ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications xxx (xxxx) xxx

stochastic processes and their applications

www.elsevier.com/locate/spa

Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise

Vasileios Maroulas[a,*], Xiaoyang Pan[a], Jie Xiong[b]

[a] *Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA*
[b] *Department of Mathematics, Southern University of Science and Technology, Shenzhen, China*

Received 2 January 2018; received in revised form 1 February 2019; accepted 21 February 2019
Available online xxxx

Abstract

In this paper, we focus on the asymptotic behavior of the optimal filter where both signal and observation processes are driven by Lévy noises. Indeed, we study large deviations for the case where the signal-to-noise ratio is small by considering weak convergence arguments. To that end, we first prove the uniqueness of the solution of the controlled Zakai and Kushner–Stratonovich equations. For this, we employ a method which transforms the associated equations into SDEs in an appropriate Hilbert space. Next, taking into account the controlled analogue of Zakai and Kushner–Stratonovich equations, respectively, the large deviation principle follows by employing the existence, uniqueness and tightness of the solutions.

Keywords: Nonlinear filtering; Lévy noise; Uniqueness; Large deviations

1. Introduction

Stochastic filtering deals with the estimation problem under partial information. Given two stochastic processes, the signal process and observation process, the filtering problem aims to estimate a functional of the signal based on the partially observed data. This paper focuses on a general model in which signal and observation processes are driven by Lévy noises. Several models in finance engineering, biology etc., e.g. see the partial list [1,8,9,12,15,17,22,23,25,26,31,32], have considered such stochastic dynamics driven by a pertinent Lévy noise for describing partially received information with discontinuities or jumps in a time interval. By the

* Corresponding author.
E-mail address: vmaroula@utk.edu (V. Maroulas).

https://doi.org/10.1016/j.spa.2019.02.009
0304-4149/

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

same token, there exists an array of studies which encounter the analysis of optimal filter in a Lévy environment, for instance, [6,7,13,20,28,29]. We investigate a small-noise large deviation principle (LDP) of the optimal filter, which is basically a solution of the Kushner–Stratonovich stochastic partial differential equation (SPDE). Such analysis is associated with the rare events of a small signal-to-noise ratio.

The early work [14] derived large deviations for the conditional density for diffusion systems in which both the signal and the observation noises were small. The study [27] established a quenched large deviation principle with small-noise observations. In a similar setup as herein, [33] took aim in a model where the signal was a diffusion process and the observation process was driven by a Brownian motion. A fractional Brownian motion model was studied in [24]. This work investigates the large deviations for the optimal filter where both signal and observation processes are driven by Lévy noises.

The strategy we apply is to prove the Laplace principle, which is equivalent to the large deviation principle, by using a weak convergence argument, proposed in [3,4]. This weak convergence method is an approach that has been increasingly used, e.g., the large deviations for a variety of SPDEs [2,3,5,10,21,30,35,36], based on variational representations of the functionals of driving Brownian motions and Poisson random measures. The novelty of such a method is that, it does not require the exponential continuity or exponential tightness, and in contrast only basic qualitative properties of existence, uniqueness of controlled analogues of the stochastic dynamical systems of interest are needed to be shown.

To that end, we first prove the uniqueness of the controlled unnormalized filtering equation, i.e. the controlled Zakai equation, and subsequently this of the controlled filtering equation, the controlled Kushner–Stratonovich equation. Some studies have been devoted to verifying the uniqueness of the filtering equation. In [7], the uniqueness was shown by the Filtered Martingale Problem approach which was proposed in [18], however, that model has a limitation that the signal and observation are driven by the same Poisson random measure having common jump times. A more general setting was suggested in [29] and the uniqueness was also proved by the approach of Filtered Martingale Problem, however, it was shown under the assumption that the correlated Poisson random measure is independent of the signal. Moreover, in these studies the uniqueness of the Filtered Martingale Problem was assumed when the Filtered Martingale Problem method was used, which requires the regularity conditions of the coefficients of the equations. In our work, based on a method using the Brownian motion semigroup, we bypass these restrictions and establish the uniqueness for the general model with a mild assumption on the coefficients of the Poissonian noise.

The paper is organized as follows. Section 2 discusses preliminaries on the introduction to the controlled Poisson random measure and a general criterion of large deviations. We prove the uniqueness of the solutions to the Zakai and Kushner–Stratonovich equations in Section 3 which are given in Theorems 3.3 and 3.4 respectively. Section 4 focuses on the establishment of the large deviation principle for the optimal filter. We start with deriving the controlled version and zero-noise version of Zakai equation, then the existence and uniqueness of these two versions of Zakai equation are verified and finally the results are concluded by demonstrating the sufficient conditions presented in Propositions 4.2 and 4.3.

2. Preliminaries

Our problem concerns with a small-noise large deviation principle for the optimal filter where the respective SPDEs are driven by a pertinent Brownian motion and a Poisson random measure. This section lists definitions and notations used later on. Most of these notations can be also found in [4] but they are presented here for the sake of completeness.

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

2.1. Definitions and conventional notations

Let $\mathbb{U}$ be a locally compact Polish space and denote by $\mathcal{M}(\mathbb{U})$ the space of all measures ν on $(\mathbb{U}, \mathcal{B}(\mathbb{U}))$, satisfying $\nu(A) < \infty$ for every compact subset A of $\mathbb{U}$, and $\mathcal{B}(\mathbb{U})$ is the Borel σ-field on $\mathbb{U}$. Endow $\mathcal{M}(\mathbb{U})$ with the weakest topology such that for every continuous function f on $\mathbb{U}$ with compact support, the function $\nu \to \int_{\mathbb{U}} f(u)\nu(du)$, $\nu \in \mathcal{M}(\mathbb{U})$ is continuous. This topology can be metrized such that $\mathcal{M}(\mathbb{U})$ is a Polish space, e.g. see [4]. For a fixed $T \in (0, \infty)$, denote by $\mathbb{M} = \mathcal{M}(\mathbb{U}_T)$ the space of measures on $\mathbb{U}_T = [0, T] \times \mathbb{U}$ and let $\nu_T = \lambda_T \otimes \nu$, λ_T is the Lebesgue measure on $[0, T]$. We recall that a general Poisson random measure (PRM) $\mathbf{n}$ on $\mathbb{U}_T$ with intensity measure ν_T is an $\mathbb{M}$-valued random variable such that for each $A \in \mathcal{B}(\mathbb{U}_T)$ with $\nu_T(A) < \infty$, $\mathbf{n}(A)$ is Poisson distributed with mean $\nu_T(A)$ and for disjoint $A_1, \ldots, A_k \in \mathcal{B}(\mathbb{U}_T)$, $\mathbf{n}(A_1), \ldots, \mathbf{n}(A_k)$ are mutually independent random variables.

For any $m \geq 1$, let $\mathbb{W}_m = C([0, T], \mathbb{R}^m)$ be the space of all continuous functions from $[0, T]$ to $\mathbb{R}^m$, and $D([0, T], \mathcal{E})$ denote the space of right continuous functions with left limits from $[0, T]$ to a Polish space $\mathcal{E}$. Take $\mathbb{V} = \mathbb{W}_m \times \mathbb{W}_n \times \mathbb{M}$ and let $\mathbb{P}$ be the probability measure on $(\mathbb{V}, \mathcal{B}(\mathbb{V}))$ such that (i) $N : \mathbb{V} \to \mathbb{M}$ is a Poisson random measure with intensity measure $\theta\nu_T$, and $\nu_T(A) < \infty$ for all $A \in \mathcal{B}(\mathbb{U}_T)$; (ii) $W : \mathbb{V} \to \mathbb{W}_m$ is a $\mathbb{R}^m$-valued Brownian motion and $B : \mathbb{V} \to \mathbb{W}_n$ is a $\mathbb{R}^n$-valued Brownian motion; and (iii) $\{W_t\}_{t\in[0,T]}$, $\{B_t\}_{t\in[0,T]}$ and $\{N([0, t] \times A), t \in [0, T]\}$ are $\mathcal{G}_t$-martingales for every $A \in \mathcal{B}(\mathbb{U})$, where the filtration $\mathcal{G}_t := \sigma\{N([0, t] \times A) - \theta t\nu(A), W_s, B_s : 0 \leq s \leq t, A \in \mathcal{B}(\mathbb{U})\}$.

To adopt the strategy of weak convergence arguments in order to prove the large deviations, we introduce a properly controlled Poisson random measure. Define $\mathbb{Y}_T = [0, T] \times \mathbb{Y}$, where $\mathbb{Y} = \mathbb{U} \times [0, \infty)$ and then denote $\bar{\mathbb{M}} = \mathcal{M}(\mathbb{Y}_T)$. Suppose $\bar{N}$ is a Poisson random measure with points on $\bar{\mathbb{V}} = \mathbb{W}_m \times \mathbb{W}_n \times \bar{\mathbb{M}}$ with intensity measure $\bar{\nu}_T = \lambda_T \otimes \nu \otimes \lambda_\infty$ where λ_∞ is the Lebesgue measure on $[0, \infty)$. Similarly abusing notations, B and W, are Brownian motions on $\bar{\mathbb{V}}$. Next define $(\bar{\mathbb{P}}, \{\bar{\mathcal{G}}_t\})$ on $(\bar{\mathbb{V}}, \mathcal{B}(\bar{\mathbb{V}}))$ analogous to $(\mathbb{P}, \{\mathcal{G}_t\})$ by replacing $(N, \theta\nu_T)$ with $(\bar{N}, \bar{\nu}_T)$. Consider the $\bar{\mathbb{P}}$-completion of the filtration $\{\mathcal{G}_t\}$ and denote it by $\{\bar{\mathcal{F}}_t\}$. We denote by $\bar{\mathcal{P}}$ the predictable σ-field on $[0, T] \times \bar{\mathbb{V}}$ with the filtration $\{\bar{\mathcal{F}}_t : 0 \leq t \leq T\}$ on $(\bar{\mathbb{V}}, \mathcal{B}(\bar{\mathbb{V}}))$. Let $\bar{\mathcal{A}}$ be the class of all $(\bar{\mathcal{P}} \otimes \mathcal{B}(\mathbb{U})) \setminus \mathcal{B}[0, \infty)$ measurable maps $\phi : \mathbb{U}_T \times \bar{\mathbb{M}} \to [0, \infty)$. For the variable $\phi \in \bar{\mathcal{A}}$ which basically controls the intensity at time s on location u, define the counting process N^ϕ on $\mathbb{U}_T$ by

$$N^\phi((0, t] \times A) = \int_{(0,t]\times A\times(0,\infty)} 1_{[0,\phi(s,u)]}(r)\bar{N}(ds, du, dr) \tag{1}$$

where $t \in [0, T]$, $A \in \mathcal{B}(\mathbb{U})$. If $\phi(s, u) = \theta$ for all $(s, u) \in \mathbb{U}_T$, then we write $N^\phi = N^\theta$, where N^θ has the same distribution on $\bar{\mathbb{M}}$ with respect to $\bar{\mathbb{P}}$ as N has on $\mathbb{M}$ with respect to $\mathbb{P}$. For any $\phi \in \bar{\mathcal{A}}$, the quantity

$$L_T(\phi) = \int_{\mathbb{U}_T} l(\phi(t, u))\ \nu(du)ds \tag{2}$$

is well-defined as a $[0, \infty]$-valued random variable where $l(r) := r \log r - r + 1, r \in [0, \infty)$. We denote by $L^2([0, T], \mathbb{R}^m)$ the Hilbert space from $[0, T]$ to $\mathbb{R}^m$ satisfying $|\psi(s)|^2 = \sum_{i=1}^m \psi_i(s)^2 < \infty$. Define $\mathcal{P}_2 =$

$$\left\{\psi = (\psi_i)_{i=1}^m : \psi_i \text{ is } \bar{\mathcal{P}} \setminus \mathcal{B}(\mathbb{R}) \text{ measurable and } \int_0^T |\psi(s)|^2 ds < \infty, \bar{\mathbb{P}} - a.s.\right\}$$

and set $\mathcal{U} = \mathcal{P}_2 \times \bar{\mathcal{A}}$. For $\psi \in \mathcal{P}_2$ define

$$\tilde{L}_T(\psi) := \frac{1}{2}\int_0^T \psi(s)^2 ds \tag{3}$$

and for $u = (\psi, \phi) \in \mathcal{U}$, set

$$\bar{L}_T(u) = L_T(\phi) + \tilde{L}_T(\psi). \tag{4}$$

2.2. A general criterion of large deviations

The theory of small-noise large deviations concerns with the asymptotic behavior of solutions of SPDEs, say $\{X^\epsilon\}, \epsilon > 0$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which converge exponentially fast as $\epsilon \to 0$. The decay rate is expressed via a rate function. An equivalent argument of the large deviations principle is the Laplace principle. A reader may refer to [11, Theorem 1.2.1 and Theorem 1.2.3].

Definition 2.1. A function $I : \mathcal{E} \to [0, \infty]$ is called a rate function on $\mathcal{E}$, if for each $M < \infty$ the level set $\{x \in \mathcal{E} : I(x) \leq M\}$ is a compact subset of $\mathcal{E}$. The family $\{X^\epsilon\}$ is said to satisfy the Laplace principle on $\mathcal{E}$ with rate function I, if for all bounded continuous functions h mapping $\mathcal{E}$ into $\mathbb{R}$,

$$\lim_{\epsilon\to 0}\left|\epsilon \log \mathbb{E}_{x_0}\left\{\exp\left[-\frac{1}{\epsilon}h(X^\epsilon)\right]\right\} + \inf_{x\in\mathcal{E}}\{h(x)+I(x)\}\right| = 0.$$

Next, a set of sufficient conditions for a uniform large deviation principle for functionals of a Brownian motion and Poisson random measure is presented. Consider the family of measurable maps, for any $\epsilon > 0$, $\mathcal{G}^\epsilon : \mathbb{W}_m \times \mathbb{M} \to \mathcal{E}$ defined as follows:

$$\{X^\epsilon \doteq \mathcal{G}^\epsilon(\sqrt{\epsilon}W, \epsilon N^{\epsilon^{-1}})\}.$$

Define the space, for some $M \in \mathbb{N}$, $\tilde{S}^M := \{\psi \in L^2([0,T], \mathbb{R}^m) : \tilde{L}_T(\psi) \leq M\}$ and $S^M := \{\phi : \mathbb{U}_T \to [0,\infty) : L_T(\phi) \leq M\}$, where L_T and $\tilde{L}_T$ are defined in Eqs. (2) and (3), respectively. For a function ϕ, define a measure $\nu_T^\phi \in \mathbb{M}$, such that

$$\nu_T^\phi(A) = \int_A \phi(s,u)\nu(du)ds, \quad A \in \mathcal{B}(\mathbb{U}_T).$$

Throughout we encapsulate the topology on S^M obtained through this identification which makes S^M a compact space. Let $\bar{S}^M = \tilde{S}^M \times S^M$ with the usual product topology and $\mathbb{S} = \cup_{M=1}^\infty \bar{S}^M$. Define the space of controls $\mathcal{U}^M = \{u = (\psi, \phi) \in \mathcal{U} : u(\omega) \in \bar{S}^M, \bar{\mathbb{P}} \text{ a.e. } \omega\}$, where $\mathcal{U} = \mathcal{P}_2 \times \bar{\mathcal{A}}$.

The following condition in [4] plays a key role in proving large deviation estimates for the filtering equation driven by a Brownian motion and an independent Poisson random measure. The corresponding rate function is given in Eq. (5).

Condition 2.1. *There exists a measurable map $\mathcal{G}^0 : \mathbb{W}_m \times \mathbb{M} \to \mathcal{E}$ such that the following holds.*

1. *For $M \in \mathbb{N}$, let $(f_n, g_n), (f, g) \in \bar{S}^M$ be such that $(f_n, g_n) \to (f, g)$ as $n \to \infty$, then*

$$\mathcal{G}^0\left(\int_0^\cdot f_n(s)ds, \nu_T^{g_n}\right) \to \mathcal{G}^0\left(\int_0^\cdot f(s)ds, \nu_T^{g}\right).$$

2. *For $M \in \mathbb{N}$, let $\xi^\epsilon = (\psi^\epsilon, \phi^\epsilon), \xi = (\psi, \phi) \in \mathcal{U}^M$ be such that $\xi^\epsilon \xrightarrow{d} \xi$, where $\xrightarrow{d}$ denotes convergence in distribution, as $\epsilon \to 0$. Then*

$$\mathcal{G}^\epsilon(\sqrt{\epsilon}W + \int_0^\cdot \psi^\epsilon(s)ds, \epsilon N^{\epsilon^{-1}\phi^\epsilon}) \xrightarrow{d} \mathcal{G}^0(\int_0^\cdot \psi(s)ds, \nu_T^\phi).$$

For $\zeta \in \mathcal{E}$, define $\mathbb{S}_\zeta = \{u \in \mathbb{S} : \zeta = \mathcal{G}^0(\int_0^\cdot \psi(s)ds, \nu_T^\phi)\}$. Let $I : \mathcal{E} \to [0, \infty]$ be defined by

$$I(\zeta) = \inf_{u \in \mathbb{S}_\zeta} \left\{ \bar{L}_T(u) \right\}. \tag{5}$$

By convention, $I(\zeta) = \infty$ if $\mathbb{S}_\zeta = \emptyset$. Under Condition 2.1, we have the following theorem shown in [4].

