The Annals of Applied Probability
1991, Vol. 1, No. 3, 379–405

DIFFUSION APPROXIMATION IN PAST DEPENDENT MODELS AND APPLICATIONS TO OPTION PRICING

By Paolo Kind,[1] Robert Sh. Liptser[2] and Wolfgang J. Runggaldier

University of Chicago, USSR Academy of Sciences and Università di Padova

We obtain a diffusion approximation result for processes satisfying equations with past-dependent coefficients. We apply this result to a model of option pricing, in which the underlying asset price volatility depends on past evolution, and obtain a generalized (asymptotic) Black and Scholes formula.

1. Introduction. The purpose of this paper is twofold. A first (theoretical) purpose is to provide a diffusion approximation result (which to the best of our knowledge is new) for a certain class of processes satisfying equations with past-dependent coefficients. We assume, in particular, that the dependence on the past is through the quadratic variation process. The main result states that the pairs, given by the processes and their quadratic variations, converge in a suitable sense to a limit pair, where the first component is the limit process and the second component is its quadratic variation. Since the limiting quadratic variation process satisfies a deterministic delay equation, if the initial condition is known (deterministic), then the entire limiting quadratic variation process is deterministic; this in turn implies that our limit process is a Gauss–Markov diffusion process.

The second purpose is more of an applied nature and concerns the problem of determining the value function for risky financial operations, in particular for European call options. Since in real world situations the volatility of asset price returns is subject to random and time-varying changes, which can often be explained by changes in past sample volatilities of asset prices, we believe that in a consistent model of option valuation one should explicitly account for the dependence of future asset price distributions on past evolution. Assuming that asset price observations occur more realistically at discrete time points (with the intervals between successive observations being small with respect to the total time interval) and that the dependence on the past is through a dependence of the asset price volatility on the quadratic variation, we recover for the asset price evolution the class of processes that is the object of our theoretical investigations. Although it is possible to derive an option valuation formula when asset prices evolve according to processes of the above type, the

Received January 1990; revised August 1990.

[1]Currently with Salomon Bros. Int. Ltd., London.

[2]Work performed during a stay in Padova supported by GNAFA/CNR.

AMS 1980 *subject classifications*. Primary 60F17; secondary 90A09.

Key words and phrases. Diffusion approximation, stochastic delay equations, option pricing, Black and Scholes formula.

actual computation is practically impossible. It turns out, however, that the limit process, obtained by letting the observation intervals tend to zero, is of the lognormal type with a deterministic quadratic variation process; for such a process one can therefore apply the standard Black and Scholes formula [Black and Scholes (1973)]. The convergence result of the theoretical part then allows us to approximate the option value of the more realistic discrete observation and past dependent case by that of the limit process, for which an explicitly computable formula is available.

In Section 1.1, we introduce the class of processes that is the object of our study; Section 1.2 states the assumptions and the main convergence result and Section 1.3 explains more closely the applied aspects of the paper. In Section 2 we prove our main convergence results. In Section 3 we describe in more detail the option pricing problem for the past-dependent setting; here we also apply the results of Section 2 to derive the approximation procedure leading to the asymptotic option valuation formula for the above more general setting. In Remark 3.3 of Section 3 we also discuss the relation of the present paper to other work concerning the question of continuous-time versus discrete-time trading. An Appendix contains two auxiliary results for Section 2.

1.1. *The model.* We start by defining the model for the sequence of stochastic processes $(X_t^\varepsilon, Y_t^\varepsilon)_{\varepsilon>0}$ that will be the main object of our study. For this purpose consider a sequence of random variables $(x_{t_k})_{-\infty<k<+\infty}$, where $(t_k)_{-\infty<k<+\infty}$ is a sequence of deterministic time points such that for a given $\varepsilon > 0$,

$$t_0 = 0, \qquad t_{k+1} - t_k = \varepsilon. \tag{1.1}$$

Given a positive number I, define for $t > -I$

$$[X^\varepsilon, X^\varepsilon]_t := \sum_{k:\ -I<t_k\le t} (x_{t_k} - x_{t_{k-1}})^2 \tag{1.2}$$

and let

$$[X^\varepsilon, X^\varepsilon]_s^t := [X^\varepsilon, X^\varepsilon]_t - [X^\varepsilon, X^\varepsilon]_s, \qquad t > s > -I.$$

We suppose that for $k \ge 0$, the sequence (x_{t_k}) satisfies the following recursive relation

$$\begin{aligned} x_{t_{k+1}} = x_{t_k} &+ \Big\{a_{t_{k+1}}\Big([X^\varepsilon, X^\varepsilon]_{t_{k+1}-I}^{t_k}\Big) + \hat{a}_{t_{k+1}}\Big([X^\varepsilon, X^\varepsilon]_{t_{k+1}-I}^{t_k}\Big)x_{t_k}\Big\} \\ &\times (t_{k+1} - t_k) + b_{t_{k+1}}\Big([X^\varepsilon, X^\varepsilon]_{t_{k+1}-I}^{t_k}\Big)(t_{k+1} - t_k)^{1/2}\xi_{k+1}, \end{aligned} \tag{1.3}$$

where $a_t(y)$, $\hat{a}_t(y)$, $b_t(y)$ are given functions in (t, y) and $(\xi_k)_{k\ge1}$ is a sequence of i.i.d. random variables, independent of the random sequence $(x_{t_k}, -\infty < k \le 0)$, with

$$E\xi_1 = 0, \qquad E(\xi_1)^2 = 1. \tag{1.4}$$

Let us now introduce the following processes:

$$X_t^\varepsilon := x_{t_k}, \qquad 0 \le t_k \le t < t_{k+1}, \tag{1.5}$$

$$A_t^\varepsilon := t_k, \qquad t_k \le t < t_{k+1}, \tag{1.6}$$

$$M_t^\varepsilon := \varepsilon^{1/2} \sum_{k:\, t_k \le t} \xi_k, \qquad M_0^\varepsilon = 0, \tag{1.7}$$

as well as the filtration $\mathscr{F}^\varepsilon = (\mathscr{F}_t^\varepsilon)_{t \ge 0}$, where

$$\mathscr{F}_t^\varepsilon = \mathscr{F}_{t_k}, \qquad t_k \le t < t_{k+1}, \tag{1.8}$$

with

$$\mathscr{F}_{t_k} = \sigma\{x_{t_j}, -I \le t_j \le 0; \xi_j, 1 \le j \le k\}. \tag{1.9}$$

From the above definitions we have immediately that A_t^ε is an increasing and right continuous process whose jumps ΔA_t^ε do not exceed ε, that is,

$$\Delta A_t^\varepsilon \le \varepsilon \quad (\text{more precisely, either } \Delta A_t^\varepsilon = 0 \text{ or } \Delta A_t^\varepsilon = \varepsilon) \tag{1.10}$$

and

$$[A^\varepsilon, A^\varepsilon]_t = \sum_{0 < s \le t} (\Delta A_s^\varepsilon)^2. \tag{1.11}$$

Furthermore,

$$A_t^\varepsilon \le t, \qquad \sup_t |A_t^\varepsilon - t| \to_{\varepsilon \to 0} 0. \tag{1.12}$$

$M^\varepsilon = (M_t^\varepsilon)_{t \ge 0}$ is an $\mathscr{F}^\varepsilon$-square integrable martingale with

$$\langle M^\varepsilon \rangle_t = A_t^\varepsilon \tag{1.13}$$

and

$$[M^\varepsilon, M^\varepsilon]_t = \sum_{0 < s \le t} (\Delta M_s^\varepsilon)^2. \tag{1.14}$$

Notice that the process $[A^\varepsilon, M^\varepsilon] = ([A^\varepsilon, M^\varepsilon]_t)_{t \ge 0}$ with

$$[A^\varepsilon, M^\varepsilon]_t = \sum_{0 < s \le t} \Delta A_s^\varepsilon \Delta M_s^\varepsilon = \int_0^t \Delta A_s^\varepsilon \, dM_s^\varepsilon \tag{1.15}$$

is [see, e.g., Problem 2.3.5 in Liptser and Shiryayev (1989)] also an $\mathscr{F}^\varepsilon$-square integrable martingale.

Taking into account (1.3), we can now give the following representation for the process $X^\varepsilon = (X_t^\varepsilon)$ defined in (1.5):

$$\begin{aligned} X_t^\varepsilon = x_0 &+ \int_0^t \{a_s([X^\varepsilon, X^\varepsilon]_{s-I}^{s-}) + \hat{a}_s([X^\varepsilon, X^\varepsilon]_{s-I}^{s-}) X_{s-}^\varepsilon\} \, dA_s^\varepsilon \\ &+ \int_0^t b_s([X^\varepsilon, X^\varepsilon]_{s-I}^{s-}) \, dM_s^\varepsilon, \end{aligned} \tag{1.16}$$

where $h_{s-} := \lim_{u \uparrow s} h_u$. From (1.16) [equivalent to (1.3)], we also obtain the following representation, for $t \ge 0$, of the process $([X^\varepsilon, X^\varepsilon]_t)$ defined according to (1.2):

$$
\begin{aligned}
[X^\varepsilon, X^\varepsilon]_t = [X^\varepsilon, X^\varepsilon]_0 &+ \int_0^t \{a_s([X^\varepsilon, X^\varepsilon]_{s-I}^{s-}) \\
&\qquad + \hat{a}_s([X^\varepsilon, X^\varepsilon]_{s-I}^{s-}) X_{s-}^\varepsilon\}^2 \, d[A^\varepsilon, A^\varepsilon]_s \\
&+ 2\int_0^t \{a_s([X^\varepsilon, X^\varepsilon]_{s-I}^{s-}) + \hat{a}_s([X^\varepsilon, X^\varepsilon]_{s-I}^{s-}) X_{s-}^\varepsilon\} \\
&\quad \times b_s([X^\varepsilon, X^\varepsilon]_{s-I}^{s-}) \, d[A^\varepsilon, M^\varepsilon]_s \\
&+ \int_0^t b_s^2([X^\varepsilon, X^\varepsilon]_{s-I}^{s-}) \, d[M^\varepsilon, M^\varepsilon]_s.
\end{aligned} \tag{1.17}
$$

Letting

$$
Y_t^\varepsilon := [X^\varepsilon, X^\varepsilon]_t \tag{1.18}
$$

we finally obtain for the sequence $(X_t^\varepsilon, Y_t^\varepsilon)_{t \ge 0}$ the model

$$
\begin{aligned}
X_t^\varepsilon = x_0 &+ \int_0^t \{a_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) + \hat{a}_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) X_{s-}^\varepsilon\} \, dA_s^\varepsilon \\
&+ \int_0^t b_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) \, dM_s^\varepsilon,
\end{aligned} \tag{1.19}
$$

$$
\begin{aligned}
Y_t^\varepsilon = Y_0^\varepsilon &+ \int_0^t \{a_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) + \hat{a}_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) X_{s-}^\varepsilon\}^2 \, d[A^\varepsilon, A^\varepsilon]_s \\
&+ 2\int_0^t \{a_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) + \hat{a}_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) X_{s-}^\varepsilon\} \\
&\quad \times b_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) \, d[A^\varepsilon, M^\varepsilon]_s \\
&+ \int_0^t b_s^2(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) \, d[M^\varepsilon, M^\varepsilon]_s.
\end{aligned} \tag{1.20}
$$

Notice that Y_t^ε is an increasing process.

1.2. *Assumptions and main results.* We shall make the following assumptions on model (1.19), (1.20).

ASSUMPTION 1. The function $\hat{a}_t(y)$ is continuous in t and y and bounded, that is, $|\hat{a}_t(y)| \le L$.

ASSUMPTION 2. $a_t(y)$ and $b_t^2(y)$ are continuous in t and Lipschitz continuous in y with Lipschitz constant not depending on t; furthermore, $\sup_t(|a_t(0)| + |b_t(0)|) < \infty$.

