The Annals of Statistics
1999, Vol. 27, No. 4, 1385–1414

OPTIMAL TESTS FOR AUTOREGRESSIVE MODELS BASED ON AUTOREGRESSION RANK SCORES

BY MARC HALLIN[1] AND JANA JUREČKOVÁ[2]

Université Libre de Bruxelles and Charles University, Prague

Locally asymptotically optimal tests based on autoregression rank scores are constructed for testing linear constraints on the structural parameters of AR processes. Such tests are asymptotically distribution free and do not require the estimation of nuisance parameters. They constitute robust, flexible and quite powerful alternatives to existing methods such as the classical correlogram-based parametric tests, the Gaussian Lagrange multiplier tests, the optimal non-Gaussian and *ranked residual* tests described by Kreiss, as well as to the aligned rank tests of Hallin and Puri. Optimality requires a nontrivial extension of existing asymptotic representation results to the case of unbounded score functions (such as the Gaussian quantile function). The problem of testing AR($p-1$) against AR(p) dependence is considered as an illustration. Asymptotic local powers and asymptotic relative efficiencies are explicitly computed. In the special case of van der Waerden scores, the asymptotic relative efficiency with respect to optimal correlogram-based procedures is uniformly larger than one.

1. Introduction. Koenker and Bassett (1978) introduced the concept of *regression quantile* extending the classical notion of sample quantile to the linear regression model. Ruppert and Carroll (1980) showed that these regression quantiles admit an asymptotic representation analogous to that of Bahadur (1966), a result which was completed by Jurečková (1984), where the exact rate for this representation is provided. Portnoy (1991) obtained their asymptotic representation for linear models with dependent errors, under the assumption of *m-decomposability* in the sense of Chanda, Puri, and Ruymgaart (1990). Koul and Saleh (1995) introduced the *auto*regression quantiles as an extension of regression quantiles to the context of AR models, and derived their asymptotic representation.

Regression rank scores, which are related to regression quantiles through a duality property in the linear programming sense, were introduced by Gutenbrunner and Jurečková (1992) for the general linear model with independent observations. Regression rank score statistics, under appropriate technical conditions, are asymptotically equivalent to the corresponding simple linear

Received January 1997; revised July 1999.

[1]Supported in part by Human Capital and Mobility Contract ERB CT CHRX 963 and by the Fonds d'Encouragement à la Recherche de l'Université Libre de Bruxelles.

[2]Supported in part by Czech Republic Grants 201/96/0230 and 201/99/0264, and a *Research in Brussels* Grant of the Région de Bruxelles Capitale.

AMS 1991 *subject classifications*. 62M10, 62F35, 62G10.

Key words and phrases. Time series, autoregressive models, asymptotic tests, autoregression rank scores.

rank statistics computed from the ranks of the nonobservable errors. As such, they are *asymptotically invariant* with respect to the group of order-preserving transformations acting on the errors, hence asymptotically distribution free, and thus provide an alternative to the usual aligned rank statistics, where the ranks are those of some estimated residuals. Contrary to aligned ranks, however, regression rank scores do not rely on any estimation of nuisance parameters, and therefore are not affected by possibly poor or nonrobust estimates. Moreover, provided the appropriate technical conditions are satisfied, the asymptotic equivalence of regression rank score statistics with the corresponding rank statistic computed from the unobservable errors holds for all of them, whereas the same equivalence does not hold for aligned rank statistics unless some orthogonality conditions are satisfied. Regression rank score statistics are thus particularly suitable for testing linear constraints on the parameters of linear regression models; regression rank score tests in this context are indeed asymptotically equivalent to, hence asymptotically as efficient as, their genuinely distribution-free rank-based counterparts for the model in which the value of nuisance parameters would be specified.

Tests of linear subhypotheses based on regression rank scores first appeared in Gutenbrunner and Jurečková (1992). However, the test statistics considered there are based on bounded score-generating functions whose support is a compact subinterval of $(0, 1)$. Gutenbrunner, Jurečková, Koenker and Portnoy (1993), under appropriate assumptions on the tails of the underlying distributions, extended these tests to score functions whose support is the entire interval $(0, 1)$. A different approach [Jurečková (1999)], based on sequences of score functions which are constant outside $[\alpha_n, 1 - \alpha_n]$, $\lim_{n\to\infty} \alpha_n = 0$, avoids these conditions and allows for possibly heavy-tailed densities. In a slightly different direction, Jurečková (1991) proposes a regression rank score test of the Kolmogorov–Smirnov type.

