J. Math. Anal. Appl. 472 (2019) 1–10

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Statistical inference for the intensity in a partially observed jump diffusion

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, 518055, China

ARTICLE INFO

Article history:
Received 4 August 2018
Available online 15 October 2018
Submitted by U. Stadtmueller

Keywords:
Statistical inference
Moderate deviations
Poisson process
Jump diffusions

ABSTRACT

This article focuses on a partially observed linear diffusion with jumps described by a Poisson process. Precisely, we study an inferential problem for the intensity of the Poisson process by establishing a moderate deviation principle for the least-squares estimator of the intensity.

1. Introduction

An array of studies have been devoted to jump-diffusion models in filtering and control problems because of their applicability in finance, wireless sensor networks, biology, etc. (see, e.g., [4,15,20–22,26,23]. However, it is often of great interest to estimate the associated parameters instead of just the typical state estimation (see, e.g., [10,14,18,24,11]. Precisely, in this work, the signal process is described by

$$dX_t = aX_t dt + bdW_t + \delta dN_t, \quad X_0 = x_0 \in \mathbb{R}, \tag{1}$$

where W_t is a standard Brownian motion and N_t is an independent Poisson process with intensity λ on the filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\in[0,T]})$ for some $T > 0$ and constants a, b, δ in $\mathbb{R}$. However, Eq. (1) is not observed, but instead partial information is propagated via the equation

$$Y_t = h\int_0^t X_s ds + B_t, \quad t \in [0, T], \tag{2}$$

* Corresponding author.
E-mail addresses: vmaroula@utk.edu (V. Maroulas), pan@math.utk.edu (X. Pan), xiongj@sustc.edu.cn (J. Xiong).

https://doi.org/10.1016/j.jmaa.2018.10.026
0022-247X/

where the driving noise, B_t, is another standard Brownian motion, independent of W_t and N_t in Eq. (1), and $h \in \mathbb{R}$ is a constant that explains the relationship between the signal and its observation.

The goal of this work is the statistical inference of λ, where tail probabilities of a pertinent statistics need to be established. However, inference for partially observed jump-diffusion processes is complex and challenging because such distributions typically do not have a closed form. To that end, the typical maximum likelihood estimation method for diffusion cases (e.g., [5,8]) is rarely tractable and thus need to be estimated. Indeed, very few studies have investigated the estimation of Poisson intensity for partially observed jump-diffusion systems. Johannes et al. [18] studied a discretized model using a particle filtering approach. Vacarescu [25] investigated the maximum likelihood estimation for the intensity of a counting process using an expectation maximization algorithm, where the intensity is restricted to a pertinent dynamic process. Djouadi et al. [10] proposed a consistent least-squares estimator (LSE) for the intensity λ of the Poisson process when the drift coefficient in Eq. (1) is negative (i.e., $a < 0$) and established a central limit theorem for the LSE given in Theorem 2.1. However, the associated LSE's asymptotic variance is a function of the true value of the intensity, and thus the approximation cannot be used to resolve the inference problem of the intensity of the signals' jumps.

To that end, the moderate deviation principle of the LSE is considered herein as an alternative convergence scheme influenced by the early studies of Chernoff [7] and Bahadur [2] and more recent developments of large deviations and moderate deviations in statistics [1,3,12,13,19,17]. The moderate deviations provide us with the rates of convergence of the LSE and prescribe a strategy for approximating the power of the associated hypothesis testing as well as computing the $(1-\alpha_0)\cdot 100\%$ confidence interval for the intensity.

The remainder of this article is organized as follows. In Section 2, we state the problem and introduce the LSE of the Poisson intensity. Some motivating elements of large deviation theory are also presented. The inferential results, the power of hypothesis testing, and the confidence interval are presented in Section 3 by our proving a moderate deviation principle.

