to\infty} \mathbb{E}|\delta_t^{(1)n} - \delta_t^{(1)}| \\ &\quad\le 3 \lim_{n\to\infty} \mathbb{E}\left(\int_0^T \int_{\mathfrak{Z}_1} |h_s^{(1)n}(\mathfrak{z}) - h_s^{(1)}(\mathfrak{z})|^2\ \nu_1(d\mathfrak{z})ds\right)^{1/2} = 0, \end{aligned}$$

which by virtue of (3.31) and (3.32) finishes the proof of the first equation in (3.6). The second equation in (3.6) can be obtained similarly. □

Remark 3.1 We have that almost surely

$$\int_0^t \int_{\mathfrak{Z}_i} |{}^o h_s^{(i)}(\mathfrak{z})|^2\ \nu_i(d\mathfrak{z})\, ds \le \int_0^t {}^o(|h_s^{(i)}|^2_{L_2(\mathfrak{Z}_i)})\, ds \quad \text{(a.s.) for } i = 0, 1$$

for all $t \in [0, T]$. Thus

$$\int_0^T {}^o(|h_t^{(i)}|^2_{L_2(\mathfrak{Z}_i)})\, dt < \infty \quad \text{(a.s.) for } i = 0, 1 \tag{3.33}$$

implies the assumption on $h^{(i)}$ in (3.27).

Proof *(Proof of Remark 3.1)* Let $i \in \{0, 1\}$ be fixed and let $(A_n)_{n=1}^\infty$ be an increasing sequence of sets from $\mathfrak{Z}_i$ such that $\cup_{n=1}^\infty A_n = \mathfrak{Z}_i$ and $\nu_i(A_n) < \infty$ for every $n \ge 1$. Set

$$h_t^{i,n}(\mathfrak{z}) := (-n) \vee (\mathbf{1}_{A_n} h_t^{(i)}(\mathfrak{z})) \wedge n.$$

Then by Jensen's inequality for the optional projections we have $|{}^o h_s^{i,n}(\mathfrak{z})|^2 \le {}^o(|h_s^{i,n}(\mathfrak{z})|^2)$ for every $\mathfrak{z} \in \mathfrak{Z}_i$, and by an application of Corollary 3.3 we obtain

$$\begin{aligned} \int_0^t \int_{\mathfrak{Z}_i} |{}^o h_s^{i,n}(\mathfrak{z})|^2\ \nu_i(d\mathfrak{z})ds &\le \int_0^t \int_{\mathfrak{Z}_i} {}^o(|h_s^{i,n}(\mathfrak{z})|^2)\ \nu_i(d\mathfrak{z})ds \\ &= \mathbb{E}\left(\int_0^t \int_{\mathfrak{Z}_i} |h_s^{i,n}(\mathfrak{z})|^2\ \nu_i(d\mathfrak{z})ds \Big| \mathcal{G}_t\right) \\ &= \int_0^t {}^o(|h_s^{i,n}|^2_{L_2(\mathfrak{Z}_i)})\, ds \le \int_0^t {}^o(|h_s^{(i)}|^2_{L_2(\mathfrak{Z}_i)})\, ds. \end{aligned}$$

Letting here $n \to \infty$ and using the Monotone Convergence Theorem and the properties of extended optional projections on the left-hand side of the first inequality, we finish the proof of the remark. □

Let $\mathbb{P}(\mathbb{R}^d)$ be the space of of probability measures on the Borel sets of $\mathbb{R}^d$, equipped with the topology of weak convergence of measures. Recall that $C_b(\mathbb{R}^d)$ denotes the space of bounded continuous real functions on $\mathbb{R}^d$, and as before, let $(\mathfrak{Z}, \mathcal{Z})$ be a separable measurable space.

Lemma 3.4 *Let $(\Omega, \mathcal{F}, P)$ be a complete probability space equipped with a right-continuous filtration $(\mathcal{G}_t)_{t\geq 0}$, $\mathcal{G}_t \subset \mathcal{F}$ for $t \geq 0$, such that $\mathcal{G}_0$ contains all P-zero sets of $\mathcal{F}$. Let $(X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)$-measurable cadlag process. Then the following statements hold.*

(i) *There is a $\mathbb{P}(\mathbb{R}^d)$-valued weakly cadlag process $(P_t)_{t\geq 0}$ such that for every bounded real-valued Borel function φ on $\mathbb{R}^d$ and for each $t \geq 0$*

$$P_t(\varphi) = \mathbb{E}(\varphi(X_t)|\mathcal{G}_t) \quad \text{(a.s.)}. \tag{3.34}$$

(ii) *Let $(P_t)_{t\geq 0}$ be the measure-valued process from (i). Assume $f = f(\omega, t, \mathfrak{z}, x)$ is a $\mathcal{O}_{\mathcal{G}} \otimes \mathcal{Z} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable real function on $\Omega \times \mathbb{R}_+ \times \mathfrak{Z} \times \mathbb{R}^d$. Define*

$$P_t(f(t,\mathfrak{z})) := \begin{cases} \int_{\mathbb{R}^d} f(t,\mathfrak{z},x)\, P_t(dx), \; \textit{for } (t,\omega,\mathfrak{z}), \\ \textit{if } \int_{\mathbb{R}^d} |f(t,\mathfrak{z},x)|\, P_t(dx) < \infty \\ \infty \; \textit{elsewhere.} \end{cases}$$

Then $P_t(f(t,\mathfrak{z}))$ is an $\mathcal{O}_{\mathcal{G}} \otimes \mathcal{Z}$-measurable (extended) function of $(\omega, t, \mathfrak{z})$ such that

$$\mathbb{E}(f(t,\mathfrak{z},X_t)|\mathcal{G}_t) = P_t(f(t,\mathfrak{z})) \text{ (a.s.)} \quad \textit{for each } (\mathrm{t},\mathfrak{z}) \in \mathbb{R}_+ \times \mathfrak{Z}, \tag{3.35}$$

whenever one of the expressions is finite (a.s.). Moreover, $P_{t_0}(f(t_0,\mathfrak{z}_0))$ is finite (a.s.) for a pair $(t_0,\mathfrak{z}_0) \in \mathbb{R}_+ \times \mathfrak{Z}$ if $f(t_0,\mathfrak{z}_0,X_{t_0})$ is σ-integrable relative to $\mathcal{G}_{t_0}$.