A fairly complete theory of tests based on regression rank scores thus exists in the linear regression model with independent observations. In view of their flexibility, their robustness, their asymptotic distribution freeness and their excellent asymptotic performances, such tests are extremely attractive and certainly could be recommended to practitioners. In particular, the asymptotic relative efficiencies of regression rank score tests based on van der Waerden (normal) scores, with respect to their normal-theory counterparts based on traditional Student or Fisher–Snedecor procedures, are uniformly larger than one; see Chernoff and Savage (1958).

All these properties sound even more attractive in the context of time series, where the need for robust and non-Gaussian procedures has been stressed by many authors. Also, adequately defined tests based on analogues of regression rank scores can be expected to retain most of the advantages just described in the context of AR and other time series models. A first step towards the application of these techniques in a time-series context was taken by Koul and Saleh (1995) who introduce the concept of autoregression rank scores and establish some of their basic properties similar to those in

Gutenbrunner and Jurečková (1992). However, these authors do not use them for testing purpose, and their assumptions rule out the construction of locally asymptotically optimal tests corresponding to unbounded score functions, such as the normal one. The results in this paper allow for such scores, the practical importance of which is substantiated by a time-series version of Chernoff and Savage's (1958) result [see Hallin (1994), or Hallin and Werker (1998)]; the testing methods based on van der Waerden (normal score) rank-based autocorrelations uniformly outperform, under finite Fisher information conditions, their traditional counterparts, based on traditional correlograms.

The main objective of the present paper is to show that locally asymptotically optimal tests based on autoregression rank scores and possibly unbounded score functions can be constructed for testing linear hypotheses on the parameters of AR processes. These tests enjoy all the attractive features of regression rank score tests in regression models with independent observations. They are natural competitors of the classical correlogram-based parametric tests such as Gaussian Lagrange multipliers [see Godfrey (1979) or Hosking (1980)], of the optimal non-Gaussian parametric and the *ranked residual* tests described by Kreiss (1990), and of the aligned rank tests of Hallin and Puri (1994). Whereas their performances are comparable, or even uniformly better [see Section 4.2 for asymptotic relative efficiencies (AREs)], they do not require any consistent estimation of nuisance parameters.

Autoregression quantiles and autoregression rank scores are briefly defined in Section 2. Our definition, which includes an arbitrary $p \times q$ matrix $\mathbf{Q}$, is slightly more general than the traditional one; the presence of this matrix $\mathbf{Q}$, as we shall see, is motivated by the nature of the testing problem under study. Section 3 presents the main theoretical results on the asymptotic behavior of linear autoregression rank score statistics with possibly unbounded scores. These results constitute an extension to the unbounded score case of those obtained by Koul and Saleh (1995) in the bounded score case. These statistics then are used in the derivation of the optimal testing procedures described in Section 4. The general problem of testing arbitrary linear constraints on the parameter $\boldsymbol{\theta}$ of an AR(p) model is addressed in Section 4.1. Section 4.2 concentrates on the particular case of testing AR($p-1$) against AR(p) dependence, with obvious implications in the context of order identification. Section 4.3 discusses the asymptotic relative efficiencies of our autoregression rank score tests with respect to their main parametric and nonparametric competitors.

The numerical performance of the tests we are proposing is illustrated in Hallin, Jurečková, Kalvová, Picek and Zahaf (1997) and Kalvová, Jurečková, Nemešová and Picek (1998).

