

## EVALUATING HEDGING ERRORS: AN ASYMPTOTIC APPROACH

TAKAKI HAYASHI

*Department of Statistics, Columbia University*

PER A. MYKLAND

*Department of Statistics, University of Chicago*

We propose a methodology for evaluating the hedging errors of derivative securities due to the discreteness of trading times or the observation times of market prices, or both. Utilizing a weak convergence approach, we derive the asymptotic distributions of the hedging errors as the discreteness disappears in several situations. First, we examine the hedging error due to discrete-time trading when the true strategy is known, which generalizes the result of Bertsimas, Kogan, and Lo (2000) to continuous Itô processes. Then we consider a data-driven strategy, when the true strategy is unknown. This strategy is free of parametric model assumptions, therefore it is expected to serve as a benchmark for the evaluation of parametric strategies. Finally, we consider a case study of the Black-Scholes delta-hedging strategy when the volatility is unknown in the proposed framework. The results obtained give us a prospect for further developments of the framework under which various parametric strategies could be compared in a unified manner.

KEY WORDS: delta hedging, weak convergence, incomplete market, model uncertainty, nonparametric regression

### 1. INTRODUCTION

It is well understood that the perfect *replication* of a financial derivative security by any self-financing strategy is not possible generally in incomplete markets. In this paper, we are particularly interested in the incompleteness when only discrete-time hedging or observations, or both, are allowed. We develop a theoretical framework for measuring the hedging error of a derivative security under those circumstances. For that we utilize an asymptotic approach, as the time interval of the discreteness goes to zero.

First, we investigate the hedging errors solely due to discrete-time hedging, when perfect knowledge about the true dynamics of the price processes is available to the hedger. Defining the hedging error as a stochastic process, we derive its limiting distribution as the time interval goes to zero. In particular, we show that the asymptotic distribution is characterized by a stochastic integral with respect to a new independent Brownian

We are grateful to Professors Ioannis Karatzas, Jussi Keppo, Shigeo Kusuoka, Steven Lalley, Seongjoo Song, Nakahiro Yoshida, and the referee for helpful comments and suggestions. All remaining errors are our own. This research was partially supported by National Science Foundation grants DMS-9626266, DMS-9971738, and DMS-0204639.

*Manuscript received November 2002; final revision received April 2004.*

Address correspondence to Takaki Hayashi, Department of Statistics, Columbia University, MC4690, 1255 Amsterdam Avenue, New York, NY 10027, USA; e-mail: hayashi@stat.columbia.edu.

motion, which is the evidence of the incompleteness resulting from the discrete hedging. This result has the same spirit as that of Bertsimas, Kogan, and Lo (2000). The first contribution of the paper, however, is to generalize their result to continuous Itô processes driven by a multidimensional Brownian motion. Thus the result covers important classes such as stochastic volatility models and non-Markovian models.

Second, we consider a hedger who lacks perfect knowledge about the true market model but who is able to partially access the market prices. The hedger would conduct a data-driven, nonparametric regression strategy in discrete time, based on the price data available to the hedger. We are interested in evaluating its hedging error relative to the “true” strategy, which could have been adopted if perfect knowledge had been provided. The limiting distribution is found as the discreteness disappears. It is shown to have the same distribution as the first one, up to a constant multiplier. This type of result is new in the literature, so far as we are aware.

We combine the first two results, assuming both sampling of prices and trading are done at discrete times. We evaluate the hedging error of a *historical* delta-hedging strategy—a regression strategy based on the discrete observations. Since the regression-based strategy is free of parametric model assumptions about the securities, it is robust to misspecification of models and their parameters. Thus it is expected to serve as a benchmark for evaluating various parametric strategies; any existing parametric strategies would be justifiable only if they “outperform” this benchmark strategy when there is uncertainty about the true model.

Finally, we consider a case study where the Black-Scholes model is correct but the volatility is unknown in the proposed framework, and we examine the hedging error of the Black-Scholes delta-hedging strategy with the implied volatility.

In the literature, consequences of discrete-time hedging operation have been considered, usually and naturally in conjunction with the existence of transaction costs, because then hedgers would like to trade as least frequently as possible. Pioneered by Leland (1985), asymptotic approaches—as the size of trading intervals shrinks to zero—have been well utilized (see, e.g., Kabanov and Safarian 1997; Ahn et al. 1998; and Grandits and Schachinger 2001). Most of the existing research studies that are interested in the magnitude of resulting errors, however, have evaluated their (approximate) second moments *at the expiration* of options. In the meantime, Liang (1997) and Bertsimas et al. (2000) independently computed the asymptotic error distribution *over the entire life* of an option in a one-dimensional diffusion model (with little or no reference to transaction costs). Other than those works, the studies closest to ours are those by Gobet and Temam (2001) and Temam (2003). We comment on them in Remark 3.2 to Theorem 3.1.

Measuring hedging errors due to model misspecification has drawn more attention in recent years. The main difficulty here is that formulating “model misspecification” is an arduous task; therefore, the hedging error due to it is hard to quantify. Among the latest results in this newly emerging branch of stochastic finance are those of Bouleau (2003), Mahayni (2003), and Talay and Zheng (2003). Bouleau proposed an *error calculus* (which is an extension of the Malliavin calculus based on Dirichlet forms) in order to study the sensitivity of pricing and hedging to changes of parameters and to perturbations of  $S$  processes. The problem is different from ours in that that study was not concerned with discrete-time trading. In Mahayni, expected costs of delta hedging due to discrete-time trading as well as model misspecification (in the sense of applying strategies based on continuous-time models to discrete-time trading) were evaluated in a log-normal process, without resorting to asymptotics. The author was essentially

interested in biases of hedging errors (that are caused by asset price trends), not in variances or distributions. Talay and Zheng, together with their preceding paper, proposed a methodology to measure the magnitude of the discretization error, induced by the Euler scheme, of quantiles of the distribution of one component of a multidimensional diffusion process at expiration date  $T$ , with suitable conditions specified in terms of Malliavin covariance matrices. They applied their methodology to evaluating a model risk due to a misspecified hedging strategy in a Heath-Jarrow-Morton model. Since their quantity of interest was different from ours, it is hard to discuss mutual relevance. As mentioned, our approach is that, instead of trying to measure directly the hedging errors due to “model uncertainty” whose formulation could be ad hoc and controversial, we will measure those of a *least informative* strategy—a nonparametric regression strategy.

Nonparametric/semiparametric hedging strategies, as an alternative to the existing parametric model-based strategies, have been proposed by, for example, by Hutchinson, Lo, and Poggio (1994) and Ait-Sahalia and Lo (1998). Those studies estimated pricing formulas (i.e., the price of the claim as a nonlinear function of the price of the underlying asset and other parameters) by nonparametric or semiparametric statistical methods such as neural networks and kernel regression, and they measured their performance numerically or empirically. In the meantime, our data-driven scheme is closer, both in motivation and methodology, to that of Bossaerts and Hillion (1997, 2003), who proposed a locally linear, “semiparametric” regression strategy based on *time-series* data of prices, utilizing the local “curvature” information of the Black-Scholes model. They intended to incorporate available information more efficiently than our “nonparametric” strategy, assuming the Black-Scholes model to be locally correct. Although they investigated numerical and empirical performance, they did not provide analytic considerations. In the meantime, we compute the asymptotic distribution of the hedging error process of a nonparametric strategy, which could show us its approximate structure, together with the rate of convergence.

The main workhorse in this study is the theory of weak convergence for stochastic processes. Jacod and Shiryaev (1987) is an indispensable reference in the area. A nice overview of applications of the theory in financial markets has been recently provided by Prigent (2003).

## 2. SETUPS

Let  $W := (W^1, W^2)'$  be independent Brownian motions. Let  $T < \infty$  be a fixed terminal time. We have three types of securities: bond, stock, and derivative security. For simplicity, there is no transaction cost.

Assume the bond has a constant price; that is, assume the interest rate is zero. All the results extend easily to the case of deterministic interest rates.<sup>1</sup> Assume the stock price follows a positive Itô process,

$$(2.1) \quad dS_t = \mu_t dt + f_t dW_t^1$$

with  $S_0$  a constant, where  $\mu$  and  $f$  are adapted, suitably integrable processes, and  $f_t > 0$  for all  $t$ .

<sup>1</sup> In particular, due to the Numéraire Invariance principle (e.g., see Duffie 1996, p. 102), self-finance strategies considered in the paper would be still applicable even when the interest rate is nonzero.

Assume that the European derivative security  $C$ , which will yield a nonnegative,  $\mathcal{F}_T$ -measurable payoff  $C_T \geq 0$  at  $T$ , is created via “gain from the trades” in  $S$ ; that is,

$$dC_t = \theta_t dS_t$$

associated with a given *self-financing* strategy  $\theta$ , which is an adapted, suitably integrable process.  $C_0$  is a constant.<sup>2</sup> (Note: There is no gain/loss from the trades in the bond.) To avoid obvious arbitrage opportunities, we must require  $C_t \geq 0$  for all  $t$ .

REMARK 2.1. Notice that  $C$  has been defined without reference to any prespecified payoff to be hedged. Recall that in incomplete markets not all contingent claims are replicable by a self-financing strategy. Let  $\eta \geq 0$  be an arbitrary  $\mathcal{F}_T$ -measurable claim. In the best scenario,  $\theta$  can be the perfect replicating strategy for  $\eta$ ; in other words,  $C_T = C_0 + \int_0^T \theta_t dS_t = \eta$ , a.s., if the market is complete or if  $\eta$  is specifically hedgeable so that such  $\theta$  exists in an incomplete market. Otherwise,  $\theta$  can be a *superreplicating* strategy; that is,  $C_T \geq \eta$ , a.s., or any other (e.g., Föllmer and Leukert 1999 and Mykland 2000), so far as the use of such a strategy is justifiable and the conditions elaborated in this study are met. Allowing for potential lacking of the perfect hedge for  $\eta$ , we instead focus on the portfolio value process  $C$  with its generating strategy  $\theta$ . Skipping an arduous discussion of the superiority of a choice of  $\theta$  against another, we begin with the situation that  $\theta$  has been selected by a certain criterion already. Throughout the paper, we will treat  $\theta$  as the “true” strategy and the corresponding  $C$  as the security of interest whose market price shall be targeted by the hedgers.

A standard market model that has the dynamics

$$(2.2) \quad dS_t = a_t S_t dt + \sigma_t S_t dW_t^1$$

for some suitably regular processes  $a$  and  $\sigma$  (e.g., Karatzas and Shreve 1998, p. 3), is covered, and is in general non-Markovian. Of particular importance are one-dimensional diffusion models with  $a_t = a(t, S_t)$  and  $\sigma_t = \sigma(t, S_t)$  for some suitable  $\mathbb{R}$ -valued functions  $a(\cdot, \cdot)$  and  $\sigma(\cdot, \cdot)$ . Another important class includes stochastic volatility models. Throughout the paper, stochastic differential equations such as those are meant to have *strong* solutions.

We suppose that the “true” trading strategy  $\theta$  is also an Itô process,

$$(2.3) \quad d\theta_t = \theta_t^{[0]} dt + \underline{\theta}_t^{[1]} dW_t$$

with  $\theta_0$  a constant, where  $\theta^{[0]}$  and  $\underline{\theta}^{[1]}$  are adapted, suitably integrable,  $\mathbb{R}$ - and  $\mathbb{R}^2$ -valued processes. Since  $\theta$  defines the value of  $C$ , which shall yield a “reasonable” payoff, we will always consider strategies such that the associated gain from trades,  $\int_0^t \theta_u dS_u$ , is a.s. bounded from below by some real constant (e.g., Duffie 1996, pp.103–105).

REMARK 2.2. If  $S$  is driven only by  $W^1$  and  $\theta$  is driven by  $W^1$  and  $W^2$ , then  $C$  will depend on both  $W^1$  and  $W^2$ . Thus  $C$  would not be replicable by continuous-time trading only in  $S$ , *but for* the knowledge about  $\theta$ . In this case,  $C$  is not a *redundant* security.

<sup>2</sup> In incomplete markets the claim price  $C_0$  may not be uniquely determined. However, we assume that the “price” is given, without questioning the direction.

EXAMPLE 2.1. True strategies  $\theta$ :

- (i) *European (path-dependent) option C* (the general case). So far as continuous observations are available, the hedging strategy  $\theta$  can be “retrieved” by

$$(2.4) \quad \theta_t = \frac{d\langle S, C \rangle_t}{d\langle S, S \rangle_t},$$

where  $\langle \cdot, \cdot \rangle$  is the quadratic variation. Among all (2.4), we restrict our attention to the processes having the decomposition (2.3).

- (ii) *European path-independent option C*. If the option price is of the form  $C_t = C(S_t, t)$  for a smooth function  $C(\cdot, \cdot)$ , then Itô’s formula implies

$$dC_t = \dot{C}_S(S_t, t) dS_t + \dot{C}_t(S_t, t) dt + \frac{1}{2} \ddot{C}_{SS}(S_t, t) d\langle S, S \rangle_t.$$

Hence,

$$d\langle S, C \rangle_t = \dot{C}_S(S_t, t) d\langle S, S \rangle_t$$

and

$$\theta_t = \dot{C}_S(S_t, t).$$

(Note that because the interest rate is zero we have  $\dot{C}_t(S_t, t) dt + \frac{1}{2} \ddot{C}_{SS}(S_t, t) \cdot d\langle S, S \rangle_t = 0$  by the no-arbitrage condition.)

By Itô’s formula applied to  $\theta_t = \dot{C}_S(S_t, t)$ , we can identify in (2.3)

$$(2.5) \quad \theta_t^{[0]} = \ddot{C}_{St} + \frac{1}{2} \ddot{C}_{SSS} f_t^2 + \ddot{C}_{SS} \mu_t; \quad \theta_t^{[1]} = \ddot{C}_{SS} f_t.$$

For instance, if  $C$  is a European call option in the Black-Scholes market ( $f_t = \sigma S_t$  and  $\mu_t = aS$  for constant  $\sigma > 0$  and  $a$  in (2.1)), then the option *delta* is given by  $\theta_t = \Phi(d_t)$ , where  $\Phi$  is the standard normal cumulative distribution function (cdf), with  $d_t := \frac{\ln(S_t/K) + (\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$ . If it is a *binary* call,  $C_T = 1_{\{S_T > K\}}$ ,  $\theta_t = \frac{\phi(d_t)}{\sigma S_t \sqrt{T-t}}$ , where  $\phi$  is the standard normal probability density function. Their respective  $\theta_t^{[0]}$  and  $\theta_t^{[1]}$  can be computed explicitly.

There are certainly cases that are not covered in the current set-up. An example of such is a lookback (*Russian*) option  $C_T = (M_T - K)^+$  in the Black-Scholes market, where  $M_t = \max_{0 \leq s \leq t} S_s$  (see Shepp and Shiryaev 1993). A result is yet to be obtained for cases when (2.3) is not satisfied.

Let  $\pi$  be a feasible strategy that a hedger could take. We assume that hedging operations are done only at (nonrandom) discrete times,  $t_i, i = 0, \dots, n - 1$ , that are equally spaced;  $t_{i+1} - t_i \equiv \frac{T}{n}$ , in particular. Also, we assume that  $\pi = (\pi_{t_i})$  is self-financing. In other words, at the beginning of each trading period between  $t_i$  and  $t_{i+1}$ , the hedger holds  $\pi_{t_i}$  shares of the stock with all remaining wealth put in the risk-free bond. When  $\pi_{t_i}$  is computed, the hedger can only access the information set available up to time  $t_i$ , in addition to the prior knowledge that the hedger has.

Accordingly, the gain from the trades by the discrete-time strategy  $\pi := (\pi_{t_i})$  is defined by

$$V_j^\pi := \sum_{i=0}^{j-1} \pi_{t_i} \Delta S_{t_i}, \quad j = 0, 1, \dots, n,$$

where  $\Delta S_{t_i} := S_{t_{i+1}} - S_{t_i}$  is the change in price from  $t_i$  to  $t_{i+1}$ . Here,  $\sum_{i=0}^{-1} = 0$  by convention.

For measuring the performance of  $\pi$  relative to the true (continuous-time)  $\theta$ , it is natural to define the hedging error  $L^n = (L_{t_j}^n)$ , by

$$L_{t_j}^n := C_{t_j} - C_0 - V_{t_j}^\pi = \int_0^{t_j} \theta_u dS_u - \sum_{i=0}^{j-1} \pi_{t_i} \Delta S_{t_i}, \quad j = 0, 1, \dots, n,$$

which will be referred to as the *total* hedging error. Another type of error, which measures the performance of  $\pi$  relative to  $\theta$  after the discretized effect has been allowed for, is defined by

$$L_{2,t_j}^n := V_{t_j}^\pi - V_{t_j}^{\pi(n)} = \sum_{i=0}^{j-1} (\theta_{t_i} - \pi_{t_i}) \Delta S_{t_i}, \quad j = 0, 1, \dots, n.$$

This will be referred to as the *relative* hedging error.

Some questions worthwhile addressing include the following. What is the impact of discrete-time trading on the hedging error when knowing the true model, or at least having complete access to the continuous record of the market prices? What if we lack knowledge about the true model but are able to develop a data-driven hedging scheme based on the observations available to us? In particular, how good or bad are such strategies relative to the “true” ones? Are we able to utilize those answers to develop a more general framework for evaluating parametric strategies under the model uncertainty? We shall try to give partial answers to these questions by exploring *small interval asymptotics*, the meaning of which shall be evident.

Throughout the paper, we shall further assume that the price dynamics introduced above have a hierarchical structure; that is, the coefficients evolve as stochastic processes of a similar kind, as stated in the following.

