The Annals of Statistics
1996, Vol. 24, No. 3, 1025–1052

ADAPTIVE ESTIMATION IN A RANDOM COEFFICIENT AUTOREGRESSIVE MODEL

By Hira L. Koul[1] and Anton Schick[2]

Michigan State University and State University of New York at Binghamton

This paper proves the local asymptotic normality of a stationary and ergodic first order random coefficient autoregressive model in a semiparametric setting. This result is used to show that Stein's necessary condition for adaptive estimation of the mean of the random coefficient is satisfied if the distributions of the innovations and the errors in the random coefficients are symmetric around zero. Under these symmetry assumptions, a locally asymptotically minimax adaptive estimator of the mean of the random coefficient is constructed. The paper also proves the asymptotic normality of generalized M-estimators of the parameter of interest. These estimators are used as preliminary estimators in the above construction.

1. Introduction. The construction of estimators that are asymptotically efficient in the presence of infinite dimensional nuisance parameters has been the focus of numerous researchers in the last three decades. For example, see the recent monograph by Bickel, Klaassen, Ritov and Wellner (1993) and the references therein. The present paper is concerned with the construction of such estimators in the first order random coefficient autoregression (RCAR) model. In fact, the estimators of this paper are *adaptive* in the sense that they are asymptotically as efficient as if the nuisance parameters were known.

In the RCAR model one observes $X_0, \dots, X_n$, where the sequence $\{X_j: j \in \mathbb{Z}\}$ satisfies

$$X_j = \big(\theta + Z_j\big) X_{j-1} + \varepsilon_j, \qquad j \in \mathbb{Z}, \tag{1.1}$$

for some unknown θ in $\mathbb{R}$, and for independent sequences $\{\varepsilon_j: j \in \mathbb{Z}\}$ and $\{Z_j: j \in \mathbb{Z}\}$ of independent and identically distributed random variables with distribution functions F and G, respectively. Here $\mathbb{Z}$ denotes the set of all integers. The importance of general RCAR models in time series analysis is illustrated in the lecture notes by Nicholls and Quinn (1982) and in the book by Tong (1990).

Received March 1994; revised February 1995.
[1]Research supported in part by NSF Grants DMS-91-02041 and DMS-94-02904.
[2]Research supported in part by NSF Grant DMS-92-06138.
AMS 1991 *subject classifications*. 62G05, 62M10.
Key words and phrases. Locally asymptotically minimax adaptive, semiparametric, stationary, ergodic, generalized M-estimator.

Let $\mathfrak{D}$ denote the class of all distribution functions with zero means and finite variances. Let σ_D^2 denote the variance of a D in $\mathfrak{D}$. Throughout the paper it is assumed that F and G belong to $\mathfrak{D}$ and that

$$\theta^2 + \sigma_G^2 < 1. \tag{1.2}$$

In view of Theorems 2.1 and 2.7 of Nicholls and Quinn (1982), under these assumptions the process $\{X_j: j \in \mathbb{Z}\}$ satisfying (1.1) is strictly stationary and ergodic. It can be constructed so that

$$X_j = \varepsilon_j + \sum_{i=1}^{\infty} \varepsilon_{j-i} \prod_{k=j-i+1}^{j} (\theta + Z_k) \tag{1.3}$$

almost surely and in mean square. We denote the underlying probability measure by $P_{\theta,F,G}$ and the corresponding expectation by $E_{\theta,F,G}$. From (1.3) one obtains that

$$E_{\theta,F,G} X_0 = 0 \quad \text{and} \quad E_{\theta,F,G} X_0^2 = \frac{\sigma_F^2}{1 - \theta^2 - \sigma_G^2}. \tag{1.4}$$

The problem of interest is the construction of adaptive estimators of θ in the presence of the nuisance parameter (F, G). If G is degenerate at 0, then the model (1.1) reduces to the first order autoregression model. Kreiss (1987a) provides adaptive estimates for parameters in ARMA models when the error distribution is symmetric and has finite second moment and finite Fisher information for location, and Kreiss (1987b) constructs adaptive estimates of parameters in AR models without the symmetry assumption. The latter result has been improved and generalized by Schick (1993) to AR models with unknown regression. See also Koul and Pflug (1990) for adaptive estimation in explosive autoregression.

From the general asymptotic theory for adaptive estimation in locally asymptotically normal (LAN) families it follows that adaptive estimation is not always possible. Necessary conditions for adaptive estimation for these families are given by Fabian and Hannan (1982). A general method of constructing adaptive estimates was originally proposed by Bickel (1982). His method has been generalized and improved by Schick (1986, 1987). This theory is developed for the case of independent (and identically distributed) random variables. No general theory for the construction of adaptive estimates for dependent random variables exists at the present stage.

The paper is organized as follows. Section 2 addresses the problem of constructing preliminary estimators of θ. It is shown that generalized M-estimators are $\sqrt{n}$-consistent for θ under fairly general assumptions on (F, G) and the underlying weight and score functions. This class of estimators includes the least squares and the least absolute deviations estimators. Section 3 discusses local asymptotic normality for the above RCAR semiparametric model. Section 4 addresses the question of efficient and adaptive estimation of θ. First, a locally asymptotically minimax (LAM) estimator of θ is given when F and G are known. Then the necessary condition for adaptive estimation is verified for the model when both F and G are symmetric about zero. Finally, an LAM-adaptive estimator of θ is constructed. The Appendix contains technical details that may be of independent interest also.

Throughout this paper, θ, F and G are fixed and $\Theta = \{\vartheta\colon \vartheta^2 + \sigma_G^2 < 1\}$. For convenience, $P_{\vartheta, F, G}$, is abbreviated by P_ϑ and the corresponding expectation is denoted by E_ϑ. For a sequence $\{\theta_n\}$ in Θ and a sequence $\{a_n\}$ of positive numbers, $o_{\theta_n}(a_n)$ $[O_{\theta_n}(a_n)]$ denotes a sequence of random variables $\{X_n\}$ such that $a_n^{-1}X_n$ converges to 0 (is bounded) in P_{θ_n}-probability. The distribution of a random variable X under a probability measure P is denoted by $\mathfrak{L}(X \mid P)$.

2. Generalized M-estimators of $\boldsymbol{\theta}$. This section discusses the asymptotic distributions of a class of generalized M-estimators. These results are of independent interest. In addition, any one of these estimators can be used as a preliminary estimator in the construction of efficient estimators of θ. To define generalized M-estimators, let g be a measurable function from $\mathbb{R}$ to $\mathbb{R}$ such that $xg(x) \geq 0$ for all $x \in \mathbb{R}$, and let ψ be a monotone function from $\mathbb{R}$ to $\mathbb{R}$. Set

$$T_n(t) \coloneqq \frac{1}{\sqrt{n}} \sum_{j=1}^{n} g(X_{j-1})\psi(X_j - tX_{j-1}), \qquad t \in \mathbb{R}.$$

A generalized M-estimator $\hat{\theta}_n$ of θ is defined as a solution of the equation $T_n(t) = 0$.

In this section, it is assumed that

$$\begin{aligned} &E_\theta(\psi(X_1 - \theta X_0) \mid X_0) = 0 \quad \text{and} \\ &0 < \gamma \coloneqq E_\theta g^2(X_0)\psi^2(X_1 - \theta X_0) < \infty. \end{aligned} \tag{2.1}$$

These assumptions imply that $T_n(\theta)$ is a mean zero square integrable martingale which satisfies the conditions of Corollary 3.1 of Hall and Heyde (1980). Thus

$$\mathfrak{L}(T_n(\theta) \mid P_\theta) \Rightarrow N(0, \gamma). \tag{2.2}$$

REMARK 2.1. If ε_1 and Z_1 are symmetrically distributed around zero, then the conditional distribution of $X_1 - \theta X_0 = \varepsilon_1 + Z_1 X_0$, given X_0, is symmetric around zero. Hence, $E_\theta(\psi(X_1 - \theta X_0) \mid X_0) = 0$ for all odd functions ψ with $E_\theta|\psi(X_1 - \theta X_0)| < \infty$.

The asymptotic distribution of $\hat{\theta}_n$ can now be obtained via the following lemma.

LEMMA 2.2. *Suppose* (2.1) *holds and*

$$T_n(\theta + n^{-1/2}t) - T_n(\theta) = -tb + o_p(1) \tag{2.3}$$

for all $t \in \mathbb{R}$ *and some* $b \neq 0$. *Then*

$$\mathfrak{L}\left(\sqrt{n}\left(\hat{\theta}_n - \theta\right) \mid P_\theta\right) \Rightarrow N(0, \gamma b^{-2}).$$

PROOF. By the monotonicity of ψ, the map $t \mapsto T_n(t)$ is monotone. Thus (2.3) implies

$$\sup_{|t| \le c} \left| T_n(\theta + n^{-1/2} t) - T_n(\theta) + tb \right| = o_\theta(1)$$

for every finite c. From this one concludes in a routine fashion that

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) = b^{-1} T_n(\theta) + o_\theta(1).$$

The desired result follows from this and (2.2). □

The following two lemmas give sufficient conditions for (2.3).

LEMMA 2.3. *Suppose ψ is absolutely continuous with a.e.-derivative ψ' satisfying*

$$E_\theta | X_0 g(X_0) \psi'(X_1 - \theta X_0)| < \infty \tag{2.4}$$

and

$$\lim_{s \to 0} E_\theta | X_0 g(X_0)| \, |\psi'(X_1 - \theta X_0 - s X_0) - \psi'(X_1 - \theta X_0)| = 0. \tag{2.5}$$

Then (2.3) *holds with*

$$b = E_\theta X_0 g(X_0) \psi'(X_1 - \theta X_0).$$

PROOF. Let $U_j = X_j - \theta X_{j-1}$, $j = 1, \ldots, n$. By (2.4) and the ergodic theorem,

$$\frac{1}{n} \sum_{j=1}^{n} X_{j-1} g(X_{j-1}) \psi'(U_j) \to b \quad \text{a.s.}$$

Using stationarity, the absolute continuity of ψ, Fubini's theorem and (2.5), we obtain

$$\begin{aligned}
&E_\theta \left| T_n(\theta + n^{-1/2} t) - T_n(\theta) + \frac{t}{n} \sum_{j=1}^{n} X_{j-1} g(X_{j-1}) \psi'(U_j) \right| \\
&\quad \le n^{1/2} E_\theta |g(X_0)| \left| \psi\left(U_1 - \frac{tX_0}{\sqrt{n}}\right) - \psi(U_1) + \frac{tX_0}{\sqrt{n}} \psi'(U_1) \right| \\
&\quad \le |t| \int_0^1 E_\theta |X_0 g(X_0)| \left| \psi'(U_1) - \psi'\left(U_1 - z \frac{tX_0}{\sqrt{n}}\right) \right| dz \to 0.
\end{aligned}$$

Combining the above yields the desired result. □

EXAMPLE 2.4 (Least squares estimator). Let $g(x) = \psi(x) = x$, $x \in \mathbb{R}$. The resulting estimator is the least squares estimator. Assume that $E_\theta X_0^4 < \infty$ and $\sigma_F^2 > 0$. Then (2.1) holds with $\gamma = E_\theta X_0^2 E_\theta \varepsilon_1^2 + E_\theta Z_1^2 E_\theta X_0^4$ and the assumptions of Lemma 2.3 hold with $b = E_\theta X_0^2$. Applying Lemma 2.2 yields $\mathfrak{L}(n^{1/2}(\hat{\theta}_n - \theta) \mid P_\theta) \Rightarrow N(0, \tau)$, where

$$\tau = \frac{\sigma_F^2}{E_\theta X_0^2} + \frac{\sigma_G^2 E_\theta(X_0^4)}{\left(E_\theta(X_0^2)\right)^2}.$$

This is a special case of a result of Nicholls and Quinn (1982).

EXAMPLE 2.5 (Modified least squares estimator). This estimator is obtained upon taking $\psi(x) = x$ and $g(x) = xI[|x| \le c] + c\,\mathrm{sgn}(x) I[|x| > c]$, where c is a known positive constant. If $\sigma_F^2 > 0$, then (2.1) holds with $\gamma = \sigma_F^2 E_\theta g^2(X_0) + \sigma_G^2 E_\theta X_0^2 g^2(X_0)$ and the assumptions of Lemma 2.3 hold with $b = E_\theta X_0 g(X_0)$. Consequently, $\mathfrak{L}(n^{1/2}(\hat{\theta}_n - \theta) \mid P_\theta) \Rightarrow N(0, \tau)$, where

$$\tau = \frac{\sigma_F^2 E_\theta g^2(X_0) + \sigma_G^2 E_\theta X_0^2 g^2(X_0)}{\left(E_\theta X_0 g(X_0)\right)^2}.$$

Unlike the least squares estimator, this estimator does not require the finiteness of the fourth moment of X_0.

