Adv. Appl. Prob. **47,** 902–918 (2015)
Printed in Northern Ireland

NONLINEAR FILTERING OF STOCHASTIC DYNAMICAL SYSTEMS WITH LÉVY NOISES

HUIJIE QIAO,* *Southeast University*
JINQIAO DUAN,** *Illinois Institute of Technology*

Abstract

Nonlinear filtering is investigated in a system where both the signal system and the observation system are under non-Gaussian Lévy fluctuations. Firstly, the Zakai equation is derived, and it is further used to derive the Kushner–Stratonovich equation. Secondly, by a filtered martingale problem, uniqueness for strong solutions of the Kushner–Stratonovich equation and the Zakai equation is proved. Thirdly, under some extra regularity conditions, the Zakai equation for the unnormalized density is also derived in the case of α-stable Lévy noise.

Keywords: Nonlinear filtering; Lévy process; stochastic partial differential equations with jumps; Zakai equation

2010 Mathematics Subject Classification: Primary 60G52; 60G35; 60H15

1. Introduction

Consider a partially observed system (X, Y) on a complete filtered probability space $(\Omega, \mathfrak{F}, \{\mathfrak{F}_t\}_{t\in[0,T]}, \mathbb{P})$ with $T > 0$ a fixed time. Here X_t stands for the unobservable component of the process, referred to as the signal process, whereas Y_t is the observable part, called the observation process. Given a Borel measurable function F, the nonlinear filtering problem leads to evaluating the 'filter' $\mathbb{E}[F(X_t) \mid \mathfrak{F}_t^Y]$, where $\mathfrak{F}_t^Y$ is the σ-algebra generated by $\{Y_s, 0 \leq s \leq t\}$ and $\mathbb{E}|F(X_t)| < \infty$ for $t \in [0, T]$.

Filtering problems arise from various contexts in engineering, information science, and finance. Filtering problems for systems with Gaussian noise have been widely studied; see [3], [19], and the references therein. When X and Y are continuous diffusion processes, Rozovskii [19] formulated a nonlinear filtering theory by means of stochastic evolution systems. See [20] for more recent results under more general conditions.

Nonlinear filtering problems with jump diffusion signal processes or observation processes are considered by some authors. In [14], Meyer-Brandis and Proske studied a nonlinear filtering problem with continuous diffusion signals and mixed observations, modeled by a Brownian motion and a generalized Cox process, whose jump intensity is given in terms of a Lévy measure (see Section 2.2). Later, Mandrekar *et al.* [12] added jumps to signal processes and obtained the corresponding Zakai equation. Popa and Sritharan [15] derived Zakai and Kushner–Stratonovich equations for nonlinear filtering problems with Itô–Lévy diffusion signal processes and continuous diffusion observation processes. In [4] and [5], Ceci and Colaneri dealt with a filtering problem of a jump diffusion process X and a correlated jump diffusion

Received 28 February 2014; revision received 17 June 2014.
* Postal address: Department of Mathematics, Southeast University, Nanjing, Jiangsu, 211189, China.
Email address: hjqiaogean@seu.edu.cn
** Postal address: Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616, USA.
Email address: duan@iit.edu

process Y in the one-dimensional case. They derived Kushner–Stratonovich equations by the innovation method in [4] and Zakai equations by the change of measure in [5]. These two methods are major approaches to investigate nonlinear filtering problems.

In this paper, we combine Itô–Lévy diffusion signal processes in [15] with mixed observations in [14] and study a type of nonlinear filtering problems for the multidimensional case. Compared with the results in [4] and [5], our driven processes are more general and the assumption conditions are weaker. We derive the Zakai and Kushner–Stratonovich equations and then represent the filter by an integral. It is worthwhile to mention that the integral characterization of the unnormalized filter was considered in [7] under the same assumption to Assumption 5.1 (see Section 5). Since signals are continuous diffusion processes in [7], our Lemma 5.1 is a generalization of [7, Theorem 2.2]. Note that Grigelionis and Mikulevicius [8] studied a nonlinear filtering problem with jump diffusion signal processes and observation processes and derived the Zakai and Kushner–Stratonovich equations. However, because their observation processes are very general, the integral representation for the filter was not obtained in [8].

This paper is arranged as follows. In Section 2 we introduce Sobolev spaces and α-stable Lévy processes. Zakai and Kushner–Stratonovich equations are derived in Section 3. In Section 4, by a filtered martingale problem (FMP), uniqueness for strong solutions of the Kushner–Stratonovich equation and the Zakai equation is proved. In Section 5, we prove the existence of the unnormalized and normalized densities under some regularity conditions.

The following convention will be used throughout this paper: C with or without indices will denote different positive constants (depending on the indices) whose values may change from one place to another.

2. Preliminary

2.1. Sobolev spaces

Let $\mathcal{C}_0(\mathbb{R}^n)$ be the space of continuous functions f on $\mathbb{R}^n$ satisfying $\lim_{|x|\to\infty} f(x) = 0$ with norm $\|f\|_{\mathcal{C}_0(\mathbb{R}^n)} = \sup_{x\in\mathbb{R}^n} |f(x)|$. Let $\mathcal{C}_0^2(\mathbb{R}^n)$ be the set of $f \in \mathcal{C}_0(\mathbb{R}^n)$ such that f is twice differentiable and the partial derivatives of f with order ≤ 2 belong to $\mathcal{C}_0(\mathbb{R}^n)$. Let $\mathcal{C}_c^k(\mathbb{R}^n)$ stand for the space of all k-times differentiable functions on $\mathbb{R}^n$ with compact support. Let $S(\mathbb{R}^n)$ be the Schwartz space of all rapidly decreasing real-valued $\mathcal{C}^\infty$ functions on $\mathbb{R}^n$ and $S'(\mathbb{R}^n)$ the space of all tempered distributions on $\mathbb{R}^n$. Let $\hat{f}$ and $\check{f}$ be the Fourier transform and Fourier inversion transform of $f \in S'(\mathbb{R}^n)$, respectively. That is,

$$\hat{f}(u) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}\langle u,x\rangle} f(x)\,\mathrm{d}x, \qquad \check{f}(u) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}\langle u,x\rangle} f(x)\,\mathrm{d}x$$

for all $u \in \mathbb{R}^n$. We introduce the following Sobolev space:

$$\mathbb{H}^{\lambda,2}(\mathbb{R}^n) := \{f \in S'(\mathbb{R}^n) \colon \|f\|_{\lambda,2} < \infty\}$$

for any $\lambda \in \mathbb{R}$, where

$$\|f\|_{\lambda,2}^2 := \int_{\mathbb{R}^n} (1 + |u|^2)^{\lambda} |\hat{f}(u)|^2\,\mathrm{d}u.$$

In particular, $\mathbb{H}^{0,2}(\mathbb{R}^n) = L^2(\mathbb{R}^n)$.

2.2. Symmetric α-stable Lévy processes

Definition 2.1. A process $L = (L_t)_{t\geq 0}$ with $L_0 = 0$ almost surely (a.s.) is a n-dimensional Lévy process if

(i) L has independent increments; i.e. $L_t - L_s$ is independent of $L_v - L_u$ if $(u, v) \cap (s, t) = \varnothing$;

(ii) L has stationary increments; i.e. $L_t - L_s$ has the same distribution as $L_v - L_u$ if $t - s = v - u > 0$;

(iii) L_t is stochastically continuous;

(iv) L_t is right-continuous with left limit.

The characteristic function of L_t is given by

$$\mathbb{E}[\exp\{\mathrm{i}\langle z, L_t\rangle\}] = \exp\{t\Psi(z)\}, \qquad z \in \mathbb{R}^n.$$

The function $\Psi\colon \mathbb{R}^n \to \mathcal{C}$ is called the characteristic exponent of the Lévy process L. By the Lévy–Khintchine formula, there exist a nonnegative-definite $n \times n$ matrix $\boldsymbol{Q}$, a measure ν on $\mathbb{R}^n$ satisfying

$$\nu(\{0\}) = 0, \qquad \int_{\mathbb{R}^n\setminus\{0\}} (|u|^2 \wedge 1)\nu(\mathrm{d}u) < \infty,$$

and $\boldsymbol{\gamma} \in \mathbb{R}^n$ such that

$$\Psi(z) = \mathrm{i}\langle z, \boldsymbol{\gamma}\rangle - \frac{1}{2}\langle z, \boldsymbol{Q}z\rangle + \int_{\mathbb{R}^n\setminus\{0\}} (\mathrm{e}^{\mathrm{i}\langle z,u\rangle} - 1 - \mathrm{i}\langle z, u\rangle\, \mathbf{1}_{\{|u|\leq 1\}})\nu(\mathrm{d}u).$$

The measure ν is called the Lévy measure.

Definition 2.2. For $\alpha \in (0, 2)$. An n-dimensional symmetric α-stable process L^α is a Lévy process such that its characteristic exponent Ψ is given by

$$\Psi(z) = -C_1(n, \alpha)|z|^\alpha \quad \text{for } z \in \mathbb{R}^n$$

with $C_1(n, \alpha) := \pi^{-1/2}\Gamma((1 + \alpha)/2)\Gamma(n/2)/\Gamma((n + \alpha)/2)$.

Thus, for an n-dimensional symmetric α-stable process L^α, the diffusion matrix $\boldsymbol{Q} = 0$, the drift vector $\boldsymbol{\gamma} = 0$, and the Lévy measure ν is given by

$$\nu(\mathrm{d}u) = \frac{C_2(n, \alpha)}{|u|^{n+\alpha}}\,\mathrm{d}u \quad \text{with } C_2(n, \alpha) := \frac{\alpha\Gamma((n + \alpha)/2)}{2^{1-\alpha}\pi^{n/2}\Gamma(1 - \alpha/2)}.$$

Define

$$(\mathcal{L}_\alpha f)(x) := \int_{\mathbb{R}^n\setminus\{0\}} (f(x+u) - f(x) - \langle \partial_x f(x), u\rangle\, \mathbf{1}_{\{|u|\leq 1\}})\frac{C_2(n, \alpha)}{|u|^{n+\alpha}}\,\mathrm{d}u \quad \text{for } f \in \mathcal{C}_0^2(\mathbb{R}^n).$$

Then $\mathcal{L}_\alpha$ extends uniquely to the infinitesimal generator of L_t^α and by [2, Example 3.3.8, p. 166], for every $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$,

$$(\mathcal{L}_\alpha f)(x) = C_1(n, \alpha)[-(-\Delta)^{\alpha/2} f](x).$$

Moreover, the following result is well known (see [1]).