2. Preliminaries

Consider a latent signal described by the jump diffusion given in Eq. (1) and observed by the observation process in Eq. (2). The asymptotic behavior of the LSE $\hat{\lambda}_n$ of the intensity λ is considered when the drift coefficient $a < 0$ in Eq. (1). According to [10], $a \geq 0$ does not yield interesting behavior, and one should notice that the trajectory of X_t blows up when $a > 0$. Assume that Y_t is observable at discrete times $\{t_0^n, t_1^n, ..., t_n^n\} \in [0, T]$ and precisely $t_i^n = i$, $i = 1, ..., n$, such that $T = n \to \infty$ without causing any loss of generality. Define the innovation process $Z_{t_i^n} := Y_{t_i^n} - Y_{t_{i-1}^n} = h\int_{t_{i-1}^n}^{t_i^n} X_s ds + \varepsilon_i$, where ε_i denotes the normally distributed and independent Brownian increments $B_{t_i^n} - B_{t_{i-1}^n} \sim$ i.i.d. $N(0, \Delta t_{i-1}^n)$, $\Delta t_{i-1}^n = t_i^n - t_{i-1}^n$, $i = 1, ..., n$. Let Θ denote the parameter space of the intensity of the driving Poisson process. The LSE $\hat{\lambda}_n$ of λ minimizes the residual sum of squares, $\hat{\lambda}_n = \arg\min_{\lambda\in\Theta} Q(\lambda)$, where $Q(\lambda) = \sum_{i=1}^{n}(Z_{t_i^n} - \mathbb{E}(Z_{t_i^n}))^2$. The solution of the optimization problem derived in [10] yields an unbiased and consistent LSE of the intensity λ:

$$\hat{\lambda}_n = \frac{a}{h\delta\sum_{i=1}^{n}(\frac{1}{a}d_i - 1)^2}\sum_{i=1}^{n}\left(Z_i - \frac{hx_0}{a}d_i\right)\left(\frac{1}{a}d_i - 1\right), \tag{3}$$

where $d_i = e^{ai} - e^{ai-a}$. The following lemma provides an alternative definition of the LSE in Eq. (3), the asymptotic normality of which is shown in Theorem 2.1. The reader should refer to [10] for the proofs of the statements.

Lemma 2.1. *Consider the partially observed signal process X_t as described in Eq. (1) and the associated observation Y_t in Eq. (2). Suppose that data are received in time horizon $[0,T)$ such that $T \to \infty$. Then $\hat{\lambda}_n$ is unbiased. Furthermore, define*

$$\tilde{\lambda}_n := \frac{|a|}{n\delta} \int_0^n X_s ds + \frac{|a|}{hn\delta} \sum_{i=1}^n \varepsilon_i. \tag{4}$$

Then the LSE can be written as

$$\hat{\lambda}_n = \frac{n}{n+k_n} \left(\tilde{\lambda}_n + \frac{1}{n} \xi_n \right) + \frac{\ell_n}{n}, \tag{5}$$

where the sequence of (k_n, ℓ_n) converges to (k, ℓ), and the random sequence ξ_n almost surely converges to a finite random variable ξ.

Theorem 2.1. *The LSE $\hat{\lambda}_n$ in Eq. (3) satisfies the following central limit theorem:*

$$\sqrt{n} \left(\hat{\lambda}_n - \lambda \right) \xrightarrow{L} N \left(0, \sigma^2 \right) \qquad \text{as } n \to \infty, \tag{6}$$

where $\sigma^2 = \frac{b^2}{\delta^2} + \lambda + \frac{a^2}{h^2\delta^2}$.

Theorem 2.1 provides the limit distribution of the LSE, $\hat{\lambda}_n$; however, its asymptotic variance depends on the true intensity. To that end, different convergence arguments, precisely a moderate deviation principle for $\hat{\lambda}_n$, need to be established. The moderate deviation principle is related to the large deviations; therefore we present below large and moderate deviation formalisms. The reader may refer to [9] for more details.

Definition 2.1. A function $I : \mathbb{R} \to [0, \infty]$ is called a *rate function* if it is lower semicontinuous, for any $l > 0$, the level sets $I_l = \{x \in \mathbb{R}; I(x) \le l\}$ is a closed set. Further, a rate function is said to be good if every level is compact in $\mathbb{R}$.

Assumption 2.1. *Let $\{\gamma_n\}$ be a sequence of random variables with topological state space Γ. For each $\mu \in \mathbb{R}$ and a sequence $c_n \to \infty$ as $n \to \infty$, the logarithmic moment generating function $\Lambda(\mu)$, defined as the limit*

$$\Lambda(\mu) = \lim_{n\to\infty} \frac{1}{c_n} \log \mathbb{E} \exp(\mu c_n \gamma_n), \tag{7}$$

exists as in $(-\infty, +\infty]$. Further, the origin belongs to the interior $\mathcal{D}_\Lambda^\circ$ of the domain $\mathcal{D}_\Lambda = \{\mu \in \mathbb{R} : \Lambda(\mu) < \infty\}$ of the function $\Lambda(\mu)$.