ASSUMPTION 3. $\limsup_{\varepsilon \to 0} EY_0^\varepsilon < \infty$.

ASSUMPTION 4. $E|x_0| < \infty$.

We assume, furthermore, that for a given nondecreasing deterministic process $\tilde{Y}_u$, $-I \le u \le 0$, we have

$$\sup_{-I \le u \le 0} |Y_u^\varepsilon - \tilde{Y}_u| \to_P 0. \tag{1.21}$$

Consider now the process (Y_t) defined, for $t \ge 0$, by

$$Y_t = \tilde{Y}_0 + \int_0^t b_s^2(Y_s - Y_{s-I})\, ds \tag{1.22}$$

with $Y_{s-I} = \tilde{Y}_{s-I}$ for $s \le I$, and notice that by the Lipschitz property of $b_t^2(y)$ (Assumption 2), the equation (1.22) has a unique solution. Furthermore, let (X_t) be given by

$$\begin{aligned} X_t = x_0 &+ \int_0^t \{a_s(Y_s - Y_{s-I}) + \hat{a}_s(Y_s - Y_{s-I}) X_{s-}\}\, ds \\ &+ \int_0^t b_s(Y_s - Y_{s-I})\, dW_s \end{aligned} \tag{1.23}$$

with $W = (W_t)$ a Wiener process.

Our main results consist of the following two theorems, which will be proved in Section 2, and where $\to_P$ denotes convergence in probability while $\to_L$ stands for convergence in distribution (law).

THEOREM 1.1. *Consider Assumptions* 1–4 *as well as* (1.21). *We then have for arbitrary* $T > 0$,

$$\sup_{-I \le t \le T} |Y_t^\varepsilon - Y_t| \to_P 0 \quad \text{for } \varepsilon \to 0. \tag{1.24}$$

THEOREM 1.2. *Again, consider Assumptions* 1–4 *and* (1.21); *then* $X^\varepsilon \to_L X$.

It is worth remarking that in our case the quadratic variation of the semimartingale X_t is

$$[X, X]_t = \int_0^t b_s^2(Y_s - Y_{s-I})\, ds \tag{1.25}$$

so that by (1.22),

$$Y_t = Y_0 + [X, X]_t, \tag{1.26}$$

that is, the process (Y_t) is the quadratic variation process of (X_t).

Since $\tilde{Y}_u$, $-I \le u \le 0$, is deterministic, by (1.22), Y_t, $t \ge 0$, is also deterministic so that X_t in (1.23) is a Gauss–Markov diffusion process. The main purpose of proving Theorems 1.1 and 1.2 is therefore equivalent to giving a diffusion approximation result for a process X^ε that [see (1.16)] depends on its past over a finite interval of length I.

1.3. *A financial application.* As already mentioned, the model (1.19) [equivalent to that of (1.3)] is motivated [see Kind (1988)] by the problem of determining the value function for risky operations, in particular by that of pricing European call options.

The classical result for option pricing is the celebrated Black and Scholes formula [Black and Scholes (1973)] which is based on a continuous-time asset price evolution model of the lognormal type, where the price volatility (standard deviation) is deterministic. In reality, although it may be reasonable to assume that asset prices evolve in continuous time, they can be observed only at discrete time points (generally with short intervals between successive observations), and there is some kind of nonstationarity present in their distribution, in particular in the standard deviation. It turns out that changes in volatility of asset price returns can often be explained by changes in past asset price sample volatilities. In a consistent model of option valuation, the specification of the asset price process should thus explicitly account for the dependence of the future asset price distribution on past evolution. [For a related discussion see also Section 7.4 of Aase (1988), where a Monte Carlo-type simulation is suggested to compute the value of the option.] It will be shown in Section 3.1 below that a special case of model (1.19) may well accomplish the above purpose, thereby providing the basis for a theory of option pricing under more general and possibly more realistic assumptions than those underlying the Black and Scholes model.

The greater generality of the underlying model raises a first question concerning the existence of a perfect hedge portfolio for the option. It is clear, however, that the possibility of pricing contingent claims on the basis of arbitrage considerations alone should not depend on the specific mathematical model for the underlying asset price process, but only on economic considerations. In fact, as shown already in Harrison and Pliska (1981) and (1983), a perfect hedge portfolio exists also under very general assumptions. Subsequent work by Bensoussan (1984) and Karatzas (1989) show such a property more specifically for models that include the ones considered here. All of the above works also provide a formula for the computation of the value of a contingent claim that is a direct generalization of the Black and Scholes formula stating that the value of the claim is equal to the expected discounted value of the cumulative payoff under the risk-neutral probability measure. Such a formula will be described in more detail in Section 3.1, but its explicit computation in the more general nonlognormal case will be practically impossible. An approximation procedure is therefore in order.

The main result of the paper concerning the convergence of the discrete time process $(X^\varepsilon, Y^\varepsilon)$ to (X, Y) provides the basis for such an approximation procedure that will be the main subject of Section 3.2. Since Y is deterministic (and consequently X is a Gauss–Markov diffusion process), the central feature of this approximation procedure is as follows.

Assume the asset price past trajectory is given; then, if the time intervals between the discrete asset price observations are small, the distribution (under the risk-neutral probability measure) of any future asset price is close to a lognormal distribution with a time-varying variance that can be computed explicitly in terms of asset price past returns. With the limiting lognormal asset price distribution, one can then compute the standard Black and Scholes formula. In this way, for the above more general situation, one obtains an asymptotically (for the discrete time observation intervals tending to zero) exact closed-form option valuation formula given by a modified version of the standard Black and Scholes formula, where in place of a constant asset price volatility one has to substitute a function of the asset price past returns.

Besides the important qualification concerning the approximation character of the Black and Scholes formula for the limit process X, there is also the important aspect that it provides a closed-form option valuation formula that explicitly accounts for the fact that different realizations of the underlying asset price process may affect the value of an option not only through a direct underlying asset price effect, but also through an indirect conditional distribution effect.

2. Proofs of the main results.

2.1. *Proof of Theorem* 1.1. The proof of Theorem 1.1 follows immediately from Lemma 2.1 and Lemma 2.2. We have:

LEMMA 2.1. *Let the process* $\hat{Y}_t^\varepsilon$ *be defined by*

$$\hat{Y}_t^\varepsilon = Y_0^\varepsilon + \int_0^t b_s^2\left(\hat{Y}_{s-}^\varepsilon - \hat{Y}_{s-I}^\varepsilon\right) d[M^\varepsilon, M^\varepsilon]_s, \tag{2.1}$$

where $\hat{Y}_{s-I}^\varepsilon = Y_{s-I}^\varepsilon$ *for* $s \le I$. *Then, under the assumptions of Theorem* 1.1,

$$\sup_{t \le T} |\hat{Y}_t^\varepsilon - Y_t| \to_P 0 \quad \textit{for } \varepsilon \to 0. \tag{2.2}$$

PROOF. We have

$$\begin{aligned}
\hat{Y}_t^\varepsilon - Y_t &= Y_0^\varepsilon - \tilde{Y}_0 + \int_0^t b_s^2\left(\hat{Y}_{s-}^\varepsilon - \hat{Y}_{s-I}^\varepsilon\right) d[M^\varepsilon, M^\varepsilon]_s - \int_0^t b_s^2(Y_{s-} - Y_{s-I})\, ds \\
&= \left(Y_0^\varepsilon - \tilde{Y}_0\right) + \int_0^t \left\{ b_s^2\left(\hat{Y}_{s-}^\varepsilon - \hat{Y}_{s-I}^\varepsilon\right) - b_s^2(Y_{s-}^\varepsilon - Y_{s-I})\right\} d[M^\varepsilon, M^\varepsilon]_s \qquad (2.3)\\
&\quad + \int_0^t b_s^2(Y_{s-} - Y_{s-I})\, d([M^\varepsilon, M^\varepsilon]_s - s).
\end{aligned}$$

By the Lipschitz property of $b_t^2(y)$ (Assumption 2), we have for $s \ge 0$:

$$\begin{aligned}
&|b_s^2\left(\hat{Y}_{s-}^\varepsilon - \hat{Y}_{s-I}^\varepsilon\right) - b_s^2(Y_{s-} - Y_{s-I})| \\
&\quad \le L\left(|\hat{Y}_{s-}^\varepsilon - Y_{s-}| + |\hat{Y}_{s-I}^\varepsilon - Y_{s-I}|\right) \\
&\quad = L\left(|\hat{Y}_{s-}^\varepsilon - Y_{s-}| + I(s \le I)|Y_{s-I}^\varepsilon + \tilde{Y}_{s-I}| + I(s > I)|\hat{Y}_{s-I}^\varepsilon - Y_{s-I}|\right) \qquad (2.4)\\
&\quad \le 2L\left\{ \sup_{0 \le u \le s} |\hat{Y}_{u-}^\varepsilon - Y_{u-}| + \sup_{-I \le u \le 0} |Y_u^\varepsilon - \tilde{Y}_u| \right\}
\end{aligned}$$

with $I(\cdot)$ denoting the indicator function. Thus, for $t \le T$,

$$\begin{aligned}(2.5)\qquad &\sup_{u\le t}\left|\int_0^u \Big\{b_s^2\big(\hat{Y}_{s-}^\varepsilon - \hat{Y}_{s-I}^\varepsilon\big) - b_s^2(Y_{s-} - Y_{s-I})\Big\}\,d[M^\varepsilon, M^\varepsilon]_s\right| \\ &\quad\le 2L \sup_{-I\le u\le 0}|Y_u^\varepsilon - \tilde{Y}_u|[M^\varepsilon, M^\varepsilon]_T \\ &\qquad + 2L\int_0^t \sup_{u\le s}|\hat{Y}_{u-}^\varepsilon - Y_{u-}|\,d[M^\varepsilon, M^\varepsilon]_s.\end{aligned}$$

From (2.3) and (2.5) we obtain that

$$\begin{aligned}\sup_{0\le t\le T}|\hat{Y}_t^\varepsilon - Y_t| \le |Y_0^\varepsilon - \tilde{Y}_0| &+ 2L\sup_{-I\le u\le 0}|Y_u^\varepsilon - \tilde{Y}_u|[M^\varepsilon, M^\varepsilon]_T \\ &+ 2L\int_0^T \sup_{u\le s}|\hat{Y}_{u-}^\varepsilon - Y_{u-}|\,d[M^\varepsilon, M^\varepsilon]_s \\ &+ \sup_{u\le T}\left|\int_0^u b_s^2(Y_{s-} - Y_{s-I})\,d([M^\varepsilon, M^\varepsilon]_s - s)\right|.\end{aligned}$$

Define

$$\begin{aligned}(2.6)\qquad \alpha^\varepsilon = |Y_0^\varepsilon - \tilde{Y}_0| &+ 2L\sup_{-I\le u\le 0}|Y_u^\varepsilon - \tilde{Y}_u|[M^\varepsilon, M^\varepsilon]_T \\ &+ \sup_{u\le T}\left|\int_0^u b_s^2(Y_{s-} - Y_{s-I})\,d([M^\varepsilon, M^\varepsilon]_s - s)\right|.\end{aligned}$$

Then, for $t \le T$,

$$\sup_{0\le s\le t}|\hat{Y}_s^\varepsilon - Y_s| \le \alpha^\varepsilon + 2L\int_0^t \sup_{u\le s}|\hat{Y}_{u-}^\varepsilon - Y_{u-}|\,d[M^\varepsilon, M^\varepsilon]_s$$

and so from Liptser and Shiryayev [(1989), Theorem 2.4.3] we have for any $T > 0$,

$$(2.7)\qquad \sup_{0\le t\le T}|\hat{Y}_t^\varepsilon - Y_t| \le \alpha^\varepsilon \exp(2L[M^\varepsilon, M^\varepsilon]_T).$$

