Risk Minimization for a Filtering Micromovement Model of Asset Price*

Kiseop Lee † and Yong Zeng ‡

August 11, 2009

Abstract

The classical option hedging problems have been studied mostly under continuous-time or equally spaced discrete time models, which ignore two important components in actual price: random trading times and market microstructure noise . In this paper, we study optimal hedging strategies for European derivatives based on a filtering micromovement model of asset prices with the two commonly-ignored characteristics. We employ the local risk minimization criterion to develop optimal hedging strategies under full information. Then, we project the hedging strategies on the observed information to obtain hedging strategies under partial information. Furthermore, we develop a related nonlinear filtering technique under the minimal martingale measure for the computation of such hedging strategies.

JEL Classification: C61, G13

Key Words: risk minimization, minimal martingale measure, filtering, counting process, and high frequency data.

---

*We are grateful to Tom Kurtz for an insightful comment simplifying a representation, to an anonymous referee for constructive comments considerably improving the quality of the paper, and to Xia Chen, Tyrone Duncan, Yaozhong Hu, Jin Ma, David Nualart, Philip Protter, Balram Rajput, Jan Rosinski, Richard Stockbridge, Jie Xiong and Ziyu Zheng for helpful discussions and comments. We thank participants at the conference of "Stochastic Control and Numerics" (2005) held at University of Wisconsin at Milwaukee and at the 2006 Fall Central Section Meeting of AMS at University of Cincinnati, and thank seminar participants at Purdue University, University of Kansas, University of Southern California, and University of Tennessee for comments. Yong Zeng gratefully acknowledges the financial support for his research from the National Science Foundation under grant DMS-0604722.

†Department of Mathematics, University of Louisville, Louisville, KY 40292, USA. Tel: (502) 852 6292. Fax: (502) 852 7132. Email: kiseop.lee@louisville.edu.

‡Department of Mathematics and Statistics, University of Missouri at Kansas City, Kansas City, MO 64110, USA. Tel: (816) 235 5850. Fax: (816) 235 5517. Email: zengy@umkc.edu. Website: http://mendota.umkc.edu/.

1 Introduction

The literature on option pricing and hedging has been prolific since the seminal Black-Scholes (BS) model was published in the early 1970s. Early research focuses on generalizing the BS model to a more general setup such as diffusion models, stochastic volatility models, jump diffusion models or time changed Levy processes (Carr and Wu 2004; Huang and Wu 2004). Recent research focuses on incorporating default risk and credit risk (Duffie and Singleton 2003) or liquidity risk (Cetin, Jarrow and Protter 2006).

In another front, modeling the micromovement of asset prices has become important to the econometrics population, because micromovement more accurately captures the price behaviors at the micro level referred to as transaction or trade-by trade price behavior. Micromovement data is named *ultra-high frequency* (UHF) data by Engle (2000) because of their ultimate disaggregation nature.

UHF data have two characteristics which are different from the continuous-time models in asset pricing, or the price macromovement referred to as the equally-spaced daily, or weekly closing price behavior in the econometric literature. First, the micromovement observations occur at varying random times. Engle (2000) treats UHF data as an irregularly-spaced time series and proposes a general framework for jointly modeling time and price. He further develops a UHF-GARCH model, extending the autoregressive conditional duration model in Engle and Russell (1998). Engle and Russell (2005) provides a survey on such an approach.

The second stylized characteristic is that market microstructure noise exists in prices. In contrast with information which has a long-term impact on price, noise only has a short-term impact on price (Hasbrouck 1996). However, it is well-known that noise plays a fundamental role in asset pricing, especially when modeling UHF data. The topics of the two recent presidential addresses to the American Finance Association were "Noise" (Black 1986) and "Friction" (Stoll 2000). Both address treated noise as an essential matter. Therefore, noise must be incorporated in any suitable micromovement model.

Recently, extending the fast-growing literature on realized volatility (RV) models, Zhang, Mykland and Aït-Sahalia (2005), Aït-Sahalia, Mykland and Zhang (2005), Bandi and Russell (2006), Diebold (2006), and Fan and Wang (2007) develop different realized volatility estimators under the two-time-scale frameworks which incorporate market microstructure noises. Especially, Li and Mykland (2007) shows that rounding noise in UHF data may severely distort even the two-scale estimators of realized volatility, and the error could be infinite.

In derivative pricing literature, the general assumption under most asset pricing research is continuous-time trading. It is no surprise because of technical tractability. However, this assumption obviously contradicts the two aforementioned stylized facts of UHF data. The research for the models equipped with such characteristics is rare on option pricing and hedging. We surmise that this is because those micromovement characteristics generate difficulties in modeling, and further, generate the market incompleteness, making the hedging and option pricing problem even more difficult .

One of the few exceptions is considered by Frey (2000). In Frey's model, the security price is modeled as an exponential martingale of a doubly stochastic Poisson process, whose stochastic intensity depends on an unobservable state-variable process[1]. Clearly, Frey's model is able to capture the random arrival time nature of high frequency transaction data. Due to the presence of jumps and stochastic intensity, the market is incomplete. Frey used the risk minimization criterion proposed by Föllmer and Sondermann (1986) to determine hedging strategies. Along this line, Ceci (2006) studies risk minimizing hedging for a similar model. However, market microstructure noise, a central piece in micromovement, is missing in these two models.

Zeng (2003) proposes a general filtering micromovement model (FM model, as we simply call it) for asset price, where the two stylized facts of micromovements are incorporated in a consistent manner. Namely, the FM model accommodates not only random trading times as Frey's model, but also market microstructure noise in UHF data. In the FM model, there is an unobservable intrinsic value process for an asset, which corresponds to the usual price process in the option pricing literature. The intrinsic value process is the permanent component and has a long-term impact on price. Prices are observed only at random trading times which are driven by a conditional Poisson process, whose intensity may depend on the intrinsic value. Prices are distorted observations of the intrinsic value process at the trading times by noise, which is the transient component and only has a short-term impact (when a trade happens) on price. Financial noise is explicitly modeled by the transition probability from the intrinsic value to the price at trading time. This way is flexible enough to accommodate the rounding noise aforementioned in Li and Mykland (2007) and more complicated noise (see examples in Zeng 2003 and Spalding, Tsui and Zeng 2006).

In this setup, the price process is a semimartingale rather than a martingale under the physical measure, which is a more realistic hypothesis. One important feature of the FM model is that trade-by-trade prices can be viewed as a collection of counting processes of price levels and the model can be framed as a filtering problem with counting process observations. Then, the filtering equation characterizing the evolution of the conditional distribution of the unobservable intrinsic process is derived. The related numerical solution and the Bayes estimation via filtering for the intrinsic value process and the related parameters in the model are developed. Furthermore, Bayesian hypothesis testing or model selection via filtering for this class of models has been developed in Kouritzin and Zeng (2005).

In this paper, we study derivative pricing and hedging when making use of a FM model, where the intrinsic value process is assumed to be a diffusion process. The contingent claim is written upon the observable price process only as in the real world instead of including the unobservable intrinsic value process. The market under the FM model has jumps with only partial information and is clearly incomplete. We have to choose alternative approaches (see Schweizer 2001 for a survey) such as mean-variance hedging, local risk minimization or utility maximization, to obtain a reasonable hedge. We follow Frey (2000) to adopt the risk minimization criterion. This approach has the important benefit of offering a simple form of an optimal hedging strategy, allowing fairly explicit computations. However, unlike the

[1] We will describe it further in Section 2.1

models in Frey (2000) and Ceci (2006), the price in the FM model is a semimartingale, not a martingale. Then, the risk minimization criterion proposed by Föllmer and Sondermann (1986) (as used in Frey 2000 and Ceci 2006) does not exist. Instead, we use the *local* risk minimization criterion developed in Föllmer and Schweizer (1991), which requires a notion of *minimal martingale measure*. Föllmer and Schweizer (1991) uses a continuous semimartingale. Their results are not obviously correct if the price process has jumps (as in the FM model) because of some technical issues. Lee and Protter (2008) overcomes those technical difficulties and extends the results to the case with jumps under mild conditions. In this paper, we employ the same approach in Lee and Protter (2008) to prove that the optimal strategy can be again expressed in terms of the minimum martingale measure.

