We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter 17

Recent Advances in Nonlinear Filtering with a Financial Application to Derivatives Hedging under Incomplete Information

Claudia Ceci and Katia Colaneri

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.70060

Abstract

In this chapter, we present some recent results about nonlinear filtering for jump diffusion signal and observation driven by correlated Brownian motions having common jump times. We provide the Kushner-Stratonovich and the Zakai equation for the normalized and the unnormalized filter, respectively. Moreover, we give conditions under which pathwise uniqueness for the solutions of both equations holds. Finally, we study an application of nonlinear filtering to the financial problem of derivatives hedging in an incomplete market with partial observation. Precisely, we consider the risk-minimizing hedging approach. In this framework, we compute the optimal hedging strategy for an informed investor and a partially informed one and compare the total expected squared costs of the strategies.

Keywords: nonlinear filtering, jump diffusions, risk minimization, Galtchouk-Kunita-Watanabe decomposition, partial information

1. Introduction

Bayesian inference and stochastic filtering are strictly related, since in both approaches, one wants to estimate quantities which are not directly observable. However, while in Bayesian inference, all uncertainty sources are considered as random variables, stochastic filtering refers to stochastic processes. It also covers many situations, from linear to nonlinear case, with various types of noises.

The objective of this chapter is to present nonlinear filtering results for Markovian partially observable systems where the state and the observation processes are described by jump diffusions with correlated Brownian motions and common jump times. We also aim at applying this

INTECH
open science | open minds

© 2017 The Author(s). Licensee InTech. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

theory to the financial problem of derivatives hedging for a trader who has limitative information on the market.

A filtering model is characterized by a signal process, denoted by X, which cannot be observed directly, and an observation process denoted by Y whose dynamics depends on X. The natural filtration of Y, $\mathbb{F}^Y = \{\mathcal{F}_t^Y, t \in [0, T]\}$, represents the available information. The goal of solving a filtering problem is to determine the best estimation of the signal X_t from the knowledge of $\mathcal{F}_t^Y$. Similar to optimal Bayesian filtering, we seek for the best estimation of the signal according to the minimum mean-squared error criterion, which corresponds to compute the posterior distribution of X_t given the available observations up to time t.

Historically, the first example of continuous-time filtering problem is the well-known Kalman-Bucy filter which concerns the case where Y gives the observation of X in additional Gaussian noise and both processes X and Y are modeled by linear stochastic differential equations. In this case, one ends up with a filter having finite-dimensional realization. Since then, the problem has been extended in many directions. To start, a number of authors including Refs. [1–3] studied the nonlinear case in the setting of additional Gaussian noise. Other references in a similar framework are given, for instance, by Refs. [4–8]. Subsequently also the case of counting process or marked point process observation has been considered (see Refs. [9–14] and reference therein). A more recent literature contains the case of mixed-type observations (marked point processes and diffusions or jump-diffusion processes), see, for, example, Refs. [15–18].

There are two major approaches to nonlinear filtering problems: the innovations method and the reference probability method. The latter is usually employed when it is possible to find an equivalent probability measure that makes the state X and the observations Y independent. This technique may appear problematic when, for instance, signal and observation are correlated and present common jump times. Therefore, in this chapter, we use the innovations approach which allows circumventing the technical issues arising in the reference probability method. By characterizing the innovation process and applying a martingale representation theorem, we can derive the dynamics of the filter as the solution of the Kushner-Stratonovich equation, which is a nonlinear stochastic partial integral differential equation. By considering the unnormalized version of the filter, it is possible to simplify this equation and make it at least linear. The resulting equation is called the Zakai equation, and due to its linear nature, it is of particular interest in many applications. We also compute the dynamics of the unnormalized filter, and we investigate pathwise uniqueness for the solutions of both equations. Normalized and unnormalized filters are probability measure and finite measure-valued processes, respectively, and therefore in general infinite-dimensional. Due to this, various recursive algorithms for statistical inference have come in to address this intractability, such as extended Kalman filter, statistical linearization, or particle filters. These algorithms intend to estimate both state and parameters. For the parameter estimation, we also mention the expectation maximization (EM) algorithm which enables to estimate parameters in models with incomplete data, see, for example, Ref. [19].

The success of the filtering theory over the years is due to its use in a great variety of problems arising from many disciplines such as engineering, informational sciences and mathematical finance. Specifically, in this chapter, we have a financial application in view. In real financial

http://dx.doi.org/10.5772/intechopen.70060

markets, it is reasonable that investors cannot fully know all the stochastic factors that may influence the prices of negotiated assets, since these factors are usually associated with economic quantities which are hard to observe. Filtering theory represents a way to measure, in some sense, this uncertainty. A consistent part of the literature over the last years has considered stochastic factor models under partial information for analyzing various financial problems, as, for example, pricing and hedging of derivatives, optimal investment, credit risk, and insurance modeling. A list, definitely nonexhaustive, is given by Refs. [15, 16, 20–26]).

In the following, we consider the problem of a trader who wants to determine the hedging strategy for a European-type contingent claim with maturity T in an incomplete financial market where the investment possibilities are given by a riskless asset, assumed to be the numéraire, and a risky asset with price dynamics given by a geometric jump diffusion, modeled by the process Y. We assume that the drift, as well as the intensity and the jump size distribution of the price process, is influenced by an unobservable stochastic factor X, modeled as a correlated jump diffusion with common jump times. By common jump times, we intend to take into account catastrophic events which affect both the asset price and the hidden state variable driving its dynamics. The agent knows the asset prices, since they are publicly available, and trades on the market by using the available information $\mathbb{F}^Y$.

Partial information easily leads to incomplete financial markets as clearly the number of random sources is larger than the number of tradeable risky asset. Therefore, the existence of a self-financing strategy that replicates the payoff of the given contingent claim at maturity is not guaranteed. Here, we assume that the risky asset price is modeled under a martingale measure, and we choose the risk-minimization approach as hedging criterion, see, for example, Refs. [27, 28].

According to this method, the optimal hedging strategy is the one that perfectly replicates the claim at maturity and has minimum cost in the mean-square sense. Equivalently, we say that it minimizes the associated risk defined as the conditional expected value of the squared future costs, given the available information (see Refs. [28, 29] and references therein).

The risk-minimizing hedging strategy under restricted information is strictly related to Galtchouk-Kunita-Watanabe decomposition of the random variable representing the payoff of the contingent claim in a partial information setting. Here, we provide a characterization of the risk-minimizing strategy under partial information via this orthogonal decomposition and obtain a representation in terms of the corresponding risk-minimizing hedging strategy under full information (see, e.g., Refs. [29, 30]) via predictable projections on the available information flow by means of the filter. Finally, we investigate the difference of expected total risks associated with the optimal hedging strategies under full and partial information.

The chapter has the following structure. In Section 2, we introduce the general framework. In Section 3, we study the filtering equations. In particular, we derive the dynamics for both normalized and unnormalized filters, and we investigate uniqueness of the solutions of the Kushner-Stratonovich and the Zakai equation. In Section 4, we analyze a financial application to risk minimization by computing the optimal hedging strategies for a European-type contingent claim under full and partial information and providing a comparison between the corresponding expected squared total costs.

2. The setting

We consider a pair of stochastic processes (X,Y), with values on $\mathbb{R} \times \mathbb{R}$ and càdlàg trajectories, on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, where $\mathbb{F} = \{\mathcal{F}_t, t \in [0, T]\}$ is a filtration satisfying the usual condition of right continuity and completeness, and T is a fixed time horizon. The pair (X, Y) represents a partially observable system, where X is a signal process that describes a phenomenon which is not directly observable and Y gives the observation of X, and it is modeled by a process correlated with the signal, having possibly common jump times.

Remark 1. *In view of the financial application discussed in Section 4, Y represents the price of some risky asset, while X is an unknown stochastic factor, which may describe the activity of other markets, macroeconomic factors or microstructure rules that influences the dynamics of the stock price process.*

We define the *observed history* as the natural filtration of the observation process Y, that is, $\mathbb{F}^Y = \{\mathcal{F}_t^Y\}_{t \in [0,T]}$, where $\mathcal{F}_t^Y := \sigma(Y_s, 0 \le s \le t)$. The σ-algebra $\mathcal{F}_t^Y$ can be interpreted as the information available from observations up to time t. We aim to compute the best estimate of the signal X from the available information, in the quadratic sense. In other terms, this corresponds to determine the filter which furnishes the conditional distribution of X_t given $\mathcal{F}_t^Y$, for every $t \in [0, T]$.

Let $\mathcal{M}(\mathbb{R})$ be the space of finite measures over $\mathbb{R}$ and $\mathcal{P}(\mathbb{R})$ the subspace of the probability measures over $\mathbb{R}$. Given $\mu \in \mathcal{M}(\mathbb{R})$, for any bounded measurable function f, we write

$$\mu(f) = \int_{\mathbb{R}} f(x)\mu(\mathrm{d}x). \tag{1}$$

Definition 2. *The filter is the $\mathbb{F}^Y$-càdlàg process π taking values in $\mathcal{P}(\mathbb{R})$ defined by*

$$\pi_t(f) := \mathbb{E}\left[f(t, X_t)|\mathcal{F}_t^Y\right] = \int_{\mathbb{R}} f(t, x)\pi_t(\mathrm{d}x), \tag{2}$$

for all bounded and measurable functions $f(t, x)$ on $[0, T] \times \mathbb{R}$.

In the sequel, we denote by π_{t-} the left version of the filter and for all functions $F(t, x, y)$ such that $\mathbb{E}|F(t, X_t, Y_t)| < \infty$ (resp. $\mathbb{E}|F(t, X_{t-}, Y_{t-})| < \infty$) for every $t \in [0,T]$, we use the notation $\pi_t(F) := \pi_t(F(t, \cdot, Y_t))$ (resp. $\pi_{t-}(F) := \pi_{t-}(F(t, \cdot, Y_{t-}))$).

In this paper, we wish to consider the filtering problem for a partially observable system (X, Y) described by the following pair of stochastic differential equations:

$$\begin{cases} \mathrm{d}X_t = b_0(t, X_t)\mathrm{d}t + \sigma_0(t, X_t)\mathrm{d}W_t^0 + \int_Z K_0(t, X_{t-}; \zeta)N(\mathrm{d}t, \mathrm{d}\zeta); X_0 = x_0 \in \mathbb{R} \\ \mathrm{d}Y_t = b_1(t, X_t, Y_t)\mathrm{d}t + \sigma_1(t, Y_t)\mathrm{d}W_t^1 + \int_Z K_1(t, X_{t-}, Y_{t-}; \zeta)N(\mathrm{d}t, \mathrm{d}\zeta); Y_0 = y_0 \in \mathbb{R} \end{cases} \tag{3}$$

http://dx.doi.org/10.5772/intechopen.70060

where W^0 and W^1 are correlated $(\mathbb{F}, \mathbf{P})$-Brownian motions with correlation coefficient $\rho \in [-1,1]$ and $N(\mathrm{d}t, \mathrm{d}\zeta)$ is a Poisson random measure on $\mathbb{R}^+ \times Z$ whose intensity $\nu(\mathrm{d}\zeta)\mathrm{d}t$ is a σ – finite measure on a measurable space $(Z, \mathcal{Z})$. Here, b_0, b_1, σ_0, σ_1, K_0, and K_1 are $\mathbb{R}$-valued and measurable functions of their arguments. In particular, $\sigma_0(t, x)$ and $\sigma_1(t, x, y)$ are strictly positive for every $(t, x, y) \in [0, T] \times \mathbb{R}^2$.

