The Annals of Statistics
1992, Vol. 20, No. 2, 623–654

ASYMPTOTIC EXPANSIONS AND BOOTSTRAPPING DISTRIBUTIONS FOR DEPENDENT VARIABLES: A MARTINGALE APPROACH[1]

BY PER ASLAK MYKLAND

University of Chicago and University of California, Berkeley

The paper develops a one-step triangular array asymptotic expansion for continuous martingales which are asymptotically normal. Mixing conditions are not required, but the quadratic variations of the martingales must satisfy a law of large numbers and a central limit type condition. From this result we derive expansions for the distributions of estimators in asymptotically ergodic differential equation models, and also for the bootstrapping estimators of these distributions.

1. Introduction. Edgeworth-expansions are a useful tool in inference. Many methods are directly based on expansions, ranging from the conditional and parametric [Barndorff-Nielsen (1980, 1983, 1984, 1986a, b, 1988), Cox (1980), Hinkley (1980), McCullagh (1984, 1987), McCullagh and Cox (1986)] to the unconditional or nonparametric [Hall (1983), Withers (1983), Abramovitch and Singh (1985), Bhattacharya and Qumsiyeh (1989)]. Expansions are also useful for procedures which originate from considerations other than Edgeworth correction, such as bootstrapping [Singh (1981), Beran (1982, 1987, 1988a, b), Hall (1986a, b, 1988), Efron (1987), Liu and Singh (1987)]. The references given here are, of course, only a small sample of the work done in the areas concerned.

Until recently, however, asymptotic expansions were only known for estimators based on independent observations [Bhattacharya and Ranga Rao (1976, 1986), Bhattacharya and Ghosh (1978)]. This situation is currently in the process of changing. Research has been conducted on three different types of inference situations:

1. Parametric models: Independence assumptions are often not required in the papers on conditional inference cited above.
2. Models satisfying weak dependence/mixing assumptions [Goetze and Hipp (1983), Bose (1986c, 1987, 1988), Jensen (1986)].
3. Markov models [Malinovskii (1987), Jensen (1989)].

Received January 1990; revised January 1991.
[1]Research supported by Grant D.08.40.121 from the Norwegian General Research Council (NAVF) and partially also by NSF Grants DMS-87-01426 and DMS-89-02667.
AMS 1980 *subject classifications*. Primary 62E20, 62M05, 62M09; secondary 60F99, 60G44, 60H10.
Key words and phrases. Bootstrapping, differential equations, Edgeworth-expansions, martingales.

There are also some papers which are specific for ARMA-processes, such as Taniguchi (1984) and Tanaka (1986).

The one technique which is conspicuously absent in the literature on expansions for dependent variables is the use of martingales. Martingales have been extremely powerful for proving central limit theorems for estimators [see Hall and Heyde (1980), Rebolledo (1980), Helland (1982), Jeganathan (1982) and many others] and one would expect this to carry over to expansions. The program of this paper is to pursue this line of investigation.

We shall in the following present a one-step Edgeworth expansion for continuous martingales which are asymptotically normal (Section 2). The results will be for triangular arrays of such martingales. This generality is useful, inter alia, for a second order analysis of estimators under local alternatives and of bootstrapping (as in Section 3.2).

As an application, we investigate (in Section 3) how this expansion turns out in the context of unconditional inference in a class of (2, 1) exponential family models. We study both the Edgeworth expansion for the sampling distribution and for the bootstrapping estimator of this distribution. The class of models to be considered are described by having likelihoods of the form

$$\frac{dP_\theta}{dP_0} \text{ at time } t = \exp\left\{\theta u_t - \frac{1}{2}\theta^2 I_t\right\}, \tag{1.1}$$

where (u_t, I_t) is a sufficient statistic for some process up to time t. We shall impose the condition that the time variable t is continuous, that (u_t, I_t) is continuous in t and that I_t is nondecreasing. Inference situations covered by this model include stochastic differential equations of the form

$$X_t = X_0 + \theta\int_0^t a_s(X)\,ds + \int_0^t v_s(X)\,ds + \int_0^t \gamma_s(X)\,dB_s, \qquad t \geq 0, \tag{1.2}$$

where (B_t) is a Wiener process; see Section 4.

Comparing to the i.i.d. case, our results generalize the asymptotic expansions in Goetze and Hipp (1978). The expansions do not hold pointwise—they hold in a test function topology. This circumvents the problems which have been connected to finding Berry-Esseen bounds for martingales with a rate of convergence of $n^{-1/2}$. This is a rate which is often achieved in our martingale expansions (see Sections 3.3 and 4), but uniform bounds with such a rate have only been found under very strong conditions or in special cases. Work in this area includes Bolthausen (1982), Lipster and Shiryaev (1982), Mishra and Prakasa Rao (1985), Bose (1986a, b) and Haeusler (1988).

2. Asymptotic expansions for martingales.

2.1. *Main result*: *Nonrandom norming.* Our purpose is to find an asymptotic expansion for a triangular array $l_t^{(T)}$, $0 \leq t \leq T$, $T \geq 0$, of continuous martingales. $(l_t^{(T)}, 0 \leq t \leq T)$ is defined on a filtered probability space with probability measure $P^{(T)}$. This filtered space can vary with T. A key role will be played by the (total) quadratic variation of $l_t^{(T)}$ up to time T, $I_T = \langle l^{(T)} \rangle_T = \lim_{\Delta \to 0} \Sigma (l_{t_{i+1}}^{(T)} - l_{t_i}^{(T)})^2$, where the t_i's define a partition of $[0, T]$ and $\Delta = \max(t_{i+1} - t_i)$. Note that this I_T is not yet the same as the one in (1.1). In Sections 3 and 4, however, they will be the same.

The reason why I_T is such an important object can be seen from the central limit theorem for continuous martingales. The Lindeberg condition being automatic for continuous martingales, it has the following form [cf. Hall and Heyde (1980), Theorem A (page 100) and Dambis (1965), Theorem 7]: If there are nonrandom constants c_T and b^2, with $b > 0$ and $c_T \to \infty$ as $T \to \infty$, so that

$$\frac{I_T}{c_T} \to_P b^2, \tag{2.1}$$

then

$$\frac{l_T^{(T)}}{c_T^{1/2}} \to N(0, b^2) \quad \text{in law.} \tag{2.2}$$

In other words, a law of large numbers for I_T implies a CLT for $l_T^{(T)}$. The interesting thing is that a central limit condition for I_T is the main assumption needed for the asymptotic expansion. In each case, the martingale structure permits you to move one step further in asymptotic refinement from what you can otherwise show.

To be precise, the conditions are the following:

THE CENTRAL LIMIT CONDITION FOR I_T. There are stochastic variables (Z, ξ) so that

$$\left(\frac{l_T^{(T)}}{b c_T^{1/2}}, \frac{1}{r_T} \left(\frac{I_T}{c_T} - b^2 \right) \right) \to (Z, \xi) \quad \text{in law, as } T \to \infty, \tag{2.3}$$

where r_T is another normalizing constant, $r_T \to 0$.

THE INTEGRABILITY CONDITION FOR I_T. There are constants $\underline{k}$ and $\bar{k}$, $0 \le \underline{k} < b^2 < \bar{k} \le \infty$, so that

$$\frac{1}{r_T} \left(\frac{I_T}{c_T} - b^2 \right) \chi_{D_T} \quad \text{is uniformly integrable,} \tag{2.4}$$

χ_{D_T} being the indicator function of the set D_T, and

$$P^{(T)}(\tilde{D}_T) = o(r_T), \tag{2.5}$$

where $\tilde{}$ denotes complementation and

$$D_T = \left\{ \underline{k} \le \frac{I_T}{c_T} \le \bar{k} \right\}. \tag{2.6}$$

So far, we have said nothing about the joint distribution of (Z, ξ). In general, it is only known that the marginal distribution of Z is $N(0, 1)$; see

(2.1)–(2.2). Further answering this question will be a prime concern later on in the paper.

The main expansion result is as follows.

THEOREM 2.1. *Suppose* $\mathbf{T}$ *is a subset of* $\mathbf{R}^+$ *with* $\sup \mathbf{T} = \infty$. *Suppose that* $(l_t^{(T)}, 0 \le t \le T)$, $T \in \mathbf{T}$, *is a triangular array of zero mean martingales and that the central limit and integrability conditions above are satisfied as* $T \to \infty$ *through* $\mathbf{T}$.

Suppose that C *is a collection of twice continuously differentiable functions* $g\colon \mathbf{R} \to \mathbf{R}$ *satisfying*: (i) *there is an* $M > 0$ *such that* $|g(x)| \le M$, $|g'(x)| \le M$ *and* $|g''(x)| \le M$ *for all* x *and all* $g \in C$, *and* (ii) $\{g''\colon g \in C\}$ *is equicontinuous almost everywhere with respect to Lebesgue measure. Then, uniformly in* $g \in C$, *as* $T \to \infty$ *through* $\mathbf{T}$,

$$Eg\left(\frac{l_T^{(T)}}{c_T^{1/2}}\right) - Eg(N(0, b^2)) = r_T \frac{1}{2} E(\xi g''(bZ)) + o(r_T). \tag{2.7}$$

Note that a.e. equicontinuity serves the same purpose here as in weak convergence; see Pollard (1984), Example IV.19 (page 73).

2.2. *Random norming, test function convergence and local martingales.* It is desirable to consider an expansion for the law of $l_T^{(T)}/\sigma_T^{1/2}$, where σ_T is a random variable. For example, it could be an estimate of the variance of $l_T^{(T)}$. Theorem 2.1 extends easily to this case, subject to some conditions on σ_T. In this section, we shall state this extension and then discuss an alternative representation for the expansion. Finally, we shall look at how this all works for local martingales.

The conditions on σ_T which are needed for a result on random norming are as follows:

THE CENTRAL LIMIT CONDITION FOR σ_T. There is a $b_* > 0$ and there are random variables (Z, ξ^*) so that

$$\left(\frac{l_T^{(T)}}{b c_T^{1/2}}, \frac{1}{r_T}\left(\frac{\sigma_T}{c_T} - b_*^2\right)\right) \to (Z, \xi^*) \quad \text{in law, as } T \to \infty, \tag{2.8}$$

where b^2, c_T and r_T are the same as in (2.3).

THE INTEGRABILITY CONDITION FOR σ_T. There are measurable sets D_T^* and a $\delta > 0$ so that

$$\sup_{T \in \mathbf{T}} E^{(T)}\left[\left|\frac{1}{r_T}\left(\frac{\sigma_T}{c_T} - b_*^2\right)\right|^{1+\delta} \chi_{D_T^*}\right] < \infty, \tag{2.9}$$

$\chi_{D_T^*}$ being the indicator function of the set D_T^*, $E^{(T)}$ being the expectation operator of the probability measure $P^{(T)}$ and

$$P^{(T)}(\tilde{D}_T^*) = o(r_T). \tag{2.10}$$

We are now ready to state the extended expansion result.

THEOREM 2.2. *Suppose* $\mathbf{T}$ *is a subset of* $\mathbf{R}^+$ *with* $\sup \mathbf{T} = \infty$. *Suppose that* $(l_t^{(T)}, 0 \le t \le T)$, $T \in \mathbf{T}$, *is a triangular array of zero mean martingales and that the central limit and integrability conditions for* I_T *and* σ_T *are satisfied as* $T \to \infty$ *through* $\mathbf{T}$. *Further assume that* $P^{(T)}(\sigma_T > 0) > 0$ *for all* $T \in \mathbf{T}$. *Set*

$$F_T(x) = P^{(T)}\left(\frac{l_T^{(T)}}{\sigma_T^{1/2}} \le x \,\middle|\, \sigma_T > 0 \right). \tag{2.11}$$

Suppose that C *is a collection of twice continuously differentiable functions* g: $\mathbf{R} \to \mathbf{R}$ *satisfying conditions* (i) *and* (ii) *of Theorem* 2.1.