2. Autoregression rank scores.

2.1. *Autoregression quantiles and autoregression rank scores.* Consider the autoregressive model under which $(X_{-p+1}, \dots, X_0, X_1, \dots, X_n)'$, an observed series of length $(n+p)$, satisfies the linear recursion of order p [AR(p) model],

$$X_t = \theta_1 X_{t-1} + \cdots + \theta_p X_{t-p} + \varepsilon_t, \qquad t \in \mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}, \tag{2.1}$$

where $\{\varepsilon_t,\ t \in \mathbb{Z}\}$ is a sequence of i.i.d. random variables with distribution function F and density f (the *innovation density*). We do not assume that f is known; we only require that it belongs to a family $\mathscr{F}$ of densities satisfying

$$\int_{-\infty}^{\infty} x\, dF(x) = 0, \qquad 0 < \int_{-\infty}^{\infty} x^2\, dF(x) = \sigma^2 < \infty, \tag{2.2}$$

and a few other assumptions [(F1(a), (b) and (F2); see Section 3.1]. As usual, the autoregression parameter $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)' \in \mathbb{R}^p$ is supposed to be such that the polynomial

$$\theta(z) := 1 - \sum_{i=1}^{p} \theta_i z^i, \qquad z \in \mathbb{C} \tag{2.3}$$

has no root within the unit disk (the classical *causality* assumption; see, e.g., Brockwell and Davis (1991)). The set of parameter values satisfying this assumption is denoted as $\boldsymbol{\Theta}$.

The causality assumption ensures the existence and unicity of a strictly stationary solution of the stochastic difference equation (2.1). Throughout, we assume that the observed series is a finite realization of that stationary solution. Our objective is to test the null hypothesis that the parameter $\boldsymbol{\theta}$ lies in some linear subspace of $\boldsymbol{\Theta}$, while the innovation density f remains unspecified. This subspace can be characterized as the intersection of $\boldsymbol{\Theta}$ and the linear subspace $\mathscr{L}(\mathbf{Q})$ of $\mathbb{R}^p$ spanned by the columns of some $p \times q$ matrix $\mathbf{Q}$ of maximal rank q $(q \le p)$. Equivalently, denoting by $\mathbf{Q}^{\perp}$ a $p \times (p-q)$ matrix whose columns constitute a basis of the $(p-q)$-dimensional subspace of $\mathbb{R}^p$ orthogonal to $\mathscr{L}(\mathbf{Q})$, this hypothesis can be characterized by the $p-q$ linear constraints on $\boldsymbol{\theta} \in \boldsymbol{\Theta}$,

$$\mathbf{Q}^{\perp\prime}\boldsymbol{\theta} = \mathbf{0}. \tag{2.4}$$

In the context of linear models with independent observations, such hypotheses easily reduce to a *canonical form* [see, e.g., Scheffé (1959) Section 2.6] by means of some adequate linear reparametrization. Under this new parametrization, $\mathbf{Q}'$ has the simpler form $\mathbf{Q}_0' = (\mathbf{I}_q \vdots \mathbf{0})$, where $\mathbf{I}_q$ stands for the $q \times q$ identity matrix. Such reparametrizations also can be considered (see Section 2.2) in the autoregressive context, but they do not preserve the autoregressive form of the model. This is why a generalized form of autoregression quantiles and autoregression rank scores involving the matrix $\mathbf{Q}$ is introduced here.

Denote by $\mathbf{Q}$ a full-rank $(p \times q)$ $(p > q)$ matrix of constants, and by $\mathscr{L}(\mathbf{Q})$ the linear q-dimensional subspace of $\mathbb{R}^p$ spanned by the columns of $\mathbf{Q}$: the null hypotheses considered in Section 4 are of the form $\boldsymbol{\theta} \in \boldsymbol{\Theta} \cap \mathscr{L}(\mathbf{Q})$. In order to test such linear hypotheses, we need slightly more general concepts of autoregression quantiles and autoregression rank scores than in Koul and

Saleh (1995). Letting

$$\mathbf{y}_t^* := (X_t, \dots, X_{t-p+1})' \quad \text{and} \quad \mathbf{y}_t := (1, \mathbf{y}_t^{*\prime})', \qquad t = 0, \dots, n-1, \tag{2.5}$$

consider the [$n \times p$ and $n \times (p+1)$, respectively] random matrices,

$$\mathbf{Y}_n^* := (\mathbf{y}_0^*, \dots, \mathbf{y}_{n-1}^*)' \quad \text{and} \quad \mathbf{Y}_n := (\mathbf{y}_0, \dots, \mathbf{y}_{n-1})'$$

and define the *α-autoregression quantile associated with* $\mathbf{Q}$ as a solution

$$\hat{\boldsymbol{\rho}}_{\mathbf{Q};n}(\alpha) := (\hat{\rho}^0_{\mathbf{Q};n}(\alpha), \hat{\boldsymbol{\rho}}^{1\prime}_{\mathbf{Q};n}(\alpha))', \qquad \hat{\rho}^0_{\mathbf{Q};n} \in \mathbb{R}, \hat{\boldsymbol{\rho}}^1_{\mathbf{Q};n} \in \mathbb{R}^q$$

