The Annals of Statistics
1997, Vol. 25, No. 5, 2127–2163

ON GUARANTEED ESTIMATION OF THE MEAN OF AN AUTOREGRESSIVE PROCESS[1]

BY V. KONEV AND S. PERGAMENSHCHIKOV

Tomsk University

This paper considers the problem of sequential point estimation of the mean of a stable autoregressive process with unknown scale and autoregressive parameters. The construction of a sequential procedure makes use of special stopping rules and some modifications of least-squares estimates. The procedure enables estimating the mean with prescribed mean-square accuracy uniformly in nuisance parameters. The uniform asymptotic normality and the asymptotic minimaxity of the sequential estimate are established. The asymptotic formula for the mean sample size is obtained.

1. Introduction. Consider the autoregressive AR(p) process

$$x_n - \mu = \lambda_1(x_{n-1} - \mu) + \cdots + \lambda_p(x_{n-p} - \mu) + d\varepsilon_n, \tag{1.1}$$

where $\{\varepsilon_n\}$ is a sequence of independent identically distributed (i.i.d.) random variables with $E\varepsilon_n = 0$ and $E\varepsilon_n^2 = 1$. The vector of initial values $\zeta = (x_0, \dots, x_{-p+1})'$ is assumed to be stochastically independent of the sequence $\{\varepsilon_n\}$; the prime denotes the transposition. The constant parameters $\mu, \lambda_1, \dots, \lambda_p, d$ are unknown with the vector $\lambda = (\lambda_1, \dots, \lambda_p)'$ belonging to the stability region Λ of the process (1.1); that is, all roots of the characteristic equation

$$z^p - \lambda_1 z^{p-1} - \cdots - \lambda_p = 0 \tag{1.2}$$

lie inside the unit circle. The problem is to estimate the parameter μ with preassigned mean-square accuracy in the presence of nuisance parameters $\lambda_1, \dots, \lambda_p, d$ by observations of the process x_n. The mean μ in the model (1.1) is customarily estimated by the sample average

$$\hat{\mu}_n = \frac{\sum_{i=1}^n x_i}{n}. \tag{1.3}$$

Asymptotic properties of estimate (1.3) for model (1.1) and more general stationary processes have been studied in detail [see, e.g., Anderson (1984)]. This estimate has several merits: it is unbiased; there are explicit equations for its variance; it is asymptotically normal. One can note, however, that there are some difficulties connected with using the estimate (1.3) if the parameters

---

Received September 1995; revised November 1996.

[1]Research supported by Russian Fund of Fundamental Investigations Grant 96-01-00189.

AMS 1991 *subject classifications*. 62L12, 62F12.

Key words and phrases. Autoregression, sequential estimation, fixed-precision estimators, nuisance parameters.

$\lambda_1, \ldots, \lambda_p, d$ are unknown. For example, in the case of an autoregression of the first order, the variance of $\hat{\mu}_n$ is

$$E_\theta(\hat{\mu}_n - \mu)^2 = \frac{d^2 \sum_{l=1}^n \sum_{k=1}^n \lambda^{|l-k|}}{n^2(1-\lambda^2)}, \tag{1.4}$$

$\theta = (\mu, \lambda, d)$; E_θ denotes the average by the distribution of the process (1.1) with given parameters μ, λ and d. When the parameter λ is unknown, one should use estimate (1.3) very carefully because it is no longer guaranteed in the sense that its mean-square accuracy is unknown. Furthermore, (1.4) implies that

$$\sup_{|\lambda|<1} E_\theta(\hat{\mu}_n - \mu)^2 = \infty, \tag{1.5}$$

that is, for any $n \geq 1$, one cannot guarantee the quality of estimate (1.3) of parameter μ without additional a priori information about parameter λ.

There arises the following natural question: is it possible in principal to estimate the mean of an autoregressive process AR(p) with preassigned mean-square accuracy when parameters $\lambda_1, \ldots, \lambda_p$ and d are unknown? A similar problem arose, probably for the first time, when estimating the mean of a normal population with unknown variance. The solution of this problem has been obtained by Stein (1945) who proposed the well-known two-stage sequential procedure. Since then estimation theory has been a fast-growing subject and many fundamental results have been proved [we refer to Govindarajulu (1987), Sen (1982), Siegmund (1985) and Woodroofe (1982)]. We cannot go into detail here and would like only to stress that most of the available literature considers the situation when observations are independent and there is very little known for processes with dependent values. Though the results on sequential estimation for the i.i.d. observation scheme cannot be applied directly to the above-stated problem, they are helpful in realizing that dependence of the mean-square accuracy of estimates of the mean μ on nuisance parameters $\lambda_1, \ldots, \lambda_p$ and d can be overcome only by employing sequential procedures. This paper proposes one of the approaches to estimation of the parameter μ in the process (1.1), which makes use of the results on point sequential estimates of parameters λ_i obtained by Konev and Pergamenshchikov (1981). The results on sequential estimation of parameters λ_i in autoregressive schemes and further references can be founded in Borisov and Konev (1977), Lai and Siegmund (1983), Sriram (1988), Konev and Lai (1995) and Dmitrienko and Konev (1994). Different approaches to estimation of the mean of an autoregressive process, taking into account the cost of observation, were proposed by Aras (1990), Fakhre-Zakeri and Lee (1993) and Sriram (1987).

The remainder of this paper is arranged as follows. In Section 2 we consider fixed precision estimators of the parameter μ of model (1.1) with unknown parameters $\lambda_1, \ldots, \lambda_p, d$. The sequential plan of estimating μ is a pair $(T(H),\ \tilde{\mu}(H))$, in which $T(H)$ is a special stopping rule, $\tilde{\mu}(H)$ is a sequential estimate of the parameter μ at the moment $T(H)$ and H is a positive parameter. It is shown (Theorem 2.1) that this plan has the following properties: for any $0 < L < \infty$ and $0 < H < \infty$,

$$T(H) < \infty, \qquad P_\theta \text{ a.s.}, \tag{1.6}$$
$$\sup_{\theta\in\Theta} E_\theta\big(\tilde{\mu}(H) - \mu\big)^2 \le \kappa/H,$$

where $\Theta = \{\theta = (\mu, \lambda, d)\colon |\mu - \mu_0| \le q,\ \lambda \in \Lambda,\ d \neq 0\}$, κ and μ_0 are known constants.

By the second inequality (1.6), parameter μ in model (1.1) can be estimated with any preassigned mean-square accuracy by making use of sequential plan $(T(H),\ \tilde{\mu}(H))$ for an appropriate value of the procedure parameter H. The construction of sequential estimates in Section 2 is rather cumbersome but there is one excuse for it. By (1.1), information about the parameter μ has to be extracted from observations x_n, satisfying the following equations:

$$x_n = m + \lambda_1 x_{n-1} + \cdots + \lambda_p x_{n-p} + d\varepsilon_n, \tag{1.7}$$
$$m = \mu(1 - \lambda_1 - \cdots - \lambda_p).$$

From here one can see that information about the parameter μ in the sample $(x_1, \dots, x_n)$ may be small if the sum $\lambda_1 + \cdots + \lambda_p$ is close to 1. Therefore even good estimates of parameters $\lambda_1, \dots, \lambda_p$ do not ensure by themselves fixed precision estimation of μ and specific stopping rules are needed.

In Section 3 the asymptotic behavior of the mean duration of the proposed sequential plan has been studied under the assumption that the parameter λ belongs to some compact set in the stability region Λ (Theorem 3.1). We refer the reader to Shiryaev and Spokoiny (1993) for the other results on asymptotic properties of stopping moments associated with sequential estimation of parameters in autoregression. In Section 4 uniform asymptotic normality of the estimate $\tilde{\mu}(H)$ is shown (Theorem 4.1). In Section 5 we prove asymptotic optimality in the minimax sense of the proposed estimate in a certain class of sequential and nonsequential procedures (Theorem 5.1).

2. Fixed-accuracy estimate for parameter μ.

We consider model (1.1) assuming that

$$\mu \in (\mu_0 - q, \mu_0 + q), \tag{2.1}$$

where $-\infty < \mu_0 < \infty$, $q > 0$, μ_0 and q are known constants; the distribution function of the noise ε_n has a bounded density $f(x)$, that is,

$$f(x) \le f^*, \qquad x \in (-\infty,\ \infty). \tag{2.2}$$

The mean μ will be estimated by observations $y_n = x_n - \mu_0$, which according to (1.7) satisfy the equations

$$y_n = m + \lambda_1 y_{n-1} + \cdots + \lambda_p y_{n-p} + d\varepsilon_n, \tag{2.3}$$

$$m = (\mu - \mu_0)(1 - 1'_p\lambda); \tag{2.4}$$

$1_p = (1, \dots, 1)'$ is the vector of dimension p. The basic idea of estimation of μ is to sample until enough information is gathered and to use least-squares estimates of parameters m and λ_i, $i = \overline{1, \dots, p}$. To realize it, we need a system of guaranteed estimates of these parameters. At first we assume that there exists a sequence $\tilde{a}_k = (\tilde{m}_k, \ \tilde{\lambda}'_k)'$, $k = 1, 2, \dots$ of estimates of the vector $a = (m, \lambda')'$ such that

$$\sup_{-\infty<\mu<\infty,\ \lambda\in\Lambda,\ d\neq 0} E_\theta \|\tilde{a}_k - a\|^2 \le e_k, \tag{2.5}$$

where $\{e_k\}$ is a numerical sequence with

$$\rho_1 = \sum_{k\ge 1} e_k < \infty; \tag{2.6}$$

$\|a\|^2 = a'a$. Then (2.4) may be rewritten as

$$\tilde{m}_k = (\mu - \mu_0)\tilde{\beta}_k + \tilde{\Delta}_k, \tag{2.7}$$

where

$$\begin{aligned} \tilde{\beta}_k &= 1 - 1'_p \tilde{\lambda}_k, \\ \tilde{\Delta}_k &= \tilde{m}_k - m + (\mu - \mu_0) 1'_p (\tilde{\lambda}_k - \lambda). \end{aligned} \tag{2.8}$$

Here $\tilde{\Delta}_k$ is an unobservable sequence (noise). Now we apply the sequential least-squares method to estimate the parameter μ from (2.7). Define a stopping rule

$$\begin{gathered} \nu = \nu(H) = \inf\left\{ l \ge 1\colon \sum_{k=1}^{l} \tilde{\beta}_k^2 \ge H \right\}, \qquad 0 < H < \infty, \\ (\inf\{\varnothing\} = +\infty) \end{gathered} \tag{2.9}$$

and the sequential estimate for the parameter μ at the moment ν (on the set where $\nu < +\infty$)

$$\tilde{\mu} = \tilde{\mu}(H) = \sum_{k=1}^{\nu} \tilde{\beta}_k \tilde{m}_k \Big/ \sum_{k=1}^{\nu} \tilde{\beta}_k^2 + \mu_0. \tag{2.10}$$

According to (2.9) the sample volume is measured in terms of accumulated energy of "the useful signal" $\tilde{\beta}_k$.

PROPOSITION 2.1. *Let a sequence of estimates* $\tilde{a}_k = (\tilde{m}_k, \tilde{\lambda}'_k)'$, $k = 1, 2, \dots$ *of the vector* $a = (m, \lambda')'$ *in* (2.3) *satisfy conditions* (2.5), (2.6).

Then for any $0 < q < \infty$, $-\infty < \mu_0 < \infty$ *and* $0 < H < \infty$, *the sequential plan* (2.9), (2.10) *has the following properties*:

$$\begin{gathered} \nu(H) < \infty, \qquad P_\theta\text{-a.s.}, \qquad \lambda \in \Lambda; \\ \sup_{|\mu-\mu_0|\le q,\ \lambda\in\Lambda} E_\theta (\tilde{\mu}(H) - \mu)^2 \le \kappa / H, \qquad \kappa = (1 + pq^2)\rho_1. \end{gathered} \tag{2.11}$$

PROOF. Under conditions (2.5) and (2.6), the estimate $\tilde{a}_k$ is strongly consistent and for all $\mu \in (-\infty, \infty)$ and $\lambda \in \Lambda$,

$$\lim_{k\to\infty} \tilde{\beta}_k = 1 - \lambda_1 - \cdots - \lambda_p > 0 \quad \text{a.s.} \tag{2.12}$$

Therefore the first inequality, (2.11), holds. From (2.9) and (2.10) by the Cauchy–Bunyakovskii inequality, we obtain

$$E_\theta\big(\tilde{\mu}(H) - \mu\big)^2 = E_\theta\bigg(\sum_{k=1}^{\nu} \tilde{\beta}_k \tilde{\Delta}_k \bigg/ \sum_{k=1}^{\nu} \tilde{\beta}_k^2\bigg)^2$$
$$\leq E_\theta \sum_{k=1}^{\nu} \tilde{\Delta}_k^2/H \leq \kappa/H, \qquad |\mu - \mu_0| \leq q, \ \lambda \in \Lambda.$$

Hence Proposition 2.1. □

To complete the construction of a sequential estimate for the parameter μ, it remains to choose a sequence of estimates $\{\tilde{a}_k,\ k \geq 1\}$ satisfying conditions (2.5), (2.6). Note that ordinary least-square estimates of parameters $m, \lambda_1, \dots, \lambda_p$ do not fit, because in the general case their mean-square accuracies depend on the values of these parameters. We will make a choice of estimates $\tilde{a}_k$ with the required properties by exploiting the method of guaranteed estimation of autoregressive parameters proposed by Konev and Pergamenshchikov (1981) and its modification for the case of unknown variance [Dmitrienko and Konev (1994)]. This method is realized in two steps.

STEP 1. Let us introduce a system of stopping moments $\{\tau(z),\ z > 0\}$ by

$$\tau(z) = \inf\bigg\{l > r(z)\colon \sum_{j=r(z)+1}^{l} \|Y\|^2 \geq zR(z)\bigg\}, \qquad z > 0,$$
$$Y_j = (1, y_{j-1}, \dots, y_{j-p})', \tag{2.13}$$
$$R(z) = \sum_{j=1}^{r(z)} y_j^2/r(z) + 1, \qquad r(z) = 6 + [z/l(z)],$$

where $[a]$ denotes the whole part of a number a. Here $l\colon [0, \infty) \to [0, \infty)$ is a slowly increasing function in the sense that

$$\lim_{z\to\infty} l(z) = \infty, \qquad \lim_{z\to\infty} l(z)/z^\gamma = 0 \tag{2.14}$$

for any $\gamma > 0$ and the function $z/l(z)$ is also assumed to be increasing [for example, $l(z) = \ln(z + e)$].

