POLITECNICO DI MILANO

Scuola di Ingegneria Industriale e dell'Informazione
Corso di Laurea Magistrale in Ingegneria Matematica

Tesi di Laurea Magistrale

FILTERING OF PURE JUMP MARKOV PROCESSES WITH NOISE-FREE OBSERVATION

RELATORE:
Prof. Marco Alessandro Fuhrman

TESI DI LAUREA DI:
Alessandro Calvia
Matr. n. 781602

ANNO ACCADEMICO 2013–2014

Ai miei genitori Nicolao e Adriana

Lorsque l'on expose devant un public de mathématiciens [. . .] on peut supposer que chacun connaît les variétés de Stein ou les nombres de Betti d'un espace topologique; mais si l'on a besoin d'une intégrale stochastique, on doit définir à partir de zéro les filtrations, les processus prévisibles, les martingales, etc. Il y a là quelque chose d'anormal. Les raisons en sont bien sûr nombreuses, à commencer par le vocabulaire ésotérique des probabilistes. . .

— Laurent Schwartz

CONTENTS

ABSTRACT

This aim of this Thesis is to describe and analyze the problem of stochastic filtering of continuous-time pure jump Markov processes with noise-free observation.

A couple of continuous-time stochastic processes X_t and Y_t, defined on some probability space and with values in two measurable spaces $(I, \mathcal{I})$ and $(O, \mathcal{O})$ respectively, is given. We assume that the process X_t is a pure jump Markov process of known rate transition measure $\lambda(x, dy)$. Moreover, the observation process Y_t is not directly affected by noise. Finally, a deterministic function $h\colon I \to O$ relates the two processes, in the sense that $Y_t = h(X_t)$, $t \geqslant 0$.

We will derive an explicit equation for the filtering process $\Pi_t(A) = \mathbb{P}\left(X_t \in A \mid \mathcal{F}_t^Y\right)$, $A \in \mathcal{I}$, $t \geqslant 0$, where $\mathcal{F}_t^Y$ denotes the natural filtration of the process Y_t. This task will be accomplished by writing the processes X_t and Y_t in terms of two marked point processes. Then martingale calculus will be applied to the latter processes and will provide the explicit form of the aforementioned equation.

Keywords: nonlinear filtering; marked point processes; pure jump Markov processes.

SOMMARIO

L'obiettivo di questa Tesi è la descrizione e l'analisi del problema di filtraggio di un processo di Markov di puro salto a tempo continuo, con osservazioni prive di rumore.

Come noto, il filtraggio di processi stocastici riveste un'importanza fondamentale in molte aree scientifiche. In particolare, se si pensa all'ambito ingegneristico, è sufficiente menzionare i rami dell'automatica, dell'elettronica e dell'informatica. La sua rilevanza è dovuta al fatto che molti problemi concreti si presentano come problemi di filtraggio e controllo. In questi contesti, si vuole intraprendere delle azioni di controllo su taluni processi stocastici non osservabili direttamente; pertanto, attraverso osservazioni relative a un altro processo stocastico, si intende stimare lo stato del processo non osservabile e compiere le azioni di controllo sulla base di tali stime.

In questo lavoro, il dato con cui si ha a che fare è una coppia di processi stocastici X_t e Y_t a tempo continuo, definiti su uno stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e a valori in due spazi misurabili $(I, \mathcal{I})$ e $(O, \mathcal{O})$, rispettivamente. Il processo X_t sarà detto *processo non osservabile* e il processo Y_t sarà chiamato *processo osservabile*.

Lo scopo del problema di filtraggio è la descrizione e l'analisi delle proprietà del *processo di filtraggio*

$$\Pi_t(A) = \mathbb{P}\left(X_t \in A \mid \mathcal{F}_t^Y\right), \quad A \in \mathcal{I}, \quad t \geqslant 0,$$

ove con $\mathcal{F}_t^Y$ si è indicata la filtrazione naturale del processo Y_t, ossia la famiglia di σ-algebre $\left\{\sigma(Y_s, 0 \leqslant s \leqslant t)\right\}_{t \geqslant 0}$. Spesso tale processo soddisfa delle opportune equazioni differenziali e può essere caratterizzato come soluzione unica di tali equazioni, dette *equazioni di filtraggio*.

In letteratura, il modello di gran lunga più analizzato e quello in cui X_t è un generico processo markoviano e Y_t è un processo a valori in $O = \mathbb{R}^m$ della forma

$$Y_t = \int_0^t f(X_s)\, ds + W_t, \quad t \geqslant 0,$$

ove W_t è un moto browniano standard a valori in $\mathbb{R}^m$ definito sul medesimo spazio di probabilità su cui sono definiti X_t e Y_t; la funzione $f \colon I \to R^m$ è una funzione assegnata. Nel modello si considera, dunque, un processo osservabile che è un funzionale del processo non osservabile e su cui agisce un rumore non-degenere. Il lettore interessato a questo caso, può consultare i testi [1, 2, 18] per una sua trattazione generale. Un'esposizione dettagliata della soluzione a tale problema tramite i classici approcci con filtri di Kalman o Wonham, si trova in [11, 15].

Un modello che ha destato attenzione solo recentemente è quello nel quale si suppone

$$Y_t = h(X_t), \quad t \geqslant 0,$$

ove $h\colon I \to O$ è una funzione data. È evidente, in tal caso, che sul processo osservabile non agisce direttamente un rumore. Vi agisce solo indirettamente attraverso il processo non osservabile, che contiene in sé tutte le fonti di casualità. Questa classe di problemi è stata analizzata in relazione a diversi modelli di filtraggio, come in [6, 14, 16], oppure in casi particolari. Ad esempio, in [12] si pone $I = \mathbb{R}^n$, $O = \mathbb{R}^m$ e si fanno ipotesi specifiche sulla funzione h.

In questo contesto, il problema di filtraggio si inserisce nella più ampia classe dei modelli *Hidden Markov*, molto usata nelle applicazioni e a tutt'oggi molto studiata. Si rimanda il lettore al testo [10] per una trattazione esaustiva di questi modelli, sia a tempo discreto che a tempo continuo e sia con spazi di stato discreti o più generali.

Uno studio sistematico del problema di filtraggio nel caso di osservazioni prive di rumore è ancora assente in letteratura. Le sole opere che trattano questo specifico problema a tempo continuo sono [12] e [5]. Quest'ultimo è il lavoro su cui questa Tesi è basato. Ivi gli autori suppongono che gli spazi $(I, \mathcal{I})$ e $(O, \mathcal{O})$ siano di dimensione finita e che $h\colon I \to O$ sia una funzione suriettiva. Il processo X_t è una catena di Markov omogenea nel tempo a valori in I di cui è nota la matrice dei ratei di transizione Λ. Con queste assunzioni, il processo di filtraggio prende una forma più semplice, essendo completamente specificato da un numero finito di processi a valori reali, vale a dire

$$\Pi_t(i) = \mathbb{P}\left(X_t = i \mid \mathcal{F}_t^Y\right), \quad i \in I, \quad t \geqslant 0.$$

Inoltre per ricavare la forma dell'equazione di filtraggio è stato usato un metodo basato su approssimazioni discrete.

In questo lavoro, invece, si adotterà un punto di vista differente, basato sui processi di punto marcati e sulla teoria delle martingale. Un processo di punto marcato è una collezione di coppie di variabili aleatorie $(T_n, \xi_n)_{n \geqslant 1}$, definite su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, a valori in $\big([0, +\infty] \times E\big)$, ove $(E, \mathcal{E})$ è uno spazio misurabile. Questi processi sono atti a descrivere, da un punto di vista applicativo, sequenze di fenomeni fisici distribuiti nel tempo di cui si registrano dei valori d'interesse. Essi sono usati in svariati ambiti, ad esempio la teoria delle code.

Dei due approcci tipicamente usati nella descrizione di questi processi, il primo basato sulla teoria della misura, il secondo sul concetto di *intensità stocastica*, sarà quest'ultimo a essere adottato nel lavoro presente. L'intensità stocastica, se esiste, riassume in sé una misura, a un certo istante fissato, del potenziale che un processo di punto ha di generare un evento nell'immediato futuro, nota una certa quantità di informazione che includa almeno la conoscenza dei valori passati del processo stesso.

Di questa quantità si dà una definizione tramite la teoria delle martingale, come noto una teoria molto sviluppata, di cui, dunque, si potranno sfruttare i risultati. Il calcolo stocastico che ne deriva, fornirà degli strumenti adeguati e flessibili per una trattazione dei modelli basati su processi di punto sotto un'ottica dinamica. Questo modo di procedere è analogo al caso di sistemi governati da moti browniani, con cui si potranno ravvisare diverse similarità.

Il contributo principale di questa Tesi è di applicare le tecniche appena descritte al modello precedentemente presentato, il quale possiede senza dubbio una natura dinamica. Questo approccio ci consentirà di fornire dimostrazioni più semplici rispetto a quelle riportate in [5]. In più, potremo introdurre un'ulteriore novità. Saranno, infatti, indebolite le ipotesi introdotte in [5], supponendo che $(I, \mathcal{I})$ sia uno spazio metrico completo e separabile. Il processo X_t sarà, dunque, un processo di Markov di puro salto omogeneo nel tempo e a valori in I, di cui assumeremo note la misura dei ratei di transizione $\lambda(x, dy)$ e la distribuzione iniziale $\mu(dx)$.

Sintetizziamo, brevemente, i risultati originali contenuti in questo lavoro. L'equazione di filtraggio ottenuta è data da

$$\begin{aligned}
\hat{Z}_t(\omega, A) &= H_{Y_0(\omega)}[\mu](A) + \\
&+ \int_0^t \left\{ \int_I \lambda\Big(x, A \cap h^{-1}\big(Y_s(\omega)\big)\Big)\, \hat{Z}_{s-}(\omega, dx) - \int_A \lambda(x)\hat{Z}_{s-}(\omega, dx) + \right. \\
&\qquad \left. + \hat{Z}_{s-}(\omega, A) \int_I \lambda\Big(x, h^{-1}\big(Y_s(\omega)\big)^c\Big)\, \hat{Z}_{s-}(\omega, dx) \right\} ds + \\
&\qquad\qquad + \sum_{0<\tau_n(\omega)\leqslant t} \Big\{ H_{Y_{\tau_n}(\omega)}[\mu_n](A) - \hat{Z}_{\tau_n -}(\omega, A) \Big\},
\end{aligned}$$

per ogni $\omega \in \Omega$ e per ogni $t \geqslant 0$. Essa può anche essere scritta in una forma leggermente diversa, mostrata nella formula (2.30) contenuta nell'osservazione finale del Capitolo 2. La sua struttura, all'apparenza complessa, è in realtà semplice e a breve sarà discussa.

Il processo $\hat{Z}_t(\omega, A)$, $A \in \mathcal{I}$, altresì indicato con $\hat{Z}_t(A)$, è una versione del processo di filtraggio $\mathbb{P}(X_t \in A \mid \mathcal{O}_t)$. Ciò significa che, per ogni $t \geqslant 0$ e ogni $A \in \mathcal{I}$, $\hat{Z}_t(\omega, A) = \mathbb{P}(X_t \in A \mid \mathcal{O}_t)$, $\mathbb{P}-$q.c.. La filtrazione $\mathcal{O}_t$ è associata al processo osservabile e, nel nostro caso, si ha $\mathcal{F}_t^Y \equiv \mathcal{O}_t$. Pertanto, il processo presentato ora è parimenti una versione del processo di filtraggio scritto in precedenza.

Sia $A \in \mathcal{I}$ un insieme fissato. Possiamo scomporre il processo $\hat{Z}_t(A)$ in tre termini principali, visibili rispettivamente nella prima riga, nelle due righe centrali e nell'ultima riga della precedente equazione:

(1) Il valore iniziale, corrispondente a $\mathbb{P}(X_0 \in A \mid Y_0)$. L'operatore H che compare nell'equazione agisce sulla misura di probabilità μ trasformandola nella suddetta probabilità condizionale.

(II) La parte deterministica, composta di una parte lineare (seconda riga) e di una parte quadratica (terza riga), un aspetto comune a molti processi di filtraggio. È facile vedere che, tra i tempi di salto $(\tau_n)_{n\in\mathbb{N}}$ del processo Y_t, cioè per $t \in [\tau_n, \tau_{n+1})$, $n \in \mathbb{N}$, il processo evolve seguendo la dinamica deterministica prescritta dai termini summenzionati.

(III) La parte salto. Al tempo di salto τ_n, il solo termine $H_{Y_{\tau_n}}[\mu_n](A)$, $n \in \mathbb{N}$, determina il valore del processo. Esso corrisponde alla probabilità condizionale che il processo X_t prenda valori nell'insieme A successivamente all'*n*-esimo salto, rispetto alla σ-algebra $\mathcal{O}_{\tau_n}$.

Per quanto concerne la misura dei ratei di transizione, la notazione usata è $\lambda(x, A)$ anziché $\int_A \lambda(x, dy)$, $A \in \mathcal{I}$ e $\lambda(x)$ in luogo di $\int_I \lambda(x, dy)$, onde evitare troppi simboli d'integrazione nella formula. Infine, osserviamo che tutti gli integrali che compaiono, sono calcolati rispetto alla misura di probabilità condizionale $\hat{Z}_{s-}(dx)$, la quale è definita dallo stesso processo di filtraggio.

È evidente che il caso discusso in [5] è un particolare esempio di quanto analizzato in questo lavoro. Di conseguenza, è naturale notare la somiglianza tra [5, eq. 2.5] e la presente equazione. Ciò che differirà sarà il percorso seguito per ottenere questa formula.

Inizieremo con lo scrivere i processi X_t e Y_t come processi di punto marcati. Per la precisione, il processo osservabile sarà scritto come un processo di punto K-variato, ove con $K \in \mathbb{N}$ si denota la dimensione dell'insieme O. Introdurremo, poi, il processo $Z_t(A) = \mathbb{1}_A(X_t)$, $t \geqslant 0$, ove A sarà un boreliano di I fissato. Scriveremo, dunque, la sua rappresentazione di semimartingala. Questo sarà un compito agevole, poiché è noto il generatore infinitesimale $\mathcal{L}$ del processo X_t, dunque la formula di Dynkin fornirà il risultato richiesto. Dopo aver verificato le opportune ipotesi, potremo applicare il teorema di filtraggio e ottenere l'equazione di filtraggio in una prima formulazione:

$$\hat{Z}_t(A) = \hat{Z}_0(A) + \int_0^t \hat{f}_s \, ds + \hat{m}_t, \quad t \geqslant 0,$$

ove

- $\hat{Z}_0(A) = \mathbb{P}\left(X_0 \in A \mid Y_0\right)$,
- $\hat{f}_t = \mathbb{E}\left[\mathcal{L}\varphi(X_t) \mid \mathcal{O}_t\right], \quad \varphi(x) = \mathbb{1}_A(x)$,
- $\hat{m}_t = \sum_{k=1}^K \int_0^t K_s(k, A)\,[dN_s^Y(k) - \hat{\lambda}_s^Y(k)\,ds]$.

Il termine $dN_s^Y(k) - \hat{\lambda}_s^Y(k)\,ds$ è la cosiddetta *misura compensata* associata al *k*-esimo processo di conteggio $N_t^Y(k)$, $k = 1, \dots, K$. Tutte queste quantità sono collegate al processo osservabile e saranno analizzate in dettaglio nel Capitolo 1.

La parte saliente di questa formulazione dell'equazione di filtraggio è data dal *processo di guadagno* $K_t(k, A)$. Per ogni $k = 1, \dots, K$, esso può essere scritto come somma di tre termini, $\Psi_{1,t}(k, A)$, $\Psi_{2,t}(k, A)$ e $\Psi_{3,t}(k, A)$. Ciascuno di questi ultimi è definito come derivata di Radon-Nikodym di ben precise misure e ne potremo esplicitare la forma. Infine, dopo una serie di calcoli. giungeremo all'equazione di filtraggio nella sua formulazione definitiva prima mostrata.

La Tesi sarà strutturata come segue:

- Nel Capitolo 1 descriveremo i processi di punto marcati, dandone anche qualche semplice esempio, e illustreremo le tecniche di filtraggio con osservazioni costituite da processi di punto.
- Nel Capitolo 2 esporremo dettagliatamente il problema di filtraggio sintetizzato in precedenza e ne dimostreremo la formula di filtraggio.
- Nella sezione conclusiva riassumeremo i principali risultati ottenuti e daremo uno sguardo alle potenziali estensioni di questo lavoro e agli sviluppi futuri sull'argomento.

Una breve esposizione sui principali concetti riguardanti la teoria dei processi stocastici (alcuni dei quali sono già stati incontrati), come filtrazioni, tempi d'arresto, martingale e processi prevedibili, verrà fornita nell'Appendice A.

Tutti i risultati richiamati nel prossimo Capitolo o in Appendice sono dati senza dimostrazione. Le uniche dimostrazioni contenute nella presente Tesi sono quelle originali, riguardanti i risultati presentati nel Capitolo 2.

PAROLE CHIAVE: filtraggio non lineare; processi di punto marcati; processi di Markov di puro salto.

INTRODUCTION

The aim of this work is to describe and analyze the problem of stochastic filtering of continuous-time pure jump Markov processes with noise-free observation.

Filtering of stochastic processes is of foremost importance in many scientific areas, in particular in engineering fields such as automation, electronics and informatics. Its prominence is due to the nature of many real-world problems, where controlled processes cannot be observed directly and control actions are instead performed on the basis of another observed process.

To make more explicit this setting, let us introduce the classical formulation of the problem, that will be restated and further detailed in Chapter 2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(X_t)_{t \geqslant 0}$ and $(Y_t)_{t \geqslant 0}$ be a pair of stochastic processes with values in two measurable spaces $(I, \mathcal{I})$ and $(O, \mathcal{O})$, respectively. X_t is called the *unobserved* (or *signal*) process and Y_t is called the *observation process*. We can then define a *filtering process* as

$$\Pi_t(A) = \mathbb{P}\left(X_t \in A \mid \mathcal{F}_t^Y\right), \quad A \in \mathcal{I}, \quad t \geqslant 0, \tag{1}$$

where $\mathcal{F}_t^Y = \sigma(Y_s, 0 \leqslant s \leqslant t)$ are the σ-algebras of the natural filtration of the observation process.

In a general context, the filtering problem addresses the issue of describing the process (1) and finding its key properties. It is often the case that this process satisfies some differential equations and it can be characterized as the unique solution of such equations. These are called the *filtering equations*.

In the literature, the most common case is to take X_t as a Markov process and Y_t as an $\mathbb{R}^m$-valued process (i.e. $O = \mathbb{R}^m$) of the form

$$Y_t = \int_0^t f(X_s)\, ds + W_t, \quad t \geqslant 0, \tag{2}$$

where W_t is a standard $\mathbb{R}^m$-valued Wiener process defined on the same probability space as X_t and Y_t and $f\colon I \to R^m$ is a given function. In words, the observation process is a functional of the unobserved one and a non-degenerate noise is acting on it. The interested reader is referenced to [1, 2, 18] for a general treatment of this situation and to [11, 15] for a detailed exposition of the solution to this case with the classical approach of Kalman or Wonham filters.

A different model addressed recently by several authors is the following:

$$Y_t = h(X_t), \quad t \geqslant 0, \tag{3}$$

where $h\colon I \to O$ is a given function. This is the case where Y_t is a noise-free observation, i.e. it is not directly affected by noise. Thus, all the sources of randomness are included in the unobserved process. This kind of problems has been considered in connection with different filtering models, as in [6, 14, 16], or in special cases, as in [12] where $I = \mathbb{R}^n$, $O = \mathbb{R}^m$ and the function h in (3) bears some special assumptions.

In this setting, the filtering problem is an instance of a Hidden Markov model. This is a broader class of stochastic filtering models, greatly used in the applications and that is still the subject of intense investigation. A comprehensive exposition of these models, both in discrete- and continuous-time and with discrete and general state spaces, can be found in [10].

However, filtering in the noise-free case has not been yet systematically studied. To the best of my knowledge, the only works covering this issue in continuous time are [12] and [5]. The latter forms the basis for this work.[1] It is assumed there that I and O are two finite sets, $h\colon I \to O$ is a surjective function and X_t is a time-homogeneous Markov chain with values in I and known rate transition matrix Λ. Then, with the observation process defined as in (3), the filtering process is specified by a finite set of scalar processes, namely

$$\Pi_t(i) = \mathbb{P}\left(X_t = i \mid \mathcal{F}_t^Y\right), \quad i \in I, \quad t \geqslant 0. \tag{4}$$

To some extent, this framework simplifies the model and the filtering equation presented there is proven with a method based on discrete approximation.

In this work we will use, instead, a different approach, based on marked point processes and the martingale theory. To fix some key ideas (that will be thoroughly exposed in Chapter 1), a marked point process is basically a collection $(T_n, \xi_n)_{n \geqslant 1}$ of pairs of random variables, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with values in $\big([0, +\infty] \times E\big)$, where $(E, \mathcal{E})$ is a measurable space. From the point of view of applications, the nth couple of random variables can be thought of as the nth occurrence of a given physical phenomenon, of which we record the time of occurrence and some related attributes, e. g. the spike time and spike amplitude of the activity of a nervous fiber. Point-process models are widely used in various fields, one for all queuing theory in operations research.

Two main approaches are used to describe these processes:

- a measure-theoretical one, where they are viewed as discrete random measures;
- a dynamical one, via the concept of *stochastic intensity*.

It is the second one that will be adopted here. Roughly speaking, the stochastic intensity summarizes at a given instant the potential that a

1 An example of filtering equations is also present in [3, Proposition 3.2], albeit in the specific case of a Markov chain with four states and transition rates equal to 0 or 1.

point process has to generate an event in the immediate future, given a certain amount of information available at that time including the knowledge of the past of the process itself.

The martingale definition of stochastic intensity enables to use the results known from the deeply developed *martingale theory*. The *martingale calculus* for point processes provides flexible instruments for a treatment of point process models from a dynamical point of view. The reader acquainted with these arguments in the case of systems driven by brownian motion, will find the exposition very familiar.

The main contribution of this Thesis is to apply these techniques to the noise-free model earlier presented, clearly possessing a dynamical nature. This will permit us to give simpler proofs than those reported in [5] and even to introduce another point of novelty in the analysis of this subject. We will, in fact, weaken the assumptions made earlier and suppose that $(I, \mathcal{I})$ is a complete separable metric space. The process X_t is, then, an I-valued time-homogeneous pure jump Markov process of known rate transition measure $\lambda(x, dy)$ and initial distribution $\mu(dx)$.