Then, uniformly in $g \in C$, *as* $T \to \infty$ *through* $\mathbf{T}$,

$$\begin{aligned} &\int_{-\infty}^{\infty} g(x)\, dF_T(x) - \int_{-\infty}^{\infty} g(x)\, d\Phi(\beta^{-1}x) \\ &\qquad = r_T \tfrac{1}{2} E\left[b^{-2}\xi\beta^2 g''(\beta Z) - b_*^{-2}\xi^*\beta Z g'(\beta Z) \right] + o(r_T), \end{aligned} \tag{2.12}$$

where β *is the asymptotic standard deviation of* $l_T^{(T)}/\sigma_T^{1/2}$, *that is,*

$$\beta = b b_*^{-1}. \tag{2.13}$$

Φ *is the c.d.f. of the standard normal distribution. Furthermore, for all* $g \in C$,

$$\begin{aligned} &\left| \frac{1}{r_T} \int_{-\infty}^{\infty} g(x)\, d\left[F_T(x) - \Phi(\beta^{-1}x) \right] \right| \\ &\qquad \le \frac{1}{2} M E^{(T)} \left| \frac{1}{r_T} \left(\frac{I_T}{c_T} - b^2 \right) \chi_{D_T} \right| \\ &\qquad\quad + Mk \left[E^{(T)} \left| \frac{1}{r_T} \left(\frac{\sigma_T}{c_T} - b_*^2 \right) \chi_{D_T^*} \right|^{1+\delta} \right]^{1/(1+\delta)} \\ &\qquad\quad + Mk \frac{1}{r_T} P^{(T)}(\tilde{D}_T) + 3M \frac{1}{r_T} P^{(T)}(\tilde{D}_T^*), \end{aligned} \tag{2.14}$$

where k *depends on* δ, $\bar{k}$, $\underline{k}$, b^2 *and* b_*^2.

REMARK 2.3. To see that the result above is indeed an asymptotic expansion in a test function topology, consider the following convergence type.

DEFINITION. If $\{G_T, T \in \mathbf{T}\}$ is a set of functions of finite variation, set

$$G_T(x) = o_2(r_T) \quad \text{as } T \to \infty \text{ through } \mathbf{T}$$

if

$$\sup_{g \in C} \left| \frac{1}{r_T} \int_{-\infty}^{\infty} g(x)\, dG_T(x) \right| \to 0 \quad \text{as } T \to \infty \text{ through } \mathbf{T}$$

for every set C satisfying the description in Theorem 2.1. If the G_T's are random objects on some space with probability measure Q, we write

$$G_T(x) = o_{Q,2}(r_T) \quad \text{as } T \to \infty \text{ through } \mathbf{T}$$

if the convergence holds in Q-probability. The phrase "as $T \to \infty$ through $\mathbf{T}$" will be suppressed when $\mathbf{T} = \mathbf{R}^+$.

Now let ψ and ψ_* be Borel-measurable functions so that

$$\psi(Z) = b^{-2}E(\xi|Z) \quad \text{a.s.,} \tag{2.15}$$

$$\psi_*(Z) = b_*^{-2}E(\xi^*|Z) \quad \text{a.s.} \tag{2.16}$$

If we assume that ψ is absolutely continuous, that ψ' and ψ_* are of bounded variation on finite intervals and that $\psi'(x)\exp(-x^2/2) \to 0$, $x\psi(x) \exp(-x^2/2) \to 0$ and $x\psi_*(x)\exp(-x^2/2) \to 0$ as $|x| \to \infty$, the result of the theorem can be written as

$$\begin{aligned} &F_T(x) - \Phi(\beta^{-1}x) - r_T \tfrac{1}{2}\lambda(\beta^{-1}x)\phi(\beta^{-1}x) \\ &\qquad = o_2(r_T) \quad \text{as } T \to \infty \text{ through } \mathbf{T}, \end{aligned} \tag{2.17}$$

where

$$\lambda(x) = \psi'(x) + \psi(x)x + \psi_*(x)x, \tag{2.18}$$

ϕ being the density of $N(0,1)$.

To see that (2.12) is equivalent to (2.17) under the assumptions on ψ and ψ_*, one needs the fact that Z is $N(0,1)$. The result then follows by integration by parts.

REMARK 2.4. It is sufficient for the result to hold that $(l_t^{(T)}, 0 \le t \le T)$ be an array of local martingales in the sense that for each T there is a sequence of stopping times $\{\tau_N\}$, $\tau_N \le T$, so that $l_{\min(\tau_N, t)}^{(T)}$ is a martingale for each N, and so that if we set $\tau = \lim_{N \to \infty} \tau_N$ and we define I_T to be the quadratic variation of $l_t^{(T)}$ on $[0, \tau)$, then either $T = \tau$ or $I_T = \infty$ (or both).

One then has to require that $\mathbf{T}$ be such that

$$P(I_T < \infty \text{ and } \sigma_T > 0) > 0 \quad \text{for all } T \in \mathbf{T}, \tag{2.19}$$

and to replace (2.11) by

$$F_T(x) = P^{(T)}\left(\frac{l_T^{(T)}}{\sigma_T^{1/2}} \le x \,\middle|\, I_T < \infty \text{ and } \sigma_T > 0 \right). \tag{2.20}$$

This extension is highly relevant in connection with inference problems; see Sections 3 and 4. Our proof (in Section 5) of Theorems 2.1 and 2.2 is actually the proof of this remark.

2.3. *Some initial examples.* In the following, we illustrate Theorem 2.1 with some simple cases. Further illustration is provided by Sections 3 and 4.

EXAMPLE. Let (W_t) be a Wiener process and let I_t be a process which is independent of (W_t). Set $l_t = W_{I_t}$. If $T^{1/2}(I_T/T - b^2) \to \xi$, then Z and ξ are independent and $\psi(z) = b^{-2}E\xi$ for all z.

EXAMPLE. Let $(l_t) = (l_t^{\mathrm{OU}})$ be the martingale which is the derivative of the log likelihood when estimating θ in the Ornstein–Uhlenbeck process $dX_t = \theta X_t\,dt + \gamma\,dB_t$, where $\theta < 0$. From (4.13) below, it follows that $b^2 = 1/2|\theta|$ and that $\xi = 2b^3Z$, so that here there is perfect dependence between Z and ξ. In this case, $\psi(z) = 2bz$.

EXAMPLE. Let (l_t^{OU}) be as in the previous example and let I_t^{OU} be its quadratic variation. By Dambis (1965) and Dubins and Schwartz (1965), there is a Brownian motion (W_t) so that $l_t^{\mathrm{OU}} = W_{I_t^{\mathrm{OU}}}$. Now set (for some constant A)

$$I_t = \max\left(I_t^{\mathrm{OU}}, b^2t + b^2A\sqrt{t}\right) \tag{2.21}$$

and

$$l_t = W_{I_t}. \tag{2.22}$$

I_t is now the quadratic variation of (l_t) and the joint limit of $(l_T/b\sqrt{T}, \sqrt{T}(I_T/T - b^2))$ is, in view of the preceding example, $(Z, \max(2b^3Z, b^2A))$. Hence

$$\psi(z) = \max(2bz, A). \tag{2.23}$$

EXAMPLE. In the previous example, let A be random and independent of the underlying Ornstein–Uhlenbeck process. Also suppose that $E|A| < \infty$. Then, from the above,

$$\psi(z) = 2bz + \int_{2bz}^{+\infty} P(A > x)\,dx. \tag{2.24}$$

3. Application to (2, 1) exponential family models: Sampling distributions and bootstrapping distributions.

3.1. *Introduction.* For the (2, 1) exponential families defined by (1.1), the derivative of the log likelihood is given by

$$l_t(\theta) = u_t - \theta I_t. \tag{3.1}$$

$(l_t(\theta))$ is a (local) martingale under P_θ [in analogy to Chapter 6 of Hall and Heyde (1980)]. I_T is the quantity from (1.1), but it is also the I_T which is the quadratic variation of $(l_t(\theta))$ (this is seen by using Itô's formula on the log likelihood and then using that a continuous martingale of finite total variation is constant).

The maximum likelihood estimator in this model is $\hat{\theta}_T = u_T/I_T$. Hence

$$(\hat{\theta}_T - \theta) = \frac{l_T(\theta)}{I_T}. \tag{3.2}$$

We shall analyze the approximate pivot ("root") $\sqrt{s_T}(\hat{\theta}_T - \theta)$, where s_T can be one of several normalizations:

$$\begin{aligned} \text{nonrandom: } s_T &= c_T, \\ \text{observed information: } s_T &= I_T, \\ \text{estimated expected information: } s_T &= E_{\hat{\theta}_T} I_T. \end{aligned} \tag{3.3}$$

Both s_T and I_T can take on unacceptable values, such as 0 and $+\infty$. We shall discuss this technical difficulty in Section 3.4. At this point, it suffices to mention that we shall consider the distribution function

$$F_T(x;\theta) = P_\theta\Big(\sqrt{s_T}\big(\hat{\theta}_T - \theta\big) \le x \Big| 0 < I_T < \infty \text{ and } s_T < \infty\Big). \tag{3.4}$$

By the remarks at the beginning of Section 2.1, if

$$\frac{I_T}{c_T} \to_{P_\theta} b^2 \quad \text{and} \quad \frac{s_T}{c_T} \to_{P_\theta} b_{**}^2, \tag{3.5}$$

then $F_T(x;\theta)$ converges weakly to the c.d.f. of $N(0,\beta^2)$, where $\beta = b^{-1}b_{**}$. It is also easy to see that this is also the limit of $F_T(x;\theta_T)$ provided $\sqrt{c_T}(\theta_T - \theta) = O(1)$. Since $\sqrt{c_T}(\hat{\theta}_T - \theta) = O_{P_\theta}(1)$, it further follows [cf. Pollard (1984), Theorem IV.13 (page 71)] that $F_T(x;\hat{\theta}_T)$ converges in probability to $N(0,\beta^2)$, whence the bootstrap distribution is a consistent estimator of the underlying distribution.

Thus, both $F_T(x;\theta_T)$ and $F_T(x;\hat{\theta}_T)$ are close to $F_T(x;\theta)$ as T becomes large. Turning to second order considerations, we can compare the Edgeworth expansions for $F_T(x;\theta)$, $F_T(x;\theta_T)$ and $F_T(x;\hat{\theta}_T)$. These expansions are obtained by using Theorem 2.2 or Remark 2.4 on the (local) martingale in (3.1).

We shall do this in two instalments. In the next section, we shall give a discussion of the expansion terms for the various choices for s_T. We begin, however, by stating formally an expansion result for the root $\sqrt{I_T}(\hat{\theta}_T - \theta)$. For $\sqrt{c_T}(\hat{\theta}_T - \theta)$, the conditions are the same, but the correction term is different (see Section 3.2). For $\sqrt{E_{\hat{\theta}_T} I_T}(\hat{\theta}_T - \theta)$, the correction term is (usually) the same, but stronger conditions have to be imposed (cf. Section 3.2 and the derivation of Table 1 in Section 6.1).

All the conditions imposed will be under measure P_θ, so they do not involve triangular arrays. We shall consider convergence through subsets $\mathbf{T}$ of $\mathbf{R}^+$ because this will be useful in Section 3.3.

PROPOSITION 3.1. *Assume that* $\mathbf{T}$ *is an unbounded subset of* $\mathbf{R}^+$ *for which there are constants* $b = b(\theta) > 0$, c_T *and* r_T, $c_T \to \infty$ *and* $r_T \to 0$ *as* $T \to \infty$

through $\mathbf{T}$, *so that*

$$\xi_T(\theta) = \frac{1}{r_T}\left(\frac{I_T}{c_T} - b^2(\theta)\right) \tag{3.6}$$

satisfies:

(i) *There is a* $\delta > 0$ *such that*

$$\sup_{T \in \mathbf{T}} E_\theta |\xi_T(\theta)|^{1+\delta} < \infty. \tag{3.7}$$

(ii) *There is a Borel measurable function* $\psi_{\theta,\mathbf{T}}(z)$ *so that whenever*

$$\left(\frac{l_T(\theta)}{I_T^{1/2}}, \xi_T(\theta)\right) \to (Z, \xi), \tag{3.8}$$

in law under P_θ *through some subset of* $\mathbf{T}$, *then*

$$\psi_{\theta,\mathbf{T}}(Z) = b^{-2}(\theta) E_\theta(\xi | Z) \qquad P_\theta - a.s. \tag{3.9}$$

Then, if F_T *is given by* (3.4), *with* $s_T = I_T$, *and if* $b\sqrt{c_T}(\theta_T - \theta) \to \alpha$,

$$\int g(x)\, d[F_T(x;\theta_T) - \Phi(x)] - r_T \tfrac{1}{2} \int \psi_{\theta,\mathbf{T}}(x+\alpha)\, d[\phi(x) g'(x)] = o(r_T) \quad \text{as } T \to \infty \text{ through } \mathbf{T}, \tag{3.10}$$

the convergence being uniform in sets of functions g described in Theorem 2.1. *If* $|g(x)|$, $|g'(x)|$ *and* $|g''(x)|$ *are all bounded by* M, *then for all* $T \in \mathbf{T}$,

$$\left|\frac{1}{r_T}\int_{-\infty}^{\infty} g(x)\, d[F_T(x;\theta_T) - \Phi(x)]\right| \le k \exp\left\{\left(\frac{1+2\delta}{2\delta(1+\delta)} + \eta\right) b^2 c_T (\theta_T - \theta)^2\right\} \tag{3.11}$$

for all $\eta > 0$, *where* k *depends on* η, M, δ *and* $b^{-2(1+\delta)} \sup_T E_\theta |\xi_T|^{1+\delta}$ *only (and in particular not on* θ *or* $\{\theta_T\}_{T \in \mathbf{T}}$).