EXAMPLE 2.6 (Huber estimator). This estimator is obtained upon taking $g(x) = x$ and $\psi(x) = xI[|x| \le c] + c\,\mathrm{sgn}(x) I[|x| > c]$, $x \in \mathbb{R}$, where c is a known positive constant. Assume that F and G are symmetric around zero, F is continuous and $E_\theta X_0^2[F(c - X_0 Z_1) - F(-c - X_0 Z_1)] > 0$. Then one verifies (2.1), (2.4) and (2.5) and obtains from Lemmas 2.2 and 2.3 that $\mathfrak{L}(n^{1/2}(\hat{\theta}_n - \theta) \mid P_\theta) \Rightarrow N(0, \tau)$, where

$$\tau = \frac{E_\theta X_0^2 \psi^2(X_1 - \theta X_0)}{\left(E_\theta X_0^2[F(c - X_0 Z_1) - F(-c - X_0 Z_1)]\right)^2}.$$

EXAMPLE 2.7 (Arctan estimator). Let $g(x) = x$ and $\psi(x) = \arctan(x)$, $x \in \mathbb{R}$. Assume that F and G are symmetric around zero and that $\sigma_F^2 > 0$. Then one verifies (2.1), (2.4) and (2.5) and obtains from Lemmas 2.2 and 2.3 that $\mathfrak{L}(n^{1/2}(\hat{\theta}_n - \theta) \mid P_\theta) \Rightarrow N(0, \tau)$, where

$$\tau = \frac{E_\theta X_0^2 \arctan^2(\varepsilon_1 + Z_1 X_0)}{\left(E_\theta X_0^2 \arctan'(\varepsilon_1 + Z_1 X_0)\right)^2}.$$

LEMMA 2.8. *Suppose ψ is bounded, F has a bounded and continuous Lebesgue density f and $E_\theta g^2(X_0) < \infty$. Then* (2.3) *holds with*

$$b = E_\theta X_0 g(X_0) \int f(u - Z_1 X_0)\, d\psi(u).$$

PROOF. Without loss of generality assume that ψ is nondecreasing. Fix $t \in \mathbb{R}$. With U_j's as in the previous proof, let

$$D_{n,1} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} g(X_{j-1}) \left(\psi\left(U_j - \frac{tX_{j-1}}{\sqrt{n}} \right) - \psi(U_j) \right),$$

$$D_{n,2} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} g(X_{j-1}) E_\theta \left(\psi\left(U_j - \frac{tX_{j-1}}{\sqrt{n}} \right) - \psi(U_j) \mid X_{j-1} \right),$$

$$D_{n,3} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} g(X_{j-1}) \frac{tX_{j-1}}{\sqrt{n}} \int E_\theta\big(f(u - Z_j X_{j-1}) \mid X_{j-1} \big)\, d\psi(u).$$

Observe that $D_{n,1} - D_{n,2}$ is a mean zero square integrable stationary martingale. Thus,

$$E_\theta (D_{n,1} - D_{n,2})^2 \le E_\theta g^2(X_0) \left(\psi\left(U_1 - \frac{tX_0}{\sqrt{n}} \right) - \psi(U_1) \right)^2.$$

Let

$$\xi_n(u) = F(u - Z_1 X_0) - F\left(u - Z_1 X_0 + \frac{tX_0}{\sqrt{n}} \right), \qquad u \in \mathbb{R}.$$

Using the fact that $\psi(x) - \psi(y) = \int (I[u \le x] - I[u \le y])\, d\psi(u)$ and Fubini's theorem, one obtains from the above inequality that

$$E_\theta (D_{n,1} - D_{n,2})^2 \le 2 \sup_{x \in \mathbb{R}} |\psi(x)| \int E_\theta g^2(X_0) |\xi_n(u)|\, d\psi(u).$$

By the Lebesgue dominated convergence theorem and the continuity of F, this upper bound tends to zero. Hence

$$E_\theta (D_{n,1} - D_{n,2})^2 \to 0. \tag{2.6}$$

Using Fubini's theorem and a conditioning argument together with the underlying independence between ε_i, Z_i and X_{i-1}, we obtain

$$\begin{aligned} &E_\theta |D_{n,2} + D_{n,3}| \\ &\quad \le n^{1/2} \int E_\theta \left(|g(X_0)| \left| \xi_n(u) + \frac{tX_0}{\sqrt{n}} f(u - Z_1 X_0) \right| \right) d\psi(u) \to 0, \end{aligned} \tag{2.7}$$

by the Lebesgue dominated convergence theorem and the assumed properties of f. Finally, by the ergodic theorem,

$$D_{n,3} = tb + o_\theta(1). \tag{2.8}$$

Combining (2.6)–(2.8) yields (2.3). □

EXAMPLE 2.9 (Least absolute deviation estimator). Let $g(x) = x$ and $\psi(x) = \operatorname{sgn}(x)$. The resulting estimator is the least absolute deviation estimator. Assume now that F has a continuous, bounded and even density f, G

is symmetric around zero and $E_\theta(X_0^2 f(Z_1 X_0)) > 0$. Then (2.1) holds and the assumptions of Lemma 2.8 are met. Thus Lemma 2.2 implies that $\mathfrak{L}(n^{1/2}(\hat{\theta}_n - \theta) | P_\theta) \Rightarrow N(0, \tau_1)$, where

$$\tau_1 = \frac{E_\theta\big(X_0^2\big)}{4\big(E_\theta X_0^2 f(Z_1 X_0)\big)^2}.$$

We need the continuity of f everywhere to conclude this result. This is unlike the ordinary autoregression model, where f needs to be continuous at zero only. See, for example, Chapter 7 of the monograph by Koul (1992).

3. Local asymptotic normality. In this section we obtain the LAN condition for the RCAR model when the distributions of ε_1 and Z_1 are allowed to depend on a finite dimensional parameter. This result is basic to the characterization of efficient estimators and the ensuing discussion of adaptation.

From now on we assume that F has finite Fisher information for location, that is, F has an absolutely continuous density f with a.e.-derivative f' and

$$J(f) = \int \left(\frac{f'(x)}{f(x)} \right)^2 dF(x) < \infty. \tag{3.1}$$

Then, under P_ϑ, X_0 has a Lebesgue density q_ϑ and the conditional distribution $X_1 - \vartheta X_0$ given X_0 has a density p given by

$$p(x_1, x_2) = \int f(x_2 - z x_1)\, dG(z), \qquad x_1, x_2 \in \mathbb{R},$$

for $\vartheta \in \Theta$. Consequently, a joint density π_ϑ of $(X_0, X_1 - \vartheta X_0)$ under P_ϑ is given by

$$\pi_\vartheta(x_1, x_2) = q_\vartheta(x_1) p(x_1, x_2), \qquad x_1, x_2 \in \mathbb{R}, \vartheta \in \Theta. \tag{3.2}$$

Define now

$$p'(x_1, x_2) = \int f'(x_2 - z x_1)\, dG(z), \qquad x_1, x_2 \in \mathbb{R},$$

$$L(x_1, x_2) = \frac{p'(x_1, x_2)}{p(x_1, x_2)}, \qquad x_1, x_2 \in \mathbb{R},$$

$$W(\theta) = \int x_1^2 L^2(x_1, x_2) p(x_1, x_2) q_\theta(x_1)\, dx_1\, dx_2$$

and

$$l_j(\vartheta) = -X_{j-1} L\big(X_{j-1}, X_j - \vartheta X_{j-1}\big), \qquad j = 1, \dots, n, \vartheta \in \mathbb{R}.$$

Observe that $\sum_{j=1}^n l_j(\theta)$ and $nW(\theta)$ are the score and the Fisher information for the estimation of θ, respectively, when F and G are known.

An application of the Cauchy–Schwarz inequality and Fubini's theorem shows that

$$\int L^2(x_1, x_2)p(x_1, x_2)\,dx_2 \le J(f), \qquad x_1 \in \mathbb{R}. \tag{3.3}$$

Therefore, $W(\theta) \le J(f)E_\theta X_0^2 < \infty$. It is also easy to check that $W(\theta) > 0$.

We now introduce a parametrization of the distributions of (ε_1, Z_1).

DEFINITION 3.1. By a *path* we mean a map $\eta \mapsto (F_\eta, G_\eta)$ from $(-1,1)$ into $\mathfrak{D} \times \mathfrak{D}$ satisfying $(F_0, G_0) = (F, G)$ and $\theta^2 + \int z^2\,dG_\eta(z) < 1 - \alpha_*$ for all $\eta \in (-1,1)$ and some $\alpha_* > 0$.

Fix a path $\eta \mapsto (F_\eta, G_\eta)$. Then there exist positive numbers α_0 and α_1 such that

$$(\vartheta + t)^2 + \int z^2\,dG_\eta(z) < 1$$

for all $\vartheta \in \Theta_0 = [\theta - \alpha_0, \theta + \alpha_0]$ and $(t, \eta) \in \Delta_0 = (-\alpha_1, \alpha_1) \times (-1, 1)$. For $\vartheta \in \Theta_0$ and $\delta = (t, \eta) \in \Delta_0$, let $Q^n_{\vartheta,\delta}$ denote the restriction of $P_{(\vartheta+t, F_\eta, G_\eta)}$ to $\mathfrak{F}_n$ and let $\Lambda_n(\vartheta, \delta)$ denote the log likelihood of $Q^n_{\vartheta,\delta}$ with respect to $Q^n_{\vartheta,0}$. Here and in the sequel, $\mathfrak{F}_j$ denotes the σ-field generated by $(X_0, \dots, X_j)$, $j = 0, 1, \dots, n$.

DEFINITION 3.2. We say the path $\eta \mapsto (F_\eta, G_\eta)$ is *regular* if the following conditions hold.

(R.1) The map $\eta \mapsto \int x^2\,dF_\eta(x)$ is continuous at 0.

(R.2) For each η, F_η has Lebesgue density f_η and there exists a measurable function ζ from $\mathbb{R}$ to $\mathbb{R}$ such that $\int \zeta^2(x)f(x)\,dx < \infty$ and

$$\int \left(\sqrt{f_\eta(x)} - \sqrt{f(x)} - \frac{\eta}{2}\zeta(x)\sqrt{f(x)}\right)^2 dx = o(\eta^2).$$

(R.3) There exists a measurable function ξ from $(0,1)$ to $\mathbb{R}$ such that $\int_0^1 \xi^2(u)\,du < \infty$ and

$$\int_0^1 \left(G_\eta^{-1}(u) - G^{-1}(u) - \eta\xi(u)\right)^2 du = o(\eta^2),$$

where $G_\eta^{-1}(u) = \inf\{t\colon G_\eta(t) \ge u\}$.

In this case we set, for $\vartheta \in \Theta$,

$$\zeta_j(\vartheta) = \frac{\int \zeta\big(X_j - (\vartheta + z)X_{j-1}\big) f\big(X_j - (\vartheta + z)X_{j-1}\big)\,dG(z)}{\int f\big(X_j - (\vartheta + z)X_{j-1}\big)\,dG(z)}, \tag{3.4}$$

$$\xi_j(\vartheta) = -X_{j-1} \frac{\int_0^1 \xi(u) f'\big(X_j - (\vartheta + G^{-1}(u))X_{j-1}\big)\, du}{\int f\big(X_j - (\vartheta + z)X_{j-1}\big)\, dG(z)}, \tag{3.5}$$

$$S_j(\vartheta) = \begin{pmatrix} l_j(\vartheta) \\ \zeta_j(\vartheta) + \xi_j(\vartheta) \end{pmatrix}, \qquad j = 1, 2, \ldots,$$

and

$$V(\theta) = E_\theta\big[S_1(\theta) S_1^T(\theta)\big].$$

We are now ready to state the LAN result.

