



Working Paper Series  
Department of Economics  
University of Verona

Continuous time filtering for a class of marked doubly stochastic  
Poisson processes

Marco Minozzo

WP Number: 23

December 2011

ISSN: 2036-2919 (paper), 2036-4679 (online)

# Continuous Time Filtering for a Class of Marked Doubly Stochastic Poisson Processes

Silvia Centanni<sup>a</sup>, Marco Minozzo<sup>a,\*</sup>, Paola Tardelli<sup>b</sup>

<sup>a</sup>*Department of Economics, University of Verona*

<sup>b</sup>*Department of Electrical Engineering, University of L'Aquila*

---

## Abstract

We model a sequence of events by using a class of marked doubly stochastic Poisson processes where the intensity is given by a generalization of the classical shot noise process, specified as a positive function of another non-explosive marked point process. To filter the unobservable intensity, a time recursion is constructed to characterize a sequence of filtering distributions, that is, the conditional distributions of the intensity, given the past observations, evaluated at opportunely chosen time instants. To approximate this sequence, we consider a discrete approximation with random support by implementing a particle filter, in which we draw recursively from each filtering distribution. In the case in which the pair formed by the marked point process and by the intensity is a Markov process, this filtering recursion can be related to the classical filtering theory.

*Keywords:* Cox process, Marked point process, Particle filtering, Sequential Monte Carlo method, Shot noise process

*2010 MSC:* 62M20, 60G55, 65C05

---

## 1. Introduction

Let us consider a sequence of  $N$  events characterized by the values  $Z_1, \dots, Z_N$  assumed by a certain quantity of interest and by the times  $T_1, \dots, T_N$  at which they took place. Data of this kind can be naturally modeled as marked point

---

\*Department of Economics, University of Verona, Via dell'Artigliere 19, 37129 Verona, Italy

*Email address:* marco.minozzo@univr.it (Marco Minozzo)

processes (MPP) which can be seen as sequences of random times where each time is complemented with a random vector (the mark) taking values in some measurable space. In this setting, unmarked point processes are recovered as a particular case by letting the mark space to contain only one element. An important subclass of these processes is the class of marked doubly stochastic Poisson processes (DSPP) which are characterized by having the number  $N(s, t)$  of events in any given time interval  $(s, t]$  to be conditionally Poisson distributed, given another positive stochastic process  $\lambda$  called intensity (see Brémaud 1980; Cox and Isham 1980; Last and Brandt 1995). Given  $\lambda$ , the number of events in disjoint time intervals are also conditionally independent. In this article, as in Centanni and Minozzo (2006b) and in Gerardi and Tardelli (2006), we consider a class of marked DSPP in which the intensity process is a given function  $h$  of a non-explosive positive jump process that we characterize through the distributions of jump times and sizes. Such an intensity process can be viewed as a generalization of the classical shot noise process, as defined in Cox and Isham (1980). Marked DSPP with shot noise intensity have found applications in many fields such as quantum electronics (Teich and Saleh 2000), insurance (Dassios and Jang 2003) and finance (Duffie and Garleanu 2001).

Marked point processes have recently found a prominent natural application in the modeling of ultra-high-frequency (UHF) financial data. Since the nineties, UHF data, also known as tick-by-tick data, have been available for most exchange markets of financial assets such as the New York Stock Exchange, the Paris Bourse and the Frankfurt Stock Exchange. Moreover, for currency trading, the Swiss consultancy Olsen & Associates has now collected several years of such data in its databases. Whereas standard financial databases usually provide information on a daily or weekly basis (usually, for stocks, the closing prices and traded volumes), UHF databases provide a much finer information, recording the time at which each market event, such as a trade or a quote update by a market maker, takes place, together with its characteristics, for instance, the price and volume of the transaction (see Guillaume et al. 1999). In the literature, two main classes of models for MPP have so far been proposed for UHF data: the class based on the autoregressive conditional duration (ACD) models of Engle and Russel (1998) and the broad class based on DSPP with marks (see, for example, Frey 2000; Frey and Runggaldier 2001; Rydberg and Shephard 2000; Rogers and Zane 1998). In this latter class, Frey and Runggaldier (2001) allow the logprice process of a financial asset to change only at some random times  $T_1 < T_2 < \dots < T$ ,

where it equals the value of a ‘shadow’ logprice process modeled as a diffusion driven by a standard Brownian motion, and where the diffusion coefficient is influenced by an unobserved Markovian state process. In turn the random times  $T_n$  are modeled as the jump times of a counting process whose intensity depends on the level of the state process and is interpreted as the rate at which important news are absorbed by the market. In the same class of models, Rydberg and Shephard (2000) model the times of trades as a DSPP with unknown Markovian intensity.

A central problem faced when modeling with these kind of DSPP processes is the filtering of the underlying, and typically unobservable, intensity  $\lambda$ . This is a nonlinear problem which does not admit an explicit solution, and which, in similar circumstances, is usually tackled by approximation or simulation techniques. Whereas in Frey and Runggaldier (2001) it is proposed a filtering technique based on the reference probability method, in Rydberg and Shephard (2000) it is proposed a particle filtering method based on the auxiliary sampling importance resampling algorithm. A different solution is proposed in Centanni and Minozzo (2006b), where the intensity process has a particular form which can be interpreted in terms of the effect that major unexpected economic news have on the price formation process (see, among others, Kalev, Liu, Pham and Jarnecic 2004). There, the filtering problem, that is, the problem of finding the conditional distribution of the whole intensity, in a given time interval, given a realized trajectory of the DSPP  $(T_n, Z_n)_{n \in \mathbf{N}}$ , is solved by a simulation algorithm based on the reversible jump Markov chain Monte Carlo technique.

In this article, to solve this nonlinear filtering problem for a model similar to the one in Centanni and Minozzo (2006b), a time recursion is constructed to characterize the filtering distribution at certain time instants in terms of past filtering distributions. Then, to approximate this filter, we draw recursively samples from each filtering distribution. These samples can be viewed, in the optic of particle filters (see, for example, Doucet, De Freitas and Gordon 2001), as discrete approximations of the distributions of interest. From these approximate distributions, other quantities of interest can be derived such as the conditional expectation of the intensity, given the observations, at any given time instant. This filtering method does not require the intensity to be a Markov process, however, in the case in which the pair formed by the DSPP  $(T_n, Z_n)_{n \in \mathbf{N}}$  and by the intensity  $\lambda$  is indeed a Markov process, this method has interesting relations with the classical filtering theory (see, for example, Jazwinski 1970). In this case, under some additional con-

ditions, it can be shown that the recursive filtering distribution of  $\lambda$  is the (unique) solution of a suitable Kushner-Stratonovich equation and that the particle filter is strongly related to the Feynman-Kač representation of the unnormalized version of the equation.