[in short, $\hat{\boldsymbol{\rho}}(\alpha) = (\hat{\rho}_0(\alpha), \hat{\boldsymbol{\rho}}_1'(\alpha))'$] of

$$\sum_{t=1}^{n} h_\alpha(X_t - r_0 - \mathbf{y}_{t-1}^{*\prime}\mathbf{Q}\mathbf{r}_1) := \min, \tag{2.6}$$

where the minimum is taken with respect to $\mathbf{r} = (r_0, \mathbf{r}_1')' \in \mathbb{R}^{q+1}$, and

$$h_\alpha(u) := |u|(\alpha I[u > 0] + (1-\alpha) I[u \le 0]), \qquad u \in \mathbb{R}, \alpha \in [0,1].$$

Subscripts n and $\mathbf{Q}$ are omitted whenever no confusion is possible. It follows from the definition that $(\hat{\rho}_0(\alpha), (\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha))')'$ formally coincides with Koenker and Bassett's (1978) α-regression quantile in the linear model with design matrix $\mathbf{Y}_n$ and response vector $\mathbf{X}_n := (X_1, \dots, X_n)'$. Accordingly, $\hat{\boldsymbol{\rho}}(\alpha)$ coincides with the first $(q+1)$ components, $\hat{\mathbf{r}} = (\hat{r}_0, \hat{\mathbf{r}}_1')'$, of the optimal solution $(\hat{\mathbf{r}}, \hat{\boldsymbol{\mu}}^+, \hat{\boldsymbol{\mu}}^-)$ of the parametric linear programming problem,

$$\begin{aligned} &\alpha \mathbf{1}_n' \boldsymbol{\mu}^+ + (1-\alpha)\mathbf{1}_n'\boldsymbol{\mu}^- := \min \\ &\mathbf{X}_n - \mathbf{1}_n r_0 - \mathbf{Y}_n^* \mathbf{Q}\mathbf{r}_1 = \boldsymbol{\mu}^+ - \boldsymbol{\mu}^-, \\ &r_0 \in \mathbb{R}, \mathbf{r}_1 \in \mathbb{R}^q, \boldsymbol{\mu}^\pm \in \mathbb{R}^n_+, \qquad 0 \le \alpha \le 1, \end{aligned} \tag{2.7}$$

where $\mathbf{1}_n$ stands for the n-dimensional vector $(1, \dots, 1)'$.

The dual program associated with (2.7) is

$$\begin{aligned} &\mathbf{X}_n' \mathbf{a} := \max, \\ &\mathbf{1}_n' \mathbf{a} = n(1-\alpha), \\ &\mathbf{Q}'\mathbf{Y}_n^{*\prime}(\mathbf{a} - (1-\alpha)\mathbf{1}_n) = \mathbf{0}, \\ &\mathbf{a} \in [0,1]^n, \qquad 0 \le \alpha \le 1. \end{aligned} \tag{2.8}$$

Following Jurečková and Gutenbrunner (1992) and Koul and Saleh (1995), we call the components of the optimal solution $\hat{\mathbf{a}}_{\mathbf{Q};n}(\alpha) := (\hat{a}_{\mathbf{Q};n;1}(\alpha), \dots, \hat{a}_{\mathbf{Q};n;n}(\alpha))'$, $0 \le \alpha \le 1$, of (2.8) the *auto*regression rank scores associated with $\mathbf{Q}$.