Proof Statement (i) is shown in [28]. Thus (ii) holds if $f = g(t,\mathfrak{z})\varphi(x)$ for bounded $\mathcal{O}_{\mathcal{G}} \otimes \mathcal{Z}$-measurable functions g on $\Omega \times \mathbb{R}_+ \times \mathfrak{Z}$ and bounded Borel functions φ on $\mathbb{R}^d$. Hence by a standard monotone class argument we get (ii) under the additional assumption that f is bounded. In the general case, the set $A \subset \Omega \times \mathbb{R}_+ \times \mathfrak{Z}$ where

$$\int_{\mathbb{R}^d} |f(t,\mathfrak{z},x)| P_t(dx) = \infty$$

is in $\mathcal{O}_{\mathcal{G}} \otimes \mathcal{Z}$. Consequently, $P_t(f(t,\mathfrak{z}))$ is $\mathcal{O}_{\mathcal{G}} \otimes \mathcal{Z}$-measurable in $(\omega, t, \mathfrak{z})$. We have

$$\mathbb{E}(|f(t,\mathfrak{z},X_t)| \wedge n|\mathcal{G}_t) = \int_{\mathbb{R}^d} |f(t,\mathfrak{z},x)| \wedge n \, P_t(dx) \quad \text{(a.s.)}$$

for every integer $n \geq 1$. Letting here $n \to \infty$ we get

$$\mathbb{E}(|f(t, \mathfrak{z}, X_t)||\mathcal{G}_t) = \int_{\mathbb{R}^d} |f(t, \mathfrak{z}, x)| \, P_t(dx) \quad \text{(a.s.)}, \tag{3.36}$$

that implies (3.35). If $f(t_0, \mathfrak{z}_0, X_{t_0})$ is σ-integrable relative to $\mathcal{G}_{t_0}$, then there is an increasing sequence $(\Omega_n)_{n=1}^{\infty}$ such that $\Omega_n \in \mathcal{G}_{t_0}$, $P(\cup_{n=1}^{\infty} \Omega_n) = 1$, and

$$\mathbf{1}_{\Omega_n} \int_{\mathbb{R}^d} f(t_0, \mathfrak{z}_0, x) \, P_{t_0}(dx) = \mathbb{E}(\mathbf{1}_{\Omega_n} f_n(t_0, \mathfrak{z}_0, X_{t_0})|\mathcal{G}_{t_0})$$

is almost surely finite for every $n \geq 1$. □

Corollary 3.5 *Let $(\Omega, \mathcal{F}, P, (\mathcal{G}_t)_{t\geq 0})$ and $(X_t)_{t\in[0,T]}$ be a filtered probability space and a stochastic process, respectively, satisfying the conditions in Lemma 3.4. Let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration such that $\mathcal{G}_t \subset \mathcal{F}_t \subset \mathcal{F}$ for $t \geq 0$. Let Q be a probability measure on $\mathcal{F}$ such that $dQ = \gamma_T dP$ for a $\mathcal{F}_T$-measurable positive random variable γ_T. Then the following statements hold.*

(i) *There is an $\mathbb{M}(\mathbb{R}^d)$-valued weakly cadlag stochastic process $(\mu_t)_{t\in[0,T]}$ such that for every bounded real-valued Borel function φ on $\mathbb{R}^d$ and for every $t \in [0, T]$*

$$\mu_t(\varphi) = \mathbb{E}_Q(\gamma_T^{-1}\varphi(X_t)|\mathcal{G}_t) = \mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t)|\mathcal{G}_t) \text{ (a.s.)}. \tag{3.37}$$

(ii) *Let $f = f(\omega, t, \mathfrak{z}, x)$ be an $\mathcal{O}_\mathcal{G} \otimes \mathcal{Z} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable real function on $\Omega \times [0, T] \times \mathfrak{Z} \times \mathbb{R}^d$. Define*

$$\mu_t(f(t, \mathfrak{z})) := \begin{cases} \int_{\mathbb{R}^d} f(t, \mathfrak{z}, x) \, \mu_t(dx), \ \text{for } (t, \omega, \mathfrak{z}), \ \text{if } \int_{\mathbb{R}^d} |f(t, \mathfrak{z}, x)| \, \mu_t(dx) < \infty \\ \infty \ \text{elsewhere.} \end{cases}$$

Then $\mu_t(f(t, \mathfrak{z}))$ is an $\mathcal{O}_\mathcal{G} \otimes \mathcal{Z}$-measurable function such that for each $(t, \mathfrak{z})$ we have

$$\mathbb{E}_Q(\gamma_T^{-1} f(t, \mathfrak{z}, X_t)|\mathcal{G}_t) = \mathbb{E}_Q(\gamma_t^{-1} f(t, \mathfrak{z}, X_t)|\mathcal{G}_t) = \mu_t(f(t, \mathfrak{z})) \quad \textit{(a.s.)}, \tag{3.38}$$

whenever one of the expressions is finite (a.s.). Moreover, $\mu_{t_0}(f(t_0, \mathfrak{z}_0))$ is finite (a.s.) for a pair $(t_0, \mathfrak{z}_0) \in [0, T] \times \mathfrak{Z}$ if $\gamma_{t_0}^{-1} f(t_0, \mathfrak{z}_0, X_{t_0})$ is σ-integrable (with respect to Q) relative to $\mathcal{G}_{t_0}$, or equivalently, if $f(t_0, \mathfrak{z}_0, X_{t_0})$ is σ-integrable with respect to P relative to $\mathcal{G}_{t_0}$.

Proof Considering $(\mathcal{F}_{t+})_{t\geq 0}$ in place of $(\mathcal{F}_t)_{t\geq 0}$ we may assume in the proof that $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous. By Doob's theorem there is a cadlag $\mathcal{F}_t$-martingale, $(\gamma_t)_{t\in[0,T]}$, such that $\gamma_t = \mathbb{E}_P(\gamma_T|\mathcal{F}_t)$ (P-a.s) for each $t \in [0, T]$. Clearly, almost surely $\gamma_t > 0$ for all $t \in [0, T]$ since

$$0 = \mathbb{E}_P(\mathbf{1}_{\gamma_t=0}\gamma_t) = \mathbb{E}_P(\mathbf{1}_{\gamma_t=0}\gamma_T)$$

implies $P(\gamma_t = 0) = 0$ for every $t \in [0, T]$. Thus $(\gamma_t^{-1})_{t\in[0,T]}$ is a cadlag process, and it is an $\mathcal{F}_t$-martingale under Q, because

$$\mathbb{E}_Q(\gamma_T^{-1}|\mathcal{F}_t) = 1/\mathbb{E}_P(\gamma_T|\mathcal{F}_t) = \gamma_t^{-1} \quad \text{almost surely for } t \in [0, T].$$

Since, $\gamma = (\gamma)_{t\in[0,T]}$ is a (cadlag) $\mathcal{F}_t$-martingale under P, the set $\{\gamma_\tau\}$ for $\mathcal{F}_t$-stopping times $\tau \le T$ is uniformly P-integrable, and hence one knows that ${}^o\gamma$, the $\mathcal{G}_t$-optional projection of γ under P, is a cadlag process. Due to $\gamma > 0$, we have ${}^o\gamma > 0$ (a.s.). Define $\mu_t := ({}^o\gamma_t)^{-1}P_t$ for $t \in [0, T]$, where $(P_t)_{t\in[0,T]}$ is the $\mathbb{P}(\mathbb{R}^d)$-valued $\mathcal{G}_t$-adapted cadlag process (in the topology of weak convergence of measures) by Lemma 3.4. Hence, $(\mu_t)_{t\in[0,T]}$ is a $\mathcal{G}_t$-adapted cadlag $\mathbb{M}(\mathbb{R}^d)$-valued process, and by (3.34) for every bounded Borel function φ on $\mathbb{R}^d$ we have

$$\begin{aligned}&\mathbb{E}_Q(\gamma_T^{-1}\varphi(X_t)|\mathcal{G}_t) = \mathbb{E}_P(\varphi(X_t)|\mathcal{G}_t)/\mathbb{E}_P(\gamma_T|\mathcal{G}_t)\\ &\quad = \mathbb{E}_P(\varphi(X_t)|\mathcal{G}_t)({}^o\gamma_t)^{-1} = ({}^o\gamma_t)^{-1}P_t(\varphi) = \mu_t(\varphi) \quad \text{(a.s) for each } t \in [0, T].\end{aligned}$$