If $\psi_{\theta,\mathbf{T}}$ satisfies the differentiability and growth conditions described in Remark 2.3, then the expansion in (3.10) reads

$$F_T(x;\theta_T) - \Phi(x) - r_T \tfrac{1}{2} \psi'_{\theta,\mathbf{T}}(x+\alpha)\phi(x) = o_2(r_T) \quad \text{as } T \to \infty \text{ through } \mathbf{T}. \tag{3.12}$$

Note that the bound (3.11) is useful in deriving results on coverage probabilities for bootstrap- and Edgeworth-based confidence intervals; see Part II of Mykland (1989).

3.2. *Expressions for the expansion term for different choices of* s_T. Let $\theta_T = \theta + \alpha b^{-1} c_T^{-1/2} + o(c_T^{-1/2})$, where α is a local parameter in a contiguity neighborhood of θ. Under conditions to be discussed below, the asymptotic

TABLE 1
Form of λ for different roots

Root	β	Form of $\lambda_{\theta,\alpha}(x)$ under (3.16)	$F_T(x;\hat{\theta}_T) - F_T(x;\theta) = o_{P_\theta,2}(r_T)$?
$\sqrt{c_T}(\hat{\theta}_T - \theta)$	b^{-1}	$k(\theta)(1 - \alpha x - x^2) - \kappa(\theta)x$	Typically no
$\sqrt{I_T}(\hat{\theta}_T - \theta)$	1	$k(\theta)$	Yes
$\sqrt{E_{\hat{\theta}_T} I_T}(\hat{\theta}_T - \theta)$	1	$k(\theta)$	Yes

expansion for $F_T(x;\theta_T)$ is

$$F_T(x;\theta_T) - \Phi(\beta^{-1}x) - r_T \tfrac{1}{2}\lambda_{\theta,\alpha}(\beta^{-1}x)\phi(\beta^{-1}x) = o_2(r_T), \tag{3.13}$$

where $\lambda_{\theta,\alpha}$ is given by Table 1. The expansion for $F_T(x;\hat{\theta}_T)$ is then immediate [cf. Theorem IV.13 (page 71) of Pollard (1984)]. The last column in Table 1 answers the question of whether

$$F_T(x;\hat{\theta}_T) - F_T(x;\theta) = o_{P_\theta,2}(r_T), \tag{3.14}$$

which, when true, may suggest that the bootstrapping distribution asymptotically is a better estimator of $F_T(x;\theta)$ than is the asymptotic distribution.

The two last roots in Table 1 have asymptotic distribution (i.e., β) independent of θ, while this is not true for the first one. The relationship between (3.14) and the independence between β and θ is consistent with the conclusions from the iid case, see, for example, Beran (1987) and Beran (1988b). Since this is a parametric model, it is also intimately related to results on a.e. automatic invariance, like those in Le Cam [(1986), Chapter 8]. We pursue this further in the next section.

The condition for (3.13) to hold is, clearly, that $l_T^{(T)} = l_T(\theta_T)$, I_T and $\sigma_T = I_T^2/s_T$ satisfy the assumptions of Theorem 2.2 or Remark 2.4. For the case $s_T = I_T$, sufficient conditions are stated explicitly in Proposition 3.1.

The form of $\lambda_{\theta,\alpha}$ can be related to that of $\lambda_{\theta,0}$, since an argument akin to Le Cam's third lemma (see Section 6.1) yields (in similar notation) that

$$\psi_{\theta,\alpha}(x) = \psi_{\theta,0}(x+\alpha). \tag{3.15}$$

Table 1 then follows if we assume that $\psi_{\theta,0}$ has the form

$$\psi_{\theta,0}(x) = k(\theta)x + \kappa(\theta). \tag{3.16}$$

This is, in fact, typical, as will be clear from Section 3.3 and Section 4. The table, together with (3.15), is derived in Section 6.1.

3.3. *The form of ψ_θ and r_T: Uniformity conditions.* It is usually the case that one can take $r_T = c_T^{-1/2}$ and

$$\psi_{\theta,\mathrm{T}}(z) = zb(\theta)^{-3}\frac{d}{d\theta}b^2(\theta) + \text{constant}, \tag{3.17}$$

where $b^2 = b^2(\theta)$ is the asymptotic variance of $l_T(\theta)/c_T^{1/2}$ under P_θ. This explains our assumption (3.16). Also, it means that the convergence rate in this expansion is of the usual $n^{-1/2}$ form: If $c_T = T$ (which it is in asymptotically ergodic situations), then $r_T = T^{-1/2}$. This rate has been very difficult to obtain for martingale Berry–Esseen bounds (cf. the remarks at the end of Section 1).

In the asymptotically ergodic case, therefore, the expansion has a very standard appearance. For example, if $F_T(x;\theta)$ is the distribution of $I_T^{1/2}(\hat{\theta}_T - \theta)$ under P_θ [as in (3.4)], then

$$F_T(x;\theta_T) = \Phi(x) + \tfrac{1}{2}T^{-1/2}\psi'_{\theta,\mathbf{T}}\phi(x) + o_2(T^{-1/2}). \tag{3.18}$$

It is not necessary that $c_T = T$. In particular, c_T can be anything that is continuous and increases to infinity. Any such rate c_T can be obtained from any other by a nonrandom time change. For a natural example of this kind of time change, see Proposition 4.2. However, (3.18) remains true, with $c_T^{-1/2}$ replacing $T^{-1/2}$.

To back up the above statements with theorems, we shall do two things. One is to show that they are true for specific differential equation models, as in Section 4. The other is to impose some uniformity in θ and use the kind of argument that, for example, leads to the Hájek–Le Cam convolution theorem or the a.e. automatic invariance of limits. See, for example, Le Cam [(1986), Chapter 8]. That is what we shall do in this section.

We rely on the following assumption:

Assumption H. (c_T) and (r_T) are independent of $\theta \in \Theta$ and there is a $\delta > 0$ such that, for some T_0,

$$\sup_{T \geq T_0,\, \theta \in \Theta} E_\theta |\xi_T(\theta)|^{1+\delta} < \infty. \tag{3.19}$$

Proposition 3.2. *Let Θ be an open interval satisfying Assumption* H. *Then, if $\liminf_{T\to\infty} r_T\sqrt{c_T} < \infty$, we have that $b^2(\theta)$ is Lipschitz (and hence absolutely) continuous on Θ; if $\liminf_{T\to\infty} r_T\sqrt{c_T} = 0$, then $b^2(\theta)$ is constant on Θ.*

Further, assume that $\mathbf{T}$ is such that $d = \lim_{T\to\infty, T\in\mathbf{T}} r_T\sqrt{c_T}$ exists (in $[0,\infty]$) and such that $\psi_{\theta,\mathbf{T}}$ exists for all $\theta \in \Theta$. Then, for almost all $\theta \in \Theta$ (under Lebesgue measure), there is a version of $\psi_{\theta,\mathbf{T}}$ satisfying

$$\psi_{\theta,\mathbf{T}}(z) = d^{-1}k(\theta)z + E\psi_{\theta,\mathbf{T}}(N(0,1)), \tag{3.20}$$

where $\infty \cdot 0 = 0$ and

$$k(\theta) = \begin{cases} b(\theta)^{-3}\dfrac{d}{d\theta}b^2(\theta), & \text{if } \liminf\limits_{T\to\infty} r_T\sqrt{c_T} < \infty, \\ 0, & \text{otherwise}, \end{cases} \tag{3.21}$$

whenever the right-hand side is well defined. If $\psi_{\theta,\mathbf{T}}(N(0,1))$ is continuous in *law as a function of* θ, *then* (3.20) *holds for all* $\theta \in \Theta$ *for which* $k(\theta)$ *is well defined.*

It is, of course, unfortunate that Proposition 3.2 does not give any information about the pointwise form of $\psi_{\theta,\mathbf{T}}$. For this we need some uniform convergence.

PROPOSITION 3.3. *Let* Θ *be an open interval around* θ *satisfying Assumption* H *and suppose that* $k(\theta)$ *is well defined (for this specific* θ). *Assume that*

$$\lim_{T\to\infty}\left[E_{\theta_T}\xi_T(\theta_T) - E_\theta \xi_T(\theta)\right] = 0 \tag{3.22}$$

whenever $\sqrt{c_T}(\theta_T - \theta) = O(1)$. *Then, whenever* $\mathbf{T}$ *is such that* $\psi_{\theta,\mathbf{T}}$ *and* $d = \lim_{T\to\infty, T\in\mathbf{T}} r_T\sqrt{c_T}$ *exist,* $\psi_{\theta,\mathbf{T}}$ *has the form* (3.20).

If Assumption H and the conclusion of Proposition 3.3 hold, then Proposition 3.1 yields that

$$F_T(x;\theta_T) = \Phi(x) + \tfrac{1}{2}c_T^{-1/2}k(\theta)\phi(x) + o_2(r_T). \tag{3.23}$$

Note that the assumptions of this section can be weakened somewhat (particularly in Proposition 3.3), at the expense of greater technical complication.

3.4. *Some technical issues: Describing the* CL(2, 1)EF. The θ's for which

$$E_0 \exp\{\theta u_T - \tfrac{1}{2}\theta^2 I_T\} = 1 \tag{3.24}$$

(E_0 being the expectation operator at $\theta = 0$), may not be the entire real line. Hence (1.1) may not define the family $\{P_\theta\}_{\theta\in\mathbf{R}}$. This is a somewhat complicating factor, both when carrying out maximum likelihood estimation and when trying to say something about local alternatives or bootstrapping distributions. It raises the question of whether a natural extension of the family (1.1) exists even when (3.24) does not hold.

Such an extension should, at least, provide the correct probability distribution for the solution of the stochastic differential equation

$$X_t = X_0 + \theta\int_0^t a_s(X)\,ds + \int_0^t v_s(X)\,ds + \int_0^t \gamma_s(X)\,dB_s, \qquad t \ge 0, \tag{3.25}$$

irrespective of whether (3.24) is true or not. [If (3.24) is satisfied, P_θ is given by (1.1), with u_T and I_T given by (4.2)–(4.3) below; cf. Chapter 7 of Liptser and Shiryayev (1977) and also our discussion in Section 7.1.]

Such an extension does exist, and we shall call such models *continuous local* (2, 1) *exponential families*—CL(2, 1)EF for short. A similar approach has been used by Sørensen (1983).

DEFINITION. A CL(2, 1)EF is a family $\{P_\theta\}_{\theta\in\mathbf{R}}$ of probability distributions (on a filtered probability space) which are mutially absolutely continuous up to (stopping) times τ_N, $N = 1, 2, \dots$, so that

$$\text{(3.26)} \qquad \frac{dP_\theta}{dP_0} \text{ at time } \min(t, \tau_N) = \exp\left\{\theta u_{\min(t,\tau_N)} - \frac{1}{2}\theta^2 I_{\min(t,\tau_N)}\right\},$$

with (u_t, I_t) continuous on $[0, \tau_N]$ for all N and $u_0 = I_0 = 0$. It is required of τ_N to satisfy that if $\tau = \lim_{N\to\infty} \tau_N$, then either $\tau = \infty$ or $I_\tau = \infty$ (or both).

For τ_N one can, for example, take

$$\text{(3.27)} \qquad \tau_N = \inf\{t: I_t \geq N, |l_t| \geq N\}.$$

Families of the type (1.1) are a special case by taking $\tau_N = \infty$ for all N. The CL(2, 1)EF, however, extends to all stochastic differential equation models (3.25) satisfying a unique solution condition, see Section 4.1.

The distribution under P_θ of $\hat{\theta}_T$ given $I_T > 0$ is now defined for all θ provided $P_0(I_T > 0) \neq 0$, see (3.28) below. The condition $P_0(I_T > 0) \neq 0$ is natural because otherwise, for all θ, $dP_\theta/dP_0 = 1$ a.s. at time T [since then $(u_t)_{0\leq t\leq T}$ is a martingale with quadratic variation zero, and hence is zero].