Note that $\hat{a}_{\mathbf{Q};n;t}(\cdot)$ [in short, $\hat{a}_t(\cdot)$] is a continuous, piecewise linear function, with $\hat{a}_t(0) = 1$, $\hat{a}_t(1) = 0$, $t = 1, \dots, n$; moreover, letting $\mathscr{T}(\alpha) := \{t | X_t = \hat{\rho}_0(\alpha) + \mathbf{y}_{t-1}^{*\prime}\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha)\}$,

$$\hat{a}_t(\alpha) = \begin{cases} 1, & \text{if } X_t > \hat{\rho}_0(\alpha) + \mathbf{y}_{t-1}^{*\prime}\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha), \\ 0, & \text{if } X_t < \hat{\rho}_0(\alpha) + \mathbf{y}_{t-1}^{*\prime}\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha), \end{cases} \qquad t \notin \mathscr{T}(\alpha), \tag{2.9}$$

whereas, for $t \in \mathscr{T}(\alpha)$, $\hat{a}_t(\alpha)$ is obtained as the solution of the system of $q+1$ linear equations

$$\sum_{t \in \mathscr{T}(\alpha)} \begin{pmatrix} 1 \\ \mathbf{Q}'\mathbf{y}^*_{t-1} \end{pmatrix} \hat{a}_t(\alpha) = (1-\alpha) \sum_{t=1}^{n} \begin{pmatrix} 1 \\ \mathbf{Q}'\mathbf{y}^*_{t-1} \end{pmatrix} - \sum_{t=1}^{n} \begin{pmatrix} 1 \\ \mathbf{Q}'\mathbf{y}^*_{t-1} \end{pmatrix} I\left[X_t > \hat{\rho}_0(\alpha) + \mathbf{y}^{*\prime}_{t-1}\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha)\right].$$

One of the most appealing properties of the autoregression rank scores $\hat{\mathbf{a}}_{\mathbf{Q};\,n}(\alpha)$ is that they are *autoregression invariant*. More precisely, (2.8) implies that $\hat{\mathbf{a}}_{\mathbf{Q};\,n}(\alpha)$ formally also could be written as a solution of the linear program

$$\begin{aligned} &\boldsymbol{\varepsilon}'\mathbf{a} := \max, \\ &\mathbf{1}'_n\mathbf{a} = n(1-\alpha), \\ &\mathbf{Q}'\mathbf{Y}^{*\prime}_n(\mathbf{a} - (1-\alpha)\mathbf{1}_n) = \mathbf{0}, \\ &\mathbf{a} \in [0,1]^n, \qquad 0 \le \alpha \le 1, \end{aligned} \tag{2.10}$$

where $\boldsymbol{\varepsilon} := (\varepsilon_1, \dots, \varepsilon_n)'$ is the unobservable white noise process. This autoregression-invariance property is essential to our approach of the testing problem considered here. It follows indeed from (2.10) that $\hat{\mathbf{a}}_{\mathbf{Q};\,n}(\alpha)$, hence our test statistic (4.1), does not explicitly depend on the unknown parameters $\theta_1, \dots, \theta_p$, so that inference based on $\hat{\mathbf{a}}_{\mathbf{Q};\,n}(\alpha)$ does not require any preliminary estimation of $\boldsymbol{\theta}$.

The algebraic relations between the autoregression quantiles and autoregression rank scores, proved in the following lemma, help to understand the structure of these concepts.

LEMMA 2.1. (i) *Let* $0 < \alpha_1 < \alpha_2 < 1$. *Then*

$$\begin{aligned} (\alpha_2 - \alpha_1)&\left[\hat{\rho}_0(\alpha_1) + \left(n^{-1}\sum_{t=1}^{n}\mathbf{y}^*_{t-1}\right)'\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha_1)\right] \\ &\le -\frac{1}{n}\sum_{t=1}^{n} X_t(\hat{a}_t(\alpha_2) - \hat{a}_t(\alpha_1)) \\ &\le (\alpha_2 - \alpha_1)\left[\hat{\rho}_0(\alpha_2) + \left(n^{-1}\sum_{t=1}^{n}\mathbf{y}^*_{t-1}\right)'\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha_2)\right]. \end{aligned} \tag{2.11}$$

(ii) *If* $\alpha \in (0,1)$ *is a continuity point of* $\hat{\boldsymbol{\rho}}(\alpha)$, *then*

$$\hat{\rho}_0(\alpha) + \left(n^{-1}\sum_{t=1}^{n}\mathbf{y}^*_{t-1}\right)'\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha) = -\sum_{t=1}^{n} X_t \frac{d}{d\alpha}\hat{a}_t(\alpha). \tag{2.12}$$

(iii) $\hat{\rho}_0(\alpha) + (n^{-1}\sum_{t=1}^{n}\mathbf{y}^*_{t-1})'\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha)$, *hence also* $-\sum_{t=1}^{n} X_t(d/d\alpha)\hat{a}_t(\alpha)$, *are nondecreasing step-functions of* $\alpha \in [0,1]$.