To solve the option pricing and hedging problem for the FM model, we first identify the minimal martingale measure. Then, we derive the pricing equation and the locally risk-minimizing hedging strategy with complete information. In the case of incomplete information, we project the hedging strategy under full information to the partially-observed prices and we show the computation of such hedging strategy is closely connected to the filtering technique developed in Zeng (2003). Since the price process is a semi-martingale instead of a martingale under the physical measure, the optimality of the projected hedging strategy remains open.

The outline of the paper is as follows. Section 2 briefly reviews the FM model in two equivalent representations and presents the integral (and differential) form for the evolution of the most recent price, which is crucial for option pricing. Section 3 provides a concise summary of local risk minimization theory. Section 4 studies option pricing based on the FM model under the local risk minimization and Section 5 concludes.

2 The Micromovement Models

After briefly describing Frey's model, this section focuses on presenting a sub-model of Zeng (2003), whose option pricing and hedging is considered in this paper. There are two equivalent representations of the model, described in sections 2.2 and 2.3, respectively. Section 2.4 presents an integral form of the most recent price, which plays a crucial rule in Section 4.

2.1 Frey's model

Let Z_n be the percentage return in the nth trade and $\{Z_n\}$ are i.i.d. random variables on $(-1, +\infty)$ with mean zero. Note that Z_n has a lower bound of -1, 100% loss, due to the limited liability of stocks. Let N_t be a conditional Poisson process with the conditional stochastic intensity function $a(X_t, t)$, and X_t is a stochastic factor following a diffusion: $dX_t = \mu(X_t)dt + \sigma(X_t)dB_t$ where B_t is a standard Brownian motion. Frey interprets X_t as a kind of stochastic volatility. With these notations, Frey assumes the stock price, S_t, follows

$$dS_t = S_{t-}dR_t \quad \text{with an initial price } S_0, \tag{1}$$

where $R_t = \sum_{n=1}^{N_t} Z_n$ is the return process.

In the models of Ceci (2006) or Cvitnic, Lipster and Rozovsky (2006), X_t or N_t, the point process of observation times, are more general processes. However, it is clear that all of these models do not incorporate financial noise.

2.2 Representation I: Construction of Price from Intrinsic Value

The FM model is predicated on a simple intuition that the price is formed from the intrinsic value process of an asset corrupted by trading (or market microstructure) noise. Let $X = (X_t)_{0\le t\le T}$ be the intrinsic value process. We use $Y = Y(t)$ or $\vec{Y}(t)$ for $t \in [0, T]$ to denote the price process.

Throughout the rest of Section 2 and in Section 4, we assume that (X, Y) or $(X, \vec{Y})$ is on a filtered complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, P)$ with the usual hypotheses (see, for example, Protter 2004), the σ-field is $\mathcal{F} = \mathcal{F}_T$ and P is the physical probability measure. All processes used in this paper are adapted to this filtration. We define subfiltrations $\mathcal{F}_t^{X,Y} = \sigma((X_s, Y_s) : 0 \le s \le t)$, and $\mathcal{F}_t^Y = \sigma(Y_s : 0 \le s \le t)$ for hedging and pricing purposes, which we will use in Section 4.3.

Assumption 1. *X_t, the intrinsic value process of an asset follows a diffusion:*

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t,$$

and X_t has a unique weak solution.

The infinitesimal generator of X is given by

$$\mathbf{A}f(x) = \mu(x)\frac{\partial f}{\partial x}(x) + \frac{1}{2}\sigma^2(x)\frac{\partial^2 f}{\partial x^2}(x). \tag{2}$$

Obviously, Assumption 1 includes geometric Brownian motion, used in the Black-Scholes model. The intrinsic value process can not be observed directly, but can be partially observed through the price process, Y. Due to price discreteness, Y is in a discrete state space given by the multiples of *tick*, the minimum price variation set by trading regulation. Y is a distorted observation of X at some random times.

There are three general steps in constructing Y from X.

- Step 1 : Specify $X(t)$ as in Assumption 1.
- Step 2 : Specify the trading times $t_1, t_2, \ldots, t_i, \ldots$, which are driven by a conditional Poisson process with a conditional intensity function $a(X(t), t)$.
- Step 3 : $Y(t_i)$, the price at time t_i, is specified by

  $$Y(t_i) = F(X(t_i)),$$

  where $y = F(x)$ is a random transformation with the transition probability $p(y|x)$, modeling trading noise.

Under this construction, information affects $X(t)$, the intrinsic value of an asset, and has a permanent influence on the price while noise affects $F(x)$, the random transformation, and has only a transitory influence on price. This formulation is similar to the time series structural models applied in many market microstructure papers (see Hasbrouck (1996), a survey paper, and Hasbrouck (2002)) in that $X(t)$ is the permanent component and $F(x)$ is the transient component. Moreover, the formulation is closely related to the two-time-scale frameworks incorporating market microstructure noises in the fast growing literature of realized volatility estimators. See Zhang, Mykland and Aït-Sahalia (2005), Aït-Sahalia, Mykland and Zhang (2005), Bandi and Russell (2006), Diebold (2006), Fan and Wang (2007), and especially Li and Mykland (2007).

Examples of $F(x)$(or $p(y|x)$) are given in Zeng (2003), Zeng (2004), and Spalding, Tsui and Zeng (2006). These aforementioned examples well accommodate the three types of well-documented noise in financial literature: discrete noise, clustering noise, and non-clustering noise. Especially, Spalding et. at. (2006) applied a simple model of Zeng (2003) with new measures of trading noises and trading cost to further support the important findings of Christie and Schultz (1994), Christie, Harris and Schultz (1994) and Barclay, Christie, Harris, Kandel and Schultz (1999) [2].

2.3 Representation II: Filtering with Counting Process Observations

Alternatively, we can view the transaction prices in the levels of price due to price discreteness. That is, we can view the prices as a collection of counting processes in the following form:

$$\vec{Y}(t) = \begin{pmatrix} N_1(\int_0^t \lambda_1(s, X(s))ds) \\ N_2(\int_0^t \lambda_2(s, X(s))ds) \\ \vdots \\ N_n(\int_0^t \lambda_n(s, X(s))ds) \end{pmatrix}, \tag{3}$$

where $Y_k(t) = N_k(\int_0^t \lambda_k(s, X(s))ds)$ is the counting process recording the cumulative number of trades that have occurred at the kth price level (denoted by y_k) up to time t.

The following four mild assumptions are invoked so that the probability of simultaneous trades is zero, and more importantly, this representation is equivalent to the previous one in distribution. The equivalence ensures that the statistical analysis based on the latter specification can be applied to the former and the equivalence is proven in Zeng (2005).

Assumption 2. $\{N_k\}_{k=1}^n$ *are unit Poisson processes under* P.

Assumption 3. $X, N_1, N_2, \ldots, N_n$ *are independent under* P.

[2] The results of Christie and Schultz (1994) led to regulatory investigations, legal activities, and numerous academic studies. This culminated with the Securities and Exchange Commission imposing a series of market reforms in NASDAQ. Barclay et. al. (1999) documented effects of the NASDAQ market reforms.

Assumption 4. *The intensity at price level k, $\lambda_k(t,x) = a(t,x)p(y_k|x)$, where $a(t,x)$ is the total trading intensity at time t with $x = X(t)$ and $p(y_k|x)$ is the transition probability from x to y_k, the kth price level. Furthermore, $|y_k - R[x, \frac{1}{M}]| \le C_1$, where $R[x, \frac{1}{M}]$ rounds x to the closest $\frac{1}{M}$ and the positive C_1 is achievable, namely, the support of the possible y given x is from $R[x, \frac{1}{M}] - C_1$ to $R[x, \frac{1}{M}] + C_1$.*

Remark 1. Note that $p(y_k|x)$ is the same as $p(y|x)$ for $F(x)$ of Representation I. The structure of λ_k implies that $a(t, X(t))$ specifies when the trade might occur while $p(y_k|x)$ specifies at which price level the trade might occur. Since the impact of trading noise in price formation is limited, it is reasonable to assume the price is in a bounded range of the intrinsic value.