The diffusion coefficient  $\underline{\theta}^{[1]}$  of the true hedge ratio  $\theta$  is a two-dimensional Itô process such that

$$(2.6) \quad d\underline{\theta}_t^{[1]} = \underline{\theta}_t^{[10]} dt + \underline{\theta}_t^{[11]} dW_t,$$

where  $\underline{\theta}^{[10]}$  and  $\underline{\theta}^{[11]}$  are adapted, suitably integrable,  $\mathbb{R}^2$ - and  $\mathbb{R}^{2 \times 2}$ -valued. Also, the diffusion coefficient  $f$  of  $S$  process is an Itô process such that

$$df_t = f_t^{[0]} dt + \underline{f}_t^{[1]} dW_t,$$

where  $f^{[0]}$  and  $\underline{f}^{[1]}$  are adapted, suitably integrable,  $\mathbb{R}$ - and  $\mathbb{R}^2$ -valued. Similarly, the further coefficients  $\underline{f}^{[0]}$ , and  $\underline{f}^{[1]}$  are Itô processes with the corresponding forms.

EXAMPLE 2.2. Coefficient dynamics of  $S$ :

(i) *The Black-Scholes model.* Suppose  $S$  follows the geometric Brownian motion

$$dS_t = S_t(a dt + \sigma dW_t^1)$$

for constant  $a$  and  $\sigma > 0$ . Since  $f_t = \sigma S_t$ ,

$$df_t = a\sigma S_t dt + \sigma^2 S_t dW_t^1 = f_t(a dt + \sigma dW_t^1),$$

it is the same diffusion as  $S$  (up to constant). So are  $f_t^{[0]} = a\sigma S_t$ , and  $\underline{f}_t^{[1]} = \sigma^2 S_t$ .

(ii) *Stochastic volatility model.* Suppose  $S$  is a stochastic volatility model of the form

$$\begin{aligned} dS_t &= S_t(a dt + \sigma_t dW_t^1), \\ d\sigma_t &= \alpha(\sigma_t) dt + \beta(\sigma_t) dW_t^{(2)} \end{aligned}$$

for some suitable functions  $\alpha, \beta$ , where  $W^{(2)} := \rho W^1 + \sqrt{1 - \rho^2} W^2$ .

Since  $f_t = \sigma_t S_t$ , by Itô's formula,

$$\begin{aligned} df_t &= \sigma_t dS_t + S_t d\sigma_t + d\langle \sigma, S \rangle_t \\ &= (a\sigma_t + \alpha(\sigma_t) + \rho\sigma_t\beta(\sigma_t))S_t dt \\ &\quad + (\sigma_t^2 + \rho\beta(\sigma_t))S_t dW_t^1 + \sqrt{1 - \rho^2}\beta(\sigma_t)S_t dW_t^2 \\ &= f^{[0]}(f_t, S_t) dt + f^{[11]}(f_t, S_t) dW_t^1 + f^{[12]}(f_t, S_t) dW_t^2, \end{aligned}$$

where  $f^{[0]}(x, y) := ax + y\alpha(x/y) + \rho x\beta(x/y)$ ,  $f^{[11]}(x, y) := x^2/y + \rho y\beta(x/y)$ ,  $f^{[12]}(x, y) := \sqrt{1 - \rho^2}y\beta(x/y)$ . Together with  $S, f$  is a diffusion. So are  $f_t^{[0]} = f^{[0]}(f_t, S_t)$  and  $f_t^{[1]} = (f^{[11]}(f_t, S_t), f^{[12]}(f_t, S_t))'$ .

EXAMPLE 2.3. *The hedge ratio dynamics of a European path-independent  $C$*  (continued from Example 2.1(ii)): Recall from (2.5) that

$$\theta_t^{[1]} = \ddot{C}_{SS}f_t.$$

Itô's formula implies that

$$\theta_t^{[10]} = \ddot{C}_{SSS}f_t\mu_t + \ddot{C}_{SS}f_t + \ddot{C}_{SS}f_t^{[0]}; \quad \underline{\theta}_t^{[11]} = (\ddot{C}_{SSS}f_t^2 + \ddot{C}_{SS}f_t^{[11]}, \ddot{C}_{SS}f_t^{[12]})$$

in (2.6), where  $f^{[1]k}$  is the  $k$ th element of  $f^{[1]}$ ,  $k = 1, 2$ .

These examples show that in many instances of practical importance diffusion (and drift) coefficients inherit their process “types” in the hierarchical structure as assumed.

The notations and definitions in this paper are essentially those of Jacod and Shiryaev (1987). In what follows, “ $\xrightarrow{P}$ ” denotes convergence in probability, and “ $\xrightarrow{\mathcal{L}}$ ” stands for convergence in law in a space of continuous functions, endowed with the uniform topology, unless otherwise stated.<sup>3</sup>

In this setup, we derive asymptotic distributions of the hedging errors of strategies under various situations, as will be seen in the following sections.

### 3. HEDGING ERROR DISTRIBUTION DUE TO DISCRETE-TIME TRADING

In this section, we examine the impact of discrete-time trading. The true strategy  $\theta$  is known but trading times are restricted to  $\{t_i\}$ , which are regularly spaced,  $t_{i+1} - t_i \equiv \frac{T}{n}$ , for each  $i$ .<sup>4</sup> Suppose the hedger wishes to hedge a European derivative security whose value evolves in continuous time as

<sup>3</sup> A sequence  $(X^n)$  of random variables that take values in a space of continuous functions,  $\mathbb{C}$ , defined possibly on different probability spaces,  $(\Omega^n, \mathcal{F}^n, P^n)$ , is said to *converge in law* to  $X$ , a random variable that takes values in  $\mathbb{C}$  defined on some probability space  $(\Omega, \mathcal{F}, P)$ , if  $E^n[f(X^n)] \rightarrow E[f(X)]$  for all bounded continuous functions  $f$  on  $\mathbb{C}$ . Here,  $E^n[\cdot]$  is the expectation taken with respect to  $P^n$ .

<sup>4</sup> Liang (1997) studied the asymptotic distribution of the hedging error due to nonregularly spaced trading in the Black-Scholes model.

$$(3.1) \quad dC_t = \theta_t dS_t,$$

with  $C_t \geq 0$  for all  $t$ . In this case, the natural discrete-time strategy for the hedger is to take  $\pi = \theta$ .

The total hedging error due to the discrete-time hedging at  $\{t_i\}$  is given by

$$L_{1,t_j}^n := \int_0^{t_j} \theta_u dS_u - \sum_{i=0}^{j-1} \theta_{t_i} \Delta S_{t_i}.$$

We derive the limiting process of  $L_1^n$  as  $n \rightarrow \infty$ , with an appropriate rescaling factor. The Euclidean norm is denoted by  $\|\cdot\|$ .

**THEOREM 3.1.** (Hedging error of discrete-time trading.) *Under suitable regularity conditions,*

$$(3.2) \quad \sqrt{n} L_1^n \xrightarrow{\mathcal{L}} \sqrt{\frac{T}{2}} \int_0^\cdot \|\underline{\theta}_u^{[1]}\| f_u dW_u^*$$

as  $n \rightarrow \infty$ , where  $W^*$  is a Brownian motion, independent of  $W = (W^1, W^2)$ .

*Proof.* All the theorems are proved in the Appendix. □

Thus, we have established that the asymptotic distribution of the hedging error is characterized by the stochastic integral with respect to some new independent Brownian motion  $W^*$ , appearing in the first-order sense. This result manifests the *market incompleteness* due to the discrete hedging. This is still true even when  $C$  is a *redundant* security, in the sense that the single  $W^1$  drives both continuous-time processes  $(S, C)$  (see Remark 2.2). On the other hand, according to the proof of Theorem 3.1, we can see that the addition of independent Brownian motions ( $W^2$  or even others) would have no effect. That is, it would not matter whether or not  $(S, C)$  would form a complete market even if they were traded in continuous time.

This theorem endorses the importance of “gamma”  $\underline{\theta}^{[1]}$ —and here we abuse this financial jargon, the actual meaning of which is the partial derivative (or the sensitivity) of the delta with respect to the underlying stock price—which plays an essential role in determining the magnitude of the hedging error. The larger gamma in magnitude the derivative security has, the larger hedging error in magnitude the hedger will incur, which makes sense from practitioners’ viewpoints.

Now, let us denote the limit process as  $L_1 := \sqrt{\frac{T}{2}} \int_0^\cdot \|\underline{\theta}_u^{[1]}\| f_u dW_u^*$ . An obvious corollary is given as follows.

**COROLLARY 3.1.** *Conditioning on the whole paths of  $(f_u, \underline{\theta}_u^{[1]}, 0 \leq u \leq T)$ ,  $L_1$  is a continuous, centered Gaussian process with covariance function  $\gamma(s, t) := \frac{T}{2} \int_0^{s \wedge t} (\|\underline{\theta}_u^{[1]}\| f_u)^2 du$ .*

**REMARK 3.1.**

- (a) This result extends that of Bertsimas et al. (2000) to continuous Itô processes (especially, non-Markovian models), as they assume  $S$  is a diffusion driven by a single Brownian motion.
- (b) When  $S$  and  $C$  are observed continuously, the price volatility  $f_t = (\frac{d\langle S, S \rangle_t}{dt})^{\frac{1}{2}}$  and the hedge ratio  $\theta_t = \frac{d\langle S, C \rangle_t}{d\langle S, S \rangle_t}$  are also known to the hedger. On the other hand,



the drift  $\mu$  is in general uncertain. (For instance, even in the simplest case where  $\mu_t = aS_t$  for some unknown constant  $a$ , such  $a$  may still be uncertain at the finite expiration  $T < \infty$ .) However, the drift uncertainty turns out to be irrelevant for the hedging error, at least in the first-order sense.

- (c) In the one-dimensional case, the limit may be represented alternatively as  $L_1 = \sqrt{\frac{T}{2}} \int_0^T \theta_u^{[1]} f_u dW_u^*$ .

EXAMPLE 3.1. The Black-Scholes delta-hedging strategy of European path-independent options: Since  $f_t = \sigma S_t$  and  $\theta_t^{[1]} = \ddot{C}_{SS}(\sigma, S_t)\sigma S_t$ ,

$$L_1 = \sqrt{\frac{T}{2}} \int_0^T \ddot{C}_{SS}(\sigma, S_u)\sigma^2 S_u^2 dW_u^*,$$

where  $C = C(t, x)$  is the Black-Scholes price of an option at time  $t$  with the underlying price  $x$ .

REMARK 3.2.

- (a) Rootzén (1980) obtained a general convergence result for stochastic integrals, which could have various applications in stochastic finance. Indeed, Gobet and Temam (2001) saw its connection to evaluation problems of hedging errors. They have stated weak convergence of the hedging error (at expiration  $T$ ) due to discrete-time delta hedging of a European option to a time-changed Brownian motion in a driftless, time-homogeneous diffusion model (with  $a_t = 0$  and  $\sigma_t = \sigma(S_t)$  for some regular  $\sigma(\cdot)$  in (2.2)):

$$\sqrt{n}L_{1,T}^n \xrightarrow{\mathcal{L}} \hat{W}_\tau$$

where  $\tau := \frac{T}{2} \int_0^T \ddot{C}_{SS}(\sigma, S_u)^2 \sigma^4 S_u^4 du$  and  $\hat{W}$  is a Brownian motion independent of  $\tau$  (see their Theorems 3, p. 360). See also Remark A.1 in Section A.1 of the Appendix.

- (b) Gobet and Temam (2001) computed the *asymptotic variances* of hedging errors at expiration date due to discrete-time hedging strategies in a time-homogeneous diffusion model, and, with the aid of Malliavin calculus, Temam (2003) extended the result to a multidimensional model. Their general finding was that the rates of convergence are  $1/\sqrt{n}$  for options with Lipschitz payoff and slower for options with “irregular” payoff; for example,  $1/n^{1/4}$  for digital options, as opposed to the rate  $1/\sqrt{n}$  for the weak convergence. Thus, it is a nontrivial example that convergence in law does not necessarily imply convergence of variance. (See also a remark after Theorems 1 and 2 of Gobet and Temam 2001.) In the meantime, Geiss (2002) and Geiss and Geiss (2004) have shown that the rate for digital options can be improved up to  $1/\sqrt{n}$ , by an optimal choice of nonequidistant, deterministic trading times. We have not yet investigated nonequidistant cases, however it would be of interest for future research.<sup>5</sup>

*Consequences.* As an important consequence of the weak convergence result (3.2), we may apply the continuous mapping theorem. For instance, the maximum hedging error through the life of the derivative security is given by

<sup>5</sup> We are grateful to the referee for pointing out the references in this Remark, especially Rootzén (1980).

$$\sqrt{n} \max_{0 \leq t \leq T} L_{1,t}^n \xrightarrow{\mathcal{L}} \max_{0 \leq t \leq T} L_{1,t}.$$

*Multiasset Cases.* As in Bertsimas et al. (2000, pp. 194–195), we may generalize the result without difficulty to multiasset cases. Let us focus only on the two-asset case. Let  $S^{(1)}$  be the first stock price modeled by a process of type (2.1), driven by  $W^{(1)} := W^1$  with the coefficients  $\mu^{(1)}$  and  $f^{(1)}$ . Suppose  $S^{(2)}$  is another stock price of type (2.1), driven by a Brownian motion,  $W^{(2)} := \rho_t W^1 + \sqrt{1 - \rho_t^2} W^2$ , with the corresponding coefficients,  $-1 \leq \rho_t \leq 1$ . For simplicity, the correlation  $\rho_t$  is assumed to be adapted and continuous. Let  $L_1^{(l)n}$  be the hedging error of an option  $C^{(l)}$  hedged by the  $l$ th stock  $S^{(l)}$  as per the strategy  $\underline{\theta}^{(l)}$  with the corresponding coefficients  $\underline{\theta}^{(l)[1]}$  and  $\theta^{(l)[0]}$ ,  $l = 1, 2$ .

COROLLARY 3.2. *As  $n \rightarrow \infty$ ,*

$$(\sqrt{n}L_1^{(1)n}, S^{(2)}) \xrightarrow{\mathcal{L}} \left( \sqrt{\frac{T}{2}} \int_0^\cdot \|\underline{\theta}_u^{(1)[1]}\| f_u^{(1)} dW_u^*, S^{(2)} \right).$$

Note that, conditionally on  $\underline{\theta}_u^{(1)[1]}$ ,  $f_u^{(1)}$ , and  $S_u^{(2)}$ , for  $u \in [0, T]$ , the normalized hedging error has zero expectation asymptotically. This implies that  $\sqrt{n}L_1^{(1)n}$  and  $S^{(2)}$  are asymptotically uncorrelated.<sup>6</sup> By the same token, even when there are more than two assets (hence, with more than two Brownian motions), we could show that  $\sqrt{n}L_1^{(1)n}$  is asymptotically uncorrelated with any asset in the market.

Moreover, we have the following corollary.

COROLLARY 3.3. *As  $n \rightarrow \infty$ ,*

$$(\sqrt{n}L_1^{(1)n}, \sqrt{n}L_1^{(2)n}) \xrightarrow{\mathcal{L}} \left( \sqrt{\frac{T}{2}} \int_0^\cdot \|\underline{\theta}_u^{(1)[1]}\| f_u^{(1)} d\tilde{W}_u^1, \sqrt{\frac{T}{2}} \int_0^\cdot \|\underline{\theta}_u^{(2)[1]}\| f_u^{(2)} d\tilde{W}_u^2 \right),$$

where  $\tilde{W} = (\tilde{W}_u^1, \tilde{W}_u^2)$  is a new Brownian motion independent of  $W = (W^1, W^2)$ , and

$$d\langle \tilde{W}^1, \tilde{W}^2 \rangle_u = 1_{\{\|\underline{\theta}_u^{(1)[1]}\| \|\underline{\theta}_u^{(2)[1]}\| \neq 0\}} \rho_u \frac{\underline{\theta}_u^{(1)[1]} \cdot \underline{\theta}_u^{(2)[1]}}{\|\underline{\theta}_u^{(1)[1]}\| \|\underline{\theta}_u^{(2)[1]}\|} du.$$

EXAMPLE 3.2. *European path-independent options in two-dimensional diffusion markets:* Suppose each of  $(S^{(1)}, S^{(2)})$  follows one-dimensional diffusion; that is,

$$dS_t^{(l)} = \mu(t, S_t^{(l)})S_t^{(l)} dt + \sigma(t, S_t^{(l)})S_t^{(l)} dW_t^{(l)} \quad l = 1, 2.$$

Note that there is no direct interaction between the two stocks. Suppose  $C^{(1)}$  is a European path-independent option on  $S^{(1)}$ ,  $C^{(2)}$  on  $S^{(2)}$ , both of which are the same type. Hence,

$$d\theta_t^{(l)} = \theta_t^{(l)[0]} dt + \theta_t^{(l)[1]} dW_t^{(l)} =: \theta_t^{(l)[0]} dt + \underline{\theta}_t^{(l)[1]} dW_t \quad l = 1, 2,$$

where  $\theta_t^{(1)[1]} = \ddot{C}_{SS}\sigma(t, S_t^{(1)})S_t^{(1)}$ ,  $\theta_t^{(2)[1]} = \ddot{C}_{SS}\sigma(t, S_t^{(2)})S_t^{(2)}$ ,  $\underline{\theta}_t^{(1)[1]} = (\theta_t^{(1)[1]}, 0)$ , and  $\underline{\theta}_t^{(2)[1]} = \theta_t^{(2)[1]}(\rho_t, \sqrt{1 - \rho_t^2})$ . In this case,  $d\langle \tilde{W}^1, \tilde{W}^2 \rangle_u = \rho_u^2 du$ .