For the rest of the paper, we assume that strong existence and uniqueness for system Eq. (3) holds. Sufficient conditions are collected, for instance, in Ref. [18, Appendix]. These assumptions also imply Markovianity for the pair (X, Y).

Remark 3. *Note that the quadratic variation process of* Y *defined by*

$$[Y]_t = Y_t^2 - 2\int_0^t Y_{u^-}\mathrm{d}Y_u, \quad t \in [0, T], \tag{4}$$

is $\mathbb{F}^Y$-adapted and $[Y]_t = \int_0^t \sigma_1^2(u, Y_u)\mathrm{d}u + \sum_{u \le t}(\Delta Y_u)^2$, where $\Delta Y_t := Y_t - Y_{t-}$. Therefore, it is natural to assume that the signal X does not affect the diffusion coefficient in the dynamics of Y. If Y describes the price of a risky asset, this implies that the volatility of the stock price does not depend on the stochastic factor X.

The jump component of Y can be described in terms of the following integer-valued random measure on $[0, T] \times \mathbb{R}$:

$$m(\mathrm{d}t, \mathrm{d}z) = \sum_{s:\Delta Y_s \neq 0} \delta_{\{s, \Delta Y_s\}}(\mathrm{d}t, \mathrm{d}z), \tag{5}$$

where δ_a denotes the Dirac measure at point a. Note that the following equality holds:

$$\int_0^t\int_{\mathbb{R}} z m(\mathrm{d}s, \mathrm{d}z) = \int_0^t\int_Z K_1(s, X_{s^-}, Y_{s^-}; \zeta)N(\mathrm{d}s, \mathrm{d}\zeta). \tag{6}$$

For all $t \in [0, T]$, for all $A \in \mathcal{B}(\mathbb{R})$, we define the following sets:

$$d^0(t, x) := \{\zeta \in Z : K_0(t, x; \zeta) \neq 0\}, \quad d^1(t, x, y) := \{\zeta \in Z : K_1(t, x, y; \zeta) \neq 0\}, \tag{7}$$

$$d^A(t, x, y) := \{\zeta \in Z : K_1(t, x, y; \zeta) \in A \setminus \{0\}\} \subseteq d^1(t, x, y), \tag{8}$$

$$D_t^A := d^A(t, X_{t-}, Y_{t-}) \subseteq D_t := d^1(t, X_{t-}, Y_{t-}), \quad D_t^0 := d^0(t, X_{t-}). \tag{9}$$

Typically, we have $D_t^0 \cap D_t \neq \emptyset$ $\mathbf{P}$ $-$ *a.s.*, which means that state and observation may have common jump times. This characteristic is particularly meaningful in financial applications to model catastrophic events that produce jumps in both the stock price and the underlying stochastic factor that influences its dynamics.

To ensure existence of the first moment for the pair (X, Y) and non-explosiveness for the jump process governing the dynamics of X and Y, we make the following assumption:

Assumption 4.

$$\mathbb{E}\left[\int_0^T |b_0(t, X_t)| + \sigma_0^2(t, X_t) + \int_Z |K_0(t, X_{t-}; \zeta)|\nu(\mathrm{d}\zeta)\mathrm{d}t\right] < \infty, \tag{10}$$

$$\mathbb{E}\left[\int_0^T |b_1(t, X_t, Y_t)| + \sigma_1^2(t, Y_t) + \int_Z |K_1(t, X_{t-}, Y_{t-}; \zeta)|\nu(\mathrm{d}\zeta)\mathrm{d}t\right] < \infty, \tag{11}$$

$$\mathbb{E}\left[\int_0^T \nu(D_t^0 \cup D_t)\mathrm{d}t\right] < \infty. \tag{12}$$

Denote by $\eta^{\mathbf{P}}(\mathrm{d}t, \mathrm{d}z)$ the $(\mathbb{F}, \mathbf{P})$ compensator of $m(\mathrm{d}t, \mathrm{d}z)$ (see, e.g., Refs. [9, 31] for the definition).

Then, in Ref. [14, Proposition 2.2], it is proved that

$$\eta^{\mathbf{P}}(\mathrm{d}t, \mathrm{d}z) = \lambda(t, X_{t-}, Y_{t-})\phi(t, X_{t-}, Y_{t-}, \mathrm{d}z)\mathrm{d}t, \tag{13}$$

where

$$\lambda(t, x, y)\phi(t, x, y, \mathrm{d}z) = \int_{d^1(t,x,y)} \delta_{K^1(t,x,y;\zeta)}(\mathrm{d}z)\nu(\mathrm{d}\zeta) \tag{14}$$

and in particular $\lambda(t, x, y) = \nu(d^1(t, x, y))$.

Remark 5. *Let us observe that both the local jump characteristics* $(\lambda(t, X_{t-}, Y_{t-}), \phi(t, X_{t-}, Y_{t-}, \mathrm{d}z))$ *depend on X and, for all* $A \in \mathcal{B}(\mathbb{R})$, $\lambda(t, X_{t-}, Y_{t-})\phi(t, X_{t-}, Y_{t-}, A) = \nu(D_t^A)$ *provides the* $(\mathbb{F}, \mathbf{P})$ *-intensity of the point process* $N_t(A) := m((0, t] \times A)$. *According to this, the process* $\lambda(t, X_{t-}, Y_{t-}) = \nu(D_t)$ *is the* $(\mathbb{F}, \mathbf{P})$ *-intensity of the point process* $N_t(\mathbb{R})$ *which counts the total number of jumps of* Y *until time t*.

2.1. The innovation process

To derive the filtering equation, we use the innovations approach. This method requires to introduce a pair (I, m^π), called the *innovation process*, consisting of the $(\mathbb{F}^Y, \mathbf{P})$-Brownian motion and the $(\mathbb{F}^Y, \mathbf{P})$-compensated jump measure that drive the dynamics of the filter. The innovation also represents the building block of $(\mathbb{F}^Y, \mathbf{P})$ -martingales.

To introduce the first component of the innovation process, we assume that

$$\mathbb{E}\left[\exp\left\{\frac{1}{2}\int_0^T \left(\frac{b_1(t, X_t, Y_t)}{\sigma_1(t, Y_t)}\right)^2 \mathrm{d}t\right\}\right] < \infty, \tag{15}$$

and define

$$I_t := W_t^1 + \int_0^t \left(\frac{b_1(s, X_s, Y_s)}{\sigma_1(s, Y_s)} - \frac{\pi_s(b_1)}{\sigma_1(s, Y_s)}\right)\mathrm{d}s, \quad t \in [0, T]. \tag{16}$$

The process I is an $(\mathbb{F}^Y, \mathbf{P})$-Brownian motion (see, e.g., Ref. [4]) and the $(\mathbb{F}^Y, \mathbf{P})$-compensated jump martingale measure is given by

$$m^{\pi}(\mathrm{d}t, \mathrm{d}z) = m^{\pi}(\mathrm{d}t, \mathrm{d}z) - \pi_{t-}(\lambda\phi(\mathrm{d}z))\mathrm{d}t, \tag{17}$$

See, e.g. Ref. [14]. The following theorem provides a characterization of the $(\mathbb{F}^Y, \mathbf{P})$-martingale in terms of the innovation process.

Theorem 6 (A martingale representation theorem). *Under Assumption 4 and the integrability condition Eq. (15), every $(\mathbb{F}^Y, \mathbf{P})$-local martingale M admits the following decomposition:*

$$M_t = M_0 + \int_0^t \int_{\mathbb{R}} w_s(z) m^{\pi}(\mathrm{d}s, \mathrm{d}z) + \int_0^t h_s \mathrm{d}I_s, \quad t \in [0, T], \tag{18}$$

where $w(z) = \{w_t(z),\ t \in [0, T]\}$ is an $\mathbb{F}^Y$-predictable process indexed by z, and $h = \{h_t,\ t \in [0, T]\}$ is an $\mathbb{F}^Y$-adapted process such that

$$\int_0^T \int_{\mathbb{R}} |w_t(z)| \pi_{t-}(\lambda\phi(\mathrm{d}z))\mathrm{d}t < \infty, \quad \int_0^T h_t^2 \mathrm{d}t < \infty \quad \mathbf{P} - a.s.. \tag{19}$$

Proof. The proof is given in Ref. [17, Proposition 2.4]. Note that here condition (15) implies that $\mathbb{E}\left[\int_0^T \left(\frac{b_1(t, X_t, Y_t)}{\sigma_1(t, Y_t)}\right)^2 \mathrm{d}t\right] < \infty$, and also that the process L defined by

$$L_t = \exp\left(-\int_0^t \frac{b_1(s, X_s, Y_s)}{\sigma_1(s, Y_s)} \mathrm{d}W_s^1 - \frac{1}{2}\int_0^t \left(\frac{b_1(s, X_s, Y_s)}{\sigma_1(s, Y_s)}\right)^2 \mathrm{d}s\right), \tag{20}$$

for every $t \in [0, T]$, is an $(\mathbb{F}, \mathbf{P})$-martingale.

3. The filtering equations

Theorem 7 (The Kushner-Stratonovich equation). *Under Assumptions 4 and condition (15), the filter π solves the following Kushner-Stratonovich equation, that is, for every $f \in C_b^{1,2}([0, T] \times \mathbb{R})$:*

$$\pi_t(f) = f(0, x_0) + \int_0^t \pi_s(\mathcal{L}^X f)\mathrm{d}s + \int_0^t \int_{\mathbb{R}} w_s^{\pi}(f, z) m^{\pi}(\mathrm{d}s, \mathrm{d}z) + \int_0^t h_s^{\pi}(f)\mathrm{d}I_s, \quad t \in [0, T] \tag{21}$$

where

$$w_t^{\pi}(f, z) = \frac{\mathrm{d}\pi_{t-}(\lambda\phi f)}{\mathrm{d}\pi_{t-}(\lambda\phi)}(z) - \pi_{t-}(f) + \frac{\mathrm{d}\pi_{t-}(\overline{\mathcal{L}}f)}{\mathrm{d}\pi_{t-}(\lambda\phi)}(z), \tag{22}$$

$$h_t^{\pi}(f) = \sigma_1^{-1}(t)[\pi_t(b_1 f) - \pi_t(b_1)\pi_t(f)] + \rho\pi_t\left(\sigma_0 \frac{\partial f}{\partial x}\right). \tag{23}$$

Here, by $\frac{\mathrm{d}\pi_{t-}(\lambda\phi f)}{\mathrm{d}\pi_{t-}(\lambda\phi)}(z)$ and $\frac{\mathrm{d}\pi_{t-}(\overline{\mathcal{L}}f)}{\mathrm{d}\pi_{t-}(\lambda\phi)}(z)$, we mean the Radon-Nikodym derivatives of the measures $\pi_{t-}(\lambda f\phi(\mathrm{d}z))$ and $\pi_{t-}(\overline{\mathcal{L}}f)(\mathrm{d}z)$, with respect to $\pi_{t-}\left(\lambda\phi(\mathrm{d}z)\right)$. Moreover, the operator $\overline{\mathcal{L}}$ defined by $\overline{\mathcal{L}}_t f(\mathrm{d}z) := \overline{\mathcal{L}}f(.,\ Y_{t-}, \mathrm{d}z)$ is such that for every $A \in \mathcal{B}(\mathbb{R})$,

$$\overline{\mathcal{L}}f(t,x,y,A)=\int_{d^A(t,x,y)}[f(t,x+K_0(t,x;\zeta))-f(t,x)]\nu(\mathrm{d}\zeta) \tag{24}$$

takes into account common jump times between the signal X and the observation Y.