Theorem 2.1. *For $\epsilon > 0$, let $X^\epsilon \doteq \mathcal{G}^\epsilon(\sqrt{\epsilon}W, \epsilon N^{\epsilon^{-1}})$, and suppose that Condition 2.1 holds. Then $I : \mathcal{E} \to [0, \infty]$, defined by Eq. (5), is a rate function on $\mathcal{E}$ and the family $\{X^\epsilon\}$ satisfies the large deviation principle on $\mathcal{E}$ with rate function I.*

3. Existence and uniqueness of filtering equations

3.1. Existence

This section first presents the filtering model and filtering equations for the system driven by Lévy noise. Consider the following signal-observation system (X_t, Y_t) on $\mathbb{R}^d \times \mathbb{R}^m$:

$$dX_t = b_1(X_t)dt + \sigma_1(X_t)dB_t + \int_{\mathbb{U}_1} f_1(X_{t-}, u)\tilde{N}_p(dt, du) + \int_{\mathbb{U}\setminus\mathbb{U}_1} g_1(X_{t-}, u)N_p(dt, du), \tag{6a}$$

$$dY_t = b_2(t, X_t)dt + \sigma_2(t)dW_t + \int_{\mathbb{U}_2} f_2(t, u)\tilde{N}_\lambda(dt, du) + \int_{\mathbb{U}\setminus\mathbb{U}_2} g_2(t, u)N_\lambda(dt, du), \tag{6b}$$

where B_t, W_t are n-dimensional and m-dimensional Brownian motions respectively defined on the filtered probability space $(\mathbb{V}, \mathcal{B}(\mathbb{V}), \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$. N_p is a Poisson random measure such that $\mathbb{E}N_p([0, t] \times A) = t\nu_1(A)$, for any $A \in \mathcal{B}(\mathbb{U})$ with $\nu_1(A) < \infty$, and $\nu_1(\mathbb{U} \setminus \mathbb{U}_1) < \infty$, $\int_{\mathbb{U}_1} \|u\|_{\mathbb{U}}^2 \nu_1(du) < \infty$ where $\|\cdot\|_{\mathbb{U}}$ denotes the norm on the measurable space $(\mathbb{U}, \mathcal{B}(\mathbb{U}))$ and $\mathbb{U}_1 \subset \mathbb{U}$. Denote the compensated measure $\tilde{N}_p([0, t] \times du) = N_p([0, t], du) - t\nu_1(du)$. Let $N_\lambda(dt \times du)$ be an integer-valued random measure and its predictable compensator is given by $\lambda(t, X_{t-}, u)dt\nu_2(du)$, where the function $\lambda(t, x, u) \in [l, 1), 0 < l < 1$, and $\tilde{N}_\lambda([0, t] \times A) = N_\lambda([0, t] \times A) - \int_0^t \int_A \lambda(s, X_{s-}, u)\nu_2(du)ds$ such that for each $A \in \mathcal{B}(\mathbb{U})$, $\nu_2(A) < \infty$, and $\nu_2(\mathbb{U} \setminus \mathbb{U}_2) < \infty$, $\int_{\mathbb{U}_2} \|u\|_{\mathbb{U}}^2 \nu_2(du) < \infty$ with $\mathbb{U}_2 \subset \mathbb{U}$. Moreover, B_t, W_t, N_p, N_λ are mutually independent.

We assume that the mappings $b_1 : \mathbb{R}^d \to \mathbb{R}^d$, $b_2 : [0, T] \times \mathbb{R}^d \to \mathbb{R}^m$, $\sigma_1 : \mathbb{R}^d \to \mathbb{R}^{d\times n}$, $\sigma_2 : [0, T] \to \mathbb{R}^{m\times m}$, $f_1 : \mathbb{R}^d \times \mathbb{U}_1 \to \mathbb{R}^d$, $f_2 : [0, T] \times \mathbb{U}_2 \to \mathbb{R}^m$, $g_1 : \mathbb{R}^d \times (\mathbb{U} \setminus \mathbb{U}_1) \to \mathbb{R}^d$, and $g_2 : [0, T] \times (\mathbb{U} \setminus \mathbb{U}_2) \to \mathbb{R}^m$ are all Borel measurable, and satisfy the following conditions:

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

Assumption 3.1. For each $x_1, x_2 \in \mathbb{R}^d$, there exists a constant $K > 0$ such that

$$|b_1(x_1) - b_1(x_2)|^2 + |\sigma_1(x_1) - \sigma_1(x_2)|^2 + \int_{\mathbb{U}_1} |f_1(x_1, u) - f_1(x_2, u)|^2 \nu_1(du) \leq K|x_1 - x_2|^2,$$

$$\int_{\mathbb{U}\setminus\mathbb{U}_1} |g_1(x_1, u) - g_1(x_2, u)|\nu_1(du) \leq K|x_1 - x_2|,$$

where $|\cdot|$ denotes the Hilbert–Schmidt norm for a matrix and the length for a vector.

Assumption 3.2. $\sigma_2(t)$ is invertible for $t \in [0, T]$, and for each $x \in \mathbb{R}^d$,

$$|\sigma_1(x)|^2 + \int_{\mathbb{U}_1} |f_1(x, u)|^2 \nu_1(du) + |b_2(t, x)|^2 + \left|\sigma_2^{-1}(t)\right|^2 + \int_{\mathbb{U}_2} |f_2(t, u)|^2 \nu_2(du) \leq K.$$

$$\int_{\mathbb{U}\setminus\mathbb{U}_2} |g_2(t, u)|\nu_2(du) \leq K.$$

Note that $Y_t = f_2(t, p_\lambda(t)) + g_2(t, p_\lambda(t))1_{D_\lambda}(t)$ is observable, where D_λ and p_λ are the jumping times and locations of the random measure N_λ. As g_2 describes the large jumps while f_2 the small ones, we assume they have disjoint ranges and, for each t, $f_2(t, \cdot)$ and $g_2(t, \cdot)$ are invertible functions. Namely, we assume that N_λ is observable. Let $\tilde{Y}_t$ be given by $d\tilde{Y}_t = b_2(t, X_t)dt + \sigma_2(t)dW_t$. Then, $\mathcal{F}_t^Y = \mathcal{F}_t^{\tilde{Y}} \vee \mathcal{F}_t^{N_\lambda}$. Based on the discussion above, we make the following assumption throughout this article.

Assumption 3.3. For each t, $f_2(t, \cdot)$ and $g_2(t, \cdot)$ are invertible functions with disjoint ranges.

Set $\pi_t(F) := \mathbb{E}(F(X_t)|\mathcal{F}_t^Y)$, for any $F \in C_b^2(\mathbb{R}^d)$, where $C_b^2(\mathbb{R}^d)$ denotes the set of all bounded functions $F : \mathbb{R}^d \to \mathbb{R}^d$ that have continuous second order derivative. Taking into account the Radon–Nikodym derivative there is an equivalent probability measure $\tilde{\mathbb{P}}$ such that, $d\tilde{\mathbb{P}} = \Lambda_T^{-1} d\mathbb{P}$, where

$$\Lambda_T^{-1} = \exp\left(-\int_0^T \sigma_2^{-1}(s) b_2(s, X_s)^* dW_s - \frac{1}{2}\int_0^T \left|\sigma_2^{-1}(s) b_2(s, X_s)\right|^2 ds \right. \\ \left. - \int_0^T \int_{\mathbb{U}} \log \lambda(s, X_{s-}, u) N_\lambda(ds, du) - \int_0^T \int_{\mathbb{U}} (1 - \lambda(s, X_s, u))\nu_2(du)ds \right). \tag{7}$$

where $*$ denotes the transpose operator. In turn, the Kallianpur–Striebel formula [34] gives $\pi_t(F) = \frac{\mu_t(F)}{\mu_t(1)}$, where $\mu_t(F) := \tilde{\mathbb{E}}(F(X_t)\Lambda_t|\mathcal{F}_t^Y)$ and $\tilde{\mathbb{E}}$ denotes expectation under the measure $\tilde{\mathbb{P}}$. Now we are ready to derive Zakai and Kushner–Stratonovich equations.

Proposition 3.1 (*Zakai Equation*)**.** *Assume Assumptions 3.1–3.3. For $F \in D(\mathcal{L})$, the Zakai equation of Eq. (6) is given by*

$$\mu_t(F) = \mu_0(F) + \int_0^t \mu_s(\mathcal{L}F)ds + \sum_{i=1}^m \int_0^t \mu_s(F(\sigma_2^{-1}(s)b_2(s, \cdot))^i)d\tilde{W}_s^i \\ + \int_0^t \int_{\mathbb{U}} \mu_{s-}(F(\lambda(s, \cdot, u) - 1))\tilde{N}(ds, du), \tag{8}$$

where $\tilde{W}_t = W_t + \int_0^t \sigma_2^{-1}(s) b_2(s, X_s) ds$ *and* $\tilde{N}(dt, du) = N_\lambda(dt, du) - dt\nu_2 d(u)$, *and for any* $f \in C_b^2(\mathbb{R}^d)$, *the infinitesimal generator,* $\mathcal{L}$, *is given by*

$$\begin{aligned}\mathcal{L}f(x) = &\sum_{i=1}^d \frac{\partial f(x)}{\partial x_i} b_1^i(x) + \frac{1}{2}\sum_{i,j=1}^d \sum_{k=1}^n \frac{\partial^2 f(x)}{\partial x_i \partial x_j}\sigma_1^{ik}(x)\sigma_1^{jk}(x)\\ &+ \int_{\mathbb{U}\setminus\mathbb{U}_1} [f(x+g_1(x,u)) - f(x)]\nu_1(du)\\ &+ \int_{\mathbb{U}_1} [f(x+f_1(x,u)) - f(x) - \sum_{i=1}^d \frac{\partial f(x)}{\partial x_i} f_1^i(x,u)]\nu_1(du),\end{aligned} \tag{9}$$

where σ_1^{ik} *is the* (i,k)*th entry of the diffusion coefficient* σ_1.

Proof. By Itô's formula, we have

$$d\Lambda_t = \Lambda_t \sigma_2^{-1}(t) b_2(t, X_t) d\tilde{W}_t + \int_{\mathbb{U}} \Lambda_{t-}(\lambda(t, X_{t-}, u) - 1)\tilde{N}(dt, du),$$

and

$$\begin{aligned}dF(X_t) = &\mathcal{L}F(X_t)dt + \nabla F(X_t)\sigma_1(X_t)dB_t\\ &+ \int_{\mathbb{U}\setminus\mathbb{U}_1} [F(x+g_1(x,u)) - F(x)]\tilde{N}_p(dt, du)\\ &+ \int_{\mathbb{U}_1} [F(x+f_1(x,u)) - F(x)]\tilde{N}_p(dt, du).\end{aligned}$$

By Itô's formula again, we have

$$\begin{aligned}d(\Lambda_t F(X_t)) = &\Lambda_t F(X_t)\sigma_2^{-1}(t) b_2(t, X_t) d\tilde{W}_t\\ &+ \int_{\mathbb{U}} \Lambda_{t-} F(X_{t-})(\lambda(t, X_{t-}, u) - 1)\tilde{N}(dt, du)\\ &+ \Lambda_t \mathcal{L}F(X_t)dt + \nabla F(X_t)\sigma_1(X_t)dB_t\\ &+ \int_{\mathbb{U}\setminus\mathbb{U}_1} \Lambda_{t-}[F(x+g_1(x,u)) - F(x)]\tilde{N}_p(dt, du)\\ &+ \int_{\mathbb{U}_1} \Lambda_{t-}[F(x+f_1(x,u)) - F(x)]\tilde{N}_p(dt, du).\end{aligned}$$

Taking conditional expectations on both sides, we arrive at (8). ■

Using the Kallianpur–Striebel formula, we obtain the following Kushner–Stratonovich equation.

Proposition 3.2. *Assume Assumptions 3.1–3.3. For* $F \in D(\mathcal{L})$, *the solution of the following equation exists*

$$\begin{aligned}\pi_t(F) = &\pi_0(F) + \int_0^t \pi_s(\mathcal{L}F)ds\\ &+ \sum_{i=1}^m \int_0^t (\pi_s(F(\sigma_2^{-1}(s)b_2(s,\cdot))^i) - \pi_s(F)\pi_s(\sigma_2^{-1}(s)b_2(s,\cdot))^i) d\hat{W}_s^i\\ &+ \int_0^t \int_{\mathbb{U}} \frac{\pi_{s-}(F\lambda(s,\cdot,u)) - \pi_{s-}(F)\pi_{s-}(\lambda(s,\cdot,u))}{\pi_{s-}(\lambda(s,\cdot,u))}\hat{N}(ds, du),\end{aligned} \tag{10}$$

where $\hat{W}_t = \tilde{W}_t - \int_0^t \pi_s(\sigma_2^{-1}(s)b_2(s,\cdot))ds$ *is the innovation process and* $\hat{N}(dt,du) = N_\lambda(dt,du) - \pi_{t-}(\lambda(t,\cdot,u))\nu_2(du)dt$.

3.2. Uniqueness

In this section, we prove the uniqueness for the solutions to the Zakai and Kushner–Stratonovich equations for the signal-observation model (6). Although the uniqueness was investigated in [29], it was assumed that the Poisson noise in the observation is independent of the signal, i.e., $\lambda(t,x,u) = \lambda(t,u)$. This reduces the complexity of the Zakai equation; that is, the Zakai equation is independent of the Poisson noise. This makes the problem more tractable since the Poissonian part in Eq. (10) vanishes, see [29, Section 4]. Furthermore, therein, the uniqueness of the Filtered Martingale Problem is assumed, and regularity conditions on the coefficients of the signal and observation processes are required (see [29, Remark 4.1] and [18]). Next, we show the uniqueness of Zakai and Kushner–Stratonovich equations by bypassing the above restrictive assumptions and instead imposing the following mild assumption.

Assumption 3.4. $\left|\det\left(J_{f_1}+I\right)\right| > \frac{1}{C}$ and $\left|\det\left(J_{g_1}+I\right)\right| > \frac{1}{C}$ for a constant $C > 0$, where J_{f_1} and J_{g_1} are the Jacobian matrices of f_1 and g_1 with respect to x, respectively.

The uniqueness for the solution to Zakai equation is proved by transforming it to an SDE in a pertinent Hilbert space and by making use of estimates based on Hilbert-space techniques, which was studied in [19,34]. Recall that the optimal filter $\mathbb{E}(F(X_t)|\mathcal{F}_t^Y)$ is the solution to the filtering Eq. (10) characterized by the conditional probability π_t. Denote $\mathcal{P}(\mathbb{R}^d)$ the collection of all Borel probability measures on $\mathbb{R}^d$ such that $\pi_t \in \mathcal{P}(\mathbb{R}^d)$. Denote by $\langle \nu, F\rangle$ the integral of a function F with respect to a measure ν, e.g., for any $F \in C_b(\mathbb{R}^d)$, $\mathbb{E}(F(X_t)|\mathcal{F}_t^Y) = \langle \pi_t, F\rangle$. Let $\mathcal{M}_F(\mathbb{R}^d)$ be the collection of all finite Borel measures on $\mathbb{R}^d$ such that the unnormalized filter μ_t is an $\mathcal{M}_F(\mathbb{R}^d)$-valued process. Let $H_0 = L^2(\mathbb{R}^d)$ be the Hilbert space consisting of square-integrable functions on $\mathbb{R}^d$ with the usual L^2-norm and the inner product given by $\|\phi\|_0^2 = \int_{\mathbb{R}^d}|\phi(x)|^2dx$ and $\langle\phi,\psi\rangle_0 = \int_{\mathbb{R}^d}\phi(x)\psi(x)dx$. We introduce an operator to transform a measure-valued process to an H_0-valued process. Denote by $\mathcal{M}_G(\mathbb{R}^d)$ the space of finite signed measures on $\mathbb{R}^d$. For any $\nu \in \mathcal{M}_G(\mathbb{R}^d)$ and $\delta > 0$, let

$$(T_\delta\nu)(x) = \int_{\mathbb{R}^d} G_\delta(x-y)\nu(dy), \tag{11}$$

where G_δ is the heat kernel given by $G_\delta(x) = (2\pi\delta)^{-\frac{d}{2}}\exp\left(-\frac{|x|^2}{2\delta}\right)$. We use the same notation as in Eq. (11) for the Brownian motion semigroup on H_0, i.e., for $t \geq 0$, define operator $T_t : H_0 \to H_0$ by $T_t\phi(x) = \int_{\mathbb{R}^d} G_t(x-y)\phi(y)dy$, for any $\phi \in H_0$. Then Lemma 3.1 obtains bounds for the partial derivative of T_δ, and Lemma 3.2 is directly applied to Theorem 3.1. The reader should refer to [34] for the proofs of these two lemmas.

Lemma 3.1.

(i) The family of operators $\{T_t : t \geq 0\}$ *forms a contraction semigroup on* H_0, *i.e. for any* $t, s \geq 0$ *and* $\phi \in H_0$, *we have* $T_{t+s} = T_tT_s$ *and* $\|T_t\phi\|_0 \leq \|\phi\|_0$.

(ii) If $\nu \in \mathcal{M}_G(\mathbb{R}^d)$ *and* $\delta > 0$, *then* $T_\delta\nu \in H_0$.

