CIRJE-F-914

Optimal Hedging for Fund & Insurance Managers with Partially Observable Investment Flows

Masaaki Fujii
The University of Tokyo

Akihiko Takahashi
The University of Tokyo

January 2014; Revised in July 2014

Optimal Hedging for Fund & Insurance Managers with Partially Observable Investment Flows *

Masaaki Fujii[†], Akihiko Takahashi[‡]

July 28, 2014

Abstract

All the financial practitioners are working in incomplete markets full of unhedgeable risk-factors. Making the situation worse, they are only equipped with the imperfect information on the relevant processes. In addition to the market risk, fund and insurance managers have to be prepared for sudden and possibly contagious changes in the investment flows from their clients so that they can avoid the over- as well as under-hedging. In this work, the prices of securities, the occurrences of insured events and (possibly a network of) the investment flows are used to infer their drifts and intensities by a stochastic filtering technique. We utilize the inferred information to provide the optimal hedging strategy based on the mean-variance (or quadratic) risk criterion. A BSDE approach allows a systematic derivation of the optimal strategy, which is shown to be implementable by a set of simple ODEs and the standard Monte Carlo simulation. The presented framework may also be useful for manufactures and energy firms to install an efficient overlay of dynamic hedging by financial derivatives to minimize the costs.

Keywords : Insurance, Mean-variance hedging, BSDE, Filtering, Queueing, Jackson's network, random measure

*Forthcoming in *Quantitative Finance*. All the contents expressed in this research are solely those of the authors and do not represent any views or opinions of any institutions. The authors are not responsible or liable in any manner for any losses and/or damages caused by the use of any contents in this research.

†Graduate School of Economics, The University of Tokyo. e-mail: mfujii@e.u-tokyo.ac.jp

‡Graduate School of Economics, The University of Tokyo. e-mail: akihikot@e.u-tokyo.ac.jp

1 Introduction

In this paper, we discuss the optimal hedging strategy based on the mean-variance criterion for the fund and insurance managers in the presence of incompleteness as well as imperfect information in the market. If an unhedgeable risk-factor exists, the fund and insurance managers are forced to work in the physical measure and resort to a certain optimization technique to decide their trading strategies. In the physical measure, however, they soon encounter the problem of *imperfect information* which is usually hidden in the traditional risk-neutral world.

One of the most important factors in the financial optimizations is the drift term in the price process of a financial security. In fact, many of the financial decisions consist of taking a careful balance between the expected return, i.e. drift, and the size of risk. However, the observation of a drift term is always associated with a noise, and we need to adopt some statistical inference method. In a large number of existing works on the mean-variance hedging problem, which usually adopt the *duality method*, Pham (2001) [26], for example, studied the problem in this *partially observable drift* context. In spite of a great amount of literature [1], results with explicit solutions which can be directly implementable by practitioners have been quite rare thus far. When the explicit forms are available, they usually require various simplifying assumptions on the dependence structure among the underlying securities and their risk-premium processes, and also on the form of the hedging target, which make the motivations somewhat obscure from a practical point of view.

A new approach was proposed by Mania & Tevzadze (2003) [22], where the authors studied a minimization problem for a convex cost function and showed that the optimal value function follows a backward stochastic partial differential equation (BSPDE). They were able to decompose it into three backward stochastic differential equations (BSDEs) when the cost function has a quadratic form. Although the relevant equations are quite complicated, their approach allows a systematic derivation for a generic setup in such a way that it can be linked directly to the dynamic programming approach yielding HJB equation. In Fujii & Takahashi (2013) [9], we have studied their BSDEs to solve the mean-variance hedging problem with partially observable drifts. In the setup where Kalman-Bucy filtering scheme is applicable, we have shown that a set of simple ordinary differential equations (ODEs) and the standard Monte Carlo simulation are enough to implement the optimal strategy. We have also derived its approximate analytical expression by an asymptotic expansion method, with which we were able to simulate the distribution of the hedging error.

The problem of imperfect information is not only about the drifts of securities. Fund and insurance managers have to deal with stochastic investment flows from their clients. In particular, the timings of buy/sell orders are unpredictable and their intensities can be only statistically inferred. The same is true for loan portfolios and possibly their securitized products. It is, in fact, a well-known story in the US market that the prepayments of residential mortgages have a big impact on the residential mortgage-backed security (RMBS) price, which in turn induces significant hedging demand on interest rate swaps and swaptions. See [25], for example, as a recent practical review on the real estate finance.

[1] See Schweizer (2010) [31] as a brief survey.

In this paper, we extend [9] to incorporate the stochastic investment flows with *partially observable intensities* [2]. In the first half of the paper, where we introduce two counting processes to describe the in- and outflow of the investment units, we provide the mathematical preparations necessary for the filtering procedures. Then, we explain the solution technique for the relevant BSDEs in detail, which gives the optimal hedging strategy by means of a set of simple ODEs and the standard Monte Carlo simulation. In the latter half of the paper, we further extend the framework so that we can deal with a portfolio of insurance products. We provide a method to differentiate the effects on the demand for insurance after the insured events based on their loss severities. Furthermore, we explain how to utilize Jackson's network that is often adopted to describe a network of computers in the Queueing analysis. We show that it is quite useful for the modeling of a general network of investment flows, such as the one arising from a group of funds within which investors can switch a fund to invest.

Although we are primarily interested in providing a flexible framework for the portfolio management, the presented framework may be applicable to manufacturers and energy firms operating multiple lines of production. For example, they can use it to install an efficient overlay of dynamic hedging by financial derivatives, such as commodity and energy futures, in order to minimize the stochastic production as well as storage costs.

2 The financial market

We consider the market setup quite similar to the one used in [9] except the introduction of the stochastic investment/order flows with partially observable intensities. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\mathbb{F} = \{\mathcal{F}_t, 0 \leq t \leq T\}$ where T is a fixed time horizon. We put $\mathcal{F} = \mathcal{F}_T$ for simplicity. We assume that $\mathbb{F}$ satisfies the *usual conditions* and is big enough in a sense that it makes all the processes we introduce are adapted to this filtration.

We consider the financial market with one risk-free asset, d tradable stocks or any kind of securities, and $m := (n-d)$ non-tradable indexes or otherwise state variables relevant for stochastic volatilities, etc. For simplicity of presentation, we assume that the risk-free interest rate r is zero. Using a vector notation, the dynamics of the securities' prices $S = \{S_i\}_{1\leq i\leq d}$ and the non-tradable indexes $Y = \{Y_j\}_{d+1\leq j\leq n}$ are assumed to be given by the following diffusion processes:

$$\begin{aligned} dS_t &= \sigma(t, S_t, Y_t)\Big(dW_t + \theta_t dt\Big) \\ dY_t &= \bar{\sigma}(t, S_t, Y_t)\Big(dW_t + \theta_t dt\Big) + \rho(t, S_t, Y_t)\Big(dB_t + \alpha_t dt\Big) \ . \end{aligned} \tag{2.1}$$

Here, (W, B) are the standard $(\mathbb{P}, \mathbb{F})$-Brownian motions independent of each other and valued in $\mathbb{R}^d$ and $\mathbb{R}^m$, respectively. The known functions $\sigma(t,s,y)$, $\bar{\sigma}(t,s,y)$ and $\rho(t,s,y)$ are measurable and smooth mappings from $[0,T] \times \mathbb{R}^d \times \mathbb{R}^m$ into $\mathbb{R}^{d\times d}$, $\mathbb{R}^{m\times d}$ and $\mathbb{R}^{m\times m}$, respectively. The risk premium $z_t := \begin{pmatrix} \theta_t \\ \alpha_t \end{pmatrix}$ is assumed to follow a mean-reverting linear

[2] Note that the standard setup with the perfect observation can be treated as a special case of our framework.

Gaussian process:

$$dz_t = [\mu_t - F_t z_t]dt + \delta_t dV_t \tag{2.2}$$

where μ, F and δ are continuous and deterministic functions of time taking values in $\mathbb{R}^n$, $\mathbb{R}^{n\times n}$ and $\mathbb{R}^{n\times p}$. V is a p-dimensional standard $(\mathbb{P},\mathbb{F})$-Brownian motion independent from W as well as B.

Let us now discuss the dynamics of the investment flows. We introduce the two counting processes A and D, i.e. right-continuous integer valued increasing processes with jumps of at most 1. (A_t, D_t) represent, respectively, the total inflow and outflow of investors or investment-units[3] for an interested fund in the time interval $(0,t]$ with $A_0 = D_0 = 0$. For simplicity, we assume that they do not jump simultaneously. The total number of investment-units for the fund at time t is denoted by Q_t, which is given by

$$Q_t = Q_0 + A_t - D_t\ . \tag{2.3}$$

In this way, we model the change of the investment-units by a simple Queueing system with a single server. Later, we shall make use of a special type of Queueing network to allow investors to switch within a group of funds, which typically bundles Money-Reserve, Bond, Equity, Bull-Bear, or regional equity indexes. See [2] as a standard textbook on Queueing systems.

We assume that the counting processes have $(\mathbb{P},\mathbb{F})$-compensators, i.e.

$$\begin{aligned} \check{A}_t &:= A_t - \int_0^t \lambda^A(s, X_{s-})ds \\ \check{D}_t &:= D_t - \int_0^t \lambda^D(s, X_{s-})\mathbf{1}_{\{Q_{s-}>0\}}ds \end{aligned} \tag{2.4}$$

are $(\mathbb{P},\mathbb{F})$-martingales. Here, the intensity processes are modulated by a finite-state Markov-chain process X which takes its value in one of the N unit-vectors, $E = \{\vec{e}_1, \cdots, \vec{e}_N\}$. The dynamics of X is assumed to be given by

$$X_t = X_0 + \int_0^t R_s X_{s-} ds + U_t\ . \tag{2.5}$$

Here $\{R_t, 0 \le t \le T\}$ is a deterministic $\mathbb{R}^{N\times N}$-valued continuous function with $[R_t]_{i,j}$ denoting the rate of transition from state j to state i. U is a bounded $\mathbb{R}^N$-valued $(\mathbb{P},\mathbb{F})$-martingale independent of W, B, V, A and D.

We assume that the fund manager can continuously observe S, $\{Y\}^{\rm obs} \subset \{Y_j\}_{d+1\le j\le n}$, and the flows of investments, i.e. A and D. Q_0, which is the initial number of investment-units, is known for the manager at $t=0$. We introduce $\mathbb{G} = \{\mathcal{G}_t, 0 \le t \le T\}$ that is the $\mathbb{P}$-augmented filtration generated by the observable processes $(S, \{Y\}^{\rm obs}, A, D)$. Q_0 $(\in \mathbb{R})$ is assumed to be $\mathcal{G}_0$-measurable. As one can see from the definition of (A,D), the timing of an each investment flow is totally inaccessible for the fund manager. For the fixed-term

[3] For practical use, one may need the appropriate rescaling to make Q have tractable size.

contracts, the manager can know exactly the timing of expiries given the knowledge of the initiation dates of the contracts. However, we think that it is rather unrealistic to seek the optimal control based on the knowledge of a specific date of expiry of an each investment-unit. In our setup, the manager partially knows (i.e. statistically infer) the rate of the investment flow but cannot tell its timing at all.

$\{Y\}^{\rm obs}$ are intended to be any index processes continuously observable in the market but nontradable for the manager, which possibly include financial indexes but non-tradable for the manager by regulatory or some other reasons. $\{Y\}^{\rm obs}$ can also represent various characteristics of investors which affect the dynamics of the investment flows. They can be very important non-financial factors for the modeling of residential mortgages and life/health insurance, for example. Various aggregations of individual data at a portfolio level can be used to construct (approximately) real-time composite indexes, which then can be used as non-tradable indexes included in $\{Y\}^{\rm obs}$. If the process turns out to be rather stable, then, it can be simply added as a deterministic function.

Remark 1 : It is straightforward to introduce a stochastic interest rate if we assume that the short-rate process r is perfectly observable. In particular, if r follows a (quadratic) Gaussian process, we lose no analytical tractability for BSDEs relevant for the mean-variance hedging. The contracts of Futures written on interest rates, commodities, energies etc., which have the cycles of enlists and delists, can also be embedded into exactly the same framework. Full details are available in the extended version of our previous work [10]. ■

Assumption (A1)

(i) The stochastic differential equations (SDEs) given in (2.1) have the unique strong solutions for S and Y.
(ii) Every Y_j $(d+1 \leq j \leq n)$ is adapted to the observable filtration $\mathbb{G}$.
(iii) The matrices σ and ρ are always invertible.

Let us make a comment on the assumption *(ii)*. Through the observation of the quadratic (co)variations of $(S, \{Y\}^{obs})$, we can recover the values of $\sigma_t\sigma_t^\top$, $\bar{\sigma}_t^{obs}\sigma_t^\top$ and $(\bar{\sigma}_t\bar{\sigma}_t^\top + \rho_t\rho_t^\top)^{obs}$. We can satisfy *(ii)* by assuming the maps $(\sigma, \bar{\sigma}, \rho)$ are constructed in such a way that they allow to fix the values of all the remaining $Y_k \in \{Y\}_{d+1 \leq j \leq n} \backslash \{Y\}^{obs}$ uniquely from these quantities at any time $t \in [0,T]$ [4].

As a result, we can see that $\mathbb{G}$ is in fact the augmented filtration generated by (S, Y, A, D), and we express this fact by $\mathbb{G} = \mathbb{F}^{S,Y,A,D}$. If necessary, we can extend the model of (S, Y) in such a way that $(\sigma, \bar{\sigma}, \rho)$ can be generic $\mathbb{G}$-predictable processes, and hence can be dependent on the past history of (A, D), as long as Assumption (A1) is satisfied. This may represent a possible feedback from the investment flows to the financial market.

[4] In the case of $d = m = 1$, it is automatically satisfied by many stochastic volatility models where σ^2 depends on Y monotonically.

Assumption (A2)

(i) For every $\vec{e} \in E$, $\{\lambda^A(s,\vec{e}), 0 \le s \le T\}$ and $\{\lambda^D(s,\vec{e}), 0 \le s \le T\}$ are strictly positive $\mathbb{G}$-predictable processes.
(ii) $\mathbb{E}\left[\int_0^T \lambda^A(s,X_{s-})ds\right] + \mathbb{E}\left[\int_0^T \lambda^D(s,X_{s-})ds\right] < \infty$.

The assumption (ii) simply guarantees $\check{A}$ and $\check{D}$ are true $(\mathbb{P},\mathbb{F})$-martingales. Note that the assumption (i) allows $(\lambda_t^A,\, \lambda_t^D)$ to be dependent on $(S_t, Y_t, A_{t-}, D_{t-})$ and possibly on their past history. This flexibility is crucial for the practical use, where the first step to describe the flow of investments is regressing it by various observable quantities. We are going to model remaining unobservable effects by the hidden Markov-chain X. Note that this setup can incorporate the self-exiting jump processes (Cohen & Elliott (2013) [3]), which may be useful when there exist strong clusterings in the buy/sell orders from the investors. See also [8] for various techniques and applications of *hidden Markov models*.