Theorem 2.1. *Let $\mathcal{L}_\alpha$ be as above for $\alpha \in (0, 2)$ and $\mathcal{L}_2 = \Delta$, as defined on $\mathcal{C}_c^\infty(\mathbb{R}^n)$ in $L^2(\mathbb{R}^n)$. Then $\mathcal{L}_\alpha$, $0 < \alpha \leq 2$, has a unique closed extension to a self-adjoint negative operator on the domain $\mathbb{H}^{\alpha,2}(\mathbb{R}^n)$.*

3. Nonlinear filtering

In this section we study nonlinear filtering for a non-Gaussian signal-observation system, and derive Zakai and Kushner–Stratonovich equations.

Consider the following signal-observation (X_t, Y_t) system on $\mathbb{R}^n \times \mathbb{R}^m$:

$$\begin{aligned} \mathrm{d}X_t &= b_1(X_t)\,\mathrm{d}t + \sigma_1(X_t)\,\mathrm{d}B_t + \int_{\mathbb{U}_1} f_1(X_{t-}, u)\tilde{N}_p(\mathrm{d}t,\ \mathrm{d}u) \\ &\quad + \int_{\mathbb{U}\setminus\mathbb{U}_1} g_1(X_{t-}, u)N_p(\mathrm{d}t,\ \mathrm{d}u), \qquad (3.1\mathrm{a}) \\ \mathrm{d}Y_t &= b_2(t, X_t)\,\mathrm{d}t + \sigma_2(t)\,\mathrm{d}W_t + \int_{\mathbb{U}_2} f_2(t, u)\tilde{N}_\lambda(\mathrm{d}t,\ \mathrm{d}u) \\ &\quad + \int_{\mathbb{U}\setminus\mathbb{U}_2} g_2(t, u)N_\lambda(\mathrm{d}t,\ \mathrm{d}u), \qquad 0 \le t \le T, \qquad (3.1\mathrm{b}) \end{aligned}$$

where B, W are d-dimensional and m-dimensional Brownian motion, respectively, and p is a stationary Poisson point process of the class (quasi left-continuous) defined on $(\Omega, \mathfrak{F}, \{\mathfrak{F}_t\}_{t\in[0,T]}, \mathbb{P})$ with values in $\mathbb{U}$ and characteristic measure ν_1. Here ν_1 is a σ-finite measure defined on a finite-dimensional, measurable normed space $(\mathbb{U}, \mathfrak{U})$ with the norm $\|\cdot\|_{\mathbb{U}}$. Fix $\mathbb{U}_1 \in \mathfrak{U}$ with $\nu_1(\mathbb{U}\setminus\mathbb{U}_1) < \infty$ and $\int_{\mathbb{U}_1} \|u\|_{\mathbb{U}}^2\, \nu_1(\mathrm{d}u) < \infty$. Let $N_p((0,t],\ \mathrm{d}u)$ be the counting measure of p_t such that $\mathbb{E}N_p((0,t], A) = t\nu_1(A)$ for $A \in \mathfrak{U}$. Denote

$$\tilde{N}_p((0,t],\ \mathrm{d}u) := N_p((0,t],\ \mathrm{d}u) - t\nu_1(\mathrm{d}u),$$

the compensated measure of p_t. Let N_λ be an integer-valued random measure and its predictable compensator is given by $\lambda(t, X_{t-}, u)t\nu_2(\mathrm{d}u)$, where the function $\lambda(t, x, u) \in [\iota, 1)$, $0 < \iota < 1$, and ν_2 is another σ-finite measure defined on $\mathbb{U}$ with $\nu_2(\mathbb{U}\setminus\mathbb{U}_2) < \infty$ and $\int_{\mathbb{U}_2} \|u\|_{\mathbb{U}}^2\, \nu_2(\mathrm{d}u) < \infty$ for $\mathbb{U}_2 \in \mathfrak{U}$. That is, $\tilde{N}_\lambda((0,t], \mathrm{d}u) := N_\lambda((0,t],\ \mathrm{d}u) - \lambda(t, X_{t-}, u)t\nu_2(\mathrm{d}u)$ is its compensated measure. Moreover, B_t, W_t, N_p, N_λ are mutually independent. The initial value X_0 is assumed to be a random variable independent of Y_0, B_t, W_t, N_p, and N_λ. Under $\mathbb{P}$, the initial value X_0 is assumed to have a density function ρ with values in $L^2(\mathbb{R}^n)$.

The mappings $b_1 \colon \mathbb{R}^n \mapsto \mathbb{R}^n$, $b_2 \colon [0,T]\times\mathbb{R}^n \mapsto \mathbb{R}^m$, $\sigma_1 \colon \mathbb{R}^n \mapsto \mathbb{R}^{n\times d}$, $\sigma_2 \colon [0,T] \mapsto \mathbb{R}^{m\times m}$, $f_1 \colon \mathbb{R}^n \times \mathbb{U}_1 \mapsto \mathbb{R}^n$, $f_2 : [0,T]\times\mathbb{U}_2 \mapsto \mathbb{R}^m$, $g_1 \colon \mathbb{R}^n \times (\mathbb{U}\setminus\mathbb{U}_1) \mapsto \mathbb{R}^n$, and $g_2 \colon [0,T]\times(\mathbb{U}\setminus\mathbb{U}_2) \mapsto \mathbb{R}^m$ are all Borel measurable. We make the following assumptions, in order to guarantee existence and uniqueness for the solution of (3.1).

Assumption 3.1. (i) ($\boldsymbol{H}^1_{b_1,\sigma_1,f_1}$). For $x_1, x_2 \in \mathbb{R}^n$,

$$\begin{aligned} |b_1(x_1) - b_1(x_2)| &\le L_1|x_1 - x_2|\kappa_1(|x_1 - x_2|), \\ \|\sigma_1(x_1) - \sigma_1(x_2)\|^2 &\le L_1|x_1 - x_2|^2\kappa_2(|x_1 - x_2|), \\ \int_{\mathbb{U}_1} |f_1(x_1, u) - f_1(x_2, u)|^{p'}\nu_1(\mathrm{d}u) &\le L_1|x_1 - x_2|^{p'}\kappa_3(|x_1 - x_2|), \end{aligned}$$

hold for $p' = 2$ and 4, where $|\cdot|$ denotes the length of a vector in $\mathbb{R}^n$ and $\|\cdot\|$ the Hilbert–Schmidt norm from $\mathbb{R}^d$ to $\mathbb{R}^n$. Here L_1 is a constant and κ_i is a positive continuous function, bounded on $[1, \infty)$ and it satisfies

$$\lim_{x\downarrow 0} \frac{\kappa_i(x)}{\log x^{-1}} = \delta_i < \infty, \qquad i = 1, 2, 3,$$

where $\delta_i \ge 0$, $i = 1, 2, 3$, are constants.

(ii) $(\boldsymbol{H}^2_{b_1,\sigma_1,f_1})$. For $x \in \mathbb{R}^n$,

$$|b_1(x)|^2 + \|\sigma_1(x)\|^2 + \int_{\mathbb{U}_1} |f_1(x,u)|^2 \nu_1(\mathrm{d}u) \le L_1(1+|x|)^2.$$

(iii) $(\boldsymbol{H}^1_{b_2,\sigma_2,f_2})$. It holds that $\sigma_2(t)$ is invertible for $t \in [0,T]$, $b_2, \sigma_2, \sigma_2^{-1}$ are bounded by a positive constant L_2, and

$$\int_0^T \int_{\mathbb{U}_2} |f_2(s,u)|^2 \nu_2(\mathrm{d}u)\,\mathrm{d}s < \infty.$$

By [17, Theorem 1.2], (3.1) has a pathwise unique strong solution denoted by (X_t, Y_t). Set

$$\begin{aligned}\Lambda_t^{-1} := \exp\bigg\{ &-\int_0^t (\sigma_2^{-1}(s)b_2(s,X_s))^i\,\mathrm{d}W_s^i - \frac{1}{2}\int_0^t |\sigma_2^{-1}(s)b_2(s,X_s)|^2\,\mathrm{d}s \\ &- \int_0^t \int_{\mathbb{U}_2} \log\lambda(s,X_{s-},u) N_\lambda(\mathrm{d}s,\,\mathrm{d}u) - \int_0^t \int_{\mathbb{U}_2} (1-\lambda(s,X_s,u))\nu_2(\mathrm{d}u)\,\mathrm{d}s \bigg\}.\end{aligned}$$

Throughout, we use the convention that repeated indices imply summation.