THEOREM 3.3. *Suppose the path $\eta \mapsto (F_\eta, G_\eta)$ is regular and the sequence $\{\theta_n\}$ in Θ_0 is such that $\sqrt{n}\,(\theta_n - \theta)$ is bounded. Then*

$$\Lambda_n\left(\theta_n, \frac{1}{\sqrt{n}} u_n\right) - \frac{1}{\sqrt{n}} \sum_{j=1}^n u_n^T S_j(\theta_n) + \frac{1}{2} u_n^T V(\theta) u_n \to 0 \tag{3.6}$$

in $Q^n_{\theta_n, 0}$-probability for every bounded sequence $\{u_n\}$ in $\mathbb{R}^2$ and

$$\mathfrak{L}\left(\frac{1}{\sqrt{n}} \sum_{j=1}^n S_j(\theta_n) \,|\, Q^n_{\theta_n, 0}\right) \Rightarrow N(0, V(\theta)).$$

The proof of this theorem is facilitated by the following two lemmas, proofs of which are given in the Appendix. Let $q_{\vartheta,\delta}$ and $p_{\vartheta,\delta}$, respectively, denote the stationary and transition densities under $Q^n_{\vartheta,\delta}$ if the path is regular. In this case define also a map $\dot{s}_\vartheta$ from $\mathbb{R}^2$ to $\mathbb{R}^2$ such that

$$\dot{s}_\vartheta(X_0, X_1) = S_1(\vartheta)\sqrt{p(X_0, X_1 - \vartheta X_0)} \quad \text{a.s. } P_\vartheta.$$

Observe that $p_{\vartheta,\delta}(X_0, X_1) = p_{0,\delta}(X_0, X_1 - \vartheta X_0)$ and $\dot{s}_\vartheta(X_0, X_1) = \dot{s}_0(X_0, X_1 - \vartheta X_0)$.

LEMMA 3.4. *Suppose the path $\eta \mapsto (F_\eta, G_\eta)$ is regular. Then*

$$\sup_{\vartheta \in \Theta_0} \int (1 + x^2)\Big(\sqrt{q_{\vartheta,\delta}(x)} - \sqrt{q_{\vartheta,0}(x)}\Big)^2 dx \to 0 \quad \text{as } \delta \to 0. \tag{3.7}$$

LEMMA 3.5. *Suppose the path $\eta \mapsto (F_\eta, G_\eta)$ is regular. Then*

$$\sup_{\vartheta \in \Theta_0} E_\vartheta \int \Big(\sqrt{p_{\vartheta,\delta}(X_0, x)} - \sqrt{p_{\vartheta,0}(X_0, x)} - \tfrac{1}{2}\delta^T \dot{s}_\vartheta(X_0, x)\Big)^2 dx = o(\|\delta\|^2)$$

and

$$\sup_{\vartheta \in \Theta_0} E_\vartheta \int \|\dot{s}_{\vartheta+t}(X_0, x) - \dot{s}_\vartheta(X_0, x)\|^2 dx \to 0 \quad \text{as } t \to 0.$$

PROOF OF THEOREM 3.3. Let $Z_{n,j} = n^{-1/2} S_j(\theta_n)$, $j = 1, \ldots, n$. Our proof utilizes the martingale central limit theorem [Corollary 3.1 in Hall and

Heyde (1980)] and a proper application of Theorem 3.10 in Fabian and Hannan (1987). More precisely, we shall apply their theorem with $\Theta_n = \Delta_0$, $\theta = 0$, $E_{n,\delta}(\cdot) = \int \cdot \, dQ^n_{\theta_n,\delta}$, $U_{n,j} = Z_{n,j}$ and $M_n = nI_2$, where I_2 is the 2×2 identity matrix. In view of these results it suffices to verify

$$E_{\theta_n}\big(Z_{n,j} \mid \mathfrak{F}_{j-1}\big) = 0, \qquad j = 1, \dots, n, \text{ a.s. } P_{\theta_n}, \tag{3.8}$$

$$L_n(a) = \sum_{j=1}^n E_{\theta_n}\Big(\|Z_{n,j}\|^2 I\big[\|Z_{n,j}\| > a\big] \mid \mathfrak{F}_{j-1}\Big) = o_{\theta_n}(1), \qquad a > 0, \tag{3.9}$$

$$\sum_{j=1}^n E_{\theta_n}\big(Z_{n,j} Z_{n,j}^T \mid \mathfrak{F}_{j-1}\big) = V(\theta) + o_{\theta_n}(1), \tag{3.10}$$

and, for every sequence $\{\delta_n\} = \{n^{-1/2} u_n\}$ in Δ_0 with $\{u_n\}$ bounded,

$$\begin{aligned} W_n &= \int \Big(\sqrt{q_{\theta_n,\delta_n}(x)} - \sqrt{q_{\theta_n,0}(x)}\Big)^2 \, dx \\ &\quad + \sum_{j=1}^n \int w^2_{\theta_n,\delta_n}(X_{j-1}, y) \, dy = o_{\theta_n}(1), \end{aligned} \tag{3.11}$$

where $w_{\vartheta,\delta} = \sqrt{p_{\vartheta,\delta}} - \sqrt{p_{\vartheta,0}} - \frac{1}{2}\delta^T \dot{s}_\vartheta$.

It follows from Lemma 3.5 that

$$\int \dot{s}_\vartheta(X_0, x) \sqrt{p_{\vartheta,0}(X_0, x)} \, dx = 0 \quad \text{a.s. } P_\vartheta, \ \vartheta \in \Theta_0,$$

which implies (3.8). Verify that for $a > 0$,

$$E_{\theta_n}(L_n(a)) = \int\int \|\dot{s}_0(u,v)\|^2 I\Big[\|\dot{s}_0(u,v)\| > a\sqrt{n}\sqrt{p(u,v)}\Big] q_{\theta_n}(u) \, du \, dv.$$

Analogous to (3.3), one verifies $\int \|\dot{s}_0(u,v)\|^2 \, dv < 2(1 + u^2)[J(f)(1 + \int_0^1 \zeta^2(t)\,dt) + \int \xi^2 \, dF]$. This, together with Lemma 3.4, yields (3.9). Next, Lemmas 3.4 and 3.5 imply that

$$\begin{aligned} E_{\theta_n}(W_n) &= \int \Big(\sqrt{q_{\theta_n,\delta_n}(x)} - \sqrt{q_{\theta_n,0}(x)}\Big)^2 \, dx \\ &\quad + n E_{\theta_n} \int \Big(\sqrt{p_{\theta_n,\delta_n}(X_0, x)} - \sqrt{p_{\theta_n,0}(X_0, x)} \\ &\qquad\qquad - \tfrac{1}{2}\delta_n^T \dot{s}_{\theta_n}(X_0, x)\Big)^2 \, dx \to 0 \end{aligned}$$

for every sequence $\{\delta_n\} = \{n^{-1/2} u_n\}$ in Δ_0 with $\{u_n\}$ bounded. This gives (3.11).

Verify that

$$\sum_{j=1}^n E_{\theta_n}\big(Z_{n,j} Z_{n,j}^T \mid \mathfrak{F}_{j-1}\big) = \frac{1}{n} \sum_{j=1}^n \int \dot{s}_0(X_{j-1}, v) \dot{s}_0^T(X_{j-1}, v) \, dv.$$

Thus (3.10) follows from the ergodic theorem if $\theta_n = \theta$ for all n. This shows that the theorem holds for the constant sequence $\{\theta\}$. From this one concludes that $\{Q^n_{\theta_n,0}\}$ is contiguous to $\{Q^n_{\theta,0}\}$ and hence (3.10). This completes the proof. □

COROLLARY 3.6. *Let P^n_ϑ denote the restriction of P_ϑ to $\mathfrak{F}_n$ for $\vartheta \in \Theta$. Let $\{\theta_n\}$ be a sequence in Θ_0 such that $\sqrt{n}(\theta_n - \theta)$ is bounded and let $\{t_n\}$ be a bounded sequence in $\mathbb{R}$. Then*

$$\log \frac{dP^n_{\theta_n + n^{-1/2}t_n}}{dP^n_{\theta_n}} - \frac{t_n}{\sqrt{n}} \sum_{j=1}^n l_j(\theta_n) + \frac{1}{2} t_n^2 W(\theta) = o_{\theta_n}(1) \tag{3.12}$$

and

$$\mathfrak{L}\left(\frac{1}{\sqrt{n}} \sum_{j=1}^n l_j(\theta_n) \mid P_{\theta_n} \right) \Rightarrow N(0, W(\theta)). \tag{3.13}$$

Consequently,

$$\frac{1}{\sqrt{n}} \sum_{j=1}^n l_j(\theta_n) - \frac{1}{\sqrt{n}} \sum_{j=1}^n l_j(\theta) + \sqrt{n}(\theta_n - \theta) W(\theta) = o_{\theta_n}(1). \tag{3.14}$$

REMARK 3.7. The results (3.12) and (3.13) together become the usual LAN condition for the joint log-likelihood ratios if $\theta_n = \theta$. This result is different from the LAN-type expansion for the *conditional* log-likelihood ratios, given X_0, of Hwang and Basawa (1993).

4. Efficient and adaptive estimates. This section discusses efficient estimation of θ. First, consider the case when F and G are *known*. As can be seen from Corollary 3.6, the parametric experiment associated with this case satisfies the LAN condition. Thus, it follows from the Hájek–Le Cam theory for LAN experiments that an estimator $\{\hat{\theta}_n\}$ is LAM at θ for bounded loss functions, herein called *efficient* for θ, if

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) - \frac{1}{\sqrt{n}\, W(\theta)} \sum_{j=1}^n l_j(\theta) = o_\theta(1). \tag{4.1}$$

See Fabian and Hannan (1982). To exhibit such an estimator, let

$$\mathbb{Z}_n(\vartheta) = \vartheta + \frac{\sum_{j=1}^n l_j(\vartheta)}{\sum_{j=1}^n l_j^2(\vartheta)}, \qquad \vartheta \in \mathbb{R}.$$

The following construction of efficient estimators uses discretized $\sqrt{n}$-consistent preliminary estimators of θ. The idea of discretization goes back to Le Cam (1960) and has become an important technical tool in the construction of efficient estimators in semiparametric models; see Bickel, Klaasen, Ritov and Wellner (1993) and references therein. In our problem, $\sqrt{n}$-consistent preliminary estimators can be chosen from the class of generalized M-estimators of Section 2.

THEOREM 4.1. *Suppose* (3.1) *holds. Then, for every sequence* $\{\theta_n\}$ *in* Θ *such that* $\sqrt{n}(\theta_n - \theta)$ *is bounded,*

$$\sqrt{n}\,(\mathbb{Z}_n(\theta_n) - \theta) - \frac{1}{\sqrt{n}\,W(\theta)} \sum_{j=1}^{n} l_j(\theta) = o_\theta(1). \tag{4.2}$$

Consequently, if $\{\hat{\theta}_n\}$ *is a discrete* $\sqrt{n}$*-consistent estimator of* θ, *then* $\{\mathbb{Z}_n(\hat{\theta}_n)\}$ *is efficient for* θ.

PROOF. Fix a sequence $\{\theta_n\}$ in Θ such that $\sqrt{n}(\theta_n - \theta)$ is bounded. From (3.8) to (3.10) one obtains that

$$\frac{1}{n}\sum_{j=1}^{n} l_j^2(\theta_n) = \frac{1}{n}\sum_{j=1}^{n} E_{\theta_n}\big(l_j^2(\theta_n)\,|\,\mathfrak{F}_{j-1}\big) + o_{\theta_n}(1) = W(\theta) + o_{\theta_n}(1).$$

This, together with (3.14), gives (4.2). □

Our next goal is to construct estimates of θ that satisfy (4.1) when F and G are *unknown*. Such estimates are LAM-adaptive for θ in the class of all LAN-subproblems that satisfy Stein's (1956) necessary condition for adaptation; see Fabian and Hannan (1982). We now show that this necessary condition is satisfied in our model if F and G are symmetric around zero. To see this consider a regular path $\eta \mapsto (F_\eta, G_\eta)$, where both F_η and G_η are symmetric around zero for all $\eta \in (-1,1)$. For such a path, the conditional density p_η of $X_1 - \theta X_0$, given X_0, satisfies

$$\begin{aligned} p_\eta(X_0, -y) &= \int f_\eta(-y - zX_0)\,dG_\eta(z) \\ &= \int f_\eta(y - zX_0)\,dG_\eta(z) = p_\eta(X_0, y), \qquad y \in \mathbb{R}. \end{aligned}$$

This implies that $\zeta_1(\theta) + \xi_1(\theta)$ [cf. (3.4) and (3.5)] is an even function in the variable $X_1 - \theta X_0$ and $l_1(\theta)$ is an odd function in the variable $X_1 - \theta X_0$. Consequently,

$$V_{12}(\theta) = E_\theta l_1(\theta)\{\zeta_1(\theta) + \xi_1(\theta)\} = 0,$$

which is Stein's necessary condition for adaptation for the given path.