Finally, the operator $\mathcal{L}^X$ *given by*

$$\mathcal{L}^Xf(t,x)=\frac{\partial f}{\partial t}+b_0(t,x)\frac{\partial f}{\partial x}+\frac{1}{2}\sigma_0^2(t,x)\frac{\partial^2 f}{\partial x^2}+\int_Z\{f(t,x+K_0(t,x;\zeta))-f(t,x)\}\nu(\mathrm{d}\zeta). \tag{25}$$

denotes the generator of the Markov process X.

Proof. The theorem is proved in Ref. [17, Theorem 3.1].

Example 8 (Observation dynamics driven by independent point processes with unobservable intensities). In the sequel, we provide an example where the Kushner-Stratonovich equation simplifies and the Radon-Nikodym derivatives appearing in the dynamics of $\pi(f)$ reduce to ratios. Suppose that there exists a finite set of measurable functions $K_1^i(t,y)\neq 0$ for all $(t,y)\in[0,T]\times\mathbb{R}$, for $i\in\{1,\ldots,n\}$, such that the dynamics of Y is given by

$$\mathrm{d}Y_t=b_1(t,X_t,Y_t)\mathrm{d}t+\sigma_1(t,Y_t)\mathrm{d}W_t^1+\sum_{i=1}^n K_1^i(t,Y_{t-})\mathrm{d}N_t^i,\quad Y_0=y_0\in\mathbb{R}, \tag{26}$$

where N^i are independent counting processes with $(\mathbb{F},\mathbf{P})$ intensities $\lambda^i(t,X_{t-},Y_{t-})$.

For simplicity, in this example, we assume that X and Y have no common jump times. Then, the filtering Eq. (21) reads as

$$\begin{aligned}\pi_t(f)=&f(0,x_0)+\int_0^t\pi_s(\mathcal{L}^Xf)\mathrm{d}s+\int_0^t\left\{\sigma_1(s)^{-1}[\pi_s(b_1f)-\pi_s(b_1)\pi_s(f)]+\rho\pi_s\left(\sigma_0\frac{\partial f}{\partial x}\right)\right\}\mathrm{d}I_s\\&+\sum_{i=1}^n\int_0^t\mathbf{1}_{\pi_{s^-}(\lambda^i)>0}\frac{\pi_{s^-}(\lambda^if)-\pi_{s^-}(f)\pi_{s^-}(\lambda^i)}{\pi_{s^-}(\lambda^i)}\Big(\mathrm{d}N_s^i-\pi_{s^-}(\lambda^i)\mathrm{d}s\Big),\quad t\in[0,T].\end{aligned} \tag{27}$$

Note that Eq. (21) has an equivalent expression in terms of the operator $\mathcal{L}_0^X$, given by

$$\begin{aligned}\mathcal{L}_0^Xf(t,x,y)&=\mathcal{L}^Xf(t,x)-\overline{\mathcal{L}}f(t,x,y,\mathbb{R})\\&=\frac{\partial f}{\partial t}(t,x)+b_0(t,x)\frac{\partial f}{\partial x}+\frac{1}{2}\sigma_0^2(t,x)\frac{\partial^2 f}{\partial x^2}+\int_{d_t^1(t,x,y)^c}\{f(t,x+K_0(t,x,\zeta))-f(t,x)\}\nu(d\zeta),\end{aligned} \tag{28}$$

where $d^1(t,x,y)^c=\{\zeta\in Z:K_1(t,x,y,\zeta)=0\}$. Indeed, we get

$$\mathrm{d}\pi_t(f)=\{\pi_t(\mathcal{L}_0^Xf)+\pi_t(f)\pi_t(\lambda)-\pi_t(\lambda f)\}\mathrm{d}t+h_t^\pi\mathrm{d}I_t+\int_{\mathbb{R}}w^\pi(t,z)m(\mathrm{d}t,\mathrm{d}z). \tag{29}$$

Moreover, the filter has a natural recursive structure. To show this, define the sequence $\{T_n,Z_n\}_{n\in\mathbb{N}}$ of jump times and jump sizes of Y, that is, $Z_n=Y_{T_n}-Y_{T_n^-}$. These are observable

http://dx.doi.org/10.5772/intechopen.70060

data. Then, between two consecutive jump times the filter is governed by a diffusion process, that is, for $t \in (T_n \wedge T, T_{n+1} \wedge T)$

$$\pi_t(f) = \pi_{T_n}(f) + \int_{T_n}^{t} \{\pi_s(\mathcal{L}_0^X f) + \pi_s(f)\pi_s(\lambda) - \pi_s(\lambda f)\}\mathrm{d}s + \int_{T_n}^{t} h_s^{\pi}(f)\mathrm{d}I_s, \tag{30}$$

and at any jump time T_n occurring before time T, it is given by

$$\pi_{T_n}(f) = \frac{\mathrm{d}\pi_{T_n^-}(\lambda\phi f)}{\mathrm{d}\pi_{T_n^-}(\lambda\phi)}(Z_n) + \frac{\mathrm{d}\pi_{T_n^-}(\overline{\mathcal{L}}f)}{\mathrm{d}\pi_{T_n^-}(\lambda\phi)}(Z_n), \tag{31}$$

which implies that $\pi_{T_n}(f)$ is completely determined by the observed data (T_n, Z_n) and the knowledge of $\pi_t(f)$ in the time interval $[T_{n-1}, T_n)$, since $\pi_{T_n^-}(f) = \lim_{t \to T_n^-} \pi_t(f)$.

Note that the Kushner-Stratonovich equation is an infinite-dimensional nonlinear stochastic differential equation. Often, it is possible to characterize the filter in terms of a simpler equation, known as the Zakai equation which provides the dynamics of the unnormalized version of the filter. Although the Zakai equation is still infinite-dimensional, it has the advantage to be linear.

The idea for getting the dynamics of the unnormalized filter consists of performing an equivalent change of probability measure defined by

$$\left.\frac{\mathrm{d}\mathbf{P}_0}{\mathrm{d}\mathbf{P}}\right|_{\mathcal{F}_t} = Z_t, \quad t \in [0, T] \tag{32}$$

for a suitable strictly positive $(\mathbb{F}, \mathbf{P})$-martingale Z, in such a way that the so-called unnormalized filter p is the $\mathcal{M}(\mathbb{R})$-valued process defined by

$$p_t(f) := \mathbb{E}^0\left[Z_t^{-1} f(t, X_t) | \mathcal{F}_t^Y\right], \quad t \in [0, T], \tag{33}$$

Remark 9. *By the Kallianpur-Striebel formula, we get that*

$$\pi_t(f) = \frac{\mathbb{E}^0\left[f(t, X_t) Z_t^{-1} | \mathcal{F}_t^Y\right]}{\mathbb{E}^0\left[Z_t^{-1} | \mathcal{F}_t^Y\right]} = \frac{p_t(f)}{p_t(1)}, \quad t \in [0, T], \tag{34}$$

where $p_t(1) := \mathbb{E}^0\left[Z_t^{-1} | \mathcal{F}_t^Y\right]$. *This provides the relation between the filter and its unnormalized version*.

In order to compute the Zakai equation, we make the following assumption.

Assumption 10. *Suppose that there exists a transition function* $\eta^0(t, y, \mathrm{d}z)$ *such that the* $(\mathbb{F}^Y, \mathbf{P})$*-predictable measure* $\eta^0(t, Y_{t-}, \mathrm{d}z)$ *is equivalent to* $\lambda(t, X_{t-}, Y_{t-})\phi(t, X_{t-}, Y_{t-}, \mathrm{d}z)$ *and*

$$\mathbb{E}\left[\int_0^T \eta^0(t, Y_{t^-}, \mathbb{R})\mathrm{d}t\right] < \infty. \tag{35}$$

Remark 11. *In Ref. [18], a weaker assumption is considered. That condition allows to introduce an equivalent probability measure on* $(\Omega, \mathcal{F}_T^Y)$ *which is not necessarily the restriction on* $\mathcal{F}_T^Y$ *of an equivalent probability measure on* $(\Omega, \mathcal{F}_T)$.

Remark 12. *In the context of Example 8, Assumption 10 is satisfied if, for instance,* $\lambda^i(t, X_{t-}, Y_{t-}) > 0$ **P***-a.s. for every* $t \in [0, T]$.

Assumption 10 equivalently means that there exists an $(\mathbb{F}^Y, \mathbf{P})$-predictable process $\Psi(t, X_{t-}, Y_{t-}, z)$ such that

$$\lambda(t, X_{t-}, Y_{t-})\phi(t, X_{t-}, Y_{t-}, \mathrm{d}z)\mathrm{d}t = (1 + \Psi(t, X_{t-}, Y_{t-}, z))\eta^0(t, Y_{t-}, \mathrm{d}z)\mathrm{d}t \tag{36}$$

and $1 + \Psi(t, X_{t^-}, Y_{t^-}, z) > 0$ **P**-a.s. for every $t \in [0, T],\ z \in \mathbb{R}$. Setting

$$U(t, z) := \frac{1}{1 + \Psi(t, X_{t^-}, Y_{t^-}, z)} - 1, \tag{37}$$

we also assume that the following integrability condition holds:

$$\mathbb{E}\left[\exp\left\{\frac{1}{2}\int_0^T \left(\frac{b_1(s, X_s, Y_s)}{\sigma_1(s, Y_s)}\right)^2 \mathrm{d}s + \int_0^T \int_{\mathbb{R}} U^2(s, z)\lambda(s, X_{s^-}, Y_{s^-})\phi(s, X_{s^-}, Y_{s^-}, \mathrm{d}z)\mathrm{d}s\right\}\right] < \infty. \tag{38}$$

The subsequent proposition provides a useful version of the Girsanov Theorem that fits to our setting.