The paper is organized as follows. In Section 2 we present our modeling framework, that is, a class of DSPP with intensity driven by MPP. Then in Section 3 we derive the recursion of filtering distributions, which is approximated by the particle filter in Section 3.3. In Section 4 we present some simulation studies, whereas in Section 5 we give some conclusions. We discuss our filtering algorithm in relation with the classical filtering theory in Appendix A

## 2. Marked DSPP with intensity driven by MPP

On a filtered probability space  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ , let  $\Phi = (T_n, Z_n)_{n \in \{0, 1, 2, \dots\}}$  be an adapted MPP where  $T_n$  are positive random variables such that  $T_0 = 0$  and  $T_n < +\infty \Rightarrow T_n < T_{n+1}$ ; and  $Z_n$  are  $\mathcal{Z}$ -valued random variables such that  $Z_0 = 0$ , where  $\mathcal{Z}$  is a measurable subset of  $\mathbf{R}$ . Let us also denote with  $N$  the counting process defined by  $N_t = \sum_{n \in \{0, 1, 2, \dots\}} \mathbb{1}_{\{T_n \leq t\}}$ , where  $\mathbb{1}_A$  is the indicator function of the set  $A$ . In a financial context, for instance, the random variables  $T_n$  and  $Z_n$  could represent the time and the size of the  $n$ th logreturn change and  $N_t$  represent the number of changes occurred up to time  $t$ . Let us notice that, as in Last and Brandt (1995), we identify an MPP with the random counting measure that it induces on  $\mathbf{R}^+ \times \mathcal{Z}$  by  $\Phi(C) = \sum_{n \geq 1} \mathbb{1}_{\{(T_n, Z_n) \in C\}}$ , for  $C \in \mathcal{B}(\mathbf{R}^+) \otimes \mathcal{B}(\mathcal{Z})$ . In this optic,  $N_t = \Phi((0, t] \times \mathcal{Z})$ . On the other hand, an MPP can also be seen as a mapping from  $\Omega$  to the space  $N_{\mathcal{Z}}$  of sequences of the type  $\varphi = (t_n, z_n)_{n \in \{0, 1, 2, \dots\}}$  such that  $z_n \in \mathcal{Z}$  and  $t_n < +\infty \Rightarrow t_n < t_{n+1}$ .

On the above filtered probability space, let us also introduce another MPP  $\Psi = (\tau_j, X_j)_{j \in \{0, 1, 2, \dots\}}$ , with  $\tau_0 = 0$ , where  $X_j$ , for all  $j \in \{0, 1, 2, \dots\}$ , take values in a measurable subset  $\mathcal{X} \subseteq \mathbf{R}$ ; we let also  $N'_t = \sum_{j \in \{0, 1, 2, \dots\}} \mathbb{1}_{\{\tau_j \leq t\}}$ . For  $\Psi$  we assume that for each  $j \in \{0, 1, 2, \dots\}$ ,  $\mathbb{P}(\tau_j < +\infty) = 1$ ,  $\tau_j < \tau_{j+1}$ , and that  $\lim_{j \rightarrow \infty} \tau_j = +\infty$  (that is, that  $\Psi$  is non explosive).

In the following we will assume that the filtration  $\{\mathcal{F}_t\}_{t \geq 0}$  is given by  $\mathcal{F}_t = \sigma(\Psi, \Phi_t)$ , where we denote with  $\Phi_t$  the restriction of  $\Phi$  to  $[0, t] \times \mathcal{Z}$ , and also that  $\Phi$  is a marked DSPP with respect to  $\{\mathcal{F}_t\}_{t \geq 0}$  (see Last and Brandt 1995, Chapter 6). This means that there exists an  $\mathcal{F}_0$ -measurable random

measure  $\nu$  on  $\mathbf{R}^+ \times \mathcal{Z}$  such that

$$\mathbb{P}(\Phi((s, t] \times A) = k | \mathcal{F}_s) = \frac{(\nu((s, t] \times A))^k}{k!} e^{-\nu((s, t] \times A)},$$

almost surely, for every  $s < t$  and for  $A \in \mathcal{B}(\mathcal{Z})$ , and that  $\nu$  is an  $\{\mathcal{F}_t\}$ -compensator of  $\Phi$ . It is implicit from the definition that  $\nu$  is an  $\{\mathcal{F}_t\}$ -compensator of  $\Phi$ , that is, an  $\{\mathcal{F}_t\}$ -predictable random measure such that

$$\mathbb{E} \left( \int_0^t \int_{\mathbf{R}} f(s, z) \Phi(ds, dz) \right) = \mathbb{E} \left( \int_0^t \int_{\mathbf{R}} f(s, z) \nu(ds, dz) \right), \quad (1)$$

for all predictable  $f : \Omega \times \mathbf{R}^+ \times \mathcal{Z} \rightarrow \mathbf{R}$ . Also, the process  $\Phi$  has a finite number of points in bounded intervals and no fixed point of jump, and the compensator  $\nu$  admits the disintegration  $\nu(dt, dz) = \bar{\nu}(dt)K(t, dz)$ , where  $\bar{\nu}(\cdot) = \nu(\cdot \times \mathbf{R})$  and  $K$  is an  $\mathcal{F}_0 \otimes \mathcal{B}(\mathbf{R}^+)$ -measurable stochastic kernel from  $(\Omega \times \mathbf{R}^+)$  to  $\mathcal{Z}$  (see Last and Brandt 1995, Appendix A2).

To completely characterize our framework, let us assume that  $\Lambda_t = \bar{\nu}((0, t])$  has the form  $\Lambda_t = \int_0^t \lambda_s ds$ . Then the counting process  $N$  is a DSPP with intensity  $\lambda$  and, given the whole history of  $\lambda$ , the number of points in any time interval  $(s, t]$  is Poisson distributed (independently of  $\mathcal{F}_s$ ) with mean  $\Lambda_t - \Lambda_s$ . In particular,

$$\mathbb{P}(T_{N_s+1} > t | \mathcal{F}_s) = \exp \left( - \int_s^t \lambda_u du \right) \quad (2)$$

and

$$\mathbb{P}(Z_{N_s+1} \in B | \mathcal{F}_s, T_{N_s+1}) = \int_B K(T_{N_s+1}, dz), \quad (3)$$

for all  $B \in \mathcal{B}(\mathcal{Z})$ . As far as the form of the intensity process is concerned, we will assume that the intensity  $\lambda$  is given by  $\lambda_t = h(t, \Psi_t)$ , where  $h(\cdot)$  is a measurable function from  $\mathbf{R}^+ \times N_{\mathcal{X}}$  to  $\mathbf{R}^+$ , where  $N_{\mathcal{X}}$  is the space of sequences of the type  $\psi = (\tau_j, x_j)_{j \in \{0, 1, 2, \dots\}}$  such that  $x_j \in \mathcal{X}$ ,  $\tau_j < +\infty$ ,  $\tau_j < \tau_{j+1}$ , and  $\lim_{j \rightarrow \infty} \tau_j = +\infty$ . In other words, this means that the intensity process can be written as a function of time and of the realization, up to time  $t$ , of another non-explosive MPP. This latter is said to drive the intensity process  $\lambda$  of the marked DSPP  $\Phi$ .

Although we are referring to a particular subclass of DSPP with marks, let us remark that the probability framework just outlined allows a wide range of models to be specified to account for different hypotheses on the

times and sizes of the intensity jumps, on the effect of these on the arrival mechanism of the events of the DSPP, as well as on its marks. For example, a particular specification for the intensity  $\lambda$ , which will be considered in later sections, could be

$$\lambda_t = h(t, \Psi_t) = a(t) + b \sum_{j=1}^{N'_t} X_j e^{-k(t-\tau_j)}, \quad (4)$$

where  $b$  and  $k$  are positive parameters and  $a(\cdot)$  is a deterministic integrable function which, in a financial context, might be used to model seasonal patterns. In particular, when  $a(\cdot) \equiv 0$  the function  $h$  defines a shot noise process (see Cox and Isham, 1980).