In order to construct adaptive estimates of θ, F and G will be assumed to be symmetric around zero for the remainder of this section. To describe our estimates, let k denote the logistic density. Let $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ denote sequences of positive integers which converge to zero. Let $\{d_n\}$ be a sequence of positive numbers such that $d_n \to \infty$ and let χ_n denote the map from $\mathbb{R}$ to $\mathbb{R}$ defined by $\chi_n(x) = xI[|x| \le d_n] + d_n\,\mathrm{sgn}(x)I[|x| > d_n]$, $x \in \mathbb{R}$. Define functions k_n and k_n' from $\mathbb{R}^2 \times \mathbb{R}^2$ to $\mathbb{R}$ by

$$k_n(u,v) = \frac{1}{2a_n b_n} k\left(\frac{u_1 - v_1}{b_n}\right)\left(k\left(\frac{u_2 - v_2}{a_n}\right) + k\left(\frac{-u_2 - v_2}{a_n}\right)\right)$$

and

$$k_n'(u,v) = \frac{1}{2a_n^2 b_n} k\left(\frac{u_1 - v_1}{b_n}\right)\left(k'\left(\frac{u_2 - v_2}{a_n}\right) - k'\left(\frac{-u_2 - v_2}{a_n}\right)\right),$$

$u = (u_1, u_2) \in \mathbb{R}^2$, $v = (v_1, v_2) \in \mathbb{R}^2$. For $\vartheta \in \mathbb{R}$ and $j = 1, \dots, n$, set

$$Y_j(\vartheta) = \left(X_{j-1}, X_j - \vartheta X_{j-1}\right)$$

and

$$\hat{l}_{n,j}(\vartheta) = -\chi_n(X_{j-1}) \frac{(1/n)\sum_{i=1}^n k_n'\left(Y_j(\vartheta), Y_i(\vartheta)\right)}{c_n + (1/n)\sum_{i=1}^n k_n\left(Y_j(\vartheta), Y_i(\vartheta)\right)}.$$

Finally, set

$$\hat{\mathbb{Z}}_n(\vartheta) = \vartheta + \frac{(1/n)\sum_{j=1}^n \hat{l}_{n,j}(\vartheta)}{(1/n)\sum_{j=1}^n \hat{l}_{n,j}^2(\vartheta)}, \qquad \vartheta \in \mathbb{R}.$$

THEOREM 4.2. *Suppose F and G are symmetric and F satisfies* (3.1). *Assume that the sequences $\{a_n\}, \{b_n\}, \{c_n\}, \{d_n\}$, in addition, satisfy*

$$n^{1-\alpha} a_n^4 b_n^2 c_n^2 \to \infty \quad \text{for some } \alpha > 0, \tag{4.3}$$

$$a_n \geq b_n d_n. \tag{4.4}$$

Then for every sequence $\{\theta_n\}$ in Θ such that $\sqrt{n}(\theta_n - \theta)$ is bounded,

$$\sqrt{n}\left(\hat{\mathbb{Z}}_n(\theta_n) - \theta\right) - \frac{1}{\sqrt{n}\, W(\theta)} \sum_{j=1}^n l_j(\theta) = o_\theta(1). \tag{4.5}$$

Consequently, if $\{\hat{\theta}_n\}$ is a discrete $\sqrt{n}$-consistent estimate of θ, then $\{\hat{\mathbb{Z}}_n(\hat{\theta}_n)\}$ is LAM-adaptive for θ.

PROOF. For $\vartheta \in \mathbb{R}$, let

$$\bar{\mathbb{Z}}_n(\vartheta) = \vartheta + \frac{(1/n)\sum_{j=1}^n \bar{l}_{n,j}(\vartheta)}{(1/n)\sum_{j=1}^n \bar{l}_{n,j}^2(\vartheta)},$$

where

$$\bar{l}_{n,j}(\vartheta) = -\frac{\chi_n(X_{j-1}) \int k_n'\left(Y_j(\vartheta), v\right)\pi_\vartheta(v)\, dv}{c_n + \int k_n\left(Y_j(\vartheta), v\right)\pi_\vartheta(v)\, dv}, \qquad j = 1, \dots, n,$$

and π_ϑ is as in (3.2).

Now fix a sequence $\{\theta_n\}$ in Θ such that $\sqrt{n}(\theta_n - \theta)$ is bounded. It suffices to prove the following two statements:

$$\sqrt{n}\left(\bar{\mathbb{Z}}_n(\theta_n) - \mathbb{Z}_n(\theta_n)\right) = o_{\theta_n}(1), \tag{4.6}$$

$$\sqrt{n}\left(\hat{\mathbb{Z}}_n(\theta_n) - \bar{\mathbb{Z}}_n(\theta_n)\right) = o_{\theta_n}(1). \tag{4.7}$$

As $n^{-1/2}\sum_{j=1}^{n} l_j(\theta_n)$ and $(1/n)\sum_{j=1}^{n} l_j^2(\theta_n)$ are bounded in P_{θ_n}-probability, (4.6) and (4.7) follow from the following four statements:

$$\Delta_{n,1} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left(\bar{l}_{n,j}(\theta_n) - l_j(\theta_n) \right) = o_{\theta_n}(1), \tag{4.8}$$

$$\Delta_{n,2} = \frac{1}{n} \sum_{j=1}^{n} \left(\bar{l}_{n,j}(\theta_n) - l_j(\theta_n) \right)^2 = o_{\theta_n}(1), \tag{4.9}$$

$$\Delta_{n,3} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left(\hat{l}_{n,j}(\theta_n) - \bar{l}_{n,j}(\theta_n) \right) = o_{\theta_n}(1), \tag{4.10}$$

$$\Delta_{n,4} = \frac{1}{n} \sum_{j=1}^{n} \left(\hat{l}_{n,j}(\theta_n) - \bar{l}_{n,j}(\theta_n) \right)^2 = o_{\theta_n}(1). \tag{4.11}$$

Let $Y_{n,j} = Y_j(\theta_n)$ and $\mathbf{Y}_n = (Y_{n,1},\dots,Y_{n,n})$. Define maps $\hat{k}_n$ and $\hat{k}'_n$ from $\mathbb{R}^2 \times \mathbb{R}^{2n}$ into $\mathbb{R}$ as follows:

$$\hat{k}_n(u, y_1,\dots,y_n) = \frac{1}{n}\sum_{j=1}^{n} k_n(u, y_j),$$

$$\hat{k}'_n(u, y_1,\dots,y_n) = \frac{1}{n}\sum_{j=1}^{n} k'_n(u, y_j), \qquad u, y_1,\dots,y_n \in \mathbb{R}^2.$$

Let K denote the logistic distribution function. Define maps $\bar{k}_n^*$, $\bar{k}_n$ and $\bar{k}'_n$ from $\mathbb{R}^2$ to $\mathbb{R}$ by

$$\bar{k}_n^*(u) = \int q_{\theta_n}(u_1 - b_n v_1)\, p(u_1 - b_n v_1, u_2)\, dK(v_1),$$

$$\bar{k}_n(u) = \int\int q_{\theta_n}(u_1 - b_n v_1)\, p(u_1 - b_n v_1, u_2 - a_n v_2)\, dK(v_1)\, dK(v_2),$$

$$\bar{k}'_n(u) = \int\int q_{\theta_n}(u_1 - b_n v_1)\, p'(u_1 - b_n v_1, u_2 - a_n v_2)\, dK(v_1)\, dK(v_2)$$

for $u = (u_1, u_2) \in \mathbb{R}^2$. Verify that

$$\bar{k}_n(u) = \int k_n(u, y)\pi_{\theta_n}(y)\, dy \quad \text{and} \quad \bar{k}'_n(u) = \int k'_n(u, y)\pi_{\theta_n}(y)\, dy, \qquad u \in \mathbb{R}^2.$$

The following facts, proved in the Appendix, are used to establish (4.8) and (4.9):

$$J_{n,1} = b_n^{-2} \int\int \left(\sqrt{\bar{k}_n^*(x_1, x_2)} - \sqrt{\pi_{\theta_n}(x_1, x_2)} \right)^2 dx_1\, dx_2 \to 0, \tag{4.12}$$

$$J_{n,2} = \int\int x_1^2 \left(\frac{\bar{k}'_n(x_1, x_2)}{\sqrt{\bar{k}_n(x_1, x_2)}} - \frac{\sqrt{q_{\theta_n}(x_1)}\, p'(x_1, x_2)}{\sqrt{p(x_1, x_2)}} \right)^2 dx_1\, dx_2 \to 0, \tag{4.13}$$

$$(4.14)\quad J_{n,3}=a_n^{-2}\iint\left(1+x_1^2\right)\left(\sqrt{\bar{k}_n(x_1,x_2)}-\sqrt{\bar{k}_n^*(x_1,x_2)}\right)^2 dx_1\,dx_2\to 0.$$

Now set

$$R_n(u)=\frac{\bar{k}'_n(u)}{\bar{k}_n(u)},\qquad \bar{L}_n(u)=\frac{\bar{k}'_n(u)}{c_n+\bar{k}_n(u)},\qquad \hat{L}_n(u,y)=\frac{\hat{k}'_n(u,y)}{c_n+\hat{k}_n(u,y)}$$

for $u\in\mathbb{R}^2$ and $y\in\mathbb{R}^{2n}$. Verify that, for $j=1,\dots,n$,

$$\hat{l}_{n,j}(\theta_n)=-\chi_n(X_{j-1})\hat{L}_n(Y_{n,j},\mathbf{Y}_n)\quad\text{and}\quad \bar{l}_{n,j}(\theta_n)=-\chi_n(X_{j-1})\bar{L}_n(Y_{n,j}).$$

We shall now summarize various properties of R_n, $\bar{L}_n$ and $\hat{L}_n$ needed in proofs of (4.8)–(4.11). The logistic kernel satisfies

$$(4.15)\quad k(x)\le 1/4,\qquad |k'(x)|\le k(x),\qquad |k''(x)|\le k(x),\qquad x\in\mathbb{R}.$$

From this and (4.4) one obtains the following properties. For each $u\in\mathbb{R}^2$, $y\in\mathbb{R}^{2n}$ and $t\in\mathbb{R}$,

$$(\text{L.1})\quad \left|\bar{L}_n(u)\right|\le\left|R_n(u)\right|\le\frac{1}{a_n}\quad\text{and}\quad\left|\frac{\partial}{\partial u_i}\bar{L}_n(u)\right|\le\frac{1}{a_nb_n},\qquad i=1,2;$$

$$(\text{L.2})\qquad \left|\hat{L}_n(u,y)\right|\le\frac{1}{a_n};$$

$$(\text{L.3})\qquad \left|\frac{\partial}{\partial u_i}\hat{L}_n(u,y)\right|\le\frac{1}{a_nb_n},\qquad i=1,2;$$

$$(\text{L.4})\qquad \left|\frac{\partial}{\partial y_i}\hat{L}_n(u,y)\right|\le\frac{1}{na_n^2b_n^2c_n},\qquad i=1,\dots,2n;$$

$$(\text{L.5})\qquad \left|\hat{L}_n(u,y+te_i)-\hat{L}_n(u,y)\right|\le\frac{1}{na_n^2b_nc_n},\qquad i=1,\dots,2n,$$

where $e_1,\dots,e_{2n}$ denotes the standard basis in $\mathbb{R}^{2n}$. Note also that

$$(\text{L.6})\quad \bar{L}_n((u_1,-u_2))=-\bar{L}_n(u)\quad\text{and}\quad \hat{L}_n((u_1,-u_2),y)=-\hat{L}_n(u,y)$$

for $u=(u_1,u_2)\in\mathbb{R}^2$ and $y\in\mathbb{R}^{2n}$.