Remark 2.1. By the Hölder inequality, $\Lambda(\mu)$ is a convex function. Define the *Fenchel–Legendre transform*

$$I(x) = \sup_{\mu\in\mathbb{R}} \{\mu x - \Lambda(\mu)\}, \ x \in \mathbb{R}. \tag{8}$$

Then the function $I(x)$ is a good rate function; see [9, Section 4.5].

Definition 2.2. A convex function $\Lambda : \mathbb{R} \to (-\infty, \infty]$ is essentially smooth if

(1) $\mathcal{D}_\Lambda^\circ = (a, b)$ is nonempty for some $-\infty \le a < b \le \infty$.

(2) $\Lambda(\mu)$ is differentiable in $\mathcal{D}_\Lambda^\circ$.
(3) $\Lambda(\cdot)$ is steep; that is, $\lim_{\mu\to a^+}\Lambda'(\mu) = \lim_{\mu\to b^-}\Lambda'(\mu) = \infty$.

The following theorem is known as the *Gartner–Ellis theorem* (see, e.g., [6,9]) on large deviations, which shows the correspondence between the logarithmic moment function and large deviations.

Theorem 2.2. *Assume Assumption 2.1 holds. For any closed set $F \subset \mathbb{R}$,*

$$\limsup_{n\to\infty} \frac{1}{c_n} \log \mathbb{P}\left(\gamma_n \in F\right) \leq -\inf_{x\in F} I(x).$$

If we further assume that the logarithmic moment function $\Lambda(\mu)$ is essentially smooth, then for any open set $G \subset \mathbb{R}$,

$$\liminf_{n\to\infty} \frac{1}{c_n} \log \mathbb{P}\left(\gamma_n \in G\right) \geq -\inf_{x\in G} I(x).$$

Next, the motivating large deviation principle for the LSE, $\hat{\lambda}_n$, is presented in Theorem 2.3, whose proof is given in the Appendix.

Theorem 2.3. *Consider the signal described by Eq. (1) and the associated observation process in Eq. (2). The sequence of least-squares estimates $\{\hat{\lambda}_n, n \geq 1\}$ for the intensity, given by Eq. (3), satisfies the large deviation principle in $\mathbb{R}$, with rate function $I_\ell(x) = \sup_{\mu\in\mathbb{R}}\{\mu x - \Lambda_\ell(\mu)\}$, where the logarithmic moment generating function $\Lambda_\ell(\mu)$ defined with $c_n = n$ in Eq. (7) is given by*

$$\Lambda_\ell(\mu) = \frac{(h^2b^2 + a^2)\mu^2}{2h^2\delta^2} + \lambda e^{\mu} - \lambda, \quad \mu \in \mathbb{R}. \tag{9}$$

Theorem 2.3 provides the rate function of convergence, I_ℓ; however, it depends on the true value of the unknown parameter (as the asymptotic variance in Theorem 2.1). Therefore Theorem 2.3 is insufficient to conduct statistical inference relying on this asymptotic tail probability. To that end, moderate deviations need to be examined instead.

3. Statistical inference

3.1. Moderate deviations

We are now concerned with the weak convergence of the quantity $n^\alpha(\hat{\lambda}_n - \lambda)$ for $0 < \alpha < \frac{1}{2}$ (i.e., the moderate deviation principle) using the logarithmic moment generating function $\frac{1}{c_n}\log\mathbb{E}\exp\left[c_n\mu n^\alpha(\hat{\lambda}_n - \lambda)\right]$, where $c_n = n^\beta$ for some fixed $\beta > 0$. The result is presented in the following theorem.