### 3. The filtering of the intensity process

Let us now assume that only the process  $\Phi$  is observable, that is, let us restrict ourselves to the filtration generated by the observations  $\mathcal{G}_t = \sigma(\Phi_t)$ , where  $\mathcal{G}_t \subseteq \mathcal{F}_t$ , for all  $t \geq 0$ . Indeed, with respect to our financial interpretation, it is reasonable to assume that market agents are restricted to observe only the price process of a financial asset, while the intensity process remains unobservable. The problem is then to compute the conditional distribution of the process  $\Psi$  driving the intensity  $\lambda$ , as we already noticed, at any time instant  $t$ , given the observations of  $\Phi$  up to the same time instant. In the case in which the pair formed by the DSPP  $(T_n, Z_n)_{n \in \mathbf{N}}$  and by the intensity  $\lambda$  is a Markov process, and under some additional conditions, it can be shown that the filtering problem can be solved using classical filtering techniques and that the filtering distribution is the solution of a suitable Kushner-Stratonovich equation (see Appendix A). On the other hand, here, without making the assumption of Markovianity, we tackle the filtering problem in a more general setting by resorting to sequential Monte Carlo filtering. Let us note that an approximate solution for this problem is unavoidable since the target conditional distribution does not have a closed form.

#### 3.1. Particle filters for discrete time processes

In the context of discrete time ( $i = 1, 2, \dots$ ) dynamical models, particle filters are iterative algorithms which, at each iteration  $i$ , provide a Monte Carlo approximation of a distribution  $\pi_i$  on a space of interest  $\Theta_i$ . This is achieved

by constructing at each iteration a set of weighted particles  $(\theta_i^m, w_i^m)_{m=1, \dots, M}$  for which, for each measurable function  $f$  the expected value  $E_{\pi_i}(f)$  is approximated by the quantity

$$\frac{\sum_{m=1}^M w_i^m f(\theta_i^m)}{\sum_{m=1}^M w_i^m},$$

in this case the set  $(\theta_i^m, w_i^m)_{m=1, \dots, M}$  is said to target  $\pi_i$  (for further details, see Chopin, 2004).

If at each time step  $i$  there exists a ‘prior density’  $q_i(\theta_i|\theta_{i-1})$ , the construction of an appropriate set of particles (with the corresponding weights) can be performed with the following three steps. Denoting with  $(\hat{\theta}_{i-1}^m, 1)$  a set of particles of unit weights targeting  $\pi_{i-1}$ :

- i) (mutation step) produce the set  $(\theta_i^m, 1)_{m=1 \dots M}$  of new weighted particles, each from the density  $q_i(\theta_i|\hat{\theta}_{i-1}^m)$ ,  $m = 1, \dots, M$ , which targets the distribution

$$\tilde{\pi}_i(\cdot) = \int \pi_{i-1}(\theta_{i-1}) q_i(\cdot|\theta_{i-1}) d(\theta_{i-1});$$

- ii) (correction step) Compute new weights  $w_i^m \propto v_i(\theta_i^m)$ , where

$$v_i(\theta_i^m) = \frac{\pi_i(\theta_i^m)}{\tilde{\pi}_i(\theta_i^m)},$$

obtaining a set  $(\theta_i^m, w_i^m)_{m=1, \dots, M}$  which targets  $\pi_i$ ;

- iii) (selection step) replace the current set of particles  $(\theta_i^m, w_i^m)_{m=1, \dots, M}$  with an uniformly weighted vector  $(\hat{\theta}_i^m, 1)_{m=1, \dots, M}$  containing a (random) number  $n_m \geq 0$  of replicates of particle  $\theta_i^m$ , such that

$$\sum_{m=1}^M n_m = M \quad \text{and} \quad E(n_m) = M \frac{w_i^m}{\sum_{m=1}^M w_i^m}.$$

This can be accomplished, for example, with a multinomial resampling, that is by drawing independently  $M$  particles from the set  $(\theta_t^m, w_t^m)_{m=1, \dots, M}$  seen as a multinomial distribution.

### 3.2. Recursion of filtering distributions for continuous time DSPP

Turning now to the setting described in Section 2, let us assume that  $K(t, dz)$  is given by  $K(t, dz) = k(\Psi_t, t, z)dz$ , where  $k$  is a measurable function from  $N_{\mathcal{X}} \times \mathbf{R}^+ \times \mathcal{Z}$  to  $\mathbf{R}^+$ . This means that, given any event time  $T_n$ , the distribution (admitting a density) of the corresponding mark  $Z_n$  is allowed to depend on the position in time of the mark, as well as on the MPP driving the intensity. In this case, the compensator of  $\Phi$  is absolutely continuous with respect to the Lebesgue measure on  $\mathbf{R}^+ \times \mathcal{Z}$ . Thus, for each  $n \geq 0$ , and for  $T_n \leq t < T_{n+1}$ , we can write:

$$\begin{aligned} \mathbb{P}(\Phi_t \in d\phi | \mathcal{F}_{T_n}) &= \mathbb{I}_{\{\Phi_{T_n} \in d\phi\}} \mathbb{P}(T_{n+1} > t | \Psi, \Phi_{T_n}) \\ &= \mathbb{I}_{\{\Phi_{T_n} \in d\phi\}} \exp\left(-\int_{T_n}^t h(s, \Psi_s) ds\right); \end{aligned} \quad (5)$$

$$\begin{aligned} \mathbb{P}(\Phi_{T_{n+1}} \in d\phi | \mathcal{F}_t) &= \mathbb{E}\left(\mathbb{I}_{\{\Phi_{T_{n+1}} \in d\phi\}} | \Psi, \Phi_t\right) \\ &= \iint \mathbb{I}_{\{s > t\}} \mathbb{I}_{\{\Phi_t + \delta_{(s,z)} \in d\phi\}} h(s, \Psi_s) k(\Psi_s, s, z) \cdot \exp\left(-\int_t^s h(u, \Psi_u) du\right) ds dz. \end{aligned} \quad (6)$$

Moreover, defining  $\mathcal{F}_{T_{n+1}}^- = \sigma(\Psi, \Phi_{T_{n+1}-})$ , we have

$$\begin{aligned} \mathbb{P}(\Phi_{T_{n+1}} \in d\phi | \mathcal{F}_{T_{n+1}}^-) &= \mathbb{E}\left(\mathbb{I}_{\{\Phi_{T_{n+1}} \in d\phi\}} | \Psi, \Phi_{T_{n+1}-}\right) \\ &= \int \mathbb{I}_{\{\Phi_{T_n} + \delta_{(T_{n+1}, z)} \in d\phi\}} h(T_{n+1}, \Psi_{T_{n+1}}) k(\Psi_{T_{n+1}}, T_{n+1}, z) dz. \end{aligned} \quad (7)$$

We can observe that from (5) and (6) it follows that

$$\mathbb{P}(\Phi_t \in d\phi | \mathcal{F}_{T_n}) = \mathbb{P}(\Phi_t \in d\phi | \Psi_t, \Phi_{T_n}), \quad (8)$$

and

$$\mathbb{P}(\Phi_{T_{n+1}} \in d\phi | \mathcal{F}_t) = \mathbb{P}(\Phi_{T_{n+1}} \in d\phi | \Psi_{T_{n+1}}, \Phi_t). \quad (9)$$