(iii) If $\nu \in \mathcal{M}_G(\mathbb{R}^d)$ *and* $\delta > 0$, *then* $\|T_{2\delta}|\nu|\|_0 \leq \|T_\delta|\nu|\|_0$, *where* $|\nu|$ *is the total variation measure of* ν.

Lemma 3.2. *For any $\delta > 0$, $\nu \in \mathcal{M}_G(\mathbb{R}^d)$ and $\phi \in H_0$, we have*

(i)

$$\langle T_\delta \nu, \phi\rangle_0 = \langle \nu, T_\delta \phi\rangle. \tag{12}$$

(ii) If, in addition, $\partial_i \phi \in H_0$, where $\partial_i \phi = \frac{\partial \phi}{\partial x_i}$, then

$$\partial_i T_\delta \phi = T_\delta \partial_i \phi. \tag{13}$$

The next theorem presents an expression for the expectation of the transformation applying to the solution to Zakai equation. The proof is delegated to Appendix A.

Theorem 3.1. *Let $\mu_t \in \mathcal{M}_F(\mathbb{R}^d)$ be a solution to Zakai equation (8) and let $Z_t^\delta = T_\delta \mu_t$. Considering the probability measure $\tilde{\mathbb{P}}$ defined by (7), the following holds.*

$$\tilde{\mathbb{E}}\|Z_t^\delta\|_0^2 = A_1 - 2A_2 + A_3 + 2A_4 + A_5 + A_6 + A_7, \tag{14}$$

where

$$\begin{aligned}
A_1 &= \|Z_0^\delta\|_0^2, \quad A_2 = \sum_{i=1}^d \int_0^t \tilde{\mathbb{E}}\langle Z_s^\delta, \partial_i T_\delta(b_1^i \mu_s)\rangle_0 ds, \\
A_3 &= \sum_{i,j=1}^d \sum_{k=1}^n \int_0^t \tilde{\mathbb{E}}\langle Z_s^\delta, \partial_{ij}^2 T_\delta(\sigma_1^{ik}\sigma_1^{kj}\mu_s)\rangle_0 ds, \\
A_4 &= \int_0^t \int_{\mathbb{U}\setminus\mathbb{U}_1} [\tilde{\mathbb{E}}\langle Z_s^\delta, \tilde{Z}_s^{\delta,g_1}\rangle_0 - \tilde{\mathbb{E}}\|Z_s^\delta\|_0^2]\nu_1(du)ds, \\
A_5 &= \int_0^t \int_{\mathbb{U}_1} [\tilde{\mathbb{E}}\langle Z_s^\delta, \tilde{Z}_s^{\delta,f_1}\rangle_0 - \tilde{\mathbb{E}}\|Z_s^\delta\|_0^2 \\
&\quad - \sum_{i=1}^d \tilde{\mathbb{E}}\langle Z_s^\delta, \partial_i T_\delta(f_1^i(\cdot,u)\mu_s)\rangle_0]\nu_1(du)ds, \\
A_6 &= \int_0^t \tilde{\mathbb{E}}\|\langle T_\delta(\sigma_2^{-1}(s)b_2(s,\cdot)\mu_s)\rangle\|_0^2 ds, \\
A_7 &= \int_0^t \int_{\mathbb{U}} \tilde{\mathbb{E}}\|T_\delta\left((\lambda(s,\cdot,u)-1)\mu_s\right)\|_0^2 \nu_2(du)ds;
\end{aligned} \tag{15}$$

and

$$\langle Z_t^\delta, \tilde{Z}_t^{\delta,f_1}\rangle_0 = \sum_\eta \langle Z_t^\delta, \eta\rangle_0 \langle Z_t^\delta, \eta(\cdot + f_1(\cdot,u))\rangle_0, \tag{16}$$

$$\langle Z_t^\delta, \tilde{Z}_t^{\delta,g_1}\rangle_0 = \sum_\eta \langle Z_t^\delta, \eta\rangle_0 \langle Z_t^\delta, \eta(\cdot + g_1(\cdot,u))\rangle_0, \tag{17}$$

here the set of functions $\{\eta\}$ is a complete orthonormal system (CONS) of H_0.

In order to get estimates of the terms as defined in Eq. (15), we proceed with the following lemmas.

Lemma 3.3. *Suppose Assumption 3.4 holds. Then there exists a constant $C_0 > 0$ such that*

$$\left|\langle Z_t^\delta, \tilde{Z}_t^{\delta,f_1}\rangle_0\right| \le C_0\|Z_t^\delta\|_0^2, \tag{18}$$

and

$$\left|\langle Z_t^\delta, \tilde{Z}_t^{\delta,g_1}\rangle_0\right| \le C_0\|Z_t^\delta\|_0^2. \tag{19}$$

Proof. We only need to show inequality (18). Let $y = x + f_1(x,u)$ and assume $x = h(y,u)$. Denote the Jacobian of h with respect to y by $|\partial_y h(y,u)|$. According to Assumption 3.4, $\left|\det\left(I + J_{f_1}\right)\right| > \frac{1}{C}$, then

$$\left|\partial_y h(y,u)\right| = \left|\det\left(I + J_{f_1}\right)\right|^{-1} \le C. \tag{20}$$

Note that

$$\begin{aligned}\langle Z_t^\delta, \eta(\cdot + f_1(\cdot,u))\rangle_0 =& \int_{\mathbb{R}^d} Z_t^\delta(h(y,u))\eta(y)|\partial_y h(y,u)|dy\\ =&\langle Z_t^\delta(h(\cdot,u))|\partial_y h(\cdot,u)|, \eta\rangle_0.\end{aligned}$$

Furthermore, summing over $\{\eta\}$ in a CONS of H_0, we have

$$\sum_\eta \langle Z_t^\delta(h(\cdot,u))|\partial_y h(\cdot,u)|, \eta\rangle_0 \langle Z_t^\delta, \eta\rangle_0 = \langle Z_t^\delta(h(\cdot,u))|\partial_y h(\cdot,u)|, Z_t^\delta\rangle_0.$$

Hence, by Eq. (20) we get

$$\begin{aligned}\|Z_t^\delta(h(y,u))|\partial_y h(y,u)|\|_0^2 =& \int_{\mathbb{R}^d} |T_\delta\mu_t(h(y,u))|^2\,|\partial_y h(y,u)|^2 dy \\ =& \int_{\mathbb{R}^d} |T_\delta\mu_t(x)|^2\,|\partial_y h(y,u)| dx \le C\|Z_t^\delta\|_0^2,\end{aligned} \tag{21}$$

and the bound of Eq. (18) then follows from the Cauchy–Schwarz inequality. ■

Lemma 3.4, verified in [19, Lemma 3.2], is useful to estimate the transformation T_δ and the derivatives of T_δ in Lemma 3.5 and Theorem 3.2.

Lemma 3.4. *Let $(H, \mathcal{H}, \eta)$ be a measure space and $\mathbb{H} = Ł_2(\eta)$. Let $\phi_i : \mathbb{R}^d \to \mathbb{H}, i = 1,2$ such that there exists a constant $K > 0$, for any $x \in \mathbb{R}^d$, $\|\phi_i(x)\|_{\mathbb{H}} \le K$. Let $\zeta \in \mathcal{M}_G(\mathbb{R}^d)$. Then there exists a constant $K_1 \equiv K_1(\phi_i)$ such that*

$$\|\|T_\delta(\phi_1\zeta)\|_0\|_{\mathbb{H}} \le K_1\|T_\delta(|\zeta|)\|_0. \tag{22}$$

If, additionally, (we are interested in $H = \mathbb{U}$ with $\eta = \nu_1$) $\|\phi_i(x) - \phi_i(y)\|_{\mathbb{H}} \le K|x-y|$, then

$$|\langle T_\delta(\phi_2\zeta), \partial_i T_\delta(\phi_1\zeta)\rangle_{H_0\otimes\mathbb{H}}| \le K_1\||T_\delta(|\zeta|)\|_0^2, \tag{23}$$

where $\otimes$ denotes convolution.

The next lemma gives a bound of the derivatives of the transformation on σ_1. The proof is shown in Appendix B.

Lemma 3.5. *There exists constant K_1 such that for any $\zeta \in \mathcal{M}_G(\mathbb{R}^d)$, we have*

$$\sum_{i,j=1}^d \langle T_\delta\zeta, \partial_{ij}^2 T_\delta(\sigma_1\sigma_1^*)_{ij}\zeta\rangle_0 + \sum_{i=1}^n \left\|\sum_{i=1}^d \partial_i T_\delta(\sigma_1^{ik}\zeta)\right\|_0^2 \le K_1\|T_\delta(|\zeta|)\|_0^2. \tag{24}$$

Now applying Lemmas 3.3–3.5 yields the following theorem.

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

Theorem 3.2. *If μ is a measure-valued solution of the Zakai equation (8) and $Z^\delta = T_\delta \mu$, then*

$$\tilde{\mathbb{E}}\|Z_t^\delta\|_0^2 \leq \|Z_0^\delta\|_0^2 + K_1 \int_0^t \tilde{\mathbb{E}}\|T_\delta(|\mu_s|)\|_0^2 ds, \tag{25}$$

where K_1 is a suitable constant.

Proof. Consider Eq. (14) such that $\tilde{\mathbb{E}}\|Z_t^\delta\|_0^2 = A_1 - 2A_2 + A_3 + 2A_4 + A_5 + A_6 + A_7$ and $A_1, \ldots, A_7$ are defined in Eq. (15). Then by inequality (23), A_2 is bounded by a constant times $\|T_\delta(|\mu_s|)\|_0^2$. The bound for A_3 follows from inequality (24). A_4 follows from Lemma 3.3 and inequality (22), and A_5 is bounded by Lemma 3.3, inequalities (22) and (23). The bounds for A_6 and A_7 follow from inequality (22). ■

Corollary 3.1. *If μ is a measure-valued solution to Eq. (8) and $\mu_0 \in H_0$, then $\mu_t \in H_0$ a.s. and $\tilde{\mathbb{E}}\|\mu_t\|_0^2 < \infty$, for all $t \geq 0$.*

Proof. Eq. (25) yields that $\tilde{\mathbb{E}}\|Z_t^\delta\|_0^2 \leq \|Z_0^\delta\|_0^2 + K_1 \int_0^t \tilde{\mathbb{E}}\|Z_s^\delta\|_0^2 ds$. By Gronwall's inequality, we get $\tilde{\mathbb{E}}\|Z_t^\delta\|_0^2 \leq \|Z_0^\delta\|_0^2 e^{K_1 t}$. Note that $\lim_{\delta\to 0} \langle Z_t^\delta, F\rangle_0 = \lim_{\delta \to 0} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} G_\delta(x-y)F(x)dx\mu_t(dy) = \langle \mu_t, F\rangle$. Let $\{\phi_j\}$ be a CONS of H_0 such that $\phi_j \in C_b(\mathbb{R})$. Then by Fatou's lemma

$$\tilde{\mathbb{E}}\left(\sum_j \langle \mu_t, \phi_j\rangle^2\right) = \tilde{\mathbb{E}}\left(\sum_j \lim_{\delta\to 0}\langle Z_t^\delta, \phi_j\rangle_0^2\right) \leq \liminf_{\delta \to 0} \tilde{\mathbb{E}}\|Z_t^\delta\|_0^2 \leq \|Z_0^\delta\|_0^2 e^{K_1 t}.$$

Let $\tilde{\mu}_t = \sum_j \langle \mu_t, \phi_j\rangle \phi_j$. Then $\tilde{\mu}_t \in H_0$ and $\langle \tilde{\mu}_t, F\rangle_0 = \sum_j \langle \mu_t, \phi_j\rangle\langle F, \phi_j\rangle_0 = \langle \mu_t, F\rangle$. Hence, $\mu_t \in H_0$ and $\tilde{\mathbb{E}}\|\mu_t\|_0^2 < \infty$. ■

Theorem 3.3. *Suppose that $\mu_0 \in H_0$. Then the solution of Zakai equation (8) is unique.*

Proof. Let μ_t^1 and μ_t^2 be two measure-valued solution with the same initial value μ_0. By Corollary 3.1, μ_t^1 and $\mu_t^2 \in H_0$ a.s. Let $D_t = \mu_t^1 - \mu_t^2$. Then $D_t \in H_0$ and $\tilde{\mathbb{E}}\|T_\delta D_t\|_0^2 \leq K_1 \int_0^t \tilde{\mathbb{E}}\|T_\delta(|D_s|)\|_0^2 ds$. Note that by Lemma 3.1, $\int_0^t \tilde{\mathbb{E}}\|T_\delta(|D_s|)\|_0^2 ds \leq \int_0^t \tilde{\mathbb{E}}\|D_s\|_0^2 ds < \infty$. Then letting $\delta \to 0$, by dominated convergence we have $\tilde{\mathbb{E}}\|D_t\|_0^2 \leq K_1 \int_0^t \tilde{\mathbb{E}}\|D_s\|_0^2 ds$, and Gronwall's inequality yields $D_t = 0$. ■

Consequently, the uniqueness of Kushner–Stratonovich equation then follows from Theorem 3.3.

Theorem 3.4. *The solution of Kushner–Stratonovich equation (10) is unique.*

Proof. Let π_t^1 and π_t^2 be two solutions to the Kushner–Stratonovich equation (10). Note that for $i = 1, 2$ and $F \in C_b(\mathbb{R})$ we have $\pi_t^i(F)\mu_t^i(1) = \mu_t^i(F)$, where μ_t^i are the corresponding solutions to Zakai equation. From Theorem 3.3, we have $\mu_t^1 = \mu_t^2$, $a.s.$ for all $t \geq 0$. Hence for all $t \geq 0$, $\pi_t^1(F) = \frac{\mu_t^1(F)}{\mu_t^1(1)} = \frac{\mu_t^2(F)}{\mu_t^2(1)} = \pi_t^2(F)$, $a.s.$ ■

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

4. Large deviation principle

4.1. LDP result for the optimal filter

We study the limiting behavior of the optimal filter with a small signal-to-noise ratio, i.e., consider the signal given in Eq. (6a) and the observation process below, for $\epsilon \downarrow 0$,

$$Y_t^\epsilon = \sqrt{\epsilon}\int_0^t b_2(s, X_s)ds + \int_0^t \sigma_2(s)dW_s \\ + \int_0^t\int_{\mathbb{U}_2} f_2(s,u)\tilde{N}_{\lambda^\epsilon}(ds,du) + \int_0^t\int_{\mathbb{U}\setminus\mathbb{U}_2} g_2(s,u)N_{\lambda^\epsilon}(ds,du), \tag{26}$$

where $N_{\lambda^\epsilon}(dt,du)$ is a Poisson random measure with intensity $\lambda^\epsilon(t,x,u)\nu_2(du)dt$ and $\lambda^\epsilon(t,x,u) = \epsilon\lambda(t,x,u)+1-\epsilon$. For any test function $F \in C_b^2(\mathbb{R}^d)$, set $\pi_t^\epsilon(F) = \mathbb{E}(F(X_t)|\mathcal{F}_t^{Y^\epsilon})$ and define similarly to Eq. (7) an equivalent probability measure which makes the signal and observation processes independent, i.e.,

$$\Lambda_t^\epsilon = \exp\left\{\sqrt{\epsilon}\int_0^t (\sigma_2^{-1}(s)b_2(s,X_s))^* dW_s \right.\\ + \frac{\epsilon}{2}\int_0^t |\sigma_2^{-1}(s)b_2(s,X_s)|^2 ds + \int_0^t\int_{\mathbb{U}}(1-\lambda(s,X_s,u))\nu_2(du)ds \\ \left. + \int_0^t\int_{\mathbb{U}} \log(\epsilon\lambda(s,X_{s-},u)+1-\epsilon)N_{\lambda^\epsilon}^{\epsilon^{-1}}(ds,du)\right\},$$

where $N_{\lambda^\epsilon}^{\epsilon^{-1}}(dt,du)$ is a Poisson random measure with intensity $\epsilon^{-1}\ \lambda^\epsilon(t,X_t,u)\ \nu_2(du)dt$. Consider $\tilde{\mathbb{P}}^\epsilon$ such that $d\tilde{\mathbb{P}}^\epsilon/d\mathbb{P} = (\Lambda_T^\epsilon)^{-1}$ and $\mu_t^\epsilon(F) = \tilde{\mathbb{E}}^\epsilon(\Lambda_t^\epsilon F(X_t)|\mathcal{F}_t^{Y^\epsilon})$, where $\tilde{\mathbb{E}}^\epsilon$ denotes the expectation under the measure $\tilde{\mathbb{P}}^\epsilon$. First, we establish the existence of the small-noise optimal filter, given in Proposition 4.1. Its proof is delegated to Appendix C.