<sup>6</sup> As Bertsimas et al. (2000) put it (p. 194), this does *not* imply that the asymptotic joint distribution of  $(\sqrt{n}L_1^{(1)n}, S^{(2)})$  does not depend on the correlation  $\rho_t$  between  $W^{(1)}$  and  $W^{(2)}$ . Indeed, it *does*, since this correlation determines the joint distribution of  $S^{(1)}$  and  $S^{(2)}$ .

Therefore, the observation found by Bertsimas et al. (2000) can be ascertained:

- (i) The correlation of the two Brownian motions describing the asymptotic behavior of the two rescaled hedging errors is always nonnegative, regardless of the sign of the covariation of the original  $W^{(1)}$  and  $W^{(2)}$ .
- (ii) When  $C^{(1)}$  and  $C^{(2)}$  have convex payoffs, the hedging errors resulting from hedging on the two stocks  $S^{(1)}$  and  $S^{(2)}$  are asymptotically positively correlated, even if the returns on  $S^{(1)}$  and  $S^{(2)}$  are negatively correlated.

REMARK 3.3. Indeed, when the payoff of a derivative security is a function of multiple assets (e.g., a basket option on stocks), the limiting law of the corresponding hedging error shall depend on “cross-Gammas” (so far as the underlying assets are correlated), especially in Markovian models,  $\frac{\partial^2 \varphi(x,t)}{\partial x_i \partial x_j}$ , where  $\varphi(x, t)$  is the price of a European option as a function of ingredient current stock prices  $x = (x_1, \dots, x_m)$  and the current time  $t$ . See Zhang (1999, Prop. 3.1.6, p. 100) and Temam (2003, Thm. 2.2), who compute the asymptotic variances of the hedging errors (due to discrete-time delta hedging) at terminal time  $T$  for such European options in multidimensional diffusion stock markets. Our approach is inherently capable of dealing with such situations.

#### 4. DATA-DRIVEN REGRESSION STRATEGY AND ITS RELATIVE PERFORMANCE

Suppose the hedger does not know the true hedge ratio  $\theta$ , but needs to hedge the security  $C$  (3.1) at any rate.<sup>7</sup>

The general form  $\theta_t = \frac{d\langle C, S \rangle_t}{d\langle S, S \rangle_t}$  in (2.4) of the true hedge ratio may suggest the following strategy: At every trading time  $t_i$ , hold  $S$  according to

$$(4.1) \quad \tilde{\theta}_i^n := \frac{\langle S, C \rangle_{t_i} - \langle S, C \rangle_{t_{i-1}}}{\langle S, S \rangle_{t_i} - \langle S, S \rangle_{t_{i-1}}}, \quad i = 1, \dots, n - 1,$$

with  $\tilde{\theta}_{t_0}^n := \theta_0$  (known constant) for simplicity. We set  $\tilde{\theta}_i^n := \tilde{\theta}_{t_{i-1}}^n$  whenever the denominator is zero. This strategy is feasible as long as the hedger can observe the cumulative “variance”  $\langle S, S \rangle$  of  $S$  and “covariance”  $\langle S, C \rangle$  of  $S$  and  $C$  at each trading time. Thus, it is a *data-driven* regression strategy. It is a reasonable strategy to take when the size of the trading intervals  $(t_{i+1} - t_i)$  is sufficiently small.

In this case, it is natural to consider the “relative” hedging error of the discrete-time hedging ( $\tilde{\theta}_i^n$ ), in comparison with the true discrete-time strategy ( $\theta_{t_i}$ ) that was considered in the previous section. That is,

$$L_{2,t_j}^n := \sum_{i=0}^{j-1} (\theta_{t_i} - \tilde{\theta}_{t_i}^n) \Delta S_{t_i}, \quad j = 0, 1, \dots, n.$$

As the discreteness disappears, we find its asymptotic distribution as follows.

THEOREM 4.1. *Under suitable regularity conditions,*

$$(\sqrt{n}L_1^n, \sqrt{n}L_2^n) \xrightarrow{\mathcal{L}} \left( \sqrt{\frac{T}{2}} \int_0^\cdot \|\theta_u^{[1]}\| f_u dW_u^{*1}, \sqrt{\frac{T}{3}} \int_0^\cdot \|\theta_u^{[1]}\| f_u dW_u^{*2} \right),$$

<sup>7</sup> Consideration of the interrelations among heterogeneous groups (in light of information) in the market, especially arbitrage opportunities among them, is beyond the scope of this paper.

as  $n \rightarrow \infty$ , where  $W^{*1}$  and  $W^{*2}$  are mutually independent Brownian motions, independent of  $W = (W^1, W^2)$ .

Hence, (i)  $\sqrt{n}L_1^n$  and  $\sqrt{n}L_2^n$  have the same marginal limit law up to a constant multiplier, and (ii) they are asymptotically uncorrelated. The total hedging error of the data-driven strategy  $\tilde{\theta}^n$  can be obtained by the continuous mapping theorem.

REMARK 4.1. When  $\theta_0$  is unknown, the hedger needs to use some estimate  $\tilde{\theta}_0^n$ , the estimation error of which may introduce an additional term.

### 5. REGRESSION STRATEGY WITH DISCRETE OBSERVATIONS

In this section, suppose that prices  $(S, C)$  are observed only at some discrete-time points; in other words, only “snapshots” of the underlying continuous process  $(S_t, C_t)_{0 \leq t \leq T}$  are available. In particular, we consider a fairly liquid market, where the observation times are more frequent than the trading times. For ease of discussion, assume each trade is conducted at every  $m$  observing times. Denote the trading times by  $\{t_i, i = 0, \dots, n\}$  with  $t_0 = 0$  and  $t_n = T$  as before, and the observation times—which subdivide  $\{t_i\}$ —by  $\{t_i^j, j = 0, \dots, m - 1\}$  with  $t_i^0 \equiv t_i$  and  $t_i^m \equiv t_{i+1}^0 \equiv t_{i+1}$ . Equally spaced observation and trading are assumed; the length of each observation interval is  $\Delta t := \frac{T}{mm}$ , and the length of each trading interval is  $h := \frac{T}{n}$ ; that is,  $h = m\Delta t$ .

Due to the discreteness of the observation times, the fair price of the derivative security and its “true” strategy  $\theta$  may be uncertain, and the degree of such uncertainty may depend on the number of observation times  $m$  for each trade. To incorporate this idea, suppose that, at each  $t$ , the “true” value  $C_t^\infty := C_t = C_0 + \int_0^t \theta_t dS_t$  is contaminated by an additive “noise”  $N_t^m$ , which is an adapted continuous semimartingale such that  $\langle S, N^m \rangle_t \equiv 0$  for every  $m \in \mathbb{N}$ . That is,  $C_t^m$  is generated by

$$dC_t^m = \theta_t dS_t + dN_t^m$$

associated with a given, adapted true strategy  $\theta$ . We require  $C_t^m \geq 0$  for all  $t$ —or at least all  $t_i$ —to avoid obvious arbitrage opportunities. Apparently, this covers the previous case (3.1), by setting  $N_t^m \equiv 0$ . In what follows, we consider a family of the models indexed by  $m$  and  $n$ . Note that true  $\theta$  is assumed to be fixed in terms of  $m$ .<sup>8</sup>

For simplicity, the initial values  $\theta_0$  and  $N_0^m$  are constants. The fair value  $C^\infty$  is not observable at any time except at  $t = 0$  when the derivative security is written. Neither is  $N^m$ . (Note that whether  $C_T^\infty$  [or  $N_T^m$ ] is observable will depend on the problem; see Section 6 for example. We do not need to impose such a requirement here.) As  $C^\infty$ , we require  $N^m$  to be bounded from below (w.p.1). For every  $m$ , the hedger cannot hedge  $C^m$  perfectly due to the existence of  $N^m$ .<sup>9</sup>

<sup>8</sup> Market-makers are often said to “lean on the liquid market.” Usually prices of the underlying securities are considerably more reliable; therefore, they quote and trade derivatives while constantly monitoring the underlying, liquid markets so as to execute any of their subsequent hedging operations with these monitored prices. To incorporate this practice, we assume that  $S$  is priced and observed correctly at every discrete-time point  $t_i$  (although market participants may not necessarily know the true, continuous model for  $S$ ), while the option price is perturbed to  $C^m$ , which reflects the price uncertainty. Moreover, so far as hedging problems are concerned, it is sensible to assume that the price of the hedging instrument  $S$  is “given.”

<sup>9</sup> The case when  $C_T^m$  is a square-integrable random variable and  $N^m$  is a square-integrable martingale may be of importance, and might remind readers of the so-called Föllmer-Schweizer decomposition (for every fixed  $m$ ); see Föllmer and Schweizer 1991.

It may be an interesting question whether an equivalent martingale measure,  $Q$ , exists for  $(S, C^m)$  in either discrete or continuous time (especially when  $N_T^m = 0$ ). We do not discuss the issue here. Rather, we are interested in the hedging errors not under  $Q$  but under  $P$  for risk-management purposes, so there will be no need to refer to  $Q$  in the following discussion.

Under this setup, let us consider the “historical” delta-hedging strategy defined by

$$(5.1) \quad \hat{\theta}_i^{n,m} := \frac{\sum_{j=0}^{m-1} (S_{t_{i-1}^{j+1}} - S_{t_{i-1}^j})(C_{t_{i-1}^{j+1}}^m - C_{t_{i-1}^j}^m)}{\sum_{j=0}^{m-1} (S_{t_{i-1}^{j+1}} - S_{t_{i-1}^j})^2}, \quad i = 1, 2, \dots, n-1,$$

with  $\hat{\theta}_0 = \theta_0$  (known constant) for every  $m$  and  $n$ . We will set  $\hat{\theta}_i^{n,m} := \hat{\theta}_{i-1}^{n,m}$  whenever the denominator is zero. This is a modified version of the regression strategy (4.1) introduced in the previous section; it is completely feasible so far as the discretely sampled prices  $(S_{t_i}, C_{t_i}^m)$  are available. In the econometrics literature, the ratio of this form is recently referred to as the “realized beta” (Andersen, *et al.* 2004) or the “realized regression” (Barndorff-Nielsen and Shephard 2004). In our context, (5.1) may alternatively be referred to as the *realized delta* (hedging strategy).<sup>10</sup>

Notice that the regression strategy is “nonoverlapping” in the sense that for each trade at  $t_i \equiv t_i^0$ ,  $i = 1, 2, \dots, n-1$ , only  $(m+1)$  pairs of  $(S, C^m)$  observed in the preceding trading interval  $[t_{i-1}, t_i]$  are at most utilized to compute the new hedge ratio (i.e., there are  $m$  pairs of  $(\Delta S, \Delta C^m)$  to be used). For example, the Black-Scholes delta-hedging strategy with the implied volatility is nonoverlapping since it utilizes only the latest pair of  $(S, C^m)$  for each trade; see Section 6.

It can be shown easily that the “relative” hedging error of the discrete-time hedging ( $\hat{\theta}_i^{n,m}$ ) relative to ( $\hat{\theta}_i^n$ ) of the previous section,  $\sum_{i=0}^{j-1} (\hat{\theta}_i^n - \hat{\theta}_i^{n,m}) \Delta S_{t_i}$ ,  $j = 0, \dots, n$ , is of  $o_P(1)$  as  $m \rightarrow \infty$ , for every  $n \in \mathbb{N}$  (Proposition A.1 in the Appendix). By that reason, Theorem 4.1 essentially carries over. Formally, we have obtained the following asymptotic result of the rescaled hedging error  $\sqrt{n}L^{n,m}$ , as  $m \rightarrow \infty$  and  $n \rightarrow \infty$ .

**THEOREM 5.1.** (Hedging error of data-driven strategy with discrete observations.) *Suppose  $N^m = o_P(1)$  as  $m \rightarrow \infty$ . Under suitable regularity conditions, as  $m \rightarrow \infty$  followed by  $n \rightarrow \infty$ , we have*

$$\sqrt{n}L^{n,m} \xrightarrow{\mathcal{L}} \sqrt{\frac{5}{6}T} \int_0^\cdot \|\hat{\theta}_u^{[1]}\| f_u dB_u,$$

where  $B$  is a new Brownian motion, independent of  $W = (W^1, W^2)$ .

As in Theorem 3.1, we find that the new Brownian motion  $B$  appears in the first-order sense, which is the evidence of market incompleteness (at least, from the hedger’s perspective). This is still true, even when  $C^m$  is a redundant security, in the sense that only  $W^1$  drives the prices  $(S, C^m)$  (see Remark 2.2). We also note that Remark 4.1 to Theorem 4.1 is applicable here.

<sup>10</sup> Hedge ratios are to be interpreted as regression coefficients. See, for example, Avellaneda (1998) or Jacod, Méléard, and Protter (2000). See also Frey and Runggaldier (1999), who consider a risk-minimizing strategy (e.g., Föllmer and Schweizer 1991) in a stochastic volatility model when prices are observed only at discrete random times.

REMARK 5.1. The condition  $N^m = o_P(1)$  as  $m \rightarrow \infty$  means that the noise is sufficiently small. But for the condition the *rescaled* hedging error  $\sqrt{n}L^{n,m}$  would diverge. Because  $C^m \xrightarrow{P} C^\infty$  as  $m \rightarrow \infty$ ,  $C^m$  considered in the statement is a type of such derivative securities that the more active the market ( $m \rightarrow \infty$ ) the more accurately one is able to observe its true value  $C^\infty$  and true ratio  $\theta$ .

We could establish corollaries similar to the ones stated after Theorem 3.1—for example, the following corollary.

COROLLARY 5.1. *Given the whole paths of  $(\theta^{[1]}, f)$ , the limit process  $L$  is a continuous Gaussian process with mean zero and covariance function  $\gamma(s, t) := \int_0^{s \wedge t} \frac{5}{6} T(\|\theta_u^{[1]}\| f_u)^2 du$ .*

The historical (or realized) delta is completely *model-free* regarding both stock and derivative security price processes. This gives the hope that it is robust to misspecification of models and their parameters. Moreover, this strategy is adapted to discrete observations of  $S$  and  $C^m$  in the sense that it intends to keep up with the latest information contained in the data, except for some time lag to collect data and do a reasonable estimation. Any of the existing parametric strategies should outperform this naive strategy in terms of the hedging error, if their use is justifiable. In this sense, the benchmark strategy may be well-utilized to construct the minimum criterion for model selection in practice. We expect that this scheme may serve as a building block for evaluating the hedging errors of a broad class of hedging strategies and for possible comparison among them when the true model is uncertain.

REMARK 5.2. Bermin (2003) investigated a relationship between delta-hedge ratios and Malliavin derivatives in a complete diffusion model (with  $\sigma_t = \sigma(S_t)$  for some suitable  $\sigma(\cdot)$  in (2.2)). The result may indicate a potential usage of Malliavin calculus for the problem we have dealt with. Indeed, recently emerging studies on measuring hedging errors due to discrete-time trading or model misspecification are often seen to utilize Malliavin calculus; for instance, Temam (2001, 2003), Bouleau (2003), and Talay and Zheng (2003).

### 6. THE BLACK-SCHOLES MODEL WITH UNKNOWN VOLATILITY

In this section, we briefly consider the Black-Scholes delta-hedging strategy when the underlying volatility is *uncertain* to the hedger in the proposed framework.

As in Section 5, prices are observed discretely as “snapshots” of the continuous processes. We assume that  $S$  follows a generalized geometric Brownian motion; that is,  $\mu_t = a_t S_t$  and  $f_t = \sigma_t S_t$  in (2.1) with deterministic, time-varying functions  $a_t$  and  $\sigma_t$ .

A European path-independent option is contaminated by an additive noise,  $C_t^m = C_t^\infty + N_t^m$ , with the terminal payoff  $C_T^m = \eta(S_T)$  for some nonnegative, convex payoff  $\eta(\cdot)$ .  $N^m$  is such that  $N_0^m = 0$  and  $N_T^m = 0$  and that it converges in probability to some continuous limit  $N$  (that possesses a certain integrability condition) at the rate of convergence  $\gamma^m$ ; that is,

$$\frac{1}{\gamma^m} N^m \xrightarrow{P} N \quad \text{as } m \rightarrow \infty.$$

This condition will be used to identify the limiting hedging error. The other assumptions in the previous sections may remain untouched. Due to the discreteness,  $C^\infty$  and  $N^m$  are unknown (except when  $t = 0$  and  $T$ ), let alone  $\theta$  is unknown.

Suppose that the fair value is determined by the Black-Scholes formula; in other words,  $C_t^\infty = \varphi(S_t, \Xi_t)$ , where  $\Xi_t := \int_t^T \sigma_u^2 du$  is the cumulative variance for the remaining term and

$$\varphi(S, \Xi) := E \left[ \eta \left( S \exp \left( -\frac{\Xi}{2} + \sqrt{\Xi} Z \right) \right) \right]$$

with  $Z$  a standard normal random variable. Note that  $\varphi(x, t, \cdot)$  is increasing in  $\Xi$  due to the convexity of  $\eta$  (among others, see El Karoui, Jeanblanc-Picque, and Shreve 1998; Hobson 1998). In this case, the hedge ratio process is nothing more than the *Black-Scholes delta*,  $\theta_t = \dot{\varphi}_S(S_t, \Xi_t) =: \theta(\Xi_t)$ . Given observed pairs of prices  $(S_t, C_t^m)$ , the (cumulative) *implied volatility* at time  $t$  is defined to be the unique solution  $\hat{\Xi}_t$  to  $C_t^m = \varphi(S_t, \hat{\Xi}_t)$ , denoted as

$$\hat{\Xi}_t^m := (\varphi(S_t, \cdot))^{-1}(C_t^m) =: h(C_t^m).$$

We suppose that the  $\hat{\Xi}_t^m$  exists for all  $0 \leq t < T$ , w.p.1 (see Remark A.3 in the Appendix).