Using successive conditionings (see, for example, Last and Brandt 1995, Theorem A3.21) we obtain

$$\mathbb{P}((\Psi_t, \Phi_t) \in d(\psi, \phi) | \Psi_{T_n}, \Phi_{T_n}) = \mathbb{P}(\Psi_t \in d\psi | \Psi_{T_n}) \cdot \mathbb{P}(\Phi_t \in d\phi | \Psi_t, \Phi_{T_n}), \quad (10)$$

$$\begin{aligned} \mathbb{P}((\Psi_{T_{n+1}}, \Phi_{T_{n+1}}) \in d(\psi, \phi) | \Psi_t, \Phi_t) & \quad (11) \\ &= \mathbb{P}(\Psi_{T_{n+1}} \in d\psi | \Psi_t) \cdot \mathbb{P}(\Phi_{T_{n+1}} \in d\phi | \Psi_{T_{n+1}}, \Phi_t). \end{aligned}$$

Finally, by Bayes Theorem, using (8) and (9) and letting  $\mathbb{P}(\Psi_0 \in d\psi | \Phi_0) = \mathbb{P}(\Psi_0 \in d\psi)$ , we can write, for  $T_n \leq t < T_{n+1}$ , and for  $n \geq 1$ :

$$\begin{aligned} \mathbb{P}(\Psi_t \in d\psi | \Phi_t) &= \frac{\exp\left(-\int_{T_n}^t h(s, \psi) ds\right) \cdot \mathbb{P}(\Psi_t \in d\psi | \Phi_{T_n})}{\int \exp\left(-\int_{T_n}^t h(s, \chi) ds\right) \cdot \mathbb{P}(\Psi_t \in d\chi | \Phi_{T_n})}; \quad (12) \\ \mathbb{P}(\Psi_{T_{n+1}} \in d\psi | \Phi_{T_{n+1}}) &= \frac{h(T_{n+1}, \phi)k(\phi, T_{n+1}, Z_{n+1}) \mathbb{P}(\Phi_{T_{n+1}} \in d\phi | \Phi_t)}{\int h(T_{n+1}, \chi)k(\chi, T_{n+1}, Z_{n+1}) \mathbb{P}(\Phi_{T_{n+1}} \in d\chi | \Phi_t)}; \quad (13) \end{aligned}$$

$$\mathbb{P}(\Psi_t \in d\psi | \Phi_{T_n}) = \mathbb{P}(\Psi_{T_n} \in d\psi_{T_n} | \Phi_{T_n}) \cdot \mathbb{P}(\Psi_t \in d\psi | \Psi_{T_n} = \psi_{T_n}); \quad (14)$$

$$\mathbb{P}(\Psi_{T_{n+1}} \in d\psi | \Phi_t) = \mathbb{P}(\Psi_t \in d\psi_t | \Phi_t) \cdot \mathbb{P}(\Psi_{T_{n+1}} \in d\psi | \Psi_t = \psi_t). \quad (15)$$

### 3.3. Particle approximation

Though the above filtering distributions are not computable in closed form, these can be numerically evaluated by a particle filter. The idea is to approximate, for each  $n$ , the distributions in (12)–(15) by a set of particle distributions, by building a sequence of discrete distributions with random support, obtained by drawing recursively a sample (the set of particles) of size  $M$  from the distributions of interest. The drawings are made, for each  $n$ , at an arbitrarily chosen point between  $T_n$  and  $T_{n+1}$  (that could also be  $T_{n+1}^-$ , that is, the time instant ‘just before’  $T_{n+1}$ ), and at the jump time  $T_{n+1}$ . The algorithm can be described as follows. To start, we draw a sample  $\psi_0^1, \dots, \psi_0^M$  from  $\mathbb{P}(\Psi_0 \in d\psi | \Phi_0)$  and we treat  $\widehat{\mathbb{P}}(\Psi_0 \in d\psi | \Phi_0)$ , that is, the distribution assigning mass  $1/M$  to each particle  $\psi_0^m$ ,  $m = 1, \dots, M$ , as if it were the true distribution. Then, for each  $n = 1, \dots, N$ :

- i)* at an arbitrary time  $t \in (T_n, T_{n+1})$ , we sample  $\psi_t^m$  from  $\mathbb{P}(\Psi_t \in d\psi | \Psi_{T_n} = \psi_{T_n}^m)$ , for  $m = 1, \dots, M$ , and assign to  $\psi_t^m$  a weight  $W_t^m$  computed with (12);
- ii)* at  $T_{n+1}$ , we sample  $\psi_{T_{n+1}}^m$  from  $\mathbb{P}(\Psi_{T_{n+1}} \in d\psi | \Psi_t = \psi_t^m)$ , for  $m = 1, \dots, M$ , and assign to  $\psi_{T_{n+1}}^m$  a weight  $W_{T_{n+1}}^m$  computed with (13).

## 4. Simulation studies and application to UHF data

To give an illustration of the potentials of our particle filtering algorithm, we performed some simulation studies assuming the setting that is described in Section 2. Furthermore, we also considered an UHF data set relative to the prices of the S&P 500 index future.

### 4.1. A subclass of models used in applications

For our purposes, we consider here a subclass of models, which we call the basic class, defined by the following assumptions:

- A1. The intensity process has the form given in Equation (4), with  $a(\cdot) = 0$  and  $b = 1$ , that is,

$$\lambda_t = \sum_{j=1}^{N'_t} X_j e^{-k(t-\tau_j)}.$$

- A2.  $N'$  is a Poisson process with constant intensity  $\nu$ .
- A3.  $X_j$  are independently exponentially distributed with mean  $1/\gamma$ .
- A4.  $\lambda_0$  has a Gamma distribution with parameters  $\nu/k$  and  $\gamma$  (that is,  $E(\lambda_0) = \nu/(k\gamma)$ ).
- A5.  $Z_i$  are independently and identically distributed.

For the evaluation of the sequential Monte Carlo algorithm, specified in Section 3.3, for the filtering of the intensity process we run, under the basic class, some simulation experiments with the following sets of parameters:  $k = .0074$ ,  $\gamma = 2.25$ ,  $\nu = 1/60$ ; and  $k = .0011$ ,  $\gamma = 30$ ,  $\nu = 1/120$ . In a financial context, these parameter values might correspond to different hypotheses regarding the speed of transaction (or quotation changes) of the asset and the variability of the intensity process. We chose a time horizon of  $T = 2,400$  minutes corresponding to one week of market activity (8 hours per day for 5 days).

#### 4.2. Simulation study

For each set of parameters, we run the sequential Monte Carlo algorithm on a simulated trajectory of jump times  $T_n$  (times of price change), obtained by simulating a trajectory of  $\Psi$  and  $\Phi$ . Recall that, since in the basic class the marks (logreturn changes) are independent from  $\lambda$ , the set of jump times contains all the relevant information. For these runs we chose  $M = 10000$  particles for each  $n = 1, \dots, N$ , where  $N = 2478$  and  $N = 545$ , for the first and second parameter sets, respectively. The filtering expectations have been evaluated with the arithmetic means  $\sum_{m=1}^M h(T_n, \psi_{T_n}^m) \cdot W_{T_n}^m$ , for each  $n = 1, \dots, N$ . By independently replicating for a few times the filtering algorithm on the two simulated trajectories of jump times  $T_n$ , we found that the number of particles considered was large enough to guarantee a very good approximation to the true conditional expectation  $E(\lambda_{T_n} | \Phi_{T_n})$  of the intensity. Indeed, over different replicates we practically obtained the same simulated filtering expectations. Figure 1 represents, for the two sets of parameters, the simulated trajectory of the intensity (grey line), used to simulate the jump times  $T_n$ , and the corresponding simulated filtering expectations (black line), given the observations up to time  $T_n$ . As it can be observed, for the first set of parameter values the conditional expectation is much closer to the realized trajectory of the intensity.