Proposition 13. *Let Assumptions 4 and 10, and condition (38) hold and define the process* $Z_t := \mathcal{E}\left(-\int_0^t \frac{b_1(s, X_s, Y_s)}{\sigma_1(s, Y_s)}\mathrm{d}W_s^1 + \int_0^t \int_{\mathbb{R}} U(s, z)\Big(m(\mathrm{d}s, \mathrm{d}z) - \lambda(s, X_{s^-}, Y_{s^-})\phi(s, X_{s^-}, Y_{s^-}, \mathrm{d}z)\mathrm{d}s\Big)\right)$, *for every* $t \in [0, T]$, *where* $\mathcal{E}(M)$ *denotes the Doléans-Dade exponential of a martingale M. Then, Z is a strictly positive* $(\mathbb{F}, \mathbf{P})$ *-martingale. Let* $\mathbf{P}^0$ *be the probability measure equivalent to* **P** *given by*

$$\left.\frac{\mathrm{d}\mathbf{P}^0}{\mathrm{d}\mathbf{P}}\right|_{\mathcal{F}_t} = Z_t, \quad t \in [0, T]. \tag{39}$$

Then, the process

$$\widetilde{W}_t^1 := W_t^1 + \int_0^t \frac{b_1(s, X_s, Y_s)}{\sigma_1(s, Y_s)}\mathrm{d}s, \quad t \in [0, T] \tag{40}$$

is an $(\mathbb{F}, \mathbf{P}^0)$*-Brownian motion, and the* $(\mathbb{F}, \mathbf{P}^0)$*-predictable projection of the integer-valued random measure* $m(\mathrm{d}t, \mathrm{d}z)$ *is given by* $\eta^0(t, Y_{t^-}, \mathrm{d}z)\mathrm{d}t$.

Proof. [32, Theorem 9] ensures that Z is a martingale under Assumptions 10, 4 and integrability condition Eq. (38). Then the proof follows by Ref. [31, Chapter III, Theorem 3.24].

Note that, by Eq. (16), we get that the process $\widetilde{W^1}$ can also be written as

$$\widetilde{W}_t^1 = I_t + \int_0^t \pi_s\left(\frac{b_1}{\sigma_1}\right)\mathrm{d}s, \quad t \in [0, T] \tag{41}$$

which implies that $\widetilde{W}^1$ is also an $(\mathbb{F}^Y, \mathbf{P}^0)$-Brownian motion. Moreover, since $\eta^0(t, Y_{t^-}, \mathrm{d}z)$ is $\mathbb{F}^Y$ predictable, it provides the $(\mathbb{F}^Y, \mathbf{P}^0)$-predictable projection of the measure $m(\mathrm{d}t, \mathrm{d}z)$ and the observation process Y satisfies $\mathrm{d}Y_t = \sigma_1(t, Y_t)\mathrm{d}\tilde{W}_t^1 + \int_{\mathbb{R}} z m(\mathrm{d}t, \mathrm{d}z)$. In particular, $\eta_t^0(\mathbb{R}) := \eta^0(t, Y_{t^-}, \mathbb{R})$ is the $(\mathbb{F}^Y, \mathbf{P}^0)$-intensity of the point process which counts the total jumps of Y until time t.

Theorem 14 (The Zakai equation). *Under Assumptions 4 and 10 and condition* (38), *let* $\mathbf{P}^0$ *be the probability measure defined in Proposition 13. For every* $f \in C_b^{1,2}([0, T] \times \mathbb{R})$, *the unnormalized filter defined in Eq*. (33) *satisfies the equation*

$$\begin{aligned}\mathrm{d}p_t(f) =& \{p_t(\mathcal{L}_0^X f) - p_t(\lambda f) + \eta_t^0(\mathbb{R})p_t(f)\}\mathrm{d}t + \left\{\frac{p_t(b_1 f)}{\sigma_1(t, Y_t)} + \rho\, p_t\left(\sigma_0 \frac{\partial f}{\partial x}\right)\right\}\mathrm{d}\widetilde{W}_t^1 \\ &+ \int_{\mathbb{R}}\left\{p_{t^-}(f\Psi)(z) + \frac{\mathrm{d}p_{t^-}(\overline{\mathcal{L}}f)}{\mathrm{d}\eta_t^0}(z)\right\} m(\mathrm{d}t, \mathrm{d}z).\end{aligned} \tag{42}$$

See Ref. [18, Theorem 3.6] for the proof.

3.1. Uniqueness of the filtering equations

In this section, we show pathwise uniqueness for the solution of the Kushner-Stratonovich and the Zakai equations. The first result provides the equivalence of uniqueness of the solutions to the filtering Eqs. (21) and (42).

Theorem 15. *Let Assumptions 4 and 10 and condition* (38) *hold*.

i. *Assume strong uniqueness for the solution to the Zakai equation, let* μ *be a* $\mathcal{P}(\mathbb{R})$*-valued process which is a strong solution of the Kushner-Stratonovich equation. Then* $\mu_t = \pi_t$ $\mathbf{P} - a.s.$ *for all* $t \in [0, T]$.

ii. *Conversely, suppose that pathwise uniqueness for the solution of the Kushner-Stratonovich equation holds and let* ξ *be an* $\mathcal{M}(\mathbb{R})$*-valued process which is a strong solution of the Zakai equation. Then* $\xi_t = p_t$ $\mathbf{P} - a.s.$ *for all* $t \in [0, T]$.

Proof. The proof follows by Ref. [18, Theorems 4.5 and 4.6]. Here, note that Assumption 10 implies that the measures $\mu_{t^-}(\lambda\phi(\mathrm{d}z))$ and $\pi_{t^-}(\lambda\phi(\mathrm{d}z))$ are equivalent.

Finally, strong uniqueness for the solution of both filtering equations is established in the subsequent theorems.

Theorem 16. *Let* (X, Y) *be the partially observed system defined in Eq*. (3), *and assume in addition to Assumptions 4 and 10 and condition* (15) *that*

$$\sup_{t, x, y}\int_Z \{|K_0(t, x; \zeta)| + |K_1(t, x, y; \zeta)|\}\nu(d\zeta) < \infty. \tag{43}$$

*Let μ be a strong solution of the Kushner-Stratonovich equation. Then $\mu_t = \pi_t$ **P**-a.s. for every $t \in [0, T]$.*

Proof. See Ref. [17, Theorem 3.3].

Theorem 17. *Let (X, Y) be the partially observed system in Eq.* (3). *Under Assumptions 4 and 10 and conditions* (38) *and* (43), *let ξ be a strong solution to the Zakai equation, then $\xi_t = p_t$ **P**-a.s. for every $t \in [0, T]$.*

Proof. The proof follows by Ref. [18, Theorem 4.7], after noticing that under Assumption 10 the measures $\xi_{t-}(\lambda\phi(\mathrm{d}z))$ and $p_{t-}(\lambda\phi(\mathrm{d}z))$ are equivalent.

4. A financial application to risk minimization

In the current section, we focus on a financial application. We consider a simple financial market where agents may invest in a risky asset whose price is described by the process Y given in Eq. (3) and a riskless asset with price process B. Without loss of generality, we assume that $B_t = 1$ for every $t \in [0, T]$. We also assume throughout the section the following dynamics for the process Y:

$$\mathrm{d}Y_t = Y_t\Big(\sigma(t, Y_t)\mathrm{d}W_t^1 + \int_Z K(t, X_{t-}, Y_{t-}; \zeta)\Big(N(\mathrm{d}t, \mathrm{d}\zeta) - \nu(\mathrm{d}\zeta)\mathrm{d}t\Big)\Big), \quad Y_0 = y_0 \in \mathbb{R}^+ \tag{44}$$

for some functions $\sigma(t, y)$ and $K(t, x, y; \zeta)$ such that $\sigma(t, y) > 0$ and $K(t, x, y; \zeta) > -1$.

This choice for the dynamics of Y has a double advantage. On one side assuming a geometric form, together with the condition that $K(t, x, y; \zeta) > -1$ guarantees nonnegativity which is desirable when talking about prices. On the other hand, we are modeling Y directly under a martingale measure, and by Assumption 18, it turns out to be a square integrable $(\mathbb{F}, \mathbf{P})$-martingale.

Considering Eq. (44) corresponds to take in system (3)

$$\begin{aligned} b_1(t, x, y) &= -y\int_Z K(t, x, y; \zeta)\nu(\mathrm{d}\zeta) \\ \sigma_1(t, y) = y\sigma(t, y), &\quad K_1(t, x, y; \zeta) = yK(t, x, y; \zeta). \end{aligned} \tag{45}$$

In addition, we me make the following assumption.

Assumption 18.

$$0 < c_1 < \sigma(t, y) < c_2, \quad |K(t, x, y; \zeta)| < c_3, \quad \nu(D_t) < c_4, \tag{46}$$

for every $(t, x, y) \in [0, T] \times \mathbb{R} \times \mathbb{R}^+$, $\zeta \in Z$ and for some positive constants c_1, c_2, c_3, c_4.

Remark 19. *In the sequel, it might be useful to specify the dynamics of* Y *also in terms of the jump measure $m(\mathrm{d}t, \mathrm{d}z)$. Recalling Eqs. (6) and (14), we have*

$$\mathrm{d}Y_t = Y_t\sigma(t, Y_t)\mathrm{d}W_t^1 + \int_{\mathbb{R}} z\Big(m(\mathrm{d}t, \mathrm{d}z) - \lambda(t, X_{t^-}, Y_{t^-})\phi(t, X_{t^-}, Y_{t^-}, \mathrm{d}z)\mathrm{d}t\Big). \tag{47}$$

http://dx.doi.org/10.5772/intechopen.70060

The stochastic factor X which affects intensity and jump size distribution of Y may represent the state of the economy and is not directly observable by market agents. This is a typical situation arising in real financial markets.

We model by $\mathbb{F}^Y$ the available information to investors. Since Y is $\mathbb{F}^Y$ adapted, it is in particular an $(\mathbb{F}^Y, \mathbf{P})$-martingale with the following decomposition:

$$Y_t = y_0 + \int_0^t Y_s \sigma(s, Y_s) \mathrm{d}I_s + \int_0^t \int_{\mathbb{R}} z\Big(m(\mathrm{d}s, \mathrm{d}z) - \pi_{s^-}(\lambda\phi(\mathrm{d}z))\mathrm{d}s\Big), \quad t \in [0, T]. \tag{48}$$

By Eqs. (14) and (45), in this setting the first component of the innovation process I defined in Eq. (16) is given by $I_t = W_t^1 + \int_0^t \frac{1}{Y_s \sigma(s, Y_s)} \int_{\mathbb{R}} z\Big(\lambda(s, X_s, Y_s)\phi(s, X_s, Y_s, \mathrm{d}z) - \pi_s(\lambda\phi(\mathrm{d}z))\Big)\mathrm{d}s$.