For each moment $\tau(z)$ we define the modified least-square estimate as

$$a(z) = \begin{pmatrix} m(z) \\ \lambda(z) \end{pmatrix} = G^+(z)\bigg(\sum_{j=r(z)+1}^{\tau(z)-1} Y_j y_j + \alpha(z) Y_{\tau(z)} y_{\tau(z)}\bigg), \tag{2.15}$$

where

$$G(z) = \sum_{j=r(z)+1}^{\tau(z)-1} Y_j Y_j' + \alpha(z) Y_{\tau(z)} Y_{\tau(z)}';$$

G^+ denotes the inverse of the matrix G if $\det G > 0$ and $G^+ = 0$ otherwise; $\alpha(z)$ is a weight multiplier which is determined from the equation

$$\sum_{j=r(z)+1}^{\tau(z)-1} \|Y_j\|^2 + \alpha(z) \|Y_{\tau(z)}\|^2 = zR(z). \tag{2.16}$$

Note that $0 < \alpha(z) \le 1$.

REMARK. The estimate (2.15) for the vector $a = (m, \lambda')'$ is based on the sample $(Y_1, \ldots, Y_\tau)$ of random size τ. One can see that the initial sample $(Y_1, \ldots, Y_r)$ of nonrandom size $r(z)$ is used only to calculate the function $R(z)$ in (2.13). This function is due to the unknown variance of the noise in model (1.1). [see Dmitrienko and Konev (1994)]. The estimate (2.15) coincides with the corresponding LSE of the vector a based on the sample $(Y_{r+1}, \ldots, Y_\tau)$ if $\alpha(z) = 1$. The weight multiplier $\alpha(z)$ and the function $l(z)$ are needed to ensure some nice properties of the estimates.

The system of estimates (2.13), (2.15) is the basis for construction of the sequential estimate for the vector a. We need also the following function:

$$b(z) = \begin{cases} R^{-2}(z) z^{-2} \|G^{-1}(z)\|^{-2}, & \text{if } \det G(z) > 0; \\ 0, & \text{otherwise}; \end{cases} \tag{2.17}$$

$\|A\|^2 = \operatorname{tr} AA'$.

STEP 2. Let $\{c_k\}$ be a nondecreasing sequence of positive constants such that

$$\rho_2 = \sum_{k \ge 1} 1/c_k < \infty. \tag{2.18}$$

By putting $z = c_k$ in (2.13), (2.15) and (2.17), we obtain

$$\tau_k = \tau(c_k), \qquad a_k = a(c_k), \qquad b_k = b(c_k), \qquad r_k = r(c_k). \tag{2.19}$$

Define the sequential estimate of the vector a as a weighted average

$$a^*(h) = \frac{\sum_{k=1}^{\sigma} b_k a_k}{\sum_{k=1}^{\sigma} b_k}, \tag{2.20}$$

where σ determines the number of estimates a_k used in the average. Define σ as

$$\sigma = \sigma(h) = \inf\left\{ l \ge 1 \colon \sum_{k=1}^{l} b_k \ge h \right\}, \qquad 0 < h < \infty. \tag{2.21}$$

The sample size which is required for calculating the estimate (2.20) is

$$N(h) = \tau(c_\sigma). \tag{2.22}$$

Further, we shall need the property of guaranteed accuracy of estimates (2.20), (2.22) contained in the following assertion.

PROPOSITION 2.2. *Let process* (1.1) *be stable, that is,* $\lambda \in \Lambda$, *and* $\{\varepsilon_n\}$ *be a sequence of i.i.d. random variables with* $E\varepsilon_n = 0$, $E\varepsilon_n^2 = 1$ *and the density* $f(x)$ *satisfying* (2.2).

Then for any $h > 0$

$$N(h) < \infty, \qquad P_\theta\text{-a.s.}, \tag{2.23}$$

$$\sup_{-\infty<\mu<\infty,\, d\neq 0,\, \lambda\in\Lambda} E_\theta \|a^*(h) - a\|^2 \le \rho_3/h, \tag{2.24}$$

where

$$\begin{aligned} &\rho_3 = \pi(f^*)^2 \sum_{k\ge 1} g_k/c_k, \qquad g_k = \frac{r_k^2}{(\Gamma(r_k/2))^{2/r_k}(r_k-2)}, \\ &\Gamma(a) = \int_0^\infty t^{a-1}\exp(-t)\,dt, \qquad a > 0. \end{aligned} \tag{2.25}$$

PROOF. It is well known [see, for example, Anderson (1984)] that for $-\infty < \mu < \infty$ and $\lambda \in \Lambda$, the limiting relationship

$$\lim_{N\to\infty} \frac{\sum_{i=1}^N Y_i Y_i'}{N} = F \tag{2.26}$$

holds. Here

$$\begin{aligned} &F = F(\theta) = \begin{bmatrix} 1 & \vdots & (\mu-\mu_0)1_p' \\ \cdots & \vdots & \cdots \\ (\mu-\mu_0)1_p & \vdots & (\mu-\mu_0)^2 1_p 1_p' + F_0 \end{bmatrix}, \\ &F_0 = F_0(\theta) = \sum_{k\ge 0} A^k B (A')^k, \\ &A = \begin{bmatrix} \lambda_1 & \cdots & \lambda_p \\ I_{p-1} & ; & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} d^2 & 0\ldots 0 \\ \cdots & \cdots \\ 0 & 0\ldots 0 \end{bmatrix}; \end{aligned} \tag{2.27}$$

I_p is a unit matrix of order p; F is a positive definite matrix. From (2.17), (2.19) and (2.26) we have

$$\begin{aligned} &\sum_{i\ge 1} \|Y_i\|^2 = \infty, \qquad P_\theta\text{-a.s.}, \\ &\lim_{z\to\infty} b(z) = b_0 = b_0(\theta) = \left(\operatorname{tr} F \|F^{-1}\|\right)^{-2}, \qquad P_\theta\text{-a.s.} \end{aligned} \tag{2.28}$$

From here and (2.19), (2.21) and (2.22), we obtain (2.23). From (2.15), (2.19) and (2.3) it follows that

$$\|a^* - a\|^2 \le \left(\sum_{k=1}^{\sigma} b_k \|a(c_k) - a\|/h \right)^2 \\ \le \sum_{k=1}^{\sigma} b_k \|a(c_k) - a\|^2 \Big/ h \le d^2 \sum_{k=1}^{\infty} \|M(c_k)\|^2 / c_k^2 R^2(c_k), \tag{2.29}$$

where

$$M(z) = \sum_{j=r(z)+1}^{\tau(z)-1} Y_j \varepsilon_j + \alpha(z) Y_{\tau(z)} \varepsilon_{\tau(z)}. \tag{2.30}$$

Taking into account (2.16) we have

$$E_\theta\big(\|M(z)\|^2 \big| Y_1, \dots, Y_{r(z)+1}\big) \\ \le E_\theta\left(\sum_{j=r(z)+1}^{\tau(z)-1} \|Y_j\|^2 + \alpha(z) \|Y_{\tau(z)}\|^2 \middle| Y_1, \dots, Y_{r(z)+1} \right) = zR(z).$$

Combining this estimate with (2.28) we obtain

$$E_\theta \|a^* - a\|^2 \le \frac{1}{h} \sum_{k\ge 1} \frac{1}{c_k} E_\theta \frac{d^2}{R(c_k)} \\ \le \frac{1}{h} \sum_{k\ge 1} \frac{r_k}{c_k} E_\theta \frac{d^2}{\sum_{j=1}^{r_k} y_j^2} \le \frac{\pi (f^*)^2 \sum_{k\ge 1} g_k / c_k}{h}. \tag{2.31}$$

The last inequality is due to condition (2.2) and can be proved along the lines of Proposition 2 in Dmitrienko and Konev (1994). Hence Proposition 2.2. □

REMARK. The sequence g_k in (2.25) is convergent:

$$\lim_{k\to\infty} g_k = 2e$$

and is bounded:

$$\max_{k\ge 1} g_k \le g_0, \qquad g_0 = 3\exp(1 + 1/4\pi e). \tag{2.32}$$

Now we can define the sequence of estimates $\{a_k\}$ satisfying conditions (2.5), (2.6). Let

$$\tilde{a}_k = (\tilde{m}_k, \tilde{\lambda}_k')' = a^*(h_k) = \big(m^*(h_k), (\lambda^*(h_k))'\big)', \tag{2.33}$$

where $\{h_k\}$ is a nondecreasing sequence of positive numbers such that

$$\rho_4 = \sum_{k\ge 1} 1/h_k < \infty. \tag{2.34}$$

The sequence of estimates (2.33) satisfies (2.5), (2.6) with

$$e_k = \rho_3 / h_k. \tag{2.35}$$

If one uses the estimate of the mean μ of the process (1.1), defined by (2.9), (2.10) and (2.33), then the total duration of the estimation procedure is

$$T(H) = N(h_{\nu(H)}) = \tau(c_{\sigma(H)}), \qquad H > 0. \tag{2.36}$$

THEOREM 2.1. *Let process* (1.1) *be stable and* $\{\varepsilon_n\}$ *be a sequence of i.i.d. random variables with density* $f(x)$ *satisfying* (2.2) *and* $E\varepsilon_n = 0$, $E\varepsilon_n^2 = 1$. *Then the sequential plan* $(T(H),\ \tilde{\mu}(H))$ *defined by* (2.9), (2.10), (2.33) *and* (2.36) *possesses the properties*: *for any* $0 < H < \infty$ *and* $0 < q < \infty$,

$$T(H) < \infty, \qquad P_\theta\text{-}a.s., \qquad \lambda \in \Lambda, \tag{2.37}$$

$$\sup_{\theta \in \Theta} E_\theta \big(\tilde{\mu}(H) - \mu\big)^2 \le \kappa_1 / H, \tag{2.38}$$

where $\Theta = \{\theta = (\mu, \lambda', d) \colon |\mu - \mu_0| \le q,\ \lambda \in \Lambda,\ d \ne 0\}$, $\kappa_1 = (1 + pq^2)\rho_3\rho_4$.

The assertion of Theorem 2.1 follows directly from Propositions 2.1, 2.2.

3. The mean duration of the procedure.

In this section we will study asymptotic properties of the moment (2.36) as $H \to \infty$. In addition to conditions imposed in Sections 1 and 2 on the distribution of the noise ε_n in (1.1), we assume that

$$E\varepsilon_1^8 < \infty, \qquad E\|\zeta\|^8 < \infty. \tag{3.1}$$

Also the sequences $\{c_k\}$ and $\{h_k\}$ will be chosen as functions of the parameter H. Let

$$\begin{gathered} c_k = c_k(H) = \begin{cases} H, & k \le n_0(H); \\ c_k^*, & k > n_0(H); \end{cases} \\ h_k = h_k(H) = c_k(H); \end{gathered} \tag{3.2}$$

where $c_k^* = k^{1+\delta}$, $0 < \delta < \sqrt{2} - 1$; $n_0(H) = [Hl(H)]$ and the function $l(H)$ satisfies (2.14). Such choice of these sequences is associated with the asymptotic behavior of the basic least-squares estimates of the vector a and their sequential modifications (2.15). As a result, the two-step procedure (2.20), (2.22) transforms, as will be shown below, into the sequential estimate (2.15) with $z = H$ and for large H,

$$a^*(H) \approx a(H), \qquad T(H) \approx \tau(H).$$

The following theorem assesses the performance of the procedure duration $T(H)$.

THEOREM 3.1. *Under the conditions of Theorem* 2.1 *and* (3.1), *the stopping moment* (2.36) *satisfies the limiting relationship*

$$\lim_{H \to \infty} \sup_{\theta \in \mathscr{K}} E_\theta \left| \frac{T(H)}{H} - \frac{d_1}{\operatorname{tr} F} \right| = 0, \tag{3.3}$$

for any $-\infty < \mu_0 < \infty$, *and* $q > 0$, *where* $\mathscr{K}$ *is an arbitrary compact set in the region*

$$\Theta = \{\theta = (\mu, \lambda', d): |\mu - \mu_0| \le q,\ \lambda \in \Lambda,\ d \ne 0\}; \tag{3.4}$$
$$d_1 = 1 + \langle F \rangle_{22}.$$

F *is a positive definite matrix defined by* (2.27); $\langle F \rangle_{ij}$ *is the* (i, j)*th element of the matrix* F.

REMARK. By Theorem 3.1 for any fixed $\mathscr{K} \subset \Theta$ and sufficiently large H,

$$E_\theta T(H) \approx \frac{d_1 H}{\operatorname{tr} F}.$$

To prove Theorem 3.1 we need several auxiliary propositions. First we will establish some properties of stopping moments $\sigma(H)$ and $\nu(H)$. The moment $\sigma(H)$, defined in (2.21), depends on the function $b(z)$, satisfying (2.28). The following lemma gives the rate of convergence of $b(z)$ to its limit.

LEMMA 3.1. *Under the conditions of Theorem* 3.1,

$$\limsup_{z \to \infty} \sup_{\theta \in \mathscr{K}} z^2 E_\theta (b(z) - b_0(\theta))^4 < \infty. \tag{3.5}$$

Proof of Lemma 3.1 is given in the Appendix.

PROPOSITION 3.1. *Let the conditions of Theorem* 3.1 *hold. Then for sufficiently large* H *and for all* $m \ge n_0(H)$,

$$\sup_{\theta \in \mathscr{K}} P_\theta(\sigma(H) > m) \le L l(H)(n_0^3(H)/H m^4 + m^{-2-\delta}), \tag{3.6}$$

where $0 < L < \infty$, $l(H)$ *is defined in* (2.14), δ *is the same as in* (3.2).