To study the sampling variability of the particle filter, we independently run the algorithm 50 times with a much smaller number of particles. For each of the two sets of parameters:

1. For the same simulated trajectory of the intensity and of jump times  $T_n$  obtained above, we considered 50 independent runs of the filtering algorithm with  $M = 1000$  particles.
2. For each run  $j$ ,  $j = 1, \dots, 50$ , we calculated the arithmetic means

$$\widehat{E}(\lambda_{T_n} | \Phi_{T_n})^{(j)} = \sum_{m=1}^{1000} h(T_n, \psi_{T_n}^{m,j}) \cdot W_{T_n}^{m,j},$$

for each  $n = 1, \dots, N$ .

Figure 2 shows the conditional expectations (black lines)  $E(\lambda_{T_n} | \Phi_{T_n})$  obtained with  $M = 10000$  particles, together with the 50 independent replicates of the algorithm, each based on  $M = 1000$  particles (grey lines).

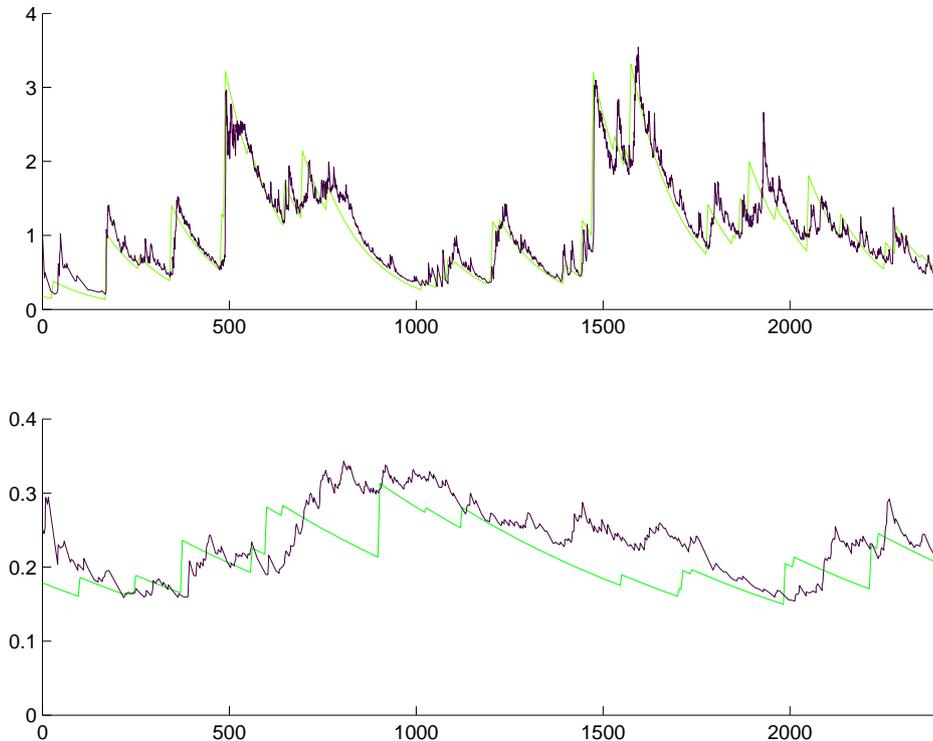


Figure 1: Filtering of the intensity by sequential Monte Carlo. True (simulated) trajectories of the intensity process  $\lambda$  (grey lines) and corresponding filtering expectations  $E(\lambda_{T_n}|\Phi_{T_n})$  of the intensity (black lines), for the sets of parameters:  $k = .0074$ ,  $\gamma = 2.25$ ,  $\nu = 1/60$  (top);  $k = .0011$ ,  $\gamma = 30$ ,  $\nu = 1/120$  (bottom). The simulated filtering expectations were obtained using  $M = 10000$  particles for each  $n = 1, \dots, N$ , where  $N = 2478$  (top) and  $N = 545$  (bottom).

Figure 3 shows the mean square error (MSE) of the above arithmetic means  $\widehat{E}(\lambda_{T_n}|\Phi_{T_n})^{(j)}$  as estimates of  $E(\lambda_{T_n}|\Phi_{T_n})$ , computed as

$$\text{MSE}(T_n) = \sqrt{\frac{1}{50} \sum_{j=1}^{50} \left( \widehat{E}(\lambda_{T_n}|\Phi_{T_n})^{(j)} - E(\lambda_{T_n}|\Phi_{T_n}) \right)^2},$$

for each  $n = 1, \dots, N$ , where  $E(\lambda_{T_n}|\Phi_{T_n})$  are the simulated filtering expectations obtained with  $M = 10000$ , that we consider as the true conditional expectations. It can be observed that for the two sets of parameter values, the MSE does not increase with time.

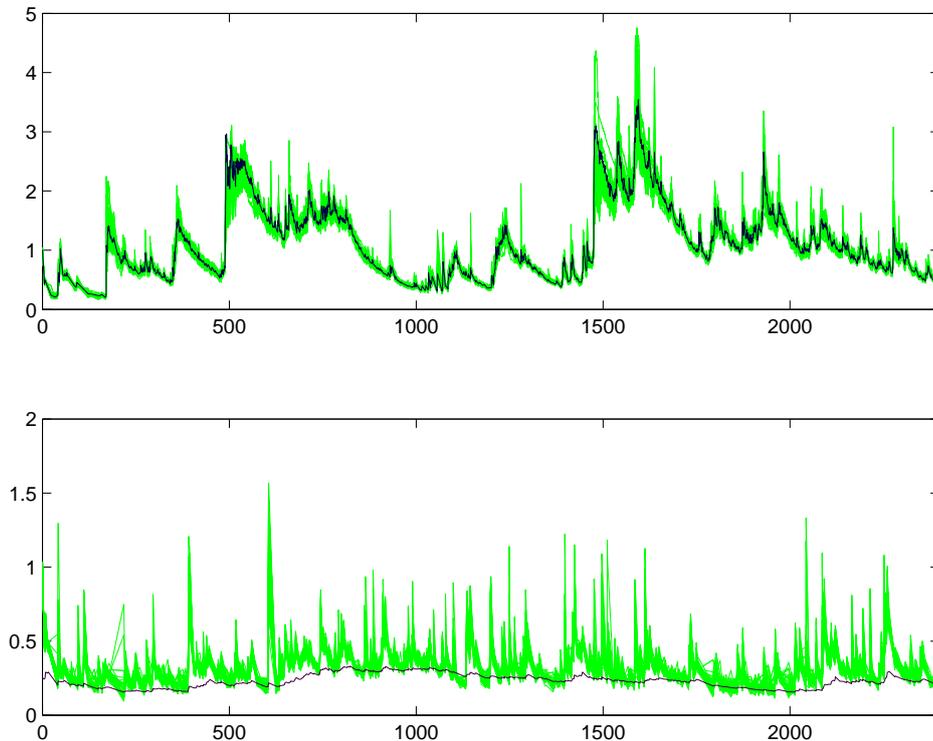


Figure 2: Variability of the sequential Monte Carlo filter. The black lines are the conditional expectations  $E(\lambda_{T_n}|\Phi_{T_n})$  obtained with  $M = 10000$  particles. The grey lines have been obtained running 50 independent replicates of the algorithm, each based on  $M = 1000$  particles. The two plots are relative to the sets of parameters:  $k = .0074$ ,  $\gamma = 2.25$ ,  $\nu = 1/60$  (top);  $k = .0011$ ,  $\gamma = 30$ ,  $\nu = 1/120$  (bottom).

