Filtering with Marked Point Process Observations via Poisson Chaos Expansion

Wei Sun, Yong Zeng, and Shu Zhang*

Abstract

We study a general filtering problem with marked point process observations. The motivation comes from modeling financial ultra-high frequency data. First, we rigorously derive the unnormalized filtering equation with marked point process observations under mild assumptions, especially relaxing the bounded condition of stochastic intensity. Then, we derive the Poisson chaos expansion for the unnormalized filter. Based on the chaos expansion, we establish the uniqueness of solutions of the unnormalized filtering equation. Moreover, we derive the Poisson chaos expansion for the unnormalized filter density under additional conditions. To explore the computational advantage, we further construct a new consistent recursive numerical scheme based on the truncation of the chaos density expansion for a simple case. The new algorithm divides the computations into those containing solely system coefficients and those including the observations, and assign the former off-line.

MSC: 60G35, 93E11, 60G55

Keywords: Poisson chaos expansion; Filtering; marked point processes; random measures; ultra-high frequency data.

---

*Wei Sun and Shu Zhang are with Department of Mathematics and Statistics, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, Quebec, Canada, H3G 1M8. For Wei Sun, Tel: (514) 848 2424 ext. 5218; Fax: (514) 848 2831; Email: wsun@mathstat.concordia.ca; and website: http://alcor.concordia.ca/∼wsun/. For Shu Zhang, Email: zhangshuisme@hotmail.com. *Correspondence Address to* Yong Zeng, Department of Mathematics and Statistics, University of Missouri at Kansas City, Kansas City, MO 64110, USA. Tel: (816) 235 5850; Fax: (816) 235 5517; Email: zengy@umkc.edu; and website: http://z.web.umkc.edu/zengy/.

1. Introduction

The filtering problem with counting process observations was originated from optical signal processing and was first introduced by Synder in [45]. Then, such a filtering problem was subsequently studied with applications in communication and control in [14], [3], [35, §19.3] (the first edition in 1978), [5, §6.3], [24], [28] and [16].

Recently, there has been considerable interest in filtering with marked point process (MPP) observations, which is a generalization of filtering with counting process observations. An insurance claim process can be naturally modeled as an MPP, which can be expressed as a random measure. Motivated by inferring the underlining state of economy, Elliott et al. in [17] studied the problem of filtering a Markov modulated random measure. Our motivation to consider a more general filtering problem with MPP observations comes from modeling financial ultra-high frequency data, namely, time-stamped transactions data of asset price. Engle in [18] called such data as ultra high frequency (UHF) data, because they achieve the maximum level of disaggregation. UHF data hold the most accurate information of the price evolution, and the direct modeling and study of these data is important for comprehending and studying market microstructure theory.

UHF data have two stylized characteristics distinguishing from the continuous-time models in asset pricing, or the equally-spaced daily or weekly closing price behavior in the econometric literature. First, the time-stamped UHF data occur at varying random time intervals. Two early influential works, [19] and [18], attempted to model these time-stamped data from the viewpoint of irregularly-spaced time series. Many papers follow this viewpoint. See a survey paper [41] and the references therein. The second stylized characteristic is that UHF data contain microstructure (or trading) noise due to price discreteness, price clustering, ask and bid bounce and other market microstructure issues. In contrast with information which has a long-term impact on price, noise only has a short-term impact on price. However, it is well-known that noise (or friction) plays a fundamental role in asset pricing (see for example [44], a recent presidential addresses to AFA), and noise should be incorporated in any suitable model of UHF data.

From the standpoint of stochastic process, a general nonlinear *filtering model with counting process observations* for UHF data of asset price was proposed in [46]. In this paper, we first propose a general filtering model with MPP observations in the same spirit as the one for UHF data in [46]. There is an unobservable signal process, which can be interpreted as the intrinsic value process for an asset in the setup of UHF data. The intrinsic value process corresponds to the continuous-time price process in the option pricing literature in a frictionless market or in the empirical econometric literature for equally-spaced daily closing prices. Observations (or prices) are observed only at random trading times which are modeled by a conditional Poisson process. Trading (or market microstructure) noise is explicitly modeled and prices are distorted observations of the intrinsic value process at the trading times. The proposed filtering model is capable of matching the two stylized features of UHF data as well as many other features in equally-spaced time series data, because of the general assumption on the intrinsic value process. Technically, the new filtering model extends the one studied in [46], mainly by considering the extended generator of a Markov process, the more general MPP observations, and by relaxing the bounded condition on the stochastic intensity of the observations.

For the extended filtering framework with MPP observations, we derive the unnormalized filtering equation, which corresponds to the Duncan-Mortensen-Zakai (DMZ) equation in classical filtering of Wiener process and the Synder equation with counting process observations. We focus on the unnormalized filtering equation because it can compute the normalized filters and can avoid the possibility of reciprocating zero in the normalized filters. The derivation contains two steps. The first step is to prove the equivalence of the weak and mild solutions and the second one is to derive the mild solution. Note that our derivation of the filtering equation is different from those given in literature and has its own interest.

Wiener chaos expansion for the DMZ equation in the classical nonlinear filtering has been obtained (cf. e.g. [36]). Correspondingly, we obtain the Poisson chaos expansion for the unnormalized filtering equation via a new filtering technique recently developed in [26] for the classical nonlinear filtering case. As Wiener chaos expansion is an approach to prove the uniqueness of solutions of classical filtering equations (cf. e.g. [29]), we further prove the uniqueness of solutions of the unnormalized filtering equation via Poisson chaos expansion. To the best of our knowledge, this is the first time that this kind of approach is applied in the framework of filtering with marked point process observations. The main contribution of this paper is that we investigate the general measure-valued solutions of filtering equations and, more importantly, we prove its existence and uniqueness without assuming that the stochastic intensity of the observations is bounded. The relaxation of the bounded condition on the stochastic intensity makes the chaos expansion method more attractive from the viewpoint of applications and we provide concrete, interesting practical examples with unbounded intensity functions.

Numerical solution of the unnormalized filter is important because it provides the numerical solution to the marginal likelihood and the posterior distribution, both of which are important in statistical analysis. Based on the Wiener chaos expansion and using the orthonormal Hermite polynomial system, Lototsky, Mikulevicius and Rozovskii developed in [36, 38] a consistent recursive algorithm to numerically solve the DMZ equation. A major advantage of the algorithm is that the computations can be separated into two parts: off-line and online. A large amount of computation related solely to the system coefficients can be carried out off-line, namely, precomputed and stored. The online computation involving the observations can be computed as needed such as the final moment, and can be easily parallelized. Thus, the algorithm greatly increases the speed for online computation. Correspondingly, we first derive the Poisson chaos expansion for the unnormalized filter density of a multi-dimensional diffusion signal. Then, using the orthonormal Hermite polynomial system, we develop a new consistent recursive algorithm for computing the unnormalized filter density for a simple case. The new recursive algorithm shares the same desirable properties of the corresponding one based on the Wiener chaos expansion.

Related filtering models for UHF data, not incorporating market microstructure noise and focusing on estimating price volatility, are [21], [13], [12] and [10]. The innovation approach was used in [21], [12] and [10] to characterize the unique solution to the Kushner-Stratonovich equation.

The rest of the paper is organized as follows. The next section presents the filtering model with MPP observations for UHF data. Section 3 derives the two equivalent forms of the unnormalized filtering equation for the extended filtering model. Using the newly-developed filtering technique, Section 4 derives the Poisson chaos expansion and applies it to prove the uniqueness of solutions of the unnormalized filtering equation. Section 5 derives the Poisson chaos expansion for the unnormalized filter density and develops a consistent recursive algorithm further using Hermite polynomials. Section 6 concludes.

2 The Filtering Model

In this section, we present the signal and the MPP observations of the filtering model as well as an equivalent representation. Let X be the signal and Y be the MPP observations. We assume that (X, Y) is on a complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ with the usual hypothesis.

2.1 The Signal

In the filtering model of [46], the signal is the latent intrinsic value process of an asset, X, taking values in $\mathbb{R}$. Here, we consider the state space of X as E, a complete separable metric space with Borel σ-algebra $\mathcal{B}(E)$. In this paper, we only consider time-homogeneous X. However, the results obtained in this paper can be directly applied to the parametric and non-homogeneous model of [46], simply by enlarging the dimension of X and noticing the general assumption on E. Below we first invoke a mild assumption on X.

Assumption 2.1 $(X(t))_{t\geq 0}$ is a càdlàg time-homogeneous Markov process taking values in E and with the *extended (infinitesimal) generator* defined in (2.1) below.

Let $B_b(E)$ be the family of all bounded Borel measurable functions on E and $P(t,x,\Gamma)$ $(t \geq 0, x \in E, \Gamma \in \mathcal{B}(E))$ be the transition function for X. Then, we have the transition semigroup $(T_t)_{t\geq 0}$ defined by $T_t f(x) = \int_E f(y)P(t,x,dy)$ for $f \in B_b(E)$. Set $\mathcal{C} := \{f \in B_b(E) : bp - \lim_{t\downarrow 0} T_t f(x) = f(x), \forall x \in E\}$, where bp-lim stands for bounded pointwise limit. Then $\mathcal{C} \supset C_b(E)$, the family of all bounded continuous functions on E. Define

$$\mathcal{D} := \left\{ f \in \mathcal{C} : \exists\ \mathbf{A}f \in \mathcal{C} \ \text{ such that } \ T_t f(x) = f(x) + \int_0^t (T_s \mathbf{A} f)(x)ds, \forall x \in E \right\}, \tag{2.1}$$

where $\mathbf{A}$ is called the *extended (infinitesimal) generator* of X and $\mathcal{D}$ is its domain (cf. [27, pp. 1102]). Note that $\mathcal{D}$ is bounded pointwise dense in $B_b(E)$ and hence measure determining.

Let $\mathcal{N}$ be the collection of P-null sets and define $\mathcal{F}_t^X := \sigma(\mathcal{N}, X(s), 0 \leq s \leq t)$, $t \geq 0$. Then by [27, Lemma A.1], $M_f(t) = f(X(t)) - \int_0^t \mathbf{A}f(X(s))ds$ is an $\mathcal{F}_t^X$-martingale under P for $f \in \mathcal{D}$. This shows that our Assumption 2.1 implies the assumption of [46] on X.

The generator and martingale problem approach (see, for example, [20, Chapter 4]) offers a useful tool for the characterization of Markov processes. The above assumption is general, subsuming most of popular stochastic processes such as diffusion, stochastic volatility, jump and Lévy processes employed in asset pricing theory. When $\mathbf{A}$ is the generator for a Markov chain, our filtering model becomes the one considered in [17].

2.2 The Observations

The observation Y has two representations. One representation is a random measure Y, which is a direct generalization of the counting process observations in [46]. The other representation is $\{Y(T_i)\}_{i\geq 1}$, or equivalently, $\{(T_i, Y_i)\}_{i\geq 1}$ where (T_i, Y_i) is a marked point with T_i as the event time and $Y_i = Y(T_i)$ as the mark describing the event occurred at time T_i. We first describe the random measure one.

2.2.1 Random Measure Observations

Let U be the mark space of observations with the triple $(U, \mathcal{U}, \mu)$, which is a measure space where μ is a finite measure. The random counting measure Y is specified by $\{Y(A,t) : A \in \mathcal{U}, t \in [0,\infty)\}$, where $Y(A,t)$ is a counting process recording the cumulative number of events that have occurred up to time t with the marks in set A. Before we present the random measure observations, we make two more assumptions.