Theorem 3.1. *Consider the signal described by Eq. (1) and the associated observation process in Eq. (2). The least-squares estimate, $\{\hat{\lambda}_n\}$, given by Eq. (3), satisfies a moderate deviation principle. Precisely, for any $\alpha \in (0, \frac{1}{2})$, the sequence $\{n^\alpha(\hat{\lambda}_n - \lambda), n \geq 1\}$ satisfies the large deviation principle in $\mathbb{R}$ with speed $n^{1-2\alpha}$ and the rate function*

$$I_m(x) = \frac{x^2}{2\kappa^2 + 2}, \quad \textit{where} \quad \kappa^2 = \frac{h^2b^2 + a^2}{h^2\delta^2}. \tag{10}$$

Proof. Given Theorem 2.2, we need only to consider for some $\beta > 0$, $\alpha \in (0, \frac{1}{2})$

$$\frac{1}{n^\beta} \log \mathbb{E} \exp \left[\mu n^{\beta+\alpha}(\hat{\lambda}_n - \lambda)\right]. \tag{11}$$

According to Eqs. (4) and (11), we have

$$\begin{aligned} &\mathbb{E} \exp \left[\mu n^{\beta+\alpha} \left(\frac{|a|}{n\delta} \int_0^n X_s ds + \frac{|a|}{hn\delta} B_T - \lambda \right)\right] \\ =&\mathbb{E} \exp \left(\delta^{-1} |a| \mu n^{\beta+\alpha-1} \int_0^n X_s ds \right) + \mathbb{E} \exp \left(\mu n^{\beta+\alpha-1} \frac{|a|}{h\delta} B_T \right) - \exp\left(\mu\lambda n^{\alpha+\beta}\right). \end{aligned} \tag{12}$$

Applying Itô's formula [16], we write the solution of the signal process as

$$X_t = e^{at}(x_0 + b \int_0^t e^{-as} dW_s + \delta \int_0^t e^{-as} dN_s), \tag{13}$$

and consequently,

$$\int_0^n X_r dr = \frac{b}{|a|} \int_0^n (1 - e^{an-as}) dW_s + \frac{\delta}{|a|} \int_0^n (1 - e^{an-as}) dN_s + \frac{1}{|a|} x_0 (1 - e^{an}). \tag{14}$$

If we let $S_n := \delta^{-1} \mu n^{\beta+\alpha-1}$, Eq. (14) yields

$$\begin{aligned} \mathbb{E} \exp \left(|a| S_n \int_0^n X_s ds \right) =&\mathbb{E} \exp \left[b S_n \int_0^n (1 - e^{a(n-s)}) dW_s \right] \\ &\times \mathbb{E} \exp \left[\delta S_n \int_0^n (1 - e^{a(n-s)}) dN_s \right] \\ &\times \mathbb{E} \exp \left[S_n x_0 (1 - e^{an}) \right]. \end{aligned} \tag{15}$$

Recall that B_T is a standard Brownian motion at time T, and hence it is easy to show that

$$\mathbb{E} \exp \left(\mu n^{\beta+\alpha-1} \frac{|a|}{h\delta} B_T \right) = \exp \left(\frac{a^2 \mu^2}{2h^2 \delta^2} n^{2\beta+2\alpha-1} \right). \tag{16}$$

Note that $bS_n \int_0^n (1 - e^{a(n-s)}) dW_s$ follows a normal distribution $N(0, \tau_n^2)$ for a pertinent variance τ_n^2, $n \in \mathbb{N}$. Hence

$$\mathbb{E} \exp \left[b S_n \int_0^n (1 - e^{a(n-s)}) dW_s \right] = \exp \left(\frac{1}{2} \tau_n^2 \right), \tag{17}$$

where $\tau_n^2 = \mathbb{E} \left[bS_n \int_0^n (1 - e^{a(n-s)}) dW_s \right]^2 = b^2 S_n^2 \int_0^n (1 - e^{a(n-s)})^2 ds$. By Itô's formula [16, Rule 4.23], for a suitable function g, we have the stochastic differential equation

$$d\left[\exp\left(\int_0^t g(s)dN_s\right)\right] = \exp\left(\int_0^t g(s)dN_s\right)\left(e^{g(t)}-1\right)dN_t. \tag{18}$$

Taking the expectation on the integral form of Eq. (18) we obtain

$$\mathbb{E}\exp\left(\int_0^t g(s)dN_s\right) = 1+\lambda\int_0^t \left(e^{g(r)}-1\right)\mathbb{E}\exp\left(\int_0^r g(s)dN_s\right)dr.$$