Let us put $w_t := \begin{pmatrix} W_t \\ B_t \end{pmatrix}$ and introduce the following process:

$$\begin{aligned}
\widetilde{\xi}_t &:= 1 - \int_0^t \widetilde{\xi}_{s-} z_{s-}^\top dw_s + \int_0^t \widetilde{\xi}_{s-}\left(\frac{1}{\lambda^A(s,X_{s-})} - 1\right) d\check{A}_s \\
&\quad + \int_0^t \widetilde{\xi}_{s-}\left(\frac{1}{\lambda^D(s,X_{s-})} - 1\right) d\check{D}_s
\end{aligned} \tag{2.6}$$

which yields

$$\begin{aligned}
\widetilde{\xi}_t &= \exp\left(-\int_0^t z_s^\top dw_s - \frac{1}{2}\int_0^t ||z_s||^2 ds\right) \\
&\quad \times \exp\left(\int_0^t (\lambda^A(s,X_{s-}) - 1)ds + \int_0^t (\lambda^D(s,X_{s-}) - 1)\mathbf{1}_{\{Q_{s-}>0\}} ds\right) \\
&\quad \times \prod_{s\in(0,t]} \left[\frac{1}{\lambda^A(s,X_{s-})}\right]^{\Delta A_s} \prod_{s\in(0,t]} \left[\frac{1}{\lambda^D(s,X_{s-})}\right]^{\Delta D_s}.
\end{aligned} \tag{2.7}$$

We also define

$$\begin{aligned}
\widetilde{\xi}_{1,t} &:= 1 - \int_0^t \widetilde{\xi}_{1,s} z_s^\top dw_s \\
&= \exp\left(-\int_0^t z_s^\top dw_s - \frac{1}{2}\int_0^t ||z_s||^2 ds\right)
\end{aligned} \tag{2.8}$$

$$\begin{aligned}
\widetilde{\xi}_{2,t} &:= 1 + \int_0^t \widetilde{\xi}_{2,s-}\left(\frac{1}{\lambda^A(s,X_{s-})} - 1\right) d\check{A}_s + \int_0^t \widetilde{\xi}_{2,s-}\left(\frac{1}{\lambda^D(s,X_{s-})} - 1\right) d\check{D}_s \\
&= \exp\left(\int_0^t (\lambda^A(s,X_{s-}) - 1)ds + \int_0^t (\lambda^D(s,X_{s-}) - 1)\mathbf{1}_{\{Q_{s-}>0\}} ds\right) \\
&\quad \times \prod_{s\in(0,t]} \left[\frac{1}{\lambda^A(s,X_{s-})}\right]^{\Delta A_s} \prod_{s\in(0,t]} \left[\frac{1}{\lambda^D(s,X_{s-})}\right]^{\Delta D_s}.
\end{aligned} \tag{2.9}$$

We can show that $\{\widetilde{\xi}_{1,t}, 0 \le t \le T\}$ is a true $(\mathbb{P}, \mathbb{F})$-martingale due to the linear Gaussian nature of z and Lemma 3.9 in [1].

Assumption (A3)

(i) $\{\widetilde{\xi}_t, 0 \le t \le T\}$ *is a true* $(\mathbb{P}, \mathbb{F})$*-martingale.*
(ii) $\{\widetilde{\xi}_{2,t}, 0 \le t \le T\}$ *is a true* $(\mathbb{P}, \mathbb{F})$*-martingale.*

Under Assumption $(A3)$, we can define the three probability measures $\widetilde{\mathbb{P}}$, $\widetilde{\mathbb{P}}_1$ and $\widetilde{\mathbb{P}}_2$ equivalent to $\mathbb{P}$ on $(\Omega, \mathcal{F})$:

$$\left.\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \widetilde{\xi}_t, \qquad 0 \le t \le T \tag{2.10}$$

$$\left.\frac{d\widetilde{\mathbb{P}}_1}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \widetilde{\xi}_{1,t}, \quad 0 \le t \le T \tag{2.11}$$

$$\left.\frac{d\widetilde{\mathbb{P}}_2}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \widetilde{\xi}_{2,t}, \quad 0 \le t \le T\ . \tag{2.12}$$

Then, by Girsanov-Maruyama theorem (see, for example, [28]), one can show that

$$\widetilde{W}_t \ := \ W_t + \int_0^t \theta_u du \tag{2.13}$$

$$\widetilde{B}_t \ := \ B_t + \int_0^t \alpha_u du \tag{2.14}$$

are the standard $(\widetilde{\mathbb{P}}, \mathbb{F})$ as well as $(\widetilde{\mathbb{P}}_1, \mathbb{F})$-Brownian motions, and that

$$\widetilde{A}_t \ := \ A_t - t \tag{2.15}$$

$$\widetilde{D}_t \ := \ D_t - \int_0^t \mathbf{1}_{\{Q_{s-}>0\}} ds \tag{2.16}$$

are $(\widetilde{\mathbb{P}}, \mathbb{F})$ as well as $(\widetilde{\mathbb{P}}_2, \mathbb{F})$-martingales. The following lemma tells us that the filtration $\mathbb{G}$ can be generated by these simple martingales, too. This is crucial for the filtering technique we shall use below.

Lemma 1 *The filtration* $\mathbb{G} = \mathbb{F}^{S,Y,A,D}$ *is the augmented filtration generated by* $(\widetilde{W}, \widetilde{B}, \widetilde{A}, \widetilde{D})$.

Proof: Since σ and ρ are assumed to be always invertible, we can write

$$\widetilde{W}_t = \int_0^t \sigma^{-1}(u, S_u, Y_u) dS_u \tag{2.17}$$

$$\widetilde{B}_t = \int_0^t \rho^{-1}(u, S_u, Y_u)\Big(dY_u - \bar{\sigma}(u, S_u, Y_u)\sigma^{-1}(u, S_u, Y_u) dS_u\Big)\ . \tag{2.18}$$

In addition,

$$\widetilde{A}_t = A_t - t \tag{2.19}$$

$$\widetilde{D}_t = D_t - \int_0^t \mathbf{1}_{\{Q_0+A_{s-}-D_{s-}>0\}} ds \tag{2.20}$$

and $Q_0 \in \mathcal{G}_0$. Hence it is clear that $\mathbb{F}^{\widetilde{W},\widetilde{B},\widetilde{A},\widetilde{D}} \subset \mathbb{G}$. On the other hand, we have

$$\begin{aligned}
S_t &= S_0 + \int_0^t \sigma(u, S_u, Y_u) d\widetilde{W}_u \\
Y_t &= Y_0 + \int_0^t \bar{\sigma}(u, S_u, Y_u) d\widetilde{W}_u + \int_0^t \rho(u, S_u, Y_u) d\widetilde{B}_u \\
A_t &= \widetilde{A}_t + t \\
D_t &= \widetilde{D}_t + \int_0^t \mathbf{1}_{\{Q_0+A_{u-}-D_{u-}>0\}} du
\end{aligned} \tag{2.21}$$

and hence $\mathbb{G} \subset \mathbb{F}^{\widetilde{W},\widetilde{B},\widetilde{A},\widetilde{D}}$. $\square$

3 Filtering equations

In order to obtain tractable filtering equations for the unobservable processes (θ, α, X), we want to use the method of the "reference" measure where every increment of the stochastic factors becomes independent from the past filtration. The following lemmas are modifications of Proposition 3.15 in [1] to our setup.

Lemma 2 *Let Ψ_t be an integrable $\mathcal{F}_t$-measurable ($t \in [0, T]$) random variable. Then,*

$$\mathbb{E}^{\widetilde{\mathbb{P}}}\big[\Psi_t|\mathcal{G}_T\big] = \mathbb{E}^{\widetilde{\mathbb{P}}}\big[\Psi_t|\mathcal{G}_t\big] \ . \tag{3.1}$$

Proof: Let us put

$$\mathcal{G}_{t,T} = \sigma\Big(\widetilde{W}_u - \widetilde{W}_t, \widetilde{B}_u - \widetilde{B}_t, \widetilde{A}_u - \widetilde{A}_t, \widetilde{D}_u - \widetilde{D}_t; u \in [t, T]\Big) \ , \tag{3.2}$$

and then

$$\mathcal{G}_T = \mathcal{G}_t \vee \mathcal{G}_{t,T} := \sigma(\mathcal{G}_t \cup \mathcal{G}_{t,T}). \tag{3.3}$$

If $\mathcal{G}_{t,T}$ is independent of $\mathcal{F}_t$ under the measure $\widetilde{\mathbb{P}}$, it is clear that (3.1) holds as explained in [1]. Unfortunately, this is not the case in our setup due to the information carried by the jump intensity of $\widetilde{D}$, which is $\mathbf{1}_{\{Q_->0\}}$. However, in measure $\widetilde{\mathbb{P}}$, (A, D, Q) consists of a completely decoupled Queueing system with a single server, where the entrance of new queue is given by the Poisson process with unit intensity and the service (or exit) intensity is also 1 unless the queue is empty. Thus, all the information dependent on $\mathcal{F}_t$ contained in $\mathcal{G}_{t,T}$ is restricted to the Queueing system $\{(A_s, D_s, Q_s), t < s \leq T\}$. Since it is irrelevant for Ψ_t, (3.1) holds true. $\square$

Let $\mathbb{D}$ ($\mathbb{C}$) be the set of all E-valued càdlàg ($\mathbb{R}^n$-valued continuous) functions in the time interval $[0,T]$, respectively.

Lemma 3 *Let Ψ be a map $\Psi : [0,T]\times\Omega\times\mathbb{D}\to\mathbb{R}$ in such a way that $\{\Psi_t(x), 0\le t\le T\}$ is an integrable $\mathbb{G}$-predictable process for any given step function $x\in\mathbb{D}$. Then, using the hidden Markov-chain X in (2.5), we have*

$$\mathbb{E}^{\widetilde{\mathbb{P}}_2}\Bigl[\Psi_t(\{X_s, 0\le s\le t\})\Big|\mathcal{G}_T\Bigr] = \mathbb{E}^{\widetilde{\mathbb{P}}_2}\Bigl[\Psi_t(\{X_s, 0\le s\le t\})\Big|\mathcal{G}_t\Bigr] \ . \tag{3.4}$$

Proof: (A,D,Q) consists of a completely decoupled Queueing system with unit entrance and service intensities also in measure $\widetilde{\mathbb{P}}_2$. Although $(\widetilde{W},\widetilde{B})$ carries non trivial information through its drift $z=\begin{pmatrix}\theta\\ \alpha\end{pmatrix}$, it does not affect the dynamics of X by the model setup. □

Similarly, we also need the following lemma.

Lemma 4 *Let Ψ be a map $\Psi : [0,T]\times\Omega\times\mathbb{C}\to\mathbb{R}$ in such a way that $\{\Psi_t(x), 0\le t\le T\}$ is an integrable $\mathbb{G}$-predictable process for any given continuous function $x\in\mathbb{C}$. Then, using the hidden process z in (2.2), we have*

$$\mathbb{E}^{\widetilde{\mathbb{P}}_1}\Bigl[\Psi_t(\{z_s, 0\le s\le t\})\Big|\mathcal{G}_T\Bigr] = \mathbb{E}^{\widetilde{\mathbb{P}}_1}\Bigl[\Psi_t(\{z_s, 0\le s\le t\})\Big|\mathcal{G}_t\Bigr] \ . \tag{3.5}$$

Proof: In measure $\widetilde{\mathbb{P}}_1$, $(\widetilde{W},\widetilde{B})$ becomes a n-dimensional standard Brownian motion and hence the information generated by its increments is independent of $\mathcal{F}_t$. On the other hand, the observation of A and D provides non-trivial information through their intensities, $(\lambda^A(s,X_{s-}),\lambda^D(s,X_{s-}))$. However, by Assumption (A2) (i), any available information on diffusions can only appear in the form generated by $(\widetilde{W},\widetilde{B})$ and X is irrelevant for z. □

We would like to obtain the filtering equations for

$$\hat{\theta}_t := \mathbb{E}\bigl[\theta_t|\mathcal{G}_t\bigr], \quad \hat{\alpha}_t := \mathbb{E}\bigl[\alpha_t|\mathcal{G}_t\bigr] \tag{3.6}$$

and

$$\hat{X}_t := \mathbb{E}\bigl[X_t|\mathcal{G}_t\bigr] \ . \tag{3.7}$$

Since X_t is valued in $E=\{\vec{e}_1,\cdots,\vec{e}_N\}$, we have

$$\begin{aligned}\hat{\lambda}^A_t &:= \mathbb{E}\bigl[\lambda^A(t,X_{t-})|\mathcal{G}_t\bigr] = \mathbb{E}\bigl[\lambda^A(t,X_{t-})|\mathcal{G}_{t-}\bigr]\\ &= \bigl(\lambda^A(t,\vec{e})\cdot\hat{X}_{t-}\bigr) \ ,\end{aligned} \tag{3.8}$$

and similarly for $\hat{\lambda}_t^D$. Here, we have used the inner product defined by

$$\left(\lambda^A(t,\vec{e})\cdot\hat{X}_{t-}\right) := \sum_{i=1}^N \lambda^A(t,\vec{e}_i)\hat{X}_{t-}^i \tag{3.9}$$

where $\hat{X}^i$ is the i-th element of $\hat{X}$.

