The Annals of Statistics
1990, Vol. 18, No. 2, 939–960

WEAKLY ADAPTIVE ESTIMATORS IN EXPLOSIVE AUTOREGRESSION

BY HIRA L. KOUL[1] AND GEORG CH. PFLUG[2]

Michigan State University and Justus-Liebig University

Consider the model $X_i = \rho X_{i-1} + \varepsilon_i$, $|\rho| > 1$, where $X_0, \varepsilon_1, \varepsilon_2, \dots$ are independent random variables with $\varepsilon_1, \varepsilon_2, \dots$ having common density ψ. This paper gives sufficient conditions under which the sequence of experiments induced by $\{X_0, X_1, \dots, X_n\}$ has a weak limit in the sense of Le Cam. In general, the limiting experiment is translation invariant and neither LAN nor LAMN. The paper further shows that the sequence of Pitman-type estimators of ρ at a given ψ converges weakly to T, where T is minimax for the limiting experiment under a weighted squared error loss function. Finally, for an unknown ψ, a sequence of Pitman-type estimators that converges weakly to T is constructed. These estimators are called *weakly adaptive*. The class of error densities for which these results hold include some that may not have finite Fisher information.

1. Introduction and summary. The construction of estimators that are asymptotically efficient over a range of nuisance parameters, otherwise known as adaptive, has been the focus of many researchers. See, e.g., Stein (1956), Levit (1975), Koshevnik and Levit (1976), Beran (1976, 1978), Bickel (1982), Fabian and Hannan (1982), Schick (1986) and Kreiss (1987), among others. In all of these and related papers the observations are either independent or stationary and, more importantly, the underlying experiments are locally asymptotically normal (LAN) in the sense of Le Cam (1960, 1986). Relatively little is known about the construction of adaptive estimators in experiments that may be neither LAN nor locally asymptotically mixed normal (LAMN).

This paper considers the problem of efficient estimation for an explosive autoregression model. To describe this model in its simplest form, let ψ be a density with support on the entire real line $\mathbb{R}$ and $\varepsilon_1, \varepsilon_2, \dots$ be independent and identically distributed (i.i.d.) ψ random variables (r.v.'s). Let X_0 be an observable r.v. independent of ε_i, $i \geq 1$. The observable process $\{X_i\}$ is such that for a $|\rho| > 1$,

$$X_i = \rho X_{i-1} + \varepsilon_i, \qquad i \geq 1. \tag{1.1}$$

The parameter of interest is ρ and the nuisance parameter is ψ.

The process $\{X_i\}$ is nonstationary. If $E \log^+ |\varepsilon_1| < \infty$, then $|X_i| \to \infty$ a.s. as $i \to \infty$, thereby rendering $\{X_i\}$ explosive. Even when ψ is a normal density, the sequence of experiments induced by $\{X_0, X_1, \dots, X_n\}$ is not LAN as noted in Basawa and Koul (1979). See also Lemma 2.3 below for the nature of the experiment in this case.

Received June 1988; revised February 1989.
[1]Research partly supported by Deutsche Forschungsgemeinschaft.
[2]Currently at the Institut für Statistik and Informatik, Universität Wien, Vienna, Austria.
AMS 1980 *subject classifications*. Primary 62G05; secondary 62G20, 62G10.
Key words and phrases. Pitman estimators, minimax estimators, nonstationary.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics*.
www.jstor.org

For a general ψ, Theorem 2.1 gives a set of sufficient conditions on ψ under which the sequence of these experiments has a weak limit in the sense of Le Cam. As noted in Section 2, the limiting experiment is translation invariant and neither LAN nor LAMN in general. The class of densities for which Theorem 2.1 holds includes some that may have neither finite first moment nor finite Fisher information.

Under some additional moment conditions, Lemma 2.2 shows that a sequence of Pitman-type estimators at a given error density ψ_0 has weak limit T. Using Strasser (1985, Lemma 39.28), T is shown to be minimax for an appropriately normalized squared error loss function in Theorem 2.2. In the process of the proof of Lemma 2.2, we essentially show that the sequence of experiments induced by $X_0, X_1, \dots, X_n$ is 1 times Lebesgue uniformly integrable. The structure of the limiting experiment is also discussed in Section 2.

Finally, in Section 3, a sequence of Pitman-type estimators corresponding to kernel-type density estimators is constructed and shown to converge weakly to T for a large class of error densities. This property of a given sequence of estimators is termed *weak adaptivity*. The class of densities for which we have a weakly adaptive sequence of estimators included some that may have infinite Fisher information.

In the sequel, ε will stand for a r.v. having the same distribution as ε_1. The symbol $\Rightarrow$ denotes the weak convergence, $\to_{\text{f.d.}}$ denotes the weak convergence of the finite dimensional distributions and $=_{\mathscr{L}}$ denotes the equivalence in distribution.

2. Sufficient conditions and limiting experiment. Let $P^n_{\rho,\psi}$ be the distribution of $(X_0, X_1, \dots, X_n)$ given by (1.1) when ψ is the density of ε and ρ is the true parameter. In order to find the corresponding asymptotic experiment in the sense of Le Cam we first study the limit of the log-likelihood process

$$\Lambda_n(t|\psi) := \log\left(dP^n_{\rho+\delta_n t,\psi}/dP^n_{\rho,\psi}\right), \tag{2.1}$$

where the shrinking factor δ_n must be chosen in such a way that the process $\Lambda_n(t|\psi)$ has a nondegenerate limit.

The following assumption is used in determining the magnitude of δ_n and the limit of $\Lambda_n(\cdot|\psi)$ under P^n_{ρ,ψ_0} where ψ_0 is a density on $\mathbb{R}$. In the sequel, P_0 denotes the probability measure of ε under ψ_0.

ASSUMPTION A. (i) For any $r > 1$, $\sum_{i\geq 1} P_0(|\varepsilon| > r^i) < \infty$.

(ii) $\log\psi$ is locally Lipschitz in the following sense: There are a nonnegative measurable function h and positive constants α, θ such that

$$|\log\psi(y) - \log\psi(x)| \leq h(x)|y-x|^{\alpha} \quad \text{for } |x-y| \leq \theta$$

with

$$\sum_{i=1}^{\infty} P_0\left(h(\varepsilon) > r^j\right) < \infty, \quad \text{for any } r > 1. \tag{2.2}$$

Assumption A(i) is equivalent to $E\log^+|\varepsilon| < \infty$.

To determine the magnitude of δ_n in the present case it is helpful to understand the behavior of $\{X_i\}$. From (1.1) we readily have that

$$\rho^{-i}X_i = X_0 + \sum_{j=1}^{i} \rho^{-j}\varepsilon_j, \qquad i \geq 1. \tag{2.3}$$

This motivates the following somewhat general lemma which will be used repeatedly in the proof of Theorem 2.1 with different η's.

LEMMA 2.1. *Let $\eta, \eta_1, \eta_2, \dots,$ be i.i.d. r.v.'s and η_0 be a r.v. independent of $\eta, \eta_1, \eta_2, \dots$. Assume that for any $r > 1$,*

$$\sum_{i=1}^{\infty} P(|\eta| > r^i) < \infty. \tag{2.4}$$

Let $|\gamma| > 1$ and define

$$Y_m = \sum_{i=0}^{m} \gamma^{-i}\eta_i, \qquad Y = \sum_{i=0}^{\infty} \gamma^{-i}\eta_i. \tag{2.5}$$

Then

$$Y \text{ is a.s. finite.} \tag{2.6}$$

Moreover, as $m \to \infty$,

$$|s^m(Y_m - Y)| \to 0, \quad a.s., \text{ for all } 1 \leq |s| < |\gamma| \tag{2.7}$$

and

$$|\gamma^m(Y_m - Y)| = O_p(1). \tag{2.8}$$

PROOF. Fix an s with $1 \leq |s| < |\gamma|$. Choose an $r > 1$ such that $|s|r < |\gamma|$, $r < |\gamma|$. Let

$$A_i = \{|\eta_i| \leq r^i\}, \qquad i \geq 1, \qquad B_n = \bigcap_{i \geq n} A_i.$$

By (2.4) and the Borel–Cantelli lemma, $P(\bigcup_n B_n) = 1$. But on B_n, for $m > n$,

$$|s^m(Y_m - Y)| = |s^m| \left| \sum_{i=m+1}^{\infty} \gamma^{-i}\eta_i \right| \leq \left| \frac{sr}{\gamma} \right|^m \frac{r}{|\gamma| - r} \to 0, \quad \text{as } m \to \infty. \tag{2.9}$$

This implies (2.7). Choose $s = 1$ in (2.9) to conclude (2.6). To prove (2.8), note that

$$\gamma^m(Y_m - Y) = -\sum_{i=m+1}^{\infty} \gamma^{m-i}\eta_i = -\sum_{i=1}^{\infty} \gamma^{-i}\eta_{m+i} =_{\mathscr{L}} -\sum_{i=1}^{\infty} \gamma^{-i}\eta_i = -Y.$$