Proposition 4.1. *Let the signal be defined as in Eq. (6a) and the observation process, Y_t^ϵ, be as in Eq. (26). Then we have the following small-noise Zakai equation, for any $F \in \mathcal{D}(\mathcal{L})$*

$$\mu_t^\epsilon(F) = \mu_0(F) + \int_0^t \mu_s^\epsilon(\mathcal{L}F)ds + \sqrt{\epsilon}\sum_{i=1}^m \int_0^t \mu_s^\epsilon(F(\sigma_2^{-1}(s)b_2(s,\cdot))^i)d\tilde{W}_s^{\epsilon,i} \\ + \int_0^t\int_{\mathbb{U}} \mu_{s-}^\epsilon(F(\lambda(s,\cdot,\mu)-1))(\epsilon N_{\lambda^\epsilon}^{\epsilon^{-1}}(ds,du) - \nu_2(du)ds). \tag{27}$$

The corresponding small-noise Kushner–Stratonovich equation is given by

$$\pi_t^\epsilon(F) = \pi_0(F) + \int_0^t \pi_s^\epsilon(\mathcal{L}F)ds \\ + \sqrt{\epsilon}\sum_{i=1}^m \int_0^t (\pi_s^\epsilon(F(\sigma_2^{-1}(s)b_2(s,\cdot))^i) - \pi_s^\epsilon(F)\pi_s^\epsilon(\sigma_2^{-1}(s)b_2(s,\cdot))^i)d\hat{W}_s^{\epsilon,i} \\ + \int_0^t\int_{\mathbb{U}} \frac{\pi_{s-}^\epsilon(F\lambda(s,\cdot,u)) - \pi_{s-}^\epsilon(F)\pi_{s-}^\epsilon(\lambda(s,\cdot,u))}{\pi_{s-}^\epsilon(\epsilon\lambda(s,\cdot,u)+1-\epsilon)}\epsilon\hat{N}^{\epsilon^{-1}}(ds,du), \tag{28}$$

where $\hat{W}_t^\epsilon = \tilde{W}_t^\epsilon - \sqrt{\epsilon}\int_0^t \pi_s^\epsilon(\sigma_2^{-1}(s)b_2(s,\cdot))ds$ *is the innovation process and*

$$\epsilon\hat{N}^{\epsilon^{-1}}(dt,du) = \epsilon N_{\lambda^\epsilon}^{\epsilon^{-1}}(dt,du) - \pi_{t-}^\epsilon(\epsilon\lambda(s,\cdot,u)+1-\epsilon)\nu_2(du)dt.$$

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

What it follows verifies Condition 2.1 such that we show the LDP for the signal described in Eq. (6a) and observation process as in Eq. (26). To proceed with the demonstration of the LDP, the assumption on the boundedness of the infinitesimal generator $\mathcal{L}$ is necessary. However, this condition does not contradict with the well-posedness of the optimal filtering framework, as seen in [29] and the proof of Proposition 4.1.

Assumption 4.1. The test function F has continuous and bounded derivatives up to order 2.

The images of $\mathcal{G}^\epsilon, \mathcal{G}^0$ considered in Condition 2.1 are solutions of versions of Zakai equation with or without noise respectively. Recall that $\mathcal{M}_F(\mathbb{R}^d)$ denotes the collection of finite Borel measure on $\mathbb{R}^d$, and μ_t^ϵ is an $\mathcal{M}_F(\mathbb{R}^d)$-valued process. For each $\epsilon > 0$, let $\mathcal{G}^\epsilon : C([0,T],\mathbb{R}^m) \times \mathbb{M}(\mathbb{U}_T) \to D([0,T],\mathcal{M}_F(\mathbb{R}^d))$ be a measurable map, such that

$$\mu^\epsilon := \mathcal{G}^\epsilon(\sqrt{\epsilon}\tilde{W}^\epsilon, \epsilon N_{\lambda^\epsilon}^{\epsilon^{-1}}). \tag{29}$$

Adopting the arguments of Theorem 3.3, the following holds.

Theorem 4.1. *Under Assumptions 3.1–3.4, the unnormalized filtered μ^ϵ defined in Eq. (29) is the unique solution of the Zakai equation (27).*

Let $\xi = (\psi, \phi) \in \mathcal{U}^M$. The controlled version of Eq. (27) for all $F \in \mathcal{D}(\mathcal{L})$, is given by

$$\begin{aligned}
\mu_t^{\epsilon,\xi}(F) &= \mu_0(F) + \int_0^t \mu_s^{\epsilon,\xi}(\mathcal{L}F)ds + \sqrt{\epsilon}\sum_{i=1}^m \int_0^t \mu_s^{\epsilon,\xi}(F(\sigma_2^{-1}(s)b_2(s,\cdot))^i)d\tilde{W}_s^{\epsilon,i} \\
&\quad + \int_0^t \mu_s^{\epsilon,\xi}(F(\sigma_2^{-1}(s)b_2(s,\cdot)))\psi(s)ds \\
&\quad + \int_0^t \int_{\mathbb{U}} \mu_{s-}^{\epsilon,\xi}(F(\lambda(s,\cdot,\mu)-1))(\epsilon N_{\lambda^\epsilon}^{\epsilon^{-1}\phi}(ds,du) - \nu_2(du)ds).
\end{aligned} \tag{30}$$

Let $\mu_t^{0,\xi}$ be the solution of the noise-free controlled version of Eq. (30), i.e.

$$\begin{aligned}
\mu_t^{0,\xi}(F) &= \mu_0(F) + \int_0^t \mu_s^{0,\xi}(\mathcal{L}F)ds + \int_0^t \mu_s^{0,\xi}(F(\sigma_2^{-1}(s)b_2(s,\cdot)))\psi(s)ds \\
&\quad + \int_0^t \int_{\mathbb{U}} \mu_s^{0,\xi}(F(\lambda(s,\cdot,\mu)-1))(\phi(s,u)-1)\nu_2(du)ds.
\end{aligned} \tag{31}$$

For $g : \mathcal{B}(\mathbb{U}_T) \to [0,\infty)$, define $\nu_2^g(A) = \int_A g(s,u)\nu_2(du)ds$ for any $A \in \mathcal{B}(\mathbb{U}_T)$ where ν_2 is the intensity measure of N_{λ^ϵ} in the observation process. Let $\mathcal{G}^0 : \mathcal{M}_F(\mathbb{R}^d) \times C([0,T],\mathbb{R}^m) \times \mathbb{M}(\mathbb{U}_T) \to C([0,T],\mathcal{M}_F(\mathbb{R}^d))$ be a measurable map such that $\mathcal{G}^0(\mu_0, w, m) = \mu_t^{0,\xi}$ if $(w,m) = (\int_0^\cdot \psi(s)ds, \nu_2^\phi) \in C([0,T],\mathbb{R}^m) \times \mathbb{M}(\mathbb{U}_T)$ for $\xi = (\psi,\phi)$, otherwise $\mathcal{G}^0 = 0$.

For $\mu \in C([0,T],\mathcal{M}_F(\mathbb{R}^d))$ define

$$I_1 \equiv I_1(\mu) := \inf_{\{\xi=(\psi,\phi)\in\mathbb{S}_\mu : \mu = \mathcal{G}^0(\int_0^\cdot \psi(s)ds, \nu_2^\phi)\}} \bar{L}_T(\xi), \tag{32}$$

where $\bar{L}_T$ is defined in Eq. (4) by replacing ν with ν_2. The following theorem is the main one that establishes the uniform large deviations for the unnormalized filter. Its proof is delegated to the next section.

Theorem 4.2. *Let μ^ϵ be as in Eq. (29). Then I_1, defined in (32), is a rate function and $\{\mu^\epsilon\}$ satisfies the large deviation principle on $D([0,T],\mathcal{M}_F(\mathbb{R}^d))$ with the rate function I_1.*

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

Recall that $\mathcal{P}(\mathbb{R}^d)$ denotes the collection of the optimal filter π_t. We define $\tilde{\mathcal{G}}^0 : L^2([0,T],\mathbb{R}^m) \times [0,\infty) \to C([0,T],\mathcal{P}(\mathbb{R}^d))$ a measurable function and suppose that $\pi^{0,\xi} = \tilde{\mathcal{G}}^0(\psi,\phi)$, where $\pi^{0,\xi}$ is the solution of the following controlled equation.

$$\pi_t^{0,\xi}(F) = \pi_0(F) + \int_0^t \pi_s^{0,\xi}\left\{\mathcal{L}F + \left(\sigma_2^{-1}(s)b_2(s,\cdot) - \pi_s^{0,\xi}(\sigma_2^{-1}(s)b_2(s,\cdot))\right)\psi(s)F\right\}ds$$
$$+ \int_0^t \int_{\mathbb{U}} \left[\pi_s^{0,\xi}(F\lambda(s,\cdot,u)) - \pi_s^{0,\xi}(F)\pi_s^{0,\xi}(\lambda(s,\cdot,u))\right](\phi(s,u)-1)\nu_2(du)dt. \tag{33}$$

For $\pi \in C([0,T],\mathcal{P}(\mathbb{R}^d))$ define

$$I_2 \equiv I_2(\pi) := \inf_{\{\xi=(\psi,\phi)\in\mathbb{S}_\pi : \pi=\mathcal{G}^0(\int_0^\cdot \psi(s)ds, \nu_2^\phi)\}} \bar{L}_T(\xi). \tag{34}$$

Lemma 4.1. *For $\xi = (\psi,\phi) \in \mathcal{U}^M$, let*

$$M_t^\xi = \exp\left(\int_0^t \pi_s(\sigma_2^{-1}(s)b_2(s,\cdot)\psi(s))ds + \int_0^t \int_{\mathbb{U}} \pi_s(\lambda(s,\cdot,\mu)-1)(\phi(s,u)-1)\nu_2(du)ds\right).$$

Then $\mu_t^{0,\xi}(F) = \pi_t^{0,\xi}(F)M_t^\xi$ satisfies the noise-free Zakai equation (31).

Proof. We first note that

$$dM_t^\xi = M_t^\xi\left(\pi_t(\sigma_2^{-1}(t)b_2(t,\cdot)\psi(t))dt + \int_{\mathbb{U}} \pi_t(\lambda(t,\cdot,\mu)-1)(\phi(t,u)-1)\nu_2(du)dt\right).$$

Differentiating the product, $\pi_t^{0,\xi}(F)M_t^\xi$, we have

$$\begin{aligned}
&\pi_t^{0,\xi}(F)M_t^\xi \pi_t(\sigma_2^{-1}(t)b_2(t,\cdot)\psi(t))dt \\
&+ \pi_t^{0,\xi}(F)M_t^\xi \int_{\mathbb{U}} \pi_t(\lambda(t,\cdot,\mu)-1)(\phi(t,u)-1)\nu_2(du)dt \\
&+ \pi_t^{0,\xi}(\mathcal{L}F)M_t^\xi dt + \pi_t(F\sigma_2^{-1}(t)b_2(t,\cdot)\psi(t))M_t^\xi dt \\
&- \pi_t^{0,\xi}(F)M_t^\xi \pi_t(\sigma_2^{-1}(t)b_2(t,\cdot)\psi(t))dt \\
&+ \int_{\mathbb{U}} M_t^\xi\left[\pi_s^{0,\xi}(F\lambda(s,\cdot,u)) - \pi_s^{0,\xi}(F)\pi_s^{0,\xi}(\lambda(s,\cdot,u))\right](\phi(s,u)-1)\nu_2(du)dt.
\end{aligned} \tag{35}$$

Regroup the right-hand side of Eq. (35) and then the integral form coincides with Eq. (31). ■

The following theorem establishes a uniform large deviation principle for the optimal filtering defined in Eq. (28).

Theorem 4.3. *Suppose π_t^ϵ is the optimal filter described by the Kushner–Stratonovich equation (28). Then $\{\pi^\epsilon\}$ satisfies the large deviation principle on $D([0,T],\mathcal{P}(\mathbb{R}^d))$ with the rate function I_2, defined in Eq. (34).*

Proof. Define a map $G : D([0,T],\mathcal{M}_F(\mathbb{R}^d)\setminus\{0\}) \to D([0,T],\mathcal{P}(\mathbb{R}^d))$ such that $(G\mu)_t = \frac{\mu_t(F)}{\mu_t(1)}$. Then, by the contraction principle, $\{\pi^\epsilon = G(\mu^\epsilon)\}$ satisfies the large deviation principle

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

with rate function $I_2'(\pi) = \inf\{I_1(\mu) : G(\mu) = \pi\}$. Suppose $I_2'(\pi) < \infty$, then for all $\delta > 0$ there exists μ such that $G(\mu) = \pi$ and $I_1(\mu) < I_2'(\pi) + \delta$. Choose a control $\xi \in \mathbb{S}_\pi$ such that $\gamma(\xi) = \mu$, where γ is the solution of Eq. (31) and $\bar{L}_T(\xi) < I_1(\mu) + \delta$. Taking $\tilde{\gamma} \doteq G \circ \gamma$ we have $\tilde{\gamma}(\xi) = \xi$, where $\tilde{\gamma}$ is the solution of Eq. (33), and $\bar{L}_T(\xi) < I_2'(\pi) + 2\delta$. Thus, by definition of I_2 we have

$$I_2(\pi) \leq I_2'(\pi). \tag{36}$$

Now if $I_2(\pi) < \infty$. Then for all $\delta > 0$, there exists $\xi \in \mathbb{S}_\pi$ such that $\tilde{\gamma}(\xi) = \pi$ and $\bar{L}_T(\xi) < I_2(\pi) + \delta$. By Lemma 4.1, we have $\mu_t = M_t^\xi \pi_t$. Then $\mu = \gamma(\xi)$ is the solution of Eq. (31) and $G(\mu) = \pi$. Hence

$$I_2'(\pi) \leq I_1(\mu) \leq \bar{L}_T(\xi) < I_2(\pi) + \delta. \tag{37}$$

Eqs. (36) and (37) give $I_2 = I_2'$ and then the result follows. ∎

4.2. Proof of Theorem 4.2

In the subsection we verify Condition 2.1 in order to show the large deviation estimates for the unnormalized filter. The following theorem establishes the existence and uniqueness of the controlled version of Zakai equation given in Eq. (30).

Theorem 4.4. *Suppose $\mathcal{G}^\epsilon$ is given by $\mu^\epsilon = \mathcal{G}^\epsilon(\sqrt{\epsilon}\tilde{W}^\epsilon, \epsilon N_{\lambda^\epsilon}^{\epsilon^{-1}})$, and $\xi = (\psi, \phi) \in \mathcal{U}^M$ for some $M > 0$. For $\epsilon > 0$, define $\mu^{\epsilon,\xi} = \mathcal{G}^\epsilon(\sqrt{\epsilon}\tilde{W}^\epsilon + \int_0^\cdot \psi(s)ds, \epsilon N_{\lambda^\epsilon}^{\epsilon^{-1}\phi})$. Then $\mu^{\epsilon,\xi}$ is the unique solution of Eq. (30).*

Proof. Take a control $\xi = (\psi, \phi) \in \mathcal{U}^M$, and consider

$$\begin{aligned} M_t^{\epsilon,\xi} := \exp\Bigg\{ &-\frac{1}{\sqrt{\epsilon}} \int_0^t \psi(s) d\tilde{W}_s^\epsilon + \int_0^t \int_{\mathbb{U}\times[0,\epsilon^{-1}]} \log\phi(s,u) \bar{N}_{\lambda^\epsilon}(ds,du,dr) \\ &- \frac{1}{2\epsilon}\int_0^t |\psi(s)|^2 ds + \int_0^t \int_{\mathbb{U}\times[0,\epsilon^{-1}]} (1-\phi(s,u))\bar{\nu}_2(ds,du,dr) \Bigg\}, \end{aligned}$$

where $\bar{N}_{\lambda^\epsilon}$ and the corresponding intensity $\bar{\nu}_2$ satisfy the definitions of $\bar{N}$ and $\bar{\nu}_T$ in Section 2.1 respectively. By Itô's formula,

$$\begin{aligned} dM_t^{\epsilon,\xi} &= M_t^{\epsilon,\xi}\Bigg(-\frac{1}{\sqrt{\epsilon}}\psi(t) d\tilde{W}_t^\epsilon - \frac{1}{2\epsilon}|\psi(t)|^2 dt \\ &\quad + \int_{\mathbb{U}\times[0,\epsilon^{-1}]} (1-\phi(s,u))\bar{\nu}_2(ds,du,dr) \Bigg) \\ &\quad + \frac{1}{2\epsilon} M_t^{\epsilon,\xi}|\psi(t)|^2 dt + \int_{\mathbb{U}\times[0,\epsilon^{-1}]} M_{t-}^{\epsilon,\xi}(\phi(t,u)-1)\bar{N}_{\lambda^\epsilon}(ds,du,dr). \end{aligned}$$

Thus,

$$\begin{aligned} M_t^{\epsilon,\xi} &= 1 - \frac{1}{\sqrt{\epsilon}}\int_0^t M_s^{\epsilon,\xi}\psi(s) d\tilde{W}_s^\epsilon \\ &\quad + \int_0^t \int_{\mathbb{U}} M_{s-}^{\epsilon,\xi}(\phi(s,u)-1)\left(\bar{N}_{\lambda^\epsilon}(ds,du,dr) - \bar{\nu}_2(ds,du,dr)\right). \end{aligned}$$

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

Define a new probability measure $\mathbb{P}^{\epsilon,\xi}$ by $\frac{d\mathbb{P}^{\epsilon,\xi}}{d\tilde{\mathbb{P}}^{\epsilon}} = M_T^{\epsilon,\xi}$, then $\mathbb{P}^{\epsilon,\xi}$ is an equivalent probability measure with respect to $\tilde{\mathbb{P}}^{\epsilon}$. By Girsanov's theorem, $W_t^{\epsilon,\xi} := \tilde{W}_t^{\epsilon} + \frac{1}{\sqrt{\epsilon}} \int_0^t \psi(s)ds$ is a $\mathbb{P}^{\epsilon,\xi}$-Brownian motion and $N^{\epsilon,\xi}(dt, du) := N_{\lambda^\epsilon}(dt, du) - \phi(t, u)\nu_2(du)dt$ is a $\mathbb{P}^{\epsilon,\xi}$-Poisson martingale measure. Hence, the desired result follows from replacing $(\tilde{W}, N_{\lambda^\epsilon})$ in Eq. (27) with $(W^{\epsilon,\xi}, N^{\epsilon,\xi})$ and Theorem 4.1. ■

We now verify the existence and uniqueness of the solution of the zero-noise controlled Zakai equation (31).