On the other hand, by well-known properties of conditional expectations

$$\begin{aligned}&\mathbb{E}_Q(\gamma_T^{-1}\varphi(X_t)|\mathcal{G}_t) = \mathbb{E}_Q(\mathbb{E}_Q(\gamma_T^{-1}\varphi(X_t)|\mathcal{F}_t)|\mathcal{G}_t)\\ &\quad = \mathbb{E}_Q(\varphi(X_t)\mathbb{E}_Q(\gamma_T^{-1}|\mathcal{F}_t)|\mathcal{G}_t) = \mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t))|\mathcal{G}_t),\end{aligned}$$

which completes the proof of (i). To prove (ii), note that the function $\mu_t(f(t, \mathfrak{z}))$ is $\mathcal{O}_\mathcal{G} \otimes \mathcal{Z}$-measurable in $(\omega, t, \mathfrak{z})$, and by (3.35) for each $(t, \mathfrak{z})$ almost surely

$$\begin{aligned}&\mu_t(f(t, \mathfrak{z})) = ({}^o\gamma_t)^{-1}P_t(f(t, \mathfrak{z})) = \mathbb{E}(({}^o\gamma_t)^{-1}f(t, \mathfrak{z}, X_t)|\mathcal{G}_t)\\ &\quad = \mathbb{E}_Q(\gamma_T^{-1}({}^o\gamma_t)^{-1}f(t, \mathfrak{z}, X_t)|\mathcal{G}_t)/\mathbb{E}_Q(\gamma_T^{-1}|\mathcal{G}_t) = \mathbb{E}_Q(\gamma_T^{-1}({}^o\gamma_t)^{-1}f(t, \mathfrak{z}, X_t)|\mathcal{G}_t){}^o\gamma_t\\ &\quad = \mathbb{E}_Q(\gamma_T^{-1}f(t, \mathfrak{z}, X_t)|\mathcal{G}_t) = \mathbb{E}_Q(\gamma_t^{-1}f(t, \mathfrak{z}, X_t)|\mathcal{G}_t),\end{aligned}$$

where the last equation holds because γ^{-1} is an $\mathcal{F}_t$-martingale under Q. We finish the proof with the obvious observation that $\gamma_{t_0}^{-1}f(t_0, \mathfrak{z}_0, X_{t_0})$ is σ-integrable with respect to Q relative to $\mathcal{G}_{t_0}$ if $f(t_0, \mathfrak{z}_0, X_{t_0})$ is σ-integrable with respect to P relative to $\mathcal{G}_{t_0}$. □

4 Proof of Theorem 2.1

Recall that by Assumption 2.2 the measure Q, defined by $dQ = \gamma_T dP$ is a probability measure, equivalent to P, and by Girsanov's theorem under Q the process $(W_t, \tilde{V}_t)_{t\in[0,T]}$, where $(\tilde{V}_t)_{t\in[0,T]}$ is defined by (2.5), is a $d_1 + d'$-dimensional $\mathcal{F}_t$-Wiener process. Moreover, under Q the random measures $\tilde{N}_0$ and $\tilde{N}_1$ remain independent $\mathcal{F}_t$-Poisson martingale measures, with characteristic measures ν_0 and ν_1, respectively. Clearly, $(\gamma_t)_{t\in[0,T]}$ is an $\mathcal{F}_t$-martingale under P. By Itô's formula

$$d\gamma_t^{-1} = \gamma_t^{-1}B_t^l(X_t)\,d\tilde{V}_t^l, \tag{4.1}$$

the process $\gamma^{-1} = (\gamma_t^{-1})_{t\in[0,T]}$ is an $\mathcal{F}_t$-local martingale under Q. Hence, taking into account $\mathbb{E}_Q\gamma_T^{-1} = 1$, we get that γ^{-1} is an $\mathcal{F}_t$-martingale under Q. Thus the Bayes formula for bounded Borel functions φ on $\mathbb{R}^d$ gives

$$\mathbb{E}(\varphi(X_t)|\mathcal{F}_t^Y) = \frac{\mathbb{E}_Q(\gamma_T^{-1}\varphi(X_t)|\mathcal{F}_t^Y)}{\mathbb{E}_Q(\gamma_T^{-1}|\mathcal{F}_t^Y)} = \frac{\mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t)|\mathcal{F}_t^Y)}{\mathbb{E}_Q(\gamma_t^{-1}|\mathcal{F}_t^Y)} \text{ (a.s.)}, \tag{4.2}$$

often also referred as Kallianpur-Striebel formula in the literature. Using $\tilde{V}$ we can rewrite system (1.1) in the form

$$\begin{aligned} dX_t &= b(t, Z_t)\,dt + \sigma(t, Z_t)\,dW_t + \rho(t, Z_t)\,dV_t \\ &\quad + \int_{\mathfrak{Z}_0} \eta(t, Z_{t-}, \mathfrak{z})\,\tilde{N}_0(d\mathfrak{z}, dt) + \int_{\mathfrak{Z}_1} \xi(t, Z_{t-}, \mathfrak{z})\,\tilde{N}_1(d\mathfrak{z}, dt), \\ dY_t &= d\tilde{V}_t + \int_{\mathfrak{Z}_1} \mathfrak{z}\,\tilde{N}_1(dt, d\mathfrak{z}), \end{aligned} \tag{4.3}$$

which shows, in particular, that $(Y_t)_{t\in[0,T]}$ is a Lévy process under Q, and hence it is well-known that the filtration $(\mathcal{F}_t^Y)_{t\in[0,T]}$ is right-continuous. Thus we can apply Lemma 3.4 and Corollary 3.5 with the unobservable process $(X_t)_{t\in[0,T]}$ and the filtration $(\mathcal{G}_t)_{t\in[0,T]} = (\mathcal{F}_t^Y)_{t\in[0,T]}$ to have a $\mathbb{P}$-valued and $\mathbb{M}$-valued weakly cadlag $\mathcal{F}_t^Y$-adapted processes $P_t(dx)$ and $\mu_t(dx)$, respectively, such that for every bounded Borel function φ on $\mathbb{R}^d$ for each $t \in [0, T]$ we have

$$P_t(\varphi) = \mathbb{E}(\varphi(X_t)|\mathcal{F}_t^Y), \quad \mu_t(\varphi) = \mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t)|\mathcal{F}_t^Y) \text{ (a.s.)},$$

and by (4.2) it follows that almost surely $P_t = \mu_t/\mu_t(\mathbf{1})$ for all $t \in [0, T]$. To get an equation for $d\mu_t(\varphi)$ for sufficiently smooth functions we calculate first the stochastic differential $d(\gamma_t^{-1}\varphi(X_t))$.