For notational simplicity, let us put

$$\hat{z}_t := \mathbb{E}[z_t|\mathcal{G}_t] = \begin{pmatrix}\mathbb{E}[\theta_t|\mathcal{G}_t]\\ \mathbb{E}[\alpha_t|\mathcal{G}_t]\end{pmatrix} . \tag{3.10}$$

Using Kallianpur-Striebel formula, we have

$$\hat{z}_t = \frac{\mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[\xi_{1,t}z_t|\mathcal{G}_t\right]}{\mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[\xi_{1,t}|\mathcal{G}_t\right]} \tag{3.11}$$

and

$$\hat{X}_t = \frac{\mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\xi_{2,t}X_t|\mathcal{G}_t\right]}{\mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\xi_{2,t}|\mathcal{G}_t\right]} \tag{3.12}$$

where $\xi_{1,t} := 1/\widetilde{\xi}_{1,t}$ and $\xi_{2,t} := 1/\widetilde{\xi}_{2,t}$. Note that $\{\xi_{1,t}, 0 \le t \le T\}$ and $\{\xi_{2,t}, 0 \le t \le T\}$ are $(\widetilde{\mathbb{P}}_1, \mathbb{F})$ and $(\widetilde{\mathbb{P}}_2, \mathbb{F})$ martingales, respectively. This fact can be easily proved by Bayes formula and Assumption $(A3)$. They define the inverse measure-change by:

$$\left.\frac{d\mathbb{P}}{d\widetilde{\mathbb{P}}_1}\right|_{\mathcal{F}_t} = \xi_{1,t}, \quad \left.\frac{d\mathbb{P}}{d\widetilde{\mathbb{P}}_2}\right|_{\mathcal{F}_t} = \xi_{2,t} . \tag{3.13}$$

Remark 2 : Of course, $(\hat{z}_t, \hat{X}_t)$ can also be given by the Bayes formula with $\mathbb{E}^{\widetilde{\mathbb{P}}}[\cdot|\mathcal{G}_t]$ and a $(\widetilde{\mathbb{P}}, \mathbb{F})$-martingale $\xi_t := 1/\widetilde{\xi}_t$ which defines

$$\left.\frac{d\mathbb{P}}{d\widetilde{\mathbb{P}}}\right|_{\mathcal{F}_t} = \xi_t , \tag{3.14}$$

or any other equivalent probability measures with the corresponding Radon-Nikodym densities. However, other choices do not lead to a tractable filtering equation since z and X appear together in a single equation, or the properties proved in Lemma 3 and 4 do not hold which then mixes the filter and the smoother of the unobservables. ■

Applying Itô formula, one can easily find

$$\begin{aligned}\xi_{1,t} &= 1 + \int_0^t \xi_{1,s} z_s^\top d\widetilde{w}_s \\ &= \exp\left(\int_0^t z_s^\top d\widetilde{w}_s - \frac{1}{2}\int_0^t ||z_s||^2 ds\right)\end{aligned} \tag{3.15}$$

where we have used the shorthand notation, $\widetilde{w}_t := \begin{pmatrix} \widetilde{W}_t \\ \widetilde{B}_t \end{pmatrix}$. Similarly,

$$\begin{aligned}
\xi_{2,t} &= 1 + \int_0^t \xi_{2,s-}\big(\lambda^A(s, X_{s-}) - 1\big)d\widetilde{A}_s + \int_0^t \xi_{2,s-}\big(\lambda^D(s, X_{s-}) - 1\big)d\widetilde{D}_s \\
&= \exp\left(-\int_0^t \big(\lambda^A(s, X_{s-}) - 1\big)ds - \int_0^t \big(\lambda^D(s, X_{s-}) - 1\big)\mathbf{1}_{\{Q_{s-}>0\}}ds\right) \\
&\quad \times \prod_{s\in(0,t]} \Big[\lambda^A(s, X_{s-})\Big]^{\Delta A_s} \prod_{s\in(0,t]} \Big[\lambda^D(s, X_{s-})\Big]^{\Delta D_s} , \qquad (3.16)
\end{aligned}$$

and, of course, $\xi_t = \xi_{1,t}\xi_{2,t}$. Now, we need the following two lemmas.

Lemma 5 *Let f and h be the maps $f : [0, T] \times \Omega \times \mathbb{D} \to \mathbb{R}$ and $h : [0, T] \times \Omega \times \mathbb{D} \to \mathbb{R}^N$ in such a way that $\{f_t(x), 0 \le t \le T\}$ and $\{h_t(x), 0 \le t \le T\}$ are $\mathbb{G}$-predictable processes for any given step function $x \in \mathbb{D}$. For each $t \in [0, T]$, $f_t(x)$ and $h_t(x)$ depend on x only in the corresponding time interval $[0, t)$. In addition, let suppose they satisfy*

$$\mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\int_0^T |f_s(X)|ds\right] + \mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\int_0^T ||h_s(X)||ds\right] < \infty . \qquad (3.17)$$

Then, the following relations hold:

$$\mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\int_0^t f_s(X)ds\Big|\mathcal{G}_t\right] = \int_0^t \mathbb{E}^{\widetilde{\mathbb{P}}_2}\big[f_s(X)|\mathcal{G}_{s-}\big]ds \qquad (3.18)$$

$$\mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\int_0^t f_s(X)d\widetilde{A}_s\Big|\mathcal{G}_t\right] = \int_0^t \mathbb{E}^{\widetilde{\mathbb{P}}_2}\big[f_s(X)|\mathcal{G}_{s-}\big]d\widetilde{A}_s \qquad (3.19)$$

$$\mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\int_0^t f_s(X)d\widetilde{D}_s\Big|\mathcal{G}_t\right] = \int_0^t \mathbb{E}^{\widetilde{\mathbb{P}}_2}\big[f_s(X)|\mathcal{G}_{s-}\big]d\widetilde{D}_s \qquad (3.20)$$

$$\mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\int_0^t h_s(X)^\top dU_s\Big|\mathcal{G}_t\right] = 0 . \qquad (3.21)$$

Proof: Let us prove the first relation. Suppose that f is simple, i.e.

$$f_s(X) = \sum_{i=1}^k f_i(X)\mathbf{1}_{(a_i,b_i]}(s) \qquad (3.22)$$

where $(a_i, b_i], i = 1, \cdots, k$ are the disjoint intervals of $[0, t]$ and $f_i(X)$ is $\mathcal{F}_{a_i}$-measurable.

We have

$$\begin{aligned}
\mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\int_0^t f_s(X)ds\Big|\mathcal{G}_t\right] &= \sum_{i=1}^k \mathbb{E}^{\widetilde{\mathbb{P}}_2}\big[f_i(X)(b_i-a_i)|\mathcal{G}_t\big] \\
&= \sum_{i=1}^k \mathbb{E}^{\widetilde{\mathbb{P}}_2}\big[f_i(X)|\mathcal{G}_{a_i}\vee\mathcal{G}_{a_i,t}\big](b_i-a_i) \\
&= \sum_{i=1}^k \mathbb{E}^{\widetilde{\mathbb{P}}_2}\big[f_i(X)|\mathcal{G}_{a_i}\big](b_i-a_i) \\
&= \int_0^t \mathbb{E}^{\widetilde{\mathbb{P}}_2}\big[f_s(X)|\mathcal{G}_{s-}\big]ds\ , \qquad (3.23)
\end{aligned}$$

where, in the third equality, we have used Lemma 3. For general f, we can use the decomposition $f = f^+ - f^-$ and the monotone convergence of increasing sequence of simple functions.

Now, let us move to the second relation. We know that $\{\widetilde{A}_t, 0 \le t \le T\}$ is a pure jump $(\widetilde{\mathbb{P}}_2,\mathbb{F})$-martingale with unit intensity. By (3.17), we see

$$\left\{\int_0^t f_s(X)d\widetilde{A}_s, 0\le t\le T\right\} \qquad (3.24)$$

is a $(\widetilde{\mathbb{P}}_2,\mathbb{F})$-martingale. Let us suppose $\{\varphi_s, 0 \le s \le T\}$ is an arbitrary bounded $\mathbb{G}$-predictable process. Then,

$$\begin{aligned}
\mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\int_0^t \varphi_s f_s(X)dA_s\right] &= \mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\int_0^t \varphi_s f_s(X)ds\right] \\
&= \mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\int_0^t \varphi_s \mathbb{E}^{\widetilde{\mathbb{P}}_2}\big[f_s(X)|\mathcal{G}_{s-}\big]ds\right] \\
&= \mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\int_0^t \varphi_s \mathbb{E}^{\widetilde{\mathbb{P}}_2}\big[f_s(X)|\mathcal{G}_{s-}\big]dA_s\right]\ , \qquad (3.25)
\end{aligned}$$

where, in the second equality, we have used the result of the first part of the proof. Since the relation holds true for an arbitrary φ, the second claim of Lemma needs to hold. The third relation with $\widetilde{D}$ can be proved exactly in the same way. The last relation is trivial since U is a bounded martingale independent from the filtration $\mathcal{G}$. $\square$

Lemma 6 *Let f, g and h be the maps $f : [0,T]\times\Omega\times\mathbb{C}\to\mathbb{R}$, $g : [0,T]\times\Omega\times\mathbb{C}\to\mathbb{R}^n$ and $h : [0,T]\times\Omega\times\mathbb{C}\to\mathbb{R}^p$ in such a way that $\{f_t(x), 0\le t\le T\}$, $\{g_t(x), 0\le t\le T\}$ and $\{h_t(x), 0\le t\le T\}$ are $\mathbb{G}$-predictable processes for any given continuous function $x\in\mathbb{C}$. For each $t\in[0,T]$, $f_t(x)$, $g_t(x)$ and $h_t(x)$ depend on x only in the corresponding time interval $[0,t]$. In addition, let suppose they satisfy*

$$\mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[\int_0^T |f_s(z)|ds\right] + \mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[\int_0^T ||g_s(z)||^2 ds\right] + \mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[\int_0^T ||h_s(z)||^2 ds\right] < \infty\ . \qquad (3.26)$$

Then, the following relations hold:

$$\mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[\int_0^t f_s(z)ds\Big|\mathcal{G}_t\right] = \int_0^t \mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[f_s(z)|\mathcal{G}_s\right]ds \tag{3.27}$$

$$\mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[\int_0^t g_s(z)^\top d\widetilde{w}_s\Big|\mathcal{G}_t\right] = \int_0^t \mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[g_s(z)|\mathcal{G}_s\right]^\top d\widetilde{w}_s \tag{3.28}$$

$$\mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[\int_0^t h_s(z)^\top dV_s\Big|\mathcal{G}_t\right] = 0\ . \tag{3.29}$$

Proof: It can be proved similarly as Lemma 5 using the result of Lemma 4. See the proof of Lemma 5.4 in [32] for detail. □

Using Lemma 6 and Kallianpur-Striebel formula, we can apply the well-known Kalman-Bucy filter for z. Saying that, applying Lemma 6 is non-trivial due to the unbounded nature of the Gaussian process z. Fortunately, however, the discussion in Chapter 3 in [1] shows Lemma 6 can still be applied, and also guarantees that the famous Zakai and Kushner-Stratonovich equations hold true.

Let us suppose that the prior distribution of z is a Gaussian distribution with a mean z_0 and a covariance Σ_0. Then, the dynamics of the conditional expectation is known to follow

$$d\hat{z}_t = \left[\mu_t - F_t\hat{z}_t\right]dt + \Sigma(t)dn_t, \quad \hat{z}_0 = z_0 \tag{3.30}$$

where n_t is the shorthand notation of $n_t = \begin{pmatrix} N_t \\ M_t \end{pmatrix}$, and $\Sigma(t)$ is the solution for the following ODE:

$$\frac{d\Sigma(t)}{dt} = \delta_t\delta_t^\top - F_t\Sigma(t) - \Sigma(t)F_t^\top - \Sigma(t)^2, \quad \Sigma(0) = \Sigma_0 \tag{3.31}$$

Here,

$$\begin{aligned} N_t &:= \widetilde{W}_t - \int_0^t \hat{\theta}_s ds \\ M_t &:= \widetilde{B}_t - \int_0^t \hat{\alpha}_s ds \end{aligned} \tag{3.32}$$

are called the innovation processes, which are independent $(\mathbb{P}, \mathbb{G})$-Brownian motions. For detail of the derivation, see Section 6 in [1].

Now, let us move to the filtering equation for X. We follow the arguments of derivation given in [7, 3]. Firstly, we want to derive the unnormalized filter of X:

$$q_t := \mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\xi_{2,t}X_t|\mathcal{G}_t\right]\ . \tag{3.33}$$

Applying Itô-formula, one obtains

$$\xi_{2,t}X_t = X_0 + \int_0^t \xi_{2,s-}R_sX_{s-}ds + \int_0^t \xi_{2,s-}dU_s$$
$$+ \int_0^t \xi_{2,s-}X_{s-}\Big[\big(\lambda^A(s,X_{s-})-1\big)d\widetilde{A}_s + \big(\lambda^D(s,X_{s-})-1\big)d\widetilde{D}_s\Big] \ . \qquad (3.34)$$

Lemma 7 *The dynamics of q_t is given by the following equation:*

$$q_t = q_0 + \int_0^t R_sq_{s-}ds + \int_0^t \big(\Lambda_s^A - \mathbb{I}\big)q_{s-}d\widetilde{A}_s + \int_0^t \big(\Lambda_s^D - \mathbb{I}\big)q_{s-}d\widetilde{D}_s \ , \qquad (3.35)$$

where

$$\Lambda_s^A = \mathrm{diag}\Big(\lambda^A(s,\vec{e}_1),\cdots,\lambda^A(s,\vec{e}_N)\Big), \ 0 \le s \le T \qquad (3.36)$$
$$\Lambda_s^D = \mathrm{diag}\Big(\lambda^D(s,\vec{e}_1),\cdots,\lambda^D(s,\vec{e}_N)\Big), \ 0 \le s \le T \qquad (3.37)$$

are $\mathbb{G}$-predictable processes valued in $(n \times n)$ diagonal matrices.

Proof: Take the conditional expectation $\mathbb{E}^{\widetilde{\mathbb{P}}_2}[\cdot|\mathcal{G}_t]$ in the both hands of (3.34). Due to the bounded nature of X and Assumption (A2), we can apply Lemma 5. In particular, one can see

$$\mathbb{E}^{\widetilde{\mathbb{P}}_2}\left[\int_0^t \xi_{2,s-}\Big|\lambda^A(s,X_{s-})-1\Big|ds\right] = \mathbb{E}^{\mathbb{P}}\left[\int_0^t \Big|\lambda^A(s,X_{s-})-1\Big|ds\right] < \infty. \qquad (3.38)$$

Using the fact that $\lambda^a(s,X_s)X_s = \Lambda_s^a X_s$ for $a = A, \ D$, one obtains the desired result. □

Since $(\mathbf{1}\cdot X_t) \equiv 1$, we obtain

$$\hat{X}_t = \frac{q_t}{(\mathbf{1}\cdot q_t)} \ , \qquad (3.39)$$

where $\mathbf{1} = (1,\cdots,1)^\top$ is a N-dimensional vector. Now, the filtered intensities $(\hat{\lambda}^A, \hat{\lambda}^D)$ can be obtained by (3.8). We can show by Assumption (A2) that

$$\hat{A}_t = A_t - \int_0^t \hat{\lambda}_s^A ds$$
$$\hat{D}_t = D_t - \int_0^t \hat{\lambda}_s^D \mathbf{1}_{\{Q_{s-}>0\}}ds \qquad (3.40)$$

are $(\mathbb{P},\mathbb{G})$-martingales.