The distribution of $\hat{\theta}_T$ conditional on $I_T > 0$ is obtained by using the following:

$$\text{(3.28)} \qquad \begin{aligned} &P_\theta\big(\hat{\theta}_T \in A \text{ and } I_T > 0\big) \\ &\quad = P_\theta\big(\hat{\theta}_T \in A \text{ and } I_T > 0 \text{ and } \tau \leq T\big) \\ &\qquad + P_\theta\big(\hat{\theta}_T \in A \text{ and } I_T > 0 \text{ and } \tau > T\big) \\ &\quad = \chi\{\theta \in A\} P_\theta(\tau \leq T) \\ &\qquad + \lim_{N\to\infty} P_\theta\big(\hat{\theta}_T \in A \text{ and } I_T > 0 \text{ and } \tau_N > T\big) \\ &\quad = \chi\{\theta \in A\} P_\theta(\tau \leq T) \\ &\qquad + \lim_{N\to\infty} E_0\bigg(\chi\big\{\hat{\theta}_T \in A \text{ and } I_T > 0 \text{ and } \tau_N > T\big\} \\ &\qquad\qquad \times \frac{dP_\theta}{dP_0} \text{ at time } \min(T, \tau_N)\bigg). \end{aligned}$$

Note that $\hat{\theta}_T = \theta$ when $\tau \leq T$ because $l_T(\theta)/I_T \to 0$ as $I_T \to \infty$. This follows from Lepingle (1978), see also Sørensen (1983).

Finally, note that for a CL(2, 1)EF, $(l_t(\theta_T), 0 \leq t \leq T)$ is always an array of P_{θ_T}-local martingales (in the sense of Remark 2.4), with quadratic variation I_T (still using the convention that $I_T = \infty$ when $\tau < T$).

4. The case of stochastic differential equations.

4.1. *Introduction*. In the following, we shall consider the differential equation

$$\text{(4.1)} \quad X_t = X_0 + \theta\int_0^t a_s(X)\, ds + \int_0^t v_s(X)\, ds + \int_0^t \gamma_s(X)\, dB_s, \qquad t \geq 0,$$

where (B_t) is a Wiener process. $a_t(x)$, $v_t(x)$ and $\gamma_t(x)$ are measurable functionals on the space of continuous functions x on $[0, \infty)$, satisfying that $a_t(x)$, $v_t(x)$ and $\gamma_t(x)$ only depend on $(x_s, 0 \leq s \leq t)$. For simplicity we shall write $a_t = a_t(X)$, $v_t = v_t(X)$ and $\gamma_t = \gamma_t(X)$.

For references on the application of such models, see, for example, Malliaris and Brock (1982) (economics/finance) and Viterbi (1966) and Kushner [(1984), Chapters 8–10] (engineering).

The relationship to (2, 1) exponential families is that (4.1) is a CL(2, 1)EF provided there is a unique solution, in a sense to be defined below. The connection is that

$$u_T = \int_0^T a_s \gamma_s^{-2} (dX_s - v_s\, ds) \tag{4.2}$$

[the integral is the stochastic integral over an Itô-process; see Liptser and Shiryayev (1977), Chapter 4] and

$$I_T = \int_0^T a_s^2 \gamma_s^{-2}\, ds. \tag{4.3}$$

The exact definitions of existence and uniqueness (weak uniqueness) of solution of (4.1) are relegated to Section 7.1. The essential content of the concept is that (X_t) is a solution if (4.1) is satisfied from time 0 until *termination time*, which is either time $+\infty$ or the time when I_T becomes $+\infty$, whichever comes first (we do not, for example, permit processes which run off to infinity in finite time and leave behind finite information). The solution is weakly unique if the probability law of the solution (X_t) is unique (up to termination time).

Existence and uniqueness of solution of (4.1) can be proved under Lipschitz and/or integrability assumptions on the coefficients; see Liptser and Shiryayev (1977) or the references mentioned at the beginning of Section 7.1. A formal statement of the CL(2, 1)EF-ness of the differential equation is as follows.

LEMMA. *If* (4.1) *has a solution for one* θ, *it has a solution for all* $\theta \in \mathbf{R}$. *If the solution is weakly unique for one* θ, *it is weakly unique for all* $\theta \in \mathbf{R}$. *In this case, the family of processes defined by* (4.1) *is a CL*(2, 1)*EF, with* u_T *and* I_T *given by* (4.2)–(4.3).

The above lemma generalizes similar results (where *termination time* = $+\infty$); see, for example, Liptser and Shiryayev (1977, 1978), Basawa and Prakasa Rao [(1980), Chapter 9.5], Elliot (1982) or Kutoyants (1984) and also the results of Sørensen (1983).

An important problem when dealing with stochastic differential equations is that it is very hard to get at u_T and I_T, both when observing the process and when simulating it. Usually they can, however, be approximated arbitrarily well by observing/simulating on a grid. This is usually consistent as the grid gets finer, see, for example, Stroock and Varadhan [(1979), chapter 11.2, page 266–272] and Jacod and Shiryaev [(1987), Chapter IX.4b, pages 516–523].

In the following we shall verify existence/uniqueness and the conditions for Proposition 3.1 and find $\psi_{\theta,T}$ for first order stochastic differential equations. The conditions can also be satisfied for higher order linear differential equations, see Mykland (1989), Section I-4.3.

In both examples we have used Markov methods. One can also use mixing conditions to check the assumptions of Proposition 3.1, since these also give central limit theorems. Mixing has previously been shown to yield good expansion results directly [see Goetze and Hipp (1983)].

4.2. *First order differential equations.* Consider the equation

$$dX_t = \theta a(X_t)\,dt + v(X_t)\,dt + \gamma(X_t)\,dB_t, \tag{4.4}$$

and assume that γ is continuous and positive and that a, v, γ and γ^{-1} are locally bounded. We shall state conditions under which $\psi_{\theta,T}$ is affine and on the form given in Section 3.3.

Also assume that

$$\int_0^x \exp\left\{-\int_0^z \frac{2(\theta a(y) + v(y))}{\gamma(y)^2}\,dy\right\} dz \to \begin{cases} +\infty, & \text{as } x \to +\infty, \\ -\infty, & \text{as } x \to -\infty, \end{cases} \tag{4.5}$$

and that

$$\int_{-\infty}^{\infty} \gamma(z)^{-2} \exp\left\{\int_0^z \frac{2(\theta a(y) + v(y))}{\gamma(y)^2}\,dy\right\} dz < \infty. \tag{4.6}$$

Under these conditions (see the beginning of Section 7.2), (4.4) has a unique solution and the process is strong Markov and asymptotically ergodic with limiting distribution given by

$$dm_\theta(z) = D_\theta \gamma(z)^{-2} \exp\left\{\int_0^z \frac{2(\theta a(y) + v(y))}{\gamma(y)^2}\,dy\right\} dz, \tag{4.7}$$

where D_θ is a constant which normalizes m to a probability measure. If we further impose the condition that

$$b(\theta)^2 = \int_{-\infty}^{\infty} a(z)^2 \gamma(z)^{-2}\,dm_\theta(z) < \infty, \tag{4.8}$$

then $I_T/T \to b(\theta)^2$ in probability [see Rogers and Williams (1987), Theorem (53.1), page 300] and hence $\sqrt{I_T}(\hat{\theta}_T - \theta)$ is asymptotically $N(0,1)$ under P_θ. As far as expansions are concerned, we have the following result:

PROPOSITION 4.1. *Assume the above conditions. Set*

$$\begin{aligned} h_\theta(x) = D_\theta^{-1} \exp&\left\{-\int_0^x \frac{2(\theta a(y) + v(y))}{\gamma(y)^2}\,dy\right\} \\ &\times \int_{-\infty}^{x} 2\left[a(z)^2 \gamma(z)^{-2} - b(\theta)^2\right] dm_\theta(z), \end{aligned} \tag{4.9}$$

and assume that for some $\delta \geq 1$,

$$\int_{-\infty}^{\infty} |h_\theta(x)\gamma(x)|^{1+\delta}\, dm_\theta(z) < \infty \tag{4.10}$$

and

$$\int_{-\infty}^{\infty} \left| \int_0^x h_\theta(z)\, dz \right|^{1+\delta} dm_\theta(x) < \infty. \tag{4.11}$$

Also assume that X_0 *is nonrandom. Then the conditions of Proposition* 3.1 *are satisfied in the following way*: $\mathbf{T} = [1, \infty)$, $c_T = T$, $r_T = 1/\sqrt{T}$ *and*

$$\sup_{T \geq 1} E_\theta |\xi_T(\theta)|^{1+\delta} < +\infty. \tag{4.12}$$

Further, $(l_T(\theta)/b(\theta)\sqrt{T}, \xi_T(\theta))$ *converge jointly in law under* P_θ *to a normal distribution with mean* 0 *and covariance matrix*

$$\begin{bmatrix} 1 & -\dfrac{1}{b(\theta)} \displaystyle\int_{-\infty}^{\infty} a(x) h_\theta(x)\, dm_\theta(x) \\ -\dfrac{1}{b(\theta)} \displaystyle\int_{-\infty}^{\infty} a(x) h_\theta(x)\, dm_\theta(x) & \displaystyle\int_{-\infty}^{\infty} h_\theta(x)^2 \gamma(x)^2\, dm_\theta(x) \end{bmatrix}. \tag{4.13}$$

In particular, $\psi_\theta = \psi_{\theta, \mathbf{T}}$ *is well defined and we can take*

$$\psi_\theta(z) = -z b(\theta)^{-3} \int_{-\infty}^{\infty} a(x) h_\theta(x)\, dm_\theta(x). \tag{4.14}$$

A small bit of algebra shows that $\psi_\theta(z) = zk(\theta)$, where k is given by (3.21).

4.3. *Some remarks on nonhomogeneous equations.* To illustrate the fact that (c_T) can have any form so long as it is continuous and increasing and tends to $+\infty$, we shall study time changes in the model

$$dX_t = \theta a(X_t, t)\, dt + v(X_t, t)\, dt + \gamma(X_t, t)\, dB_t. \tag{4.15}$$

Assume that (X_t) follows (4.15) and that $Y_t = X_{A(t)}$, where $A(t)$ is a nonrandom, increasing and differentiable function. It is easy to see that Y_t satisfies

$$dY_t = \theta \tilde{a}(Y_t, t)\, dt + \tilde{v}(Y_t, t)\, dt + \tilde{\gamma}(Y_t, t)\, d\tilde{B}_t, \tag{4.16}$$

where $(\tilde{B}_t)$ is a Brownian motion and where

$$\tilde{a}(x,t) = a(x, A(t)) A'(t), \qquad \tilde{v}(x,t) = v(x, A(t)) A'(t) \tag{4.17}$$

and

$$\tilde{\gamma}(x,t) = \gamma(x, A(t)) \sqrt{A'(t)}\,. \tag{4.18}$$

If I_T and $\tilde{I}_T$ are the informations for (4.15) and (4.16), respectively, then $\tilde{I}_T = I_{A(t)}$. This immediately yields:

PROPOSITION 4.2. *Assume that the model* (4.15) *has a weakly unique solution and satisfies condition* (i) *of Proposition* 3.1 *with some choice of* $\mathbf{T}$, (c_T) *and* (r_T). *Then* (4.16) *satisfies the same conditions with the choices* $A^{-1}(\mathbf{T})$, $(c_{A(t)})$ *and* $(r_{A(t)})$. *If* $\psi_{\theta,\mathbf{T}}$ *exists for the model* (4.15), *then* $\psi_{\theta, A^{-1}(T)}$ *exists for the model* (4.16) *and they are the same.*

We have not investigated when processes of the type (4.15) satisfy the conditions of Proposition 3.1. It seems, however, that if sufficiently many conditions are imposed, the asymptotically homogeneous case can presumably be tackled by methods akin to those used in Takeyama (1985).

5. Proof of Theorems 2.1 and 2.2.

5.1. *Reformulation and truncation.* This is actually the proof of the more general version described in Remark 2.4. Assume for simplicity that $b = 1$. The proof is similar in the general case. For now we assume that $\sigma_T = c_T$. We derive the case $\sigma_T \neq c_T$ in Section 5.4. It is assumed that $T \in \mathbf{T}$.

Truncation will be done with constants $\underline{k}$ and $\bar{k}$ from the statements of the theorems. Truncated processes are given by

$$\tilde{I}_T = \max\bigl(\min(I_T, \bar{k}c_T), \underline{k}c_T\bigr) \tag{5.1}$$

and

$$\tilde{\xi}_T = \frac{1}{r_T}\left(\frac{\tilde{I}_T}{c_T} - 1\right). \tag{5.2}$$

By the Dambis–Dubins–Schwartz theorem [see Dambis (1965), Theorem 7 or Dubins and Schwartz (1965)], there is a filtration $(\mathbf{G}_t^{(T)})$, possibly on an extension of the space $(\Omega^{(T)}, \mathbf{F}^{(T)}, P^{(T)})$, so that there is a $(\mathbf{G}_t^{(T)}, P^{(T)})$-Wiener process $(W_t^{(T)})$ so that $W_{I_t}^{(T)} = l_t^{(T)}$ for $0 \le t \le T$ on $\{I_T < \infty\}$.