Theorem 4.5. *For a pair of well-defined control $\xi = (\psi, \phi) \in \mathcal{U}^M$, $M > 0$, there is a unique solution $\mu_t^{0,\xi} \in C([0, T], \mathcal{M}_F(\mathbb{R}^d))$ of Eq. (31).*

Proof. We start with showing the existence of the solution. Let

$$\begin{aligned}\Lambda_t^{0,\xi} = \exp\Bigg(&\int_0^t \sigma_2^{-1}(s) b_2(s, X_s)\psi(s)ds \\ &+ \int_0^t \int_{\mathbb{U}} (\lambda(s, X_s, \mu) - 1)(\phi(s, u) - 1)\nu_2(du)ds \Bigg).\end{aligned}$$

Next, we define $\mu_t^{0,\xi} = \mathbb{E}(\Lambda_t^{0,\xi} F(X_t))$ and we show that it is a solution of Eq. (31). First of all,

$$d\Lambda_t^{0,\xi} = \Lambda_t^{0,\xi}\left(\sigma_2^{-1}(t) b_2(t, X_t)\psi(t)dt + \int_{\mathbb{U}} (\lambda(t, X_t, \mu) - 1)(\phi(t, u) - 1)\nu_2(du)dt\right).$$

Note that for any $F \in \mathcal{D}(\mathcal{L})$ we have $M_t^F \equiv F(X_t) - \int_0^t \mathcal{L}F(X_s)ds$ is a martingale. Then by Itô's formula, we have

$$\begin{aligned}d\Lambda_t^{0,\xi} F(X_t) = \Lambda_t^{0,\xi}\left(dM_t^F + \mathcal{L}F(X_t)dt\right) &+ F(X_t)\Lambda_t^{0,\xi}\sigma_2^{-1}(t) b_2(t, X_t)\psi(t)dt \\ &+ \int_{\mathbb{U}} F(X_t)\Lambda_t^{0,\xi}(\lambda(t, X_t, \mu) - 1)(\phi(t, u) - 1)\nu_2(du)dt.\end{aligned}$$

Hence,

$$\begin{aligned}\Lambda_t^{0,\xi} F(X_t) = F(X_0) + \int_0^t \Lambda_s^{0,\xi}(\mathcal{L}F(X_s) + \sigma_2^{-1}(s) b_2(s, X_s)\psi(s))ds. \qquad (38)\\ + \int_0^t \int_{\mathbb{U}} \Lambda_s^{0,\xi} F(X_s)(\lambda(s, X_s, \mu) - 1)(\phi(s, u) - 1)\nu_2(du)ds + \int_0^t \Lambda_s^{0,\xi} dM_s^F.\end{aligned}$$

Taking expectation on both sides of Eq. (38) yields Eq. (31). Thus $\mu_t^{0,\xi}(F) = \mathbb{E}(\Lambda_t^{0,\xi} F(X_t))$ is a solution of Eq. (31).

Next, we verify the uniqueness by a similar strategy to Theorem 3.3. Recall that for any $\mu \in \mathcal{M}_F(\mathbb{R}^d)$, $\langle \mu, F\rangle = \int_{\mathbb{R}} F(x)\mu(dx)$. Then Eq. (31) is written as

$$\begin{aligned}\langle \mu_t^{0,\epsilon}, F\rangle = \langle \mu_0, F\rangle + \int_0^t \langle \mu_s^{0,\epsilon}, \mathcal{L}F\rangle ds &+ \int_0^t \langle \mu_s^{0,\epsilon}, F(\sigma_2^{-1}(s) b_2(s, \cdot)\psi(s))\rangle ds \\ &+ \int_0^t \int_{\mathbb{U}} \langle \mu_s^{0,\epsilon}, F(\lambda(s, \cdot, u) - 1)(\phi(s, u) - 1)\rangle \nu_2(du)ds.\end{aligned}$$

Let $Z_t^\delta = T_\delta \mu_t^{0,\epsilon}$ be defined as in Eq. (11). Replacing F by $T_\delta F$ and noting that $\langle T_\delta \nu, F\rangle = \langle \nu, T_\delta F\rangle$, we have that $\langle Z_t^\delta, F\rangle_0$ equals to

$$\begin{aligned}&\langle \mu_0, T_\delta F\rangle + \int_0^t \langle \mu_s^{0,\epsilon}, \mathcal{L} T_\delta F\rangle ds + \int_0^t \langle \mu_s^{0,\epsilon}, T_\delta F(\sigma_2^{-1}(s) b_2(s,\cdot)\psi(s))\rangle ds \\ &+ \int_0^t \int_{\mathbb{U}} \langle \mu_s^{0,\epsilon}, T_\delta F(\lambda(s,\cdot,u)-1)(\phi(s,u)-1)\rangle \nu_2(du) ds. \end{aligned} \tag{39}$$

Employing Lemma 3.2, Eq. (39) is given by

$$\begin{aligned}\langle Z_t^\delta, F\rangle_0 = {}& \langle Z_0^\delta, F\rangle + \int_0^t \langle \partial_i T_\delta(b_1^i \mu_s^{0,\epsilon}), F\rangle_0 ds + \frac{1}{2}\int_0^t \langle \partial_{ij}^2 T_\delta(\sigma_1^{ik}\sigma_1^{kj}\mu_s^{0,\epsilon}), F\rangle_0 ds \\ &+ \int_0^t \langle T_\delta(\sigma_2^{-1}(s) b_2(s,\cdot)\psi(s)\mu_s^{0,\epsilon}), F\rangle_0 ds \\ &+ \int_0^t \int_{\mathbb{U}\setminus\mathbb{U}_1} \left[\langle T_\delta \mu_s^{0,\epsilon}, F(\cdot + g_1(\cdot,u))\rangle_0 - \langle T_\delta \mu_s^{0,\epsilon}, F\rangle_0\right] \nu_1(du) ds \\ &+ \int_0^t \int_{\mathbb{U}_1} \big[\langle T_\delta \mu_s^{0,\epsilon}, F(\cdot + f_1(\cdot,u))\rangle_0 - \langle T_\delta \mu_s^{0,\epsilon}, F\rangle_0 \\ &- \langle \partial_i T_\delta(f_1^i(\cdot,u)\mu_s^{0,\epsilon}), F\rangle_0\big] \nu_1(du) ds \\ &+ \int_0^t \int_{\mathbb{U}} \langle T_\delta(\lambda(s,\cdot,u)-1)((\phi(s,u)-1)\mu_s^{0,\epsilon}), F\rangle_0 \nu_2(du) ds. \end{aligned}$$

Furthermore, summing $\langle Z_t^\delta, \eta\rangle_0^2 = \langle Z_0^\delta, \eta\rangle_0^2 + \int_0^t 2\langle Z_s^\delta, \eta\rangle_0 d\langle Z_s^\delta, \eta\rangle_0$ over η in a CONS of L^2-space, there exists a constant $K_1 > 0$ such that

$$\|Z_t^\delta\|_0^2 \le \|Z_0^\delta\|_0^2 + K_1 \int_0^t \|T_\delta \mu_s^{0,\epsilon}\|_0^2 ds. \tag{40}$$

Now let $\mu_{1,t}^{0,\epsilon}$ and $\mu_{2,t}^{0,\epsilon}$ be two solutions with the same initial value μ_0. Then by Corollary 3.1 and Eq. (40), we get $\|T_\delta(\mu_{1,t}^{0,\epsilon} - \mu_{2,t}^{0,\epsilon})\|_0^2 \le K_1 \int_0^t \|T_\delta(\mu_{1,s}^{0,\epsilon} - \mu_{2,s}^{0,\epsilon})\|_0^2 ds$. Taking $\delta \to 0$ gives $\|\mu_{1,t}^{0,\epsilon} - \mu_{2,t}^{0,\epsilon}\|_0^2 \le K_1 \int_0^t \|\mu_{1,s}^{0,\epsilon} - \mu_{2,s}^{0,\epsilon}\|_0^2 ds$, and by Gronwall's inequality, we have $\mu_{1,t}^{0,\epsilon} - \mu_{2,t}^{0,\epsilon} = 0$. ■

The following proposition demonstrates the first statement in Condition 2.1.

Proposition 4.2. *For $M < \infty$, suppose that $\xi_n = (\psi_n, \phi_n)$, $\xi = (\psi, \phi) \in \mathcal{U}^M$ such that $\xi_n \to \xi$, as $n \to \infty$. Then*

$$\mathcal{G}^0\Big(\int_0^\cdot \psi_n(s) ds, \nu_2^{\phi_n}\Big) \to \mathcal{G}^0\Big(\int_0^\cdot \psi(s) ds, \nu_2^{\phi}\Big), \text{ as } n \to \infty.$$

Proof. Consider

$$\begin{aligned}\Lambda_t^{0,\xi} = \exp\Bigg(&\int_0^t \sigma_2^{-1}(s) b_2(s, X_s)\psi(s) ds \\ &+ \int_0^t \int_{\mathbb{U}} (\lambda(s, X_s, \mu) - 1)(\phi(s,u) - 1)\nu_2(du) ds\Bigg). \end{aligned}$$

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

The proof of the existence in Theorem 4.5 shows that $\rho_t F = \int_{\mathbb{R}^d} \mu_0(dx)\mathbb{E}_x \Lambda_t^{0,\xi} F(X_t)$ is a solution of zero-noise version of Zakai equation (31). Furthermore, the solution is unique according to Theorem 4.5 and hence we get $\mu_t^{0,\xi} = \rho_t$. Notice that, by the boundedness of σ_2^{-1}, b_2 and $(\psi_n, \phi_n) \in \mathcal{U}^M$, there exists a constant $K > 0$ such that

$$\begin{aligned}&\int_0^T |\sigma_2^{-1}(s)b_2(s, X_s)\psi_n(s)|ds\\&\leq \left(\int_0^T |\sigma_2^{-1}(s)b_2(s, X_s)|^{\,2}ds\right)^{1/2}\left(\int_0^T |\psi_n(s)|^2 ds\right)^{1/2} \leq (KT)^{1/2}M^{1/2}.\end{aligned}$$

In addition, by the inequality, for any $\phi \in [0, \infty)$,

$$|\phi - 1| \leq 2 + \phi \log \phi - \phi + 1, \tag{41}$$

we have

$$\begin{aligned}&\int_0^T \int_{\mathbb{U}} |(\lambda(s, X_s, \mu) - 1)(\phi_n(s, u) - 1)|\nu_2(du)ds\\&\leq 2\int_0^T \int_{\mathbb{U}} (2 + \phi_n(s, u)\log \phi_n(s, u) - \phi_n(s, u) + 1)\,\nu_2(du)ds\\&\leq 4T\nu_2(\mathbb{U}) + 2M.\end{aligned}$$

The dominated convergence theorem yields the convergence below, as $n \to \infty$,

$$\begin{aligned}\mu_t^{0,\xi_n}(F) = &\int_{\mathbb{R}} \mu_0(dx)\mathbb{E}_x F(X_t)\exp\left(\int_0^t \sigma_2^{-1}(s)b_2(s, X_s)\psi_n(s)ds\right.\\&\qquad\qquad \left.+ \int_0^t \int_{\mathbb{U}} (\lambda(s, X_s, \mu) - 1)(\phi_n(s, u) - 1)\nu_2(du)ds\right)\\&\to \int_{\mathbb{R}} \mu_0(dx)\mathbb{E}_x F(X_t)\exp\left(\int_0^t \sigma_2^{-1}(s)b_2(s, X_s)\psi(s)ds\right.\\&\qquad\qquad \left.+ \int_0^t \int_{\mathbb{U}} (\lambda(s, X_s, \mu) - 1)(\phi(s, u) - 1)\nu_2(du)ds\right)\\&= \mu_t^{0,\xi}(F).\end{aligned}$$

Then the proof is completed. ■

The next proposition verifies the second part of Condition 2.1.

Proposition 4.3. *For $M < \infty$, let $\xi^\epsilon = (\psi^\epsilon, \phi^\epsilon)$, $\xi = (\psi, \phi) \in \mathcal{U}^M$ be such that $\xi^\epsilon \xrightarrow{d} \xi$ as $\epsilon \to 0$. Then*

$$\mathcal{G}^\epsilon(\sqrt{\epsilon}\tilde{W}^\epsilon + \int_0^\cdot \psi^\epsilon(s)ds, \epsilon N_{\lambda^\epsilon}^{\epsilon^{-1}\phi^\epsilon}) \xrightarrow{d} \mathcal{G}^0(\int_0^\cdot \psi(s)ds, \nu_2^\phi),\ \textit{as } \epsilon \to 0.$$

Proof. First, we prove that the family

$$\{\mu^{\epsilon,\xi^\epsilon} = \mathcal{G}^\epsilon(\sqrt{\epsilon}\tilde{W}^\epsilon + \int_0^\cdot \psi^\epsilon(s)ds, \epsilon N_{\lambda^\epsilon}^{\epsilon^{-1}\phi^\epsilon}), \epsilon \in (0, \epsilon_0)\}$$

is tight in $D([0,T],\mathcal{M}_F(\mathbb{R}^d))$ for some $\epsilon_0>0$. Note that

$$\begin{aligned}\mu_t^{\epsilon,\xi^\epsilon}(1) = \mu_0(1) + \sqrt{\epsilon}\sum_{i=1}^m \int_0^t \mu_s^{\epsilon,\xi^\epsilon,\mu_0^\epsilon}((\sigma_2^{-1}(s)b_2(s,\cdot))^i)d\tilde{W}_s^{\epsilon,i}\\ + \int_0^t \mu_s^{\epsilon,\xi^\epsilon}(\sigma_2^{-1}(s)b_2(s,\cdot))\psi^\epsilon(s)ds\\ + \int_0^t\int_{\mathbb{U}} \mu_s^{\epsilon,\xi^\epsilon}(\lambda(s,\cdot,\mu)-1)\,(\phi^\epsilon(s,u)-1)\,\nu_2(du)ds\\ +\epsilon\int_0^t\int_{\mathbb{U}} \mu_{s-}^{\epsilon,\xi^\epsilon}(\lambda(s,\cdot,\mu)-1)(N_{\lambda^\epsilon}^{\epsilon^{-1}\phi^\epsilon}(ds,du)-\epsilon^{-1}\phi^\epsilon(s,u)\nu_2(du)ds).\end{aligned}$$

By Itô's formula, we get $\mu_t^{\epsilon,\xi^\epsilon}(1)^2 = A_t^1+A_t^2+A_t^3+A_t^4+A_t^5+A_t^6+A_t^7$ where

$$\begin{aligned}A_t^1 &= \mu_0^{\epsilon,\xi^\epsilon}(1)^2, \quad A_t^2 = 2\int_0^t \mu_s^{\epsilon,\xi^\epsilon}(1)\mu_s^{\epsilon,\xi^\epsilon}(\sigma_2^{-1}(s)b_2(s,\cdot))\psi^\epsilon(s)ds,\\ A_t^3 &= 2\int_0^t\int_{\mathbb{U}} \mu_s^{\epsilon,\xi^\epsilon}(1)\mu_s^{\epsilon,\xi^\epsilon}(\lambda(s,\cdot,\mu)-1)\,(\phi^\epsilon(s,u)-1)\,\nu_2(du)ds,\\ A_t^4 &= \epsilon\int_0^t \left|\mu_s^{\epsilon,\xi^\epsilon}(\sigma_2^{-1}(s)b_2(s,\cdot))\right|^2 ds,\\ A_t^5 &= \int_0^t\int_{\mathbb{U}} \epsilon^2\mu_s^{\epsilon,\xi^\epsilon}(\lambda(s,\cdot,\mu)-1)^2\phi^\epsilon(s,u)\nu_2(du)ds,\\ A_t^6 &= 2\sqrt{\epsilon}\sum_{i=1}^m\int_0^t \mu_s^{\epsilon,\xi^\epsilon}(1)\mu_s^{\epsilon,\xi^\epsilon}((\sigma_2^{-1}(s)b_2(s,\cdot))^i)d\tilde{W}_s^{\epsilon,i},\\ A_t^7 &= \int_0^t\int_{\mathbb{U}}\Big[2\epsilon\mu_{s-}^{\epsilon,\xi^\epsilon}(1)\mu_{s-}^{\epsilon,\xi^\epsilon}(\lambda(s,\cdot,\mu)-1),\\ &\quad +\epsilon^2\mu_{s-}^{\epsilon,\xi^\epsilon}((\lambda(s,\cdot,\mu)-1))^2\Big]\left(N_{\lambda^\epsilon}^{\epsilon^{-1}\phi^\epsilon}(ds,du)-\epsilon^{-1}\phi^\epsilon(s,u)\nu_2(du)(ds)\right).\end{aligned}$$