We now show that the right side of the inequality (2.7) tends to zero in probability for $\varepsilon \to 0$. For this purpose we first show that

$$(2.8)\qquad \lim_{\varepsilon\to 0}\sup_{t\le T}|[M^\varepsilon, M^\varepsilon]_t - t| = 0 \qquad P \text{ a.s. for all } T > 0.$$

From the definition of M^ε [see (1.7)] one easily obtains

$$[M^\varepsilon, M^\varepsilon]_t = \varepsilon\sum_{k=1}^{[t/\varepsilon]}\xi_k^2 = t\frac{[t/\varepsilon]}{t/\varepsilon}\frac{1}{[t/\varepsilon]}\sum_{k=1}^{[t/\varepsilon]}\xi_k^2.$$

On the other hand, by Birkhoff–Khinchin's theorem [see, e.g., Stout (1973)],

$$\lim_{\varepsilon\to 0}\frac{1}{[t/\varepsilon]}\sum_{k=1}^{[t/\varepsilon]}\xi_k^2 = E\xi_1^2 = 1, \qquad P \text{ a.s.}$$

Therefore, for all t,

$$[M^\varepsilon, M^\varepsilon]_t \mapsto t \qquad \text{for } \varepsilon \to 0, \qquad P \text{ a.s.},$$

and, consequently, (2.8) holds by Problem 5.3.2 in Liptser and Shiryayev (1989). Thus in (2.7), $\exp(2L[M^\varepsilon, M^\varepsilon]_T) \to \exp(2LT)$, P a.s., and it remains to prove that

$$\alpha^\varepsilon \to_P 0 \quad \text{for } \varepsilon \to 0. \tag{2.9}$$

From (2.8) and the given assumptions, it follows that the first two terms on the right-hand side of (2.6) tend to zero in probability for $\varepsilon \to 0$ so that (2.9) holds if

$$\sup_{u \le T} \left| \int_0^u b_s^2(Y_{s-} - Y_{s-I})\, d([M^\varepsilon, M^\varepsilon]_s - s) \right| \to_P 0 \quad \text{for } \varepsilon \to 0, \tag{2.10}$$

which in turn follows from Lemma A.2 in the Appendix by taking $G_t^\varepsilon = [M^\varepsilon, M^\varepsilon]_t$, $G_t = t$, using furthermore (2.8) and noticing that, by Assumption 2 and the definition of Y_t in (1.22), the function $b_s^2(Y_{s-} - Y_{s-I})$ is continuous in s. □

LEMMA 2.2. *With $\hat{Y}_t^\varepsilon$ given in* (2.1), *we have under the assumptions of Theorem* 1.1,

$$\sup_{t \le T} |Y_t^\varepsilon - \hat{Y}_t^\varepsilon| \to_P 0 \quad \textit{for } \varepsilon \to 0. \tag{2.11}$$

PROOF. From (1.20) and (2.1), it follows that for $t \le T$,

$$|Y_t^\varepsilon - \hat{Y}_t^\varepsilon| \le \beta^\varepsilon + \int_0^t \left|b_s^2(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) - b_s^2\left(\hat{Y}_{s-}^\varepsilon - \hat{Y}_{s-I}^\varepsilon\right)\right| d[M^\varepsilon, M^\varepsilon]_s, \tag{2.12}$$

where

$$\begin{aligned} \beta^\varepsilon = &\int_0^T \{a_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) + \hat{a}_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) X_{s-}^\varepsilon\}^2\, d[A^\varepsilon, A^\varepsilon]_s \\ &+ 2 \sup_{0 \le t \le T} \bigg| \int_0^t \{a_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) + \hat{a}_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) X_{s-}^\varepsilon\} \\ &\qquad \times b_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon)\, d[A^\varepsilon, M^\varepsilon]_s \bigg|. \end{aligned} \tag{2.13}$$

Analogously to (2.4) we obtain

$$\left|b_s^2(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) - b_s^2\left(\hat{Y}_{s-}^\varepsilon - \hat{Y}_{s-I}^\varepsilon\right)\right| \le 2L \sup_{0 \le u \le s} |Y_{u-}^\varepsilon - \hat{Y}_{u-}^\varepsilon|,$$

so that from (2.12) it follows that for $t \le T$,

$$\sup_{0 \le u \le t} |Y_u^\varepsilon - \hat{Y}_u^\varepsilon| \le \beta^\varepsilon + 2L \int_0^t \sup_{0 \le u \le s} |Y_{u-}^\varepsilon - \hat{Y}_{u-}^\varepsilon|\, d[M^\varepsilon, M^\varepsilon]_s$$

and thus by Liptser and Shiryayev [(1989), Theorem 2.4.3]

$$\sup_{0 \le u \le t} |Y_u^\varepsilon - \hat{Y}_u^\varepsilon| \le \beta^\varepsilon \exp(2L[M^\varepsilon, M^\varepsilon]_T), \qquad t \le T.$$

Since by (2.8), we have $[M^\varepsilon, M^\varepsilon]_T \to T$, P a.s., we obtain (2.11) if

$$(2.14) \qquad \beta^\varepsilon \to_P 0 \quad \text{for } \varepsilon \to 0.$$

According to the definition (2.13), $\beta^\varepsilon = \beta_1^\varepsilon + \beta_2^\varepsilon$. We now show that

$$(2.15) \qquad \beta_i^\varepsilon \to_P 0 \quad \text{for } \varepsilon \to 0, \qquad i = 1, 2.$$

For this purpose define

$$(2.16) \qquad \Gamma^\varepsilon := \{Y_T^\varepsilon \le r_1\} \cap \left\{\sup_{s \le T} |X_s^\varepsilon| \le r_2\right\}.$$

Due to the fact that the Lipschitz condition implies linear growth, it follows from Assumptions 1 and 2 that on the set Γ^ε we have the following inequality, uniformly in $s \le T$:

$$(2.17) \quad \{a_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) + \hat{a}_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon)X_{s-}^\varepsilon\}^2 b_s^2(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) \le k,$$

where the constant k depends only on r_1, r_2 and the L appearing in Assumptions 1 and 2.

Notice that an inequality of the type (2.17) holds also without the factor $b_s^2(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon)$ on the left. Since furthermore [see (1.10) and (1.11)]

$$[A^\varepsilon, A^\varepsilon]_T = \sum_{0 < s \le T} (\Delta A_s^\varepsilon)^2 \le \varepsilon \sum_{0 < s \le T} (\Delta A_s^\varepsilon) = \varepsilon A_T^\varepsilon \le \varepsilon T,$$

we then first have that

$$\beta_1^\varepsilon \le \varepsilon k T,$$

so that

$$\lim_{\varepsilon \to 0} P(\beta_1^\varepsilon > \delta, \Gamma^\varepsilon) = 0 \quad \text{for all } \delta > 0.$$

Consequently,

$$\begin{aligned} P(\beta_1^\varepsilon > \delta) &\le P(\beta_1^\varepsilon > \delta, \Gamma^\varepsilon) + P(\Omega \setminus \Gamma^\varepsilon) \\ &\le P(\beta_1^\varepsilon > \delta, \Gamma^\varepsilon) + P(Y_T^\varepsilon > r_1) + P\Big(\sup_{s \le T} |X_s^\varepsilon| > r_2\Big) \to 0 \end{aligned}$$

upon taking $\lim_{(r_1, r_2) \to \infty} \limsup_{\varepsilon \to 0}$, since by Lemma A.1 of the Appendix we have

$$(2.18) \qquad \begin{aligned} &\lim_{r_1 \to \infty} \limsup_{\varepsilon \to 0} P(Y_T^\varepsilon > r_1) = 0 \quad \text{and} \\ &\lim_{r_2 \to \infty} \limsup_{\varepsilon \to 0} P\Big(\sup_{s \le T} |X_s^\varepsilon| > r_2\Big) = 0. \end{aligned}$$

Thus we have (2.15) for $i = 1$. Let us establish (2.15) for $i = 2$. As already remarked [see (1.15)], $[A^\varepsilon, M^\varepsilon]_t = \int_0^t \Delta A_s^\varepsilon \, dM_s^\varepsilon$ is a square integrable martingale. Furthermore, the process

$$\{a_s(Y^\varepsilon_{s-} - Y^\varepsilon_{s-I}) + \hat{a}_s(Y^\varepsilon_{s-} - Y^\varepsilon_{s-I})X^\varepsilon_{s-}\}b_s(Y^\varepsilon_{s-} - Y^\varepsilon_{s-I})$$

is left-continuous and, consequently, locally bounded. From this property of local boundedness and the inequality $0 \le A^\varepsilon_s \le \varepsilon$ it follows that the process

$$\mathscr{N}^\varepsilon_t = \int_0^t \{a_s(Y^\varepsilon_{s-} - Y^\varepsilon_{s-I}) + \hat{a}_s(Y^\varepsilon_{s-} - Y^\varepsilon_{s-I})X^\varepsilon_{s-}\}b_s(Y^\varepsilon_{s-} - Y^\varepsilon_{s-I})\, d[A^\varepsilon, M^\varepsilon]_s$$

is a locally square integrable martingale with $\mathscr{N}^\varepsilon_0 = 0$ and predictable quadratic variation [see (1.13), (1.15)]

$$\begin{aligned}\langle \mathscr{N}^\varepsilon\rangle_t &= \int_0^t \{a_s + \hat{a}_s X^\varepsilon_{s-}\}^2 b_s^2(\Delta A^\varepsilon_s)^2\, d\langle M^\varepsilon\rangle_s\\ &= \int_0^t \{a_s + \hat{a}_s X^\varepsilon_{s-}\}^2 b_s^2(\Delta A^\varepsilon_s)^2\, dA^\varepsilon_s.\end{aligned}$$

From the implication $\langle \mathscr{N}^\varepsilon\rangle_T \to_P 0 \Rightarrow \sup_{t\le T}|\mathscr{N}^\varepsilon_t| \to_P 0$ as $\varepsilon \to 0$ [see, e.g., Problem 1.9.2 in Liptser and Shiryayev (1989)] it follows that we need only to prove $\langle \mathscr{N}^\varepsilon\rangle_T \to_P 0$ as $\varepsilon \to 0$ or, in our situation,

$$\lim_{\varepsilon\to 0} P(\langle \mathscr{N}^\varepsilon\rangle_T \ge \gamma, \Gamma^\varepsilon) = 0 \tag{2.19}$$

for any fixed γ as well as r_1 and r_2 (see the definition of the set Γ^ε). On the other hand, from (2.17) it follows that on the set Γ^ε, we have $\langle \mathscr{N}^\varepsilon\rangle_T \le k\varepsilon^2 T$, and thus (2.19) holds. □

2.2. *Proof of Theorem* 1.2. The proof follows in a straightforward fashion [see, e.g., Problem 6.1.2 in Liptser and Shiryayev (1989)] from Lemma 2.3 and Lemma 2.4 below. We have:

LEMMA 2.3. *Let the process $\hat{X}^\varepsilon_t$ be defined by*

$$\begin{aligned}\hat{X}^\varepsilon_t &= x_0 + \int_0^t \{a_s(Y_{s-} - Y_{s-I}) + \hat{a}_s(Y_{s-} - Y_{s-I})\hat{X}^\varepsilon_{s-}\}\, dA^\varepsilon_s\\ &\quad + \int_0^t b_s(Y_{s-} - Y_{s-I})\, dM^\varepsilon_s,\end{aligned} \tag{2.20}$$

where Y_s is the solution of (1.22). *Then, under the assumptions of Theorem* 1.2,

$$\hat{X}^\varepsilon \to_L X. \tag{2.21}$$

PROOF. [In what follows we shall drop the arguments $(Y_{s-} - Y_{s-I})$ of the functions $a_s(Y_{s-} - Y_{s-I})$, $\hat{a}_s(Y_{s-} - Y_{s-I})$, $b_s(Y_{s-} - Y_{s-I})$.]