Assumption 5. *The total intensity, $a(t,x)$, is uniformly bounded above; i.e., there exists a constant, C_2, such that $a(t,x) \le C_2$ for all $t > 0$ and x.*

Remark 2. Under this representation, $X(t)$ becomes the signal process, which cannot be observed directly, and $\vec{Y}(t)$ becomes the observation process, which is corrupted by noise. Hence, $(X, \vec{Y})$ is framed as a filtering problem with counting process observations.

For readers not familiar with counting processes, note that $Y_k(t) = N_k(\int_0^t \lambda_k(s, X(s))ds)$ has a Poisson distribution with parameter $\int_0^t \lambda_k(s, X(s))ds$ when the path of X, $\{X(s) : 0 \le s \le t\}$, is given. Actually, Y_k is a random time-change of the unit Poisson process N_k. Also, although $N_1, ..., N_n$ are independent unit Poisson processes, $Y_1, ..., Y_n$ are not independent, because their intensity functions all depends on X.

Under Representation II, with Assumptions 1 - 5, the filtering equations and a consistent computational algorithm are developed in Zeng (2003). We review them in Section 4.4 in order to compute the hedging strategy under partial information.

2.4 An Integral Form of Price

To solve the option pricing and hedging problem, the differential or integral equation (SDE) of the most recent price is needed. However, the previous two representations do not provide such a form for the price. We fill this hole with the two-step construction below.

• Step 1 : We define a random counting measure. Let $U = \{0, \frac{1}{M}, \frac{2}{M}, \cdots\}$ be all the possible price levels, and $u = \frac{j}{M}$ be a generic point in U. Define $m(A,t) = \sum_{u\in A} Y_{uM}(t)$, for $A \in U$. Then, $m(A,t)$ counts the cumulative number of trades whose price levels are in A up to time t. Note that m is a random counting measure.

A random measure is fully characterized by its compensator. To describe the compensator, we denote η as the usual counting measure on U with the following two properties: For $A \in U$, $\eta(A) = \int_U I_A \eta(du)$ (i.e. $\eta(A)$ counts the number of elements in A) and $\int_A f(u)\eta(du) = \sum_{u\in A} f(u)$. Then, $\gamma_m(A,t)$, the compensator of $m(A,t)$ with respect to $\mathcal{F}_t^X$, can be written as

$$\gamma_m(A,t) = \int_0^t \int_A p(u|X(s))a(s, X(s))\eta(du)ds$$

$$= \sum_{u \in A} \int_0^t p(u|X(s))a(s, X(s))ds$$

and

$$\gamma_m(du, dt) = p(u|X_t)a(t, X_t)\eta(du)dt.$$

We are in the position to define the integral form needed.

• Step 2 : Let $Y(t)$ be the price of the most recent transaction at or before time t. Then,

$$Y(t) = Y(0) + \int_{[0,t]\times U} (u - Y(s-))m(du, ds). \tag{4}$$

Observe that $m(du, ds)$ is zero most of the time and becomes one only at trading time t_i with $u = Y(t_i)$, the trading price. The above expression is but a telescoping sum: $Y(t) = Y(0) + \sum_{t_i<t}(Y(t_i) - Y(t_{(i-1)}))$.

Correspondingly, there is a differential form for $Y(t)$:

$$dY(t) = \int_U (u - Y(t-))m(du, dt). \tag{5}$$

To help explain the above differential equation, if there is a price change from $Y(t-)$ to u which occurs at time t, then $Y(t) - Y(t-) = (u - Y(t-))$ implying $Y(t) = u$.

With the above evolution of price, we study option pricing and hedging for the FM model through the local risk minimizing criterion in Section 4. But, we first review risk minimization in Section 3.

3 A Brief Review of Risk Minimization

The general theory of the risk minimization approach was first developed in Föllmer and Sondermann (1986) and then extended to the local risk minimization in Föllmer and Schweizer (1991). This brief review is based on these two papers.

3.1 Setup and Mean Self-Financing Strategies

Suppose that the stock price, $Y = (Y_t)_{0\le t\le T}$, is a continuous square-integrable semimartingale on some filtered probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{0\le t\le T}, P)$. Y has the Doob-Meyer decomposition: $Y = Y_0 + M + A$ where M is a local martingale and A is a predictable process with finite variation paths. Suppose that the random variable $H \in \mathcal{L}^2(\Omega, \mathcal{G}_T, P)$ is a contingent claim at time T. A hedging strategy against this claim is a portfolio strategy consisting of the stock Y and a riskless bond $D \equiv 1$. Define ξ_t and ζ_t as the number of shares in stock and bond, respectively, held at time t. ξ_t is assumed to be predictable (i.e. roughly, $\mathcal{G}_{t-}$-measurable) and ζ_t to be adapted (i.e. $\mathcal{G}_t$-measurable). Then, $V_t = \xi_t Y_t + \zeta_t$ is the value of the portfolio at time t and $C_t = V_t - \int_0^t \xi_s dY_s$ is the cost of the portfolio at time t with respect to the strategy (ξ, ζ). With the following integrability condition

$$E\Big[\int_0^T \xi_s^2 d\langle Y\rangle_s + \Big(\int_0^T |\xi_s| d|A|_s\Big)^2\Big] < \infty, \tag{6}$$

the process of stochastic integral in C_t is well-defined. Here, $\langle Y\rangle$ (or $\langle V, Y\rangle$) is the conditional (co-)quadratic variation (see Protter 2004).

A *mean self-financing* strategy $(\xi_t, \zeta_t)_{0\le t\le T}$ is an admissible strategy whose C_t is a martingale(with $V_T = H$). Note that mean self-financing strategies include self-financing strategies (i.e. $C_t = C_0$, a constant cost).

3.2 Complete Market Case

When the market is complete and has no arbitrage, the contingent claim, H, admits an Itô representation. That is, there exists an adapted ξ^H such that: $H = H_0 + \int_0^T \xi_s^H dY_s$. Then, let $\xi := \xi^H$, $\zeta := V - \xi Y$, and $V_t := H_0 + \int_0^T \xi_s^H dY_s$, for $0 \le t \le T$. In such a case, it is clear that such a strategy is admissible and moreover, self-financing, that is, $C_t = C_T = H_0$.

Also in this standard case, no arbitrage is equivalent to the existence of the unique equivalent martingale measure (EMM), P', under which Y becomes a martingale. Then, the above strategy is uniquely determined as follows: $V_t = E'[H|\mathcal{G}_t]$, and $\xi^H = \frac{d\langle V,Y\rangle}{d\langle Y\rangle}$, the Radon-Nikodym derivative of the conditional quadratic covariation of V and Y with respect to the conditional quadratic variation of Y.

3.3 Incomplete Market Case I

In the ideal complete market, the hedging completely gets rid of the risk in a contingent claim. It is no longer possible in an incomplete market. A common claim has an *intrinsic risk*, the smallest possible risk. The problem is to find a strategy that decreases the extant risk to the intrinsic part. Case I assumes that Y is already a martingale under P and Föllmer and Sondermann (1986) studies this problem.

Define the *remaining risk* among the admissible strategies at each time t as $E[(C_T - C_t)^2|\mathcal{G}_t]$. Define the *risk minimizing strategy* at each time t as the optimal strategy among admissible strategies that minimizes the remaining risk at each time t. Then, the *intrinsic risk* is the minimum remaining risk achieved by the risk minimization strategy.

Suppose $\mathcal{G}_T$-measurable H has the unique *Kunita-Watanabe* decomposition under P,

$$H = H_0 + \int_0^T \xi_s^H dY_s + L_T^H$$

where ξ^H is predictable and L^H is a martingale orthogonal to Y. Then, the risk-minimizing strategy is uniquely given by $\xi = \xi^H = \frac{d\langle V,Y\rangle}{d\langle Y\rangle}$, and $\zeta = V - \xi Y$ with $V_t = E[H|\mathcal{G}_t] = H_0 + \int_0^t \xi_s^H dY_s + L_t^H$.