Assumption 3.2. *Let*

$$\mathbb{E}\bigg[\exp\bigg\{\int_0^T \int_{\mathbb{U}_2} \frac{(1-\lambda(s,X_s,u))^2}{\lambda(s,X_s,u)} \nu_2(\mathrm{d}u)\,\mathrm{d}s\bigg\}\bigg] < \infty.$$

Set

$$M_t := -\int_0^t (\sigma_2^{-1}(s)b_2(s,X_s))^i\,\mathrm{d}W_s^i + \int_0^t \int_{\mathbb{U}_2} \frac{1-\lambda(s,X_{s-},u)}{\lambda(s,X_{s-},u)} \tilde{N}_\lambda(\mathrm{d}s,\,\mathrm{d}u),$$

and then under Assumption 3.2, M is a locally square integrable martingale. Moreover, $M_t - M_{t-} > -1$ a.s. and

$$\begin{aligned}&\mathbb{E}\big[\exp\big\{\tfrac{1}{2} < M^{\mathrm{c}}, M^{\mathrm{c}} >_T + < M^{\mathrm{d}}, M^{\mathrm{d}} >_T\big\}\big] \\ &\quad = \mathbb{E}\bigg[\exp\bigg\{\frac{1}{2}\int_0^T |\sigma_2^{-1}(s)b_2(s,X_s)|^2\,\mathrm{d}s + \int_0^T \int_{\mathbb{U}_2} \bigg(\frac{1-\lambda(s,X_s,u)}{\lambda(s,X_s,u)}\bigg)^2 \\ &\qquad\qquad \times \lambda(s,X_s,u)\nu_2(\mathrm{d}u)\,\mathrm{d}s\bigg\}\bigg] \\ &\quad < \infty,\end{aligned}$$

where M^{c} and M^{d} are continuous and purely discontinuous martingale parts of M, respectively. Thus, from [16, Theorem 6], it follows that Λ_t^{-1}, the Doléans–Dade exponential of M, is a martingale. Define a measure $\tilde{\mathbb{P}}$ via

$$\frac{\mathrm{d}\tilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}} = \frac{1}{\Lambda_T}.$$

https://doi.org/10.1239/aap/1444308887 Published online by Cambridge University Press

By the Girsanov theorem for Brownian motions and random measures, it follows that under the measure $\tilde{\mathbb{P}}$ *(3.1) is transformed as*

$$\mathrm{d}X_t = b_1(X_t)\,\mathrm{d}t + \sigma_1(X_t)\,\mathrm{d}B_t + \int_{\mathbb{U}_1} f_1(X_{t-}, u)\tilde{N}_p(\mathrm{d}t,\ \mathrm{d}u)$$
$$+ \int_{\mathbb{U}\setminus\mathbb{U}_1} g_1(X_{t-}, u)N_p(\mathrm{d}t,\ \mathrm{d}u),$$
$$\mathrm{d}Y_t = \sigma_2(t)\,\mathrm{d}\tilde{W}_t + \int_{\mathbb{U}_2} f_2(t, u)\tilde{N}(\mathrm{d}t,\ \mathrm{d}u) + \int_{\mathbb{U}\setminus\mathbb{U}_2} g_2(t, u)N_\lambda(\mathrm{d}t,\ \mathrm{d}u),$$

where

$$\tilde{W}_t := W_t + \int_0^t \sigma_2^{-1}(s)b_2(s, X_s)\,\mathrm{d}s, \qquad \tilde{N}(\mathrm{d}t,\ \mathrm{d}u) := N_\lambda(\mathrm{d}t,\ \mathrm{d}u) - \mathrm{d}t\,\nu_2(\mathrm{d}u).$$

We collect some properties of Λ_t, $\tilde{W}_t$, and $\tilde{N}$ in the next lemma.

Lemma 3.1. (i) Λ_t *satisfies the following equation:*

$$\Lambda_t = 1 + \int_0^t \Lambda_s(\sigma_2^{-1}(s)b_2(s, X_s))^i\,\mathrm{d}\tilde{W}_s^i + \int_0^t\int_{\mathbb{U}_2} \Lambda_{s-}(\lambda(s, X_{s-}, u) - 1)\tilde{N}(\mathrm{d}s,\ \mathrm{d}u).$$

(ii) *Under the measure* $\tilde{\mathbb{P}}$, $\tilde{W}_t$ *is a Brownian motion and* $\tilde{N}$ *is a Poisson martingale measure.*

Proof. For (i), by the Itô formula, we obtain

$$\begin{aligned}\Lambda_t &= 1 + \int_0^t \Lambda_s(\sigma_2^{-1}(s)b_2(s, X_s))^i\,\mathrm{d}W_s^i + \frac{1}{2}\int_0^t \Lambda_s|\sigma_2^{-1}(s)b_2(s, X_s)|^2\,\mathrm{d}s\\ &\quad + \int_0^t\int_{\mathbb{U}_2} \Lambda_{s-}(\lambda(s, X_{s-}, u) - 1)N_\lambda(\mathrm{d}s,\ \mathrm{d}u)\\ &\quad + \int_0^t\int_{\mathbb{U}_2} \Lambda_{s-}(1 - \lambda(s, X_{s-}, u))\nu_2(\mathrm{d}u)\,\mathrm{d}s + \frac{1}{2}\int_0^t \Lambda_s|\sigma_2^{-1}(s)b_2(s, X_s)|^2\,\mathrm{d}s\\ &= 1 + \int_0^t \Lambda_s(\sigma_2^{-1}(s)b_2(s, X_s))^i\,\mathrm{d}\tilde{W}_s^i + \int_0^t\int_{\mathbb{U}_2} \Lambda_{s-}(\lambda(s, X_{s-}, u) - 1)\tilde{N}(\mathrm{d}s,\ \mathrm{d}u).\end{aligned}$$

For (ii), we use [10, Theorem 3.17].

Set

$$\tilde{\mathbb{P}}_t(F) := \tilde{\mathbb{E}}[F(X_t)\Lambda_t \mid \mathfrak{F}_t^Y],$$

where $\tilde{\mathbb{E}}$ denotes expectation under the measure $\tilde{\mathbb{P}}$. The equation satisfied by $\tilde{\mathbb{P}}_t(F)$ is called the Zakai equation.

In order to derive the Zakai equation, we need the following lemma.

Lemma 3.2. *Let* $\mathfrak{F}_t^{\tilde{W}}$, $\mathfrak{F}_t^{\tilde{N}}$ *be the* σ*-algebras generated by* $\{\tilde{W}_s, 0 \leq s \leq t\}$, $\{\tilde{N}((0, s], A), 0 \leq s \leq t, A \in \mathfrak{U}\}$, *respectively. Then*

$$\mathfrak{F}_t^Y = \mathfrak{F}_t^{\tilde{W}} \vee \mathfrak{F}_t^{\tilde{N}} \vee \mathfrak{F}_0^Y.$$

Since its proof is similar to that in [19, Lemma 4, p. 228], we omit it.

As in [18], introduce the infinitesimal generator $\mathcal{L}$ of X_t,

$$
\begin{aligned}
(\mathcal{L}\varphi)(x) = {} & \frac{\partial\varphi(x)}{\partial x_i} b_1^i(x) + \frac{1}{2}\frac{\partial^2\varphi(x)}{\partial x_i \partial x_j}\sigma_1^{ik}(x)\sigma_1^{kj}(x) \\
& + \int_{\mathbb{U}\setminus\mathbb{U}_1} [\varphi(x + g_1(x,u)) - \varphi(x)]\nu_1(\mathrm{d}u) \\
& + \int_{\mathbb{U}_1} \left[\varphi(x + f_1(x,u)) - \varphi(x) - \frac{\partial\varphi(x)}{\partial x_i} f_1^i(x,u)\right]\nu_1(\mathrm{d}u) \quad \text{for } \varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n),
\end{aligned}
$$

and denote its domain by $\mathcal{D}(\mathcal{L})$.

Now, we are ready to obtain the Zakai equation for $\tilde{\mathbb{P}}_t(F)$.

Theorem 3.1. (Zakai equation.) *For $F \in \mathcal{D}(\mathcal{L})$, the Zakai equation of (3.1) is given by*

$$
\begin{aligned}
\tilde{\mathbb{P}}_t(F) = \tilde{\mathbb{P}}_0(F) + \int_0^t \tilde{\mathbb{P}}_s(\mathcal{L}F)\,\mathrm{d}s + \int_0^t \tilde{\mathbb{P}}_s(F(\sigma_2^{-1}(s)b_2(s,\cdot))^i)\,\mathrm{d}\tilde{W}_s^i \\
+ \int_0^t \int_{\mathbb{U}_2} \tilde{\mathbb{P}}_{s-}(F(\lambda(s,\cdot,u) - 1))\tilde{N}(\mathrm{d}s,\,\mathrm{d}u).
\end{aligned} \tag{3.2}
$$

Proof. Applying the Itô formula to X_t, we have

$$
\begin{aligned}
F(X_t) = {} & F(X_0) + \int_0^t \frac{\partial F(X_s)}{\partial x_i} b_1^i(X_s)\,\mathrm{d}s + \int_0^t \frac{\partial F(X_s)}{\partial x_i}\sigma_1^{ik}(X_s)\,\mathrm{d}B_s^k \\
& + \int_0^t \int_{\mathbb{U}_1} [F(X_{s-} + f_1(X_{s-},u)) - F(X_{s-})]\tilde{N}_p(\mathrm{d}s,\,\mathrm{d}u) \\
& + \int_0^t \int_{\mathbb{U}\setminus\mathbb{U}_1} [F(X_{s-} + g_1(X_{s-},u)) - F(X_{s-})]N_p(\mathrm{d}s,\,\mathrm{d}u) \\
& + \int_0^t \int_{\mathbb{U}_1} \left[F(X_{s-} + f_1(X_{s-},u)) - F(X_{s-}) - \frac{\partial F(X_s)}{\partial x_i} f_1^i(X_{s-},u)\right]\nu_1(\mathrm{d}u)\,\mathrm{d}s \\
& + \frac{1}{2}\int_0^t \frac{\partial^2 F(X_s)}{\partial x_i \partial x_j}\sigma_1^{ik}(X_s)\sigma_1^{kj}(X_s)\,\mathrm{d}s.
\end{aligned}
$$

By the mutual independence of B_t, $\tilde{N}_p(\mathrm{d}t,\,\mathrm{d}u)$, W_t, and $\tilde{N}(\mathrm{d}t,\,\mathrm{d}u)$, it is clear that

$$
F(X_t)\Lambda_t = F(X_0) + \int_0^t F(X_{s-})\,\mathrm{d}\Lambda_s + \int_0^t \Lambda_{s-}\,\mathrm{d}F(X_s).
$$

Taking the conditional expectation on both sides of this equation, we obtain

$$
\begin{aligned}
\tilde{\mathbb{E}}[F(X_t)\Lambda_t \mid \mathfrak{F}_t^Y] &= \tilde{\mathbb{E}}[F(X_0) \mid \mathfrak{F}_t^Y] + \tilde{\mathbb{E}}\left[\int_0^t F(X_{s-})\,\mathrm{d}\Lambda_s \,\middle|\, \mathfrak{F}_t^Y\right] \\
&\quad + \tilde{\mathbb{E}}\left[\int_0^t \Lambda_{s-}\,\mathrm{d}F(X_s) \,\middle|\, \mathfrak{F}_t^Y\right] \\
&=: I_1 + I_2 + I_3.
\end{aligned} \tag{3.3}
$$