This completes the proof of the lemma. □

COROLLARY 2.1. *Let* (1.1) *hold for a* $|\rho| > 1$. *Assume that the error r.v.'s* $\{\varepsilon_i\}$ *satisfy* A(i). *Then, as* $m \to \infty$,

$$|s^m(\rho^{-m}X_m - Z)| \to 0, \quad a.s., \text{ for every } 1 \le |s| < |\rho|, \tag{2.10}$$

$$\rho^m(\rho^{-m}X_m - Z) =_{\mathscr{L}} -Z, \tag{2.11}$$

where

$$Z = X_0 + \sum_{i=1}^{\infty} \rho^{-i}\varepsilon_i \quad \text{is a.s. finite.} \tag{2.12}$$

PROOF. Follows from (2.3) and Lemma 2.1 upon taking $\eta_i \equiv \varepsilon_i$, $\eta_0 = X_0$ and $\gamma = \rho$ in there. □

COROLLARY 2.2. *Let* (1.1) *hold. Assume that the error r.v.'s satisfy* A(i). *Let* $|u_n|$ *be a bounded sequence and*

$$X_{i,n} := (\rho + u_n\rho^{-n})^i X_0 + \sum_{j=1}^{i} (\rho + u_n\rho^{-n})^{i-j}\varepsilon_j.$$

Then, as $n \to \infty$,

$$|s|^n \sup_{1 \le i \le n} |\rho^{-i}(X_{i,n} - X_i)| \to 0, \quad a.s., \text{ for } 1 \le |s| < |\rho|$$

and hence

$$|s|^i|\rho^{-i}X_{i,n} - Z| \to 0, \quad a.s. \text{ as } n \to \infty, i \to \infty, \text{ for } 1 \le |s| < |\rho|.$$

PROOF. Since $|(1 + a/m)^m - 1| \le 2|a|$ for all $|a| \le 1$ and for all $m \ge 1$, it readily follows that for n sufficiently large such that $|nu_n\rho^{-n-1}| \le 1$,

$$\begin{aligned}
&\max_{1 \le i \le n} |s|^n|\rho^{-i}(X_{i,n} - X_i)| \\
&\quad \le |s|^n \max_{1 \le i \le n}\left(\sum_{j=1}^{i} |\rho|^{-j}|\varepsilon_j|\left|(1 + \rho^{-n-1}u_n)^{i-j} - 1\right|\right. \\
&\qquad\qquad \left. + |X_0|\left|(1 + \rho^{-n-1}u_n)^i - 1\right|\right) \\
&\quad \le 2\left|\frac{s}{\rho}\right|^n n|u_n|\left(\sum_{j=1}^{\infty} |\rho|^{-j}|\varepsilon_j| + |X_0|\right) \to 0, \quad \text{a.s., as } n \to \infty
\end{aligned}$$

because by Lemma 2.1 applied to $\eta = |\varepsilon|$, $\gamma = |\rho|$, the r.v. $\sum_{j=1}^{\infty}|\rho|^{-j}|\varepsilon_j| < \infty$ a.s. The last assertion of the corollary follows from this, the triangle inequality and (2.10). □

Now, (2.10) suggests that δ_n in (2.1) should be of the order ρ^{-n}. Thus our standardized likelihood process for the model (1.1) is

$$\Lambda_n(t|\psi) = \sum_{i=1}^{n} \log\{\psi(\varepsilon_i - \rho^{-n}tX_{i-1})/\psi(\varepsilon_i)\}. \tag{2.13}$$

Theorem 2.1 below gives the weak limit of $\Lambda_n(t|\psi)$ when the underlying error density in (1.1) is ψ_0, $\psi_0 \neq \psi$. Since the distribution of X_0 does not appear in Λ_n, from now on we shall take $X_0 = 0$ without loss of generality.

THEOREM 2.1. *Let* (1.1) *hold. Let* ψ_0 *be the density of* ε_i, $i \geq 1$, *and* ψ *be another density on* $\mathbb{R}$. *Assume that* A(i) *and* A(ii) *hold. Then, under* P^n_{ρ,ψ_0},

$$\Lambda_n(\cdot|\psi) \to_{\text{f.d.}} \Lambda(\cdot|\psi),$$

where

$$\Lambda(t|\psi) := \sum_{i=1}^{\infty} \log\{\psi(\varepsilon_i' - t\rho^{-i}Z)/\psi(\varepsilon_i')\}, \tag{2.14}$$

with $\{\varepsilon_i'\}$ *i.i.d.* ψ_0 *r.v.'s, independent of* Z.

PROOF. We first show that the process (2.14) is well defined. Let θ be as in A(ii). There is an N such that $|t \cdot \rho^{-N}Z| < \theta$, on the set $\{|Z| < K\}$. Therefore, by A(ii),

$$\sum_{i=N}^{\infty} \left| \log \frac{\psi(\varepsilon_i' - t\rho^{-i}Z)}{\psi(\varepsilon_i')} \right| \leq |tZ|^{\alpha} \sum_{i=N}^{\infty} h(\varepsilon_i')|\rho|^{-i\alpha} \quad \text{on } \{|Z| < K\}.$$

This bound together with Lemma 2.1 applied to $\eta_i \equiv h(\varepsilon_i')$, $\eta_0 = 0$, $\gamma = |\rho|^{\alpha}$ and the arbitrariness of K ensure that $\Lambda(\cdot|\psi) < \infty$, a.s. To prove the main assertion, let $m < n$ and rewrite

$$\begin{aligned} \Lambda_n(t|\psi) = &\sum_{i=n-m+1}^{n} \log \frac{\psi(\varepsilon_i - t\rho^{i-2n+m-1}X_{n-m})}{\psi(\varepsilon_i)} \\ &+ \sum_{i=1}^{n-m} \log \frac{\psi(\varepsilon_i - t\rho^{-n}X_{i-1})}{\psi(\varepsilon_i)} + \sum_{i=n-m+1}^{n} \log \frac{\psi(\varepsilon_i - t\rho^{-n}X_{i-1})}{\psi(\varepsilon_i - t\rho^{i-2n+m-1}X_{i-1})} \\ = &\; U_{nm} + R_{nm1} + R_{nm2} \quad \text{(say)}: \end{aligned} \tag{2.15}$$

Thus it suffices to prove the following three statements:

$$U_{nm} \to_{\text{f.d.}} \sum_{i=1}^{m} \log \frac{\psi(\varepsilon_i' - t\rho^{-i}Z)}{\psi(\varepsilon_i')} \quad \text{as } n \to \infty, \text{ for all } m, \tag{2.16}$$

$$|R_{nm1}| = o_p(1), \quad \text{as } n \to \infty,\ m \to \infty, \tag{2.17}$$

$$|R_{nm2}| = o_p(1), \quad \text{as } n \to \infty, \text{ for all } m. \tag{2.18}$$

To prove (2.16), note that

$$U_{n,m} = \sum_{j=1}^{m} \log \frac{\psi\left(\varepsilon_{n-j+1} - t\rho^{m-n-j}X_{n-m}\right)}{\psi(\varepsilon_{n-j+1})} =_{\mathscr{L}} \sum_{j=1}^{m} \log \frac{\psi\left(\varepsilon_j' - t\rho^{m-n-j}X_{n-m}\right)}{\psi(\varepsilon_j')},$$

where the r.v.'s $\{\varepsilon_i'\}$ are independent of the r.v.'s $\{\varepsilon_i\}$ and have the same distribution as that of $\{\varepsilon_i\}$. It follows from the continuity of ψ and (2.10) that as $n \to \infty$,

$$\sum_{j=1}^{m} \left| \log \frac{\psi\left(\varepsilon_j' - t\rho^{m-n-j}X_{n-m}\right)}{\psi(\varepsilon_j')} - \log \frac{\psi\left(\varepsilon_j' - t\rho^{-j}Z\right)}{\psi(\varepsilon_j')} \right| \to 0,$$

a.s., as $n \to \infty$, for each m.