Let us now briefly summarize the original results that will be thoroughly discussed in Chapter 2. The filtering equation to be obtained will present itself in the final form

$$\begin{aligned} \hat{Z}_t(\omega, A) &= H_{Y_0(\omega)}[\mu](A) + \\ &+ \int_0^t \left\{ \int_I \lambda\Big(x, A \cap h^{-1}\left(Y_s(\omega)\right)\Big)\, \hat{Z}_{s-}(\omega, dx) - \int_A \lambda(x) \hat{Z}_{s-}(\omega, dx) + \right. \\ &\quad \left. + \hat{Z}_{s-}(\omega, A) \int_I \lambda\Big(x, h^{-1}\left(Y_s(\omega)\right)^c\Big)\, \hat{Z}_{s-}(\omega, dx) \right\} ds + \\ &\quad + \sum_{0 < \tau_n(\omega) \leqslant t} \left\{ H_{Y_{\tau_n}(\omega)}[\mu_n](A) - \hat{Z}_{\tau_n -}(\omega, A) \right\}, \end{aligned} \tag{5}$$

for all $\omega \in \Omega$ and for all $t \geqslant 0$. It can also be stated in a slightly different form, that will be shown by equation (2.30) in the final Remark of Chapter 2. Though it may seem daunting at a first glance, this equation has a simple structure, that we will shortly discuss. Let us, first, explain the symbols that appear there.

The process $\hat{Z}_t(\omega, A)$, $A \in \mathcal{I}$, equivalently indicated by $\hat{Z}_t(A)$, is a version of the filtering process $\mathbb{P}\left(X_t \in A \mid \mathcal{O}_t\right)$. This means that for all $t \geqslant 0$ and all $A \in \mathcal{I}$, $\hat{Z}_t(\omega, A) = \mathbb{P}\left(X_t \in A \mid \mathcal{O}_t\right)$, $\mathbb{P}-a.s.$. The filtration $\mathcal{O}_t$ is the so called *observed history* and is associated to the observation process. We will see that in our case $\mathcal{F}_t^Y \equiv \mathcal{O}_t$, so the process in equation (5) is also a version of the filtering process presented earlier in (1).

Let us now fix a set $A \in \mathcal{I}$. The process $\hat{Z}_t(A)$ is composed of three main terms:

(1) The summand in the first line of equation (5) is equal to the conditional probability $\mathbb{P}\left(X_0 \in A \mid Y_0\right)$. It is, then, the starting value

of the process. The operator H acts on the probability measure μ transforming it into the conditional probability just written.

(II) The summands in the second and third line of equation (5) are relative to the deterministic part of the filtering process. It is easy to see that, between the jump times $(\tau_n)_{n\in\mathbb{N}}$ of the process Y_t, i.e. for $t \in [\tau_n, \tau_{n+1})$, $n \in \mathbb{N}$, the process evolves according to the deterministic dynamic given by these terms. They are composed of a linear part (the second line of the equation) and of a quadratic part (the third line of the equation), a feature shared by various filtering processes.

(III) The summand in the last line of equation (5) represents the jump component of the filtering process. At the *n*th jump time, the term $H_{Y_{\tau_n}}[\mu_n](A)$, $n \in \mathbb{N}$, alone determines the value of the process. This one is the conditional probability that the process X_t will take a value in the set A after the *n*th jump, with respect to the observed history up to that time.

Another quantity that appears in equation (5) is the rate transition measure $\lambda(x, dy)$. To avoid too many integrals in it, we have adopted the notations $\lambda(x, A)$ for $\int_A \lambda(x, dy)$, $A \in \mathcal{I}$, and $\lambda(x)$ for $\int_I \lambda(x, dy)$. Finally, we notice that all the integrals that appear are computed with respect to the conditional probability measure $\hat{Z}_{s-}(dx)$ defined by the process itself.

The filtering equation (5) is very similar to [5, eq. 2.5], the latter being a special case of the former. However, the path that we will follow to derive it will be very different.

We will begin by writing the processes X_t and Y_t as a marked point process. To be precise, the observation process will be written as a K-variate point process, where $K \in \mathbb{N}$ is the dimension of the set O. Then, we will introduce the process $Z_t(A) = \mathbb{1}_A(X_t)$, $t \geqslant 0$, where A is a fixed borelian subset of I. We will write down its semimartingale representation, an easy task to do since the infinitesimal generator $\mathcal{L}$ of the process X_t is known, so an immediate application of Dynkin's formula will yield the result. After checking the relevant hypotheses, we will be in a position to apply the filtering theorem and reach the filtering equation, in a first "rough" form:

$$\hat{Z}_t(A) = \hat{Z}_0(A) + \int_0^t \hat{f}_s \, ds + \hat{m}_t, \quad t \geqslant 0, \tag{6}$$

where

- $\hat{Z}_0(A) = \mathbb{P}(X_0 \in A \mid Y_0)$,
- $\hat{f}_t = \mathbb{E}[\mathcal{L}\varphi(X_t) \mid \mathcal{O}_t], \quad \varphi(x) = \mathbb{1}_A(x)$,
- $\hat{m}_t = \sum_{k=1}^K \int_0^t K_s(k, A)\,[dN_s^Y(k) - \hat{\lambda}_s^Y(k)\,ds]$.

The term $dN_s^Y(k) - \hat{\lambda}_s^Y(k)\,ds$ is the so called *compensated measure* associated to the *k*th counting process $N_t^Y(k)$, $k = 1, \dots, K$. All these objects are related to the observed process and will be analyzed in detail in Chapter 1.

The core of that formula is represented by the *innovations gain* processes $K_t(k, A)$. For each $k = 1, \dots, K$, it can be written as the sum of three terms, $\Psi_{1,t}(k, A)$, $\Psi_{2,t}(k, A)$ and $\Psi_{3,t}(k, A)$. Each of them is defined as a Radon-Nikodym derivative of specified measures and the martingale calculus for point processes will allow us to explicit their form. After all the due computations, we will arrive to the filtering formula (5).

Let us now conclude this brief presentation of the work by detailing the structure of the thesis.

- In Chapter 1 we will describe the notion of marked point process, with a few simple but meaningful examples, and the stochastic filtering techniques with point process observations.

- In Chapter 2 we will present in full detail the filtering problem that we briefly summarized before and we will derive the corresponding filtering equation.

- A final section will be devoted to synthesizing the main results obtained here and will give a hint of what the future extensions and further developments on this subject could be.

A few concepts concerning stochastic processes (some of which we have already encountered), such as filtrations, stopping times, martingales and predictability, will be reviewed in Appendix A.

All the results presented in the next Chapter or in Appendix A are stated without proof. The only proofs contained in this Thesis are the original ones, concerning the discussion made in Chapter 2.

1 MARKED POINT PROCESSES

The class of marked point processes plays a central role in this work. The filtering problem that we will address in Chapter 2 will be completely described by suitably defined marked point processes.

We recall that, in order to provide a solution to it, we shall adopt a dynamical point of view on such processes, i. e. operate through the associated counting measures and intensity kernels. Martingale theory will then provide us with key results, mainly the integral representation of point-process martingales: this is the fundamental theorem to be used in the development of the filtering techniques and will be discussed in the final section of this Chapter.

For these reasons, this Chapter will be loosely based on the approach presented by Brémaud in [4]. For the sake of completeness, every proposition or theorem here exposed will feature a precise reference to that source.

Before starting, we recall some useful notations.

- The indicator function will often be denoted by $\mathbb{1}(x \in A)$, instead of the classical $\mathbb{1}_A(x)$.
- $\mathcal{B}(A)$ are the Borel subsets of $A \subset \mathbb{R}$.
- The set $[0, +\infty)$ will be indicated by $\mathbb{R}_+$ and, correspondingly, $\mathcal{B}_+ = \mathcal{B}\big([0, +\infty)\big)$.
- h^{-1} denotes the pre-image of a set under the function h.

Throughout this Chapter we will assume defined a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a measurable space $(E, \mathcal{E})$.

1.1 POINT PROCESSES AND STOCHASTIC INTENSITY

Let us first define the class of point processes. They can be thought of as the *n*th occurrence of a given physical phenomenon. Their relevance here is to help us introduce important objects related to them, such as counting processes and stochastic intensities. These objects will be later generalized to marked point processes, of which point processes are particular and simpler examples (as we will see, their so called *mark space* reduces to a single point).

Definition 1.1 (Point Process): Let $(T_n)_{n \in \mathbb{N}}$ be a sequence of random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $[0, +\infty]$, such that

$$T_0 = 0, \tag{1.1a}$$

$$T_n < +\infty \Rightarrow T_n < T_{n+1},\ \mathbb{P} - a.s., \forall n \in \mathbb{N}. \tag{1.1b}$$

Then the sequence $(T_n)_{n \in \mathbb{N}}$ is called a *point process*. It is said to be $\mathbb{P}$-*nonexplosive* if

$$T_\infty = \lim_{n \to +\infty} T_n = +\infty \quad \mathbb{P} - a.s. \tag{1.2}$$

REMARK: Henceforward we will assume that all point processes are $\mathbb{P}$-nonexplosive.

A *counting process* can be associated to a point process, simply defining

$$N_t = \sum_{n \geqslant 1} \mathbb{1}(T_n \leqslant t). \tag{1.3}$$

This process is also called a point process, by abuse of notation, since T_n and N_t carry the same information.
Moreover, the process N_t is said to be *integrable* if

$$\mathbb{E}[N_t] < \infty, \quad \forall t \geqslant 0. \tag{1.4}$$

Naturally linked to a point process is the concept of stochastic intensity. The following examples will help us to introduce it, before giving its general definition.

EXAMPLE 1.1 (Homogeneous Poisson Process): Let N_t be a point process adapted to a filtration $\mathcal{F}_t$ and let λ be a nonnegative constant.
If for all $0 \leqslant s \leqslant t$ and all $u \in \mathbb{R}$

$$\mathbb{E}\left[e^{iu(N_t - N_s)} \mid \mathcal{F}_s\right] = \exp\{\lambda(t-s)(e^{iu} - 1)\}, \tag{1.5}$$

then N_t is called a $\mathcal{F}_t$-*homogeneous Poisson process* with intensity λ.
The condition (1.5) implies that for all $0 \leqslant s \leqslant t$ the increments $N_t - N_s$ are $\mathbb{P}$-independent of $\mathcal{F}_s$ given $\mathcal{F}_0$. Moreover, it leads to the usual formula

$$\mathbb{P}(N_t - N_s = k \mid \mathcal{F}_s) = e^{-\lambda(t-s)} \frac{(\lambda(t-s))^k}{k!} \quad \forall k \in \mathbb{N}. \tag{1.6}$$

A simple calculation using formula (1.6) shows that $\mathbb{E}[N_t] = \lambda t$. This allows us to interpret the intensity of the process N_t as the expected number of "events" that occur per unit time and identifying it with λ. This reasoning can be further generalized in order to consider a wider class of processes that are still related to the Poisson distribution, as shown in the following example.

EXAMPLE 1.2 (Conditional Poisson Process): Let N_t be a point process adapted to a filtration $\mathcal{F}_t$ and let λ_t be a nonnegative measurable

process.
Suppose that the following conditions hold:

$$\lambda_t \text{ is } \mathcal{F}_0\text{-measurable}, \quad \forall t \geqslant 0, \tag{1.7a}$$

$$\int_0^t \lambda_s \, ds < \infty \quad \mathbb{P} - a.s., \quad \forall t \geqslant 0, \tag{1.7b}$$

$$\mathbb{E}\left[e^{iu(N_t - N_s)} \mid \mathcal{F}_s\right] = \exp\left\{(e^{iu} - 1) \int_s^t \lambda_r \, dr\right\}. \tag{1.7c}$$

Then N_t is called a $\mathcal{F}_t$-*conditional Poisson process* with the stochastic intensity λ_t.[1]

Equations (1.5) and (1.7) provide the very definition of stochastic intensity for the counting process N_t. However, the conditions previously stated give a very peculiar probabilistic structure to that process, e. g. conditionally independent and Poisson-distributed increments (even stationary in the former example). If we want to define the $\mathcal{F}_t$-intensity of a point process N_t in the general case, we cannot resort to those conditions.

Nonetheless, it is possible to provide a definition, by using the hypotheses contained in the Watanabe's characterization theorem for conditional Poisson processes.[2]

Definition 1.2 (Stochastic Intensity): Let N_t be a point process adapted to a filtration $\mathcal{F}_t$, and let λ_t be a nonnegative $\mathcal{F}_t$-progressive process such that for all $t \geqslant 0$

$$\int_0^t \lambda_s \, ds < \infty \quad \mathbb{P} - a.s. \tag{1.8}$$

If for all nonnegative $\mathcal{F}_t$-predictable processes C_t the equality

$$\mathbb{E}\left[\int_0^\infty C_s \, dN_s\right] = \mathbb{E}\left[\int_0^\infty C_s \, \lambda_s \, ds\right] \tag{1.9}$$

is verified, then we say that N_t *admits the* $\mathcal{F}_t$-*stochastic intensity* λ_t.

It is important to remark, at this point, that the stochastic intensity may fail to exist. The object of which we can grant the existence is the *dual predictable projection* of the point process N_t. Before stating its existence theorem, we remember the so called *usual conditions* for a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geqslant 0}, \mathbb{P})$ as defined by Dellacherie in [9]:

(1) $\mathcal{F}$ is $\mathbb{P}$-complete,

1 Other terminologies for these processes are *doubly stochastic Poisson processes* or *Cox processes*.

2 For the sake of precision, the theorem presented by Watanabe in 1964[17] concerns Poisson processes. The generalized version for conditional Poisson processes can be found in [4, p. 25]

(2) $\mathcal{F}_t$ is right-continuous,

(3) $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets of $\mathcal{F}_t$.

THEOREM 1.1 (Existence of the Dual Predictable Projection [4, T12, p. 245]): *Let N_t be a point process adapted to a filtration $\mathcal{F}_t$ and assume that for the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geqslant 0}, \mathbb{P})$ the usual conditions are verified. Then, there exists a unique (up to $\mathbb{P}$-indistinguishability) right-continuous $\mathcal{F}_t$-predictable nondecreasing process A_t, with $A_0 = 0$, such that for all nonnegative $\mathcal{F}_t$-predictable processes C_t,*

$$\mathbb{E}\left[\int_0^\infty C_s \, dN_s\right] = \mathbb{E}\left[\int_0^\infty C_s \, dA_s\right]. \tag{1.10}$$

The process A_t is called the dual $\mathcal{F}_t$-predictable projection of N_t.

REMARK: If the process A_t is absolutely continuous with respect to the Lebesgue measure, in the sense that there exists a $\mathcal{F}_t$-progressive nonnegative process λ_t such that

$$A_t = \int_0^t \lambda_s \, ds, \quad t \geqslant 0, \tag{1.11}$$

then the stochastic intensity exists. In our situation, we will always be able to show its existence in a direct way.

The stochastic intensity permits us to link the martingale theory to point processes, via the following theorem.

THEOREM 1.2 (Integration Theorem [4, T8, p. 27]): *If the $\mathcal{F}_t$-adapted point process N_t admits the $\mathcal{F}_t$-intensity λ_t, then N_t is $\mathbb{P}$-nonexplosive and*

(1) $M_t = N_t - \int_0^t \lambda_s \, ds$ *is a $\mathcal{F}_t$-local martingale;*

(2) *if X_t is a $\mathcal{F}_t$-predictable process such that*

$$\mathbb{E}\left[\int_0^t |X_s| \lambda_s \, ds\right] < \infty, \forall t \geqslant 0, \tag{1.12}$$

then $\int_0^t X_s \, dM_s$ is a $\mathcal{F}_t$-martingale;

(3) *if X_t is a $\mathcal{F}_t$-predictable process such that*

$$\int_0^t |X_s| \lambda_s \, ds < \infty, \mathbb{P} - a.s., \forall t \geqslant 0, \tag{1.13}$$

then $\int_0^t X_s \, dM_s$ is a $\mathcal{F}_t$-local martingale.

The following characterization of the stochastic intensity is of great importance in the applications. It exploits the martingale relation presented in the preceding theorem.

THEOREM 1.3 (Martingale Characterization of Intensity [4, T9, p. 28]): *Let* N_t *be a nonexplosive point process adapted to a filtration* $\mathcal{F}_t$ *and let* $(T_n)_{n\geqslant 1}$ *be the sequence of its jump times.*
Suppose that for some nonnegative $\mathcal{F}_t$*-progressive process* λ_t *and for all* $n \geqslant 1$,

$$N_{t\wedge T_n} - \int_0^{t\wedge T_n} \lambda_s \, ds \text{ is an } \mathcal{F}_t\text{-martingale} \tag{1.14}$$

Then λ_t *is the* $\mathcal{F}_t$*-intensity of* N_t.

We conclude this Section by pointing out that, in general, more than one $\mathcal{F}_t$-intensity can be exhibited for a point process N_t. However, we can always find a predictable version of the intensity and if we constrain the intensity to be predictable, then it is essentially unique. This is the content of the following theorem.

THEOREM 1.4 (Existence and Uniqueness of Predictable Versions of the Intensity [4, T12 and T13, p. 31]): *Let* N_t *be a point process adapted to a filtration* $\mathcal{F}_t$*, admitting an* $\mathcal{F}_t$*-intensity* λ_t*, and let* $(T_n)_{n\geqslant 1}$ *be the sequence of its jump times.*
Then an $\mathcal{F}_t$*-predictable version of* λ_t *exists. Moreover, if* $\hat{\lambda}_t$ *and* $\tilde{\lambda}_t$ *are two* $\mathcal{F}_t$*-predictable intensities of* N_t*, then*

$$\hat{\lambda}_t(\omega) = \tilde{\lambda}_t(\omega) \quad \mathbb{P}(d\omega)dN_t(\omega)\text{-a.e.} \tag{1.15}$$

In particular, $\mathbb{P} - \text{a.s.}$,

$$\hat{\lambda}_{T_n} = \tilde{\lambda}_{T_n} \quad \text{on } \{T_n < \infty\}, \quad \forall n \geqslant 1, \tag{1.16a}$$

$$\hat{\lambda}_t(\omega) = \tilde{\lambda}_t(\omega) \quad \hat{\lambda}_t(\omega)dt \text{ and } \tilde{\lambda}_t(\omega)dt\text{-a.e.,} \tag{1.16b}$$

$$\hat{\lambda}_{T_n} > 0 \quad \text{on } \{T_n < \infty\}, \quad \forall n \geqslant 1. \tag{1.16c}$$

1.2 MARKED POINT PROCESSES AND INTENSITY KERNELS

We can now generalize the concept of point process in the following way. Let N_t be a point process and let $(T_n)_{n\geqslant 1}$ be the sequence of its jump times. We can associate to these jump times a sequence of E-valued random variables $(\xi_n)_{n\geqslant 1}$, defined in the same probability space as the point process.
To give a practical meaning to the situation described above, we can think of T_n as the *n*th occurrence of a specific physical phenomenon being described by the value ξ_n of some attributes. For instance, the phenomenon could be the *n*th lightning occurring during a storm at time T_n, whose magnitude is recorded and described in a suitable way by ξ_n.

DEFINITION 1.3 (Marked Point Process): Let there be defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ a point process N_t and a sequence of E-valued random variables $(\xi_n)_{n\geqslant 1}$.

1. The sequence $(T_n, \xi_n)_{n\geqslant 1}$ is called an E-*marked point process*.
2. The measurable space $(E, \mathcal{E})$ on which the sequence $(\xi_n)_{n\geqslant 1}$ takes its values is the *mark space*.

Example 1.3 (Point Process): As anticipated in the beginning of section 1.1, a point process can be thought as a marked point process whose mark space E is reduced to a single point. Then the sequences $(T_n)_{n\geqslant 1}$ and $(T_n, \xi_n)_{n\geqslant 1}$ can be obviously identified.

Example 1.4 (Multivariate Point Process): A peculiar kind of marked point process is the so called *K-variate point process*. In this case the mark space E consists of $K \in \mathbb{N}$ points $z_1, \dots, z_K$. We can simplify the notation by defining, for each $k = 1, \dots, K$,

$$N_t(k) = \sum_{n\geqslant 1} \mathbb{1}(\xi_n = z_k)\, \mathbb{1}(T_n \leqslant t), \quad t \geqslant 0. \tag{1.17}$$

Then, we obtain a collection of K point processes $\big(N_t(1), \dots, N_t(K)\big)$, that have no common jumps. This property means that, for all $t \geqslant 0$ and $\mathbb{P} - a.s.$,

$$\Delta N_t(i)\, \Delta N_t(j) = 0, \quad \forall i \neq j, \quad i, j \in \{1, \dots, K\}, \tag{1.18}$$

where $\Delta N_t(i) = N_t(i) - N_{t-}(i)$. These processes will be central in Chapter 2.

We can associate to any measurable set $A \in \mathcal{E}$ the counting process $N_t(A)$ defined by

$$N_t(\omega, A) = \sum_{n\geqslant 1} \mathbb{1}(\xi_n(\omega) \in A)\, \mathbb{1}(T_n(\omega) \leqslant t). \tag{1.19}$$

In particular, $N_t(E) = N_t$. Through this process we can define another important object, the *counting measure*, given by

$$p(\omega, (0, t] \times A) = N_t(\omega, A), \quad t \geqslant 0,\ A \in \mathcal{E} \tag{1.20}$$

It is a transition measure from $(\Omega, \mathcal{F})$ into $\big((0, \infty) \times E, \mathcal{B}(0, \infty) \otimes \mathcal{E}\big)$, i.e.

(1) $p(\omega, \cdot)$ is a measure on $\big((0, \infty) \times E, \mathcal{B}(0, \infty) \otimes \mathcal{E}\big)$, for all $\omega \in \Omega$;

(2) $\omega \mapsto p(\omega, B)$ is $\mathcal{F}$-measurable, for all $B \in \mathcal{B}(0, \infty) \otimes \mathcal{E})$

Remark: For ease of notation, as we have implicitly done in the previous section, in the sequel we will frequently drop the ω in the notation of all the random quantities.

