

When Is Time Continuous?*

Dimitris Bertsimas, Leonid Kogan, and Andrew W. Lo[†]

This Draft: August 3, 1998

Abstract

Continuous-time stochastic processes have become central to many disciplines, yet the fact that they are approximations to physically realizable phenomena is often overlooked. We quantify one aspect of the approximation errors of continuous-time models by investigating the replication errors that arise from delta hedging derivative securities in discrete time. We characterize the asymptotic distribution of these replication errors and their joint distribution with other assets as the number of discrete time periods increases. We introduce the notion of *temporal granularity* for continuous-time stochastic processes, which allows us to quantify the extent to which discrete-time implementations of continuous-time models can track the payoff of a derivative security. We show that granularity is a function of the contract specifications of the derivative security, and of the degree of market completeness. We derive closed form expressions for the granularity of geometric Brownian motion and of an Ornstein-Uhlenbeck process for call and put options, and perform Monte Carlo simulations that illustrate the practical relevance of granularity.

*This research was partially supported by the MIT Laboratory for Financial Engineering, the National Science Foundation (Grant No. SBR-9709976) and a Presidential Young Investigator Award (Grant No. DDM-9158118) with matching funds from Draper Laboratory. We thank Michael Brandt, George Martin, Matthew Richardson, Bill Schwert, Jiang Wang, an anonymous referee, and seminar participants at Columbia, the Federal Reserve Board, the London School of Economics, the MIT finance lunch group, the MIT LIDS Colloquium, the NBER Asset Pricing Program, NYU, Renaissance Technologies, the Wharton School, the University of Chicago, the University of Massachusetts at Amherst, the University of Southern California, and the 1997 INFORMS Applied Probability Conference for valuable discussions and comments.

[†]MIT Sloan School of Management. Corresponding author: Andrew Lo, 50 Memorial Drive, E52-432, Cambridge, MA 02142-1347, (617) 253-0920 (voice), (617) 258-5727 (fax), a1omit.edu (email).

Contents

1	Introduction	1
2	Defining Temporal Granularity	4
2.1	Delta-Hedging in Complete Markets	6
2.2	Asymptotic Behavior of the Tracking Error and RMSE	10
2.3	Interpretation of Granularity	13
2.4	Granularity in Incomplete Markets	14
3	Applications	17
3.1	Granularity of Geometric Brownian Motion	17
3.2	Granularity of a Mean-Reverting Process	19
4	Monte Carlo Analysis	20
4.1	Accuracy of the Asymptotics	20
4.2	Qualitative Behavior of the RMSE	21
5	Extensions and Generalizations	26
5.1	Sample-Path Properties of Tracking Errors	27
5.2	Joint Distributions of Tracking Errors and Prices	28
5.3	A PDE Characterization of the Tracking Error	30
5.4	Alternative Measures of the Tracking Error	31
6	Conclusions	33
A	Appendix	35
A.1	Proof of Theorem 1	35
A.1.1	Theorem 1(a)	35
A.1.2	Theorem 1(b)	37
A.1.3	Theorem 1(c)	38
A.2	Proof of Theorem 2	39
A.3	Proof of Theorem 3	44
A.4	Proof of Theorem 4	45
A.5	Proof of Theorem 5	46

1 Introduction

Since Wiener’s (1923) pioneering construction of Brownian motion and Itô’s (1951) theory of stochastic integrals, continuous-time stochastic processes have become indispensable to many disciplines ranging from chemistry and physics to engineering to biology to financial economics. In fact, the application of Brownian motion to financial markets pre-dates Wiener’s contribution by almost a quarter century (see Bachelier (1900)), and Merton’s (1973) seminal derivation of the Black and Scholes (1973) option-pricing formula in continuous time—and, more importantly, his notion of delta hedging and dynamic replication—is often cited as the foundation of today’s multi-trillion dollar derivatives industry.

Indeed, the mathematics and statistics of Brownian motion have become so intertwined with so many scientific theories that we often forget the fact that continuous-time processes are only approximations to physically realizable phenomena. In fact, for the more theoretically inclined, Brownian motion may seem more “real” than discrete-time discrete-valued processes. Of course, whether time is continuous or discrete is a theological question best left for philosophers. But a more practical question remains: under what conditions are continuous-time models good approximations to specific physical phenomena, i.e., when does time seem “continuous” and when does it seem “discrete”?

In this paper, we provide a concrete answer to this question in the context of continuous-time derivative-pricing models, e.g., Merton (1973), by characterizing the replication errors that arise from delta hedging derivatives in discrete time.

Delta-hedging strategies play a central role in the theory of derivatives and in our understanding of dynamic notions of spanning and market completeness. In particular, delta-hedging strategies are recipes for replicating the payoff of a complex security by sophisticated dynamic trading of simpler securities. When markets are dynamically complete (see, for example, Harrison and Kreps (1979) and Duffie and Huang (1985)) and continuous trading is feasible, it is possible to replicate certain derivative securities perfectly. However, when markets are not complete or when continuous trading is not feasible, e.g., trading frictions or periodic market closings, perfect replication is not possible and the usual delta-hedging strategies exhibit *tracking* errors. These tracking errors comprise the focus of our attention.

Specifically, we characterize the asymptotic distribution of the tracking errors of delta-

hedging strategies using continuous-record asymptotics, i.e., we implement these strategies in discrete time and let the number of time periods increase while holding the time span fixed. Since the delta-hedging strategies we consider are those implied by continuous-time models like Merton (1973), it is not surprising that tracking errors arise when such strategies are implemented in discrete time, nor is it surprising that these errors disappear in the limit of continuous time. However, by focusing on the continuous-record asymptotics of the tracking error, we can quantify the discrepancy between the discrete-time hedging strategy and its continuous-time limit, answering the question “When is time continuous?” in the context of replicating derivative securities.

We show that the normalized tracking error converges weakly to a particular stochastic integral and that the root-mean-squared tracking error is of order $N^{-1/2}$ where N is the number of discrete time periods over which the delta hedging is performed. This provides a natural definition for *temporal granularity*: it is the coefficient that corresponds to the $\mathcal{O}(N^{-1/2})$ term. We derive a closed-form expression for the temporal granularity of a diffusion process paired with a derivative security, and propose this as a measure of the “continuity” of time. The fact that granularity is defined with respect to a derivative-security/price-process pair underscores the obvious: a need for specificity in quantifying the approximation errors of continuous-time processes. It is impossible to tell how good an approximation a continuous-time process is to a physical process without specifying the nature of the physical process.

We also characterize the behavior of delta-hedging tracking errors in the case of dynamically *incomplete* markets, markets where, even in the continuous-time limit, derivative securities cannot be perfectly replicated. Although standard delta-hedging strategies are not necessarily optimal in incomplete markets, nevertheless this is still an important case because it provides us with an opportunity to distinguish between two very different economic sources of tracking errors: the inability to trade continuously (due to market frictions or institutional restrictions) and the inability to span states of nature (due to an insufficient number of securities). Of course, these two sources of tracking errors are deeply related,¹ but they are distinct and have different implications for temporal granularity. In particular,

¹They can be traded off to some degree, as Merton (1982a) conjectured and Duffie and Huang (1985) proved.

we find that the root-mean-squared-error of delta-hedging strategies in incomplete markets is comprised of two terms: a constant term and a term that is of order $N^{-1/2}$. This provides a natural decomposition of granularity into two components, one that is fixed, and the other that declines with the number of hedging intervals.

In addition to the general usefulness of a measure of temporal granularity for continuous-time stochastic processes, our results have other, more immediate applications. For example, for a broad class of derivative securities and price processes, our measure of granularity provides a simple method for determining the approximate number of hedging intervals N^* needed to achieve a target root-mean-squared-error δ : $N^* = g^2/\delta^2$ where g is the granularity coefficient of the derivative-security/price-process pair. This expression shows that to halve the root-mean-squared-error of a typical delta-hedging strategy, the number of hedging intervals must be increased approximately fourfold.

Moreover, for some special cases, e.g., the Black-Scholes case, the granularity coefficient can be obtained in closed form, and these cases shed considerable light on several aspects of derivatives replication. For example, in the Black-Scholes case, does an increase in volatility make it easier or more difficult to replicate a simple call option? Common intuition suggests that the tracking error increases with volatility, but the closed-form expression for granularity (3.2) shows that the granularity achieves a maximum as a function of σ and that beyond this point, it becomes a decreasing function of σ . The correct intuition is that at lower levels of volatility, tracking error is an increasing function of volatility because an increase in volatility implies more price movements and a greater likelihood of hedging errors in each hedging interval. But at higher levels of volatility, price movements are so extreme that an increase in volatility in this case implies that prices are less likely to fluctuate near the strike price where delta-hedging errors are the largest, hence granularity is a decreasing function of σ . In other words, at sufficiently high levels of volatility, the nonlinear payoff function of a call option “looks” approximately linear and is therefore easier to hedge. Similar insights can be gleaned from other closed-form expressions of granularity (see, for example, Section 3.2).

In Section 2, we provide a complete characterization of the asymptotic behavior of the tracking error for delta hedging an arbitrary derivative security, and formally introduce the notion of granularity for both the complete- and incomplete-markets cases. To illustrate

the practical relevance of granularity, in Section 3 we obtain closed-form expressions for granularity in two specific cases: call options under geometric Brownian motion, and under a mean-reverting process. In Section 4 we check the accuracy of our continuous-record asymptotic approximations by presenting Monte Carlo simulation experiments for the two examples of Section 3 and comparing them to the corresponding analytical expressions. We present other extensions and generalizations in Section 5 such as a characterization of the sample-path properties of tracking errors, the joint distributions of tracking errors and prices, a PDE characterization of the tracking error, and more general loss functions than root-mean-squared tracking error. We conclude in Section 6.

2 Defining Temporal Granularity

The relationship between continuous-time and discrete-time models in economics and finance has been explored in a number of studies. One of the earliest examples is Merton (1969), in which the continuous-time limit of the budget equation of a dynamic portfolio choice problem is carefully derived from discrete-time considerations (see also Merton (1975, 1982b)). Foley's (1975) analysis of "beginning-of-period" versus "end-of-period" models in macroeconomics is similar in spirit, though quite different in substance.

More recent interest in this issue stems primarily from two sources. On the one hand, it is widely recognized that continuous-time models are useful and tractable approximations to more realistic discrete-time models. Therefore, it is important to establish that key economic characteristics of discrete-time models converge properly to the characteristics of their continuous-time counterparts. A review of recent research along these lines can be found in Duffie and Protter (1992).

On the other hand, while discrete-time and discrete-state models such as those based on binomial and multinomial trees, e.g., Cox, Ross, and Rubinstein (1979), He (1990, 1991), and Rubinstein (1994), may not be realistic models of actual markets, nevertheless they are convenient computational devices for analyzing continuous-time models. Willinger and Taqqu (1991) formalize this notion and provide a review of this literature.

For derivative-pricing applications, the distinction between discrete-time and continuous-time models is a more serious one. For all practical purposes, trading takes place at dis-

crete intervals, and a discrete-time implementation of Merton's (1973) continuous-time delta-hedging strategy cannot perfectly replicate an option's payoff. The tracking error that arises from implementing a continuous-time hedging strategy in discrete time has been studied by several authors.

One of the first studies was conducted by Boyle and Emanuel (1980), who consider the statistical properties of "local" tracking errors. At the beginning of a sufficiently small time interval, they form a hedging portfolio comprised of options and stock according to the continuous-time Black-Scholes/Merton delta-hedging formula. The composition of this hedging portfolio is held fixed during this time interval, which gives rise to a tracking error (in continuous time, the composition of this portfolio would be adjusted continuously to keep its dollar value equal to zero). The dollar-value of this portfolio at the end of the interval is then used to quantify the tracking error.

More recently, Toft (1996) shows that a closed-form expression for the variance of the cash flow from a discrete-time delta-hedging strategy can be obtained for a call or put option in the special case of geometric Brownian motion. However, he observes that this expression is likely to span several pages and is therefore quite difficult to analyze.

But perhaps the most relevant literature for our purposes is Leland's (1985) investigation of discrete-time delta-hedging strategies motivated by the presence of transactions costs, an obvious but important motivation (why else would one trade discretely?) that spurred a series of studies on option pricing with transactions costs, e.g., Bensaïd et al. (1992), Boyle and Vorst (1992), Edirisinghe, Naik, and Uppal (1993), Henrotte (1993), Avellaneda and Paras (1994), and Grannan and Swindle (1996). This strand of the literature provides compelling economic motivation for discrete delta-hedging—trading continuously would generate infinite transactions costs. However, the focus of these studies is the pricing and optimal replication of derivative securities, not the statistical behavior of tracking errors. Although we too have investigated these issues (see Bertsimas, Kogan, and Lo [1997]), our focus here is strictly on the approximation errors that arise from applying continuous-time models discretely, without regard to the specific motivation for discrete trading.

Specifically, we investigate the discrete-time implementation of continuous-time delta-hedging strategies and characterize the stochastic properties of the tracking error in considerable generality by appealing to continuous-record asymptotics. We introduce the notion of

temporal granularity which is central to the issue of when time may be considered continuous, i.e., when continuous-time models are good approximations to discrete-time phenomena. In Section 2.1, we describe the complete-markets framework in which our delta-hedging strategy will be implemented and define the tracking error and related quantities. In Section 2.2, we characterize the continuous-record asymptotic behavior of the tracking error and define the notion of temporal granularity. We provide an interpretation of granularity in Section 2.3 and discuss its implications, and in Section 2.4 we extend our definition of granularity to the incomplete-markets case.

2.1 Delta-Hedging in Complete Markets

We begin by specifying the market environment. For simplicity, we assume that there are only two traded securities: a riskless asset (bond) and a risky asset (stock). Time t is normalized to the unit interval so that trading takes place from $t=0$ to $t=1$. In addition, we assume

(A1) *Markets are frictionless, i.e., there are no taxes, transactions costs, shortsales restrictions, or borrowing restrictions.*

(A2) *The riskless borrowing and lending rate is 0.²*

(A3) *The price P_t of the risky asset follows a diffusion process*

$$\frac{dP_t}{P_t} = \mu(t, P_t) dt + \sigma(t, P_t) dW_t \quad , \quad \sigma(t, P_t) \geq \sigma_0 > 0 \quad (2.1)$$

where the coefficients μ and σ satisfy standard regularity conditions that guarantee existence and uniqueness of the strong solution of (2.1) and market completeness (see Duffie (1996)).