We now evaluate I_1, I_2, and I_3 one by one. First, by the independence of X_0 and $\mathfrak{F}_t^Y$, we have

$$
I_1 = \tilde{\mathbb{E}}[F(X_0)] = \tilde{\mathbb{E}}[F(X_0) \mid \mathfrak{F}_0^Y] = \tilde{\mathbb{P}}_0(F). \tag{3.4}
$$

https://doi.org/10.1239/aap/1444308887 Published online by Cambridge University Press

By Lemma 3.1, it holds that

$$\begin{aligned}I_2 &= \tilde{\mathbb{E}}\bigg[\int_0^t \Lambda_s F(X_s)(\sigma_2^{-1}(s)b_2(s, X_s))^i \,\mathrm{d}\tilde{W}_s^i \,\bigg|\, \mathfrak{F}_t^Y\bigg] \\ &\quad + \tilde{\mathbb{E}}\bigg[\int_0^t \int_{\mathbb{U}_2} \Lambda_{s-} F(X_{s-})(\lambda(s, X_{s-}, u) - 1)\tilde{N}(\mathrm{d}s, \mathrm{d}u) \,\bigg|\, \mathfrak{F}_t^Y\bigg] \\ &=: I_{21} + I_{22}.\end{aligned}$$

For I_{21}, by Lemma 3.2 and [19, Theorem 1.4.7], we obtain

$$I_{21} = \int_0^t \tilde{\mathbb{E}}[\Lambda_s F(X_s)(\sigma_2^{-1}(s)b_2(s, X_s))^i \mid \mathfrak{F}_s^Y]\,\mathrm{d}\tilde{W}_s^i.$$

For I_{22}, using the same method to that in the proof of [19, Theorem 1.4.7], we can prove by the definition of stochastic integrations with respect to random measures (see [9]) and Lemma 3.2 that

$$I_{22} = \int_0^t \int_{\mathbb{U}_2} \tilde{\mathbb{E}}[\Lambda_{s-} F(X_{s-})(\lambda(s, X_{s-}, u) - 1) \mid \mathfrak{F}_s^Y]\tilde{N}(\mathrm{d}s, \mathrm{d}u).$$

Note that $\tilde{\mathbb{E}}[\Lambda_{s-} F(X_{s-})(\lambda(s, X_{s-}, u) - 1) \mid \mathfrak{F}_s^Y]$ is not predictable. Using the method of [13, Theorem 2.3], we show that it admits a predictable modification.

Thus,

$$I_2 = \int_0^t \tilde{\mathbb{P}}_s(F(\sigma_2^{-1}(s)b_2(s, \cdot))^i)\,\mathrm{d}\tilde{W}_s^i + \int_0^t \int_{\mathbb{U}_2} \tilde{\mathbb{P}}_{s-}(F(\lambda(s, \cdot, u) - 1))\tilde{N}(\mathrm{d}s, \mathrm{d}u). \tag{3.5}$$

For I_3, from the independence of B, $\tilde{N}_p$, and $\mathfrak{F}^Y$, it follows that

$$\begin{aligned}I_3 &= \tilde{\mathbb{E}}\bigg[\int_0^t \Lambda_s \frac{\partial F(X_s)}{\partial x_i} b_1^i(X_s)\,\mathrm{d}s \,\bigg|\, \mathfrak{F}_t^Y\bigg] + \tilde{\mathbb{E}}\bigg[\int_0^t \Lambda_s \frac{\partial F(X_s)}{\partial x_i} \sigma_1^{ik}(X_s)\,\mathrm{d}B_s^k \,\bigg|\, \mathfrak{F}_t^Y\bigg] \\ &\quad + \tilde{\mathbb{E}}\bigg[\int_0^t \int_{\mathbb{U}_1} \Lambda_s[F(X_{s-} + f_1(X_{s-}, u)) - F(X_{s-})]\tilde{N}_p(\mathrm{d}s, \mathrm{d}u) \mid \mathfrak{F}_t^Y\bigg] \\ &\quad + \tilde{\mathbb{E}}\bigg[\int_0^t \int_{\mathbb{U}\setminus\mathbb{U}_1} \Lambda_s[F(X_{s-} + g_1(X_{s-}, u)) - F(X_{s-})]N_p(\mathrm{d}s, \mathrm{d}u) \mid \mathfrak{F}_t^Y\bigg] \\ &\quad + \tilde{\mathbb{E}}\bigg[\frac{1}{2}\int_0^t \Lambda_s \frac{\partial^2 F(X_s)}{\partial x_i \partial x_j} \sigma_1^{ik}(X_s)\sigma_1^{kj}(X_s)\,\mathrm{d}s \,\bigg|\, \mathfrak{F}_t^Y\bigg] \\ &\quad + \tilde{\mathbb{E}}\bigg[\int_0^t \int_{\mathbb{U}_1} \Lambda_s\bigg[F(X_{s-} + f_1(X_{s-}, u)) - F(X_{s-}) - \frac{\partial F(X_s)}{\partial x_i} f_1^i(X_{s-}, u)\bigg] \\ &\qquad\qquad \times \nu_1(\mathrm{d}u)\,\mathrm{d}s \,\bigg|\, \mathfrak{F}_t^Y\bigg] \end{aligned} \tag{3.6}$$

$$\begin{aligned} &= \int_0^t \tilde{\mathbb{E}}\bigg[\Lambda_s \frac{\partial F(X_s)}{\partial x_i} b_1^i(X_s) \,\bigg|\, \mathfrak{F}_s^Y\bigg]\mathrm{d}s + \frac{1}{2}\int_0^t \tilde{\mathbb{E}}\bigg[\Lambda_s \frac{\partial^2 F(X_s)}{\partial x_i \partial x_j} \sigma_1^{ik}(X_s)\sigma_1^{kj}(X_s) \,\bigg|\, \mathfrak{F}_s^Y\bigg]\mathrm{d}s \\ &\quad + \tilde{\mathbb{E}}\bigg[\int_0^t \int_{\mathbb{U}\setminus\mathbb{U}_1} \Lambda_s[F(X_{s-} + g_1(X_{s-}, u)) - F(X_{s-})]\nu_1(\mathrm{d}u)\,\mathrm{d}s \mid \mathfrak{F}_t^Y\bigg] \\ &\quad + \int_0^t \int_{\mathbb{U}_1} \tilde{\mathbb{E}}\bigg[\Lambda_s\bigg[F(X_{s-} + f_1(X_{s-}, u)) - F(X_{s-}) - \frac{\partial F(X_s)}{\partial x_i} f_1^i(X_{s-}, u)\bigg] \,\bigg|\, \mathfrak{F}_s^Y\bigg] \\ &\qquad \times \nu_1(\mathrm{d}u)\,\mathrm{d}s \end{aligned}$$

https://doi.org/10.1239/aap/1444308887 Published online by Cambridge University Press

$$
\begin{aligned}
&= \int_0^t \tilde{\mathbb{E}}[\Lambda_s \mathcal{L} F(X_s) \mid \mathfrak{F}_s^Y] \,\mathrm{d}s \\
&= \int_0^t \tilde{\mathbb{P}}_s(\mathcal{L} F) \,\mathrm{d}s,
\end{aligned} \tag{3.7}
$$

where we have used the property of random measures (see [9]).

Combining (3.3) with (3.4), (3.5), and (3.7), we obtain the Zakai equation (3.2).

Thus, by Theorem 3.1, $\tilde{\mathbb{P}}_t(1)$ satisfies the following equation:

$$
\begin{aligned}
\tilde{\mathbb{P}}_t(1) &= 1 + \int_0^t \tilde{\mathbb{P}}_s((\sigma_2^{-1}(s) b_2(s,\cdot))^i) \,\mathrm{d}\tilde{W}_s^i + \int_0^t \int_{\mathbb{U}_2} \tilde{\mathbb{P}}_{s-}(\lambda(s,\cdot,u) - 1) \tilde{N}(\mathrm{d}s, \,\mathrm{d}u) \\
&= 1 + \int_0^t \tilde{\mathbb{P}}_s(1) \mathbb{P}_s((\sigma_2^{-1}(s) b_2(s,\cdot))^i) \,\mathrm{d}\tilde{W}_s^i \\
&\quad + \int_0^t \int_{\mathbb{U}_2} \tilde{\mathbb{P}}_{s-}(1) \mathbb{P}_{s-}(\lambda(s,\cdot,u) - 1) \tilde{N}(\mathrm{d}s, \,\mathrm{d}u).
\end{aligned} \tag{3.8}
$$

Set

$$
\tilde{M}_t := \int_0^t \mathbb{P}_s((\sigma_2^{-1}(s) b_2(s,\cdot))^i) \,\mathrm{d}\tilde{W}_s^i + \int_0^t \int_{\mathbb{U}_2} \mathbb{P}_{s-}(\lambda(s,\cdot,u) - 1) \tilde{N}(\mathrm{d}s, \,\mathrm{d}u).
$$

Since $\iota < \lambda(t,x,u) < 1$, by $(\boldsymbol{H}^1_{b_2,\sigma_2,f_2})$, the Jensen inequality, and Assumption 3.2, we obtain $\tilde{\mathbb{P}}_t(1) = \mathcal{E}(\tilde{M})_t$, the Doléans–Dade exponential of $\tilde{M}$. Thus, $\tilde{\mathbb{P}}_t(1) > 0$.

Besides, set

$$
\mathbb{P}_t(F) := \mathbb{E}[F(X_t) \mid \mathfrak{F}_t^Y],
$$

and then, by the Kallianpur–Striebel formula, the following holds:

$$
\mathbb{P}_t(F) = \mathbb{E}[F(X_t) \mid \mathfrak{F}_t^Y] = \frac{\tilde{\mathbb{E}}[F(X_t)\Lambda_t \mid \mathfrak{F}_t^Y]}{\tilde{\mathbb{E}}[\Lambda_t \mid \mathfrak{F}_t^Y]} = \frac{\tilde{\mathbb{P}}_t(F)}{\tilde{\mathbb{P}}_t(1)}. \tag{3.9}
$$

By using (3.8), the Zakai equation (3.2), and the Itô formula, together with a similar argument as in the proof of Theorem 3.1, we obtain the Kushner–Stratonovich equation satisfied by $\mathbb{P}_t(F)$.