3.4 Incomplete Market Case II

Case II assumes that P itself is no longer a martingale measure and Föllmer and Schweizer (1991) studies this general incomplete case, under which the risk minimizing strategy does not exist. However, they proposed a *local risk minimization* criterion. A *locally risk-minimizing* strategy is the one minimizing the remaining risk under all infinitesimal perturbations of the strategy at time t. This definition is equivalent to the associated cost, C, of the locally risk-minimizing strategy being a square-integrable martingale, orthogonal to M under P. In another way, the existence of a locally risk-minimizing strategy is equivalent to H having the *Föllmer and Schweizer* decomposition:

$$H = H_0 + \int_0^T \xi_s^H dY_s + L_T^H$$

under P with $H_0 \in \mathcal{L}^2(\Omega, \mathcal{G}_0, P)$, where ξ^H satisfies Equation(6) and L^H is a square-integrable martingale orthogonal to M, the martingale part of Y in its Doob-Meyer decomposition.

To identify the above Föllmer and Schweizer decomposition, the *minimal martingale measure* provides an important tool. A martingale measure Q is called *minimal* if $Q = P$ on $\mathcal{G}_0$, and if any square-integrable P-martingale L that satisfies $\langle L, M\rangle = 0$ remains a martingale under Q. Note that the minimal martingale measure preserves orthogonality, namely, if $\langle L, M\rangle = 0$ under P, then $\langle L, X\rangle = 0$ under Q.

Under the above setup, they show that the minimal martingale measure and the locally risk-minimizing strategy are uniquely determined. More importantly, define V_t as $V_t = E^Q[H|\mathcal{F}_t]$, and the optimal strategy is given by $\xi^H = \frac{d\langle V,Y\rangle}{d\langle Y\rangle}$ and $\zeta = V - \xi Y$.

4 Risk Minimization for the FM model

In this section, we use the methodology of Föllmer and Schweizer (1991) to study the risk minimizing hedging problem of the FM model described in Section 2. For the most recent price process, we first obtain the Doob-Meyer decomposition, which helps to specify the minimal martingale measure. Note that the FM model is not a continuous semimartingale and the results reviewed in Section 3 are not obviously applicable if the price process has jumps (as in the FM model). Lee and Protter (2008) shows that the results of the local risk minimization still hold for a class of semimartingales allowing feedback jumps.

In the FM model, there are not only jumps, but also there is partial information. Namely, the hedging is not based on $\mathcal{F}_t^{X,Y} = \sigma((X_s, Y_s) : 0 \le s \le t)$, but based only on $\mathcal{F}_t^Y = \sigma(Y_s : 0 \le s \le t)$. Nevertheless, first under full information, we are able to adopt the argument in Lee and Protter (2008) to prove the existence of the minimal martingale measure, and derive the optimal hedging strategy. Then, we obtain the filtered hedging strategy under partial information by taking the conditional expectation and develop a related filtering technique for the computation of such hedging strategies.

4.1 The Doob-Meyer Decomposition of the Price Process

To describe the Doob-Meyer decomposition, we first define a couple of functions which also play an important role in the specification of the minimal martingale measure. Recall in Assumption 4 that $|y - R[x, \frac{1}{M}]| \le C_1$, where C_1 is achievable. Let

$$\begin{aligned} g_1(X(t), Y(t-)) &= \int_U (u - Y(t-))p(u|X(t))\eta(du) \\ &= \sum_{j=M(R[X(t),\frac{1}{M}]-C_1)}^{M(R[X(t),\frac{1}{M}]+C_1)} (u - Y(t-))p(u|X(t)). \end{aligned} \tag{7}$$

and

$$\begin{aligned} g_2(X(t), Y(t-)) &= \int_U (u - Y(s-))^2 p(u|X(t))\eta(du) \\ &= \sum_{j=M(R[X(t),\frac{1}{M}]-C_1)}^{M(R[X(t),\frac{1}{M}]+C_1)} (u - Y(t-))^2 p(u|X(t)). \end{aligned} \tag{8}$$

Then, $g_1(X(t), Y(t-))$ and $g_2(X(t), Y(t-))$ are the first two conditional moments of $Y(t) - Y(t-)$ given $Y(t-)$ and $X(t)$ and if a trade occurred at time t. We observe that

$$\begin{aligned} \langle Y \rangle_t &= \int_0^t \int_U (u - Y(s-))^2 \gamma_m(du, ds) \\ &= \int_0^t \Big[\int_U (u - Y(s-))^2 p(u|X(s))\eta(du) \Big] a(s, X(s))ds \\ &= \int_0^t g_2(X(s), Y(s-))a(s, X(s))dt. \end{aligned} \tag{9}$$

Now, we can write the Doob-Meyer decomposition in a nice form. Recall that the most recent price in the model of Section 2 in differential form is

$$dY_t = \int_U (u - Y(s-))m(du, ds)$$

where the compensator of $m(du, ds)$ is $\gamma_m(du, dt) = p(u|X_t)a(t, X_t)\eta(du)dt$. Denote the compensated $m(du, ds)$ as $\tilde{m}(du, ds) = m(du, ds) - \gamma_m(ds, ds)$. Hence, the Doob-Meyer Decomposition can be written as $Y_t = M_t + A_t$ where

$$M_t = Y(0) + \int_0^t \int_U (u - Y(s-))\tilde{m}(du, ds), \tag{10}$$

is the local martingale and

$$\begin{aligned} A_t &= \int_0^t \int_U (u - Y(s-))\gamma_m(du, ds) \\ &= \int_0^t \Big[\int_U (u - Y(s-))p(u|X_s)\eta(du) \Big] a(s, X_s)ds \\ &= \int_0^t g_1(X(s), Y(s-))a(s, X(s))ds \\ &= \int_0^t \frac{g_1(X(s), Y(s-))}{g_2(X(s), Y(s-))} d\langle Y \rangle_s. \end{aligned} \tag{11}$$

is the predictable Finite Variation (FV) part. The above equation provides a key component in identifying the Radon-Nikodym density that determines the minimal martingale measure.

4.2 Minimal Martingale Measure

One more assumption is needed.

Assumption 6. *For all $t \in [0, T]$, $g_2(X(t), Y(t-)) \leq C_3$ and*

$$g_2(X(t), Y(t-)) - g_1(X(t), Y(t-))(Y(t) - Y(t-)) > 0. \tag{12}$$

Remark 3. One simple sufficient condition for Assumption 6 is that $g_1(X(t), Y(t-)) = 0$. That is, the expected difference between $Y(t)$ and $Y(t-)$ is zero given $Y(t-)$ and $X(t)$ and if a trade occurred at time t.

Below, we present a theorem on the minimal martingale measure.

Theorem 1. *Under Assumptions 1 - 6, define Z as*

$$Z_t = 1 - \int_0^t Z_{s-} \frac{g_1(X(s), Y(s-))}{g_2(X(s), Y(s-))} \int_U (u - Y(s-))\tilde{m}(dz, ds).$$

Then, $Z_t > 0$ and $E(Z_t) = 1$ for all $t \in [0, T]$. Furthermore, Q defined by $\frac{dQ}{dP} = Z_T$ is the minimal martingale measure of Y.

Proof. The proof consists of three parts. The first one shows that Z defines an equivalent probability measure, the second part shows the measure Q defined by Z is a martingale measure, and the third part shows Q has the minimal property.

• Step 1 : For the first part, it suffices to show that $Z_t > 0$ and $E(Z_t) = 1$ for all t. Notice that

$$Z_t = 1 + \int_0^t Z_{s-} dK_s,$$

where

$$K_t = -\int_0^t \frac{g_1(X(s), Y(s-))}{g_2(X(s), Y(s-))} \int_U (u - Y(s-))\tilde{m}(du, ds). \tag{13}$$

Using the Stochastic Exponential formula, we can solve it for Z as

$$Z_t = \exp(K_t - \frac{1}{2}[K, K]_t^c) \prod_{0<s\leq t} (1 + \Delta K_s) \exp(-\Delta K_s).$$

To obtain $Z_t > 0$, we should show that $(1 + \Delta K_t) > 0$ for all $t \in (0, T]$, which is equivalent to

$$1 - \frac{g_1(X(t), Y(t-))}{g_2(X(t), Y(t-))}(Y(t) - Y(t-)) > 0$$

for all $t \in (0, T]$. Assumption 6 ensures the above condition is always satisfied. Therefore $Z_t > 0$.