We now introduce a European derivative security on the stock that pays $F(P_1)$ dollars at time $t=1$. We will call $F(\cdot)$ the payoff function of the derivative. The equilibrium price of

²This entails little loss of generality since we can always renormalize all prices by the price of a zero-coupon bond with maturity at time 1 (see, for example, Harrison and Kreps [1979]). This assumption does rule out the case of a stochastic interest rate. However, we are able to accommodate this and other forms of market incompleteness in Section 2.4.

the derivative, $H(t, P_t)$, satisfies the following partial differential equation (PDE) (see Cox, Ingersoll, and Ross (1985) and Duffie (1996)):

$$\frac{\partial H(t, x)}{\partial t} + \frac{1}{2}\sigma^2(t, x)x^2\frac{\partial^2 H(t, x)}{\partial x^2} = 0 \quad (2.2)$$

with the boundary condition

$$H(1, x) = F(x) . \quad (2.3)$$

This is a generalization of the standard Black-Scholes model which can be obtained as a special case when the coefficients of the diffusion process (2.1) are constant, i.e., $\mu(t, P_t) = \mu$, $\sigma(t, P_t) = \sigma$, and the payoff function $F(P_1)$ is given by $\text{Max}[P_1 - K, 0]$ or $\text{Min}[P_1, K]$.

The delta-hedging strategy was introduced by Black and Scholes (1973) and Merton (1973) and when implemented continuously on $t \in [0, 1]$, the payoff of the derivative at expiration can be replicated perfectly by a portfolio of stocks and riskless bonds. This strategy consists of forming a portfolio at time $t = 0$ containing only stocks and bonds with an initial investment of $H(0, P_0)$, and rebalancing it continuously in a *self-financing* manner—all long positions are financed by short positions and no money is withdrawn or added to the portfolio—so that at all times $t \in [0, 1]$ the portfolio contains $\partial H(t, P_t)/\partial P_t$ shares of the stock. The value of such a portfolio at time $t = 1$ is exactly equal to the payoff, $F(P_1)$, of the derivative. Therefore, the price, $H(t, P_t)$, of the derivative can also be considered the production cost of replicating the derivative's payoff $F(P_1)$ starting at time t .

Such an interpretation becomes important when continuous-time trading is not feasible. In this case, $H(t, P_t)$ can no longer be viewed as the equilibrium price of the derivative. However, the function $H(t, P_t)$, defined formally as a solution of (2.2)–(2.3), can still be viewed as the production cost $H(0, P_0)$ of an approximate replication of the derivative's payoff, and may be used to define the production process itself (we formally define a discrete-time delta-hedging strategy below).³ Therefore, when we refer to $H(t, P_t)$ as the derivative's

³The term “approximate replication” indicates the fact that when continuous trading is not feasible, the difference between the payoff of the derivative and the end-of-period dollar-value of the replicating portfolio will be nonzero. See Bertsimas, Kogan, and Lo (1997) for a discussion of derivative replication in discrete

“price” below, we shall have in mind this more robust interpretation of production cost and approximate replication strategy.⁴

More formally, we assume:

(A4) *Trading takes place only at N regularly spaced times t_i , $i = 1, \dots, N$, where*

$$t_i \in \left\{ 0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N} \right\}.$$

Under (A4), the difference between the payoff of the derivative and the end-of-period dollar-value of the replicating portfolio—the *tracking error*—will be nonzero.

Following Hutchinson, Lo, and Poggio (1994), let $V_{t_i}^{(N)}$ be the value of the replicating portfolio at time t_i . Since the replicating portfolio consists of shares of the stock and the bond, we can express $V_{t_i}^{(N)}$ as

$$V_{t_i}^{(N)} = V_{S,t_i}^{(N)} + V_{B,t_i}^{(N)} \tag{2.4}$$

where $V_{S,t_i}^{(N)}$ and $V_{B,t_i}^{(N)}$ denote the dollar amount invested in the stock and the bond, respectively, in the replicating portfolio at time t_i . At time $t=0$ the total value of the replicating portfolio is equal to the price (production cost) of the derivative

$$V_0^{(N)} = H(0, P_0) \tag{2.5}$$

and its composition is given by

$$V_{S,0}^{(N)} = \left. \frac{\partial H(t, P_t)}{\partial P_t} \right|_{t=0} P_0, \quad V_{B,0}^{(N)} = V_0^{(N)} - V_{S,0}^{(N)}, \tag{2.6}$$

time and the distinction between production cost and equilibrium price.

⁴Alternatively, we can conduct the following equivalent thought experiment: while some market participants can trade costlessly and continuously in time and thus ensure that the price of the derivative is given by the solution of (2.2)–(2.3), we will focus our attention on other market participants who can trade only a finite number of times.

hence the portfolio contains $\partial H(t, P_t)/\partial P_t|_{t=0}$ shares of stock. The replicating portfolio is rebalanced at time periods t_i so that

$$V_{S,t_i}^{(N)} = \left. \frac{\partial H(t, P_t)}{\partial P_t} \right|_{t=t_i} P_{t_i} \quad , \quad V_{B,t_i}^{(N)} = V_{t_i}^{(N)} - V_{S,t_i}^{(N)} . \quad (2.7)$$

Between time periods t_i and t_{i+1} , the portfolio composition remains unchanged. This gives rise to non-zero tracking errors $\epsilon_{t_i}^{(N)}$:

$$\epsilon_{t_i}^{(N)} \equiv H(t_i, P_{t_i}) - V_{t_i}^{(N)} . \quad (2.8)$$

The value of the replicating portfolio at time $t = 1$ is denoted by $V_1^{(N)}$ and the end-of-period tracking error is denoted by $\epsilon_1^{(N)}$.

The sequence of tracking errors contains a great deal of information about the approximation errors of implementing a continuous-time hedging strategy in discrete time, and in Sections 2.2–2.4 and 5 we provide a complete characterization of the continuous-time limiting distribution of $\epsilon_1^{(N)}$ and $\{\epsilon_{t_i}\}$. However, because tracking errors also contain noise, we also investigate the properties of the root-mean-squared-error (RMSE) of the end-of-period tracking error ϵ_1 (see Hutchinson, Lo, and Poggio (1994) for other alternatives):

$$\text{RMSE}^{(N)} = \sqrt{\mathbf{E}_0[(\epsilon_1^{(N)})^2]} . \quad (2.9)$$

where $\mathbf{E}_0[\cdot]$ denotes the conditional expectation, conditional on information available at time $t=0$. Whenever exact replication of the derivative's payoff is impossible, $\text{RMSE}^{(N)}$ is positive.

Of course, root-mean-squared-error is only one of many possible summary statistics of the tracking error. A more general specification is the expected loss of the tracking error

$$\mathbf{E}_0 [U(\epsilon_1^{(N)})]$$

where $U(\cdot)$ is a general *loss function*, and we consider this case explicitly in Section 5.4.

2.2 Asymptotic Behavior of the Tracking Error and RMSE

We characterize analytically the asymptotic behavior of the tracking error and RMSE by appealing to continuous-record asymptotics, i.e., by letting the number of trading periods N increase without bound while holding the time span fixed. This characterization provides several important insights into the behavior of the tracking error of general European derivative securities that previous studies have only hinted at indirectly (and only for simple put and call options).⁵ A by product of this characterization is a useful definition for the temporal granularity of a continuous-time stochastic process (relative to a specific derivative security).

We begin with the case of smooth payoff functions $F(P_1)$:

Theorem 1 *Let the derivative's payoff function $F(x)$ in (2.3) be six times continuously differentiable and all of its derivatives be bounded, and suppose there exists a positive constant K such that functions $\mu(\tau, x)$ and $\sigma(\tau, x)$ in (2.1) satisfy*

$$\left| \frac{\partial^{\beta+\gamma}}{\partial \tau^\beta \partial x^\gamma} \mu(\tau, x) \right| + \left| \frac{\partial^{\beta+\gamma}}{\partial \tau^\beta \partial x^\gamma} \sigma(\tau, x) \right| + \left| \frac{\partial^\alpha}{\partial x^\alpha} (x\sigma(\tau, x)) \right| \leq K \quad (2.10)$$

where $(\tau, x) \in [0, 1] \times [0, \infty)$, $1 \leq \alpha \leq 6$, $0 \leq \beta \leq 1$, $0 \leq \gamma \leq 3$, and all partial derivatives are continuous. Then under Assumptions (A1)–(A4):

(a) *The RMSE of the discrete-time delta-hedging strategy (2.7) satisfies*

$$\text{RMSE}^{(N)} = \mathbf{O} \left(\frac{1}{\sqrt{N}} \right). \quad (2.11)$$

(b) *The normalized tracking error satisfies:*

$$\sqrt{N} \epsilon_1^{(N)} \Rightarrow G$$

where

$$G \equiv \frac{1}{\sqrt{2}} \int_0^1 \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW_t', \quad (2.12)$$

⁵See, for example, Boyle and Emanuel (1980) Hutchinson, Lo, and Poggio (1994), Leland (1985), and Toft (1996).

W'_t is a Wiener process independent of W_t , and “ \Rightarrow ” denotes convergence in distribution.

(c) The RMSE of the discrete-time delta-hedging strategy (2.7) satisfies

$$\text{RMSE}^{(N)} = \frac{g}{\sqrt{N}} + \mathbf{O}\left(\frac{1}{N}\right) \quad (2.13)$$

where

$$g = \sqrt{\mathbf{E}_0[\mathcal{R}]}, \quad (2.14)$$

$$\mathcal{R} = \frac{1}{2} \int_0^1 \left(\sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} \right)^2 dt. \quad (2.15)$$

Proof: See Appendix A.1.

Theorem 1 shows that the tracking error is asymptotically equal in distribution to G/\sqrt{N} (up to $\mathbf{O}(N^{-1})$ terms), where G is a random variable given by (2.12). The expected value of G is zero by the martingale property of stochastic integrals. Moreover, the independence of the Wiener processes W'_t and W_t implies that the asymptotic distribution of the normalized tracking error is symmetric, i.e., in the limit of frequent trading, positive values of the normalized tracking error are just as likely as negative values of the same magnitude.

This result might seem somewhat counterintuitive at first, especially in light of Boyle and Emanuel’s (1980) finding that in the Black-Scholes framework the distribution of the local tracking error over a short trading interval is significantly skewed. However, Theorem 1(b) describes the asymptotic distribution of the tracking error over the *entire life* of the derivative, not over short intervals. Such an aggregation of local errors leads to a symmetric asymptotic distribution, just as a normalized sum of random variables will have a Gaussian distribution asymptotically under certain conditions, e.g., the conditions for a functional central limit theorem to hold.

Note that Theorem 1 applies to a wide class of diffusion processes (2.1) and to a variety of derivative payoff functions $F(P_1)$. In particular, it holds when the stock price follows a diffusion process with constant coefficients, as in Black and Scholes (1973).⁶ However, the requirement that the payoff function $F(P_1)$ is smooth—six times differentiable with bounded derivatives—is violated by the most common derivatives of all: simple puts and calls. In the

⁶For this case, the formula for the RMSE (2.14)–(2.15) was first derived by Grannan and Swindle (1996).

next theorem, we extend our results to cover this most basic set of payoff functions.

Theorem 2 *Let the payoff function $F(P_1)$ be continuous and piecewise linear, and suppose (2.10) holds. In addition, let*

$$\left| x^2 \frac{\partial^\alpha \sigma(\tau, x)}{\partial x^\alpha} \right| \leq K_2 \quad (2.16)$$

for $(\tau, x) \in [0, 1] \times [0, \infty)$, $2 \leq \alpha \leq 6$, and some positive constant K_2 . Then under Assumptions (A1)–(A4):

(a) *The RMSE of the discrete-time delta-hedging strategy (2.7) satisfies*

$$\text{RMSE}^{(N)} = \frac{g}{\sqrt{N}} + \mathfrak{o}\left(\frac{1}{\sqrt{N}}\right)$$

where g is given by (2.14)–(2.15).⁷

(b) *The normalized tracking error satisfies*

$$\sqrt{N} \epsilon_1^{(N)} \Rightarrow \frac{1}{\sqrt{2}} \int_0^1 \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW_t' \quad (2.17)$$

where W_t' is a Wiener process independent of W_t .⁸

Proof: See Appendix A.2. By imposing an additional smoothness condition (2.16) on the

diffusion coefficient $\sigma(\tau, x)$, Theorem 2 assures us that the conclusions of Theorem 1 also hold for the most common types of derivatives, those with piecewise linear payoff functions.

Theorems 1 and 2 allows us to define the coefficient of *temporal granularity* g for any combination of continuous-time process $\{P_t\}$ and derivative payoff function $F(P_1)$ —it is the constant associated with the leading term of the RMSE’s continuous-record asymptotic

⁷It is easy to show, using Hölder’s inequality, that (A.15) implies that $g < \infty$. Details of the argument are given in Appendix A.2.

⁸The fact that $g < \infty$ implies that the stochastic integral in (2.17) is well defined.

expansion:

$$g \equiv \sqrt{\frac{1}{2}\mathbf{E}_0 \left[\int_0^1 \left(\sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} \right)^2 dt \right]} \quad (2.18)$$

where $H(t, P_t)$ satisfies (2.2) and (2.3).

2.3 Interpretation of Granularity

The interpretation for temporal granularity is clear: it is a measure of the approximation errors that arise from implementing a continuous-time delta-hedging strategy in discrete time. A derivative-pricing model—recall that this is comprised of a payoff function $F(P_1)$ and a continuous-time stochastic process for P_t —with high granularity requires a larger number of trading periods to achieve the same level of tracking error as a derivative-pricing model with low granularity. In the former case, time is “grainier”, calling for more frequent hedging activity than the latter case. More formally, according to Theorems 1 and 2, to a first-order approximation the RMSE of an N -trade delta-hedging strategy is g/\sqrt{N} . Therefore, if we desire the RMSE to be within some small value δ , we require

$$N \approx \frac{g^2}{\delta^2}$$

trades in the unit interval. For a fixed error δ , the number of trades needed to reduce the RMSE to within δ grows quadratically with granularity. If one derivative-pricing model has twice the granularity of another, it would require four times as many delta-hedging transactions to achieve the same RMSE tracking error.

From (2.18) is it clear that granularity depends on the derivative-pricing formula $H(t, P_t)$ and the price dynamics P_t in natural ways. Equation (2.18) formalizes the intuition that derivatives with higher volatility and higher “gamma” risk (large second derivative with respect to stock price) are more difficult to hedge, since these cases imply larger values for the integrand in (2.18). Accordingly, derivatives on less volatile stocks are easier to hedge. Consider a stock price process which is almost deterministic, i.e., $\sigma(t, P_t)$ is very small. This implies a very small value for g , hence derivatives on such a stock can be replicated almost

perfectly, even if continuous trading is not feasible. Alternatively, such derivatives require relatively few rebalancing periods N to maintain small tracking errors.