We are interested in the relative performance of the Black-Scholes delta-hedging strategy against the true one,  $L_{2,t}^{n,m} := \sum_{i=0}^{j-1} (\theta_{t_i} - \hat{\theta}_{t_i}^m) \Delta S_{t_i}$ , where  $\hat{\theta}_t^m := \theta(\hat{\Xi}_t^m)$ .

REMARK 6.1. If indeed,  $N_t^m \equiv 0$ , or, equivalently,  $C_t^m \equiv C_t^\infty$ , the hedger can always make a perfect guess for the future cumulative variance,  $\hat{\Xi}_t^m \equiv \Xi_t$ ; therefore, the plug-in strategy with the implied volatility is optimal; that is, its relative hedging error  $L_{2,t}^{n,m} \equiv 0$ .

REMARK 6.2. The setup here may be generalized to the one provided by Mykland (2000, 2003a, 2003b), where  $a_t$  and  $\sigma_t$  are stochastic.

PROPOSITION 6.1. (Relative hedging error of Black-Scholes delta-hedging strategy.) *Under suitable regularity conditions, the relative hedging error of the Black-Scholes delta-hedging strategy with the implied volatility converges,*

$$(6.1) \quad \frac{1}{\gamma^m} L_{2,t}^{n,m} \xrightarrow{P} - \int_0^t \frac{\ddot{\varphi}_{S\Xi}(S_u, \Xi_u)}{\dot{\varphi}_\Xi(S_u, \Xi_u)} N_u dS_u,$$

as  $m \rightarrow \infty$  then  $n \rightarrow \infty$ .

The performance depends on the path of  $\frac{\ddot{\varphi}_{S\Xi}(S, \Xi)}{\dot{\varphi}_\Xi(S, \Xi)} N$ . If we let  $v_t := \dot{\varphi}_\Xi(S_t, \Xi_t)$ , the Black-Scholes ‘‘Vega,’’ then the coefficient  $\frac{\ddot{\varphi}_{S\Xi}(S_t, \Xi_t)}{\dot{\varphi}_\Xi(S_t, \Xi_t)} = \frac{\partial \ln v_t}{\partial S}$  is its logarithmic sensitivity to  $S$ ; the larger in magnitude, the larger impact the incorrect input for  $\Xi_t$  yields.

EXAMPLE 6.1. *European call options:* In this case,  $v_t = \frac{S_t \phi(d_t)}{2\sqrt{\Xi_t}}$ , where  $\phi(\cdot)$  the standard normal p.d.f. and  $d_t := \frac{\ln(S_t/K) + \Xi_t/2}{\sqrt{\Xi_t}}$ , so  $\frac{\ddot{\varphi}_{S\Xi}(S_t, \Xi_t)}{\dot{\varphi}_\Xi(S_t, \Xi_t)} = -\frac{\ln(S_t/K) - \Xi_t/2}{S_t \Xi_t}$  (recall that the interest rate is zero). We need to note that the sensitivity coefficient  $\frac{\ddot{\varphi}_{S\Xi}(S_t, \Xi_t)}{\dot{\varphi}_\Xi(S_t, \Xi_t)}$  diverges a.s. as  $t \rightarrow T$ ; hence  $N_t$  must be such that it goes to zero sufficiently fast so that the limit in (6.1) exists (see Remark A.3 in the Appendix). The proposition says that, if  $S$  stays in the neighborhood of the near-at-the-money curve,  $S_t = Ke^{\Xi_t/2}$ , around which the coefficient changes its sign, and if the pricing noise  $N$  fluctuates around the origin, then the error would be relatively small in magnitude (at least, until  $t$  gets closer to  $T$ ). On the other hand, if  $S$  stays either in-the-money or out-of-the-money, and if the market price  $C^m$  is always underpriced or overpriced relative to the true  $C^\infty$ , the error magnitude tends to be large.

In this setup, the order of the relative hedge error of the delta-hedging strategy is small, compared to the discrete trading error. Hence, the Black-Scholes-IV strategy produces the overall hedging error,  $L^{BS}$ , approximately (in law)

$$L_t^{BS} \simeq \sqrt{\frac{T}{2n}} \int_0^t \ddot{\psi}_{SS}(S_u, \Xi_u) \sigma_u^2 S_u^2 dB_u,$$

according to Theorem 3.1; on the other hand, the regression strategy yields the overall hedging error,  $L^{RG}$ , approximately

$$L_t^{RG} \simeq \sqrt{\frac{5T}{6n}} \int_0^t \ddot{\psi}_{SS}(S_u, \Xi_u) \sigma_u^2 S_u^2 dB'_u,$$

according to Theorem 5.1. Here,  $B$  and  $B'$  are Brownian motions independent of  $W = (W^1, W^2)$ . The limits are the same in law, up to a constant multiplier. Hence, in the liquid market under consideration their difference is not substantial. Another case—when the relative error of the delta hedging and the discretization error are comparable in magnitude—requires further investigation.

## 7. CONCLUSIONS

We have proposed a methodology for evaluating the hedging errors of derivative securities due to the discreteness of trading times or observation times of market prices, or both. The setup covers important cases such as stochastic volatility models or non-Markovian models, as well as one-dimensional diffusion models. The study is an initial attempt toward the development of a new framework under which various hedging strategies could be evaluated in a unified manner when the true model is unknown.

There are several directions in which we may proceed. The approach developed here could be integrated into a *model selection* problem. To this purpose we need to provide utility functions and sensible optimality criteria. The development of “better” trading strategies is challenging but interesting for both theoretical and practical purposes.

The framework could naturally incorporate transaction costs (as a relevant work, see Pergamenschikov 2003). It would be interesting to develop “optimal” strategies in the presence of transaction costs.

Some empirical studies point out that stock prices jump. Although allowing for jumps induces further technical complexity, the relaxation of the continuous path assumption (to general semimartingales) may also be worth pursuing.

## APPENDIX

### A.1. Limit Theorem

Before providing proofs of the main results of the paper, we establish a version of martingale central limit theorem, which will be invoked in those proofs.<sup>11</sup>

The notations and definitions in this paper are essentially those of Jacod and Shiryaev (1987), hereafter referred to as J-S. Let  $T < \infty$  be a terminal time. We assume that there is

<sup>11</sup> See Jacod (1997) for relevant convergence results. See also Genon-Catalot and Jacod (1993) and Yoshida (1997, sect. 6.3) for statistical inference problems for diffusion-type processes. We are grateful to Professors Shigeo Kusuoka and Nakahiro Yoshida for some technical comments.



a  $d$ -dimensional Brownian motion  $W := (W^1, \dots, W^d)'$  on a probability space  $(\Omega, \mathcal{F}, P)$ , endowed with  $\mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]}$ , the natural filtration of  $W$ . Let  $M^n$  be an  $\mathbb{R}^m$ -valued continuous local martingale starting from  $M_0^n := 0$ , with  $K^n := \langle M^n, M^n \rangle$ , which is an  $\mathbb{R}^{m \times m}$ -valued predictable process with components  $\langle M^n, M^n \rangle^{i,j} = \langle M^{n,i}, M^{n,j} \rangle$ .

THEOREM A.1.

(1) Suppose that

(a)  $K_t^n \xrightarrow{P} K_t$  for every  $t \in [0, T]$ , where  $K$  is  $\mathbb{R}^{m \times m}$ -valued, symmetric, and semi-positive definite;

(b)  $\langle M^n, W \rangle_t \xrightarrow{P} 0$  for every  $t \in [0, T]$ ;

(c)  $K = \int_0^\cdot k_s ds$  with  $k$  predictable,  $\mathbb{R}^{m \times m}$ -valued.

Then,  $(M^n, W) \xrightarrow{\mathcal{L}} (M, W)$ , where  $M$  is a continuous local martingale, with the representation

$$M = \int_0^\cdot z_s dW_s^*,$$

where  $z_s$  is an  $\mathbb{R}^{m \times m}$ -valued “square-root matrix of  $k$ ” process such that  $zz' = k$ , and  $W^*$  is an  $m$ -dimensional Brownian motion independent of  $W$ .

(2) Suppose further that  $A^n$  and  $A$  are  $\mathbb{R}^m$ -valued,  $\mathbb{F}$ -adapted, continuous processes such that  $\sup_{t \in [0, T]} |A_t^n - A_t| \xrightarrow{P} 0$ . Then

$$(A^n + M^n, W) \xrightarrow{\mathcal{L}} (A + M, W),$$

where  $M$  is a continuous local martingale, with the representation appearing in (1).

In both (1) and (2), the convergence takes place on the space of continuous functions,  $\mathbb{C}([0, T] \rightarrow \mathbb{R}^{m+d})$ , endowed with the uniform topology.

Proof of Theorem A.1.

*Proof of (1):* We are going to adopt a subsequence argument. For this we will apply various results from J-S; hence, as in J-S, convergence will be discussed first on the space of càdlàg functions,  $D([0, T] \rightarrow \mathbb{R}^{m+d})$ , endowed with the Skorohod topology (J-S, Thm. VI.1b, pp. 291ff). Note that since  $M^n$  and  $W$  are continuous processes, their optional and predictable quadratic variations coincide. Let  $G^n := \text{trace}(K^n) = \sum_{j \leq m} \langle M^{n,j}, M^{n,j} \rangle$  and  $G := \text{trace}(K)$ .

(I) First show that  $(M^n, W)$  are jointly tight. Note by the setup that  $(G^n)$  is continuous, increasing, and  $G_0^n = 0$ ; so is  $G$  by (a) and (c). Besides, since by (a)  $(G^n)$  converges in probability to  $G$  for some dense subset of  $[0, T]$ , J-S’s Theorem VI.2.15(c) implies that  $(G^n)$  converges in probability to  $G$ . Therefore,  $(G^n)$  converges in law to  $G$ .

Therefore,  $(G^n)$  is tight. Moreover, since  $G$  is continuous, it is  $C$ -tight (using the terminology of J-S, Thm. VI.3.25, p. 315). Hence, by J-S Theorem VI.4.13, p. 322,  $(M^n)$  is tight since  $(M_0^n) \equiv (0)$  is trivially tight.

Finally,  $(M^n, W)$  is jointly tight, since  $(W)$  is “fixed.”

(II) We will show that  $(M^n, W)$  have a weak limit, jointly with the quadratic variations  $[M^n, M^n], [M^n, W]$ , and  $[W, W]$ . Since  $(M^n, W)$  is tight, for all subsequences  $N' \subset \mathbb{N}$  there exists a further subsequence  $N'' \subset N'$  such that  $(M^n, W), n \in N''$ , converges in law to  $(M, W)$ , where  $M$  is a limit of  $M^n$  corresponding to the further subsequence  $N''$ . Then, J-S Theorem VI.6.7 (p. 342) implies that

$$(A.1) \quad (M^n, W, [M^n, M^n], [M^n, W], [W, W]) \xrightarrow{\mathcal{L}} (M, W, [M, M], [M, W], [W, W]),$$

for such  $n \in N''$ . Let  $A^n := (M^n, W, [M^n, M^n], [M^n, W], [W, W])$  and  $A$  be the limit.

We denote by  $(\bar{\Omega}, \bar{\mathcal{F}}_T, \bar{P})$  a probability space on which  $A$  (is defined and) has the distribution equivalent to the limiting law in (A.1). This might be the canonical space consisting of the appropriate number of càdlàg functions. For the time being we will write deliberately the random elements on  $\bar{\Omega}$  as  $\bar{M}, \bar{W}$ , or  $\bar{A}$ , for example, so as to distinguish them from the ones on  $\Omega$ . Let the space endowed with  $(\bar{\mathcal{F}}_t)$ , be the filtration generated by  $\bar{A}$ . Let  $(\bar{\mathcal{G}}_t)$  be the filtration generated by  $\bar{W}$ .

(III) We will show absolute continuity of the quadratic variations in the limiting law. Note from (a) and (c) that

$$(A.2) \quad (W, [M^n, M^n]) \xrightarrow{P} (W, K), \quad n \in \mathbb{N},$$

and

$$[M, M]_t = K_t = \int_0^t k_s ds, \quad 0 \leq t \leq T,$$

$P$ -a.s. on  $\Omega$ . On the other hand, (A.1) implies that

$$(A.3) \quad (W, [M^n, M^n]) \xrightarrow{\mathcal{L}} (\bar{W}, [\bar{M}, \bar{M}]), \quad n \in N''.$$

Set  $\bar{K} := [\bar{M}, \bar{M}]$ . Then, the uniqueness of the limiting law in (A.3) (along with  $N''$ ), together with (A.2), implies the equivalence of the joint law,

$$(A.4) \quad \mathcal{L}(W, K | P) = \mathcal{L}(\bar{W}, \bar{K} | \bar{P}).$$

If we write  $K =: F(W)$ , where  $F$  is a functional of the Brownian motion  $W$ , then we must have  $\bar{K} = F(\bar{W})$ . Therefore,  $\bar{K}$  must be absolutely continuous, especially,  $[\bar{M}, \bar{M}]_t = \int_0^t \bar{k}_s ds, 0 \leq t \leq T, \bar{P}$ -a.s., for  $(\bar{\mathcal{G}}_t)$ -predictable,  $R^m$ -valued  $\bar{k}$  on  $(\bar{\Omega}, \bar{\mathcal{F}}_T, \bar{P})$ .

At the same time, by (b) it is obvious that  $[\bar{M}, \bar{W}]_t = 0$ , which is the limit of  $[M^n, W]_t, n \in N''$ , and  $[\bar{W}, \bar{W}]_t = tI_d, 0 \leq t \leq T (\bar{P}$ -a.s.).

For notational convenience, throughout the rest of the proof we will express all the random elements on  $\bar{\Omega}$  as  $M, W$ , and so forth, as there is no likelihood of confusion.

(IV) We will show that  $M = (M^j)_{1 \leq j \leq d}$  is a continuous local martingale with respect to  $(\bar{\mathcal{F}}_t)$  on  $(\bar{\Omega}, \bar{\mathcal{F}}_T, \bar{P})$ . Since  $[M^j, M^j] = \int_0^t k_u^{j,j} du, \Delta[M^j, M^j] = 0$ , therefore,  $(\Delta M^j)^2 = 0$  by J-S Theorem I.4.47(c) (p. 52).

Since  $M^{n,j}$  is a (one-dimensional) local martingale and is continuous, and  $A^n \xrightarrow{\mathcal{L}} A (n \in N'')$ ,  $M^j$  is a (one-dimensional) local martingale with respect to the filtration generated by  $A$ , by invoking J-S Prop. IX.1.17 (p. 485).

It follows that  $M = (M^j)_{1 \leq j \leq m}$  is a continuous local martingale with respect to the filtration generated by  $A$ —that is, by  $(\bar{\mathcal{F}}_t)$ .

- (V) We will show that on an extended probability space,  $M_t$  has the following representation:

$$(A.5) \quad M = \int_0^\cdot z_s dW_s^*,$$

where  $z$  is an  $\mathbb{R}^{m \times m}$ -valued “square-root matrix of  $k$ ” process such that  $z_t z_t' = k_t$ , for all  $t$ , and  $W^*$  is an  $m$ -dimensional Brownian motion independent of  $W$ .

To this end, first we diagonalize  $k$ , writing  $k_t = \pi_t' \lambda_t \pi_t$ , where  $\pi_t = (\pi_t^{i,j})$  is an orthogonal  $m \times m$  matrix-valued process and  $\lambda_t$  is a diagonal  $m \times m$  matrix-valued process whose  $(i, i)$ th element is  $\lambda_t^i$ ; both are predictable with respect to  $(\bar{\mathcal{G}}_t)$ .

We can always enlarge the space  $\bar{\Omega}$  and the filtration  $(\bar{\mathcal{F}}_t)$  so that there is a Brownian motion  $Z = (Z^j)_{1 \leq j \leq m}$  that is independent of  $A$ .

Set  $W^* = (W_t^{*i})_{1 \leq i \leq m}$  by

$$(A.6) \quad W_t^{*i} := \int_0^t 1_{\{\lambda_s^i > 0\}} (\lambda_s^i)^{-\frac{1}{2}} \sum_{j=1}^m \pi_s^{j,i} dM_s^j + \int_0^t 1_{\{\lambda_s^i = 0\}} dZ_s^i, \quad 1 \leq i \leq m.$$

Obviously,  $W_t^* = (W_t^{*i})_{1 \leq i \leq m}$  is a vector of continuous local martingales with respect to  $(\bar{\mathcal{F}}_t, \bar{P})$ . (Recall that  $M$  in (A.6) is the weak limit of  $M^n$  along the subsequence chosen  $N''$ .)

By standard calculations, we can check that

$$\langle W^{*i}, W^{*j} \rangle_t = \delta_{i,j} t$$

and

$$\langle W^{*i}, W^j \rangle_t = 0,$$

so that, by Levy’s Theorem (e.g., see Karatzas and Shreve 1991, Thm. 3.16, p. 157), on  $(\bar{\Omega}, \bar{\mathcal{F}}_T, (\bar{\mathcal{F}}_t), \bar{P})$ ,  $W^* = (W^{*i})_{1 \leq i \leq m}$  is an  $m$ -dimensional Brownian motion, independent of  $W = (W^j)_{1 \leq j \leq d}$ .