If g is as specified by Theorem 2.1,

$$\begin{aligned} &\int g(x)\, dF_T(x) - \int g(x)\, d\Phi(x) \\ &\qquad = E^{(T)}\Bigl[g\bigl(c_T^{-1/2}W_{\tilde{I}_T}^{(T)}\bigr) - g\bigl(c_T^{-1/2}W_{c_T}^{(T)}\bigr)\Bigr] + R_T, \end{aligned} \tag{5.3}$$

where

$$\begin{aligned} R_T = {} & E^{(T)}\Bigl[g\bigl(c_T^{-1/2}W_{I_T}^{(T)}\bigr) - g\bigl(c_T^{-1/2}W_{\tilde{I}_T}^{(T)}\bigr)\Big| I_T < \infty\Bigr] \\ & -\Bigl\{E^{(T)}\Bigl[g\bigl(c_T^{-1/2}W_{\tilde{I}_T}^{(T)}\bigr)\Big| I_T = \infty\Bigr] \\ & -E^{(T)}\Bigl[g\bigl(c_T^{-1/2}W_{\tilde{I}_T}^{(T)}\bigr)\Big| I_T < \infty\Bigr]\Bigr\}P(I_T = \infty). \end{aligned} \tag{5.4}$$

Now,

$$|R_T| \le 2MP^{(T)}\big(\tilde{I}_T \ne I_T \big| I_T < \infty\big) + 2MP^{(T)}(I_T = \infty) \le 4MP^{(T)}\big(\tilde{D}_T\big) = o(r_T) \tag{5.5}$$

by (2.5). Furthermore, we have:

LEMMA 5.1. *Assume that g is twice continuously differentiable and that $|g|$, $|g'|$ and $|g''|$ are bounded by M. Then*

$$\frac{1}{r_T} E^{(T)}\Big[g\big(c_T^{-1/2} W_{\tilde{I}_T}^{(T)}\big) - g\big(c_T^{-1/2} W_{c_T}^{(T)}\big)\Big] = E^{(T)} \Xi_T(g), \tag{5.6}$$

where

$$\Xi_T(g) = \frac{1}{r_T} \int_{c_T}^{\tilde{I}_T} \frac{1}{2} \frac{1}{c_T} g''\big(c_T^{-1/2} W_s^{(T)}\big)\, ds. \tag{5.7}$$

PROOF OF LEMMA 5.1. Since $(W_t^{(T)})$ is a $(\mathbf{G}_t^{(T)})$-Wiener process, Itô's lemma [see, e.g., Theorem 4.4 (page 122) of Liptser and Shiryayev (1977)] applied to $g(\cdot / \sqrt{c_T})$ yields

$$g\big(c_T^{-1/2} W_{\tilde{I}_T}^{(T)}\big) - g\big(c_T^{-1/2} W_{c_T}^{(T)}\big) = r_T \Xi_T(g) + \int_{c_T}^{\tilde{I}_T} c_T^{-1/2} g'\big(c_T^{-1/2} W_s^{(T)}\big)\, dW_s^{(T)}. \tag{5.8}$$

Taking expectations (under $P^{(T)}$), the second term on the right-hand side of (5.8) vanishes. This follows from the optional stopping theorem [see, e.g., Theorem V.11 (page 8) of Dellacherie and Meyer (1980)] since $\tilde{I}_T$ is a $(\mathbf{G}_t)$-stopping time and since

$$E^{(T)} \int_{c_T}^{\tilde{I}_T} \big[(c_T^{-1/2}) g'\big(c_T^{-1/2} W_s^{(T)}\big)\big]^2\, ds \le M^2 \int_{c_T \underline{k}}^{c_T \bar{k}} c_T^{-1}\, ds < \infty. \tag{5.9}$$

Hence Lemma 5.1 is proved. □

5.2. *Limit for $E^{(T)}\Xi_T$.* Define the measure M_T^+ on $\mathbf{R}$ by

$$M_T^+(B) = E^{(T)} \frac{1}{r_T} \int_{c_T}^{\max(c_T, \tilde{I}_T)} \frac{1}{2} \frac{1}{c_T} \chi\big(c_T^{-1/2} W_s^{(T)} \in B\big)\, ds, \tag{5.10}$$

where χ is the indicator function. If M_T^- is defined analogously and $M_T = M_T^+ - M_T^-$, then

$$E^{(T)} \Xi_T(g) = \int_{-\infty}^{\infty} g''(x)\, dM_T(x). \tag{5.11}$$

For an arbitrary continuous and bounded h, set

$$\lambda_s^{(T)} = \tfrac{1}{2} c_T^{-1} h\big(c_T^{-1/2} W_s^{(T)}\big). \tag{5.12}$$

Note that

$$(5.13)\qquad \int h(x)\, dM_T^+(x) = \frac{1}{r_T} E^{(T)} \int_{c_T}^{\max(c_T, \tilde{I}_T)} \lambda_s^{(T)}\, ds.$$

On the set $\{\tilde{I}_T \geq c_T\}$, define ω_T by

$$(5.14)\qquad \omega_T = \inf\left\{ t \geq c_T : \left(\tilde{I}_T - c_T\right)\lambda_t^{(T)} = \int_{c_T}^{\tilde{I}_T} \lambda_s^{(T)}\, ds \right\}.$$

By the mean value theorem, ω_T exists,

$$(5.15)\qquad \omega_T \in \operatorname{int}\left[c_T, \tilde{I}_T\right]$$

and

$$(5.16)\qquad \frac{1}{r_T} \int_{c_T}^{\max(c_T, \tilde{I}_T)} \lambda_s^{(T)}\, ds = \frac{1}{r_T}\left(\tilde{I}_T - c_T\right)^+ \lambda_{\omega_T}^{(T)} = \frac{1}{2}\tilde{\xi}_T^+ h\left(\frac{W_{\omega_T}^{(T)}}{\sqrt{c_T}}\right).$$

Furthermore, by the continuity of $\lambda_s^{(T)}$, ω_T is a random variable (but it is not, in general, a stopping time). $W_{\omega_T}^{(T)}$ is then also a random variable (since the Wiener process is continuous). Thus, the equality in (5.16) is also true for the expected values of each side.

Expressions (5.13) and (5.16) will then give us the needed limit in view of the following result.

Lemma 5.2. *Let ω_T be a random variable, $\omega_T \in \operatorname{int}[c_T, \tilde{I}_T]$ $P^{(T)}$-a.s. Then, whenever the convergence in* (2.3) *holds,*

$$(5.17)\qquad \left(\frac{W_{\omega_T}^{(T)}}{\sqrt{c_T}}, \tilde{\xi}_T\right) \to (Z, \xi) \quad \text{as } T \to \infty \text{ through } \mathbf{T},$$

in law under $P^{(T)}$, where $\tilde{\xi}_T$ is defined by (5.2) *and Z and ξ are defined in Section* 2.1.

Proof of Lemma 5.2. In view of the Burkholder–Davis–Gundy inequality [see, e.g., Theorem (2.34) in Jacod (1979)],

$$(5.18)\qquad E^{(T)}\left[\sup_{s \in \operatorname{int}[c_T, \tilde{I}_T]} \left(\frac{W_s - W_{c_T}}{\sqrt{c_T}}\right)^2\right] \leq k E^{(T)}\left|\frac{\tilde{I}_T}{c_T} - 1\right|,$$

where k is some constant. In view of condition (2.4), this yields that

$$(5.19)\qquad \sup_{s \in \operatorname{int}[c_T, \tilde{I}_T]} \left|\frac{W_s - W_{c_T}}{\sqrt{c_T}}\right| \to 0 \quad \text{as } T \to \infty$$

in probability. The result now follows from $\omega_T \in \operatorname{int}[c_T, \tilde{I}_T]$, the representation $l_T^{(T)} = W_{I_T}$ and from (2.3). □

If we define the finite signed measure M by $M(B) = E\xi\chi(Z \in B)/2$, then (2.4), (5.13), (5.16) and Lemma 5.2 yield that M_T^+ converges weakly to M^+. Similarly, M_T^- converges weakly to M^-. Since M is absolutely continuous w.r.t. Lebesgue measure (the density being $\psi\phi/2$ in the notation of Remark 2.3), the following result follows from Example IV.19 (page 73) in Pollard (1984).

LEMMA 5.3. *Let $T_n \to \infty$ as $n \to \infty$ and let h_n be a sequence of bounded functions which converges to h pointwise and which is equicontinuous a.e. under Lebesgue measure. Then*

$$\int h_n(x)\, dM_{T_n} \to \int h(x)\, dM. \tag{5.20}$$

5.3. *Tying the case $\sigma_T = c_T$ together.* The statement (2.14) with the second and fourth terms equal to zero follows from (5.3), (5.5) and (5.6).

As also follows from (5.3), (5.5) and (5.6), to see (2.12) it is enough to show that, uniformly in $g \in C$, as $T \to \infty$ through $\mathbf{T}$,

$$E^{(T)}\Xi_T(g) \to E\tfrac{1}{2}\xi g''(Z). \tag{5.21}$$

To prove (5.21), it is obviously enough to show that for every sequence $\{T_n, g_n\}$, $T_n \to \infty$ through $\mathbf{T}$ as $n \to \infty$ and $g_n \in C$ for all n, there is a subsequence $\{T_{n_k}, g_{n_k}\}$ so that

$$E^{(T_{n_k})}\Xi_{T_{n_k}}(g_{n_k}) - E\tfrac{1}{2}\xi g''_{n_k}(Z) \to 0 \quad \text{as } k \to \infty. \tag{5.22}$$

In view of $\{g''_n\}$ being bounded, there is a subsequence $\{T_{n_k}, g_{n_k}\}$ and a g so that $g''_{n_k} \to g''$ pointwise and so that $\{T_{n_k}\} \subseteq \mathbf{T}$. This will be our choice of subsequence.

First, note that

$$E^{(T)}\tfrac{1}{2}\xi g''_{n_k}(Z) \to E\tfrac{1}{2}\xi g''(Z) \quad \text{as } k \to \infty. \tag{5.23}$$

This is because the conditions of the theorem imply that ξ is integrable, so (5.23) follows from dominated convergence.

Second, (5.11) and Lemma 5.3 show that if $\{g''_{n_k}\}$ is equicontinuous a.e. (Lebesgue), then as $k \to \infty$,

$$E^{(T_{n_k})}\Xi_{T_{n_k}}(g_{n_k}) \to \tfrac{1}{2}E\xi g''(Z). \tag{5.24}$$

Hence the result is proved for the case $c_T = \sigma_T$.

5.4. *The case $\sigma_T \neq c_T$.* We assume that $b_* = 1$. The general case follows from altering c_T and ξ^*.

A one-step Taylor expansion yields

$$g\left(\frac{l_T^{(T)}}{\sigma_T^{1/2}}\right) - g\left(\frac{l_T^{(T)}}{c_T^{1/2}}\right) = -\frac{1}{2}\left(\frac{\sigma_T}{c_T} - 1\right) z_T^{-3/2}\,\frac{l_T^{(T)}}{c_T^{1/2}} g'(z_T^*), \tag{5.25}$$

where

$$z_T \in \operatorname{int}\left[\frac{\sigma_T}{c_T}, 1\right] \quad \text{and} \quad z_T^* \in \operatorname{int}\left[\frac{l_T^{(T)}}{\sigma_T^{1/2}}, \frac{l_T^{(T)}}{c_T^{1/2}}\right]. \tag{5.26}$$

Set, for some $k_* \in (0, 1)$,

$$D_T^{**} = D_T^* \cap \left\{\frac{\sigma_T}{c_T} \geq k_*\right\}. \tag{5.27}$$

Using the Hölder inequality on (5.25), for $r \geq 1$, we get, where $p^{-1} + q^{-1} = 1$,

$$\begin{aligned}
&\left\|\left[g\left(\frac{l_T^{(T)}}{\sigma_T^{1/2}}\right) - g\left(\frac{l_T^{(T)}}{c_T^{1/2}}\right)\right]\chi_{D_T \cap D_T^{**}}\right\|_r \\
&\quad \leq \frac{1}{2}M\left\|\left(\frac{\sigma_T}{c_T} - 1\right)\chi_{D_T^{**}}\right\|_{rp} k_*^{-3/2}\left\|\frac{l_T^{(T)}}{c_T^{1/2}}\chi_{D_T}\right\|_{rq} \\
&\quad \leq \frac{1}{2}M\left\|\left(\frac{\sigma_T}{c_T} - 1\right)\chi_{D_T^*}\right\|_{rp} k_*^{-3/2}C_{rq}\bar{k}
\end{aligned} \tag{5.28}$$

since $D_T^{**} \subseteq D_T^*$ and by the Burkholder–Davis–Gundy inequality [see Jacod (1979), Theorem (2.34)]. [If $\bar{k} = \infty$, make it less than ∞ by the reasoning used in (5.29) below.] C_{rq} is an absolute constant. Also,

$$\begin{aligned}
P(\tilde{D}_T^{**}) &= P\left(\left\{\frac{\sigma_T}{c_T} < k_*\right\} \cap D_T^*\right) + P(\tilde{D}_T^*) \\
&\leq (1 - k_*)^{-1}\left\|\left(\frac{\sigma_T}{c_T} - 1\right)\chi_{D_T^*}\right\|_1 + P(\tilde{D}_T^*),
\end{aligned} \tag{5.29}$$

by Chebyshev's inequality.