Also, a derivative with a particularly simple payoff function should be easier to hedge than derivatives on the same stock with more complicated payoffs. For example, consider a derivative with the payoff function $F(P_1) = P_1$. This derivative is identical to the underlying stock, and can always be replicated perfectly by buying a unit of the underlying stock at time $t = 0$ and holding it until expiration. The tracking error for this derivative is always equal to zero, no matter how volatile the underlying stock is. This intuition is made precise by Theorem 1, which describes exactly how the error depends on the properties of the stock price process and the payoff function of the derivative: it is determined by the behavior of the integral \mathcal{R} , which tends to be large when stock prices “spend more time” in regions of the domain that imply high volatility and high convexity or gamma of the derivative.

We will investigate the sensitivity of g to the specification of the stock price process in Sections 3 and 4.

2.4 Granularity in Incomplete Markets

When markets are dynamically complete, as we assume in Sections 2.1–2.3, the only source of tracking errors is the fact that delta hedging occurs at discrete intervals. However, in dynamically incomplete markets, tracking errors are non-zero even in the limit of continuous delta hedging. For example, derivative-pricing models with stochastic volatility (Hull and White (1987), Wiggins (1987)) are not arbitrage-based models since the presence of a second state variable—stochastic volatility—renders perfect replication impossible with only stocks and bonds as hedging instruments, even with frictionless continuous trading. This suggests two sources of granularity: one from discreteness, and one from the inability to hedge certain risks. While one source of granularity may depend on the frequency of delta hedging, the other should not.⁹ We formalize this intuition in this section.

Let $H(t, \mathbf{Z}_t)$ denote the price of a derivative with the payoff function $F(\mathbf{Z}_1)$, which now

⁹Of course, in incomplete markets, delta hedging need not be the preferred method for hedging a derivative security (in particular, Bertsimas, Kogan, and Lo (1997) provide an alternative based on the minimization of mean-squared-error using stochastic dynamic programming). Nevertheless, an analysis of the delta-hedging errors in incomplete markets provides considerable insight into the focus of this paper: the approximation errors inherent in continuous-time models.

depends on the value of a vector of $J + 1$ state variables \mathbf{Z}_t at time $t = 1$.¹⁰ For notational compactness, let the entries in \mathbf{Z}_t be indexed from $j = 0, \dots, J$ and denote by the first state variable Z_t^0 the price of the traded asset (stock). We replace Assumption (A3) with

(A3') *The state variables Z_t^j are governed by a vector diffusion process*

$$dZ_t^j = \mu_j(t, \mathbf{Z}_t) Z_t^j dt + Z_t^j \sum_{k=0}^J \sigma_{jk}(t, \mathbf{Z}_t) dW_t^k, \quad j = 0, \dots, J \quad (2.19)$$

where W_t^k are independent Wiener processes, and the coefficients μ_j and σ_{jk} satisfy standard regularity conditions that guarantee existence and uniqueness of the strong solution of (2.19). State variables Z_t^j , $j = 1, \dots, J$ do not correspond to the prices of any traded assets.

Under Assumption (A3'), the equilibrium price of the derivative is given by the fundamental PDE:

$$\frac{\partial H(t, \mathbf{Z}_t)}{\partial t} + \sum_{j=1}^J \hat{\mu}_j(t, \mathbf{Z}_t) Z_t^j \frac{\partial H(t, \mathbf{Z}_t)}{\partial Z_t^j} + \frac{1}{2} \sum_{j,k=0}^J a_{jk}(t, \mathbf{Z}_t) Z_t^j Z_t^k \frac{\partial^2 H(t, \mathbf{Z}_t)}{\partial Z_t^j \partial Z_t^k} = 0 \quad (2.20)$$

with the boundary condition

$$H(1, \mathbf{Z}_1) = F(\mathbf{Z}_1) \quad (2.21)$$

where $a_{jk} \equiv \sum_{l=0}^J \sigma_{jl} \sigma_{kl}$ and $\hat{\mu}_j(t, \mathbf{Z}_t)$ are risk-adjusted drift rates.¹¹

Since, by assumption, the state variables Z_t^j , $j = 1, \dots, J$ do not correspond to the prices of any traded assets, the market is dynamically incomplete. Nevertheless, a delta-hedging strategy can still be applied by the holder of the short position in the derivative. Specifically, we assume that the number of shares of stock held in the replicating portfolio at time t is given by

$$\frac{\partial H(t, \mathbf{Z}_t)}{\partial Z_t^0} + \sum_{j=1}^J \lambda_j(t, \mathbf{Z}_t) \frac{\partial H(t, \mathbf{Z}_t)}{\partial Z_t^j} = \sum_{j=0}^J \lambda_j(t, \mathbf{Z}_t) \frac{\partial H(t, \mathbf{Z}_t)}{\partial Z_t^j} \quad (2.22)$$

¹⁰Vectors and matrices are denoted by boldface roman characters.

¹¹Note that the existence of risk-adjusted drift rates in incomplete markets is far from straightforward, and requires additional structure and significant restrictions on the underlying model of economic equilibrium. See, for example, Merton (1973), Garman (1976), He and Pearson (1991), Huang (1987), and Hull and White (1987).

where $\lambda_0(t, \mathbf{Z}_t) \equiv 1$ and the functions $\lambda_j(t, \mathbf{Z}_t)$, $j = 1, \dots, J$, determine the extent to which the traded asset is used to hedge the risk associated with other (non-traded) state variables.

We can now characterize the behavior of the tracking error that arises from the discrete-time delta-hedging strategy (2.22). First, define the following notation to be used in the next theorem:

$$\frac{\partial^{(\alpha)}}{\partial \mathbf{x}^{(\alpha)}} \equiv \frac{\partial^{\alpha_0 + \dots + \alpha_J}}{\partial x_0^{\alpha_0} \dots \partial x_J^{\alpha_J}} \quad , \quad |\alpha| \equiv \alpha_0 + \dots + \alpha_J .$$

Then we have:

Theorem 3 *Let the payoff function $F(\mathbf{Z})$ be six times continuously differentiable with all of its derivatives bounded. Suppose there exists a positive constant K such that functions $\mu_j(s, \mathbf{Z})$, $\lambda_j(s, \mathbf{Z})$ and $\sigma_{jk}(s, \mathbf{Z})$, $j, k = 0, \dots, J$ satisfy*

$$\begin{aligned} & \left| \frac{\partial^{\beta+(\gamma)}}{\partial \tau^\beta \partial \mathbf{Z}^{(\gamma)}} \mu_j(\tau, \mathbf{Z}) \right| + \left| \frac{\partial^{\beta+(\gamma)}}{\partial \tau^\beta \partial \mathbf{Z}^{(\gamma)}} \lambda_j(\tau, \mathbf{Z}) \right| + \\ & \left| \frac{\partial^{\beta+(\gamma)}}{\partial \tau^\beta \partial \mathbf{Z}^{(\gamma)}} \sigma_{jk}(\tau, \mathbf{Z}) \right| + \left| \frac{\partial^{(\alpha)}}{\partial \mathbf{Z}^{(\alpha)}} (x_j \sigma_{jk}(\tau, \mathbf{Z})) \right| \leq K \end{aligned} \quad (2.23)$$

for $(\tau, \mathbf{Z}) \in [0, 1] \times [0, \infty)^{J+1}$, $1 \leq |\alpha| \leq 6$, $0 \leq \beta \leq 1$, $0 \leq |\gamma| \leq 3$, and all partial derivatives are continuous. Then under Assumptions (A1), (A2), (A3), and (A4) the RMSE of the tracking error satisfies

$$\text{RMSE}^{(N)} = a + \frac{g}{\sqrt{N}} + \mathbf{O}\left(\frac{1}{N}\right) \quad (2.24)$$

where

$$a = \mathbf{E}_0 \left[(I_1 + I_2)^2 \right]^{\frac{1}{2}} \quad (2.25)$$

$$I_1 = \int_0^1 \sum_{j=1}^J \frac{\partial H(t, \mathbf{Z}_t)}{\partial Z_t^j} \left[\left(\mu_j(t, \mathbf{Z}_t) - \hat{\mu}_j(t, \mathbf{Z}_t) \right) Z_t^j - \lambda_j(t, \mathbf{Z}_t) \mu_0(t, \mathbf{Z}_t) Z_t^0 \right] dt \quad (2.26)$$

$$I_2 = \int_0^1 \sum_{j=1}^J \frac{\partial H(t, \mathbf{Z}_t)}{\partial Z_t^j} \sum_{k=0}^J \left(Z_t^j \sigma_{jk}(t, \mathbf{Z}_t) - \lambda_j(t, \mathbf{Z}_t) Z_t^0 \sigma_{0k}(t, \mathbf{Z}_t) \right) dW_k . \quad (2.27)$$

and g is a positive constant.

Proof: See Appendix A.3.

In addition to the $\mathbf{O}(N^{-1/2})$ term, the expression for the RMSE in Theorem 3 now contains an $\mathbf{O}(1)$ term a . This confirms the intuition we started with: tracking errors arise from both market incompleteness and from discrete hedging, and while the latter declines in the limit of continuous trading, the former does not. Moreover, (2.26) and (2.27) show that the magnitude of a is determined by the nature of market incompleteness. If none of the non-traded state variables enter the derivative-pricing function H , then $I_1 = I_2 = 0$ and we are back to the complete-markets case of Theorems 1 and 2.

Finally, Theorem 3 can also be extended to derivatives with continuous, piecewise linear payoffs as in Theorem 2, but we omit this straightforward extension to conserve space.

3 Applications

To develop further intuition for our measure of temporal granularity, in this section we derive closed-form expressions for g in two important special cases: the Black-Scholes option pricing model with geometric Brownian motion, and the Black-Scholes model with a mean-reverting (Ornstein-Uhlenbeck) process.

3.1 Granularity of Geometric Brownian Motion

Suppose that stock price dynamics are given by:

$$\frac{dP_t}{P_t} = \mu dt + \sigma dW_t . \tag{3.1}$$

where μ and σ are constants. Under this assumption we obtain the following explicit characterization of the granularity g .

Theorem 4 *Under Assumptions (A1)–(A4), stock price dynamics (3.1), and the payoff function of simple call and put options, the granularity g in (2.13) is given by*

$$g = K\sigma \left(\int_0^1 \frac{\exp \left[-\frac{[\mu t + \ln(\frac{P_0}{K}) - \sigma^2/2]^2}{\sigma^2(1+t)} \right]}{4\pi\sqrt{1-t^2}} dt \right)^{\frac{1}{2}} \quad (3.2)$$

where K is the option's strike price.

Proof: See Appendix A.4.

It is easy to see that $g = 0$ if $\sigma = 0$ and g increases with σ in the neighborhood of zero. When σ increases without bound, the granularity g decays to zero, which means that it has at least one local maximum as a function of σ . The granularity g also decays to zero when P_0/K approaches zero or infinity. In the important special case of $\mu = 0$, we conclude by direct computation that g is a unimodal function of P_0/K , that achieves its maximum at $P_0/K = \exp(\sigma^2/2)$.

The fact that granularity is not monotone increasing in σ may seem counterintuitive at first—after all, how can delta-hedging errors become smaller for larger values of σ ? The intuition follows from the fact that at small levels of σ , an increase in σ leads to larger granularity because there is a greater chance that the stock price will fluctuate around regions of high gamma (where $\partial^2 H(t, P_t)/\partial P_t^2$ is large, i.e., near the money), leading to greater tracking errors. However, at very high levels of σ , prices fluctuate so wildly that an increase in σ will decrease the probability that the stock price stays in regions of high gamma for very long—in these extreme cases, the payoff function “looks” approximately linear hence granularity becomes a decreasing function of σ .

Also, we show below that g is not very sensitive to changes in μ when σ is sufficiently large. This implies that, for an empirically relevant range of parameter values, g , as a function of the initial stock price, achieves its maximum close to the strike price, i.e., at $P_0/K \approx 1$. These observations are consistent with the behavior of the tracking error for finite values of N that we see in the Monte Carlo simulations of Section 4.

When stock prices follow a geometric Brownian motion, expressions similar to (3.2) can be obtained for derivatives other than simple puts and calls. For example, for a straddle,

consisting of one put and one call option with the same strike price K , the constant g is twice as large as for the put or call option alone.

3.2 Granularity of a Mean-Reverting Process

Let $p_t \equiv \ln(P_t)$ and suppose

$$dp_t = \left(-\gamma(p_t - (\alpha + \beta t)) + \beta \right) dt + \sigma dW_t \quad (3.3)$$

where $\beta = \mu - \sigma^2/2$ and α is a constant. This is an Ornstein-Uhlenbeck process with a linear time trend, and the solution of (3.3) is given by

$$p_t = (p_0 - \alpha)e^{-\gamma t} + (\alpha + \beta t) + \sigma \int_0^t e^{-\gamma(t-s)} dW_s . \quad (3.4)$$

Under these price dynamics, we have:

Theorem 5 *Under Assumptions (A1)–(A4), stock price dynamics (3.3), and the payoff function of simple call and put options, the granularity g in (2.13) is given by*

$$g = K\sigma \left(\int_0^1 \frac{\sqrt{\gamma} \exp \left[-\frac{\gamma [\alpha + \mu t + (\ln(\frac{P_0}{K}) - \alpha) \exp(-\gamma t) - \sigma^2/2]^2}{\sigma^2 [\gamma(1-t) + 1 - \exp(-2\gamma t)]} \right]}{4\pi \sqrt{1-t} \sqrt{\gamma(1-t) + 1 - \exp(-2\gamma t)}} dt \right)^{\frac{1}{2}} \quad (3.5)$$

where K is the option's strike price.

Proof: See Appendix A.5.

Expression (3.5) is a direct generalization of (3.2): when the mean-reversion parameter γ is set to zero, the process (3.3) becomes a geometric Brownian motion and (3.5) reduces to (3.2). Theorem 5 has some interesting qualitative implications for the behavior of the tracking error in presence of mean-reversion. We will discuss them in detail in the next section.

4 Monte Carlo Analysis

Since our analysis of granularity is based entirely on continuous-record asymptotics, we must check the quality of these approximations by performing Monte Carlo simulation experiments for various values of N . The results of these Monte Carlo simulations are reported in Section 4.1. We also use Monte Carlo simulations to explore the qualitative behavior of the RMSE for various parameter values of the stock price process, and these simulations are reported in Section 4.2.

4.1 Accuracy of the Asymptotics

We begin by investigating the distribution of the tracking error $\epsilon^{(N)}$ for various values of N . We do this by simulating the hedging strategy of Section 2.2 for call and put options,¹² assuming that price dynamics are given by a geometric Brownian motion (3.1).¹³ We set the parameters of the stock price process to $\mu = 0.1$, $\sigma = 0.3$, $P_0 = 1.0$ and let the strike price be $K = 1$. We consider $N = 10, 20, 50, 100$, and simulate the hedging process 250,000 times for each value of N .