STEP 1. Since $|\hat{a}_s| \le L$, let ε_0 be such that $L\varepsilon_0 \le \frac{1}{2}$ and take $\varepsilon \le \varepsilon_0$. Put

$$B^\varepsilon_t = \int_0^t \hat{a}_s\, dA^\varepsilon_s, \tag{2.22}$$

and notice that $|\Delta B_t^\varepsilon| \le \frac{1}{2}$. Defining the Doleans-Dade exponential [see Doleans-Dade (1970)]

$$\mathscr{E}_t(B^\varepsilon) = e^{B_t^\varepsilon} \prod_{s \le t} (1 + \Delta B_s^\varepsilon) e^{-\Delta B_s^\varepsilon}, \tag{2.23}$$

let us first show that for any $t \le T$,

$$\sup_{t \le T} \left| \mathscr{E}_t(B^\varepsilon) - \exp\left[\int_0^t \hat{a}_s \, ds\right] \right| \to 0 \quad \text{for } \varepsilon \to 0. \tag{2.24}$$

For this purpose notice first that

$$\sup_{t \le T} \left| B_t^\varepsilon - \int_0^t \hat{a}_s \, ds \right| \to 0 \quad \text{for } \varepsilon \to 0. \tag{2.25}$$

In fact, write $|B_t^\varepsilon - \int_0^t \hat{a}_s \, ds| = |\int_0^t \hat{a}_s (dA_s^\varepsilon - ds)|$ and, taking (1.12) into account, use Lemma A.2. Next we have

$$\sup_{t \le T} \left| \prod_{s \le t} (1 + \Delta B_s^\varepsilon) e^{-\Delta B_s^\varepsilon} - 1 \right| \to 0 \quad \text{for } \varepsilon \to 0. \tag{2.26}$$

Indeed, using $|\Delta B_s^\varepsilon| \le \frac{1}{2}$, the fact that for $|x| \le \frac{1}{2}$ it follows $\ln(1 + x) - x \le x^2 \sum_{k=2}^\infty |x^{k-2}| \le x^2 \sum_{k=0}^\infty (\frac{1}{2})^k = 2x^2$, and that $\sum |\Delta B_s^\varepsilon| = \sum |\hat{a}_s| |\Delta A_s^\varepsilon| \le L \sum |\Delta A_s^\varepsilon|$, we obtain

$$\begin{aligned} &\sup_{t \le T} \left| \prod_{s \le t} (1 + \Delta B_s^\varepsilon) e^{-\Delta B_s^\varepsilon} - 1 \right| \\ &\quad = \sup_{t \le T} \left| \exp\left(\sum_{s \le t} \ln(1 + \Delta B_s^\varepsilon) - \Delta B_s^\varepsilon \right) - 1 \right| \\ &\quad \le \left(\exp\left[2 \sum_{s \le T} (\Delta B_s^\varepsilon)^2 \right] - 1 \right) \le \left(\exp\left[2L \sup_{s \le T} |\Delta B_s^\varepsilon| A_T^\varepsilon \right] - 1 \right) \\ &\quad \le \left(\exp\left[2L^2 \varepsilon A_T^\varepsilon \right] - 1 \right) \to 0 \quad \text{for } \varepsilon \to 0. \end{aligned}$$

Relation (2.24) now follows from (2.25) and (2.26). Similarly we obtain

$$\sup_{t \le T} \left| \int_0^t \mathscr{E}_s^{-1}(B^\varepsilon) a_s \, dA_s^\varepsilon - \int_0^t \exp\left[-\int_0^s \hat{a}_u \, du \right] a_s \, ds \right| \to 0 \quad \text{for } \varepsilon \to 0. \tag{2.27}$$

Step 2. From the definition of $\mathscr{E}_t(B^\varepsilon)$ [see (2.23)], it follows that $\mathscr{E}_t(B^\varepsilon) = \prod_{s \le t}(1 + \Delta B_s^\varepsilon)$ and consequently

$$\mathscr{E}_t^{-1}(B^\varepsilon) = \prod_{s \le t} \left(1 - \frac{\Delta B_s^\varepsilon}{1 + \Delta B_s^\varepsilon} \right).$$

[We suppose that $\varepsilon \le \varepsilon_0$ and so (see Step 1), $|\Delta B_s^\varepsilon| \le \frac{1}{2}$.] Thus

$$\begin{aligned} \mathscr{E}_t^{-1}(B^\varepsilon) &\le \exp\left(\sum_{s \le t} \ln\left(1 + \frac{|\Delta B_s^\varepsilon|}{1 - |\Delta B_s^\varepsilon|} \right) \right) \\ &\le \exp\left(2 \int_0^t |\hat{a}_s| \, dA_s^\varepsilon \right) \le \exp(2LT). \end{aligned}$$

Using this estimate and the continuity property (in s) of the deterministic function $b_s = b_s(Y_{s-} - Y_{s-I})$, we now consider the square integrable martingale $\mathscr{N}^\varepsilon = (\mathscr{N}_t^\varepsilon)_{t\ge 0}$, where

$$\mathscr{N}_t^\varepsilon = \int_0^t \mathscr{E}_s^{-1}(B^\varepsilon) b_s \, dM_s^\varepsilon, \tag{2.28}$$

$$\langle \mathscr{N}^\varepsilon \rangle_t = \int_0^t \left(\mathscr{E}_s^{-1}(B^\varepsilon) b_s\right)^2 d\langle M^\varepsilon \rangle_s = \int_0^t \mathscr{E}_s^{-2}(B^\varepsilon) b_s^2 \, dA_s^\varepsilon, \tag{2.29}$$

and show that

$$\mathscr{N}^\varepsilon \to_L \mathscr{N} \quad \text{for } \varepsilon \to 0, \tag{2.30}$$

where

$$\mathscr{N}_t = \int_0^t \exp\left[-\int_0^s \hat{a}_u \, du\right] b_s \, dW_s,$$

with $W = (W_t)_{t\ge 0}$ a Wiener process. For this purpose it is sufficient to show [see Theorem 5.5.6 in Liptser and Shiryayev (1989)] that

$$\langle \mathscr{N}^\varepsilon \rangle_t \to_P \int_0^t \exp\left[-2\int_0^s \hat{a}_u \, du\right] b_s^2 \, ds \quad \text{for all } t > 0 \tag{2.31}$$

and that

$$E \sum_{s\le t} (\Delta \mathscr{N}_s^\varepsilon)^2 I(|\Delta \mathscr{N}_s^\varepsilon| > \delta) \to 0 \quad \text{with } \varepsilon \to 0 \text{ for all } \delta > 0 \text{ and } t > 0. \tag{2.32}$$

In order to establish (2.31), notice that for $t \le T$ with $T > 0$,

$$\begin{aligned} &\left| \langle \mathscr{N}^\varepsilon \rangle_t - \int_0^t \exp\left[-2\int_0^s \hat{a}_u \, du\right] b_s^2 \, ds \right| \\ &\quad \le \int_0^T \left| \mathscr{E}_s^{-2}(B^\varepsilon) - \exp\left[-2\int_0^s \hat{a}_u \, du\right] \right| b_s^2 \, dA_s^\varepsilon \\ &\qquad + \sup_{t\le T} \left| \int_0^t \exp\left[-2\int_0^s \hat{a}_u \, du\right] b_s^2 \, d(A_s^\varepsilon - s) \right|. \end{aligned} \tag{2.33}$$

Using Assumption 2 and the fact that $A_T^\varepsilon \le T$, the first term on the right in (2.33) is bounded above by

$$\sup_{s\le T} \left| \mathscr{E}_s^{-2}(B^\varepsilon) - \exp\left[-2\int_0^s \hat{a}_u \, du\right] \right| L(1 + Y_T) T$$

and this bound tends to zero in probability [by (2.24) and the fact that Y_T is a given deterministic quantity]. The second term on the right in (2.33) tends to zero in probability by Lemma A.2 with $G_t^\varepsilon = A_t^\varepsilon$ and $G_t = t$.

Let us now establish (2.32). For this purpose notice that by Assumption 2 and the fact that Y_t is deterministic we have $\sup_{s\le t}|b_s| \le 1 + \sup_{s\le t} b_s^2 \le$

$(L+1) + LY_t < \infty$. Furthermore,

$$\sup_{s \le t} |\mathscr{E}_s^{-1}(B^\varepsilon)| \le \exp\left[\int_0^T |\hat{a}_u|\, du\right] + \sup_{s \le T}\left|\mathscr{E}_s^{-1}(B^\varepsilon) - \exp\left[-\int_0^s \hat{a}_u\, du\right]\right|,$$

where by (2.24) the second term on the right tends to zero for $\varepsilon \to 0$. From (2.28) we then have for small ε and $s \le T$,

$$|\Delta \mathscr{N}_s^\varepsilon| = \mathscr{E}_s^{-1}(B^\varepsilon)|b_s|\,|\Delta M_s^\varepsilon| \le K|\Delta M_s^\varepsilon|,$$

where $K = \exp(2LT)\sup_{s \le T}|b_s|$. Consequently, (2.32) holds if for $\varepsilon \to 0$,

$$E\sum_{s \le t} (\Delta M_s^\varepsilon)^2 I(|\Delta M_s^\varepsilon| > \delta) \to 0 \quad \text{for all } \delta > 0 \text{ and } t > 0. \tag{2.34}$$

But

$$\begin{aligned} E\sum_{s \le t} (\Delta M_s^\varepsilon)^2 I(|\Delta M_s^\varepsilon| > \delta) &= E\sum_{k=1}^{[t/\varepsilon]} \varepsilon \xi_k^2 I\left(|\xi_k| > \frac{\delta}{\sqrt{\varepsilon}}\right) \\ &= \varepsilon\left[\frac{t}{\varepsilon}\right] E\xi_1^2 I\left(|\xi_1| > \frac{\delta}{\sqrt{\varepsilon}}\right) \\ &\le tE\xi_1^2 I\left(|\xi_1| > \frac{\delta}{\sqrt{\varepsilon}}\right) \to 0 \quad \text{for } \varepsilon \to 0. \end{aligned}$$

STEP 3. Rewrite (2.20) in the form

$$\hat{X}_t^\varepsilon = x_0 + \int_0^t \hat{a}_s(Y_{s-} - Y_{s-I})\hat{X}_{s-}^\varepsilon\, dA_s^\varepsilon + S_t^\varepsilon, \tag{2.35}$$

where

$$S_t^\varepsilon = \int_0^t b_s(Y_{s-} - Y_{s-I})\, dM_s^\varepsilon + \int_0^t a_s(Y_{s-} - Y_{s-I})\, dA_s^\varepsilon \tag{2.36}$$

is a semimartingale. For $\varepsilon > 0$ sufficiently small ($\varepsilon < L^{-1}$) and using the Doleans–Dade exponential (2.23), we then obtain

$$\hat{X}_t^\varepsilon = \mathscr{E}_t(B^\varepsilon)\left(x_0 + \int_0^t \mathscr{E}_s^{-1}(B^\varepsilon) a_s\, dA_s^\varepsilon + \mathscr{N}_t^\varepsilon\right), \tag{2.37}$$

where B_t^ε is as in (2.22) and $\mathscr{N}_t^\varepsilon$ as defined in (2.28). From (2.24), (2.27) and (2.30) it then follows that

$$\hat{X}^\varepsilon \to_L \hat{X}, \tag{2.38}$$

where

$$\begin{aligned} \hat{X}_t = \exp\left[\int_0^s \hat{a}_u\, du\right]\bigg(x_0 &+ \int_0^t \exp\left[-\int_0^s \hat{a}_u\, du\right] a_s\, ds \\ &+ \int_0^t \exp\left[-\int_0^s \hat{a}_u\, du\right] b_s\, dW_s\bigg) \end{aligned} \tag{2.39}$$

which is the unique solution to the linear Itô equation

$$\hat{X}_t = x_0 + \int_0^t \{a_s + \hat{a}_s \hat{X}_s\}\, ds + \int_0^t b_s\, dW_s \tag{2.40}$$

so that [see (1.23)] $\hat{X} = X$ and thus $\hat{X}^\varepsilon \to_L X$. □

LEMMA 2.4. *Under the assumptions of Theorem* 1.2 *we have*

$$\sup_{t \le T} |X_t^\varepsilon - \hat{X}_t^\varepsilon| \to_P 0 \quad \text{for } \varepsilon \to 0 \text{ and all } T > 0, \tag{2.41}$$

where $\hat{X}_t^\varepsilon$ *is as defined in* (2.20).