Let $G(t) := \mathbb{E}\exp\left(\int_0^t g(s)dN_s\right)$. Solving the ordinary differential equation $G'(t) = \lambda G(t)\left(e^{g(t)}-1\right)$ yields

$$G(t) = G(0)\exp\left(\lambda\int_0^t \tilde{g}(s)ds\right), \tag{19}$$

where $\tilde{g}(s) := e^{g(s)} - 1$. Consequently, substitution of $g(s) := \delta S_n(1-e^{a(n-s)})$ into Eq. (19) gives

$$\begin{aligned}&\mathbb{E}\exp\left[\delta S_n\int_0^n (1-e^{a(n-s)})dN_s\right]\\ &=\exp\left(\lambda\int_0^n\left\{\exp\left[\mu n^{\beta+\alpha-1}(1-e^{a(n-s)})\right]-1\right\}ds\right).\end{aligned} \tag{20}$$

Recall that $a<0$; thus the term $\int_0^n \exp(\mu n^{\beta+\alpha-1}e^{an-as})ds$ decays exponentially. If we choose $\beta+\alpha-1<0$, by Taylor's formula, Eq. (20) becomes

$$\exp\left[\lambda\mu n^{\beta+\alpha}+\frac{1}{2}\mu^2 n^{2\beta+2\alpha-1}+o(n^{2\beta+2\alpha-1})\right]. \tag{21}$$

Then plugging Eqs. (16), (17), and (21) into Eq. (15) and taking into account Eq. (12), we can express the logarithmic moment generating function of Eq. (11) as

$$\begin{aligned}&\frac{\mu^2 b^2 n^{\beta+2\alpha-2}}{4a\delta^2}(2an+3-4e^{an}+e^{2an})+\frac{1}{2}\mu^2 n^{\beta+2\alpha-1}+o(n^{\beta+2\alpha-1})\\ &+\frac{\mu x_0}{\delta}n^{\alpha-1}(1-e^{an})+\frac{a^2\mu^2}{2h^2\delta^2}n^{\beta+2\alpha-1}.\end{aligned} \tag{22}$$

If we let $\beta = 1-2\alpha$, Eq. (22) converges to a constant as $n\to\infty$. Next we claim that $\frac{1}{n}\xi_n$ converges to 0 exponentially fast; that is,

$$\limsup_{n\to\infty}\frac{1}{n^\beta}\log\mathbb{P}\left(\left|\frac{1}{n}\xi_n\right|>\epsilon\right] = -\infty. \tag{23}$$

By Markov's inequality

$$\mathbb{P}\left(\frac{1}{n}|\xi_n|>\epsilon\right) = \mathbb{P}\left[\left|\frac{(1-e^{-a})}{h\delta}\sum_{i=1}^n e^{ai}Z_i\right|>n\epsilon\right]$$

$$\leq e^{-n^2\epsilon^2}\mathbb{E}\exp\left\{\left[\frac{(1-e^{-a})}{h\delta}\sum_{i=1}^{n}e^{ai}Z_i\right]^2\right\},$$

and then the claim follows from Lemma 2.1. Thus, together with Lemma 2.1, Eqs. (22) and (A.3) imply that the logarithmic moment generating function

$$\Lambda_m(\mu) := \lim_{n\to\infty}\frac{1}{n^{1-2\alpha}}\log\mathbb{E}\exp\left[\mu n^{1-\alpha}(\hat{\lambda}_n-\lambda)\right] = \frac{(h^2b^2+a^2)\mu^2}{2h^2\delta^2}+\frac{1}{2}\mu^2$$

for any $\alpha\in(0,\frac{1}{2})$. The rate function is then given by the Fenchel–Legendre transform (8):

$$I_m(x) = \sup_{\mu\in\mathbb{R}}\{\mu x-\Lambda_m(\mu)\} = \frac{x^2}{2\kappa^2+2},$$

with $\kappa^2 = \frac{h^2b^2+a^2}{h^2\delta^2}$. □

3.2. Hypothesis testing and confidence interval

Consider the null and alternative hypotheses

$$H_0:\lambda=\lambda_0 \text{ and } H_1:\lambda\neq\lambda_0,$$

and the test statistic $T_n := \sqrt{n}(\hat{\lambda}_n-\lambda_0)$. For $0<\alpha_0<1$, denote the rejection region for testing the null hypothesis H_0 against H_1 to be $\mathcal{R}=\{|T_n|\geq c(\alpha_0)\}$, where $c(\alpha_0)$ is a positive constant in $\mathbb{R}$ such that $\alpha_0/2 = 1-\Phi(\frac{c(\alpha_0)}{\sigma_0})$, with $\sigma_0^2=\frac{b^2}{\delta^2}+\lambda_0+\frac{a^2}{h^2\delta^2}$, and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution. We consider the *power*, $1-\beta_n$, of the hypothesis testing (i.e., the probability of correctly rejecting a false null hypothesis), where β_n is the probability of type II error given by

$$\beta_n = \mathbb{P}\left(|T_n|<c(\alpha_0)\;\Big|\;\lambda=\lambda_1\neq\lambda_0 \text{ under } H_1\right).$$