At this point, (5.28)–(5.29) yield, in view of (2.5), (2.9) and (2.10), that (2.14) holds. This is because it holds when $\sigma_T = c_T$ with the second and fourth terms equal to zero (see beginning of Section 5.3). Also, we can conclude that (5.25) is uniformly integrable. The result then follows from (2.8) in the same way as we proved the result in Section 5.3.

6. Proofs for Section 3.

6.1. *Proofs of Proposition* 3.1, *Table* 1 *and* (3.15).

PROOF OF PROPOSITION 3.1. In the conditions for Theorem 2.2, let $\sigma_T = I_T$ and $D_T^* = D_T$. Assume that $\underline{k}$ and $\bar{k}$ are chosen so that $0 < \underline{k} < b^2 < \bar{k} < \infty$. Let $0 \leq \varepsilon < \delta$. Let $\tilde{I}_T$, $W_t^{(T)}$ be as in Section 5.

We need bounds on the quantities in (2.14).

$$\left[E_{\theta_T}\left|\frac{1}{r_T}\left(\frac{I_T}{c_T}-b^2\right)\chi_{D_T}\right|^{1+\varepsilon}\right]^{1/(1+\varepsilon)}$$

$$\leq\left[E_\theta\left|\frac{1}{r_T}\left(\frac{I_T}{c_T}-b^2\right)\chi_{D_T}\right|^{1+\varepsilon}\right.$$

$$\left.\times\exp\left\{(\theta_T-\theta)W_{I_T}-\frac{1}{2}(\theta_T-\theta)^2 I_T\right\}\right]^{1/(1+\varepsilon)}$$

(in view of the form of dP_{θ_T}/dP_θ)

$$\leq\left[E_\theta\left|\frac{1}{r_T}\left(\frac{I_T}{c_T}-b^2\right)\chi_{D_T}\right|^{1+\varepsilon}\right.$$

$$\left.\times\exp\left\{(\theta_T-\theta)W_{\bar{k}c_T}-\frac{1}{2}(\theta_T-\theta)^2\bar{k}c_T\right\}\right]^{1/(1+\varepsilon)}$$

(by the martingale property of $\exp\{(\theta_T-\theta)W_t-(\theta_T-\theta)^2t/2\}$)

$$\leq\left\|\frac{1}{r_T}\left(\frac{I_T}{c_T}-b^2\right)\chi_{D_T}\right\|_{p(1+\varepsilon)}$$

$$\times\left\|\exp\left\{(\theta_T-\theta)W_{\bar{k}c_T}-\frac{1}{2}(\theta_T-\theta)^2\bar{k}c_T\right\}\right\|_q^{1/(1+\varepsilon)}$$

$$(6.1)\qquad\leq\left\|\frac{1}{r_T}\left(\frac{I_T}{c_T}-b^2\right)\chi_{D_T}\right\|_{p(1+\varepsilon)}\left(\frac{q}{q-1}\right)^{1/(1+\varepsilon)}$$

$$\times\exp\left\{\frac{1}{2}\bar{k}\frac{q^2-1}{q(1+\varepsilon)}c_T(\theta_T-\theta)^2\right\},$$

since $W_{\bar{k}c_T}$ is $N(0,\bar{k}c_T)$, where $p^{-1}+q^{-1}=1$ and $\|\cdot\|_p=(E_\theta|\cdot|^p)^{1/p}$.

On the other hand, by the same reasoning as in (6.1),

$$\frac{1}{r_T}P_{\theta_T}(\tilde{D}_T)=\frac{1}{r_T}(1-P_{\theta_T}(D_T))$$

$$\leq\frac{1}{r_T}\left(1-E_\theta\chi_{D_T}\exp\left\{(\theta_T-\theta)W_{\bar{k}c_T}-\frac{1}{2}(\theta_T-\theta)^2\bar{k}c_T\right\}\right)$$

$$\leq m\left\|\frac{1}{r_T}\left(\frac{I_T}{c_T}-b^2\right)\chi_{\tilde{D}_T}\right\|_p\left\|\exp\left\{(\theta_T-\theta)W_{\bar{k}c_T}-\frac{1}{2}(\theta_T-\theta)^2\bar{k}c_T\right\}\right\|_q$$

[by Chebyshev's inequality, with $m^{-1} = \min(\bar{k} - b^2, b^2 - \underline{k})$]

$$(6.2) \qquad \le m \left\| \frac{1}{r_T}\left(\frac{I_T}{c_T} - b^2\right)\chi_{\tilde{D}_T} \right\|_p \left(\frac{q}{q-1}\right)\exp\left\{\frac{1}{2}\bar{k}\frac{q^2-1}{q}c_T(\theta_T - \theta)^2\right\}.$$

Now choose some $\eta > 0$. It is easy to see that there are ε, $0 < \varepsilon \le \delta$, p, q $(p^{-1} + q^{-1} = 1)$ and $\bar{k} > b^2$ so that

$$(6.3) \qquad \bar{k}\frac{q^2-1}{q} = \left(\frac{1+2\delta}{\delta(1+\delta)} + 2\eta\right)b^2$$

and

$$(6.4) \qquad p(1+\varepsilon) = 1 + \delta.$$

Thus, with ε replacing δ in the conditions of Theorem 2.2, the integrability assumptions of this theorem are satisfied and (2.14) implies (3.11). Proposition 3.1 then follows from Theorem 2.2 and Remark 2.4 in view of (6.5)–(6.6) below. □

PROOF OF FORMULA (3.15). It follows from the joint limit law of

$$(6.5) \qquad \left(\frac{l_T^{(T)}(\theta_T)}{b\sqrt{c_T}}, \frac{1}{r_T}\left(\frac{I_T}{c_T} - b^2\right), \frac{dP_{\theta_T}}{dP_\theta}\right),$$

under P_θ, which is that of

$$(6.6) \qquad \left(Z - \alpha, \xi, \exp\left\{\alpha Z - \frac{1}{2}\alpha^2\right\}\right). \qquad □$$

PROOF OF TABLE 1. The cases $s_T = I_T$ and $s_T = c_T$ follow from (2.18) in view of (3.15). As far as the case $s_T = E_{\hat{\theta}_T} I_T$ is concerned, we need slightly stronger conditions than in Proposition 3.1 to satisfy the integrability conditions. It is clear from Propositions 3.2 and 3.3 that Assumption H in Section 3.3 is sufficient, since then

$$(6.7) \qquad P_\theta\left(\left|\hat{\theta}_T - \theta\right| > \varepsilon\right) = o(r_T)$$

(using Proposition 3.1 with c_T instead of I_T). Under (3.15), if ψ has the form (3.16) and if $b\sqrt{c_T}(\theta_T - \theta) \to \alpha$,

$$(6.8) \qquad \frac{1}{r_T}\left(\frac{E_{\theta_T} I_T}{c_T} - b^2\right) \to b^2 E_\theta \psi_{\theta,\alpha}(Z) = b^2[k(\theta)\alpha + \kappa(\theta)]$$

as $T \to \infty$. Thus, by Theorem IV.13 (page 71) in Pollard (1984),

$$(6.9) \qquad \frac{1}{r_T}\left(\frac{s_T}{c_T} - b^2\right) = b^2\left[k(\theta) b\sqrt{c_T}\left(\hat{\theta}_T - \theta\right) + \kappa(\theta)\right] + o_p(1),$$

whence, in terms of Section 2.2, $\xi^* = k(\theta)Z + \kappa(\theta)$. The statement in the table then follows from (2.18). □

6.2. *Proofs for Section* 3.3. Suppose Assumption H.

(i) Let $\tilde{I}_T$ have the same meaning as in (5.1), with $\underline{k} = 0$ and $\bar{k} > \sup_{\theta \in \Theta} b(\theta)^2$. Set

$$\tilde{\xi}_T(\theta) = \frac{1}{r_T}\left(\frac{\tilde{I}_T}{c_T} - b^2(\theta)\right). \tag{6.10}$$

Also do the same construction as in Section 5 to let $(W_t(\theta))$ be the Brownian motion satisfying $W_{I_t}(\theta) = l_t(\theta)$. Analogously to the derivation of Table 1 in the previous section, if $\psi_{\theta,\mathbf{T}}$ exists and $T \to \infty$ through $\mathbf{T}$,

$$E_{\theta_T}\tilde{\xi}_T(\theta) \to b^2(\theta)E_\theta\psi_{\theta,\mathbf{T}}(Z)\exp\left\{\alpha b(\theta)Z - \tfrac{1}{2}\alpha^2 b(\theta)^2\right\} \tag{6.11}$$

if $\sqrt{c_T}(\theta_T - \theta) \to \alpha$ and

$$E_\theta\tilde{\xi}_T(\theta)\frac{W_{\tilde{I}_t}(\theta)}{\sqrt{c_T}} \to b(\theta)^3 E_\theta\psi_{\theta,\mathbf{T}}(Z)Z, \tag{6.12}$$

with the left-hand sides bounded by a constant which only depends on $\sup_{\theta \in \Theta, T} E|\xi_T(\theta)|^{1+\delta}$ and δ.

If $\theta, \theta_0 \in \Theta$, then by the form of dP_{θ_T}/dP_θ and by Fubini,

$$\begin{aligned}\frac{\partial}{\partial\theta}E_\theta\frac{\tilde{I}_T}{c_T} &= \frac{\partial}{\partial\theta}E_{\theta_0}\frac{\tilde{I}_T}{c_T}\exp\left\{(\theta-\theta_0)W_{\tilde{I}_T}(\theta_0) - \frac{1}{2}(\theta-\theta_0)^2\tilde{I}_T\right\}\\ &= E_{\theta_0}\frac{\tilde{I}_T}{c_T}\left(W_{\tilde{I}_T}(\theta_0) - (\theta-\theta_0)\tilde{I}_T\right)\\ &\qquad\times\exp\left\{(\theta-\theta_0)W_{\tilde{I}_T}(\theta_0) - \frac{1}{2}(\theta-\theta_0)^2\tilde{I}_T\right\}\\ &= r_T\sqrt{c_T}E_\theta\tilde{\xi}_T(\theta)\frac{W_{\tilde{I}_T}(\theta)}{\sqrt{c_T}}.\end{aligned} \tag{6.13}$$

In other words, if $\theta_0, \theta_1 \in \Theta$,

$$E_{\theta_1}\frac{\tilde{I}_T}{c_T} - E_{\theta_0}\frac{\tilde{I}_T}{c_T} = r_T\sqrt{c_T}\int_{\theta_0}^{\theta_1}E_\theta\tilde{\xi}_T(\theta)\frac{W_{\tilde{I}_T}(\theta)}{\sqrt{c_T}}\,d\theta. \tag{6.14}$$

If $\liminf r_T\sqrt{c_T} < \infty$, there is a sequence $\{T_n\}$ such that $d = \lim_{n\to\infty} r_{T_n}\sqrt{c_{T_n}}$ is finite and such that $\psi_{\theta,(T_n)}$ exists for all $\theta \in \Theta$ (by Tychonoff's product compactness theorem). Since the l.h.s. of (6.14) converges to $b^2(\theta_1) - b^2(\theta_0)$, (6.12) and (6.14) yield

$$b^2(\theta_1) - b^2(\theta_0) = d\int_{\theta_0}^{\theta_1}b(\theta)^3E_\theta\psi_{\theta,(T_n)}(Z)Z\,d\theta. \tag{6.15}$$

This is because the convergence in (6.12) is dominated. As it is, in fact, dominated by a function of θ which is constant on Θ, it also follows that $b^2(\cdot)$ is Lipschitz continuous. If $d = 0$ (which can be arranged for some subsequence if $\liminf r_T\sqrt{c_T} = 0$), $b^2(\cdot)$ is constant. Hence the first part of Proposition 3.2 is proved.