The first assumption specifies a Poisson random measure on a three dimension space, $U \times [0,\infty) \times [0,\infty)$, where the first $[0,\infty)$ is for time and the second one for intensity. Let $\mathcal{B}[0,\infty)$ be the Borel σ-algebra of $[0,\infty)$, and m be the Lebesgue measure on $[0,\infty)$.

Assumption 2.2 *ξ is a Poisson random measure on $\mathcal{U}\times\mathcal{B}[0,\infty)\times\mathcal{B}[0,\infty)$ with the mean measure $\mu \times m \times m$ under P.*

We give two simple examples to illustrate two important properties of the Poisson random measure. If $A \in \mathcal{U}$, then $\xi(A\times[0,t]\times[0,a])$ has a Poisson distribution with mean $\mu(A)\times m([0,t])\times m([0,a]) = \mu(A)ta$. If $C_i \in \mathcal{U} \times \mathcal{B}[0,\infty) \times \mathcal{B}[0,\infty)$ for $i = 1,2$ and $C_1 \cap C_2 = \phi$, then $\xi(C_1)$ and $\xi(C_2)$ are independent Poisson random variables.

The second assumption specifies the independence of the signal and the Poisson random measure.

Assumption 2.3 *X and ξ are independent under P.*

A random measure can be characterized by its stochastic intensity kernels (See Theorems 8.3.3 and 8.2.2 in [33]). Let $\lambda(u, X(t))$ be the stochastic intensity kernel of Y under P at time t for $u \in U$. Using the projection method of ξ first introduced in [30] and further developed in [22], we can express the random measure observations, $Y(A,t)$, in the following integral form: for $A \in \mathcal{U}$,

$$Y(A,t) = Y(A \times [0,t]) = \int_{A\times[0,t]\times[0,+\infty)} \mathbf{I}_{[0,\lambda(u,X(s))]}(v)\xi(du \times ds \times dv).$$

The random measure Y can be characterized by the fact that

$$Y(A,t) - \int_{A\times[0,t]\times[0,+\infty)} \mathbf{I}_{[0,\lambda(u,X(s))]}(v)\mu(du)dsdv = Y(A,t) - \int_0^t \int_A \lambda(u, X(s))\mu(du)ds$$

is an $\mathcal{F}_t$-martingale. The counting process observations in [46] is a special case with $U = \{1, 2, \ldots, n\}$ and $A = \{i\}$. Then, $Y(A, t) = Y(\{i\}, t) = Y_i(t) = N_i(\int_0^t \lambda_i(X(s))ds)$ and $Y_i(t) - \int_0^t \lambda_i(X(s))ds$ is a martingale.

The next assumption on the stochastic intensity kernel ensures the existence of a desirable reference measure.

Assumption 2.4 $\lambda : U \times E \to (0, \infty)$ *is* $\mathcal{U} \times \mathcal{B}(E)$*-measurable and satisfies*

$$\int_{[0,t]\times U} E[\lambda(u, X(s))]\mu(du)ds < \infty, \quad \forall\ t > 0. \tag{2.2}$$

Note that Assumption 2.4 relaxes the bounded assumption in [46] and implies that

$$\int_0^t \int_U \lambda(u, X(s))\mu(du)ds < \infty, \quad P - a.s., \quad \forall\ t > 0. \tag{2.3}$$

2.2.2 Marked Point Process Observations

Our main results only require the above four assumptions. In order to obtain the equivalence (in the sense of having the same probability law) of the above random counting measure representation with a heuristic representation from UHF data, we make the following additional assumption similar to the one made in [46]. Let $a(X(t)) = \int_U \lambda(u, X(t))\mu(du)$ be the total stochastic intensity at time t.

Assumption 2.5 *Under P, the stochastic intensity kernel of Y for (u, t) has the following form:*

$$\lambda(u, X(t)) = a(X(t))p(u|X(t)),$$

where $p(u|X(t))$ is the transition probability from $X(t)$ to u.

The structure of $\lambda(u, X(t))$ suggests that $a(X(t))$ stipulates when an event might occur while $p(u|X(t))$ stipulates where the mark might occur.

With Assumption 2.5, we can present the observations as an MPP, $\{(T_i, Y_i)\}_{i\geq 1}$, constructed directly from the signal X. Below, we describe the representation in the context of modeling UHF data. First, we specify the intrinsic value process X as in Assumption 2.1. For UHF data, prices are observed at random trading times and contain trading noises. So we need two more steps. First, we stipulates the trading (or sampling) times $T_1, T_2, \ldots, T_i, \ldots$, as a doubly stochastic Poisson process with stochastic intensity $a(X(t))$. Then, $Y_i = Y(T_i)$, the noisy trading price at time T_i, is given by: $Y(T_i) = F(X(T_i))$, where $y = F(x)$ is a random transformation specified by the transition probability $p(y|x)$. It is straightforward to show that $\{(T_i, Y_i)\}_{i\geq 1}$ and the Poisson random measure Y have the same stochastic intensity kernel with Assumptions 2.1 - 2.5 (see Proposition 2.1 in [47]). Thus, the two representations have the same probability law and are equivalent.

The above formulation has a notable financial implication: Price is affected by information and noise. Information has an effect on the intrinsic value of an asset, X, and has a long-run impact on the price. However, noise, stipulated by the random transformation $F(x)$, has only a short-run impact on price, because $F(x)$ does not have an effect on the intrinsic value.

Note that the intrinsic value process is assumed to be a general Markov process. Thus, the proposed filtering model connects to another econometric literature on operator approach for continuous-time Markov processes surveyed in [1]. Especially, the new filtering model is a natural extension of the setup of estimating Markov process sampled at random-time intervals in [15] where no market microstructure noise is considered.

Moreover, the filtering model has the structure similar to two classes of models. One class is the time series structural VAR (Vector AutoRegressive) models developed in many early market microstructure papers (see the recent book [25]). The other class of models is the recent two-scale frameworks, incorporating market microstructure noises, in the exploding literature of realized volatility estimators. The papers on this topic include [48] and [2].

3. The Unnormalized Filtering Equation

In this section, we employ the reference measure approach. To relax the bounded condition of intensity, we begin with identifying the reference measure under condition (2.3). Then, we prove the equivalence of the mild solution and the usual (weak) solution of the unnormalized filtering equation. Finally, we obtain the main result of this section, Theorem 3.4, establishing the *existence* of solutions of the unnormalized filtering equation.

3.1 The Reference Measure

Note that X and ξ are independent under P. However, this does not imply that X and Y are independent under P. In the following, we will show that there exists a probability measure Q on $(\Omega, \mathcal{F})$ such that X and Y become independent and Y becomes a unit Poison random measure with the mean measure $\mu \times m$. Namely, under the reference measure Q, for $A \in \mathcal{U}$,

$$Y(A,t) = \int_{A\times[0,t]\times[0,+\infty)} \mathbf{I}_{[0,1]}(v)\xi(du \times ds \times dv),$$

and

$$Y(A,t) - \int_{A\times[0,t]\times[0,+\infty)} \mathbf{I}_{[0,1]}(v)\mu(du)dsdv = Y(A,t) - \mu(A)t$$

is a martingale.

Lemma 3.1 Suppose that (X,Y) satisfies Assumptions 2.1 - 2.3, $\lambda : U \times E \to (0,\infty)$ is $\mathcal{U}\times\mathcal{B}(E)$-measurable and satisfies (2.3). Let

$$\kappa(u,t) = \frac{1}{\lambda(u, X(t))}$$

and define the process $(\hat{L}(t))_{t\geq 0}$ by

$$\hat{L}(t) := \exp\left\{\int_0^t \int_U \log\kappa(u,s-)Y(du,ds) + \int_0^t \int_U (1-\kappa(u,s))\lambda(u,X(s))\mu(du)ds\right\}. \tag{3.1}$$

Then $(\hat{L}(t))_{t\geq 0}$ is a $(P, \mathcal{F}_t)$-martingale.

Proof. Observe that

$$\hat{L}(t) = 1 + \int_0^t \int_U \hat{L}(s-)(\kappa(u,s-)-1)\left(Y(du,ds) - \lambda(u,X(s))\mu(du)ds\right). \tag{3.2}$$

By (2.3) one finds that $(\hat{L}(t))_{t\geq 0}$ is a $(P, \mathcal{F}_t)$-local martingale. Since $(\hat{L}(t))_{t\geq 0}$ is also nonnegative, $(\hat{L}(t))_{t\geq 0}$ is a $(P, \mathcal{F}_t)$-supermartingale. To show that $(\hat{L}(t))_{t\geq 0}$ is a $(P, \mathcal{F}_t)$-martingale, it suffices to show that $E[\hat{L}(t)] = 1$ for $t \geq 0$. Let P^X and P^ξ be the marginal probabilities w.r.t. (with respect to) X and ξ, respectively. Note that $\hat{L}$ is pathwisely defined. We need to show that $E^\xi[\hat{L}(t)] = 1$, P^X-a.s., for $t \geq 0$. (3.2) implies that $E^\xi[\hat{L}(t)] \leq 1$, P^X-a.s. By (2.3) and the assumption that μ is a finite measure on U, we get

$$\int_0^t \int_U E^\xi[\hat{L}(s)] \cdot |(\kappa(u,s)-1)\lambda(u,X(s))|\mu(du)ds < \infty, \quad P^X - a.s.$$

Then $\hat{L}$ is a $(P^\xi, \mathcal{F}_t)$-martingale, P^X-a.s. and hence $E^\xi[\hat{L}(t)] = 1$, P^X-a.s., for $t \geq 0$. Therefore, $E[\hat{L}(t)] = 1$ for $t \geq 0$. The proof is complete. $\square$

Now we can define $\frac{dQ}{dP}|_{\mathcal{F}_t} = \hat{L}(t)$ for $t \geq 0$. Then, Q can be extended to be a probability measure on $(\Omega, \mathcal{F})$. Under Q, Y is independent of X and Y is a unit Poisson random measure with the mean measure $\mu \times m$ (cf. [5, VIII.T10, pp. 241]). Define

$$L(t) = 1/\hat{L}(t) \tag{3.3}$$

and denote by E^Q the expectation w.r.t. Q. Then $\frac{dP}{dQ}|_{\mathcal{F}_t} = L(t)$ and $E^Q[L(t)] = 1$ for $t \geq 0$. For $f \in B_b(E)$, we define

$$\pi_t(f) = E[f(X(t))|\mathcal{F}_t^Y].$$

By Bayes' theorem (cf. [5, VI.L5, pp. 171]), we get

$$\pi_t(f) = \frac{E^Q[f(X(t))L(t)|\mathcal{F}_t^Y]}{E^Q[L(t)|\mathcal{F}_t^Y]} := \frac{\sigma_t(f)}{\sigma_t(1)}. \tag{3.4}$$

Following the literature, we call $(\pi_t)_{t\geq 0}$ the *normalized filter* and $(\sigma_t)_{t\geq 0}$ the *unnormalized filter*.

3.2 The Derivation

To obtain the unnormalized filtering equation for $(\sigma_t)_{t\geq 0}$ (see Theorem 3.4 below) without assuming the bounded condition of intensity, we first prove the equivalence of the weak and mild solutions (see Lemma 3.2 below) under such setup by applying a truncation technique via stopping times. Fix a constant $T > 0$ and denote by $\mathcal{M}_b(E)$ the family of all finite signed measures on $(E, \mathcal{B}(E))$.