PROOF OF (4.8) AND (4.9). Using (L.6), one obtains that

$$E_{\theta_n}\left(\bar{l}_{n,j}(\theta_n)-l_j(\theta_n)\mid X_{j-1}\right)=0,\qquad j=1,\dots,n.$$

Thus (4.8) and (4.9) will follow if we show that

$$\begin{aligned}I_n&=E_{\theta_n}\left(\bar{l}_{n,1}(\theta_n)-l_1(\theta_n)\right)^2\\&=\iint\left(\chi_n(x_1)\bar{L}_n(x_1,x_2)-x_1L(x_1,x_2)\right)^2\pi_{\theta_n}(x_1,x_2)\,dx_1\,dx_2\to 0.\end{aligned}$$

However, we can bound I_n by $5(I_{n,1} + I_{n,2} + I_{n,3} + I_{n,4} + I_{n,5})$, where

$$I_{n,1} = \int\int (\chi_n(x_1) - x_1)^2 L^2(x_1, x_2)\pi_{\theta_n}(x_1, x_2)\,dx_1\,dx_2,$$

$$I_{n,2} = \int\int \chi_n^2(x_1)\left(\frac{\bar{k}_n'(x_1, x_2)}{\sqrt{\bar{k}_n(x_1, x_2)}} - \frac{\sqrt{q_{\theta_n}(x_1)}\,p'(x_1, x_2)}{\sqrt{p(x_1, x_2)}}\right)^2 dx_1\,dx_2,$$

$$I_{n,3} = \int\int \chi_n^2(x_1)R_n^2(x_1, x_2)\Big(\sqrt{\bar{k}_n(x_1, x_2)} - \sqrt{\bar{k}_n^*(x_1, x_2)}\Big)^2 dx_1\,dx_2$$

$$I_{n,4} = \int\int \chi_n^2(x_1)R_n^2(x_1, x_2)\Big(\sqrt{\bar{k}_n^*(x_1, x_2)} - \sqrt{\pi_{\theta_n}(x_1, x_2)}\Big)^2 dx_1\,dx_2$$

and

$$\begin{aligned} I_{n,5} &= \int\int \chi_n^2(x_1)\big(\bar{L}_n(x_1, x_2) - R_n(x_1, x_2)\big)^2 \pi_{\theta_n}(x_1, x_2)\,dx_1\,dx_2 \\ &= \int\int\left(\frac{c_n}{c_n + \bar{k}_n(x_1, x_2)}\right)^2 \chi_n^2(x_1)R_n^2(x_1, x_2)\pi_{\theta_n}(x_1, x_2)\,dx_1\,dx_2. \end{aligned}$$

It follows from (3.3) and Lemma 3.4 that $I_{n,1} \to 0$. Bound $I_{n,2}$ by $J_{n,2}$ to conclude $I_{n,2} \to 0$, by (4.13). Use (L.1) to bound $I_{n,3}$ by $J_{n,3}$. Thus $I_{n,3} \to 0$ by (4.14). As $|\chi_n(x_1)R_n(x_1, x_2)| \le d_n a_n^{-1}$ and $d_n a_n^{-1} \le b_n^{-1}$ by (4.4), $I_{n,4} \le J_{n,1}$. Hence, by (4.12), $I_{n,4} \to 0$. Using Lemma 3.4 and the fact $I_{n,1} + \cdots + I_{n,4} \to 0$, one obtains

$$I_{n,5} = \int\int\left(\frac{c_n}{c_n + \bar{k}_n(x_1, x_2)}\right)^2 x_1^2 L^2(x_1, x_2)\pi_\theta(x_1, x_2)\,dx_1\,dx_2 + o(1).$$

By (4.12), (4.14) and Lemma 3.4, $\bar{k}_n \to \pi_\theta$ in measure. Thus, $I_{n,5} \to 0$ by the Lebesgue dominated convergence theorem. This proves that $I_n \to 0$ and obtains (4.8) and (4.9). □

Proof of (4.10) and (4.11). Without loss of generality, assume that the underlying probability space $(\Omega, \mathscr{A}, P_{\theta_n})$ is rich enough so that there exist independent sequences $\{\hat{X}_{n,j}: j \in \mathbb{Z}\}$, $\{\hat{Y}_{n,j}: j \in \mathbb{Z}\}$ and $\{(\varepsilon_{n,j}, Z_{n,j}): j \in \mathbb{Z}\}$ of independent and identically distributed random vectors such that $\hat{X}_{n,1}$ has the same distribution as X_1, $\hat{Y}_{n,1}$ has the same distribution as $Y_{n,1}$, $(\varepsilon_{n,1}, Z_{n,1})$ has distribution $F \times G$ and

$$X_j = \varepsilon_{n,j} + \sum_{a=1}^{\infty} \varepsilon_{n,j-a} \prod_{\nu=j+1-a}^{j} (\theta_n + Z_{n,\nu})$$

almost surely P_{θ_n}. From this one also obtains the P_{θ_n}-a.s. representation

$$X_j = \varepsilon_{n,j} + \sum_{a=1}^{j-i-1} \varepsilon_{n,j-a} \prod_{\nu=j+1-a}^{j} (\theta_n + Z_{n,\nu}) + X_i \prod_{\nu=i+1}^{j} (\theta_n + Z_{n,\nu}), \qquad i < j. \tag{4.16}$$

Define now

$$X_{j,i} = \varepsilon_{n,j} + \sum_{a=1}^{j-i-1} \varepsilon_{n,j-a} \prod_{\nu=j+1-a}^{j} (\theta_n + Z_{n,\nu}) + \hat{X}_{n,i} \prod_{\nu=i+1}^{j} (\theta_n + Z_{n,\nu}), \qquad i < j.$$

Then $X_{j,i}$ has the same distribution as X_j, and $X_{j,i}$ is independent of X_r and of $X_{r,s}$ for $s < r \le i < j$. The proof below exploits this independence.

Now let $\{m_n\}$ be a sequence of positive integers such that $m_n \sim n^{\alpha/2}$, where α is as in (4.3), and let $\rho_n = \theta_n^2 + \sigma_G^2$. As $\theta_n \to \theta$, $\rho_n \to \theta^2 + \sigma_G^2$, which is less than 1 by (1.2). Therefore, by (4.3),

$$\frac{n \rho_n^{m_n}}{a_n^4 b_n^4 c_n^2} \to 0 \quad \text{and} \quad \frac{m_n^2}{n a_n^4 b_n^2 c_n^2} \to 0. \tag{4.17}$$

Let

$$\mathbf{Y}_n^* = (Y_{n,1}^*, \dots, Y_{n,n}^*), \quad \text{where } Y_{n,j}^* = \big(X_{j-1,j-m_n}, X_{j,j-m_n} - \theta_n X_{j-1,j-m_n}\big).$$

For a subset A of $\{1, \dots, n\}$, let $\mathbf{Y}_n^*(A)$ denote the vector obtained from $\mathbf{Y}_n^*$ by replacing $Y_{n,i}^*$ by $\hat{Y}_{n,i}$ for each $i \in A$. Abbreviate $\mathbf{Y}_n^*(A)$ by $\mathbf{Y}_{n,j}^*$ if $A = \{a\colon a = 1, \dots, n, |a - j| < m_n\}$ and by $\mathbf{Y}_{n,j,i}^*$ if $A = \{a\colon a = 1, \dots, n, |a - j| < m_n, |a - i| < m_n\}$.

Observe that

$$\Delta_{n,3} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} D_n(Y_{n,j}, \mathbf{Y}_n) \quad \text{and} \quad \Delta_{n,4} = \frac{1}{n} \sum_{j=1}^{n} D_n^2(Y_{n,j}, \mathbf{Y}_n),$$

where D_n is the map from $\mathbb{R}^2 \times \mathbb{R}^{2n}$ to $\mathbb{R}$ defined by

$$D_n(u, y) = \chi_n(u_1)\big(\hat{L}_n(u, y) - \bar{L}_n(u)\big), \qquad u = (u_1, u_2) \in \mathbb{R}^2,\ y \in \mathbb{R}^{2n}.$$

To prove (4.10) it suffices to verify the following two statements:

$$T_{n,1} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \big(D_n(Y_{n,j}, \mathbf{Y}_n) - D_n(Y_{n,j}^*, \mathbf{Y}_{n,j}^*)\big) = o_{\theta_n}(1), \tag{4.18}$$

$$T_{n,2} = E_{\theta_n}\bigg(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} D_n(Y_{n,j}^*, \mathbf{Y}_{n,j}^*)\bigg)^2 \to 0. \tag{4.19}$$

By the construction of $X_{j,i}$,

$$E_{\theta_n}(X_j - X_{j,i})^2 \le 2\rho_n^{j-i} E_{\theta_n}(X_0^2), \qquad j > i.$$

Therefore,

$$S_{n,1} = \frac{1}{n}\sum_{j=1}^{n} |X_{j-1,j-m_n} - X_{j-1}|^2 = O_{\theta_n}(\rho_n^{m_n})$$

and

$$\sum_{j=1}^{n} \|Y_{n,j} - Y_{n,j}^*\|^2 = O_{\theta_n}(n\rho_n^{m_n}).$$

Utilizing (L.1) one obtains

$$S_{n,2} = \sum_{j=1}^{n} \left(\bar{L}_n(Y_{n,j}) - \bar{L}_n(Y_{n,j}^*)\right)^2 \le \frac{2}{a_n^2 b_n^2}\sum_{j=1}^{n} \|Y_{n,j} - Y_{n,j}^*\|^2 = O_{\theta_n}\left(\frac{n\rho_n^{m_n}}{a_n^2 b_n^2}\right),$$

and utilizing (L.3)–(L.5) one verifies

$$\begin{aligned} S_{n,3} &= \sum_{j=1}^{n} \left(\hat{L}_n(Y_{n,j}, \mathbf{Y}_n) - \hat{L}_n(Y_{n,j}^*, \mathbf{Y}_{n,j}^*)\right)^2 \\ &\le \frac{6}{a_n^2 b_n^2}\sum_{j=1}^{n} \|Y_{n,j} - Y_{n,j}^*\|^2 + \frac{6}{a_n^4 b_n^4 c_n^2}\sum_{j=1}^{n} \|Y_{n,j} - Y_{n,j}^*\|^2 + \frac{12 m_n^2}{n a_n^4 b_n^2 c_n^2} \\ &= O_{\theta_n}\left(\frac{n\rho_n^{m_n}}{a_n^4 b_n^4 c_n^2} + \frac{m_n^2}{n a_n^4 b_n^2 c_n^2}\right). \end{aligned}$$

Using (L.1) and (L.2) one obtains

$$T_{n,1}^2 \le \frac{12n}{a_n^2} S_{n,1} + \frac{3}{n}\sum_{j=1}^{n} X_{j-1,j-m_n}^2 (S_{n,2} + S_{n,3}).$$

In view of the above bounds, (4.18) follows from (4.17).

Since $Y_{n,i}^*$ and $Y_{n,j}^*$ are independent for $|i-j| \ge m_n$, one obtains with the aid of (4.15) and the identity $E_{\theta_n} k_n(u, Y_{n,1}^*) = E_{\theta_n} k_n(u, Y_{n,1}) = \bar{k}_n(u)$ that

$$\begin{aligned} E_{\theta_n}&\left(\hat{k}_n(u, \mathbf{Y}_n^*) - \bar{k}_n(u)\right)^2 \\ &= \frac{1}{n^2}\sum_{|i-j|<m_n} E_{\theta_n}\left(k_n(u, Y_{n,j}^*) - \bar{k}_n(u)\right)\left(k_n(u, Y_{n,i}^*) - \bar{k}_n(u)\right) \\ &\le \frac{2m_n}{n} E_{\theta_n}\left(k_n(u, Y_{n,1}) - \bar{k}_n(u)\right)^2 \\ &\le \frac{2m_n}{n} E_{\theta_n} k_n^2(u, Y_{n,1}) \le \frac{m_n}{n a_n b_n}\bar{k}_n(u) \end{aligned}$$

and, similarly,

$$E_{\theta_n}\left(\hat{k}_n'(u, \mathbf{Y}_n^*) - \bar{k}_n'(u)\right)^2 \le \frac{2m_n}{n} E_{\theta_n}\left(\hat{k}_n'(u, Y_{n,1}^*)\right)^2 \le \frac{m_n}{n a_n^3 b_n}\bar{k}_n(u).$$

The same arguments give

$$E_{\theta_n} \frac{\left(\hat{k}_n(u, \mathbf{Y}^*_{n,j}) - \bar{k}_n(u)\right)^2}{\bar{k}_n(u)} \leq \frac{m_n}{n a_n b_n}$$

and

$$E_{\theta_n} \frac{\left(\hat{k}'_n(u, \mathbf{Y}^*_{n,j}) - \bar{k}'_n(u)\right)^2}{\bar{k}_n(u)} \leq \frac{m_n}{n a_n^3 b_n}.$$

This shows that

$$\begin{aligned} &E_{\theta_n}\left(\hat{L}_n(u, \mathbf{Y}^*_{n,j}) - \bar{L}_n(u)\right)^2 \\ &\quad = E_{\theta_n}\left(\hat{L}_n(u, \mathbf{Y}^*_{n,j}) \frac{\bar{k}_n(u) - \hat{k}_n(u, \mathbf{Y}^*_{n,j})}{c_n + \bar{k}_n(u)} + \frac{\hat{k}'_n(u, \mathbf{Y}^*_{n,j}) - \bar{k}'_n(u)}{c_n + \bar{k}_n(u)}\right)^2 \\ &\quad \leq \frac{2}{a_n^2 c_n} \frac{m_n}{n a_n b_n} + \frac{2 m_n}{n a_n^3 b_n c_n} = \frac{4 m_n}{n a_n^3 b_n c_n}. \end{aligned}$$