PROOF. By definition of $\sigma(H)$ and (2.28) we have

$$\begin{aligned} P_\theta(\sigma(H) > m) &\le P_\theta\left(\sum_{k=1}^{m} b(c_k) < H\right) \\ &\le P_\theta\left(\sum_{k=1}^{m} |b(c_k) - b_0(\theta)| > m b_* - H\right), \end{aligned} \tag{3.7}$$

where

$$b_* = \inf_{\theta \in \mathscr{K}} b_0(\theta).$$

Further note that for sufficiently large H and all $m \ge n_0(H)$,

$$m b_*/2 - H \ge 0.$$

This inequality follows from (3.2) and (2.14). From here and (3.7), with the help of the Chebyshev and Hölder inequalities, we obtain

$$\begin{aligned}
P_\theta(\sigma(H) > m) &\le P_\theta\left(\sum_{k=1}^{m} |b(c_k) - b_0(\theta)| > \frac{mb_*}{2}\right)\\
&\le \frac{16E_\theta(\sum_{k=1}^{m} |b(c_k) - b_0(\theta)|)^4}{(b_* m)^4}\\
&\le \frac{16}{(b_* m)^4}\left(\sum_{k=1}^{m}\left(\frac{1}{c_k}\right)^{1/3}\right)^3 \sum_{k=1}^{m} \frac{E_\theta(\sqrt{c}_k |b(c_k) - b_0(\theta)|)^4}{c_k}.
\end{aligned}$$

By applying Lemma 3.1 and taking into account the choice of the sequence c_k in (3.2) we come to (3.6). Hence Proposition 3.1. □

In order to prove a similar property for the moment $\nu(H)$, we need the following result.

LEMMA 3.2. *Under the conditions of Theorem* 3.1 *the fourth moment of the deviation of the estimate* (2.20) *satisfies the inequality*

$$\sup_{\theta \in \mathscr{K}} E_\theta \|a^*(h) - a\|^4 \le Ll^2(H)/h^2 \tag{3.8}$$

for all $h > 0$ *and sufficiently large* H.

The proof of Lemma 3.2 is given in the Appendix.

PROPOSITION 3.2. *Under the conditions of Theorem* 3.1 *for sufficiently large* H *and all* $m \ge n_0(H)$,

$$\sup_{\theta \in \mathscr{K}} P_\theta\{\nu(H) > m\} \le Ll^3(H)\big(n_0^3(H)/m^4 H + m^{-2-\delta}\big). \tag{3.9}$$

PROOF. From the definition of the stopping moment $\nu(H)$ in (2.9),

$$\begin{aligned}
P_\theta\{\nu(H) > m\} &= P_\theta\left\{\sum_{k=1}^{m} \tilde\beta_k^2 < H\right\} = P_\theta\left\{m\beta^2 + \sum_{k=1}^{m} (\tilde\beta_k^2 - \beta^2) < H\right\}\\
&= P_\theta\left\{m\beta^2 + \sum_{k=1}^{m} (\tilde\beta_k - \beta)^2 + 2\beta \sum_{k=1}^{m} (\tilde\beta_k - \beta) + < H\right\}\\
&\le P_\theta\left\{m\beta^2 + 2\beta \sum_{k=1}^{m} (\tilde\beta_k - \beta) < H\right\}\\
&\le P_\theta\left\{2\beta \sum_{k=1}^{m} \big|1_p'(\tilde\lambda_k - \lambda)\big| > m\beta_* - H\right\},
\end{aligned}$$

where

$$\beta_* = \inf_{\lambda \in \mathscr{K}_1} \beta^2(\lambda), \qquad \mathscr{K}_1 = \big\{\lambda \in \Lambda\colon \theta = (\theta, \lambda', d) \in \mathscr{K}\big\};$$

$\beta(\lambda)$ and $\tilde{\lambda}_k$ are defined by (2.12) and (2.33). Note that in view of (2.8) and (2.14) for sufficiently large H and all $m \geq n_0(H)$,

$$m\beta_* - H \geq m\beta_*/2.$$

Therefore we have

$$P_\theta\{\nu(H) > m\} \leq P_\theta\left\{2\sqrt{p}\sum_{k=1}^{m}\|\tilde{\lambda}_k - \lambda\| > m\beta_*/2\right\}$$
$$\leq P_\theta\left\{4\sqrt{p}\sum_{k=1}^{m}\|a^*(h_k) - a\| > m\beta_*\right\}.$$

By applying the Chebyshev inequality, Lemma 3.1 and substituting (3.2) we obtain (3.9). Hence Proposition 3.2. □

Further, we need the following technical result.

LEMMA 3.3. *Let y_n be defined by* (2.3). *Then for any $n \geq 1$ and all $\theta \in \mathscr{K}$,*

$$\sum_{k=1}^{n} y_k^2 \leq L\left(\sum_{k=1}^{n}\varepsilon_k^2 + n + \|\zeta\|^2\right), \tag{3.10}$$

where L is some constant which does not depend on θ.

Proof of Lemma 3.3 is given in the Appendix.

PROPOSITION 3.3. *Under the conditions of Theorem* 3.1 *for some $\gamma > 0$,*

$$\limsup_{H\to\infty}\sup_{\theta\in\mathscr{K}} E_\theta(T(H)/H)^{(1+\gamma)} < \infty. \tag{3.11}$$

PROOF. For notational simplicity we shall write $v = v(H)$ instead of $c_{\sigma(h_{\nu(H)})}(H)$. By (2.13),

$$\tau(H) \leq HR(H) + r(H) + 1 \tag{3.12}$$

and, hence, by (2.36),

$$T(H) \leq vR(v) + r(v).$$

Therefore it suffices to show that

$$\limsup_{H\to\infty}\sup_{\theta\in\mathscr{K}} E_\theta(r(v)/H)^{(1+\gamma)} < \infty, \tag{3.13}$$

$$\limsup_{H\to\infty}\sup_{\theta\in\mathscr{K}} E_\theta(vR(v)/H)^{(1+\gamma)} < \infty \tag{3.14}$$

for some $\gamma > 0$.

Since $r(z) \leq z + 6$, then the inequality (3.13) holds if

$$\limsup_{H\to\infty}\sup_{\theta\in\mathscr{K}} E_\theta(v(H)/H)^{(1+\gamma)} < \infty. \tag{3.15}$$

The proof of this inequality is given in the Appendix. Consider (3.14). By definition of $R(v)$ in (2.13) and Lemma 3.2,

$$vR(v) \le L\left(v \sum_{j=1}^{r(v)} \varepsilon_j^2 / r(v) + v + \|\zeta\|^2 v / r(v) \right) \le L\left(l(v) \sum_{j=1}^{r(v)} \varepsilon_j^2 + v + \|\zeta\|^2 v / r(v) \right).$$

Therefore it suffices to verify that

$$\limsup_{H\to\infty} \sup_{\theta\in\mathscr{K}} E_\theta \left(l(v) \sum_{j=1}^{r(v)} \varepsilon_j^2 / H \right)^{(1+\gamma)} < \infty, \tag{3.16}$$

$$\limsup_{H\to\infty} \sup_{\theta\in\mathscr{K}} E_\theta \bigl(l(v) \|\zeta\|^2 / H \bigr)^{(1+\gamma)} < \infty. \tag{3.17}$$

We have

$$\begin{aligned} E_\theta \left(l(v) \sum_{j=1}^{r(v)} \varepsilon_j^2 \right)^{(1+\gamma)} &= (1+\gamma) \int_0^\infty t^\gamma P_\theta \left\{ l(v) \sum_{j=1}^{r(v)} \varepsilon_j^2 > t \right\} dt = I_1 + I_2; \\ I_1 &= (1+\gamma) \int_0^\infty t^\gamma P_\theta \left\{ l(v) \sum_{j=1}^{r(v)} \varepsilon_j^2 > t,\ v \le t/2 \right\} dt; \\ I_2 &= (1+\gamma) \int_0^\infty t^\gamma P_\theta \{ v > t/2 \}\, dt. \end{aligned} \tag{3.18}$$

Taking into account (2.14), we obtain

$$\begin{aligned} P_\theta \left\{ l(v) \sum_{j=1}^{r(v)} \varepsilon_j^2 > t,\ v \le t/2 \right\} &\le P_\theta \left\{ l(t/2) \sum_{j=1}^{r(t/2)} \varepsilon_j^2 > t \right\} \\ &= P_\theta \left\{ r(t/2) l(t/2) + l(t/2) \sum_{j=1}^{r(t/2)} (\varepsilon_j^2 - 1) > t \right\} \\ &\le P_\theta \left\{ t/2 + 6 l(t/2) + l(t/2) \sum_{j=1}^{r(t/2)} (\varepsilon_j^2 - 1) > t \right\} \\ &\le \frac{2^4 l^4(t/2) E_\theta \bigl(\sum_{j=1}^{r(t/2)} (\varepsilon_j^2 - 1) \bigr)^4}{(t - 12 l(t/2))^4} \\ &\le \frac{2^4 l^4(t/2) r^2(t/2) E_\theta (\varepsilon_1^2 - 1)^4}{(t - 12 l(t/2))^4} \\ &\le \frac{2^4 l^2(t/2) (t + 6 l(t/2))^2 E_\theta (\varepsilon_1^2 - 1)^4}{(t - 12 l(t/2))^4}. \end{aligned}$$

Let $t_0 > 0$ be such that

$$t - 12 l(t/2) \ge t/2$$

for all $t \geq t_0$. Then I_1 can be estimated as

$$I_1 \leq (1+\gamma)\int_0^{t_0} t^{\gamma}\,dt + L\int_{t_0}^{\infty} \frac{l^2(t)(t^{2+\gamma}+t^{\gamma}l^2(t))}{t^4}\,dt < \infty \tag{3.19}$$

for $0 < \gamma < 1$. Further we note that

$$I_2 = 2^{1+\gamma} E_\theta v^{1+\gamma}. \tag{3.20}$$

By making use of estimates (3.19), (3.20) in (3.18) and applying (3.15) we obtain inequality (3.16). The next inequality (3.17) follows from conditions (2.14) and (3.1). Therefore (3.14) holds. Hence Proposition 3.3. □

This completes the proof of the auxiliary propositions.

PROOF OF THEOREM 3.1. By Proposition 3.3, the family $\{T(H)/H,\ H > 0\}$ is uniformly integrable. Therefore it suffices to show that for any $\delta > 0$,

$$\lim_{H\to\infty} \sup_{\theta\in\mathscr{K}} P_\theta\big\{|T(H)/H - d_1/\operatorname{tr} F| > \delta\big\}.$$

From the definition of $T(H)$ in (2.36) and from (3.1) it follows that

$$P_\theta\big\{|T(H)/H - d_1/\operatorname{tr} F| > \delta\big\} \leq P_\theta\big\{|\tau(H)/H - d_1/\operatorname{tr} F| > \delta\big\} + P_\theta\{\sigma(H) > n_0(H)\} + P_\theta\{\nu(H) > n_0(H)\}.$$

In view of Propositions 3.1 and 3.2, it remains to verify that

$$\lim_{H\to\infty} \sup_{\theta\in\mathscr{K}} P_\theta\{|\tau(H)/H - d_1/\operatorname{tr} F| > \delta\} = 0. \tag{3.21}$$

This relationship is proved in the Appendix. Hence Theorem 3.1. □

4. Uniform asymptotic normality of estimates (2.10), (2.33). The following is a key result for studying the asymptotic distribution of the sequential estimates $\tilde{\mu}(H)$.

PROPOSITION 4.1 [Lai and Siegmund (1983)]. *Let $x_n,\ \varepsilon_n,\ n = 0, 1, \ldots$ be random variables adapted to the increasing sequence of σ-algebras $\mathscr{F}_n,\ n = 0, 1, \ldots$. Let $\{P_\theta,\ \theta \in \Theta\}$ be a family of probability measures such that under every P_θ the following hold*:

(i) *$\varepsilon_1, \varepsilon_2, \ldots$ are i.i.d. with $E_\theta\varepsilon_1 = 0,\ E_\theta\varepsilon_1^2 = 1$;*
(ii) *$\sup_{\theta\in\Theta} E_\theta\{\varepsilon_1^2,\ |\varepsilon_1| > a\} \to 0,$ as $a \to \infty$;*
(iii) *ε_n is independent of $\mathscr{F}_{n-1}$ for each $n \geq 1$;*
(iv) *$P_\theta\{\sum_{i\geq 1} x_i^2 = \infty\} = 1$;*
(v) *$\sup_{\theta\in\Theta} P_\theta\{x_n^2 > a\} \to 0$ as $a \to \infty$ for each $n \geq 0$;*
(vi) *for each $\delta > 0$*

$$\lim_{m\to\infty} \sup_{\theta\in\Theta} P_\theta\left\{x_n^2 \geq \delta \sum_{i=1}^{n-1} x_i^2 \text{ for some } n \geq m\right\} = 0.$$

For $c > 0$ let

$$N_c = \inf\left\{n \geq 1\colon \sum_{i=1}^{n} x_{i-1}^2 \geq c\right\}, \qquad \inf\{\varnothing\} = \infty.$$

Then uniformly in $\theta \in \Theta$ and $-\infty < t < \infty$,

$$P_\theta\left\{c^{-1/2}\sum_{i=1}^{N_c} x_{i-1}\varepsilon_i \leq t\right\} \to \Phi(t) \quad \text{as } c \to \infty,$$

where $\Phi(t)$ is the standard normal distribution function.

First we shall prove asymptotic normality of the sequential least-squares estimate (2.15) of the vector of unknown parameters $a = (m, \lambda')'$.

PROPOSITION 4.2. *Let process* (1.1) *be stable, $\{\varepsilon_n\}$ be i.i.d. with $E\varepsilon_1 = 0$, $E\varepsilon_1^2 = 1$, $E\varepsilon_1^8 < \infty$ and* (2.2) *be satisfied. Then*

$$\lim_{H\to\infty} \sup_{\theta\in\mathscr{K}} |P_\theta\{\sqrt{H}V'(a(H) - a) \leq t\} - \Phi(t/d_2(V))| = 0, \tag{4.1}$$

where $\mathscr{K}$ is an arbitrary compact set in the region Θ, $V \in R^{p+1}$; $d_2^2(V) = d^2 V' F^{-1} V \operatorname{tr} F/d_1$.

The proof of Proposition 4.2 is given in the Appendix.

Now we can show uniform asymptotic normality of the estimate (2.10), (2.33).