We also easily get

$$(A.7) \quad M_t^i = \int_0^t \sum_{j=1}^m \pi_s^{j,i} \sqrt{\lambda_s^j} dW_s^{*j}, \quad 1 \leq i \leq m.$$

Hence, the representation (A.5) for the  $m$ -dimensional continuous  $(\bar{\mathcal{F}}_t, \bar{P})$ -local martingales  $M = (M^j)_{1 \leq j \leq m}$  is established (by putting  $z_s := (z_s^{i,j})_{1 \leq i,j \leq m}$  with  $z_s^{i,j} := \pi_s^{j,i} \sqrt{\lambda_s^j}$ ).

- (VI) We next establish the (global) uniqueness of the weak limit. Since  $z_s$  is  $(\bar{\mathcal{G}}_t)$ -adapted, it is independent of  $W^*$ . It follows that, on  $(\bar{\Omega}, \bar{\mathcal{F}}_T, (\bar{\mathcal{F}}_t), \bar{P})$ , conditional on  $\bar{\mathcal{G}}_T$ ,  $M$  is a Gaussian continuous martingale with (deterministic) bracket  $[M, M] = \int_0^\cdot k_u du$ . Trivially,  $[M, W] = 0$  and  $[W, W] = \cdot I_d$  are  $\bar{\mathcal{G}}_T$ -measurable. Therefore, the law  $\mathcal{L}(A | \bar{P})$  is uniquely determined.

Since every subsequence  $(M^n, W)$ ,  $n \in N' (\subset \mathbb{N})$ , has a further subsequence  $N'' \subset N'$  converging in law to the unique  $(M, W)$ , we must have that  $(M^n, W)$  converges in law to  $(M, W)$ .

Finally, since both the prelimiting processes  $(M^n, W)$  and the limit  $(M, W)$  are continuous, the convergence indeed takes place on the space of continuous functions,  $C([0, T] \rightarrow \mathbb{R}^{m+d})$ , endowed with the uniform topology.

*Proof of (2).* For convenience, put  $X^n := (M^n, W)$  as well as  $X := (M, W)$ . Notice that  $(A^n, X^n) = (A, X^n) + (A^n - A, 0)$ . Then, noting that  $A$  is an  $\mathbb{F}$ -adapted, “fixed” process, by the same argument as the proof of (1) above, we can show that  $(A, X^n) \xrightarrow{\mathcal{L}} (A, X)$ , while by hypothesis  $(A^n - A, 0) \xrightarrow{P} (0, 0)$ . Hence, by invoking J-S Lemma VI.3.31 (p. 316),

$$(A^n, X^n) \xrightarrow{\mathcal{L}} (A, X).$$

Therefore, by the continuous mapping theorem, we conclude that

$$A^n + M^n \xrightarrow{\mathcal{L}} A + \int_0^\cdot z_s dW_s^*,$$

jointly with  $W$ , where  $W^*$  is a new Brownian motion independent of  $W$ . □

REMARK A.1. The convergence result given by Rootzén (1980)—convergence to a time-changed Brownian motion—is based on the martingale representation as a time-changed Brownian motion (e.g., Karatzas and Shreve 1991, Thm. 3.4.6, p. 174), whereas ours is based on the representation as a stochastic integral with respect to Brownian motion. For the former, one only requires  $K(t) = \langle M, M \rangle_t$  to be continuous and not *absolutely* continuous. Although having the stochastic integral representation seems more desirable in finance, it is worthwhile to check whether one can use Rootzén’s techniques to improve and shorten the arguments here.

### A.2. Preliminaries for the Proofs in Sections 3-6

Let  $T < \infty$  be a terminal time. Let  $\mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]}$  be the natural filtration generated by a two-dimensional Brownian motion  $W := (W^1, W^2)'$ , satisfying the usual conditions. Unless otherwise stated, convergence is considered on the space of continuous functions, endowed with the uniform topology.

The total hedging error  $L^n := (L_t^n)$  of an arbitrary self-financing discrete-time trading strategy  $\pi$  is defined, as a discrete-time process, by

$$L_t^n := \int_0^{t_j} \theta_u dS_u - \sum_{i=0}^{j-1} \pi_{t_i} \Delta S_{t_i}, \quad j = 0, 1, \dots, n.$$

In order to treat the error in continuous time, we interpolate it as a continuous process:

$$(A.8) \quad L_t^n := \int_0^t \theta_u dS_u - \sum_{i=0}^{n-1} \pi_{t_i} \Delta S_{t_i \wedge t},$$

where  $\Delta S_{t_i \wedge t} := S_{t_{i+1} \wedge t} - S_{t_i \wedge t}$ . Note that in Section 5 we will use  $\{t_i^j, i = 0, \dots, n-1, j = 0, \dots, m-1\} \cup \{t_n\}$  in place of  $\{t_i, i = 0, \dots, n\}$  as the observing times. The relative hedging error  $L_2^n := (L_{2, t_i}^n)$  in discrete time can be interpolated similarly. Here,  $\sum_{i=0}^{-1} := 0$  by convention.

Although the continuous version (A.8) is  $\mathbb{F}$ -adapted, it is not  $\bar{\mathbb{F}} := (\bar{\mathcal{F}}_t)$ -adapted, where  $\bar{\mathbb{F}} := (\bar{\mathcal{F}}_t)_{t \in [0, T]}$  with  $\bar{\mathcal{F}}_t := \mathcal{F}_{t_i}, t \in [t_i, t_{i+1})$ , is the “piecewise-constant” filtration. Indeed,  $(L_t^n)$  is “observed” by the hedger only at every  $t_i$ .  $(L_t^n)$  may be interpreted as the “real-time,” “market-to-market” value of the hedging error.

Let

$$\mathcal{L}^1 := \{f : \mathbb{F}\text{-progressively measurable and } \int_0^T |f_u| du < \infty \text{ a.s.}\},$$

$$\mathcal{L}^2 := \{f : \mathbb{F}\text{-progressively measurable and } \int_0^T f_u^2 du < \infty \text{ a.s.}\}.$$

Here is the summary of the basic assumptions stated previously:

- (i) The bond price is constant; that is, the interest rate is zero.
- (ii) The stock price  $S$  is a positive Itô process such that

$$dS_t = \mu_t dt + f_t dW_t^1,$$

where  $S_0$  is a constant,  $\mu \in \mathcal{L}^1$ , and  $f \in \mathcal{L}^2$  with  $f > 0$ .

- (iii) The true strategy  $\theta$  is an Itô process such that

$$(A.9) \quad d\theta_t = \theta_t^{[0]} dt + \theta_t^{[11]} dW_t^1 + \theta_t^{[12]} dW_t^2 (\equiv \theta_t^{[0]} dt + \underline{\theta}_t^{[1]} dW_t),$$

where  $\theta_0$  is a constant,  $\theta^{[0]} \in \mathcal{L}^1$ , and  $\theta^{[1]} = (\theta^{[11]}, \theta^{[12]}) \in (\mathcal{L}^2)^2$ .

- (iv) The European derivative security  $C$  associated with  $\theta$  is defined by

$$dC_t = \theta_t dS_t,$$

with  $C_t \geq 0$  for all  $t$ , and  $C_0$  is a constant.

- (v) The exogenous noise  $N^m$  is an  $\mathbb{F}$ -adapted, continuous semimartingale with  $N_0^m$  constant such that  $\langle S, N^m \rangle = 0$ , for every  $m \in \mathbb{N}$ . (This is used in the proofs for Sections 5 and 6.)

We have further assumed that the coefficients are also Itô processes.

- (vi)  $\underline{\theta}^{[1]}$  is a two-dimensional Itô process such that

$$(A.10) \quad d\underline{\theta}_t^{[1]} = \underline{\theta}_t^{[10]} dt + \underline{\theta}_t^{[11]} dW_t,$$

where  $\underline{\theta}_0^{[1]}$  is a constant,  $\underline{\theta}^{[10]} \in (\mathcal{L}^1)^2$ , and  $\underline{\theta}^{[11]} \in (\mathcal{L}^2)^{2 \times 2}$ .

- (vii)  $f$  is an Itô process such that

$$(A.11) \quad df_t = f_t^{[0]} dt + \underline{f}_t^{[1]} dW_t,$$

where  $f_0$  is a constant,  $f^{[0]} \in \mathcal{L}^1$ , and  $\underline{f}^{[1]} \in (\mathcal{L}^2)^2$ . Also,  $f^{[0]}$ , and  $\underline{f}^{[1]}$  are Itô processes with the corresponding forms.

In particular, when those processes are defined as solutions to stochastic differential equations, they ought to be *strong* solutions; that is, they are expressible as functionals of the given  $W$ . To avoid trivial cases, assume that  $\int_0^t \|\underline{\theta}_u^{[10]}\| du + \int_0^t \|\underline{\theta}_u^{[11]}\|^2 du > 0$  and  $\int_0^t |f_u^{[0]}| du + \int_0^t \|\underline{f}_u^{[1]}\|^2 du > 0$ , for  $t > 0$ . The Euclidean norm of a vector  $x \in \mathbb{R}^d$  is denoted by  $\|x\| := \sqrt{\sum_{i=1}^d (x^i)^2}$ .  $O_P(\cdot)$  and  $o_P(\cdot)$  may mean to take place *uniformly* in  $t \in [0, T]$  when they are referring to stochastic processes.

Furthermore, assume also that  $\theta^{[0]}$ ,  $\underline{\theta}^{[10]}$ ,  $\underline{\theta}^{[11]}$ , and  $\mu$  are locally bounded. In this setup, by localization, we are able to make all the above processes bounded. Therefore, in what follows we may and will treat them as bounded. This is without loss of generality since our convergence is implicitly stable.

A.3. Proofs in Sections 3 to 6

A.3.1. Proof of Theorem 3.1. We will explore the limit of the continuously interpolated version of rescaled  $(L_{1,t}^n)$ ,

$$\begin{aligned}
 \text{(A.12)} \quad \sqrt{n}L_{1,t}^n &= \sqrt{n} \left( \int_0^t \theta_u dS_u - \sum_{i=0}^{n-1} \theta_{t_i} \Delta S_{t_i \wedge t} \right) \\
 &= \sqrt{n} \sum_{i=0}^{n-1} \int_{t_i \wedge t}^{t_{i+1} \wedge t} (\theta_u - \theta_{t_i \wedge t}) \mu_u du + \sqrt{n} \sum_{i=0}^{n-1} \int_{t_i \wedge t}^{t_{i+1} \wedge t} (\theta_u - \theta_{t_i \wedge t}) f_u dW_u^1 \\
 &= \int_0^t \psi_u^n \eta_u du + \int_0^t \psi_u^n dW_u^1 =: A_{1,t}^n + M_{1,t}^n,
 \end{aligned}$$

where  $\Delta S_{t_i \wedge t} := S_{t_{i+1} \wedge t} - S_{t_i \wedge t}$ ,  $\psi_u^n := \sqrt{n} \sum_{i=0}^{n-1} (\theta_t - \theta_{t_i}) f_t 1_{\{t_i < t \leq t_{i+1}\}}$ , and  $\eta_t := \frac{\mu_t}{f_t}$ . Consider first the martingale part  $M_1^n$  in (A.12).

(I) (i) We identify the limit of its quadratic variation process as

$$K^n(t) := \langle M_1^n, M_1^n \rangle_t = \frac{T}{2} \int_0^t (\|\theta_u^{[1]}\| f_u)^2 du + o_P(1).$$

Notice that

$$M_{1,t}^n = \sqrt{n} \sum_i \int_{t_i \wedge t}^{t_{i+1} \wedge t} (\theta_u - \theta_{t_i \wedge t}) f_u dW_u^1$$

is a continuous martingale. Consider its quadratic variation,

$$\begin{aligned}
 \langle M_1^n, M_1^n \rangle_t &= n \sum_i \int_{t_i \wedge t}^{t_{i+1} \wedge t} (\theta_u - \theta_{t_i \wedge t})^2 f_{t_i \wedge t}^2 du + o_P(1) \\
 &= n \sum_i \int_{t_i \wedge t}^{t_{i+1} \wedge t} \left( \int_{t_i \wedge t}^u \underline{\theta}_s^{[1]} dW_s + \int_{t_i \wedge t}^u \theta_s^{[0]} ds \right)^2 f_{t_i \wedge t}^2 du + o_P(1) \\
 &= n \sum_i \int_{t_i \wedge t}^{t_{i+1} \wedge t} \left( \int_{t_i \wedge t}^u \underline{\theta}_s^{[1]} dW_s \right)^2 f_{t_i}^2 du + o_P(1).
 \end{aligned}$$

Because  $\theta$  and  $f$  are Itô processes ((A.9) and (A.11)), these computations can be carried out by decomposing terms appropriately, then by evaluating the magnitude of each by means of Burkholder’s inequality, and so on. By applying Lemma A.4 to the principal term, one has

$$\langle M_1^n, M_1^n \rangle_t = n \sum_{i=0}^{n-1} \frac{1}{2} (\|\underline{\theta}_{t_i}^{[1]}\| f_{t_i})^2 (\Delta t_i \wedge t)^2 + o_P(1)$$

with  $\Delta t_i \wedge t := t_{i+1} \wedge t - t_i \wedge t$ .

Therefore,

$$\text{(A.13)} \quad K^n(t) := \langle M_1^n, M_1^n \rangle_t \xrightarrow{P} K(t) := \frac{T}{2} \int_0^t (\|\theta_u^{[1]}\| f_u)^2 du,$$

since  $\theta^{[1]}$  and  $f$  are continuous on  $[0, T]$ . (Recall that  $\Delta t_i \equiv \frac{T}{n}$ .) For the time being, assume we have proved

$$\text{(A.14)} \quad \int_0^t \psi_u^n du = o_P(1).$$

(ii) For the covariation of  $M_1^n$  and  $W = (W^1, W^2)$ , it is immediate that

$$\langle M_1^n, W^1 \rangle_t = \sqrt{n} \sum_{i=0}^{n-1} \int_{t_i \wedge t}^{t_{i+1} \wedge t} (\theta_u - \theta_{t_i \wedge t}) f_u du = \int_0^t \psi_u^n du = o_P(1).$$

Also, since  $W_t^1$  and  $W_t^2$  are independent,

$$\langle M_1^n, W^2 \rangle_t = \sqrt{n} \sum_{i=0}^{n-1} \int_{t_i \wedge t}^{t_{i+1} \wedge t} (\theta_u - \theta_{t_i \wedge t}) f_u d\langle W^1, W^2 \rangle_u = 0.$$

(II) Next we consider  $A_1^n$ . We claim that  $A_1^n = o_P(1)$ ; that is,

$$(A.15) \quad \sup_{0 \leq t \leq T} \left| \int_0^t \psi_u^n \eta_u du \right| \xrightarrow{P} 0.$$

For showing this, observe first that  $\eta_u \equiv \frac{\mu_u}{J_u}$  satisfies  $\int_0^T \eta_u^2 du < \infty$ , a.s. Note also that (A.15) holds for  $\eta$  of the form

$$\eta^{(k)} := \sum_{j=0}^{\lfloor \frac{k}{T} \rfloor - 1} \eta_j 1_{\left\{ \frac{j}{k} \leq t < \frac{j+1}{k} \right\}}, \quad k = 1, 2, \dots,$$

because, for every fixed  $k$ ,

$$\begin{aligned} \sup_{0 \leq t \leq T} \left| \int_0^t \psi_u^n \eta_u^{(k)} du \right| &= \sup_{0 \leq t \leq T} \left| \sum_{j=0}^{\lfloor \frac{k}{T} \rfloor - 1} \eta_j \int_{\frac{j}{k}}^{\frac{j+1}{k}} \psi_u^n du \right| \leq \sup_{0 \leq t \leq T} \sum_{j=0}^{\lfloor \frac{k}{T} \rfloor - 1} \left| \eta_j \int_{\frac{j}{k}}^{\frac{j+1}{k}} \psi_u^n du \right| \\ &= \sum_{j=0}^{k-1} |\eta_j| \left| \int_{\frac{j}{k}}^{\frac{j+1}{k}} \psi_u^n du \right| \xrightarrow{P} 0, \end{aligned}$$

as  $n \rightarrow \infty$ , by (A.14).

Further, for arbitrary  $\eta$  with  $\int_0^T \eta_u^2 du < \infty$ , a.s., it is possible to find a sequence  $\{\eta^{(k)}\}$  of this form such that  $\int_0^t (\eta_u - \eta_u^{(k)})^2 du \xrightarrow{P} 0$  as  $k \rightarrow \infty$ .

Thus, by Cauchy-Schwarz's inequality and (A.13), as  $n \rightarrow \infty$ ,

$$\begin{aligned} \sup_{0 \leq t \leq T} \left| \int_0^t \psi_u^n \eta_u du \right| &\leq \sup_{0 \leq t \leq T} \left| \int_0^t \psi_u^n \eta_u^{(k)} du \right| + \sup_{0 \leq t \leq T} \left| \int_0^t \psi_u^n (\eta_u - \eta_u^{(k)}) du \right| \\ &\leq \sup_{0 \leq t \leq T} \left| \int_0^t \psi_u^n \eta_u^{(k)} du \right| + \left( \sup_{0 \leq t \leq T} \int_0^t (\psi_u^n)^2 du \right)^{\frac{1}{2}} \\ &\quad \times \left( \sup_{0 \leq t \leq T} \int_0^t (\eta_u - \eta_u^{(k)})^2 du \right)^{\frac{1}{2}} \\ &\xrightarrow{P} 0 + \left( K(T) \cdot \int_0^T (\eta_u - \eta_u^{(k)})^2 du \right)^{\frac{1}{2}}. \end{aligned}$$

Letting  $k \rightarrow \infty$  proves (A.15).