Proposition 4.1 *Let $\varphi \in C_b^2(\mathbb{R}^d)$. Then for the stochastic differential of $\gamma_t^{-1}\varphi(X_t)$ we have*

$$\begin{aligned} d\left(\gamma_t^{-1}\varphi(X_t)\right) &= \gamma_t^{-1}\mathcal{L}_t\varphi(X_t)\,dt + \gamma_t^{-1}\mathcal{M}_t^l\varphi(X_t)\,d\tilde{V}_t^l + \gamma_t^{-1}\sigma_t^{ik}(X_t)D_i\varphi(X_t)\,dW_t^k \\ &\quad + \gamma_t^{-1}\int_{\mathfrak{Z}_0} I_t^\eta\varphi(X_{t-})\,\tilde{N}_0(d\mathfrak{z}, dt) + \gamma_t^{-1}\int_{\mathfrak{Z}_1} I_t^\xi\varphi(X_{t-})\,\tilde{N}_1(d\mathfrak{z}, dt) \\ &\quad + \gamma_t^{-1}\int_{\mathfrak{Z}_0} J_t^\eta\varphi(X_t)\,\nu_0(d\mathfrak{z})\,dt + \gamma_t^{-1}\int_{\mathfrak{Z}_1} J_t^\xi\varphi(X_t)\,\nu_1(d\mathfrak{z})\,dt. \end{aligned} \tag{4.4}$$

Proof By Itô's formula, see for example in [1] or [16], for $\varphi \in C_b^2(\mathbb{R}^d)$ we have

$$
\begin{aligned}
d\varphi(X_t) = &\Big(\mathcal{L}_t\varphi(X_t) - \rho_t^{il} B_t^l(X_t) D_i\varphi(X_t)\Big)\, dt \\
&+ \sigma_t^{ik}(X_t) D_i\varphi(X_t)\, dW_t^k + \rho_t^{il}(X_t) D_i\varphi(X_t)\, d\tilde{V}_t^l \\
&+ \int_{\mathfrak{Z}_0} I_t^\eta \varphi(X_{t-})\, \tilde{N}_0(dt, d\mathfrak{z}) + \int_{\mathfrak{Z}_1} I_t^\xi \varphi(X_{t-})\, \tilde{N}_1(dt, d\mathfrak{z}) \\
&+ \int_{\mathfrak{Z}_0} J_t^\eta \varphi(X_{t-})\, \nu_0(d\mathfrak{z})dt + \int_{\mathfrak{Z}_1} J_t^\xi \varphi(X_{t-})\, \nu_1(d\mathfrak{z})dt,
\end{aligned}
$$

where we use the notations introduced before the formulation of Theorem 2.1. Hence using (4.1) and the stochastic differential rule for products,

$$
d(\gamma_t^{-1}\varphi(X_t)) = \gamma_t^{-1} d\varphi(X_t) + \varphi(X_{t-})\, d\gamma_t^{-1} + d\gamma_t^{-1} d\varphi(X_t),
$$

where

$$
d\gamma_t^{-1} d\varphi(X_t) = \gamma_t^{-1} \rho_t^{il} B_t^l(X_t) D_i\varphi(X_t)\, dt,
$$

we obtain (4.4). □

To calculate the conditional expectation (under Q) of the terms in the equation for $\gamma_t^{-1}\varphi(X_t)$, given $\mathcal{F}_t^Y$, we describe below the structure of $\mathcal{F}_t^Y$. For each $t \geq 0$ we denote by $\mathcal{F}_t^{\tilde{N}}$ the P-completion of the σ-algebra generated by the random variables $N_1((0, s] \times \Gamma)$ for $s \in (0, t]$ and $\Gamma \in \mathcal{Z}_1$ such that $\nu_1(\Gamma) < \infty$.

Lemma 4.2 *For every $t \in [0, T]$ we have*

$$
\mathcal{F}_t^Y = \mathcal{F}_0^Y \vee \mathcal{F}_t^{\tilde{V}} \vee \mathcal{F}_t^{\tilde{N}_1},
$$

where $\mathcal{F}_0^Y \vee \mathcal{F}_t^{\tilde{V}} \vee \mathcal{F}_t^{\tilde{N}_1}$ denotes the P-completion of the smallest σ-algebra containing $\mathcal{F}_0^Y$, $\mathcal{F}_t^{\tilde{V}}$ and $\mathcal{F}_t^{\tilde{N}_1}$.

Proof From (4.3) it immediately follows that

$$
\mathcal{F}_t^Y \subseteq \mathcal{F}_0^Y \vee \mathcal{F}_t^{\tilde{V}} \vee \mathcal{F}_t^{\tilde{N}_1}.
$$

To prove the reversed inclusion, we claim

$$
N^Y((0, t] \times A) = N_1((0, t] \times A) \quad \text{almost surely for all } t \in [0, T] \tag{4.5}
$$

for every $A \in \mathcal{Z}_1$, where N^Y is the measure of jumps for the process Y. Clearly, $N^Y(d\mathfrak{z}, dt) = N^M(d\mathfrak{z}, dt)$, where N^M is the measure of jumps for the process

$$
M_t = \int_0^t \int_{\mathfrak{Z}_1} \mathfrak{z}\, \tilde{N}_1(d\mathfrak{z}, dt), \quad t \geq 0,
$$

i.e.,

$$N^Y((0,t]\times A)=\sum_{0<s\le t}\mathbf{1}_A(\Delta Y_s)=\sum_{0<s\le t}\mathbf{1}_A(\Delta M_s)\quad A\in\mathcal{Z}_1.$$

To show (4.5) let $A=A_0$ be a set from $\mathcal{Z}_1$ such that $\nu_1(A_0)<\infty$. Then

$$M_t^{A_0}:=\int_0^t\int_{A_0}\mathfrak{z}\,\tilde{N}_1(d\mathfrak{z},ds)=\sum_{0<s\le t}p_s\mathbf{1}_{A_0}(p_s)-t\int_{A_0}\mathfrak{z}\,\nu_1(d\mathfrak{z}),$$

where $(p_t)_{t\in[0,T]}$ is the Poisson point process associated with N_1. Note that due to $\nu_1(A_0)<\infty$, the decomposition in the last displayed formula is well-defined, the sum appearing in the right-hand side has only finitely many terms, and the integral there is finite almost surely. Hence for $N^0(d\mathfrak{z},dt)$, the measure of jumps of the process M^{A_0}, we have that almost surely

$$N^0((0,t]\times A_0)=N_1((0,t]\times A_0)\quad\text{for all } t\in[0,T]. \tag{4.6}$$

It is not difficult to see that $N^0((0,t]\times A_0)=N^M((0,t]\times A_0)$. Hence (4.5) for $A=A_0$ follows.

Since ν_1 is σ-finite, for an arbitrary $B\in\mathcal{Z}_1$ there is a sequence $(B_n)_{n=1}^\infty$ of disjoint sets $B_n\in\mathcal{Z}_1$ such that $B=\bigcup_{n=1}^\infty B_n$ and $\nu_1(B_n)<\infty$ for each $n\ge 1$. Thus for each integer $n\ge 1$ we have (4.5) with B_n in place of A, and summing this up over $n\ge 1$ and using the σ-additivity of N^Y and N_1 we obtain (4.5) with B in place of A. Noting that $\mathcal{F}_t^Y$ contains the σ-algebra generated by $N^Y((0,s]\times B)$ for each $s\le t$ and $B\in\mathcal{Z}$, we see that $\mathcal{F}_t^Y\supset\mathcal{F}_t^{\tilde{N}_1}$. Clearly, $\mathcal{F}_t^Y\supset\mathcal{F}_0^Y$, and taking into account

$$Y_t-Y_0-\int_0^t\int_{\mathfrak{Z}_1}\mathfrak{z}\,\tilde{N}_1(d\mathfrak{z},ds)=\tilde{V}_t,\quad\text{for } t\in[0,T],$$

we get $\mathcal{F}_t^Y\supset\mathcal{F}_t^{\tilde{V}}$. Consequently,

$$\mathcal{F}_t^Y\supseteq\mathcal{F}_0^Y\vee\mathcal{F}_t^{\tilde{V}}\vee\mathcal{F}_t^{\tilde{N}},$$

that completes the proof. □

The above lemma is an essential tool in obtaining the filtering equations. A similar lemma in a more general setting in some directions is presented in [26] and [27] to obtain the filtering equations for the model considered in these papers. It seems to us, however, that this lemma, Lemma 3.2 in [26], used as well in [27, p.4], may not hold under the general conditions formulated in these papers, since it is not true in the simple case of vanishing coefficients in front of the random measures in the observation process. It is worth noticing that when instead of the integrand $\mathfrak{z}$ a stochastic integrand depending on $Z_t=(X_t,Y_t)$ is considered in the observation process Y, the integral of such a term against a Poisson random measure may fail to be a Lévy process, as it may not have independent increments, which is a crucial property for the filtration generated by the observation.