Suppose that we are given a European-type contingent claim whose final payoff is a square integrable $\mathcal{F}_T^Y$-measurable random variable ξ, that is, $\xi \in L^2(\mathcal{F}_T^Y)$ where

$$L^2(\mathcal{F}_T^Y) := \{\text{random variables } \Gamma \in \mathcal{F}_T^Y : \mathbb{E}\big[\Gamma^2\big] < \infty\}. \tag{49}$$

The objective of the agent is to find the optimal hedging strategy for this derivative. Since the number of random sources exceeds the number of tradeable risky assets, the market is incomplete. It is well known that in this setting, perfect replication by self-financing strategies is not feasible. Then, we suppose that the investor intends to pursue the risk-minimization approach. Risk minimization is a quadratic hedging method that allows determining a dynamic investment strategy that replicates perfectly the claim with minimal cost. Let us properly introduce the objects of interest. We start with the following notation. For any pair of $\mathbb{F}$-adapted (respectively, $\mathbb{F}^Y$-adapted) processes Ψ^1, Ψ^2 we refer to $\langle \Psi^1, \Psi^2 \rangle^{\mathbb{F}}$ for the predictable covariation computed with respect to filtration $\mathbb{F}$ (respectively, $\langle \Psi^1, \Psi^2 \rangle^{\mathbb{F}^Y}$ for the predictable covariation computed with respect to filtration $\mathbb{F}^Y$). Note that

$$\begin{aligned} \langle Y \rangle_t^{\mathbb{F}} &= \int_0^t Y_s^2 \Big(\sigma^2(s, Y_{s^-}) + \int_Z K^2(s, X_{s^-}, Y_{s^-}; \zeta)\nu(\mathrm{d}\zeta)\Big)\mathrm{d}s \\ &= \int_0^t \Big(Y_s^2 \sigma^2(s, Y_{s^-}) + \int_{\mathbb{R}} z^2 \lambda(s, X_{s^-}, Y_{s^-})\phi(s, X_{s^-}, Y_{s^-}, \mathrm{d}z)\Big)\mathrm{d}s, \quad t \in [0, T], \end{aligned} \tag{50}$$

and since Y is also $\mathbb{F}^Y$ adapted, we also have

$$\langle Y \rangle_t^{\mathbb{F}^Y} = \int_0^t \Big(Y_s^2 \sigma^2(s, Y_{s^-}) + \int_{\mathbb{R}} z^2 \pi_{s^-}(\lambda\phi(\mathrm{d}z))\Big)\mathrm{d}s, \quad t \in [0, T]. \tag{51}$$

We stress that, due to the presence of a jump component, the predictable quadratic variations of Y with respect to filtrations $\mathbb{F}$ and $\mathbb{F}^Y$ are different.

Now we introduce a technical definition of two spaces, $\Theta(\mathbb{F})$ and $\Theta(\mathbb{F}^Y)$

Definition 20. *The space $\Theta(\mathbb{F}^Y)$ (respectively, $\Theta(\mathbb{F})$) is the space of all $\mathbb{F}^Y$-predictable (respectively, $\mathbb{F}$-predictable) processes θ such that*

$$\mathbb{E}\left[\int_0^T \theta_u^2 \mathrm{d}\langle Y\rangle_u^{\mathbb{F}^Y}\right] < \infty \quad \left(\textit{respectively} \quad \mathbb{E}\left[\int_0^T \theta_u^2 \mathrm{d}\langle Y\rangle_u^{\mathbb{F}}\right] < \infty\right). \tag{52}$$

We observe that for every $\theta \in \Theta(\mathbb{F}^Y)$, thanks to $\mathbb{F}^Y$-predictability, we have

$$\mathbb{E}\left[\int_0^T \theta_u^2 \mathrm{d}\langle Y\rangle_u^{\mathbb{F}}\right] = \mathbb{E}\left[\int_0^T \theta_u^2 \mathrm{d}\langle Y\rangle_u^{\mathbb{F}^Y}\right] < \infty, \tag{53}$$

which implies that $\Theta(\mathbb{F}^Y) \subseteq \Theta(\mathbb{F})$.

Since we have two different levels of information represented by the filtrations $\mathbb{F}$ and $\mathbb{F}^Y$, we may define two classes of admissible strategies.

Definition 21. *An $\mathbb{F}^Y$-strategy (respectively, $\mathbb{F}$-strategy) is a pair $\psi = (\theta, \eta)$ of stochastic processes, where θ represents the amount invested in the risky asset and η is the amount invested in the riskless asset, such that $\theta \in \Theta(\mathbb{F}^Y)$ (respectively, $\theta \in \Theta(\mathbb{F})$) and η is $\mathbb{F}^Y$-adapted (respectively, $\mathbb{F}$-adapted).*

This definition reflects the fact that investor's choices should be adapted to her/his knowledge of the market. The value of a strategy $\psi = (\theta, \eta)$ is given by

$$V_t(\psi) = \theta_t Y_t + \eta_t, \quad t \in [0, T], \tag{54}$$

and its cost is described by the process

$$C_t(\psi) = V_t(\psi) - \int_0^t \theta_u \mathrm{d}Y_u, \quad t \in [0, T]. \tag{55}$$

In other terms, the cost of a strategy is the difference between the value process and the gain process. For a self-financing strategy, the value and the gain processes coincide, up to the initial wealth V_0, and therefore the cost is constant and equal to $C_t = V_0$, for every $t \in [0, T]$. We continue by defining the risk process, in the partial information setting.

Definition 22. *Given an $\mathbb{F}^Y$-strategy (respectively, an $\mathbb{F}$-strategy) $\psi = (\theta, \eta)$, we denote by $R^{\mathcal{F}^Y}(\psi)$ (respectively, $R^{\mathcal{F}}(\psi)$) the associated risk process defined as*

$$R_t^{\mathcal{F}^Y}(\psi) := \mathbb{E}\left[\Big(C_T(\psi) - C_t(\psi)\Big)^2 | \mathcal{F}_t^Y\right], \quad \left(\textit{respectively } R_t^{\mathcal{F}}(\psi) := \mathbb{E}\left[\Big(C_T(\psi) - C_t(\psi)\Big)^2 | \mathcal{F}_t\right]\right), \tag{56}$$

for every $t \in [0, T]$.

Then, we have the following definition of risk-minimizing strategy under partial information.

http://dx.doi.org/10.5772/intechopen.70060

Definition 23. *An $\mathbb{F}^Y$-strategy ψ is risk minimizing if*

i. $V_T(\psi) = \xi$,

ii. *for any other $\mathbb{F}^Y$ -strategy $\tilde{\psi}$ we have $R_t^{\mathcal{F}^Y}(\psi) \leq R_t^{\mathcal{F}^Y}(\tilde{\psi})$, for every $t \in [0, T]$.*

The corresponding definitions of risk process and risk-minimizing strategy under full information can be obtained replacing $\mathbb{F}^Y$ and $R_t^{\mathcal{F}^Y}$ with $\mathbb{F}$ and $R_t^{\mathcal{F}}$ in Definition 23. To differentiate, when it is necessary, we use the terms $\mathbb{F}^Y$-risk-minimizing strategy or $\mathbb{F}$-risk-minimizing strategy. The criterion (ii) in Definition 23 can be also written as

$$\min_{\psi \in \Theta(\mathbb{F}^Y)} \mathbb{E}\Big[(C_T(\psi) - C_t(\psi))^2\Big], \quad t \in [0, T], \tag{57}$$

which intuitively means that a strategy is risk minimizing if it minimizes the variance of the cost. This equivalent definition allows to obtain a nice property of risk-minimizing strategies which turn out to be *self-financing on average*, that is, the cost process C is a martingale and therefore has constant expectation (see, e.g., Ref. [27, Lemma 2] or [28, Lemma 2.3]).

In the sequel, we aim to characterize the optimal hedging strategy for the contingent claim ξ under full and partial information, that is, the $\mathbb{F}$- and the $\mathbb{F}^Y$-risk-minimizing strategies. To this, we introduce two orthogonal decompositions known as the Galtchouk-Kunita-Watanabe decompositions under full and partial information (see, e.g., [30]). To understand better the relevance of these decompositions, we assume for a moment completeness of the market and full information. Then, it is well known that for every European-type contingent claim with final payoff ξ, there exists a self-financing strategy $\psi = (\theta, \eta)$ such that

$$\xi = V_0 + \int_0^T \theta_u \mathrm{d}Y_u, \quad \mathbf{P} - \text{a.s.} \tag{58}$$

that is, a replicating portfolio is uniquely determined by the initial wealth and the investment in the risky asset. When the market is incomplete, decomposition Eq. (58) does not hold in general. Intuitively, this implies that we might expect additional terms in Eq. (58), and according to the risk-minimization criterion, this additional terms need to be such that the final cost does not deviate too much from the average cost, in the quadratic sense. Specifically, we have the following decomposition of the random variable ξ:

$$\xi = V_0 + \int_0^T \theta_u \mathrm{d}Y_u + G_T, \quad \mathbf{P} - \text{a.s.} \tag{59}$$

where G_T is the value at time T of a suitable process G. The minimality criterion requires that G is a martingale orthogonal to Y. We refer the reader to Ref. [28] for a detailed survey. Under suitable hypothesis, the above decomposition takes the name of Galtchouk-Kunita-Watanabe decomposition.

Now we wish to be more formal, and we introduce the following definitions:

Consider a random variable $\xi \in L^2(\mathcal{F}_T^Y)$. Since $\mathcal{F}_T^Y \subseteq \mathcal{F}_T$, we can define the following decompositions for ξ.

Definition 24. *a. The Galtchouk-Kunita-Watanabe decomposition of* $\xi \in L^2(\mathcal{F}_T^Y)$ *with respect to* Y *and* $\mathbb{F}$ *is given by*

$$\xi = U_0^{\mathcal{F}} + \int_0^T \theta_u^{\mathcal{F}} \mathrm{d}Y_u + G_T^{\mathcal{F}} \quad \mathbf{P} - a.s., \tag{60}$$

where $U_0^{\mathcal{F}} \in L^2(\mathcal{F}_0)$, $\theta^{\mathcal{F}} \in \Theta(\mathbb{F})$ *and* $G^{\mathcal{F}}$ *is a square integrable* $(\mathbb{F}, \mathbf{P})$*-martingale, with* $G_0^{\mathcal{F}} = 0$, *orthogonal to Y, that is,* $\langle G^{\mathcal{F}}, Y \rangle_t^{\mathbb{F}} = 0$ *for every* $t \in [0, T]$.

b. The Galtchouk-Kunita-Watanabe decomposition of $\xi \in L^2(\mathcal{F}_T^Y)$ *with respect to* Y *and* $\mathbb{F}^Y$ *is given by*

$$\xi = U_0^{\mathcal{F}^Y} + \int_0^T \theta_u^{\mathcal{F}^Y} \mathrm{d}Y_u + G_T^{\mathcal{F}^Y} \quad \mathbf{P} - a.s., \tag{61}$$

where $U_0^{\mathcal{F}^Y} \in L^2(\mathcal{F}_0^Y)$, $\theta^{\mathcal{F}^Y} \in \Theta(\mathbb{F}^Y)$ *and* $G^{\mathcal{F}^Y}$ *is a square integrable* $(\mathbb{F}^Y, \mathbf{P})$ *-martingale, With* $G_0^{\mathcal{F}^Y} = 0$, *strongly orthogonal to Y, that is,* $\langle G^{\mathcal{F}}, Y \rangle_t^{\mathbb{F}^Y} = 0$ *for every* $t \in [0, T]$.