Remark 3 : Let us comment on how to simulate (A,D) in the physical measure $(\mathbb{P},\mathbb{G})$.

q_t can be expressed as

$$q_t = q_0 + \int_0^t R_s q_{s-} ds - \int_0^t \Big\{ \big(\Lambda_s^A - \mathbb{I}\big) + \big(\Lambda_s^D - \mathbb{I}\big)\mathbf{1}_{\{Q_{s-}>0\}} \Big\} q_{s-} ds$$
$$+ \int_0^t \big(\Lambda_s^A - \mathbb{I}\big) q_{s-} dA_s + \int_0^t \big(\Lambda_s^D - \mathbb{I}\big) q_{s-} dD_s \ . \tag{3.41}$$

Thus, between any two jumps, q follows a $\mathbb{G}$-predictable continuous process given by the first line of (3.41). When there is a jump, we have

$$q_t = \Lambda_t^A q_{t-} \Delta A_t + \Lambda_t^D q_{t-} \Delta D_t \ . \tag{3.42}$$

In $(\mathbb{P}, \mathbb{G})$, A and D are counting processes whose intensities are $\hat{\lambda}_t^A = \big(\lambda^A(t, \vec{e}) \cdot \hat{X}_{t-}\big)$ and $\hat{\lambda}_t^D = \big(\lambda^D(t, \vec{e}) \cdot \hat{X}_{t-}\big)\mathbf{1}_{\{Q_{t-}>0\}}$ respectively, where $\hat{X}_t$ is given by (3.39). Thus, based on these formulas, we can carry out random draw for A and D by running the q's process in parallel. At the jump, $(\hat{\lambda}^A, \hat{\lambda}^D)$ also jumps due to the jump of q given by (3.42). In fact, it is well-known that these jumps in intensities are crucial to reproduce strong clusterings of events observed in defaults, rating migrations, and other herding behaviors among investors. It may be also the case for natural disasters affected by the global climate change. ■

For later purpose, let us define

$$\xi_t^{\mathcal{G}} = \mathbb{E}^{\widetilde{\mathbb{P}}}\big[\xi_t | \mathcal{G}_t\big] \tag{3.43}$$

which is $(\widetilde{\mathbb{P}}, \mathbb{G})$-martingale specifying the measure change conditional on $\mathcal{G}_t$:

$$\frac{d\mathbb{P}}{d\widetilde{\mathbb{P}}}\Big|_{\mathcal{G}_t} = \xi_t^{\mathcal{G}} \ . \tag{3.44}$$

Then, the inverse measure change is similarly given by using $\widetilde{\xi}_t^{\mathcal{G}} := 1/\xi_t^{\mathcal{G}}$ as

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\Big|_{\mathcal{G}_t} = \widetilde{\xi}_t^{\mathcal{G}} \ . \tag{3.45}$$

4 Mean-Variance (Quadratic) Hedging

We suppose that the manager wants to minimize the square difference between the liability and the value of the hedging portfolio. The terminal liability $H = H(S_u, Y_u, A_u, D_u;\ 0 \leq u \leq T)$, which is assumed to be $\mathcal{G}_T$-measurable random variable, would depend on the performance of tradable and/or non-tradable indexes as well as the number of investment-units. It can contain not only the payments to the investors but also the target profit for the management company.

In addition to the terminal liability, we assume that there also exist cash flows associated with the payments of dividends, principles for unwound units, and the receipts of management fees, penalties for early terminations and the initial proceeds, etc. It is convenient for us to include the stream of cash flows into the wealth dynamics as

$$\mathcal{W}_t^\pi(s,w) = w + \int_s^t \pi_u^\top dS_u$$
$$+ \int_s^t \kappa_u Q_u du + \int_s^t e_u dA_u - \int_s^t g_u dD_u \tag{4.1}$$

where $\big(\kappa_t, e_t, g_t,\ 0 \le t \le T\big)$ are $\mathbb{G}$-predictable processes representing various cash flows just explained. Here, $\{\pi_t \in \mathbb{R}^d, 0 \le t \le T\}$ is a $\mathbb{G}$-predictable trading strategy for the tradable securities. We suppose that the goal of the fund manager is to solve

$$V(t,w) = \operatorname*{ess\,inf}_{\pi\in\Pi} \mathbb{E}\left[\Big(H - \mathcal{W}_T^\pi(t,w)\Big)^2 \Big| \mathcal{G}_t\right] . \tag{4.2}$$

Here, we denote Π is the set of $\mathbb{G}$-predictable trading strategies satisfying the $\mathbb{E}[(\mathcal{W}_T^\pi)^2] < \infty$. For the problem being well-posed, we assume H and the intermediate cash flows $(\kappa_u, e_u, g_u, 0 \le u \le T)$ satisfy the square integrability condition.

Assumption (A4)

$$\mathbb{E}\left[|H|^2 + \int_0^T \Big(|\kappa_u|^2 Q_u^2 + |e_u|^2 \lambda_u^A + |g_u|^2 \lambda_u^D\Big) du\right] < \infty . \tag{4.3}$$

We also make the following assumption in order to obtain the predictable representation in terms of the set of innovation processes:

Assumption (A5)

Every $(\widetilde{P}, \mathbb{G})$-local martingale $\widetilde{m} = (\widetilde{m}_t)_{t\ge 0}$ has the integral form

$$\widetilde{m}_t = \widetilde{m}_0 + \int_0^t \widetilde{\phi}_s^\top d\widetilde{w}_s + \int_0^t \widetilde{J}_s^A d\widetilde{A}_s + \int_0^t \widetilde{J}_s^D d\widetilde{D}_s \tag{4.4}$$

with appropriate $\mathbb{G}$-predictable coefficients $(\widetilde{\phi}, \widetilde{J}^A, \widetilde{J}^D)$.

Note that $\mathbb{G}$ is the augmented filtration generated by $(\widetilde{w}, \widetilde{A}, \widetilde{D})$. If the indicator function $\mathbf{1}_{\{Q_- > 0\}}$ is absent from the predictable part of $\widetilde{D}$, the above assumption is indeed satisfied by Theorem 4.34 in Chapter III of [15]. Or, if $\widetilde{m}$ is square integrable, then one can use Theorem 44 in Chapter IV of [28] to show the assumption holds true.

Lemma 8 *Let m be any $(\mathbb{P}, \mathbb{G})$-local martingale with $m_0 = 0$. Under Assumption (A5), there exist $\mathbb{G}$-predictable processes $(\phi_t \in \mathbb{R}^n, J_t^A \in \mathbb{R}, J_t^D \in \mathbb{R}, 0 \le t \le T)$ such that*

$$m_t = \int_0^t \phi_s^\top dn_s + \int_0^t J_s^A d\hat{A}_s + \int_0^t J_s^D d\hat{D}_s,\ 0 \le t \le T . \tag{4.5}$$

Proof: The proof is very similar to that of Lemma 4.1 in [27]. Suppose m is a $(\mathbb{P}, \mathbb{G})$-local martingale. Then, the Bayes formula tells us that the process

$$\widetilde{m}_t = m_t \xi_t^{\mathcal{G}}, \ 0 \leq t \leq T \tag{4.6}$$

is a $(\widetilde{\mathbb{P}}, \mathbb{G})$-local martingale.

By Assumption (A5), we have an integral form

$$\widetilde{m}_t = \int_0^t \widetilde{\phi}_s^{\top} d\widetilde{w}_s + \int_0^t \widetilde{J}_s^A d\widetilde{A}_s + \int_0^t \widetilde{J}_s^D d\widetilde{D}_s \tag{4.7}$$

with some appropriate $\mathbb{G}$-predictable coefficients. Since $m_t = \widetilde{m}_t \widetilde{\xi}_t^{\mathcal{G}}$, the application of Itô formula yields

$$\begin{aligned} dm_t = \widetilde{\xi}_{t-}^{\mathcal{G}} \Bigg\{ & [\widetilde{\phi}_t - \widetilde{m}_{t-} \hat{z}_t]^{\top} dn_t + \frac{1}{\hat{\lambda}_t^A} \big[\widetilde{J}_t^A - (\hat{\lambda}_t^A - 1) \widetilde{m}_{t-} \big] d\hat{A}_t \\ & + \frac{1}{\hat{\lambda}_t^D} \big[\widetilde{J}_t^D - (\hat{\lambda}_t^D - 1) \widetilde{m}_{t-} \big] d\hat{D}_t \Bigg\}, \end{aligned} \tag{4.8}$$

which proves our claim with

$$\begin{aligned} &\phi_t = \widetilde{\xi}_{t-}^{\mathcal{G}} [\widetilde{\phi}_t - \widetilde{m}_{t-} \hat{z}_t] \\ &J_t^A = \widetilde{\xi}_{t-}^{\mathcal{G}} \frac{1}{\hat{\lambda}_t^A} \big[\widetilde{J}_t^A - (\hat{\lambda}_t^A - 1) \widetilde{m}_{t-} \big], \quad J_t^D = \widetilde{\xi}_{t-}^{\mathcal{G}} \frac{1}{\hat{\lambda}_t^D} \big[\widetilde{J}_t^D - (\hat{\lambda}_t^D - 1) \widetilde{m}_{t-} \big] \ . \ \square \end{aligned} \tag{4.9}$$

Let us now follow the methodology proposed by Mania&Tevzadze (2003) [22] and extend it to derive the set of BSDEs for the optimal hedging strategy with jump processes.

Firstly, let us remind **the optimality principle** (see, Proposition A.1 in [22]):
(i) For all $w \in \mathbb{R}$, $\pi \in \Pi$ and $s \in [0, T]$, the process $\{V(t, \mathcal{W}_t^{\pi}(s, w)), s \leq t \leq T\}$ is a $(\mathbb{P}, \mathbb{G})$-submartingale.
(ii) π^* is optimal if and only if $\{V(t, \mathcal{W}_t^{\pi^*}(s, w)), s \leq t \leq T\}$ is a $(\mathbb{P}, \mathbb{G})$-martingale.

By Lemma 8, we can express

$$\begin{aligned} V(t, w) = V(0, w) + \int_0^t a(u, w) du + \int_0^t Z(u, w)^{\top} dN_u + \int_0^t \Gamma(u, w)^{\top} dM_u \\ + \int_0^t J^A(u, w) d\hat{A}_u + \int_0^t J^D(u, w) d\hat{D}_u \end{aligned} \tag{4.10}$$

with appropriate $\mathbb{G}$-predictable processes (a, Z, Γ, J^A, J^D) for a given $w \in \mathbb{R}$. More precisely, predictable jump components can exist, for example if there exist discrete coupon payments in the process $\mathcal{W}$. The necessary extension can be done straightforwardly. Assuming that $V(t, w)$ is twice continuously differentiable with respect to w for all (ω, t), we can apply Itô-Ventzell formula. Details of the Itô-Ventzell formula are available in

Theorem 3.3.1 of [20] as well as in Theorem 3.1 of [24]. Note that the *forward integral* with respect to the random measure used in [24] simply coincides with the Itô integral when the integrands are predictable processes as in the current problem.

Now, the dynamics of $V(t, \mathcal{W}_t^\pi(s,w))$ is given by

$$\begin{aligned}
V(t,\mathcal{W}_t^\pi) &= V(s,w) + \int_s^t a(u,\mathcal{W}_{u-}^\pi)du + \int_s^t Z(u,\mathcal{W}_{u-}^\pi)^\top dN_u + \int_s^t \Gamma(u,\mathcal{W}_{u-}^\pi)^\top dM_u \\
&+ \int_s^t V_w(u,\mathcal{W}_{u-}^\pi)d\mathcal{W}_u^{\pi,c} + \int_s^t d\Big\langle V_w^c(\cdot,\mathcal{W}_\cdot^\pi),\mathcal{W}_\cdot^{\pi,c}\Big\rangle_u + \frac{1}{2}\int_s^t V_{ww}(u,\mathcal{W}_{u-}^\pi)d\big\langle\mathcal{W}^{\pi,c}\big\rangle_u \\
&+ \int_s^t J^A(u,\mathcal{W}_{u-}^\pi)d\hat{A}_u^c + \int_s^t J^D(u,\mathcal{W}_{u-}^\pi)d\hat{D}_u^c \\
&+ \int_s^t \Big[V(u,\mathcal{W}_u^\pi) + J^A(u,\mathcal{W}_u^\pi) - V(u,\mathcal{W}_{u-}^\pi)\Big]dA_u \\
&+ \int_s^t \Big[V(u,\mathcal{W}_u^\pi) + J^D(u,\mathcal{W}_u^\pi) - V(u,\mathcal{W}_{u-}^\pi)\Big]dD_u \ . \qquad (4.11)
\end{aligned}$$

Here the superscript c denotes the continuous part of the process. Arranging the drift term and completing the square in terms of π so that it satisfies the conditions for the optimality principle, one can find

$$\begin{aligned}
a(t,w) &+ \inf_{\pi\in\Pi}\left\{\frac{1}{2}V_{ww}(t,w)\Big|\Big|\sigma_t^\top\pi_t + \frac{[Z_w(t,w)+V_w(t,w)\hat{\theta}_t]}{V_{ww}(t,w)}\Big|\Big|^2 - \frac{||Z_w(t,w)+V_w(t,w)\hat{\theta}_t||^2}{2V_{ww}(t,w)}\right\} \\
&+V_w(t,w)\kappa_t Q_t + \Big[J^A(t,w+e_t) - J^A(t,w) + V(t,w+e_t) - V(t,w)\Big]\hat{\lambda}_t^A \\
&+\Big[J^D(t,w-g_t) - J^D(t,w) + V(t,w-g_t) - V(t,w)\Big]\hat{\lambda}_t^D\mathbf{1}_{\{Q_{t-}>0\}} = 0 \ . \qquad (4.12)
\end{aligned}$$

Assuming that there exist $\pi^* \in \Pi$ making $||\cdot||^2$ vanish, which is the first term inside the $\{\ \}$ of (4.12), the value function is given by the following backward stochastic PDE:

$$\begin{aligned}
V(t,w) &= (H-w)^2 - \int_t^T\left\{\frac{||Z_w(s,w)+V_w(s,w)\hat{\theta}_s||^2}{2V_{ww}(s,w)} - V_w(s,w)\kappa_s Q_s\right\}ds \\
&+ \int_t^T \Big[J^A(s,w+e_s) - J^A(s,w) + V(s,w+e_s) - V(s,w)\Big]\hat{\lambda}_s^A ds \\
&+ \int_t^T \Big[J^D(s,w-g_s) - J^D(s,w) + V(s,w-g_s) - V(s,w)\Big]\hat{\lambda}_s^D\mathbf{1}_{\{Q_{s-}>0\}}ds \\
&- \int_t^T Z(s,w)^\top dN_s - \int_t^T \Gamma(s,w)^\top dM_s - \int_t^T J^A(s,w)d\hat{A}_s - \int_t^T J^D(s,w)d\hat{D}_s \ . \qquad (4.13)
\end{aligned}$$

Although the above BSPDE looks much more complicated than that appears in [22] with continuous underlyings, we can still exploit the quadratic nature of the problem. By inserting

$$
\begin{aligned}
V(t,w) &= w^2V_2(t) - 2wV_1(t) + V_0(t) \\
Z(t,w) &= w^2Z_2(t) - 2wZ_1(t) + Z_0(t), \quad \Gamma(t,w) = w^2\Gamma_2(t) - 2w\Gamma_1(t) + \Gamma_0(t) \\
J^A(t,w) &= w^2J_2^A(t) - 2wJ_1^A(t) + J_0^A(t), \quad J^D(t,w) = w^2J_2^D(t) - 2wJ_1^D(t) + J_0^D(t)
\end{aligned}
\tag{4.14}
$$

into (4.13), we can decompose the BSPDE into the following three w-independent BSDEs:

$$
V_2(t) = 1 - \int_t^T \frac{||Z_2(s) + V_2(s)\hat{\theta}_s||^2}{V_2(s)} ds - \int_t^T Z_2(s)^\top dN_s - \int_t^T \Gamma_2(s)^\top dM_s \tag{4.15}
$$

$$
\begin{aligned}
V_1(t) = H &- \int_t^T \frac{[Z_2(s) + V_2(s)\hat{\theta}_s]^\top [Z_1(s) + V_1(s)\hat{\theta}_s]}{V_2(s)} ds \\
&- \int_t^T \left\{ \left[\kappa_s Q_s + e_s\hat{\lambda}_s^A - g_s\hat{\lambda}_s^D \mathbf{1}_{\{Q_{s-}>0\}}\right] V_2(s) \right\} ds \\
- \int_t^T Z_1(s)^\top dN_s &- \int_t^T \Gamma_1(s)^\top dM_s - \int_t^T J_1^A(s)d\hat{A}_s - \int_t^T J_1^D(s)d\hat{D}_s
\end{aligned}
\tag{4.16}
$$

$$
\begin{aligned}
V_0(t) = H^2 &- \int_t^T \left\{ \frac{||Z_1(s) + V_1(s)\hat{\theta}_s||^2}{V_2(s)} + 2\kappa_s Q_s V_1(s) \right\} ds \\
&+ \int_t^T \left[e_s^2 V_2(s) - 2e_s\big(J_1^A(s) + V_1(s)\big) \right] \hat{\lambda}_s^A ds \\
&+ \int_t^T \left[g_s^2 V_2(s) + 2g_s\big(J_1^D(s) + V_1(s)\big) \right] \hat{\lambda}_s^D \mathbf{1}_{\{Q_{s-}>0\}} ds \\
- \int_t^T Z_0(s)^\top dN_s &- \int_t^T \Gamma_0(s)^\top dM_s - \int_t^T J_0^A(s)d\hat{A}_s - \int_t^T J_0^D(s)d\hat{D}_s \, .
\end{aligned}
\tag{4.17}
$$

In the derivation, we have used the fact that both J_2^A and J_2^D are identically zero due to the continuity of the risk-premium process $\hat{z}$.