Now we are going to get an equation for $\mu(\varphi)$ by noting that by Proposition 4.1 we have

$$\gamma_t^{-1}\varphi(X_t) = \varphi(X_0) + \alpha_t + \alpha_t^0 + \alpha_t^1 + \beta_t^0 + \beta_t^1 + \delta_t^0 + \delta_t^1, \quad t \in [0,T], \quad (4.7)$$

where

$$\begin{aligned}
\alpha_t &:= \int_0^t \gamma_s^{-1}\mathcal{L}_s\varphi(X_s)\,ds,\\
\alpha_t^0 &:= \int_0^t\int_{\mathfrak{Z}_0} \gamma_s^{-1}J_s^{\eta}\varphi(X_s)\,\nu_0(d\mathfrak{z})ds, \quad \alpha_t^1 := \int_0^t\int_{\mathfrak{Z}_1} \gamma_s^{-1}J_s^{\xi}\varphi(X_s)\,\nu_1(d\mathfrak{z})ds,\\
\beta_t^0 &:= \int_0^t \gamma_s^{-1}\sigma_s^{ik}(X_s)D_i\varphi(X_s)\,dW_s^k, \quad \beta_t^1 := \int_0^t \gamma_s^{-1}\mathcal{M}_s^l\varphi(X_s)\,d\tilde{V}_s^l,\\
\delta_t^0 &:= \int_0^t\int_{\mathfrak{Z}_0} \gamma_s^{-1}I_s^{\eta}\varphi(X_{s-})\,\tilde{N}_0(d\mathfrak{z},ds), \quad \delta_t^1 := \int_0^t\int_{\mathfrak{Z}_1} \gamma_s^{-1}I_s^{\xi}\varphi(X_{t-})\,\tilde{N}_1(d\mathfrak{z},ds),
\end{aligned}$$

for $\varphi \in C_b^2(\mathbb{R}^d)$. We want to take the conditional expectation of both sides of Eq. (4.7) for each $t \in [0,T]$, under Q, given $\mathcal{F}_t^Y$. In order to apply Corollary 3.3, we should verify that the random variables

$$\begin{aligned}
G &:= \int_0^T \gamma_s^{-1}|\mathcal{L}_t\varphi(X_s)|\,ds,\\
G^{(0)} &:= \int_0^T\int_{\mathfrak{Z}_0} \gamma_s^{-1}|J_s^{\eta}\varphi(X_s)|\,\nu_0(d\mathfrak{z})\,ds,\\
G^{(1)} &:= \int_0^T\int_{\mathfrak{Z}_1} \gamma_s^{-1}|J_s^{\xi}\varphi(X_s)|\,\nu_1(d\mathfrak{z})\,ds,\\
F^{(0)} &:= \left(\int_0^T \gamma_s^{-2}|\sigma_s^i(X_s)D_i\varphi(X_s)|^2\,ds\right)^{1/2},\\
F^{(1)} &:= \left(\int_0^T \gamma_s^{-2}\sum_l|\mathcal{M}_s^l\varphi(X_s)|^2\,ds\right)^{1/2},\\
H^{(0)} &:= \left(\int_0^T\int_{\mathfrak{Z}_0} \gamma_s^{-2}|I_s^{\eta}\varphi(X_{s-})|^2\,\nu_0(d\mathfrak{z})ds\right)^{1/2},\\
H^{(1)} &:= \left(\int_0^T\int_{\mathfrak{Z}_1} \gamma_s^{-2}|I_s^{\xi}\varphi(X_{s-})|^2\,\nu_1(d\mathfrak{z})ds\right)^{1/2}
\end{aligned}$$

are σ-integrable with respect to Q relative to $\mathcal{F}_0^Y$, and that (3.27) holds for ${}^Qf^{(i)}$ in place of ${}^of$, and for ${}^Qh^{(i)}$ in place of ${}^oh^{(i)}$, where ${}^Qf^{(0)}$, ${}^Qf^{(1)}$, ${}^Qh^{(0)}$ and ${}^Qh^{(1)}$ are the

$\mathcal{F}_t^Y$-optional projection under Q of

$$f^{k(0)} := (\gamma_s^{-1}\sigma_s^{ik}(X_s)D_i\varphi(X_s))_{s\in[0,T]}, \quad f^{l(1)} := (\gamma_s^{-1}\mathcal{M}_s^l\varphi(X_s))_{s\in[0,T]},$$
$$h^{(0)} := (\gamma_s^{-1}I_s^\eta\varphi(X_{s-}))_{s\in[0,T]} \quad \text{and} \quad h^{(1)} := (\gamma_s^{-1}I_s^\xi\varphi(X_{s-}))_{s\in[0,T]},$$

respectively for each fixed $k = 1, 2, \dots, d_1$ and $l = 1, \dots, d'$. For a fixed integer $n \geq 1$ let $\Omega_n = \{\omega \in \Omega : |Y_0| \leq n\}$. Then due to Assumption 2.1, the martingale property of $(\gamma_t)_{t\in[0,T]}$ and (2.2) we have

$$\begin{aligned}
\mathbb{E}_Q(\mathbf{1}_{\Omega_n}G) &\leq N\mathbb{E}\left(\gamma_T\int_0^T \gamma_t^{-1}(K_0 + K_1\mathbf{1}_{\Omega_n}|Z_t| + K_2\mathbf{1}_{\Omega_n}|Z_t|^2)\,dt\right)\\
&= N\int_0^T \mathbb{E}(\gamma_T\gamma_t^{-1}(K_0 + K_1\mathbf{1}_{\Omega_n}|Z_t| + K_2\mathbf{1}_{\Omega_n}|Z_t|^2)\,dt\\
&= N\int_0^T \mathbb{E}(K_0 + K_1\mathbf{1}_{\Omega_n}|Z_t| + K_2\mathbf{1}_{\Omega_n}|Z_t|^2)\,dt\\
&\leq N'(K_0 + K_1\mathbb{E}|X_0| + K_1\mathbb{E}(\mathbf{1}_{\Omega_n}|Y_0|) + K_2\mathbb{E}|X_0|^2 + K_2\mathbb{E}(\mathbf{1}_{\Omega_n}|Y_0|^2)) < \infty
\end{aligned}$$

with constants N and N', which shows that G is σ-integrable with respect to Q relative to $\mathcal{F}_0^Y$. Similarly, using the estimate