In the sequel, we refer to Eqs. (60) and (61) as the Galtchouk-Kunita-Watanabe decompositions under full information and under partial information, respectively. Since Y is a square integrable martingale with respect to both filtrations $\mathbb{F}$ and $\mathbb{F}^Y$, decompositions Eqs. (60) and (61) exist.

Next proposition provides a relation between the integrands $\theta^{\mathcal{F}}$ and $\theta^{\mathcal{F}^Y}$ of decompositions Eqs. (60) and (61) in terms of predictable projections. For any $(\mathbb{F}, \mathbf{P})$-predictable process A of finite variation, we denote by $A^{p,\mathbb{F}^Y}$ its $(\mathbb{F}^Y, \mathbf{P})$-dual-predictable projection.[1]

Proposition 25. *The integrands in decompositions Eqs. (60) and (61) satisfy the following relation:*

$$\theta_t^{\mathcal{F}^Y} = \frac{\mathrm{d}\left(\int_0^t \theta_u^{\mathcal{F}} \mathrm{d}\langle Y \rangle_u^{\mathbb{F}}\right)^{p,\mathbb{F}^Y}}{\mathrm{d}\langle Y \rangle_t^{p,\mathbb{F}^Y}}, \quad t \in [0, T]. \tag{62}$$

Here, $\langle Y \rangle^{p,\mathbb{F}^Y}$ *denotes the* $(\mathbb{F}^Y, \mathbf{P})$*-dual-predictable projection of* $\langle Y \rangle^{\mathbb{F}}$ *and it is given by*

[1]We call $(\mathbb{F}^Y, \mathbf{P})$- *dual predictable projection* of a process A the $\mathbb{F}^Y$-predictable finite variation process $A^{p,\mathbb{F}^Y}$ such that for any $\mathbb{F}^Y$-predictable-bounded process ϕ we have

$$\mathbb{E}\left[\int_0^T \phi_s \mathrm{d}A_s\right] = \mathbb{E}\left[\int_0^T \phi_s \mathrm{d}A_s^{p,\mathbb{F}^Y}\right]$$

$$\langle Y\rangle_t^{p,\mathbb{F}^Y} = \langle Y\rangle_t^{\mathbb{F}^Y} = \int_0^t Y_s^2\sigma^2(s, Y_{s^-})\mathrm{d}s + \int_0^t\int_{\mathbb{R}} z^2\pi_{s^-}(\lambda\phi(\mathrm{d}z))\mathrm{d}s, \quad t\in[0,T]. \tag{63}$$

Proof. First note that the $(\mathbb{F}^Y, \mathbf{P})$-dual-predictable projection of the process $\langle Y\rangle^{\mathbb{F}}$ coincides with the predictable quadratic variation of the process Y itself, computed with respect to its internal filtration, given in Eq. (51), since for any $(\mathbb{F}^Y, \mathbf{P})$-predictable-(bounded) process ϕ, we have that $\mathbb{E}\left[\int_0^T \phi_t \mathrm{d}\langle Y\rangle_t^{\mathbb{F}}\right] = \mathbb{E}\left[\int_0^T \phi_t \mathrm{d}\langle Y\rangle_t^{\mathbb{F}^Y}\right]$. This proves Eq. (63).

Let

$$\theta_t := \frac{\mathrm{d}\left(\int_0^t \theta_u^{\mathcal{F}}\mathrm{d}\langle Y\rangle_u^{\mathbb{F}}\right)^{p,\mathbb{F}^Y}}{\mathrm{d}\langle Y\rangle_t^{p,\mathbb{F}^Y}}, \quad t\in[0,T]. \tag{64}$$

By the Galtchouk-Kunita-Watanabe decomposition Eq. (60), we can write

$$\xi = U_0^{\mathcal{F}} + \int_0^T \theta_u \mathrm{d}Y_u + G_T^{\mathcal{F}} + \widetilde{G}_T \quad \mathbf{P}-\text{ a.s.}, \tag{65}$$

where $\widetilde{G}_t := \int_0^t (\theta_u^{\mathcal{F}} - \theta_u)\mathrm{d}Y_u$, for every $t\in[0,T]$. We observe that for every $\mathbb{F}^Y$-predictable process ϕ the following holds:

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T \phi_u\theta_u \mathrm{d}\langle Y\rangle_u^{\mathbb{F}}\right] = \mathbb{E}\left[\int_0^T \phi_u\theta_u \mathrm{d}\langle Y\rangle_u^{\mathbb{F}^Y}\right] \\
&= \mathbb{E}\left[\int_0^T \phi_u(\theta_u^{\mathcal{F}}\mathrm{d}\langle Y\rangle_u^{\mathbb{F}})^{p,\mathbb{F}^Y}\right] = \mathbb{E}\left[\int_0^T \phi_u\theta_u^{\mathcal{F}}\mathrm{d}\langle Y\rangle_u^{\mathbb{F}}\right].
\end{aligned} \tag{66}$$

By choosing $\phi = \theta$ and applying the Cauchy-Schwarz inequality, we obtain

$$\mathbb{E}\left[\int_0^T (\theta_u)^2 \mathrm{d}\langle Y\rangle_u^{\mathbb{F}^Y}\right] \le \mathbb{E}\left[\int_0^T (\theta_u^{\mathcal{F}})^2 \mathrm{d}\langle Y\rangle_u^{\mathbb{F}}\right] < \infty. \tag{67}$$

This implies that $\theta\in\Theta(\mathbb{F}^Y)\subseteq\Theta(\mathbb{F})$ and that $\widetilde{G}$ is an $(\mathbb{F}, \mathbf{P})$-martingale. Taking the conditional expectation with respect to $\mathcal{F}_T^Y$ in Eq. (65) leads to

$$\xi = \mathbb{E}\left[U_0^{\mathcal{F}}|\mathcal{F}_T^Y\right] + \int_0^T \theta_u \mathrm{d}Y_u + G_T^{\mathcal{F}} + \widetilde{G}_T = \mathbb{E}\left[U_0^{\mathcal{F}}|\mathcal{F}_0^Y\right] + \int_0^T \theta_u \mathrm{d}Y_u + \widehat{G}_T^{\mathcal{F}^Y} \quad \mathbf{P}-\text{a.s.} \tag{68}$$

where

$$\widehat{G}_t^{\mathcal{F}^Y} := \mathbb{E}\left[U_0^{\mathcal{F}}|\mathcal{F}_t^Y\right] - \mathbb{E}\left[U_0^{\mathcal{F}}|\mathcal{F}_0^Y\right] + \mathbb{E}\left[G_T^{\mathcal{F}}|\mathcal{F}_t^Y\right] + \mathbb{E}\left[\widetilde{G}_T|\mathcal{F}_t^Y\right], \quad t\in[0,T], \tag{69}$$

which provides the Galtchouk-Kunita-Watanabe decomposition Eq. (61) if we can show that the $(\mathbb{F}^Y, \mathbf{P})$-martingale $\widehat{G}^{\mathcal{F}^Y}$ is strongly orthogonal to Y, that is, if for any $(\mathbb{F}^Y, \mathbf{P})$-predictable-(bounded) process ϕ the following holds:

$$\mathbb{E}\left[\widehat{G}_T^{\mathcal{F}^Y}\int_0^T \phi_u \mathrm{d}Y_u\right] = 0. \tag{70}$$

Note that orthogonality of the term $\mathbb{E}\left[U_0^{\mathcal{F}}|\mathcal{F}_t^Y\right] - \mathbb{E}\left[U_0^{\mathcal{F}}|\mathcal{F}_0^Y\right] + \mathbb{E}\left[G_T^{\mathcal{F}}|\mathcal{F}_t^Y\right]$ follows by the orthogonality of $G^{\mathcal{F}}$ and Y. Moreover, we have

$$\mathbb{E}\left[\mathbb{E}\left[\widetilde{G}_T|\mathcal{F}_T^Y\right]\int_0^T \phi_u \mathrm{d}Y_u\right] = \mathbb{E}\left[\widetilde{G}_T\int_0^T \phi_u \mathrm{d}Y_u\right] = \mathbb{E}\left[\int_0^T \phi_u(\theta_u^{\mathcal{F}} - \theta_u)\mathrm{d}\langle Y\rangle_u^{\mathbb{F}}\right], \tag{71}$$

and by Eq. (64)

$$\begin{aligned}\mathbb{E}\left[\int_0^T \phi_u\theta_u \mathrm{d}\langle Y\rangle_u^{\mathbb{F}}\right] &= \mathbb{E}\left[\int_0^T \phi_u\theta_u \mathrm{d}\langle Y\rangle_u^{\mathbb{F}^Y}\right] \\ &= \mathbb{E}\left[\int_0^T \phi_u \mathrm{d}(\int_0^u \theta_r^{\mathcal{F}}\mathrm{d}\langle Y\rangle_r)^{p,\mathbb{F}^Y}\right] = \mathbb{E}\left[\int_0^T \phi_u\theta_u^{\mathcal{F}}\mathrm{d}\langle Y\rangle_u^{\mathbb{F}}\right],\end{aligned} \tag{72}$$

which proves strong orthogonality.

Theorem 26 shows the relation between the Galtchouk-Kunita-Watanabe decompositions and the optimal strategies under full and partial information.

Theorem 26. *i. Every contingent claim $\xi \in L^2(\mathcal{F}_T^Y, \mathbf{P})$ admits a unique $\mathbb{F}$-risk-minimizing strategy $\psi^{*,\mathcal{F}} = (\theta^{*,\mathcal{F}}, \eta^{*,\mathcal{F}})$, explicitly given by*

$$\theta^{*,\mathcal{F}} = \theta^{\mathcal{F}}, \quad \eta^{*,\mathcal{F}} = V(\psi^{*,\mathcal{F}}) - \theta^{*,\mathcal{F}}Y, \tag{73}$$

where $V_t(\psi^{*,\mathcal{F}}) = \mathbb{E}[\xi|\mathcal{F}_t]$ for every $t \in [0,T]$, with minimal cost

$$C_t(\psi^{*,\mathcal{F}}) = U_0^{\mathcal{F}} + G_t^{\mathcal{F}}, \quad t \in [0,T]. \tag{74}$$

Here, $\theta^{\mathcal{F}}$, $U_0^{\mathcal{F}}$, and $G^{\mathcal{F}}$ are given in Definition 24 part a.

ii. Moreover, it also admits a unique $\mathbb{F}^Y$-risk-minimizing strategy $\psi^{*,\mathcal{F}} = (\theta^{*,\mathcal{F}}, \eta^{*,\mathcal{F}^Y})$, explicitly given by

$$\theta^{*,\mathcal{F}^Y} = \theta^{\mathcal{F}^Y}, \quad \eta^{*,\mathcal{F}^Y} = V(\psi^{*,\mathcal{F}^Y}) - \theta^{*,\mathcal{F}^Y}Y, \tag{75}$$

where $V_t(\psi^{*,\mathcal{F}^Y}) = \mathbb{E}\left[\xi|\mathcal{F}_t^Y\right]$ for every $t \in [0,T]$, with minimal cost

$$C_t(\psi^{*,\mathcal{F}^Y}) = U_0^{\mathcal{F}^Y} + G_t^{\mathcal{F}^Y}, \quad t \in [0,T], \tag{76}$$

and $\theta^{\mathcal{F}^Y}$, $U_0^{\mathcal{F}^Y}$ *and* $G^{\mathcal{F}^Y}$ *are given in Definition 24 part b*.