(ii) Now

$$\begin{aligned}\sup_{\theta\in\Theta} E_\theta|\xi_T(\theta) - \tilde{\xi}_T(\theta)| \\ \text{(6.16)} \qquad \le \sup_{\theta\in\Theta} E_\theta|\xi_T(\theta)|\chi\{|\xi_T(\theta)| \ge r_T^{-1}\cdot \text{constant}\} \\ \to 0 \quad \text{as } T\to\infty,\end{aligned}$$

since uniform integrability follows from Assumption H. Combining (6.16) with (6.11) yields that if $\sqrt{c_T}(\theta_T - \theta) \to \alpha$ for some $\theta \in \Theta$ and some $\alpha \in \mathbf{R}$ and if $\psi_{\theta,\mathbf{T}}$ exists, then

$$\text{(6.17)} \qquad \begin{aligned}&\lim_{T\to\infty, T\in\mathbf{T}} \left[E_{\theta_T}\xi_T(\theta) - E_\theta\xi_T(\theta)\right] \\ &\quad = b^2(\theta)E_\theta\psi_{\theta,\mathbf{T}}(Z)\left[\exp\{\alpha b(\theta)Z - \tfrac{1}{2}\alpha^2b^2(\theta)\} - 1\right].\end{aligned}$$

On the other hand, if $k(\theta)$ is well defined and $d = \lim_{T\to\infty, T\in\mathbf{T}} r_T\sqrt{c_T}$ exists, then

$$\text{(6.18)} \qquad \frac{b^2(\theta_T) - b^2(\theta)}{r_T} \to \begin{cases}\alpha d^{-1}\dfrac{d}{d\theta}b^2(\theta), & \text{if } \liminf_{T\to\infty} r_T\sqrt{c_T} < \infty,\\ 0, & \text{otherwise},\end{cases}$$

still with the convention that $0\cdot\infty = 0$. Hence, under the conditions which lead to (6.17) and (6.18),

$$\text{(6.19)} \qquad \lim_{T\to\infty, T\in\mathbf{T}} \left[E_{\theta_T}\xi_T(\theta_T) - E_\theta\xi_T(\theta)\right] = f(\theta,\alpha),$$

where

$$\begin{aligned}f(\theta,\alpha) &= b^2(\theta)E_\theta\psi_{\theta,\mathbf{T}}(Z)\left[\exp\{\alpha b(\theta)Z - \tfrac{1}{2}\alpha^2b^2(\theta)\} - 1\right] - \alpha d^{-1}b^3(\theta)k(\theta) \\ \text{(6.20)} \quad &= b^2(\theta)E_\theta\left[\psi_{\theta,\mathbf{T}}(Z) - E_\theta\psi_{\theta,\mathbf{T}}(Z) - d^{-1}k(\theta)Z\right] \\ &\qquad\times\exp\{\alpha b(\theta)Z - \tfrac{1}{2}\alpha^2b^2(\theta)\},\end{aligned}$$

since $EZ\exp(\alpha bZ - (\alpha b)^2/2) = \alpha b$. Obviously, if $f(\theta,\alpha) = 0$ for almost all α (under Legesgue measure), then $\psi_{\theta,\mathbf{T}}$ has the form (3.20).

(iii) If $\mathbf{T}$ is as described in the statement of Proposition 3.2, then (6.19) holds whenever $f(\theta,\alpha)$ is defined, that is, almost everywhere [in view of (ii)]. On the other hand, Assumption H yields that, by dominated convergence,

$$\text{(6.21)} \qquad \int_{\theta_0}^{\theta_1}\left[E_{\theta+\alpha c_T^{-1/2}}\xi_T(\theta + \alpha c_T^{-1/2}) - E_\theta\xi_T(\theta)\right]d\theta \to 0$$

as $T\to\infty$ through $\mathbf{T}$, for $\theta_0, \theta_1 \in \Theta$. Combining this with (6.19) gives that, for all α,

$$\text{(6.22)} \qquad \int_{\theta_0}^{\theta_1} f(\theta,\alpha)\,d\theta = 0.$$

From this, $f(\theta, \alpha) = 0$ for almost all $\alpha \in \mathbf{R}$ for almost all $\theta \in \Theta$, whence (3.20) follows for almost all $\theta \in \Theta$. Since

$$\int_{\theta_0}^{\theta_1} f(\theta, \alpha)\, d\theta$$

$$= -\alpha B(\theta_1) + \int_{\theta_0}^{\theta_1} b^2(\theta) E_\theta \psi_{\theta,\mathbf{T}}(Z)\left[\exp\{\alpha b(\theta) Z - \tfrac{1}{2}\alpha^2 b^2(\theta)\} - 1\right],$$

with

$$B(\theta_1) = \begin{cases} b^2(\theta_1) - b^2(\theta_0), & \text{if } \liminf_{T\to\infty} r_T\sqrt{c_T} < \infty, \\ 0, & \text{otherwise}, \end{cases}$$

continuity in law of $\psi_{\theta,\mathbf{T}}(Z)$ and Assumption H combine with (6.22) to yield $f(\theta, \alpha) = 0$ for all $\alpha \in \mathbf{R}$ for all $\theta \in \Theta$ for which $k(\theta)$ is well defined, whence the last statement of Proposition 3.2 is proved.

(iv) If $\mathbf{T}$ is as described in Proposition 3.3, (3.22) combined with (6.19) shows that (3.20) is satisfied whenever $k(\theta)$ is defined. □

7. Proofs concerning stochastic differential equations.

7.1. *Comments and proof for Section* 4.1. The most convenient way of looking at (4.1) is to consider it as a martingale problem on a canonical space, see, for example, Jacod (1979), Stroock and Varadhan (1979), Jacod and Mémin (1981), Lebedev (1983). In this setup, Ω is taken to be the set of functions x from subsets of $[0, \infty)$ into $\mathbf{R}$ for which there is an $S \in (0, \infty]$ so that (i) x is continuous on $[0, S)$; and (ii) $S = +\infty$, or $\lim_{t\to S^-} x_t$ either does not exist or is infinite.

Since there can be only one such S for each x, set $S = S(x)$. $(\mathbf{F}_t)$ is the smallest right continuous filtration with respect to which the identity process (x_t) is adapted for $x \in \Omega$ and $\mathbf{F}_\infty = \bigvee_t \mathbf{F}_t$, the smallest σ-field containing all the $\mathbf{F}_t$. Nothing is lost by this assumption, for there is a one-to-one correspondence between this canonical system and the relevant part of any other filtered measurable space [see, e.g., Jacod (1979), Chapter IV-4-a]. For all $x \in \Omega$, set, for $0 \le t < S(x)$,

$$m_t^\theta(x) = x_t - \theta \int_0^t a_s(x)\, ds - \int_0^t v_s(x)\, ds, \tag{7.1}$$

$$\sigma_t(x) = \int_0^t \gamma_s(x)^2\, ds, \tag{7.2}$$

$$I_t(x) = \int_0^t a_s(x)^2 \gamma_s(x)^{-2}\, ds \tag{7.3}$$

[where $a_s(x)^2\gamma_s(x)^{-2}$ is set to 0 if both a and γ are 0 and to $+\infty$ if γ is 0 and a is not]. Also set (inf of empty set being equal to ∞)

$$\tau(x) = \inf\{s \le S(x): I_s(x) = \infty\}, \tag{7.4}$$

$$A_t(x) = \inf\{s \le S(x): I_s(x) > t\}, \tag{7.5}$$

$$\mathbf{G}_t = \mathbf{F}_{A_t} \tag{7.6}$$

and

$$\mathbf{G}_\infty = \mathbf{F}_{\tau-}. \tag{7.7}$$

S, A_t and τ are stopping times; see Dellacherie [(1972), Chapter III.3 (pages 52–57)] or Rogers and Williams [(1987), Chapter VI.17 (pages 343–346)] for the definition of (7.6)–(7.7). Note that the definitions above are made without any reference to any probability, and that, except for m_t^θ, the objects are always well defined. Also note that

$$\tau(x) = \lim_{t\to\infty} A_t(x), \tag{7.8}$$

whence [see Theorem T35 (page 55) of Dellacherie (1972) or Lemma 17.9 (page 345) of Rogers and Williams (1987)]

$$\mathbf{F}_{\tau-} = \vee_t \mathbf{F}_{A_t}. \tag{7.9}$$

The equation (4.1) is defined by θ, the functionals a_t, v_t and γ_t and by a distribution for X_0. A solution (weak solution) of this equation under θ is a probability measure P_θ on $\mathbf{F}_{\tau-}$ (or, equivalently, on $\mathbf{F}_\infty$) which satisfies that (i) $S(x) \ge \tau(x)$ P_θ-a.s.; (ii) x_0 has the prespecified distribution under P_θ; and (iii) (m_t^θ) exists and is a local martingale under P_θ with quadratic variation σ_t for $0 \le t < \tau$. The solution is weakly unique if two solutions P_θ and P'_θ must coincide on $\mathbf{F}_{\tau-}$.

Proof of Lemma: Existence. Let (4.1) have a solution P_θ for θ. We shall show that it has one for $\tilde\theta$.

Let $l_t(\theta) = u_t - \theta I_t$. Set, for all $T \in \mathbf{R}^+$,

$$l_t^{(T)} = l_{\min(t, A_T)}(\theta) \tag{7.10}$$

and

$$f_T = \exp\Big\{(\tilde\theta - \theta) l_{A_T}(\theta) - \tfrac{1}{2}(\tilde\theta - \theta)^2 I_{A_T}\Big\}. \tag{7.11}$$

Since $I_{A_T} \le T$ (since I_T does not jump), $E_\theta f_T = 1$, and we can define on $\mathbf{F}_{A_T}$,

$$dQ_T = f_T\, dP_\theta. \tag{7.12}$$

By Theorem 1.1.9 (page 17) in Stroock and Varadhan (1979) [condition (1.1.9) in that book is verified in roughly the same fashion as in the first half of the proof of Stroock and Varadhan's Theorem 1.3.5 (page 34)], there is a unique probability measure Q on $\mathbf{F}_{\tau-}$ so that

$$Q(E) = Q_T(E) \tag{7.13}$$

for all $E \in \mathbf{F}_{A_T}$.

Since P_θ and Q are equivalent on $\mathbf{F}_{A_T}$, $S(x) \geq A_T(x)$ Q-a.s., and hence $S(x) \geq \tau(x)$ Q-a.s. Also, the generalized Girsanov's theorem [see Jacod and Mémin (1976) or Chapter IV.38 (pages 79–83) of Rogers and Williams (1987)] yields that $m_t^{\tilde{\theta}}$ is a local martingale under Q_T for $0 \leq t \leq A_T$ with quadratic variation σ_t. By (7.8), this also holds for $0 \leq t < \tau$. Since P_θ and Q coincide on $\mathbf{F}_0$, this proves that Q is a solution for $\tilde{\theta}$. □

Proof of Lemma: Uniqueness. Assume that the Q above is the unique solution for $\tilde{\theta}$. By the construction of Q, any solution P_θ for θ is mutually absolutely continuous with Q on $\mathbf{F}_{A_T}$, with

$$\frac{dQ}{dP_\theta} = f_T \quad \text{on } \mathbf{F}_{A_T}. \tag{7.14}$$

Hence if P_θ and P'_θ are two solutions for θ, they must coincide on $\mathbf{F}_{A_T}$. Since this holds for all $T \in \mathbf{R}^+$, they must coincide on $\mathbf{F}_{\tau-}$ by (7.9) and the uniqueness part of Caratheodory's extension theorem. This proves weak uniqueness. □

7.2. *Comments and proof for Section* 4.2. The solution to (4.4) exists and is unique and asymptotically ergodic for the following reason. Set $G_n = [0, +\infty) \times (-n, n)$ and use the construction in (10.1.5)–(10.1.6) (page 250) in Stroock and Varadhan (1979) along with Theorem 7.2.1 (page 187) and Corollary 10.1.2 (page 250) in the same work to establish that the solution of (4.4) is unique provided it exists (in the more usual sense of Stroock and Varadhan (1979), which includes not running off to infinity). On the other hand, Theorem I.16 (page 46) of Skorokhod (1989) shows that the process does indeed exist (in this sense) and that it is asymptotically ergodic. By Theorem 10.1.1 (page 249) of Stroock and Varadhan (1979), the solution is also strong Markov in view of its uniqueness in law.