It remains to show (A.14). By the same principle mentioned in (I)(i) above, via approximations  $(\theta_u - \theta_{t_i}) \simeq \int_{t_i}^u \underline{\theta}_s^{[1]} dW_s$  and  $f_u \simeq f_{t_i}$  for  $u \in [t_i, t_{i+1}]$  due to (A.9) and (A.11), one can carry out the computation as

$$\begin{aligned} \int_0^t \psi_u^n du &= \sqrt{n} \sum_i \int_{t_i \wedge t}^{t_{i+1} \wedge t} (\theta_u - \theta_{t_i \wedge t}) f_u du \\ &= \sqrt{n} \sum_i \int_{t_i \wedge t}^{t_{i+1} \wedge t} \left( \int_{t_i \wedge t}^u \underline{\theta}_s^{[1]} dW_s \right) f_{t_i \wedge t} du + o_P(1). \end{aligned}$$

By direct application of Lemma A.3 to the principal term, one can ascertain that

$$\int_0^t \psi_u^n du = o_P(1).$$

Therefore, (A.14) is shown.

(III) Finally, putting together the results of (I) and (II), we can conclude, via Theorem A.1(2), that

$$\sqrt{n} L_1^n \equiv A_1^n + M_1^n \xrightarrow{\mathcal{L}} \sqrt{\frac{T}{2}} \int_0^t \|\theta_u^{[1]}\| f_u dW_u^*,$$

jointly with  $W = (W^1, W^2)$ , where  $W^*$  is a new Brownian motion independent of  $W$ . □

*A.3.2. Proof of Corollary 3.3.* Consider cases when  $W := (W^1, W^2)$  and  $\underline{\theta}^{(1)[1]}$  and  $\underline{\theta}^{(2)[1]}$  are two-dimensional vector valued. Since

$$M_{1,t}^{(l)n} := \sqrt{n} \sum_i \int_{t_i \wedge t}^{t_{i+1} \wedge t} (\theta_u^{(l)} - \theta_{t_i \wedge t}^{(l)}) f_u^{(l)} dW_u^{(l)}, \quad l = 1, 2,$$

where  $W_t^{(1)} := W_t^1$ ,  $W_t^{(2)} := \rho_t W_t^1 + \sqrt{1 - \rho_t^2} W_t^2$ , and  $\rho_t$  is continuous and adapted, we can show that

$$\begin{aligned} \langle M_1^{(1)n}, M_1^{(2)n} \rangle_t &= n \sum_i \int_{t_i \wedge t}^{t_{i+1} \wedge t} (\theta_u^{(1)} - \theta_{t_i \wedge t}^{(1)}) (\theta_u^{(2)} - \theta_{t_i \wedge t}^{(2)}) f_u^{(1)} f_u^{(2)} \rho_u du \\ &= \frac{T}{2} \int_0^t \underline{\theta}_u^{(1)[1]} \cdot \underline{\theta}_u^{(2)[1]} f_u^{(1)} f_u^{(2)} \rho_u du + o_P(1), \end{aligned}$$

by the same methodology as above.

Therefore,

$$\begin{aligned} K^n &:= \begin{pmatrix} \langle M_1^{(1)n}, M_1^{(1)n} \rangle & \langle M_1^{(1)n}, M_1^{(2)n} \rangle \\ \langle M_1^{(2)n}, M_1^{(1)n} \rangle & \langle M_1^{(2)n}, M_1^{(2)n} \rangle \end{pmatrix} \\ \xrightarrow{P} K &:= \begin{pmatrix} \frac{T}{2} \int_0^t \|\underline{\theta}_u^{(1)[1]}\|^2 f_u^{(1)2} du & \frac{T}{2} \int_0^t \underline{\theta}_u^{(1)[1]} \cdot \underline{\theta}_u^{(2)[1]} f_u^{(1)} f_u^{(2)} \rho_u du \\ \frac{T}{2} \int_0^t \underline{\theta}_u^{(1)[1]} \cdot \underline{\theta}_u^{(2)[1]} f_u^{(1)} f_u^{(2)} \rho_u du & \frac{T}{2} \int_0^t \|\underline{\theta}_u^{(2)[1]}\|^2 f_u^{(2)2} du \end{pmatrix}. \end{aligned}$$



Note that the martingale representation (A.7) in Theorem A.1(1) implies  $dM_t = \pi'_t \lambda_t^{\frac{1}{2}} dW_t^*$ ; in matrix notation this is,

$$d \begin{pmatrix} M_t^1 \\ M_t^2 \end{pmatrix} = \begin{pmatrix} \pi_t^{1,1} & \pi_t^{2,1} \\ \pi_t^{1,2} & \pi_t^{2,2} \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_t^1} & 0 \\ 0 & \sqrt{\lambda_t^2} \end{pmatrix} d \begin{pmatrix} W_t^{*1} \\ W_t^{*2} \end{pmatrix}.$$

Define a  $2 \times 2$  matrix of the normalizing coefficients,

$$\Lambda_t := \begin{pmatrix} (\pi_t^{1,1})^2 \lambda_t^1 + (\pi_t^{2,1})^2 \lambda_t^2 & 0 \\ 0 & (\pi_t^{1,2})^2 \lambda_t^1 + (\pi_t^{2,2})^2 \lambda_t^2 \end{pmatrix} = \begin{pmatrix} k_t^{1,1} & 0 \\ 0 & k_t^{2,2} \end{pmatrix},$$

due to the identity  $(\pi_t^{1,i})^2 \lambda_t^1 + (\pi_t^{2,i})^2 \lambda_t^2 = k_t^{i,i}$  (recall  $\pi'_t \lambda_t \pi_t = k_t$ ). Then, we may construct the two-dimensional correlated Brownian motion  $\tilde{W}_t := (\tilde{W}_t^1, \tilde{W}_t^2)'$  by

$$d\tilde{W}_t := 1_{\{\lambda_t^1 > 0 \text{ or } \lambda_t^2 > 0\}} \Lambda_t^{-\frac{1}{2}} \pi'_t \lambda_t^{\frac{1}{2}} dW_t^* + 1_{\{\lambda_t^1 = 0, \lambda_t^2 = 0\}} d\tilde{Z}_t,$$

with  $\tilde{Z}_t := (\tilde{Z}_t^1, \tilde{Z}_t^2)'$  another independent, standard Brownian motion. Hence,

$$\begin{aligned} d\langle \tilde{W}^1, \tilde{W}^2 \rangle_t &= 1_{\{\lambda_t^1 > 0 \text{ or } \lambda_t^2 > 0\}} \frac{1}{\sqrt{k_t^{1,1}} \sqrt{k_t^{2,2}}} (\pi_t^{1,1} \pi_t^{1,2} \lambda_t^1 + \pi_t^{2,1} \pi_t^{2,2} \lambda_t^2) dt \\ &= 1_{\{\lambda_t^1 > 0 \text{ or } \lambda_t^2 > 0\}} \frac{1}{\sqrt{k_t^{1,1}} \sqrt{k_t^{2,2}}} k_t^{1,2} dt \end{aligned}$$

as  $\pi_t^{1,1} \pi_t^{1,2} \lambda_t^1 + \pi_t^{2,1} \pi_t^{2,2} \lambda_t^2 = k_t^{1,2}$ .

Since  $k_t^{i,i} = \frac{T}{2} \|\underline{\theta}_t^{(i)[1]}\|^2 f_t^{(i)2}$  (instantaneous variance of  $M^i$ ) and  $k_t^{1,2} = \frac{T}{2} \underline{\theta}_t^{(1)[1]'} \underline{\theta}_t^{(2)[1]'} f_t^{(1)} f_t^{(2)} \rho_t$  (instantaneous covariance of  $M^1$  and  $M^2$ ),

$$d\langle \tilde{W}^1, \tilde{W}^2 \rangle_t = 1_{\{\|\underline{\theta}_t^{(1)[1]}\| \|\underline{\theta}_t^{(2)[1]}\| > 0\}} \frac{\underline{\theta}_t^{(1)[1]'} \cdot \underline{\theta}_t^{(2)[1]'} f_t^{(1)} f_t^{(2)} \rho_t}{\|\underline{\theta}_t^{(1)[1]}\| f_t^{(1)} \|\underline{\theta}_t^{(2)[1]}\| f_t^{(2)}} dt.$$

Finally, in light of  $\tilde{W}_t = (\tilde{W}_t^1, \tilde{W}_t^2)'$ , the limiting martingale is represented as

$$\begin{aligned} M &= \int_0^\cdot \pi'_s \lambda_s^{\frac{1}{2}} dW_s^* = \int_0^\cdot \Lambda_s^{\frac{1}{2}} d\tilde{W}_s = \int_0^\cdot \begin{pmatrix} \sqrt{k_s^{1,1}} & 0 \\ 0 & \sqrt{k_s^{2,2}} \end{pmatrix} d \begin{pmatrix} \tilde{W}_s^1 \\ \tilde{W}_s^2 \end{pmatrix} \\ &= \sqrt{\frac{T}{2}} \int_0^\cdot \begin{pmatrix} \|\underline{\theta}_s^{(1)[1]}\| f_s^{(1)} & 0 \\ 0 & \|\underline{\theta}_s^{(2)[1]}\| f_s^{(2)} \end{pmatrix} d \begin{pmatrix} \tilde{W}_s^1 \\ \tilde{W}_s^2 \end{pmatrix}, \end{aligned}$$

noting that  $\int_0^\cdot 1_{\{\lambda_s^1 = 0, \lambda_s^2 = 0\}} \Lambda_s^{\frac{1}{2}} d\tilde{Z}_s = 0$  and that the martingale  $\int_0^\cdot 1_{\{\lambda_s^1 = 0, \lambda_s^2 = 0\}} \Lambda_s^{\frac{1}{2}} d\tilde{W}_s$ , with the quadratic variation  $\int_0^\cdot 1_{\{\lambda_s^1 = 0, \lambda_s^2 = 0\}} \Lambda_s ds = 0$ , is identically zero.

Apply Theorem A.1(2) with  $A^n := (\sqrt{n} A_1^{(1)n}, \sqrt{n} A_1^{(2)n})'$  and  $A := (0, 0)'$  to obtain the assertion.  $\square$

*A.3.3. Proof of Theorem 4.1.* If we define the (piecewise) random “weight” process  $\pi$  associated with  $S$  by  $\pi_u := \frac{h}{f_{t_{i-1}}^i} \frac{d\langle S, S \rangle_u}{du} = \frac{f_u^2}{\frac{1}{h} \int_{t_{i-1}}^{t_i} f_u^2 du}$  on  $[t_{i-1}, t_i)$ ,  $i = 1, \dots, n - 1$ ,

with  $\pi_u := 1$  on  $[t_{n-1}, t_n]$  (we also let  $\pi_u := 1$  whenever the denominator is zero), it is reasonable to impose that

$$(A.16) \quad \sup_{u \in [0, T]} \|\pi_u\|_p < \infty, \quad 1 \leq p < \infty,$$

in what follows since the positive Itô process  $(\frac{1}{f_t})$  is locally bounded.

Recall that the regression strategy is given by, for each  $n$  and  $1 \leq i \leq n - 1$ ,

$$\tilde{\theta}_i^n := \frac{\langle S, C \rangle_{t_i} - \langle S, C \rangle_{t_{i-1}}}{\langle S, S \rangle_{t_i} - \langle S, S \rangle_{t_{i-1}}} = \frac{\int_{t_{i-1}}^{t_i} \theta_u d\langle S, S \rangle_u}{\int_{t_{i-1}}^{t_i} d\langle S, S \rangle_u},$$

and  $\tilde{\theta}_{t_0}^n := \theta_0$ . The continuously interpolated version of the rescaled relative hedging error  $(L_{2,t_i}^n)$  is decomposed as

$$(A.17) \quad \begin{aligned} \sqrt{n}L_{2,t}^n &:= \sqrt{n} \sum_{i=0}^{n-1} (\theta_{t_i} - \tilde{\theta}_{t_i}^n) \Delta S_{t_i \wedge t} \\ &= \sqrt{n} \sum_{i=0}^{n-1} (\theta_{t_i} - \tilde{\theta}_{t_i}^n) \int_{t_i \wedge t}^{t_{i+1} \wedge t} \mu_u du + \sqrt{n} \sum_{i=0}^{n-1} (\theta_{t_i} - \tilde{\theta}_{t_i}^n) \int_{t_i \wedge t}^{t_{i+1} \wedge t} f_u dW_u^1 \\ &= \int_0^t \psi_u^n \eta_u du + \int_0^t \psi_u^n dW_u^1 =: A_{2,t}^n + M_{2,t}^n, \end{aligned}$$

where  $\Delta S_{t_i \wedge t} := S_{t_{i+1} \wedge t} - S_{t_i \wedge t}$ ,  $\psi_u^n := \sqrt{n} \sum_{i=0}^{n-1} (\theta_{t_i} - \tilde{\theta}_{t_i}^n) f_t 1_{\{t_i < t \leq t_{i+1}\}}$ , and  $\eta_t := \frac{\mu_t}{f_t}$ .

Under condition (A.16), together with the facts that  $\theta^{[1]}$  and  $f$  are Itô ((A.10) and (A.11)) and  $\theta^{[0]}$  is (locally) bounded, it is not difficult to show that, for all large  $n$ ,

$$(A.18) \quad \max_{0 \leq i \leq n-1} \|\theta_{t_i} - \tilde{\theta}_{t_i}^n\|_p \leq Cn^{-\frac{1}{2}}$$

and

$$(A.19) \quad \max_{0 \leq i \leq n-1} \|(\theta_{t_i} - \tilde{\theta}_{t_i}^n) - \zeta_{t_i}\|_p \leq Cn^{-1}, \quad 1 \leq p < \infty,$$

where  $\zeta_{t_i} := \frac{\theta_{t_{i-1}}^{[1]}}{T} \frac{n}{T} \int_{t_{i-1}}^{t_i} (u - t_{i-1}) dW_u$ ,  $i = 1, 2, \dots, n - 1$ ,  $\zeta_{t_0} := 0$ ,  $C$  is a constant that does not depend on  $n$ .

- (I) (i) For the martingale part  $M_2^n$  in (A.17), consider its quadratic variation. After decomposing it in terms of  $f^2$ , together with (A.18) and (A.19), we can show that

$$\begin{aligned} \langle M_2^n, M_2^n \rangle_t &= n \sum_i (\theta_{t_i} - \tilde{\theta}_{t_i}^n)^2 f_{t_i}^2 \Delta(t_i \wedge t) + o_P(1) \\ &= n \sum_i \zeta_{t_i}^2 f_{t_i}^2 \Delta(t_i \wedge t) + o_P(1). \end{aligned}$$

Then, by invoking Lemma A.2 and the path continuity of  $\theta^{[1]}$  and  $f$ ,

$$\begin{aligned} \langle M_2^n, M_2^n \rangle_t &= \frac{T}{3} \sum_{i=0}^{n-1} (\|\underline{\theta}_{t_{i-1}}^{[1]}\| f_{t_i})^2 \Delta(t_i \wedge t) + o_P(1) \\ &= \frac{T}{3} \int_0^t (\|\underline{\theta}_u^{[1]}\| f_u)^2 du + o_P(1), \end{aligned}$$

as  $n \rightarrow \infty$ . (Note that  $\theta$  is driven by  $W := (W^1, W^2)$ . Also,  $\Delta(t_i \wedge t) \equiv t_{i+1} \wedge t - t_i \wedge t$  and  $\theta_{t_{i-1}}^{[1]} \equiv 0$ .)

(ii) Consider next the covariations. Due to the approximations,  $(\theta_u - \theta_{t_i}) \simeq \int_{t_i}^u \underline{\theta}_s^{[1]} dW_s$  and  $f_u^2 \simeq f_{t_i}^2$  for  $u \in [t_i, t_{i+1}]$ ,

$$\begin{aligned} \langle M_1^n, M_2^n \rangle_t &= n \sum_{i=0}^{n-1} (\theta_{t_i} - \tilde{\theta}_{t_i}^n) \int_{t_i \wedge t}^{t_{i+1} \wedge t} (\theta_u - \theta_{t_i \wedge t}) f_u^2 du \\ &= n \sum_{i=0}^{n-1} (\theta_{t_i} - \tilde{\theta}_{t_i}^n) \int_{t_i \wedge t}^{t_{i+1} \wedge t} \left( \int_{t_i \wedge t}^u \underline{\theta}_s^{[1]} dW_s \right) f_{t_i \wedge t}^2 du + o_P(1). \end{aligned}$$

Then, applying Lemma A.3 to the principal term, one has  $\langle M_1^n, M_2^n \rangle = o_P(1)$  owing to (A.18). In the meantime,

$$\begin{aligned} \langle M_2^n, W^1 \rangle_t &= \sqrt{n} \sum_{i=0}^{n-1} (\theta_{t_i} - \tilde{\theta}_{t_i}^n) f_{t_i \wedge t} \Delta(t_i \wedge t) + o_P(1) \\ &= \sqrt{n} \sum_{i=0}^{n-1} \{ (\theta_{t_i} - \tilde{\theta}_{t_i}^n) - \zeta_{t_i} \} f_{t_i} \Delta(t_i \wedge t) \\ &\quad + \sqrt{n} \sum_{i=0}^{n-1} \zeta_{t_i} f_{t_i} \Delta(t_i \wedge t) + o_P(1). \end{aligned}$$

Now, (A.19) and Lemma A.1 respectively imply that the first and second terms are of  $o_P(1)$  so that  $\langle M_2^n, W^1 \rangle = o_P(1)$ .