Proof. The proof of part *i*. is given, for example, in Ref. [28, Theorem 2.4]. For part *ii*., note that using the martingale representation of Y with respect to its inner filtration given in Eq. (48) and the fact that $\xi \in L^2(\mathcal{F}_T^Y)$, it is possible to reduce the partial information case to full information and apply again [28, Theorem 2.4]. □

Proposition 25 helps us in the computation of the optimal strategy under partial information. Indeed, it is sufficient to compute the corresponding strategy $\theta^{*,\mathcal{F}}$ under full information and the Radon-Nikodym derivative given in Eq. (62). To get more explicit representations, we assume that the payoff of the contingent claim has the form $\xi = H(T, Y_T)$, for some function $H : [0, T] \times \mathbb{R}^+ \to \mathbb{R}$. Let $\mathcal{L}^{X,Y}$ denote the Markov generator of the pair (X, Y), that is

$$\begin{aligned}\mathcal{L}^{X,Y} f(t,x,y) = &\frac{\partial f}{\partial t} + b_0(t,x)\frac{\partial f}{\partial x} + b_1(t,x,y)\frac{\partial f}{\partial y} + \frac{1}{2}\sigma_0^2(t,x)\frac{\partial^2 f}{\partial x^2} + \rho y \sigma_0(t,x)\sigma(t,y)\frac{\partial^2 f}{\partial x \partial y} \\ &+ \frac{1}{2}y^2\sigma^2(t,y)\frac{\partial^2 f}{\partial y^2} + \int_Z \Delta f(t,x,y;\zeta)\nu(\mathrm{d}\zeta)\end{aligned} \tag{77}$$

for every $f \in C_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R}^+)$, where

$$\Delta f(t,x,y;\zeta) := f(t, x + K_0(t,x;\zeta), y(1 + K(t,x,y;\zeta))) - f(t,x,y). \tag{78}$$

By the Markov property, we have that for any $t \in [0, T]$ there exists a measurable function $h(t, x, y)$ such that

$$h(t, X_t, Y_t) = \mathbb{E}[H(T, Y_T)|\mathcal{F}_t]. \tag{79}$$

If the function h is sufficiently regular, for instance $h \in C_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R}^+)$, we can apply Itô's formula and get that

$$h(t, X_t, Y_t) = h(0, X_0, Y_0) + \int_0^t \mathcal{L}^{X,Y} h(s, X_s, Y_s)ds + M_t^h \tag{80}$$

where M^h is the $(\mathbb{F}, \mathbf{P})$-martingale given by

$$\begin{aligned}\mathrm{d}M_t^h = &\int_0^t \frac{\partial h}{\partial x}(s, X_s, Y_s)\sigma_0(s, X_s)\mathrm{d}W_s^0 + \int_0^t \frac{\partial h}{\partial y}(s, X_s, Y_s)Y_s\sigma(s, Y_s)\mathrm{d}W_s^1 \\ &+ \int_0^t \int_Z \Delta h(s, X_{s^-}, Y_{s^-}; \zeta)\Big(N(\mathrm{d}s, \mathrm{d}\zeta) - \nu(\mathrm{d}\zeta)\mathrm{d}s\Big).\end{aligned} \tag{81}$$

By Eq. (79), the process $\{h(t, X_t, Y_t),\ t \in [0, T]\}$ is an $(\mathbb{F}, \mathbf{P})$-martingale. Then, the finite variation term vanishes, which means that the function h satisfies $\mathcal{L}^{X,Y} h(t, X_t, Y_t) = 0$, **P**-a.s. and for almost every $t \in [0, T]$. The next proposition provides the risk-minimizing strategy under partial information.

Proposition 27. *Assume $h \in C_b^{1,2,2}([0,T] \times \mathbb{R} \times \mathbb{R}^+)$. Then the first components $\theta^{*,\mathcal{F}}$ and $\theta^{*,\mathcal{F}^Y}$ of the risk-minimizing strategies under full and partial information are given by*

$$\theta_t^{*,\mathcal{F}} = \frac{g(t, X_{t-}, Y_{t-})}{Y_{t-}^2 \sigma^2(t, Y_{t-}) + \int_{\mathbb{R}} z^2 \lambda(t, X_{t-}, Y_{t-}) \phi(t, X_{t-}, Y_{t-}, \mathrm{d}z)}, \quad t \in [0,T] \tag{82}$$

$$\theta_t^{*,\mathcal{F}^Y} = \frac{\pi_{t-}(g)}{Y_{t-}^2 \sigma(t, Y_{t-}) + \int_{\mathbb{R}} z^2 \pi_{t-}(\lambda \phi(\mathrm{d}z))}, \quad t \in [0,T] \tag{83}$$

respectively, where the function $g(t, x, y)$ is

$$g(t,x,y) = \rho\, \sigma_0(t,x) y \sigma(t,y) \frac{\partial h}{\partial x} + y^2 \sigma^2(t,y) \frac{\partial h}{\partial y} + \int_Z y K(t,x,y;\zeta) \Delta h(t,x,y;\zeta) \nu(d\zeta). \tag{84}$$

Proof. Consider decomposition Eq. (60) for $\xi = H(T, Y_T)$. Then, conditioning on $\mathcal{F}_t$ we get

$$h(t, X_t, Y_t) = U_0 + \int_0^t \theta_s^{*,\mathcal{F}} \mathrm{d}Y_s + G_t^{\mathcal{F}}. \tag{85}$$

Taking the covariation with respect to Y and $\mathbb{F}$, we obtain

$$\langle h(\cdot, X, Y), Y \rangle_t^{\mathbb{F}} = \int_0^t \theta_s^{*,\mathcal{F}} \mathrm{d}\langle Y \rangle_s^{\mathbb{F}}. \tag{86}$$

On the other hand, $h(t, X_t, Y_t) = M_t^h$, then taking Eqs. (81) and (44) into account we get that

$$\langle h(\cdot, X, Y), Y \rangle_t^{\mathbb{F}} = \int_0^t g(s, X_s, Y_s) \mathrm{d}s, \tag{87}$$

where $g(t, x, y)$ is given in Eq. (84). Hence, by Eqs. (50) and (87), we may represent $\theta^{*,\mathcal{F}}$ as

$$\theta_t^{*,\mathcal{F}} = \frac{\mathrm{d}\langle h(\cdot, X, Y), Y \rangle_t^{\mathbb{F}}}{\mathrm{d}\langle Y \rangle_t^{\mathbb{F}}} = \frac{g(t, X_{t-}, Y_{t-})}{Y_{t-}^2 \sigma^2(t, Y_{t-}) + \int_{\mathbb{R}} z^2 \lambda(t, X_{t-}, Y_{t-}) \phi(t, X_{t-}, Y_{t-}, \mathrm{d}z)} \tag{88}$$

Note that by Eq. (51) and

$$\left(\int_0^t \theta_u^{*,\mathcal{F}} \mathrm{d}\langle Y \rangle_u^{\mathbb{F}} \right)^{p,\mathbb{F}^Y} = \left(\int_0^t g(s, X_s, Y_s) \mathrm{d}s \right)^{p,\mathbb{F}^Y} = \int_0^t \pi_s(g) \mathrm{d}s, \tag{89}$$

applying Eq. (62) we get representation Eq. (83).

Our ultimate objective in this section is to investigate on the relation between costs of the $\mathbb{F}$-optimal strategy and the $\mathbb{F}^Y$-optimal strategy, or equivalently the associated risk processes.

It clearly holds that $\theta^{*,\mathcal{F}^Y} \in \Theta(\mathbb{F})$, and then the $\mathbb{F}^Y$-risk-minimizing strategy is also an $\mathbb{F}$-strategy. Considering the corresponding risks, we have

$$\begin{aligned}\mathbb{E}\left[\left(C_T(\psi^{*,\mathcal{F}^Y}) - C_t(\psi^{*,\mathcal{F}^Y})\right)^2 | \mathcal{F}_t^Y\right] &= \mathbb{E}\left[\mathbb{E}\left[\left(C_T(\psi^{*,\mathcal{F}^Y}) - C_t(\psi^{*,\mathcal{F}^Y})\right)^2 | \mathcal{F}_t\right] | \mathcal{F}_t^Y\right] \\ \geq \mathbb{E}\left[\mathbb{E}\left[\left(C_T(\psi^{*,\mathcal{F}}) - C_t(\psi^{*,\mathcal{F}})\right)^2 | \mathcal{F}_t\right] | \mathcal{F}_t^Y\right] &= \mathbb{E}\left[\left(C_T(\psi^{*,\mathcal{F}}) - C_t(\psi^{*,\mathcal{F}})\right)^2 | \mathcal{F}_t^Y\right],\end{aligned} \tag{90}$$

and then $\mathbb{E}\left[R_t^{\mathcal{F}}(\psi^{*,\mathcal{F}})\right] \leq \mathbb{E}\left[R_t^{\mathcal{F}^Y}(\psi^{*,\mathcal{F}^Y})\right]$, for every $t \in [0, T]$. In the remaining part of the paper, we assume that $\mathcal{F}_0^Y = \mathcal{F}_0 = \{\Omega, \emptyset\}$, and we wish to measure the difference in the total risk taken by an informed investor, endowed with a filtration $\mathbb{F}$, and a partially informed investor, whose information is described by $\mathbb{F}^Y$. Precisely, we compute the difference $R_0^{\mathcal{F}^Y}(\psi^{*,\mathcal{F}^Y}) - R_0^{\mathcal{F}}(\psi^{*,\mathcal{F}})$. By decompositions Eqs. (60) and (61), we have that $C_T(\psi^{*,\mathcal{F}}) - C_0(\psi^{*,\mathcal{F}}) = G_T^{\mathcal{F}}$ and $C_T(\psi^{*,\mathcal{F}^Y}) - C_0(\psi^{*,\mathcal{F}^Y}) = G_T^{\mathcal{F}^Y}$ and also

$$G_T^{\mathcal{F}^Y} = U_0^{\mathcal{F}} - U_0^{\mathcal{F}^Y} + \int_0^T (\theta_r^{*,\mathcal{F}} - \theta_r^{*,\mathcal{F}^Y}) \mathrm{d}Y_r + G_T^{\mathcal{F}}, \tag{91}$$

since $\mathcal{F}_0^Y = \mathcal{F}_0 = \{\Omega, \emptyset\}$, $U_0^{\mathcal{F}} = U_0^{\mathcal{F}^Y}$. Then computing the square of $G_T^{\mathcal{F}^Y}$ and taking the expectation we get

$$\mathbb{E}\left[(G_T^{\mathcal{F}^Y})^2\right] = \mathbb{E}\left[(G_T^{\mathcal{F}})^2\right] + \mathbb{E}\left[\left(\int_0^T (\theta_r^{*,\mathcal{F}} - \theta_r^{*,\mathcal{F}^Y}) \mathrm{d}Y_r\right)^2\right] + 2\mathbb{E}\left[G_T^{\mathcal{F}} \int_0^T (\theta_r^{*,\mathcal{F}} - \theta_r^{*,\mathcal{F}^Y}) \mathrm{d}Y_r\right]. \tag{92}$$