To prove (2.17), let $Y_i = \rho^{-i}X_i$, $i \ge 1$. An application of Lemma 2.1 with $\eta_0 = 0$, $\gamma = |\rho|$, $\eta = |\varepsilon|$ and A(i) readily yields that

$$V^p := \max_{i \ge 1} |Y_i|^p \le \left[\sum_{j=1}^{\infty} |\rho|^{-j} |\varepsilon_j| \right]^p < \infty, \quad \text{a.s., for any } p > 0.$$

Also $|t\rho^{-n}X_{i-1}| \le |t\rho^{-n+i-1}|V$, $i \ge 1$. Therefore, there is an n, sufficiently large, such that $|t\rho^{-n}X_{i-1}| \le \theta$ for all $1 \le i \le n - m$, on the set $[V < K]$, $K > 0$. Hence, by A(ii),

$$\begin{aligned} |R_{nm1}| &\le |t|^{\alpha} \sum_{i=1}^{n-m} |\rho|^{-\alpha(n-i+1)} h(\varepsilon_i) |Y_{i-1}|^{\alpha} \\ &\le |t|^{\alpha} V^{\alpha} \sum_{i=1}^{n-m} |\rho|^{-\alpha(n-i+1)} h(\varepsilon_i) = o_p(1) \quad \text{on } [V < K] \end{aligned} \tag{2.19}$$

because

$$\sum_{i=1}^{n-m} |\rho|^{-\alpha(n-i+1)} h(\varepsilon_i) =_{\mathscr{L}} \sum_{i=m+1}^{n} |\rho|^{-\alpha i} h(\varepsilon_i') \to 0, \quad \text{a.s., as } n \to \infty,\ m \to \infty$$

by Lemma 2.1 applied to $\gamma = |\rho|^{\alpha}$, $\eta = h(\varepsilon_i')$, $\eta_0 = 0$. Because K is arbitrary, this completes the proof of (2.17).

To prove (2.18), rewrite

$$R_{nm2} = \sum_{j=1}^{m} \log\left\{\psi\left(\varepsilon_{n-j+1} - t\rho^{-j}Y_{n-j}\right) / \psi\left(\varepsilon_{n-j+1} - t\rho^{-j}Y_{n-m}\right)\right\}.$$

By (2.10), for every j and each m, $|t\rho^{-j}(Y_{n-j} - Y_{n-m})| \to 0$, a.s., as $n \to \infty$. This together with the continuity of $\log \psi$ and the fact $\varepsilon_{n-j+1} - t\rho^{-j}Y_{n-m}$ and $\varepsilon_{n-j+1} - t\rho^{-j}Y_{n-j}$ are bounded in probability for each j and m yield (2.18), thereby completing the proof of the theorem. □

Remark 2.1. By standard embedding technique, we may find a version of the process from the LHS (2.16) which converges a.s. to a version of the limit process, uniformly for t's lying in a bounded set. From (2.19) and the proof of

(2.18), it readily follows then that the convergence of the finite-dimensional distributions of $\Lambda_n(\cdot|\psi)$ is uniform for these t's.

REMARK 2.2. In order to conclude that the sequence of experiments $\mathscr{E}_n = (\mathbb{R}_n, \mathscr{B}_n, \{P^n_{\rho+\rho^{-n}t,\psi_0}\})$ has a weak limit in the sense of Le Cam, we need to prove that under $P^n_{\rho+\rho^{-n}u,\psi_0}$,

$$(2.20)\quad dP^n_{\rho+\rho^{-n}t,\psi_0}/dP^n_{\rho+\rho^{-n}u,\psi_0} \to_{\text{f.d.}} \exp(\Lambda(t-u|\psi_0)) \quad \text{for every } u,t \text{ in } \mathbb{R}.$$

But, under $P^n_{\rho+\rho^{-n}u,\psi_0}$, the model (1.1) is equivalent to

$$X_{i,n} = (\rho+\rho^{-n}u)X_{i-1,n} + \varepsilon_i, \qquad X_{i,n} = \sum_{j=1}^{\infty}(\rho+\rho^{-n}u)^{i-j}\varepsilon_j, \qquad i \geq 1.$$

Moreover, for every u and t in $\mathbb{R}$, under $P^n_{\rho+\rho^{-n}u,\psi_0}$,

$$\frac{dP^n_{\rho+\rho^{-n}t,\psi_0}}{dP^n_{\rho+\rho^{-n}u,\psi_0}} = \prod_{i=1}^{n}\frac{\psi_0(X_{i,n}-(\rho+\rho^{-n}t)X_{i-1,n})}{\psi_0(X_{i,n}-(\rho+\rho^{-n}u)X_{i-1,n})}$$
$$= \prod_{i=1}^{n}\left\{\frac{\psi_0(\varepsilon_i-\rho^{-n}(t-u)\cdot X_{i-1,n})}{\psi_0(\varepsilon_i)}\right\}.$$

Now mimic the proof of Theorem 2.1 after replacing X_i by $X_{i,n}$, t by $t-u$ and ψ by ψ_0 in there to conclude (2.20). Use Corollary 2.2 in place of Corollary 2.1 whenever needed.

We now turn to an efficient estimation of ρ at a given ψ_0. For this we need the following additional assumption.

ASSUMPTION B. (i) The density ψ_0 satisfies A(ii) with h symmetric, $h(x)$ nondecreasing in $|x|$.

(ii) $E\log^+ h(|\varepsilon|+K) < \infty$, for all $0 < K < \infty$.

(iii) With $b(y) := \sup_{|x|\geq y}\psi_0(x)$, $\int_0^\infty y b(y)\,dy < \infty$.

Let k_n be a sequence of integers, $1 \leq k_n \leq n-2$, $k_n \to \infty$ as $n \to \infty$. Define

$$(2.21)\quad \tilde{\rho}_n := \int u \prod_{i=k_n+1}^{n}\psi_0(X_i-uX_{i-1})\,du \Big/ \int \prod_{i=k_n+1}^{n}\psi_0(X_i-uX_{i-1})\,du,$$

$$(2.22)\quad T_n := \int u \prod_{i=k_n+1}^{n}\psi_0(\varepsilon_i-u\rho^{-n}X_{i-1})\,du \Big/ \int \prod_{i=k_n+1}^{n}\psi_0(\varepsilon_i-u\rho^{-n}X_{i-1})\,du.$$

By a change of variable,

$$(2.23)\qquad \rho^n(\tilde{\rho}_n-\rho) = T_n.$$

With $Y_i = \rho^{-i}X_i$, define

$$T_n' := \int u \prod_{i=k_n+1}^{n}\psi_0\left(\varepsilon_i-u\rho^{-n+i-1}Y_{k_n}\right)du \Big/ \int \prod_{i=k_n+1}^{n}\psi_0\left(\varepsilon_i-u\rho^{-n+i-1}Y_{k_n}\right)du.$$

Direct calculations show that

(2.24)
$$
\begin{aligned}
T_n' &= \frac{\rho\varepsilon_n}{Y_{k_n}} - \frac{\rho}{Y_{k_n}} E\Big(\varepsilon_n \Big| \varepsilon_{k_n+1} - \rho^{-n+k_n+1}\varepsilon_n, \dots, \varepsilon_{n-1} - \rho^{-1}\varepsilon_n\Big) \quad \text{a.s.} \\
&=_{\mathscr{L}} \frac{\varepsilon_1'\rho}{Y_{k_n}} - \frac{\rho}{Y_{k_n}} E\Big(\varepsilon_1' \Big| \varepsilon_2' - \rho^{-1}\varepsilon_1', \varepsilon_3 - \rho^{-2}\varepsilon_1', \dots, \varepsilon_{n-k_n}' - \rho^{-(n-k_n-1)}\varepsilon_1'\Big) \\
&= S_n, \quad \text{say.}
\end{aligned}
$$

Note that under B(iii) all of the above entities are well defined. Moreover, because ε is a continuous r.v., $P(Y_{k_n} = 0) = 0$. Now, by (2.10) and the martingale convergence theorem, as long as $n - k_n \to \infty$,

$$S_n \to T \quad \text{a.s.,} \tag{2.25}$$

where

$$T = \varepsilon_1'\rho Z^{-1} - \rho Z^{-1} E\big(\varepsilon_1' | \varepsilon_2' - \rho^{-1}\varepsilon_1', \varepsilon_3 - \rho^{-2}\varepsilon_1', \dots\big). \tag{2.26}$$

Thus under B(iii), from (2.24) and (2.25),

$$T_n' \Rightarrow T. \tag{2.27}$$

Next we have,

Lemma 2.2. *Assume* (1.1) *holds with the error density* ψ_0 *satisfying the Assumption* B. *Then*

$$\rho^n(\tilde{\rho}_n - \rho) \Rightarrow T. \tag{2.28}$$

Proof. Because of (2.23) and (2.27), it suffices to show that

$$\big|T_n - T_n'\big| = o_p(1) \qquad \big(P^n_{\rho,\psi_0}\big). \tag{2.29}$$

For a $c > 0$, define the event

(2.30)
$$
\begin{aligned}
A_{n,c} := \Bigg\{ &\Big(c\big|\rho^{-1}Y_{n-1}\big| - 1\Big) \wedge \Big(c\big|\rho^{-1}Y_{k_n}\big| - 1\Big) \ge |\varepsilon_n|, \big|Y_{k_n}\big| \le c, \\
&\max_{k_n < i \le n} |\rho|^{k_n+1} \left| \sum_{j=k_n+1}^{i} \rho^{-j}\varepsilon_j \right| \le c, \psi_0(\varepsilon_n) > c^{-1} \Bigg\}.
\end{aligned}
$$