Trivially,  $\langle M_2^n, W^2 \rangle = 0$  due to the independence of  $W^1$  and  $W^2$ .

(II) We are going to apply Theorem A.1(1). Since

$$\begin{aligned} K^n &:= \begin{pmatrix} \langle M_1^n, M_1^n \rangle & \langle M_1^n, M_2^n \rangle \\ \langle M_2^n, M_1^n \rangle & \langle M_2^n, M_2^n \rangle \end{pmatrix} \\ &= \begin{pmatrix} \frac{T}{2} \int_0^t \|\underline{\theta}_u^{[1]}\|^2 f_u^2 du + o_P(1) & o_P(1) \\ o_P(1) & \frac{T}{3} \int_0^t \|\underline{\theta}_u^{[1]}\|^2 f_u^2 du + o_P(1) \end{pmatrix}, \end{aligned}$$

its limit has the derivative  $dK_t = k_t = z_t z_t'$ , for all  $t$ , with

$$z_t = \begin{pmatrix} \sqrt{\frac{T}{2}} \|\underline{\theta}_t^{[1]}\| f_t & 0 \\ 0 & \sqrt{\frac{T}{3}} \|\underline{\theta}_t^{[1]}\| f_t \end{pmatrix}.$$

Therefore, the limiting martingale is represented as

$$M = \int_0^\cdot z_s dW_s^* = \int_0^\cdot \begin{pmatrix} \sqrt{\frac{T}{2}} \|\theta_s^{[1]}\| f_s & 0 \\ 0 & \sqrt{\frac{T}{3}} \|\theta_s^{[1]}\| f_s \end{pmatrix} d \begin{pmatrix} W_s^{*1} \\ W_s^{*2} \end{pmatrix},$$

by use of Theorem A.1(1).

- (III) For the finite variation part  $A_2^n$  in (A.17) it can be shown that  $A_2^n = \int_0^\cdot \psi_u^n \eta_u du = o_P(1)$  by the same token as (A.15) in (II) of Section A.3.1 above, since we have already shown that  $\int_0^\cdot \psi_u^n du = \langle M_2^n, W^1 \rangle = o_P(1)$ .
- (IV) Finally, by means of Theorem A.1(2) with  $A^n := (\sqrt{n} A_1^n, \sqrt{n} A_2^n)'$  and  $A := (0, 0)'$ , the assertion of the theorem is obtained.  $\square$

A.3.4. *Proposition A.1 and Its Proof.* Let  $L_{3,t}^{n,m} := \sum_{i=0}^{n-1} (\tilde{\theta}_t^n - \hat{\theta}_t^{n,m}) \Delta S_{t_i \wedge t}$ , for every  $n, m \in \mathbb{N}, 0 \leq t \leq T$ .

PROPOSITION A.1. As  $m \rightarrow \infty, L_3^{n,m} = o_P(1)$ , for every  $n \in \mathbb{N}$ .

*Proof.* For simplicity of notation, we denote  $\Delta S_{j,u} := S_u - S_{t_{i-1}^j}$ , for  $t_{i-1}^j \leq u \leq t_i^{j+1}$ , and so forth. For fixed  $m$  and  $n$ , and for each  $i = 1, \dots, n - 1$ ,

$$\hat{\theta}_t^{n,m} = \frac{\sum_j \left( \int_{t_{i-1}^j}^{t_i^{j+1}} dS_u \right) \left( \int_{t_{i-1}^j}^{t_i^{j+1}} \theta_u dS_u \right)}{\sum_j \left( \int_{t_{i-1}^j}^{t_i^{j+1}} dS_u \right)^2} + \frac{\sum_j \left( \int_{t_{i-1}^j}^{t_i^{j+1}} dS_u \right) \left( \int_{t_{i-1}^j}^{t_i^{j+1}} dN^m \right)}{\sum_j \left( \int_{t_{i-1}^j}^{t_i^{j+1}} dS_u \right)^2} =: J_1 + J_2.$$

By means of Itô's Lemma followed by Taylor's expansion, for fixed  $n$ ,

$$\begin{aligned} J_1 &= \frac{\sum_j \int_{t_{i-1}^j}^{t_i^{j+1}} \Delta(\theta \cdot S)_{j,u} dS_u + \sum_j \int_{t_{i-1}^j}^{t_i^{j+1}} \Delta S_{j,u} \theta_u dS_u + \sum_j \int_{t_{i-1}^j}^{t_i^{j+1}} \theta_u d\langle S, S \rangle_u}{\sum_j \int_{t_{i-1}^j}^{t_i^{j+1}} d\langle S, S \rangle_u + 2 \sum_j \int_{t_{i-1}^j}^{t_i^{j+1}} \Delta S_{j,u} dS_u} \\ &= \frac{\int_{t_{i-1}^i}^{t_i^i} \theta_u d\langle S, S \rangle_u + o_P(1)}{\int_{t_{i-1}^i}^{t_i^i} d\langle S, S \rangle_u + o_P(1)} = \frac{\int_{t_{i-1}^i}^{t_i^i} \theta_u d\langle S, S \rangle_u}{\int_{t_{i-1}^i}^{t_i^i} d\langle S, S \rangle_u} + o_P(1) \end{aligned}$$

as well as

$$\begin{aligned} J_2 &= \frac{\sum_j \int_{t_{i-1}^j}^{t_i^{j+1}} \Delta N_{j,u}^m dS_u + \sum_j \int_{t_{i-1}^j}^{t_i^{j+1}} \Delta S_{j,u} dN_u^m + \sum_j \int_{t_{i-1}^j}^{t_i^{j+1}} d\langle N^m, S \rangle_u}{\sum_j \int_{t_{i-1}^j}^{t_i^{j+1}} d\langle S, S \rangle_u + 2 \sum_j \int_{t_{i-1}^j}^{t_i^{j+1}} \Delta S_{j,u} dS_u} \\ &= \frac{o_P(1)}{\int_{t_{i-1}^i}^{t_i^i} d\langle S, S \rangle_u + o_P(1)} = o_P(1), \end{aligned}$$

as  $m \rightarrow \infty$ , for  $i = 1, \dots, n - 1$ , noting that  $N^m$  is a continuous semimartingale with  $\langle N^m, S \rangle = 0$ . Hence, for  $i = 1, \dots, n$ ,

$$\tilde{\theta}_t^n - \hat{\theta}_t^{n,m} = o_P(1),$$

as  $m \rightarrow \infty$ , for every  $n$ ; consequently,  $L_3^{n,m} \equiv \sum_{i=0}^{n-1} (\tilde{\theta}_t^n - \hat{\theta}_t^{n,m}) \Delta S_{t_i \wedge t} = o_P(1)$  as  $m \rightarrow \infty$ , for every  $n$ .  $\square$

A.3.5. *Proof of Theorem 5.1.* Decompose the continuously interpolated hedging error,

$$L_t^{n,m} := C_t^m - C_0^m - \sum_{i=0}^{n-1} \hat{\theta}_t^{n,m} \Delta S_{t_i \wedge t} = L_{0,t}^m + L_{1,t}^n + L_{2,t}^n + L_{3,t}^{n,m},$$

where  $L_{0,t}^m := C_t^m - C_0^m - \int_0^t \theta_u dS_u$ , recalling that

$$L_{1,t}^n = \int_0^t \theta_u dS_u - \sum_{i=0}^{n-1} \theta_{t_i} \Delta S_{t_i \wedge t}, \quad L_{2,t}^n = \sum_{i=0}^{n-1} (\theta_{t_i} - \tilde{\theta}_{t_i}^n) \Delta S_{t_i \wedge t},$$

$$L_{3,t}^{n,m} = \sum_{i=0}^{n-1} (\tilde{\theta}_{t_i}^n - \hat{\theta}_{t_i}^{n,m}) \Delta S_{t_i \wedge t},$$

and  $\tilde{\theta}_{t_i}^n = \frac{\int_{t_{i-1}}^{t_i} \theta_u d\langle S, S \rangle_u}{\int_{t_{i-1}}^{t_i} d\langle S, S \rangle_u}$ ,  $i = 1, \dots, n - 1$ , with  $\tilde{\theta}_0^n = \theta_0$ .

$L_1^n$  and  $L_2^n$  are taken care of by Theorems 3.1 and 4.1, respectively.  $L_3^{n,m}$  is treated by Proposition A.1.  $L_0^m$  is also trivial. Indeed, by definition,

$$L_{0,t}^m = C_t^m - C_0^m - (C_t^\infty - C_0^\infty) = N_t^m - N_0^m,$$

which is of  $o_p(1)$  uniformly in  $t$  as  $m \rightarrow \infty$  by assumption.

Let  $L_i$ ,  $i = 0, \dots, 3$ , be the limit points of the rescaled hedging errors,  $\sqrt{n}L_0^m$ ,  $\sqrt{n}L_1^n$ ,  $\sqrt{n}L_2^n$ , and  $\sqrt{n}L_3^{n,m}$ . To obtain the joint convergence of the four components, we can apply Theorem A.1(2) by putting

$$M^{n,m} := (0, M_1^n, M_2^n, 0)^\prime, \quad A^{n,m} := (\sqrt{n}L_0^m, A_1^n, A_2^n, \sqrt{n}L_3^{n,m})^\prime, \quad A := (0, 0, 0, 0)^\prime,$$

where  $M_k^n$  and  $A_k^n$ ,  $k = 1, 2$ , are as defined as previously, and by letting  $m \rightarrow \infty$  then  $n \rightarrow \infty$ . (Note that in this case both *first* and *last* eigenvalues of the  $4 \times 4$  derivatives matrix  $k$  of  $K$  are 0, which is harmless in the martingale representation (A.7) of  $M$ .) Then the continuous mapping theorem implies that  $\sqrt{n}L^{n,m}$  converges in law to  $L_1 + L_2$ .

Finally, construct a new Brownian motion from the two independent Brownian motions  $W^{*1}$  and  $W^{*2}$  (which appeared in Theorem 4.1) by

$$B := \rho W^{*1} + \sqrt{1 - \rho^2} W^{*2},$$

where  $\rho := \frac{\sqrt{\frac{T}{3}}}{\sqrt{(\sqrt{\frac{T}{3}})^2 + (\sqrt{\frac{T}{3}})^2}} = \sqrt{\frac{3}{5}}$ . Note that this remains independent of  $W^1$  and  $W^2$ .

Therefore, we may rewrite as

$$L_1 + L_2 = \sqrt{\frac{5}{6}} T \int_0^\cdot \|\theta_u^{[1]}\| f_u dB_u.$$

The assertion of Theorem 5.1 follows. □

REMARK A.2. The decomposition above divides the whole hedging error ( $L^{n,m}$ ) into four components: systematic error ( $L_0^m$ ), discrete-hedging error ( $L_1^n$ ), estimation-bias error ( $L_2^n$ ), and sampling error ( $L_3^{n,m}$ ). The last two together form the *relative* error of the strategy (the “strategy-specific” part).

A.3.6. *Proof of Proposition 6.1.* Recall that  $h(C_t^m) \equiv (\varphi(S_t, \cdot))^{-1}(C_t^m)$ , where  $\varphi(S, \Xi)$  is the Black-Scholes price of a European option (not necessarily a call) at stock price  $S$  and cumulative variance  $\Xi$ .  $\theta(\cdot) = \dot{\varphi}_S(S, \cdot)$  is the Black-Scholes delta, a function of  $\Xi$ .  $\Xi_t := \int_t^T \sigma_u^2 du$  is the cumulative variance for the remaining period.

Assume the European option price  $\varphi$  has continuous  $\ddot{\varphi}_{S\Xi\Xi}$  and  $\dot{\varphi}_\Xi > 0$  in  $(0, T)$ , as is the case for European calls. (Then we have continuous  $\ddot{h}_{CC} (= -\ddot{\varphi}_{\Xi\Xi}(\dot{\varphi}_\Xi)^{-3})$  and  $\dot{\theta}_{\Xi\Xi} (= \ddot{\varphi}_{S\Xi\Xi})$ .)

We set  $\psi_t := \frac{\ddot{\varphi}_{S\Xi}(S_t, \Xi_t)}{\dot{\varphi}_\Xi(S_t, \Xi_t)}$ , with an additional assumption that the limiting process  $N$  satisfies  $\int_0^T |\psi_t N_t| \mu_t S_t dt + \int_0^T (\psi_t N_t)^2 \sigma_t^2 S_t^2 dt < \infty$ , a.s. Let  $L_{2,t} := \int_0^t \psi_u N_u dS_u$ .

We are going to show that

$$\sup_{t \in [0, T]} \left| \frac{1}{\gamma^m} L_{2,t}^{n,m} - L_{2,t} \right| \leq \sup_{t \in [0, T]} \left| \frac{1}{\gamma^m} L_{2,t}^{n,m} - L_{2,t}^n \right| + \sup_{t \in [0, T]} |L_{2,t}^n - L_{2,t}|$$

goes to zero in probability as  $m \rightarrow \infty$  then  $n \rightarrow \infty$ , where  $L_{2,t}^n := \sum_{i=0}^{n-1} \psi_{t_i} N_{t_i} \Delta S_{t_i \wedge t}$ .

Since the second term vanishes as  $n \rightarrow \infty$  by the continuity of  $\psi$  and  $N$  (e.g. J-S, Thm. I.4.44, p. 51), it suffices to show that the first term goes to zero as  $m \rightarrow \infty$  (for every fixed  $n$ ). To this end, apply the Taylor theorem to  $\hat{\Xi}_t^m$  around  $\Xi_t$ ,

$$\begin{aligned} \text{(A.20)} \quad \hat{\Xi}_t^m - \Xi_t &= \dot{h}_C(C_t) \cdot (C_t^m - C_t) + \frac{1}{2} \ddot{h}_{CC}(\tilde{C}_t) \cdot (C_t^m - C_t)^2 \\ &= \dot{h}_C(C_t) \cdot N_t^m + \frac{1}{2} \ddot{h}_{CC}(\tilde{C}_t) \cdot (N_t^m)^2, \end{aligned}$$

for some  $\tilde{C}_t \in [C_t^m \wedge C_t, C_t^m \vee C_t]$ . In the meantime, by another usage of the Taylor theorem applied to  $\theta(\hat{\Xi}_t^m)$ ,

$$\text{(A.21)} \quad \theta(\Xi_t) - \theta(\hat{\Xi}_t^m) = \dot{\theta}_\Xi(\Xi_t) \cdot (\Xi_t - \hat{\Xi}_t^m) + \frac{1}{2} \ddot{\theta}_{\Xi\Xi}(\tilde{\Xi}_t) \cdot (\Xi_t - \hat{\Xi}_t^m)^2,$$

for some  $\tilde{\Xi}_t \in [\Xi_t \wedge \hat{\Xi}_t^m, \Xi_t \vee \hat{\Xi}_t^m]$ . Hence, by substituting (A.20) into (A.21),

$$\begin{aligned} \text{(A.22)} \quad \frac{1}{\gamma^m} L_{2,t}^{n,m} &= \frac{1}{\gamma^m} \sum_{i=0}^{n-1} (\theta(\Xi_{t_i}) - \theta(\hat{\Xi}_{t_i}^m)) \Delta S_{t_i \wedge t} \\ &= - \sum_{i=0}^{n-1} \dot{\theta}_\Xi(\Xi_{t_i}) \dot{h}_C(C_{t_i}) \frac{N_{t_i}^m}{\gamma^m} \Delta S_{t_i \wedge t} + R, \end{aligned}$$

where  $R = \gamma^m(-\frac{1}{2}A + \frac{1}{2}B) + \frac{1}{2}\gamma^{m^2}C + \frac{1}{8}\gamma^{m^3}D$ , with

$$\begin{aligned} A &= \sum_{i=0}^{n-1} \dot{\theta}_\Xi(\Xi_{t_i}) \ddot{h}_{CC}(\tilde{C}_{t_i}) \left( \frac{N_{t_i}^m}{\gamma^m} \right)^2 \Delta S_{t_i \wedge t}, \\ B &= \sum_{i=0}^{n-1} \ddot{\theta}_{\Xi\Xi}(\tilde{\Xi}_{t_i}) \dot{h}_C^2(C_{t_i}) \left( \frac{N_{t_i}^m}{\gamma^m} \right)^2 \Delta S_{t_i \wedge t}, \\ C &= \sum_{i=0}^{n-1} \ddot{\theta}_{\Xi\Xi}(\tilde{\Xi}_{t_i}) \dot{h}_C(C_{t_i}) \ddot{h}_{CC}(\tilde{C}_{t_i}) \left( \frac{N_{t_i}^m}{\gamma^m} \right)^3 \Delta S_{t_i \wedge t}, \\ D &= \sum_{i=0}^{n-1} \ddot{\theta}_{\Xi\Xi}(\tilde{\Xi}_{t_i}) \ddot{h}_{CC}^2(\tilde{C}_{t_i}) \left( \frac{N_{t_i}^m}{\gamma^m} \right)^4 \Delta S_{t_i \wedge t}. \end{aligned}$$

(Recall that  $\gamma^m \rightarrow 0$  as  $m \rightarrow \infty$ .) Due to smoothness conditions on  $h$  and  $\theta$ ,  $R = O_P(\gamma^m)$  (uniformly in  $t$ ) as  $m \rightarrow \infty$  is obtained.