#### 4.3. Case study: the S&P 500 index future

We consider here an UHF data set relative to the prices of the S&P 500 index future (SPU01). We consider all price changes (39,889) from the 9th to the 27th of July 2001 (15 trading days). The market was open from 8.30 in the morning to 15.15 in the afternoon. The data set shows successive logreturns not to be autocorrelated and also not correlated with the waiting times  $T_n - T_{n-1}$ . Moreover, the time between successive price changes shows to be exponentially distributed.

To account for the intraday seasonality in the data, we consider an intensity process given by Equation (4), in which the deterministic part is given by

$$a(t) = a_0 + a_1 t + a_2 t^2,$$

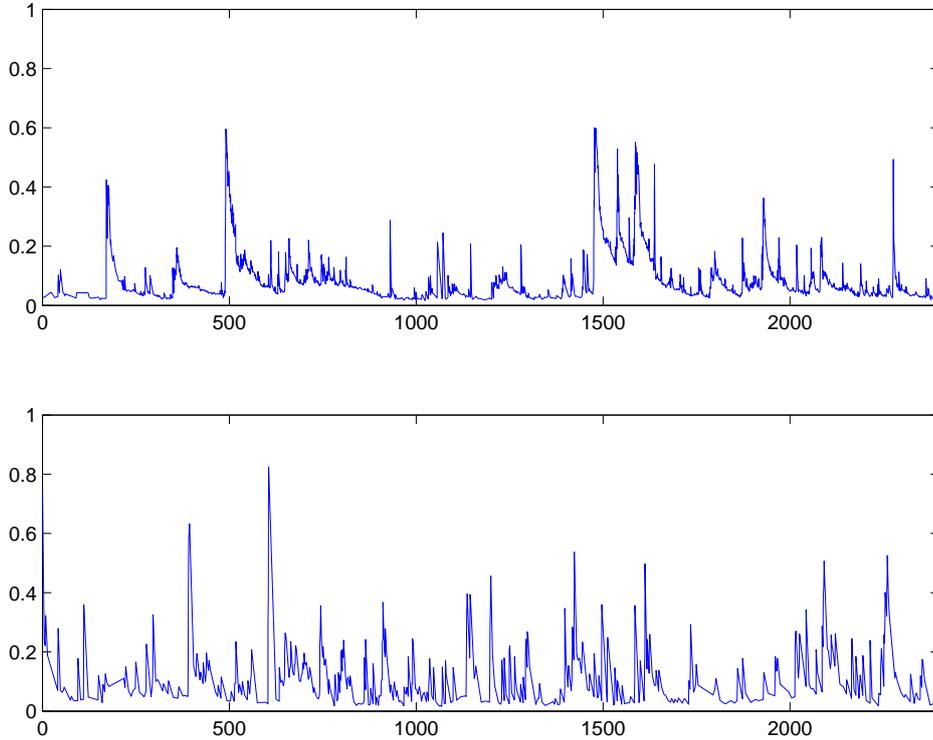


Figure 3: Mean square error of  $\widehat{\mathbf{E}}(\lambda_{T_n}|\Phi_{T_n})$ . MSE of the particle filtering computed over 50 independent replicates of the algorithm, each based on  $M = 1000$ . The two plots are relative to the sets of parameters:  $k = .0074$ ,  $\gamma = 2.25$ ,  $\nu = 1/60$  (top);  $k = .0011$ ,  $\gamma = 30$ ,  $\nu = 1/120$  (bottom). The value of the MSE in the two plots does not diverge with time.

where  $b = 1$ , and where the stochastic part of  $\lambda_t$  is assumed to satisfy Assumptions A2–A5 of the basic class.

Let us note that the chosen shape for the deterministic part of the intensity is standard in UHF data analysis. For the parameters of this part we set  $a_0 = 6.7226$ ,  $a_1 = -.0306$ ,  $a_2 = .0001$ , whereas point estimates of  $k$ ,  $\gamma$  and  $\nu$  are provided by a running of the stochastic EM algorithm, giving  $\hat{k} = .006$ ,  $\hat{\gamma} = 2.875$  and  $\hat{\nu} = .050$  (see Centanni, Minozzo (2006b) and Minozzo, Centanni 2008).

The filtering of the intensity was performed by running the particle filtering algorithm assuming the above estimates as the true parameter values. The graph at the top of Figure 4 shows the filtering expectation of the stochastic part of the intensity, for each time  $T_n$  of price change. The graph

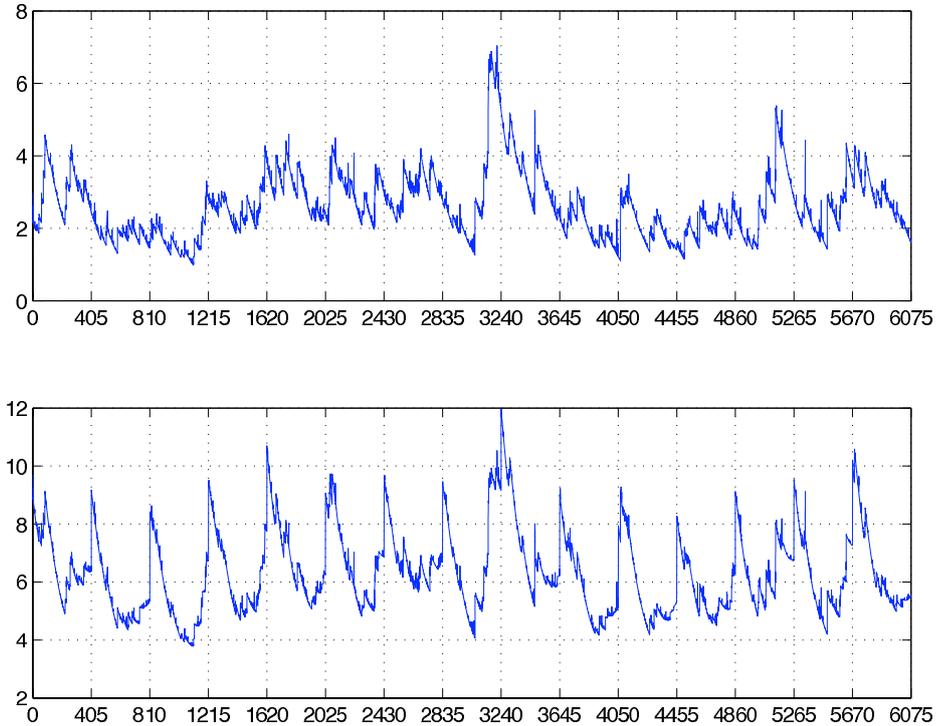


Figure 4: Filtering of the intensity with the particle filtering algorithm for the S&P 500 data set. The graph at the top shows the filtering expectation of the stochastic part of the intensity, for each time  $T_n$  of price change, under the model described in the text, assuming the parameter values:  $k = .006$ ,  $\gamma = 2.875$ ,  $\nu = .050$ . The graph at the bottom shows the total (filtered) intensity comprehensive also of the deterministic part.

at the bottom shows the total (filtered) intensity comprehensive also of the deterministic part. The peaks in the top graph show periods of high trading activity, not accounted for by the intraday seasonality, which may correspond to the perturbations in the market due to the arrivals of relevant news.