The counting measure can be also written as

$$p(dt \times dx) = \sum_{n\geqslant 1} \delta_{(T_n, \xi_n)}(dt \times dx)\, \mathbb{1}(T_n < \infty), \tag{1.21}$$

where δ is the *Dirac measure*. It is, then, obvious that the sequence $(T_n, \xi_n)_{n \geqslant 1}$ and the counting measure $p(dt \times dx)$ can be identified and both called E-*marked point process*.
For this reason, the natural filtration of $(T_n, \xi_n)_{n \geqslant 1}$, defined by

$$\mathcal{F}_t^p = \sigma(N_t(A);\ 0 \leqslant s \leqslant t,\ A \in \mathcal{E}). \tag{1.22}$$

is indicated using the superscript ^{p}.

Before proceeding, we introduce an important class of processes.

Definition 1.4 (Indexed Predictable Process): Let $p(dt \times dx)$ be a E-marked point process and let $\mathcal{F}_t$ be a filtration such that

$$\mathcal{F}_t \supset \mathcal{F}_t^p, \quad \forall t \geqslant 0. \tag{1.23}$$

Let $\tilde{\mathcal{P}}(\mathcal{F}_t)$ be the σ-field defined on $(0, \infty) \times \Omega \times E$ as

$$\tilde{\mathcal{P}}(\mathcal{F}_t) = \mathcal{P}(\mathcal{F}_t) \otimes \mathcal{E}, \tag{1.24}$$

where $\mathcal{P}(\mathcal{F}_t)$ is the predictable σ-field on $(0, \infty) \times \Omega$.
Any $\tilde{\mathcal{P}}(\mathcal{F}_t)$-measurable mapping $H\colon (0, \infty) \times \Omega \times E \to \mathbb{R}$ is called a $\mathcal{F}_t$*-predictable process indexed by* E.

Remark: It is worth noting that the σ-field $\tilde{\mathcal{P}}(\mathcal{F}_t)$ is generated by the mappings H of the form:

$$H_t(\omega, x) = C_t(\omega)\, \mathbb{1}_A(x), \tag{1.25}$$

where C_t is a $\mathcal{F}_t$-predictable process and $A \in \mathcal{E}$.

We can introduce the following notation, that gives a precise meaning to the integration of predictable processes with respect to the counting measure $p(dt \times dx)$:

$$\int_0^t \int_E H_s(x)\, p(ds \times dx) = \sum_{n=1}^{\infty} H_{T_n}(\xi_n)\, \mathbb{1}(T_n \leqslant t), \tag{1.26}$$

where the symbol $\int_a^b$ is to be interpreted as $\int_{(a,b]}$ if $b < \infty$, and $\int_{(a,b)}$ if $b = \infty$.

We present now another fundamental quantity for marked point processes, that is the analogous of the stochastic intensity defined in Section 1.1.

Definition 1.5 (Intensity Kernel): Let $p(dt \times dx)$ be a $\mathcal{F}_t$-adapted E-marked point process. Suppose that for each $A \in \mathcal{E}$, $N_t(A)$ admits the $\mathcal{F}_t$-predictable intensity $\lambda_t(A)$, where $\lambda_t(\omega, dx)$ is a transition measure from $(\Omega \times [0, \infty), \mathcal{F} \otimes \mathcal{B}_+)$ into $(E, \mathcal{E})$.
We say that $p(dt \times dx)$ *admits the* $\mathcal{F}_t$*-intensity kernel* $\lambda_t(dx)$.

THEOREM 1.5 (Projection Theorem [4, T3, p. 235]): *Let* $p(dt \times dx)$ *be a* E*-marked point process with* $\mathcal{F}_t$*-intensity kernel* $\lambda_t(dx)$*. Then for each nonnegative* $\mathcal{F}_t$*-predictable* E*-indexed process* H

$$\mathbb{E}\left[\int_0^\infty \int_E H_s(x)\, p(ds \times dx)\right] = \mathbb{E}\left[\int_0^\infty \int_E H_s(x)\, \lambda_s(dx)\, ds\right]. \quad (1.27)$$

THEOREM 1.6 (Integration Theorem [4, C4, p. 235]): *Let* $p(dt \times dx)$ *be a* E*-marked point process with* $\mathcal{F}_t$*-intensity kernel* $\lambda_t(dx)$*. Let* H *be a* $\mathcal{F}_t$*-predictable* E*-indexed process.*

(1) *If, for all* $t \geqslant 0$ *we have*

$$\int_0^t \int_E |H_s(x)|\, \lambda_s(dx)\, ds < \infty \quad \mathbb{P}-a.s. \quad (1.28)$$

then $\int_0^t \int_E H_s(x)\, \tilde{p}(ds \times dx)$ *is a* $\mathcal{F}_t$*-local martingale.*

(2) *If, for all* $t \geqslant 0$ *we have*

$$\mathbb{E}\left[\int_0^t \int_E |H_s(x)|\, \lambda_s(dx)\, ds\right] < \infty \quad (1.29)$$

then $\int_0^t \int_E H_s(x)\, \tilde{p}(ds \times dx)$ *is a* $\mathcal{F}_t$*-martingale,*

where $\tilde{p}(ds \times dx) = p(ds \times dx) - \lambda_s(dx)\, ds$.

REMARK: The measure $\tilde{p}(dt \times dx) = p(dt \times dx) - \lambda_t(dx)\, dt$ is usually referred to as the *compensated measure* associated to the marked point process $p(dt \times dx)$. The term $\lambda_t(dx)\, dt$ is commonly called the *compensator*.

To summarize, what we have done is a generalization of the results of the preceding Section. With the last two theorems, we have linked marked point processes to the martingale theory, via the concept of intensity kernel, as in Section 1.1 we have done so by using stochastic intensities.

We can further characterize the intensity kernel and achieve a better understanding of this concept.

DEFINITION 1.6 (Local Characteristics): Let $p(dt \times dx)$ be a E-marked point process with $\mathcal{F}_t$-intensity kernel $\lambda_t(dx)$ of the form

$$\lambda_t(dx) = \lambda_t \Phi_t(dx), \quad (1.30)$$

where λ_t is a nonnegative $\mathcal{F}_t$-predictable process and $\Phi_t(\omega, dx)$ is a probability transition kernel from $(\Omega \times [0, \infty), \mathcal{F} \otimes \mathcal{B}_+)$ into $(E, \mathcal{E})$. The pair $\big(\lambda_t, \Phi_t(dx)\big)$ is called the $\mathcal{F}_t$*-local characteristics of* $p(dt \times dx)$.

Since $\Phi_t(dx)$ is a probability, we have that $\Phi_t(E) = 1$, for all $t \geqslant 0$. We can then identify $\lambda_t \equiv \lambda_t(E)$ with the $\mathcal{F}_t$-intensity of the underlying point process $N_t = N_t(E)$.

An interpretation for the kernel $\Phi_t(dx)$ is given by this theorem.

THEOREM 1.7 ([4, T6, p. 236]): *Let* $p(dt \times dx)$ *be a* E*-marked point process with* $\mathcal{F}_t$*-local characteristics* $\big(\lambda_t, \Phi_t(dx)\big)$*. If the filtration* $\mathcal{F}_t$ *is of the form*

$$\mathcal{F}_t = \mathcal{F}_0 \vee \mathcal{F}_t^p, \tag{1.31}$$

then for all $n \geqslant 1$ *and all* $A \in \mathcal{E}$*,*

$$\Phi_{T_n}(A) = \mathbb{P}\left(Z_n \in A \mid \mathcal{F}_{T_n-}\right) \quad \mathbb{P}-a.s. \quad \text{on } \{T_n < \infty\}, \tag{1.32}$$

where $(T_n)_{n\geqslant 1}$ *is the sequence of the jump times of the underlying point process* $N_t = N_t(E)$*.*

As in the case of the stochastic intensity discussed in the previous Section, we cannot always grant the existence of the local characteristics of a marked point process. The following theorem ensures that the *generalized local characteristics* of a marked point process always exist.

THEOREM 1.8 (Existence of the Generalized Local Characteristics [4, T14, p. 246]): *Let* $p(dt \times dx)$ *be a marked point process adapted to a filtration* $\mathcal{F}_t$ *and assume that the usual conditions are verified for the filtered probability space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$*. Suppose that the mark space* $(E, \mathcal{E})$ *is such that* E *is a Borel subset of a compact metric space and* $\mathcal{E} = \mathcal{B}(E)$*. Then, there exists*

(1) *a unique (up to* $\mathbb{P}$*-indistinguishability) right-continuous* $\mathcal{F}_t$*-predictable nondecreasing process* A_t*, with* $A_0 = 0$*,*

(2) *a probability transition measure* $\Phi_t(\omega, dx)$ *from* $(\Omega \times [0, \infty), \mathcal{F} \otimes \mathcal{B}_+)$ *into* $(E, \mathcal{E})$*, such that, for all* $n \geqslant 1$*,*

$$\mathbb{E}\left[\int_0^\infty \int_E H_s(x)\, p(ds \times dx)\right] = \mathbb{E}\left[\int_0^\infty \int_E H_s(x)\, \Phi_s(dx)\, dA_s\right],$$

for all nonnegative $\mathcal{F}_t$*-predictable* E*-indexed process* H*.*

The pair $\big(A_t, \Phi_t(dx)\big)$ *is called the* generalized $\mathcal{F}_t$-local characteristics of $p(dt \times dx)$.

Nonetheless, but under strict conditions, we can find an explicit form of the local characteristics of a marked point process. This is the content of the following theorem.

THEOREM 1.9 ([4, T7, p. 238]): *Let* $p(dt \times dx)$*, equivalently* $(T_n, \xi_n)_{n\geqslant 1}$*, be a* E*-marked point process. Let* $\mathcal{F}_t$ *be a filtration of the form* $\mathcal{F}_t = \mathcal{F}_0 \vee \mathcal{F}_t^p$*. Suppose that, for each* $n \geqslant 1$*, there exists a regular conditional distribution of* (S_{n+1}, ξ_{n+1}) *given* $\mathcal{F}_{T_n}$ *of the form*

$$\mathbb{P}\left(S_{n+1} \in A, \xi_{n+1} \in C \mid \mathcal{F}_{T_n}\right) = \int_A g^{(n+1)}(s, C)\, ds, \tag{1.33}$$

where $A \in \mathcal{B}_+$*,* $C \in \mathcal{E}$*,* $S_{n+1} = T_{n+1} - T_n$ *and* $g^{(n+1)}(\omega, s, C)$ *is a finite kernel from* $\big(\Omega \times [0, \infty), \mathcal{F}_{T_n} \otimes \mathcal{B}_+\big)$ *into* $(E, \mathcal{E})$*, that is to say:*

(1) $(\omega, s) \mapsto g^{(n+1)}(\omega, s, C)$ *is* $\mathcal{F}_{T_n} \otimes \mathcal{B}_+$*-measurable, for all* $C \in \mathcal{E}$,

(2) *for all* $(\omega, s) \in \Omega \times [0, \infty)$, $C \mapsto g^{(n+1)}(\omega, s, C)$ *is a finite measure on* $(E, \mathcal{E})$.

Then $p(dt \times dx)$ *admits the* $\mathcal{F}_t$*-local characteristics* $\big(\lambda_t, \Phi_t(dx)\big)$ *defined by*

$$\lambda_t(C) = \frac{g^{(n+1)}(t - T_n, C)}{1 - \int_0^{t-T_n} g^{(n+1)}(s, E)\, ds}, \text{ on } (T_n, T_{n+1}], \tag{1.34a}$$

$$\lambda_t = \lambda_t(E), \tag{1.34b}$$

$$\Phi_t(C) = \frac{\lambda_t(C)}{\lambda_t(E)}. \tag{1.34c}$$

We conclude this section by giving the central result of this chapter.

THEOREM 1.10 (Integral Representation of Marked Point Process Martingales [4, T8, p. 239]): *Let* $\mathcal{F}_t$ *be a filtration of the form* $\mathcal{F}_t = \mathcal{F}_0 \vee \mathcal{F}_t^p$ *and for the filtered probability space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geqslant 0}, \mathbb{P})$ *let the usual conditions hold. Let* $p(dt \times dx)$ *be a* E*-marked point process with* $\mathcal{F}_t$*-local characteristics* $\big(\lambda_t, \Phi_t(dx)\big)$. *Then any* $\mathcal{F}_t$*-martingale* M_t *admits the stochastic integral representation*

$$M_t = M_0 + \int_0^t \int_E H_s(x)\, \tilde{p}(ds \times dx) \quad \mathbb{P} - a.s., \quad \forall t \geqslant 0, \tag{1.35}$$

where H *is an* E*-indexed* $\mathcal{F}_t$*-predictable process such that*

$$\int_0^t \int_E |H_s(x)|\, \lambda_s(dx)\, ds < \infty \quad \mathbb{P} - a.s., \quad \forall t \geqslant 0. \tag{1.36}$$

The E*-indexed* $\mathcal{F}_t$*-predictable process* H *in the above representation is essentially unique with respect to the measure* $\mathbb{P}(d\omega)\lambda_t(\omega, dx)$ *on* $\big(\Omega \times [0, \infty) \times E, \tilde{\mathcal{P}}(\mathcal{F}_t)\big)$.
Moreover, if M_t *is square-integrable,* H *satisfies a stronger condition than* (1.36), *namely*

$$\mathbb{E}\left[\int_0^t \int_E |H_s(x)|^2\, \lambda_s(dx)\, ds\right] < \infty \quad \forall t \geqslant 0. \tag{1.37}$$

1.3 FILTERING WITH MARKED POINT PROCESS OBSERVATIONS

Stochastic filtering techniques address the issue of estimating the state at time t of a given dynamical stochastic system, based on the available information at the same time t. A similar problem can be faced when the information is available up to time $t - a$, where a is a strictly positive constant. It then assumes the name of *prediction* of the system's state. Instead, if the observations can be retrieved up to time $t + a$, then the problem is one of *smoothing*.

In the context of second-order stationary processes, two approaches have mainly been used:

- Frequency spectra analysis (Kolmogorov-Wiener).
- Time-domain analysis (Kalman).

Due to the dynamical nature of the problem addressed in this work and the martingale point of view adopted so far, we will use tools that are based on Kalman's innovations theory.

There are two main objects of interest: a *state process* and an *observed process*. The former is an unobserved stochastic process; we are interested in the estimation of its state or, more generally, of the state of a process that depends solely on it. The latter is an observed process, at our disposal to calculate this estimate.

Having in mind this setting, we will proceed along this path:

1. Find the innovating representation of the state process and then project this representation on the natural filtration of the observed process, i. e. the so called *observed history*.
2. Search for filtering formulas, expressed in terms of the innovations gain and of the innovating part, using the representation of the martingales with respect to the observed history.
3. Use the martingale calculus to identify the innovations gain.

1.3.1 *The Innovating Structure of the Filter*

Let X_t and Y_t be two $(E, \mathcal{E})$ valued processes and let $Z_t = h(Y_t)$ be a real-valued process, with h being a measurable function from $(E, \mathcal{E})$ to $(\mathbb{R}, \mathcal{B})$. We interpret X_t as the observation process, Y_t as the state process and Z_t as the process that we aim to filter.

Let $\mathcal{F}_t^X$ and $\mathcal{F}_t^Y$ be the natural filtrations of the processes X_t and Y_t respectively. With the notation

$$\mathcal{F}_t = \mathcal{F}_t^X \vee \mathcal{F}_t^Y,$$
$$\mathcal{O}_t = \mathcal{F}_t^Y,$$

we indicate the *global history* and the *observed history* respectively.

In the sequel we suppose that the process Z_t satisfies the equation

$$Z_t = Z_0 + \int_0^t f_s \, ds + m_t, \quad \mathbb{P} - a.s., \quad \forall t \geqslant 0, \tag{1.38}$$

where

(1) f_t is an $\mathcal{F}_t$-progressive process such that

$$\int_0^t |f_s| \, ds < \infty \quad \mathbb{P} - a.s. \quad \forall t \geqslant 0, \tag{1.39}$$

(2) m_t is a zero mean $\mathcal{F}_t$-local martingale.

Equation (1.38) is called the *semi-martingale representation* of Z_t. In most cases of practical interest, the existence of this representation can be directly exhibited as shown in the following examples.

EXAMPLE 1.5 (Signal Corrupted by a White Noise[3]): Let Y_t be the real-valued process

$$Y_t = Y_0 + \int_0^t S_r \, dr + W_t, \tag{1.40}$$

where

- S_t is a measurable process adapted to $\mathcal{F}_t$ such that

$$\int_0^t |S_r| \, dr < \infty \quad \mathbb{P} - a.s. \quad \forall t \geqslant 0,$$

- W_t is a $\mathcal{F}_t$-Wiener process.

Let $h\colon \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function and let $Z_t = h(Y_t)$. Then, application of Ito's differentiation rule yields

$$Z_t = Z_0 + \int_0^t \left(\frac{\partial h}{\partial y}(Y_r) S_r + \frac{1}{2}\frac{\partial^2 h}{\partial y^2}(Y_r) \right) ds + \int_0^t \frac{\partial h}{\partial y}(Y_r) \, dW_r, \tag{1.41}$$

where the last term in the sum is an Ito's integral. Formula (1.41) is a representation for the process Y_t of type (1.38) with

$$f_t = \frac{\partial h}{\partial y}(Y_r) S_r + \frac{1}{2}\frac{\partial^2 h}{\partial y^2}(Y_r),$$
$$m_t = \int_0^t \frac{\partial h}{\partial y}(Y_r) \, dW_r.$$

EXAMPLE 1.6 (Markov Processes With a Generator): Let Y_t be a E-valued homogeneous $\mathcal{F}_t$-Markov process with the $\mathcal{F}_t$-transition semi-group $(P_t)_{t \geqslant 0}$. We recall this means that, for all $t \geqslant 0$, P_t is a mapping from $b(E)$ into itself[4], such that

$$P_t f = 1, \quad \forall t \geqslant 0, \quad \text{whenever } f(x) = 1 \quad \forall x \in E; \tag{1.42a}$$
$$P_0 = I \text{ (identity)}; \tag{1.42b}$$
$$P_t P_s = P_{t+s}, \quad \forall t \geqslant 0, \forall s \geqslant 0. \tag{1.42c}$$

If we assume that the semigroup $(P_t)_{t \geqslant 0}$ has an infinitesimal generator $\mathcal{L}$ of domain $\mathcal{D}(\mathcal{L})$, then for any $f \in \mathcal{D}(\mathcal{L})$, by application of Dynkin's formula, we obtain

$$f(X_t) = f(X_0) + \int_0^t \mathcal{L} f(X_s) \, ds + m_t, \tag{1.43}$$

where m_t is an $\mathcal{F}_t$-martingale. The representation (1.43) is clearly of the form (1.38) and will be used in Chapter 2.

3 For a background in stochastic processes driven by Wiener-processes, see [13].
4 $b(E)$ is the set of bounded measurable functions from $(E, \mathcal{E})$ into $(\mathbb{R}, \mathcal{B})$.

As previously stated, the first step in the innovations method consists in projecting the semi-martingale representation equation (1.38) on the observed history $\mathcal{O}_t$. This is the content of the following theorem.

THEOREM 1.11 (Projection of the State [4, T1, p. 87]): *Let Z_t be an integrable real-valued process with the semi-martingale representation:*

$$Z_t = Z_0 + \int_0^t f_s \, ds + m_t,$$

where

(I) f_t *is an $\mathcal{F}_t$-progressive process such that*

$$\mathbb{E}\left[\int_0^t |f_s| \, ds\right] < \infty \quad \forall t \geqslant 0,$$

(II) m_t *is a zero mean $\mathcal{F}_t$-martingale.*

Let $\mathcal{O}_t$ be a filtration such that $\mathcal{O}_t \subset \mathcal{F}_t$, $\forall t \geqslant 0$. Then

$$\mathbb{E}\left[Z_t \mid \mathcal{O}_t\right] = \mathbb{E}\left[Z_0 \mid \mathcal{O}_0\right] + \int_0^t \hat{f}_s \, ds + \hat{m}_t, \tag{1.44}$$

where

(1) $\hat{m}_t$ *is a zero mean $\mathcal{O}_t$-martingale,*

(2) $\hat{f}_t$ *is a $\mathcal{O}_t$-progressive process defined by*

$$\mathbb{E}\left[\int_0^t C_s \, f_s \, ds\right] = \mathbb{E}\left[\int_0^t C_s \, \hat{f}_s \, ds\right], \tag{1.45}$$

for all nonnegative bounded $\mathcal{O}_t$-progressive processes C_t.

REMARK: The rather abstract definition of the process $\hat{f}_t$ might seem daunting at a first glance, and we may wonder how to explicitly calculate it or even if it exists. The following remarks will help clarifying these issues.

(A) The existence of $\hat{f}_t$ is always granted by the Radon-Nikodym derivative theorem. Let μ_1 and μ_2 be two measures defined on $(\Omega \times (0, \infty), \mathrm{prog}\mathcal{O}_t)$ by

$$\mu_1(d\omega \times dt) = \mathbb{P}(d\omega) \, dt,$$
$$\mu_2(d\omega \times dt) = \mathbb{P}(d\omega) \, f_t(\omega) dt,$$

Then $\hat{f}_t(\omega)$ is the Radon-Nikodym derivative of the measure μ_2 with respect to the measure μ_1. Moreover, two versions of $\hat{f}_t$ differ only on a set of μ_1-measure zero.

(B) Suppose that there exists, for all $t \geqslant 0$, a version of $\mathbb{E}[f_t \mid \mathcal{O}_t] = \tilde{f}_t$ such that the mapping $(\omega, t) \mapsto \tilde{f}_t(\omega)$ is $\mathcal{O}_t$-progressively measurable. Then, by setting $\hat{f}_t(\omega) = \tilde{f}_t(\omega)$, we satisfy the requirements for the definition of the process $\hat{f}_t$. Indeed, applying the Fubini theorem, we obtain

$$\mathbb{E}\left[\int_0^t C_s\, f_s\, ds\right] = \int_0^t \mathbb{E}\left[C_s\, f_s\right]\, ds = \int_0^t \mathbb{E}\left[\mathbb{E}\left[C_s\, f_s \mid \mathcal{O}_s\right]\right]\, ds =$$
$$\int_0^t \mathbb{E}\left[C_s\, \mathbb{E}\left[f_s \mid \mathcal{O}_s\right]\right]\, ds = \int_0^t \mathbb{E}\left[C_s\, \tilde{f}_s\right]\, ds = \mathbb{E}\left[\int_0^t C_s\, \tilde{f}_s\, ds\right].$$

In the applications, this version of $\mathbb{E}[f_t \mid \mathcal{O}_t]$ usually exists, but cannot be granted in general, because *a priori* nothing is known about the measurability in t of $\mathbb{E}[f_t \mid \mathcal{O}_t]$.