Since Z is a P-local martingale and $Z_0 = 1$, it suffices to show that $E([Z]_t) < \infty$ to obtain $E(Z_t) = 1$ for all t, because Z becomes a martingale if $E([Z]_t) < \infty$, where $[Z]_t$ is the quadratic variation process of Z. We have

$$\begin{aligned} E\{[Z]_t\} &= E\int_0^t Z_{s-}^2 d[K]_s = E\int_0^t Z_{s-}^2 \left(\frac{g_1}{g_2}\right)^2 \int_U (u - Y(s-))^2 m(du, ds) \\ &\leq E\int_0^t Z_{s-}^2 \left(\frac{g_1}{g_2}\right)^2 g_2 \ a(s, X(s))ds \leq C^* \int_0^t E\{Z_s^2\}ds, \end{aligned}$$

for some large enough C^* by Assumption 6 and the boundedness of $\frac{|g_1|}{g_2}$, $a(t, X(t))$ and g_2. Further by the same boundedness and from the direct evaluation of the expected value the under Poisson distribution, we obtain

$$E\{Z_t^2\} \leq E\exp(2K_t) < C^{**}$$

for some constant C^{**}.

• Step 2 : To show Q is a martingale measure, we apply the Girsanov-Meyer Theorem (Theorem 3.36 on page 133 of Protter (2004)). Then, it suffices to show that the FV part $A_t = -\int_0^t \frac{1}{Z_{s-}} d\langle M, Z\rangle_s$, where M is the local martingale in the Doob-Meyer decomposition of Y and $\langle M, Z\rangle_t$ is the conditional quadratic co-variation process (or angle bracket process) of M and Z. It is true as shown below:

$$
\begin{aligned}
& -\int_0^t \frac{1}{Z_{s-}} d\langle M, Z\rangle_s \\
& = \int_0^t \frac{1}{Z_{s-}} d\Bigg\langle \int_0^t \int_U (u - Y(s-))\tilde{m}(du, ds), \\
& \int_0^t Z_{s-} \frac{g_1(X(s), Y(s-))}{g_2(X(s), Y(s-))} \int_U (u - Y(s-))\tilde{m}(dz, ds) \Bigg\rangle \\
& = \int_0^t \frac{g_1(X(s), Y(s-))}{g_2(X(s), Y(s-))} \Big[\int_U (u - Y(s-))^2 p(u|X(s))\eta(du) \Big] a(s, X(s)) ds \\
& = \int_0^t g_1(X(s), Y(s-)) a(s, X(s)) ds = A_t.
\end{aligned}
$$

For the second equality, we apply a well-known property of the angle bracket process: Suppose that $f_1(t, u, \omega)$ and $f_2(t, u, \omega)$ are $\mathcal{F}_t^{X,Y}$-predictable functions. Then,

$$
\begin{aligned}
& \Bigg\langle \int_0^\cdot \int_U f_1(s, u, \omega)\tilde{m}(du, ds), \int_0^\cdot \int_U f_2(s, u, \omega)\tilde{m}(du, ds) \Bigg\rangle_t \\
& = \int_0^t \int_U f_1(u, s, \omega) f_2(u, s, \omega) \gamma_m(du, ds).
\end{aligned} \tag{14}
$$

• Step 3 : We show that the measure Q satisfies the minimal condition. Let M' be a P-martingale such that $\langle M', M\rangle = 0$. Then, again by Girsanov-Meyer theorem, we should show that M' is a Q-martingale, which is true since

$$
\begin{aligned}
\langle M', Z\rangle_t &= \int_0^t Z_{s-} d\langle M', K\rangle_s \\
&= -\int_0^t Z_{s-} \frac{g_1(X(s), Y(s-))}{g_2(X(s), Y(s-))} d\langle M', M\rangle_s \\
&= 0.
\end{aligned}
$$

Note that the conditional quadratic co-variation is calculated under P. □

Having established a minimal martingale measure Q, we study the change of the Brownian motion and the random measure under Q. The next theorem states that the Brownian motion remains invariant under Q, but the random measure has a new compensator, which is explicitly provided.

Theorem 2. *Under the minimal martingale measure Q defined in Theorem 1, $B_t^* = B_t$ and* *the compensator of* $m^*(du,dt)$ *is given by*

$$\begin{aligned}&\gamma_{m^*}(du,dt)\\=&\Big(1-\frac{g_1(X(t),Y(t-))}{g_2(X(t),Y(t-))}(u-Y(t-)\Big)\gamma_m(du,dt)\\=&\Big(1-\frac{g_1(X(t),Y(t-))}{g_2(X(t),Y(t-))}(u-Y(t-)\Big)p(u|X(t))\eta(du)a(X(t-),t)dt.\end{aligned}\tag{15}$$

Remark 4. Assumption 6 ensures the nonnegativity of the compensator of m^*.

Since the proof contains some technical concepts of random measure, those concepts and the proof are provided in Appendix A.

4.3 Local Risk Minimization under Full Information

Recall the filtrations $\mathcal{F}_t^{X,Y}=\sigma(X_s,Y_s:0\le s\le t)$, the full information, and $\mathcal{F}_t^{Y}=\sigma(Y_s:0\le s\le t)$, the partial information observed by agents. Let H be a European style contingent claim of Y and define $H_t=E^Q[H(Y_T)|\mathcal{F}^{X,Y}]$. Note that X is Markovian, but Y itself is not. Nevertheless, we have the lemma below with its proof in Appendix A.

Lemma 1. *The triple* (t,X,Y) *forms a three-dimensional Markov process under both* P *and* Q*.*

Hence,

$$H_t=E^Q[H(Y_T)|\mathcal{F}_t^{X,Y}]=E^Q[H(Y_T)|X_t,Y_t]=f(t,X_t,Y_t)$$

for some function f.

The following theorem presents the PDE for $f(t,x,y)$, the martingale representation for H_t and the local risk-minimizing hedging strategy under the full information.

Theorem 3. *The integro-differential equation for* $f(t,x,y)\in C^{1,2,2}$ *where* $x=X(s)$ *and* $y=Y(s-)$ *is*

$$\begin{aligned}0=&f_t(t,x,y)+\mu(x)f_x(t,x,y)+\frac{1}{2}\sigma^2(x)f_{xx}(t,x,y)\\&+a(t,x)\sum_u\kappa(u;t,x,y)\Big(1-\frac{g_1(y,x)}{g_2(y,x)}(u-y)\Big)p(u|x)\end{aligned}\tag{16}$$

where $\kappa(u;t,x,y)=f(t,x,u)-f(t,x,y)$ *with the terminal condition:*

$$f(T,X(T),Y(T))=H(Y(T)).\tag{17}$$

Moreover, the martingale representation for H_t *is:*

$$\begin{aligned}H_t=&H_0+\int_0^t f_x(s,X_s,Y_{s-})\sigma(X_s)dB_s\\&+\int_0^t\int_U\kappa(u;s,X_s,Y_{s-})[m^*(du,ds)-\gamma_{m^*}(du,ds)].\end{aligned}\tag{18}$$

Let $g_3(X(s), Y(s-)) = \int_U \kappa(u; s, X_s, Y_{s-})(u - Y_{s-})p(u|X(s))\eta(du)$ *and let*

$$\xi_t^{X,Y} = \frac{d\langle H, Y\rangle}{d\langle Y\rangle} = \frac{g_3(X(t), Y(t-))}{g_2(X(t), Y(t-))}. \tag{19}$$

Then, the local risk minimizing hedging strategy under full information is given by $\xi^{X,Y}$ *and* $\zeta^{X,Y} = H - \xi^{X,Y} Y$.