Considering Assumption 3.2 and Cauchy–Schwarz inequality, A_t^3 is bounded by

$$\begin{aligned}&2K\int_0^t \mu_s^{\epsilon,\xi^\epsilon}(1)^2\,|\psi^\epsilon(s)|\,ds \\ &\le 2K\left(\int_0^t \mu_s^{\epsilon,\xi^\epsilon}(1)^2\,|\psi^\epsilon(s)|^2\,ds\right)^{1/2}\left(\int_0^t \mu_s^{\epsilon,\xi^\epsilon}(1)^2 ds\right)^{1/2}\\ &\le K\int_0^t \mu_s^{\epsilon,\xi^\epsilon}(1)^2\,|\psi^\epsilon(s)|^2\,ds + K\int_0^t \mu_s^{\epsilon,\xi^\epsilon}(1)^2 ds,\end{aligned} \tag{42}$$

and $A_t^3+A_t^4+A_t^5$ is bounded by

$$\begin{aligned}&4\int_0^t\int_{\mathbb{U}} \mu_s^{\epsilon,\xi^\epsilon}(1)^2\,|\phi^\epsilon(s,u)-1|\,\nu_2(du)ds\\ &+K\epsilon\int_0^t \mu_s^{\epsilon,\xi^\epsilon}(1)^2 ds + 4\epsilon^2\int_0^t\int_{\mathbb{U}} \mu_s^{\epsilon,\xi^\epsilon}(1)^2|\phi^\epsilon(s,u)|\nu_2(du)ds.\end{aligned} \tag{43}$$

Set

$$A_t^{7,1} = \int_0^t \int_{\mathbb{U}} 2\epsilon \mu_{s-}^{\epsilon,\xi^\epsilon}(1)\mu_{s-}^{\epsilon,\xi^\epsilon}(\lambda(s,\cdot,\mu)-1)\left(N_{\lambda^\epsilon}^{\epsilon^{-1}\phi^\epsilon}(ds,du) - \epsilon^{-1}\phi^\epsilon(s,u)\nu_2(du)(ds)\right) \tag{44}$$

and

$$A_t^{7,2} = \int_0^t \int_{\mathbb{U}} \epsilon^2 \mu_{s-}^{\epsilon,\xi^\epsilon}((\lambda(s,\cdot,\mu)-1))^2\left(N_{\lambda^\epsilon}^{\epsilon^{-1}\phi^\epsilon}(ds,du) - \epsilon^{-1}\phi^\epsilon(s,u)\nu_2(du)(ds)\right). \tag{45}$$

Combining Eqs. (42) and (43) implies that

$$\begin{aligned}\mu_t^{\epsilon,\xi^\epsilon}(1)^2 &\le \mu_0^{\epsilon,\xi^\epsilon}(1)^2 + \sup_{0\le t\le T}|A_t^6| + \sup_{0\le t\le T}|A_t^{7,1}| + \sup_{0\le t\le T}|A_t^{7,2}| \\ &+ \int_0^t \mu_s^{\epsilon,\xi^\epsilon}(1)^2\left(K|\psi^\epsilon(s)|^2 + K + \int_{\mathbb{U}}\left(|\phi^\epsilon(s,u)-1| + K\epsilon + 4\epsilon^2\phi^\epsilon(s,u)\right)\nu_2(du)\right)ds.\end{aligned}$$

Furthermore, inequality (41) gives

$$\begin{aligned}&\int_0^T\int_{\mathbb{U}} |\phi^\epsilon(s,u)|\,\nu_2(du)ds \le \int_0^T\int_{\mathbb{U}} (|\phi^\epsilon(s,u)-1|+1)\,\nu_2(du)(ds) \\ &\le \int_0^T\int_{\mathbb{U}} (3+\phi^\epsilon(s,u)\log\phi^\epsilon(s,u) - \phi^\epsilon(s,u)+1)\,\nu_2(du)(ds) \\ &\le 3T\nu_2(\mathbb{U}) + M.\end{aligned} \tag{46}$$

Then by Gronwall's inequality and Eq. (46), we have

$$\mu_t^{\epsilon,\xi^\epsilon}(1)^2 \le C_0\left(\mu_0^{\epsilon,\xi^\epsilon}(1)^2 + \sup_{0\le t\le T}|A_t^6| + \sup_{0\le t\le T}|A_t^{7,1}| + \sup_{0\le t\le T}|A_t^{7,2}|\right), \tag{47}$$

where $C_0 = \exp\left(KM + KT + K\epsilon T + 2T\nu_2(\mathbb{U}) + M + 4\epsilon^2(3T\nu_2(\mathbb{U})+M)\right)$. Next, we have

$$\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\le t\le T}|A_t^6|\right) \le 8K\sqrt{\epsilon}\tilde{\mathbb{E}}^\epsilon\sqrt{\int_0^T \mu_s^{\epsilon,\xi^\epsilon}(1)^4 ds} \le 8K\sqrt{T}\sqrt{\epsilon}\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\le t\le T}\mu_s^{\epsilon,\xi^\epsilon}(1)^2\right). \tag{48}$$

For Eq. (44), we have

$$\begin{aligned}&\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\le t\le T}|A_t^{7,1}|\right) \\ &\le 8\epsilon\tilde{\mathbb{E}}^\epsilon\left(\int_0^T\int_{\mathbb{U}} \mu_{s-}^{\epsilon,\xi^\epsilon}(1)^2\mu_{s-}^{\epsilon,\xi^\epsilon}(\lambda(s,\cdot,\mu)-1)^2 N_{\lambda^\epsilon}^{\epsilon^{-1}\phi^\epsilon}(ds,du)\right)^{\frac{1}{2}} \\ &\le 16\epsilon\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\le t\le T}\mu_t^{\epsilon,\xi^\epsilon}(1)\cdot\left(\int_0^T\int_{\mathbb{U}}\mu_{s-}^{\epsilon,\xi^\epsilon}(1)^2 N_{\lambda^\epsilon}^{\epsilon^{-1}\phi^\epsilon}(ds,du)\right)^{\frac{1}{2}}\right) \\ &\le \frac{1}{C}\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\le t\le T}\mu_t^{\epsilon,\xi^\epsilon}(1)^2\right) \\ &\quad + 64\epsilon^2 C\tilde{\mathbb{E}}^\epsilon\left(\int_0^T\int_{\mathbb{U}}\mu_{s-}^{\epsilon,\xi^\epsilon}(1)^2 N_{\lambda^\epsilon}^{\epsilon^{-1}\phi^\epsilon}(ds,du)\right) \\ &= \frac{1}{C}\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\le t\le T}\mu_t^{\epsilon,\xi^\epsilon}(1)^2\right)\end{aligned} \tag{49}$$

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

$$
\begin{aligned}
&+ 64\epsilon C\left(\int_0^T \int_{\mathbb{U}} \mu_s^{\epsilon,\xi^\epsilon}(1)^2 \phi^\epsilon(s,u)\nu_2(du)ds\right)\\
&\leq \left(\frac{1}{C} + 64\epsilon C(3T\nu_2(\mathbb{U}) + M)\right)\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\leq t\leq T}\mu_t^{\epsilon,\xi^\epsilon}(1)^2\right),
\end{aligned}
$$

where C is any positive number. Eq. (44) is bounded by

$$
\begin{aligned}
&\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\leq t\leq T}|A_t^{7,2}|\right) \qquad (50)\\
&\leq \tilde{\mathbb{E}}^\epsilon\left(\int_0^T\int_{\mathbb{U}} \epsilon^2 \mu_{s-}^{\epsilon,\xi^\epsilon}((\lambda(s,\cdot,\mu)-1))^2 N_{\lambda^\epsilon}^{\epsilon^{-1}\phi^\epsilon}(ds,du)\right)\\
&\quad + \epsilon\tilde{\mathbb{E}}^\epsilon\left(\int_0^T\int_{\mathbb{U}} \mu_s^{\epsilon,\xi^\epsilon}((\lambda(s,\cdot,\mu)-1))^2\phi^\epsilon(s,u)\nu_2(du)(ds)\right)\\
&\leq 8\epsilon\tilde{\mathbb{E}}^\epsilon\left(\int_0^T\int_{\mathbb{U}} \mu_s^{\epsilon,\xi^\epsilon}(1)^2\phi^\epsilon(s,u)\nu_2(du)(ds)\right)\\
&\leq 8\epsilon\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\leq t\leq T}\mu_t^{\epsilon,\xi^\epsilon}(1)^2\right)\left(\int_0^T\int_{\mathbb{U}}\phi^\epsilon(s,u)\nu_2(du)(ds)\right)\\
&\leq \epsilon\,(24T\nu_2(\mathbb{U}) + 8M)\,\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\leq t\leq T}\mu_t^{\epsilon,\xi^\epsilon}(1)^2\right),
\end{aligned}
$$

where inequality (46) was used. By Eqs. (48), (49) and (50), Eq. (47) turns out to be

$$
\begin{aligned}
&\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\leq t\leq T}\mu_t^{\epsilon,\xi^\epsilon}(1)^2\right) \leq C_0\mu_0^{\epsilon,\xi^\epsilon}(1)^2\\
&+ C_0\left(8K\sqrt{T}\sqrt{\epsilon} + \epsilon\,(24T\nu_2(\mathbb{U}) + 8M) + \frac{1}{C} + 64\epsilon C(3T\nu_2(\mathbb{U}) + M)\right).
\end{aligned}
$$

Recall that the constant C can be arbitrarily large, we hence select C and ϵ_0 small enough such that

$$
C_0\left(8K\sqrt{T}\sqrt{\epsilon} + \epsilon\,(24T\nu_2(\mathbb{U}) + 8M) + \frac{1}{C} + 64\epsilon C(3T\nu_2(\mathbb{U}) + M)\right) < \frac{1}{2}.
$$

Therefore, there is a constant K_1 such that

$$
\sup_{0<\epsilon\leq\epsilon_0}\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\leq t\leq T}\mu_t^{\epsilon,\xi^\epsilon}(1)^2\right) \leq K_1. \qquad (51)
$$

We now establish the tightness of $\{\mu^{\epsilon,\xi^\epsilon}\}$. It is well-known, e.g. see [16] that we only need to prove the tightness of $\{\mu_t^{\epsilon,\xi^\epsilon}(F)\}$ in $D([0,T],\mathbb{R})$ for every test function F in $\mathcal{D}(\mathcal{L})$. By Eq. (51) and the boundness of F, we have

$$
\sup_{0<\epsilon\leq\epsilon_0}\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\leq t\leq T}\mu_t^{\epsilon,\xi^\epsilon}(F)^2\right) \leq K_2. \qquad (52)
$$

Denote

$$
\begin{aligned}
A_t^\epsilon &= \int_0^t \mu_s^{\epsilon,\xi^\epsilon}(\mathcal{L}F)ds + \int_0^t \mu_s^{\epsilon,\xi^\epsilon}(F(\sigma_2^{-1}(s)b_2(s,\cdot)))\psi^\epsilon(s)ds\\
&\quad + \int_0^t\int_{\mathbb{U}} \mu_s^{\epsilon,\xi^\epsilon}(F(\lambda(s,\cdot,\mu)-1))\,(\phi^\epsilon(s,u)-1)\,\nu_2(du)ds.
\end{aligned}
$$

Let $M_t^\epsilon = M_t^{\epsilon,1} + M_t^{\epsilon,2}$, where

$$M_t^{\epsilon,1} = \sqrt{\epsilon}\sum_{i=1}^m \int_0^t \mu_s^{\epsilon,\xi^\epsilon}(F(\sigma_2^{-1}(s)b_2(s,\cdot))^i) d\tilde{W}_s^{\epsilon,i}$$

and

$$M_t^{\epsilon,2} = \int_0^t \int_{\mathbb{U}} \mu_{s-}^{\epsilon,\xi^\epsilon}(F(\lambda(s,\cdot,\mu)-1))\left(\epsilon N_{\lambda^\epsilon}^{\epsilon^{-1},\phi^\epsilon}(ds,du) - \phi^\epsilon(s,u)\nu_2(du)ds\right).$$

To verify the tightness of A_t^ϵ in $D([0,T],\mathbb{R})$, it suffices to show (see Lemma 6.1.2 of [16]) that for all $\delta > 0$, there exists $\tau = \tau_\delta > 0$ such that

$$\sup_{0\le\epsilon\le\epsilon_0} \tilde{\mathbb{P}}^\epsilon\left(\sup_{0<t_1<t_2<\tau} |A_{t_1}^\epsilon - A_{t_2}^\epsilon| > \delta\right) < \delta.$$

Then for arbitrary $\tau > 0$ and a fixed $\delta > 0$,

$$\begin{aligned}
&\sup_{0\le\epsilon\le\epsilon_0} \tilde{\mathbb{P}}^\epsilon\left(\sup_{0<t_1<t_2<\tau} |A_{t_1}^\epsilon - A_{t_2}^\epsilon| > \delta\right) \qquad (53)\\
&\le \sup_{0\le\epsilon\le\epsilon_0} \tilde{\mathbb{P}}^\epsilon\left(\sup_{0<t_1<t_2<\tau} \left|\int_{t_1}^{t_2} \mu_s^{\epsilon,\xi^\epsilon}(\mathcal{L}F)ds\right| > \frac{\delta}{3}\right)\\
&+ \sup_{0\le\epsilon\le\epsilon_0} \tilde{\mathbb{P}}^\epsilon\left(\sup_{0<t_1<t_2<\tau} \left|\int_{t_1}^{t_2} \mu_s^{\epsilon,\xi^\epsilon}(F(\sigma_2^{-1}(s)b_2(s,\cdot)))\psi^\epsilon(s)ds\right| > \frac{\delta}{3}\right)\\
&+ \sup_{0\le\epsilon\le\epsilon_0} \tilde{\mathbb{P}}^\epsilon\left(\sup_{0<t_1<t_2<\tau} \left|\int_{t_1}^{t_2}\int_{\mathbb{U}} \mu_s^{\epsilon,\xi^\epsilon}(F(\lambda(s,\cdot,\mu)-1))\right.\right.\\
&\left.\left.\times(\phi^\epsilon(s,u)-1)\,\nu_2(du)ds\right| > \frac{\delta}{3}\right)\\
&:= D_1 + D_2 + D_3.
\end{aligned}$$

The term D_1 in Eq. (53) is bounded by

$$\begin{aligned}
&\sup_{0\le\epsilon\le\epsilon_0} \frac{9}{\delta^2}\tilde{\mathbb{E}}^\epsilon\left(\sup_{0<t_1<t_2<\tau}\int_{t_1}^{t_2} \mu_s^{\epsilon,\xi^\epsilon}(\mathcal{L}F)ds\right)^2 \qquad (54)\\
&\le \sup_{0\le\epsilon\le\epsilon_0} \frac{9}{\delta^2}\tilde{\mathbb{E}}^\epsilon\left(\delta^2 \sup_{0\le t\le T} \mu_t^{\epsilon,\xi^\epsilon}(\mathcal{L}F)^2\right).
\end{aligned}$$

By Eq. (52) and Assumption 4.1, we can find $\tau_1 > 0$ such that for all $\tau \le \tau_1$, Eq. (54) is bounded by $\delta/3$.

For D_2 in Eq. (53), we note that

$$\begin{aligned}
&\tilde{\mathbb{E}}^\epsilon\left|\int_{t_1}^{t_2} \mu_s^{\epsilon,\xi^\epsilon}(F(\sigma_2^{-1}(s)b_2(s,\cdot)))\psi^\epsilon(s)ds\right|^2\\
&\le \tilde{\mathbb{E}}^\epsilon\left(\int_{t_1}^{t_2}\left|\mu_s^{\epsilon,\xi^\epsilon}(F(\sigma_2^{-1}(s)b_2(s,\cdot)))\right|^2 ds\int_{t_1}^{t_2}|\psi^\epsilon(s)|^2 ds\right) \le MK_2|t_2-t_1|.
\end{aligned}$$

Thus, we can find $\tau_2 > 0$ such that for all $\tau \le \tau_2$ the second term D_2 in Eq. (53) is bounded by $\delta/3$.