Figure 1a shows the empirical probability density function (PDF) of $\epsilon_1^{(N)}$ for each N . As expected, the distribution of the tracking error becomes tighter as the trading frequency increases. It is also apparent that the tracking error can be significant even for $N = 100$. Figure 1b contains the empirical PDFs of the normalized tracking error, $\sqrt{N}\epsilon_1^{(N)}$, for the same values of N . These PDFs are compared to the PDF of the asymptotic distribution (2.17), which is estimated by approximating the integral in (2.17) using a first-order Euler scheme. The functions in Figure 1b are practically identical and indistinguishable, which suggests that the asymptotic expression for the distribution of $\sqrt{N}\epsilon_1^{(N)}$ in Theorem 1(b) is an excellent approximation to the finite-sample PDF for values of N as small as 10.

To evaluate the accuracy of the asymptotic expression g/\sqrt{N} for finite values of N , we compare g/\sqrt{N} to the actual RMSE from Monte Carlo simulations of the delta-hedging

¹²According to Theorem 1, the asymptotic expressions for the tracking error and the RMSE are the same for put and call options since these options have the same second partial derivative of the option price with respect to the current stock price. Moreover, it is easy to verify, using the put-call parity relation, that these options give rise to identical tracking errors.

¹³When the stock price process P_t follows a geometric Brownian motion, the stock price at time t_{i+1} is distributed (conditional on the stock price at time t_i) as $P_{t_i} \exp((\mu - \sigma^2/2)\Delta t + \sigma\sqrt{\Delta t}\eta)$, where $\eta \sim \mathcal{N}(0, 1)$. We use this relation to simulate the delta-hedging strategy.

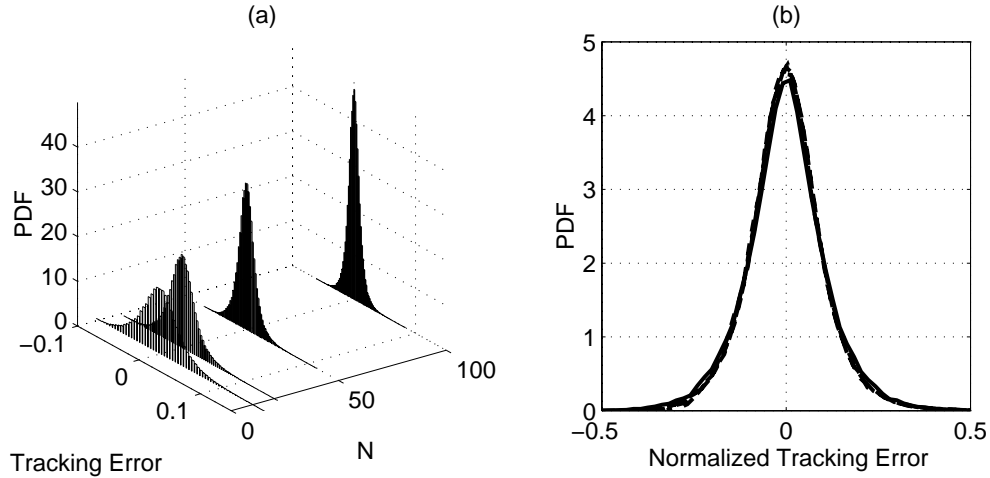


Figure 1: Empirical probability density functions of (a) the tracking error and (b) the normalized tracking error (dashed line), are plotted for different values of the trading frequency N . Figure (b) also shows the empirical probability density function of the asymptotic distribution (2.17) (solid line). The stock price process is given by (3.1) with parameters $\mu = 0.1$, $\sigma = 0.3$, $P_0 = 1.0$. The option is a European call (put) option with strike price $K = 1$.

strategy of Section 2.2. Specifically, we simulate the delta-hedging strategy for a set of European put and call options with strike price $K = 1$ under geometric Brownian motion (3.1) with different sets of parameter values for (σ, μ, P_0) . The tracking error is tabulated as a function of these parameters and the results are summarized in Tables 1, 2, and 3.

Tables 1–3 show that g/\sqrt{N} is an excellent approximation to the RMSE across a wide range of parameter values for (μ, σ, P_0) , even for as few as $N = 10$ delta-hedging periods.

4.2 Qualitative Behavior of the RMSE

The Monte Carlo simulations of Section 4.1 show that RMSE increases with the diffusion coefficient σ in an empirically relevant range of parameter values (see Table 2), and that the RMSE is not very sensitive to the drift rate μ of the stock price process when σ is sufficiently large (see Table 3). These properties are illustrated in Figures 2a and 3. In Figure 2a the logarithm of RMSE is plotted against the logarithm of trading periods N for $\sigma = 0.1, 0.2, 0.3$ —as σ increases, the locus of points shifts upward. Figure 3 shows that granularity g is not a monotone function of σ and goes to zero as σ increases without bound.

Table 1: The sensitivity of the RMSE as a function of the initial price P_0 . The RMSE is estimated using Monte Carlo simulation. Options are European calls and puts with strike price $K = 1$. 250,000 simulations are performed for every set of parameter values. The stock price follows a geometric Brownian motion (3.1). The drift and diffusion coefficients of the stock price process are $\mu = 0.1$ and $\sigma = 0.3$. $\text{RMSE}^{(N)}$ is compared to the asymptotic approximation $gN^{-1/2}$ in (2.13)–(3.2). The relative error (RE) of the asymptotic approximation is defined as $|gN^{-1/2} - \text{RMSE}^{(N)}|/\text{RMSE}^{(N)} \times 100\%$.

Parameters		$gN^{-1/2}$	$\text{RMSE}^{(N)}$	R.E.	Call Option		Put Option	
N	P_0				$H(0, P_0)$	$\frac{\text{RMSE}^{(N)}}{H}$	$H(0, P_0)$	$\frac{\text{RMSE}^{(N)}}{H}$
10	0.50	0.0078	0.0071	8.9%	7E-4	9.64	0.501	0.014
20	0.50	0.0055	0.0052	7.9%	7E-4	6.88	0.501	0.010
50	0.50	0.0035	0.0033	5.9%	7E-4	4.43	0.501	0.007
100	0.50	0.0025	0.0024	3.1%	7E-4	3.22	0.501	0.005
10	0.75	0.0259	0.0248	3.8%	0.023	1.08	0.273	0.091
20	0.75	0.0183	0.0177	2.9%	0.023	0.760	0.273	0.065
50	0.75	0.0116	0.0113	2.6%	0.023	0.490	0.273	0.041
100	0.75	0.0082	0.0082	2.3%	0.023	0.345	0.273	0.029
10	1.00	0.0334	0.0327	4.1%	0.119	0.269	0.119	0.269
20	1.00	0.0236	0.0227	3.4%	0.119	0.192	0.119	0.192
50	1.00	0.0149	0.0145	2.3%	0.119	0.122	0.119	0.122
100	1.00	0.0106	0.0104	1.9%	0.119	0.087	0.119	0.087
10	1.25	0.0275	0.0263	5.8%	0.294	0.088	0.044	0.588
20	1.25	0.0194	0.0187	3.9%	0.294	0.064	0.044	0.423
50	1.25	0.0123	0.0120	2.7%	0.294	0.041	0.044	0.271
100	1.25	0.0087	0.0087	1.8%	0.294	0.029	0.044	0.195
10	1.50	0.0181	0.0169	7.7%	0.515	0.033	0.015	1.130
20	1.50	0.0128	0.0122	5.3%	0.515	0.024	0.015	0.816
50	1.50	0.0081	0.0076	2.9%	0.515	0.015	0.015	0.528
100	1.50	0.0057	0.0056	3.0%	0.515	0.011	0.015	0.373

Table 2: The sensitivity of the RMSE as a function of volatility σ . The RMSE is estimated using Monte Carlo simulation. Options are European calls and puts with strike price $K = 1$. 250,000 simulations are performed for every set of parameter values. The stock price follows a geometric Brownian motion (3.1). The drift coefficient of the stock price process is $\mu = 0.1$, and the initial stock price is $P_0 = 1.0$. $\text{RMSE}^{(N)}$ is compared to the asymptotic approximation $gN^{-1/2}$ in (2.13)–(3.2). The relative error (RE) of the asymptotic approximation is defined as $|gN^{-1/2} - \text{RMSE}^{(N)}|/\text{RMSE}^{(N)} \times 100\%$.

Parameters		$gN^{-1/2}$	Call and Put Options			
N	σ		$\text{RMSE}^{(N)}$	R.E.	$H(0, P_0)$	$\frac{\text{RMSE}^{(N)}}{H}$
10	0.3	0.0334	0.0327	4.1%	0.119	0.269
20	0.3	0.0236	0.0227	3.4%	0.119	0.192
50	0.3	0.0149	0.0145	2.3%	0.119	0.122
100	0.3	0.0106	0.0104	1.9%	0.119	0.087
10	0.2	0.0219	0.0212	3.4%	0.080	0.266
20	0.2	0.0155	0.0151	3.0%	0.080	0.189
50	0.2	0.0098	0.0096	2.1%	0.080	0.121
100	0.2	0.0069	0.0068	1.7%	0.080	0.086
10	0.1	0.0100	0.0102	1.6%	0.040	0.255
20	0.1	0.0071	0.0071	0.04%	0.040	0.177
50	0.1	0.0045	0.0044	1.1%	0.040	0.111
100	0.1	0.0032	0.0031	0.9%	0.040	0.078

Table 3: The sensitivity of the RMSE as a function of the drift μ . The RMSE is estimated using Monte Carlo simulation. Options are European calls and puts with strike price $K = 1$. 250,000 simulations are performed for every set of parameter values. The stock price follows a geometric Brownian motion (3.1). The diffusion coefficient of the stock price process is $\sigma = 0.3$, and the initial stock price is $P_0 = 1.0$, the number of trading periods is $N = 20$. $\text{RMSE}^{(N)}$ is compared to the asymptotic approximation $gN^{-1/2}$ in (2.13)–(3.2). The relative error (RE) of the asymptotic approximation is defined as $|gN^{-1/2} - \text{RMSE}^{(N)}|/\text{RMSE}^{(N)} \times 100\%$.

Parameters		Call and Put Options			
μ	$gN^{-1/2}$	$\text{RMSE}^{(N)}$	R.E.	$H(0, P_0)$	$\frac{\text{RMSE}^{(N)}}{H}$
0.0	0.0235	0.0226	4.3%	0.119	0.189
0.1	0.0236	0.0229	3.4%	0.119	0.192
0.2	0.0230	0.0226	1.7%	0.119	0.190
0.3	0.0218	0.0220	1.0%	0.119	0.184

Figure 2b plots the RMSE as a function of the initial stock price P_0 . RMSE is a unimodal function of P_0/K (recall that the strike price has been normalized to $K = 1$ in all our calculations), achieving its maximum around 1 and decaying to zero as P_0/K approaches zero or infinity (see Table 1). This confirms the common intuition that close-to-the-money options are the most difficult to hedge—they exhibit the largest RMSE.

Finally, the relative importance of the RMSE can be measured by the ratio of the RMSE to the option price: $\text{RMSE}^{(N)}/H(0, P_0)$. This quantity is the root-mean-squared error per dollar invested in the option. Table 1 shows that this ratio is highest for out-of-the-money options, despite the fact that the RMSE is highest for close-to-the-money options. This is due to the fact that the option price decreases faster than the RMSE as the stock moves away from the strike.

Now consider the case of mean-reverting stock price dynamics (3.3). Recall that under these dynamics, the Black-Scholes formula still holds.¹⁴ Nevertheless, the behavior of gran-

¹⁴However, the numerical value for σ may be different than that of a geometric Brownian motion because the presence of mean-reversion can affect conditional volatility, holding unconditional volatility fixed. See Lo and Wang (1995) for further discussion.

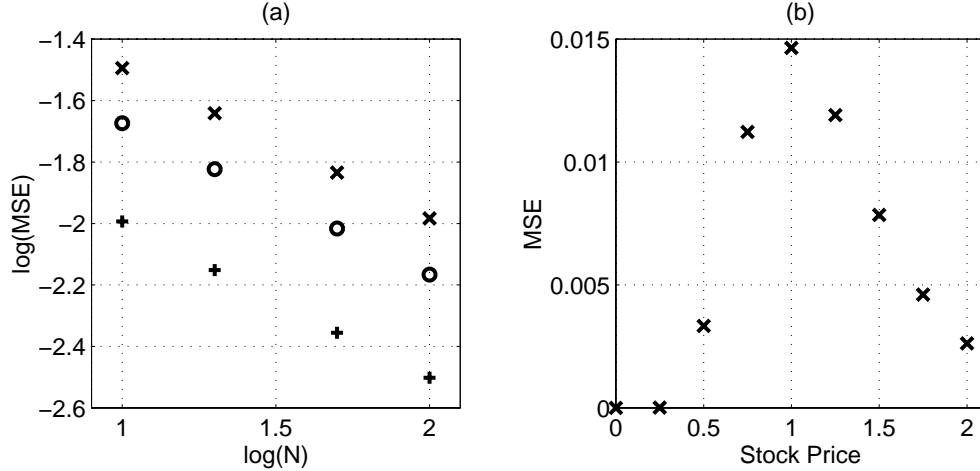


Figure 2: (a) The logarithm of the root-mean-squared error $\log_{10}(\text{RMSE}^{(N)})$ is plotted as a function of the logarithm of the trading frequency $\log_{10}(N)$. The option is a European call (put) option with the strike price $K = 1$. The stock price process is given by (3.1) with parameters $\mu = 0.1$, $P_0 = 1.0$. The diffusion coefficient of the stock price process takes values $\sigma = 0.3$ (x's), $\sigma = 0.2$ (o's) and $\sigma = 0.1$ (+'s). (b) The root-mean-squared error RMSE is plotted as a function of the initial stock price P_0 . The option is a European put option with the strike price $K = 1$. Parameters of the stock price process are $\mu = 0.1$, $\sigma = 0.3$.

ularity and RMSE is quite different in this case. Figure 4 plots the granularity g of call and put options for the Ornstein-Uhlenbeck process (3.3) as a function of α and P_0 . Figure 4a assumes a value of 0.1 for the mean-reversion parameter γ and Figure 4b assumes a value of 3.0. It is clear from these two plots that the degree of mean reversion γ has an enormous impact on granularity. When γ is small, Figure 4a shows that the RMSE is highest when P_0 is close to the strike price and is not sensitive to α . But when γ is large, Figure 4b suggests that the RMSE is highest when $\exp(\alpha)$ is close to the strike price and is not sensitive to P_0 .

The influence of γ on the granularity can be understood by recalling that granularity is closely related to the option's gamma (see Section 2.3). When γ is small, the stock price is more likely to spend time in the neighborhood of the strike price—the region with the highest “gamma” or $\partial^2 H(t, P_t) / \partial P_t^2$ —when P_0 is close to K . However, when γ is large, the stock price is more likely to spend time in a the neighborhood of $\exp(\alpha)$, thus g is highest when $\exp(\alpha)$ is close to K .