1.3.2 *Filtering Equations*

We now assume that the observation process is an E-marked point process $p(dt \times dx)$, adapted to the filtration $\mathcal{F}_t$. The observed history has the form $\mathcal{O}_t = \mathcal{G}_0 \vee \mathcal{F}_t^p$, where $\mathcal{F}_t^p$ is the natural filtration of the marked point process and $\mathcal{G}_0 \subset \mathcal{F}_0$. Moreover, we suppose that $p(dt \times dx)$ admits the $\mathcal{F}_t$-local characteristics $\left(\lambda_t, \Phi_t(dx)\right)$ and the $\mathcal{O}_t$-local characteristics $\left(\hat{\lambda}_t, \hat{\Phi}_t(dx)\right)$.

For technical reasons, the usual conditions stated in section 1.1 are assumed to hold for the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and for all the filtrations here specified.

Let Z_t be a real-valued state process satisfying the conditions stated in theorem 1.11. We add the following

ASSUMPTION (H): The semi-martingale representation of Z_t is such that

(H1) $m_t = m_t^d + m_t^c$, where m_t^d is a $\mathcal{F}_t$-martingale of integrable variation over finite intervals and m_t^c is a continuous $\mathcal{F}_t$-martingale.

(H2) $Z_t - m_t^c$ is a bounded process.

We are now in a position to state the central result of this section. In fact, recalling the representation theorem 1.10, we can express in a more precise form the $\mathcal{O}_t$-martingale $\hat{m}_t$ that figures in equation (1.44). This can be done since the filtration $\mathcal{O}_t$ is, apart from the initial σ-algebra $\mathcal{G}_0$, the natural filtration of the marked point process $p(dt \times dx)$.

THEOREM 1.12 (Filtering Theorem [4, T9, p. 240]): *Let the conditions stated in this subsection and the assumption (H) hold. Then for all* $t \geqslant 0$ *and* $\mathbb{P}-a.s.$

$$\hat{Z}_t = \mathbb{E}\left[Z_t \mid \mathcal{O}_t\right] = \mathbb{E}\left[Z_0 \mid \mathcal{O}_0\right] + \int_0^t \hat{f}_s\, ds + \\ + \int_0^t \int_E K_s(x)\left[p(ds \times dx) - \hat{\lambda}_s \hat{\Phi}_s(dx)\, ds\right]. \quad (1.46)$$

The process $K_t(x)$ *is a* $\mathcal{O}_t$*-predictable process indexed by* E*, that is defined* $\mathbb{P}(d\omega)p(dt \times dx)$*-essentially uniquely by*

$$K_t(x) = \Psi_t^1(x) - \Psi_t^2(x) + \Psi_t^3(x). \quad (1.47)$$

The processes $\Psi_t^i(x)$*,* $i = 1, 2, 3$*, are* $\mathcal{O}_t$*-predictable processes indexed by* E *and are* $\mathbb{P}(d\omega)\, p(dt \times dx)$*-essentially uniquely defined by the following equalities holding for all* $t \geqslant 0$ *and for all bounded* $\mathcal{O}_t$*-predictable processes* $C_t(x)$ *indexed by* E*:*

$$\begin{aligned} \mathbb{E}\left[\int_0^t \int_E \Psi_s^1(x) C_s(x) \hat{\lambda}_s(dx) ds\right] &= \mathbb{E}\left[\int_0^t \int_E Z_s C_s(x) \lambda_s(dx) ds\right], \\ \mathbb{E}\left[\int_0^t \int_E \Psi_s^2(x) C_s(x) \hat{\lambda}_s(dx) ds\right] &= \mathbb{E}\left[\int_0^t \int_E Z_s C_s(x) \hat{\lambda}_s(dx) ds\right], \quad (1.48) \\ \mathbb{E}\left[\int_0^t \int_E \Psi_s^3(x) C_s(x) \hat{\lambda}_s(dx) ds\right] &= \mathbb{E}\left[\int_0^t \int_E \Delta Z_s C_s(x) p(ds \times dx)\right]. \end{aligned}$$

REMARK: The existence of the processes $\Psi_t^1(x)$, $\Psi_t^2(x)$ and $\Psi_t^3(x)$, and in turn of the process $K_t(x)$, is granted again by the Radon-Nikodym derivative theorem. In fact:

(1) $\Psi_t^1(x)$ is the Radon-Nikodym derivative of the measure $\mu_1^1(d\omega \times dt \times dx)$ with respect to the measure $\mu_2^1(d\omega \times dt \times dx)$, where

$$\mu_1^1(d\omega \times dt \times dx) = \mathbb{P}(d\omega)\, Z_t(\omega)\, \lambda_t(\omega, dx)\, dt,$$
$$\mu_2^1(d\omega \times dt \times dx) = \mathbb{P}(d\omega)\, \hat{\lambda}_t(\omega, dx)\, dt.$$

Both measures are defined on $\left(\Omega \times (0, \infty) \times E, \tilde{\mathcal{P}}(\mathcal{O}_t)\right)$. The first one is a signed measure, is σ-finite by the assumption of boundedness of the process Z_t, and is absolutely continuous with respect to the second one. Moreover, being a Radon-Nikodym derivative, the process $\Psi_t^1(x)$ is $\mathcal{P}(\mathcal{O}_t)$-measurable, i. e. it is a $\mathcal{G}_t$-predictable process.

(2) $\Psi_t^2(x)$ is the Radon-Nikodym derivative of the measure $\mu_1^2(d\omega \times dt \times dx)$ with respect to the measure $\mu_2^2(d\omega \times dt \times dx)$, where

$$\mu_1^2(d\omega \times dt \times dx) = \mathbb{P}(d\omega)\, Z_t(\omega)\, \hat{\lambda}_t(\omega, dx)\, dt,$$
$$\mu_2^2(d\omega \times dt \times dx) = \mathbb{P}(d\omega)\, \hat{\lambda}_t(\omega, dx)\, dt.$$

Similar considerations to the ones made for the process $\Psi_t^1(x)$ apply to this process.

(3) $\Psi_t^3(x)$ is the Radon-Nikodym derivative of the measure $\mu_1^3(d\omega \times dt \times dx)$ with respect to the measure $\mu_2^3(d\omega \times dt \times dx)$, where

$$\mu_1^3(d\omega \times dt \times dx) = \mathbb{P}(d\omega)\, dZ_t(\omega)\, p(dt \times dx),$$
$$\mu_2^3(d\omega \times dt \times dx) = \mathbb{P}(d\omega)\, \hat{\lambda}_t(\omega, dx)\, dt.$$

Both measures are defined on $\big(\Omega \times (0,\infty) \times E, \tilde{\mathcal{P}}(\mathcal{O}_t)\big)$. The first one is a signed measure, is σ-finite since Z_t and hence $|\Delta Z_t|$ is bounded, and is absolutely continuous with respect to the second one, because on the space of definition of these measures, $\mathbb{P}(d\omega)\, \hat{\lambda}_t(\omega, dx)\, dt = \mathbb{P}(d\omega)\, p(dt \times dx)$. The $\mathcal{G}_t$-predictability of the process $\Psi_t^3(x)$ comes from the same arguments applied to the processes $\Psi_t^1(x)$ and $\Psi_t^2(x)$.

We end this chapter with a consideration very useful in the application of the filtering formula (1.46). The process $\Psi_t^2(x)$ is $\mathbb{P}-a.s.$ equal to the process $\hat{Z}_{t-}$. Indeed, we can develop the second relation in (1.48) using the Fubini theorem, like this:

$$\begin{aligned}
&\mathbb{E}\left[\int_0^t \int_E C_s(x)\, Z_s\, \hat{\lambda}_s(dx)\, ds\right] = \int_0^t \int_E \mathbb{E}\left[C_s(x)\, Z_s\, \hat{\lambda}_s(dx)\right] ds = \\
&= \int_0^t \int_E \mathbb{E}\left[\mathbb{E}\left[C_s(x)\, Z_s\, \hat{\lambda}_s(dx) \mid \mathcal{O}_s\right]\right] ds = \\
&= \int_0^t \int_E \mathbb{E}\left[C_s(x)\, \hat{\lambda}_s(dx)\, \mathbb{E}\left[Z_s \mid \mathcal{O}_s\right]\right] ds = \\
&= \int_0^t \int_E \mathbb{E}\left[C_s(x)\, \hat{\lambda}_s(dx)\, \hat{Z}_s\right] ds = \mathbb{E}\left[\int_0^t \int_E C_s(x)\, \hat{Z}_s\, \hat{\lambda}_s(dx)\, ds\right] = \\
&= \mathbb{E}\left[\int_0^t \int_E C_s(x)\, \hat{Z}_{s-}\, \hat{\lambda}_s(dx)\, ds\right],
\end{aligned}$$

since $\hat{Z}_s\, ds = \hat{Z}_{s-}\, ds$. Hence one always has

$$\Psi_t^2(x) = \hat{Z}_{t-} \tag{1.49}$$

2 FILTERING EQUATION IN THE NOISE-FREE MODEL

In this Chapter we are going to address the object of study of this Thesis: the stochastic filtering of a time-homogeneous pure jump Markov process with noise-free observation. In the sequel we will assume defined:

- A complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
- A complete separable metric space $(I, \mathcal{I})$, where $\mathcal{I} = \mathcal{B}(I)$,
- A measurable space $(O, \mathcal{O})$, where O is a finite set of cardinality $K \in \mathbb{N}$ and $\mathcal{O} = 2^O$ is the power set of O,
- A surjective measurable function $h \colon I \to O$.

We immediately notice that the function h creates a partition on the set I, given by the pre-images of the points of O. In fact, if we denote by $a_1, \dots, a_K$ the elements of the set O, then the $\mathcal{I}$-measurable sets $A_k = h^{-1}(a_k)$, $k = 1, \dots, K$, are such that

(i) $A_k \neq \emptyset$, for all $k = 1, \dots, K$,

(ii) $A_i \cap A_j = \emptyset$, for all $i \neq j$, $i, j = 1, \dots, K$,

(iii) $\bigcup_{k=1}^K A_k = I$,

so they are a partition of the set I. This is an important property that will be crucial in the sequel.

2.1 THE NOISE-FREE MODEL

In the noise-free model that we are going to analyze there are two main objects to define: the *unobserved process* X_t and the *observed process* Y_t.

The unobserved process X_t is a I-valued pure-jump Markov process defined, for all $t \geqslant 0$ and all $\omega \in \Omega$, by

$$X_t(\omega) = \sum_{n \in \mathbb{N}} \xi_n(\omega) \mathbb{1}\big(T_n(\omega) \leqslant t < T_{n+1}(\omega)\big) \mathbb{1}\big(T_n(\omega) < +\infty\big). \quad (2.1)$$

The random quantities that appear in (2.1) are:

(i) The sequence of jump times $\big(T_n(\omega)\big)_{n \in \mathbb{N}}$, where

- $T_0(\omega) = 0$, for all $\omega \in \Omega$,

- $T_n : \Omega \to (0, +\infty]$, $n = 1, 2, \ldots$, are measurable random variables, such that, for all $\omega \in \Omega$ and all $n = 1, 2, \ldots$, the condition (1.1b) holds.

We will denote by $T_\infty(\omega) = \lim_{n\to\infty} T_n(\omega)$, $\omega \in \Omega$, the explosion point of the process X_t. We suppose that $T_\infty = +\infty$ $\mathbb{P} - a.s.$, i. e. that the unobserved process is $\mathbb{P} - a.s.$-nonexplosive.

(II) The sequence of random values of the process X_t, a collection of measurable random variables $\xi_n : \Omega \to I$, $n \in \mathbb{N}$. The law of the initial value is known and we will denote it by $\mu(dx)$, i. e. $\xi_0 \sim \mu$.

We will indicate the natural filtration of the process X_t by $\mathcal{F}_t^X$. Other known objects, that are related to the unobserved process, are:

(A) The *rate transition function* $\lambda : I \to [0, +\infty)$. It is a $\mathcal{I}$-measurable function that determines the rate parameter of the exponential distribution that characterizes the holding times of the process X_t, i. e.

$$\mathbb{P}\left(T_{n+1} - T_n > t \mid \mathcal{F}_{T_n}^X\right) = e^{-\lambda(\xi_n)t}, \quad t \geqslant 0, \quad n \in \mathbb{N}. \tag{2.2}$$

(B) The *probability transition kernel* $q(x, dy)$, a function such that:

- $x \mapsto q(x, A)$ is $\mathcal{I}$-measurable, for all $A \in \mathcal{I}$,
- $A \mapsto q(x, A)$ is a probability measure on $(I, \mathcal{I})$, for all $x \in I$.

It characterizes the distribution of the random values of the process X_t, in the sense that

$$\mathbb{P}\left(\xi_{n+1} \in A \mid \mathcal{F}_{T_n}^X\right) = q(\xi_n, A), \quad \forall A \in \mathcal{I}, \quad n \in \mathbb{N}. \tag{2.3}$$

(C) The *rate transition measure* $\lambda(x, dy) = \lambda(x)q(x, dy)$. We adopt the notations $\lambda(x, A)$ for $\int_A \lambda(x, dy)$, $A \in \mathcal{I}$, and $\lambda(x)$ for $\int_I \lambda(x, dy)$.

The observed process is simply defined by

$$Y_t(\omega) = h\big(X_t(\omega)\big), \quad \forall t \geqslant 0, \quad \forall \omega \in \Omega. \tag{2.4}$$

As in the case of the unobserved process, we can define:

(I) The sequence of jump times $\big(\tau_n(\omega)\big)_{n\in\mathbb{N}}$ of the process Y_t. They are related to the jump times of the process X_t, that is to say $\tau_n(\omega) = T_k(\omega)$ for some $k \geqslant n$, $k = k(\omega)$, $\omega \in \Omega$.

(II) The sequence of random values of the process Y_t, denoted by $\big(\zeta_n(\omega)\big)_{n\in\mathbb{N}}$. They are related to the random values of the process X_t, that is to say $\zeta_n(\omega) = h\big(\xi_k(\omega)\big)$ for some $k \geqslant n$, $k = k(\omega)$, $\omega \in \Omega$.

We will denote the natural filtration of the process Y_t by $\mathcal{F}_t^Y$.

In the rest of the work we will suppose that

$$\Lambda = \sup_{x \in I} \lambda(x) < +\infty. \qquad (2.5)$$

This hypothesis is important, since it eliminates the need to explicitly assume that the process X_t is $\mathbb{P}-a.s.$-nonexplosive. In fact, (2.5) implies that $T_\infty = +\infty$ $\mathbb{P}-a.s.$. Moreover, all the filtrations defined here and in the sequel are supposed to have been properly modified in order to satisfy the usual conditions stated in Section 1.1.

The model is now completely defined. As we saw, few quantities are needed to specify it, namely the unobserved process X_t, the observed process Y_t, the functions $\lambda(x)$ and $h(x)$, the probability transition kernel $q(x, dy)$ and the initial distribution $\mu(dx)$ of the process X_t.

2.1.1 *The Marked Point Process Formulation*

To apply the martingale techniques earlier announced, we have to formulate the noise-free model in terms of marked point processes.

To start, we notice that the pairs $(T_n, \xi_n)_{n \geqslant 1}$ naturally define a marked point process of mark space I associated to X_t. Then, we can link to X_t a counting process, defining for all $t \geqslant 0$, all $A \in \mathcal{I}$ and all $\omega \in \Omega$

$$N_t(\omega, A) = \sum_{n \geqslant 1} \mathbb{1}(\xi_n(\omega) \in A)\, \mathbb{1}(T_n(\omega) \leqslant t). \qquad (2.6)$$

As we explained in Section 1.2, a random measure is naturally related to this counting process, namely

$$p(\omega, (0, t] \times A) = N_t(\omega, A), \quad t \geqslant 0,\ A \in \mathcal{I},\ \omega \in \Omega. \qquad (2.7)$$

We will, then, identify the marked point process $(T_n, \xi_n)_{n \geqslant 1}$ with the measure $p(dt \times dy)$.

It is a known fact that the marked point process $p(dt \times dy)$, defined in connection with the pure-jump Markov process X_t of known rate transition measure $\lambda(x, dy)$, admits the pair $\big(\lambda_t, \Phi_t(dy)\big)$ as its $\mathcal{F}_t^X$-local characteristics given by:

- $\lambda_t = \lambda(X_{t-}), \quad t \geqslant 0,$
- $\Phi_t(dy) = q(X_{t-}, dy), \quad t \geqslant 0.$

This is shown by a simple application of Theorem 1.9. The hypotheses are all satisfied, since the natural filtration is of the form $\mathcal{F}_t^X = \sigma(X_0) \vee \mathcal{F}_t^p$ and the regular conditional distribution of (S_{n+1}, ξ_{n+1}) given $\mathcal{F}_{T_n}$ is a finite kernel from $\big(\Omega \times [0, \infty), \mathcal{F}_{T_n} \otimes \mathcal{B}_+\big)$ into $(I, \mathcal{I})$. To find its expression, it suffices to remember the conditional independence of the holding times S_n and the random values ξ_n. Then, applying formulas (2.2) and (2.3), for all $A \in \mathcal{I}$ and all $n \in \mathbb{N}$ we obtain

$$g^{(n+1)}(t, A) = \lambda(X_{T_n})\, e^{-\lambda(X_{T_n})t}\, q(X_{T_n}, A), \quad t \in (T_n, T_{n+1}]. \tag{2.8}$$

Then the equations (1.34) yield the announced result.

We can, then, define the compensated measure associated to the marked point process $p(dt \times dy)$ as

$$\tilde{p}(dt \times dy) = p(dt \times dy) - \lambda_t(dy)\, dt, \tag{2.9}$$

where $\lambda_t(dy) = \lambda_t\, \Phi_t(dy)$.

Regarding the observed process Y_t, its sequence $(\tau_n, \zeta_n)_{n\geqslant 1}$ of jump times and random values defines, too, a marked point process, of mark space O. Due to the finiteness of its mark space, it presents itself in the simpler form of a K-variate point process. Thus, remembering Example 1.4, we define $N_t^Y = \big(N_t^Y(1), \dots, N_t^Y(K)\big)$, where for all $k = 1, \dots, K$

$$N_t^Y(k) = \sum_{n\geqslant 1} \mathbb{1}(\zeta_n = a_k)\, \mathbb{1}(\tau_n \leqslant t), \quad t \geqslant 0. \tag{2.10}$$

The marked point process formulation of the noise-free model is now complete. With the processes $p(dt \times dy)$, of known local characteristics $\big(\lambda_t, \Phi_t(dy)\big)$, and N_t^Y we can proceed in the analysis of the filtering problem.

2.2 THE FILTERING PROBLEM

Having in mind the previously defined setting, we are now going to address the filtering problem. Before starting, let us define the following filtrations:

$$\mathcal{F}_t^k = \sigma(N_s^Y(k), 0 \leqslant s \leqslant t), \quad k = 1, \dots, K, \quad t \geqslant 0, \tag{2.11a}$$

$$\mathcal{G}_t = \bigvee_{k=1}^{K} \mathcal{F}_t^k, \quad t \geqslant 0, \tag{2.11b}$$

$$\mathcal{O}_t = \mathcal{G}_t \vee \sigma(Y_0), \quad t \geqslant 0, \tag{2.11c}$$

$$\mathcal{F}_t = \mathcal{F}_t^X \vee \mathcal{O}_t, \quad t \geqslant 0. \tag{2.11d}$$

In the sequel, the relevant filtrations will be $\mathcal{F}_t$ and $\mathcal{O}_t$, named respectively the *global history* and the *observed history* using the terminology adopted in Section 1.3.

We notice that, in our model, the filtration $\mathcal{O}_t$ is such that $\mathcal{O}_t \subset \mathcal{F}_t^X$, for all $t \geqslant 0$. By construction, the filtration $\mathcal{O}_t$ coincides with the natural filtration $\mathcal{F}_t^Y$ of the observed process Y_t (simply recall how the point process N_t^Y is defined). Finally, being X_t and Y_t linked by the surjective function h, it is clear that $\mathcal{F}_t^Y \subset \mathcal{F}_t^X$, for all $t \geqslant 0$. This, in turn, implies that $\mathcal{F}_t \equiv \mathcal{F}_t^X$, for all $t \geqslant 0$. Albeit having the same meaning, we will prefer the easier notation $\mathcal{F}_t$ just introduced. It is important, though, to keep in mind this equivalence.

Let us introduce the real-valued process

$$Z_t(A) = \mathbb{1}(X_t \in A), \quad t \geqslant 0, \quad A \in \mathcal{I}. \tag{2.12}$$

The filtering problem consists in finding an explicit expression for the *filtering process*

$$\hat{Z}_t(A) = \mathbb{E}\left[Z_t(A) \mid \mathcal{O}_t\right] = \mathbb{P}\left(X_t \in A \mid \mathcal{O}_t\right), \quad \forall t \geqslant 0, \quad \forall A \in \mathcal{I}. \tag{2.13}$$

Henceforward, we will assume fixed, once and for all, a set $A \in \mathcal{I}$. We have now to focus on verifying the hypotheses of the filtering theorem 1.12 in order to apply it and achieve the desired result.

2.2.1 *Application of the Filtering Theorem*

Following the lines of Subsection 1.3.2, we will now check the hypotheses stated in Theorem 1.12. We immediately notice that, being the mark space O a finite set, the counting measure $\pi(dt \times dz)$ associated to the observed process takes a simpler form. This allows us to use an easier notation, namely

$$\pi(dt \times dz) = \pi(dt \times \{a_k\}) = dN_t^Y(k), \quad t \geqslant 0, \quad k = 1, \dots, K. \tag{2.14}$$

Further simplifications will derive from the finiteness of the set O, that will be progressively exhibited.

Concerning the measurability of the process Y_t, it is clearly $\mathcal{F}_t$- and $\mathcal{O}_t$-adapted and the filtration $\mathcal{O}_t$ has the requested form, given in (2.11c).