THEOREM 4.1. *Let the conditions of Theorem* 2.1 *and* (3.3) *hold and the sequences c_k and h_k be defined by* (3.1). *Then*

$$\lim_{H\to\infty} \sup_{\theta\in\Theta} \sup_{-\infty<t<\infty} \left|P_\theta\left\{\sqrt{E_\theta T(H)}(\tilde{\mu}(H) - \mu) \leq t\right\} - \Phi((1 - \lambda' 1_p)t/d)\right| = 0; \tag{4.2}$$

$\mathscr{K}$ is an arbitrary compact set in Θ.

PROOF. By (3.1) the estimate $\tilde{\mu}$, defined by (2.10) and (2.33), can be written on the intersection of sets

$$\Gamma(H) = \{\sigma(H) \leq n_0(H)\} \cap \{\nu(H) \leq n_0(H)\} \tag{4.3}$$

as follows:

$$\tilde{\mu}(H) = \frac{m(H)}{1 - 1_p'\lambda(H)}.$$

From here and (2.15) we have on Γ :

$$\sqrt{E_\theta T(H)}(\tilde{\mu}(H) - \mu) = \sqrt{E_\theta T(H)} \frac{V_0'(a(H) - a)}{1 - 1_p'\lambda(H)}, \tag{4.4}$$

where $V_0 = \big(1, (\mu - \mu_0)1_p'\big)'$.

Propositions 3.1 and 3.2 [under $m = n_0(H)$] imply

$$\lim_{H\to\infty} \sup_{\theta\in\Theta} P_\theta\{\Gamma^c(H)\} = 0 \tag{4.5}$$

for any compact set $\mathscr{K} \in \Theta$. By applying Proposition 4.2 and Theorem 3.1 to (4.4) with $V = V_0$ and noting that $d_2^2(V_0) = d^2$, we obtain (4.2). Hence Theorem 4.1. □

5. Asymptotic minimaxity of the sequential procedure. In this section we shall show that the estimate of the mean μ of the process (1.1) defined by (2.10), (2.33) is optimal in a minimax sense in some class of sequential and nonsequential estimation procedures. In addition to the conditions imposed on the distribution of the noise in (1.1), (2.2) and (3.1), we assume that it has a piecewise continuously differentiable density $f(x)$ such that

$$J(f) = \int_{\infty}^{\infty} \frac{(f')^2}{f(x)}\, dx < \infty. \tag{5.1}$$

Let $\mathscr{P}$ denote the class of all noise distributions satisfying these conditions. First by making use of the sequential estimate $(T(H),\ \tilde{\mu}(H))$ defined by (2.9), (2.10) and (2.33), we construct a sequential procedure $\tilde{\mu}$. Let $\{H_n\}$ be some nondecreasing sequences of positive numbers H_n with $\lim H_n = +\infty$. Define the sequential procedure as

$$\tilde{\mu} = \{(T(H_n),\ \tilde{\mu}(H_n)),\ n \geq 1\}. \tag{5.2}$$

The performance of this procedures depends both on the unknown parameter μ and the nuisance parameters λ, d. For each number $\mu_1 \in (\mu_0 - q, \mu_0 + q)$ we introduce a special class of estimation procedures $\mathscr{M}$. Let $\hat{\mu} = \{(t_n, \hat{\mu}_n),\ n \geq 1\}$ be an arbitrary estimation procedure, where t_n is a Markovian moment with respect to the process $\{x_k\}$, and $\tilde{\mu}_n$ is a Borelian function of the observations $x_0, \dots, x_{t_n},\ n \geq 1$. Denote by $\mathscr{M}$ the class of all procedures $\hat{\mu}$ satisfying the following condition: there exists an interval $U(\mu_1) \subset (\mu_0 - q,\ \mu_0 + q)$ such that

$$E_\theta t_n \leq E_\theta T(n) \tag{5.3}$$

for all $\theta \in \{\theta = (\mu, \lambda', d)\colon \mu \in U(\mu_1),\ \Lambda,\ d \neq 0\}$ under any noise distribution from the class $\mathscr{P}$. For each procedure $\hat{\mu} \in \mathscr{M}$ we define the risk as

$$R_{\mu_1}(\hat{\mu}) = \sup_{d\neq 0} \sup_{\lambda\in\Lambda} \sup_{f\in\mathscr{P}} \lim_{\delta\to 0} \sup \limsup_{n\to\infty} \sup_{|\mu-\mu_1|<\delta} E_\theta |\phi(n)(\tilde{\mu}_n - \mu)|^\gamma, \tag{5.4}$$

where $0 < \gamma \leq 2;\ \phi(n) = w(E_{\theta_1} T(n))$,

$$w(x) = \sqrt{x}(1 - 1_p'\lambda)/d, \qquad \theta_1 = (\mu_1, \lambda', d). \tag{5.5}$$

Further we shall need some definitions and results from the theory of local asymptotic normality (LAN).

Let the parametric family of finite-dimensional distributions $\{Q_{\mu,n}\colon \mu \in S,\ n \geq 1\}$ be given, where S is an open parametric set on the line $(-\infty, +\infty)$.

We denote $dQ_{\mu_2,n}/dQ_{\mu_1,n}$ the Radon–Nikodim derivative of the absolutely continuous component of the measure $Q_{\mu_2,n}$ with respect to the measure $Q_{\mu_1,n}$.

DEFINITION. The family $\{Q_{\mu,n}: \mu \in S,\ n \geq 1\}$ is said to have the property LAN at the point $\mu_1 \in S$ if for some deterministic function $v(k)$ with $\lim v(k) = +\infty$ and for each $-\infty < u < \infty$,

$$\ln \frac{dQ_{\mu_1+uv_n^{-1}}}{dQ_{\mu_1,n}} = u\Delta_n - \frac{u^2}{2} + \psi_n(\mu_1, u),$$

where Δ_n is asymptotically (as $n \to \infty$) normal with parameters (0,1) with respect to the distribution $Q_{\mu_1,n}$, that is,

$$\lim_{n\to\infty} Q_{\mu_1,n}\{\Delta_n \leq z\} = \Phi(z), \qquad -\infty < z < \infty,$$

and $\psi_n(\mu_1, n)$ converges to zero by distribution $Q_{\mu_1,n}$, that is, for $\delta > 0$,

$$\lim_{n\to\infty} Q_{\mu_1,n}\{|\psi(\mu_1, u)| > \delta\} = 0.$$

PROPOSITION 5.1. *Let a family of distributions $\{Q_{\mu_1,n}\}$ have the property LAN at point $\mu_1 \in S$ and $\{(t_n, \hat{\mu}_n),\ n \geq 1\}$ be a sequential estimation procedure for μ. Then*

$$\kappa = \liminf_{\delta\to 0} \liminf_{n\to\infty} \sup_{|\mu-\mu_1|<\delta} E_\mu |v(n)(\hat{\mu}_n - \mu)|^\gamma \geq I(\mu_1), \tag{5.6}$$

$$I(\mu_1) = \liminf_{\delta\to 0} \liminf_{n\to\infty} \left| \frac{v(n)}{\sup_{|\mu-\mu_1|<\delta} E_\mu v(t_n)} \right|^\gamma E|\xi|^\gamma; \tag{5.7}$$

ξ is a Gaussian random variable with parameters (0,1).

This result follows directly from Theorem 1 of Efroimovich (1980).

Let $p(x_0, \dots, x_n; \mu, \lambda, d)$ be the probability density of a vector $(x_0, \dots, x_n)$ consisting of the values of the autoregressive process (1.1).

PROPOSITION 5.2. *The family of probability densities*

$$\big\{p(x_0, \dots, x_n; \mu, \lambda, d):\ \mu \in (\mu_0 - q, \mu_0 + q),\ n \geq 1\big\} \tag{5.8}$$

has the LAN property for $\lambda \in \Lambda$, $d \neq 0$ with $v(n) = \sqrt{J(f)}w(n)$.

[For proof we refer the reader to Beran (1976) and Akritas and Johnson (1982).] Now we can state the basic result of this section.

THEOREM 5.1. *For each $\mu_1 \in (\mu_0 - q, \mu_0 + q)$ the sequential procedure* (5.2) *is optimal in the minimax sense in the class $\mathscr{M}$, that is, the risk* (5.4) *satisfies the equality*

$$R_{\mu_1}(\tilde{\mu}) = \inf_{\hat{\mu}\in\mathscr{M}} R_{\mu_1}(\hat{\mu}). \tag{5.9}$$

PROOF. First we derive the lower bound for the risk (5.4) of Proposition 5.1. for the family (5.8). By Theorem 3.1 we have

$$\lim_{n\to\infty} \frac{E_{\theta_1} T(n)}{n} = \frac{d_1}{\operatorname{tr} F},$$
$$\lim_{\delta\to 0} \limsup_{n\to\infty} \sup_{|\mu-\mu_1|<\delta} \left| \frac{E_\theta T(n)}{E_{\theta_1} T(n)} - 1 \right| = 0. \tag{5.10}$$

By (5.4), (5.10) the risk $R_{\mu_1}(\hat{\mu})$ and the left-hand side of (5.6) satisfy inequality $f \in \mathscr{P}$ by the formula

$$R_{\mu_1}(\hat{\mu}) \geq \kappa \left(\frac{d_1}{J(f) \operatorname{tr} F(\theta_1)} \right)^{\gamma/2}. \tag{5.11}$$

Further, in view of (5.2), (5.10) we obtain

$$\begin{aligned} I(\mu_1) &\geq \liminf_{\delta\to 0} \liminf_{n\to\infty} \left(\frac{\sqrt{n}}{\sup_{|\mu-\mu_1|<\delta} E_\theta \sqrt{t_n}} \right)^{\gamma} E|\xi|^\gamma \\ &\geq E|\xi|^\gamma \liminf_{\delta\to 0} \liminf_{n\to\infty} \inf_{|\mu-\mu_1|<\delta} \left(\frac{n}{E_\theta T(n)} \right)^{\gamma/2} \\ &= E|\xi|^\gamma \liminf_{\delta\to 0} \liminf_{n\to\infty} \inf_{|\mu-\mu_1|<\delta} \left(\frac{E_{\theta_1} T(n)}{E_\theta T(n)} \frac{n}{E_{\theta_1} T(n)} \right)^{\gamma/2} \\ &= E|\xi|^\gamma \left(\frac{d_1}{\operatorname{tr} F(\theta_1)} \right)^{-\gamma/2}. \end{aligned} \tag{5.12}$$

From here and (5.6), (5.11) it follows that

$$R_{\mu_1}(\hat{\mu}) \geq \sup_{f\in\mathscr{P}} \frac{E|\xi|^\gamma}{(J(f))^{\gamma/2}} = E|\xi|^\gamma \tag{5.13}$$

for each $\hat{\mu} \in \mathscr{M}$. The last equality holds because $J(f) \geq 1$ for all $f \in \mathscr{P}$ and $J(f) = 1$ for the Gaussian density $f \in \mathscr{P}$. It remains to verify that

$$R_{\mu_1}(\hat{\mu}) = E|\xi|^\gamma, \qquad \mu_1 \in (\mu_0 - q, \mu_0 + q). \tag{5.14}$$

For this we need the following assertion.

LEMMA 5.1. *Under the conditions of Theorem* 5.1, *for some* $r > 2$,

$$\limsup_{H\to\infty} \sup_{|\mu-\mu_1|<\delta} E_\theta \left| \sqrt{H} (\tilde{\mu}(H) - \mu) \right|^r < \infty,$$

where $\mu_1 \in (\mu_0 - q, \mu_0 + q)$, $0 < \delta < 1$, $\lambda \in \Lambda$.

The proof of Lemma 5.1 is given in the Appendix. By applying Lemma 5.1 and Theorem 4.1 to the risk (5.3) we obtain (5.14). Hence Theorem 5.1.

APPENDIX

In the sequel we shall need the Burkholder inequality in the following form.

LEMMA A.1 [Shiryaev (1984), Liptser and Shiryaev (1986)]. *If* $X=(X_n, \mathscr{F}_n)$ *is a martingale and* T *is a Markovian moment, then for each* $p > 1$ *there exists the universal constant* C_p (*independent of* X *and* T) *such that*

$$E|X_T|^p \le C_p E\left(\sum_{k=1}^{T} (X_k - X_{k-1})^2 \right)^{p/2}, \qquad X_0 = 0.$$

A.1. Proof of Lemma 3.1. By (2.17)

$$\begin{aligned} E_\theta z^2 (b(z) - b_0)^4 &= z^2 E_\theta (b(z) - b_0)^4 \chi_{\{\det G(z) > 0\}} \\ &\quad + z^2 b_0^4 P_\theta\{\det G(z) = 0\}, \end{aligned} \tag{A.1}$$

where χ_A denotes the indicator of a set A. Since $b(z) \le 1$ and $b_0 \le 1$, then on the set $\{\det G(z) > 0\}$,

$$|b(z) - b_0| \le 2 \left| \frac{1}{zR(z)\|G^{-1}(z)\|} - \frac{1}{\operatorname{tr} F \|F^{-1}\|} \right| \le 2 \left\| \frac{G(z)}{zR(z)} - \frac{F}{\operatorname{tr} F} \right\|. \tag{A.2}$$

Now we show that

$$\lim_{z \to \infty} \sup_{\theta \in \Theta} z^2 E_\theta \left\| \frac{G(z)}{zR(z)} - \frac{F}{\operatorname{tr} F} \right\|^4 < \infty. \tag{A.3}$$

From the definition of $G(z)$ in (2.15), it follows that

$$\begin{aligned} &\left\| \frac{G(z)}{zR(z)} - \frac{1}{\operatorname{tr} F} \right\| \\ &\quad = \left\| \frac{D(z)}{zR(z)} - \frac{(1 - \alpha(z)) Y_{\tau(z)} Y'_{\tau(z)}}{zR(z)} + F\left(\frac{\tau(z) - r(z)}{zR(z)} - \frac{F}{\operatorname{tr} F} \right) \right\|, \end{aligned} \tag{A.4}$$

where

$$D(z) = \sum_{k=r(z)+1}^{\tau(z)} (Y_k Y'_k - F). \tag{A.5}$$

By (2.13), (2.16) we have

$$\left| \frac{\tau(z) - r(z)}{zR(z)} - \frac{F}{\operatorname{tr} F} \right| \le \frac{\sqrt{p+1}\|D(z)\| + \|Y_{\tau(z)}\|^2}{zR(z) \operatorname{tr} F}. \tag{A.6}$$

By (A.4) and (A.6),

$$\left\| \frac{G(z)}{zR(z)} - \frac{F}{\operatorname{tr} F} \right\| \le \frac{(\sqrt{p+1} + 1)\|D(z)\| + 2\|Y_{\tau(z)}\|^2}{zR(z)}$$

and, hence, taking into account that $R(z) \geq 1$,

$$\left\| \frac{G(z)}{zR(z)} - \frac{F}{\operatorname{tr} F} \right\|^4 \leq 8 \frac{(\sqrt{p+1}+1)^4 \|D(z)\|^4 + 16\|Y_{\tau(z)}\|^8}{(zR(z))^4} \leq 8 \frac{(\sqrt{p+1}+1)^4 \|D(z)\|^4 + 16\|Y_{\tau(z)}\|^8}{z^4}. \tag{A.7}$$

Now we show that

$$\sup_{\theta \in \mathscr{K}} E_\theta \|Y_{\tau(z)}\|^8 \leq L(1+z), \tag{A.8}$$

$$\sup_{\theta \in \mathscr{K}} E_\theta \|D(z)\|^4 \leq L(1+z)^2. \tag{A.9}$$

Write (2.3) in vector form:

$$X_n = (\mu - \mu_0)1_p + \xi_n, \qquad \xi_n = A\xi_{n-1} + d\eta_n, \qquad \xi_0 = \zeta - (\mu - \mu_0)1_p, \tag{A.10}$$

where $X_n = (y_n, \dots, y_{n-p+1})'$, $\eta_n = (\varepsilon_n, 0, \dots, 0)'$ and the matrix A is defined in (2.27).