It is difficult to give the general conditions which guarantee the existence and uniqueness of the solutions for (4.15), (4.16) and (4.17). In particular, the unboundedness of $\hat{z}$ due to its Gaussian nature, makes the problem complicated. However, the following lemma is a simple consequence of the *optimality principle.*

Lemma 9 *Suppose that the three BSDEs (4.15), (4.16) and (4.17) have well-defined solutions and*

$$
\pi_t^* = (\sigma^{-1})^\top(t, S_t, Y_t)\frac{1}{V_2(t)}\Big\{ \big[Z_1(t) + V_1(t)\hat{\theta}_t\big] - \mathcal{W}_t^{\pi^*}\big[Z_2(t) + V_2(t)\hat{\theta}_t\big] \Big\} \tag{4.18}
$$

is an admissible strategy i.e. $\pi^* \in \Pi$. *Then,* π^* *is the optimal hedging strategy and the value function is given by the solutions of these BSDEs by* $V(t,w) = w^2V_2(t) - 2wV_1(t) + V_0(t)$.

Furthermore, if there exists the optimal strategy π^*, we can show that it is unique due to the strict convexity of the cost function. (See, Remark 2.2 of [22].) Note that the form of the optimal hedging strategy π^* in (4.18) can be easily found from (4.12) and the decomposition (4.14). The variance optimal measure used in the duality approach is closely related to V_2. See Propositions 1.5.2 and 1.5.3 of Mania & Tevzadze (2008) [23].

Although the three BSDEs (4.15), (4.16) and (4.17) look very complicated at first sight, they have the following nice properties which make the mean-variance (or quadratic) hedging particularly useful for a large scale portfolio management:

- *Only V_2 follows a non-linear BSDE.*
- *V_2 (and hence Z_2) is independent from the hedging target and the cash-flow streams.*
- *V_1 depends on the hedging target and the cash-flow streams, but follows a linear BSDE.*
- *V_1 (and hence Z_1) depends only linearly on the hedging target and the cash-flow streams.*

These properties are stemming from the fact that the optimal strategy is given by the projection of the hedging target in $L^2(\mathbb{P})$ on the space spanned by the tradable securities [30]. From (4.18), we can see that the optimal hedging strategy is linear in the hedging target as well as the other cash-flow streams for a given horizon T. This means that, for a given wealth $\mathcal{W}_t$ at time t, the optimal hedging positions can be evaluated for each portfolio component separately. Therefore, sharing the information about the overall wealth $\mathcal{W}_t$, a large scale portfolio can be controlled systematically by arranging desks in such a way that each desk is responsible for evaluating and hedging a certain sector of portfolio, such as equity-related and commodity-related sub-portfolios, etc.

5 A solution technique for the optimal strategy

5.1 Solving V_2 by ODEs

From the discussion in the last section, it becomes clear that solving the BSDE for V_2 (4.15) is the key. Although the existence and uniqueness of the solution for (4.15) are proven for the case with a bounded risk-premium process by Kobylanski (2000) [18] and Kohlmann & Tang (2002) [19], this is not the case in the current setup since $(\hat{\theta}, \hat{\alpha})$ arising from the Kalman-Bucy filter are Gaussian and hence unbounded. Although the general conditions are not known, we have a very useful method to directly solve it under certain conditions, which are likely to hold in most of the plausible situations [9].

Firstly, let us define the following change of variables:

$$\begin{aligned} V_L(t) &:= \log V_2(t) \\ Z_L(t) &:= Z_2(t)/V_2(t) \\ \Gamma_L(t) &:= \Gamma_2(t)/V_2(t) \ . \end{aligned} \tag{5.1}$$

Then, (4.15) can equivalently be given by a quadratic-growth BSDE

$$V_L(t) = -\int_t^T \left\{ \frac{1}{2}\Big(||Z_L(s)||^2 - ||\Gamma_L(s)||^2 \Big) + 2\hat{\theta}_s^\top Z_L(s) + ||\hat{\theta}_s||^2 \right\} ds$$
$$- \int_t^T Z_L(s)^\top dN_s - \int_t^T \Gamma_L(s)^\top dM_s \ . \tag{5.2}$$

We introduce a $(n \times n)$ matrix-valued deterministic function defined by

$$\Xi(t) := \big(\Sigma_d^\top \Sigma_d\big)(t) - \big(\Sigma_m^\top \Sigma_m\big)(t) \tag{5.3}$$

where $\Sigma_d(t)$ ($\Sigma_m(t)$) are $d \times n$ ($m \times n$) matrices obtained by restricting to the first d (last m) rows of $\Sigma(t)$. Furthermore, we use $\mathbf{1}_{(d,0)}$ to represent a $(n \times n)$ diagonal matrix whose first d elements are 1 and the others zero.

Lemma 10 *Consider the following matrix-valued ODEs for $a^{[2]}(t) \in \mathbb{R}^{n\times n}$, $a^{[1]}(t) \in \mathbb{R}^n$ and $a^{[0]}(t) \in \mathbb{R}$,*

$$\dot{a}^{[2]}(t) = 2\mathbf{1}_{(d,0)} + a^{[2]}(t)\Xi(t)a^{[2]}(t)$$
$$+F_t^\top a^{[2]}(t) + a^{[2]}(t)F_t + 2\Big(\mathbf{1}_{(d,0)}\Sigma(t)a^{[2]}(t) + a^{[2]}(t)\Sigma(t)\mathbf{1}_{(d,0)}\Big) \tag{5.4}$$
$$\dot{a}^{[1]}(t) = -a^{[2]}(t)\mu_t + \Big(F_t^\top + a^{[2]}(t)\Xi(t) + 2\mathbf{1}_{(d,0)}\Sigma(t)\Big)a^{[1]}(t) \tag{5.5}$$
$$\dot{a}^{[0]}(t) = -\mu_t^\top a^{[1]}(t) - \frac{1}{2}\mathrm{tr}\Big(a^{[2]}(t)\Sigma^2(t)\Big) + \frac{1}{2}a^{[1]}(t)^\top \Xi(t) a^{[1]}(t) \tag{5.6}$$

with terminal conditions

$$a^{[2]}(T) = a^{[1]}(T) = a^{[0]}(T) = 0 \ . \tag{5.7}$$

Suppose that the above ODEs have a bounded solution for $a^{[2]}$ (and hence also for $a^{[1]}$ and $a^{[0]}$) for a given time interval $[0,T]$. Then, the solution of the BSDE (5.2) is given by

$$V_L(t) = \frac{1}{2}\hat{z}_t^\top a^{[2]}(t)\hat{z}_t + a^{[1]}(t)^\top \hat{z}_t + a^{[0]}(t) \tag{5.8}$$
$$\begin{pmatrix} Z_L(t) \\ \Gamma_L(t) \end{pmatrix} = \Sigma(t)\Big(a^{[1]}(t) + a^{[2]}(t)\hat{z}_t\Big) \tag{5.9}$$

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

Proof: Consistency between (5.8) and (5.9) can be checked easily by Itô-formula. One can match the dynamics of V_L implied by (5.9) and (5.2), and the dynamics obtained from Itô-formula applied to the hypothesized solution (5.8). See Section 5 of [9] for detailed calculation.

The ODE for $a^{[2]}$ given in (5.4) is a Riccati matrix differential equation. Because of the quadratic term, the existence of bounded solution is not guaranteed and it may possibly blow up in finite time. The sufficient conditions for a bounded solution for an arbitrary time interval can be found, for example, in [14, 17]. In our setting, it requires $\Xi(t)$ to be always negative semidefinite for $t \in [0,T]$, which is not satisfied unfortunately. However, it is clear that the solutions remain finite in a short enough interval $[t,T]$ because of the continuity of the ODE. Furthermore, since $\Xi(t)$ has the order of $\mathcal{O}(\Sigma(t)^2)$, where Σ is the covariance of the signal processes (θ, α), it is naturally expected to be quite small. As long as $\int_t^T \Big|\Xi(s)\Big| ds \ll \mathcal{O}(1)$, we can expect a bounded solution. Although we may not have a bounded solution if the risk-premium processes have very large volatilities, but then, a sensible fund manager is likely to avoid using those instruments for his/her hedging in the first place. Since one can easily analyze the ODEs numerically in $(a^{[2]} \to a^{[1]} \to a^{[0]})$ order, one can directly check if the condition is satisfied in any case.

Assumption (A6)

There exists a bounded solution of $(a^{[2]}, a^{[1]}, a^{[0]})$ *for the relevant time interval* $[0,T]$.

For the case where S itself follows a jump process or more generally a semimartingale, see a recent work by Jeanblanc et.al.(2012) [16] and the references therein. They have shown that we can still characterize the optimal strategy in terms of the three BSDEs. Unfortunately though, the BSDE for V_2 becomes much more complicated and its solution is not yet known except very simplistic examples.

5.2 V_1 and the optimal hedging strategy

In a differential form, the BSDE for V_1 in (4.16) is given by

$$\begin{aligned} dV_1(t) &= \big[||\hat{\theta}_t||^2 + Z_L(t)^\top \hat{\theta}_t\big] V_1(t)dt + e^{V_L(t)}\big[\kappa_t Q_t + e_t \hat{\lambda}_t^A - g_t \hat{\lambda}_t^D \mathbf{1}_{\{Q_{t-}>0\}}\big]dt \\ &+ Z_1(t)^\top \Big(dN_t + \big[Z_L(t) + \hat{\theta}_t\big]dt\Big) + \Gamma_1(t)^\top dM_t + J_1^A(t) d\hat{A}_t + J_1^D(t) d\hat{D}_t \end{aligned} \tag{5.10}$$

with the terminal condition $V_1(T) = H$. Now, let us define

$$\begin{aligned} \xi_t^{\mathcal{A}} &:= 1 - \int_0^t \xi_s^{\mathcal{A}} \big[Z_L(s) + \hat{\theta}_s\big]^\top dN_s \\ &= \exp\left(-\int_0^t \big[Z_L(s) + \hat{\theta}_s\big]^\top dN_s - \frac{1}{2}\int_0^t ||Z_L(s) + \hat{\theta}_s||^2 ds\right) . \end{aligned} \tag{5.11}$$

By Lemma 3.9 in [1], $\{\xi_t^{\mathcal{A}}, 0 \le t \le T\}$ is a true $(\mathbb{P}, \mathbb{G})$-martingale. Thus, we can define a probability measure $\mathbb{P}^{\mathcal{A}}$ equivalent to $\mathbb{P}$ on $(\Omega, \mathcal{G})$ by

$$\frac{d\mathbb{P}^{\mathcal{A}}}{d\mathbb{P}}\bigg|_{\mathcal{G}_t} = \xi_t^{\mathcal{A}} . \tag{5.12}$$

By Girsanov-Maruyama theorem,

$$N_t^{\mathcal{A}} := N_t + \int_0^t \big[Z_L(s) + \hat{\theta}_s\big] ds \tag{5.13}$$

and M form the standard $(\mathbb{P}^{\mathcal{A}}, \mathbb{G})$-Brownian motions. Although

$$\begin{aligned}
\hat{A}_t &= A_t - \int_0^t \hat{\lambda}_s^A ds \\
\hat{D}_t &= D_t - \int_0^t \hat{\lambda}_s^D \mathbf{1}_{\{Q_{s-}>0\}} ds
\end{aligned} \tag{5.14}$$

remain $(\mathbb{P}^{\mathcal{A}}, \mathbb{G})$-martingales, their intensities are changed indirectly through the dependence on (S, Y).

Then, one can easily evaluate V_1 as

Lemma 11 *V_1 is given by*

$$\begin{aligned}
V_1(t) = \mathbb{E}^{\mathcal{A}} \Bigg[& e^{-\int_t^T \eta_s ds} H\Big(S_u, Y_u, A_u, D_u; 0 \le u \le T\Big) \\
& - \int_t^T e^{-\int_t^s \eta_u du} \Big(\kappa_s Q_s + e_s \hat{\lambda}_s^A - g_s \hat{\lambda}_s^D \mathbf{1}_{\{Q_{s-}>0\}}\Big) V_2(s) ds \Big| \mathcal{G}_t \Bigg]
\end{aligned} \tag{5.15}$$

if the expectation in the right hand side exists. Here, $\mathbb{E}^{\mathcal{A}}[\ ]$ denotes the expectation under the measure $\mathbb{P}^{\mathcal{A}}$, and $\eta_s := ||\hat{\theta}_s||^2 + Z_L(s)^\top \hat{\theta}_s$.