Lemma 3.2 Suppose that (X, Y) satisfies Assumptions 2.1 - 2.3, $\lambda : U \times E \to (0, \infty)$ is $\mathcal{U} \times \mathcal{B}(E)$-measurable and satisfies (2.3). Let $(v_t)_{0\leq t\leq T}$ be an $\mathcal{M}_b(E)$-valued càdlàg process satisfying

$$Q\left[\int_0^T \int_U |v_t|(|\lambda(u) - 1|)\mu(du)dt < \infty\right] = 1.$$

Hereafter $\lambda(u) := \lambda(u, \cdot)$, which is a Borel measurable function on E. Then

$$\begin{aligned} v_t(f) &= v_0(f) + \int_0^t v_s(\mathbf{A}f)ds + \int_0^t \int_U v_{s-}((\lambda(u) - 1)f)(Y(du, ds) - \mu(du)ds), \\ &\qquad Q-a.s., \quad \forall f \in \mathcal{D} \end{aligned} \tag{3.5}$$

if and only if

$$\begin{aligned} v_t(f) &= v_0(T_t f) + \int_0^t \int_U v_{s-}((\lambda(u) - 1)T_{t-s}f)(Y(du, ds) - \mu(du)ds), \quad Q-a.s., \\ &\qquad \forall f \in B_b(E). \end{aligned} \tag{3.6}$$

Remark 3.3 In the literature, the solution of equation (3.5) is called weak solution while the solution of equation (3.6) is called mild solution. Lemma 3.2 established the equivalence of these two types of solutions.

Proof. We will prove that (3.5) and (3.6) are equivalent. For $f \in \mathcal{D}$, we define

$$\begin{aligned}
v(t,f) &= v_0(T_t f) + \int_0^t \int_U v_{s-}((\lambda(u)-1)T_{t-s}f)(Y(du,ds)-\mu(du)ds) \\
&\quad - \left\{ v_0(f) + \int_0^t v_s(\mathbf{A}f)ds + \int_0^t \int_U v_{s-}((\lambda(u)-1)f)(Y(du,ds)-\mu(du)ds) \right\} \\
&= [v_0(T_t f) - v_0(f)] + \\
&\quad \left\{ \int_0^t \int_U v_{s-}((\lambda(u)-1)T_{t-s}f)(Y(du,ds)-\mu(du)ds) \right. \\
&\qquad \left. - \int_0^t \int_U v_{s-}((\lambda(u)-1)f)(Y(du,ds)-\mu(du)ds) \right\} \\
&\quad - \int_0^t v_s(\mathbf{A}f)ds. \qquad (3.7)
\end{aligned}$$

Note that $T_t f - f = \int_0^t T_s \mathbf{A} f ds$. We obtain by Fubini's Theorem that

$$v(t,f) = \int_0^t v_0(T_s\mathbf{A}f)ds - \int_0^t v_s(\mathbf{A}f)ds + I(t,\mathbf{A}f), \qquad (3.8)$$

where

$$I(t,f) := \int_0^t \int_U \int_0^{t-s} v_{s-}((\lambda(u)-1)T_r f)dr(Y(du,ds)-\mu(du)ds).$$

For $n \in \mathbb{N}$, define the stopping time T_n by

$$T_n := \inf\left\{ 0 \leqslant t \leqslant T : \int_0^t \int_U |v_s|(|\lambda(u)-1|)\mu(du)ds > n \right\}. \qquad (3.9)$$

Then $T_n \uparrow T$ as $n \to \infty$, Q-a.s. We also define

$$I(t,f,n) := \int_0^{t\wedge T_n} \int_U \int_0^{t-s} v_{s-}((\lambda(u)-1)T_r f)dr(Y(du,ds)-\mu(du)ds),$$

$$v(t,f,n) := \int_0^t v_0(T_s\mathbf{A}f)ds - \int_0^t v_s(\mathbf{A}f)ds + I(t,\mathbf{A}f,n). \qquad (3.10)$$

Part I First, suppose that (3.5) holds. Let $f \in \mathcal{D}$. By the stochastic Fubini's theorem (cf. [42, Theorem IV. 46]), (3.9) and (3.5), we get

$$\begin{aligned}
I(t,f,n) &= \int_0^t \int_0^{t-r} \int_U I_{\{s\leqslant T_n\}} v_{s-}((\lambda(u)-1)T_r f)(Y(du,ds)-\mu(du)ds)dr \\
&= \int_0^t \left[v_{(t-r)\wedge T_n}(T_r f) - v_0(T_r f) - \int_0^{(t-r)\wedge T_n} v_s(\mathbf{A}T_r f)ds \right] dr. \qquad (3.11)
\end{aligned}$$

The last term of (3.11) is equal to

$$\begin{aligned}\int_0^t\int_0^t I_{\{s\leqslant t-r\}}I_{\{s\leqslant T_n\}}v_s(\mathbf{A}T_rf)dsdr &= \int_0^t\int_0^t I_{\{s\leqslant T_n\}}I_{\{r\leqslant t-s\}}v_s(\mathbf{A}T_rf)drds\\ &= \int_0^{t\wedge T_n} v_s\left(\int_0^{t-s}\mathbf{A}T_rfdr\right)ds\\ &= \int_0^{t\wedge T_n} v_s(T_{t-s}f)ds-\int_0^{t\wedge T_n} v_s(f)ds.\end{aligned}$$

Consequently,

$$\begin{aligned}I(t,f,n) &= \int_0^t v_{s\wedge T_n}(T_{t-s}f)ds-\int_0^t v_0(T_sf)ds\\ &\quad -\int_0^{t\wedge T_n} v_s(T_{t-s}f)ds+\int_0^{t\wedge T_n} v_s(f)ds.\end{aligned}$$

We complete the proof through the following three steps.

(a) Let $f=R_\alpha\varphi$ for some $\alpha>0$ and $\varphi\in\mathcal{D}$. Hereafter $R_\alpha:=\int_0^\infty e^{-\alpha t}T_tdt$, $\alpha>0$, is the resolvent of $(T_t)_{t\geq 0}$. Then we obtain by (3.10) that

$$v(t,f,n)=\int_0^t I_{\{s>T_n\}}v_{T_n}(T_{t-s}\mathbf{A}f)ds-\int_0^t I_{\{s>T_n\}}v_s(\mathbf{A}f)ds.$$

Since $T_n\uparrow T$ as $n\to\infty$, Q-a.s., we get $v(t,f)=\lim_{n\to\infty}v(t,f,n)=0$, Q-a.s.

(b) Let $f\in\mathcal{D}$. Define $f_k:=kR_kf$, $k\in\mathbb{N}$. Note that $f_k\to f$ boundedly and pointwise as $k\to\infty$. By the bounded convergence theorem, $v_0(T_tf_k)\to v_0(T_tf)$ and $v_t(f_k)\to v_t(f)$ as $k\to\infty$. Moreover, by (3.9), we get

$$\begin{aligned}E^Q\Bigg[\Bigg|&\int_0^{t\wedge T_n}\int_U v_{s-}((\lambda(u)-1)T_{t-s}f_k)(Y(du,ds)-\mu(du)ds)\\ &-\int_0^{t\wedge T_n}\int_U v_{s-}((\lambda(u)-1)T_{t-s}f)(Y(du,ds)-\mu(du)ds)\Bigg|\Bigg]\to 0\end{aligned}$$

as $k\to\infty$ for each n. Then $v(t,f)=0$, Q-a.s, on $\{t\leq T_n\}$. Thus $v(t,f)=0$, Q-a.s. Therefore (3.6) holds for any $f\in\mathcal{D}$ by (3.7) and (3.5).

(c) Let $f\in B_b(E)$. Since $\mathcal{D}$ is bounded pointwise dense in $B_b(E)$, there exists a sequence $\{f_n\}_{n\geq 1}\subset\mathcal{D}$ such that $\sup_{n\geq 1}\|f_n\|_\infty<\infty$ and $\lim_{n\to\infty}f_n=f$ boundedly and pointwise. Therefore (3.6) holds for f by (b), the dominated convergence theorem and the stopping time argument (cf. (b)).

Part II Conversely, suppose that (3.6) holds for any $f\in B_b(E)$. Let $f\in\mathcal{D}$. We obtain that

$$\int_0^t v_s(\mathbf{A}f)ds=\int_0^t v_0(T_s\mathbf{A}f)ds+\int_0^t\int_0^r\int_U v_{s-}((\lambda(u)-1)T_{r-s}\mathbf{A}f)(Y(du,ds)-\mu(du)ds)dr.$$

Hence, by (3.8), we get

$$\begin{aligned}
v(t,f) &= I(t,\mathbf{A}f) - \int_0^t\int_0^r\int_U v_{s-}((\lambda(u)-1)T_{r-s}\mathbf{A}f)(Y(du,ds)-\mu(du)ds)dr \\
&= \int_0^t\int_U\int_s^t v_{s-}((\lambda(u)-1)T_{r-s}\mathbf{A}f)dr(Y(du,ds)-\mu(du)ds) \\
&\quad -\int_0^t\int_U\int_0^r v_{s-}((\lambda(u)-1)T_{r-s}\mathbf{A}f)(Y(du,ds)-\mu(du)ds)dr \\
v(t,f,n) &= \int_0^t\int_U\int_s^t I_{\{s\leqslant T_n\}}v_{s-}((\lambda(u)-1)T_{r-s}\mathbf{A}f)dr(Y(du,ds)-\mu(du)ds) \\
&\quad -\int_0^t\int_0^r\int_U v_{s-}((\lambda(u)-1)T_{r-s}\mathbf{A}f)(Y(du,ds)-\mu(du)ds)dr \\
&= \int_0^t\int_0^{r\wedge T_n}\int_U v_{s-}((\lambda(u)-1)T_{r-s}\mathbf{A}f)(Y(du,ds)-\mu(du)ds)dr \\
&\quad -\int_0^t\int_0^r\int_U v_{s-}((\lambda(u)-1)T_{r-s}\mathbf{A}f)(Y(du,ds)-\mu(du)ds)dr \to 0 \quad \text{as } n\to\infty, \quad Q-a.s.
\end{aligned}$$

Therefore (3.5) holds. □

Theorem 3.4 Suppose that (X,Y) satisfies Assumptions 2.1 - 2.4. Then, for any $t>0$ and $f\in\mathcal{D}$, we have

$$\sigma_t(f)=\sigma_0(f)+\int_0^t\sigma_s(\mathbf{A}f)ds+\int_0^t\int_U\sigma_{s-}((\lambda(u)-1)f)(Y(du,ds)-\mu(du)ds), \quad Q-a.s. \tag{3.12}$$

Proof. Note that (2.2) is equivalent to

$$\int_0^t\int_U E^Q[L(s)|\lambda(u,X(s))-1|]\mu(du)ds<\infty, \quad \forall\ t>0,$$

and, under Q, Y is independent of X and Y is a unit Poisson random measure with the mean measure $\mu\times m$. By (3.1) and (3.3), we get

$$L(t)=1+\int_0^t\int_U L(s-)(\lambda(u,X(s))-1)(Y(du,ds)-\mu(du)ds).$$

Then, for $f\in B_b(E)$, we have

$$f(X(t))L(t)=f(X(t))+\int_0^t\int_U f(X(t))L(s-)(\lambda(u,X(s))-1)(Y(du,ds)-\mu(du)ds). \tag{3.13}$$

By taking conditional expectations on both sides of (3.13) and using the integral representation of martingales (cf. [5, VIII.T8, pp. 239]), as in [5, VI.T7 and R9, pp.173 and 177], we find that

$$\begin{aligned}
E^Q[f(X(t))L(t)|\mathcal{F}_t^Y] &= E^Q[f(X(t))] \\
&\quad +\int_0^t\int_U E^Q[f(X(t))L(s-)(\lambda(u,X(s))-1)|\mathcal{F}_s^Y](Y(du,ds)-\mu(du)ds).
\end{aligned}$$

Further, by the Q-independence of X and Y and the Markovian property of X, we have

$$\begin{aligned}&\int_0^t \int_U E^Q[f(X(t))L(s-)(\lambda(u,X(s))-1)|\mathcal{F}_s^Y](Y(du,ds)-\mu(du)ds)\\&\qquad\qquad = \int_0^t \int_U E^Q[L(s-)(\lambda(u,X(s))-1)T_{t-s}f(X_s)|\mathcal{F}_s^Y](Y(du,ds)-\mu(du)ds).\end{aligned}$$

(cf. [5, VI.T7 and R9, pp. 173 and 177]). Then, we have

$$\sigma_t(f) = \sigma_0(T_t f) + \int_0^t \int_U \sigma_{s-}((\lambda(u)-1)T_{t-s}f)(Y(du,ds)-\mu(du)ds), \quad Q-a.s., \quad \forall f \in B_b(E).$$

Therefore, the proof is completed by Lemma 3.2. □

In the literature,(3.12) is called the unnormalized filtering equation, its *uniqueness* will be discussed in the next section.