Since

$$\xi_k = A^k \xi_0 + d \sum_{j=1}^{k} A^{k-j} \eta_j, \tag{A.11}$$

then there exists some constant L for the compact set $\mathscr{K}$ that for all $\theta \in \mathscr{K}$,

$$\|Y_k\| \leq L(1 + \|\xi_{k-1}\|). \tag{A.12}$$

Therefore

$$\sup_{k \geq 1} \sup_{\theta \in \mathscr{K}} E_\theta \|Y_k\|^8 < \infty. \tag{A.13}$$

Further, by the definition of $\tau(z)$ we have

$$E_\theta \tau(z) \leq z E_\theta R(z) + r(z) + 1 \leq \frac{z \sum_{k=1}^{r(z)+1} E_\theta \|Y_k\|^2}{r(z)} + r(z) + 1.$$

From here it follows that

$$E_\theta \tau(z) \leq L(1+z), \qquad z > 0. \tag{A.14}$$

By (A.11) and Hölder's inequality,

$$\|\xi_k\|^8 \leq 2^7 \left(\|A^k\|^8 \|\xi_0\|^8 + \left(\sum_{j \geq 1} \|A^j\|^{8/7} \right)^7 \sum_{j=1}^{k} |\varepsilon_j|^8 \right).$$

Taking into account (A.14) and applying the Wald identity, we obtain

$$\sup_{\theta \in \mathscr{K}} E_\theta \|\xi_{\tau(z)}\|^8 \leq L(1+z). \tag{A.15}$$

This inequality and (A.12) imply (A.8). Now we verify (A.9). By (A.5) and (A.10)

$$\|D(z)\|^2 = 2\left\|\sum_{k=r+1}^{\tau(z)} \xi_{k-1}\right\|^2 + \left\|\sum_{k=r+1}^{\tau(z)} (X_{k-1}X'_{k-1} - (\mu-\mu_0)^2 1_p 1'_p - F_0)\right\|^2$$

$$\leq 2(1+4(\mu-\mu_0)^2 p)\left\|\sum_{k=r+1}^{\tau(z)} \xi_{k-1}\right\|^2 + 2\|S(z)\|^2,$$

where

$$S(z) = \sum_{k=r+1}^{\tau(z)} (\xi_{k-1}\xi'_{k-1} - F_0). \tag{A.16}$$

Hence,

$$\|D(z)\|^4 \leq 8(1+4q^2p)^2\left\|\sum_{k=r+1}^{\tau(z)} \xi_{k-1}\right\|^4 + 2\|S(z)\|^4. \tag{A.17}$$

Since by (A.10),

$$\sum_{k=r+1}^{\tau(z)} \xi_{k-1} = (I-A)^{-1}(\xi_r - \xi_{\tau(z)}) + d(I-A)^{-1}\sum_{k=r+1}^{\tau(z)} \eta_k,$$

then for all $\theta \in \mathscr{K}$,

$$E_\theta\left\|\sum_{k=r+1}^{\tau(z)} \xi_{k-1}\right\|^4 \leq L\left(E_\theta\|\xi_r\|^4 + E_\theta\|\xi_{\tau(z)}\|^4 + E_\theta\left(\sum_{k=r+1}^{\tau(z)} \varepsilon_k^2\right)^2\right)$$

$$\leq L\left(E_\theta\|\xi_r\|^4 + E_\theta\|\xi_{\tau(z)}\|^4 + E_\theta\left(\sum_{k=r+1}^{\tau(z)} (\varepsilon_k^2 - 1)\right)^2 + E_\theta(\tau(z) - r(z))^2\right)$$

$$\leq L\left(E_\theta\|\xi_r\|^4 + E_\theta\|\xi_{\tau(z)}\|^4 + E_\theta(\tau(z)-r(z)) + z^2E_\theta R^2(z) + 1\right)$$

$$\leq L\left(E_\theta\|\xi_r\|^4 + E_\theta\|\xi_{\tau(z)}\|^4 + z + z^2(r+1)\sum_{j=1}^{r+1} E_\theta\|Y_j\|^4/r^2 + 1\right).$$

From here, (A.13) and (A.15) it follows that

$$\sup_{\theta\in\mathscr{K}} E_\theta\left\|\sum_{k=r+1}^{\tau(z)} \xi_{k-1}\right\|^4 \leq L(1+z^2). \tag{A.18}$$

Further we note that the matrix (A.16) satisfies the equations

$$S(z) - AS(z)A' = \xi_r\xi'_r - \xi_{\tau(z)}\xi'_{\tau(z)} + dAM_1(z) + dM'_1(z)A' + M_2(z),$$

where

$$M_1(z) = \sum_{k=r+1}^{\tau(z)} \xi_{k-1}\eta'_k, \qquad M_2(z) = \sum_{k=r+1}^{\tau(z)} (d^2\eta_k\eta'_k - B).$$

Therefore for $\theta \in \mathscr{K}$,

$$E_\theta\|S(z)\|^4 \le L\big(E_\theta\|\xi_r\|^8 + E_\theta\|\xi_{\tau(z)}\|^8 + E_\theta\|M_1(z)\|^4 + E_\theta\|M_2(z)\|^4\big). \tag{A.19}$$

From here, (A.15), (A.17) and (A.18) we have

$$E_\theta\|D(z)\|^4 \le L\big(1 + z^2 + E_\theta\|M_1(z)\|^4 + E_\theta\|M_2(z)\|^4\big). \tag{A.20}$$

Show that

$$\sup_{\theta\in\mathscr{K}} E_\theta\|M_1(z)\|^4 \le (1+z)^2, \qquad z > 0, \tag{A.21}$$

$$\sup_{\theta\in\mathscr{K}} E_\theta\|M_2(z)\|^4 \le (1+z)^2, \qquad z > 0. \tag{A.22}$$

By applying Lemma A.1 we obtain

$$E_\theta\|M_1(z)\|^4 \le LE_\theta\bigg(\sum_{k=r+1}^{\tau(z)} \|\xi_{k-1}\|^2\varepsilon_k^2\bigg)^2.$$

By (A.10),

$$\begin{aligned}\|\xi_{k-1}\|^2 &\le (|\mu-\mu_0|\sqrt{p} + \|X_{k-1}\|)^2 \\ &\le 2q^2p(1+\|X_{k-1}\|^2) = 2q^2p\|Y_k\|^2.\end{aligned} \tag{A.23}$$

Therefore, taking into account (2.13), we have inequality

$$\begin{aligned}E_\theta\|M_1(z)\|^4 &\le LE_\theta\bigg(\sum_{k=r+1}^{\tau(z)} \|Y_k\|^2\varepsilon_k^2\bigg)^2 \\ &\le L\Bigg(E_\theta\bigg(\sum_{k=r+1}^{\tau(z)-1} \|Y_k\|^2\varepsilon_k^2\bigg)^2 + E_\theta\|Y_{\tau(z)}\|^4\varepsilon_{\tau(z)}^4\Bigg) \\ &\le L\Bigg(E_\theta \sum_{k=r+1}^{\tau(z)-1} \|Y_k\|^2 \sum_{k=r+1}^{\tau(z)-1} \|Y_k\|^2\varepsilon_k^4 + E_\theta\|Y_{\tau(z)}\|^4\varepsilon_{\tau(z)}^4\Bigg) \\ &\le L\Bigg(z^2E_\theta R(z)E_\theta\bigg(\sum_{k=r+1}^{\tau(z)-1} \|Y_k\|^2\varepsilon_k^4\bigg|Y_1,\dots,Y_{r+1}\bigg) + E_\theta\|Y_\tau\|^4\Bigg) \\ &\le L\big(z^2E_\theta R^2(z) + E_\theta\|Y_\tau\|^4\big) \le L\Bigg(z^2\sum_{j=1}^{r+1} E_\theta\|Y_j\|^4/r + E_\theta\|Y_\tau\|^4\Bigg).\end{aligned}$$

From here and (A.8), (A.13), (A.21) follows. Inequality (A.22) can be proved similarly to (A.18). Combining (A.20)–(A.22) yields (A.9). From (A.7)–(A.9) inequality (A.3) follows.

It remains to prove that for the second addend in (A.1),

$$\limsup_{z\to\infty}\sup_{\theta\in\mathscr{K}} z^2 P_\theta\{\det G(z)=0\}<\infty. \tag{A.24}$$

Note that

$$\begin{aligned}P_\theta\{\det G(z)=0\}&=P_\theta\Big\{\inf_{\|v\|=1} v'G(z)v/zR(z)=0\Big\}\\&=P_\theta\Big\{\inf_{\|v\|=1}\big(v'Fv/\operatorname{tr}F+v'(G(z)/zR(z)-F/\operatorname{tr}F)v\big)=0\Big\}.\end{aligned}$$

Let

$$\phi_*=\inf_{\theta\in\mathscr{K}}\inf_{\|v\|=1} v'F(\theta)v/\operatorname{tr}F(\theta),$$

where the matrix $F(\theta)$ is defined by (2.27). By making use of the Chebyshev inequality we have

$$\begin{aligned}P_\theta\{\det G(z)=0\}&\le P_\theta\big\{\|G(z)/zR(z)-F/\operatorname{tr}F\|>\phi_*\big\}\\&\le\frac{E_\theta\|G(z)/zR(z)-F/\operatorname{tr}F\|^4}{\phi_*^4}.\end{aligned}$$

From here by applying (A.3), we obtain (A.24). Hence Lemma 3.1. □

A.2. Proof of Lemma 3.2. By (2.17), (2.20) and (2.3), taking into account that $b(z)=0$ if $\det G(z)=0$, we have

$$\begin{aligned}a^*(h)-a&=\frac{\sum_{k=1}^{\sigma(h)} b(c_k)(a(c_k)-a)}{\sum_{k=1}^{\sigma(h)} b(c_k)}\\&=d\frac{\sum_{k=1}^{\sigma(h)} b(c_k)G^{-1}(c_k)M(c_k)\chi_{\{\det G(c_k)>0\}}}{\sum_{k=1}^{\sigma(h)} b(c_k)},\end{aligned}$$

where $M(z)$ is defined in (2.30). By the Cauchy–Bunyakovskii inequality,

$$\begin{aligned}E_\theta\|a^*(h)-a\|^4&\le d^4E_\theta\bigg(\frac{\sum_{k=1}^{\sigma(h)}\sqrt{b(c_k)}\|M(c_k)\|/c_kR(c_k)}{\sum_{k=1}^{\sigma(h)} b(c_k)}\bigg)^4\\&\le d^4E_\theta\frac{(\sum_{k=1}^{\sigma(h)} b(c_k))^2(\sum_{k=1}^{\sigma(h)}\|M(c_k)\|^2/c_k^2R^2(c_k))^2}{(\sum_{k=1}^{\sigma(h)} b(c_k))^4}\\&\le d^4\frac{\Big(\sum_{k\ge1}\sqrt{E_\theta\|M(c_k)\|^4/c_k^4R^4(c_k)}\Big)^2}{h^2}\\&\le d^4\frac{\Big(\sum_{k\ge1}\sqrt{E_\theta\|M(c_k)\|^4/c_k^4}\Big)^2}{h^2}.\end{aligned}\tag{A.25}$$

(The last inequality holds because $R(c_j)\ge1$.)