Proof. Assume $f(t, x, y)$ is in $C^{1,2,2}$. Recall $X(t)$ is a diffusion by Assumption 1. Itô's formula gives

$$\begin{aligned} &f(t, X_t, Y_t) \\ =&f(0, X_0, Y_0) + \int_0^t f_x(s, X_s, Y_{s-})\sigma(X_s)dB_s + \int_0^t \int_U \kappa(u; s, X_s, Y_{s-})\tilde{m}^*(du, ds) \\ &+ \int_0^t \Big(f_t(s, X_s, Y_{s-}) + f_x(s, X_s, Y_{s-})\mu(X_s) + \frac{1}{2} f_{xx}(s, X_s, Y_{s-})\sigma^2(X_s)\Big)ds \\ &+ \int_0^t \int_U \kappa(u; s, X_s, Y_{s-})\gamma_{m^*}(du, ds) \end{aligned} \tag{20}$$

where $\tilde{m}^*(dz, dt) = m^*(du, dt) - \gamma_{m^*}(dz, ds)$, the compensated $m^*(du, dt)$ under Q. Recall $\gamma_{m^*}(du, dt)$ in Equation(15). So, the last term of the above equation is a term of ds. Since $f(t, X_t, Y_t)$ is a martingale under Q, the ds term must vanish. With the last term merging into ds, we obtain the integro-differential equation of Equation(16) for $f(t, x, y)$ where $x = X(s)$ and $y = Y(s-)$. Note that in Equation(17) we have used $Y(T) = Y(T-)$ a.s. Furthermore, we obtain the martingale representation for H_t as in Equation(18).

Föllmer and Schweizer (1991) showed that the hedging strategy under full information $\xi^{X,Y}$ can be obtained as the $\mathcal{F}_t^{X,Y}$-predictable density of the conditional quadratic covariation of H and Y with respect to the conditional quadratic variation of Y. Namely,

$$\xi^{X,Y} = \frac{d\langle H, Y\rangle}{d\langle Y\rangle}, \tag{21}$$

where the angle brackets are calculated under P and with respect to $\mathcal{F}_t^{X,Y}$.

The above result is not obviously applicable when the FM model has jumps. The key point to note is that the minimal martingale measure preserves orthogonality when passing from P to Q, whereas the above result rests on the converse implication. Fortunately, the above result still hold for $\mathcal{H}^2$ semimartingales which includes the FM model (see Equation (11) in Section 4 of Lee and Song 2007).

Next, we compute $\langle H, Y\rangle$. We will use the bilinearity of the angle brackets and the well known property in Equation(14).

Note that

$$\begin{aligned}\langle H, Y\rangle_t &= \int_0^t \int_U \kappa(u; s, X_s, Y_{s-})(u - Y_{s-})\gamma_m(du, ds) \\ &= \int_0^t \int_U \kappa(u; s, X_s, Y_{s-})(u - Y_{s-})p(u|X(s))\eta(du)a(s, X(s))ds \qquad (22) \\ &= \int_0^t g_3(X(s), Y(s-))a(s, X(s))ds.\end{aligned}$$

Recall $d\langle Y\rangle_t = g_2(X(t), Y(t-))a(t, X(t))dt$. So the hedging strategy under full information is given by

$$\xi_t^{X,Y} = \frac{g_3(X(t), Y(t-))}{g_2(X(t), Y(t-))}. \qquad (23)$$

Observe that the above $\xi_t^{X,Y}$ can be viewed as an infinitesimal projection of ΔH_t on ΔY_t. $\square$

4.4 Filtered Local Risk Minimization with Partial Information

We now turn to a more realistic partial information case where the hedger only observes the prices at trading times, namely, $\mathcal{F}_t^Y$, excluding the continuous path of the intrinsic value process of the asset.

With the partial information, one way to produce a reasonable hedging strategy is to project the optimal strategy of Section 4.3 on the space spanned by the partial information. Namely, we propose to use the conditional expectation of the optimal strategy under full information given the observed trading prices.

Since the strategy is predictable, it is defined for time $t-$. Let

$$\xi_{t-}^Y = E^Q\Big[\frac{g_3(X(t), Y(t-))}{g_2(X(t), Y(t-))}|\mathcal{F}_{t-}^Y\Big],$$

$$H_{t-}^Y = E^Q[H_{t-}|\mathcal{F}_t^Y] = E^Q[f(t, X(t), Y(t-)|\mathcal{F}_{t-}^Y],$$

and

$$\zeta_{t-}^Y = H_{t-}^Y - \xi_{t-}^Y Y_{t-}.$$

Hence, the hedging strategy under partial information is given by (ξ_-^Y, ζ_-^Y).

In the next two subsections, we employ the filtering technique developed in Zeng (2003) to develop the computation of the above hedging strategy under Q. For such a connection with filtering, we call our hedge strategy a ”filtered locally risk-minimizing hedging strategy”.

4.4.1 The Filtering Equation under Q

The filtering technique enables us to compute the conditional distribution of $X(t)$ given $\mathcal{F}_t^Y$ as well as that of $X(t)$ given $\mathcal{F}_{t-}^Y$. Note that the filtering equation derived in Zeng (2003) is under P, not Q, and hence, is not suitable for computing the filtered locally risk-minimizing hedging strategy. However, the technique in Zeng (2003) can be used to derive the related filtering equation under Q.

The filtering equation under Q is to be derived for Representation II as in Zeng (2003) and we present Representation II under Q as shown below:

$$\vec{Y}(t) = \begin{pmatrix} N_1(\int_0^t \hat{\lambda}_1(s, X(s), Y(s-))ds) \\ N_2(\int_0^t \hat{\lambda}_2(s, X(s), Y(s-))ds) \\ \vdots \\ N_n(\int_0^t \hat{\lambda}_n(s, X(s), Y(s-))ds) \end{pmatrix}, \tag{24}$$

where $\hat{\lambda}_k$ depends on Y, the most recent price given in Equation (4) under Q. Precisely,

$$\hat{\lambda}_k(t, (X(t), Y(t-)) = a(t, X(t))\Big(1 - \frac{g_1(X(t), Y(t-))}{g_2(X(t), Y(t-))}(y_k - Y(t-)\Big)p(y_k|X(t)). \tag{25}$$

Let $\hat{\pi}_t$ be the conditional distribution of $X(t)$ given $\mathcal{F}_t^{\vec{Y}}$ under Q and let

$$\hat{\pi}(f, t) = E^Q[f(X(t))|\mathcal{F}_t^{\vec{Y}}] = \int f(x)\hat{\pi}_t(dx).$$

As long as we know the conditional distribution of $X(t)$ given $\mathcal{F}_t^Y$ or $\mathcal{F}_{t-}^Y$ under Q, we can compute

$$E^Q[b(t, X(t), Y(t-))|\mathcal{F}_{t-}^Y] = \int b(t, x, Y(t-))\hat{\pi}_t(dx)$$

for any given function $b(t, X(t), Y(t-))$. Then, we can compute the filtered hedging strategy under partial information. The conditional distribution is characterized by the filtering equation in the theorem below.

Theorem 4. *Suppose that $(X, \vec{Y})$ satisfies Assumptions 1 - 6 and Q is the minimal martingale measure defined in Theorem 1. Then, $\hat{\pi}_t$ is the unique measure-valued solution of the normalized filtering equation*

$$\begin{aligned} \hat{\pi}(f, t) &= \hat{\pi}(f, 0) + \int_0^t \hat{\pi}(\mathbf{A}f, s)ds \\ &+ \sum_{k=1}^n \int_0^t \Big[\frac{\hat{\pi}(f\hat{\lambda}_k, s-)}{\hat{\pi}(\hat{\lambda}_k, s-)} - \hat{\pi}(f, s-)\Big] d\big(Y_k(s) - \hat{\pi}(f\hat{\lambda}_k, s)s\big) \end{aligned} \tag{26}$$

for $t > 0$ and $f \in D(\mathbf{A})$, the domain of generator $\mathbf{A}$ of X given by Equation (2), where $\hat{\lambda}_k = \hat{\lambda}_k(t, X(t), Y(t-))$ given by Equation(25), and $p_k = p(y_k|(X(t))$ is the transition probability from $X(t)$ to y_k.

The proof of Theorem 4 employs the reference probability measure approach and is a slight modification of that of Theorem 3.1 in Zeng (2003). Thus the proof is omitted.

Remark 5. Equation (26) can be separated in two parts. The first is called the *propagation equation*, describing the evolution without trades and the second is called the *updating equation*, describing the update when a trade occurs. Let $t_1, t_2, \ldots,$ be the trading times. The propagation equation has no random component and is written as

$$\hat{\pi}(f, t_{i+1}-) = \hat{\pi}(f, t_i) + \int_{t_i}^{t_{i+1}-} \Big[\hat{\pi}(\mathbf{A}f, s) - \sum_{k=1}^{m} \big(\hat{\pi}(f\hat{\lambda}_k, s) - \hat{\pi}(f, s)\hat{\pi}(\hat{\lambda}_k, s)\big)\Big] ds. \tag{27}$$

This implies that when there are no trades, the posterior evolves deterministically.