Theorem 3.2. (Kushner–Stratonovich equation.) *For $F \in \mathcal{D}(\mathcal{L})$, $\mathbb{P}_t(F)$ solves the following equation:*

$$
\begin{aligned}
\mathbb{P}_t(F) &= \mathbb{P}_0(F) + \int_0^t \mathbb{P}_s(\mathcal{L} F) \,\mathrm{d}s + \int_0^t (\mathbb{P}_s(F(\sigma_2^{-1}(s) b_2(s,\cdot))^i) \\
&\quad - \mathbb{P}_s(F) \mathbb{P}_s((\sigma_2^{-1}(s) b_2(s,\cdot))^i)) \,\mathrm{d}\bar{W}_s^i \\
&\quad + \int_0^t \int_{\mathbb{U}_2} \frac{\mathbb{P}_{s-}(F\lambda(s,\cdot,u)) - \mathbb{P}_{s-}(F)\mathbb{P}_{s-}(\lambda(s,\cdot,u))}{\mathbb{P}_{s-}(\lambda(s,\cdot,u))} \bar{N}(\mathrm{d}s, \,\mathrm{d}u),
\end{aligned} \tag{3.10}
$$

where $\bar{W}_t := \tilde{W}_t - \int_0^t \mathbb{P}_s(\sigma_2^{-1}(s) b_2(s,\cdot)) \,\mathrm{d}s$ is the innovation process and

$$
\bar{N}(\mathrm{d}t, \,\mathrm{d}u) = N_\lambda(\mathrm{d}t, \,\mathrm{d}u) - \mathbb{P}_{t-}(\lambda(t,\cdot,u)) \nu_2(\mathrm{d}u) \,\mathrm{d}t.
$$

The Kushner–Stratonovich equation (3.10) corresponds to [8, Equation (3.2)].

4. Uniqueness for the Kushner–Stratonovich equation and the Zakai equation

In this section we first show uniqueness for strong solutions to the Kushner–Stratonovich equation by means of an FMP (see [11]), and then apply the relation between the Kushner–Stratonovich equation and the Zakai equation to prove uniqueness for the Zakai equation.

In this section we assume that $g_1(x,u) = 0$, $g_2(t,u) = 0$, and $\lambda(t,x,u) = \lambda(t,u)$. And then (3.1) is written as

$$\mathrm{d}\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} b_1(X_t) \\ b_2(t, X_t) \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \sigma_1(X_t)0 \\ 0\sigma_2(t) \end{pmatrix} \mathrm{d}\begin{pmatrix} B_t \\ W_t \end{pmatrix} + \int_{\mathbb{U}_1} \begin{pmatrix} f_1(X_{t-}, u) \\ 0 \end{pmatrix} \tilde{N}_p(\mathrm{d}t,\, \mathrm{d}u) + \int_{\mathbb{U}_2} \begin{pmatrix} 0 \\ f_2(t, u) \end{pmatrix} \tilde{N}_\lambda(\mathrm{d}t,\, \mathrm{d}u). \tag{4.1}$$

Lemma 4.1. *Suppose that* $(\boldsymbol{H}^1_{b_1,\sigma_1,f_1})$, $(\boldsymbol{H}^2_{b_1,\sigma_1,f_1})$, *and* $(\boldsymbol{H}^1_{b_2,\sigma_2,f_2})$ *are satisfied. Then the infinitesimal generator of (4.1) is given by*

$$\begin{aligned}\mathcal{L}^{X,Y} H(x,y) = {} & \frac{\partial H(x,y)}{\partial x_i} b_1^i(x) + \frac{1}{2}\frac{\partial^2 H(x,y)}{\partial x_i \partial x_j}\sigma_1^{ik}(x)\sigma_1^{kj}(x) \\ & + \int_{\mathbb{U}_1}\bigg[H(x+f_1(x,u), y) - H(x,y) - \frac{\partial H(x,y)}{\partial x_i} f_1^i(x,u)\bigg]\nu_1(\mathrm{d}u) \\ & + \frac{\partial H(x,y)}{\partial y_l} b_2^l(t,x) + \frac{1}{2}\frac{\partial^2 H(x,y)}{\partial y_l \partial y_q}\sigma_2^{lk}(t)\sigma_2^{kq}(t) \\ & + \int_{\mathbb{U}_2}\bigg[H(x, y+f_2(t,u)) - H(x,y) - \frac{\partial H(x,y)}{\partial y_l} f_2^l(t,u)\bigg]\lambda(t,u)\nu_2(\mathrm{d}u)\end{aligned}$$

for $H \in \mathcal{D}(\mathcal{L}^{X,Y})$.

By the Itô formula and the definition of the infinitesimal generator, it is easy to prove the above lemma. Therefore, we omit its proof.

Before introducing an FMP, we define several notations. Let $\mathcal{P}(\mathbb{R}^n)$ denote the set of the probability measures on $\mathbb{R}^n$ and $\mathcal{M}^+(\mathbb{R}^n)$ denote the set of positive bounded Borel measures on $\mathbb{R}^n$. For a process π-valued in $\mathcal{P}(\mathbb{R}^n)$ or $\mathcal{M}^+(\mathbb{R}^n)$, $\pi_t(F) \equiv \int_{\mathbb{R}^n} F(x)\pi_t(\cdot,\, \mathrm{d}x)$, $F \in \mathcal{C}_0^2(\mathbb{R}^n)$.

Definition 4.1. A process (π, U) defined on a probability space $(\Omega, \mathfrak{F}, \{\mathfrak{F}_t\}_{t\in[0,T]}, \mathbb{P})$, with càdlàg trajectories and values in $\mathcal{P}(\mathbb{R}^n) \times \mathbb{R}^m$, is a solution of the FMP $(\mathcal{L}^{X,Y}, X_0, Y_0)$ if π is $\mathfrak{F}_t^U$-adapted and, for all $H \in \mathcal{D}(\mathcal{L}^{X,Y})$,

$$\pi_t(H(\cdot, U_t)) - \int_0^t \pi_s(\mathcal{L}^{X,Y} H(\cdot, U_s))\,\mathrm{d}s$$

is a $(\mathbb{P}, \mathfrak{F}_t^U)$-martingale and $\mathbb{E}[\pi_0(H(\cdot, U_0))] = \mathbb{E}[H(X_0, Y_0)]$.

Definition 4.2. Uniqueness for the FMP $(\mathcal{L}^{X,Y}, X_0, Y_0)$ means that if (π, U) is a solution of the FMP $(\mathcal{L}^{X,Y}, X_0, Y_0)$, for each $t \in [0, T]$ there exists a Borel measurable $\mathcal{P}(\mathbb{R}^n)$-valued function $\tilde{H}_t$ satisfying

$$\pi_t = \tilde{H}_t(U), \qquad \mathbb{P}_t = \tilde{H}_t(Y), \qquad \mathbb{P}\text{-a.s.},$$

and (π, U) has the same distribution as $(\mathbb{P}, Y)$.

Definition 4.3. A strong solution for (3.10) is a $\mathfrak{F}_t^Y$-adapted, càdlàg, $\mathscr{P}(\mathbb{R}^n)$-valued process $\{\pi_t\}_{t\in[0,T]}$ such that $\{\pi_t\}_{t\in[0,T]}$ solves (3.10), i.e. for $F \in \mathcal{D}(\mathcal{L})$,

$$\pi_t(F) = \mathbb{P}_0(F) + \int_0^t \pi_s(\mathcal{L}F)\,\mathrm{d}s + \int_0^t (\pi_s(F(\sigma_2^{-1}(s)b_2(s,\cdot))^i) \\ - \pi_s(F)\pi_s((\sigma_2^{-1}(s)b_2(s,\cdot))^i))\,\mathrm{d}\hat{W}_s^i,$$

where $\hat{W}_t := \tilde{W}_t - \int_0^t \pi_s(\sigma_2^{-1}(s)b_2(s,\cdot))\,\mathrm{d}s$.

Definition 4.4. A strong solution for (3.2) is a $\mathfrak{F}_t^Y$-adapted, càdlàg, $\mathcal{M}^+(\mathbb{R}^n)$-valued process $\{\mu_t\}_{t\in[0,T]}$ such that $\{\mu_t\}_{t\in[0,T]}$ solves (3.2), i.e. for $F \in \mathcal{D}(\mathcal{L})$,

$$\mu_t(F) = \tilde{\mathbb{P}}_0(F) + \int_0^t \mu_s(\mathcal{L}F)\,\mathrm{d}s + \int_0^t \mu_s(F(\sigma_2^{-1}(s)b_2(s,\cdot))^i)\,\mathrm{d}\tilde{W}_s^i \\ + \int_0^t \int_{\mathbb{U}_2} \mu_{s-}(F)(\lambda(s,u)-1)\tilde{N}(\mathrm{d}s,\,\mathrm{d}u).$$

Now we prove uniqueness for solutions of the Kushner–Stratonovich equation.