$$|J^\eta\varphi(X_t)| \leq \sup_{x\in\mathbb{R}^d}|D_{ij}\varphi(x)||\eta_t^i(X_t)||\eta_t^j(X_t)|,$$

we get

$$\begin{aligned}
\mathbb{E}_Q(\mathbf{1}_{\Omega_n}G^{(0)}) &= \int_0^T \mathbb{E}\int_{\mathfrak{Z}_0}\mathbf{1}_{\Omega_n}|J_s^\eta\varphi(X_s)|\,\nu_0(d\mathfrak{z})ds\\
&\leq N\int_0^T \mathbb{E}\int_{\mathfrak{Z}_0}\mathbf{1}_{\Omega_n}|\eta(s, Z_s, \mathfrak{z})|^2\,\nu_0(d\mathfrak{z})\,ds\\
&\leq N'\int_0^T \mathbb{E}(K_0 + K_2\mathbf{1}_{\Omega_n}|Z_s|^2)\,ds < \infty
\end{aligned}$$

with constants N and N'. In the same way we get $\mathbb{E}_Q(\mathbf{1}_{\Omega_n}G^{(1)}) < \infty$. To prove that $F^{(i)}$ and $H^{(i)}$ are σ-integrable (with respect to Q) relative to F_0^Y, we claim first that

$$A_n := \mathbb{E}_Q\mathbf{1}_{\Omega_n}\sup_{t\leq T}\gamma_t^{-1} < \infty \quad \text{for every integer } n \geq 1. \tag{4.8}$$

To prove this we repeat a method used in proof of Theorem 2.2. From (4.1) by using the Davis inequality and then Young's inequality we get

$$\mathbb{E}_Q \mathbf{1}_{\Omega_n} \sup_{t\in[0,T]} \gamma^{-1}_{t\wedge\tau_k} \le 1 + 3\mathbb{E}\left(\int_0^{T\wedge\tau_k} \mathbf{1}_{\Omega_n}\gamma_t^{-2}|B(t,Z_t)|^2\,dt\right)^{1/2}$$
$$\le 1 + \tfrac{1}{2}\mathbb{E}_Q\mathbf{1}_{\Omega_n}\sup_{t\in[0,T]}\gamma^{-1}_{t\wedge\tau_k} + 5\mathbb{E}\int_0^T \mathbf{1}_{\Omega_n}\gamma_t^{-1}|B(t,Z_t)|^2\,dt$$

for stopping times

$$\tau_k = \inf\{t\in[0,T] : \gamma_t^{-1} \ge k\}, \quad \text{for integers } k \ge 1.$$

Rearranging this inequality and then letting $k\to\infty$ by Fatou's lemma we obtain

$$\mathbb{E}_Q\mathbf{1}_{\Omega_n}\sup_{t\in[0,T]}\gamma_t^{-1} \le 2 + 10\int_0^T \mathbb{E}_Q\mathbf{1}_{\Omega_n}\gamma_t^{-1}|B(t,Z_t)|^2\,dt.$$

Hence we get (4.8) by noticing that using the martingale property of γ, the estimate in (2.2) and $K_2\mathbb{E}|X_0|^2 < \infty$, for every $t\in[0,T]$ we have

$$\mathbb{E}_Q\mathbf{1}_{\Omega_n}\gamma_t^{-1}|B(t,Z_t)|^2 = \mathbb{E}\mathbf{1}_{\Omega_n}\gamma_T\gamma_t^{-1}|B(t,Z_t)|^2 = \mathbb{E}\mathbf{1}_{\Omega_n}|B(t,Z_t)|^2$$
$$\le K_0 + K_2\mathbb{E}\mathbf{1}_{\Omega_n}|Z_t|^2 \le K_0 + K_2N\mathbb{E}(1+|X_0|^2+\mathbf{1}_{\Omega_n}|Y_0|^2) < \infty.$$

Consequently,

$$\mathbb{E}_Q(\mathbf{1}_{\Omega_n}F^{(0)}) \le \mathbb{E}_Q\left(\mathbf{1}_{\Omega_n}\sup_{s\le T}\gamma_s^{-1/2}\left(\int_0^T \mathbf{1}_{\Omega_n}\gamma_s^{-1}|\sigma_s^i(X_s)D_i\varphi(X_s)|^2\,ds\right)^{1/2}\right)$$
$$\le A_n + B_n,$$

with $A_n < \infty$, and for a constant N,

$$B_n := \mathbb{E}_Q\int_0^T \mathbf{1}_{\Omega_n}\gamma_s^{-1}|\sigma_s^i(X_s)D_i\varphi(X_s)|^2\,ds = \int_0^T \mathbb{E}(\mathbf{1}_{\Omega_n}\gamma_T\gamma_s^{-1}|\sigma_s^i(X_s)D_i\varphi(X_s)|^2)\,ds$$
$$= \int_0^T \mathbb{E}|\mathbf{1}_{\Omega_n}\sigma_s^i(X_s)D_i\varphi(X_s)|^2\,ds \le N\int_0^T \mathbb{E}(K_0 + K_2\mathbf{1}_{\Omega_n}|Z_s|^2)\,ds < \infty.$$

We get $\mathbb{E}_Q(\mathbf{1}_{\Omega_n}F^{(1)}) < \infty$ in the same way. Similarly, $\mathbb{E}_Q(\mathbf{1}_{\Omega_n}H^{(0)}) \le A_n + C_n$, with A_n given in (4.8) and

$$C_n := \mathbb{E}_Q\int_0^T \gamma_s^{-1}\int_{\mathfrak{Z}_0}\mathbf{1}_{\Omega_n}|I_s^\eta\varphi(X_s)|^2\,\nu_0(d\mathfrak{z})ds = \int_0^T \mathbb{E}\int_{\mathfrak{Z}_0}\mathbf{1}_{\Omega_n}|I_s^\eta\varphi(X_s)|^2\,\nu_0(d\mathfrak{z})ds$$
$$\le N\int_0^T \mathbb{E}\int_{\mathfrak{Z}_0}\mathbf{1}_{\Omega_n}|\eta(s,Z_s,\mathfrak{z})|^2\,\nu_0(d\mathfrak{z})ds \le N'\int_0^T \mathbb{E}(K_0 + K_2\mathbf{1}_{\Omega_n}|Z_s|^2)\,ds < \infty$$

with constants N and N', where we use that by Taylor's formula we have

$$|I_s^{\eta}\varphi(X_s)| \leq \sup_{x\in\mathbb{R}^d} |D_i\varphi(x)||\eta_s^i(X_s)|.$$