Therefore, by the independence of $Y^*_{n,j}$ and $\mathbf{Y}^*_{n,j}$,

$$E_{\theta_n} D_n^2(Y^*_{n,j}, \mathbf{Y}^*_{n,j}) = \int E_{\theta_n} D_n^2(u, \mathbf{Y}^*_{n,j}) \pi_{\theta_n}(u)\, du \leq \frac{4 m_n}{n a_n^3 b_n c_n} E_{\theta_n} X_0^2 \tag{4.20}$$

and

$$\frac{1}{n} \sum_{|i-j|<m_n} E_{\theta_n} D_n(Y^*_{n,j}, \mathbf{Y}^*_{n,j}) D_n(Y^*_{n,i}, \mathbf{Y}^*_{n,i}) \leq \frac{8 m_n^2}{n a_n^3 b_n c_n} E_{\theta_n} X_0^2. \tag{4.21}$$

As $D_n((s,-t), y) = -D_n((s,t), y)$ and $p(s,-t) = p(s,t)$ we have

$$\int D_n((s,t), y) p(s,t)\, dt = 0.$$

If $|i-j| \geq m_n$, then $Y^*_{n,j}$ and $(\mathbf{Y}^*_{n,j}, \mathbf{Y}^*_{n,j,i}, Y^*_{n,i})$ are independent and hence

$$E_{\theta_n} D_n(Y^*_{n,j}, \mathbf{Y}^*_{n,j}) D_n(Y^*_{n,i}, \mathbf{Y}^*_{n,j,i}) = 0 \quad \text{and}$$
$$E_{\theta_n} D_n(Y^*_{n,j}, \mathbf{Y}^*_{n,j,i}) D_n(Y^*_{n,i}, \mathbf{Y}^*_{n,j,i}) = 0.$$

Thus for $|i-j| \geq m_n$, we find with the help of (L.5),

$$\begin{aligned} &\left|E_{\theta_n} D_n(Y^*_{n,j}, \mathbf{Y}^*_{n,j}) D_n(Y^*_{n,i}, \mathbf{Y}^*_{n,i})\right| \\ &\quad = \left|E_{\theta_n}\left(D_n(Y^*_{n,j}, \mathbf{Y}^*_{n,j}) - D_n(Y^*_{n,j}, \mathbf{Y}^*_{n,j,i})\right)\left(D_n(Y^*_{n,i}, \mathbf{Y}^*_{n,i}) - D_n(Y^*_{n,i}, \mathbf{Y}^*_{n,j,i})\right)\right| \\ &\qquad \leq E_{\theta_n}\left|X_{j-1,j-m_n} X_{i-1,i-m_n}\right| \left(\frac{2 m_n}{n a_n^2 b_n c_n}\right)^2 \leq \frac{4 m_n^2}{n^2 a_n^4 b_n^2 c_n^2} E_{\theta_n} X_0^2. \end{aligned}$$

Consequently,

$$(4.22)\qquad \frac{1}{n}\sum_{|i-j|\ge m_n} E_{\theta_n} D_n(Y^*_{n,j},\mathbf{Y}^*_{n,j}) D_n(Y^*_{n,i},\mathbf{Y}^*_{n,i}) \le \frac{4m_n^2}{na_n^4 b_n^2 c_n^2} E_{\theta_n} X_0^2.$$

Combine (4.21) and (4.22) with (4.17) to obtain (4.19).

We are left to verify (4.11). In view of (4.20) and (4.17), it suffices to show that

$$T_{n,3} = \frac{1}{n}\sum_{j=1}^{n}\left(D_n(Y_{n,j},\mathbf{Y}_n) - D_n(Y^*_{n,j},\mathbf{Y}^*_{n,j})\right)^2 = o_{\theta_n}(1).$$

However, this follows from the bound

$$T_{n,3} \le 12 a_n^{-2} S_{n,1} + \frac{3d_n^2}{n}\sum_{j=1}^{n}\left(X_{j-1,j-m_n}\right)^2 (S_{n,2} + S_{n,3})$$

and the above calculations. □

APPENDIX

This Appendix provides the proofs of Lemmas 3.4 and 3.5 and the statements (4.12)–(4.14). To do this we review some facts about L_2-convergence and discuss uniform L_2-differentiability. Let $(S, \mathscr{S}, \rho)$ be a measure space and let $h_0, h_1, h_2, \dots$ be measurable functions from S to $\mathbb{R}$.

DEFINITION A.1. We say h_n *converges to* h_0 *in* ρ*-measure on sets of finite measure* and write $h_n \overset{\rho,\mathrm{fin}}{\to} h_0$ if $\rho(\{|h_n - h_0| > a\} \cap B) \to 0$ for every $a > 0$ and every $B \in \mathscr{S}$ with $\rho(B) < \infty$.

The notion of convergence in measure on sets of finite measure is helpful in proving convergence in $L_2(\rho)$ as is evidenced in the following lemma which can be deduced from Theorems 4.8.6 and 4.8.11 in Fabian and Hannan [(1985), pages 199–201].

LEMMA A.2. *Suppose* $h_0, h_1, h_2, \dots$ *are* ρ*-square integrable and* $h_n \overset{\rho,\mathrm{fin}}{\to} h_0$. *Then the following statements are equivalent*:

(i) $\int (h_n - h_0)^2 \, d\rho \to 0$.
(ii) $\int h_n^2 \, d\rho \to \int h_0^2 \, d\rho$.
(iii) *The sequence* $\{h_n\}$ *is uniformly* ρ*-square integrable.*

As a consequence we obtain the following result which will be used repeatedly in the sequel. Let λ denote the Lebesgue measure on $\mathbb{R}$.

LEMMA A.3. *Let $r_0, r_1, \dots$ be measurable functions from S to $\mathbb{R}$, let $\zeta_0, \zeta_1, \dots$ be elements of $L_2(\lambda)$, let $\xi_0, \xi_1, \dots$ be elements of $L_2(\rho)$ and let $\varphi_0, \varphi_1, \dots$ be the elements of $L_2(\lambda \times \rho)$ defined by*

$$\varphi_n(x, s) = \zeta_n(x - r_n(s))\xi_n(s), \qquad x \in \mathbb{R}, s \in S.$$

Suppose that ρ is σ-finite, $\xi_n \to \xi_0$ in $L_2(\rho)$, $\zeta_n \to \zeta_0$ in $L_2(\lambda)$ and $r_n \overset{\rho, \mathrm{fin}}{\to} r_0$. Then $\varphi_n \to \varphi_0$ in $L_2(\lambda \times \rho)$.

PROOF. Define a map D from $\mathbb{R}$ to $\mathbb{R}$ by

$$D(t) = \int (\zeta_0(x - t) - \zeta_0(x))^2 \, d\lambda(x), \qquad t \in \mathbb{R}.$$

Then D is bounded by $2\int \zeta_0^2 \, d\lambda$ and is uniformly continuous; see Theorem 9.5 in Rudin (1974) for the latter. The desired result follows from this Lemma A.2 and the bound

$$\begin{aligned}
&\tfrac{1}{3}\int (\varphi_n - \varphi_0)^2 \, d(\lambda \times \rho) \\
&\quad \le \int (\zeta_n - \zeta_0)^2 \, d\lambda \int \xi_n^2 \, d\rho + \int \zeta_0^2 \, d\lambda \int (\xi_n - \xi_0)^2 \, d\rho + \int D \circ (r_n - r_0)\xi_0^2 \, d\rho.
\end{aligned}$$

□

LEMMA A.4. *Assume the path $\eta \mapsto \{F_\eta, \zeta_\eta\}$ is regular. Then for every $\vartheta \in \Theta_0$, $\int |q_{\vartheta,\delta} - q_{\vartheta,0}| \, d\lambda \to 0$ as $\delta \to 0$.*

PROOF. Fix $\vartheta \in \Theta_0$. For $\delta = (t, \eta) \in \Delta_0$, define a map r_δ from $(0, 1)$ to $\mathbb{R}$ by

$$r_\delta(u) = \vartheta + t + G_\eta^{-1}(u), \qquad u \in (0, 1),$$

and set $\beta_\delta = \int_0^1 |r_\delta(u)| \, du$, $\bar{\beta}_\delta = \max\{\beta_\delta, \beta_0\}$ and $\tau_\delta = \int |x| q_{\vartheta,\delta}(x) \, dx$.
Fix a positive integer k. We shall show that

$$\lim_{\delta \to 0} \int |q_{\vartheta,\delta} - q_{\vartheta,0}| \, d\lambda \le 2k\tau_0 \beta_0^k. \tag{A.1}$$

The desired result follows from this as $\beta_0 < 1$. Let

$$\begin{aligned}
\gamma_\delta(y, s) = \int_{\mathbb{R}^{k-1}} \int_{(0,1)^k} f_\eta\left(y - \sum_{j=1}^{k-1} y_j \prod_{i=1}^{j} r_\delta(u_i) - s \prod_{i=1}^{k} r_\delta(u_i) \right) \\
\times \prod_{i=1}^{k-1} f_\eta(y_i) \, dy_i \prod_{i=1}^{k} du_i.
\end{aligned}$$

Then one verifies

$$q_{\vartheta,\delta}(y) = \int_{-\infty}^{\infty} \gamma_\delta(y, s) q_{\vartheta,\delta}(s) \, ds \quad \text{for } \lambda\text{-almost all } y \in \mathbb{R}.$$

Using this one finds that $\int |q_{\vartheta,\delta} - q_{\vartheta,0}|\, d\lambda \le A_\delta + B_\delta$, where

$$\begin{aligned} A_\delta &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\gamma_\delta(y,s) - \gamma_0(y,s)|\, dy\, q_{\vartheta,\delta}(s)\, ds \\ &\le k\int |f_\eta - f|\, d\lambda \\ &\quad + \int |f'|\, d\lambda \left(\sum_{i=1}^{k-1} \sigma_F i\big(\bar\beta_\delta\big)^{i-1} + k\big(\bar\beta_\delta\big)^{k-1}\tau_\delta \right) \int_0^1 |r_\delta(u) - r_0(u)|\, du \end{aligned}$$

and

$$\begin{aligned} B_\delta &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} \gamma_0(y,s)\big(q_{\vartheta,\delta}(s) - q_{\vartheta,0}(s)\big)\, ds \right| dy \\ &\le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\gamma_0(y,s) - \gamma_0(y,0)|\big(q_{\vartheta,\delta}(s) + q_{\vartheta,0}(s)\big)\, ds\, dy \\ &\le \int |f'|\, d\lambda\, (\tau_\delta + \tau_0)\beta_0^k. \end{aligned}$$

Taking limits shows that (A.1) holds. □

Proof of Lemma 3.4. Fix $\vartheta \in \Theta_0$. It follows from (1.4) and the properties (R.1) and (R.3) of a regular path that $\int (1+x^2) q_{\vartheta,\delta}(x)\, dx \to \int (1+x^2) q_{\vartheta,0}(x)\, dx$ as $\delta \to 0$, and from Lemma A.4 that $q_{\vartheta,\delta_n} \overset{\lambda,\mathrm{fin}}{\to} q_{\vartheta,0}$ as $\delta_n \to 0$. Consequently, by Lemma A.2,

$$\int (1+x^2)\Big(\sqrt{q_{\vartheta,\delta}(x)} - \sqrt{q_{\vartheta,0}(x)}\Big)^2\, dx \quad \text{as } \delta \to 0.$$

The desired result follows from this, the compactness of Θ_0 and the fact that $q_{\vartheta+s,(t,\eta)} = q_{\vartheta,(s+t,\eta)}$ for s in a neighborhood of 0. □

Let Γ be a compact metric space and let Δ be a subset of $\mathbb{R}^m$ which contains 0 and has 0 as accumulation point.