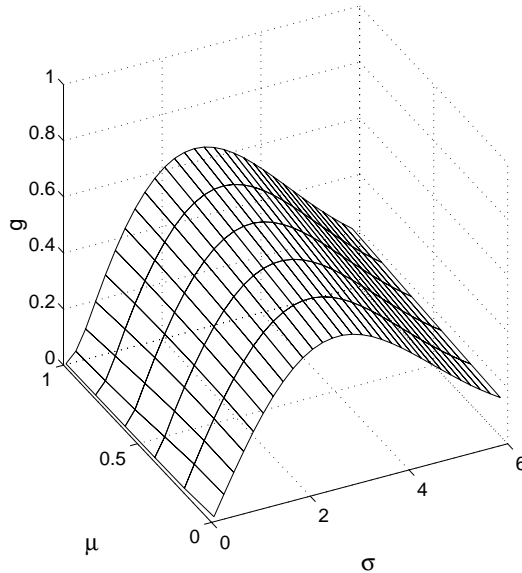


Figure 3: The granularity g is plotted as a function of σ and μ . The option is a European call (put) option with strike price $K = 1$. The stock price process is geometric Brownian motion and initial stock price $P_0 = 1$.

5 Extensions and Generalizations

The analysis of Section 2 can be extended in a number of directions, and we briefly outline four of the most important of these extensions here. In Section 5.1, we show that the normalized tracking error converges in a much stronger sense than simply in distribution, and that this stronger “sample-path” notion of convergence—called, ironically, “weak” convergence—can be used to analyze the tracking error of American-style derivative securities. In Section 5.2 we characterize the asymptotic joint distributions of the normalized tracking error and asset prices, a particularly important extension for investigating the tracking error of delta hedging a portfolio of derivatives. In Section 5.3, we provide another characterization of the tracking error, one that relies on PDE’s, that offers important computational advantages. And in Section 5.4, we consider alternatives to mean-squared-error loss functions and show that for quite general loss functions, the behavior of the expected loss of the tracking error is characterized by the same stochastic integral (2.17) as in the mean-squared-error case.

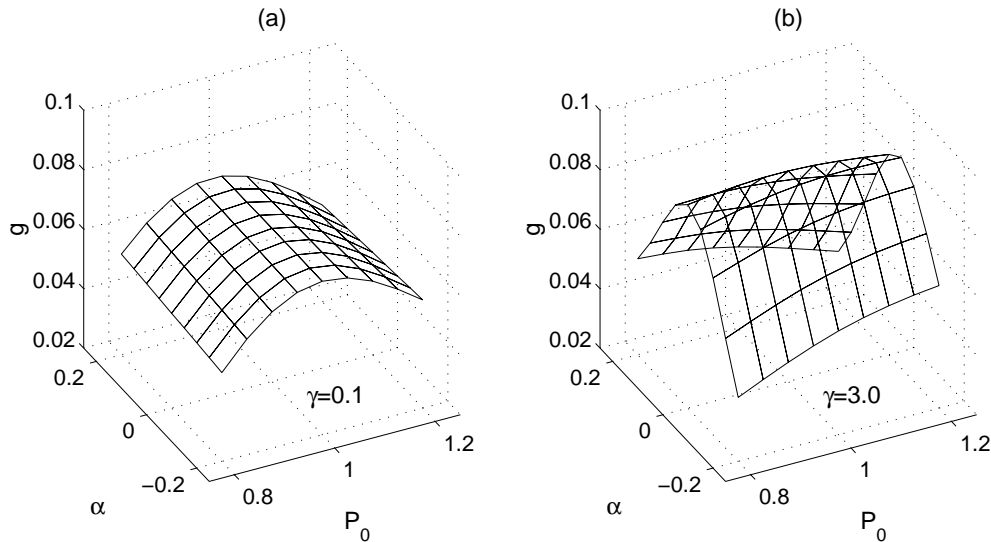


Figure 4: Granularity g is plotted as a function of P_0 and α . The option is a European call (put) option with the strike price $K = 1$. Parameters of the stock price process are $\sigma = 0.2$, $\mu = 0.05$. The stock price process is given by (3.4). Mean-reversion parameter γ takes two values: (a) $\gamma = 0.1$ and (b) $\gamma = 3.0$.

5.1 Sample-Path Properties of Tracking Errors

Recall that the normalized tracking error *process* is defined as:

$$\sqrt{N} \epsilon_t^{(N)} = \sqrt{N} (H(t, P_t) - V_t^{(N)}) \quad , \quad t \in [0, 1] \quad .$$

It can be shown that $\sqrt{N} \epsilon_t^{(N)}$ converges weakly to the stochastic process G_t , characterized by the stochastic integral in (2.12) as a function of its upper limit:¹⁵

$$G_t = \frac{1}{\sqrt{2}} \int_0^t \sigma^2(t, P_s) P_s^2 \frac{\partial^2 H(s, P_s)}{\partial P_s^2} dW'_s \quad .$$

This stronger notion of convergence yields stronger versions of Theorem 1 and 2 that can be used to analyze a number of sample-path properties of the tracking error by appealing

¹⁵The proof of this result consists of two steps. The first step is to establish that the sequence of measures induced by $\sqrt{N} \epsilon_t^{(N)}$ is tight (relatively compact). This can be done by verifying local inequalities for the moments of processes $\sqrt{N} \epsilon_t^{(N)}$ using the machinery developed in the first part of Appendix A.2 (we must use Burkholder's inequality instead of the isometric property and Hölder's inequality instead of Schwarz's inequality throughout). The second step is to characterize the limiting process. Such a characterization follows from the proof in Appendix A.1.2 and the fact that the results in Duffie and Protter (1992) guarantee weak convergence of stochastic processes, not just convergence of their one-dimensional marginal distributions.

to the Continuous Mapping Theorem (see Billingsley [1968]). This well-known result shows that the asymptotic distribution of any continuous functional $\xi(\cdot)$ of the normalized tracking error is given by $\xi(G_t)$. For example, the maximum of the normalized tracking error over the entire life of the derivative security, $\max_t \sqrt{N} \epsilon_t^{(N)}$, is distributed as $\max_t G_t$ asymptotically.

These results can be applied to the normalized tracking errors of American-style derivatives in a straightforward manner. Such derivatives differ from European derivatives in one respect: they can be exercised prematurely. Therefore, the valuation of these derivatives consists of both computing the derivative price function $H(t, P_t)$ and the optimal exercise schedule, which can be represented as a stopping time τ . Then the tracking error at the moment when the derivative is exercised behaves asymptotically as G_τ / \sqrt{N} .¹⁶ The tracking error, conditional on the derivative not being exercised prematurely, is distributed asymptotically as $(G_1 / \sqrt{N} \mid \tau = 1)$.

5.2 Joint Distributions of Tracking Errors and Prices

Theorems 1 and 2 provide a complete characterization of the tracking error and RMSE for individual derivatives, but what is often of more practical interest is the behavior of a portfolio of derivatives. Delta-hedging a portfolio of derivatives is typically easier because of the effects of diversification—as long as tracking errors are not perfectly correlated across derivatives, the portfolio tracking error will be less volatile than the tracking error of individual derivatives.

To address portfolio issues, we require the joint distribution of tracking errors for multiple stocks, as well as the joint distribution of tracking errors and prices. Consider another stock with price $P_t^{(2)}$ governed by the diffusion equation

$$\frac{dP_t^{(2)}}{P_t^{(2)}} = \mu^{(2)}(t, P_t^{(2)}) dt + \sigma^{(2)}(t, P_t^{(2)}) dW_t^{(2)} \quad (5.1)$$

where $W_t^{(2)}$ can be correlated with W_t . According to the proof of Theorem 1(b) (see Appendix A.1.2), since the random variables $(W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i)$ and $W_{t_{i+1}}^{(2)} - W_{t_i}^{(2)}$ are

¹⁶Some technical regularity conditions, e.g., the smoothness of the exercise boundary, are required to ensure convergence. See, for example, Kushner and Dupuis (1992).

uncorrelated,¹⁷ the Wiener processes W'_t and $W_t^{(2)}$ are independent. Therefore, as N increases without bound the pair of random variables $(\sqrt{N} \epsilon_1^{(N)}, P_1^{(2)})$ converges in distribution to:

$$\left(\sqrt{N} \epsilon_1^{(N)}, P_1^{(2)} \right) \Rightarrow \left(\frac{1}{\sqrt{2}} \int_0^1 \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW'_t, P_1^{(2)} \right) \quad (5.2)$$

where W'_t is independent of W_t and $W_t^{(2)}$.

An immediate corollary of this result is that the normalized tracking error is uncorrelated with any asset in the economy. This follows easily from (5.2) since, conditional on the realization of P_t and $P_t^{(2)}$, $t \in [0, 1]$, the normalized tracking error has zero expected value asymptotically. However, this does not imply that the asymptotic joint distribution of $(\sqrt{N} \epsilon_1^{(N)}, P_1^{(2)})$ does not depend on the correlation between W_t and $W_t^{(2)}$ —it does, since this correlation determines the joint distribution of P_t and $P_t^{(2)}$.

The above argument applies without change when the price of the second stock follows a diffusion process different from (5.1), e.g., (2.19). It can also easily be extended to the case of multiple stocks.

To derive the joint distribution of the normalized tracking errors for multiple stocks, we consider the case of two stocks since the generalization to multiple stocks is obvious. Let W_t and $W_t^{(2)}$ have mutual variation $dW_t dW_t^{(2)} = \rho(t, P_t, P_t^{(2)}) dt$, where $\rho(\cdot)$ is a continuously differentiable function with bounded first-order partial derivatives. We have already established that the asymptotic distribution of the tracking error is characterized by the stochastic integral (2.12). To describe the asymptotic joint distribution of two normalized tracking errors, it is sufficient to find the mutual variation of the Wiener processes in the corresponding stochastic integrals. According to the proof of Theorem 1(b) (Appendix A.1.2), this amounts to computing the expected value of the product

$$\left((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i) \right) \left((W_{t_{i+1}}^{(2)} - W_{t_i}^{(2)})^2 - (t_{i+1} - t_i) \right).$$

Using Itô's formula, it is easy to show that the expected value of the above expression is

¹⁷This follows from the fact that, for every pair of standard normal random variables X and Y with correlation ρ , $X = \rho Y + \sqrt{1 - \rho^2} Z$, where Z is a standard normal random variable, independent of Y . Thus X and $Y^2 - 1$ are uncorrelated.

equal to

$$\mathbf{E}_0 \left[2\rho^2(t, P_t, P_t^{(2)}) \right] (\Delta t)^2 + \mathbf{O} \left((\Delta t)^{\frac{5}{2}} \right) .$$

This implies that $\rho^2(t, P_t, P_t^{(2)})$ is the mutual variation of the two Wiener processes in the stochastic integrals (2.12) that describe the asymptotic distributions of the normalized tracking errors of the two stocks. Together with Theorem 1(b), this completely determines the asymptotic joint distribution of the two normalized tracking errors.¹⁸

Note that the correlation of two Wiener processes describing the asymptotic behavior of two normalized tracking errors is always nonnegative, regardless of the sign of the mutual variation of the original Wiener processes W_t and $W_t^{(2)}$. In particular, when two derivatives have convex price functions, this means that even if the returns on the two stocks are negatively correlated, the tracking errors resulting from delta hedging derivatives on these stocks are asymptotically positively correlated.

5.3 A PDE Characterization of the Tracking Error

It is possible to derive an alternative characterization of the tracking error using the intimate relationship between diffusion processes and PDE's. Although this may seem superfluous given the analytical results of Theorems 1–3, the numerical implementation of a PDE representation is often computationally more efficient.

To illustrate our approach, we begin with the RMSE. According to Theorem 1(c), the RMSE can be completely characterized asymptotically if g is known (see (2.14)). Using the Feynman-Kac representation of the solutions of PDE's (see Karatzas and Shreve (1991, Proposition 4.2.)), we conclude that $g^2 = u(0, P_0)$, where $u(t, x)$ solves the following:

$$\left[\frac{\partial}{\partial t} + \mu(t, x)x \frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(t, x)x^2 \frac{\partial^2}{\partial x^2} \right] u(t, x) + \frac{1}{2} \left(\sigma^2(t, x)x^2 \frac{\partial^2 H(t, x)}{\partial x^2} \right)^2 = 0 \quad (5.3)$$

$$u(1, x) = 0, \quad \forall x . \quad (5.4)$$

¹⁸This result generalizes the findings of Boyle and Emanuel (1980).

The PDE (5.3)–(5.4) is of the same degree of difficulty as the fundamental PDE (2.2)–(2.3) that must be solved to obtain the derivative-pricing function $H(t, P_t)$. This new representation of the RMSE can be used to implement an efficient numerical procedure for calculating RMSE without resorting to Monte Carlo simulation.¹⁹

Summary measures of the tracking error with general loss functions can also be computed numerically along the same lines, using the Kolmogorov backward equation. The probability density function of the normalized tracking error $\sqrt{N}\epsilon_1^{(N)}$ can be determined numerically as a solution of the Kolmogorov forward equation (see, for example, Karatzas and Shreve (1991, pp. 368–369)).

5.4 Alternative Measures of the Tracking Error

As we observed in Section 2.2, the root-mean-squared error is only one of many possible summary measures of the tracking error. An obvious alternative is the L_p -norm:

$$\mathbf{E}_0 \left[\left| \epsilon_1^{(N)} \right|^p \right]^{\frac{1}{p}} \quad (5.5)$$

where p is chosen so that the expectation is finite (otherwise the measure will not be particularly informative). More generally, the tracking error can be summarized by

$$\mathbf{E}_0 \left[U(\epsilon_1^{(N)}) \right] \quad (5.6)$$

where $U(\cdot)$ is an arbitrary loss function.

Consider the set of measures (5.5) first and assume for simplicity that $p \in [1, 2]$. From (2.17), it follows that

$$\mathbf{E}_0 \left[\left| \epsilon_1^{(N)} \right|^p \right]^{\frac{1}{p}} \sim N^{-1/2} \mathbf{E}_0 \left[\left| \frac{1}{\sqrt{2}} \int_0^1 \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW_t' \right|^p \right]^{\frac{1}{p}}. \quad (5.7)$$

hence the moments of the stochastic integral in (2.17) describe the asymptotic behavior of the moments of the tracking error. Conditional on the realization of $\{P_t\}$, $t \in [0, 1]$, the stochastic integral on the right side of (5.7) is normally distributed with zero mean and

¹⁹Results of some preliminary numerical experiments provide encouraging evidence of the practical value of this new representation.

variance

$$\int_0^1 \left(\sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} \right)^2 dt$$

which follows from Hull and White (1987). The intuition is that, conditional on the realization of the integrand, the stochastic integral behaves as an integral of a deterministic function with respect to the Wiener process which is a normal random variable. Now let m_p denote an L_p -norm of the standard normal random variable.²⁰ Then (5.7) can be rewritten as:

$$\mathbf{E}_0 \left[\left| \epsilon_1^{(N)} \right|^p \right]^{\frac{1}{p}} \sim \frac{m_p}{\sqrt{N}} \mathbf{E}_0 \left[\mathcal{R}^{\frac{p}{2}} \right]^{\frac{1}{p}} \quad (5.8)$$

where \mathcal{R} is given by (2.15).