Proposition 3.1. *Consider the signal described by Eq. (1) and the associated observation process in Eq. (2). Then the power of the hypothesis testing tends to* 1 *with exponential speed* $\exp(-rn^{1-2\alpha})$ *for any* $\alpha\in(0,1/2)$ *and any* $r>0$*. In other words,*

$$\lim_{n\to\infty}\frac{1}{n^{1-2\alpha}}\log\beta_n=-\infty.$$

Proof. Note that by the triangle inequality, $|\hat{\lambda}_n-\lambda_0|\geq|\lambda_1-\lambda_0|-|\hat{\lambda}_n-\lambda_1|$. Then we have

$$\begin{aligned}\beta_n &= \mathbb{P}\left(|T_n|<c(\alpha_0)\;\Big|\;\lambda=\lambda_1\right)\\ &\leq \mathbb{P}\left(n^{\alpha}|\hat{\lambda}_n-\lambda_1|\geq n^{\alpha}|\lambda_1-\lambda_0|-\frac{c(\alpha_0)}{n^{\frac{1}{2}-\alpha}}\;\Big|\;\lambda=\lambda_1\right).\end{aligned}$$

Then Theorem 3.1 implies that $\lim_{n\to\infty}\frac{1}{n^{1-2\alpha}}\log\beta_n=-\infty$. □

Next, a direct application of the asymptotic normality derived in Theorem 2.1 yields an approximate $(1-\alpha_0)\cdot 100\%$ confidence interval for λ as $\left(\hat{\lambda}_n-\frac{\sigma}{\sqrt{n}}Z_{\alpha_0/2},\hat{\lambda}_n+\frac{\sigma}{\sqrt{n}}Z_{\alpha_0/2}\right)$, where $Z_{\alpha_0/2}$ is a critical value of the standard random variable Z such that $\mathbb{P}(Z>Z_{\alpha_0/2})=\alpha_0/2$, and σ^2 is defined in Theorem 2.1. Consider the inequality $-Z_{\alpha_0/2}\leq\frac{\sqrt{n}(\hat{\lambda}_n-\lambda)}{\sigma}\leq Z_{\alpha_0/2}$, which is equivalent to

$$\lambda^2-\left(2\hat{\lambda}_n+\frac{Z^2_{\alpha_0/2}}{n}\right)\lambda+\hat{\lambda}_n^2-\left(\frac{b^2}{n\delta^2}+\frac{a^2}{nh^2\delta^2}\right)Z^2_{\alpha_0/2}\le 0. \tag{24}$$

Thus, solving inequality (24) gives the $(1-\alpha_0)\cdot 100\%$ confidence interval by the weak convergence

$$\left\{\hat{\lambda}_n+\frac{Z^2_{\alpha_0/2}}{2n}\pm\frac{1}{2}\left[\left(2\hat{\lambda}_n+\frac{Z^2_{\alpha_0/2}}{n}\right)^2-4\left(\frac{b^2}{n\delta^2}+\frac{a^2}{nh^2\delta^2}\right)Z^2_{\alpha_0/2}\right]^{1/2}\right\}. \tag{25}$$

We can see from Eq. (25) that the confidence interval obtained by the central limit theorem has large errors when the true value and accordingly the unbiased point estimate $\hat{\lambda}_n$ is large. We now apply Theorem 3.1 to construct the confidence interval for the intensity λ.