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

Now let us consider D_3 in Eq. (53). Using the fundamental inequality, for $a, b \in (0,\infty)$ and any $c \in [1,\infty)$, $ab \le e^{ac} + \frac{1}{c}(b\log b - \log b + 1) = e^{ac} + \frac{1}{c}l(b)$, twice (once with $b = \phi^\epsilon$ and once with $b = 1$), for any $C_1 \in (1,\infty)$ we have

$$\left|\int_{t_1}^{t_2}\int_{\mathbb{U}} \mu_s^{\epsilon,\xi^\epsilon}(F(\lambda(s,\cdot,\mu)-1))\,(\phi^\epsilon(s,u)-1)\,\nu_2(du)ds\right| \tag{55}$$
$$\le 2\sup_{0\le t\le T}\mu_t^{\epsilon,\xi^\epsilon}(F)\int_{t_1}^{t_2}\int_{\mathbb{U}}(\phi^\epsilon(s,u)+1)\nu_2(du)ds$$
$$\le 2\sup_{0\le t\le T}\mu_t^{\epsilon,\xi^\epsilon}(F)\left(2\tau\nu_2(\mathbb{U})e^{C_1} + \frac{M}{C_1}\right).$$

Given any $\epsilon > 0$, we can choose C_1 such that $\frac{M}{C_1} \le \epsilon$, and choose $\tau > 0$ such that $2\tau\nu_2(\mathbb{U})e^{C_1} \le \epsilon$. Hence combining Eqs. (52) and (55) we proved that

$$\lim_{\tau\to 0}\sup_{0<t_1<t_2<\tau}\left|\int_{t_1}^{t_2}\int_{\mathbb{U}}\mu_s^{\epsilon,\xi^\epsilon}(F(\lambda(s,\cdot,\mu)-1))\,(\phi^\epsilon(s,u)-1)\,\nu_2(du)ds\right| = 0.$$

Thus we find $\tau_3 > 0$ such that for $\tau \le \tau_3$, D_3 in Eq. (53) is bounded by $\delta/3$. Consequently, the tightness of $\{A^\epsilon\}_{\epsilon\le\epsilon_0}$ follows from Eq. (53) by taking $\tau = \min\{\tau_1,\tau_2,\tau_3\}$.

Next, considering M^ϵ we have $\tilde{\mathbb{E}}^\epsilon\langle M^{\epsilon,1}\rangle_T \le \epsilon K\tilde{\mathbb{E}}^\epsilon\int_0^T \mu_t^{\epsilon,\xi^\epsilon}(F)^2 ds$ and

$$\tilde{\mathbb{E}}^\epsilon\langle M^{\epsilon,2}\rangle_T = \epsilon\tilde{\mathbb{E}}^\epsilon\left(\int_0^T\int_{\mathbb{U}}\mu_t^{\epsilon,\xi^\epsilon}(F(\lambda(s,\cdot,\mu)-1))^2\phi^\epsilon(s,u)\nu_2(du)ds\right)$$
$$\le \epsilon 4\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\le t\le T}\mu_t^{\epsilon,\xi^\epsilon}(F)^2\int_0^T\int_{\mathbb{U}}\phi^\epsilon(s,u)\nu_2(du)ds\right).$$

By Eqs. (46) and (52), we have $\tilde{\mathbb{E}}^\epsilon\sup_{0\le t\le T}\langle M^\epsilon\rangle_t$ converges to 0 as $\epsilon\to 0$. Then according to Theorem 6.1.1 in [16], for any $F \in \mathcal{D}(\mathcal{L})$, the sequence of semimartingales $\mu_t^{\epsilon,\xi^\epsilon}(F) = \mu_0^{\epsilon,\xi^\epsilon}(F) + A_t^\epsilon + M_t^\epsilon$ is tight in $D([0,T],\mathbb{R})$.

Next we show that $\mathcal{G}^0(\int_0^\cdot\psi(s)ds, \nu_2^\phi)$ is the weak limit of $\mu^{\epsilon,\xi^\epsilon}$. Note that

$$\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\le t\le T}\left|M_t^{\epsilon,1}\right|^2\right) \le \epsilon K\tilde{\mathbb{E}}^\epsilon\int_0^T\mu_t^{\epsilon,\xi^\epsilon}(F)^2 ds$$

and

$$\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\le t\le T}\left|M_t^{\epsilon,2}\right|^2\right) \le \epsilon K\tilde{\mathbb{E}}^\epsilon\left(\sup_{0\le t\le T}\mu_t^{\epsilon,\xi^\epsilon}(F)^2\int_0^T\int_{\mathbb{U}}\phi^\epsilon(s,u)\nu_2(du)ds\right).$$

Therefore, $M_t^{\epsilon,1}$ and $M_t^{\epsilon,2}$ converge to 0 in distribution as $\epsilon\to 0$. Let $(\mu_t^{0,\xi},\psi,\phi,0,0)$ be any limit point of the tight sequence $(\mu_t^{\epsilon,\xi^\epsilon},\psi^\epsilon,\phi^\epsilon,M_t^{\epsilon,1},\ M_t^{\epsilon,2})$ for $\epsilon \in (0,\epsilon_0)$. Without loss of generality, we assume that the convergence is almost sure by using the Skorokhod representation theorem. Note that for any test function F in $\mathcal{D}(\mathcal{L})$, we have

$$\mu_t^{\epsilon,\xi^\epsilon}(F) = \mu_0^{\epsilon,\xi^\epsilon}(F) + \int_0^t\mu_s^{\epsilon,\xi^\epsilon}(\mathcal{L}F)ds + M_t^{\epsilon,1}$$
$$+ \int_0^t\mu_s^{\epsilon,\xi^\epsilon}(F(\sigma_2^{-1}(s)b_2(s,\cdot)))\psi^\epsilon(s)ds$$
$$+M_t^{\epsilon,2} + \int_0^t\int_{\mathbb{U}}\mu_s^{\epsilon,\xi^\epsilon}(F(\lambda(s,\cdot,\mu)-1))\,(\phi^\epsilon(s,u)-1)\,\nu_2(du)ds.$$

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

Taking $\epsilon \to 0$, along the lines of Theorem 4.5 and Proposition 4.2 we see that $\mu^{0,\xi}$ must solve Eq. (31). Then the uniqueness of the solution to Eq. (31) completes the proof. ■

Acknowledgments

The authors would like to thank an anonymous reviewer for their comments which substantially improved the manuscript. VM's research has been partially supported by the Army Research Office, W911NF-17-1-0313 and the National Science Foundation, MCB-1715794 and DMS-1821241. JX's research is supported partially by Southern University of Science and Technology Start up fund Y01286120 and National Science Foundation of China grants 61873325 and 11831010.

Appendix A. Proof of Theorem 3.1

Proof. Considering Eq. (12) and replacing F by $T_\delta F$ in the Zakai equation (8), we have that $\langle Z_s^\delta, F\rangle_0$ equals to

$$\langle \mu_0, T_\delta F\rangle + \int_0^t \langle \mu_s, \mathcal{L}T_\delta F\rangle ds + \sum_{i=1}^m \int_0^t \langle \mu_s, (T_\delta F)(\sigma_2^{-1}(s)b_2(s,\cdot))^i\rangle d\tilde{W}_s^i$$
$$+ \int_0^t \int_{\mathbb{U}} \langle \mu_{s-}, (T_\delta F)(\lambda(s,\cdot,u)-1)\rangle \tilde{N}(ds,du). \tag{A.1}$$

Note that for any $\nu \in \mathcal{M}_F(\mathbb{R}^d)$,

$$\langle \nu, \mathcal{L}T_\delta F\rangle = \langle \nu, \sum_{i=1}^d \partial_i(T_\delta F)b_1^i + \frac{1}{2}\sum_{i,j=1}^d \sum_{k=1}^n \partial_{ij}^2(T_\delta F)\sigma_1^{ik}\sigma_1^{jk}\rangle$$
$$+\langle \nu, \int_{\mathbb{U}\setminus\mathbb{U}_1} [T_\delta F(x+g_1(x,u)) - T_\delta F(x)]\nu_1(du)\rangle \tag{A.2}$$
$$+\langle \nu, \int_{\mathbb{U}_1} [T_\delta F(x+f_1(x,u)) - T_\delta F(x) - \sum_{i=1}^d \partial_i(T_\delta F)f_1^i(x,u)]\nu_1(du)\rangle.$$

Furthermore, by Lemma 3.2, we have

$$\langle \nu, \sum_{i=1}^d \partial_i(T_\delta F)b_1^i + \frac{1}{2}\sum_{i,j=1}^d \sum_{k=1}^n \partial_{ij}^2(T_\delta F)\sigma_1^{ik}\sigma_1^{jk}\rangle \tag{A.3}$$
$$= \sum_{i=1}^d \langle b_1^i\nu, T_\delta\partial_i F\rangle + \frac{1}{2}\sum_{i,j=1}^d \sum_{k=1}^n \langle \sigma_1^{ik}\sigma_1^{jk}\nu, T_\delta\partial_{ij}^2 F\rangle$$
$$= \sum_{i=1}^d \langle T_\delta(b_1^i\nu), \partial_i F\rangle_0 + \frac{1}{2}\sum_{i,j=1}^d \sum_{k=1}^n \langle T_\delta(\sigma_1^{ik}\sigma_1^{jk}\nu), \partial_{ij}^2 F\rangle_0$$
$$= -\sum_{i=1}^d \langle \partial_i T_\delta(b_1^i\nu), F\rangle_0 + \frac{1}{2}\sum_{i,j=1}^d \sum_{k=1}^n \langle \partial_{ij}^2 T_\delta(\sigma_1^{ik}\sigma_1^{jk}\nu), F\rangle_0.$$

In addition, $\langle \nu, (T_\delta F)(\sigma_2^{-1}(s)b_2(s,\cdot))^i\rangle = \langle T_\delta(\sigma_2^{-1}(s)b_2(s,\cdot)^i\nu), F\rangle_0$, where $(T_\delta F)(\sigma_2^{-1}(s)$ $b_2(s,\cdot))^i$ is the ith entry of vector $(T_\delta F)(\sigma_2^{-1}(s)b_2(s,\cdot))$, as well as $\langle \nu, (T_\delta F)(\lambda(s,\cdot,u)-1)\rangle$

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

$= \langle T_\delta((\lambda(s,\cdot,u)-1)v), F\rangle_0$. Suppose that $F \geq 0$, then the case for general F follows from the linearity. By Fubini's theorem,

$$\begin{aligned}\langle \nu, \int_{\mathbb{U}\setminus\mathbb{U}_1} T_\delta F(x+g_1(x,u))\nu_1(du)\rangle &= \int_{\mathbb{U}\setminus\mathbb{U}_1}\int_{\mathbb{R}^d} T_\delta F(x+g_1(x,u))\nu(dx)\nu_1(du)\\ &= \int_{\mathbb{U}\setminus\mathbb{U}_1}\langle \nu, T_\delta F(x+g_1(x,u))\rangle \nu_1(du). \end{aligned} \tag{A.4}$$

Consequently, taking into consideration Eqs. (A.2), (A.3) and (A.4) into Eq. (A.1), one can get $\langle Z_s^\delta, F\rangle_0$ equals to

$$\begin{aligned}&\langle Z_0^\delta, F\rangle_0 - \sum_{i=1}^d \int_0^t \langle \partial_i T_\delta(b_1^i\mu_s), F\rangle_0 ds + \frac{1}{2}\sum_{i,j=1}^d\sum_{k=1}^n \int_0^t \langle \partial_{ij}^2 T_\delta(\sigma_1^{ik}\sigma_1^{jk}\mu_s), F\rangle_0 ds\\
&+ \int_0^t \int_{\mathbb{U}\setminus\mathbb{U}_1} \left[\langle T_\delta\mu_s, F(\cdot+g_1(\cdot,u))\rangle_0 - \langle T_\delta\mu_s, F\rangle_0\right]\nu_1(du)ds\\
&+ \sum_{i=1}^m \int_0^t \langle T_\delta(\sigma_2^{-1}(s)b_2(s,\cdot))^i\mu_s, F\rangle_0 d\tilde{W}_s^i\\
&+ \int_0^t\int_{\mathbb{U}} \langle T_\delta\left((\lambda(s,\cdot,u)-1)\mu_{s-}\right), F\rangle_0 \tilde{N}(ds,du)\\
&+ \int_0^t\int_{\mathbb{U}_1}\left[\langle T_\delta\mu_s, F(\cdot+f_1(\cdot,u))\rangle_0 - \langle T_\delta\mu_s, F\rangle_0 - \sum_{i=1}^d \langle \partial_i T_\delta(f_1^i(\cdot,u)\mu_s), F\rangle_0\right]\\
&\times\ \nu_1(du)ds.\end{aligned}$$

Then by Itô's formula, we have $\langle Z_t^\delta, F\rangle_0^2$ equals to

$$\begin{aligned}&\langle Z_0^\delta, F\rangle_0^2 + 2\int_0^t \langle Z_s^\delta, F\rangle_0 \Bigg\{-\sum_{i=1}^d \langle \partial_i T_\delta(b_1^i\mu_s), F\rangle_0 + \frac{1}{2}\sum_{i,j=1}^d\sum_{k=1}^n \langle \partial_{ij}^2 T_\delta(\sigma_1^{ik}\sigma_1^{jk}\mu_s), F\rangle_0\\
&+ \int_{\mathbb{U}\setminus\mathbb{U}_1}\left[\langle Z_s^\delta, F(\cdot+g_1(\cdot,u))\rangle_0 - \langle Z_s^\delta, F\rangle_0\right]\nu_1(du)\\
&+ \int_{\mathbb{U}_1}\left[\langle Z_s^\delta, F(\cdot+f_1(\cdot,u))\rangle_0 - \langle Z_s^\delta, F\rangle_0 - \langle \partial_i T_\delta(f_1^i(\cdot,u)\mu_s), F\rangle_0\right]\nu_1(du)\Bigg\}ds\\
&+ 2\sum_{i=1}^m\int_0^t \langle Z_s^\delta, F\rangle_0 \langle T_\delta((\sigma_2^{-1}(s)b_2(s,\cdot))^i\mu_s), F\rangle_0 d\tilde{W}_s^i\\
&+ \int_0^t \left|\langle T_\delta(\sigma_2^{-1}(s)b_2(s,\cdot)\mu_s), F\rangle_0\right|^2 ds\\
&+ \int_0^t\int_{\mathbb{U}}\left[\left(\langle Z_{s-}^\delta, F\rangle_0 + \langle T_\delta\left((\lambda(s,\cdot,u)-1)\mu_{s-}\right), F\rangle_0\right)^2 - \langle Z_{s-}^\delta, F\rangle_0^2\right]\tilde{N}(ds,du)\\
&+ \int_0^t\int_{\mathbb{U}}\Big[\left(\langle Z_s^\delta, F\rangle_0 + \langle T_\delta\left((\lambda(s,\cdot,u)-1)\mu_s\right), F\rangle_0\right)^2 - \langle Z_s^\delta, F\rangle_0^2\\
&-2\langle T_\delta\left((\lambda(s,\cdot,u)-1)\mu_s\right), F\rangle_0\langle Z_s^\delta, F\rangle_0\Big]\nu_2(du)ds.\end{aligned}$$

Summing over η in a CONS of H_0 we have

$$\sum_\eta \langle Z_0^\delta, \eta\rangle_0^2 = \|Z_0^\delta\|_0^2, \qquad \sum_\eta \langle Z_t^\delta, \eta\rangle_0 \langle \partial_i T_\delta(b_1^i \mu_t), \eta\rangle_0 = \langle Z_t, \partial_i T_\delta(b_1^i\mu_t)\rangle_0.$$

Similarly, we eventually get that $\|Z_t^\delta\|_0^2$ equals to

$$\begin{aligned}
&\|Z_0^\delta\|_0^2 - 2\sum_{i=1}^d \int_0^t \langle Z_s^\delta, \partial_i T_\delta(b_1^i\mu_s)\rangle_0 ds + \sum_{i,j=1}^d \sum_{k=1}^n \int_0^t \langle Z_s^\delta, \partial_{ij}^2 T_\delta(\sigma_1^{ik}\sigma_1^{jk}\mu_s)\rangle_0 ds \\
&+ 2\int_0^t \int_{\mathbb{U}\setminus\mathbb{U}_1} [\langle Z_s^\delta, \tilde{Z}_s^{\delta,g_1}\rangle_0 - \|Z_s^\delta\|_0^2]\nu_1(du)ds \qquad (A.5)\\
&+ \int_0^t \int_{\mathbb{U}_1} [\langle Z_s^\delta, \tilde{Z}_s^{\delta,f_1}\rangle_0 - \|Z_s^\delta\|_0^2 - \sum_{i=1}^d \langle Z_s^\delta, \partial_i T_\delta(f_1^i(\cdot,u)\mu_s)\rangle_0]\nu_1(du)ds \\
&+ 2\sum_{i=1}^m \int_0^t \langle Z_s^\delta, T_\delta((\sigma_2^{-1}(s)b_2(s,\cdot))^i\mu_s)\rangle_0 d\tilde{W}_s^i + \int_0^t \|T_\delta(\sigma_2^{-1}(s)b_2(s,\cdot)\mu_s)\|_0^2 ds \\
&+ \int_0^t \int_{\mathbb{U}} [2\langle Z_{s-}^\delta, T_\delta((\lambda(s,\cdot,u)-1)\mu_s)\rangle_0 + \|T_\delta((\lambda(s,\cdot,u)-1)\mu_{s-})\|_0^2]\tilde{N}(ds,du) \\
&+ \int_0^t \int_{\mathbb{U}} \|T_\delta((\lambda(s,\cdot,u)-1)\mu_s)\|_0^2 \nu_2(du)ds,
\end{aligned}$$

provided Z_t^{δ,g_1}, Z_t^{δ,f_1} are defined as Eqs. (16) and (17) respectively. Eventually, taking expectation on Eq. (A.5) gives Eq. (14). ■