Now we estimate $E_\theta\|M(c_k)\|^4$. Denoting $\langle Y\rangle_i$ the ith coordinate of vector Y, we have

$$E_\theta\|M(c_k)\|^4 = E_\theta\Bigg(\sum_{j=1}^{p+1}\langle M(c_k)\rangle_j^2\Bigg)^2 \le (p+1)\sum_{j=1}^{p+1} E_\theta\langle M(c_k)\rangle_j^4 = (p+1)\sum_{j=1}^{p+1} E_\theta (X_{\tau_k}^{(j)})^4,$$

where

$$X_n^{(j)} = \sum_{l=1}^{n} v_l\langle Y_l\rangle_j\varepsilon_l, \qquad v_l = \chi_{\{r_k+1\le l<\tau_k\}} + \alpha(c_k)\chi_{\{\tau_k=l\}}.$$

Note that v_l is a function depending on $y_1,\dots,y_{l-1}$. By applying Lemma A.1 to martingales $\{X_n^{(j)}\}$ with Markovian moment (2.19) and the Cauchy–Bunyakovskii inequality we obtain

$$\begin{aligned}
E_\theta\|M(c_k)\|^4 &\le LE_\theta\Bigg(\sum_{l=r_k+1}^{\tau_k-1}\|Y_l\|^2\varepsilon_l^2 + \alpha^2(c_k)\|Y_{\tau_k}\|^2\varepsilon_{\tau_k}^2\Bigg)^2 \\
&\le LE_\theta\Bigg(\sum_{l=r_k+1}^{\tau_k-1}\|Y_l\|^2 + \alpha^2(c_k)\|Y_{\tau_k}\|^2\Bigg) \\
&\qquad \times\Bigg(\sum_{l=r_k+1}^{\tau_k-1}\|Y_l\|^2\varepsilon_l^4 + \alpha^2(c_k)\|Y_{\tau_k}\|^2\varepsilon_{\tau_k}^4\Bigg) \\
&\le Lc_kE_\theta R(c_k) \\
&\qquad \times E_\theta\Bigg(\sum_{l=r_k+1}^{\tau_k-1}\|Y_l\|^2\varepsilon_l^4 + \alpha^2(c_k)\|Y_{\tau_k}\|^2\varepsilon_{\tau_k}^4\Bigg|Y_1,\dots,Y_{r_k+1}\Bigg) \\
&\le Lc_k^2E_\theta R^2(c_k)E\varepsilon_1^4 \le LE\varepsilon_1^4c_k^2(r_k+1)\sum_{j=1}^{r_k+1}E_\theta\|Y_j\|^4/r_k^2.
\end{aligned} \tag{A.26}$$

By (A.13),

$$E_\theta\|M(c_k)\|^4 \le Lc_k^2.$$

From here, (A.25) and (3.2),

$$\sup_{\theta\in\mathscr{K}} E_\theta\|a^*(h)-a\|^4 \le L\frac{(\sum_{k\ge 1}1/c_k)^2}{h^2} = \frac{L}{h^2}\Bigg(\frac{n_0(H)}{H} + \sum_{k>n_0(H)}\bigg(\frac{1}{k}\bigg)^{1+\delta}\Bigg)^2$$

By virtue of (2.14) we establish (3.8) for sufficiently large H. Hence Lemma 3.2. □

A.3. Proof of inequality (3.14). In view of (3.2) we have

$$
\begin{aligned}
E_\theta\big(c_{\sigma(h_{\nu(H)})}\big)^{1+\gamma} &= (1+\gamma)\int_0^\infty t^\gamma P_\theta\big\{c_{\sigma(h_{\nu(H)})} > t\big\}\,dt \\
&\le (1+\gamma)\int_0^H t^\gamma dt + (1+\gamma) \\
&\quad \times \int_0^\infty t^\gamma P_\theta\big\{\sigma(h_{\nu(H)}) > t^{1/(1+\delta)},\ \sigma(h_{\nu(H)}) > n_0(H)\big\}\,dt \\
&= H^{1+\gamma} + (1+\gamma)(1+\delta) \\
&\quad \times \int_0^\infty s^{\delta+\gamma(1+\delta)} P_\theta\big\{\sigma(h_{\nu(H)}) > s,\ \sigma(h_{\nu(H)}) > n_0(H)\big\}\,ds \\
&= H^{1+\gamma} + I_1(H,\theta) + (1+\gamma)(1+\delta)(I_2(H,\theta) + I_3(H,\theta)),
\end{aligned}
\tag{A.27}
$$

where

$$
\begin{aligned}
I_1(H,\theta) &= (n_0)^{(1+\gamma)(1+\delta)} P_\theta\big\{\sigma(h_{\nu(H)} > n_0(H)\big\}, \\
I_2(H,\theta) &= \int_{n_0(H)}^\infty s^{\delta+\gamma(1+\delta)} P_\theta\{\sigma(H) > s\}\,ds, \\
I_3(H,\theta) &= \int_{n_0(H)}^\infty s^{\delta+\gamma(1+\delta)} P_\theta\{\sigma(h_{\nu(H)}) > s,\ \nu(H) > n_0(H)\}\,ds.
\end{aligned}
\tag{A.28}
$$

By applying Propositions (3.1) and (3.2), we obtain the inequality

$$
\begin{aligned}
P_\theta\{\sigma(h_{\nu(H)} > n_0(H)\} &\le P_\theta\{\sigma(H) > n_0(H)\} + P_\theta\{\nu(H) > n_0(H)\} \\
&\le Ll^2(H)/H^2
\end{aligned}
$$

for sufficiently large H. Therefore

$$
\lim_{H\to\infty} \sup_{\theta\in\mathscr{K}} I_1(H,\theta) = 0 \tag{A.29}
$$

for $0 < \gamma < (1-\delta)/(1+\delta)$.

Further, by Proposition 3.1,

$$
\begin{aligned}
I_2(H,\theta) &\le L\frac{l(H)n^3(H)}{H}\int_{n_0(H)}^\infty \frac{ds}{s^{4-\delta-\gamma(1+\delta)}} + Ll(H)\int_{n_0(H)}^\infty \frac{ds}{s^{2-\gamma(1+\delta)}} \\
&= L\frac{l(H)(n_0(H))^{\delta+\gamma(1+\delta)}}{(3-\delta-\gamma(1+\delta))H} + L\frac{l(H)}{(1-\gamma(1+\delta))(n_0(H))^{1-\gamma(1+\delta)}}.
\end{aligned}
$$

From here and (2.14) it follows that

$$
\lim_{H\to\infty} \sup_{\theta\in\mathscr{K}} I_2(H,\theta) = 0 \tag{A.30}
$$

for $0 < \gamma < (1-\delta)/(1+\delta)$.

In order to estimate $I_3(H,\theta)$, we examine the probability in the integrand. By (2.21) for sufficiently large H and $s \ge n_0(H)$,

$$\begin{aligned} &P_\theta\{\sigma(h_{\nu(H)}) > s,\ \nu(H) > n_0(H)\} \\ &\quad = P_\theta\Bigg\{\sum_{k=1}^{[s]} b(c_k) < h_{\nu(H)},\ \nu(H) > n_0(H)\Bigg\} \\ &\quad \le P_\theta\Bigg\{\sum_{k=1}^{[s]} b(c_k) < b_0(\theta)[s]/2\Bigg\} \\ &\qquad + P_\theta\{h_{\nu(H)} > b_0(\theta)[s]/2, \nu(H) > n_0(H)\}, \end{aligned} \tag{A.31}$$

where $b_0(\theta)$ is defined by (2.28). Note that

$$\begin{aligned} P_\theta\Bigg\{\sum_{k=1}^{[s]} b(c_k) < \frac{b_0(\theta)[s]}{2}\Bigg\} &\le P_\theta\Bigg\{\sum_{k=1}^{[s]} |b(c_k) - b_0(\theta)| > \frac{b_* s}{4}\Bigg\} \\ &\le \frac{4^4 E_\theta(\sum_{k=1}^{[s]} |b(c_k) - b_0(\theta)|)^4}{(b_* s)^4} \\ &\le \frac{4^4}{b_*^4}\frac{(\sum_{k=1}^{[s]} 1/c_k^{1/3})^3 \sum_{k=1}^{[s]} c_k E_\theta|b(c_k) - b_0|^4}{s^4}; \end{aligned}$$

b_* is defined in (3.7). By applying Lemma 3.1 and taking into account (3.2), we obtain

$$P_\theta\Bigg\{\sum_{k=1}^{[s]} b(c_k) < b_0(\theta)[s]/2\Bigg\} \le Ll(H)\bigg(\frac{n_0^3(H)}{Hs^4} + \frac{1}{s^{2+\delta}}\bigg). \tag{A.32}$$

Now we estimate the second addend in the right-hand side of (A.31). In view of (3.2), we have

$$\begin{aligned} &P_\theta\{h_{\nu(H)} > b_0(\theta)[s]/2,\ \nu(H) > n_0(H)\} \\ &\qquad \le P_\theta\{\nu(H) > s^{1/(1+\delta)}\tilde{b},\ \nu(H) > n_0(H)\}, \end{aligned}$$

where $\tilde{b} = (b_*/4)^{1/(1+\delta)}$. By making use of this inequality and (A.31), (A.32), we can estimate $I_3(H,\theta)$:

$$\begin{aligned} I_3(H,\theta) &\le Ll(H)\frac{n_0^3(H)}{H}\int_{n_0(H)}^\infty \frac{ds}{s^{4-\delta-\gamma(1+\delta)}} \\ &\qquad + Ll(H)\int_{n_0(H)}^\infty \frac{ds}{s^{2-\gamma(1+\delta)}} + I_4(H,\theta) \\ &= Ll(H)\frac{(n_0(H))^{\delta+\gamma(1+\delta)}}{(3-\delta-\gamma(1+\delta))H} \\ &\qquad + Ll(H)\frac{1}{(1-\gamma(1+\delta))(n_0(H))^{1-\gamma(1+\delta)}} + I_4(H,\theta), \end{aligned} \tag{A.33}$$

where

$$I_4(H,\theta) = \int_{n_0(H)}^{\infty} s^{\delta+\gamma(1+\delta)} P_\theta\{\nu(H) > \tilde{b} s^{1/(1+\delta)},\ \nu(H) > n_0(H)\}\, ds.$$

This quantity can be estimated as

$$\begin{aligned} I_4(H,\theta) &\le \frac{1+\delta}{\tilde{b}^{1+\delta_1}} \int_0^\infty t^{\delta_1} P_\theta\{\nu(H) > t,\ \nu(H) > n_0(H)\}\, dt \\ &= \frac{1+\delta}{(1+\delta_1)\tilde{b}^{1+\delta_1}} \biggl((n_0(H))^{1+\delta_1} P_\theta\{\nu(H) > n_0(H)\} \\ &\qquad + \int_{n_0(H)}^\infty t_1^{\delta} P_\theta\{\nu(H) > t\}\, dt \biggr), \end{aligned}$$

$\delta_1 = (1+\delta)(\delta + \gamma(1+\delta)) + \delta$. From here and Proposition 3.2 it follows that for sufficiently large H,

$$I_4(H,\theta) \le L\left(\frac{l^2(H)(n_0(H))^{1+\delta_1}}{H^2} + \frac{l^3(H)}{(n_0(H))^{1+\delta-\delta_1}} \right).$$

Since $0 < \delta < \sqrt{2} - 1$ then for all $0 < \gamma < (1 - 2\delta - \delta^2)/(1+\delta)^2$,

$$\lim_{H\to\infty} \sup_{\theta\in\mathscr{K}} I_4(H,\theta) = 0$$

and by making use of (A.33) we obtain

$$\lim_{H\to\infty} \sup_{\theta\in\mathscr{K}} I_3(H,\theta) = 0.$$

Combining this limiting relationship and (A.27), (A.29) and (A.30) yields inequality (3.14). □

A.4. Proof of Lemma 3.3. By (2.13) and (A.12),

$$\sum_{k=1}^{n} y_k^2 \le \sum_{k=1}^{n+1} \|Y_k\|^2 \le L \sum_{k=1}^{n} (1 + \|\xi_k\|^2). \tag{A.34}$$

In view of (A.11),

$$\begin{aligned} \|\xi_k\|^2 &\le 2\left(\|A^k\|^2 \|\xi_0\|^2 + d^2 \left(\sum_{j=1}^{k} \|A^{k-j}\| |\varepsilon_j| \right)^2 \right) \\ &\le 2\left(\|A^k\|^2 \|\xi_0\|^2 + d^2 \sum_{j=1}^{k} \|A^{k-j}\| \sum_{j=1}^{k} \|A^{k-j}\| (\varepsilon_j)^2 \right) \\ &\le 2\left(\|A^k\|^2 \|\xi_0\|^2 + d^2 \sum_{j\ge 0} \|A^{j}\| \sum_{j=0}^{k-1} \|A^{j}\| (\varepsilon_{k-j})^2 \right). \end{aligned} \tag{A.35}$$

Since $\sum_{j\geq 0}\|A^j\|$ is bounded for $\theta\in\mathscr{K}$, then

$$\sum_{k=1}^{n}\|\xi_k\|^2 \leq 2\|\xi_0\|^2\sum_{k=1}^{n}\|A^k\|^2 + 2d^2\sum_{j\geq 1}\|A^j\|\sum_{j=0}^{n-1}\|A^j\|\sum_{k=j+1}^{n}(\varepsilon_{k-j})^2$$
$$\leq \max_{\theta\in\mathscr{K}} Q(\theta)\left(\|\xi_0\|^2+\sum_{k=1}^{n}\varepsilon_k^2\right),$$

where

$$Q(\theta)=2\max\left(\sum_{j\geq 1}\|A^j\|^2,\ d^2\left(\sum_{j\geq 0}\|A^j\|\right)^2\right).$$

From here and (A.34), taking into account the definition of ξ_0 in (A.10), we obtain (3.10). Hence Lemma 3.3. □

A.5. Proof of relationship (3.21). By making use of (A.6) we obtain

$$\begin{aligned}\left|\frac{\tau(H)}{H}-\frac{d_1}{\operatorname{tr}F}\right| &\leq R(H)\left|\frac{\tau(H)-r(H)}{HR(H)}-\frac{1}{\operatorname{tr}F}\right|\\ &\quad+\frac{|R(H)-d_1|}{\operatorname{tr}F}+\frac{r(H)}{H}\\ &\leq L\left(\frac{\|D(H)\|}{H}+\frac{\|Y_{\tau(H)}\|^2}{H}+|R(H)-d_1|+\frac{r(H)}{H}\right).\end{aligned}\tag{A.36}$$

By the definition of $R(H)$ in (2.13),

$$\begin{aligned}|R(H)-d_1| &= \left|\frac{\sum_{k=1}^{r}\langle Y_{k+1}Y'_{k+1}\rangle_{22}}{r(H)}-\langle F\rangle_{22}\right|\\ &\leq \frac{1}{r(H)}\left|\left\langle\sum_{k=2}^{r(H)+1}(Y_kY'_k-F)\right\rangle_{22}\right|\\ &\leq \frac{1}{r(H)}\left\|\sum_{k=2}^{r(H)+1}(Y_kY'_k-F)\right\|.\end{aligned}\tag{A.37}$$

By an argument similar to the proof of inequality (A.9) for the matrix (A.5), one can show that

$$\sup_{\theta\in\mathscr{K}} E_\theta\left\|\sum_{k=1}^{r(H)+1}(Y_kY'_k-F)\right\|^4 \leq L(1+r(H))^2.$$

From here and from (A.36), (A.37), (A.8) and (A.9), it follows that

$$\sup_{\theta\in\mathscr{K}} E_\theta\left|\frac{\tau(H)}{H}-\frac{d_1}{\operatorname{tr}F}\right|^4 \leq L\left(\frac{1}{H^2}+\frac{1}{H^3}+\frac{1}{H^4}+\frac{1}{(r(H))^2}+\frac{1}{(r(H))^4}+\frac{r(H)}{H}\right).$$

This inequality by virtue of (2.13), (2.14), completes the proof of (3.21). □

A.6. Proof of Proposition 4.2. By substituting (2.3) in (2.15) we obtain on the set $\{\det G(H) > 0\}$,

$$\zeta(H) = \sqrt{H}V'(a(H) - a) = d_2(V)\zeta_0(H) + \Delta(H), \tag{A.38}$$

where

$$\zeta_0(H) = \sum_{j=1}^{\tau_0} g_j\varepsilon_j/\sqrt{H},\ g_j = \frac{V'F^{-1}Y_j\sqrt{\operatorname{tr} F}}{\sqrt{d_1 V'F^{-1}V}}\chi_{\{j>r\}},$$
$$\tau_0 = \tau_0(H) = \inf\left\{k \geq 1\colon \sum_{j=1}^{k} g_j^2 \geq H\right\},$$
$$\Delta(H) = \Delta_1(H) + \Delta_2(H) + \Delta_3(H),$$
$$\Delta_1(H) = -d(1 - \alpha(H))\frac{V'F^{-1}Y_\tau\varepsilon_\tau \operatorname{tr} F}{d_1\sqrt{H}},$$
$$\Delta_2(H) = dV'(HG^{-1}(H) - F^{-1}\operatorname{tr} F/d_1)M(H)/\sqrt{H},$$
$$\Delta_3(H) = d_2(V)\left(\sum_{j=1}^{\tau} g_j\varepsilon_j - \sum_{j=1}^{\tau_0} g_j\varepsilon_j\right)/\sqrt{H};$$

$M(H)$ and τ are defined by (2.30) and (2.13).