PROOF. From (2.20) and (1.19), we have

$$\begin{aligned} X_t^\varepsilon - \hat{X}_t^\varepsilon &= \int_0^t [a_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) - a_s(Y_{s-} - Y_{s-I})]\, dA_s^\varepsilon \\ &\quad + \int_0^t \hat{a}_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon)(X_{s-}^\varepsilon - \hat{X}_s^\varepsilon)\, dA_s^\varepsilon \\ &\quad + \int_0^t [\hat{a}_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) - \hat{a}_s(Y_{s-} - Y_{s-I})] \hat{X}_{s-}^\varepsilon\, dA_s^\varepsilon \\ &\quad + \int_0^t [b_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) - b_s(Y_{s-} - Y_{s-I})]\, dM_s^\varepsilon. \end{aligned} \tag{2.42}$$

Define

$$\begin{aligned} \gamma^\varepsilon &= \int_0^T |a_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) - a_s(Y_{s-} - Y_{s-I})|\, dA_s^\varepsilon \\ &\quad + \int_0^T |\hat{a}_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) - \hat{a}_s(Y_{s-} - Y_{s-I})|\, |\hat{X}_{s-}^\varepsilon|\, dA_s^\varepsilon \\ &\quad + \sup_{0 \le t \le T} \left| \int_0^t [b_s(Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon) - b_s(Y_{s-} - Y_{s-I})]\, dM_s^\varepsilon \right|, \end{aligned}$$

then from (2.43), taking into account Assumption 1, it follows that for all $t \le T$,

$$\sup_{u \le t} |X_u^\varepsilon - \hat{X}_u^\varepsilon| \le \gamma^\varepsilon + L \int_0^t \sup_{u \le s} |X_{u-}^\varepsilon - \hat{X}_{u-}^\varepsilon|\, dA_s^\varepsilon.$$

By Theorem 2.4.3 in Liptser and Shiryayev (1989), we then have

$$\sup_{u \le T} |X_u^\varepsilon - \hat{X}_u^\varepsilon| \le \gamma^\varepsilon \exp[L A_T^\varepsilon].$$

Since $A_T^\varepsilon \le T$, it is therefore enough to show that $\gamma^\varepsilon \to_P 0$ for $\varepsilon \to 0$. But $\gamma^\varepsilon = \gamma_1^\varepsilon + \gamma_2^\varepsilon + \gamma_3^\varepsilon$ and so it remains to show that

$$\gamma_i^\varepsilon \to_P 0 \quad \text{for } \varepsilon \to 0 \text{ and } i = 1, 2, 3. \tag{2.43}$$

By Theorem 1.1 and the continuity of $a_s(y)$ (Assumption 2) we have

$$\sup_{s \le T} |a_s(Y^\varepsilon_{s-} - Y^\varepsilon_{s-I}) - a_s(Y_{s-} - Y_{s-I})| \to_P 0,$$

so that (2.43) holds for $i = 1$.

Again by Theorem 1.1 and the continuity this time of $\hat{a}_s(y)$ (Assumption 1), it follows that

$$\gamma^\varepsilon_2 I\left(\sup_{s \le T} |\hat{X}^\varepsilon_s| \le r\right) \to_P 0 \quad \text{for all } r > 0.$$

On the other hand, by Lemma 2.3 we have $\hat{X}^\varepsilon \to_L X$ so that

$$\lim_{r \to \infty} \limsup_{\varepsilon \to 0} P\left\{\sup_{s \le T} |\hat{X}^\varepsilon_s| > r\right\} = 0,$$

thus (2.43) holds also for $i = 2$.

To obtain (2.43) for $i = 3$, using Problem 1.9.2 in Liptser and Shiryayev (1989), it suffices to show that

$$\int_0^T [b_s(Y^\varepsilon_{s-} - Y^\varepsilon_{s-I}) - b_s(Y_s - Y_{s-I})]^2 \, d\langle M^\varepsilon \rangle_s \to_P 0 \quad \text{for } \varepsilon \to 0.$$

Since $\langle M^\varepsilon \rangle_t = A^\varepsilon_t$ and $A^\varepsilon_T \le T$, this latter property is obtained by complete analogy to the case $i = 1$, taking into account the continuity of b_s (Assumption 2). □

3. Application to option pricing.

3.1. *A past-dependent option pricing model.* In this section we introduce an option pricing model for European call options where the distribution of future asset prices depends on past evolution. We shall also present an option valuation formula for such a situation. For simplicity we shall consider only options on a single risky asset assuming furthermore that there are no dividends and that the interest rate is constant over time. With slightly more complicated derivations, our methods can, however, also handle the more general cases.

Given a probability space $(\Omega, \mathscr{F}, P)$ with a Wiener process $W = (W_t)$ and a filtration $\mathscr{F}_t := \sigma\{W_s, s \le t\}$, in our model we let the price S_t of the risky asset evolve according to the Itô equation

$$dS_t = S_t(\mu_t \, dt + \sigma_t \, dW_t). \tag{3.1}$$

The initial condition S_0 is given, deterministic and positive, while μ_t and σ_t are bounded and $(\mathscr{F}_t)$-adapted processes with $\sigma_t \ge \bar{\sigma} > 0$. In addition to the risky asset let there be given a nonrisky asset (bond) whose price S^0_t evolves according to

$$dS^0_t = \alpha S^0_t \, dt, \qquad S^0_0 = 1, \tag{3.2}$$

with α, the interest rate, being a given constant.

Due to the practical impossibility of a continuous-time asset price observation, we assume that the values of S_t and S_t^0 are observed only at discrete time points t_k satisfying (1.1) for a given $\varepsilon > 0$. More specifically, we shall assume that the coefficients are adapted in the sense that they depend only on the discrete-time past observations of S_{t_k}, $t_k \le t$. Since we may vary the length ε of the observation interval $t_{k+1} - t_k$ in (1.1), in what follows we shall use the notation μ_t^ε and σ_t^ε instead of simply μ_t and σ_t and consider instead of a fixed price process (S_t, S_t^0) given by (3.1), (3.2), the family of processes $(S_t^\varepsilon, S_t^{\varepsilon 0})_{\varepsilon > 0}$ satisfying

$$dS_t^\varepsilon = S_t^\varepsilon(\mu_t^\varepsilon\, dt + \sigma_t^\varepsilon\, dW_t), \tag{3.3}$$

$$dS_t^{\varepsilon 0} = \alpha S_t^{\varepsilon 0}\, dt, \qquad S_0^{\varepsilon 0} = 1, \tag{3.4}$$

where, for all $\varepsilon > 0$, $S_0^\varepsilon = S_0$, $\sigma_t^\varepsilon \ge \bar{\sigma} > 0$, and α is the same constant interest rate as in (3.2).

To complete the description of the option pricing problem we have to specify the claim that will be represented by a random payout $h^\varepsilon(T) \ge 0$, adapted to $\mathscr{F}_T$, at a fixed maturity date T. Being interested in European call options, we shall take

$$h^\varepsilon(T) = \max[0, S_T^\varepsilon \wedge C - K] \tag{3.5}$$

with $K > 0$ denoting the strike (exercise) price and C a sufficiently large positive constant. We use the minimum $S_T^\varepsilon \wedge C$ between S_T^ε and C instead of simply S_T^ε, only for technical reasons, such as to guarantee the existence of the expectation of $h^\varepsilon(T)$, and others that will become apparent below. From a practical point of view, however, this will hardly be any restriction.

As mentioned in the Introduction, in the past-dependent setting described here, there exists a perfect hedge portfolio and a closed-form option valuation formula. For this we refer to the literature cited in Section 1.3, recalling here only the following facts:

1. For each $\varepsilon > 0$, there exists a unique martingale measure Q^ε (risk-neutral probability measure), equivalent to P, such that under Q^ε the price S_t^ε of the risky asset evolves according to

$$dS_t^\varepsilon = \alpha S_t^\varepsilon\, dt + \sigma_t^\varepsilon\, d\tilde{W}_t^\varepsilon \tag{3.6}$$

   with the same initial condition S_0^ε as in (3.3) and where $\tilde{W}_t^\varepsilon$ is a Wiener process under Q^ε, while α is the same interest rate as in (3.4).
2. There exists one and only one option valuation function u_t^ε that can be expressed as

$$u_t^\varepsilon = E^\varepsilon\{h^\varepsilon(T)e^{-\alpha(T-t)} | \mathscr{F}_t\}, \tag{3.7}$$

   where E^ε denotes expectation with respect to Q^ε.

We remark here that the fact that (3.6) represents the price evolution in a risk-neutral world can be seen from the relation (recall that α denotes the interest rate)

$$E^{\varepsilon}\{S_t^{\varepsilon}|S_0^{\varepsilon}\} = S_0^{\varepsilon}e^{\alpha t} \tag{3.8}$$

which is an immediate consequence of (3.6).

We shall now reexpress (3.6) in a more suitable form. Defining

$$X_t^{\varepsilon} := \ln S_t^{\varepsilon} \tag{3.9}$$

and using Itô's rule, we find that under Q^{ε} the process X_t^{ε} satisfies

$$dX_t^{\varepsilon} = \left(\alpha - \tfrac{1}{2}(\sigma_t^{\varepsilon})^2\right)dt + \sigma_t^{\varepsilon}\, d\tilde{W}_t^{\varepsilon}, \qquad X_0^{\varepsilon} = \ln S_0^{\varepsilon}. \tag{3.10}$$

Recall that, by $S_0^{\varepsilon} = S_0$ for all $\varepsilon > 0$, we also have $X_0^{\varepsilon} = X_0$ for all $\varepsilon > 0$ with X_0 deterministic and known. Since the instantaneous asset price variance $(\sigma_t^{\varepsilon})^2$ was assumed to depend only on $S_{t_k}^{\varepsilon}$ for $t_k \le t$, it is constant on each of the intervals $[t_k, t_{k+1})$. Therefore, at the points t_k the solution to (3.10) has the same distribution as the discrete time process $X_{t_k}^{\varepsilon}$ defined by

$$X_{t_{k+1}}^{\varepsilon} = X_{t_k}^{\varepsilon} + \left(\alpha - \tfrac{1}{2}\left(\sigma_{t_k}^{\varepsilon}\right)^2\right)\varepsilon + \sigma_{t_k}^{\varepsilon}\varepsilon^{1/2}z_k, \tag{3.11}$$

where (z_k) is an i.i.d. sequence of standard Gaussian random variables.