As in the case of a quadratic loss function, \mathcal{R} plays a fundamental role here in describing the behavior of the tracking error. When $p = 2$, \mathcal{R} enters (5.8) linearly and closed-form expressions can be derived for special cases. However, even when $p \neq 2$, the qualitative impact of \mathcal{R} on the tracking error is the same as for $p = 2$ and our discussion of the qualitative behavior of the tracking error applies to this case as well.

For general loss functions $U(\cdot)$ that satisfy certain growth conditions and are sufficiently smooth near the origin, the delta-method can be applied and we obtain:

$$\mathbf{E}_0 \left[U \left(\epsilon_1^{(N)} \right) \right] \sim \frac{1}{N} |U''(0)| g^2 = \frac{1}{N} |U''(0)| \mathbf{E}_0[\mathcal{R}]. \quad (5.9)$$

When $U(\cdot)$ is not differentiable at 0, the delta method cannot be used. However, we can use the same strategy as in our analysis of L_p -norms to tackle this case. Suppose that $U(\cdot)$ is dominated by a quadratic function. Then

$$\mathbf{E}_0 \left[U \left(\epsilon_1^{(N)} \right) \right] \approx \mathbf{E}_0 \left[U \left(\frac{1}{\sqrt{2N}} \int_0^1 \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW_t' \right) \right]. \quad (5.10)$$

²⁰If X is a standard normal random variable, then $m_p = \mathbf{E}[|X|^p]^{1/p}$.

Now let

$$m_U(x) = \mathbf{E}[U(x\eta)] \quad , \quad \eta \sim \mathcal{N}(0, 1) .$$

Then

$$\mathbf{E}_0 [U(\epsilon_1^{(N)})] \approx \mathbf{E}_0 \left[m_U \left(\sqrt{\mathcal{R}/N} \right) \right] . \quad (5.11)$$

When the loss function $U(\cdot)$ is convex, $m_U(\cdot)$ is an increasing function (by second-order stochastic dominance). Therefore, the qualitative behavior of the measure (5.6) is also determined by \mathcal{R} and is the same as that of the RMSE.

6 Conclusions

We have argued that continuous-time models are meant to be approximations to physical phenomena, and as such, their approximation errors should be better understood. In the specific context of continuous-time models of derivative securities, we have quantified the approximation error through our definition of temporal granularity. The combination of a specific derivative security and a stochastic process for the underlying asset’s price dynamics can be associated with a measure of how “grainy” the passage of time is. This measure is related to the ability to replicate the derivative security through a delta-hedging strategy implemented in discrete time. Time is said to be very granular if the replication strategy does not work well—in such cases, time is not continuous. If, however, the replication strategy is very effective, time is said to be very smooth or continuous.

We consider the complete-markets case first, and under the assumption of general Markov diffusion price dynamics, we show that that the tracking errors for derivatives with sufficiently smooth or continuous piecewise linear payoff functions behave asymptotically (in distribution) as G/\sqrt{N} . We characterize the distribution of the random variable G as a stochastic integral, and also obtain the joint distribution of G with prices of other assets and with other tracking errors. We demonstrate that the root-mean-squared error behaves asymptotically as g/\sqrt{N} , where the constant g is what we call the coefficient of *temporal granularity*. For two special cases—call or put options on geometric Brownian motion and on an Ornstein-

Uhlenbeck process—we are able to evaluate the coefficient of granularity explicitly.

The incomplete-markets case provided an interesting contrast to the complete-markets case. In this case, the root-mean-squared error of the tracking error behaves asymptotically as $a + g/\sqrt{N}$, the first term arising from market incompleteness and hence independent of the number of trading periods, and the second term arising from discrete trading.

We also consider a number of extensions of our analysis, including an extension to alternative loss functions, a demonstration of the weak convergence of the tracking error process, a derivation of the joint distribution of tracking errors and prices, and an alternative characterization of the tracking error in terms of PDE's that can be used for efficient numerical implementation.

Because these results depend so heavily on continuous-record asymptotics, we perform Monte Carlo simulations to check the quality of our asymptotics. For the case of European puts and calls with geometric Brownian motion price dynamics, our asymptotic approximations are excellent, providing extremely accurate inferences over the range of empirically relevant parameter values, even with a small number of trading periods.

Now our definition of granularity is not invariant to the derivative security, the underlying asset's price dynamics, and other variables. But we regard this as a positive feature of our approach, not a drawback. After all, any plausible definition of granularity must be a relative one, balancing the coarseness of changes in the time domain against the coarseness of changes in the “space” or price domain. Although the title of this paper suggests that time is the main focus of our analysis, it is really the relation between time and price that determines whether or not continuous-time models are good approximations to physical phenomena. It is our hope that the definition of granularity proposed in this paper is one useful way of tackling this very complex issue.

A Appendix

The essence of these proofs involves the relation between the delta-hedging strategy and mean-square approximations of solutions of systems of stochastic differential equations described in Milstein (1974, 1987, 1995). Readers interested in additional details and intuition should consult these references directly.

A.1 Proof of Theorem 1

A.1.1 Theorem 1(a)

First we observe that the regularity conditions (2.10) imply the existence of a positive constant K_1 such that

$$\left| \frac{\partial^{\beta+\gamma}}{\partial \tau^\beta \partial x^\gamma} H(\tau, x) \right| \leq K_1 \quad (\text{A.1})$$

for $(\tau, x) \in [0, 1] \times [0, \infty)$, $0 \leq \beta \leq 1$, $1 \leq \gamma \leq 4$, and all partial derivatives are continuous.²¹

Next, by Itô's formula,

$$\begin{aligned} H(1, P_1) &= H(0, P_0) + \int_0^1 \left(\frac{\partial H(t, P_t)}{\partial t} + \frac{1}{2} \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} \right) dt + \\ &\quad \int_0^1 \frac{\partial H(t, P_t)}{\partial P_t} dP_t. \end{aligned} \quad (\text{A.2})$$

According to (2.2), the first integral on the right-hand side of (A.2) is equal to zero. Thus,

$$H(1, P_1) = H(0, P_0) + \int_0^1 \frac{\partial H(t, P_t)}{\partial P_t} dP_t \quad (\text{A.3})$$

²¹Since the price of the derivative $H(\tau, x)$ is defined as a solution of (2.2), it is equal to the expectation of $F(P_1)$ with respect to the equivalent martingale measure (see Duffie (1996)), i.e.,

$$H(\tau, x) = \mathbf{E}_{(t=\tau, P_t^*=x)}[F(P_1^*)]$$

where

$$\frac{dP_t^*}{P_t^*} = \sigma(t, P_t^*) dW_t^* .$$

and W_t^* is a Brownian motion under the equivalent martingale measure. Equation (A.1) now follows from Friedman (1975; Theorems 5.4 and 5.5, p. 122). The same line of reasoning is followed in He (1989, p. 68). Of course, one could derive (A.1) using purely analytic methods, e.g. Friedman (1964; Theorem 10, p. 72, Theorem 11, p. 24; and Theorem 12, p. 25).

which implies that $H(t, P_t)$ can be characterized as a solution of the system of stochastic differential equations

$$\begin{cases} dX_t &= \frac{\partial H(t, P_t)}{\partial P_t} \mu(t, P_t) P_t dt + \frac{\partial H(t, P_t)}{\partial P_t} \sigma(t, P_t) P_t dW_t \\ dP_t &= \mu(t, P_t) P_t dt + \sigma(t, P_t) P_t dW_t \end{cases}. \quad (\text{A.4})$$

At the same time, $V_1^{(N)}$ is given by

$$V_1^{(N)} = H(0, P_0) + \sum_{i=0}^{N-1} \frac{\partial H(t, P_t)_{t=t_i}}{\partial P_t} (P_{t_{i+1}} - P_{t_i}), \quad (\text{A.5})$$

which can be interpreted as a solution of the following approximation scheme of (A.4) (as defined in Milstein (1987)):

$$\begin{cases} \bar{X}_{t_{i+1}} - X_{t_i} &= \frac{\partial H(t, P_t)_{t=t_i}}{\partial P_t} (P_{t_{i+1}} - P_{t_i}) \\ \bar{P}_{t_{i+1}} - P_{t_i} &= P_{t_{i+1}} - P_{t_i}, \end{cases} \quad (\text{A.6})$$

where \bar{X} and \bar{P} denote approximations to X and P , respectively. We now compare (A.6) to the Euler approximation scheme in Milstein (1995)

$$\begin{cases} \bar{X}_{t_{i+1}} - X_{t_i} &= \frac{\partial H(t, P_t)_{t=t_i}}{\partial P_t} \mu(t_i, P_{t_i}) P_{t_i} (t_{i+1} - t_i) + \\ &\quad \frac{\partial H(t, P_t)_{t=t_i}}{\partial P_t} \sigma(t_i, P_{t_i}) P_{t_i} (W_{t_{i+1}} - W_{t_i}) \\ \bar{P}_{t_{i+1}} - P_{t_i} &= \mu(t_i, P_{t_i}) (t_{i+1} - t_i) + \\ &\quad \sigma(t_i, P_{t_i}) (W_{t_{i+1}} - W_{t_i}) \end{cases}. \quad (\text{A.7})$$

Regularity conditions (2.10) and (A.1) allow us to conclude (see Milstein (1995, Theorem 2.1)) that a one-step version of the approximation scheme (A.7) has order-of-accuracy 2 in expected deviation and order-of-accuracy 1 in mean-square deviation (see Milstein (1987), Milstein (1995) for definitions and discussion). It is easy to check that the approximation scheme (A.6) exhibits this same property. Milstein (1995, Theorem 1.1) relates the one-step order-of-accuracy of the approximation scheme to its order-of-accuracy on the whole interval (see also Milstein (1987)). We use this theorem to conclude that (A.6) has mean-square

order-of-accuracy 1/2, i.e.,

$$\sqrt{\mathbf{E}_0 \left[\left(X(1, P_1) - \bar{X}(1, P_1) \right)^2 \right]} = \mathbf{O} \left(\frac{1}{\sqrt{N}} \right). \quad (\text{A.8})$$

We now recall that $X(t, P_t) = H(t, P_t)$ and $\bar{X}(1, P_1) = V_1^{(N)}$ and conclude that

$$\sqrt{\mathbf{E}_0 \left[\left(H(1, P_1) - V_1^{(N)} \right)^2 \right]} = \mathbf{O} \left(\frac{1}{\sqrt{N}} \right) \quad (\text{A.9})$$

which completes the proof. ■

A.1.2 Theorem 1(b)

We follow the same line of reasoning as in the proof of Theorem 1(a), but we use the Milstein approximation scheme for (A.4) instead of the Euler scheme:

$$\left\{ \begin{array}{l} \bar{X}_{t_{i+1}} - X_{t_i} = \frac{\partial H(t, P_t)_{t=t_i}}{\partial P_t} \mu(t_i, P_{t_i}) P_{t_i} (t_{i+1} - t_i) + \\ \quad \frac{\partial H(t, P_t)_{t=t_i}}{\partial P_t} \sigma(t_i, P_{t_i}) P_{t_i} (W_{t_{i+1}} - W_{t_i}) + \\ \quad \left(\frac{\partial^2 H(t, P_t)_{t=t_i}}{\partial P_t^2} \sigma(t_i, P_{t_i}) P_{t_i} + \frac{\partial H(t, P_t)_{t=t_i}}{\partial P_t} \frac{\partial(\sigma(t, P_t) P_t)_{t=t_i}}{\partial P_t} \right) \times \\ \quad \frac{1}{2} \sigma(t_i, P_{t_i}) P_{t_i} \left((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i) \right) \\ \bar{P}_{t_{i+1}} - P_{t_i} = \mu(t_i, P_{t_i}) P_{t_i} (t_{i+1} - t_i) + \sigma(t_i, P_{t_i}) P_{t_i} (W_{t_{i+1}} - W_{t_i}) + \\ \quad \frac{1}{2} \sigma(t_i, P_{t_i}) P_{t_i} \frac{\partial(\sigma(t, P_t) P_t)_{t=t_i}}{\partial P_t} \left((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i) \right). \end{array} \right. \quad (\text{A.10})$$

According to Milstein (1974) (see also Milstein (1995, Theorem 2.1)), this one-step scheme has order-of-accuracy 2 in expected deviation and 1.5 in mean-square deviation. It is easy to check by comparison that the scheme

$$\left\{ \begin{array}{l} \bar{X}_{t_{i+1}} - X_{t_i} = \frac{\partial H(t, P_t)_{t=t_i}}{\partial P_t} (P_{t_{i+1}} - P_{t_i}) + \\ \quad \frac{1}{2} \sigma(t_i, P_{t_i})^2 P_{t_i}^2 \frac{\partial^2 H(t, P_t)_{t=t_i}}{\partial P_t^2} \left((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i) \right) \\ \bar{P}_{t_{i+1}} - P_{t_i} = P_{t_{i+1}} - P_{t_i} \end{array} \right. \quad (\text{A.11})$$

has the same property. We now use Milstein (1995, Theorem 1.1) to conclude that

$$H(1, P_1) - V_1^{(N)} = \sum_{i=0}^{N-1} \frac{1}{2} \sigma(t_i, P_{t_i})^2 P_{t_i}^2 \frac{\partial^2 H(t, P_t)_{t=t_i}}{\partial P_t^2} \times \left((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i) \right) + \mathbf{O}\left(\frac{1}{N}\right) \quad (\text{A.12})$$

where $f = \mathbf{O}\left(\frac{1}{N}\right)$ means that $\lim_{N \rightarrow \infty} N \sqrt{\mathbf{E}_{t=0}[f^2]} < \infty$. By Slutsky's theorem, we can ignore the $\mathbf{O}\left(\frac{1}{N}\right)$ term in considering the convergence in distribution of $\sqrt{N}(H(1, P_1) - V_1^{(N)})$, since $\sqrt{N} \mathbf{O}\left(\frac{1}{N}\right)$ converges to zero in mean-squared and, therefore, also in probability. Observe now that, since $W_{t_{i+1}} - W_{t_i}$ and $W_{t_{j+1}} - W_{t_j}$ are independent for $i \neq j$, $(W_{t_{i+1}} - W_{t_i})^2$ and $W_{t_{i+1}} - W_{t_i}$ are uncorrelated, $\mathbf{E}_0[(W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i)] = 0$ and $\mathbf{E}_0[((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i))^2] = 2/(t_{i+1} - t_i)^2$, by the functional central limit theorem (see Ethier and Kurtz (1986)), a piecewise constant martingale

$$\sqrt{N/2} \sum_{i=0}^{[Nt]-1} \left((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i) \right) \quad (\text{A.13})$$

converges weakly on $[0, 1]$ to a standard Brownian motion W_t' , which is independent of W_t .²² We complete the proof by applying Duffie and Protter (1992, Lemma 5.1 and Corollary 5.1). ■

A.1.3 Theorem 1(c)

Equation (2.13) follows immediately from Theorem 1(a) and the proof of Theorem 1(b).²³ Combined with Theorem 1(b), (2.13) implies that

$$g = \sqrt{\frac{1}{2} \mathbf{E}_0 \left[\left(\int_0^1 \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW_t' \right)^2 \right]}. \quad (\text{A.14})$$

Equation (2.14) follows from (A.14) using the isometric property of stochastic integrals. ■

²²The notation $[Nt]$ denotes the integer part of Nt and we use the convention $\sum_0^{-1} = 0$.