We have now to search for the $\mathcal{F}_t$- and $\mathcal{O}_t$-local characteristics of the marked point process N_t^Y. We will denote them by $\left(\lambda_t^Y, \Phi_t^Y(dz)\right)$ and $\left(\hat{\lambda}_t^Y, \hat{\Phi}_t^Y(dz)\right)$, respectively. They take a simpler form, because $\Phi_t^Y(dz)$ is a discrete probability measure on $(O, \mathcal{O})$. If we concentrate our attention on the single atom $\{a_k\} \in \mathcal{O}$, $k \in \{1, \dots, K\}$, it is not difficult to see that $\lambda_t^Y \cdot \Phi_t^Y(\{a_k\})$ is the $\mathcal{F}_t$-stochastic intensity of the point process $N_t^Y(k)$. Then we have to identify the $\mathcal{F}_t$- and $\mathcal{O}_t$-stochastic intensities of the K-variate point process $N_t^Y = \left(N_t^Y(1), \dots, N_t^Y(K)\right)$, for all $k = 1, \dots, K$. We will denote them by $\left(\lambda_t^Y(1), \dots, \lambda_t^Y(K)\right)$ and $\left(\hat{\lambda}_t^Y(1), \dots, \hat{\lambda}_t^Y(K)\right)$, respectively.

Let us, first, show this simple but useful Lemma.

Lemma 2.1: *Let* X_t *be the pure-jump* $\mathcal{F}_t$*-Markov process defined in* (2.1) *and let* $N_t^X(A, B)$ *be the point process*

$$N_t^X(A, B) = \sum_{0 < s \leqslant t} \mathbb{1}(X_{s-} \in A)\, \mathbb{1}(X_s \in B), \quad t \geqslant 0, \quad A, B \in \mathcal{I}. \tag{2.15}$$

Then it admits $\mathcal{F}_t$- and $\mathcal{O}_t$-stochastic intensities, respectively given by

$$\lambda_t^X(A, B) = \mathbb{1}(X_{t-} \in A)\, \lambda(X_{t-}, B), \quad t \geqslant 0, \tag{2.16a}$$

$$\hat{\lambda}_t^X(A, B) = \int_A \lambda(x, B)\, \hat{Z}_{t-}(dx), \quad t \geqslant 0, \tag{2.16b}$$

where $\lambda(x, dy)$ is the rate transition measure of the process X_t.

Proof. Let $t \geqslant 0$ and $A, B \in \mathcal{I}$ be fixed. Recalling that $p(dt \times dy)$ is the I-marked point process associated to X_t, the point process $N_t^X(A, B)$ can be written as

$$\begin{aligned} N_t^X(A, B) &= \sum_{0<s\leqslant t} \mathbb{1}(X_{s-} \in A)\, \mathbb{1}(X_s \in B) \\ &= \int_0^t \int_I \mathbb{1}(X_{s-} \in A)\, \mathbb{1}(y \in B)\, p(ds \times dy) \\ &= \int_0^t \int_I \mathbb{1}(X_{s-} \in A)\, \mathbb{1}(y \in B)\, \big[\tilde{p}(ds \times dy) + \lambda_s(dy)\, ds\big], \end{aligned}$$

where $\lambda_t(dy)dt$ is the compensator of $p(dt \times dy)$ and $\tilde{p}(dt \times dy)$ is its compensated measure.

The process $\mathbb{1}(X_{t-} \in A)\, \mathbb{1}(y \in B)$ is a I-indexed $\mathcal{F}_t$-predictable process, with I-indexed part given by $\mathbb{1}(y \in B)$ and $\mathcal{F}_t$-predictable part $\mathbb{1}(X_{t-} \in A)$. Moreover, remembering that $\lambda_t(dy) = \lambda(X_{t-}, dy)$, we have

$$\begin{aligned} &\mathbb{E}\left[\int_0^t \int_I \Big|\mathbb{1}(X_{s-} \in A)\, \mathbb{1}(y \in B)\Big|\, \lambda_s(dy)\, ds\right] = \\ &= \mathbb{E}\left[\int_0^t \mathbb{1}(X_{s-} \in A) \int_I \mathbb{1}(y \in B)\, \lambda(X_{s-}, dy)\, ds\right] = \\ &= \mathbb{E}\left[\int_0^t \mathbb{1}(X_{s-} \in A)\lambda(X_{s-}, B) ds\right] \leqslant \mathbb{E}\left[\int_0^t \lambda(X_{s-}) ds\right] \leqslant \Lambda t < +\infty, \end{aligned}$$

where Λ is finite by assumption (2.5).

This enables us to apply the Integration Theorem 1.6 and obtain that

$$\int_0^t \int_I \mathbb{1}(X_{s-} \in A)\, \mathbb{1}(y \in B)\, \tilde{p}(dt \times dy)$$

is a $\mathcal{F}_t$-martingale, which means that the process

$$N_t^X(A, B) - \int_0^t \mathbb{1}(X_{s-} \in A)\, \lambda(X_{s-}, B)\, ds,$$

is a $\mathcal{F}_t$-martingale.

Finally, by virtue of Theorem 1.3, we identify the $\mathcal{F}_t$-predictable process

$$\lambda_t^X(A, B) = \mathbb{1}(X_{t-} \in A)\, \lambda(X_{t-}, B)$$

as the $\mathcal{F}_t$-stochastic intensity of the point process $N_t^X(A, B)$.

Concerning the $\mathcal{O}_t$-stochastic intensity $\hat{\lambda}_t^X(A,B)$, we can use directly Definition 1.2. In fact, we have just proved the equivalence stated in equation (1.9), i. e.

$$\mathbb{E}\left[\int_0^\infty C_s \, dN_s^X(A,B)\right] = \mathbb{E}\left[\int_0^\infty C_s \, \lambda_s^X(A,B)\, ds\right],$$

for all nonnegative $\mathcal{F}_t$-predictable processes C_t. If, in particular, we restrict our attention to the subset formed by $\mathcal{O}_t$ predictable processes C_t, the last equality still holds, but the process $\lambda_t^X(A,B)$ is not $\mathcal{O}_t$-progressive as required by Definition 1.2. However, observing that $X_{t-}\, dt = X_t\, dt$, a simple application of the Fubini-Tonelli Theorem shows that

$$\begin{aligned}
&\mathbb{E}\left[\int_0^\infty C_s\,\lambda_s^X(A,B)\,ds\right] = \int_0^\infty \mathbb{E}\left[C_s\,\lambda_s^X(A,B)\right] ds = \\
&= \int_0^\infty \mathbb{E}\left[\mathbb{E}\left[C_s\,\lambda_s^X(A,B) \mid \mathcal{O}_s\right]\right] ds = \int_0^\infty \mathbb{E}\left[C_s\,\mathbb{E}\left[\lambda_s^X(A,B) \mid \mathcal{O}_s\right]\right] ds = \\
&= \int_0^\infty \mathbb{E}\left[C_s\,\mathbb{E}\left[\mathbb{1}(X_{s-} \in A)\,\lambda(X_{s-},B) \mid \mathcal{O}_s\right]\right] ds = \\
&= \int_0^\infty \mathbb{E}\left[C_s\,\mathbb{E}\left[\mathbb{1}(X_s \in A)\,\lambda(X_s,B) \mid \mathcal{O}_s\right]\right] ds.
\end{aligned}$$

We notice that we can express the conditional expectation appearing in the last formula, as the integral of the bounded $\mathcal{I}$-measurable function $\mathbb{1}(x \in A)\,\lambda(x,B)$ on the set I, with respect to the conditional law of the random variable X_s given $\mathcal{O}_s$. Indeed, the boundedness comes from assumption (2.5) and the existence of the aforementioned conditional law is granted thanks to the hypothesis that $(I, \mathcal{I})$ is a complete separable metric space. Considering that and $\hat{Z}_{t-}(dx)\,dt = \hat{Z}_t(dx)\,dt$, we can write

$$\begin{aligned}
&\int_0^\infty \mathbb{E}\left[C_s\,\mathbb{E}\left[\mathbb{1}(X_s \in A)\,\lambda(X_s,B) \mid \mathcal{O}_s\right]\right] ds = \\
&= \int_0^\infty \mathbb{E}\left[C_s \int_I \mathbb{1}(x \in A)\,\lambda(x,B)\,\hat{Z}_s(dx)\right] ds = \\
&= \mathbb{E}\left[\int_0^\infty C_s \int_A \lambda(x,B)\,\hat{Z}_s(dx)\,ds\right] = \mathbb{E}\left[\int_0^\infty C_s \int_A \lambda(x,B)\,\hat{Z}_{s-}(dx)\,ds\right].
\end{aligned}$$

Defining

$$\hat{\lambda}_t^X(A,B) = \int_A \lambda(x,B)\,\hat{Z}_{t-}(dx), \quad t \geqslant 0,$$

we observe that the process $\hat{\lambda}_t^X(A,B)$ is $\mathcal{O}_t$-adapted by definition and the last formula shows that its trajectories are left-continuous, due to the presence of the term $\hat{Z}_{t-}(\cdot)$. A known result from the theory of stochastic processes[1] ensures that it is a $\mathcal{O}_t$-predictable process and, therefore, $\mathcal{O}_t$-progressive.

1 See Appendix A, Theorem A.1

Finally, recalling that $\hat{Z}_{t-}(dx)$ is a probability, we can verify condition (1.8) in Definition 1.2. Indeed, for all $t \geqslant 0$ and $\mathbb{P}-a.s.$, we have

$$\int_0^t \hat{\lambda}_t^X(A,B) = \int_0^t \int_A \lambda(x,B)\,\hat{Z}_{t-}(dx) \leqslant \int_0^t \lambda(x) \int_A \hat{Z}_{t-}(dx) \leqslant$$
$$\leqslant \int_0^t \lambda(x) \int_I \hat{Z}_{t-}(dx) \leqslant \int_0^t \lambda(x) \leqslant \Lambda t < +\infty.$$

We can, then, affirm that (2.16b) is the $\mathcal{O}_t$-stochastic intensity of the point-process $N_t^X(A,B)$. □

Going back to the main topic, we can now explicitly write down the $\mathcal{F}_t$- and $\mathcal{O}_t$-stochastic intensities of the K-variate point process N_t^Y. Henceforth, the superscript ^{c} will denote the complement set with respect to I.

Proposition 2.1: *The* K*-variate point process* N_t^Y *defined in* (2.10) *admits* $\big(\lambda_t^Y(1),\ldots,\lambda_t^Y(K)\big)$ *and* $\big(\hat{\lambda}_t^Y(1),\ldots,\hat{\lambda}_t^Y(K)\big)$ *as its* $\mathcal{F}_t$*- and* $\mathcal{O}_t$*-stochastic intensities respectively, given by*

$$\lambda_t^Y(k) = \mathbb{1}(X_{t-} \in A_k^c)\,\lambda(X_{t-},A_k), \quad t \geqslant 0, \quad k = 1,\ldots,K, \tag{2.17a}$$
$$\hat{\lambda}_t^Y(k) = \int_{A_k^c} \lambda(x,A_k)\,\hat{Z}_{t-}(dx), \quad t \geqslant 0, \quad k = 1,\ldots,K, \tag{2.17b}$$

where $(A_k)_{1\leqslant k\leqslant K}$ *is the partition induced on the set* I *by the function* h *and, for all* $t \geqslant 0$, $\hat{Z}_t(dx)$ *is a regular version of the conditional distribution of the random variable* X_t *given* $\mathcal{O}_t$*, as defined in* (2.13)*.*

Proof. A straightforward application of Lemma 2.1 yields the expressions (2.17). In fact, for fixed $t \geqslant 0$ and $k = 1,\ldots,K$, we can write

$$N_t^Y(k) = \sum_{0<s\leqslant t} \mathbb{1}(Y_{s-} \neq a_k)\,\mathbb{1}(Y_s = a_k) =$$
$$= \sum_{0<s\leqslant t} \mathbb{1}(X_{s-} \in A_k^c)\,\mathbb{1}(X_s \in A_k) = N_t^X(A_k^c, A_k).$$

Then formulas (2.16) entail that the $\mathcal{F}_t$- and $\mathcal{O}_t$-stochastic intensities of the point process $N_t^Y(k)$ are given by (2.17a) and (2.17b), respectively. □

The last step to make before applying the filtering theorem is to write down the semimartingale decomposition of the process Z_t with respect to $\mathcal{F}_t$ and to check that it satisfies the appropriate assumptions.

It is a known fact that a pure-jump Markov process of rate transition measure $\lambda(x,dy)$ admits the infinitesimal generator

$$\mathcal{L}\,\varphi(x) = \int_I \big[\varphi(y) - \varphi(x)\big]\,\lambda(x,dy), \quad x \in I, \tag{2.18}$$

whose domain is $\mathcal{D}(\mathcal{L}) = b(I)$, the set of bounded measurable functions from $(I, \mathcal{I})$ into $(\mathbb{R}, \mathcal{B})$. Then, observing that the function φ such that $Z_t(A) = \varphi(X_t)$ is simply $\varphi(x) = \mathbb{1}(x \in A)$ and recalling Example 1.6, Dynkin's formula yields

$$\begin{aligned} Z_t(A) &= Z_0(A) + \int_0^t \int_I \left[\mathbb{1}(y \in A) - Z_{s-}(A)\right] \lambda(X_{s-}, dy)\, ds + m_t = \\ &= Z_0(A) + \int_0^t \left[\lambda(X_{s-}, A) - Z_{s-}(A)\lambda(X_{s-})\right] ds + m_t = \\ &= Z_0(A) + \int_0^t f_s\, ds + m_t, \end{aligned} \tag{2.19}$$

where, again, we used the property $X_{t-}\, dt = X_t\, dt$. The process m_t is a zero mean $\mathcal{F}_t$-martingale and we have set

$$f_t = \lambda(X_{t-}, A) - Z_{t-}(A)\lambda(X_{t-}), \quad t \geqslant 0. \tag{2.20}$$

In this case we have an explicit expression for the $\mathcal{F}_t$-martingale m_t. It suffices to observe that the term $Z_t(A) - Z_0(A)$ can be written as the telescopic sum

$$\sum_{0<T_n\leqslant t} \left[Z_{T_n}(A) - Z_{T_{n-1}}(A)\right] = \int_0^t \int_I \left[\mathbb{1}(y \in A) - Z_{s-}(A)\right] p(ds \times dy). \tag{2.21}$$

Then, solving the first line of equation (2.19) by m_t and substituting (2.21), we obtain

$$m_t = \int_0^t \int_I \left[\mathbb{1}(y \in A) - Z_{s-}(A)\right] \tilde{p}(ds \times dy). \tag{2.22}$$

We are now ready to verify the assumptions on this semimartingale decomposition.

PROPOSITION 2.2: *Let f_t and m_t be the processes appearing in equations (2.20) and (2.22), respectively. Then,*

(1) *f_t is a $\mathcal{F}_t$-progressive process such that $\mathbb{E}\left[\int_0^t |f_s|\, ds\right] < \infty$, for all $t \geqslant 0$,*

(2) *m_t is a zero mean $\mathcal{F}_t$-martingale satisfying assumption (H) stated in Subsection 1.3.2.*

Proof. (1) The $\mathcal{F}_t$-progressiveness of the process f_t is granted by the very definition of semimartingale decomposition. Then it remains to show is that f_t is an integrable process in $[0, t]$, for all $t \geqslant 0$. Recalling hypothesis (2.5), we easily obtain:

$$\begin{aligned}\mathbb{E}\left[\int_0^t |f_s|\, ds\right] &= \mathbb{E}\left[\int_0^t \left|[\lambda(X_{s-}, A) - \mathbb{1}(X_{s-} \in A)\lambda(X_{s-})]\, ds\right|\right] \leqslant \\ &\leqslant \mathbb{E}\left[\int_0^t \left|\lambda(X_{s-}, A)\right| ds\right] + \mathbb{E}\left[\int_0^t \left|\mathbb{1}(X_{s-} \in A)\lambda(X_{s-})\right| ds\right] \leqslant \\ &\leqslant \mathbb{E}\left[\int_0^t \lambda(X_{s-})\, ds\right] + \mathbb{E}\left[\int_0^t \lambda(X_{s-})\, ds\right] = 2\Lambda t < +\infty,\end{aligned}$$

for all $t \geqslant 0$.

(2) In assumption (H) the $\mathcal{F}_t$-martingale m_t is decomposed in the sum of a $\mathcal{F}_t$-martingale m_t^d of locally integrable variation and of a continuous $\mathcal{F}_t$-martingale m_t^c.

In the present case $m_t^c = 0$, so part (H2) of the assumption is trivially verified, since $Z_t(\omega, A) = \mathbb{1}(X_t(\omega) \in A) \leqslant 1$, for all $\omega \in \Omega$.

Then, we have to truly check only part (H1) of the assumption, i. e. that $\mathbb{E}\left[\int_0^t |dm_s|\right] < \infty$, for all $t \geqslant 0$. Plugging equation (2.22) into this condition, we have:

$$\begin{aligned}\mathbb{E}\left[\int_0^t |dm_s|\right] &= \mathbb{E}\left[\int_0^t \left|\int_I [\mathbb{1}(y \in A) - \mathbb{1}(X_{s-} \in A)]\, \tilde{p}(ds \times dy)\right|\right] = \\ &= \mathbb{E}\left[\int_0^t \left|[\tilde{p}(ds \times A) - \mathbb{1}(X_{s-} \in A)\tilde{p}(ds \times I)]\right|\right] \leqslant \\ &\leqslant \mathbb{E}\left[\int_0^t \left|\tilde{p}(ds \times A)\right|\right] + \mathbb{E}\left[\int_0^t \left|\mathbb{1}(X_{s-} \in A)\tilde{p}(ds \times I)\right|\right].\end{aligned}$$

Expanding the expression of the compensated measure $\tilde{p}(ds \times dy)$, we can estimate the first summand by:

$$\begin{aligned}&\mathbb{E}\left[\int_0^t \left|\tilde{p}(ds \times A)\right|\right] \leqslant \mathbb{E}\left[\int_0^t \left|p(ds \times A) - \lambda(X_{s-}, A)\, ds\right|\right] \leqslant \\ &\leqslant \mathbb{E}\left[\int_0^t \left|p(ds \times A)\right|\right] + \mathbb{E}\left[\int_0^t \left|\lambda(X_{s-}, A)\, ds\right|\right] \leqslant \\ &\leqslant \mathbb{E}\left[\int_0^t p(ds \times I)\right] + \mathbb{E}\left[\int_0^t \lambda(X_{s-})\, ds\right] \leqslant \\ &\leqslant \mathbb{E}\left[p\big((0, t] \times I\big)\right] + \Lambda t < +\infty, \quad \forall t \geqslant 0,\end{aligned}$$

being the term $\mathbb{E}\left[p\big((0, t] \times I\big)\right] < +\infty$ thanks to hypothesis (2.5) that guarantees the $\mathbb{P} - a.s.$-nonexplosiveness of the marked point process $p(ds \times dy)$. Similarly we obtain, for the second summand:

$$\begin{aligned}&\mathbb{E}\left[\int_0^t \left|\mathbb{1}(X_{s-} \in A)\tilde{p}(ds \times I)\right|\right] \leqslant \mathbb{E}\left[\int_0^t \left|\tilde{p}(ds \times I)\right|\right] \leqslant \\ &\leqslant \mathbb{E}\left[\int_0^t \left|p(ds \times I) - \lambda(X_{s-})\, ds\right|\right] \leqslant \\ &\leqslant \mathbb{E}\left[\int_0^t \left|p(ds \times I)\right|\right] + \mathbb{E}\left[\int_0^t \left|\lambda(X_{s-})\, ds\right|\right] \leqslant \\ &\leqslant \mathbb{E}\left[p\big((0, t] \times I\big)\right] + \Lambda t < +\infty, \quad \forall t \geqslant 0.\end{aligned}$$

Thus $\mathbb{E}\left[\int_0^t |dm_s|\right] < +\infty$, for all $t \geqslant 0$. □

We are now in a position to apply the Filtering Theorem 1.12. Before writing the first form of our filtering equation, we notice that the integral on the mark space of the observed process that figures in formula (1.46), becomes here a finite sum of K terms. This is due to the finiteness of the space O. Recalling the discussion on the simplification occurring to its compensated measure, the innovations gain process can also be easier denoted by K distinct processes, in a similar fashion.

The filtering equation for the noise-free model is, then, given by:

$$\hat{Z}_t(A) = \hat{Z}_0(A) + \int_0^t \hat{f}_s \, ds + \\ + \sum_{k=1}^{K} \int_0^t K_s(k, A) \, [dN_s^Y(k) - \hat{\lambda}_s^Y(k) \, ds], \quad t \geqslant 0, \quad \mathbb{P} - a.s., \tag{2.23}$$

where $\hat{Z}_0(A)$ is the starting value of the filtering process $\hat{Z}_t(A)$, $\hat{f}_t$ is a $\mathcal{O}_t$-progressive process defined as in Theorem 1.11 and $K_t(k, A)$ is the innovations gain process given by (1.47), as stated in Theorem 1.12.

We will shortly proceed in the explicit calculation of the above terms, exploiting the martingale calculus for point processes.

2.2.2 *The Explicit Form of the Filtering Equation*

In this subsection we provide the final form of the filtering equation for the noise-free model presented in Section 2.1. For the purposes concerning the next proofs, we remember, once and for all, the properties $X_{t-} \, dt = X_t \, dt$ and $\hat{Z}_{t-}(dx) \, dt = \hat{Z}_t(dx) \, dt$. Moreover, we define $\frac{0}{0} = 0$.

To start, in the following two propositions we give the explicit form of the $\mathcal{O}_t$-progressive process $\hat{f}_t$ and of the innovations gain process $K_t(k, A)$.

PROPOSITION 2.3: *For all $t \geqslant 0$ we have*

$$\hat{f}_t = \int_I \lambda(x, A) \, \hat{Z}_{t-}(dx) - \int_A \lambda(x) \, \hat{Z}_{t-}(dx), \quad \mathbb{P} - a.s., \tag{2.24}$$

where $\hat{f}_t$ is the $\mathcal{O}_t$-progressive process appearing in equation (2.23).