Definition A.5. Let $\Phi = \{\phi_{\gamma,\delta}\colon \gamma \in \Gamma,\ \delta \in \Delta\}$ be a family of elements of $L_2(\rho)$ and let $\dot\Phi = \{\dot\phi_\gamma\colon \gamma \in \Gamma\}$ be a family of elements of $L_2^m(\rho)$. Then we say Φ is *ρ-smooth for* $\dot\Phi$ if:

(i) The map $\gamma \mapsto \phi_{\gamma,0}$ is $L_2(\rho)$-continuous.
(ii) The map $\gamma \mapsto \dot\phi_\gamma$ is $L_2^m(\rho)$-continuous.
(iii) The map $\delta \mapsto \phi_{\gamma,\delta}$ has $L_2(\rho)$-derivative $\dot\phi_\gamma$ at 0 uniformly in $\gamma \in \Gamma$, that is,

$$\sup_{\gamma\in\Gamma} \int \Big(\phi_{\gamma,\delta} - \phi_{\gamma,0} - \delta^T\dot\phi_\gamma\Big)^2\, d\rho = o(\|\delta\|^2).$$

PROPOSITION A.6. *Let* $(Y, \mathscr{Y}, \nu)$ *be a* σ*-finite measure space. Let* $\{\varphi_\delta\colon \delta \in \Delta\}$ *be a class of* λ*-densities, let* $\{h_{\gamma,\delta}\colon \gamma \in \Gamma,\ \delta \in \Delta\}$ *be a class of elements of* $L_2(\nu)$ *and let* $\{r_{\gamma,\delta}\colon \gamma \in \Gamma,\ \delta \in \Delta\}$ *be a class of measurable functions from* Y *to* $\mathbb{R}$. *Let*

$$\phi_{\gamma,\delta}(x, y) = \sqrt{\varphi_\delta\big(x - r_{\gamma,\delta}(y)\big)}\, h_{\gamma,\delta}(y), \qquad x \in \mathbb{R},\ y \in Y.$$

Suppose that the following conditions hold for a class $\{\dot r_\gamma\colon \gamma \in \Gamma\}$ *of measurable functions from* Y *to* $\mathbb{R}^m$:

(a.1) *The density* φ_0 *has finite Fisher information* $J(\varphi_0)$ *for location.*
(a.2) *For some* ζ_0 *in* $L_2^m(\lambda)$, $\int(\sqrt{\varphi_\delta} - \sqrt{\varphi_0} - \delta^T\zeta_0)^2\, d\lambda = o(\|\delta\|^2)$.
(a.3) *The class* $\{h_{\gamma,\delta}\colon \gamma \in \Gamma,\ \delta \in \Delta\}$ *is* ν*-smooth for* $\{\dot h_\gamma\colon \gamma \in \Gamma\}$.
(a.4) *The class* $\{(r_{\gamma,\delta} - r_{\gamma,0})h_{\gamma,0}\colon \gamma \in \Gamma,\ \delta \in \Delta\}$ *is* ν*-smooth for* $\{\dot r_\gamma h_{\gamma,0}\colon \gamma \in \Gamma\}$.
(a.5) *For every* $\tilde\gamma \in \Gamma$, $r_{\gamma,\delta} \overset{\nu,\,\mathrm{fin}}{\to} r_{\tilde\gamma,0}$ *as* $\gamma \to \tilde\gamma$ *and* $\delta \to 0$.

Then the class $\{\phi_{\gamma,\delta}\colon \gamma \in \Gamma,\ \delta \in \Delta\}$ *is* $\lambda \times \nu$*-smooth for* $\{\dot\phi_\gamma\colon \gamma \in \Gamma\}$, *where*

$$\begin{aligned}\dot\phi_\gamma(x, y) = {}&\left(\zeta_0\big(x - r_{\gamma,0}(y)\big) - \dot r_\gamma(y)\frac{\varphi_0'}{2\sqrt{\varphi_0}}\big(x - r_{\gamma,0}(y)\big)\right)h_{\gamma,0}(y)\\ &+ \sqrt{\varphi_0\big(x - r_{\gamma,0}(y)\big)}\,\dot h_\gamma(y)\end{aligned}$$

for $x \in \mathbb{R}$ *and* $y \in Y$.

PROOF. Using (a.3)–(a.5) and Lemma A.3 one verifies the continuity of the maps $\gamma \mapsto \phi_{\gamma,0}$ and $\gamma \mapsto \dot\phi_\gamma$. Let

$$\begin{aligned} t_{\gamma,\delta}(x, y) &= \sqrt{\varphi_\delta\big(x - r_{\gamma,\delta}(y)\big)}\,\dot h_\gamma(y),\\ \zeta_{\gamma,\delta}(x, y) &= \zeta_0\big(x - r_{\gamma,\delta}(y)\big)h_{\gamma,0}(y),\\ \xi_{\gamma,\delta}(x, y) &= \dot r_\gamma(y)h_{\gamma,0}(y)\int_0^1 \frac{\varphi_0'}{2\sqrt{\varphi_0}}\big(x - v\delta^T\dot r_\gamma(y)\big)\, dv. \end{aligned}$$

It follows from (a.3)–(a.5) and Lemma A.3 that

$$t_{\gamma,\delta} \to t_{\tilde\gamma,0}, \qquad \zeta_{\gamma,\delta} \to \zeta_{\tilde\gamma,0}, \qquad \xi_{\gamma,\delta} \to \xi_{\tilde\gamma,0} \tag{A.2}$$

in $L_2^m(\lambda \times \nu)$ as $(\gamma, \delta) \to (\tilde\gamma, 0)$. As φ_0 has finite Fisher information, $\sqrt{\varphi_0}$ is absolutely continuous with derivative $\varphi_0'/2\sqrt{\varphi_0}$. This show that

$$\sqrt{\varphi_0(x - s)} - \sqrt{\varphi_0(x)} = -s\int_0^1 \frac{\varphi_0'}{2\sqrt{\varphi_0}}(x - st)\, dt, \qquad x, s \in \mathbb{R}.$$

Using this and the translation invariance of the Lebesgue measure, one derives the bound

$$
\begin{aligned}
&\tfrac{1}{6}\int\Big(\phi_{\gamma,\delta}-\phi_{\gamma,0}-\delta^T\dot\phi_\gamma\Big)^2\,d(\lambda\times\nu)\\
&\quad\le\int\Big(h_{\gamma,\delta}-h_{\gamma,0}-\delta^T\dot h_\gamma\Big)^2\,d\nu+\|\delta\|^2\int\|t_{\gamma,\delta}-t_{\gamma,0}\|^2\,d(\lambda\times\nu)\\
&\qquad+\int\Big(\sqrt{\varphi_\delta}-\sqrt{\varphi_0}-\delta^T\zeta_0\Big)^2\,d\lambda\int h^2_{\gamma,0}\,d\nu+\|\delta\|^2\int\|\zeta_{\gamma,\delta}-\zeta_{\gamma,0}\|^2\,d(\lambda\times\nu)\\
&\qquad+J(\varphi_0)\int\Big(r_{\gamma,\delta}-r_{\gamma,0}-\delta^T\dot r_\gamma\Big)^2h_\gamma\,d\nu+\|\delta\|^2\int\|\xi_{\gamma,\delta}-\xi_{\gamma,0}\|^2\,d(\lambda\times\nu).
\end{aligned}
$$

This, (A.2) and (a.2)–(a.4) establish the property (iii). □

The next result is a uniform version of a special case of Pitman's (1979) result on the preservation of Hellinger differentiability under transformations.

PROPOSITION A.7. *Let $(X,\mathscr{X},\mu)$ and $(Y,\mathscr{Y},\nu)$ be σ-finite measure spaces. Let $\Phi=\{\phi_{\gamma,\delta}\colon\gamma\in\Gamma\ \delta\in\Delta\}$ be $\mu\times\nu$-smooth for $\{\dot\phi_\gamma\colon\gamma\in\Gamma\}$. Then $\Psi=\{\psi_{\gamma,\delta}\colon\gamma\in\Gamma,\ \delta\in\Delta\}$ is μ-smooth for $\{\dot\psi_\gamma\colon\gamma\in\Gamma\}$, where*

$$
\psi_{\gamma,\delta}(x)=\sqrt{\textstyle\int\phi^2_{\gamma,\delta}(x,y)\,d\nu(y)}\qquad\textit{and}
$$

$$
\text{(A.3)}\qquad \dot\psi_\gamma(x)=\frac{\int\dot\phi_\gamma(x,y)\phi_{\gamma,0}(x,y)\,d\nu(y)}{\psi_{\gamma,0}(x)}
$$

for $x\in X$.

PROOF. For each $\gamma\in\Gamma$ and $\delta\in\Delta$, define a linear operator $A_{\gamma,\delta}$ from $L_2(\mu\times\nu)$ to $L_2(\mu)$ as follows. For $h\in L_2(\mu\times\nu)$ let $A_{\gamma,\delta}h$ denote the element of $L_2(\mu)$ defined by

$$
A_{\gamma,\delta}h(x)=\frac{\int h(x,y)\big(\phi_{\gamma,\delta}(x,y)+\phi_{\gamma,0}(x,y)\big)\,d\nu(y)}{\psi_{\gamma,\delta}(x)+\psi_{\gamma,0}(x)},\qquad x\in X.
$$

Now fix $\tilde\gamma\in\Gamma$ and $h\in L_2(\mu\times\nu)$. An application of the Cauchy–Schwarz inequality shows that

$$
\text{(A.4)}\qquad \big(A_{\gamma,\delta}h(x)\big)^2\le\int h^2(x,y)\,d\nu(y),\qquad x\in X.
$$

Therefore the class $\{A_{\gamma,\delta}h\colon\gamma\in\Gamma,\ \delta\in\Delta\}$ is uniformly μ-square-integrable. It follows from (i) and (iii) that the map $(\gamma,\delta)\mapsto\phi_{\gamma,\delta}$ is $L_2(\mu\times\nu)$-continuous at $(\tilde\gamma,0)$. This implies that the maps $(\gamma,\delta)\mapsto\psi^2_{\gamma,\delta}$ and $(\gamma,\delta)\mapsto\int h(\cdot,y)\phi_{\gamma,\delta}(\cdot,y)\,d\nu(y)$ are $L_1(\mu)$-continuous at $(\tilde\gamma,0)$. From this one concludes that $A_{\gamma,\delta}h\overset{\mu,\mathrm{fin}}{\to}A_{\tilde\gamma,0}h$ as $(\gamma,\delta)\to(\tilde\gamma,0)$. Therefore, Lemma A.2 yields that the map $(\gamma, \delta) \mapsto A_{\gamma,\delta} h$ is $L_2(\mu)$-continuous at $(\tilde{\gamma}, 0)$. Easy calculations show that for every $\gamma \in \Gamma$ and $\delta \in \Delta$, $\delta^T \dot{\psi}_\gamma = A_{\gamma,0} \delta^T \dot{\phi}_\gamma$ and

$$\psi_{\gamma,\delta} - \psi_\gamma - \delta^T \dot{\psi}_\gamma = \frac{\psi_{\gamma,\delta}^2 - \psi_{\gamma,0}^2}{\psi_{\gamma,\delta} + \psi_{\gamma,0}} - A_{\gamma,\delta} \delta^T \dot{\phi}_\gamma + A_{\gamma,\delta} \delta^T \dot{\phi}_\gamma - A_{\gamma,0} \delta^T \dot{\phi}_\gamma$$
$$= A_{\gamma,\delta}\left(\phi_{\gamma,\delta} - \phi_{\gamma,0} - \delta \dot{\phi}_\gamma\right) + A_{\gamma,\delta} \delta^T \dot{\phi}_\gamma - A_{\gamma,0} \delta^T \dot{\phi}_\gamma.$$

The desired result follows now from the properties of the operators $A_{\gamma,\delta}$ and the $\mu \times \nu$-smoothness of Φ for $\dot{\Phi}$. □