We shall now introduce an explicit form for the dependence of the instantaneous asset price variance $(\sigma_t^{\varepsilon})^2$ on $S_{t_k}^{\varepsilon}$ for $t_k \le t$. We shall in fact assume that $(\sigma_{t_k}^{\varepsilon})^2$ is, for each t_k, a weighted sum of an exogeneously given constant σ^2 and an unweighted sample variance of the logarithms of the asset prices over a past interval of length I, that is,

$$\left(\sigma_{t_k}^{\varepsilon}\right)^2 = \delta\sigma^2 + (1-\delta)\left(\hat{\sigma}_{t_k}^{\varepsilon}\right)^2, \qquad \delta \in (0,1), \tag{3.12}$$

where

$$\left(\hat{\sigma}_{t_k}^{\varepsilon}\right)^2 = I^{-1}\sum_{h=k-[I/\varepsilon]+1}^{k}\left(X_{t_h}^{\varepsilon} - X_{t_{h-1}}^{\varepsilon}\right)^2. \tag{3.13}$$

Notice now that model (3.11) together with (3.12) and (3.13) is a particular case of model (1.3) with (1.2) and therefore also of model (1.19) with (1.20), where

$$\begin{gathered} b_s^2(Y_{s-}^{\varepsilon} - Y_{s-I}^{\varepsilon}) = \delta\sigma^2 + \rho Y_{s-}^{\varepsilon} - \rho Y_{s-I}^{\varepsilon}, \\ \delta \in (0,1), \qquad \rho = \frac{1-\delta}{I}, \end{gathered} \tag{3.14}$$

$$\hat{a}_s(Y_{s-}^{\varepsilon} - Y_{s-I}^{\varepsilon}) = 0, \tag{3.15}$$

$$\begin{aligned} a_s(Y_{s-}^{\varepsilon} - Y_{s-I}^{\varepsilon}) &= \alpha - \frac{1}{2}b_s^2(Y_{s-}^{\varepsilon} - Y_{s-I}^{\varepsilon}) \\ &= \alpha - \frac{\delta}{2}\sigma^2 - \frac{\rho}{2}Y_{s-}^{\varepsilon} + \frac{\rho}{2}Y_{s-I}^{\varepsilon}, \end{aligned} \tag{3.16}$$

so that Assumptions 1 and 2 can easily be seen to be satisfied and all results obtained previously carry over to our option pricing model.

REMARK 3.1. In the description of our model we imposed the requirement that σ_t^ε be uniformly bounded, an assumption that allows the application of results in the literature concerning the existence of a perfect hedge portfolio and of a closed-form option valuation formula. Since $\sigma_t^\varepsilon = b_t(Y_{t-}^\varepsilon - Y_{t-I}^\varepsilon)$, by (3.14) this assumption would be satisfied if we knew that the process Y_t^ε was uniformly bounded almost surely. Since by our results the processes (Y_t^ε) converge uniformly to a continuous and deterministic process (Y_t) only in probability, we cannot guarantee the almost sure uniform boundedness of Y_t^ε, even for small ε and finite t. We now show that we may slightly modify our problem so as to satisfy the stronger requirement and still retain the main result of the paper. Letting $C > \sup_{t \le T} Y_t$, define two processes $\overline{X}_t^\varepsilon$ and $\overline{Y}_t^\varepsilon$ as follows: $\overline{Y}_t^\varepsilon := Y_t^\varepsilon \wedge C$, while $\overline{X}_t^\varepsilon$ is given by (1.19) with $\overline{Y}_t^\varepsilon$ replacing Y_t^ε there. On the one hand we then have $\bar{\sigma}_t^\varepsilon = b_t(\overline{Y}_{t-}^\varepsilon - \overline{Y}_{t-I}^\varepsilon)$ uniformly bounded, while on the other it is immediately seen that the results of Theorems 1.1 and 1.2 also hold for $\overline{Y}_t^\varepsilon$ and $\overline{X}_t^\varepsilon$, respectively (while the former holds by definition, the latter becomes straightforward by noticing that Lemma 2.4 holds with the same proof also for $\overline{X}_t^\varepsilon$ instead of X_t^ε).

REMARK 3.2. If we take for σ_t^ε a deterministic time function, formula (3.7) coincides with the Black and Scholes formula. In this case it follows in fact from (3.6) that, under Q^ε, the price S_T^ε has a lognormal distribution, that is, for $X_T^\varepsilon = \ln S_T^\varepsilon$, we have from (3.10):

$$X_T^\varepsilon \sim \mathscr{N}(m_T^\varepsilon, V_T^\varepsilon), \tag{3.17}$$

where the symbol $\sim$ here means distributed according to, $\mathscr{N}(m, \sigma^2)$ denotes the normal distribution with mean m and variance σ^2 and

$$m_T^\varepsilon = x_0 + \int_0^T \left(\alpha - \tfrac{1}{2}(\sigma_t^\varepsilon)^2\right) dt, \tag{3.18}$$

$$V_T^\varepsilon = \int_0^T (\sigma_t^\varepsilon)^2 \, dt. \tag{3.19}$$

With the claim given by (3.5) in the more customary form $h^\varepsilon(T) = \max[0, S_T^\varepsilon - K]$, one can then compute exactly the expectation in (3.7) obtaining for $t = 0$ rather straightforwardly [see also Remark 5.7 in Karatzas (1989)]:

$$\begin{aligned} u_0^\varepsilon = e^{X_0} N_{CF}\left((m_T^\varepsilon + V_T^\varepsilon - \ln K)(V_T^\varepsilon)^{-(1/2)}\right) \\ - Ke^{-\alpha T} N_{CF}\left((m_T^\varepsilon - \ln K)(V_T^\varepsilon)^{-(1/2)}\right), \end{aligned} \tag{3.20}$$

where

$$N_{CF}(x) = \int_{-\infty}^x (2\pi)^{-(1/2)} \exp\left(-\frac{\xi^2}{2}\right) d\xi. \tag{3.21}$$

Now (3.20) with (3.18) and (3.19) is indeed the classical Black and Scholes formula.

Notice now that, although (3.7) holds for very general price processes, its actual computation is in general prohibitive; an exception is the lognormal case mentioned in Remark 3.2. Notice, furthermore, that the proofs for such a formula, as given in the literature, are in general not constructive, being based on martingale representation results. To compute the option value in a more general situation like the past-dependent case considered here, one needs an approximation which will be the subject of the next subsection.

REMARK 3.3. Notice that the present paper differs in various respects from recent work concerning the question of continuous-versus discrete-time trading. In particular, we do not deal with the question of the validity of assuming a continuous trading model as in Denny and Suchanek (1986), nor do we address the problem of approximating an idealized continuous securities market model by a sequence of discrete-time trading models such that the latter converge and certain properties are preserved along the approximating sequence as in Willinger and Taqqu (1989) [see also Duffie and Protter (1988) as well as He (1989)]. Our models are formally all continuous trading models, but, on the one hand we use them only as convenient representations for our specific discrete-observation situation [in fact, with the coefficients in (3.6), (3.10) piecewise constant, at the discrete observation points t_k, these models are equivalent—in distribution—to the discrete time models (3.11)]. On the other hand, our limiting continuous observation model [(3.26) below] is used only as a computational tool to obtain an approximate solution to the given discrete-observation and past-dependent problems, for which the solution cannot otherwise be computed in practice. Furthermore, each of the (formally continuous) models (3.3), used as convenient representations for the more realistic situation when prices are observed in discrete time, is complete and admits a unique equivalent martingale measure as well as a perfect hedge portfolio. This has some similarity to the skeleton approach in Willinger and Taqqu (1989), but the general methodology as well as the problem itself are here rather different. Notice also that we do not deal with questions of convergence of replicating portfolio strategies, which are considered, for example, in He (1989) for a more traditional securities market model under rather strong regularity assumptions; our interest concentrates on the value of the option. Notice finally that, although we obtain a different martingale measure Q^{ε} for each discrete observation interval ε, it is possible to represent all our processes (or at least copies thereof), including the limiting continuous observation process (3.26) below, on the same probability space; this is again similar to the situation in Willinger and Taqqu (1989), where all elements in a finite market approximation are defined on the same probability space and differ from each other by the sets of trading dates, equilibrium price processes and information structures.

3.2. *An asymptotic option valuation formula.* Consider now the following option pricing problem: The logarithm $X^{\varepsilon}_{t_k}$ of the price of the risky asset evolves (under the risk-neutral measure Q^{ε}) over the discrete time observation points t_k according to (3.11)–(3.13); the claim is given by (3.5) with $C > 0$ sufficiently large. The purpose is to determine (see Theorem 3.1. below) an asymptotic (for $\varepsilon \to 0$ and $C \to \infty$) valuation formula for the given claim at time $t = 0$ assuming that the evolution of $X^{\varepsilon}_{t_k}$ before and at $t = 0$ is known. Once this purpose is accomplished, it then follows that, if the value of ε corresponding to the given process is small, that is, if the time intervals between the discrete asset price observations are small (and C is large), then the asymptotic valuation formula yields an approximation to the actual option value which is better the smaller the value of $\varepsilon > 0$.

We start by recalling from the previous subsection that the given discrete time process $X^{\varepsilon}_{t_k}$ has, for $t = t_k$, the same distributions as the process X^{ε}_t defined in (1.19) with Y^{ε}_t as in (1.20), when the coefficients are given by (3.14)–(3.16). Recall also [see (1.2)] that the initial segment Y^{ε}_u, $-I < u \le 0$ for the delay equation (1.20) is given by

$$Y^{\varepsilon}_u = \sum_{k:\ -I < t_k \le u} \left(X^{\varepsilon}_{t_k} - X^{\varepsilon}_{t_{k-1}} \right)^2, \qquad -I < u \le 0. \tag{3.22}$$

Defining

$$Z^{\varepsilon}_t := Y^{\varepsilon}_t - Y^{\varepsilon}_{t-I} \tag{3.23}$$

we may then consider $X^{\varepsilon}_{t_k}$ as being equivalently obtained, for $t = t_k$, from

$$X^{\varepsilon}_t = x_0 + \int_0^t \left(\alpha - \frac{\delta}{2}\sigma^2 - \frac{\rho}{2} Z^{\varepsilon}_{s-} \right) dA^{\varepsilon}_s + \int_0^t \left(\delta\sigma^2 + \rho Z^{\varepsilon}_{s-} \right)^{1/2} dM^{\varepsilon}_s. \tag{3.24}$$

Recall furthermore from Section 2 that, if the initial segment Y^{s}_u, $-I < u \le 0$, for (1.20) converges uniformly (in u) in probability to the initial segment $\tilde{Y}_u$ of the process Y_t defined by the delay equation (1.21), then (X^{ε}_t) converges in distribution to the process (X_t) defined by (1.23) with Y_t as in (1.22) and where the coefficients are again given by (3.14)–(3.16). More precisely, letting

$$Z_t := Y_t - Y_{t-I}, \qquad Y_t = \tilde{Y}_t \quad \text{for } -I < t \le 0, \tag{3.25}$$

such X_t, Y_t are therefore given by

$$X_t = x_0 + \int_0^t \left(\alpha - \frac{\delta}{2}\sigma^2 - \frac{\rho}{2} Z_s \right) ds + \int_0^t \left(\delta\sigma^2 + \rho Z_s \right)^{1/2} dW_s, \tag{3.26}$$

$$Y_t = \tilde{Y}_0 + \delta\sigma^2 t + \int_0^t \rho Z_s \, ds. \tag{3.27}$$

Notice that, given x_0 and $Z = (Z_t)$, the random variable X_T is conditionally Gaussian with mean $m_T(x_0, Z)$ and variance $V_T(Z)$ expressed by

$$m_T(x_0, Z) = x_0 + \int_0^T \left(\alpha - \frac{\delta}{2}\sigma^2 - \frac{\rho}{2} Z_s \right) ds, \tag{3.28}$$

$$V_T(Z) = \int_0^T \left(\delta\sigma^2 + \rho Z_s \right) ds. \tag{3.29}$$

Consequently, if $f = f(x)$ is a continuous function such that $|f(x)| \le \exp(\alpha x)$

for some constant $\alpha > 0$, then

$$E\{f(X_T)|x_0, Z\} = (2\pi)^{-1/2}\int_R f(V_T(Z)y + m_T(x_0, Z)) \times\exp\left(-\frac{y^2}{2}\right) dy = H(m_T(x_0, Z), V_T(Z)), \tag{3.30}$$

with $H = H(m, V)$ a function which is easily seen to be continuous by Lebesgue's dominated convergence theorem.