Proof of Proposition 4.1. Set $V(x) = \int_0^x h_\theta(z)\,dz$. In view of Itô's formula,

$$\begin{aligned} V(X_t) &- V(X_0) \\ &= \left[\int_0^t a(X_s)^2 \gamma(X_s)^{-2}\,ds - tb(\theta)^2\right] + \int_0^t h_\theta(X_s)\gamma(X_s)\,dB_s \\ &= t\left(\frac{I_t}{t} - b(\theta)^2\right) + \int_0^t h_\theta(X_s)\gamma(X_s)\,dB_s. \end{aligned} \tag{7.15}$$

Hence, since $\delta \geq 1$ and by the Burkholder–Davis–Gundy inequality,

$$E_{\theta, X_0 = x_0}|\xi_t(\theta)|^{1+\delta}$$

$$\leq \text{constant} \times \left[|V(x_0)|^{1+\delta} + E_{\theta, X_0=x_0}|V(X_t)|^{1+\delta} + E_{\theta, X_0=x_0}\frac{1}{t}\int_0^t |h_\theta(X_s)\gamma(X_s)|^{1+\delta}\,ds \right] \quad (7.16)$$

for $t \geq 1$. In view of (4.10)–(4.11) and Fatou's lemma, the integrands in the last two terms on the r.h.s. of (7.16) are uniformly integrable for m_θ-almost all (and hence Lebesgue-almost all) x_0. Hence the last two terms are bounded for almost all x_0. However, the local boundedness of V and $h_\theta \gamma$ implies [in view of lemma (46.1), page 273, of Rogers and Williams (1987)] that this extends to all $x_0 \in \mathbf{R}$. This yields the desired result (4.12). Furthermore, since $\int h_\theta(x)^2\gamma(x)^2\,dm_\theta(x) < \infty$, and in view of (4.8),

$$\left[\frac{1}{\sqrt{t}}\int_0^t a(X_s)\gamma(X_s)^{-1}\,dB_s, \frac{1}{\sqrt{t}}\int_0^t h_\theta(X_s)\gamma(X_s)\,dB_s \right] \quad (7.17)$$

converge jointly in law [combining Theorem (53.1) of Rogers and Williams (1987) and Appendix 2.2 of Basawa and Prakasa Rao (1980)] to a normal distribution with mean 0 and covariance matrix

$$\begin{bmatrix} b(\theta)^2 & \int_{-\infty}^{\infty} a(x)h_\theta(x)\,dm_\theta(x) \\ \int_{-\infty}^{\infty} a(x)h_\theta(x)\,dm_\theta(x) & \int_{-\infty}^{\infty} h_\theta(x)^2\gamma(x)^2\,dm_\theta(x) \end{bmatrix}. \quad (7.18)$$

Since $I_t/t \to b(\theta)^2$, the result (4.13) follows. (4.14) is then immediate. □

Acknowledgments. The author would like to thank Rudolf Beran (who was my Ph.D. advisor when this paper was originally written), David Aldous, Lucien Le Cam, Nicholas Jewell, P. Warwick Millar, Jim Pitman, Stephen Stigler, Wing Wong and two referees for invaluable help and advice.

REFERENCES

ABRAMOVITCH, L. and SINGH, K. (1985). Edgeworth corrected pivotal statistics and the bootstrap. *Ann. Statist.* **13** 116–132.

BARNDORFF-NIELSEN, O. E. (1980). Conditionality resolutions. *Biometrika* **67** 293–310.

BARNDORFF-NIELSEN, O. E. (1983). On a formula for the distribution of the maximum likelihood estimator. *Biometrika* **70** 343–365.

BARNDORFF-NIELSEN, O. E. (1984). On conditionality resolution and the likelihood ratio for curved exponential families. *Scand. J. Statist.* **11** 157–170.

BARNDORFF-NIELSEN, O. E. (1986a). Inference on full or partial parameters, based on the standardized log likelihood ratio. *Biometrika* **73** 307–322.

BARNDORFF-NIELSEN, O. E. (1986b). Likelihood and observed geometries. *Ann. Statist.* **14** 856–873.

BARNDORFF-NIELSEN, O. E. (1988). Parametric statistical models and likelihood. *Lecture Notes in Statist.* **50**. Springer, New York.

BASAWA, I. V. and PRAKASA RAO, B. L. S. (1980). *Statistical Inference for Stochastic Processes*. Academic, New York.

BERAN, R. (1982). Estimated sampling distributions: The bootstrap and its competitors. *Ann. Statist.* **10** 212–225.

BERAN, R. (1987). Prepivoting to reduce level error of confidence sets. *Biometrika* **74** 457–468.

BERAN, R. (1988a). Prepivoting test statistics. *J. Amer. Statist. Assoc.* **83** 687–697.

BERAN, R. (1988b). Discussion of "Theoretical comparison of bootstrap confidence intervals" by P. Hall. *Ann. Statist.* **16** 956–959.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 435–451.

BHATTACHARYA, R. N. and RANGA RAO, R. (1976). *Normal Approximation and Asymptotic Expansions*. Wiley, New York.

BHATTACHARYA, R. N. and RANGA RAO, R. (1986). *Normal Approximation and Asymptotic Expansions*. R. E. Krieger, Malabar, Fla.

BHATTACHARYA, R. N. and QUMSIYEH, M. (1989). Second order and L^p-comparisons between the bootstrap and empirical Edgeworth methodologies. *Ann. Statist.* **17** 160–169.

BOLTHAUSEN, E. (1982). Exact convergence rates in some martingale central limit theorems. *Ann. Probab.* **10** 672–688.

BOSE, A. (1986a). Berry–Esseen bound for the maximum likelihood estimator in the Ornstein–Uhlenbeck process. *Sankhyā Ser. A* **48** 181–187.

BOSE, A. (1986b). Certain non-uniform rates of convergence to normality for martingale differences. *J. Statist. Plann. Inference* **14** 155–168.

BOSE, A. (1986c). Asymptotic study of estimators in some discrete and continuous time models. Ph.D. dissertation, Indian Statistical Inst., Calcutta.

BOSE, A. (1987). Bootstrap in moving average models. Technical Report 87-55, Purdue Univ.

BOSE, A. (1988). Edgeworth correction by bootstrap in autoregressions. *Ann. Statist.* **16** 1709–1722.

COX, D. R. (1980). Local ancillarity. *Biometrika* **67** 279–286.

DAMBIS, K. (1965). On the decomposition of continuous sub-martingales. *Theory Probab. Appl.* **10** 401–410.

DELLACHERIE, C. (1972). *Capacités et processus stochastiques*. Springer, New York.

DELLACHERIE, C. and MEYER, P.-A. (1980). *Probabilités et potentiel, Ch. V–VIII. Théorie des martingales*. Hermann, Paris.

DUBINS, L. E. and SCHWARTZ, G. (1965). On continuous martingales. *Proc. Nat. Acad. Sci. U.S.A.* **53** 913–916.

EFRON, B. (1987). Better bootstrap confidence intervals. *J. Amer. Statist. Assoc.* **82** 171–185.

ELLIOT, R. J. (1982). *Stochastic Calculus and Applications*. Springer, New York.

GOETZE, F. and HIPP, C. (1978). Asymptotic expansions in the central limit theorem under moment conditions. *Z. Wahrsch. Verw. Gebiete* **42** 67–87.

GOETZE, F. and HIPP, C. (1983). Asymptotic expansions for sums of weakly dependent random vectors. *Z. Wahrsch. Verw. Gebiete* **64** 211–239.

HAEUSLER, E. (1988). On the rate of convergence in the central limit theorem for martingales with discrete and continuous time. *Ann. Probab.* **16** 275–299.

HALL, P. (1983). Inverting an Edgeworth expansion. *Ann. Statist.* **11** 569–576.

HALL, P. (1986a). On the bootstrap and confidence intervals. *Ann. Statist.* **14** 1431–1452.

HALL, P. (1986b). On the number of bootstrap simulations required to construct a confidence interval. *Ann. Statist.* **14** 1453–1462.

HALL, P. (1988). Theoretical comparison of bootstrap confidence intervals (with discussion). *Ann. Statist.* **16** 927–985.

HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Application*. Academic, New York.

HELLAND, I. S. (1982). Central limit theorems in martingales with discrete or continuous time. *Scand. J. Statist.* **9** 79–94.

HINKLEY, D. V. (1980). Likelihood as approximate pivotal distribution. *Biometrika* **67** 287–292.

JACOD, J. (1979). *Calcul Stochastique et Problèmes de Martingales*. *Lecture Notes in Math*. **714**. Springer, New York.

JACOD, J. and MÉMIN, J. (1976). Caractéristiques locales et conditions de continuité absolue pour les semimartingales. *Z. Wahrsch. Verw. Gebiete* **35** 1–37.

JACOD, J. and MÉMIN, J. (1981). Existence of weak solutions for stochastic differential equations with driving semimartingales. *Stochastics* **4** 317–337.

JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer, New York.

JEGANATHAN, P. (1982). A solution of the martingale central limit problem, parts I–II. *Sankhyā Ser. A* **44** 299–318, 319–340.

JENSEN, J. L. (1986). Asymptotic expansions for sums of dependent variables. Memoir 10, Dept. Theoretical Statistics, Univ. Aarhus.

JENSEN, J. L. (1989). Asymptotic expansions for strongly mixing Harris recurrent Markov chains. *Scand. J. Statist.* **16** 47–64.

KUSHNER, H. J. (1984). *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*. MIT Press.

KUTOYANTS, YU. A. (1984). *Parameter Estimation for Stochastic Processes*. Heldermann, Berlin.

LEBEDEV, V. A. (1983). On the existence of weak solutions for stochastic differential equations with driving martingales and random measures. *Stochastics* **9** 37–76.

LE CAM, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer, New York.

LEPINGLE, D. (1978). Sur le comportement asymptotique des martingales locales. *Lecture Notes in Math*. **649** 148–161. Springer, New York.

LIPTSER, R. S., and SHIRYAYEV, A. N. (1977). *Statistics of Random Processes I—General Theory*. Springer, New York.

LIPTSER, R. S. and SHIRYAYEV, A. N. (1978). *Statistics of Random Processes II—Applications*. Springer, New York.

LIPTSER, R. S. and SHIRYAEV, A. N. (1982). On the rate of convergence in the central limit theorem for semimartingales. *Theory Probab. Appl.* **27** 1–13.

LIU, R. Y. and SINGH, K. (1987). On a partial correction by the bootstrap. *Ann. Statist.* **15** 1713–1718.

MALINOVSKII, V. K. (1987). Limit theorems for Harris Markov chains. *Theory Probab. Appl.* **31** 269–285.

MALLIARIS, A. G. and BROCK, W. A. (1982). *Stochastic Methods in Economics and Finance*. North-Holland, Amsterdam.

MCCULLAGH, P. (1984). Local sufficiency. *Biometrika* **71** 233–244.

MCCULLAGH, P. (1987). *Tensor Methods in Statistics*. Chapman and Hall, London.

MCCULLAGH, P. and COX, D. R. (1986). Invariants and likelihood ratio statistics. *Ann. Statist.* **14** 1419–1430.

MISHRA, M. N. and PRAKASA RAO, B. L. S. (1985). On the Berry–Esseen bound for the maximum likelihood estimator for linear homogeneous diffusion processes. *Sankhyā Ser. A* **47** 392–398.

MYKLAND, P. A. (1989). Bootstrap and Edgeworth methods for dependent variables. Ph.D. dissertation. Univ. California, Berkeley.

POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.

REBOLLEDO, R. (1980). Central limit theorems for local martingales. *Z. Wahrsch. Verw. Gebiete* **51** 269–286.

ROGERS, L. C. G. and WILLIAMS, D. (1987). *Diffusions, Markov Processes, and Martingales*. **2**: *Itô Calculus*. Wiley, New York.

SINGH, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** 1187–1195.

SKOROKHOD, A. V. (1989). *Asymptotic Methods in the Theory of Stochastic Differential Equations*. Amer. Math. Soc., Providence, R. I.

SØRENSEN, M. (1983). On maximum likelihood estimation in randomly stopped diffusion-type processes. *Internat. Statist. Rev.* **51** 93–110.

STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes*. Springer, New York.

TAKEYAMA, O. (1985). Asymptotic properties of asymptotically homogeneous diffusion processes on a compact manifold. *J. Math. Soc. Japan* **37** 637–650.

TANAKA, K. (1986). Asymptotic expansions for time series statistics. In *Essays in Time Series and Allied Processes: Papers in Honour of E. J. Hannan* (J. Gani and M. B. Priestly, eds.) 211–227. Applied Probability Trust, Sheffield.

TANIGUCHI, M. (1984). Validity of Edgeworth expansions for statistics of time series. *J. Time Series Anal.* **5** 37–51.

VITERBI, A. J. (1966). *Principles of Coherent Communication*. McGraw-Hill, New York.

WITHERS, C. S. (1983). Expansions for the distribution and quantiles of a regular functional of the empirical distribution with applications to nonparametric confidence intervals. *Ann. Statist.* **11** 577–587.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
5734 UNIVERSITY AVENUE
CHICAGO, ILLINOIS 60637