²³Relation (A.12), established as a part of the proof of Theorem 1(b), guarantees that convergence in (2.12) occurs not only in distribution, but also in mean-squared.

A.2 Proof of Theorem 2

Before we present the proof, we establish the following result:

$$\mathbf{E}_{t=0} \left[\int_0^1 \left| \frac{\partial^2 H(t, P_t)}{\partial P_t^2} \right|^{2p} dt \right] < \infty \quad (\text{A.15})$$

for some $p > 1$. First, since $F(x)$ is a continuous piecewise linear function, it suffices to establish (A.15) for $F(x) = \max(0, x - y)$, $y \geq 0$. Using this definition of $F(x)$ and the fact that one can differentiate (2.2) with respect to P_t (see Friedman (1964, Theorem 10, p. 72)), $\partial^2 H(\tau, x)/\partial x^2$ is equal to the fundamental solution of the Cauchy problem for the following partial differential equation:

$$\frac{\partial u(\tau, x)}{\partial \tau} + \frac{1}{2} \frac{\partial^2}{\partial x^2} [\sigma^2(\tau, x) x^2 u(\tau, x)] = 0 .$$

After a logarithmic change of variables, conditions (2.10), (2.16), and $\sigma(\tau, x) \geq \sigma_0 > 0$ allow us to apply Friedman (1975, Theorem 4.5, p. 141), from which we conclude that there exist positive constants K_4 and K_5 , such that

$$\left| \frac{\partial^2 H(\tau, x)}{\partial x^2} \right| = |u(\tau, x)| \leq K_4 (1 - \tau)^{-\frac{1}{2}} \exp \left[-K_5 \frac{|\ln(x) - \ln(y)|^2}{1 - \tau} \right] .$$

By the same theorem and Friedman (1975, Theorem 5.4, p. 149), the distribution of P_t has a density, which is bounded above by $K_6 t^{-1/2}$, $K_6 > 0$. Now, by direct computation, we find that there exists a positive constant K_7 such that

$$\mathbf{E}_0 \left[\left| \frac{\partial^2 H(t, P_t)}{\partial P_t^2} \right|^{2p} \right] \leq K_7 (1 - t)^{-p + \frac{1}{2}} .$$

Condition (A.15) now follows by Fubini's theorem.

To prove part (a) of the theorem, we use Itô's formula and (2.2) to establish that

$$H(1, P_1) - V_1^{(N)} = H(1, P_1) - H(0, P_0) - \sum_{i=0}^{N-1} \frac{\partial H(t, P_t)_{t=t_i}}{\partial P_t} (P_{t_{i+1}} - P_{t_i})$$

$$\begin{aligned}
&= \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\frac{\partial H(t, P_t)}{\partial P_t} - \frac{\partial H(t, P_t)_{t=t_i}}{\partial P_t} \right) dP_t \\
&= \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^t \frac{\partial^2 H(\tau, P_\tau)}{\partial P_\tau^2} dP_\tau \right) dP_t .
\end{aligned}$$

We now use (2.1) to rewrite the last expression as a sum of four terms:

$$\begin{aligned}
&\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^t \frac{\partial^2 H(\tau, P_\tau)}{\partial P_\tau^2} \mu(\tau, P_\tau) P_\tau d\tau \right) \mu(t, P_t) P_t dt \\
&+ \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^t \frac{\partial^2 H(\tau, P_\tau)}{\partial P_\tau^2} \mu(\tau, P_\tau) P_\tau d\tau \right) \sigma(t, P_t) P_t dW_t \\
&+ \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^t \frac{\partial^2 H(\tau, P_\tau)}{\partial P_\tau^2} \sigma(\tau, P_\tau) P_\tau dW_\tau \right) \mu(t, P_t) P_t dt \\
&+ \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^t \frac{\partial^2 H(\tau, P_\tau)}{\partial P_\tau^2} \sigma(\tau, P_\tau) P_\tau dW_\tau \right) \sigma(t, P_t) P_t dW_t \\
&\equiv I_1 + I_2 + I_3 + I_4 .
\end{aligned}$$

Now we will show that, for $k = 1, 2, 3$, $\lim_{N \rightarrow \infty} N \mathbf{E}_0[I_k^2] = 0$. Consider the term I_1 first.

Using Schwarz's inequality, we conclude that

$$\begin{aligned}
&N \mathbf{E}_0[I_1^2] \\
&\leq N^2 \mathbf{E}_0 \left[\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^t \left(\frac{\partial^2 H(\tau, P_\tau)}{\partial P_\tau^2} \mu(\tau, P_\tau) P_\tau \right)^2 d\tau \cdot (t - t_i) \right) \mu^2(t, P_t) P_t^2 dt \cdot \Delta t \right] \\
&\leq \mathbf{E}_0 \left[\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^t \left(\frac{\partial^2 H(\tau, P_\tau)}{\partial P_\tau^2} \mu(\tau, P_\tau) P_\tau \right)^2 d\tau \right) \mu^2(t, P_t) P_t^2 dt \right] \\
&= \Delta t \mathbf{E}_0 \left[\frac{1}{2} \sum_{i=0}^{N-1} \left(\frac{\partial^2 H(t, P_t)_{t=\tau_i^*}}{\partial P_t^2} \right)^2 \mu^2(\tau_i^*, P_{\tau_i^*}) P_{\tau_i^*}^2 \mu^2(t_i^*, P_{t_i^*}) P_{t_i^*}^2 \Delta t \right]
\end{aligned}$$

where $\Delta t \equiv t_{i+1} - t_i = 1/N$. Random times τ_i^* and t_i^* satisfy $t_i \leq \tau_i^* \leq t_i^* \leq t_{i+1}$ and depend on a particular realization of the stock price process. Their existence is guaranteed by the

Mean-Value Theorem. The sum under the expectation sign converges to the integral

$$\frac{1}{2} \int_0^1 \left(\frac{\partial^2 H(t, P_t)}{\partial P_t^2} \mu^2(t, P_t) P_t^2 \right)^2 dt$$

so we need only to justify passing the limit through the expectation. This can be done using the Lebesgue dominated convergence theorem. Observe that the integral sum is bounded above by

$$\begin{aligned} & \frac{1}{\Delta t} \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^{t_{i+1}} \left(\frac{\partial^2 H(\tau, P_\tau)}{\partial P_\tau^2} P_\tau \right)^2 d\tau \right) \max_{(t,S) \in [0,1] \times [0,\infty]} \mu^4(t, S) \max_{t \in [0,1]} P_t^2 dt \\ &= \max_{(t,S) \in [0,1] \times [0,\infty]} \mu^4(t, S) \int_0^1 \left(\frac{\partial^2 H(t, P_t)}{\partial P_t^2} P_t \right)^2 \max_{t \in [0,1]} P_t^2 dt \end{aligned}$$

By Hölder's inequality, the expected value of the integral in the last expression is bounded above by

$$\mathbf{E}_0 \left[\int_0^1 \left(\frac{\partial^2 H(t, P_t)}{\partial P_t^2} \right)^{2p} dt \right]^{\frac{1}{p}} \mathbf{E}_0 \left[\max_{t \in [0,1]} P_t^{\frac{4p}{p-1}} \right]^{\frac{p-1}{p}}$$

for any $p > 1$. Regularity conditions on functions $\mu(t, S)$ and $\sigma(t, S)$ imply that the second term is finite (see Friedman (1975, Theorem 2.3, p. 107)). The first term is finite by (A.15). This allows us to apply the Lebesgue dominated convergence theorem.

For $k = 2$ we use the isometric property of the stochastic integral first and then apply Schwarz's inequality. As a result we obtain

$$\begin{aligned} N \mathbf{E}_0 [I_2^2] &= N \mathbf{E}_0 \left[\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^t \frac{\partial^2 H(\tau, P_\tau)}{\partial P_\tau^2} \mu(\tau, P_\tau) P_\tau d\tau \right)^2 \sigma^2(t, P_t) P_t^2 dt \right] \\ &\leq \mathbf{E}_0 \left[\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^t \left(\frac{\partial^2 H(\tau, P_\tau)}{\partial P_\tau^2} \mu(\tau, P_\tau) P_\tau \right)^2 d\tau \right) \sigma^2(t, P_t) P_t^2 dt \right]. \end{aligned}$$

This last expression converges to zero, by the same argument as in the $k = 1$ case. Similar arguments (apply Schwarz's inequality first and then the isometric property) prove the $k = 3$ case.

We now consider the last term, I_4 . Using the isometric property of the stochastic integral, we have

$$\begin{aligned}\mathbf{E}_0[I_4^2] &= \mathbf{E}_0 \left[\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\int_{t_i}^t \left(\frac{\partial^2 H(\tau, P_\tau)}{\partial P_\tau^2} \right)^2 \sigma^2(\tau, P_\tau) P_\tau^2 d\tau \right) \sigma^2(t, P_t) P_t^2 dt \right] \\ &= \mathbf{E}_0 \left[\sum_{i=0}^{N-1} \left(\frac{\partial^2 H(t, P_t)_{t=\tau_i^*}}{\partial P_t^2} \right)^2 \sigma^2(\tau_i^*, P_{\tau_i^*}) P_{\tau_i^*}^2 \sigma^2(t_i^*, P_{t_i^*}) P_{t_i^*}^2 \right] \frac{(\Delta t)^2}{2}\end{aligned}$$

for some $t_i \leq \tau_i^* \leq t_i^* \leq t_{i+1}$. Now, by the same argument as in the $k = 1$ case,

$$\begin{aligned}\lim_{N \rightarrow \infty} N \mathbf{E}_0[I_4^2] &= \lim_{N \rightarrow \infty} \mathbf{E}_0 \left[\sum_{i=0}^{N-1} \left(\frac{\partial^2 H(t, P_t)_{t=\tau_i^*}}{\partial P_t^2} \right)^2 \sigma^2(\tau_i^*, P_{\tau_i^*}) P_{\tau_i^*}^2 \sigma^2(t_i^*, P_{t_i^*}) P_{t_i^*}^2 \right] \frac{\Delta t}{2} \\ &= \frac{1}{2} \int_0^1 \mathbf{E}_0 \left[\left(\frac{\partial^2 H(t, P_t)}{\partial P_t^2} \sigma^2(t, P_t) P_t^2 \right)^2 \right] dt,\end{aligned}$$

which completes the proof of part (a) of the theorem.

Using part (a) and Markov's inequality, we conclude that the sequence of probability measures induced on the real line by $\sqrt{N} (H(1, P_1) - V_1^{(N)})$ is tight and therefore relatively compact (see Billingsley (1986)). This implies that every subsequence of the original sequence of probability measures contains a weakly converging subsequence. To prove part (b) of the theorem, we need to show that every weakly converging subsequence of the original sequence converges weakly to the stochastic integral in (2.17). Consider a weakly converging subsequence $\sqrt{N_k} (H(1, P_1) - V_1^{(N_k)})$. To show that

$$\sqrt{N_k} (H(1, P_1) - V_1^{(N_k)}) \Rightarrow \frac{1}{\sqrt{2}} \int_0^1 \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW'_t,$$

it suffices to check that for any function $\phi(x) = \exp(\sqrt{-1} \lambda x)$,

$$\mathbf{E}_0 \left[\phi \left(\sqrt{N_k} (H(1, P_1) - V_1^{(N_k)}) \right) \right] \rightarrow \mathbf{E}_0 \left[\phi \left(\frac{1}{\sqrt{2}} \int_0^1 \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW'_t \right) \right].$$

We would like to show that for any $\epsilon > 0$, there exists an integer K , such that $k > K$ implies

that

$$\begin{aligned} & \left| \mathbf{E}_0 \left[\phi \left(\sqrt{N_k} (H(1, P_1) - V_1^{(N_k)}) \right) \right] - \right. \\ & \left. \mathbf{E}_0 \left[\phi \left(\frac{1}{\sqrt{2}} \int_0^1 \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW_t' \right) \right] \right| < \epsilon. \end{aligned} \quad (\text{A.16})$$

As before, let $V_r^{(N_k)}$ denote the value of the replicating portfolio at time r . Then

$$\begin{aligned} & \left| \mathbf{E}_0 \left[\phi \left(\sqrt{N_k} (H(1, P_1) - V_1^{(N_k)}) \right) \right] - \mathbf{E}_0 \left[\phi \left(\frac{1}{\sqrt{2}} \int_0^1 \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW_t' \right) \right] \right| \\ & \leq \left| \mathbf{E}_0 \left[\phi \left(\sqrt{N_k} (H(r, P_r) - V_r^{(N_k)}) \right) \right] - \mathbf{E}_0 \left[\phi \left(\frac{1}{\sqrt{2}} \int_0^r \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW_t' \right) \right] \right| \\ & + \left| \mathbf{E}_0 \left[\phi \left(\sqrt{N_k} (H(1, P_1) - V_1^{(N_k)}) \right) \right] - \mathbf{E}_0 \left[\phi \left(\sqrt{N_k} (H(r, P_r) - V_r^{(N_k)}) \right) \right] \right| \\ & + \left| \mathbf{E}_0 \left[\phi \left(\frac{1}{\sqrt{2}} \int_0^1 \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW_t' \right) \right] - \right. \\ & \quad \left. \mathbf{E}_0 \left[\phi \left(\frac{1}{\sqrt{2}} \int_0^r \sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} dW_t' \right) \right] \right| \\ & \equiv J_1 + J_2 + J_3. \end{aligned}$$