Proposition 3.2. *Consider the signal described by Eq. (1) and the associated observation process in Eq. (2). The moderate deviation principle–based* $(1-\alpha_0)\cdot 100\%$ *confidence interval for the intensity* λ*,* $\alpha_0\in(0,1)$*, is given by*

$$\left(\hat{\lambda}_n+\sqrt{\frac{2\kappa^2+2}{n}}\left(\log\alpha_0\right)^{1/2},\hat{\lambda}_n-\sqrt{\frac{2\kappa^2+2}{n}}\left(\log\alpha_0\right)^{1/2}\right), \tag{26}$$

where κ^2 *is given in Eq. (10).*

Proof. By Theorem 3.1, for a fixed $u>0$ we have

$$\mathbb{P}\left(\left|n^\alpha(\hat{\lambda}_n-\lambda)\right|>u\right)\approx\exp\left(-n^{1-2\alpha}\inf_{|x|>u}I_m(x)\right)=\exp\left(-\frac{n^{1-2\alpha}u^2}{2\kappa^2+2}\right),$$

where $\kappa^2=\frac{h^2b^2+a^2}{h^2\delta^2}$. For a given confidence level $1-\alpha_0$, set

$$\tau_{\alpha_0}=\left(\frac{2\kappa^2+2}{n^{1-2\alpha}}\log\frac{1}{\alpha_0}\right)^{1/2}=n^\alpha\left(\frac{2\kappa^2+2}{n}\log\frac{1}{\alpha_0}\right)^{1/2}.$$

Then the $(1-\alpha_0)\cdot 100\%$ confidence interval by the moderate deviation principle for the intensity λ is approximately $\left(\hat{\lambda}_n-n^{-\alpha}\tau_{\alpha_0},\hat{\lambda}_n+n^{-\alpha}\tau_{\alpha_0}\right)$, which follows Eq. (26). □

4. Conclusion

We focused on a partially observed jump-diffusion signal. The jumps were considered from a Poisson distribution whose intensity is statistically inferred. To that end, asymptotic behavior of the consistent LSE for the Poisson intensity was investigated. The large and moderate deviation principles were proved on the basis of the Gartner–Ellis theorem. Together with the central limit theorem, these results compose the limit theorems of the LSE with different dominated scale c_n, where $c_n=n$ for large deviations and $c_n=n^\alpha$, $\alpha\in(0,\frac{1}{2})$, for moderate deviations. In turn, the moderate deviation principle was applied to find the power of the hypothesis testing and to construct an effective confidence interval. An intensity estimate for high-dimensional partially observed jump diffusions and the associated convergence rates may follow by generalization of the aforementioned large and moderate deviation results.

Acknowledgments

The authors thank an anonymous reviewer for comments. VM's research was partially supported by a Simons Foundation grant (279870) and the National Science Foundation (DMS-1821241). JX's research was partially supported by the Southern University of Science and Technology Start-up Fund (Y01286120) and the National Natural Science Foundation of China (grants 61873325 and 11831010).

Appendix A. Proof of Theorem 2.3

Using Eq. (14), we have

$$\begin{aligned}\mathbb{E}\exp\left(\frac{\mu|a|}{\delta}\int_0^n X_s ds\right) =& \mathbb{E}\exp\left[\frac{\mu b}{\delta}\int_0^n (1-e^{a(n-s)})dW_s\right] \\ &\times \mathbb{E}\exp\left[\mu\int_0^n (1-e^{a(n-s)})dN_s\right] \\ &\times \mathbb{E}\exp\left[\frac{\mu x_0}{\delta}(1-e^{an})\right]. \end{aligned} \tag{A.1}$$

Using Eqs. (A.1), (17), and (19) with $g(s) := \mu(1-e^{a(n-s)})$, we have $\frac{1}{n}\log\mathbb{E}\exp\left(\frac{\mu|a|}{\delta}\int_0^n X_s ds\right)$ equals

$$\frac{1}{n}\left[\frac{1}{2}\sigma_n^2 + \lambda\int_0^n \left(e^{\mu-\mu e^{a(n-s)}}-1\right)ds + \frac{\mu x_0}{\delta}(1-e^{an})\right] \to \frac{\mu^2 b^2}{2\delta^2} + \lambda(e^{\mu}-1). \tag{A.2}$$

Note that $\frac{\mu|a|}{h\delta}\sum_{i=1}^n \varepsilon_i \sim N(0, \frac{\mu^2 a^2 T}{h^2\delta^2})$. Therefore $\mathbb{E}\exp\left(\frac{\mu|a|}{h\delta}\sum_{i=1}^n \varepsilon_i\right) = \exp\left(\frac{\mu^2 a^2 T}{2h^2\delta^2}\right)$, and furthermore $\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\exp\left(\frac{\mu|a|}{h\delta}\sum_{i=1}^n \varepsilon_i\right) = \frac{\mu^2 a^2}{2h^2\delta^2}$. Similarly to the proof of Eq. (23), one can show that $\frac{1}{n}\xi_n$ converges to 0 exponentially fast; that is,