PROOF. For all fixed $\alpha \in [0,1]$, the duality between $\hat{\boldsymbol{\rho}}(\alpha)$ and $\hat{\mathbf{a}}(\alpha)$ implies that

$$\sum_{t=1}^{n} h_\alpha(X_t - \hat{\rho}_0(\alpha) - \mathbf{y}^{*\prime}_{t-1}\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha)) = \sum_{t=1}^{n} X_t(\hat{a}_t(\alpha) - (1-\alpha)).$$

Note that, for all $0 \le \alpha_1 \le \alpha_2 \le 1$ and $u \in \mathbb{R}$, we have $h_{\alpha_2}(u) - h_{\alpha_1}(u) = (\alpha_2 - \alpha_1)u$. Hence,

$$\begin{aligned}
(\alpha_2 - \alpha_1) &\sum_{t=1}^{n} (X_t - \hat{\rho}_0(\alpha_1) - \mathbf{y}^{*\prime}_{t-1}\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha_1)) \\
&= \sum_{t=1}^{n} \big[h_{\alpha_2}(X_t - \hat{\rho}_0(\alpha_1) - \mathbf{y}^{*\prime}_{t-1}\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha_1)) \\
&\qquad\qquad - h_{\alpha_1}(X_t - \hat{\rho}_0(\alpha_1) - \mathbf{y}^{*\prime}_{t-1}\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha_1)) \big] \\
&\ge \sum_{t=1}^{n} X_t[\hat{a}_t(\alpha_2) - \hat{a}_t(\alpha_1) + (\alpha_2 - \alpha_1)],
\end{aligned}$$

where the last inequality follows from the definition of $\hat{\boldsymbol{\rho}}(\alpha)$ as a minimizer of (2.6). Thus,

$$(\alpha_2 - \alpha_1) \sum_{t=1}^{n} (\hat{\rho}_0(\alpha_1) - \mathbf{y}^{*\prime}_{t-1}\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha_1)) \le - \sum_{t=1}^{n} X_t(\hat{a}_t(\alpha_2) - \hat{a}_t(\alpha_1)).$$

Similarly, we have

$$(\alpha_2 - \alpha_1) \sum_{t=1}^{n} (\hat{\rho}_0(\alpha_2) - \mathbf{y}^{*\prime}_{t-1}\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha_2)) \ge - \sum_{t=1}^{n} X_t(\hat{a}_t(\alpha_2) - \hat{a}_t(\alpha_1)),$$

which leads to (2.11) and entails the monotonicity of $\hat{\rho}_0(\alpha) + (n^{-1}\sum_{t=1}^{n} \mathbf{y}^{*}_{t-1})'\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha)$. On the other hand, $\hat{\boldsymbol{\rho}}(\alpha)$ is a step-function, $\hat{\mathbf{a}}(\alpha)$ is a piecewise linear function of α, and their points of discontinuity coincide. This proves (ii), and, jointly with (i), further implies (iii). □

2.2. *Linear reparametrization.* The following linear reparametrization of (2.1) will be convenient in the context of the testing problem briefly described at the beginning of Section 2.1 and will simplify the proofs. Recall that $\mathbf{Q}^{\perp}$ denotes a $p \times (p-q)$ matrix whose columns are spanning the $(p-q)$-dimensional subspace of $\mathbb{R}^p$ orthogonal to $\mathscr{L}(\mathbf{Q})$. The matrices $\mathbf{Q}$ and $\mathbf{Q}^{\perp}$ here clearly are defined up to a positive multiplicative constant, so that we safely can assume that the $p \times p$ full-rank matrix $\mathbf{A}' := (\mathbf{Q} \vdots \mathbf{Q}^{\perp})$ has modulus one: $\|\mathbf{A}\| = 1$ [all usual matrix norms are equivalent here; one may choose, for instance, the Euclidean norm $\|\mathbf{A}\| := [\sum_{i,j}(A_{i,j})^2]^{1/2}$].