Fix an arbitrary $\epsilon' > 0$. Given (A.15), we can always pick r such that

$$\mathbf{E}_0 \left[\frac{1}{2} \int_r^1 \left(\sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} \right)^2 dt \right] < \epsilon'.$$

According to Theorem 2,²⁴ there exists K_1 , such that $k > K_1$ implies that $|J_1| < \epsilon'$. Since $|\phi'(x)| \leq |\lambda|$,

$$\mathbf{E}_0 [J_2^2] \leq \lambda^2 N_k \mathbf{E}_0 \left[\left(H(1, P_1) - V_1^{(N_k)} - H(r, P_r) + V_r^{(N_k)} \right)^2 \right]. \quad (\text{A.17})$$

Using the arguments of part (a), we conclude that there exists K_2 , such that $k > K_2$ implies

²⁴Necessary regularity conditions can be established using (2.16), Friedman (1964, Theorem 10, p. 72) and Friedman (1975, Theorems 4.5 and 4.6, pp. 141–142).

that

$$\begin{aligned} & \mathbf{E}_0 \left[\left(H(1, P_1) - V_1^{(N_k)} - H(r, P_r) + V_r^{(N_k)} \right)^2 \right] \\ & < \mathbf{E}_0 \left[\frac{1}{2} \int_r^1 \left(\sigma^2(t, P_t) P_t^2 \frac{\partial^2 H(t, P_t)}{\partial P_t^2} \right)^2 dt \right] + \epsilon' < 2\epsilon'. \end{aligned}$$

Therefore, for $k > K_2$, $\mathbf{E}_0 [J_2^2] < 2\lambda^2\epsilon$. It is also clear from the previous argument that $\mathbf{E}_0 [J_3^2] < \lambda^2\epsilon$. Thus, if we set $K = \max(K_1, K_2)$ and $\epsilon' = (1 + 3\lambda^2)^{-1}\epsilon$, inequality (A.16) will hold for any $k > K$. This completes the proof of part (b) of the theorem. ■

A.3 Proof of Theorem 3

By Itô's formula,

$$\begin{aligned} H(1, \mathbf{Z}_1) &= \int_0^1 \left(\frac{\partial H(t, \mathbf{Z}_t)}{\partial t} + \frac{1}{2} \sum_{j,k=0}^J a_{jk}(t, \mathbf{Z}_t) Z_t^j Z_t^k \frac{\partial^2 H(t, \mathbf{Z}_t)}{\partial Z_t^j \partial Z_t^k} \right) dt + \\ & \int_0^1 \sum_{j=0}^J \frac{\partial H(t, \mathbf{Z}_t)}{\partial Z_t^j} dZ_t^j + H(0, \mathbf{Z}_0). \end{aligned}$$

According to (2.20), this is equivalent to

$$H(1, \mathbf{Z}_1) = H(0, \mathbf{Z}_0) + \int_0^1 \sum_{j=0}^J \frac{\partial H(t, \mathbf{Z}_t)}{\partial Z_t^j} dZ_t^j - \int_0^1 \left(\sum_{j=1}^J \hat{\mu}_j(t, \mathbf{Z}_t) Z_t^j \frac{\partial H(t, \mathbf{Z}_t)}{\partial Z_t^j} \right) dt.$$

On the other hand,

$$V_1^{(N)} = H(0, \mathbf{Z}_0) + \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(\frac{\partial H(t, \mathbf{Z}_t)_{t=t_i}}{\partial Z_t^0} + \sum_{j=1}^J \lambda_j(t, \mathbf{Z}_t) \frac{\partial H(t, \mathbf{Z}_t)_{t=t_i}}{\partial Z_t^j} \right) dZ_t^0.$$

Therefore, the tracking error can be written as

$$\epsilon_1^{(N)} = I_1 + I_2 + I_3$$

where I_1 and I_2 are given by (2.26), (2.27) and

$$I_3 = \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \sum_{j=0}^J \left(\lambda_j(t, \mathbf{Z}_t) \frac{\partial H(t, \mathbf{Z}_t)}{\partial Z_t^j} - \lambda_j(t_i, \mathbf{Z}_{t_i}) \frac{\partial H(t, \mathbf{Z}_t)_{t=t_i}}{\partial Z_t^j} \right) dZ_t^0 .$$

Now we can use the same argument as in the proofs of Theorems 1(b) and 2 to show that

$$\mathbf{E}_0[I_3^2] = \mathbf{O}\left(\frac{1}{N}\right)$$

and

$$\mathbf{E}_0[(I_1 + I_2)I_3] = \frac{g}{\sqrt{N}} + \mathbf{O}\left(\frac{1}{N}\right) .$$

Thus,

$$\begin{aligned} \text{RMSE}^{(N)} &= \mathbf{E}_0 \left[(I_1 + I_2)^2 \right]^{\frac{1}{2}} \left(1 + \frac{2A\mathbf{E}_0 \left[(I_1 + I_2)^2 \right]^{-\frac{1}{2}}}{\sqrt{N}} + \mathbf{O}\left(\frac{1}{N}\right) \right)^{\frac{1}{2}} \\ &= \mathbf{E}_0 \left[(I_1 + I_2)^2 \right]^{\frac{1}{2}} + \frac{A}{\sqrt{N}} + \mathbf{O}\left(\frac{1}{N}\right) . \end{aligned}$$

and this completes the proof. ■

A.4 Proof of Theorem 4

Equation (3.2) follows from (2.18). The closed-form expression for the option price as a function of time and stock price is given by the Black-Scholes option-pricing formula (see Black and Scholes (1973)). The diffusion coefficient of a geometric Brownian motion is constant and the price of the stock P_t has lognormal distribution. We use Fubini's theorem to change the order of integration in (3.2) and calculate expected values in closed form. ■

A.5 Proof of Theorem 5

Equation (3.5) follows from (2.18).²⁵ The closed-form expression for the option price as a function of time and stock price is given by the Black-Scholes option-pricing formula (see Black and Scholes (1973)). The diffusion coefficient of a mean-reverting process is constant and the price of the stock P_t has a lognormal distribution. We use Fubini's theorem to change the order of integration in (3.2), and we calculate expected values in closed form. ■

²⁵The regularity conditions of Theorem 2 are not satisfied here. This is not surprising, since in our derivation we assumed that the stock price process can be characterized as a strong solution of the corresponding stochastic differential equation. Therefore, the regularity conditions that we impose on the coefficients of such equation are at least as strong as those required by the existence and uniqueness theorem for stochastic differential equations. For the case of the mean-reverting process these regularity conditions are not satisfied (the growth rate of the drift coefficient is faster than linear). However, the stock price process is still well defined: there exists a unique solution of the diffusion equation (3.3) and the stock price process is obtained from it by exponentiation. Now it is straightforward to verify that our derivation of (2.18) is still valid.

References

- Avellaneda, M. and A. Paras, 1994, “Dynamic Hedging Portfolios for Derivative Securities in the Presence of Large Transactions Costs”, *Applied Mathematical Finance* 1, 165–193.
- Bachelier, L., 1900, “Theory of Speculation,” reprinted in P. Cootner, ed., 1964, *The Random Character of Stock Market Prices*. Cambridge, MA: MIT Press.
- Bensaid, B., Lesne, J., Pages, H. and J. Scheinkman, 1992, “Derivative Asset Pricing with Transaction Costs”, *Mathematical Finance* 2, 63–86.
- Billingsley, P., 1986, *Probability and Measure*, Second Edition. New York: John Wiley & Sons.
- Bertsimas, D., L. Kogan, A. Lo, 1997, “Pricing and Hedging Derivative Securities in Incomplete Markets: an ϵ -Arbitrage Approach”, working paper, Massachusetts Institute of Technology.
- Black, F., M. Scholes, 1973, “Pricing of Options and Corporate Liabilities”, *Journal of Political Economy*, 81, 637–654.
- Boyle, P., D. Emanuel, 1980, “Discretely Adjusted Option Hedges”, *Journal of Financial Economics*, 8, 259–282.
- Boyle, P. and T. Vorst, 1992, “Option Replication in Discrete Time with Transaction Costs”, *Journal of Finance* 47, 271–294.
- Cox, J., Ingersoll, J. and S. Ross, 1985, “An Intertemporal General Equilibrium Model of Asset Prices”, *Econometrica*, 36, 363–384.
- Cox, J., Ross, S. and M. Rubinstein, 1979, “Option Pricing: A Simplified Approach”, *Journal of Financial Economics* 7, 229–263.
- Duffie, D., 1996, *Dynamic Asset Pricing Theory*, Second Edition. Princeton, NJ: Princeton University Press.
- Duffie, D., P. Protter, 1992, “From Discrete- to Continuous-Time Finance: Weak Convergence of the Financial Gain Process”, *Mathematical Finance*, 2, 1–15.
- Duffie, D. and C. Huang, 1985, “Implementing Arrow-Debreu Equilibria by Continuous Trading of Few Long-Lived Securities”, *Econometrica* 53, 1337–1356.
- Eberlein, E., 1992, “On Modeling Questions in Security Valuation”, *Mathematical Finance*, 2, 17–32.
- Edirisinghe, C., Naik, V. and R. Uppal, 1993, “Optimal Replication of Options with Transaction Costs and Trading Restrictions”, *Journal of Financial and Quantitative Analysis* 28, 117–138.
- Ethier, S., T. Kurtz, 1986, *Markov Processes: Characterization and Convergence*. New York: J. Wiley & Sons.
- Foley, D., 1975, “On Two Specifications of Asset Equilibrium in Macroeconomic Models”, *Journal of Political Economy* 83, 303–324.

- Friedman, A., 1964, *Partial Differential Equations of Parabolic Type*. Englewood Cliffs, NJ: Prentice-Hall.
- Friedman, A., 1975, *Stochastic Differential Equations and Applications*, Volume 1. New York: Academic Press.
- Garman, M., 1976, “A General Theory of Asset Valuation under Diffusion State Processes”, Working Paper No. 50, University of California Berkeley.
- Grannan, E. and G. Swindle, 1996, “Minimizing Transaction Costs of Option Hedging Strategies”, *Mathematical Finance* 6, 341–364.
- Harrison, J.M., D.M. Kreps, 1979, “Martingales and Arbitrage in Multiperiod Securities Markets”, *Journal of Economic Theory*, 20, 381–408.
- He, H., 1989, “Essays in Dynamic Portfolio Optimization and Diffusion Estimations”, Ph.D. Thesis, MIT.
- He, H., 1990, “Convergence from Discrete- to Continuous- Time Contingent Claim Prices”, *Review of Financial Studies*, 3, 523–546.
- He, H., 1991, “Optimal Consumption-Portfolio Policies: A Convergence from Discrete to Continuous Time Models”, *Journal of Economic Theory*, 55, 340–363.
- He, H. and N. Pearson, 1991, “Consumption and Portfolio Policies with Incomplete Markets and Short-Sale Constraints: The Infinite-Dimensional Case”, *Journal of Economic Theory* 54, 259–304.
- Henrotte, P., 1993, “Transactions Costs and Duplication Strategies”, unpublished working paper, Stanford University.
- Huang, C., 1987, “An Intertemporal General Equilibrium Asset Pricing Model: The Case of Diffusion Information”, *Econometrica* 55, 117–142.
- Hull, J., A. White, 1987, “The Pricing of Options on Assets with Stochastic Volatilities”, *Journal of Finance*, 42, 281–300.
- Hutchinson, J., Lo, A., T. Poggio, 1994, “A Nonparametric Approach to the Pricing and Hedging of Derivative Securities Via Learning Networks”, *Journal of Finance*, 49, 851–889.
- Itô, K., 1951, “On Stochastic Differential Equations”, *Memoirs of the American Mathematical Society* 4, 1–51.
- Jerison, D., Singer, I. and D. Stroock, eds., 1997, *The Legacy of Norbert Wiener: A Centennial Symposium*. Providence, RI: American Mathematical Society.
- Karatzas, I. and S. Shreve, 1991, *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag.
- Kushner, H. and P. Dupuis, 1992, *Numerical Methods for Stochastic Control Problems in Continuous Time*. New York: Springer Verlag.
- Leland, H., 1985, “Option Pricing and Replication with Transaction Costs”, *Journal of Finance*, 40, 1283–1301.

- Lo, A. and J. Wang, 1995, “Implementing Option Pricing Models When Asset Returns Are Predictable”, *Journal of Finance* 50, 87–129.
- Merton, R., 1969, “Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case”, *Review of Economics and Statistics* 51, 247–257.
- Merton, R., 1973, “Theory of Rational Option Pricing”, *Bell Journal of Economics and Management Science* 4, 141–183.
- Merton, R., 1975, “Theory of Finance from the Perspective of Continuous Time”, *Journal of Financial and Quantitative Analysis* 10, 659–674.
- Merton, R., 1982a, “On the Microeconomic Theory of Investment Under Uncertainty”, in K. Arrow and M. Intriligator, eds., *Handbook of Mathematical Economics*, Volume II. Amsterdam: North-Holland Publishers.
- Merton, R., 1982b, “On the Mathematics and Economic Assumptions of Continuous-Time Models”, in W. Sharpe and C. Cootner, eds., *Financial Economics: Essays in Honor of Paul Cootner*. Englewood Cliffs, NJ: Prentice-Hall.
- Milstein, G.N., 1995, *Numerical Integration of Stochastic Differential Equations*. Boston, MA: Kluwer Academic Publishers.
- Milstein, G.N., 1974, “Approximate Integration of Stochastic Differential Equations”, *Theory Prob. Appl.*, 19, 557–563.
- Milstein, G.N., 1987, “A Theorem on the Order of Convergence of Mean-Square Approximations of Solutions of Systems of Stochastic Differential Equations”, *Theory Prob. Appl.*, 32, 738–741.
- Platen, E., 1981, “An Approximation Method for a Class of Itô Processes”, *Lit. Mat. Sb.*, 21, 121–133 (In Russian).
- Rubinstein, M., 1994, “Implied Binomial Trees”, *Journal of Finance* 49, 771–818.
- Stroock, D., 1993, *Probability Theory: an Analytic View*. Cambridge, UK: Cambridge University Press.
- Toft, K., 1996, “On the Mean-Variance Tradeoff in Option Replication with Transactions Costs”, *Journal of Financial and Quantitative Analysis*, 31, 233–263.
- Wiener, N., 1923, “Differential Space”, *Journal of Mathematical Physics* 2, 131–174.
- Wiggins, J., 1987, “Option Values Under Stochastic Volatility: Theory and Empirical Estimates”, *Journal of Financial Economics*, 5, 351–372.
- Willinger, W., Taqqu M. S., 1991, “Towards a Convergence Theory for Continuous Stochastic Securities Market Models”, *Mathematical Finance*, 1, 55–99.