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\left|\frac{1}{n}\xi_n\right| > \epsilon\right) = -\infty. \tag{A.3}$$

Therefore the logarithmic moment generating function (9) is obtained by Eqs. (A.3), (A.2), and (5), and the rate function follows from Theorem 2.2. □

References

[1] M.A. Arcones, Large deviations for M-estimators, Ann. Inst. Statist. Math. 58 (1) (2006) 21–52.
[2] R.R. Bahadur, On the asymptotic efficiency of tests and estimates, Sankhyā (1960) 229–252.
[3] F. Camerlenghi, E. Villa, Large and moderate deviations for kernel-type estimators of the mean density of Boolean models, Electron. J. Stat. 12 (1) (2018) 427–460.
[4] C. Ceci, K. Colaneri, Nonlinear filtering for jump diffusion observations, Adv. in Appl. Probab. 44 (03) (2012) 678–701.
[5] C.D. Charalambous, A. Logothetis, Maximum likelihood parameter estimation from incomplete data via the sensitivity equations: the continuous-time case, IEEE Trans. Automat. Control 45 (5) (2000) 928–934.
[6] X. Chen, Random Walk Intersections: Large Deviations and Related Topics, American Mathematical Society, 2010.
[7] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, Ann. Math. Stat. (1952) 493–507.
[8] A. Dembo, O. Zeitouni, Parameter estimation of partially observed continuous time stochastic processes via the EM algorithm, Stochastic Process. Appl. 23 (1) (1986) 91–113.

[9] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, 2nd edition, Springer-Verlag, New York, 1998.
[10] S.M. Djouadi, V. Maroulas, X. Pan, J. Xiong, Consistency and asymptotics of a Poisson intensity least-squares estimator for partially observed jump–diffusion processes, Statist. Probab. Lett. 123 (2017) 8–16.
[11] E. Evangelou, V. Maroulas, Sequential Empirical Bayes method for filtering dynamic spatiotemporal processes, Spatial Statist. 21 (Part A) (2017) 114–129.
[12] F. Gao, J. Xiong, X. Zhao, et al., Moderate deviations and nonparametric inference for monotone functions, Ann. Statist. 46 (3) (2018) 1225–1254.
[13] F. Gao, X. Zhao, Delta method in large deviations and moderate deviations for estimators, Ann. Statist. (2011) 1211–1240.
[14] F.B. Gonçalves, G.O. Roberts, Exact Simulation and Bayesian Inference for Jump-Diffusion Processes, Technical report, University of Warwick.
[15] B. Grigelionis, R. Mikulevicius, Nonlinear filtering equations for stochastic processes with jumps, in: The Oxford Handbook of Nonlinear Filtering, 2011, pp. 95–128.
[16] F.B. Hanson, Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis, and Computation, Society for Industrial and Applied Mathematics, 2007.
[17] T. Inglot, T. Ledwina, On probabilities of excessive deviations for Kolmogorov–Smirnov, Cramer–von Mises and chi-square statistics, Ann. Statist. (1990) 1491–1495.
[18] 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.
[19] W.C. Kallenberg, On moderate deviation theory in estimation, Ann. Statist. (1983) 498–504.
[20] 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.
[21] 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.
[22] D.R. Poklukar, Nonlinear filtering for jump-diffusions, J. Comput. Appl. Math. 197 (2) (2006) 558–567.
[23] H. Qiao, J. Duan, Nonlinear filtering of stochastic dynamical systems with Lévy noises, Adv. in Appl. Probab. 47 (03) (2015) 902–918.
[24] C.A. Ramezani, Y. Zeng, Maximum likelihood estimation of the double exponential jump-diffusion process, Ann. Finance 3 (4) (2007) 487–507.
[25] A.P. Vacarescu, Filtering and Parameter Estimation for Partially Observed Generalized Hawkes Processes, Ph.D. thesis, Stanford University, 2011.
[26] X. Yang, Information-Based Commodity Pricing and the Theory of Signal Processing with Lévy Information, Ph.D. thesis, Imperial College London, 2013.