Because of Corollary 2.1, $\forall\ \eta > 0$, $\exists\ N_\eta$ and $0 < c < \infty$ such that

$$P(A_{n,c}) \ge 1 - \eta \quad \text{for all } n > N_\eta. \tag{2.31}$$

Next, to simplify notation, define, for $u \in \mathbb{R}$, $i \ge 1$,

(2.32)
$$
\begin{aligned}
\tau_{ni}(u) &:= \psi_0(\varepsilon_i - t\rho^{-n}X_{i-1})/\psi_0(\varepsilon_i), \\
\tau_{ni}'(u) &:= \psi_0\big(\varepsilon_i - u\rho^{-n+i-1}Y_{k_n}\big)/\psi_0(\varepsilon_i).
\end{aligned}
$$

Also let

$$l_n := \prod_{i=k_n+1}^{n} \tau_{ni}, \qquad T_{nj} := \int u^j l_n(u)\, du, \qquad j = 0, 1.$$

$$l'_n := \prod_{i=k_n+1}^{n} \tau'_{ni}, \qquad T'_{nj} := \int u^j l'_n(u)\, du, \qquad j = 0, 1.$$

Note that

$$T_n = \int u l_n(u)\, du \Big/ \int l_n(u)\, du = T_{n1}/T_{n0}$$

and

$$T'_n = \int u l'_n(u)\, du \Big/ \int l'_n(u)\, du = T'_{n1}/T'_{n0}.$$

Thus, to prove (2.29), it suffices to show that

$$\left|T_{nj} - T'_{nj}\right| = o_p(1), \qquad j = 0, 1. \tag{2.33}$$

Now write, for a $0 < d < \infty$,

$$T_{nj} = \int_{|u| \le d} u^j l_n(u)\, du + \int_{|u| > d} u^j l_n(u)\, du.$$

With a similar decomposition for T'_{nj}, one readily has

$$\left|T_{nj} - T'_{nj}\right| \le d^j \int_{|u| \le d} \left|l_n(u) - l'_n(u)\right| du + \int_{|u| > d} |u|^j \left[l_n(u) + l'_n(u)\right] du,$$
$$j = 0, 1.$$

Hence, in view of (2.31), to prove (2.33), it suffices to prove the following two results:

For every $0 < d < \infty$,

$$I(A_{n,c}) \int_{|u| \le d} \left|l_n(u) - l'_n(u)\right| du = o_p(1), \qquad \text{as } n \to \infty. \tag{2.34}$$

For a sufficiently large $0 < d < \infty$,

$$EI(A_{n,c}) \int_{|u| > d} |u| \left[l_n(u) + l'_n(u)\right] du \tag{2.35}$$

can be made sufficiently small, for all $n > N_\eta$.

But on $A_{n,c}$,

$$\left|\varepsilon_n - u\rho^{-n} X_{n-1}\right| \ge \left| |u| \left|\rho^{-n} X_{n-1}\right| - |\varepsilon_n| \right|$$
$$\ge |u| \left|\rho^{-1} Y_{n-1}\right| - |\varepsilon_n| \ge |u|/c, \qquad \forall\, |u| \ge c.$$

Therefore, if $d \geq c$, then

$$I(A_{n,c})I(|u| > d)\psi_0(\varepsilon_n - u\rho^{-n}X_{n-1})/\psi_0(\varepsilon_n)$$
$$\leq I(A_{n,c})I(|u| > d)cb(|u|/c).$$

With $\mathscr{F}_i$ denoting the σ-algebra generated by $\varepsilon_1, \dots, \varepsilon_i$, note that

$$E\big[\psi_0(\varepsilon_i - u\rho^{-n}X_{i-1})/\psi_0(\varepsilon_i)|\mathscr{F}_{i-1}\big] = 1 \quad \text{for all } i \geq 1. \tag{2.36}$$

Hence

$$\begin{aligned}
&EI(A_{n,c})\int_{|u|>d}|u|l_n(u)\,du\\
&= EI(A_{n,c})\int_{|u|>d}|u|\tau_{nn}(u)\prod_{i=k_n+1}^{n-1}\tau_{ni}(u)\,du\\
&\leq c\int_{|u|>d}|u|b(|u|/c)E\left\{\prod_{i=k_n+1}^{n-1}\psi_0(\varepsilon_i - u\rho^{-n}X_{i-1})/\psi_0(\varepsilon_i)\right\}du\\
&= c\int_{|u|>d}|u|b(|u|/c)\,du,
\end{aligned}$$

which can be made arbitrarily small for d arbitrarily large by B(iii). This proves (2.35) for the first term in there. The proof for the second term of (2.35) is exactly similar.

Next, we turn to the proof of (2.34). Note that the LHS of (2.34) is bounded above by

$$I(A_{n,c})\sup_{|u|\leq d}\left|l_n(u)/l_n'(u) - 1\right|\int_{|u|\leq d}l_n'(u)\,du.$$

Moreover, (2.36) implies that $\int_{|u|\leq d}l_n'(u)\,du = O_p(1)$ for every $0 < d < \infty$. Thus to prove (2.34), it suffices to show that on $A_{n,c}$,

$$\sup_{|u|\leq d}\left|\prod_{i=k_n+1}^{n}\psi_0(\varepsilon_i - u\rho^{-n}X_{i-1})\Big/\psi_0\big(\varepsilon_i - u\rho^{-n+i-1}Y_{k_n}\big) - 1\right| = o_p(1). \tag{2.37}$$

But on $A_{n,c}$, for $k_n < i \leq n$,

$$\left|\rho^{-n}X_{i-1} - \rho^{-n+i-1}Y_{k_n}\right| = |\rho|^{-n+i-1}\left|\sum_{j=k_n+1}^{i-1}\rho^{-j}\varepsilon_j\right| \leq c|\rho|^{-n-k_n+i-1}.$$

Therefore, on $A_{n,c}$, by B(i), the

$$\begin{aligned}
\text{LHS (2.37)} &\leq \left\{\exp\left[\sum_{i=k_n+1}^{n}\sup_{|u|\leq d}h\big(\varepsilon_i - u\rho^{-n+i-1}Y_{k_n}\big)(dc)^{\alpha}|\rho|^{-\alpha(n+k_n-i+1)}\right] - 1\right\}\\
&\leq \left\{\exp\left[\sum_{i=1}^{n-k_n}h(|\varepsilon_{n-i+1}| + dc)|\rho|^{-\alpha i}|\rho|^{-\alpha k_n}(dc)^{\alpha}\right] - 1\right\} = o_p(1)
\end{aligned}$$

because, by Lemma 2.1 applied to $\eta_0 = 0$, $\eta = h(|\varepsilon| + dc)$, $\gamma = |\rho|^{-\alpha}$, the infinite series $\sum_{i=1}^{\infty} h(|\varepsilon_i| + dc)|\rho|^{-\alpha i} < \infty$ a.s., because $k_n|\rho|^{-\alpha k_n} \to 0$ and because $\{\varepsilon_{n-i+1}, i \geq 1\} =_{\mathscr{L}} \{\varepsilon_i, i \geq 1\}$. Note that we need B(ii) for this argument. This completes the proof of (2.34) and hence that of the lemma. □

The structure of the limiting experiment. Since this limiting experiment appears here to the best of our knowledge for the first time in the literature, we describe it in more detail.

Let $\Omega = (\mathbb{R} - \{0\}) \times \mathbb{R}^{\mathbb{N}}$ and let $\mu = \bar{\mu}_0 \times \mu_0 \times \mu_0 \times \mu_0 \times \cdots$ be an infinite product of measures on $\mathscr{B}_\Omega$, the product σ-algebra, where μ_0 has Lebesgue density ψ_0 and $\bar{\mu}_0$ is the distribution of

$$\sum_{i=1}^{\infty} \rho^{-i}\varepsilon_i, \quad \text{with } \varepsilon_i \text{ i.i.d. according to } \mu_0.$$

Since ε has density, Z has no mass at 0. Denote a typical point of Ω by $\omega = (z, e_1, e_2, \dots)$ and consider, for each $t \in \mathbb{R}$, the map $R_t\colon \Omega \to \Omega$ given by

$$R_t(\omega) = \left(z, e_1 + t\rho^{-1}z, e_2 + t\rho^{-2}z, \dots\right).$$

It is easily seen that $R_{s+t} = R_s \circ R_t$, hence $t \to R_t$ is a group homomorphism. Define a family of probability measures $\{P_t\}$ on $\mathscr{B}_\Omega$ by

$$P_t(A) := \mu\left(R_t^{-1}(A)\right) = \mu(R_{-t}(A)), \qquad A \in \mathscr{B}_\Omega.$$

The probabilities $\{P_t\}$ are mutually absolutely continuous with

$$\log(dP_t/dP_s)(\omega) = \sum_{i=1}^{\infty} \log\{\psi_0(e_i - t\rho^{-i}z)/\psi_0(e_i - s\rho^{-i}z)\}.$$

Since

$$\mathscr{L}\{dP_t/dP_s|P_s\} = \mathscr{L}\{(dP_t/dP_s) \circ R_s|P_0\} = \mathscr{L}\{dP_{t-s}/dP_0|P_0\},$$

the experiment $\mathscr{E} = (\Omega, \mathscr{B}_\Omega, \{P_t\}, t \in \mathbb{R})$ is translation invariant. By Remark 2.2 it follows that $\mathscr{E}$ is indeed the weak limit of the sequence of experiments $\mathscr{E}_n = (\mathbb{R}^n, \mathscr{B}^n, \{P_{\rho+\rho^{-n}t}\}, t \in \mathbb{R})$.