Appendix B. Proof of Lemma 3.5

Proof. Note that $\sum_{i,j=1}^d \langle T_\delta\zeta, \partial_{ij}^2 T_\delta(\sigma_1\sigma_1^*)_{ij}\zeta\rangle_0$ equals to

$$\sum_{i,j=1}^d \int_{\mathbb{R}^d} dx \int_{\mathbb{R}^d} \zeta(dy) G_\delta(x-y) \int_{\mathbb{R}^d} \zeta(dz) \partial_{x_i x_j}^2 G_\delta(x-z) \sum_{k=1}^n \sigma_1^{ik}(z)\sigma_1^{jk}(z). \qquad (B.1)$$

Employing the semigroup property of G_δ in Lemma 3.1, we have $\int_{\mathbb{R}^d} G_\delta(x-y)G_\delta(x-z)dx = G_{2\delta}(y-z)$. Noticing that $\partial_i G_\delta(x) = -\frac{x_i}{\delta}G_\delta(x)$, we have $\partial_{ij}^2 G_\delta(x) = \left(\frac{x_i x_j}{\delta^2} - \frac{1_{i=j}}{\delta}\right) G_\delta(x)$. Due to the fact that $\partial_{x_i x_j}^2 G_\delta(x-z) = \partial_{z_i z_j}^2 G_\delta(x-z)$, Eq. (B.1) is written as

$$\begin{aligned}
&\sum_{i,j=1}^d \int_{\mathbb{R}^d} \zeta(dy) \int_{\mathbb{R}^d} \zeta(dz) \sum_{k=1}^n \sigma_1^{ik}(z)\sigma_1^{kj}(z) \partial_{z_i z_j}^2 \int_{\mathbb{R}^d} G_\delta(x-y)G_\delta(x-z)dx \qquad (B.2)\\
&= \sum_{i,j=1}^d \int_{\mathbb{R}^d} \zeta(dy) \int_{\mathbb{R}^d} \zeta(dz) \left(\frac{(z_i-y_i)(z_j-y_j)}{4\delta^2} - \frac{1_{i=j}}{2\delta}\right) G_{2\delta}(z-y) \sum_{k=1}^n \sigma_1^{ik}(z)\sigma_1^{jk}(z).
\end{aligned}$$

By symmetry of y and z in Eq. (B.2), Eq. (B.1) is further given by

$$\begin{aligned}
&\sum_{i,j=1}^d \int_{\mathbb{R}^d} \zeta(dy) \int_{\mathbb{R}^d} \zeta(dz) \left(\frac{(z_i-y_i)(z_j-y_j)}{4\delta^2} - \frac{1_{i=j}}{2\delta}\right) \\
&\times G_{2\delta}(z-y) \frac{1}{2} \sum_{k=1}^n \left(\sigma_1^{ik}(z)\sigma_1^{jk}(z) + \sigma_1^{ik}(y)\sigma_1^{jk}(y)\right). \qquad (B.3)
\end{aligned}$$

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

Similarly, the second term on the left hand side of Eq. (24) can be expressed as

$$-\sum_{i,j=1}^{d}\int_{\mathbb{R}^d}\zeta(dy)\int_{\mathbb{R}^d}\zeta(dz)\left(\frac{(z_i-y_i)(z_j-y_j)}{4\delta^2}-\frac{1_{i=j}}{2\delta}\right)$$

$$\times\; G_{2\delta}(z-y)\frac{1}{2}\sum_{k=1}^{n}\left(\sigma_1^{ik}(y)\sigma_1^{jk}(z)+\sigma_1^{ik}(z)\sigma_1^{jk}(y)\right). \tag{B.4}$$

Now adding Eqs. (B.3) and (B.4), the left-hand side of (24) is given by

$$\sum_{i,j=1}^{d}\int_{\mathbb{R}^d}\zeta(dy)\int_{\mathbb{R}^d}\zeta(dz)\left(\frac{(z_i-y_i)(z_j-y_j)}{4\delta^2}-\frac{1_{i=j}}{2\delta}\right)$$

$$\times\; G_{2\delta}(z-y)\frac{1}{2}\sum_{k=1}^{n}\left(\sigma_1^{ik}(y)-\sigma_1^{ik}(z)\right)\left(\sigma_1^{jk}(y)-\sigma_1^{jk}(z)\right).$$

Using the identity that $G_\delta(x)=\exp\left(-\frac{|x|^2}{4\delta}\right)2^{d/2}G_{2\delta}(x)$, the Lipschitz continuity of σ_1 and Lemma 3.1, we have the quantity above estimated by

$$\sum_{i,j=1}^{d}\int_{\mathbb{R}^d}\zeta(dy)\int_{\mathbb{R}^d}\zeta(dz)\left(\frac{|z-y|^2}{4\delta^2}+\frac{1}{2\delta}\right)\exp\left(-\frac{|z-y|^2}{4\delta}\right)$$

$$\times\, 2^{d/2}G_{4\delta}(z-y)\frac{1}{2}K^2|z-y|^2$$

$$\le 4K^2\sum_{i,j=1}^{d}\int_{\mathbb{R}^d}|\zeta|(dy)\int_{\mathbb{R}^d}|\zeta|(dz)2^{d/2}G_{4\delta}(z-y)$$

$$= d^2 2^{2+d/2}K^2\|T_{2\delta}(|\zeta|)\|_0^2 \le d^2 2^{2+d/2}K^2\|T_{\delta}(|\zeta|)\|_0^2.$$

The lemma then follows with $K_1=d^2 2^{2+d/2}K^2$. ■

Appendix C. Proof of Proposition 4.1

Proof. By Itô's formula,

$$d\Lambda_t^\epsilon=\Lambda_t^\epsilon\left\{\sqrt{\epsilon}(\sigma_2^{-1}(t)b_2(t,X_t))^*dW_t\right.$$

$$\left.+\frac{\epsilon}{2}|\sigma_2^{-1}(t)b_2(t,X_t)|^2dt+\int_{\mathbb{U}}(1-\lambda(t,X_t,u))\nu_2(du)dt\right\}$$

$$+\frac{\epsilon}{2}\Lambda_t^\epsilon|\sigma_2^{-1}(t)b_2(t,X_t)|^2dt+\int_{\mathbb{U}}\Lambda_t^\epsilon(\lambda(t,X_{t-},u)-1)\epsilon N_{\lambda^\epsilon}^{\epsilon^{-1}}(dt,du).$$

Then

$$\Lambda_t^\epsilon=1+\sqrt{\epsilon}\int_0^t\Lambda_s^\epsilon(\sigma_2^{-1}(s)b_2(s,X_s))^*d\tilde{W}_s^\epsilon$$

$$+\int_0^t\int_{\mathbb{U}}\Lambda_{s-}(\lambda(s,X_{s-},u)-1)(\epsilon N_{\lambda^\epsilon}^{\epsilon^{-1}}(ds,du)-\nu_2(du)ds).$$

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.

By Girsanov's theorem, $\tilde{W}^{\epsilon}_t = W_t + \sqrt{\epsilon}\int_0^t \sigma_2^{-1}(s)b_2(s, X_s)ds$, is a $\tilde{\mathbb{P}}^{\epsilon}$-Brownian motion, and $\tilde{N}^{\epsilon^{-1}} = N^{\epsilon^{-1}}_{\lambda^{\epsilon}}(dt, du) - \epsilon^{-1}\nu_2(du)dt$, is a $\tilde{\mathbb{P}}^{\epsilon}$-Poisson martingale measure. The Zakai equation (27) is obtained by the same argument of Proposition 3.1. We now show the derivation of the Kushner–Stratonovich equation (28). Note that

$$\mu^{\epsilon}_t(1) = \mu_0(1) + \sqrt{\epsilon}\sum_{i=1}^m \int_0^t \mu^{\epsilon}_s((\sigma_2^{-1}(s)b_2(s,\cdot))^i)d\tilde{W}^{\epsilon,i}_s$$
$$+ \int_0^t \int_{\mathbb{U}} \mu^{\epsilon}_{s-}(\lambda(s,\cdot,\mu) - 1)(\epsilon N^{\epsilon^{-1}}_{\lambda^{\epsilon}}(ds, du) - \nu_2(du)ds).$$

By Itô's formula, we have that $d\frac{1}{\mu^{\epsilon}_t(1)}$ equals to

$$- \frac{\sqrt{\epsilon}}{\mu^{\epsilon}_t(1)^2}\sum_{i=1}^m \mu^{\epsilon}_t((\sigma_2^{-1}(t)b_2(t,\cdot))^i)d\tilde{W}^{\epsilon,i}_t + \frac{\epsilon}{\mu^{\epsilon}_t(1)^3}\mu^{\epsilon}_t(\sigma_2^{-1}(t)b_2(t,\cdot))dt$$
$$+ \int_{\mathbb{U}}\left(\frac{1}{\mu^{\epsilon}_t(1) + \epsilon\mu^{\epsilon}_t(\lambda(t,\cdot,u) - 1)} - \frac{1}{\mu^{\epsilon}_t(1)}\right)\left(N^{\epsilon^{-1}}_{\lambda^{\epsilon}}(dt, du) - \epsilon^{-1}\nu_2(du)dt\right)$$
$$+ \frac{1}{\epsilon}\int_{\mathbb{U}}\left(\frac{1}{\mu^{\epsilon}_t(1) + \epsilon\mu^{\epsilon}_t(\lambda(t,\cdot,u) - 1)} - \frac{1}{\mu^{\epsilon}_t(1)} + \epsilon\mu^{\epsilon}_t(\lambda(t,\cdot,u) - 1)\frac{1}{\mu^{\epsilon}_t(1)^2}\right)\nu_2(du)dt.$$

Hence, the quadratic variation, $d[\mu^{\epsilon}(F), \mu^{\epsilon}(1)^{-1}]_t$, is given by

$$- \frac{\epsilon}{\mu^{\epsilon}_t(1)^2}\sum_{i=1}^m \mu^{\epsilon}_t((\sigma_2^{-1}(t)b_2(t,\cdot))^i)\mu^{\epsilon}_t(F(\sigma_2^{-1}(t)b_2(t,\cdot))^i)dt$$
$$+ \int_{\mathbb{U}}\left(\frac{1}{\mu^{\epsilon}_t(1) + \epsilon\mu^{\epsilon}_t(\lambda(t,\cdot,u) - 1)} - \frac{1}{\mu^{\epsilon}_t(1)}\right)\epsilon\mu^{\epsilon}_{s-}(F(\lambda(s,\cdot,\mu) - 1))N^{\epsilon^{-1}}_{\lambda^{\epsilon}}(ds, du).$$

Eventually, one can easily get Eq. (28) by the following the product of Itô's formula $d\frac{\mu^{\epsilon}_t(F)}{\mu^{\epsilon}_t(1)} = \frac{1}{\mu^{\epsilon}_t(1)}d\mu^{\epsilon}_t(F) + \mu^{\epsilon}_t(F)d\frac{1}{\mu^{\epsilon}_t(1)} + d[\mu^{\epsilon}(F), \mu^{\epsilon}(1)^{-1}]_t$. ■

References

[1] F. Bao, V. Maroulas, Adaptive meshfree backward SDE filter, SIAM J. Sci. Comput. 39 (6) (2017) A2664–A2683.
[2] J. Bao, C. Yuan, Large deviations for neutral functional SDEs with jumps, Stochastics 87 (1) (2015) 48–70.
[3] A. Budhiraja, J. Chen, P. Dupuis, Large deviations for stochastic partial differential equations driven by a Poisson random measure, Stochastic Process. Appl. 123 (2) (2013) 523–560.
[4] A. Budhiraja, P. Dupuis, V. Maroulas, Variational representations for continuous time processes, Ann. Inst. Henri Poincaré Probab. Stat. 47 (3) (2011) 725–747.
[5] Y. Cai, J. Huang, V. Maroulas, Large deviations of mean-field stochastic differential equations with jumps, Statist. Probab. Lett. 96 (2015) 1–9.
[6] C. Ceci, K. Colaneri, Nonlinear filtering for jump diffusion observations, Adv. Appl. Probab. 44 (3) (2012) 678–701.
[7] C. Ceci, K. Colaneri, The Zakai equation of nonlinear filtering for jump-diffusion observations: existence and uniqueness, Appl. Math. & Optim. 69 (1) (2014) 47–82.
[8] D.D. Creal, Analysis of filtering and smoothing algorithms for Lévy-driven stochastic volatility models, Comput. Statist. Data Anal. 52 (6) (2008) 2863–2876.
[9] H. Dong, Z. Wang, D.W. Ho, H. Gao, Robust H_∞ filtering for markovian jump systems with randomly occurring nonlinearities and sensor saturation: the finite-horizon case, IEEE Trans. Signal Process. 59 (7) (2011) 3048–3057.
[10] J. Duan, A. Millet, Large deviations for the Boussinesq equations under random influences, Stochastic Process. Appl. 119 (6) (2009) 2052–2081.

[11] P. Dupuis, R.S. Ellis, A Weak Convergence Approach to the Theory of Large Deviations, Vol. 902, John Wiley & Sons, 2011.
[12] E. Evangelou, V. Maroulas, Sequential empirical Bayes method for filtering dynamic spatiotemporal processes, Spatial Statistics 21 (A) (2017) 114–129.
[13] B. Grigelionis, R. Mikulevicius, Nonlinear filtering equations for stochastic processes with jumps, in: The Oxford Handbook of Nonlinear Filtering, 2011, pp. 95–128.
[14] O. Hijab, Asymptotic Bayesian estimation of a first order equation with small diffusion, Ann. Probab. (1984) 890–902.
[15] M.S. Johannes, N.G. Polson, J.R. Stroud, Optimal filtering of jump diffusions: extracting latent states from asset prices, Rev. Financ. Stud. 22 (7) (2009) 2759–2799.
[16] G. Kallianpur, J. Xiong, Stochastic Differential Equations in Infinite Dimensional Spaces, in: Lecture Notes-Monograph Series, vol. 26, 1995, iii–342.
[17] K. Kang, V. Maroulas, I. Schizas, F. Bao, Improved distributed particle filters for tracking in a wireless sensor network, Comput. Statist. Data Anal. 117 (2018) 90–108.
[18] T.G. Kurtz, D.L. Ocone, Unique characterization of conditional distributions in nonlinear filtering, Ann. Probab. 16 (1) (1988) 80–107.
[19] T.G. Kurtz, J. Xiong, Particle representations for a class of nonlinear SPDEs, Stochastic Process. Appl. 83 (1) (1999) 103–126.
[20] V. Mandrekar, T. Meyer-Brandis, F. Proske, A Bayes formula for nonlinear filtering with Gaussian and Cox noise, J. Probab. Stat. (2011).
[21] V. Maroulas, Large deviations for infinite-dimensional stochastic systems with jumps, Mathematika 57 (1) (2011) 175–192.
[22] V. Maroulas, A. Nebenführ, Tracking rapid intracellular movements: a Bayesian random set approach, Ann. Appl. Stat. 9 (2) (2015) 926–949.
[23] V. Maroulas, X. Pan, J. Xiong, Statistical inference for the intensity in a partially observed jump diffusion, J. Math. Anal. Appl. 472 (1) (2019) 1–10.
[24] V. Maroulas, J. Xiong, Large deviations for optimal filtering with fractional Brownian motion, Stochastic Process. Appl. 123 (6) (2013) 2340–2352.
[25] T. Meyer-Brandis, F. Proske, Explicit solution of a non-linear filtering problem for Lévy processes with application to finance, Appl. Math. Optim. 50 (2) (2004) 119–134.
[26] X. Pan, S.M. Djouadi, Estimation and identification for wireless sensor network undergoing uncertain jumps, in: 2017 Proceedings of the Conference on Control and its Applications, SIAM, 2017, pp. 46–53.
[27] É. Pardoux, O. Zeitouni, Quenched large deviations for one dimensional nonlinear filtering, SIAM J. Control Optim. 43 (4) (2004) 1272–1297.
[28] S. Popa, S. Sritharan, Nonlinear filtering of Itô–Lévy stochastic differential equations with continuous observations, Commun. Stoch. Anal. 3 (3) (2009) 313–330.
[29] H. Qiao, J. Duan, Nonlinear filtering of stochastic dynamical systems with Lévy noises, Adv. Appl. Probab. 47 (03) (2015) 902–918.
[30] M. Röckner, T. Zhang, X. Zhang, Large deviations for stochastic tamed 3D Navier-Stokes equations, Appl. Math. Optim. 61 (2) (2010) 267–285.
[31] I. Sgouralis, A. Nebenführ, V. Maroulas, A Bayesian topological framework for the identification and reconstruction of subcellular motion, SIAM J. Imaging Sci. 10 (2) (2017) 871–899.
[32] P. Shi, M. Mahmoud, S.K. Nguang, A. Ismail, Robust filtering for jumping systems with mode-dependent delays, Signal Process. 86 (1) (2006) 140–152.
[33] J. Xiong, Large deviation principle for optimal filtering, Appl. Math. Optim. 47 (2) (2003) 151–165.
[34] J. Xiong, An Introduction to Stochastic Filtering Theory, Vol. 18, Oxford University Press, 2008.
[35] J. Xiong, J. Zhai, Large deviations for locally monotone stochastic partial differential equations driven by Lévy noise, arXiv preprint arXiv:1605.06618.
[36] X. Yang, J. Zhai, T. Zhang, Large deviations for SPDEs of jump type, Stoch. Dyn. 15 (04) (2015) 1550026.

Please cite this article as: V. Maroulas, X. Pan and J. Xiong, Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.02.009.