Proof. It suffices to observe that, as already stated, the fact that the measurable space $(I, \mathcal{I})$ is a complete separable metric space grants the existence of a version of $\mathbb{E}[f_t \mid \mathcal{O}_t]$, for all $t \geqslant 0$, where f_t is given in (2.20). Then, remembering the remark following Theorem 1.11, if we are able to prove the $\mathcal{O}_t$-progressiveness of such a version, the choice $\hat{f}_t = \mathbb{E}[f_t \mid \mathcal{O}_t]$ satisfies its definition given in (1.45). Using directly that definition, a simple calculation shows that, for all nonnegative bounded $\mathcal{O}_t$-progressive processes C_t, for all $t \geqslant 0$ and $\mathbb{P} - a.s.$, we have:

$$\begin{aligned}
&\mathbb{E}\left[\int_0^t C_s\, f_s\, ds\right] = \mathbb{E}\left[\int_0^t C_s \left\{\lambda(X_{s-}, A) - Z_{s-}(A)\,\lambda(X_{s-})\right\} ds\right] = \\
&= \mathbb{E}\left[\int_0^t C_s \left\{\lambda(X_s, A) - Z_s(A)\,\lambda(X_s)\right\} ds\right] = \\
&\quad = \int_0^t \mathbb{E}\left[C_s \left\{\lambda(X_s, A) - \mathbb{1}(X_s \in A)\,\lambda(X_s)\right\}\right] ds = \\
&\quad = \int_0^t \mathbb{E}\left[C_s \mathbb{E}\left[\lambda(X_s, A) - \mathbb{1}(X_s \in A)\,\lambda(X_s) \mid \mathcal{O}_s\right]\right] ds = \\
&\quad = \int_0^t \mathbb{E}\left[C_s \int_I \left\{\lambda(x, A) - \mathbb{1}(x \in A)\,\lambda(x)\right\} \hat{Z}_s(dx)\right] ds = \\
&= \mathbb{E}\left[\int_0^t C_s \int_I \left\{\lambda(x, A) - \mathbb{1}(x \in A)\,\lambda(x)\right\} \hat{Z}_s(dx)\, ds\right] = \\
&= \mathbb{E}\left[\int_0^t C_s \left\{\int_I \lambda(x, A)\, \hat{Z}_{s-}(dx) - \int_A \lambda(x)\, \hat{Z}_{s-}(dx)\right\} ds\right].
\end{aligned}$$

Defining

$$\hat{f}_t = \int_I \lambda(x, A)\, \hat{Z}_{t-}(dx) - \int_A \lambda(x)\, \hat{Z}_{t-}(dx), \quad t \geqslant 0,$$

we observe that the process $\hat{f}_t$ is $\mathcal{O}_t$-adapted by definition and the last formula shows that its trajectories are left-continuous. By Theorem A.1, it is a $\mathcal{O}_t$-predictable process and, therefore, $\mathcal{O}_t$-progressive. Then it satisfies the definition provided in (1.45) and we can choose $\hat{f}_t = \mathbb{E}[f_t \mid \mathcal{O}_t]$. □

Proposition 2.4: *For all* $t \geqslant 0$ *and all* $k = 1, \dots, K$, *the innovations gain process appearing in equation* (2.23) *is given by*

$$K_t(k, A) = \frac{\int_{A_k^c} \lambda(x, A \cap A_k)\, \hat{Z}_{t-}(dx)}{\hat{\lambda}_t^Y(k)} - \hat{Z}_{t-}(A). \tag{2.25}$$

Proof. Let $k = 1, \dots, K$ and $t \geqslant 0$ be fixed. To achieve the result it is necessary to identify the processes $\Psi_{1,t}(k, A)$, $\Psi_{2,t}(k, A)$ and $\Psi_{3,t}(k, A)$ that form $K_t(k, A)$ as in (1.47). We stress on the fact that the finiteness of the mark space O simplifies the expression of $K_t(k, A)$ and, consequently, of $\Psi_{i,t}(k, A)$, $i = 1, 2, 3$. Therefore, each of the three equations (1.48) defining them splits into K distinct equations.

To start, we recall that the process $\Psi_{2,t}$ is always given by (1.49). In this case, that formula specializes to

$$\Psi_{2,t}(k, A) = \hat{Z}_{t-}(A).$$

Regarding the process $\Psi_{1,t}(k, A)$, we elaborate the right-hand side of the first of equations (1.48) by repeatedly using the Fubini-Tonelli

Theorem and the existence of the probability distribution $\hat{Z}_t(dx)$. We then have:

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^t C_s\, Z_s(A)\,\lambda_s^Y(k)\,ds\right] = \mathbb{E}\left[\int_0^t C_s\, Z_{s-}(A)\,\lambda_s^Y(k)\,ds\right] = \\
&= \int_0^t \mathbb{E}\left[C_s\, Z_{s-}(A)\,\lambda_s^Y(k)\right] ds = \\
&= \int_0^t \mathbb{E}\left[C_s \mathbb{E}\left[Z_{s-}(A)\,\lambda_s^Y(k) \mid \mathcal{O}_s\right]\right] ds = \\
&= \int_0^t \mathbb{E}\left[C_s \mathbb{E}\left[\mathbb{1}(X_{s-} \in A)\,\mathbb{1}(X_{s-} \in A_k^c)\,\lambda(X_{s-}, A_k) \mid \mathcal{O}_s\right]\right] ds = \\
&= \int_0^t \mathbb{E}\left[C_s \mathbb{E}\left[\mathbb{1}(X_s \in A \cap A_k^c)\,\lambda(X_s, A_k) \mid \mathcal{O}_s\right]\right] ds = \\
&= \int_0^t \mathbb{E}\left[C_s \int_I \mathbb{1}(x \in A \cap A_k^c)\,\lambda(x, A_k)\,\hat{Z}_s(dx)\right] ds = \\
&= \int_0^t \mathbb{E}\left[C_s \int_{A \cap A_k^c} \lambda(x, A_k)\,\hat{Z}_{s-}(dx)\right] ds = \\
&= \mathbb{E}\left[\int_0^t C_s \left\{\int_{A \cap A_k^c} \lambda(x, A_k)\,\hat{Z}_{s-}(dx)\right\} ds\right].
\end{aligned}
$$

Then the identification with the left-hand side of the aforementioned equation provides us with

$$
\Psi_{1,t}(k, A) = \frac{\int_{A \cap A_k^c} \lambda(x, A_k)\,\hat{Z}_{t-}(dx)}{\hat{\lambda}_t^Y(k)}.
$$

With a similar reasoning, we manipulate the right-hand side of the third of equations (1.48). We observe, first, that:

$$
\Delta Z_t(A) = Z_t(A) - Z_{t-}(A) = \mathbb{1}(X_t \in A) - \mathbb{1}(X_{t-} \in A), \quad t \geqslant 0.
$$

The expression $\mathbb{E}\left[\int_0^t C_s\,\Delta Z_s(A)\,dN_s^Y(k)\right]$ can be, then, divided into two terms, given by

$$
\mathbb{E}\left[\int_0^t C_s\,\mathbb{1}(X_s \in A)\,dN_s^Y(k)\right], \tag{2.26a}
$$

$$
\mathbb{E}\left[\int_0^t C_s\,\mathbb{1}(X_{s-} \in A)\,dN_s^Y(k)\right]. \tag{2.26b}
$$

We can rewrite (2.26a) as follows:

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^t C_s\,\mathbb{1}(X_s \in A)\,dN_s^Y(k)\right] = \\
=&\mathbb{E}\left[\sum_{0<s\leqslant t} C_s\,\mathbb{1}(X_s \in A)\,\mathbb{1}(Y_{s-} \neq a_k)\,\mathbb{1}(Y_s = a_k)\right] = \\
=&\mathbb{E}\left[\sum_{0<s\leqslant t} C_s\,\mathbb{1}(X_s \in A)\,\mathbb{1}(X_{s-} \in A_k^c)\,\mathbb{1}(X_s \in A_k)\right] = \\
=&\mathbb{E}\left[\sum_{0<s\leqslant t} C_s\,\mathbb{1}(X_{s-} \in A_k^c)\,\mathbb{1}(X_s \in A \cap A_k)\right] = \\
=&\mathbb{E}\left[\int_0^t C_s\,dN_s^X(A_k^c, A \cap A_k)\right].
\end{aligned}
$$

From Lemma 2.1, we know that, for $A, B \in \mathcal{I}$, the point process $N_t^X(A, B)$ admits $\mathcal{O}_t$-stochastic intensity $\hat{\lambda}_t^X(A, B)$, given in (2.16b). Recalling that C_t is a bounded $\mathcal{O}_t$-predictable process, the very definition of stochastic intensity yields

$$\mathbb{E}\left[\int_0^t C_s\,dN_s^X(A_k^c, A \cap A_k)\right] = \mathbb{E}\left[\int_0^t C_s\,\hat{\lambda}_s^X(A_k^c, A \cap A_k)\,ds\right],$$

that together with (2.16b) allows us to write

$$\mathbb{E}\left[\int_0^t C_s \mathbb{1}(X_s \in A) dN_s^Y(k)\right] = \mathbb{E}\left[\int_0^t C_s \int_{A_k^c} \lambda(x, A \cap A_k) \hat{Z}_{s-}(dx) ds\right].$$

Concerning the term in (2.26b), it suffices to observe that the process $C_t\,\mathbb{1}(X_{t-} \in A)$ is $\mathcal{F}_t$-predictable (C_t, being $\mathcal{O}_t$-predictable, is *a fortiori* $\mathcal{F}_t$-predictable), then the definition of $\mathcal{F}_t$-stochastic intensity gives

$$\mathbb{E}\left[\int_0^t C_s\,\mathbb{1}(X_{s-} \in A)\,dN_s^Y(k)\right] = \mathbb{E}\left[\int_0^t C_s\,\mathbb{1}(X_{s-} \in A)\,\lambda_s^Y(k)\,ds\right].$$

Thus, we discover that this is the same expression considered in the computation of the term $\Psi_{1,t}(k, A)$.

Putting back together the formulas obtained for (2.26a) and (2.26b) and identifying the result with the left-hand side of the third of equations (1.48), we obtain:

$$\Psi_{3,t}(k, A) = \frac{\int_{A_k^c} \lambda(x, A \cap A_k)\,\hat{Z}_{t-}(dx)}{\hat{\lambda}_t^Y(k)} - \Psi_{1,t}(k, A).$$

Finally, recalling that

$$K_t(k, A) = \Psi_{1,t}(k, A) - \Psi_{2,t}(k, A) + \Psi_{3,t}(k, A),$$

we reach the formula in (2.25). □

We can now write an "intermediate" version of our filtering equation, more explicit than the earlier form given in (2.23). In fact, combining the results coming from Propositions 2.3 and 2.4, we obtain, for all $t \geqslant 0$ and $\mathbb{P}-a.s.$,

$$\hat{Z}_t(A) = \hat{Z}_0(A) + \int_0^t \left\{ \int_I \lambda(x, A)\, \hat{Z}_{s-}(dx) - \int_A \lambda(x)\, \hat{Z}_{s-}(dx) \right\} ds + \\ + \sum_{k=1}^{K} \int_0^t \left\{ \frac{\int_{A_k^c} \lambda(x, A \cap A_k) \hat{Z}_{s-}(dx)}{\hat{\lambda}_s^Y(k)} - \hat{Z}_{s-}(A) \right\} \left[dN_s^Y(k) - \hat{\lambda}_s^Y(k) ds \right]. \tag{2.27}$$

Before stating the final version of the filtering equation, we need to introduce a new operator.

DEFINITION 2.1 (Operator H): For all $a \in O$, the operator H is defined as a mapping $\nu \mapsto H_a[\nu]$ from the space of measures ν on $(I, \mathcal{I})$ onto itself such that, for all $A \in \mathcal{I}$

$$H_a[\nu](A) = \begin{cases} 0, & \text{if } A \cap h^{-1}(a) = \emptyset \\ \frac{\int_{A \cap h^{-1}(a)} \nu(dx)}{\int_{h^{-1}(a)} \nu(dx)}, & \text{if } A \cap h^{-1}(a) \neq \emptyset \text{ and } D > 0 \\ \rho_a, & \text{if } A \cap h^{-1}(a) \neq \emptyset \text{ and } D = 0 \end{cases} \tag{2.28}$$

where $D = \int_{h^{-1}(a)} \nu(dx)$ and ρ_a is an arbitrarily chosen probability measure on $(I, \mathcal{I})$ with support in $h^{-1}(a)$.

REMARK: If ν is a positive measure, then $H_a[\nu]$ is a probability measure on $(I, \mathcal{I})$, with support in $h^{-1}(a)$. If in addition ν is a probability measure on $(I, \mathcal{I})$, then $H_a[\nu]$ is the corresponding conditional probability measure given the event $\{x \in h^{-1}(a)\}$. We note that the exact values of the probability measure ρ_a are irrelevant.

THEOREM 2.1 (Filtering Equation): *Let $\hat{Z}_t(A)$ be the process defined by*

$$\hat{Z}_t(\omega, A) = H_{Y_0(\omega)}[\mu](A) + \\ + \int_0^t \left\{ \int_I \lambda\Big(x, A \cap h^{-1}\big(Y_s(\omega)\big)\Big)\, \hat{Z}_{s-}(\omega, dx) - \int_A \lambda(x) \hat{Z}_{s-}(\omega, dx) + \right. \\ \left. + \hat{Z}_{s-}(\omega, A) \int_I \lambda\Big(x, h^{-1}\big(Y_s(\omega)\big)^c\Big)\, \hat{Z}_{s-}(\omega, dx) \right\} ds + \\ + \sum_{0 < \tau_n(\omega) \leqslant t} \Big\{ H_{Y_{\tau_n}(\omega)}[\mu_n](A) - \hat{Z}_{\tau_n -}(\omega, A) \Big\}, \tag{2.29}$$

where, for all $n = 1, 2, \ldots,$ $\mu_n(dy) = \int_I \lambda(x, dy)\, \hat{Z}_{\tau_n -}(dx)$ is a measure on $(I, \mathcal{I})$.

Then $\hat{Z}_t(A)$ is a modification of the filtering process $\mathbb{P}\left(X_t \in A \mid \mathcal{F}_t^Y\right)$, i.e. for all $t \geqslant 0$ and all $A \in \mathcal{I}$ we have $\hat{Z}_t(A) = \mathbb{P}\left(X_t \in A \mid \mathcal{F}_t^Y\right)$, $\mathbb{P}-a.s.$.

Proof. As we observed at the beginning of this Section, the natural filtration $\mathcal{F}_t^Y$ of the process Y_t coincides with the observed history $\mathcal{O}_t$ by construction. Thus, the filtering processes $\mathbb{P}\left(X_t \in A \mid \mathcal{F}_t^Y\right)$ and $\mathbb{P}\left(X_t \in A \mid \mathcal{O}_t\right)$ are identical, and we already possess an expression for a modification of the latter, namely (2.27).

The proof will, then, explicit the terms that appear in that equation to achieve the form (2.29). It will be rather long, so we will divide it into four main blocks. Each of them will cover the four main parts of the equation, that is to say:

(1) The initial value $H_{Y_0(\omega)}[\mu](A)$;

(2) The linear term, i. e. the second line of equation (2.29);

(3) The quadratic term, i. e. the third line of equation (2.29);

(4) The jump term, i. e. the last line of equation (2.29).

In the sequel we will drop the ω in the notation of all the quantities involved. We also fix $A \in \mathcal{I}$ and $t \geqslant 0$.

(1) We can retrieve the initial value of the process $\hat{Z}_0(A)$ simply by elaborating its definition. In fact, remembering that $X_0 \sim \mu$, we have

$$\begin{aligned}\hat{Z}_0(A) &= \mathbb{P}\left(X_0 \in A \mid \mathcal{O}_0\right) = \mathbb{P}\left(X_0 \in A \mid Y_0\right) = \\ &= \mathbb{P}\left(X_0 \in A \mid Y_0 = h(Y_0)\right) = \frac{\mathbb{P}\left(X_0 \in A, X_0 \in h^{-1}(Y_0)\right)}{\mathbb{P}\left(X_0 \in h^{-1}(Y_0)\right)} = \\ &= \begin{cases} 0, & \text{if } A \cap h^{-1}(Y_0) = \emptyset \\ \frac{\int_{A \cap h^{-1}(Y_0)} \mu(dx)}{\int_{h^{-1}(Y_0)} \mu(dx)}, & \text{if } A \cap h^{-1}(Y_0) \neq \emptyset \text{ and } D > 0 \\ \rho_{Y_0}, & \text{if } A \cap h^{-1}(Y_0) \neq \emptyset \text{ and } D = 0 \end{cases}\end{aligned}$$

where $D = \int_{h^{-1}(Y_0)} \mu(dx)$ and ρ_{Y_0} an arbitrarily chosen probability measure on $(I, \mathcal{I})$ with support in $h^{-1}(Y_0)$. Then, recalling the definition of the operator H in (2.28), we can write

$$\hat{Z}_0(A) = H_{Y_0}[\mu](A).$$

(2) The linear part of equation (2.27) is given by

$$\begin{aligned}&\int_0^t \left\{\int_I \lambda(x, A)\, \hat{Z}_{s-}(dx) - \int_A \lambda(x)\, \hat{Z}_{s-}(dx)\right\} ds + \\ &\qquad - \sum_{k=1}^K \int_0^t \frac{\int_{A_k^c} \lambda(x, A \cap A_k)\, \hat{Z}_{s-}(dx)}{\hat{\lambda}_s^Y(k)}\, \hat{\lambda}_s^Y(k)\, ds.\end{aligned}$$

For ease of notation, we consider just the innermost integrals and the finite sum, i. e.

$$\int_I \lambda(x, A)\hat{Z}_{s-}(dx) - \int_A \lambda(x)\hat{Z}_{s-}(dx) - \sum_{k=1}^K \int_{A_k^c} \lambda(x, A \cap A_k)\hat{Z}_{s-}(dx),$$

where we have simplified the fraction.

Considering the peculiar structure of our model, we notice that the conditional distribution $\hat{Z}_t(dx)$ has support in the set $h^{-1}(Y_t)$, for all $t \geqslant 0$. Indeed, if at time t we observe a specific value of Y_t, then $X_t \in h^{-1}(Y_t)$, since $Y_t = h(X_t)$ and therefore the conditional distribution $\hat{Z}_t(dx)$ assigns measure zero to all the sets in $\mathcal{I}$ that are disjoint from $h^{-1}(Y_t)$. Moreover, $h^{-1}(Y_t) = A_k$, for some $k = 1, \dots, K$, because of the partition induced on I by the function h. This fact allows us to dramatically simplify the last equation.

Let us fix $s \geqslant 0$ and denote by $k^\star \in \{1, \dots, K\}$ the index such that $h^{-1}(Y_s) = A_{k^\star}$. Then, since we are considering the continuous part of the filtering process, the conditional distribution $\hat{Z}_{s-}(dx)$ has the same support as $\hat{Z}_s(dx)$, i.e. $h^{-1}(Y_s)$. This is contained in all of the sets A_k but one, precisely $A_{k^\star}$, because $A_{k^\star}^c = h^{-1}(Y_s)^c$. As a consequence, we can write

$$\sum_{k=1}^{K} \int_{A_k^c} \lambda(x, A \cap A_k) \hat{Z}_{s-}(dx) = \sum_{\substack{k=1 \\ k \neq k^\star}}^{K} \int_{A_k^c} \lambda(x, A \cap A_k) \hat{Z}_{s-}(dx).$$

For the same reasons concerning the support of the conditional distribution $\hat{Z}_{s-}(dx)$, we can extend all the integrals to the whole set I and achieve

$$\begin{aligned} \sum_{\substack{k=1 \\ k \neq k^\star}}^{K} \int_{A_k^c} \lambda(x, A \cap A_k) \hat{Z}_{s-}(dx) &= \sum_{\substack{k=1 \\ k \neq k^\star}}^{K} \int_{I} \lambda(x, A \cap A_k) \hat{Z}_{s-}(dx) = \\ = \sum_{\substack{k=1 \\ k \neq k^\star}}^{K} \int_I \int_I \mathbb{1}(y \in A \cap A_k)\, \lambda(x, dy) \hat{Z}_{s-}(dx) &= \\ = \int_I \int_I \sum_{\substack{k=1 \\ k \neq k^\star}}^{K} \mathbb{1}(y \in A \cap A_k)\, \lambda(x, dy) \hat{Z}_{s-}(dx).& \end{aligned}$$

The only terms that depend on the index k are the indicator functions. Studying them separately and considering that the sets A_k are all pairwise disjoint, we find that

$$\begin{aligned} &\sum_{\substack{k=1 \\ k \neq k^\star}}^{K} \mathbb{1}(y \in A \cap A_k) = \mathbb{1}(y \in A) \cdot \sum_{\substack{k=1 \\ k \neq k^\star}}^{K} \mathbb{1}(y \in A_k) = \\ &= \mathbb{1}(y \in A) \cdot \mathbb{1}\Bigg(y \in \bigcup_{\substack{k=1 \\ k \neq k^\star}}^{K} A_k\Bigg) = \mathbb{1}(y \in A) \cdot \mathbb{1}(y \in A_{k^\star}^c) = \\ &= \mathbb{1}(y \in A)\Big[1 - \mathbb{1}\big(y \in A_{k^\star}\big)\Big] = \mathbb{1}(y \in A)\Big[1 - \mathbb{1}\big(y \in h^{-1}(Y_s)\big)\Big] = \\ &= \mathbb{1}\big(y \in A\big) - \mathbb{1}\big(y \in A \cap h^{-1}(Y_s)\big). \end{aligned}$$

We can, then, write

$$\int_I \int_I \sum_{\substack{k=1 \\ k \neq k^\star}}^{K} \mathbb{1}(y \in A \cap A_k)\, \lambda(x, dy) \hat{Z}_{s-}(dx) =$$

$$= \int_I \int_I \Big\{ \mathbb{1}\big(y \in A\big) - \mathbb{1}\big(y \in A \cap h^{-1}(Y_s)\big) \Big\} \lambda(x, dy) \hat{Z}_{s-}(dx) =$$

$$= \int_I \lambda\big(x, A\big) \hat{Z}_{s-}(dx) - \int_I \lambda\big(x, A \cap h^{-1}(Y_s)\big) \hat{Z}_{s-}(dx).$$

The first term in this equation cancels out the first in the expression of the whole linear term. Then, putting together all the previous results, we obtain

$$\int_0^t \Big\{ \int_I \lambda\big(x, A \cap h^{-1}(Y_s)\big) \hat{Z}_{s-}(dx) - \int_A \lambda(x)\, \hat{Z}_{s-}(dx) \Big\}\, ds,$$

which is exactly the second line of (2.29).

(3) The quadratic term of equation (2.27) is

$$\sum_{k=1}^{K} \int_0^t \hat{Z}_{s-}(A)\, \hat{\lambda}_s^Y(k)\, ds = \int_0^t \hat{Z}_{s-}(A) \sum_{k=1}^{K} \int_{A_k^c} \lambda(x, A_k)\, \hat{Z}_{s-}(dx)\, ds.$$

Following the same considerations made in part (2) of the proof, we reach the expression

$$\sum_{k=1}^{K} \int_{A_k^c} \lambda(x, A_k)\, \hat{Z}_{s-}(dx) = \int_I \int_I \sum_{\substack{k=1 \\ k \neq k^\star}}^{K} \mathbb{1}(y \in A_k)\, \lambda(x, dy) \hat{Z}_{s-}(dx).$$

Then, recalling that

$$\sum_{\substack{k=1 \\ k \neq k^\star}}^{K} \mathbb{1}(y \in A_k) = 1 - \mathbb{1}\big(y \in h^{-1}(Y_s)\big) = \mathbb{1}\big(y \in h^{-1}(Y_s)^c\big),$$

we achieve

$$\int_0^t \hat{Z}_{s-}(A) \int_I \lambda\big(x, h^{-1}(Y_s)^c\big) \hat{Z}_{s-}(dx)\, ds,$$

that is precisely the third line of (2.29).