In the same way we have $\mathbb{E}_Q(\mathbf{1}_{\Omega_n}H^{(1)}) < \infty$. For processes $h = (h_t)_{t\in[0,T]}$ recall that ${}^Q h$ and ${}^o h$ denote the $\mathcal{F}_t^Y$-optional projections of h under Q and under P, respectively. Then using the formula ${}^Q h = {}^o(\gamma h)/{}^o\gamma$, well-known properties of optional projections and Remark 3.1 we have

$$\begin{aligned}|{}^Q h^{(0)}|^2_{L_2(\mathfrak{Z}_0)} &= \frac{|{}^o(I^\eta\varphi(X))|^2_{L_2(\mathfrak{Z}_0)}}{({}^o\gamma)^2} \leq \frac{{}^o\Big(|I^\eta\varphi(X))|^2_{L_2(\mathfrak{Z}_0)}\Big)}{({}^o\gamma)^2} \\ &\leq N\frac{{}^o(K_0+K_2|Z|^2)}{({}^o\gamma)^2} = N\frac{K_0}{({}^o\gamma)^2} + NK_2\frac{{}^o(|X|^2)}{({}^o\gamma)^2} + NK_2\frac{|Y|^2}{({}^o\gamma)^2}\end{aligned}$$

with a constant N. Remember that since $\gamma = (\gamma)_{t\in[0,T]}$ is a (cadlag) $\mathcal{F}_t$-martingale under P, the set $\{\gamma_\tau\}$ for $\mathcal{F}_t$-stopping times $\tau \leq T$ is uniformly P-integrable and hence due to the right-continuity of $(\mathcal{F}_t^Y)_{t\in[0,T]}$, the optional projection ${}^o\gamma$ is a cadlag process. Moreover, due to $\gamma > 0$, we have ${}^o\gamma > 0$ (a.s.). Since by (2.2)

$$K_2\mathbb{E}(\sup_{t\leq T}\mathbf{1}_{\Omega_n}|X_t|^2) < \infty \quad \text{for every } n \geq 1,$$

(and $(\mathcal{F}_t^Y)_{t\in[0,T]}$ is right-continuous), the process $K_2{}^o(|X|^2)$ is a cadlag process. Consequently, $K_0/|{}^o\gamma|^2$, $K_2{}^o(|X|^2)/|{}^o\gamma|^2$ and $|Y|^2/|{}^o\gamma|^2$ are cadlag processes. Hence

$$\int_0^T \frac{1}{({}^o\gamma_s)^2}\,ds + K_2\int_0^T \frac{{}^o(|X|^2)_s}{({}^o\gamma_s)^2}\,ds + K_2\int_0^T \frac{|Y_s|^2}{({}^o\gamma_s)^2}\,ds < \infty \quad \text{(a.s.)},$$

which proves

$$\int_0^T\int_{\mathfrak{Z}_i} |{}^Q h_s^{(i)}|^2\,\nu_i(d\mathfrak{z})\,ds < \infty \quad \text{(a.s.)}$$

for $i = 0$, and we get this for $i = 1$ in the same way. By the same argument we have

$$\int_0^T |{}^Q f_s^{(i)}|^2\,ds < \infty \quad \text{(a.s.)} \quad \text{for i} = 0, 1.$$

Thus we can apply Corollary 3.3 to the processes α, α^i, β^i and δ^i (i=0,1), and then use Corollary 3.5, to get

$$\begin{aligned}
&\mathbb{E}_Q(\alpha_t|\mathcal{F}_t^Y) = \int_0^t \mu_s(\mathcal{L}_s\varphi)\,ds,\\
&\mathbb{E}_Q(\alpha_t^0|\mathcal{F}_t^Y) = \int_0^t\int_{\mathfrak{Z}_0} \mu_s(J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds, \quad \mathbb{E}_Q(\alpha_t^1|\mathcal{F}_t^Y) = \int_0^t\int_{\mathfrak{Z}_1} \mu_s(J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds,\\
&\mathbb{E}_Q(\beta_t^0|\mathcal{F}_t^Y) = 0, \quad \mathbb{E}_Q(\beta_t^1|\mathcal{F}_t^Y) = \int_0^t \mu_s(\mathcal{M}_s^l\varphi)\,d\tilde{V}_s^l,\\
&\mathbb{E}_Q(\delta_t^0|\mathcal{F}_t^Y) = 0, \quad \mathbb{E}_Q(\delta_t^1|\mathcal{F}_t^Y) = \int_0^t\int_{\mathfrak{Z}_1} \mu_s(I_s^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z},ds)
\end{aligned}$$

for $t\in[0,T]$ and $\varphi\in C_b^2(\mathbb{R}^d)$ almost surely, where $(\mu_t)_{t\in[0,T]}$ is an $\mathbb{M}(\mathbb{R}^d)$-valued $\mathcal{F}_t^Y$-adapted weakly cadlag process such that

$$\mu_t(\varphi) := \int_{\mathbb{R}^d}\varphi(x)\,\mu_t(dx) = \mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t)|\mathcal{F}_t^Y) \quad \text{(a.s.)} \quad \text{for each t} \in [0,\text{T}],$$

for every bounded Borel function φ on $\mathbb{R}^d$. Using Lemma 3.1 with random variables $X := \varphi(X_0)$, $Y := 1$ and σ-algebras $\mathcal{G}_1 := \mathcal{F}_0^Y$, $\mathcal{G} := \mathcal{F}_0$ and $\mathcal{G}_2 := \mathcal{F}_t^{\tilde{V}} \vee \mathcal{F}_t^{\tilde{N}_1}$ we get

$$\mathbb{E}_Q(\varphi(X_0)|\mathcal{F}_t^Y) = \mathbb{E}_Q(\varphi(X_0)|\mathcal{F}_0^Y) = \mu_0(\varphi) \quad \text{(a.s.)}.$$

Consequently, taking the conditional expectation of both sides of Eq. (4.7) under Q given $\mathcal{F}_t^Y$, we see that Eq. (2.7) holds for each $t\in[0,T]$ and $\varphi\in C_b^2(\mathbb{R}^d)$ almost surely, that implies that for each $\varphi\in C_b^2(\mathbb{R}^d)$ Eq. (2.7) holds almost surely for all $t\in[0,T]$, since we have cadlag processes in both sides of equation (2.7) for each $\varphi\in C_b^2(\mathbb{R}^d)$. To prove (2.8) first notice that for $\varphi := \mathbf{1}$ Eq. (2.7) gives

$$d\mu_t(\mathbf{1}) = \mu_t(B_t^k)\,d\tilde{V}_t^k, \quad \mu_0(\mathbf{1}) = 1.$$

Since $\mu_t(\mathbf{1}) = ({}^o\gamma_t)^{-1}P_t(\mathbf{1}) = ({}^o\gamma_t)^{-1}$, $t\in[0,T]$, is a continuous process such that $\mu_t(\mathbf{1}) = \mathbb{E}_Q(\gamma_t^{-1}|\mathcal{F}_t^Y)$ (a.s.) for each $t\in[0,T]$, it is the $\mathcal{F}_t^Y$-optional projection under Q of the positive process $(\gamma_t^{-1})_{t\in[0,T]}$. Hence $\lambda_t := \mu_t(\mathbf{1})$, $t\in[0,T]$, is a positive process, and by Itô's formula

$$d\lambda_t^{-1} = -\lambda_t^{-2}\mu_t(B_t^k)\,d\tilde{V}_t^k + \lambda_t^{-3}\sum_k \mu_t^2(B_t^k)\,dt.$$

By Itô's formula for the product $P_t(\varphi) = \lambda_t^{-1}\mu_t(\varphi)$ we have

$$
\begin{aligned}
dP_t(\varphi) = P_t(\mathcal{L}_t\varphi)\,dt + P_t(\mathcal{M}_t^k\varphi)\,d\tilde{V}_t^k + \int_{\mathfrak{Z}_0} P_t(J_t^\eta\varphi)\,\nu_0(d\mathfrak{z})dt \\
+ \int_{\mathfrak{Z}_1} P_t(J_t^\xi\varphi)\,\nu_1(d\mathfrak{z})dt \\
+ \int_{\mathfrak{Z}_1} P_t(I_t^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z},dt) + \lambda_t^{-3}\mu_t(\varphi)\sum_k \mu_t^2(B_t^k)\,dt \\
-\mu_t(\varphi)\lambda_t^{-2}\mu_t(B_t^k)\,d\tilde{V}_t^k - \lambda_t^{-2}\mu_t(B_t^k)\mu_t(\mathcal{M}_t^k\varphi)\,dt
\end{aligned}
$$