## 5. Conclusions

In this paper we proposed a sequential Monte Carlo method for the filtering of DSPP with intensity driven by MPP. Traditionally particle filters have been developed for discrete time stochastic processes, in particular for Markovian state-space processes. Some applications of these filters to continuous time processes, such as stochastic volatility models and diffusions, have

been considered, avoiding only in recent works time discretization (see, for example, Fearnhead, Papaspiliopoulos, Roberts (2006) and references therein). Here we considered a filtering algorithm for DSPP that does not require discretization nor Markovianity, and that can be easily implemented, exploiting the particular structure of the intensity.

Although we compared the particle filtering algorithm to the classical filtering theory, the former can be implemented even in situations where the classical filtering theory does not apply. Let us observe that if we simulate the filter from the Feynman-Kač representation (when the stated hypotheses are satisfied and the representation does exist) we obtain the same simulation algorithm as that described in Section 3.3, that is, we obtain the same particle filter.

Though we have not given any theoretical result about the performance of the algorithm, the simulation studies gave very good indications even for a moderate number of particles. For the models considered, the variability of the algorithm resulted very modest and the MSE did not diverge with time. Future work should involve the study of a central limit theorem for particle filters for continuous time DSPP.

## Appendix A. Classical filtering theory

In the Markovian case, the filtering problem just outlined has interesting relations with the classical filtering theory. To discuss this aspect, let us assume that  $\mathcal{X} \subseteq \mathbf{R}^+$  and that

$$\lambda_t = h(t, \Psi_t) = a(t) + b \sum_{j=1}^{N'_t} X_j e^{-k(t-\tau_j)}, \quad (\text{A.1})$$

where  $b$  and  $k$  are positive parameters and  $a(\cdot)$  is a deterministic integrable function.

Let us here also assume that  $a(\cdot)$  is such that  $0 \leq a(t) \leq \bar{a} < +\infty$ , that  $J_t = \sum_{j=1}^{N'_t} X_j$  is a Markov process taking its values in a finite subset  $\mathcal{J}$  of  $\mathbf{R}^+$ ,

and that  $Y_t = \sum_{n=1}^{N_t} Z_n$  takes its values in a countable subset  $\mathcal{Y}$  of  $\mathbf{R}$ . From the Markovianity assumption, the dynamics of the process  $J$  is described by

the operator

$$L^J f(x) = \lambda^J(x) \sum_{x' \in \mathcal{J}} [f(x') - f(x)] \mu^J(x, x'),$$

for every  $x \in \mathcal{J}$ , and for  $f \in B(\mathcal{J})$ , the space of the real-valued bounded measurable functions, where  $\mu^J(x, x')$  is a transition function on  $\mathcal{J} \times \mathcal{J}$  and  $\lambda^J(x)$  is a nonnegative measurable function on  $\mathcal{J}$ , necessarily bounded. Thus, the process  $J$  is the unique solution of the *martingale problem* (see Ethier, Kurz, 1986) for the generator  $L^J$  with initial condition  $J_0 = 0$  (for more details see Gerardi and Tardelli (2006) and references therein).

As regard as the filtering problem, our aim is to find the law of  $\lambda_t$  given the history of the observed process  $Y$ . To this end, let us introduce the process  $Z_t = e^{kt} \lambda_t$ , and observe that

$$Z_t = \lambda_0 + \sum_{j \geq 0} X_j e^{kX_j} \mathbb{1}_{\{\tau_j \leq t\}} = \lambda_0 + \int_0^t e^{ks} dJ_s.$$

In other words, the process  $Z$ , taking values in  $\mathcal{Z} \subset \mathbb{R}^+$ , is a non homogeneous pure jump process with the same jump times of  $J$  and the marks of  $Z$  are deterministic functions of the marks of  $J$ , that is,

$$Z_t - Z_{t-} = e^{kt}(J_t - J_{t-}).$$

Moreover,  $\lambda$ , given by (A.1), is completely determined by  $Z$ , in fact  $\lambda_t = a(t) + be^{-kt} Z_t$ . Thus, the intensity  $\lambda$  is just a measurable function  $\lambda(t, Z_t)$  of the time  $t$  and of the process  $Z$ , and, since the dynamics of  $Z$  is driven by  $J$ , the conditional law of  $\lambda_t$  given the history of  $Y$  reduces to the conditional law of  $(J_t, Z_t)$  given the history of  $Y$ . Also, since the intensity of  $Y$  depends on  $Z$ , necessarily, the processes  $J$  and  $Y$  are not independent and  $Y$  cannot be Markovian. On the other hand, since the hidden process  $J$  is Markovian, necessarily we get that jointly  $(J, Z)$  is Markovian. Now, assuming that jointly  $(J, Z, Y)$  is Markovian and that the processes  $J$  and  $Y$  cannot have common jump times, the joint generator of  $(J, Z, Y)$ , given  $F(t, x, z, y)$ , for  $t \geq 0$ ,  $x \in \mathcal{J}$ ,  $z \in \mathcal{Z}$  and  $y \in \mathcal{Y}$ , belonging to a suitable class of real-valued functions, is given by

$$L^{J,Z,Y} F(t, x, z, y) = \frac{\partial}{\partial t} F(t, x, z, y) + L_t^{J,Z,Y} F(t, x, z, y),$$

where

$$\begin{aligned} L_t^{J,Z,Y} F(t, x, z, y) &= \lambda^J(x) \sum_{x' \in \mathcal{J}} [F(t, x', z + e^{kt}(x' - x), y) - F(t, x, z, y)] \mu^J(x, x') \\ &\quad + \lambda(t, z) \sum_{y' \in \mathcal{Y}} [F(t, x, z, y') - F(t, x, z, y)] \mu^Y(y, y'), \end{aligned}$$

and  $\mu^Y(y, y')$  is a suitable transition function according to the definition of  $Y$ . So, since under the assumptions of this section  $L_t^{X,Z,Y}$  is bounded, existence and uniqueness of the *martingale problem* associated with the operator  $L^{J,Z,Y}$ , given the initial condition  $(t_0, x_0, z_0, y_0)$ , can be achieved by applying Theorem 7.3 in Ethier and Kurtz (1986).

Let us note that the generator of  $L^{J,Z,Y}$  restricted to a function  $F$  just depending of  $x \in \mathcal{J}$  and  $z \in \mathcal{Z}$  reduces to

$$L_t^{J,Z} F(x, z) = \lambda^J(x) \sum_{x' \in \mathcal{J}} [F(x', z + e^{kt}(x' - x)) - F(x, z)] \mu^J(x, x').$$

In this case it can be shown that the cadlag version of the filter (see also Brémaud 1980, Appendix 3)  $\pi_t(F) = E[F(J_t, Z_t) | \mathcal{F}_t^Y]$ , for all  $F \in B(\mathcal{J} \times \mathcal{Z})$ , satisfies a stochastic differential equation known as Kushner-Stratonovich equation.