(4) The jump part of equation (2.27) is given by

$$\sum_{k=1}^{K} \int_0^t \left\{ \frac{\int_{A_k^c} \lambda(x, A \cap A_k)\, \hat{Z}_{s-}(dx)}{\hat{\lambda}_s^Y(k)} - \hat{Z}_{s-}(A) \right\} dN_s^Y(k) =$$

$$= \sum_{k=1}^{K} \sum_{0 < s \leqslant t} \left\{ \frac{\int_{A_k^c} \lambda(x, A \cap A_k)\, \hat{Z}_{s-}(dx)}{\hat{\lambda}_s^Y(k)} - \hat{Z}_{s-}(A) \right\} dN_s^Y(k) =$$

$$= \sum_{0 < s \leqslant t} \sum_{k=1}^{K} \left\{ \frac{\int_{A_k^c} \lambda(x, A \cap A_k)\, \hat{Z}_{s-}(dx)}{\hat{\lambda}_s^Y(k)} - \hat{Z}_{s-}(A) \right\} dN_s^Y(k).$$

We observe that, for a fixed $t \geqslant 0$ and $k \in \{1, \dots, K\}$, $dN^Y_t(k) = 1$ if and only if $\tau_n = t$ and $\zeta_n = a_k$ for some $n \geqslant 1$. Therefore, we can substitute the index s appearing in the outermost sum in the last equation with τ_n. For a fixed $n \in \mathbb{N}$ such that $0 < \tau_n \leqslant t$, we also notice that the only surviving term in the innermost sum is that relative to the index $k^\star$ such that $\zeta_n = a_{k^\star}$, since $dN^Y_{\tau_n}(k) = 0$ for all $k \in \{1, \dots, K\}$ with $k \neq k^\star$. Moreover, since $Y_{\tau_n} = a_{k^\star}$, or equivalently $A_{k^\star} = h^{-1}(Y_{\tau_n})$, the last equation reduces to

$$\sum_{0<\tau_n \leqslant t} \left\{ \frac{\int_{h^{-1}(Y_{\tau_n})^c} \lambda\big(x, A \cap h^{-1}(Y_{\tau_n})\big)\, \hat{Z}_{\tau_n -}(dx)}{\int_{h^{-1}(Y_{\tau_n})^c} \lambda\big(x, h^{-1}(Y_{\tau_n})\big)\, \hat{Z}_{\tau_n -}(dx)} - \hat{Z}_{\tau_n -}(A) \right\},$$

where we used the fact that

$$\hat{\lambda}^Y_{\tau_n}(k^\star) = \int_{A^c_{k^\star}} \lambda\big(x, A_{k^\star}\big) \hat{Z}_{\tau_n -}(dx) = \int_{h^{-1}(Y_{\tau_n})^c} \lambda\big(x, h^{-1}(Y_{\tau_n})\big) \hat{Z}_{\tau_n -}(dx).$$

Considering that $Z_{\tau_n -}(dx)$ has support in the set $h^{-1}(Y_{\tau_n -})$ and that $h^{-1}(Y_{\tau_n})^c \supset h^{-1}(Y_{\tau_n -})$, we can extend both the integrals appearing in the last expression to achieve

$$\sum_{0<\tau_n \leqslant t} \left\{ \frac{\int_I \lambda\big(x, A \cap h^{-1}(Y_{\tau_n})\big)\, \hat{Z}_{\tau_n -}(dx)}{\int_I \lambda\big(x, h^{-1}(Y_{\tau_n})\big)\, \hat{Z}_{\tau_n -}(dx)} - \hat{Z}_{\tau_n -}(A) \right\}.$$

Finally, we observe that the fraction in the last equation vanishes if $A \cap h^{-1}(Y_{\tau_n}) = \emptyset$ and also if its denominator is equal to zero, since we assumed $\frac{0}{0} = 0$. This should recall the definition of the operator H, when choosing $a = Y_{\tau_n}$ and $\rho_a \equiv 0$. Thus, we just need to identify the measure on $(I, \mathcal{I})$ on which the operator H is acting.

A final application of the Fubini-Tonelli Theorem on both the numerator and the denominator of the fraction gives

$$\frac{\int_I \lambda\big(x, A \cap h^{-1}(Y_{\tau_n})\big) \hat{Z}_{\tau_n -}(dx)}{\int_I \lambda\big(x, h^{-1}(Y_{\tau_n})\big) \hat{Z}_{\tau_n -}(dx)} = \frac{\int_I \int_{A \cap h^{-1}(Y_{\tau_n})} \lambda(x, dy) \hat{Z}_{\tau_n -}(dx)}{\int_I \int_{h^{-1}(Y_{\tau_n})} \lambda(x, dy) \hat{Z}_{\tau_n -}(dx)} =$$
$$\frac{\int_{A \cap h^{-1}(Y_{\tau_n})} \int_I \lambda(x, dy)\, \hat{Z}_{\tau_n -}(dx)}{\int_{h^{-1}(Y_{\tau_n})} \int_I \lambda(x, dy)\, \hat{Z}_{\tau_n -}(dx)} = \frac{\int_{A \cap h^{-1}(Y_{\tau_n})} \mu_n(dy)}{\int_{h^{-1}(Y_{\tau_n})} \mu_n(dy)}.$$

Then, comparing this expression with (2.28), we see that $\mu_n(dy)$ is the sought measure and we can write

$$\sum_{0<\tau_n \leqslant t} H_{Y_{\tau_n}}[\mu_n](A) - \hat{Z}_{\tau_n -}(A),$$

that gives the last term in equation (2.29). □

Remark: A simple manipulation of the filtering equation (2.29) allows us to write it in a more suggestive and "expected" way, i. e. with the use of the infinitesimal generator $\mathcal{L}$ associated to the process X_t.

It suffices to notice that the first summand of the linear term can also be expressed as

$$\begin{aligned}
&\int_0^t \int_I \lambda\big(x, A \cap h^{-1}(Y_s)\big)\hat{Z}_{s-}(dx)\, ds = \\
&= \int_0^t \int_I \int_I \mathbb{1}\big(y \in A \cap h^{-1}(Y_s)\big)\, \lambda(x, dy)\, \hat{Z}_{s-}(dx)\, ds = \\
&= \int_0^t \int_I \int_I \mathbb{1}(y \in A)\, \mathbb{1}\big(y \in h^{-1}(Y_s)\big)\, \lambda(x, dy)\, \hat{Z}_{s-}(dx)\, ds = \\
&= \int_0^t \int_I \int_I \mathbb{1}(y \in A)\, \Big[1 - \mathbb{1}\big(y \in h^{-1}(Y_s)^c\big)\Big]\, \lambda(x, dy)\, \hat{Z}_{s-}(dx)\, ds.
\end{aligned}$$

The innermost integral is then equal to

$$\int_I \mathbb{1}(y \in A)\, \lambda(x, dy) - \int_I \mathbb{1}(y \in A)\, \mathbb{1}\big(y \in h^{-1}(Y_s)^c\big)\, \lambda(x, dy).$$

Similarly the second summand of the linear term can be written as

$$\begin{aligned}
&\int_0^t \int_A \lambda(x)\, \hat{Z}_{s-}(dx)\, ds = \int_0^t \int_I \mathbb{1}(x \in A)\, \lambda(x)\, \hat{Z}_{s-}(dx)\, ds = \\
&= \int_0^t \int_I \int_I \mathbb{1}(x \in A)\, \lambda(x, dy)\, \hat{Z}_{s-}(dx)\, ds.
\end{aligned}$$

Then, putting back together the whole linear part of equation (2.29) and rearranging the terms we obtain

$$\begin{aligned}
&\int_0^t \left\{ \int_I \lambda\big(x, A \cap h^{-1}(Y_s)\big)\hat{Z}_{s-}(dx) - \int_A \lambda(x)\, \hat{Z}_{s-}(dx) \right\} ds = \\
&= \int_0^t \int_I \int_I \Big\{ \mathbb{1}(y \in A) - \mathbb{1}(x \in A) \Big\} \lambda(x, dy)\, \hat{Z}_{s-}(dx)\, ds + \\
&- \int_0^t \int_I \int_I \mathbb{1}(y \in A)\, \mathbb{1}\big(y \in h^{-1}(Y_s)^c\big)\, \lambda(x, dy)\, \hat{Z}_{s-}(dx)\, ds,
\end{aligned}$$

and we recognize the innermost integral in the second line as the infinitesimal generator $\mathcal{L}$ of the process X_t acting on the bounded measurable function $\varphi(x) = \mathbb{1}(x \in A)$. We can, therefore, write

$$\int_0^t \left\{ \int_I \mathcal{L}\varphi(x)\, \hat{Z}_{s-}(dx) - \int_I \int_{h^{-1}(Y_s)^c} \varphi(y)\, \lambda(x, dy)\, \hat{Z}_{s-}(dx) \right\} ds.$$

Finally, if we denote the terms $\hat{Z}_\cdot(A)$ and $H_\cdot[\cdot](A)$ in a slight different form, i.e. by $\hat{Z}_\cdot(\varphi)$ and $H_\cdot[\cdot](\varphi)$ respectively, we can write the filtering equation as

$$\begin{aligned}
&\hat{Z}_t(\omega, \varphi) = H_{Y_0(\omega)}[\mu](\varphi) + \\
&+ \int_0^t \left\{ \int_I \mathcal{L}\varphi(x)\hat{Z}_{s-}(\omega, dx) - \int_I \int_{h^{-1}\left(Y_s(\omega)\right)^c} \varphi(y)\lambda(x, dy)\hat{Z}_{s-}(\omega, dx) + \right. \\
&\qquad \left. + \hat{Z}_{s-}(\omega, \varphi) \int_I \lambda\Big(x, h^{-1}\big(Y_s(\omega)\big)^c\Big)\, \hat{Z}_{s-}(\omega, dx) \right\} ds + \\
&\qquad + \sum_{0 < \tau_n(\omega) \leqslant t} \Big\{ H_{Y_{\tau_n}(\omega)}[\mu_n](\varphi) - \hat{Z}_{\tau_n -}(\omega, \varphi) \Big\}. \qquad (2.30)
\end{aligned}$$

As far as this Thesis is concerned, writing φ instead of A has to be considered a mere change of notation. In a more general setting, instead, it can be proved, following the same lines that conduced us to equation (2.29) and with the proper adjustments, that (2.30) is the filtering equation for the filtering process

$$\hat{Z}_t(\varphi) = \mathbb{E}\left[\varphi(X_t) \mid \mathcal{F}_t^Y\right], \quad t \geqslant 0, \tag{2.31}$$

where φ is a real-valued bounded measurable function defined on I.

CONCLUSIONS AND FUTURE DEVELOPMENTS

Stochastic problems of a filtering nature appear today in a variety of situations. As recalled in the Introduction, they are used to analyze dynamic systems in engineering applications and they arise naturally in the description of financial and economical models. Frequently, they are the first and essential step to find a solution to optimization problems, as is the case of optimal stopping or optimal control of a stochastic process of interest. Because of their ubiquitous nature, they have been deeply analyzed, especially in the case of noisy observations, as stated in the beginning.

The purpose of this thesis was twofold: on one hand, we wanted to foster the analysis of the model presented in Chapter 2, characterized by a noise-free observation. We think it deserves attention on his own: it is not uncommon the case where either no noise is effectively acting on the observed process or it can be considered negligible with respect to the noise acting on the whole system. In these situations, all the sources of randomness are included in the unobserved process. On the other hand, we wanted to show a detailed application of marked point processes and martingale calculus to this kind of models. The power of these tools is clear: in the Introduction we stated that this model possesses a dynamical nature and that the martingale theory is well suited for the analysis of such problems. Chapter 2, where the filtering equation is derived exploiting these instruments, represents a plain and hopefully convincing explanation of this statement.

The work done in this Thesis can be, thus, summarized by three main points:

1. Investigation on the noise-free model (3) via filtering techniques based on the martingale theory, an approach not adopted so far in this context.

2. A simpler proof of the filtering formula with respect to the method adopted in [5].

3. Generalization of the noise-free model (3). We assumed the state space of the unobserved process to be a complete separable metric space, instead of the restrictive assumption of finiteness made in [5].

This Thesis aims also to be the basis for future developments on the subject. Regarding the applications of the model presented here, we notice that its simple structure allows to describe a great number of problems, for example:

- Optimal stopping and optimal control problems, where the observation is an exact datum, not corrupted by noise;
- Optimal switching, where the commutations among different system dynamics are governed by an unobserved pure jump Markov process (who may, as an example, control the drift and diffusion coefficients of an SDE driven by a Wiener process);
- Jump Markov linear systems, where a physical system can be described by a stochastic linear dynamic model whose parameters are governed by an underlying jump Markov process.

From a more purely mathematical point of view, instead, a deeper investigation on the properties of the filtering process should be carried out. In the previously cited work, the solution to the filtering equation, i. e. the filtering process, is shown to possess two important properties:

- It is a Markov process with respect to the natural filtration of the observed process.
- It is a PDP in the sense of Davis.

The class of processes named *piecewise-deterministic Markov processes* (PDPs) and introduced by M.H.A. Davis in [7], is an extensively studied one: the interested reader is referenced to [8] for a detailed exposition. A PDP has a particular structure, i. e. it has jumps at random times and otherwise it evolves deterministically. Apart from proving that the filtering process belongs to this class of processes, the relevance of this characterization is that a lot of known results from the theory of PDPs immediately apply to this case and define further properties.

It is, then, obvious to postulate that the filtering process described in this work possesses the Markov property with respect to the observed history and is a PDP. Moreover, one can wonder if it can be characterized as the unique solution to the filtering equation.

As we tried to point out, a lot of work can be done on this subject. On one hand, the apparent simplicity of the model described here hides a plethora of applications yet to be studied in a vast range of fields. New and more sophisticated models can also be built upon this one and their analysis should take advantage from the techniques explained and adopted in this work. On the other hand, the mathematical description of this problem is not over. Further generalizations, such as allowing the observed process to take values into a complete separable metric space, are yet to be explored. Various properties supposedly holding for the filtering process have yet to be proved. Other characteristics are surely yet to be discovered.

A STOCHASTIC PROCESSES

This Appendix is devoted to illustrating the fundamentals of stochastic processes theory. A selection of arguments on this vast subject has been made in order to accommodate the purposes of this Thesis, where the concepts and the terminology here recalled are constantly used. The reader needing a more complete discussion on the material covered hereinafter, can consult any of the classical textbooks on stochastic processes theory. We cite, as a reference, [4, 13, 15].

A.1 STOCHASTIC PROCESSES, FILTRATIONS AND MEASURABILITY

In this section we will review a few definitions and useful results on stochastic processes. In the sequel we will assume defined a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a measurable space $(E, \mathcal{E})$.

Definition A.1 (Continuous-time Stochastic Process): Let $(X_t)_{t\geqslant 0}$ be a collection of random variables $X_t\colon (\Omega, \mathcal{F}) \to (E, \mathcal{E})$, $t \geqslant 0$. $(X_t)_{t\geqslant 0}$ is called a E-*valued continuous-time stochastic process*.

For a fixed $\omega \in \Omega$, the mapping $t \mapsto X_t(\omega)$, $t \geqslant 0$, is called a *trajectory* or a *path* of the process X_t. If $(E, \mathcal{E})$ is also a topological space, we say that the process X_t is *continuous* (*right-continuous*, *left-continuous*), if and only if its trajectories are $\mathbb{P}-a.s.$ continuous (right-continuous, left-continuous).

Definition A.2: Let X_t and Y_t be two E-valued stochastic processes, both defined on the same probability space. They are said to be

(1) *modifications* or *versions* of one another if and only if

$$\mathbb{P}\big(\{\omega\colon X_t(\omega) \neq Y_t(\omega)\}\big) = 0, \quad \forall t \geqslant 0, \tag{A.1}$$

(2) $\mathbb{P}$-*indistinguishable* if and only if

$$\mathbb{P}\big(\{\omega\colon X_t(\omega) \neq Y_t(\omega),\, \forall t \geqslant 0\}\big) = 0, \tag{A.2}$$

that is to say, if they have identical trajectories except on a set of $\mathbb{P}$-measure zero.

Naturally linked to a stochastic process is the concept of *filtration*, that mathematically describes the idea of an increasing information pattern: as time progresses, more and more informations are revealed about the process itself or other "events".

DEFINITION A.3 (Filtration): Let $(\mathcal{F}_t)_{t\geqslant 0}$ be a family of sub-σ-fields of $\mathcal{F}$ such that

$$\mathcal{F}_s \subset \mathcal{F}_t, \quad \forall\, 0 \geqslant s \geqslant t. \tag{A.3}$$

$(\mathcal{F}_t)_{t\geqslant 0}$ is called a *filtration* or *history* on $(\Omega, \mathcal{F})$.

REMARK: For ease of notation, stochastic processes and filtrations are often denoted simply by X_t and $\mathcal{F}_t$, respectively. This should create no confusion, since is clear from the context which is the object of interest, being either the entire stochastic process or filtration, or the single random variable or σ-algebra, for a fixed $t \geqslant 0$.

DEFINITION A.4: Let $\mathcal{F}_t$ be a filtration on $(\Omega, \mathcal{F})$ and denote with $\mathcal{F}_{t+}$ the σ-algebra

$$\mathcal{F}_{t+} = \bigcap_{h>0} \mathcal{F}_{t+h}, \quad t \geqslant 0. \tag{A.4}$$

The filtration $\mathcal{F}_t$ is said to be *right-continuous* if and only if $\mathcal{F}_{t+} = \mathcal{F}_t$ for all $t \geqslant 0$.

Among the various filtrations that is possible to associate to a stochastic process X_t, a special place is occupied by the *natural filtration* (also said the *internal history*), indicated by $\mathcal{F}_t^X$ and defined by

$$\mathcal{F}_t^X = \sigma(X_s,\, 0 \leqslant s \leqslant t), \quad \forall t \geqslant 0. \tag{A.5}$$

For every $t \geqslant 0$, $\mathcal{F}_t^X$ is the σ-algebra generated by the collection of random variables $(X_s)_{s\in[0,t]}$ and it represents the stream of information generated by the process X_t itself up to time t.

We conclude this section with a brief recapitulation of the main definitions and results on measurability of stochastic processes.

We recall that a mapping between two measurable spaces $(I, \mathcal{I})$ and $(O, \mathcal{O})$, defined by $h\colon I \to O$, is said to be $\mathcal{O}/\mathcal{I}$*-measurable* if

$$h^{-1}(A) \in \mathcal{I}, \quad \forall A \in \mathcal{O}, \tag{A.6}$$

where h^{-1} denotes the pre-image of a set under h.

DEFINITION A.5: Let X_t be a E-valued stochastic process. It is said to be

(1) *measurable* if and only if the mapping

$$(t, \omega) \mapsto X_t(\omega), \quad (t, \omega) \in \mathbb{R}_+ \times \Omega \tag{A.7}$$

is $\mathcal{E}/\mathcal{B}_+ \otimes \mathcal{F}$-measurable;

(2) $\mathcal{F}_t$*-adapted*[1] if and only if, for all fixed $t \geqslant 0$, the mapping

$$\omega \mapsto X_t(\omega), \quad \omega \in \Omega \tag{A.8}$$

is $\mathcal{E}/\mathcal{F}_t$-measurable;

1 The correct notation would be "$(\mathbb{P}, \mathcal{F}_t)$"*-adapted*. However, for sake of simplicity, in the sequel we will omit to specify the probability $\mathbb{P}$, that will be always understood.

(3) $\mathcal{F}_t$-*progressive* if and only if, for all fixed $t \geqslant 0$, the mapping

$$(s, \omega) \mapsto X_s(\omega), \quad (s, \omega) \in [0, t] \times \Omega \tag{A.9}$$

is $\mathcal{E}/\mathcal{B}\big([0, t]\big) \otimes \mathcal{F}_t$-measurable;

REMARK: For the sake of clarity, we remember that $\mathcal{B}(A)$ denotes the Borel subsets of $A \subset \mathbb{R}$. The symbol $\mathbb{R}_+$ indicates the set $[0, +\infty)$ and, correspondingly, $\mathcal{B}_+ = \mathcal{B}\big([0, +\infty)\big)$.

For our purposes, the previous notions of measurability are not sufficient. We need to introduce also *predictable processes*, whose definition requires a new σ-field, namely the *predictable σ-field*.

DEFINITION A.6 (Predictable σ-Field, Predictable Process): Let $\mathcal{F}_t$ be a filtration defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{P}(\mathcal{F}_t)$ be the σ-field over $(0, +\infty) \times \Omega$ generated by the rectangles of the form

$$(s, t] \times A, \quad 0 \leqslant s \leqslant t, \quad A \in \mathcal{F}_s. \tag{A.10}$$

Then, $\mathcal{P}(\mathcal{F}_t)$ is called the $\mathcal{F}_t$-*predictable σ-field* over $(0, +\infty) \times \Omega$.

A E-valued process X_t is said to be $\mathcal{F}_t$-*predictable* if and only if X_0 is $\mathcal{F}_0$-measurable and the mapping

$$(t, \omega) \mapsto X_t(\omega), \quad (t, \omega) \in (0, +\infty) \times \Omega \tag{A.11}$$

is $\mathcal{E}/\mathcal{P}(\mathcal{F}_t)$-measurable.

REMARK: It is possible to simplify the form of the rectangles that generate a predictable σ-field by taking

$$(s, +\infty) \times A, \quad s \geqslant 0, \quad A \in \mathcal{F}_s, \tag{A.12}$$

instead of the set of generators (A.10).

In the case where the measure space $(E, \mathcal{E})$ satisfies some additional hypotheses, we can state sufficient conditions for a stochastic process to be progressive or predictable.