4. Poisson Chaos Expansion and the Uniqueness of Solutions

A general approach to prove the uniqueness of solutions of filtering equations is the method of *filtered martingale problem* developed in [32]. See [28] and [11] for filtering with counting process observations and [31] for a recent development. Here, we derive the Poisson chaos expansion for the unnormalized filtering equation and thus obtain the uniqueness of its solutions.

We first consider the special and less-demanding case that $\lambda \geq 1$. Fix a constant $T > 0$. Denote by $\mathcal{M}_+(E)$ the families of all finite measures on $(E, \mathcal{B}(E))$ and denote by ν the initial distribution of X.

Lemma 4.1 Suppose that (X,Y) satisfies Assumptions 2.1 - 2.3, $\lambda : U \times E \to [1,\infty)$ is $\mathcal{U}\times\mathcal{B}(E)$-measurable and satisfies (2.3). Let $(v_t)_{0\leq t\leq T}$ be an $\mathcal{M}_+(E)$-valued càdlàg process satisfying

(i) $\int_0^T \int_U v_t^2(\lambda(u)-1)\mu(du)dt < \infty$, Q-a.s.

(ii) For any $f \in \mathcal{D}$, $\{v_t(f)\}_{0\leq t\leq T}$ is an $\{\mathcal{F}_t^Y\}_{0\leq t\leq T}$ semi-martingale with

$$v_t(f) = v_0(f) + \int_0^t v_s(\mathbf{A}f)ds + \int_0^t \int_U v_{s-}((\lambda(u)-1)f)(Y(du,ds)-\mu(du)ds), \quad Q\text{-a.s.}$$

Let f be a function on E satisfying $E^Q[v_t^2(|f|)] < \infty$ for some $t \in [0,T]$. Then

$$v_t(f) = v_0(T_t f) + \int_0^t \int_U v_{s-}((\lambda(u)-1)T_{t-s}f)(Y(du,ds)-\mu(du)ds), \quad Q\text{-a.s.} \tag{4.1}$$

Proof. For $n \in \mathbb{N}$, we define $f_n := ((-n) \vee f) \wedge n$. Then, by Lemma 3.2, we have

$$v_t(f_n) = v_0(T_t f_n) + \int_0^t \int_U v_{s-}((\lambda(u)-1)T_{t-s}f_n)(Y(du,ds) - \mu(du)ds), \quad Q\text{-a.s.} \tag{4.2}$$

and

$$v_t(|f_n|) = v_0(T_t|f_n|) + \int_0^t \int_U v_{s-}((\lambda(u)-1)T_{t-s}|f_n|)(Y(du,ds) - \mu(du)ds), \quad Q\text{-a.s.} \tag{4.3}$$

By (4.3), Fatou's lemma and the dominated convergence theorem, we get

$$\begin{aligned} v_0^2(T_t|f|) &+ \int_0^t \int_U E^Q[v_s^2((\lambda(u)-1)T_{t-s}|f|)]\mu(du)ds \\ &\leq \varliminf_{n\to\infty} \left[v_0^2(T_t|f_n|) + \int_0^t \int_U E^Q[v_s^2((\lambda(u)-1)T_{t-s}|f_n|)]\mu(du)ds \right] \\ &= \varliminf_{n\to\infty} E^Q[v_t^2(|f_n|)] \\ &= E^Q[v_t^2(|f|)] \\ &< \infty. \end{aligned} \tag{4.4}$$

Therefore (4.1) holds by (4.2), (4.4) and the dominated convergence theorem. □

Theorem 4.2 Suppose that (X, Y) satisfies Assumptions 2.1 - 2.3, $\lambda : U \times E \to [1, \infty)$ is $\mathcal{U} \times \mathcal{B}(E)$-measurable and satisfies (2.3). Let $(v_t^i)_{0\leq t\leq T}$, $i = 1, 2$, be two $\mathcal{M}_+(E)$-valued càdlàg processes satisfying

(i) $v_0^i = \nu$ and $\int_0^T \int_U \{v_t^i(\lambda(u)-1)\}^2 \mu(du)dt < \infty$, Q-a.s.

(ii) For any $f \in \mathcal{D}$, $\{v_t^i(f)\}_{0\leq t\leq T}$ is an $\{\mathcal{F}_t^Y\}_{0\leq t\leq T}$ semi-martingale with

$$v_t^i(f) = v_0^i(f) + \int_0^t v_s^i(\mathbf{A}f)ds + \int_0^t \int_U v_{s-}^i((\lambda(u)-1)f)(Y(du,ds) - \mu(du)ds), \quad Q\text{-a.s.}$$

(iii) $E^Q[\{v_t^i(|f|)\}^2] < \infty$ for any $f \in \mathcal{D}$ and $t \in [0, T]$.

Then $v_t^1 = v_t^2$ for all $t \in [0, T]$. Moreover, we have the unique Poisson chaos expansion

$$\begin{aligned} v_t^i(f) &= \nu(T_t f) + \int_0^t \int_U \nu(T_{t_1}((\lambda(u_1)-1)T_{t-t_1}f))(Y(du_1,dt_1) - \mu(du_1)dt_1) \\ &\quad + \int_0^t \int_U \int_0^{t_1} \int_U \nu(T_{t_2}((\lambda(u_2)-1)T_{t_1-t_2}((\lambda(u_1)-1)T_{t-t_1}f))) \\ &\qquad \cdot (Y(du_2,dt_2) - \mu(du_2)dt_2)(Y(du_1,dt_1) - \mu(du_1)dt_1) \\ &\quad + \cdots, \quad Q\text{-a.s.}, \quad \forall f \in \mathcal{D}. \end{aligned} \tag{4.5}$$

Proof. Let $f \in \mathcal{D}$ and set $v_t = v_t^1$ or $v_t = v_t^2$ for $t \in [0, T]$. Then, by Lemma 4.1, we get

$$v_t(f) = v_0(T_t f) + \int_0^t \int_U v_{t_1-}((\lambda(u_1) - 1)T_{t-t_1} f)(Y(du_1, dt_1) - \mu(du_1)dt_1), \quad Q\text{-a.s.} \tag{4.6}$$

By (4.6) we get $\int_0^T \int_U E^Q[v_{t_1}^2((\lambda(u_1)-1)T_{t-t_1}|f|)]\mu(du_1)dt_1 \le E^Q[v_T^2(|f|)] < \infty$, hence $E^Q[v_{t_1}^2((\lambda(u_1)-1)T_{t-t_1}|f|)] < \infty$ for $\mu \times m$-a.e. (u_1, t_1). By Lemma 4.1, for $\mu \times m$-a.e. (u_1, t_1), we have

$$\begin{aligned} v_{t_1}((\lambda(u_1) - 1)T_{t-t_1} f) &= v_0(T_{t_1}((\lambda(u_1) - 1)T_{t-t_1} f)) \\ &+ \int_0^{t_1} \int_U v_{t_2-}((\lambda(u_2) - 1)T_{t_1-t_2}((\lambda(u_1) - 1)T_{t-t_1} f))(Y(du_2, dt_2) - \mu(du_2)dt_2), \quad Q\text{-a.s.} \end{aligned} \tag{4.7}$$

Since $E^Q[v_{t_1}^2((\lambda(u_1) - 1)T_{t-t_1}|f|)] < \infty$, similar to the proof of Lemma 4.1, we obtain by considering $((-n) \vee ((\lambda - 1)T_{t-t_1} f)) \wedge n$ and Fatou's lemma that

$$\begin{aligned} v_0^2(T_{t_1}((\lambda(u_1) - 1)T_{t-t_1}|f|)) + \int_0^{t_1} \int_U E^Q[v_{t_2}^2((\lambda(u_2) - 1)T_{t_1-t_2}((\lambda(u_1) - 1)T_{t-t_1}|f|))]\mu(du_2)dt_2 \\ \le E^Q[v_{t_1}^2((\lambda(u_1) - 1)T_{t-t_1}|f|)] < \infty. \end{aligned}$$

Then, by (4.6) and (4.7), we get

$$\begin{aligned} v_t(f) &= v_0(T_t f) + \int_0^t \int_U v_0(T_{t_1}((\lambda(u_1) - 1)T_{t-t_1} f))(Y(du_1, dt_1) - \mu(du_1)dt_1) \\ &+ \int_0^t \int_U \int_0^{t_1} \int_U v_{t_2-}((\lambda(u_2) - 1)T_{t_1-t_2}((\lambda(u_1) - 1)T_{t-t_1} f)) \\ &\cdot(Y(du_2, dt_2) - \mu(du_2)dt_2)(Y(du_1, dt_1) - \mu(du_1)dt_1), \quad Q\text{-a.s.} \end{aligned}$$

Furthermore, we obtain by induction that for each $n \ge 2$,

$$\begin{aligned} v_t(f) &= v_0(T_t f) + \int_0^t \int_U v_0(T_{t_1}((\lambda(u_1) - 1)T_{t-t_1} f))(Y(du_1, dt_1) - \mu(du_1)dt_1) \\ &+ \cdots \\ &+ \int_0^t \int_U \int_0^{t_1} \int_U \cdots \int_0^{t_{n-2}} \int_U v_0(T_{t_{n-1}}(\cdots)) \\ &\cdot(Y(du_{n-1}, dt_{n-1}) - \mu(du_{n-1})dt_{n-1}) \cdots (Y(du_2, dt_2) - \mu(du_2)dt_2)(Y(du_1, dt_1) - \mu(du_1)dt_1) \\ &+ \int_0^t \int_U \int_0^{t_1} \int_U \cdots \int_0^{t_{n-1}} \int_U v_{t_n-}((\lambda(u_n) - 1)T_{t_{n-1}-t_n}(\cdots)) \\ &\cdot(Y(du_n, dt_n) - \mu(du_n)dt_n) \cdots (Y(du_2, dt_2) - \mu(du_2)dt_2)(Y(du_1, dt_1) - \mu(du_1)dt_1), \quad Q\text{-a.s.} \end{aligned} \tag{4.8}$$