Letting $\mathbf{X}_n := (X_1, \dots, X_n)'$ and $\boldsymbol{\varepsilon}_n := (\varepsilon_1, \dots, \varepsilon_n)'$, the AR model (2.1) then also can be written as

$$\begin{aligned}
\mathbf{X}_n &= \mathbf{Y}^*_n\boldsymbol{\theta} + \boldsymbol{\varepsilon}_n = (\mathbf{Y}^*_n\mathbf{A}^{-1})(\mathbf{A}\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_n \\
&:= \check{\mathbf{Y}}^*_n\boldsymbol{\beta} + \boldsymbol{\varepsilon}_n = \check{\mathbf{Y}}^*_{I;n}\boldsymbol{\beta}_I + \check{\mathbf{Y}}_{II;n}\boldsymbol{\beta}_{II} + \boldsymbol{\varepsilon}_n,
\end{aligned} \tag{2.13}$$

with

$$\boldsymbol{\beta} := \begin{pmatrix} \boldsymbol{\beta}_I \\ \boldsymbol{\beta}_{II} \end{pmatrix} := \mathbf{A}\boldsymbol{\theta} = \begin{pmatrix} \mathbf{Q}'\boldsymbol{\theta} \\ \mathbf{Q}^{\perp\prime}\boldsymbol{\theta} \end{pmatrix}$$

and

$$\check{\mathbf{Y}}_n^* := \left(\check{\mathbf{Y}}_{I;n}^* \vdots \check{\mathbf{Y}}_{II;n}^*\right) := \mathbf{Y}_n^*\mathbf{A}^{-1} = \mathbf{Y}_n^*\left(\mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1} \vdots \mathbf{Q}^{\perp}(\mathbf{Q}^{\perp\prime}\mathbf{Q}^{\perp})^{-1}\right).$$

The null hypothesis (2.4) under this reparametrization takes the simple form $\boldsymbol{\beta}_{II} = \mathbf{0}$ or, equivalently, $\boldsymbol{\beta} \in \mathbf{A}\boldsymbol{\Theta} \cap \mathscr{L}(\mathbf{Q}_0)$, with $\mathbf{Q}_0' := (\mathbf{I}_q \vdots \mathbf{0})$; $\mathbf{I}_q$ stands for the $q \times q$ identity matrix.

If the linear program (2.7) defining the autoregression quantiles is rewritten in this new parametrization [i.e., if $\check{\mathbf{Y}}_n^*$ is substituted for $\mathbf{Y}_n^*$ and $\mathbf{Q}_0$ for $\mathbf{Q}$], the resulting quantiles, $(\check{\rho}_0(\alpha), \check{\boldsymbol{\rho}}_1'(\alpha))'$ are such that $(\check{\rho}_0(\alpha), \check{\boldsymbol{\rho}}_1'(\alpha))' = (\rho_0(\alpha), \boldsymbol{\rho}_1'(\alpha)\mathbf{Q}'\mathbf{Q})'$.

As for the dual program (2.8), its solution remains totally unchanged in the reparametrization: substituting $\check{\mathbf{Y}}_n^*$ for $\mathbf{Y}_n^*$ and $\mathbf{Q}_0$ for $\mathbf{Q}$ in (2.8) leaves the autoregression rank scores unaffected, since the residuals $\boldsymbol{\varepsilon}_n$ themselves are not affected. The advantage of this reparametrization is that the matrix $\mathbf{Q}'$ now can be assumed to be of the form $(\mathbf{I}_q \vdots \mathbf{0})$, a form that will be easier to handle in the proofs.

However, a major difference between the present AR case and the classical regression setting is that $\boldsymbol{\beta}$ in general cannot be considered as an autoregression parameter anymore. The $\check{\mathbf{y}}_t$ vectors indeed no longer consist of lagged values of the observations and, substituting $\mathbf{A}^{-1\prime}(X_t, \dots, X_{t-p})'$ for X_t would introduce a moving average (MA) component into the model.

Note that, in case the hypothesis to be tested is of the slightly more general form $\mathbf{Q}^{\perp\prime}\boldsymbol{\theta} = \mathbf{q}_0$, for some constant p-dimensional vector $\mathbf{q}_0 \notin \mathscr{L}(\mathbf{Q})$, the substitution of $\mathbf{X}_n - \mathbf{Y}_n^*\mathbf{q}_0$ for $\mathbf{X}_n$ (with unchanged $\mathbf{Y}_n^*$ and $\mathbf{Q}$) brings us back to the former situation.