Thus, the evaluation of V_1 is essentially equivalent to the pricing of an European contingent claim H with an intermediate cash-flow stream. In the measure $(\mathbb{P}^{\mathcal{A}}, \mathbb{G})$, the dynamics of the underlyings are

$$dS_t = \sigma(t, S_t, Y_t)\Big(dN_t^{\mathcal{A}} - Z_L(t)dt\Big) \tag{5.16}$$

$$dY_t = \bar{\sigma}(t, S_t, Y_t)\Big(dN_t^{\mathcal{A}} - Z_L(t)dt\Big) + \rho(t, S_t, Y_t)\Big(dM_t + \hat{\alpha}_t dt\Big) \tag{5.17}$$

$$d\hat{z}_t = \Big(\mu_t - F_t \hat{z}_t - \Sigma_d(t)^\top \big[Z_L(t) + \hat{\theta}_t\big]\Big)dt + \Sigma(t) d \begin{pmatrix} N_t^{\mathcal{A}} \\ M_t \end{pmatrix} \tag{5.18}$$

and (A, D) are counting processes with intensity $(\hat{\lambda}^A, \hat{\lambda}^D)$, which are, in turn, determined by q. The procedures to run q and these counting processes are given in Remark 3. Assuming $V_1(t)$ depends smoothly on the underlyings, it is easy to see

$$\begin{aligned}
\big[Z_1(t)\big]_j &= \sum_{i=1}^d \frac{\partial V_1(t)}{\partial S_i(t)} \big[\sigma(t, S_t, Y_t)\big]_{i,j} + \sum_{i=d+1}^n \frac{\partial V_1(t)}{\partial Y_i(t)} \big[\bar{\sigma}(t, S_t, Y_t)\big]_{i,j} \\
& \quad + \sum_{i=1}^n \frac{\partial V_1(t)}{\partial \hat{z}_i(t)} \big[\Sigma(t)\big]_{i,j}, \qquad 1 \le j \le d,
\end{aligned} \tag{5.19}$$

which is the sum of the delta sensitivity with respect to each $\mathbb{G}$-adapted diffusion process multiplied by its volatility function. One also obtains J_1^A and J_1^D as

$$\begin{aligned} J_1^A(t) &= V_1(t-; A_{t-}+1) - V_1(t-) \\ J_1^D(t) &= \Big[V_1(t-; D_{t-}+1) - V_1(t-)\Big]\mathbf{1}_{\{Q_{t-}>0\}} \end{aligned} \tag{5.20}$$

where the first term is calculated by shifting the initial value of A (D) by 1, respectively.

Therefore, the pair of (V_1, Z_1) can be estimated by using the standard Monte Carlo simulations. Combining the solution of (V_2, Z_2) obtained by the ODEs and the current value of wealth, one can completely specify the optimal hedging position π^* from (4.18). Several numerical examples are available in [9] although intermediate cash flows are not included.

6 Evaluation of V_0

Since V_0 follows a linear BSDE, it is easy to see the following:

Lemma 12 *V_0 is given by*

$$\begin{aligned} V_0(t) &= \mathbb{E}\Bigg[H^2 - \int_t^T \Big\{\frac{||Z_1(s)+V_1(s)\hat{\theta}_s||^2}{V_2(s)} + 2\kappa_s Q_s V_1(s)\Big\}ds \\ &\quad + \int_t^T \Big\{e_s^2 V_2(s) - 2e_s\big(J_1^A(s)+V_1(s)\big)\Big\}\hat{\lambda}_s^A ds \\ &\quad + \int_t^T \Big\{g_s^2 V_2(s) + 2g_s\big(J_1^D(s)+V_1(s)\big)\Big\}\hat{\lambda}_s^D \mathbf{1}_{\{Q_{s-}>0\}} ds \,\Bigg|\mathcal{G}_t\Bigg] , \end{aligned} \tag{6.1}$$

if the expectation in the right hand side exists.

The difficulty in the evaluation of V_0 is quite similar to that of CVA (Credit Valuation Adjustment), where we need to evaluate V_1 (and its martingale coefficients) in each path and at each point of time. Naive application of nested Monte Carlo simulations would be too time-consuming for the practical use. The most straightforward way is to use the *least square regression method* (LSM). If $(\kappa, e, g, \hat{\lambda}^A, \hat{\lambda}^D)$ and H included in V_1 given in (5.15) have Markovian properties with respect to $(S, Y, A, D, \hat{z}, q)$, one can write V_1 as

$$V_1(t) = f(t, S_t, Y_t, A_t, D_t, \hat{z}_t, q_t) \tag{6.2}$$

with an appropriate measurable function f. Here, it is important to include $\hat{z}$ and q to recover the Markovian property. The function f is usually approximated by a polynomial function and the associated coefficients are regressed so that the square difference from the simulated V_1 is minimized. Once the estimated function f is given, the evaluation of (V_1, Z_1, J_1^A, J_1^D) in each path is straightforward. See [21] and Section 8.6 in [13] for details on LSM.

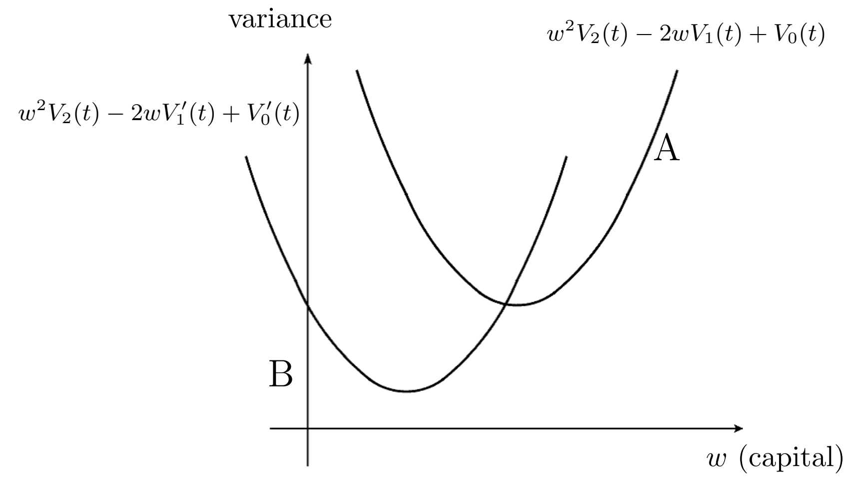

Figure 1: An example of value functions for two different service charges.

Although V_0 is unnecessary for getting the optimal hedging strategy π^*, we need it to obtain the full value function $V(t,w) = w^2V_2(t) - 2wV_1(t) + V_0(t)$. Notice that the value function $V(t,w)$ can provide valuable information to choose a profitable service-charge policy represented by (κ, e, g). For example, consider the situation given in Figure 1, where the value functions for two different cases (distinguished by V_i') of the service charges are given. Note that V_2 remains the same since it is independent from (κ, e, g). In this example, the case B is definitely better than the case A since it achieves a smaller hedging error with a smaller initial capital. If one allows (λ^A, λ^D) to depend explicitly on (κ, e, g), based on some empirical analysis for example, one can use the information of $V(t,w)$ to achieve desirable intensities of investment flows.

7 The optimal hedging for an insurance portfolio

7.1 Setup

In this section, we consider a possible extension of the framework to handle the hedging problem for an insurance portfolio. For recent applications of the mean-variance criterion for life and non-life insurance, see [4, 5] and references therein. See [29] for a general review on various control problems for the insurance industry. We shall show that one can work in a more realistic framework with imperfect information based on the method developed in the previous sections.

For the underlyings (S, Y, A, D) as well as (θ, α, X), we assume the same dynamics and the observability given in Section 2 [5]. In addition to these processes, we introduce a random measure $\mathcal{N}(dt \times dx)$. The random measure $\mathcal{N}(dt \times dx)$, which describes the occurrence of loss event and its size, is assumed to be observable to the fund manager.

[5] As mentioned before, $(\sigma, \bar{\sigma}, \rho)$ can be dependent on the past history of $(A, D, \mathcal{N})$ as long as they satisfy the listed Assumptions.

The cumulative loss process to the fund is given by

$$\int_0^t \int_K l(s,x)\mathcal{N}(ds \times dx), \tag{7.1}$$

where $K \subset (c,\infty)$ is a compact support for the jump size distribution and c (> 0) is a positive constant. $l(s,x)$ is introduced to represent the payment amount to the insured for a given loss x at time s. It can denote the minimum and/or maximum threshold, or the necessary triggers to be satisfied for the payment to the insured to occur.

We assume, for simplicity, that there is no simultaneous jump among $(A, D, \mathcal{N})$. In the current setup, the observable filtration $\mathbb{G}$ is generated by $(S, Y, A, D, \mathcal{N})$. We assume that $\{l(s,x), 0 \leq s \leq T\}$ is a $\mathbb{G}$-predictable process for any $x \in K$. $\{Y\}^{\text{obs}}$ may represent, for example, various weather related variables such as the strength of the wind, atmospheric pressure, the amount of rainfall for the insurance-covered region for non-life insurance. For life insurance, $\{Y\}^{\text{obs}}$ can contain various indexes of individual health information aggregated at a portfolio level. If the insurance portfolio contains various protections written on quite different perils, covered regions or diseases, it should be better to model each of them separately to achieve a more accurate description. For this issue, we shall discuss an extension in Section 8.

We assume that the compensated random measure in $(\mathbb{P}, \mathbb{F})$ is given by

$$\check{\mathcal{N}}(dt \times dx) = \mathcal{N}(dt \times dx) - \nu_t(x)\lambda^{\mathcal{N}}(t, X_{t-})\mathbf{1}_{\{Q_{t-}>0\}}dxdt \ . \tag{7.2}$$

Here $\lambda^{\mathcal{N}}$ is the intensity of the event occurrence, $\nu_t(\cdot)$ is the density function of the loss given the occurrence of an insured event, and it is assumed to have the compact support K for every $t \in [0,T]$. The indicator function $\mathbf{1}_{\{Q_{t-}>0\}}$ guarantees that no insured event occurs when there is no outstanding contract. The inclusion of the indicator is important to obtain the correct result for the filtering. $\lambda^{\mathcal{N}}$ is assumed to satisfy the same conditions as (λ^A, λ^D) given in Assumption (A2) and modulated by the unobservable Markov-chain process X. If there is no strong bias among the insured, one can naturally expect that $\lambda^{\mathcal{N}}$ is roughly proportional to Q.

Let us make the following assumption with regard to the density function ν:

$\{\nu_t(x), 0 \leq t \leq T\}$ is a strictly positive $\mathbb{G}$-predictable process for every $x \in K$.

Because of this assumption, the observations regarding the size of loss cannot provide any additional information on the unobservable processes, (θ, α) and X. Although it seems very hard to treat a generic situation of imperfect information, we shall discuss an extension in Section 7.4 to address the issue in a practical way.

For convenience, let us define the counting process for the insured events:

$$C_t := \sum_{u \in (0,t]} \mathbf{1}_{\{\int_K \mathcal{N}(du \times dx) \neq 0\}} \ . \tag{7.3}$$

We have $\mathbb{E}[C_T] < \infty$ due to the assumption on $\lambda^{\mathcal{N}}$. The process

$$\check{C}_t = C_t - \int_0^t \lambda^{\mathcal{N}}(s, X_{s-})\mathbf{1}_{\{Q_{s-}>0\}}ds \tag{7.4}$$

is a $(\mathbb{P}, \mathbb{F})$-martingale. If the provided insurance contract is such that it terminates when an insured event occurs (such as life insurance), we can model it easily by redefining the number of contracts as $Q_t = Q_0 + A_t - C_t - D_t$, which is a Queueing system with two exits.

7.2 Filtering

Due to the assumption on $\lambda^{\mathcal{N}}$ and ν, one can see that the filtering for the risk-premium process (θ, α) is unaffected by the observation of $\mathcal{N}$. In particular, Lemma 4 holds also in the current case. As a result, the filtered risk-premium process $\hat{z}$ has the same dynamics given in (3.30).

Let us now derive the filtering equation for X. This can be done by defining the measure $\widetilde{\mathbb{P}}_2$ by the new process

$$\widetilde{\xi}_{2,t} = 1 + \int_0^t \widetilde{\xi}_{2,s-} \left(\frac{1}{\lambda^A(s, X_{s-})} - 1 \right) d\breve{A}_s + \int_0^t \widetilde{\xi}_{2,s-} \left(\frac{1}{\lambda^D(s, X_{s-})} - 1 \right) d\breve{D}_s + \int_0^t \widetilde{\xi}_{2,s-} \left(\frac{1}{\lambda^{\mathcal{N}}(s, X_{s-})} - 1 \right) d\breve{C}_s \tag{7.5}$$

instead of (2.9). We assume that $\widetilde{\xi}_2$ is a true $(\mathbb{P}, \mathbb{F})$-martingale so that we can justify the measure change: $d\widetilde{\mathbb{P}}_2/d\mathbb{P}\Big|_{\mathcal{F}_t} = \widetilde{\xi}_{2,t}$. Then, in addition to $(\widetilde{A}, \widetilde{D})$ given in (2.15) and (2.16), we have

$$\widetilde{C}_t = C_t - \int_0^t \mathbf{1}_{\{Q_{s-}>0\}} ds \tag{7.6}$$

as a $(\widetilde{\mathbb{P}}_2, \mathbb{F})$-martingale. The inverse process $\xi_{2,t} := 1/\widetilde{\xi}_{2,t}$ is given by

$$\xi_{2,t} = 1 + \int_0^t \xi_{2,s-} \big(\lambda^A(s, X_{s-}) - 1\big) d\widetilde{A}_s + \int_0^t \xi_{2,s-} \big(\lambda^D(s, X_{s-}) - 1\big) d\widetilde{D}_s + \int_0^t \xi_{2,s-} \big(\lambda^{\mathcal{N}}(s, X_{s-}) - 1\big) d\widetilde{C}_s \tag{7.7}$$

instead of (3.16).

One can confirm that Lemma 3 holds in the current setup due to the assumption that ν_t is $\mathbb{G}$-predictable process and the fact that (A, D, C, Q) are completely decoupled from the market in measure $\widetilde{\mathbb{P}}_2$. Thus, the unnormalized filter $q_t := \mathbb{E}^{\widetilde{\mathbb{P}}_2}[\xi_{2,t} X_t | \mathcal{G}_t]$ follows

$$q_t = q_0 + \int_0^t R_s q_{s-} ds + \int_0^t \big(\Lambda_s^A - \mathbb{I}\big) q_{s-} d\widetilde{A}_s + \int_0^t \big(\Lambda_s^D - \mathbb{I}\big) q_{s-} d\widetilde{D}_s + \int_0^t \big(\Lambda_s^{\mathcal{N}} - \mathbb{I}\big) q_{s-} d\widetilde{C}_s \tag{7.8}$$

as in Lemma 7. $\Lambda^{\mathcal{N}}_{\cdot} = \mathrm{diag}\Big(\lambda^{\mathcal{N}}(\cdot, \vec{e}_1), \cdots, \lambda^{\mathcal{N}}(\cdot, \vec{e}_N)\Big)$ is a $\mathbb{G}$-predictable process similarly defined as Λ^A and Λ^D. The filtered processes, $\hat{\lambda}^A$, $\hat{\lambda}^D$ and $\hat{\lambda}^{\mathcal{N}}$ can be simulated by using q as explained in Remark 3. For later use, let us give the compensated random measure

$\hat{\mathcal{N}}$ in $(\mathbb{P}, \mathbb{G})$:

$$\hat{\mathcal{N}}(dt \times dx) = \mathcal{N}(dt \times dx) - \nu_t(x)\hat{\lambda}_t^{\mathcal{N}} \mathbf{1}_{\{Q_{t-}>0\}} dxdt \ . \tag{7.9}$$

7.3 The optimal hedging

Let us suppose that the fund manager of the insurance portfolio wants to minimize the quadratic hedging error

$$V(t, w) = \operatorname{ess}\inf_{\pi \in \Pi} \mathbb{E}\left[\left(H - \mathcal{W}_T^\pi(t, w)\right)^2 \Big| \mathcal{G}_t\right] \tag{7.10}$$

as before. Here, the wealth process $\mathcal{W}^\pi$ is defined by

$$\begin{aligned} \mathcal{W}_t^\pi(s, w) &= w + \int_s^t \pi_u^\top dS_u + \int_s^t \kappa_u Q_u du \\ &\quad + \int_s^t e_u dA_u - \int_s^t g_u dD_u - \int_s^t \int_K l(u, x)\mathcal{N}(du \times dx) \ , \end{aligned} \tag{7.11}$$

with the payout to the insured described by the last term. We assume the necessary square integrability as before

$$\mathbb{E}\left[|H|^2 + \int_0^T \left(|\kappa_u|^2 Q_u^2 + |e_u|^2 \lambda_u^A + |g_u|^2 \lambda_u^D + \Big[\int_K l(u, x)^2 \nu_u(x) dx\Big] \lambda_u^{\mathcal{N}}\right) du\right] < \infty \ . \tag{7.12}$$

We suppose that Assumption (A3) still holds true with the new definition of $\widetilde{\xi}_2$ in (7.5). We then assume the following modification of (A5).