We now discuss the efficient estimation for the limiting experiment. Let

$$U(t) = e^{\Lambda(t|\psi_0)} = \prod_{i=1}^{\infty} \psi_0(\varepsilon_i' - t\rho^{-i}Z)/\psi_0(\varepsilon_i').$$

We know from Theorem 2.1 that $U(t)$ converges a.e. By conditioning first on Z we see that $E(U(t)|P_0) = 1$. By Remark 2.1 the convergence in Theorem 2.1 is uniform for bounded t's and therefore there is a version of the process $U(t)$ which has continuous paths. Moreover, $\int(1 + |t|)U(t)\,dt < \infty$ a.e. which follows from an argument identical to that used in (2.35) and (2.36). Hence it makes sense to speak of the posterior density process

$$V(t) := U(t)\Big/\int U(s)\,ds.$$

The shape of the trajectories of the process $V(t)$ depends heavily on the density ψ_0. An explicit expression for $V(t)$ is only known for normal densities.

LEMMA 2.3. *If* $\psi_0(x) = (2\pi\sigma^2)^{-1/2}\exp(-x^2/2\sigma^2)$, *then*

$$\Lambda(t|\psi_0) = (t/\sigma^2)WZ - \big(t^2/2\sigma^2(\rho^2-1)\big)Z^2$$

and

$$V(t) = (2\pi\sigma^2)^{-1/2}|Z|(\rho^2-1)^{-1/2}\exp\left\{-\frac{1}{2}\frac{Z^2}{\sigma^2(\rho^2-1)}\left[t-\frac{W}{Z}(\rho^2-1)\right]^2\right\},$$

where $W = \sum_{i=1}^{\infty}\rho^{-i}\varepsilon_i'$. *W and Z are independent* $N(0, \sigma^2(\rho^2-1)^{-1})$ *r.v.'s. Consequently, the sequence of experiments* $\{P^n_{\rho+\rho^{-n}t}\}$, $n \geq 1$, $|\rho| > 1$, $t \in \mathbb{R}$, *is locally asymptotically mixed normal (LAMN).*

PROOF. Follows by direct calculations from Theorem 2.1. □

From the above representation we see that for the normal error density ψ_0, all posterior densities $V(t)$ are normal with mean $(W/Z)(\rho^2-1)$ and variance $\sigma^2(\rho^2-1)/Z^2$. This implies that the Pitman estimator for quasiconvex loss function coincides with the maximum likelihood (ML) estimate. We simulated the process $V(t)$ for $\rho = \frac{4}{3}$. Five trajectories are shown in Fig. 1.

If the density ψ_0 is not normal, then the situation is completely different. We simulated the process $V(t)$ also for $\psi_0(t) = \text{const.}\exp(-x^4)$ and $\rho = \frac{4}{3}$ (see Fig. 2). Here, the trajectories do not have a center of symmetry, but they are skewed. Consequently, in contrast to the mixed-normal case with symmetric trajectories, we cannot expect that the Pitman estimates for different quasiconvex loss functions coincide. Also in the ML-estimate, the mode of $V(t)$ is

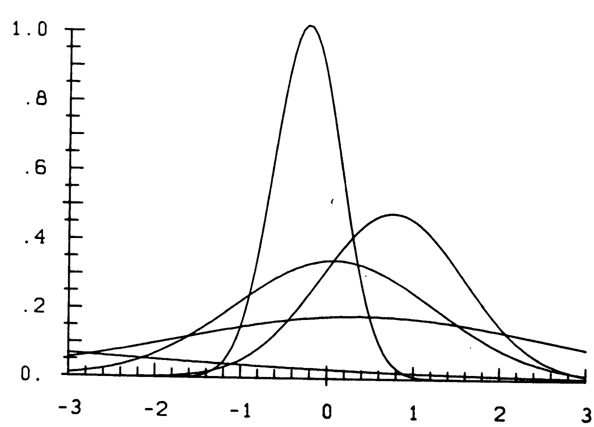

FIG. 1. *The posterior density process for normal* ψ_0.

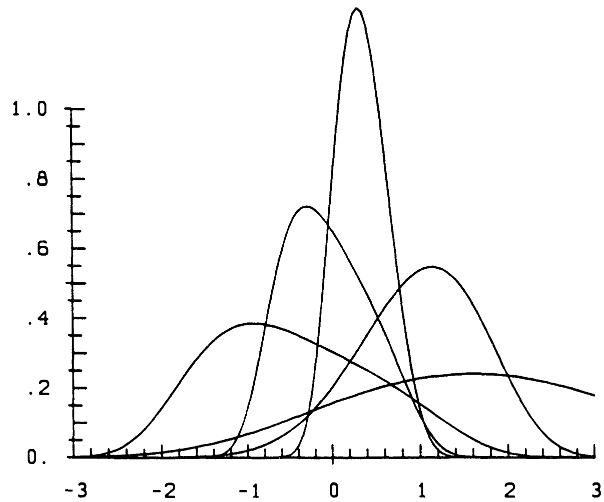

FIG. 2. *The posterior density process for* ψ_0 = const. $\exp(-x^4)$.

different from all Pitman estimates. By simulating 5000 trajectories of this process, we found that the Pitman estimate for the square loss has variance 0.95, whereas the variance of the ML-estimate did not converge and there is some evidence that it is infinity.

The question of efficient estimation of t in the limiting experiment $\mathscr{E}$ is answered by Theorem 2.2.

THEOREM 2.2. *Assume that* $E(\varepsilon^2) < \infty$. *Let*

$$T(\omega) = \varepsilon_1 \rho Z^{-1} - \rho Z^{-1} E\left(\varepsilon_1 \middle| \varepsilon_2 - \rho^{-1}\varepsilon_1, \varepsilon_3 - \rho^{-2}\varepsilon_1, \dots\right). \tag{2.38}$$

Then $T(\omega)$ *is a minimax estimate in the following sense*: *For every other estimate* T',

$$\sup_t E\left(Z^2(T'(\omega) - t)^2 \middle| P_t\right) \geq \sup_t E\left(Z^2(T(\omega) - t)^2\; P_t\right)$$
$$= \rho^2 E\left[\varepsilon_1 - E\left(\varepsilon_1 \middle| \varepsilon_2 - \rho^{-1}\varepsilon_1, \varepsilon_3 - \rho^{-2}\varepsilon_1, \dots\right)\right]^2.$$

PROOF. Let $S(\omega) = (z, 0, e_2 - \rho^{-1}e_1, e_3 - \rho^{-2}e_1, \dots)$. The transformation $S(\omega)$ is maximal invariant in the sense that $S(\omega_1) = S(\omega_2)$ iff there is a t with $R_t(\omega_1) = \omega_2$. Notice that $\varepsilon_1\rho/Z$ is an equivariant estimate. Thus, by Strasser (1985, Lemma 39.28), $[\rho\varepsilon_1/Z] - [\rho/Z]E(\varepsilon_1|S)$ is a minimax estimate for the quadratic loss, conditionally given Z. □

REMARK 2.3. The representation of the Pitman estimate in a form similar to (2.38) appears already in Girshick and Savage (1951).

3. Adaptive estimators. In this section we give a sequence of estimators of ρ that converge to T in distribution for a large class of densities. This sequence is obtained from $\tilde{\rho}_n$ of (2.21) by replacing ψ_0 in there by a density estimator. Accordingly, let g be a symmetric density with the support on $[-1, 1]$, Lipschitz of order 1 and with its almost everywhere derivative g' satisfying $\int |g'(t)|\, dt < \infty$. Let

$$\rho_n := \sum_{i=2}^{n} X_i X_{i-1} \Big/ \sum_{i=1}^{n-1} X_i^2$$

and

$$(3.1) \qquad \hat{\varepsilon}_{i,n} := X_i - \rho_n X_{i-1} = \varepsilon_i - \rho^n(\rho_n - \rho)\rho^{-n} X_{i-1}, \qquad i \geq 1.$$

Define the kernel-type density estimators

$$(3.2) \qquad \hat{\psi}_n(x) := (\sigma_n k_n)^{-1} \sum_{i=1}^{k_n} g\left(\frac{x - \hat{\varepsilon}_{i,n}}{\sigma_n}\right), \qquad x \in \mathbb{R},$$

where $0 < \sigma_n \to 0$ and k_n is as in Lemma 2.2. An adaptive sequence of estimators of ρ is defined to be

$$(3.3) \quad \hat{\rho}_n := \int u \prod_{i=k_n+1}^{n} \hat{\psi}_n(X_i - uX_{i-1})\, du \Big/ \int \prod_{i=k_n+1}^{n} \hat{\psi}_n(X_i - uX_{i-1})\, du.$$

Before proving the weak adaptivity of $\hat{\rho}_n$, we state:

LEMMA 3.1. *Let* (1.1) *hold with* $\{\varepsilon_i\}$ *i.i.d.* ψ_0. *Assume that* A(i) *holds. Then*

$$(3.4) \qquad \rho^n(\rho_n - \rho) \Rightarrow (\rho^2 - 1) W Z^{-1},$$

where W *is a r.v. independent of* Z, $W =_{\mathscr{L}} Z$, Z *as in* (2.12).