Note that $v_t(f) \in L^2(\Omega, \mathcal{F}_t^Y, Q)$. By the theory of Poisson chaos expansion (cf. e.g. [40]) we have

$$L^2(\Omega, \mathcal{F}_t^Y, Q) = \sum_{n=0}^{\infty} \bigoplus \mathcal{H}(t)_n, \tag{4.9}$$

where $\mathcal{H}(t)_n$ is the space of n-fold multiple Poisson integrals on the interval $[0,t]$. We obtain by the orthogonality of $\mathcal{H}(t)_n$, $n \in \mathbb{N}$, that

$$\begin{aligned}
&E^Q\left[|v_0(T_t f)|^2\right] + E^Q\left[\left|\int_0^t \int_U v_0(T_{t_1}((\lambda(u_1)-1)T_{t-t_1}f))(Y(du_1,dt_1)-\mu(du_1)dt_1)\right|^2\right] \\
&\quad + \cdots + E^Q\Big[\Big|\int_0^t \int_U \int_0^{t_1} \int_U \cdots \int_0^{t_{n-2}} \int_U v_0(T_{t_{n-1}}(\cdots)) \\
&\qquad \cdot (Y(du_{n-1},dt_{n-1}) - \mu(du_{n-1})dt_{n-1}) \cdots (Y(du_2,dt_2)-\mu(du_2)dt_2)(Y(du_1,dt_1)-\mu(du_1)dt_1)\Big|^2\Big] \\
&\quad \le E^Q[|v_t(f)|^2] < \infty.
\end{aligned}$$

Since n is arbitrary, the right hand side of (4.5) converges in $L^2(\Omega, \mathcal{F}_t^Y, Q)$ and is thus well-defined. By (4.8), one finds that for each $n \in \mathbb{N}$ the first n terms of the Poisson chaos expansion of $v_t(f)$ must be equal to the first n terms of the right hand side of (4.5). Hence by (4.9) we obtain the unique Poisson chaos expansion (4.5) of $v_t(f)$. Therefore, $v_t^1 = v_t^2$ for all $t \in [0,T]$ since $\mathcal{D}$ is a measure determining subset of $B_b(E)$. □

Next, we consider the general and more complicated case when $\lambda > 0$.

Theorem 4.3 Suppose that (X,Y) satisfies Assumptions 2.1 - 2.3, $\lambda : U \times E \to (0,\infty)$ is $\mathcal{U} \times \mathcal{B}(E)$-measurable and satisfies (2.3), and

$$\begin{aligned}
&\int_0^T \int_U \int_0^{t_1} \int_U \cdots \int_0^{t_{n-1}} \int_U \nu^2(T_{t_n}(|\lambda(u_n)-1| T_{t_{n-1}-t_n}(|\lambda(u_{n-1})-1| \cdots T_{t_1}|\lambda(u_1)-1|))) \\
&\qquad \cdot \mu(du_n)dt_n \cdots \mu(du_2)dt_2\mu(du_1)dt_1 < \infty, \quad \forall n \in \mathbb{N}.
\end{aligned} \tag{4.10}$$

Let $(v_t^i)_{0\le t\le T}$, $i=1,2$, be two $\mathcal{M}_b(S)$-valued càdlàg processes satisfying

i) $v_0^i = \nu$ and $\int_0^T \int_U E^Q[|v_t^i|^2(|\lambda(u)-1|)]\mu(du)dt < \infty$.

(ii) For any $f \in \mathcal{D}$, $\{v_t^i(f)\}_{0\le t\le T}$ is an $\{\mathcal{F}_t^Y\}_{0\le t\le T}$ semi-martingale with

$$v_t^i(f) = v_0^i(f) + \int_0^t v_s^i(\mathbf{A}f)ds + \int_0^t \int_U v_{s-}^i((\lambda(u)-1)f)(Y(du,ds)-\mu(du)ds), \quad Q-a.s.$$

Then $v_t^1 = v_t^2$ for all $t \in [0,T]$. Moreover, we have the unique Poisson expansion

$$\begin{aligned}
v_t^i(f) \;&=\; \nu(T_t f) + \int_0^t \int_U \nu(T_{t_1}((\lambda(u_1)-1)T_{t-t_1}f))(Y(du_1,dt_1)-\mu(du_1)dt_1) \\
&\quad + \int_0^t \int_U \int_0^{t_1} \int_U \nu(T_{t_2}((\lambda(u_2)-1)T_{t_1-t_2}((\lambda(u_1)-1)T_{t-t_1}f))) \\
&\qquad \cdot (Y(du_2,dt_2)-\mu(du_2)dt_2)(Y(du_1,dt_1)-\mu(du_1)dt_1) \\
&\quad + \cdots, \quad Q\text{-a.s.}, \quad \forall f \in B_b(E).
\end{aligned} \tag{4.11}$$

Remark 4.4 Condition (4.10) is not needed for the case when $\lambda \geq 1$, because the positiveness of the unnormalized filter makes it possible to relax condition (4.10) when $\lambda \geq 1$. This point was noticed in [26], where the classical filtering models are considered. Note that $\lambda \geq 1$ ensures that Fatou's lemma and the dominated convergence theorem can be applied to obtain equation (4.4) and hence equation (4.1) of Lemma 4.1. For general $\lambda > 0$, the positiveness of the unnormalized filters alone is not sufficient to guarantee that equation (4.4) holds. Condition (4.10) is needed so that the dominated convergence theorem can apply. In the L_2-framework, condition (4.10) is a very weak condition to ensure the full Poisson chaos expansion (4.11).

Proof. Set $v_t = v_t^1$ or $v_t = v_t^2$ for $t \in [0,T]$. Then $\int_0^T \int_U E^Q[|v_t|^2(|\lambda(u)-1|)]\mu(du)dt < \infty$ by condition (i). Let $f \in B_b(E)$. By Lemma 3.2, $v_t(f) \in L^2(\Omega, \mathcal{F}_t, Q)$ for $t \in [0,T]$ and

$$v_t(f) = v_0(T_t f) + \int_0^t \int_U v_{t_1-}((\lambda(u_1)-1)T_{t-t_1}f)(Y(du_1,dt_1) - \mu(du_1)dt_1), \quad Q-a.s.$$

Define $(\lambda-1)_n := (\lambda-1)\wedge n$ for $n \in \mathbb{N}$. Then, by the dominated convergence theorem, we get

$$v_t(f) = v_0(T_t f) + \lim_{n\to\infty}\int_0^t \int_U v_{t_1-}((\lambda(u_1)-1)_n T_{t-t_1}f)(Y(du_1,dt_1) - \mu(du_1)dt_1), \quad Q-a.s.$$

Hereafter the limit is taken in the L^2-sense.

Apply the above argument to $(\lambda-1)_n T_{t-t_1}f$. By $\int_0^T \int_U E^Q[|v_t|^2(|\lambda(u)-1|)]\mu(du)dt < \infty$, the dominated convergence theorem and (4.10), we get

$$\begin{aligned}
v_t(f) &= v_0(T_t f) + \lim_{n\to\infty}\int_0^t \int_U v_0(T_{t_1}((\lambda(u_1)-1)_n T_{t-t_1}f))(Y(du_1,dt_1) - \mu(du_1)dt_1) \\
&+ \lim_{n\to\infty}\lim_{n_1\to\infty}\int_0^t \int_U \int_0^{t_1}\int_U v_{t_2-}((\lambda(u_2)-1)_{n_1}T_{t_1-t_2}((\lambda(u_1)-1)_n T_{t-t_1}f)) \\
&\qquad \cdot(Y(du_2,dt_2)-\mu(du_2)dt_2)(Y(du_1,dt_1)-\mu(du_1)dt_1) \\
&= v_0(T_t f) + \int_0^t \int_U v_0(T_{t_1}((\lambda(u_1)-1)T_{t-t_1}f))(Y(du_1,dt_1) - \mu(du_1)dt_1) \\
&+ \lim_{n\to\infty}\lim_{n_1\to\infty}\int_0^t \int_U \int_0^{t_1}\int_U v_{t_2-}((\lambda(u_2)-1)_{n_1}T_{t_1-t_2}((\lambda(u_1)-1)_n T_{t-t_1}f)) \\
&\qquad \cdot(Y(du_2,dt_2)-\mu(du_2)dt_2)(Y(du_1,dt_1)-\mu(du_1)dt_1), \\
&\qquad\qquad\qquad Q\text{−}a.s.
\end{aligned}$$

Then, by orthogonality, the first two terms of the above summation must be the first two terms of the unique Poisson expansion of $v_t(f)$. Repeat this procedure, by induction, we obtain the unique Poisson expansion (4.11) of $v_t(f)$. Therefore, $v_t^1 = v_t^2$ for all $t \in [0,T]$. □

Theorem 4.5 Suppose that (X,Y) satisfies Assumptions 2.1 - 2.4, and

$$\sum_{n=1}^{\infty}\int_0^T\int_U\int_0^{t_1}\int_U\cdots\int_0^{t_{n-1}}\int_U \nu^2(T_{t_n}(|\lambda(u_n)-1|T_{t_{n-1}-t_n}(|\lambda(u_{n-1})-1|\cdots T_{t_1}|\lambda(u_1)-1|))) \\ \cdot\mu(du_n)dt_n\cdots\mu(du_2)dt_2\mu(du_1)dt_1<\infty. \tag{4.12}$$

Then $\{\sigma_t\}_{0\le t\le T}$ is the unique $\mathcal{M}_b(S)$-valued solution of the unnormalized filtering equation (3.12). Moreover, we have the unique Poisson chaos expansion

$$\begin{aligned}\sigma_t(f) &= \nu(T_tf)+\int_0^t\int_U \nu(T_{t_1}((\lambda(u_1)-1)T_{t-t_1}f))(Y(du_1,dt_1)-\mu(du_1)dt_1)\\ &+ \int_0^t\int_U\int_0^{t_1}\int_U \nu(T_{t_2}((\lambda(u_2)-1)T_{t_1-t_2}((\lambda(u_1)-1)T_{t-t_1}f)))\\ &\qquad\cdot(Y(du_2,dt_2)-\mu(du_2)dt_2)(Y(du_1,dt_1)-\mu(du_1)dt_1)\\ &+\cdots,\quad Q\text{-a.s.},\quad \forall f\in B_b(E).\end{aligned} \tag{4.13}$$

Proof. By Theorems 3.4 and 4.3, we only need to show that

$$E^Q\left[\int_0^T\int_U\{\sigma_t(|\lambda(u)-1|)\}^2\mu(du)dt\right]<\infty.$$

Define $(\lambda-1)_n:=(\lambda-1)\wedge n$ for $n\in\mathbb{N}$. Let $\{\sigma_t^{(\lambda-1)_n}\}_{0\le t\le T}$ be the unnormalized filtering process w.r.t. $(\lambda-1)_n$, i.e.,

$$\sigma_t^{(\lambda-1)_n}(f):=E^{Q,(\lambda-1)_n}[f(X(t))L^{(\lambda-1)_n}(t)|\mathcal{F}_t^{Y,(\lambda-1)_n}],\quad f\in B_b(E).$$