In view of (A.24), to prove the desired conclusion (4.1) it suffices to show the following.

1. The random variable ζ_0 in (A.38) is asymptotically normal uniformly in parameter $\theta \in \mathscr{K}$, that is,

$$\lim_{H\to\infty} \sup_{\theta\in\mathscr{K}} \sup_{-\infty<t<\infty} \left|P_\theta\{\zeta_0(H) \leq t\} - \Phi(t)\right| = 0; \tag{A.39}$$

2. The random variable $\Delta(H)$ converges to zero in probability uniformly in $\theta \in \mathscr{K}$ as H tends to infinity, that is, for every $\delta > 0$,

$$\lim_{H\to\infty} \sup_{\theta\in\mathscr{K}} P_\theta\{|\Delta(H)| > \delta,\ \det G(H) > 0\} = 0. \tag{A.40}$$

To prove (A.39) we apply Proposition 4.1 to the sequences g_k, ε_k and the stopping moment τ_0. Let us verify its conditions. Conditions (i)–(iii) are evident. Condition (iv) holds because by (2.26),

$$\lim_{N\to\infty} \frac{\sum_{j=1}^N g_j^2}{N} = \frac{\operatorname{tr} F}{d_1} \quad P_\theta\text{-a.s.}$$

for $\theta \in \mathscr{K}$. Condition (v) holds because $E_\theta g_n^2$ is bounded for $\theta \in \mathscr{K}$ and $n \geq 0$, due to (A.12), (A.35). In order to verify condition (vi) it suffices to show that for any $0 < \delta < 1$,

$$\sum_{n>n_1} \sup_{\theta\in\mathscr{K}} P_\theta\left\{g_n^2 > \delta \sum_{j=1}^{n-1} g_{j-1}^2\right\} < \infty, \tag{A.41}$$

where $n_1 = n_1(r, \delta) = [r/(1 - \delta)] + 1$.

We have

$$P_\theta\left\{g_n^2 > \delta \sum_{j=1}^{n-1} g_j^2\right\} \le P_\theta\left\{g_n^2 > \delta^2(n-1)\right\} + P_\theta\left\{\sum_{j=1}^{n-1} g_j^2 \le \delta(n-1)\right\}. \tag{A.42}$$

By the Chebyshev inequality and (A.13),

$$\sup_{\theta\in\mathscr{K}} P_\theta\{g_n^2 > \delta^2(n-1)\} \le \sup_{\theta\in\mathscr{K}} \frac{E_\theta g_n^2}{\delta^2(n-1)^4} \le \frac{L}{(n-1)^4}. \tag{A.43}$$

Further, we make use of the representation

$$\sum_{j=1}^{n-1} g_j^2 = \frac{\operatorname{tr} F(n-r-1)}{d_1} + \frac{V'F^{-1}(\sum_{j=r+1}^{n-1}(Y_jY'_j - F))F^{-1}V \operatorname{tr} F}{V'F^{-1}Vd_1}.$$

From here it follows that for sufficiently small $\delta > 0$,

$$\begin{aligned} &P_\theta\left\{\sum_{j=1}^{n-1} g_j^2 \le \delta(n-1)\right\} \\ &\quad\le P_\theta\left\{\frac{\|V'F^{-1}\|^2}{V'F^{-1}V}\|\sum_{j=r+1}^{n-1}(Y_jY'_j - F)\| > (n-r-1) - \tilde{\delta}(n-1)\right\} \\ &\quad\le L\frac{E_\theta\|\sum_{j=r+1}^{n-1}(Y_jY'_j - F)\|^4}{((1-\tilde{\delta})n-r)^4}, \end{aligned} \tag{A.44}$$

where $\tilde{\delta} = \delta d_1/\operatorname{tr} F$. Now we note that

$$\sup_{\theta\in\mathscr{K}} E_\theta\|\sum_{j=r+1}^{n-1}(Y_jY'_j - F)\|^4 \le Ln^2, \qquad n \ge r+2. \tag{A.45}$$

This inequality can be shown by an argument similar to that used in the proof of (A.9). Combining (A.42)–(A.45) yields (A.41). This completes the proof of (A.39). It remains to verify (A.40).

We have

$$\begin{aligned} E_\theta|\Delta_1(H)|^4 &\le \frac{d^4(\operatorname{tr} F)^4}{(d_1)^4H^2} E(\varepsilon_1)^4 E_\theta|V'F^{-1}Y_\tau|^4 \\ &\le \frac{d^4(\operatorname{tr} F)^4(V'F^{-2}V)^2E(\varepsilon_1)^4E_\theta\|Y_\tau\|^4}{(d_1)^4H^2}. \end{aligned}$$

By (A.12) and (A.15),

$$\sup_{\theta\in\mathscr{K}} E_\theta\|Y_\tau\|^4 \le L(1+H). \tag{A.46}$$

Thus

$$\lim_{H\to\infty} \sup_{\theta\in\mathscr{K}} E_\theta|\Delta_1|^4 = 0. \tag{A.47}$$

Next consider $\Delta_2(H)$. On the set $\{\det G(H) > 0\}$, we have

$$|\Delta_2(H)| \le \frac{d}{d_1}\|V\|\|G^{-1}(H)H\|\left\|\frac{G(H)}{H} - \frac{d_1 F}{\operatorname{tr} F}\right\|\frac{\|M(H)\|}{\sqrt{H}}\|F^{-1}\operatorname{tr} F\|. \tag{A.48}$$

By the definition of τ and $R(H)$ in (2.13) and (A.34), it follows that

$$\sup_{\theta\in\mathscr{K}} E_\theta \frac{\|M(H)\|^2}{H} \le \sup_{\theta\in\mathscr{K}} E_\theta R(H) \le L < \infty. \tag{A.49}$$

Further, we show that

$$\lim_{H\to\infty}\sup_{\theta\in\mathscr{K}} E_\theta \left\|\frac{G(H)}{H} - \frac{d_1 F}{\operatorname{tr} F}\right\| = 0. \tag{A.50}$$

We have

$$\left\|\frac{G(H)}{H} - \frac{d_1 F}{\operatorname{tr} F}\right\| \le R(H)\left\|\frac{G(H)}{HR(H)} - \frac{F}{\operatorname{tr} F}\right\| + \frac{\|F\|}{\operatorname{tr} F}|R(H) - d_1|. \tag{A.51}$$

By (A.7)–(A.9),

$$\sup_{\theta\in\mathscr{K}} E_\theta (R(H))^4 \left\|\frac{G(H)}{HR(H)} - \frac{F}{\operatorname{tr} F}\right\|^4 \le L\frac{(1+H)^2}{H^4}. \tag{A.52}$$

By (A.37),

$$\sup_{\theta\in\mathscr{K}} E_\theta \frac{\|F\|}{\operatorname{tr} F}|R(H) - d_1| \le L\frac{1+\sqrt{r(H)}}{r(H)}. \tag{A.53}$$

Combining this inequality and (A.51), (A.52) yields (A.50). From (A.48)–(A.50) we obtain

$$\lim_{H\to\infty}\sup_{\theta\in\mathscr{K}} P_\theta\big\{|\Delta_2| > \delta,\ \det G(H) > 0\big\} = 0. \tag{A.54}$$

To analyze $\Delta_3(H)$ we need the following relationship:

$$\limsup_{H\to\infty}\sup_{\theta\in\mathscr{K}} E_\theta \frac{\tau_0(H)}{H} < \infty. \tag{A.55}$$

By definition of τ_0 in (A.38),

$$\begin{aligned} E_\theta \tau_0(H) &\le 1 + \sum_{n\ge 1} P_\theta\{\tau_0(H) > n\} \\ &\le 1 + \sum_{n\ge 1} P_\theta\left\{\sum_{j=1}^n g_j^2 < H\right\} \\ &\le 1 + \sum_{n\ge 1}\chi_{\{n\delta < H\}} + \sum_{n\ge[H/\delta]} P_\theta\left\{\sum_{j=1}^n g_j^2 < n\delta\right\}, \qquad \delta > 0. \end{aligned} \tag{A.56}$$

Further we shall make use of (A.44). Let $0 < \delta < \inf_{\theta\in\mathscr{K}} \operatorname{tr} F/d_1$. By (2.13) there exists such a number $H_0 > 0$ that for all $n \ge [H_0/\delta]$,

$$(1-\tilde{\delta})n - r > (1-\tilde{\delta})n/2.$$

Therefore, by (A.44), (A.45),

$$P_\theta\left\{\sum_{j=1}^{n} g_j^2 < n\delta\right\} \le L\frac{n^2}{((1-\tilde{\delta})n-r)^4} \le L/n^2, \qquad n \ge \left[\frac{H}{\delta}\right],$$

for all $H \ge H_0$. This estimate and (A.56) imply (A.55). Further, we show that

$$\limsup_{H\to\infty}\sup_{\theta\in\mathscr{K}} E_\theta \frac{(g_{\tau_0})^8}{H^2} = 0. \tag{A.57}$$

By the definition of g_k in (A.38) and (A.12),

$$|g_{\tau_0}|^8 \le L\left(1+\sum_{j=1}^{\tau_0}|\varepsilon_j|^8\right).$$

In view of (3.1), applying the Wald identity we obtain (A.57). Now we can estimate the second moment of $\Delta_3(H)$. From the definitions of Δ_3 and τ_0, it follows that

$$\begin{aligned}
E_\theta(\Delta_3(H))^2 &= \frac{(d_2(V))^2}{H}E_\theta\left|\sum_{k=1}^{\tau} g_k^2 - \sum_{k=1}^{\tau_0} g_k^2\right| \\
&\le (d_2(V))^2 E_\theta\left|\frac{\operatorname{tr} F}{Hd_1 V'F^{-1}V} V'F^{-1}\sum_{k=r+1}^{\tau} Y_k Y_k' F^{-1}V - 1\right| \\
&\quad + (d_2(V))^2\frac{E_\theta(g_{\tau_0})^2}{H} \\
&= (d_2(V))^2 E_\theta\left|\frac{(\tau-r)\operatorname{tr} F}{Hd_1} - 1\right| \\
&\quad + (d_2(V))^2\frac{\operatorname{tr} F}{Hd_1}\frac{|V'F^{-1}D(H)F^{-1}V|}{V'F^{-1}V} + (d_2(V))^2\frac{E_\theta(g_{\tau_0})^2}{H},
\end{aligned} \tag{A.58}$$

where the matrix $D(H)$ is defined by (A.5). Further, applying (A.6), (A.8), (A.9) and (A.53), we have

$$\begin{aligned}
&E_\theta\left|\frac{\tau(H)-r(H)}{H}\frac{\operatorname{tr} F}{d_1} - 1\right| \\
&\quad \le \frac{\operatorname{tr} FR(H)}{d_1}E_\theta\left|\frac{\tau(H)-r(H)}{HR(H)} - \frac{1}{\operatorname{tr} F}\right| + \frac{E_\theta|R(H)-d_1|}{d_1} \\
&\quad \le \frac{\sqrt{p+1}}{d_1 H}E_\theta\|D(H)\| + \frac{E_\theta\|Y_{\tau(H)}\|^2}{d_1} + \frac{E_\theta|R(H)-d_1|}{d_1} \\
&\quad \le L\left(\frac{1+\sqrt[4]{H}+\sqrt{H}}{H} + \frac{1+\sqrt{r(H)}}{r(H)}\right), \qquad \theta\in\mathscr{K}.
\end{aligned} \tag{A.59}$$

The second term in the right-hand side of (A.58), in view of (A.9), can be estimated as

$$\sup_{\theta\in\mathscr{K}} E_\theta(d_2(V))^2 \frac{\operatorname{tr} F}{Hd_1} \frac{|V'F^{-1}D(H)F^{-1}V|}{V'F^{-1}V} \le L\frac{1+\sqrt{H}}{H}.$$

Combining this estimate with (A.57)–(A.59) yields

$$\lim_{h\to\infty} \sup_{\theta\in\mathscr{K}} E_\theta(\Delta_3(H))^2 = 0.$$

The required conclusion (A.40) follows from here and (A.47), (A.54). Hence Proposition 4.2. □