We next show how to use our main convergence result for (X_t^ε), (Y_t^ε), as well as (3.30), for the approximate computation of the value of our claim at $t = 0$, namely of [see (3.7)]

$$E^\varepsilon\{\max[0, \exp(X_T^\varepsilon \wedge C) - K]e^{-\alpha T}|x_0, Y_u^\varepsilon, -I < u \le 0\}$$

when ε is small and C is large.

Using the notation $\alpha^\varepsilon \to_{Q^\varepsilon} \alpha$ with the meaning that

$$\lim_{\varepsilon\to 0} Q^\varepsilon\{|\alpha^\varepsilon - \alpha| > \delta\} = 0$$

for all $\delta > 0$, we first have the following.

PROPOSITION 3.1. *For any continuous and bounded function $f = f(x)$ we have for $\varepsilon \to 0$,*

$$E^\varepsilon\{f(X_T^\varepsilon)|x_0, Y_u^\varepsilon, -I < u \le 0\} \to_{Q^\varepsilon} E\{f(X_T)|x_0, Z\}.$$

PROOF. Consider the process $(\hat{X}_t^\varepsilon)$, defined in the proof of Theorem 1.2 [formula (2.20)], which in our particular case (and under the measure Q^ε) is given by

$$\hat{X}_t^\varepsilon = x_0 + \int_0^t\left(\alpha - \frac{\delta}{2}\sigma^2 - \frac{\rho}{2}Z_s\right) dA_s^\varepsilon + \int_0^t(\delta\sigma^2 - \rho Z_s)\, dM_s^\varepsilon. \tag{3.31}$$

As was established in the proof of Theorem 1.2, the process $(\hat{X}_t^\varepsilon)$ has the following properties for $\varepsilon \to 0$ (see Lemma 2.4 and 2.3, respectively):

$$\sup_{t\le T}|X_t^\varepsilon - \hat{X}_t^\varepsilon| \to_{Q^\varepsilon} 0, \tag{3.32}$$

$$(\hat{X}_t^\varepsilon) \to_L (X_t). \tag{3.33}$$

From (3.32) it follows that

$$\begin{aligned} &E^\varepsilon|E^\varepsilon\{f(X_T^\varepsilon)|x_0, Y_u^\varepsilon, -I < u \le 0\} \\ &\qquad - E^\varepsilon\{f(\hat{X}_T^\varepsilon)|x_0, Y_u^\varepsilon, -I < u \le 0\}| \\ &\quad\le E^\varepsilon|f(X_T^\varepsilon) - f(\hat{X}_T^\varepsilon)| \to 0 \quad \text{as } \varepsilon \to 0. \end{aligned} \tag{3.34}$$

On the other hand, for each ε,

$$E^\varepsilon\{f(\hat{X}_T^\varepsilon)|x_0, Y_u^\varepsilon, -I < u \le 0\} = E^\varepsilon\{f(\hat{X}_T^\varepsilon)|x_0, Z\} \qquad P \text{ a.s.}$$

Furthermore, let $\hat{X}_t^{\varepsilon,0} = \hat{X}_t^{\varepsilon} - x_0$ and $X_t^0 = X_t - x_0$. From (3.31) and (3.26), it then follows that $(\hat{X}_t^{\varepsilon,0})$ and (X_t^0) are independent of x_0, and the convergence (3.33) implies $\hat{X}^{\varepsilon,0} \to_L X^0$. Therefore, for each constant $d \in R$, $E^{\varepsilon} f(\hat{X}_T^{\varepsilon,0} + d) \to Ef(X_T^0 + d)$ as $\varepsilon \to 0$. Since $E^{\varepsilon}\{f(\hat{X}_T^{\varepsilon,0} + x_0)|x_0\} = E^{\varepsilon} f(\hat{X}_T^{\varepsilon,0} + d)_{|d=x_0}$ and $E\{f(X_T^0 + x_0)|x_0\} = Ef(X_T^0 + d)_{|d=x_0}$, we then have

$$\lim_{\varepsilon \to 0} E^{\varepsilon}\left\{ f\left(\hat{X}_T^{\varepsilon,0} + x_0\right)|x_0\right\} = E\left\{ f\left(X_T^0 + x_0\right)|x_0\right\}.$$

The desired result now follows from this and (3.34), Z being deterministic. □

Let

$$f_C(x) = \max[0, \exp(x \wedge C) - K]e^{-\alpha T} \tag{3.35}$$

and let [see (3.30)] $H_C = H_C(m, V)$ correspond to f_C. By Proposition 3.1 we then have

$$\begin{aligned} &E^{\varepsilon}\left\{\max[0, \exp(X_T^{\varepsilon} \wedge C) - K]e^{-\alpha T}|x_0, Y_u^{\varepsilon}, -I < u \le 0\right\} \\ &\quad \to_{Q^{\varepsilon}} H_C(m_T(x_0, Z), V_T(Z)) \quad \text{as } \varepsilon \to 0. \end{aligned} \tag{3.36}$$

Consequently, if Z_s, $0 \le s \le T$, is known and $H_C(m, V)$ can be computed, we may compute $m_T(x_0, Z)$ and $V_T(Z)$ according to (3.28), (3.29) and take as an estimate for $E^{\varepsilon}\{\max[0, \exp(X_T^{\varepsilon} \wedge C) - K]e^{-\alpha T}|x_0, Y_u^{\varepsilon}, -I < u \le 0\}$ the value of $H_C(m_T(x_0, Z), V_T(Z))$. Concerning the computation of the function $H_C(m, V)$, notice that it may be approximated, for large C, by the function $H(m, V)$ corresponding to

$$f(x) = \max[0, \exp(x) - K]e^{-\alpha T}. \tag{3.37}$$

In fact, X_T being (conditionally) Gaussian, we have

$$\begin{aligned} &\lim_{C \to \infty} H_C(m_T(x_0, Z), V_T(Z)) \\ &\quad = \lim_{C \to \infty} E\left\{\max[0, \exp(X_T \wedge C) - K]e^{-\alpha T}|x_0, Z\right\} \\ &\quad = E\left\{\max[0, \exp(X_T) - K]e^{-\alpha T}|x_0, Z\right\} = H(m_T(x_0, Z), V_T(Z)) \end{aligned} \tag{3.38}$$

with the expectation of the right being finite and explicitly computable via a Black and Scholes formula according to Remark 3.2 [formula (3.20) with m_T^{ε}, V_T^{ε} replaced by $m_T(x_0, Z), V_T(Z)$].

On the other hand, $Z_s = Y_s - Y_{s-I}$ has to be computed using (3.27) with the unobservable initial segment $\tilde{Y}_u$, $-I < u \le 0$. We do observe, however, the random function Y_u^{ε}, $-I < u \le 0$, which is supposed to be close, for small ε, to $\tilde{Y}_u$ in the sense specified in (1.21). Therefore, defining the random function

$$\hat{Z}_t^{\varepsilon} = \hat{Y}_t^{\varepsilon} - \hat{Y}_{t-I}^{\varepsilon}, \tag{3.39}$$

where $\hat{Y}_t^{\varepsilon}$ is obtained, by analogy with (3.27), from

$$\hat{Y}_t^{\varepsilon} = Y_0^{\varepsilon} + \delta\sigma^2 t + \int_0^t \rho \hat{Z}_s^{\varepsilon}\, ds, \qquad \hat{Y}_t^{\varepsilon} = Y_t^{\varepsilon} \quad \text{for } -I < t \le 0, \tag{3.40}$$

it follows that $\sup_{s \le T} |\hat{Z}_s^\varepsilon - Z_s| \to_{Q^\varepsilon} 0$ as $\varepsilon \to 0$ and, consequently, for $\varepsilon \to 0$,

$$H\big(m_T(x_0, \hat{Z}^\varepsilon), V_T(\hat{Z}^\varepsilon)\big) \to_{Q^\varepsilon} H(m_T(x_0, Z), V_T(Z)), \tag{3.41}$$

where $m_T(x_0, \hat{Z}^\varepsilon)$ and $V_T(\hat{Z}^\varepsilon)$ are obtained from formulae (3.28) and (3.29), replacing Z_s by $\hat{Z}_s^\varepsilon$.

The convergence results (3.36), (3.38) and (3.41) now lead to the following.

THEOREM 3.1. *The exact value of the claim at* $t = 0$, *namely the value of* $E^\varepsilon\{\max[0, \exp(X_T^\varepsilon \wedge C) - K]e^{-\alpha T} | x_0, Y_u^\varepsilon, -I < u \le 0\}$ *can be arbitrarily closely approximated by the value of* $H(m_T(x_0, \hat{Z}^\varepsilon), V_T(\hat{Z}^\varepsilon))$, *provided* ε *is sufficiently small and* C *sufficiently large.*

To conclude, notice that the result of Theorem 3.1 continues to hold, if instead of $(X_t^\varepsilon, Y_t^\varepsilon)$ we consider the pair $(\overline{X}_t^\varepsilon, \overline{Y}_t^\varepsilon)$ introduced in Remark 3.1.

APPENDIX

LEMMA A.1. *Consider Assumptions* 1–4. *For any* $T > 0$, *we then have*

$$\lim_{c \to \infty} \limsup_{\varepsilon \to 0} P(Y_T^\varepsilon > c) = 0 \quad \textit{and} \quad \lim_{c \to \infty} \limsup_{\varepsilon \to 0} P\Big(\sup_{t \le T} |X_t^\varepsilon| > c\Big) = 0.$$

PROOF. Let

$$V_t^\varepsilon = Y_t^\varepsilon + \sup_{s \le t} |X_s^\varepsilon|. \tag{A.1}$$

Given the obvious inequalities $Y_T^\varepsilon \le V_T^\varepsilon$ and $\sup_{t \le T} |X_t^\varepsilon| \le V_T^\varepsilon$, it suffices to prove that

$$\lim_{c \to \infty} \limsup_{\varepsilon \to 0} P(V_T^\varepsilon > c) = 0. \tag{A.2}$$

Introduce the stopping times

$$\tau_r^\varepsilon = \inf(t \le T : V_t^\varepsilon \ge r) \wedge T, \qquad \inf(\varnothing) = \infty,$$

and notice that for any r,

$$P(V_T^\varepsilon > c) \le P\big(V_{\tau_r^\varepsilon}^\varepsilon > c\big) + P(\tau_r^\varepsilon < T). \tag{A.3}$$

Now, using the facts that $\{\tau_r^\varepsilon = 0\} = \{V_0^\varepsilon \ge r\}$ and $\{0 < \tau_r^\varepsilon < T\} = \{0 < \tau_r^\varepsilon < T\} \cap \{V_0^\varepsilon < r\} \cap \{V_{\tau_r^\varepsilon}^\varepsilon \ge r\}$ as well as Chebyshev's inequality, we obtain from (A.3):

$$P(V_T^\varepsilon > c) \le \frac{1}{c} E V_{\tau_r^\varepsilon}^\varepsilon + \frac{1}{r} E V_{\tau_r^\varepsilon}^\varepsilon + \frac{1}{r} E V_0^\varepsilon. \tag{A.4}$$

Since $EV_0^\varepsilon = EY_0^\varepsilon + E|x_0|$ and by Assumption 3, $\limsup_{\varepsilon \to 0} EY_0^\varepsilon < \infty$, we have

$$\lim_{r \to \infty} \limsup_{\varepsilon \to 0} \frac{1}{r} E V_0^\varepsilon = 0.$$