THEOREM A.1: *Let E be a metrizable topological space and let X_t be a E-valued process adapted to a filtration $\mathcal{F}_t$ on $(\Omega, \mathcal{F})$.*

(I) *If X_t is right-continuous then X_t is $\mathcal{F}_t$-progressive.*

(II) *If X_t is left-continuous then X_t is $\mathcal{F}_t$-predictable.*

Moreover, a $\mathcal{F}_t$-predictable process is $\mathcal{F}_t$-progressive.

A.2 MARKOV PROCESSES

In this section we will expose some basic notions about a fundamental class of processes, that is the object of this work, namely *Markov processes*. Here we will assume that the previously given measure space $(E, \mathcal{E})$ is also a topological space. We will denote by $b(E)$ the set of bounded measurable functions $f \colon E \to \mathbb{R}$.

A Markov process is a stochastic process X_t whose main feature is that its increments, i. e. the quantities $X_{t+h} - X_t$, where $h > 0$, depend upon the past of the process itself only through its present value X_t. Formalizing our previous statement, we give the following definition.

DEFINITION A.7 (Markov Process): Let X_t be an E-valued process defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{F}_t$-adapted, for some filtration $\mathcal{F}_t$ on $(\Omega, \mathcal{F})$. Let $\mathcal{F}_t^\infty = \sigma(X_s, s \geqslant t)$. Then, X_t is called a $\mathcal{F}_t$-*Markov process* if and only if, for all $t \geqslant 0$, $\mathcal{F}_t^\infty$ and $\mathcal{F}_t$ are independent given X_t.

In particular, if X_t is a Markov process, then for all $0 \leqslant s \leqslant t$ and for all $f \in b(E)$ the next formula holds:

$$\mathbb{E}\left[f(X_t) \mid \mathcal{F}_s\right] = \mathbb{E}\left[f(X_t) \mid X_s\right] \tag{A.13}$$

A Markov process may admit a *transition function*, an object that describes the probabilistic structure of the transitions of the process from a given state x at time s to a specified set A at a future time t.

DEFINITION A.8 (Transition Function): Let $P_{s,t}(x, A)$, $x \in E$, $A \in \mathcal{E}$, $0 \leqslant s \leqslant t$, be a function from $E \times \mathcal{E}$ into $\mathbb{R}_+$ such that:

(1) $A \mapsto P_{s,t}(x, A)$ is a probability on $(E, \mathcal{E})$ for all $x \in E$,

(2) $x \mapsto P_{s,t}(x, A)$ is $\mathcal{B}_+/\mathcal{E}$-measurable for all $A \in \mathcal{E}$,

(3) the *Chapman-Kolmogorov* equation holds, i. e. for all $0 \leqslant s \leqslant u \leqslant t$, all $x \in E$ and all $A \in \mathcal{E}$

$$P_{s,t}(x, A) = \int_E P_{s,u}(x, dy)\, P_{u,t}(y, A). \tag{A.14}$$

Then, the function $P_{s,t}(x, A)$ is called a *Markov transition function* on $(E, \mathcal{E})$.
If $P_{s,t}(x, A) = P_{t-s}(x, A)$, then the Markov transition function is said to be *homogeneous*.

DEFINITION A.9: Let X_t be a E-valued $\mathcal{F}_t$-Markov process and let $P_{s,t}(x, A)$ be a Markov transition function on $(E, \mathcal{E})$. If

$$\mathbb{E}\left[f(X_t) \mid \mathcal{F}_s\right] = \int_E f(y)\, P_{s,t}(X_s, dy) \tag{A.15}$$

for all $0 \leqslant s \leqslant t$ and all $f \in b(E)$, then the Markov process X_t is said to *admit the Markov transition function* $P_{s,t}(x, A)$.
If $P_{s,t}(x, A)$ is homogeneous, the Markov process X_t is said to be a *homogeneous* $\mathcal{F}_t$*-Markov process* admitting the Markov transition function $P_{t-s}(x, A)$.

If we take the function $f(x) = \mathbb{1}_A(x)$, $A \in \mathcal{E}$, in equation (A.15), it specializes to

$$\mathbb{P}\left(X_t \in A \mid \mathcal{F}_s\right) = P_{s,t}(X_t, A). \tag{A.16}$$

It is then clear that the function $P_{s,t}(x, A)$ is nothing but the probability that the value of the process X_t at time $t \geqslant 0$ belongs to the set $A \in E$, starting from the state $X_s = x \in E$ at time $s \in [0, t]$ and conditionally to $\mathcal{F}_s$.
The Chapman-Kolmogorov equation symbolizes the fact that we can express the transition of the process X_t from the state x to the set A as an infinite sum of infinitesimal disjoint transitions through intermediate states at times $u \in [s, t]$.

In the case of a homogeneous Markov process we can say something more about its transition function (if it admits one) $P_{t-s}(x, A)$. Let us define, for each $t \geqslant 0$, the following operator P_t, mapping $b(E)$ onto itself:

$$P_t\, f(x) = \int_E f(y)\, P_t(x, dy). \tag{A.17}$$

Then, by the Chapman-Kolmogorov equation (A.14) we obtain

$$P_t\, P_s = P_{t+s}, \quad s \geqslant 0. \tag{A.18}$$

Thus, the family $(P_t)_{t \geqslant 0}$ forms a *semigroup*, called the *transition semigroup* associated to the stochastic process X_t.
Finally, suppose that, for some function $f \in b(E)$, the limit

$$\mathcal{L}\, f(x) = \lim_{t \downarrow 0} \frac{P_t\, f(x) - f(x)}{t} \tag{A.19}$$

exists for all $x \in E$. Then, denoting by $\mathcal{D}(\mathcal{L})$ the family of functions $f \in b(E)$ such that the limit in equation (A.19) exists for all $x \in E$, the operator $\mathcal{L}$ is defined for all $f \in \mathcal{D}(\mathcal{L})$ and it is called the *infinitesimal generator* of the process X_t.

A.3 MARTINGALES

We now turn our attention to a fundamental kind of stochastic processes that are *martingales*. Before giving a rigorous definition of this concept, let us express in few words and in an informal way what this concept symbolizes.

In a number of situations, one can be interested in predicting the future value of a stochastic process X_t, given the knowledge of some past and present "events". As is well known, and recalling the concept of filtration previously introduced, a way to do this is to compute the conditional expectation of the random variable X_t with respect to the filtration $\mathcal{F}_s$, $0 \leqslant s \leqslant t$. It can happen that the amount of information included in the filtration $\mathcal{F}_t$ may help in reducing the level of uncertainty about the future outcome of the stochastic process. This is precisely not the case for a martingale, i. e. with what is known through the history $\mathcal{F}_t$ the best estimate that one can do for the future value of the stochastic process X_t is its present value.

Let us now formalize this statement and show two simple examples.

DEFINITION A.10 (Martingale): Let $\mathcal{F}_t$ be a filtration on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let X_t be a E-valued stochastic process and c a nonnegative real number. If

(I) X_t is adapted to $\mathcal{F}_t$,

(II) $\mathbb{E}[|X_t|] < \infty, \quad \forall t \in [0, c]$,

(III) $\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s, \quad \mathbb{P} - a.s., \quad \forall 0 \leqslant s \leqslant t \leqslant c$,

then the process X_t is called a $\mathcal{F}_t$*-martingale over* $[0, c]$. If it is a $\mathcal{F}_t$-martingale over $[0, c]$ for all $c \geqslant 0$, then X_t is called a $\mathcal{F}_t$*-martingale*.

EXAMPLE A.1: Let X_t be a real-valued stochastic process and let Y be a square-integrable real-valued random variable. Suppose that we are interested in constructing the best quadratic estimate of Y given the knowledge of the process X_t up to time t, i. e. given the natural filtration $\mathcal{F}_t^X$. As previously recalled, the answer to this question is to compute $\mathbb{E}\left[Y \mid \mathcal{F}_t^X\right]$. Indeed, defining $Y_t = \mathbb{E}\left[Y \mid \mathcal{F}_t^X\right]$, $t \geqslant 0$, it can be shown that

$$\mathbb{E}\left[(Y - Y_t)^2\right] \leqslant \mathbb{E}\left[(Y - Z)^2\right] \tag{A.20}$$

for all square-integrable and $\mathcal{F}_t^X$-measurable random variables Z.

The process Y_t is a simple example of $\mathcal{F}_t^X$-martingale. In fact, we have for all $0 \leqslant s \leqslant t$

$$\mathbb{E}\left[Y_t \mid \mathcal{F}_s^X\right] = \mathbb{E}\left[\mathbb{E}\left[Y \mid \mathcal{F}_t^X\right] \mid \mathcal{F}_s^X\right] = \mathbb{E}[Y \mid \mathcal{F}_s] = Y_s, \tag{A.21}$$

which is obvious, since the process Y_t is constructed solely upon the information about the process X_t.

EXAMPLE A.2 (Processes with Independent Increments): Let X_t be a real-valued process and suppose that, for all $0 \leqslant s \leqslant t$, its increments $X_t - X_s$ are independent of $\mathcal{F}_s^X$. The process X_t is then said to be with *independent increments*.

If we suppose, moreover, that $\mathbb{E}[|X_t|] < \infty$ and $\mathbb{E}[X_t] = 0$, for all $t \geqslant 0$, then X_t is a $\mathcal{F}_t^X$-martingale. It suffices to observe that, by linearity of the conditional expectation operator,

$$\mathbb{E}\left[X_t \mid \mathcal{F}_s^X\right] = \mathbb{E}\left[X_s \mid \mathcal{F}_s^X\right] + \mathbb{E}\left[X_t - X_s \mid \mathcal{F}_s^X\right] = X_s + 0. \tag{A.22}$$

A stronger type of martingale is given in the following definition.

DEFINITION A.11 (Square-Integrable Martingale): Let X_t be a $\mathcal{F}_t$-martingale over $[0, c]$, for some $c \geqslant 0$. If

$$\mathbb{E}\left[|X_c|^2\right] < \infty, \tag{A.23}$$

then X_t is called a *square-integrable* $\mathcal{F}_t$*-martingale over* $[0, c]$.

If X_t is a $\mathcal{F}_t$-martingale such that

$$\sup_{t \geqslant 0} \mathbb{E}\left[|X_t|^2\right] < \infty, \tag{A.24}$$

then X_t is called a *square-integrable* $\mathcal{F}_t$*-martingale*.

The concept of martingale can be generalized to include a larger class of processes, with the notion of *local martingale*. Its definition is tightly linked to another probabilistic object, a *stopping time*.

DEFINITION A.12 (Stopping Time): Let $\mathcal{F}_t$ be a filtration on $(\Omega, \mathcal{F})$. A random variable $\tau \colon \Omega \to [0, +\infty]$ is called a $\mathcal{F}_t$*-stopping time* if

$$\{\tau \leqslant t\} \in \mathcal{F}_t, \quad \forall t \in [0, +\infty). \tag{A.25}$$

DEFINITION A.13 (Local Martingale): Let X_t be a E-valued stochastic process adapted to a filtration $\mathcal{F}_t$ on $(\Omega, \mathcal{F})$. Let $(\tau_n)_{n \geqslant 1}$ be an increasing sequence of $\mathcal{F}_t$-stopping times such that

(I) $\lim_{n \uparrow \infty} S_n = +\infty$, $\mathbb{P} - a.s.$,

(II) for each $n \geqslant 1$, $X_{t \wedge \tau_n}$ is a $\mathcal{F}_t$-martingale.

Then X_t is called a $\mathcal{F}_t$*-local martingale*.

We conclude this Appendix with a theorem that links martingales and predictable processes with stochastic integration. In this context, all the stochastic integrals are always to be understood, if not otherwise specified, as *Lebesgue-Stieltjes* integrals.

Before stating the theorem, let us recall that a stochastic process X_t is said to be of *bounded variation* if its trajectories are $\mathbb{P} - a.s.$ of bounded variation over bounded intervals. It is said to be of *integrable variation* if the additional condition $\mathbb{E}\left[\int_0^t |dX_s|\right] < \infty$ holds for all $t \geqslant 0$.

THEOREM A.2 (Integration with Respect to Bounded Variation Martingales): *Let* M_t *be a* $\mathcal{F}_t$*-martingale of integrable bounded variation. Let* C_t *be a* $\mathcal{F}_t$*-predictable process such that*

$$\mathbb{E}\left[\int_0^1 |C_s|\,|dM_s|\right] < \infty. \tag{A.26}$$

Then the process $\int_0^t C_s \, dM_s$ *is a* $\mathcal{F}_t$*-martingale over* $[0, 1]$.

BIBLIOGRAPHY

[1] L. Aggoun and R. J. Elliott. *Measure Theory and Filtering. Introduction and Applications*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2004.

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

[3] P Baxendale, P. Chigansky, and R. Lipster. Asymptotic stability of the Wonham filter: Ergodic and nonergodic signals. *SIAM J. Control Optim.*, 43(2):643–669, 2004.

[4] P. Brémaud. *Point Processes and Queues*. Springer Series in Statistics. Springer-Verlag, New York, 1981.

[5] F. Confortola and M. Fuhrman. Filtering of continuous-time Markov chains with noise-free observation and applications. *Stochastics An International Journal of Probability and Stochastic Processes*, 85(2):216–251, 2013.

[6] D. Crisan, M. Kouritzin, and J. Xiong. Nonlinear filtering with signal dependent observation noise. *Electron. J. Probab.*, 14:1863–1883, 2009.

[7] M.H.A. Davis. Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models (with discussion). *J. Roy. Statist. Soc. Ser. B*, 46:353–388, 1984.

[8] M.H.A. Davis. *Markov Models and Optimization*, volume 49 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, London, 1993.

[9] C. Dellacherie. *Capacités et Processus Stochastiques*. Springer, Berlin, 1972.

[10] R.J. Elliott, L. Aggoun, and J.B. Moore. *Hidden Markov Models. Estimation and Control*, volume 29 of *Application of Mathematics*. Springer-Verlag, New York, 1995.

[11] S. Haykin. *Kalman filtering and neural networks*. John Wiley & Sons., New York, 2001.

[12] M. Joannides and F. LeGland. Nonlinear filtering with continuous time perfect observations and noninformative quadratic variation. In *Proceeding of the 36th IEEE Conference on Decision and Control*, pages 1645–1650, 1997.

[13] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, New York, 2nd edition, 1991.

[14] H. Korezlioglu and W. Runggaldier. Filtering for nonlinear systems driven by nonwhite noises: An approximation scheme. *Stoch. Stoch. Rep.*, 44:65–102, 1993.

[15] R. Lipster and A. N. Shiryaev. *Statistics of Random Processes*, volume I: General Theory. Springer, New York, 2nd edition, 2001.

[16] Y. Takeuchi and H. Akashi. Least-squares state estimation of systems with state-dependent observation noise. *Automatica*, 21 (3):303–313, 1985.

[17] S. Watanabe. Additive functionals of Markov processes and Lévy Systems. *Jap. J. Math.*, 34:53–79, 1964.

[18] J. Xiong. *An Introduction to Stochastic Filtering Theory*. Oxford University Press, New York, 2008.

RINGRAZIAMENTI

La parte più ardua da scrivere in una Tesi è quella dei ringraziamenti. Punto. Dato di fatto inconfutabile, almeno per me. Specialmente se, per di più, riguardano un periodo lungo e importante della vita come quello degli studi universitari. Il motivo? Una legge rigorosa e ineluttabile: inizierai a scriverli da capo almeno dieci volte e (quasi) certamente dimenticherai qualcuno. Il tesista, notoriamente in crisi a questo punto del lavoro, si scervella in ogni modo pur di sfuggirle: io ho scelto di iniziare a scrivere queste parole con qualche giorno di anticipo rispetto all'ora X. No non sto parlando della discussione della Tesi, ma della scadenza per il suo deposito. Momento meno emozionante del primo ma in grado di provocare cardiopalmi di seria entità egualmente.

Cionondimeno, a tale legge nessuno può sfuggire, qualunque precauzione si adotti. Se è vero che mi è già accaduto di abbozzare e cancellare un periodo almeno dieci volte, ancora non so se effettivamente dimenticherò qualcuno: succederà, credetemi. A scanso di equivoci, chiedo in anticipo perdono. E ora bando alle ciance!

Ringrazio per primo il mio relatore, Prof. Marco Fuhrman, che mi ha dato l'opportunità di svolgere questo lavoro: anzitutto, ha saputo indirizzarmi verso questo ambito, il filtraggio di processi stocastici, che inizialmente, lo confesso, snobbavo un po'. Non solo, è pure riuscito a farmici appassionare, non bastasse già il numero di cose a cui vorrei dedicarmi prima o dopo nella vita. In più, nonostante i suoi innumerevoli impegni, mi ha seguito prima nel procedere del lavoro e poi nella scrittura della Tesi, dedicandomi, a volte, anche mattinate intere. Un grazie, davvero, di cuore.

Arriva il momento dei dedicatari di questa Tesi, mamma e papà. Aiuto, devo cercare di mettere insieme almeno una frase di senso compiuto! Grazie, grazie, mille volte grazie: mi avete sempre sostenuto, a prescindere da ogni situazione, bella o brutta, da ogni periodo, positivo o negativo. Non avete mai perso la fiducia in me, anche quando non ne avrei, forse, meritata o quando, diverse volte, sono stato io stesso a perderla nei miei confronti. Non sarei arrivato a questo traguardo senza di voi, senza il vostro costante incoraggiamento. Perché se mi guardo indietro e volgo il pensiero a sei anni fa, quando timidamente vi chiesi, dopo *tutto*, se potevo sostenere il Test On Line per avere l'occasione di studiare qui, al Politecnico di Milano, vedo che minimamente pensavo di poter arrivare così lontano, con questa Tesi, con i sogni per il mio futuro, lì dove voi desideravate che giungessi, fin da quando ero bambino. C'era bonaccia e voi avete preso a soffiare forte, la vela si è gonfiata: siete stati più forti *d'u punenti e d'a* *tramontana*, che da sempre sferzano sui nostri corpi. Non basteranno mai i *grazie*.

Bene, ricomponiamoci, ma restando in famiglia. Grazie alle mie sorelle, Cinzia e Cristina e ai miei cognati, Aldo e Simon, perché anche voi avete sempre creduto in me e ci siete stati. Non si può non ringraziare i miei nipotini, Rossella, Enzo e Fabrizio, che, anche se lontani, con la loro semplicità e gioia mi hanno sempre strappato un sorriso e riscaldato il cuore.

È il turno degli amici (e qui sono dolori, prenderò una dose di fosforo prima di iniziare!). Grazie a Sara, Sofia e Angelina, non solo colleghe e compagne di (s)ventura in questa lunga traghettata (è il caso di dire *sulla Nave* o è troppo banale e scontato?). Mi avete regalato momenti indimenticabili, che porterò sempre con me: quante risate, quanti pranzi, parole crociate fatte mentre avremmo dovuto seguire, che so, le lezioni di Analisi II, un "che c**o!" a sproposito una mattina di dicembre davanti a una lavagnata densa di conti. Soprattutto mi avete sempre incoraggiato dicendomi "che sei bravo", quando penso che siate voi quelle più brave.

Un grazie anche per Andrea: hai condiviso i momenti di follia miei, di Angelina e di Sofia nonostante volessi seguire di più le lezioni (siamo sicuri? Non sono mica mancati i tuoi raggelanti contributi, il "pollinomio" per citarne uno). La tua caparbietà, tenacia e intelligenza mi hanno sempre colpito e mi hanno dato uno stimolo a proseguire.

Fuori dall'università: un mondo, passato, presente e futuro. Il ponte che unisce tutto ciò è senz'altro Valentina, a cui vanno milioni di grazie. Sei stata e so che continuerai a essere una preziosissima presenza nella mia vita: anche se la distanza è tanta, anche se gli impegni impediscono di sentirci quanto vorremmo, sei un'amica fantastica, capace di comprendermi nel profondo, conoscendomi nel profondo (a quanto siamo? 15 anni se non sbaglio?). Del resto, chi sono quei due pazzi che, nelle notti d'estate, fanno incursioni in *Panoramica*, mettendosi a discutere un secondo dei fatti personali e l'altro di strampalate (ma neanche troppo) teorie sulla cosmogonia e facendo discussioni pseudo-filosofico-matematico-chimico-fisiche con la costante paura (tua) degli alieni (babau)? Grazie per questi momenti e per tutti quelli dei quali sono costellati i miei ricordi. Grazie per tutto il sostegno che mi hai sempre offerto, per la tua compartecipazione autentica e sincera alle mie emozioni, per avermi ascoltato e consigliato, insomma: per tutto.

Un grazie a Carlo, Chiara ed Eleonora. Per merito di quel cialtrone che in fondo sarà menzionato (pensavi mi fossi dimenticato, eh? Malfidato!), siete diventati per me amici insostituibili, coi quali ho condiviso molta parte della mia nord-italica vita. Serate, cene, discussioni dalle più auliche alle più becere, spedizioni all'Auchan, all'Ikea, pomeriggi in piscina e chi più ne ha più ne metta, spesso in compagnia di Marcello e Michele, a cui pure va un grazie. Attimi di relax nella giungla padana, ma non solo. La cosa più importante è che siete stati prodighi di preziosi consigli e capaci di quell'ascolto che sino a oggi mi ha sostenuto.

Per avermi aiutato a capire meglio me stesso, il mio mondo e quello intorno a me, ringrazio il Dott. Marco Roscio. Per avermi regalato, per tramite della musica, giovedì, venerdì sera e svariati bellissimi momenti in questi ultimi anni, tra prove, servizi, saggi e concerti, ringrazio gli amici del Corpo Bandistico Legnanese e del Coro dei Ragazzi. Per avermi fatto sentire in tante serate e in svariate altre occasioni il calore come di una famiglia, anche se lontano dalla mia, grazie a Rossella, Roberto ed Elena.

Per tutto ciò che rappresenta la parte più importante dei miei giorni, grazie a te, Joshua. Mi è davvero difficile trovare le parole per esprimere quanto vorrei dire, mi sembra d'aver preso il tuo posto! E allora dirò alcuni dei miliardi di motivi per cui ringraziarti: grazie per avermi supportato sempre e soprattutto sopportato, per avermi aiutato a migliorare me stesso, per avermi condotto nel tuo mondo, fatto partecipe, fisicamente e mentalmente, delle tue giornate, per avermi fatto riscoprire quell'universo musicale che da sempre mi appartiene, perché hai sempre creduto in me, perché mi hai sempre regalato un sorriso anche nei momenti di maggior sconforto, perché altrimenti, come dici tu, "la mia vita sarebbe triste e noiosa". Grazie per avermi dato, con la tua semplicità e schiettezza, quell'affetto unico e speciale che mi ha permesso di arrivare fin qui.