Assuming the price at time t_{i+1} occurs at the jth price level, the updating equation is

$$\hat{\pi}(f, t_{i+1}) = \frac{\hat{\pi}(f\hat{\lambda}_k, t_{i+1}-)}{\hat{\pi}(\hat{\lambda}_k, t_{i+1}-)}. \tag{28}$$

It is random because the price level k, which is the observation, is random.

4.4.2 The Computation of the Approximate Hedging Strategies

The filtering equation is infinite-dimensional and to actually compute the conditional distribution, we develop the approximate consistent numerical scheme below.

Suppose the state space of X is discretized with ϵ as the length between lattices. Then, X_ϵ, an approximation for X, can be constructed via a Markov chain approximation. Namely, $X_\epsilon \Rightarrow X$, which means X_ϵ converges weakly to X in the Skorohod topology as $\epsilon \to 0$. Define under P,

$$\vec{Y}_\epsilon(t) = \begin{pmatrix} N_1(\int_0^t \lambda_1(s, X_\epsilon(s))ds) \\ N_2(\int_0^t \lambda_2(s, X_\epsilon(s))ds) \\ \vdots \\ N_n(\int_0^t \lambda_n(s, X_\epsilon(s))ds) \end{pmatrix}. \tag{29}$$

Theorem 5. *Suppose that $(X, \vec{Y})$ is on the probability space $(\Omega, \mathcal{F}, P)$ with Assumptions 1 - 6 and Q is the minimal martingale measure in Theorem 1. Suppose that for $\epsilon > 0$, $(X_\epsilon, \vec{Y}_\epsilon)$ is on $(\Omega_\epsilon, \mathcal{F}_\epsilon, P_\epsilon)$, Assumptions 1 - 6 also hold for $(X_\epsilon, \vec{Y}_\epsilon)$, and Q_ϵ is the corresponding minimal martingale. If $X_\epsilon \Rightarrow X$ as $\epsilon \to 0$, then*
(i) $\vec{Y}_\varepsilon \Rightarrow \vec{Y}$ as $\varepsilon \to 0$; and
(ii)$E^{Q_\epsilon}[F(X_\epsilon(t))|\mathcal{F}_t^{\vec{Y}_\epsilon}] \Rightarrow E^Q[F(X(t))|\mathcal{F}_t^{\vec{Y}}]$ as $\varepsilon \to 0$ for function F in the domain of the generator $\mathbf{A}$.

The proof of the theorem is a slight modification of that of Theorem 4.1 in Zeng (2003). Thus the proof is omitted also.

Remark 6. Let $\hat{\pi}_{\epsilon,t}$ be the conditional distribution of $X_\epsilon(t)$ given $\mathcal{F}_t^{\vec{Y}_\epsilon}$ under Q_ϵ. Theorem 5 implies that $\hat{\pi}_{\epsilon,t} \Rightarrow \hat{\pi}_t$, and further that

$$E^{Q_\epsilon}[b(t, X_\epsilon(t), Y_\epsilon(t-))|\mathcal{F}^{Y_{\epsilon,t-}}] \Rightarrow E^Q[b(t, X(t), Y(t-))|\mathcal{F}_{t-}^Y]$$

for any given function $b(t, X(t), Y(t-))$. Applying the Markov chain approximation method, a consistent recursive algorithm can be constructed to compute $\hat{\pi}_{\epsilon,t}$ as in Zeng (2003). Then, we can compute the approximate filtered hedging strategy under partial information.

5 Conclusion

In summary, we apply the local risk minimization approach to derive the explicit filtered hedging strategies of the FM model with partial information. The model incorporates the two stylized features of ultra-high frequency data: random trading times and market microstructure noise. We develop the computation of the derived hedging strategy through a nonlinear filtering technique.

There are several interesting possibilities for future works. The first one is to prove the optimality of the filtered locally risk-minimizing hedging strategy under partial information. However, it is a challenging task because the projection technique used for the martingale case fails in the semimartingale situation. The second one is to overcome the difficulty in the computation of the hedging and option pricing. Then, we will be able to compare the filtered locally risk-minimizing hedging with the usual Black-Scholes hedging and option pricing, and to understand how the trading intensity and the magnitude of market microstructure noise affect hedging and option pricing, which is an important practical problem. Third, it would be interesting to apply other alternative hedging approaches for incomplete markets such as the minimum entropy martingale measure approach, mean-variance or utility maximization approaches, to compare the different hedging strategies, and to evaluate them.

A Related Proofs and a Lemma

A.1 The Proof of Lemma 1

Since a Markov process is characterized by its infinitesimal generator, $\mathbf{A}$, (see Chapter 4 of Ethier and Kurtz 1986), it suffices to calculate the generater for $(t, X(t), Y(t))$. If the generator exists then $(t, X(t), Y(t))$ is an inhomogeneous Markov process. Recall that for $x \in R, y \in U$, the generator is defined as

$$\mathbf{A}f(t,x,y) = \lim_{s \downarrow 0} \frac{E[f(t+s, X(t+s), Y(t+s))|t, X(t)=x, Y(t)=y] - f(t,x,y)}{s}.$$

Using Equation (16) obtained from Itô's formula and with some computation, we obtain the generator under P as

$$\mathbf{A}f(t,x,y) = \frac{\partial}{\partial t} f(t,x,y) + \mu(x)\frac{\partial}{\partial x} f(t,x,y) + \frac{1}{2}\frac{\partial^2}{\partial x^2} f(t,x,y)$$
$$+a(t,x)\int_U \big[f(t,x,u) - f(t,x,y)\big]p(u|x)\eta(du),$$

and the generator under Q as

$$\mathbf{A}f(t,x,y) = \frac{\partial}{\partial t}f(t,x,y) + \mu(x)\frac{\partial}{\partial x}f(t,x,y) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}f(t,x,y)$$

$$+a(t,x)\int_U \big[f(t,x,u) - f(t,x,y)\big]\Big[1 - \frac{g_1(x,y)}{g_2(x,y)}(u-y)\Big]p(u|x)\eta(du).$$

Therefore, $(t, X(t), Y(t))$ is an inhomogeneous Markov process under P and Q. We also observe that when $a(t,x) = a(x)$ independent of time t, $\mathbf{A}f(X(t), Y(t))$ exists and $(X(t), Y(t))$ becomes a homogeneous Markov process.

A.2 Some Technical Definitions and a Lemma

Let U be the mark space which can be $\mathbb{R}$ or a countable set like $\{0, \frac{1}{M}, \frac{2}{M}, \cdots\}$ used in this paper. Suppose that $\mu(\omega;\cdot,\cdot)$ is a random counting measure on $U \times [0,T]$ such that $\mu(A,t)$ counts the number of jumps in A up to time t where A is a subset of U. Then, for a function $W : \Omega \times U \times [0,T] \to \mathbb{R}$,

$$\int_0^t \int_U W(\omega;,u,s)\mu(\omega;du,ds) = \sum_{n=1}^{\infty} W(\omega; U_n(\omega), T_n(\omega))1_{(T_n(\omega)\le t)},$$

when T_n is n-th jump time.

To prove the Theorem 2, we need the following lemma which is a part of Theorem 3.17 of Chapter III(page 157), Jacod and Shiryaev (2002) and a version of the Girsanov theorem for random measure. We first provide some technical definitions. Define $M_\mu^P = M_\mu^P(W) = E(W * \mu_\infty)$ as the positive measure for all measurable nonnegative functions W, where $W * \mu_\infty = \int_{[0,\infty)}\int_U W(\omega;u,s)\mu(\omega;du,ds)$. By the predictable σ-field $\tilde{\mathcal{P}}$, we mean a σ-field on $\Omega \times \mathbb{R}$ such that $\tilde{\mathcal{P}} = \mathcal{P} \otimes \mathcal{B}$ where $\mathcal{B}$ is a Borel σ-field. There is a notion of conditional expectation relative to M_μ^P with respect to the predictable σ-field $\tilde{\mathcal{P}}$: for every nonnegative measurable function W, let $W' = M_\mu^P(W|\tilde{\mathcal{P}})$ denote the M_μ^P-a.s. unique $\tilde{\mathcal{P}}$-measurable function such that $M_\mu^P(WU) = M_\mu^P(W'U)$ for all nonnegative $\tilde{\mathcal{P}}$-measurable U. For more details and examples, we refer to Chapter 2 and 3 of Jacod and Shiryaev (2002).