A.7. Proof of Lemma 5.1. On the event $\Gamma(H)$ defined in (4.3), (2.10) and (2.33) imply

$$\tilde{\mu}(H) - \mu = \frac{\Delta(H)}{1 - 1'_p\lambda(H)}$$

where $\Delta(H) = V_0'(a(H) - a)$, $V_0 = (1,\ (\mu-\mu_0)1'_p)'$; $a(H)$ is defined by (2.15). Observe that if $|\mu - \mu_1| < \delta$, then

$$|\Delta(H)| \le L\|G^{-1}(H)\|\|M(H)\|, \tag{A.60}$$

where $M(H)$ is defined in (2.30). Now we estimate each of the quantities

$$I_1(H) = E_\theta|\tilde{\mu}(H) - \mu|^r \chi_{\Gamma(H)}, \tag{A.61}$$

$$I_2(H) = E_\theta|\tilde{\mu}(H) - \mu|^r \chi_{\Gamma^c(H)}, \tag{A.62}$$

By (2.7), (2.9) we have on the set $\Gamma(H)$ the inequality

$$|1 - 1'_p\lambda(H)| \ge 1/\sqrt{l(H)}$$

which implies for $0 < t < (1 - 1'_p\lambda)/2$ and $\lambda \in \Lambda$ the following estimate:

$$I_1(H) \le \frac{2^r}{(1-1'_p\lambda)} E_\theta|\Delta(H)|^r\chi_{\Gamma(H)} + (l(H))^{r/2}E_\theta|\Delta(H)|^r\chi_{\{\Gamma(H)\cap\{\|\lambda(H)-\lambda\|>t/\sqrt{p}\}\}}.$$

From here and (A.60) it follows that

$$\begin{aligned} I_1(H) \le{}& \frac{L}{(1-1'_p\lambda)} E_\theta\|HG^{-1}(H)\|^r\|M(H)/H\|^r\chi_{\Gamma(H)} \\ &+ L(l(H))^{r/2}E_\theta\|HG^{-1}(H)\|^r \\ &\times \|M(H)/H\|^r\chi_{\{\Gamma(H)\cap\{\|\lambda(H)-\lambda\|>t/\sqrt{p}\}\}}. \end{aligned} \tag{A.63}$$

Let us estimate $\|HG^{-1}(H)\|^r$ on the event $\Gamma(H)$. By the definition of σ in (2.21), the function $b(H)$ in (2.17) and due to the choice of the sequence c_k in (3.2), we have the following inclusions:

$$\begin{aligned}\{\sigma(H) \le n_0(H)\} &= \{n_0(H)b(H) \ge H\}\\ &= \left\{\frac{[Hl(H)]}{H} \ge (R(H)H\|G^{-1}(H)\|)^2\right\}\\ &\subset \left\{\sqrt{l(H)} \ge R(H)H\|G^{-1}(H)\|\right\}.\end{aligned}$$

Thus on the event $\Gamma(H)$

$$\|HG^{-1}(H)\| \le \frac{\sqrt{l(H)}}{R(H)} \le \sqrt{l(H)}. \tag{A.64}$$

By making use of this, we obtain

$$\begin{aligned}\|HG^{-1}(H)\|^r &\le \left(\|F^{-1}\| \operatorname{tr} F\right)^r\\ &\quad\times\left(1 + \|HG^{-1}(H)\|\left\|\frac{G(H)}{HR(H)} - \frac{F}{\operatorname{tr} F}\right\|\right)^r\\ &\le L\left(1 + (l(H))^{r/2}\left\|\frac{G(H)}{HR(H)} - \frac{F}{\operatorname{tr} F}\right\|^r\right).\end{aligned} \tag{A.65}$$

The matrix F is defined in (2.27). Combining (A.63)–(A.65) yields

$$\begin{aligned}I_1(H) &\le LE_\theta\|M(H)/H\|^r\\ &\quad + L(l(H))^{r/2}E_\theta\left(\frac{\|M(H)\|}{H}\left\|\frac{G(H)}{HR(H)} - \frac{F}{\operatorname{tr} F}\right\|\right)^r\\ &\quad + L(l(H))^r E_\theta\|M(H)/H\|^r \chi_{\{\|\lambda(H)-\lambda\|>t/\sqrt{p}\}}.\end{aligned} \tag{A.66}$$

Next we show that for $2 < \gamma \le 8$,

$$\sup_{|\mu-\mu_1|<\delta} E_\theta\|M(H)/H\|^\gamma \le L\frac{E|\varepsilon_1|^\gamma}{H^{\gamma/2}}. \tag{A.67}$$

Taking into account the definition of $\tau(H)$ in (2.10) and applying Lemma A.1 and the Hölder inequality, we obtain

$$\begin{aligned}E_\theta\|M(H)\|^\gamma &\le LE_\theta\left(\sum_{j=r(H)+1}^{\tau(H)-1}\|Y_j\|^2(\varepsilon_j)^2 + (\alpha(H))^2\|Y_{\tau(H)}\|^2(\varepsilon_{\tau(H)})^2\right)^{\gamma/2}\\ &\le LE_\theta\left(\sum_{j=r(H)+1}^{\tau(H)-1}\|Y_j\|^2 + (\alpha(H))^2\|Y_{\tau(H)}\|^2\right)^{\gamma/2-1}\\ &\quad\times\left(\sum_{j=r(H)+1}^{\tau(H)-1}\|Y_j\|^2|\varepsilon_j|^\gamma + (\alpha(H))^2\|Y_{\tau(H)}\|^2|\varepsilon_{\tau(H)}|^\gamma\right)\\ &\le LE|\varepsilon_1|^\gamma E_\theta(HR(H))^{\gamma/2}.\end{aligned} \tag{A.68}$$

By (2.13),

$$E_\theta(R(H))^{\gamma/2} \le E_\theta\Big(1 + \sum_{j=1}^{r(H)} \|Y_j\|^2/r(H)\Big)^{\gamma/2} \le 2^{\gamma/2-1}\Big(1 + \max_{j\ge 1} E_\theta\|Y_j\|^\gamma\Big).$$

This and (A.68) imply (A.67). Next we estimate the second term in the right-hand side of (A.66) with the help of (A.67) and (A.3), assuming $2 < r \le 8/3$. We have

$$\begin{aligned} &E_\theta\left\| \frac{G(H)}{HR(H)} - \frac{F}{\operatorname{tr} F}\right\|^r \|M(H)/H\|^r \\ &\quad\le \left(E_\theta\left\| \frac{G(H)}{HR(H)} - \frac{F}{\operatorname{tr} F}\right\|^4\right)^{r/4} \big(E_\theta\|M(H)/H\|^{4r/(4-r)}\big)^{(4-r)/4} \le \frac{L}{H^r}. \end{aligned} \tag{A.69}$$

The last term in the right-hand side of (A.66) can be estimated by applying the Cauchy–Bunyakovsky inequality and (3.8):

$$\begin{aligned} &E_\theta\|M(H)/H\|^r \chi_{\{\|\lambda^*(H)-\lambda\|>t/\sqrt{p}\}} \\ &\quad\le \big(E_\theta\|M(H)/H\|^{2r}\big)^{1/2}\sqrt{P_\theta\{\|\lambda^*(H)-\lambda\| > t/\sqrt{p}\}} \le Ll(H)/H^{1+r/2}, \end{aligned}$$

when H is sufficiently large. This and (A.66), (A.67) and (A.69) imply

$$I_1(H) \le \frac{L}{H^{r/2}}\left(1 + \frac{(l(H))^{r/2}}{H^{r/2}} + \frac{(l(H))^{1+r}}{H}\right) \tag{A.70}$$

for sufficiently large H. It remains to estimate $I_2(H)$ in (A.62). We have

$$I_2(H) \le \big(E_\theta(\tilde\mu(H)-\mu)^4\big)^{r/4}\big(P_\theta\{\Gamma^c(H)\}\big)^{(4-r)/4}. \tag{A.71}$$

From (2.9), (2.10), (2.33), (3.2) and (3.8) it follows that

$$\begin{aligned} E_\theta(\tilde\mu(H)-\mu)^4 &\le \frac{\|V_0\|^4}{H^2} E_\theta\Big(\sum_{j=1}^{\nu(H)} \|a^*(h_j) - a\|^2\Big)^2 \\ &\le \frac{\|V_0\|^4}{H^2}\Big(\sum_{j\ge 1}\sqrt{E_\theta\|a^*(h_j)-a\|^4}\Big)^2 \le L\frac{(l(H))^4}{H^2} \end{aligned} \tag{A.72}$$

for $|\mu - \mu_1| < \delta$, $\lambda \in \Lambda$.

Further, by Propositions 3.1 and 3.2,

$$\lim_{H\to\infty} \sup_{|\mu-\mu_1|<\delta} HP_\theta\{\Gamma^c\} = 0, \qquad \lambda \in \Lambda. \tag{A.73}$$

Combining (A.71)–(A.73) yields

$$I_2(H) \le L\frac{(l(H))^r}{H^{1+r/2}}.$$

This estimate and (A.70) lead to (5.14) for $2 < r \le 8/3$. Hence Lemma 5.1. □

Acknowledgments. We are grateful to referees, an Associate Editor and the Editor, John Rice, for their valuable comments.

REFERENCES

AKRITAS, M. G. and JOHNSON, R. A. (1982). Efficiencies of tests and estimators for p-order autoregressive processes when the error distribution is nonnormal. *Ann. Inst. Statist. Math.* **34** 579–589.

ANDERSON, T. W. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations. *Ann. Math. Statist.* **30** 676–687.

ANDERSON, T. W. (1984). *The Statistical Analysis of Time Series*. Wiley, New York.

ARAS, G. (1990). Sequential estimation of the mean of a first-order autoregressive process. *Comm. Statist. Theory Methods* **19** 1639–1652.

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

BORISOV, V. Z. and KONEV, V. V. (1977). On sequential parameters estimation in discrete time processes. *Automat. Remote Control* **6** 58–64.

CHOW, Y. S. and MARTINSEK, A. T. (1982). Bounded regret of sequential procedure for estimation of the mean. *Ann. Statist.* **10** 909–914.

CHOW, Y. S. and POBBINS, H. E. (1965). On the asymptotic theory of fixed width sequential confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

CHOW, Y. S. and YU, K. F. (1981). The performance of a sequential procedure for the estimation of the mean. *Ann. Statist.* **9** 184–189.

DMITRIENKO, A. A. and KONEV, V. V. (1994). On guaranteed estimation of autoregression parameters when the noise variance is unknown. *Automat. Remote Control* **2** 87–99.

EFROIMOVICH, S. YU. (1980). Sequential estimation under the conditions of local asymptotic normality. *Teor. Veroyatnost i Primenen* **25** 30–43. (In Russian.)

FAKHRE-ZAKERI, I. and LEE, S. (1992). Sequential estimation of the mean of a linear process. *Sequential Anal.* **11** 181–197.

FAKHRE-ZAKERI, I. and LEE, S. (1993). Sequential estimation of the mean vector of a multivariate linear process. *J. Multivariate Anal.* **47** 196–209.

GHOSH, M. and MUKHOPADHYAY, N. (1979). Sequential point estimation of the mean when the distribution is unspecified. *Comm. Statist. Theory Meth.* **A8** 637–651.

GOVINDARAJULU, Z. (1987). *The Sequential Statistical Analysis of Hypothesis Testing. Point and Interval Estimation and Decision Theory.* American Sciences Press.

GREENWOOD, P. E. and SHIRYAEV, A. N. (1992). Asymptotic minimaxity of a sequential estimator for a first order autoregressive model. *Stochastics Stochastics Rep.* **38** 49–65.

GRENANDER, U. (1950). *Stochastic Processes and Statistical Inference.* Almqvist and Wiksells, Stockholm.

HAJEK, J. (1972). Local asymptotic minimax and admissibility in estimation. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **1** 175–194. Univ. California Press, Berkeley.

IBRAGIMOV, I. A. and HAS'MINSKI, R. Z. (1974). Sequential estimation. *Teor. Veroyatnost i Primenen* **19** 245–256. (In Russian.)

IBRAGIMOV, I. A. and HAS'MINSKI, R. Z. (1981). *Statistical Estimation Asymptotic Theory.* Springer, Berlin.

KONEV, V. V. and LAI, T. L. (1995). Estimation with prescribed precision in stochastic regression models. *Seqential Anal.* **14** 179–192.

KONEV, V. V. and PERGAMENSHCHIKOV, S. M. (1981). Sequential identification plans for dynamic systems. *Automat. Remote Control.* **7** 84–92.

LAI, T. L. and SIEGMUND, D. (1983). Fixed-accuracy estimation of an autoregressive parameter. *Ann. Statist.* **11** 478–485.

LE CAM, L. (1980). Locally asymptotically normal families of distributions. *Univ. Calif. Publ. Statist.* **3** 27–98.

LEE, S. (1994). Sequential estimation for the parameter of a stationary autoregressive model. *Sequential Anal.* **13** 301–314.

LIPTSER, R. SH. and SHIRYAEV, A. N. (1986). Theory of martingales.

SEN, P. K. (1982). *Sequential Nonparametrics: Invariance Principles and Statistical Inference*. Wiley, New York

SHIRYAEV, A. N. (1984). *Probability*. Springer, New York.

SHIRYAEV, A. N. and SPOKOINY, V. G. (1993). Asymptotic properties of the maximum likelihood estimators under random normalization for a first order autoregressive model. *Frontiers in Pure and Applied Probab.* **1** 223–227.

SIEGMUND, D. (1985). *Sequential Analysis: Test and Confidence Intervals.* Springer, New York.

SRIRAM, T. N. (1987). Sequential estimation of the mean of a first-order stationary autoregressive process. *Ann. Statist.* **15** 1079–1090.

SRIRAM, T. N. (1988). Sequential estimation of the autoregressive parameter in a first-order autoregressive process. *Sequential Anal.* **7** 53–74.

SRIRAM, T. N. and BOSE, A. (1988). Sequential strinkage estimation of the parameter in the general linear model. *Sequential Anal.* **7** 149–163.

STEIN, CH. (1945). A two-sample test for a linear hypothesis whose power is independent of the variance. *Ann. Math. Statist.* **16** 243–258.

WOODROOFE, M. (1982). Nonlinear renewal theory in sequential analysis. IMS, Hayward, CA.

ZACKS, S. (1971). *The Theory of Statistical Inference*. Wiley, New York.

DEPARTMENT OF APPLIED MATHEMATICS
AND CYBERNETICS
UNIVERSITY OF TOMSK
LENIN STREET 36
634050 TOMSK 50
RUSSIA
E-MAIL: vvkonev@vmn.tsu.tomsk.su