Assumption (A5)$'$

Every $(\widetilde{P}, \mathbb{G})$-local martingale $\widetilde{m} = (\widetilde{m}_t)_{t \geq 0}$ has the integral form

$$\widetilde{m}_t = \widetilde{m}_0 + \int_0^t \widetilde{\phi}_s^\top d\widetilde{w}_s + \int_0^t \widetilde{J}_s^A d\widetilde{A}_s + \int_0^t \widetilde{J}_s^D d\widetilde{D}_s + \int_0^t \int_K \widetilde{J}^{\mathcal{N}}(s, x)\widetilde{\mathcal{N}}(ds, dx) \tag{7.13}$$

with appropriate $\mathbb{G}$-predictable coefficients $(\widetilde{\phi}, \widetilde{J}^A, \widetilde{J}^D, \widetilde{J}^{\mathcal{N}})$.

Then, following the same arguments in Lemma 8, we can decompose the value function as

$$\begin{aligned} V(t, w) &= V(0, w) + \int_0^t a(u, w) du + \int_0^t Z(u, w)^\top dN_u + \int_0^t \Gamma(u, w)^\top dM_u \\ &\quad + \int_0^t J^A(u, w) d\hat{A}_u + \int_0^t J^D(u, w) d\hat{D}_u + \int_0^t \int_K J^{\mathcal{N}}(u, w, x)\hat{\mathcal{N}}(du \times dx) \end{aligned} \tag{7.14}$$

with appropriate $\mathbb{G}$-predictable coefficients, $(a, Z, \Gamma, J^A, J^D, J^{\mathcal{N}})$. We apply Itô-Ventzell formula given in [24] to derive the dynamics of $V(t, \mathcal{W}_t^\pi)$.

For *the optimality principle*, the condition for the drift term

$$
\begin{aligned}
&a(t,w)+\inf_{\pi\in\Pi}\left\{\frac{1}{2}V_{ww}(t,w)\Big|\Big|\sigma_t^\top\pi_t+\frac{[Z_w(t,w)+V_w(t,w)\hat{\theta}_t]}{V_{ww}(t,w)}\Big|\Big|^2-\frac{||Z_w(t,w)+V_w(t,w)\hat{\theta}_t||^2}{2V_{ww}(t,w)}\right\}\\
&\quad+V_w(t,w)\kappa_tQ_t+\Big[J^A(t,w+e_t)-J^A(t,w)+V(t,w+e_t)-V(t,w)\Big]\hat{\lambda}_t^A\\
&\quad+\Big[J^D(t,w-g_t)-J^D(t,w)+V(t,w-g_t)-V(t,w)\Big]\hat{\lambda}_t^D\mathbf{1}_{\{Q_{t-}>0\}}\\
&\quad+\int_K\Big[J^{\mathcal{N}}\big(t,w-l(t,x),x\big)-J^{\mathcal{N}}(t,w,x)\\
&\qquad\qquad+V\big(t,w-l(t,x)\big)-V(t,w)\Big]\nu_t(x)\hat{\lambda}_t^{\mathcal{N}}\mathbf{1}_{\{Q_{t-}>0\}}dx=0
\end{aligned}
\tag{7.15}
$$

needs to be satisfied. Considering a quadratic form in w, $J^{\mathcal{N}}(t,w,x)=w^2J_2^{\mathcal{N}}(t,x)-2wJ_1^{\mathcal{N}}(t,x)+J_0^{\mathcal{N}}(t,x)$ in addition to those given in (4.14), one can show that the resultant BSPDE can be decomposed into three w-independent BSDEs also in this case.

One can check that the formula of π^* is unchanged and given by (4.18). After the straightforward calculation, one obtains the same BSDE for V_2 as in (4.15), and hence (V_2,Z_2) can be solved by the same ODEs given in Lemma 10. The BSDEs for V_1 and V_0 can be found as follows:

$$
\begin{aligned}
V_1(t)&=H-\int_t^T\frac{[Z_2(s)+V_2(s)\hat{\theta}_s]^\top[Z_1(s)+V_1(s)\hat{\theta}_s]}{V_2(s)}ds\\
&\quad-\int_t^T\Big\{\kappa_sQ_s+e_s\hat{\lambda}_s^A-g_s\hat{\lambda}_s^D\mathbf{1}_{\{Q_{s-}>0\}}-\bar{L}_s\hat{\lambda}_s^{\mathcal{N}}\Big\}V_2(s)ds\\
&\quad-\int_t^TZ_1(s)^\top dN_s-\int_t^T\Gamma_1(s)^\top dM_s-\int_t^TJ_1^A(s)d\hat{A}_s-\int_t^TJ_1^D(s)d\hat{D}_s\\
&\quad-\int_t^TJ_1^{\mathcal{N}}(s,x)\hat{\mathcal{N}}(ds\times dx)\ ,
\end{aligned}
\tag{7.16}
$$

$$
\begin{aligned}
V_0(t)&=H^2-\int_t^T\left\{\frac{||Z_1(s)+V_1(s)\hat{\theta}_s||^2}{V_2(s)}+2\kappa_sQ_sV_1(s)\right\}ds\\
&\quad+\int_t^T\Big[e_s^2V_2(s)-2e_s\big(J_1^A(s)+V_1(s)\big)\Big]\hat{\lambda}_s^Ads\\
&\quad+\int_t^T\Big[g_s^2V_2(s)+2g_s\big(J_1^D(s)+V_1(s)\big)\Big]\hat{\lambda}_s^D\mathbf{1}_{\{Q_{s-}>0\}}ds\\
&\quad+\int_t^T\int_K\Big\{l(s,x)^2V_2(s)+2l(s,x)\Big(J_1^{\mathcal{N}}(s,x)+V_1(s)\Big)\Big\}\nu_s(x)\hat{\lambda}_s^{\mathcal{N}}dxds\\
&\quad-\int_t^TZ_0(s)^\top dN_s-\int_t^T\Gamma_0(s)^\top dM_s-\int_t^TJ_0^A(s)d\hat{A}_s-\int_t^TJ_0^D(s)d\hat{D}_s\\
&\quad-\int_t^T\int_KJ_0^{\mathcal{N}}(s,x)\hat{\mathcal{N}}(ds\times dx)\ .
\end{aligned}
\tag{7.17}
$$

Since the both BSDEs are linear, it is easy to solve them under the appropriate integrability conditions. In particular, one can use $\mathbb{P}^{\mathcal{A}}$ defined by (5.12) for V_1. The hedging strategy

π^* can be evaluated by the same procedures discussed in Section 5.2. For the numerical evaluation of V_0, we need

$$J_1^{\mathcal{N}}(t,x) = \Big(V_1(t-;x) - V_1(t-)\Big)\,\mathbf{1}_{\{Q_{t-}>0\}}\ , \tag{7.18}$$

where the first term represents the value in the presence of a jump with the size of x.

7.4 Introducing multiple grades of the loss severity

For insurance contracts, the hidden process X may represent various uncertainties involved in the loss-event modeling, which is updated based on each actual occurrence of an insured event. If the hidden process X is shared among $(\lambda^{\mathcal{N}}, \lambda^A, \lambda^D)$ in a nontrivial fashion, an actual occurrence (or non-occurrence) of peril is reflected by the change of $\hat{X}$, which then can induce a jump to the higher (or lower) demand for the insurance contract. These "contagious" behaviors of insurance buyers are expected to be more profound after a catastrophe which caused a significant loss to the human lives and property.

In the previous setup, we have treated every insured event equally and cannot take into account the *size effect* explained above. This problem is arising from the assumption that ν_t is $\mathcal{G}_{t-}$-measurable, which makes the size of loss unable to carry the information on X. Here, we explain a simple modeling scheme to address the issue in a practical manner:

(1)Introduce n_g independent random measures with disjoint supports for the density functions of the jump size, $\big\{(\mathcal{N}_j, \lambda^{\mathcal{N}_j}, \nu_j), j = \{1,\cdots,n_g\}\big\}$. (2)Interpret the jump in $\mathcal{N}_j$ as the occurrence of an insured event "with grade j severity" and arrange the support K_j of the density function ν_j with $1 \le j \le n_g$ accordingly. Here, each $\nu_j(x)$ is assumed to be a $\mathbb{G}$-predictable process as before. (3)Introduce X with the total number of states $N = n_f \times (n_g+1)$, which is specified by a double-index (i,j). (4)Assume $\lambda^{\mathcal{N}_k}(t, X_{t-})$ has sensitivity mainly on the states (i,j) with $j \simeq k$. The states $\{(i,0)\}$ are intended to describe the most relaxed environment. (5)Make $\big(\lambda^A(t,X_{t-}), \lambda^D(t,X_{t-})\big)$ sensitive more profoundly to the second index. (6)Arrange the transition matrix R_t so that it induces an appropriate speed of mean reversion to the calmer states.

In this way, one can at least differentiate the grades of the loss. It is straightforward to obtain the corresponding filtering equations and the BSDEs. The unnormalized filter q now follows:

$$\begin{aligned} q_t = q_0 &+ \int_0^t R_s q_{s-} ds + \int_0^t \big(\Lambda_s^A - \mathbb{I}\big) q_{s-} d\widetilde{A}_s + \int_0^t \big(\Lambda_s^D - \mathbb{I}\big) q_{s-} d\widetilde{D}_s \\ &+ \sum_{i=1}^{n_g} \int_0^t \big(\Lambda_s^{\mathcal{N}_i} - \mathbb{I}\big) q_{s-} d\widetilde{C}_{i,s} \end{aligned} \tag{7.19}$$

with obvious definitions. π^* is still given by (4.18) and the solution for (V_2, Z_2) is also unchanged. It is straightforward to see

$$V_1(t) = \mathbb{E}^{\mathcal{A}}\Bigg[e^{-\int_t^T \eta_s ds} H - \int_t^T e^{-\int_t^s \eta_u du}\Big\{ \kappa_s Q_s + e_s \hat{\lambda}_s^A - g_s \hat{\lambda}_s^D \mathbf{1}_{\{Q_{s-}>0\}} - \sum_{i=1}^{n_g} \bar{L}_{i,s} \hat{\lambda}_s^{\mathcal{N}_i} \Big\} V_2(s) ds \Big| \mathcal{G}_t \Bigg] \tag{7.20}$$

with $\bar{L}_{i,s} := \int_{K_i} l(s,x)\nu_i(x)dx$. The derivation of V_0 is simple and left for the interested readers.

Remark 5 : For the fund management, the same idea can be used to extend the modeling of the counting processes (A_t, D_t) to integer-valued random measures. By introducing $(A_t^i, D_t^i)_{1\le i\le n_g}$, one can treat the case where the inflow and outflow can jump by multiple units and differentiate the importance of information by the grades of the jump size. By making use of the $\mathbb{G}$-predictable jump distribution function for each (A^i, D^i), the filtering equations are reduced to those for the counting processes.

8 Application of Jackson's network

8.1 Setup

Asset management firms and insurers provide a wide choice of funds and insurance products. It is also rather popular to provide a financial product that consists of a set of funds among which investors can change (or switch) a fund to put their money on. Thus, the fund manager can access a large amount of information about the investment flows within the regulatory restrictions, and ultimately wants to implement the optimal hedging strategy and service-charge policy at this broader level. In particular, there is a need for the fund manager to be well prepared for the switching activities between the two extremes, such as (Bull-Bear) or (Equity-Bond), which easily incur the over- as well as under-hedging. Also, even if they are the inflows to the same fund, an investment from a new external client and the one from an existing client as an extension may carry quite different information.