PROOF OF LEMMA 3.5. For $\vartheta \in \Theta_0$ and $\delta = (t, \eta) \in \Delta_0$, define maps

$$\psi_{\vartheta,\delta}(x) = \sqrt{q_\vartheta(x_1) p_{\vartheta,\delta}(x)} \quad \text{and} \quad \dot{\psi}_\vartheta(x) = \sqrt{q_\vartheta(x_1)}\, \tfrac{1}{2} \dot{s}_\vartheta(x)$$

for $x = (x_1, x_2) \in \mathbb{R}^2$. Verify that these maps satisfy (A.3) with $\Gamma = \Theta_0$, $\Delta = \Delta_0$, $\nu = U$, the uniform distribution on $(0,1)$,

$$\phi_{\vartheta,\delta}(x, y) = \sqrt{f_\eta\left(x_2 - \left(\vartheta + t + G_\eta^{-1}(y)\right)x_1\right) q_\vartheta(x_1)}$$

and

$$\dot{\phi}_\vartheta(x, y)$$
$$= \tfrac{1}{2}\sqrt{q_\vartheta(x_1)} \begin{pmatrix} -x_1 \dfrac{f'}{\sqrt{f}}(x_2 - x_1\vartheta - x_1 G^{-1}(y)) \\ \zeta(x_2 - x_1\vartheta - x_1 G^{-1}(y)) - x_1 \xi(y) \dfrac{f'}{\sqrt{f}}(x_2 - x_1\vartheta - x_1 G^{-1}(y)) \end{pmatrix}$$

for $x = (x_1, x_2) \in \mathbb{R}^2$, $y \in (0,1)$, where ξ and ζ are as in (R.2) and (R.3). It suffices to show that $\{\psi_{\vartheta,\delta}\colon \vartheta \in \Theta_0,\ \delta \in \Delta_0\}$ is $\lambda \times \lambda$ smooth for $\{\dot{\psi}_\vartheta\colon \vartheta \in \Theta_0\}$. By Proposition A.7 this follows if we verify that $\{\phi_{\vartheta,\delta}\colon \vartheta \in \Theta_0,\ \delta \in \Delta_0\}$ is $\lambda \times \lambda \times U$ smooth for $\{\dot{\phi}_\vartheta\colon \vartheta \in \Theta_0\}$. However, this follows from Proposition A.6 applied with $Y = \mathbb{R} \times (0,1)$, $\nu = \lambda \times U$, $\Gamma = \Theta_0$, $\Delta = \Delta_0$, $\varphi_\delta = f_\eta$, $h_{\vartheta,\delta}(y) = \sqrt{q_\vartheta(y_1)}$ and $r_{\vartheta,\delta}(y) = y_1(\vartheta + t + G_\eta^{-1}(y_2))$ for $\delta = (t, \eta) \in \Delta$ and $y = (y_1, y_2) \in Y$. Note that the assumptions (a.1)–(a.5) are verified with $\dot{h}_\vartheta = 0$,

$$\zeta_0 = \tfrac{1}{2}\begin{pmatrix} 0 \\ \zeta \end{pmatrix} \quad \text{and} \quad \dot{r}_\gamma(y) = y_1 \begin{pmatrix} 1 \\ \xi(y_2) \end{pmatrix}, \qquad y \in Y,$$

relying on Lemma 3.4, (R.2) and (R.3). □

In the remainder of this Appendix we shall derive additional corollaries to Propositions A.6 and A.7. Corollaries A.10–A.12 will imply the required statements (4.12)–(4.14).

REMARK A.8. Let $\vartheta \in \Theta$. Then the density q_ϑ can be chosen to satisfy

$$q_\vartheta(x) = \int\int f(x - (\vartheta + z)w) q_\vartheta(w)\, dw\, dG(z)$$

for all $x \in \mathbb{R}$. It is easy to check that this choice is absolutely continuous with a.e.-derivative q'_ϑ given by

$$q'_\vartheta(x) = \int\int f'(x - (\vartheta + z)w) q_\vartheta(w)\, dw\, dG(z), \qquad x \in \mathbb{R}.$$

The argument used to derive (3.3) yields now that q_ϑ has finite Fisher information $J(q_\vartheta) \le J(f)$.

Now let, for $\vartheta \in \Theta$, $\beta \in \mathbb{R}$ and $x = (x_1, x_2) \in \mathbb{R}^2$,

$$\chi_{\beta,\vartheta}(x) = \sqrt{q_\vartheta(x_1 - \beta x_2)} \quad \text{and} \quad \dot{\chi}_{\beta,\vartheta}(x) = -x_2 \frac{q'_\vartheta(x_1 - \beta x_2)}{2\sqrt{q_\vartheta(x_1 - \beta x_2)}}.$$

Recall that K denotes the logistic distribution function.

COROLLARY A.9. *Let Γ be a compact subset of $\mathbb{R} \times \Theta$ and let λ_* be the measure defined by $d\lambda_*(t) = (1 + t^2)\, d\lambda(t)$. Then the following hold*:

(i) *The map $(\beta, \vartheta) \mapsto \chi_{\beta,\vartheta}$ is $L_2(\lambda_* \times K)$ continuous.*
(ii) *The class $\{\chi_{\beta+\delta,\vartheta}: (\beta, \vartheta) \in \Gamma, \ \delta \in \mathbb{R}\}$ is $\lambda \times K$-smooth for $\{\dot{\chi}_{\beta,\vartheta}: (\beta, \vartheta) \in \Gamma\}$.*

PROOF. The first statement follows from Lemmas A.3 and 3.4. To obtain the second statement, set

$$\phi_{\gamma,\delta}(x, y) = \sqrt{f(x_1 - (\beta + \delta)x_2 - (\vartheta + y_2)y_1)}\sqrt{q_\vartheta(y_1)}$$

and

$$\dot{\phi}_\gamma(x, y) = -x_2 \frac{f'(x_1 - \beta x_2 - (\vartheta + y_2)y_1)}{2\sqrt{f(x_1 - \beta x_2 - (\gamma + y_2)y_1)}}\sqrt{q_\vartheta(y_1)}$$

for $x = (x_1, x_2) \in \mathbb{R}^2$, $y = (y_1, y_2) \in \mathbb{R}^2$, $\gamma = (\beta, \vartheta) \in \Gamma$ and $\delta \in \mathbb{R}$. Then

$$\chi_{\beta+\delta,\vartheta}(x) = \int \phi^2_{\gamma,\delta}(x, y)\, d\nu(y) \quad \text{and}$$

$$\dot{\chi}_{\beta+\delta,\vartheta}(x) = \frac{\int \dot{\phi}_\gamma(x, y)\phi_{\gamma,0}(x, y)\, d\nu(y)}{\chi_{\beta,\vartheta}(x)},$$

where $\nu = \lambda \times G$. It follows from Proposition A.6 that $\{\phi_{\gamma,\delta}: \gamma \in \Gamma, \ \delta \in \mathbb{R}\}$ is $\lambda \times K \times \lambda \times G$ smooth for $\{\dot{\phi}_\gamma: \gamma \in \Gamma\}$. Now apply Proposition A.7 with $\nu = K \times \lambda \times G$ to obtain the desired result. □

For $\alpha, \beta \in \mathbb{R}$ and $\vartheta \in \Theta$, define functions $g_{\alpha,\beta,\vartheta}$ and $g'_{\alpha,\beta,\vartheta}$ by

$$g_{\alpha,\beta,\vartheta}(x_1, x_2) = \iiint f(x_2 - \alpha u - w(x_1 - \beta v))q_\vartheta(x_1 - \beta v)\, dK(u)\, dK(v)\, dG(w),$$

$$g'_{\alpha,\beta,\vartheta}(x_1, x_2) = \iiint f'(x_2 - \alpha u - w(x_1 - \beta v))q_\vartheta(x_1 - \beta v)\, dK(u)\, dK(v)\, dG(w)$$

for $x_1, x_2 \in \mathbb{R}$. Then $g_{0,0,\theta_n} = \pi_{\theta_n}$, $g'_{0,0,\theta_n}(x_1, x_2) = p'(x_1, x_2)q_{\theta_n}(x_1)$, $g_{a_n, b_n, \theta_n} = \bar{k}_n$ and $g_{0, b_n, \theta_n} = \bar{k}^*_n$. The following results are now easy consequences of Propositions A.7 and A.8 and of Corollary A.11.

COROLLARY A.10. *The class* $\Psi = \{\psi_{\gamma,\delta} = \sqrt{g_{0,\delta,\gamma}}\colon \gamma \in \Gamma,\ \delta \in \mathbb{R}\}$ *is* $\lambda \times \lambda$-*smooth for* $\{\dot{\psi}_\gamma = 0\colon \gamma \in \Gamma\}$ *for every compact subset* Γ *of* Θ. *Consequently*, (4.12) *holds*.

COROLLARY A.11. *For* $(\alpha, \beta, \vartheta) \in \mathbb{R} \times \mathbb{R} \times \Theta$ *and* $\delta \in \mathbb{R}$, *let*

$$\psi_{(\alpha,\beta,\vartheta),\delta}(x_1, x_2) = \sqrt{g_{\delta,\beta,\vartheta}(x_1, x_2 - \delta x_1)}$$

and

$$\dot{\psi}_{(\alpha,\beta,\vartheta)}(x_1, x_2) = \frac{-x_1 g'_{\alpha,\beta,\vartheta}(x_1, x_2)}{2\sqrt{g_{\alpha,\beta,\vartheta}(x_1, x_2)}},$$

where $x_1, x_2 \in \mathbb{R}$. *Then the class* $\Psi = \{\psi_{\gamma,\delta}\colon \gamma \in \Gamma,\ \delta \in \mathbb{R}\}$ *is* $\lambda \times \lambda$ *smooth for* $\dot{\Psi} = \{\dot{\psi}_\gamma\colon \gamma \in \Gamma\}$ *for every compact subset* Γ *of* $\mathbb{R} \times \mathbb{R} \times \Theta$. *Consequently*, (4.13) *holds*.

COROLLARY A.12. *For* $(\beta, \vartheta) \in \mathbb{R} \times \Theta$ *and* $\delta \in \mathbb{R}$, *let*

$$\psi_{(\beta,\vartheta),\delta}(x_1, x_2) = \sqrt{(1 + x_1^2) g_{\delta,\beta,\vartheta}(x_1, x_2)}, \qquad x_1, x_2 \in \mathbb{R}.$$

Then the class $\Psi = \{\psi_{\gamma,\delta}\colon \gamma \in \Gamma,\ \delta \in \mathbb{R}\}$ *is* $\lambda \times \lambda$-*smooth for* $\{\dot{\psi}_\gamma = 0\colon \gamma \in \Gamma\}$ *for every compact subset* Γ *of* $\mathbb{R} \times \Theta$. *Consequently*, (4.14) *holds*.

REFERENCES

BICKEL, P. J. (1982). On adaptive estimation. *Ann. Statist.* **10** 647–671.

BICKEL, P. J., KLAASSEN, C. A. J., RITOV, Y. and WELLNER, J. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins Univ. Press.

FABIAN, V. and HANNAN, J. (1982). On estimation and adaptive estimation for locally asymptotically normal families. *Z. Wahrsch. Verw. Gebiete* **59** 459–478.

FABIAN, V. and HANNAN, J. (1985). *Introduction to Probability and Mathematical Statistics*. Wiley, New York.

FABIAN, V. and HANNAN, J. (1987). Local asymptotic behavior of densities. *Statist. Decisions* **5** 105–138. [Correction (1988) **6** 195.]

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

HWANG, S. Y. and BASAWA, I. V. (1993). Asymptotic optimal inference for a class of nonlinear time series. *Stochastic Process. Appl.* **46** 91–114.

KREISS, J.-P. (1987a). On adaptive estimation in stationary ARMA processes. *Ann. Statist.* **15** 112–133.

KREISS, J.-P. (1987b). On adaptive estimation in autoregressive models when there are nuisance functions. *Statist. Decisions* **5** 59–76.

KOUL, H. L. (1992). *Weighted Empiricals and Linear Models*. IMS, Hayward, CA.

KOUL, H. L. and PFLUG, G. (1990). Weakly adaptive estimators in explosive autoregression. *Ann. Statist.* **18** 939–960.

LE CAM, L. (1960). Locally asymptotically normal families of distributions. *University of California Publications in Statistics* **3** 37–98.

NICHOLLS, D. F. and QUINN, B. G. (1982). *Random Coefficient Autoregression Models: An Introduction. Lecture Notes in Statist.* **11**. Springer, New York.

PITMAN, E. J. G. (1979). *Some Basic Theory for Statistcal Inference*. Chapman and Hall, London.

RUDIN, W. (1974). *Real and Complex Analysis*, 2nd ed. McGraw-Hill, New York.

SCHICK, A. (1986). On asymptotically efficient estimation in semi-parametric models. *Ann. Statist.* **14** 1139–1151.

SCHICK, A. (1987). A note on the construction of asymptotically linear estimators. *J. Statist. Plann. Inference* **16** 89–105. [Correction (1989) **22** 269–270.]

SCHICK, A. (1993). An adaptive estimator of the autocorrelation coefficient in a semiparametric regression model with autoregressive errors. Preprint.

STEIN, C. (1956). Efficient nonparametric testing and estimation. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 187–195. Univ. California Press, Berkeley.

TONG, H. (1990). *Nonlinear Time Series: A Dynamical Approach*. Oxford Univ. Press.

DEPARTMENT OF STATISTICS AND PROBABILITY
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48824-1027
E-MAIL: 20974AJT@ibm.msu.edu

DEPARTMENT OF MATHEMATICAL SCIENCES
SUNY BINGHAMTON
BINGHAMTON, NEW YORK 13902-6000
E-MAIL: anton@math.binghamton.edu