It follows from Itô isometry and the fact that $G^{\mathcal{F}}$ is orthogonal to Y, that

$$\mathbb{E}\left[(G_T^{\mathcal{F}^Y})^2\right] = \mathbb{E}\left[(G_T^{\mathcal{F}})^2\right] + \mathbb{E}\left[\int_0^T (\theta_r^{*,\mathcal{F}} - \theta_r^{*,\mathcal{F}^Y})^2 \langle Y \rangle_r^{\mathbb{F}}\right]. \tag{93}$$

Then the difference that we want to evaluate becomes

$$\begin{aligned}R_0^{\mathcal{F}^Y}(\psi^{*,\mathcal{F}^Y}) - R_0^{\mathcal{F}}(\psi^{*,\mathcal{F}}) &= \mathbb{E}\left[(G_T^{\mathcal{F}^Y})^2\right] - \mathbb{E}\left[(G_T^{\mathcal{F}})^2\right] = \mathbb{E}\left[\int_0^T (\theta_r^{*,\mathcal{F}} - \theta_r^{*,\mathcal{F}^Y})^2 \mathrm{d}\langle Y \rangle_r^{\mathbb{F}}\right] \\ &= \mathbb{E}\left[\int_0^T (\theta_r^{*,\mathcal{F}})^2 \mathrm{d}\langle Y \rangle_r^{\mathbb{F}}\right] + \mathbb{E}\left[\int_0^T (\theta_r^{*,\mathcal{F}^Y})^2 \mathrm{d}\langle Y \rangle_r^{\mathbb{F}}\right] - 2\mathbb{E}\left[\int_0^T \theta_r^{*,\mathcal{F}} \theta_r^{*,\mathcal{F}^Y} \mathrm{d}\langle Y \rangle_r^{\mathbb{F}}\right].\end{aligned} \tag{94}$$

Using Eq. (62) and the definition of $\mathbb{F}^Y$-dual-predictable projections, we have that

$$\mathbb{E}\left[\int_0^t \theta_r^{*,\mathcal{F}^Y} \theta_r^{*,\mathcal{F}} \mathrm{d}\langle Y \rangle_r^{\mathbb{F}}\right] = \mathbb{E}\left[\int_0^t (\theta_r^{*,\mathcal{F}^Y})^2 \mathrm{d}\langle Y \rangle_r^{\mathbb{F}^Y}\right] = \mathbb{E}\left[\int_0^t (\theta_r^{*,\mathcal{F}^Y})^2 \mathrm{d}\langle Y \rangle_r^{\mathbb{F}}\right], \tag{95}$$

which implies

$$R_0^{\mathcal{F}^Y}(\psi^{*,\mathcal{F}^Y}) - R_0^{\mathcal{F}}(\psi^{*,\mathcal{F}}) = \mathbb{E}\left[\int_0^T (\theta_r^{*,\mathcal{F}})^2 \mathrm{d}\langle Y\rangle_r^{\mathbb{F}}\right] - \mathbb{E}\left[\int_0^T (\theta_r^{*,\mathcal{F}^Y})^2 \mathrm{d}\langle Y\rangle_r^{\mathbb{F}^Y}\right]. \tag{96}$$

Plugging in the expressions for the optimal strategies given in Eqs. (82) and (83), respectively, and denoting $\Sigma(t, X_t, Y_t) := Y_t^2\left(\sigma^2(t, Y_t) + \int_Z z^2\lambda(t, X_{t-}, Y_{t-})\phi(t, X_{t-}, Y_{t-}, \mathrm{d}z)\right)$, we have

$$\begin{aligned} R_0^{\mathcal{F}^Y}(\psi^{*,\mathcal{F}^Y}) - R_0^{\mathcal{F}}(\psi^{*,\mathcal{F}}) &= \mathbb{E}\left[\int_0^T \left(\frac{g^2(t, X_t, Y_t)}{\Sigma(t, X_t, Y_t)} - \frac{\pi_t^2(g)}{\pi_t(\Sigma)}\right)\mathrm{d}t\right] \\ &\leq C\mathbb{E}\left[\int_0^T \left(g^2(t, X_t, Y_t) - \pi_t^2(g)\right)\mathrm{d}t\right] = C\mathbb{E}\left[\int_0^T \left(g(t, X_t, Y_t) - \pi_t(g)\right)^2 \mathrm{d}t\right] \end{aligned} \tag{97}$$

for some $C > 0$, where the inequality follows by Assumption 18, and in the last equality, we used $\mathbb{E}\left[\int_0^T 2g(t, X_t, S_t)\pi_t(g)\mathrm{d}t\right] = \mathbb{E}\left[\int_0^T 2\pi_t(g)^2\mathrm{d}t\right]$.

We can conclude by saying that we found an upper bound for the expected difference between the total risks taken by an informed investor and a partially informed one which is directly proportional to the mean-squared error between the process $\{g(t, X_t, S_t),\ t \in [0, T]\}$ and its filtered estimate $\pi(g) = \{\pi_t(g),\ t \in [0, T]\}$.

Author details

Claudia Ceci[1]* and Katia Colaneri[2]

*Address all correspondence to: c.ceci@unich.it

1 Department of Economics, University of Chieti-Pescara, Pescara, Italy

2 Department of Economics, University of Perugia, Perugia, Italy

References

[1] Kushner H. On the differential equations satisfied by conditional probability densities of Markov processes, with applications. Journal of the Society for Industrial and Applied Mathematics, Series A: Control. 1964;**2**(1):106-119

[2] Kushner H. Dynamical equations for optimal nonlinear filtering. Journal of Differential Equations. 1967;**3**(2):179-190

[3] Zakai M. On the optimal filtering of diffusion processes. Probability Theory and Related Fields. 1969;**11**(3):230-243

[4] Lipster RS, Shiryaev A. Statistics of Random Processes I. Springer-Verlag; Berlin Heidelberg, 1977

[5] Kallianpur G. Stochastic Filtering Theory. Springer; Springer Verlag New York, 1980

[6] Elliott RJ. Stochastic Calculus and Applications. Springer; Berlin Heidelberg New York, 1982

[7] Kurtz TG, Ocone D. Unique characterization of condition distribution in nonlinear filtering. Annals of Probability. 1988;**16**:80-107

[8] Bhatt AG, Kallianpur G, Karandikar RL. Uniqueness and robustness of solution of measure-valued equations of nonlinear filtering. The Annals of Probability. 1995;**23**(4): 1895-1938

[9] Brémaud P. Point Processes and Queues. Springer-Verlag; New York, 1980

[10] Kliemann WH, Koch G, Marchetti F. On the unnormalized solution of the filtering problem with counting process observations. IETIT. 1990;**36**:1415-1425

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

[12] Frey R, Runggaldier W. A nonlinear filtering approach to volatility estimation with a view towards high frequency data. International Journal of Theoretical and Applied Finance. 2001;**4**(2):199-210

[13] Ceci C, Gerardi A. A model for high frequency data under partial information: A filtering approach. International Journal of Theoretical and Applied Finance. 2006;**9**(4):1-22

[14] Ceci C. Risk minimizing hedging for a partially observed high frequency data model. Stochastics: An International Journal of Probability and Stochastic Processes. 2006;**78**(1): 13-31

[15] Frey R, Runggaldier W. Pricing credit derivatives under incomplete information: A nonlinear-filtering approach. Finance and Stochastics. 2010;**14**:495-526

[16] Frey R, Schimdt T. Pricing and hedging of credit derivatives via the innovation approach to nonlinear filtering. Finance and Stochastics. 2011;**16**(1):105-133

[17] Ceci C, Colaneri K. Nonlinear filtering for jump diffusion observations. Advances in Applied Probability. 2012;**44**(3):678-701

[18] Ceci C, Colaneri K. The Zakai equation of nonlinear filtering for jump-diffusion observations: Existence and uniqueness. Applied Mathematics & Optimization. 2014;**69**(1):47-82

[19] Elliott R, Malcolm W. Discrete-time expectation maximization algorithms for Markov-modulated Poisson processes. IEEE Transactions on Automatic Control. 2008;**53**(1):247-256

[20] Björk T, Davis M, Landén C. Optimal investment under partial information. Mathematical Methods of Operations Research. 2010;**71**(2):371-399

[21] Ceci C, Colaneri K, Cretarola A. Local risk-minimization under restricted information to asset prices. Electronic Journal of Probability. 2015;**20**(96):1-30

[22] Ceci C, Colaneri K, Cretarola A. Hedging of unit-linked life insurance contracts with unobservable mortality hazard rate via local risk-minimization. Insurance: Mathematics and Economics. 2015;**60**:47-60.

[23] Ceci C, Gerardi A. Pricing for geometric marked point processes under partial information: entropy approach. International Journal of Theoretical and Applied Finance. 2009;**12**:179-207

[24] Frey R. Risk minimization with incomplete information in a model for high-frequency data. Mathematical Finance. 2000;**10**(2):215-222

[25] Nagai H, Peng S. Risk-sensitive dynamic portfolio optimization with partial information on infinite time horizon. Annals of Applied Probability. 2000;**12**:173-195

[26] Bäuerle N, Rieder U. Portfolio optimization with jumps and unobservable intensity process. Mathematical Finance. 2007;**17**(2):205-224

[27] Föllmer H, Sondermann D. Hedging of non redundant contingent claims. In: Hildenbrand W, Mas-Colell A, editors. Contribution to Mathematical Economics. North Holland, Amsterdam New York Oxford Tokyo; 1986. pp. 205-223

[28] Schweizer M. A guided tour through quadratic hedging approaches. In: Jouini E, Cvitanic J, Musiela M, editors. Option Pricing, Interest Rate and Risk Management. Cambridge University Press; Cambridge, 2001. pp. 538-574

[29] Schweizer M. Risk minimizing hedging strategies under partial information. Mathematical Finance. 1994;**4**:327-342

[30] Ceci C, Cretarola A, Russo F. GKW representation theorem under restricted information. An application to risk-minimization. Stochastics and Dynamics. 2014;**14**(2):1350019 (p. 23)

[31] Jacod J, Shiryaev A. Limit Theorems for Stochastic Processes. 2nd ed. Berlin: Springer; 2003

[32] Protter P, Shimbo K, Ethier SN, Feng J, Stockbridge RH eds, No arbitrage and general semimartingales. In: Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz. Institute of Mathematical Statistics; Beachwood, Ohio, USA, 2008. pp. 267-283