Lemma 2. *Assume that $Q \ll P$ and let Z be the density process. Let $\mu = \mu(\omega;du,dt)$ be an integer-valued random measure on $U \times \mathbb{R}^+$, and let $\upsilon = \upsilon(\omega;dt,dx)$ be the P-compensator of μ. Let Y be any nonnegative version of $M_\mu^P(\frac{Z}{Z_-}1_{\{Z_->0\}}|\tilde{\mathcal{P}})$ and υ' be a version of the Q compensator of μ. Then $\upsilon'(\omega;dt,dx) = Y(\omega;t,x)\upsilon(\omega;dt,dx)$ $P-a.s.$*

A.3 The Proof of Theorem 2

Since $\langle Z, B\rangle = 0$, the Girsanov-Meyer Theorem gives $B_t^* = B_t$ under Q. This is the straightforward part. We use the above technical lemma to derive the change of the random measure under Q.

As mentioned in Section 4.1, the compensator of $m(du, ds)$ under the original measure P is $\gamma_m(du, dt) = p(u|X_t)a(t, X_t)\eta(du)dt$. Therefore, by the lemma, it suffices to find $M_m^P(\frac{Z}{Z_-}1_{\{Z_->0\}}|\tilde{\mathcal{P}})$.

Recall

$$Z_t = 1 + \int_0^t Z_{s-}dK_s,$$

where K_t is given by Equation (13). Therefore,

$$\frac{Z_t}{Z_{t-}} = 1 + \Delta K_t = 1 - \frac{g_1(X(t), Y(t-))}{g_2(X(t), Y(t-))}(Y(t) - Y(t-)).$$

Following an argument similar to that given in the proof of Theorem 4 of Lee and Song (2007), we can easily see that

$$M_m^P(\frac{Z}{Z_-}1_{\{Z_->0\}}|\tilde{\mathcal{P}})_t = \Big(1 - \frac{g_1(X(t), Y(t-))}{g_2(X(t), Y(t-))}(u - Y(t-)\Big).$$

This proves the theorem by Lemma 2.

References

[1] Aït-Sahalia, Y., Mykland P. A. and Zhang, L. (2005). *How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise*. Review of Financial Studies, 18, 351-416.

[2] Bandi, F. M. and Russell, J. R. (2006), *Separating microstructure noise from volatility.* Journal of Financial Economics, **79**, 655-692.

[3] Barclay, M., Christie, W., Harris, J., Kandel, E. and Schultz, P. H., (1999). *The effects of market reform on the trading costs and depths of NASDAQ stocks*, Journal of Finance, **54 (1)**, 1–34.

[4] Black, F., (1986). *Noise*, Journal of Finance, **41**, 529–543.

[5] Carr, P. and Wu, L., (2004). *Time-Changed Levy Processes and Option Pricing*, Journal of Financial Economics, **17(1)**, 113–141.

[6] Ceci, C. (2006). *Risk minimizing hedging for a partially observed High Frequency Data model*, Stochastics: An International Journal of Probability and Stochastic Processes, 78, 1, pp.13-31.

[7] Cetin, U., Jarrow, R., Protter, P., and Warachka, M., (2006). *Pricing Options in an Extended Black Scholes Economy with Illiquidity: Theory and Empirical Evidence*,Reviews of Financial Studies, **19**, 493–529.

[8] Christie, W., Harris, J. H. and Schultz, P. H., (1994). *Why did NASDAQ market makers stop avoiding odd-eighth quotes?*, Journal of Finance, **49**, 1841–1860.

[9] Christie, W. and Schultz, P. H., (1994). *Why do NASDAQ market makers avoid odd-eighth quotes?*, Journal of Finance, **49**, 1813–1840.

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

[11] Diebold, F.X., (2006). *On Market Microstructure Noise and Realized Volatility.* Journal of Business and Economic Statistics, 24, 181-183.

[12] Duffie, D and Singleton, K.J., (2003). *Credit Risk: Pricing, Measurement, and Management*, Princeton University Press, Princeton, NJ.

[13] Engle, R., (2000). *The econometrics of ultra-high-frequency data*, Econometrica, **68**, 1–22.

[14] Engle, R. and Russell, J., (2005). *Analysis of high frequency data*, *in* Y. Aït-Sahalia and L. P. Hansen, eds, 'Handbook of Financial Econometrics', North-Holland, Amsterdam, (to appear).

[15] Ethier, S. N. and T.G. Kurtz, (1986). *Markov processes : Characterization and convergence*. Wiley, New York.

[16] J. Fan, and Y. Wang, (2007). *Multi-scale jump and volatility analysis for high-Frequency financial data*. Journal of American Statistical Association, vol 102, 1349-1362.

[17] Föllmer, H., and M. Schweizer, (1991). *Hedging of contingent claims under incomplete information*, Applied Stochastic Analysis, Stochastic Monographs, Vol.5, Goldon and Breach, 389–414.

[18] Föllmer, H., and D. Sondermann (1986). *Hedgingn of non-redundant contingent-claims*, in "Contributions to Mathematical Economics" eds. W. Hilderbrand and A. Mas-Colell. Amsterdam: North-Holland, 147-160.

[19] Frey, R. (2000). *Risk-minimization with incomplete information in a model for high-frequency data*, Mathematical Finance **10**, 215–225.

[20] Hasbrouck, J., (1996). *Modeling market microstructure time series*, *in* G. Maddala and C. Rao, eds, 'Handbook of Statistics', Vol. 14, North-Holland, Amsterdam, pp. 647–692.

[21] Hasbrouck, J., (2002). *Stalking the "efficient price" in market microstructure specifications: an overview*, Journal of Financial Markets **5**, 329–339.

[22] Huang, J. and Wu, L., (2004). *Specification Analysis of Option Pricing Models Based on Time-Changed Levy Processes*, Journal of Finance, **59(3)**, 1405–1439.

[23] Jacod,J. and Shiryaev,N., (2002). *Limit Theorems for Stochastic Processes*, 2nd Edition, Springer-Verlag, Heidelberg.

[24] Kouritzin, M. and Zeng, Y., (2005). *Bayesian model selection via filtering for a class of micro-movement models of asset price*, International Journal of Theoretical and Applied Finance, **8**, 97–121.

[25] Lee, K., and Protter, P., (2008) *Hedging Claims with Feedback Jumps in the Price Process*, Communication on Stochastic Analysis, vol 2, (1) 125–143.

[26] Lee, K. and Song, S., (2007). *Insider's hedging in jump diffusion models*, Quantitative Finance, vol 7 (5) 537–545.

[27] Li, Y. and Mykland, P. A., (2007). *Are volatility estimators robust with respect to modeling assumptions?* Bernoulli, vol 13 (3).

[28] Protter, E.P. (2004). *Stochastic Integration and Differential Equations*, Springer, Berlin, Germany.

[29] Schweizer, M., (2001). *A guided Tour through Quadratic Hedging Approaches*, Option Pricing, Interest Rates and Risk Management, Cambridge Press, Cambridge, UK.

[30] Spalding, R., Tsui, K. W. and Zeng, Y., (2006). *A micro-movement model with Bayes estimation via filtering: Applications to measuring trading noises and trading cost*, Nonlinear Analysis: Theory, Methods and Applications, **64**, 295-309.

[31] Zeng, Y., (2003). *A partialy observed model for micromovement of asset prices with Bayes estimation via filtering*, Mathematical Finance, **13** 411-444.

[32] Zeng, Y., (2004). *Estimating stochastic volatility via filtering for the micro-movement of asset prices*, IEEE Transactions on Automatic Control, **49**, 338-348.

[33] Zeng, Y., (2005). *Bayesian inference via filtering for a class of counting processes: Application to the micromovement of asset price*, Statistical Inference for Stochastic Processes, **8**, 331 - 354.

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