Theorem 4.1. *Suppose that uniqueness holds for the FMP $(\mathcal{L}^{X,Y}, X_0, Y_0)$. Let $\{\pi_t\}_{t\in[0,T]}$ be a strong solution of (3.10). Then $\pi_t = \mathbb{P}_t$, $\mathbb{P}$-a.s. for all $t \in [0,T]$.*

Proof. For $G \in \mathcal{C}_c^\infty(\mathbb{R}^m)$, applying the Itô formula to (3.1), we obtain

$$G(Y_t) = G(Y_0) + \int_0^t \frac{\partial G(Y_s)}{\partial y} b_2(s,X_s)\,\mathrm{d}s + \frac{1}{2}\int_0^t \frac{\partial^2 G(Y_s)}{\partial y_i \partial y_j}\sigma_2^{ik}(s)\sigma_2^{kj}(s)\,\mathrm{d}s \\ + \int_0^t \int_{\mathbb{U}_2}\left[G(Y_{s-}+f_2(s,u)) - G(Y_{s-}) - \frac{\partial G(Y_{s-})}{\partial y_j} f_2^j(s,u)\right]\lambda(s,u)\nu_2(\mathrm{d}u)\,\mathrm{d}s \\ + \int_0^t \frac{\partial G(Y_s)}{\partial y}\sigma_2(s)\,\mathrm{d}W_s + \int_0^t \int_{\mathbb{U}_2} [G(Y_{s-}+f_2(s,u)) - G(Y_{s-})]\tilde{N}_\lambda(\mathrm{d}s,\,\mathrm{d}u).$$

Thus, from the above formula, Definition 4.3, and the Itô formula, it follows that, for $F \in \mathcal{D}(\mathcal{L})$,

$$\pi_t(F)G(Y_t) = \mathbb{P}_0(F)G(Y_0) + \int_0^t \pi_s[\mathcal{L}^{X,Y}(F(\cdot)G(Y_s))]\,\mathrm{d}s + \int_0^t \pi_s(F)\frac{\partial G(Y_s)}{\partial y}\sigma_2(s)\,\mathrm{d}\hat{W}_s \\ + \int_0^t G(Y_s)[\pi_s(F(\sigma_2^{-1}(s)b_2(s,\cdot))^i) - \pi_s(F)\pi_s((\sigma_2^{-1}(s)b_2(s,\cdot))^i)]\,\mathrm{d}\hat{W}_s^i \\ + \int_0^t \int_{\mathbb{U}_2} \pi_s(F)[G(Y_{s-}+f_2(s,u)) - G(Y_{s-})]\tilde{N}_\lambda(\mathrm{d}s,\,\mathrm{d}u). \tag{4.2}$$

Note that $\bar{W}_t$ is a $\mathcal{F}_t^Y$-Brownian motion (see [11]) and

$$\hat{W}_s = \bar{W}_t - \int_0^t (\pi_s(\sigma_2^{-1}(s)b_2(s,\cdot)) - \mathbb{P}_s(\sigma_2^{-1}(s)b_2(s,\cdot)))\,\mathrm{d}s.$$

Set

$$h(s) := \pi_s(\sigma_2^{-1}(s)b_2(s,\cdot)) - \mathbb{P}_s(\sigma_2^{-1}(s)b_2(s,\cdot)), \\ \tau_N := T \wedge \inf\left\{t > 0\colon \int_0^t |h(s)|^2\,\mathrm{d}s > N\right\},$$

and then τ_N is a $\mathscr{F}_t^Y$-stopping time and $\tau_N \to T$ as $N \to \infty$ by $(\boldsymbol{H}^1_{b_2,\sigma_2,f_2})$. Define the probability measure

$$\frac{\mathrm{d}\mathbb{P}_N}{\mathrm{d}\mathbb{P}} = \exp\left\{\int_0^{\tau_N} h(s)\,\mathrm{d}\bar{W}_s - \frac{1}{2}\int_0^{\tau_N} |h(s)|^2\,\mathrm{d}s\right\}.$$

Thus, from the Girsanov theorem, we have that $\hat{W}_t$ is a $\mathscr{F}_t^Y$-Brownian motion under $\mathbb{P}_N$.

Observe (4.2). Under $\mathbb{P}_N$,

$$\pi_{\tau_N\wedge t}(F)G(Y_{\tau_N\wedge t}) - \int_0^{\tau_N\wedge t} \pi_s \mathscr{L}^{X,Y}(F(\cdot)G(Y_s))\,\mathrm{d}s$$

is a $\mathscr{F}_t^Y$-martingale. Moreover, for $H \in \mathscr{D}(\mathscr{L}^{X,Y})$,

$$\pi_{\tau_N\wedge t}(H(\cdot, Y_{\tau_N\wedge t})) - \int_0^{\tau_N\wedge t} \pi_s(\mathscr{L}^{X,Y} H(\cdot, Y_s))\,\mathrm{d}s$$

is a $(\mathbb{P}_N, \mathfrak{F}_t^Y)$-martingale. By uniqueness for the FMP $(\mathscr{L}^{X,Y}, X_0, Y_0)$ and [11, Corollary 3.4], we know that there exists a $\mathscr{P}(\mathbb{R}^n)$-valued function $\tilde{H}_t$ such that

$$\pi_t\,\mathbf{1}_{\{t<\tau_N\}} = \tilde{H}_t(Y)\,\mathbf{1}_{\{t<\tau_N\}} = \mathbb{P}_t\,\mathbf{1}_{\{t<\tau_N\}}, \qquad \mathbb{P}_N\text{-a.s.}$$

Since $\mathbb{P}_N$ and $\mathbb{P}$ are equivalent, we could have

$$\pi_t\,\mathbf{1}_{\{t<\tau_N\}} = \mathbb{P}_t\,\mathbf{1}_{\{t<\tau_N\}}, \qquad \mathbb{P}\text{-a.s.}$$

Taking the limits on two sides as $N \to \infty$, we have

$$\pi_t = \mathbb{P}_t, \qquad \mathbb{P}\text{-a.s.}$$

The proof is completed.

Next, we state and prove uniqueness for solutions of the Zakai equation.

Theorem 4.2. *Suppose that uniqueness holds for the FMP $(\mathscr{L}^{X,Y}, X_0, Y_0)$. Let $\{\mu_t\}_{t\in[0,T]}$ be a strong solution of (3.2). Then $\mu_t = \tilde{\mathbb{P}}_t$, $\tilde{\mathbb{P}}$-a.s. for all $t \in [0, T]$.*

Proof. For $\varepsilon > 0$, define the stopping time

$$\tau_\varepsilon := \inf\{t > 0 \colon \mu_{t-}(1) < \varepsilon\} \wedge T.$$

Set

$$\pi_{t\wedge\tau_\varepsilon} := \frac{\mu_{t\wedge\tau_\varepsilon}}{\mu_{t\wedge\tau_\varepsilon}(1)},$$

and then $\pi_{t\wedge\tau_\varepsilon}$ is a $\mathfrak{F}_t^Y$-adapted, càdlàg, $\mathscr{P}(\mathbb{R}^n)$-valued process. For $F \in \mathscr{D}(\mathscr{L})$, applying the Itô formula to $\mu_{t\wedge\tau_\varepsilon}(F)/\mu_{t\wedge\tau_\varepsilon}(1)$, we obtain

$$\begin{aligned}\pi_{t\wedge\tau_\varepsilon}(F) = \mathbb{P}_0(F) + \int_0^{t\wedge\tau_\varepsilon} \pi_s(\mathscr{L}F)\,\mathrm{d}s + \int_0^{t\wedge\tau_\varepsilon} (\pi_s(F(\sigma_2^{-1}(s)b_2(s,\cdot))^i) \\ - \pi_s(F)\pi_s((\sigma_2^{-1}(s)b_2(s,\cdot))^i))\,\mathrm{d}\hat{W}_s^i.\end{aligned}$$

https://doi.org/10.1239/aap/1444308887 Published online by Cambridge University Press

Thus, $\{\pi_{t\wedge\tau_\varepsilon}\}_{t\in[0,T]}$ is a strong solution of (3.10). By Theorem 4.1,

$$\pi_t\, \mathbf{1}_{\{t<\tau_\varepsilon\}} = \mathbb{P}_t\, \mathbf{1}_{\{t<\tau_\varepsilon\}}, \qquad \mathbb{P}\text{-a.s.} \tag{4.3}$$

Next, observe $\mu_{t\wedge\tau_\varepsilon}(1)$ and $\tilde{\mathbb{P}}_{t\wedge\tau_\varepsilon}(1)$. Note that $\tilde{\mathbb{P}}_{t\wedge\tau_\varepsilon}(1) > 0$. By the Itô formula, it holds that

$$\frac{\mu_{t\wedge\tau_\varepsilon}(1)}{\tilde{\mathbb{P}}_{t\wedge\tau_\varepsilon}(1)} = 1 + \int_0^{t\wedge\tau_\varepsilon}\left[\frac{\mu_s((\sigma_2^{-1}(s)b_2(s,\cdot))^i)}{\tilde{\mathbb{P}}_s(1)} - \frac{\mu_s(1)}{\tilde{\mathbb{P}}_s(1)}\mathbb{P}_s((\sigma_2^{-1}(s)b_2(s,\cdot))^i)\right]\mathrm{d}\bar{W}_s^i.$$

For $s<\tau_\varepsilon$, from (4.3), it follows that

$$\frac{\mu_s(1)}{\tilde{\mathbb{P}}_s(1)}\mathbb{P}_s((\sigma_2^{-1}(s)b_2(s,\cdot))^i) = \frac{\mu_s(1)}{\tilde{\mathbb{P}}_s(1)}\pi_s((\sigma_2^{-1}(s)b_2(s,\cdot))^i) = \frac{\mu_s((\sigma_2^{-1}(s)b_2(s,\cdot))^i)}{\tilde{\mathbb{P}}_s(1)}.$$

Thus,

$$\mu_{t\wedge\tau_\varepsilon}(1) = \tilde{\mathbb{P}}_{t\wedge\tau_\varepsilon}(1), \qquad \mathbb{P}\text{-a.s.} \tag{4.4}$$

Substituting (4.3) into (4.4), we have

$$\tilde{\mathbb{P}}_t\, \mathbf{1}_{\{t<\tau_\varepsilon\}} = \mathbb{P}_t\tilde{\mathbb{P}}_t(1)\, \mathbf{1}_{\{t<\tau_\varepsilon\}} = \pi_t\mu_t(1)\, \mathbf{1}_{\{t<\tau_\varepsilon\}} = \mu_t\, \mathbf{1}_{\{t<\tau_\varepsilon\}}, \qquad \mathbb{P}\text{-a.s.}$$

Since $\tilde{\mathbb{P}}$ and $\mathbb{P}$ are equivalent,

$$\tilde{\mathbb{P}}_t\, \mathbf{1}_{\{t<\tau_\varepsilon\}} = \mu_t\, \mathbf{1}_{\{t<\tau_\varepsilon\}}, \quad \tilde{\mathbb{P}}\text{-a.s.}$$

Thus, $\tau_\varepsilon \geq \inf\{t>0\colon \tilde{\mathbb{P}}_t(1)<\varepsilon\}\wedge T$. Note that $\inf\{t>0\colon \tilde{\mathbb{P}}_t(1)<\varepsilon\}\wedge T = T$ when ε is small enough. So $\tau_\varepsilon = T$ and $\tilde{\mathbb{P}}_t = \mu_t$, $\tilde{\mathbb{P}}$-a.s. The proof is completed.