Hereafter, we use $Y^{(\lambda-1)_n}$, $\mathcal{F}_t^{Y,(\lambda-1)_n}$, $Q^{(\lambda-1)_n}$ and $E^{Q,(\lambda-1)_n}$ to denote respectively Y, $\mathcal{F}_t^Y$, Q and E^Q corresponding to the intensity function $(\lambda-1)_n$. Since X and Y are independent under Q, by Fatou's Lemma, Theorem 3.4, Theorem 4.3 and the dominated convergence theorem, we get

$$
\begin{aligned}
& E^Q\left[\int_0^T\int_U\{\sigma_t(|\lambda(u)-1|)\}^2\mu(du)dt\right] \\
= & \int_0^T\int_U\int\left(\int|\lambda(u)-1|(X(t))L(t)dQ^X\right)^2 dQ^Y\mu(du)dt \\
\leq & \varliminf_{n\to\infty}\int_0^T\int_U\int\left(\int|\lambda(u)-1|(X(t))L^{(\lambda-1)_n}(t)dQ^X\right)^2 dQ^Y\mu(du)dt \\
= & \varliminf_{n\to\infty}\int_0^T\int_U\int\left(\int|\lambda(u)-1|(X(t))L^{(\lambda-1)_n}(t)dQ^{(\lambda-1)_n,X}\right)^2 dQ^{(\lambda-1)_n,Y}\mu(du)dt \\
= & \varliminf_{n\to\infty}E^{Q,(\lambda-1)_n}\left[\int_0^T\int_U\{\sigma_t(|\lambda(u)-1|)\}^2\mu(du)dt\right] \\
\leq & \varliminf_{n\to\infty}\varliminf_{n_1\to\infty}E^{Q,(\lambda-1)_n}\left[\int_0^T\int_U\{\sigma_t(|(\lambda(u)-1)_{n_1}|\}^2\mu(du)dt\right] \\
\leq & \sum_{n=1}^{\infty}\int_0^T\int_U\int_0^{t_1}\int_U\cdots\int_0^{t_{n-1}}\int_U\nu^2(T_{t_n}(|\lambda(u_n)-1|T_{t_{n-1}-t_n}(|\lambda(u_{n-1})-1|\cdots T_{t_1}|\lambda(u_1)-1|))) \\
& \qquad\cdot\mu(du_n)dt_n\cdots\mu(du_2)dt_2\mu(du_1)dt_1 \\
< & \infty,
\end{aligned}
$$

where Q^X and Q^Y denote the marginal probabilities of X and Y w.r.t. Q, respectively. □

Remark 4.6 The assumptions of Theorems 4.3 and 4.5 are weak and can be verified for a large class of unbounded stochastic intensities λ. Below we give two examples. For simplicity, we let $U=\{1\}$ and denote $h:=\lambda-1$.

(i) Suppose that η is a σ-finite measure on $(E,\mathcal{B}(E))$ such that $(T_t)_{t\geq 0}$ can be extended to be a contraction semigroup on $L^2(E;\eta)$ (e.g. X is a symmetric Markov process or is a non-symmetric Markov process associated with a Dirichlet form), $d\nu=u_0d\eta$ with $u_0\in L^2(E;\eta)$, and $h\in L^n(E;\eta)$ for any $n\in\mathbb{N}$. Then (4.10) holds. In fact, noting that $(T_tf(x))^n=(E_x[f(X(t))])^n\leq$

$E_x[f^n(X(t))] = T_t f^n(x)$ for any $f \geq 0$ on E and $n \in \mathbb{N}$, we get

$$
\begin{aligned}
&\nu(T_{t_{n+1}}(|h|T_{t_n - t_{n+1}}(|h|T_{t_{n-1}-t_n}\cdots)))\\
&\quad = \int_E u_0 T_{t_{n+1}}(|h|T_{t_n - t_{n+1}}(|h|T_{t_{n-1}-t_n}\cdots))d\eta\\
&\quad \leq \left(\int_E u_0^2 d\eta\right)^{1/2}\left(\int_E [T_{t_{n+1}}(|h|T_{t_n-t_{n+1}}(|h|T_{t_{n-1}-t_n}\cdots))]^2 d\eta\right)^{1/2}\\
&\quad \leq \left(\int_E u_0^2 d\eta\right)^{1/2}\left(\int_E |h|^2(T_{t_n-t_{n+1}}(|h|T_{t_{n-1}-t_n}\cdots))^2 d\eta\right)^{1/2}\\
&\quad \leq \left(\int_E u_0^2 d\eta\right)^{1/2}\left(\int_E |h|^4 d\eta\right)^{1/4}\left(\int_E (T_{t_n-t_{n+1}}(|h|T_{t_{n-1}-t_n}\cdots))^4 d\eta\right)^{1/4}\\
&\quad \leq \left(\int_E u_0^2 d\eta\right)^{1/2}\left(\int_E |h|^4 d\eta\right)^{1/4}\left(\int_E (T_{t_n-t_{n+1}}(|h|^2(T_{t_{n-1}-t_n}\cdots)^2))^2 d\eta\right)^{1/4}\\
&\quad \leq \left(\int_E u_0^2 d\eta\right)^{1/2}\left(\int_E |h|^4 d\eta\right)^{1/4}\left(\int_E |h|^4(T_{t_{n-1}-t_n}\cdots)^4 d\eta\right)^{1/4}\\
&\quad \leq \left(\int_E u_0^2 d\eta\right)^{1/2}\left(\int_E |h|^4 d\eta\right)^{1/4}\left(\int_E |h|^8 d\eta\right)^{1/8}\left(\int_E (T_{t_{n-1}-t_n}\cdots)^8 d\eta\right)^{1/8}\\
&\quad \leq \cdots\\
&\quad < \infty.
\end{aligned}
$$

Hence (4.10) is satisfied.

This is a general example because of the generality of X and because the η (initial distribution of X) in ν is a general measure, which can be different from the Lebesgue measure dx in applications. For example, if $d\eta = \rho(x)dx$ where $\rho(x)$ decays at infinity faster than any power of x, then any function λ of polynomial-order belongs to every $L_p(\mathbb{R}^d; d\eta)$ and hence satisfies condition (4.10). An interesting and useful example is as follows. Let X be the Ornstein-Uhlenbeck (OU) process in $\mathbb{R}^d$:

$$dX(t) = -\frac{1}{2}X(t)dt + dW(t),$$

where W is the standard Brownian motion in $\mathbb{R}^d$ and let

$$d\eta = \frac{1}{(2\pi)^{d/2}} e^{-\frac{|x|^2}{2}} dx$$

be the standard d-dimensional normal distribution. Then, any function λ of polynomial-order belongs to every $L_p(\mathbb{R}^d; d\eta)$ and hence satisfies condition (4.10). Clearly, the unbounded condition of polynomial-order, which arises frequently in applications, is weaker than the bounded condition.

Vasicek interest rate model, which is an OU process, is commonly used in financial bond pricing. There are UHF data of bond trading price such as GovPX tick-by-tick data, which has been widely used in empirical finance literature (see, for example, [23] and [4]). A further concrete example can let the intrinsic value process of a short interest rate, $X(t)$, as a Vasicek model. When focusing on $U = \{1\}$, namely, the trading activity, we can allow λ as a polynomial such as $\lambda(u, x) = ax^2 + 1$.

(ii) Let $\{\sigma_t^{|h|+1}\}_{0\leq t\leq T}$ be the unnormalized filtering process w.r.t. $(|h| + 1)$, i.e.,

$$\sigma_t^{|h|+1}(f) := E^{Q,|h|+1}[f(X(t))L^{|h|+1}(t)|\mathcal{F}_t^{Y,|h|+1}], \quad f \in B_b(E),$$

where

$$L^{|h|+1}(t) := \exp\left\{\int_0^t \log(|h|(X(s-)) + 1)dY^{|h|+1}(s) - \int_0^t |h|(X(s))ds\right\},$$

with Y as a Poisson process. Hereafter, we use $Y^{|h|+1}$, $\mathcal{F}_t^{Y,|h|+1}$, $Q^{|h|+1}$ and $E^{Q,|h|+1}$ to denote respectively Y, $\mathcal{F}_t^Y$, Q and E^Q corresponding to the stochastic intensity $(|h| + 1)$. By Theorem 3.4, we get

$$\sigma_t^{|h|+1}(f) = \sigma_0(f) + \int_0^t \sigma_s^{|h|+1}(\mathbf{A}f)ds + \int_0^t \sigma_{s-}^{|h|+1}(|h|f)d(Y^{|h|+1}(s) - s), \quad Q^{|h|+1}\text{-a.s.}, \quad \forall f \in \mathcal{D}.$$

Moreover, $\{\sigma_t^{|h|+1}(1)\}_{0\leq t\leq T}$ is an $\{\mathcal{F}_t^{Y,|h|+1}\}_{0\leq t\leq T}$ martingale under $Q^{|h|+1}$.

Let X be a conservative Markov process. Then

$$\sigma_t^{|h|+1}(1) = 1 + \int_0^t \sigma_{s-}^{|h|+1}(|h|)d(Y^{|h|+1}(s) - s), \quad Q^{|h|+1}\text{-a.s.}$$

Further, we assume that $\{\sigma_t^{|h|+1}(1)\}_{0\leq t\leq T}$ is square-integrable, which is satisfied, e.g. if the Novikov-type exponential integrability condition

$$E\left[\exp\left\{\int_0^t h^2(X(s))ds\right\}\right] < \infty$$

is satisfied. Note that the Novikov-type condition is significantly weaker than the bounded condition, which is assumed in the setting of classical filtering models for multiple integral expansions (cf. e.g. [39, pp. 12]).

Similarly, in our setup, it is straightforward to give interesting examples satisfying the above Novikov-type condition. For example, let X be the standard Brownian motion in $\mathbb{R}^d$ or a Poisson process, and let λ be any function of polynomial-order. Then, by the Taylor expansion of the exponential function and the moment estimates of Brownian motion and Poisson process, we can check that the above Novikov-type condition is satisfied. Furthermore, by Lévy-Itô decomposition and Itô's SDEs, the above Novikov-type condition is also satisfied for some more general Lévy processes and diffusions.

Under the above conditions, we get

$$\int_0^t E^{Q,|h|+1}[\{\sigma_s^{|h|+1}(|h|)\}^2]ds = E^{Q,|h|+1}[\{\sigma_t^{|h|+1}(1)\}^2] - 1 < \infty.$$

Then, $\{\sigma_t^{|h|+1}\}_{0\leq t\leq T}$ satisfies all the conditions of Theorem 4.2. By Theorem 4.2, we obtain the Poisson chaos expansion of $\sigma_t^{|h|+1}(1)$:

$$\begin{aligned}\sigma_t^{|h|+1}(1) &= 1 + \int_0^t \nu(T_{t_1}|h|)d(Y^{|h|+1}(t_1) - t_1) \\ &\quad + \int_0^t \int_0^{t_1} \nu(T_{t_2}(|h|T_{t_1-t_2}|h|))d(Y^{|h|+1}(t_2) - t_2)d(Y^{|h|+1}(t_1) - t_1) + \cdots, \quad Q^{|h|+1}\text{-a.s.}\end{aligned}$$

Therefore

$$\sum_{n=1}^{\infty} \int_0^t \int_0^{t_1} \cdots \int_0^{t_{n-1}} \nu^2(T_{t_n}(|h|T_{t_{n-1}-t_n}(|h|\cdots T_{t_1}|h|)))dt_n \cdots dt_2 dt_1 < \infty.$$

Hence (4.12) is satisfied.

5. Poisson Chaos Expansion for the Unnormalized Filter Density and a Related Recursive Algorithm

In the classical diffusion filtering model, there are a number of papers dealing with multiple integral expansions of both normalized and unnormalized filters. See, for example, [39, 7, 8, 9]. With applying the established Poisson expansion (4.13) for practical computations in mind, we consider a time-homogeneous multi-dimensional diffusion signal in $\mathbb{R}^d$. We will derive the Poisson chaos expansions in two forms (see Theorems 5.1 and 5.2) for the unnormalized filter density with additional conditions when the observations are MPP. To make the presentation more transparent, we develop a recursive algorithm for the case with one counting process observation. The algorithm is the counterpart of that of [37, 38] with MPP observations and can be applied to more general signals and MPP observations.