Since  $\dot{\theta}_\Xi \dot{h}_C = \dot{\varphi}_{S\Xi}(\dot{\varphi}_\Xi)^{-1}$ , the principal term on the right-hand side of (A.22) goes in probability to

$$L_2^n \equiv \sum_{i=0}^{n-1} \psi_{t_i} N_{t_i} \Delta S_{t_i \wedge t},$$

as  $m \rightarrow \infty$  (for fixed  $n$ ), as expected. □

REMARK A.3. It is possible (and may be more convenient) to formulate the problem in terms of volatility “misspecification.” Suppose there is a volatility “error” process  $\delta_t^m$  such that, for each  $m$ ,

$$0 < \Xi_t + \delta_t^m < \infty$$

and

$$\varphi(S_t, \Xi_t + \delta_t^m) = C_t^m,$$

and suppose that  $\delta_t^m \rightarrow 0$  as  $m \rightarrow \infty$ , for  $0 \leq t < T$ . Since  $\varphi(S_t, \Xi_t) = C_t^\infty$ , by the Taylor theorem around  $\Xi_t$ ,

$$\varphi(S_t, \Xi_t + \delta_t^m) = \dot{\varphi}_\Xi(S_t, \Xi_t)\delta_t^m + \frac{1}{2}\ddot{\varphi}_{\Xi\Xi}(S_t, \tilde{\Xi}_t)(\delta_t^m)^2$$

for some  $\tilde{\Xi}_t$ . Therefore, the price “noise” process  $N_t^m$  is expressible as

$$\begin{aligned} N_t^m &= \dot{\varphi}_\Xi(S_t, \Xi_t)\delta_t^m + \frac{1}{2}\ddot{\varphi}_{\Xi\Xi}(S_t, \tilde{\Xi}_t)(\delta_t^m)^2 \\ &\simeq \dot{\varphi}_\Xi(S_t, \Xi_t)\delta_t^m, \end{aligned}$$

whenever  $\delta_t^m$  is sufficiently small.

In other words, the volatility misspecification is approximated linearly in  $N_t^m$  as

$$\delta_t^m \simeq \frac{N_t^m}{\dot{\varphi}_\Xi(S_t, \Xi_t)},$$

for sufficiently large  $m$ . For example, in the case of European calls,  $\dot{\varphi}_\Xi(S_t, \Xi_t) = \frac{S_t\phi(d_t)}{2\sqrt{\Xi_t}}$  with  $d_t := \frac{\ln(S_t/K) + \Xi_t/2}{\sqrt{\Xi_t}}$ , indicating that  $\dot{\varphi}_\Xi(S_t, \Xi_t)$  goes to zero exponentially fast as  $t$  goes to  $T$  almost surely. So, in general, in order for the implied volatility to exist (for all  $t \in [0, T)$ ), that is, in order for such  $\delta_t^m$  to exist,  $N_t^m$  should go to zero extremely fast as  $t \rightarrow T$ , for all large  $m$ .<sup>12</sup>

#### A.4. Approximation Lemmas

The following Lemmas A.1–A.4 have been used for identifying limits in the asymptotic calculations above. Let  $W := (W^1, \dots, W^d)$  be a  $d$ -dimensional standard Brownian motion. The  $k$ th element of a  $d$ -dimensional process  $\underline{b}$  may be written as  $b^k$ , and so forth.

Let  $n$  be an arbitrary positive integer. Let  $0 = t_0 < t_1 < \dots < t_n = T$  be a partition over  $[0, T]$  with the maximum interval length  $h := \max_{0 \leq i \leq n-1} (t_{i+1} - t_i)$ . The constants  $C$  below do not depend on  $n$  (or  $h$ ).

Denote  $\Delta(t_i \wedge t) := t_{i+1} \wedge t - t_i \wedge t$ . For notational convenience, the value of a process with a negative index shall be read as nil; for example,  $\alpha_{t_{-1}} := 0$  and  $\xi_{-1,0} := 0$ . Transpose is expressed as  $(\cdot)'$ .

The proof is left to the readers as an exercise; one needs to invoke standard tools repeatedly such as Burkholder’s inequality, Hölder’s inequality, Jensen’s inequality, and

<sup>12</sup> We are grateful to Professor Shigeo Kusuoka for comments on this point.

Minkowski’s inequality, while making full use of the fact that processes playing a major role in each approximation are Itô processes.

LEMMA A.1. Assume  $(f_i)_{0 \leq i \leq T}$  is an Itô process,

$$f_t = f_0 + \int_0^t \underline{b}_u dW_u + \int_0^t c_u du,$$

where  $(\underline{b}_i)_{0 \leq i \leq T}$  and  $(c_i)_{0 \leq i \leq T}$  are, respectively,  $(1 \times d)$ - and scalar-valued progressively measurable processes such that  $\sup_{t \in [0, T]} \max_{1 \leq k \leq d} \|b_t^k\|_8 < \infty$  and  $\sup_{t \in [0, T]} \|c_t\|_8 < \infty$ , and  $f_0$  is an  $\mathcal{F}_0$ -measurable variable. For  $s \leq t$ , let  $\xi_{s,t} := \int_s^t \int_s^u \underline{a}_v dW_v du$ .

$$\xi_{i,i+1} := \int_{t_i}^{t_{i+1}} \underline{a}_u dW_u,$$

where  $(\underline{a}_i)_{0 \leq i \leq T}$  is  $(1 \times d)$ -dimensional, progressively measurable and  $\sup_{t \in [0, T]} \max_{1 \leq k \leq d} \|a_t^k\|_4 < \infty$ . Let  $\alpha_i, i = 0, \dots, n-1$ , be respectively  $\mathcal{F}_i$ -measurable random variables such that  $\max_{0 \leq i \leq n-1} \|\alpha_i\|_8 < \infty$ .

Then, the piecewise-linear continuous process  $\Psi_t := \sum_{i=0}^{n-1} \alpha_{t_i-1} \xi_{i-1,i} f_{t_i} \Delta(t_i \wedge t), 0 \leq t \leq T$ , has a (uniform) moment bound as

$$\left\| \sup_{t \in [0, T]} |\Psi_t| \right\|_1 \leq C \max_{0 \leq i \leq n-1} \|\alpha_i\|_8 \sup_{t \in [0, T]} \max_{1 \leq k \leq d} \|a_t^k\|_4 h$$

for some  $C > 0$ , whenever  $h > 0$  is sufficiently small.

LEMMA A.2. We maintain the same hypotheses as in Lemma A.1 except that  $\sup_{t \in [0, T]} \max_{1 \leq k \leq d} \|a_t^k\|_8 < \infty$ . Then, the piecewise-linear continuous process  $\Phi_t := \sum_{i=0}^{n-1} \alpha_{t_i-1} (\xi_{i-1,i})^2 f_{t_i} \Delta(t_i \wedge t), 0 \leq t \leq T$ , is approximated by another process  $\tilde{\Phi}_t := \sum_{i=0}^{n-1} \alpha_{t_i-1} \gamma_{i-1,i} f_{t_i} \Delta(t_i \wedge t), 0 \leq t \leq T$ , where  $\gamma_{i-1,i} := \int_{t_{i-1}}^{t_i} \|\underline{a}_u\|^2 du$ , in such a way that

$$\left\| \sup_{t \in [0, T]} |\Phi_t - \tilde{\Phi}_t| \right\|_1 \leq C \max_{0 \leq i \leq n-1} \|\alpha_i\|_8 \sup_{t \in [0, T]} \max_{1 \leq k \leq d} \|a_t^k\|_8^2 h^{\frac{3}{2}}$$

for some  $C > 0$ , whenever  $h > 0$  is sufficiently small.

LEMMA A.3. Assume  $(\underline{a}_i)_{0 \leq i \leq T}$  to be a progressively measurable process with  $\sup_{t \in [0, T]} \max_{1 \leq k \leq d} \|a_t^k\|_4 < \infty$ . For  $s \leq t$ , let

$$\xi_{s,t} := \int_s^t \int_s^u \underline{a}_v dW_v du.$$

Let  $\alpha_i, i = 0, \dots, n-1$ , be, respectively,  $\mathcal{F}_i$ -measurable random variables such that  $\max_{0 \leq i \leq n-1} \|\alpha_i\|_4 < \infty$ .

Then the continuous process defined by

$$\Phi_t := \sum_{i=0}^{n-1} \alpha_i \xi_{t_i \wedge t, t_{i+1} \wedge t}, \quad 0 \leq t \leq T,$$

has a (uniform) moment bound as

$$\left\| \sup_{t \in [0, T]} |\Phi_t| \right\|_1 \leq C \max_{0 \leq i \leq n-1} \|\alpha_i\|_4 \sup_{t \in [0, T]} \max_{1 \leq k \leq d} \|a_t^k\|_4 h$$

for some  $C > 0$ , whenever  $h > 0$  is sufficiently small.



LEMMA A.4. We maintain the same hypotheses as in Lemma A.3. Assume further that  $(\underline{a}_t)_{0 \leq t \leq T}$  is a  $(1 \times d)$ -dimensional Itô process

$$\underline{a}_t = \underline{a}_0 + \left( \int_0^t \underline{b}_u dW_u \right)' + \int_0^t \underline{c}_u du, \quad 0 \leq t \leq T,$$

such that  $(\underline{b}_t)_{0 \leq t \leq T}$  and  $(\underline{c}_t)_{0 \leq t \leq T}$  are respectively  $(d \times d)$ - and  $(1 \times d)$ -valued, progressively measurable with  $\sup_{t \in [0, T]} \max_{1 \leq k, l \leq d} \|b_t^{kl}\|_4 < \infty$  and  $\sup_{t \in [0, T]} \max_{1 \leq k \leq d} \|c_t^k\|_4 < \infty$ . For  $s \leq t$ , let  $\eta_{s,t} := \int_s^t (\int_s^u \underline{a}_v dW_v)^2 du$ .

Then, the continuous process  $\Psi_t := \sum_{i=0}^{n-1} \alpha_i \eta_{t_i \wedge t, t_{i+1} \wedge t}$ ,  $0 \leq t \leq T$ , is approximated by another process  $\bar{\Psi}_t := \sum_{i=0}^{n-1} \alpha_i \|\underline{a}_{t_i}\|^2 \frac{1}{2} \Delta(t_i \wedge t)^2$ ,  $0 \leq t \leq T$ , in such a way that

$$\left\| \sup_{t \in [0, T]} |\Psi_t - \bar{\Psi}_t| \right\|_1 \leq C \max_{0 \leq i \leq n-1} \|\alpha_i\|_4 \sup_{t \in [0, T]} \max_{1 \leq k \leq d} \|a_t^k\|_8^2 h^{\frac{3}{2}}$$

for some constant  $C > 0$ , whenever  $h > 0$  is sufficiently small.

#### REFERENCES

- AHN, H., M. DAYAL, E. GRANNAN, and G. SWINDLE (1998): Option Replication with Transaction Costs: General Diffusion Limits, *Ann. Appl. Prob.* 8(3), 676–707.
- AÏT-SAHALIA, Y., and A. W. LO (1998): Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices, *J. Finance* 53(2), 499–547.
- ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, and J. WU (2004): Realized Beta: Persistence and Predictability, Center for Financial Studies Working Paper 2004/16.
- AVELLANEDA, M. (1998): Minimum-Entropy Calibration of Asset-Pricing Models. *Inter. J. Theor. Appl. Fin.* 1(4), 447–472.
- BARNDORFF-NIELSEN, O. E., and N. SHEPHARD (2004): Econometric Analysis of Realized Covariation: High-frequency Based Covariance, Regression and Correlation in Financial Economics, *Econometrica* 72(3), 885–925.
- BERMIN, H. (2003): Hedging Options: The Malliavin Calculus Approach versus The  $\Delta$ -Hedging Approach, *Math. Finance* 13(1), 73–84.
- BERTSIMAS, D., L. KOGAN, and A. W. LO (2000): When Is Time Continuous? *J. Financial Econ.* 55(2), 173–204.
- BOSSAERTS, P., and P. HILLION (1997): Parametric Analysis of Hedging in Discrete Time, *J. Econometrics* 81, 243–272.
- BOSSAERTS, P., and P. HILLION (2003): Local Parametric Analysis of Derivatives Pricing and Hedging, *J. Financial Markets* 6(4), 573–605.
- BOULEAU, N. (2003): Error Calculus and Path Sensitivity in Financial Models, *Math. Finance* 13(1), 115–134.
- DUFFIE, D. (1996): *Dynamic Asset Pricing Theory*, 2d ed. New Jersey: Princeton University Press.
- EL KAROUI, N., N. JEANBLANC-PICQUE, and S. E. SHREVE (1998): Robustness of the Black and Scholes Formula, *Math. Finance* 8(2), 93–126.
- FÖLLMER, H., and P. LEUKERT (1999): Quantile Hedging, *Finance Stoch.* 3(3), 251–273.
- FÖLLMER, H., and M. SCHWEIZER (1991): Hedging of Contingent Claims under Incomplete Information; in *Applied Stochastic Analysis*, M. Davis and R. Elliot, eds. London: Gordon and Breach, 389–414.

- FREY, R., and W. J. Runggaldier (1999): Risk-Minimization Hedging Strategies under Restricted Information: The Case of Stochastic Volatility Models Observable Only at Discrete Random Times, *Math. Methods Oper. Res.* 50(2), 339–350.
- GEISS, C., and S. GEISS (2004): On Approximation of a Class of Stochastic Integrals and Interpolation, *Stoch. Stoch. Rep.* 76(4), 339–362.
- GEISS, S. (2002): Quantitative Approximation of Certain Stochastic Integrals, *Stoch. Stoch. Rep.* 73(3–4), 241–270.
- GENON-CATALOT, V., and J. JACOD (1993): On the Estimation of the Diffusion Coefficient for Multi-Dimensional Diffusion Processes, *Ann. Inst. H. Poincaré, Prob. Statist.* 29(1), 119–151.
- GOBET, E., and E. TEMAM (2001): Discrete Time Hedging Errors for Options with Irregular Payoffs, *Finance Stoch.* 5(3), 357–367.
- GRANDITS, P., and W. SCHACHINGER (2001): Leland's Approach to Option Pricing: The Evolution of a Discontinuity, *Math. Finance* 11(3), 347–355.
- HOBSON, D. G. (1998): Volatility Misspecification, Option Pricing and Superreplication via Coupling, *Ann. Appl. Prob.* 8(1), 193–205.
- HUTCHINSON, J., A. W. LO, and T. POGGIO (1994): A Nonparametric Approach to Pricing and Hedging Derivative Securities via Learning Networks, *J. Finance* 49(3), 851–889.
- JACOD, J. (1997): On Continuous Conditional Gaussian Martingales and Stable Convergence in Law; in *Séminaire de Probabilités XXXI, Lecture Notes in Mathematics 1655*, pp. 232–246. Berlin: Springer.
- JACOD, J., S. MÉLÉARD, and P. PROTTER (2000): Explicit Form and Robustness of Martingale Representations, *Ann. Prob.* 28(4), 1747–1780.
- JACOD, J., and A. N. SHIRYAEV (1987): *Limit Theorems for Stochastic Processes*. Berlin: Springer-Verlag.
- KABANOV, Y. M., and M. M. SAFARIAN (1997): On Leland's Strategy of Option Pricing with Transactions Costs, *Finance Stoch.* 1(3), 239–250.
- KARATZAS, I., and S. E. SHREVE (1991): *Brownian Motion and Stochastic Calculus*, 2d ed. Berlin: Springer-Verlag.
- KARATZAS, I., and S. E. SHREVE (1998). *Methods of Mathematical Finance*. Berlin: Springer-Verlag.
- LELAND, H. (1985): Option Pricing and Replication with Transactions Costs, *J. Finance* 40(5), 1283–1301.
- LIANG, B. (1997): Options Pricing with Transaction Costs: An Asymptotic Approach. Technical Report 457, Department of Statistics, University of Chicago.
- MAHAYNI, A. (2003): Effectiveness of Hedging Strategies under Model Misspecification and Trading Restrictions, *Inter. J. Theor. Appl. Fin.* 6(5), 521–552.
- MYKLAND, P. A. (2000): Conservative Delta Hedging, *Ann. Appl. Prob.* 10(2), 664–683.
- MYKLAND, P. A. (2003a): Financial Options and Statistical Prediction Intervals, *Ann. Statistics* 31(5), 1413–1438.
- MYKLAND, P. A. (2003b): The Interpolation of Options, *Finance Stoch.* 7(4), 417–432.
- PERGAMENSHCHIKOV, S. (2003): Limit Theorem for Leland's Strategy, *Ann. Appl. Prob.* 13(3), 1099–1118.
- PRIGENT, J.-L. (2003): *Weak Convergence of Financial Markets*. Berlin: Springer-Verlag.
- ROOTZÉN, H. (1980): Limit Distributions for the Error in Approximations of Stochastic Integrals, *Ann. Prob.* 8(2), 241–251.
- TALAY, D., and Z. ZHENG (2003): Quantiles of the Euler Scheme for Diffusion Processes and Financial Applications, *Math. Finance* 13(1), 187–199.

- TEMAM, E. (2001): Couverture Approchée d'Options Exotiques—Pricing des Options Asiatiques. Ph.D. thesis, Université Paris VI.
- TEMAM, E. (2003): Analysis of Error with Malliavin Calculus: Application to Hedging, *Math. Finance* 13(1), 201–214.
- YOSHIDA, N. (1997): Malliavin Calculus and Asymptotic Expansion for Martingales, *Prob. Theory & Rel. Fields* 109(3), 301–342.
- ZHANG, R. (1999): Couverture Approchée des Options Européennes. Ph.D. thesis, Ecole Nationale des Ponts et Chaussées.