PROOF. The proof uses Corollary 2.1 and the Toeplitz lemma. Details are left out for the sake of brevity. □

REMARK 3.1. Under the finite second moment condition the consequence (3.4) is well known in the literature. The condition A(i) appears to be the weakest possible for (3.4) to hold. The Cauchy distribution satisfies A(i).

We now state and prove the main result that validates the adaptive nature of $\hat{\rho}_n$.

THEOREM 3.1. *Suppose that* (1.1) *holds with* $\{\varepsilon_i\}$ *i.i.d.* ψ_0. *Assume that* ψ_0 *satisfies* B *and*

(a) *for some* $0 < \delta < \frac{1}{2}$,

$$\int_0^\infty u[F_0(-u)]^{1/2}\,du < \infty, \qquad \int_0^\infty u[1 - F_0(u)]^{1/2-\delta}\,du < \infty.$$

(b) $$\int_{-\infty}^\infty h(u+\gamma)b(|u|)\,du < \infty, \quad \textit{for all } \gamma > 0.$$

About the sequences σ_n *and* k_n *assume that* $\sigma_n > 0$, $k_n \to \infty$, $(n - k_n) \to \infty$, $\sigma_n \to 0$ *and*

(c) $(n - k_n)\left[\sigma_n^{-1}|\rho|^{k_n - n} + \sigma_n^{-1}k_n^{-1/2} + \sigma_n^\alpha\right] \to 0$ *with* α *as in* A(ii).

Then

$$\rho^n(\hat{\rho}_n - \rho) \Rightarrow T, \tag{3.5}$$

where T *is as in* (2.26) *and* $\hat{\rho}_n$ *is as in* (3.3).

PROOF. By a change of variable, rewrite

$$\rho^n(\hat{\rho}_n - \rho) = \hat{T}_n := \frac{\int u \prod_{i=k_n+1}^n \hat{\psi}_n(\varepsilon_i - u\rho^{-n}X_{i-1})\,du}{\int \prod_{i=k_n+1}^n \hat{\psi}_n(\varepsilon_i - u\rho^{-n}X_{i-1})\,du}.$$

In view of (2.27) and (2.29), to prove (3.5), it suffices to show that

$$|\hat{T}_n - T_n| = o_p(1), \tag{3.6}$$

where T_n is as in (2.22). The proof of (3.6) involves several delicate calculations and nonuniform bounds. To begin with, define

$$\hat{F}_n(x) := k_n^{-1}\sum_{i=1}^{k_n} I(\hat{\varepsilon}_{in} \le x), \qquad F_n(x) := k_n^{-1}\sum_{i=1}^{k_n} I(\varepsilon_i \le x), \qquad x \in \mathbb{R}.$$

Then

$$\hat{\psi}_n(x) = \sigma_n^{-1}\int g((x-y)/\sigma_n)\,d\hat{F}_n(y).$$

Let, for $x \in \mathbb{R}$,

$$\psi_n(x) := \sigma_n^{-1}\int g((x-y)/\sigma_n)\,dF_n(y),$$

$$\tilde{\psi}_n(x) := \sigma_n^{-1}\int g((x-y)/\sigma_n)\,dF_0(y),$$

where F_0 is the distribution function corresponding to ψ_0. The idea is to approximate $\hat{\psi}_n$ by ψ_0 via $\tilde{\psi}_n$ and ψ_n.

From (3.1), Corollary 2.1 and Lemma 3.1,

$$\max_{1 \le i \le k_n} |\hat{\varepsilon}_{i,n} - \varepsilon_i| = O_p\big(|\rho|^{k_n - n}\big) \quad \text{as } n \to \infty.$$

From the well-known tightness results about the weighted empirical processes [see, e.g., Csörgő and Revesz (1981)], for $0 < \delta < \frac{1}{2}$,

$$\sup_x \left[\min(F_0(x), 1 - F_0(x))\right]^{-1/2+\delta} |F_n(x) - F_0(x)| = O_p\big(k_n^{-1/2}\big).$$

Now define, for a $c > 0$, the event

$$B_{n,c} := \Big\{ \max_{1 \le i \le k_n} |\hat{\varepsilon}_{i,n} - \varepsilon_i| \le c|\rho|^{k_n - n},\ c|\rho^{-n} X_{n-1}| - 1 \ge |\varepsilon_n|,\ \psi_0(\varepsilon_n) > c^{-1},$$

$$\sup_{x \in \mathbb{R}} \left[\min(F_0(x), 1 - F_0(x))\right]^{-1/2+\delta} |F_n(x) - F_0(x)| \le c k_n^{-1/2} \Big\} \cap A_{n,c}.$$

From the above discussion, (2.31) and Corollary 2.1, $\forall\, \eta > 0$, $\exists\, 0 < c < \infty$ and N_η such that

$$P(B_{n,c}) \ge 1 - \eta \quad \text{for all } n > N_\eta. \tag{3.7}$$

Further, choose n large such that

$$c|\rho|^{k_n - n} \le \delta/2, \qquad \sigma_n \le \delta/2. \tag{3.8}$$

From now on, we shall be working with n satisfying (3.7) and (3.8). On $B_{n,c}$,

$$F_n\big(x - c|\rho|^{k_n - n}\big) \le \hat{F}_n(x) \le F_n\big(x + c|\rho|^{k_n - n}\big), \qquad \forall\, x \in \mathbb{R}.$$

Hence for each $x \in \mathbb{R}$,

$$\begin{aligned} |\hat{F}_n(x) - F_0(x)| &\le \sup_{|y - x| \le c|\rho|^{k_n - n}} |F_n(y) - F_0(x)| \\ &\le \sup_{|y - x| \le \delta/2} |F_n(y) - F_0(y)| + \sup_{|y - x| \le c|\rho|^{k_n - n}} |F_0(y) - F_0(x)| \\ &\le c k_n^{-1/2} \{\min(F_0(x + \delta/2), 1 - F_0(x - \delta/2))\}^{1/2 - \delta} \\ &\quad + \big[F_0\big(x + c|\rho|^{k_n - n}\big) - F_0\big(x - c|\rho|^{k_n - n}\big)\big], \quad \text{on } B_{n,c}. \end{aligned} \tag{3.9}$$

Now apply this inequality with x replaced by $x - \sigma_n z$, $|z| \le 1$, to get, using $\sigma_n \le \delta/2$, that on $B_{n,c}$,

$$|\hat{F}_n(x - \sigma_n z) - F_0(x - \sigma_n z)| \le c\big[k_n^{-1/2} a_1(x) + |\rho|^{k_n - n} a_2(x)\big], \qquad x \in \mathbb{R},$$

where

$$a_1(x) := \{\min(F_0(x + \delta), 1 - F_0(x - \delta/2))\}^{1/2 - \delta}$$

$$a_2(x) := \begin{cases} b(|x| - \delta), & |x| \ge \delta, \\ b(0), & |x| < \delta. \end{cases}$$

Consequently, on $B_{n,c}$,

$$\begin{aligned}
&|\hat{\psi}_n(x) - \tilde{\psi}_n(x)| \\
&\leq \sigma_n^{-1} \int |\hat{F}_n(x - \sigma_n z) - F_0(x - \sigma_n z)||g'(x)|\,dz \qquad (3.10)\\
&\leq c^* \sigma_n^{-1}\{k_n^{-1/2} a_1(x) + |\rho|^{k_n - n} a_2(x)\}, \qquad c^* = c\int |g'|.
\end{aligned}$$

Next, use the absolute continuity of g to conclude that on $B_{n,c}$,

$$\begin{aligned}
|\hat{\psi}_n(x) - \psi_n(x)| &\leq \sigma_n^{-1} k_n^{-1} \sum_{i=1}^{k_n} \left| g\left(\frac{x - \varepsilon_i}{\sigma_n} - \frac{\hat{\varepsilon}_{i,n} - \varepsilon_i}{\sigma_n}\right) - g\left(\frac{x - \varepsilon_i}{\sigma_n}\right)\right| \qquad (3.11)\\
&\leq \sigma_n^{-2} k_n^{-1} \sum_{i=1}^{k_n} \int_{-c|\rho|^{k_n - n}}^{c|\rho|^{k_n - n}} \left| g'\left(\frac{x - \varepsilon_i + s}{\sigma_n}\right)\right| ds =: q_n(x).
\end{aligned}$$