Remark 4.1. Uniqueness of the FMP $(\mathcal{L}^{X,Y}, X_0, Y_0)$ is required in Theorems 4.1 and 4.2. In fact, by [11, Theorem 3.3], this could be assured by adding regular conditions to $b_1, b_2, \sigma_1, \sigma_2, f_1$, and f_2 such that $\mathcal{D}(\mathcal{L}^{X,Y})$ is a dense algebra in $\mathcal{C}_0(\mathbb{R}^n\times\mathbb{R}^m)$. Here we do not impose concrete conditions.

5. The Zakai equation for the unnormalized density in the case of α-stable Lévy noises

In this section we solve (3.2) and obtain the unnormalized and normalized densities under extra regularity conditions in the case of α-stable Lévy noises.

Assume that $f_1(x,u) = g_1(x,u) = u$, $\mathbb{U} = \mathbb{R}^n\setminus\{0\}$, $\mathbb{U}_1 = \{u\in\mathbb{U}\colon |u|\leq 1\}$, $\nu_1(\mathrm{d}u) = (C_2(n,\alpha)/|u|^{n+\alpha})\,\mathrm{d}u$, and $\lambda(t,x,u) = \lambda(t,u)$. Thus, (3.1) becomes

$$\begin{aligned}
\mathrm{d}X_t &= b_1(X_t)\,\mathrm{d}t + \sigma_1(X_t)\,\mathrm{d}B_t + \int_{|u|\leq 1} u\tilde{N}_p(\mathrm{d}t,\,\mathrm{d}u) + \int_{|u|>1} uN_p(\mathrm{d}t,\,\mathrm{d}u),\\
\mathrm{d}Y_t &= b_2(t,X_t)\,\mathrm{d}t + \sigma_2(t)\,\mathrm{d}W_t + \int_{\mathbb{U}_2} f_2(t,u)\tilde{N}_\lambda(\mathrm{d}t,\,\mathrm{d}u) + \int_{\mathbb{U}\setminus\mathbb{U}_2} g_2(t,u)N_\lambda(\mathrm{d}t,\,\mathrm{d}u).
\end{aligned}$$

Under the measure $\tilde{\mathbb{P}}$, by the Itô–Lévy decomposition theorem, the above system is converted to

$$\begin{aligned}
\mathrm{d}X_t &= b_1(X_t)\,\mathrm{d}t + \sigma_1(X_t)\,\mathrm{d}B_t + \mathrm{d}L_t^\alpha,\\
\mathrm{d}Y_t &= \sigma_2(t)\,\mathrm{d}\tilde{W}_t + \int_{\mathbb{U}_2} f_2(t,u)\tilde{N}(\mathrm{d}t,\,\mathrm{d}u) + \int_{\mathbb{U}\setminus\mathbb{U}_2} g_2(t,u)N_\lambda(\mathrm{d}t,\,\mathrm{d}u),
\end{aligned}$$

https://doi.org/10.1239/aap/1444308887 Published online by Cambridge University Press

where

$$L_t^\alpha := \int_0^t \int_{|u|\leq 1} u\tilde{N}_p(\mathrm{d}s,\, \mathrm{d}u) + \int_0^t \int_{|u|>1} uN_p(\mathrm{d}s,\, \mathrm{d}u)$$

is a symmetric α-stable process independent of B_t. The infinitesimal generator $\mathcal{L}$ of X_t is given by

$$\begin{aligned}(\mathcal{L}\varphi)(x) &= \frac{\partial\varphi(x)}{\partial x_i} b_1^i(x) + \frac{1}{2}\frac{\partial^2\varphi(x)}{\partial x_i \partial x_j}\sigma_1^{ik}(x)\sigma_1^{kj}(x) \\ &\quad + \int_{\mathbb{R}^n\setminus\{0\}} \left[\varphi(x+u) - \varphi(x) - \frac{\partial\varphi(x)}{\partial x_i} u^i\, \mathbf{1}_{\{|u|\leq 1\}}\right] \frac{C_2(n,\alpha)}{|u|^{n+\alpha}}\,\mathrm{d}u \\ &= \frac{\partial\varphi(x)}{\partial x_i} b_1^i(x) + \frac{1}{2}\frac{\partial^2\varphi(x)}{\partial x_i \partial x_j}\sigma_1^{ik}(x)\sigma_1^{kj}(x) + (\mathcal{L}_\alpha\varphi)(x) \quad \text{for } \varphi \in \mathcal{C}_0^2(\mathbb{R}^n).\end{aligned}$$

To solve the Zakai equation for the unnormalized density, we need some stronger assumptions.

Assumption 5.1. *For $i = 1, 2, \ldots, n$ and $k = 1, 2, \ldots, d$, b_1^i is one time differentiable in x and σ_1^{ik} is twice differentiable in x. Moreover, b_1, σ_1 and all their derivatives are bounded by L_3.*

Define the adjoint operator $\mathcal{L}^*$ of $\mathcal{L}$ for $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ as

$$\begin{aligned}(\mathcal{L}^*\psi)(x) &= -\frac{\partial}{\partial x_i}(b_1^i(x)\psi(x)) + \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_j}(\sigma_1^{ik}(x)\sigma_1^{kj}(x)\psi(x)) + (\mathcal{L}_\alpha\psi)(x) \\ &= (\mathcal{A}\psi)(x) + (\mathcal{B}\psi)(x),\end{aligned}$$

where

$$\begin{aligned}(\mathcal{A}\psi)(x) &:= \frac{\partial\psi(x)}{\partial x_i}\left[-b_1^i(x) + \frac{1}{2}\frac{\partial(\sigma_1^{ik}(x)\sigma_1^{kj}(x))}{\partial x_j} + \frac{1}{2}\frac{\partial(\sigma_1^{jk}(x)\sigma_1^{ki}(x))}{\partial x_j}\right] \\ &\quad + \frac{1}{2}\frac{\partial^2\psi(x)}{\partial x_i\partial x_j}\sigma_1^{ik}(x)\sigma_1^{kj}(x) + (\mathcal{L}_\alpha\psi)(x), \\ (\mathcal{B}\psi)(x) &:= \left[-\frac{\partial b_1^i(x)}{\partial x_i} + \frac{1}{2}\frac{\partial^2(\sigma_1^{ik}(x)\sigma_1^{kj}(x))}{\partial x_i \partial x_j}\right]\psi(x).\end{aligned}$$

Thus, by [18, Lemma 4.2], $\mathcal{A}$ generates a strongly continuous contraction semigroup on $\mathcal{C}_0(\mathbb{R}^n)$.

Consider the following stochastic partial differential equation with jumps on the filtered probability space $(\Omega, \mathfrak{F}, \{\mathfrak{F}_t^Y\}_{t\in[0,T]}, \tilde{\mathbb{P}})$:

$$\begin{aligned}\mathrm{d}\tilde{\rho}_t &= [\mathcal{A}\tilde{\rho}_t + \mathcal{B}\tilde{\rho}_t]\,\mathrm{d}t + \tilde{\rho}_t(\sigma_2^{-1}(t)b_2(t,\cdot))^i\,\mathrm{d}\tilde{W}_t^i + \int_{\mathbb{U}_2}\tilde{\rho}_{t-}(\lambda(t,u)-1)\tilde{N}(\mathrm{d}t,\,\mathrm{d}u), \\ \tilde{\rho}_0 &= \Lambda_T^{-1}\rho.\end{aligned} \tag{5.1}$$

By [13, Theorem 2.4], (5.1) admits a unique mild solution $\tilde{\rho}_t$, satisfying the following equation:

$$\begin{aligned}\tilde{\rho}_t &= \mathrm{e}^{t\mathcal{A}}\Lambda_T^{-1}\rho + \int_0^t \mathrm{e}^{(t-s)\mathcal{A}}\mathcal{B}\tilde{\rho}_s\,\mathrm{d}s + \int_0^t \mathrm{e}^{(t-s)\mathcal{A}}\tilde{\rho}_s(\sigma_2^{-1}(s)b_2(s,\cdot))^i\,\mathrm{d}\tilde{W}_s^i \\ &\quad + \int_0^t\int_{\mathbb{U}_2}\mathrm{e}^{(t-s)\mathcal{A}}\tilde{\rho}_{s-}(\lambda(s,u)-1)\tilde{N}(\mathrm{d}s,\,\mathrm{d}u).\end{aligned} \tag{5.2}$$

Moreover, $\tilde{\rho}_t$ is predictable and $\tilde{\mathbb{E}}[\sup_{t\in[0,T]}\|\tilde{\rho}_t\|_{0,2}^2] < \infty$.

https://doi.org/10.1239/aap/1444308887 Published online by Cambridge University Press

Note that if more conditions are required, we could obtain some better properties of $\tilde{\rho}_t$ (see [13, Theorem 2.7, 2.8]). Here we concentrate on providing a more general result.