To solve our filtering problem, we note that, in general,  $Y$  is not a counting process, and thus, following an idea presented in Calzolari and Nappo (2001), we introduce the multivariate point processes  $U = (U^1, U^2, \dots)$  defined as

$$U_t^y = \sum_{n \geq 1} \mathbb{1}_{\{T_n \leq t\}} \mathbb{1}_{\{Y_{T_n} = y\}}, \quad y \in \mathcal{Y}.$$

Recalling that  $\mathcal{Y}$  is a countable set and that  $T_1, T_2, \dots$  are the jump times of the process  $Y$ , by definition, for all  $y \in \mathcal{Y}$ ,  $U^y$  is a process taking only non negative integer values; more precisely,  $U^y$  counts the number of jumps bringing  $Y$  to  $y$ . Moreover, since  $N$  is the process counting the jumps of  $Y$ , we get that  $N_t = \sum_{n \geq 0} \mathbb{1}_{\{T_n \leq t\}} = \sum_{y \in \mathcal{Y}} U_t^y$ . Furthermore, the relation

$$Y_t = Y_0 + \int_0^t \sum_{y \in \mathcal{Y}} [y - Y_{s-}] dU_s^y \quad (\text{A.2})$$

(where just one term of the integrand is not null, almost surely) implies that  $\mathcal{F}_t^Y = \mathcal{F}_t^U = \sigma\{U_s^1, U_s^2, \dots, s \leq t\}$ . So, our problem reduces to find the conditional law of  $(J_t, Z_t)$  given  $\mathcal{F}_t^U$ . For the model presented at the beginning of this section, the Kushner-Stratonovich equation can be written as

$$\begin{aligned} \pi_t(F) &= \pi_0(F) \\ &+ \int_0^t \{\pi_s(L_s^{J,Z} F) + \pi_s(\lambda(s, \cdot) F) - \pi_s(\lambda(s, \cdot)) \pi_s(F)\} ds \\ &+ \int_0^t \pi_{s-}(\lambda(s, \cdot))^+ \{\pi_{s-}(\lambda(s, \cdot) F) - \pi_{s-}(\lambda(s, \cdot)) \pi_{s-}(F)\} dN_s, \end{aligned} \quad (\text{A.3})$$

and can be shown to have a unique strong solution (see Gerardi and Tardelli 2006). At any jump time  $T_n$ , the filter is uniquely determined by  $\pi_{T_n-}$ , namely, it results that at any jump time  $T_n$ , the filter is given by

$$\pi_{T_n}(F) = \frac{\pi_{T_n-}(\lambda(T_n, \cdot) F)}{\pi_{T_n-}(\lambda(T_n, \cdot))}; \quad (\text{A.4})$$

on the other hand, between two consecutive jump times the filter is the solution of a Lipschitz equation which is difficult to handle. It can be shown (see Gerardi and Tardelli 2006) that there exists a finite positive measure  $\rho_t^n$  such that  $0 < \rho_t^n(1) < 1$  and  $\frac{\rho_t^n(F)}{\rho_t^n(1)}$  coincides with the filter  $\pi_t(F)$ , for all  $t \in [T_n, T_{n+1})$ , which admits the Feynman-Kač representation

$$\rho_t^n(\lambda(t, \cdot)) = \int_{\mathcal{J}} \int_{\mathcal{Z}} \mathbb{E}^{P_{s,x,z}} \left[ \lambda(t, Z_t) \exp \left\{ - \int_s^t \lambda(v, Z_v) dv \right\} \right] \Big|_{s=T_n} \pi_{T_n}(dx, dz). \quad (\text{A.5})$$

Choosing  $F = \mathbb{1}_{\{\lambda(t, Z_t) \in \cdot\}}$ , it results that this construction is strictly linked with the particle filter described in Section 3; in particular Equations (A.4) and (A.5) can be numerically solved by simulation for each  $n$ . In this case the resulting algorithm is very similar to the one described in Section 3.3.

## References

- Brémaud P (1980) Point Processes and Queues. Springer
- Calzolari A, Nappo G (2001) The filtering problem in a model with grouped data and counting observation times. Working paper, Dipartimento di Matematica G. Castelnuovo, Università di Roma La Sapienza

- Centanni S, Minozzo M (2006a) Estimation and filtering by reversible jump MCMC for a doubly stochastic Poisson model for ultra-high-frequency financial data. *Statistical Modelling* 6: 97–118
- Centanni S, Minozzo M (2006b) A Monte Carlo approach to filtering for a class of marked doubly stochastic Poisson processes. *Journal of the American Statistical Association* 101: 1582–1597
- Chopin N (2004) Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. *The Annals of Statistics* 32: 2385–2411
- Cox DR, Isham V (1980) *Point Processes*. Chapman and Hall
- Dassios A, Jang J (2003) Pricing of catastrophe reinsurance and derivatives using the Cox process with shot noise intensity. *Finance & Stochastics* 7: 73–95
- Doucet A, De Freitas N, Gordon N (2001) *Sequential Monte Carlo Methods in Practice*. Springer
- Duffie D, Garleanu N (2001) Risk and valuation of collateralized debt obligation. *Financial Analysts Journal* 57: 41–59
- Engle RF, Russell JR (1998) Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica* 66: 1127–1162
- Ethier SN, Kurtz TG (1986) *Markov Processes: Characterization and Convergence*. Wiley
- Fearnhead P, Papaspiliopoulos O, Roberts GO (2006) Particle filtering for diffusions avoiding time-discretisations. *Nonlinear Statistical Signal Processing Workshop, 13-15 Sept 2006, IEEE*: 141–143
- Frey R (2000) Risk minimization with incomplete information in a model for high frequency data. *Mathematical Finance* 10: 215–225
- Frey R, Runggaldier WJ (2001) A nonlinear filtering approach to volatility estimation with a view towards high frequency data. *International Journal in Theoretical and Applied Finance* 4: 271–300

- Gerardi A, Tardelli P (2006) Filtering on a partially observed ultra-high-frequency data model. *Acta Applicandae Mathematicae* 91: 193–205
- Guillaume D, Dacorogna M, Davé R, Muller U, Olsen R, Pictet O (1999) From the bird’s eyes to the microscope: a survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics* 1: 95–129
- Jazwinski A (1970) *Stochastic Processes and Filtering Theory*. Academic Press
- Kalev PS, Liu WM, Pham PK, Jarnecic E (2004) Public information arrival and volatility of intraday stock returns. *Journal of Banking & Finance* 28: 1441–1467
- Last G, Brandt A (1995) *Marked Point Process on the Real Line: The Dynamic Approach*. Springer
- Minozzo M, Centanni S (2008) Modeling ultra-high-frequency data: the S&P500 index future. In: Perna C, Sibillo M (eds.) *Mathematical and Statistical Methods in Insurance and Finance*, Springer: 165–172
- Rogers LCG, Zane O (1998) Designing models for high frequency data. Working paper. University of Bath
- Rydborg TH, Shephard N (2000) A modelling framework for the prices and times of trades made on the New York stock exchange. In: Fitzgerald WJ, Smith RL, Walden AT, Young PC (eds.) *Nonlinear and Nonstationary Signal Processing*, Cambridge University Press: pp. 217–246
- Teich MC, Saleh BEA (2000) Branching processes in quantum electronics. *IEEE Journal on Selected Topics in Quantum Electronics* 6: 1450–1457