To simplify notation a little, define

$$\begin{aligned}
\hat{\tau}_{ni}(u) &:= \hat{\psi}_n(\varepsilon_i - u\rho^{-n} X_{i-1})/\psi_0(\varepsilon_i), \qquad u \in \mathbb{R},\ i \geq 1,\\
\hat{\tau}_n &:= \prod_{i=k_n+1}^{n-1} \hat{\tau}_{ni}, \qquad (3.12)\\
\tau_n &:= \prod_{i=k_n+1}^{n-1} \tau_{ni}, \quad \text{where } \{\tau_{ni}\} \text{ are as in (2.32)}.
\end{aligned}$$

Then

$$\begin{aligned}
\hat{T}_n &= \int u \hat{\psi}_n(\varepsilon_n - u\rho^{-n} X_{n-1}) \hat{\tau}_n(u)\,du \Big/ \int \hat{\psi}_n(\varepsilon_n - u\rho^{-n} X_{n-1}) \hat{\tau}_n(u)\,du\\
&= \int u \prod_{i=k_n+1}^{n} \hat{\tau}_{ni}(u)\,du \Big/ \int \prod_{i=k_n+1}^{n} \hat{\tau}_{ni}(u)\,du,\\
T_n &= \int u \psi_0(\varepsilon_n - u\rho^{-n} X_{n-1}) \tau_n(u)\,du \Big/ \int \psi_0(\varepsilon_n - u\rho^{-n} X_{n-1}) \tau_n(u)\,du\\
&= \int u \prod_{i=k_n+1}^{n} \tau_{ni}(u)\,du \Big/ \int \prod_{i=k_n+1}^{n} \tau_{ni}(u)\,du.
\end{aligned}$$

To prove (3.6), in view of (2.35) and (3.7), it suffices to prove the following two results:

For every $\eta > 0$, there is $0 < d < \infty$ and N_η such that

$$EI(B_{n,c}) \int_{|u|>d} |u| \hat{\psi}_n(\varepsilon_n - u\rho^{-n} X_{n-1}) \hat{\tau}_n(u)\,du < \eta, \quad \text{for all } n > N_\eta. \qquad (3.13)$$

For every $0 < d < \infty$,

$$(3.14)\quad I(B_{n,c})\int_{|u|\le d}\left|\prod_{i=k_n+1}^{n}\hat{\tau}_{ni}(u) - \prod_{i=k_n+1}^{n}\tau_{ni}(u)\right|du = o_p(1), \quad \text{as } n\to\infty.$$

Recall that an analogue of (3.13) for T_n was proved in (2.35). Now, to prove (3.13), observe that on $B_{n,c}$,

$$|\varepsilon_n - u\rho^{-n}X_{n-1}| \ge |u|/c \quad \text{for } |u|\ge c.$$

Therefore, if $d \ge c$ and $|u| > d$, from (3.10), on $B_{n,c}$,

$$(3.15)\quad \begin{aligned}&\hat{\psi}_n(\varepsilon_n - u\rho^{-n}X_{n-1})\\ &\le \sup_{|t|\ge |u|c^{-1}}\tilde{\psi}_n(t) + c^*\sigma_n^{-1}\{k_n^{-1/2}a_1^*(|u|c^{-1}) + |\rho|^{k_n-n}a_1^*(|u|c^{-1})\},\end{aligned}$$

where

$$a_j^*(y) := \sup_{|t|\ge y} a_j(y), \qquad j = 1,2,\ y\ge 0.$$

But, because $\sigma_n \le \delta$ [see (3.8)],

$$(3.16)\quad \sup_{|t|\ge|u|c^{-1}}\tilde{\psi}_n(t) \le \int \sup_{|t|\ge|u|c^{-1}}\psi_0(t-\sigma_n z)g(z)\,dz \le b(|u|c^{-1}-\delta).$$

Note that $|u| > d \ge c$ and $0 < \delta < \frac{1}{2}$ ensures $|u| > c\delta$.

Next, with $\mathscr{F}_i = \sigma$-field $\{\varepsilon_1,\dots,\varepsilon_i\}$ let E_i denote the conditional expectation, given $\mathscr{F}_i$, $i\ge 1$. Then, with q_n as in (3.11), by Fubini's theorem,

$$(3.17)\quad \begin{aligned}&E_{j-1}[q_n(\varepsilon_j - u\rho^{-n}X_{j-1})/\psi_0(\varepsilon_j)]\\ &\le \sigma_n^{-2}k_n^{-1}\sum_{i=1}^{k_n}\int_{-c|\rho|^{k_n-n}}^{c|\rho|^{k_n-n}}\int\left|g'\big((x-u\rho^{-n}X_{j-1}-\varepsilon_i+s)/\sigma_n\big)\right|dx\,ds\\ &= 2\sigma_n^{-1}c|\rho|^{k_n-n}\int|g'|, \qquad k_n < j \le n-1,\ u\in\mathbb{R}.\end{aligned}$$

Note that $\{\varepsilon_i,\ i\le k_n\}$ are $\mathscr{F}_{j-1}$-measurable for all $k_n < j \le n-1$. Now recall that ψ_n is $\mathscr{F}_{k_n}$-measurable. Thus once again, for $k_n < j \le n-1$ and for all $u\in\mathbb{R}$,

$$(3.18)\quad E_{j-1}[\psi_n(\varepsilon_j - u\rho^{-n}X_{j-1})/\psi_0(\varepsilon_j)] = \int\psi_n(x-u\rho^{-j}X_{j-1})\,dx = 1.$$

Now let

$$L_n(x) := [\psi_n(x) + q_n(x)], \quad \text{with } q_n \text{ as in (3.11)}.$$

Then from (3.11), (3.17) and (3.18), and a repeated conditioning argument,

$$
\begin{aligned}
EI(B_{n,c})\hat{\tau}_n(u) &= EI(B_{n,c}) \prod_{j=k_n+1}^{n-1} \hat{\psi}_n(\varepsilon_j - u\rho^{-n}X_{j-1})/\psi_0(\varepsilon_j) \\
&\le E \prod_{j=k_n+1}^{n-1} L_n(\varepsilon_j - u\rho^{-n}X_{j-1})/\psi_0(\varepsilon_j) \\
&\le E \prod_{j=k_n+1}^{n-2} \{L_n(\varepsilon_j - u\rho^{-n}X_{j-1})/\psi_0(\varepsilon_j)\}(1 + 2c^*\sigma_n^{-1}|\rho|^{k_n-n}) \\
&= (1 + 2c^*\sigma_n^{-1}|\rho|^{k_n-n})^{n-k_n} = O(1), \qquad \forall\, u \in \mathbb{R},
\end{aligned} \tag{3.19}
$$

because $(n - k_n)\sigma_n^{-1}|\rho|^{k_n-n} \to 0$ by assumption (c). Now use (3.15), (3.16) and (3.19) to conclude that for $d \ge c$, and for all n sufficiently large, for all $n > N_\eta$ of (3.7), the LHS of (3.13) is bounded above by

$$
\begin{aligned}
O(1) \cdot \int_{|u|>d} |u|\{b(|u|c^{-1} - \delta) + k_n^{-1/2}\sigma_n^{-1}a_1^*(|u|c^{-1}) \\
+ |\rho|^{k_n-n}\sigma_n^{-1}a_2^*(|u|c^{-1})\}\, du.
\end{aligned} \tag{3.20}
$$

By definition, b is a decreasing function and hence

$$a_2^*(|u|c^{-1}) = b(|u|c^{-1} - \delta) \quad \text{for all} \quad |u| > d.$$

Moreover B(iii) implies $\int b(u)\,du < \infty$. Hence by B(iii),

$$\int |u|b(|u|c^{-1} - \delta)\,du = c^2 \int |u + \delta c|b(|u|)\,du < \infty.$$

Similarly, for $u > 0$,

$$
\begin{aligned}
a_1^*(u/c) &= \sup_{|t|>u/c} [\min(F_0(t+\delta), 1 - F_0(t-\delta))]^{1/2-\delta} \\
&\le \max[F_0(-uc^{-1} + \delta), 1 - F_0(uc^{-1} - \delta)]^{1/2-\delta},
\end{aligned}
$$

so that

$$\int |u|a_1^*(|u|c^{-1})\,du \le c^2 \int |u| \max[F_0(-|u| + c\delta), 1 - F_0(|u| - c\delta)]^{1/2-\delta}\,du < \infty$$

because by assumption (a), for all $k < \infty$,

$$\int_0^\infty u \max[F_0(-u + k), 1 - F_0(u - k)]^{1/2-\delta}\,du < \infty.$$

Consequently, from (3.20) and the above discussion there is a d large enough such that (3.13) holds.