Lemma 5.1. *We have $\tilde{\rho}_t$ is density-valued and for $F \in \mathcal{C}_c^{\infty}(\mathbb{R}^n)$, the following representation holds:*

$$\tilde{\mathbb{P}}_t(F) = \int_{\mathbb{R}^n} F(x)\tilde{\rho}_t(x)\,\mathrm{d}x. \tag{5.3}$$

Proof. For $F \in \mathcal{C}_c^{\infty}(\mathbb{R}^n)$, by the Sobolev imbedding theorem (see [19, Theorem 3.4.4]), the following holds:

$$|\tilde{\mathbb{P}}_t(F)| \le \sup_{x\in\mathbb{R}^n} |F(x)|\tilde{\mathbb{P}}_t(1) \le C\|F\|_{q,2}\tilde{\mathbb{P}}_t(1),$$

where $q > n/2$. Besides, $\tilde{\mathbb{P}}_t(1)$ satisfies (3.8). By the Burkholder–Davis–Gunder inequality and case 3 of Assumption 3.1 ($\boldsymbol{H}^1_{b_2,\sigma_2,f_2}$), we obtain

$$\begin{aligned}
\tilde{\mathbb{E}}\Big[\sup_{t\in[0,T]} \tilde{\mathbb{P}}_t(1)^2\Big] &\le 3 + 3C\tilde{\mathbb{E}}\Big[\int_0^T \tilde{\mathbb{P}}_s(1)^2 \sum_{i=1}^n |\mathbb{P}_s((\sigma_2^{-1}(s)b_2(s,\cdot))^i)|^2\,\mathrm{d}s\Big] \\
&\quad + 3C\tilde{\mathbb{E}}\Big[\int_0^T \int_{\mathbb{U}_2} \tilde{\mathbb{P}}_s(1)^2|\lambda(s,u)-1|^2\nu_2(\mathrm{d}u)\,\mathrm{d}s\Big] \\
&\le 3 + 3C\int_0^T \tilde{\mathbb{E}}\Big[\sup_{s\in[0,t]} \tilde{\mathbb{P}}_s(1)^2\Big]\Big(nL_2^4 + \int_{\mathbb{U}_2} |\lambda(t,u)-1|^2\nu_2(\mathrm{d}u)\Big)\,\mathrm{d}t.
\end{aligned}$$

From the Grönwall inequality, it follows that

$$\Big(\tilde{\mathbb{E}}\Big[\sup_{t\in[0,T]} \tilde{\mathbb{P}}_t(1)\Big]\Big)^2 \le \tilde{\mathbb{E}}\Big[\sup_{t\in[0,T]} \tilde{\mathbb{P}}_t(1)^2\Big] < \infty.$$

Thus, there exists a Ω' with $\Omega' \subset \Omega$ and $\tilde{\mathbb{P}}(\Omega') = 1$, such that for $\omega \in \Omega'$, $\sup_{t\in[0,T]} \tilde{\mathbb{P}}_t(1) < \infty$. So the Riesz representation theorem implies that there exists a u_t taking values in $\mathbb{H}^{-q,2}(\mathbb{R}^n)$ and satisfying

$$\tilde{\mathbb{P}}_t(F) = \int_{\mathbb{R}^n} F(x)u_t(x)\,\mathrm{d}x. \tag{5.4}$$

Next, observe u_t. Firstly, using the method of [13, Theorem 2.3], we could show that the solution $\tilde{\mathbb{P}}_t(F)$ of (3.2) is predictable. And it follows from (5.4) that u_t is predictable. Secondly, replacing $\tilde{\mathbb{P}}_t(F)$ by $\int_{\mathbb{R}^n} F(x)u_t(x)\,\mathrm{d}x$ in (3.2), we obtain

$$\begin{aligned}
\int_{\mathbb{R}^n} F(x)u_t(x)\,\mathrm{d}x &= \int_{\mathbb{R}^n} F(x)u_0(x)\,\mathrm{d}x + \int_{\mathbb{R}^n} \mathcal{L}F(x)\Big(\int_0^t u_s(x)\,\mathrm{d}s\Big)\,\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^n} F(x)\Big(\int_0^t u_s(x)(\sigma_2^{-1}(s)b_2(s,x))^i\,\mathrm{d}\tilde{W}_s^i\Big)\,\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^n} F(x)\Big(\int_0^t \int_{\mathbb{U}_2} u_s(x)(\lambda(s,u)-1)\tilde{N}(\mathrm{d}s,\,\mathrm{d}u)\Big)\,\mathrm{d}x.
\end{aligned}$$

From this we know that $u_t(x)$ is a weak solution to (5.1). By [6, Theorem 5.4, p. 121], $u_t(x)$ is also a mild solution to (5.1). So $u_t = \tilde{\rho}_t$. The proof is completed.

By Lemma 5.1, we know that $\tilde{\rho}_t$ is the unnormalized density and (5.2) is the Zakai equation for the unnormalized density.

Now we are in the position to state the main result of this section.

Theorem 5.1. (Integral representation of the filter.) *There exists a* $\mathfrak{F}_t^Y$*-adapted density function* $\rho_t(x)$ *given by*

$$\rho_t(x) = \frac{\tilde{\rho}_t(x)}{\int_{\mathbb{R}^n} \tilde{\rho}_t(x)\,\mathrm{d}x}, \tag{5.5}$$

such that for $F \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, $\mathbb{P}_t(F) = \int_{\mathbb{R}^n} F(x)\rho_t(x)\,\mathrm{d}x$.

Proof. For $F \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, by (3.9) and (5.3), we conclude that

$$\mathbb{E}[F(X_t) \mid \mathfrak{F}_t^Y] = \mathbb{P}_t(F) = \frac{\tilde{\mathbb{P}}_t(F)}{\tilde{\mathbb{P}}_t(1)} = \frac{\int_{\mathbb{R}^n} F(x)\tilde{\rho}_t(x)\,\mathrm{d}x}{\int_{\mathbb{R}^n} \tilde{\rho}_t(x)\,\mathrm{d}x} = \int_{\mathbb{R}^n} F(x)\bigg(\frac{\tilde{\rho}_t(x)}{\int_{\mathbb{R}^n} \tilde{\rho}_t(x)\,\mathrm{d}x}\bigg)\,\mathrm{d}x.$$

Thus, the conditional distribution $\mathbb{P}[X_t \in \mathrm{d}x \mid \mathfrak{F}_t^Y]$ is absolutely continuous with respect to the Lebesgue measure with the filtering density function

$$\rho_t(x) = \frac{\tilde{\rho}_t(x)}{\int_{\mathbb{R}^n} \tilde{\rho}_t(x)\,\mathrm{d}x}.$$

The proof is completed.

This theorem says that the solution to the nonlinear filtering under Lévy noises (i.e. the ‘filter’) or the conditional expectation $\mathbb{P}_t(F) = \mathbb{E}[F(X_t) \mid \mathfrak{F}_t^Y]$, may be represented as an integral via a normalized density ρ_t. This normalized density is obtained by solving (5.2) together with the normalization (5.5).

Acknowledgements

We would like to thank Dan Crisan and Remigijus Mikulevicius for helpful discussions and we also wish to thank the anonymous referee for his/her suggestions and improvements. This work was partly supported by National Science Foundation of China (grant no. 11001051 and grant no. 11371352) and by the National Science Foundation (grant no. 1025422).

References

[1] Albeverio, S., Rüdiger, B. and Wu, J.-L. (2000). Invariant measures and symmetry property of Lévy type operators. *Potential Anal.* **13,** 147–168.

[2] Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus*, 2nd edn. Cambridge University Press.

[3] Bain, A. and Crisan, D. (2009). *Fundamentals of Stochastic Filtering*. Springer, New York.

[4] Ceci, C. and Colaneri, K. (2012). Nonlinear filtering for jump diffusion observations. *Adv. Appl. Prob.* **44,** 678–701.

[5] Ceci, C. and Colaneri, K. (2014). The Zakai equation of nonlinear filtering for jump-diffusion observations: existence and uniqueness. *Appl. Math. Optimization* **69,** 47–82.

[6] Da Prato, G. and Zabczyk, J. (1992). *Stochastic Equations in Infinite Dimensions*. Cambridge University Press.

[7] Frey, R., Schmidt, T. and Xu, L. (2013). On Galerkin approximations for the Zakai equation with diffusive and point process observations. *SIAM J. Numer. Anal.* **51,** 2036–2062.

[8] Grigelionis B. and Mikulevicius R. (2011). *Nonlinear Filtering Equations for Stochastic Processes with Jumps*. In *The Oxford Handbook of Nonlinear Filtering*, Oxford University Press, pp. 95–128.

[9] Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd edn. North-Holland, Amsterdam.

[10] Jacod, J. and Shiryaev, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer, Berlin.

[11] Kurtz, T. G. and Ocone, D. L. (1988). Unique characterization of conditional distributions in nonlinear filtering. *Ann. Prob.* **16,** 80–107.

[12] Mandrekar, V., Meyer-Brandis, T. and Proske, F. (2011). A Bayes formula for nonlinear filtering with Gaussian and Cox noise. *J. Prob. Statist.* **2011,** 259091.

[13] Marinelli, C., Prévôt, C. and Röckner, M. (2010). Regular dependence on initial data for stochastic evolution equations with multiplicative Poisson noise. *J. Functional Anal.* **258,** 616–649.

[14] Meyer-Brandis, T. and Proske, F. (2004). Explicit solution of a non-linear filtering problems for Lévy processes with application to finance. *Appl. Math. Optimization* **50,** 119–134.

[15] Popa, S. and Sritharan, S. S. (2009). Nonlinear filtering of Itô–Lévy stochastic differential equations with continuous observations. *Commun. Stoch. Anal.* **3,** 313–330.

[16] Protter, P. and Shimbo, K. (2008). No arbitrage and general semimartingales. In *Markov Processes and Related Topics,* Institute of Mathematical Statistics, Beachwood, OH, pp. 267–283.

[17] Qiao, H. (2014). Euler–Maruyama approximation for SDEs with jumps and non-Lipschitz coefficients. *Osaka J. Math.* **51,** 47–66.

[18] Qiao, H. and Duan, J. (2012). Stationary measures for stochastic differential equations with jumps. Preprint. Available at http://arxiv.org/abs/1209.0658.

[19] Rozovskii, B. L. (1990). *Stochastic Evolution Systems, Linear Theory and Applications to Nonlinear Filtering* (Math. Appl. (Soviet Ser.) **35**). Kluwer, Dordrecht.

[20] Zhang, X. (2011). Stochastic partial differential equations with unbounded and degenerate coefficients. *J. Differential Equat.* **250,** 1924–1966.