Consequently (A.2) holds if

$$\limsup_{\varepsilon \to 0} E\big(V_{\tau_r^\varepsilon}^\varepsilon\big) \le k \tag{A.5}$$

with k not depending on r. Indeed, from (A.4) it then follows that

$$\lim_{c\to\infty}\limsup_{\varepsilon\to 0} P(V_T^\varepsilon > c) \le \frac{k}{r} + \frac{1}{r}\limsup_{\varepsilon\to 0} EV_0^\varepsilon \to 0 \quad \text{for } r\to\infty.$$

To show inequality (A.5) let us use the representations (1.19), (1.20) [here and in what follows we shall drop the arguments in the functions $a_s, \hat{a}_s, b_s$]. From (1.20), it follows that

$$\begin{aligned} Y_t^\varepsilon &\le Y_0^\varepsilon + 2\int_0^t (a_s + \hat{a}_s X_{s-}^\varepsilon)^2 \Delta A_s^\varepsilon \, dA_s^\varepsilon + 2\int_0^t b_s^2 \, d[M^\varepsilon, M^\varepsilon]_s \\ &\le Y_0^\varepsilon + 4\varepsilon\int_0^t \Big[a_s^2 + \hat{a}_s^2 \sup_{u\le s}|X_{u-}^\varepsilon|^2\Big] dA_s^\varepsilon + 2\int_0^t b_s^2 \, d[M^\varepsilon, M^\varepsilon]_s. \end{aligned} \tag{A.6}$$

Furthermore, from (1.19) we deduce

$$\sup_{u\le t}|X_u^\varepsilon| \le |x_0| + \int_0^t \Big[|a_s| + |\hat{a}_s| \sup_{u\le s}|X_{u-}^\varepsilon|\Big] dA_s^\varepsilon + 1 + \sup_{u\le t}\left(\int_0^u b_s \, dM_s^\varepsilon\right)^2. \tag{A.7}$$

Using Assumptions 1, 2 and $Y_{s-}^\varepsilon - Y_{s-I}^\varepsilon \le Y_{s-}^\varepsilon$, we have

$$|a_s| \le L(1 + Y_{s-}^\varepsilon), \qquad |\hat{a}_s| \le L, \qquad b_s^2 \le L(1 + Y_{s-}^\varepsilon).$$

With these estimates we obtain from (A.6):

$$\begin{aligned} Y_t^\varepsilon &\le Y_0^\varepsilon + 4\varepsilon\int_0^t \Big[L^2(1 + Y_{s-}^\varepsilon)^2 + L^2 \sup_{u\le s}|X_{u-}^\varepsilon|^2\Big] dA_s^\varepsilon \\ &\quad + 2\int_0^t L(1 + Y_{s-}^\varepsilon)\, d[M^\varepsilon, M^\varepsilon]_s \\ &\le Y_0^\varepsilon + 8L^2\varepsilon\int_0^t \big[1 + (V_{s-}^\varepsilon)^2\big] dA_s^\varepsilon + 2L\int_0^t (1 + V_{s-}^\varepsilon)\, d[M^\varepsilon, M^\varepsilon]_s. \end{aligned} \tag{A.8}$$

Analogously we obtain from (A.7):

$$\sup_{u\le t}|X_u^\varepsilon| \le 1 + |x_0| + L\int_0^t (1 + V_{s-}^\varepsilon)\, dA_s^\varepsilon + \sup_{u\le t}\left(\int_0^u b_s \, dM_s^\varepsilon\right)^2. \tag{A.9}$$

From (A.8) and (A.9) it follows that

$$\begin{aligned} V_{t\wedge\tau_r^\varepsilon}^\varepsilon &\le 1 + V_0^\varepsilon + 8L^2\varepsilon\int_0^{t\wedge\tau_r^\varepsilon} \big[1 + (V_{s-}^\varepsilon)^2\big] dA_s^\varepsilon \\ &\quad + 2L\int_0^{t\wedge\tau_r^\varepsilon} (1 + V_{s-}^\varepsilon)\, d[M^\varepsilon, M^\varepsilon]_s \\ &\quad + L\int_0^{t\wedge\tau_r^\varepsilon} (1 + V_{s-}^\varepsilon)\, dA_s^\varepsilon + \sup_{u\le t\wedge\tau_r^\varepsilon}\left(\int_0^u b_s \, dM_s^\varepsilon\right)^2. \end{aligned} \tag{A.10}$$

Defining

$$U_t^{\varepsilon, r} = EV_{t\wedge\tau_r^\varepsilon}^\varepsilon, \tag{A.11}$$

from (A.10) we then find

$$\begin{aligned} U_t^{\varepsilon,r} \le 1 &+ EV_0^\varepsilon + 8L^2\varepsilon E\int_0^{t\wedge\tau_r^\varepsilon}\left[1 + (V_{s-}^\varepsilon)^2\right] dA_s^\varepsilon \\ &+ 2LE\int_0^{t\wedge\tau_r^\varepsilon}(1 + V_{s-}^\varepsilon)\, d\langle M^\varepsilon\rangle_s \\ &+ LE\int_0^{t\wedge\tau_r^\varepsilon}(1 + V_{s-}^\varepsilon)\, dA_s^\varepsilon + 4E\int_0^{t\wedge\tau_r^\varepsilon} b_s^2\, d\langle M^\varepsilon\rangle_s, \end{aligned} \tag{A.12}$$

where we have used Doob's inequality. From Assumption 2 we have $b_s^2 \le L(1 + Y_{s-}^\varepsilon) \le L(1 + V_{s-}^\varepsilon)$ and from (1.13) we have $A_t^\varepsilon = \langle M^\varepsilon\rangle_t$; therefore, using (A.12) and some additional straightforward estimates that take the definition of τ_r^ε into account, we obtain

$$\begin{aligned} U_t^{\varepsilon,r} &\le 1 + EV_0^\varepsilon + 7LE\int_0^{t\wedge\tau_r^\varepsilon}(1 + V_{s-}^\varepsilon)\, dA_s^\varepsilon \\ &\quad + 8L^2\varepsilon E\int_0^{t\wedge\tau_r^\varepsilon}\left[1 + (V_{s-}^\varepsilon)^2\right] dA_s^\varepsilon \\ &\le 1 + EV_0^\varepsilon + 8L^2\varepsilon(1 + r^2)A_T^\varepsilon + 7LA_T^\varepsilon + 7L\int_0^t (U_{s-}^{\varepsilon,r})\, dA_s^\varepsilon. \end{aligned} \tag{A.13}$$

From this inequality, using Theorem 2.4.3 in Liptser and Shiryayev (1989), namely an analogue of the classical Gronwall–Bellman inequality, we have

$$U_T^{\varepsilon,r} \le \left[1 + EV_0^\varepsilon + 8L^2\varepsilon(1 + r^2)A_T^\varepsilon + 7LA_T^\varepsilon\right]\exp[7LA_T^\varepsilon]$$

and, consequently, by the convergence $A_T^\varepsilon \to T$ for $\varepsilon \to 0$,

$$\limsup_{\varepsilon\to 0} U_T^{\varepsilon,r} \le \left[1 + \limsup_{\varepsilon\to 0} EY_0^\varepsilon + E|x_0| + 7LT\right]\exp[7LT] = k.$$

Thus [see definition (A.11)] we have established the required inequality (A.5). □

LEMMA A.2. *Let $G^\varepsilon = (G_t^\varepsilon)_{t\ge 0}$ be a family of right-continuous and increasing processes with $G_0^\varepsilon = 0$ such that*

$$G_t^\varepsilon \to_P G_t \quad \text{as } \varepsilon \to 0 \text{ for all } t > 0, \tag{A.14}$$

where G_t is a continuous and increasing process with $G_0 = 0$. Furthermore let $f = (f_t)$ be a continuous process. Then

$$\sup_{t\le T}\left|\int_0^t f_s\, d(G_s^\varepsilon - G_s)\right| \to_P 0 \quad \text{for } \varepsilon \to 0. \tag{A.15}$$

PROOF (Sketch only). For integer $N > 0$, define $f_s^N := f_{[Ns]/N}$ so that

$$\lim_{N\to\infty}\sup_{0\le s\le T}|f_s - f_s^N| = 0 \tag{A.16}$$

and consider

$$
\begin{aligned}
&\sup_{t \le T} \left| \int_0^t f_s \, d(G_s^\varepsilon - G_s) \right| \\
(\text{A.17}) \qquad &\le \sup_{t \le T} \left| \int_0^t f_s^N \, d(G_s^\varepsilon - G_s) \right| + \int_0^t |f_s - f_s^N| \, d(G_s^\varepsilon + G_s) \\
&\le \sup_{t \le T} \left| \int_0^t f_s^N \, d(G_s^\varepsilon - G_s) \right| + \sup_{0 \le s \le T} |f_s - f_s^N| (G_T^\varepsilon + G_T),
\end{aligned}
$$

where, to estimate the first term, one then uses the fact that (A.14) implies

$$\sup_{t \le T} |G_t^\varepsilon - G_t| \to_P 0 \quad \text{for } \varepsilon \to 0;$$

see, for example, Problem 5.3.2 in Liptser and Shiryayev (1989). □

REFERENCES

AASE, K. K. (1988). Contingent claims valuation when the security price is a combination of an Itô process and a random point process. *Stochastic Process. Appl.* **28** 185–220.

BENSOUSSAN, A. (1984). On the theory of option pricing. *Acta Appl. Math.* **2** 139–158.

BLACK, F. and SCHOLES, M. (1973). The pricing of options and corporate liabilities. *J. Political Economy* **81** 637–659.

DENNY, J. L. and SUCHANEK, G. L. (1986). On the use of semimartingales and stochastic integrals to model continuous trading. *J. Math. Economics* **15** 255–266.

DOLEANS-DADE, C. (1970). Quelques applications de la formule de changement de variable pour les semimartingales. *Z. Wahrsch. Verw. Gebiete* **16** 181–194.

DUFFIE, D. and PROTTER, P. (1988). From discrete to continuous time finance: Weak convergence of the financial gain process. Working paper, Stanford Univ.

HARRISON, J. M. and PLISKA, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Process. Appl.* **11** 215–260.

HARRISON, J. M. and PLISKA, S. R. (1983). A stochastic calculus model of continuous trading: Complete markets. *Stochastic Process. Appl.* **15** 313–316.

HE, H. (1989). Convergence from discrete to continuous time financial models. IBER Working paper 190, Univ. California, Berkeley.

KARATZAS, I. (1989). Optimization problems in the theory of continuous trading. *SIAM J. Control Appl.* **27** 1221–1259.

KIND, P. (1988). Option valuation in a continuous time conditional variance model of asset price returns. Dept. Economics, Univ. Chicago.

LIPTSER, R. SH. and SHIRYAYEV, A. N. (1989). *Theory of Martingales*. Kluwer, Dordrecht.

STOUT, W. F. (1973). *Almost Sure Convergence*. Wiley, New York.

WILLINGER, W. and TAQQU, M. S. (1989). Pathwise stochastic integration and application to the theory of continuous trading. *Stochastic Process. Appl.* **32** 254–280.

PAOLO KIND
DEPARTMENT OF ECONOMICS
UNIVERSITY OF CHICAGO
CHICAGO, ILLINOIS

ROBERT SH. LIPSTER
INSTITUTE OF CONTROL SCIENCES
USSR ACADEMY OF SCIENCES
USSR

WOLFGANG J. RUNGGALDIER
DIPARTIMENTO MATEMATICA
PURA ED APPLICATA
UNIVERSITÀ DI PADOVA
VIA BELZONI 7
I-35131 PADOVA
ITALY