Hence noting that

$$
\begin{aligned}
&\lambda_t^{-3}\mu_t(\varphi)\sum_k \mu_t^2(B_t^k) = P_t(\varphi)\sum_k P_t^2(B_t^k), \quad \mu_t(\varphi)\lambda_t^{-2}\mu_t(B_t^k) = P_t(\varphi)P_t(B_t^k) \\
&\lambda_t^{-2}\mu_t(B_t^k)\mu_t(\mathcal{M}_t^k\varphi) = P_t(B_t^k)P_t(\mathcal{M}_t^k\varphi),
\end{aligned}
$$

we obtain

$$
\begin{aligned}
dP_t(\varphi) = P_t(\mathcal{L}_t\varphi)\,dt + \Big(P_t(\mathcal{M}_t^k\varphi) - P_t(\varphi)P_t(B_t^k)\Big)\,d\tilde{V}_t^k \\
- \Big(P_t(\mathcal{M}_t^k\varphi) - P_t(\varphi)P_t(B_t^k)\Big)\,P_t(B_t^k)\,dt \\
+ \int_{\mathfrak{Z}_0} P_t(J_t^\eta\varphi)\,\nu_0(d\mathfrak{z})dt + \int_{\mathfrak{Z}_1} P_t(J_t^\xi\varphi)\,\nu_1(d\mathfrak{z})dt \\
+ \int_{\mathfrak{Z}_1} P_t(I_t^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z},dt).
\end{aligned}
$$

Since clearly,

$$
\begin{aligned}
&\Big(P_t(\mathcal{M}_t^k\varphi) - P_t(\varphi)P_t(B_t^k)\Big)\,d\tilde{V}_t^k - \Big(P_t(\mathcal{M}_t^k\varphi) - P_t(\varphi)P_t(B_t^k)\Big)\,P_t(B_t^k)\,dt \\
&= \Big(P_t(\mathcal{M}_t^k\varphi) - P_t(\varphi)P_t(B_t^k)\Big)\,d\bar{V}_t^k
\end{aligned}
$$

with the process $(\bar{V}_t)_{t\in[0,T]}$, given by $d\bar{V}_t = d\tilde{V}_t - P_t(B_t)\,dt$, $\bar{V}_0 = 0$, this gives Eq. (2.8), and finishes the proof of Theorem 2.1.

Acknowledgements The authors are grateful to an anonymous referee for very useful remarks, comments and suggestions, which improved the quality of the exposition. They are also grateful to Nicolai Krylov at the University of Minnesota for discussions and precious suggestions with regards to the topic of the paper.

Author Contributions All authors worked on the manuscript.

Declarations

Competing interests The authors declare that they have no competing interests.

References

1. Applebaum, D., Blackwood, S.: The Kalman–Bucy filter for integrable Lévy processes with infinite second moment. J. Appl. Probab. **52**, 636–648 (2015)
2. Blackwood, S.: Lévy Processes and Filtering Theory, Dissertation, The University of Sheffield (2014)
3. Cass, T., Clark, M., Crisan, D.: The filtering equations revisited. In: Stochastic Analysis and Applications. Springer, Cham, vol. 2014, pp. 129–162 (2014)
4. Ceci, C.: Risk minimizing hedging for a partially observed high frequency data model. Stochastics **78**, 13–31 (2006)
5. Ceci, C., Colaneri, K.: Nonlinear filtering for jump diffusion observations. Adv. Appl. Probab. **44–3**, 678–701 (2012)
6. Ceci, C., Colaneri, K.: The Zakai equation of nonlinear filtering for jump-diffusion observations: existence and uniqueness. Appl. Math. Optim. **69**, 47–82 (2014)
7. Crisan, D.: The stochastic filtering problem: a brief historical account. J. Appl. Probab. **51A**, 13–22 (2014)
8. Davie, A., Germ, F., Gyöngy, I.: On partially observed jump diffusions II: the filtering density. Stoch. Partial Differ. Equ. Anal. Comput. **12**(3), 1628–1698 (2024)
9. Dellacherie, C., Meyer, P.-A.: Probabilities and Potential B: Theory of Martingales. North-Holland Publishing Company (1982)
10. Fernando, B.P.W., Hausenblas, E.: Nonlinear filtering with correlated Lévy noise characterized by copulas. Braz. J. Probab. Stat. **32**, 374–421 (2018)
11. Fujisaki, M., Kallianpur, G., Kunita, H.: Stochastic differential equations for the nonlinear filtering problem. Osaka J. Math. **9–1**, 19–40 (1972)
12. Germ, F., Gyöngy, I.: On partially observed jump diffusions III: regularity of the filtering density. Stoch. Partial Differ. Equ. Anal. Comput. 1–53 (2024)
13. Grigelionis, B., Mikulevicius, R.: Nonlinear filtering equations for stochastic processes with jumps. In: The Oxford Handbook of Nonlinear Filtering. Oxford University Press (2011)
14. Heunis, A.J., Lucic, V.M.: On the innovations conjecture of nonlinear filtering with dependent data. Electron. J. Probab. **13**, 2190–2216 (2008)
15. He, S., Wang, J., Yan, J.: Semimartingale Theory and Stochastic Calculus. Taylor & Francis (1992)
16. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes. North-Holland Publishing Company (1992)
17. Krylov, N.V.: Controlled Diffusion Processes. Springer (1980)
18. Krylov, N.V., Zatezalo, A.: A direct approach to deriving filtering equations for diffusion processes. Appl. Math. Optim. **42**(3), 315–332 (2000)
19. Krylov, N.V.: On the equivalence of σ-algebras in the filtering problem of diffusion processes. Theor. Probab. Appl. **24**, 772–781 (1979)
20. Krylov, N.V.: Introduction to the Theory of Random Processes. American Mathematical Society: Graduate Studies in Mathematics, vol. 43 (2002)
21. Krylov, N.V.: A simple proof of a result of A. Novikov. arXiv:math/0207013v1 (2002)
22. Krylov, N.V.: A few comments on a result of A. Novikov and Girsanov's theorem. Stochastics **91**(8), 1186–1189 (2019)
23. Liptser, R.S., Shiryayev, A.N.: Statistics of Random Processes I, 2nd edn. Springer (2001)
24. Poklukar, D.R.: Nonlinear filtering for jump-diffusions. J. Comput. Appl. Math. **197**, 558–567 (2006)
25. Popa, S., Sritharan, S.S.: Nonlinear filtering of Itô–Lévy stochastic differential equations with continuous observations. Commun. Stoch. Anal. **3**, 1 (2009)

26. Qiao, H., Duan, J.: Nonlinear filtering of stochastic dynamical systems with Lévy noises. Adv. Appl. Probab. **47–3**, 902–918 (2015)
27. Qiao, H.: Nonlinear filtering of stochastic differential equations driven by correlated Lévy noises. Stochastics **93**(8), 1156–1185 (2021)
28. Yor, M.: Sur les théories du filtrage et de la prédiction, Séminaire de probabilités XI, pp. 257–297. Springer (1977)