5.1 Poisson Chaos Expansion for the Unnormalized Filter Density

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a complete probability space. Suppose that $b = (b^i)_{1\leq i\leq d}$ is a d-dimensional vector function on $\mathbb{R}^d$, $\sigma = (\sigma^{ij})_{1\leq i\leq d, 1\leq j\leq d_1}$ is a $d \times d_1$-dimensional matrix function on $\mathbb{R}^d$, and

$B = (B^i)_{1\le i\le d_1}$ is a d_1-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathcal{P})$. The signal $(X(t))_{t\ge 0}$ is given as follows.

$$X(t) = X(0) + \int_0^t b(X(s))ds + \int_0^t \sigma(X(s))dB(s).$$

The following conditions are assumed:

(i) The functions b and σ are infinitely differentiable and all the derivatives are bounded.

(ii) $X(0)$ is independent of $(B(t))_{t\ge 0}$, $X(0)$ has a density p w.r.t. the Lebesgue measure dx and p is a smooth rapidly decreasing function on $\mathbb{R}^d$.

We consider the filtering model of Section 2. Suppose that (X, Y) satisfies Assumptions 2.2 - 2.4, $X(0)$ is independent of ξ and, for any $u \in U$, $\lambda(u)$ is infinitely differentiable and all the derivatives are bounded on $\mathbb{R}^d$. By (3.4), for $f \in B_b(\mathbb{R}^d)$, we have

$$\sigma_t(f) = E^Q[f(X(t))L(t)|\mathcal{F}_t^Y] = E^Q[E^Q[L(t)|\sigma(X(t)) \vee \mathcal{F}_t^Y]f(X(t))|\mathcal{F}_t^Y].$$

Let $\tilde{L}(x, \mathcal{F}_t^Y)$ be a version of $E^Q[L(t)|\sigma(X(t)) \vee \mathcal{F}_t^Y]$ at $X(t) = x$ and given the path $\{Y(s) : 0 \le s \le t\}$. Define $P(t, A) := P(X(t) \in A)$ for $A \in \mathcal{B}(E)$. By condition (i) and Fubini's theorem, we get

$$E^Q[E^Q[L(t)|\sigma(X(t)) \vee \mathcal{F}_t^Y]f(X(t))|\mathcal{F}_t^Y] = \int_E f(x)\tilde{L}(x, \mathcal{F}_t^Y)P(t, dx).$$

By condition (ii), we know that $P(t, dx)$ is absolutely continuous w.r.t. dx. Denote by $p(x, t)$ the density and define $u(t, x) := \tilde{L}(x, \mathcal{F}_t^Y)p(t, x)$. Then

$$\sigma_t(f) = \int_{\mathbb{R}^d} u(t, x)f(x)dx, \quad \forall f \in B_b(\mathbb{R}^d),\ t \ge 0,\ x \in \mathbb{R}^d. \tag{5.1}$$

$u(t, x)$ is called the unnormalized filtering density function.

We consider the following equation corresponding to the adjoint operator of the generator $\mathbf{A}$:

$$\begin{aligned} \frac{\partial v(t,x)}{\partial t} &= \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2((\sigma\sigma')^{ij}(x)v(t,x))}{\partial x_i \partial x_j} - \sum_{i=1}^d \frac{\partial(b^i(x)v(t,x))}{\partial x_i}, \quad t > 0, \\ v(0, x) &= \varphi(x). \end{aligned} \tag{5.2}$$

Denote by $T_t^*\varphi$ the solution of (5.2). Let $T > 0$ be a fixed constant and consider a partition $0 = t_0 < t_1 < \cdots < t_M = T$ of $[0, T]$. Denote $\Delta_i = t_i - t_{i-1}$.

Theorem 5.1 In the setup of this section, we have

$$
\begin{aligned}
&u(t_0, x) = p(x),\\
&u(t_i, x) = T_t^* u(t_{i-1}, \cdot)(x)\\
&\quad + \sum_{k\geq 1} \int_0^{\Delta_i} \int_U \int_0^{s_1} \int_U \cdots \int_0^{s_{k-1}} \int_U T_{t-s_1}^*(\lambda(u_1)-1)\cdots T_{s_{k-1}-s_k}^*(\lambda(u_k)-1)T_{s_k}^* u(t_{i-1}, \cdot)))(x)\\
&\cdot(Y^{(i)}(du_k, ds_k) - \mu(du_k)ds_k)\cdots(Y^{(i)}(du_2, ds_2) - \mu(du_2)ds_2)(Y^{(i)}(du_1, ds_1) - \mu(du_1)ds_1)
\end{aligned}
$$

for $i = 1, \ldots, M$, where $Y^{(i)}(\cdot, t) = Y(\cdot, t + t_{i-1}) - Y(\cdot, t_{i-1})$, $0 \leq t \leq \Delta_i$.

Proof. Under the conditions of this section, all the assumptions of Theorem 4.5 are fulfilled. Therefore, the conclusions follow from Theorem 4.5, (5.1) and (5.2). □

To simplify notation, for $\mathcal{F}_{t_{i-1}}^Y$-measurable function $g = g(x, \omega)$ and $0 \leq t \leq \Delta_i$, we define

$$
\begin{aligned}
&F_0^{(i)}(t, g)(x) := T_t^* g(x)\\
&F_k^{(i)}(t, g)(x) := \int_0^t \int_U \int_0^{s_1} \int_U \cdots \int_0^{s_{k-1}} \int_U T_{t-s_1}^*(\lambda(u_1)-1)\cdots T_{s_{k-1}-s_k}^*(\lambda(u_k)-1)T_{s_k}^* u(t_{i-1}, \cdot)))(x)\\
&\quad \cdot(Y^{(i)}(du_k, ds_k) - \mu(du_k)ds_k)\cdots(Y^{(i)}(du_2, ds_2) - \mu(du_2)ds_2)(Y^{(i)}(du_1, ds_1) - \mu(du_1)ds_1), \quad k \geq 1.
\end{aligned}
$$

Then

$$
u(t_i, x) = \sum_{k\geq 0} F_k^{(i)}(\Delta_i, u(t_{i-1}, \cdot))(x), \quad i = 1, \ldots, M.
$$

Denote by $\|\cdot\|_0$ and $(\cdot, \cdot)_0$ the norm and the inner product of $L^2(\mathbb{R}^d; dx)$, respectively. Then, there exists a constant $c > 0$ such that (cf. [43])

$$
\|T_t^* \varphi\|_0 \leq e^{ct} \|\varphi\|_0.
$$

By induction, for every $t \in [0, \Delta_i]$, $i = 1, \ldots, M$ and $k \geq 0$, the operator $g \to F_k^{(i)}(t, g)$ is linear and bounded from $L^2(\Omega, Q; L^2(\mathbb{R}^d; dx))$ to itself and

$$
E^Q \|F_k^{(i)}(t, g)\|_0^2 \leq e^{ct}[(ct)^k/k!] E^Q \|g\|_0^2.
$$

This implies that $u(t_i, \cdot) \in L^2(\mathbb{R}^d; dx)$, Q-a.s.

Theorem 5.2 If $\{e_n\}$ is an orthonormal basis in $L^2(\mathbb{R}^d; dx)$ and random variables $\psi_n(i)$, $n \geq 0$, $i = 0, \ldots, M$ are defined recursively by

$$
\begin{aligned}
\psi_n(0) &= (p, e_n)_0,\\
\psi_n(i) &= \sum_{k\geq 0} \left(\sum_{l\geq 0} (F_k^{(i)}(\Delta_i, e_l), e_n)_0 \psi_l(i-1) \right), \quad i = 1, \ldots, M,
\end{aligned}
$$

then

$$u(t_i,\cdot) = \sum_{n\geq 0} \psi_n(i)e_n, \quad P-a.s.$$

Proof. This is a direct consequence of Theorem 5.1 and induction. $\square$

5.2 A Recursive Algorithm for One Counting Process Observation

Now we can use Theorem 5.2 to develop a recursive algorithm for $(\sigma_t)_{t\geq 0}$. For simplicity, we assume that $d=1$ and $U=\{1\}$, i.e., the observation $Y=(Y(t))_{t\geq 0}$ is only a Poisson process with intensity function λ on $\mathbb{R}$. Moreover, we assume that the partition of $[0,T]$ is uniform, i.e., $\Delta_i = \Delta$ for all $i=1,\ldots,M$. Let $\{e_n\}$ be the Hermite basis in $L^2(\mathbb{R};dx)$:

$$e_n(x) = \frac{1}{\sqrt{2^n \pi^{1/2} n!}} e^{-x^2/2} H_n(x),$$

where $H_n(x)$ is the nth Hermite polynomial defined by

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}, n \geq 0.$$

Recursive algorithm: Given a $\kappa \in \mathbb{N}$, we define random variables $\psi_{n,k}(i)$, $n=0,\ldots,\kappa$, $i=0,\ldots,M$, by

$$\begin{aligned}
\psi_{n,\kappa}(0) &= (p, e_n)_0, \\
\psi_{n,\kappa}(i) &= \sum_{l=0}^{\kappa} \Big((T^*_\Delta e_l, e_n)_0 + (T^*_\Delta(\lambda-1)e_l, e_n)_0 (Y(t_i) - Y(t_{i-1}) - \Delta) \\
&\qquad + (1/2)(T^*_\Delta(\lambda-1)^2 e_l, e_n)_0 ((Y(t_i)-Y(t_{i-1})-\Delta)^2 - \Delta) \Big) \psi_{n,\kappa}(i-1), \\
&\qquad\qquad i = 1,\ldots,M.
\end{aligned}$$

Then, the corresponding approximations to $u(t_i,x)$ and $\sigma_{t_i}(f)$ are

$$u_\kappa(t_i,x) = \sum_{n=0}^{\kappa} \psi_{n,\kappa}(i) e_n(x)$$

and

$$\sigma_{t_i,\kappa}(f) = \sum_{n=0}^{\kappa} \psi_{n,\kappa}(i) f_n,$$

where $f_n := \int_{\mathbb{R}} f(x)e_n(x)dx$.

Remark 5.3 By the well-known L^2-isomorphism between the Wiener space and the Poisson space, we can reproduce the following type of error bounds as given in [37, Theorem 3.1] step by step:

$$\max_{1\leq i\leq M}\sqrt{E^Q\|u_\kappa(t_i,\cdot)-u(t_i,\cdot)\|_0^2}\leq c\Delta+\frac{c(\gamma)}{\kappa^{\gamma-1/2}\Delta}$$

and

$$\max_{1\leq i\leq M}\sqrt{E^Q|\sigma_{t_i,\kappa}(f)-\sigma_{t_i}(f)|^2}\leq c\Delta+\frac{c(\gamma)}{\kappa^{\gamma-1/2}\Delta}$$

for any measurable function f on $\mathbb{R}$ such that $|f(x)|\leq L(1+|x|^l)$ for some $L>0$ and $l>0$. Here c is a positive constant depending only on the parameters of the model, i.e., b, σ, p and λ, and for any positive integer γ, $c(\gamma)$ is a positive constant depending only on γ and the parameters of the model. Then, by appropriate choice of Δ and κ, we can make the errors to be arbitrary small. This shows the *consistency* of the algorithm.