Now turn to the proof of (3.14). Since the range of integration is now a bounded set, it is enough to get a uniform bound on the integrand that goes to zero in probability. To that effect, from (3.11) and the triangle inequality, on $B_{n,c}$,

$$|\hat{\psi}_n - \psi_0| \le |\psi_n - \psi_0| + q_n \le \{|\psi_n - \tilde{\psi}_n| + |\tilde{\psi}_n - \psi_0| + q_n\}.$$

Let $D_{n,c} := \{k_n^{1/2} \sup_x |F_n(x) - F_0(x)| \le c\}$. Note that $B_{n,c} \subset D_{n,c}$ and $D_{n,c}$ is $\mathscr{F}_{k_n}$-measurable. Hence for all $k_n < i \le n$, all $u \in \mathbb{R}$,

$$
\begin{aligned}
&E_{i-1}\left[I(B_{n,c})|\hat{\psi}_n(\varepsilon_i - u\rho^{-n}X_{i-1}) - \psi_0(\varepsilon_i - u\rho^{-n}X_{i-1})|/\psi_0(\varepsilon_i)\right] \\
&\quad \le \left\{\int|\psi_n(x) - \tilde{\psi}_n(x)|\,dx + \int|\tilde{\psi}_n(x) - \psi_0(x)|\,dx + 2c^*\sigma_n^{-1}|\rho|^{k_n-n}\right\} I(D_{n,c}) \\
&\quad \le \left\{\sigma_n^{-1}\int|F_n(x - \sigma_n z) - F_0(x - \sigma_n z)||g'(z)|\,dz + 2c^*\sigma_n^{-1}|\rho|^{k_n-n}\right. \\
&\qquad\qquad \left. + \int|\tilde{\psi}_n(x) - \tilde{\psi}_0(x)|\,dx\right\} I(D_{n,c}) \qquad (3.21) \\
&\quad \le 2c^*\sigma_n^{-1}\{k_n^{-1/2} + |\rho|^{k_n-n}\} + \int|\tilde{\psi}_n(x) - \psi_0(x)|\,dx \\
&\quad =: \delta_n, \quad \text{on } D_{n,c}.
\end{aligned}
$$

Next,

$$
\begin{aligned}
&\left|\prod_{i=k_n+1}^{n} \hat{\tau}_{ni}(u) - \prod_{i=k_n+1}^{n} \tau_{ni}(u)\right| I(B_{n,c}) \\
&\quad \le \sum_{i=k_n+1}^{n} \prod_{j=k_n+1}^{i-1} \hat{\tau}_{nj}(u)|\hat{\tau}_{ni}(u) - \hat{\tau}_{ni}(u)| \prod_{j=i+1}^{n} \tau_{nj}(u) I(B_{n,c}) \\
&\quad \le \sum_{i=k_n+1}^{n} \prod_{j=k_n+1}^{i-1} \left[\{\psi_n(\varepsilon_j - u\rho^{-n}X_{j-1}) + q_n(\varepsilon_j - u\rho^{-n}X_{j-1})\}/\psi_0(\varepsilon_j)\right] \\
&\qquad\qquad \times|\hat{\tau}_{ni}(u) - \tau_{ni}(u)| \prod_{j=i+1}^{n} \tau_{nj}(u) I(B_{n,c}).
\end{aligned}
$$

Now take expectation in the above inequality, carry out a conditioning argument, use the fact that $E_{j-1}\tau_{nj}(u) = 1$ for all $j \ge 2$ and for all u real, the fact that $D_{n,c}$ is $\mathscr{F}_{k_n}$-measurable and that $B_{n,c} \subset D_{n,c}$ to get

(3.22)

$$
\begin{aligned}
&E\left|\prod_{i=k_n+1}^{n} \hat{\tau}_{ni}(u) - \prod_{i=k_n+1}^{n} \tau_{ni}(u)\right| I(B_{n,c}) \\
&\quad \le \sum_{i=k_n+1}^{n} E \prod_{j=k_n+1}^{i-1} \left[\{\psi_n(\varepsilon_j - u\rho^{-n}X_{j-1})\right. \\
&\qquad\qquad \left. + q_n(\varepsilon_j - u\rho^{-n}X_{j-1})\}/\psi_0(\varepsilon_j)\right] I(D_{n,c})\delta_n \\
&\quad \le \left[1 + 2c^*\sigma_n^{-1}|\rho|^{k_n-n}\right]^{n-k}(n - k_n)\delta_n \quad \text{[from (3.17), (3.18)]} \\
&\quad = O(1) \cdot (n - k_n)\delta_n,
\end{aligned}
$$

because $(n - k_n)\sigma_n^{-1}|\rho|^{k_n-n} \to 0$ by assumption (c).

Next, for n large enough so that $\sigma_n \leq \theta$,

$$\int |\tilde{\psi}_n(x) - \psi_0(x)| \, dx \leq \int \int |\psi_0(x - \sigma_n z) - \psi_0(x)| g(z) \, dz \, dx$$
$$\leq \sigma_n^{\alpha} \int \sup_{|z| \leq 1} \psi_0(x - \sigma_n z) h(x) \, dx \quad [\text{by A(ii)}]$$
$$\leq \sigma_n^{\alpha} \int a_2(x) h(x) \, dx = O(\sigma_n^{\alpha})$$

because $\alpha_n \leq \delta$ and because assumption (b) implies that

$$\int a_2(x) h(x) \, dx = 2\int_0^{\delta} b(0) h(x) \, dx + 2\int_0^{\infty} b(x) h(x + \delta) \, dx < \infty.$$

Hence,

$$(n - k_n)\delta_n = (n - k_n) 2c^* \sigma_n^{-1}\left[k_n^{-1/2} + |\rho|^{k_n - n}\right] + (n - k_n) O(\sigma_n^{\alpha}) \to 0$$

by assumption (c).

Therefore, the LHS of (3.22) goes to zero. Moreover, $B_{n,c} \subset D_{n,c}$, $P(D_{n,c}) \geq 1 - \eta$ for all $n > N_{\eta}$ by (3.7). This together with the above argument proves (3.14). This also completes the proof of the theorem. □

REMARK 3.2. An example of sequences k_n and σ_n fulfilling the assumption (b) of Theorem 3.1 is

$$k_n = [n - \log n], \qquad \sigma_n = [\log n]^{-2/\alpha_0}, \quad \text{where } \alpha_0 \leq \alpha.$$

An example of a density satisfying Assumption A with $\theta = 1$, $\alpha = \frac{1}{2}$ and $h = 4$, and not having finite first moment and finite Fisher information is

$$\psi(x) = c(1 + e^{-\sqrt{|x|}})/(1 + x^2), \qquad x \in \mathbb{R}.$$

An example of a family of densities satisfying all assumptions of Section 3 and not having finite Fisher information is

$$\psi(x, \beta, \gamma) = c(\beta, \gamma)\exp(-\beta|x|^{\gamma}), \qquad 0 < \gamma_0 \leq \gamma \leq \tfrac{1}{2}, \beta > 0, x \in \mathbb{R}.$$

REFERENCES

BASAWA, I. V. and KOUL, H. L. (1979). Asymptotic tests of composite hypotheses for nonergodic type processes. *Stochostic Process. Appl.* **14** 41–54.

BERAN, R. (1976). Adaptive estimates for autoregressive processes. *Ann. Inst. Statist. Math.* **28** 77–89.

BERAN, R. J. (1978). An efficient and robust adaptive estimator of location. *Ann. Statist.* **6** 292–313.

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

CSÖRGŐ, M. and REVESZ P. (1981). *Strong Approximations in Probability and Statistics.* Academic, New York.

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

GIRSCHICK, M. A. and SAVAGE, L. J. (1951). Bayes and minimax estimates for quadratic loss function. *Proc. Second Berkeley Symp. Math. Statist. Prob.* **1** 53–73. Univ. California Press.

KOSHEVNIK, YU. A. and LEVIT, B. YA. (1976). On a nonparametric analogue of the information matrix. *Theory Probab. Appl.* **21** 738–753.

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

LE CAM, L. (1960). Locally asymptotically normal families of distributions. *Univ. California Publ. Statist.* **3** 37–98.

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

LEVIT, B. J. (1975). The efficiency of a certain class of nonparametric estimators. *Theory Probab. Appl.* **20** 723–740.

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

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

STRASSER, H. (1982). Local asymptotic minimax properties of Pitman estimates. *Z. Wahrsch. Verw. Gebiete* **60** 223–247.

STRASSER, H. (1985). *Mathematical Theory of Statistics*. de Gruyter, Berlin.

DEPARTMENT OF STATISTICS
AND PROBABILITY
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48824-1027

INSTITUT FÜR STATISTIK
UND INFORMATIK
UNIVERSITÄT WIEN
VIENNA
AUSTRIA