In order to handle these situations, we make use of the Jackson's network typically used in the analysis of a Queueing system. See Section V.2 in [2] for detail. In addition to the same diffusion processes (S, Y, θ, α) and the hidden Markov-chain X, we introduce n_p funds/insurance products and the associated investment flows given in Figure 2 (for a case with two funds). The definition of each flow is given as follows:

$A_t(i)$: *The external inflow to the i-th fund.*
$D_t(i)$: *The unwind from the i-th fund.*
$F_t(i,j)$: *The switching from the i-th to the j-th fund.*
$F_t(i,i)$: *The extension of investments in the i-th fund.*
$A_t^*(i)$: *The total inflow to the i-th fund.*
$D_t^*(i)$: *The total outflow from the i-th fund.*

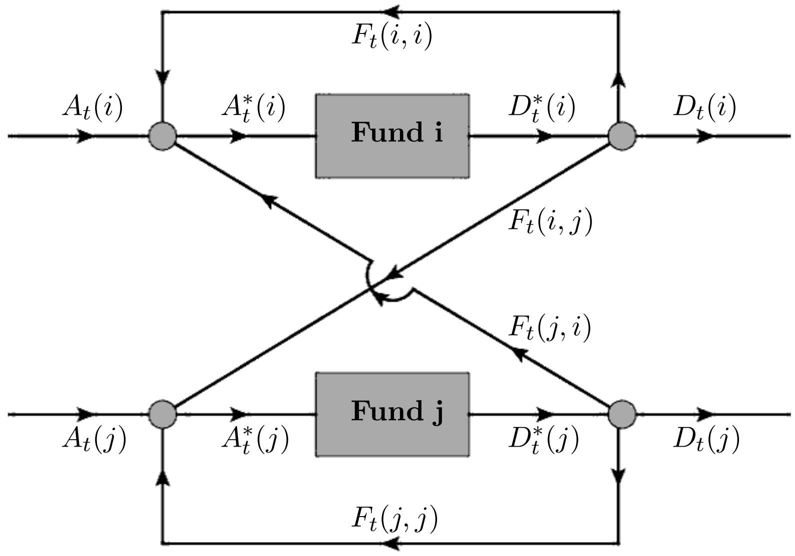

Figure 2: Jackson's network of investment flows: 2-fund's case

The following relations should be obvious

$$A^*_t(i) = A_t(i) + \sum_{j=1}^{n_p} F_t(j,i) \tag{8.1}$$

$$D^*_t(i) = D_t(i) + \sum_{j=1}^{n_p} F_t(i,j) \ . \tag{8.2}$$

Thus, the outstanding number of investment-units in the i-the fund at time t is given by

$$\begin{aligned} Q_t(i) &= Q_0(i) + A^*_t(i) - D^*_t(i) \\ &= Q_0(i) + A_t(i) - D_t(i) + \sum_{j=1}^{n_p} \Big(F_t(j,i) - F_t(i,j) \Big) \ . \end{aligned} \tag{8.3}$$

Here, all of the (A, D, F) are assumed to be the counting processes with no simultaneous jump. The associated compensated processes in $(\mathbb{P}, \mathbb{F})$ are given by

$$\check{A}_t(i) = A_t(i) - \int_0^t \lambda^A(i)(s, X_{s-}) ds \tag{8.4}$$

$$\check{D}_t(i) = D_t(i) - \int_0^t \lambda^D(i)(s, X_{s-}) \mathbf{1}_{\{Q_{s-}(i)>0\}} ds \tag{8.5}$$

and

$$\check{F}_t(i,j) = F_t(i,j) - \int_0^t \lambda^F(i,j)(s, X_{s-})\mathbf{1}_{\{Q_{s-}(i)>0\}} ds \ . \tag{8.6}$$

We also introduce n_p random measures $\{\mathcal{N}_i(dt \times dx), 1 \leq i \leq n_p\}$ to describe the occurrences of the insured events or any other contingency payouts from the corresponding fund [6]. The compensated random measure in $(\mathbb{P}, \mathbb{F})$ is given by

$$\check{\mathcal{N}}_i(dt \times dx) = \mathcal{N}_i(dt \times dx) - \nu_{i,t}(x)\lambda^{\mathcal{N}}(i)(t, X_{t-})dxdt \ , \tag{8.7}$$

where $\nu_{i,t}()$ is the density function of jump size and assumed to have a compact support $K_i \subset (c, \infty)$ with $c\ (>0)$. For convenience, we also introduce a counting process for each random measure:

$$C_t(i) = \sum_{u \in (0,t]} \mathbf{1}_{\{\int_{K_i} \mathcal{N}_i(du \times dx) \neq 0\}} \ , \tag{8.8}$$

and also the associated $(\mathbb{P}, \mathbb{F})$-compensated process

$$\check{C}_t(i) = C_t(i) - \int_0^t \lambda^{\mathcal{N}}(i)(s, X_{s-})\mathbf{1}_{\{Q_{s-}(i)>0\}} ds \ . \tag{8.9}$$

The observable filtration $\mathbb{G}$ is generated by (S, Y) and $\big(A(i), D(i), \mathcal{N}_i, F(i,j), 1 \leq i,j \leq n_p\big)$. As in Section 7, the density functions are assumed to be $\mathbb{G}$-predictable, i.e. for each $i \in \{1, \cdots, n_p\}$, $\big(\nu_{i,t}(x), 0 \leq t \leq T\big)$ is a $\mathbb{G}$-predictable process for all $x \in K_i$. We further assume that $Q_0(i) \in \mathcal{G}_0$ for all $i \in \{1, \cdots, n_p\}$ and that Assumption (A2) hold for all the relevant intensities, $(\lambda^A(i), \lambda^D(i), \lambda^F(i,j), \lambda^{\mathcal{N}}(i); 1 \leq i,j \leq n_p)$.

8.2 Filtering

It is clear that we have the same dynamics of the filtered risk-premium process $\hat{z}$ as (3.30). For the filtering of X, we define

$$\begin{aligned} \widetilde{\xi}_{2,t} = 1 &+ \sum_i \int_0^t \widetilde{\xi}_{2,s-}\left(\frac{1}{\lambda_s^A(i)} - 1\right) d\check{A}_s(i) + \sum_i \int_0^t \widetilde{\xi}_{2,s-}\left(\frac{1}{\lambda_s^D(i)} - 1\right) d\check{D}_s(i) \\ &+ \sum_{i,j} \int_0^t \widetilde{\xi}_{2,s-}\left(\frac{1}{\lambda_s^F(i,j)} - 1\right) d\check{F}_s(i,j) + \sum_i \int_0^t \widetilde{\xi}_{2,s-}\left(\frac{1}{\lambda_s^{\mathcal{N}}(i)} - 1\right) d\check{C}_s(i) \end{aligned} \tag{8.10}$$

and assume $\{\widetilde{\xi}_{2,t}, 0 \leq t \leq T\}$ is a true $(\mathbb{P}, \mathbb{F})$-martingale. We can then define an equivalent probability measure $\widetilde{\mathbb{P}}_2$ on $(\Omega, \mathcal{F})$ as (2.12). Under the measure $\widetilde{\mathbb{P}}_2$, one can see that the whole Jackson's network is completely decoupled from the external world because

$$\widetilde{A}_t(i) = A_t(i) - t \tag{8.11}$$

$$\widetilde{D}_t(i) = D_t(i) - \int_0^t \mathbf{1}_{\{Q_{s-}(i)>0\}} ds \tag{8.12}$$

[6] If necessary, one can introduce multiple grades of severity for each fund as explained in Section 7.4.

$$\widetilde{F}_t(i,j) = F_t(i,j) - \int_0^t \mathbf{1}_{\{Q_{s-}(i)>0\}} ds \tag{8.13}$$

$$\widetilde{C}_t(i) = C_t(i) - \int_0^t \mathbf{1}_{\{Q_{s-}(i)>0\}} ds \tag{8.14}$$

become $(\widetilde{\mathbb{P}}_2, \mathbb{F})$-martingales. This make Lemma 3 hold also in the current setup.

Using the $(\widetilde{\mathbb{P}}_2, \mathbb{F})$-martingale $\xi_{2,t} := 1/\widetilde{\xi}_{2,t}$, Lemma 5 and similar procedures used in Lemma 7, one obtains the dynamics of the unnormalized filter $q_t := \mathbb{E}^{\widetilde{\mathbb{P}}_2}[\xi_{2,t} X_t | \mathcal{G}_t]$:

$$\begin{aligned} q_t = q_0 + \int_0^t R_s q_{s-} ds + \sum_i \int_0^t \big(\Lambda_s^A(i) - \mathbb{I}\big) q_{s-} d\widetilde{A}_s(i) + \sum_i \int_0^t \big(\Lambda_s^D(i) - \mathbb{I}\big) q_{s-} d\widetilde{D}_s(i) \\ + \sum_{i,j} \int_0^t \big(\Lambda_s^F(i,j) - \mathbb{I}\big) q_{s-} d\widetilde{F}_s(i,j) + \sum_i \int_0^t \big(\Lambda_s^{\mathcal{N}}(i) - \mathbb{I}\big) q_{s-} d\widetilde{C}_s(i) \end{aligned} \tag{8.15}$$

where Λ's are similarly defined as in Lemma 7.

8.3 The optimal hedging

Let us suppose that the wealth process of the fund manager follows

$$\begin{aligned} \mathcal{W}_t^\pi(s,w) = w + \int_s^t \pi_u^\top dS_u + \sum_i \int_s^t \kappa_u(i) Q_u(i) du + \sum_i \int_s^t e_u(i) dA_u(i) \\ - \sum_i \int_s^t g_u(i) dD_u(i) - \sum_{i,j} \int_s^t f_u(i,j) dF_u(i,j) - \sum_i \int_s^t \int_{K_i} l_i(u,x) \mathcal{N}_i(du \times dx) \end{aligned} \tag{8.16}$$

where $f(i,j)$ denotes the cost associated with the switching from the i-th to the j-th fund, and $l_i(t,x)$ is defined as in Section 7.1 for the fund i. All the processes of coefficients $(\kappa(i), e(i), g(i), f(i,j), l_i(\cdot, x))$ are assumed to be $\mathbb{G}$-predictable and satisfy the necessary square integrability.

The fund manager's problem is to minimize the quadratic hedging error

$$V(t,w) = \operatorname*{ess\,inf}_{\pi \in \Pi} \mathbb{E}\left[\Big(H - \mathcal{W}_T^\pi(t,w) \Big)^2 \Big| \mathcal{G}_t \right] . \tag{8.17}$$

The derivation of the optimal hedging strategy π^* can be performed by a straightforward modification of those in Section 7. One can check that the BSPDE for $V(t,w)$ can still be decomposed into the three BSDEs and that the optimal hedging strategy π^* is given by the formula (4.18) with the same (V_2, Z_2) given in Lemma 10. The expressions for V_1 and V_0 can be derived easily due to their linearity as before.

9 Conclusions

In this work, the prices of securities, the occurrences of insured events and (possibly a network of) the investment flows are used to infer their drifts and intensities by a stochastic filtering technique, which are then used to determine the optimal mean-variance hedging strategy. A systematic derivation of the optimal strategy based on the BSDE approach is provided, which is also shown to be implementable by a set of simple ODEs and the standard Monte Carlo simulation.

As for the management of insurance portfolios, we have given a framework with multiple grades of loss severity, which allows a granular modeling of the change of demand for insurance products after the insured events with different sizes. We have applied the technique used in Queueing analysis to treat a complex network of the investment flows, such as those in a group of funds within which investors can switch a fund to invest.

Although a lot of problems remain unsolved especially with regard to the model specifications, the recent great developments of computer systems capable of handling the so-called *big data* and wide interests among industries in the efficient use of information may make the installation of the framework a real possibility in near future. More concrete applications to a specific product or business model using real data will be left for a future research, hopefully in a good collaboration with financial as well as non-financial institutions.

Acknowledgement

This research is partially supported by Center for Advanced Research in Finance (CARF).

References

[1] Bain, A., Crisan, D., 2009, "Fundamentals of Stochastic Filtering," Springer, NY.

[2] Brémaud, P., 1981,"Point Processes and Queues," Springer, NY.

[3] Cohen, S.N. and Elliott, R.J., 2013, "Filters and smoothers for self-exiting Markov modulated counting processes," preprint available at arXiv:1311.6257.

[4] Dahl, M. and Mϕller, T., 2006, "Valuation and hedging of life insurance liabilities with systematic mortality risk," *Insurance: Mathematics and Economics*, 39, 193-217.

[5] Delong, L. and Gerrard, R., 2007, "Mean-variance portfolio selection for a non-life insurance company," *Mathematical Methods of Operations Research*, 66, 339-367.

[6] Delong, L., 2013, "Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications," Springer, LN.

[7] Elliott, R.J. and Malcolm, W., 2005, "General Smoothing Formulas for Markov-Modulated Poisson Observations, " *IEEE Transactions on automatic control*, vol. 50, No. 8, 1123-1134.

[8] Elliott, R.J., Aggoun, L. and Moore, J.B., 2008, "Hidden Markov Models (Estimation and Control)," corrected 3rd printing, Springer, NY.

[9] Fujii, M., Takahashi, A., 2013, "Making Mean-Variance Hedging Implementable in a Partially Observable Market," forthcoming in *Quantitative Finance*.

[10] Fujii, M., Takahashi, A., 2013, "Making Mean-Variance Hedging Implementable in a Partially Observable Market-*with supplementary contents for stochastic interest rates*-," available at http://ssrn.com/abstract=2279398.

[11] Fujii, M., Takahashi, A., 2012, "Perturbative Expansion Technique for Non-linear FBSDEs with Interacting Particle Method," CARF working paper, available at at SSRN: http://ssrn.com/abstract=2038740.

[12] Fujii, M., Takahashi, A. and Sato, S., "An FBSDE approach to American Option Pricing with an Interacting Particle Method," CARF working paper series, CARF-F-302, available at http://ssrn.com/abstract=2180696.

[13] Glasserman, P., 2004, "Monte Carlo Methods in Financial Engineering," Springer, NY.

[14] Jacobson, D.H., 1970, "New conditions for boundedness of the solution of a matrix Riccati differential equation," *Journal of Differential Equations* 8, 258-263.

[15] Jacod, J. and Shiryaev, A.N., 2002, "Limit Theorems for Stochastic Processes (2nd edition)" Springer, NY.

[16] Jeanblanc, M., Mania, M., Santacroce, M. and Schweizer, M, 2012, "Mean-variance hedging via stochastic control and BSDEs for general semimartingales," *The Annals of Applied Probability*, Vol. 22, No. 6, 2388-2428.

[17] Kalman, R.E., 1960, "Contributions on the Theory of Optimal Control," Bol. Soc. Mat. Mexicana, Vol. 5, 102-119.

[18] Kobylanski, M., 2000, "Backward stochastic differential equations and partial differential equations with quadratic growth," The Annals of Probability, 28, 558-602.

[19] Kohlmann, M. and Tang, S., 2002, "Global adapted solution of one-dimensional backward stochastic Riccati equations, with application to the mean-variance hedging," *Stochastic Processes and their Applications* 97, 255-288.

[20] Kunita, H., 1990, "Stochastic flows and stochastic differential equations," Cambridge university press, UK.

[21] Longstaff, F.A., and Schwartz, E.S., 2001, "Valuing American options by simulation: a simple least-square approach," *Review of Financial Studies* 14, 113-147.

[22] Mania, M., Tevzadze, R., 2003, "Backward Stochastic PDE and Imperfect Hedging," *International Journal of Theoretical and Applied Finance*, Vol. 6, No. 7, 663-692.

[23] Mania, M. and Tevzadze, R., 2008, "Backward Stochastic Partial Differential Equations Related to Utility Maximization and Hedging, " Journal of Mathematical Science, Vol. 153, 3, 291-380.

[24] Øksendal, B. and Zhang, T., 2007, "The Itô-Ventzell formula and forward stochastic differential equations driven by Poisson random measures," *Osaka J. Math.* 44, 207-230.

[25] "Mortgage and Real Estate Finance," edited by Stefania Perrucci, 2008, Risk books, LN.

[26] Pham, H., 2001, "Mean-variance hedging for partially observed drift processes," *International Journal of Theoretical and Applied Finance*, Vol. 4, 2, 263-284.

[27] Pham, H. and Quenez, M., 2001, "Optimal Portfolio in Partially Observed Stochastic Volatility Models," *The Annals of Applied Probability*, vol. 11, No. 1, 210-238.

[28] Protter, P., 2005, "Stochastic Integration and Differential Equations-2nd edition," Springer, NY.

[29] Schmidli, H., 2008, "Stochastic Control in Insurance," Springer, LN.

[30] Schweizer, M., 1992, "Mean-variance hedging for general claims," *The Annals of Applied Probability*, 2(1), 171-179.

[31] Schweizer, M., 2010, "Mean-variance hedging" in R. Cont (ed.), "Encyclopedia of Quantitative Finance," 1177-1181, Wiley NY.

[32] Xiong, J., 2008, "An Introduction to Stochastic Filtering Theory," Oxford University press, NY.