The above algorithm looks especially promising when the parameters of the model are given. In this case, the values of $(T_\Delta^*e_l,e_n)_0$, $(T_\Delta^*(\lambda-1)e_l,e_n)_0$ and $(T_\Delta^*(\lambda-1)^2e_l,e_n)$, $n,l=1,\ldots,\kappa$, can be pre-computed and stored. So only the incorporating of the increments of the observations is required at each step of updating. This leads to the sizable speed gain in the online computation.

Finally, we would like to point out that the algorithm of this section follows [37], which does not make use of the polynomial chaos expansion of [36]. As explained in [37, §1], the type of algorithms adopted in this section is more time-saving since it does not involve evaluation of the filtering density at many spatial points and the computation of $\sigma_t(f)$ does not require its subsequent evaluation.

6 Conclusions

Motivated by the recent applications in financial econometrics and mathematical finance, we study the filtering problem with MPP observations via Poisson chaos expansion. We derive the unnormalized filtering equation under weaker conditions and derive its Poisson chaos expansion. Based on the expansion, we obtain the uniqueness of solutions of the unnormalized filtering equation. Moreover, in order to show the possibility of developing numerically efficient approximate filters based on the chaos expansion, we construct a consistent numerical scheme with desirable computational advantages for a simple case.

Since such filtering problems arose from optical signal processing, our results may apply in communication and control theory. On the other hand, since the unnormalized filtering equation characterizes the evolution of the likelihood function of the filtering model for financial UHF data and the numerical scheme can compute the likelihood function, one interesting future work is to develop the maximum-likelihood inference for the filtering model for UHF data, especially, via EM algorithm. Since the conditional distribution of $X(t)$ can be obtained by normalizing the unnormalized filter, we can compute $E[X(t)|\mathcal{F}_t^Y]$ via the constructed approximate unnormalized filter density. This leads to important applications in mathematical finance problems. For example, we can use the computed $E[X(t)|\mathcal{F}_t^Y]$ in determining the risk-minimizing hedging strategy ([34]).

Through the simple example in Section 5, we reveal the possibility of giving efficient algorithms for computing the approximate filters based on the chaos expansion method. The algorithm can be extended to the general observations of marked point process and one future work is to conduct such a more thorough investigation and to provide some numerical examples. Another future work is to study the higher-order Poisson-Charlier polynomials of the observations in the approximation of the unnormalized filtering density. The final one is to study the large deviation principle of the unnormalized filtering equation, which is an SPDE driven by a Poisson random measure. The recent work [6] may shed light on such study.

Acknowledgements We would like to thank two anonymous referees and an associate editor for their constructive comments. Wei Sun gratefully acknowledges the support from Natural Sciences and Engineering Research Council of Canada under grant No. 311945-2008, and Yong Zeng gratefully acknowledges the supports from National Science Foundation under the grants of DMS-0604722 and DMS-1228244, and a grant from the University of Missouri Research Board in 2011.

References

[1] Aït-Sahalia, Y., Hansen, L. P. and Scheinkman, J., Operator methods for continuous-time Markov processes, Y. Aït-Sahalia and L. P. Hansen eds, 'Handbook of Financial Econometrics', North-Holland, Amsterdam, 2009.

[2] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N., 'Designing realized kernels to measure the ex post variation of equity prices in the presence of noise', *Econometrica* **76**, 1481-1536, 2008.

[3] Boel, R., Varaiya, P. and Wong, E., 'Martingales on jump processes. II: Applications', *SIAM Journal of Control* **13**, 1022-1061, 1975.

[4] Brandt, M. and Kavacejz, K., ‘Price discovery in the U.S. Treasury market: The impact of order flow and liquidity on the yield curve’, *Journal of Finance* **59**, 2623-2654, 2004.

[5] Brémaud, P., *Point Processes and Queues: Martingale Dynamics*, New York: Springer-Verlag, 1981.

[6] Budhiraja, A. Chen, J. and Dupuis, P., ‘Large deviation for stochastic partial differential equations driven by a Poisson random measure’, `http://arxiv.org/abs/1203.4020`, 2012.

[7] Budhiraja, A. and Kallianpur, G., ‘Approximations to the solution of the Zakai equation using multiple Wiener and Stratonovich integral expansions’, *Stochastics Stochastics Rep.* **56**, 271-315, 1996.

[8] Budhiraja, A. and Kallianpur, G., ‘The Feynman-Stratonovich semigroup and Stratonovich integral expansions in nonlinear filtering’, *Appl. math. Optim.* **35**, 91-116, 1997.

[9] Budhiraja, A. and Kallianpur, G., ‘The generalized Hu-Meyer formula for random kernels’, *Appl. math. Optim.* **35**, 177-202, 1997.

[10] Ceci, C., ‘Risk minimizing hedging for a partially observed high frequency data model’, *Stochastics* **78**, 13-31, 2006.

[11] Ceci, C. and Gerardi, A., ‘Filtering of a Markov jump process with counting observations’, *Applied Mathematics and Optimization* **42**, 1-18, 2000.

[12] Ceci, C. and Gerardi, A., ‘A model for high frequency data under partial information: a filtering approach’, *International Journal of Theoretical and Applied Finance* **9**, 555-576, 2006.

[13] Cvitanic, J., Liptser, R. and Rozovsky, B., ‘A filtering approach to tracking volatility from prices observed at random times’, *Annals of Applied Probability* **16**, 1633-1652, 2006.

[14] Davis, M. H. A., Segall, A. and Kailath, T., ‘Nonlinear filtering with counting observations’, *IEEE Transactions on Information Theory* **21**, 143-149, 1975.

[15] Duffie, D. and Glynn, P., ‘Estimation of continuous-time Markov processes sampled at random time intervals’, *Econometrica* **72**, 1773 - 1808, 2004.

[16] Elliott, R J. and Malcolm, W. P., ‘Nonlinear filtering with counting observations’, *IEEE Transactions on Automatic Control* **50**, 1123–1134, 2005.

[17] Elliott, R. J., Siu, T.K. and Yang, H., ‘Filtering a Markov modulated random measure’, *IEEE Transactions on Automatic Control* **55**, 74–88, 2010.

[18] Engle, R. F., ‘The econometrics of ultra-high-frequency data’, *Econometrica* **68**, 1–22, 2000.

[19] Engle, R. F. and Russell, J., ‘Autoregressive conditional duration: a new model for irregularly spaced transaction data’, *Econometrica* **66**, 1127–1162, 1998.

[20] Ethier, S. and Kurtz, T., *Markov Processes: Characterization and Convergence*, Wiley, New York, 1986.

[21] Frey, R. and Runggaldier, W. J., ‘A nonlinear filtering approach to volatility estimation with a view towards high frequency data’, *International Journal of Theoretical and Applied Finance* **4**, 199–210, 2001.

[22] Garcia, N. L. and Kurtz, T. G., Spatial point process and the projection method, *In and Out of Equilibrium 2*, V. Sidoravicius and M. E. Vares, eds. *Progress in Probability* **60**, 271–298, 2008.

[23] Green, T., 'Economic news and the impact of trading on bond prices', *Journal of Finance* **59**, 1201–1233, 2004.

[24] Grigelionis, B. and Mikulevicius, R., Stochastic evolution equations and densities of the conditional distributions, *Lecture Notes in Control and Information Sciences* **49**, G. Kallianpur eds., 49–88, 1983.

[25] Hasbrouck, J., *Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading*, Oxford University Press, USA, 2007.

[26] Hu, Z. C., Ma, Z. M. and Sun W., 'Nonlinear filtering of semi-Dirichlet processes', *Stochastic Processes and Their Applications* **119**, 3890–3913, 2009.

[27] Kallianpur, G. and Karandikar, R. L., 'White noise calculus and nonlinear filtering theory', *Annals of Probability* **13**, 1033–1107, 1985.

[28] Kliemann, W., Koch, G. and Marchetti, F., 'On the unnormalized solution of the filtering problem with counting process observations', *IEEE Transactions on Information Theory* **36**, 1415–1425, 1990.

[29] Kunita, H., 'Cauchy problem for stochastic partial differential equations arising in nonlinear filtering theory', *Systems & Control Letters* **1**, 37–41, 1981.

[30] Kurtz, T., 'Stochastic processes as projections of Poisson random measures. Special invited paper at IMS meeting, Washington, D.C., unpublished, 1989.

[31] Kurtz, T. and Nappo, G. 'The filtered martingale problem', in *The Oxford Handbook of Nonlinear Filtering*, edited by D. Crisan and B. Rozovskii, 129 - 165, 2011.

[32] Kurtz, T. and Ocone, D., 'Unique characterization of conditional distributions in nonlinear filtering', *Annals of Probability* **16**, 80–107, 1988.

[33] Last, G. and Brandt, A., *Marked Point Processes on the Real Line: The Dynamic Approach*, Springer, New York, 1995.

[34] Lee, K. and Zeng, Y., 'Risk minimization for a filtering micromovement model of asset price', *Applied Mathematical Finance* **17**, 177–199, 2010.

[35] Liptser, R. and Shiryayev, A., *Statistics of Random Processes*, Vol. 2, 2 edn, Springer-Verlag, New York, 1977.

[36] Lototsky, S. V., Mikulevicius, R. and Rozovskii, B. L., 'Nonlinear filtering revisited: A spectral approach', *SIAM J. Control Optim.* **35**, 435–461, 1997.

[37] Lototsky, S. V. and Rozovskii, B. L., 'Recursive multiple Wiener integral expansion for nonlinear filtering of diffusion processes'. *Stochastic Processes and Functional Analysis*, J. Goldstein et al. eds., 199-208. Lecture Notes in Pure and Applied Math 186. New York: Marcel Dekker, 1997.

[38] Lototsky, S. V. and Rozovskii, B. L., 'Recursive nonlinear filter for a continuous-discrete time model: Separation of parameters and observations', *IEEE Transactions on Automatic Control* **43**, 1154–1158, 1998.

[39] Ocone, D., 'Multiple integral expansions for nonlinear filtering', *Stochastics* **10**, 1–30, 1983.

[40] Ogura, H., 'Orthogonal functionals of the Poisson process', *IEEE Transactions on Information Theory* **18**, 473–481, 1972.

[41] Pacurar, M., 'Autoregressive conditional duration models in finance: A survey of the theoretical and empirical literature', *Journal of Economic Surveys* **22**, 711–751, 2008.

[42] Protter P., *Stochastic Integration and Differential Equations*, Springer, 1990.

[43] Rozovskii, B. L., *Stochastic Evolution Systems*, Kluwer Academic Publishers, 1990.

[44] Stoll, H., 'Friction', *Journal of Finance* **55**, 1479–1514, 2000.

[45] Synder, D. L., 'Filtering and detection for doubly stochastic Poisson processes', *IEEE Transactions on Information Theory* **18**, 91–102, 1972.

[46] Zeng Y., 'A partially-observed model for micro-movement of asset prices with Bayes estimation via filtering', *Mathematical Finance* **13**, 411-444, 2003.

[47] Zeng Y., 'Bayesian inference via filtering for a class of counting processes: application to the micro-movement of asset price', *Statistical Inference for Stochastic Processes* **8**, 331-354, 2005.

[48] Zhang, L., Mykland, P. A. and Aït-Sahalia, Y., 'A tale of two time scales: Determining integrated volatility with noisy high frequency data', *Journal of the American Statistical Association* **100**, 1394-1411, 2005.