3. Linear autoregression rank score statistics.

3.1. *Definitions and distributional assumptions.* Let $H_f^n(\boldsymbol{\theta})$ stand for the (simple) hypothesis under which the observed series is a realization of length $(n+p)$ of a solution of (2.1), with innovation density f. The null hypothesis under which $\boldsymbol{\theta}$ belongs to the intersection $\mathscr{L}(\mathbf{Q})$ between $\boldsymbol{\Theta}$ and the q-dimensional linear subspace spanned by the columns of the $(p \times q)$ matrix $\mathbf{Q}$ will be denoted as

$$H^n(\mathbf{Q}) := \bigcup_f H_f^n(\mathbf{Q}) := \bigcup_{f \in \mathscr{F}} \bigcup_{\boldsymbol{\theta} \in \mathscr{L}(\mathbf{Q})} H_f^n(\boldsymbol{\theta}). \tag{3.1}$$

Under $H^n(\mathbf{Q})$, the innovation density f remains unspecified within the family $\mathscr{F}$ of densities satisfying (2.2) and the following conditions.

(F1a) f is positive and absolutely continuous with a.e. derivative f' and finite Fisher information for location $\mathscr{I}(f) := \int (f'(x)/f(x))^2 f(x)\,dx < \infty$; moreover, (F1b) There exists $K = K_f \geq 0$ such that, for all $|x| > K$, f has two bounded derivatives, f' and f'', respectively, and

(F2) f is monotonically decreasing to 0 as $x \to \pm\infty$ and, for some $b = b_f > 0$, $r = r_f \geq 1$,

$$\lim_{x \to -\infty} \frac{-\log F(x)}{b|x|^r} = 1 = \lim_{x \to \infty} \frac{-\log(1 - F(x))}{b|x|^r}.$$

Assumption (F2) implies that innovations ε_t with densities in $\mathscr{F}$ are exponentially tailed, with exponent r and coefficient b. This and the causality (2.3) of the AR(p) model (2.1) in turn implies that X_t, hence $\mathbf{y}_t$ are also exponentially tailed, with the same exponent r as ε_t; see Lemma A.2 for a precise statement. It follows that

$$\text{(3.2)} \qquad \mathrm{E}\Big[|\varepsilon_0|^k\Big] < \infty \quad \text{and} \quad \mathrm{E}\Big[\|\mathbf{y}_0\|^k\Big] < \infty, \qquad k \in \mathbb{N}.$$

Note that (F1) does not require that f be twice differentiable on $\mathbb{R}$ [only the tails are involved in (F1b)] and is satisfied by such nondifferentiable densities as the double exponential. Assumption (F1a) implies the quadratic mean differentiability of $f^{1/2}$, and the local asymptotic normality (LAN) of the model [Swensen (1985), Kreiss (1987), see also Koul and Schick (1997)].

Choose a nondecreasing, square integrable score generating function φ: $(0,1) \to \mathbb{R}$, such that $\varphi(1-u) = -\varphi(u)$, $0 < u < 1$, and let $K_\varphi^2 := \int_0^1 \varphi^2(u)\,du$. The *scores* generated by φ are defined by $\hat{\mathbf{b}}_n := (\hat{b}_{n;1}, \dots, \hat{b}_{n;n})'$, with (dropping the subscript n when no confusion is possible)

$$\text{(3.3)} \qquad \hat{b}_t := -\int_0^1 \varphi(u)\, d\hat{a}_t(u), \qquad t = 1, \dots, n.$$

Note that this integral, for given n and t, is finite, since φ is square-integrable, and $\hat{a}$ is piecewise linear, continuous and bounded on $[0,1]$, with a finite number of angular points. Efficient algorithms for the computation of such scores can be found in Koenker and d'Orey (1987, 1994), and Osborne (1992).

The tests we are proposing will be based on *linear autoregression rank score statistics* of the form [similar to Koul and Saleh's (1995) V_{ng}]

$$\text{(3.4)} \qquad \mathbf{S}_{\varphi;n} := n^{-1/2} \sum_{t=1}^{n} \mathbf{y}_{t-1}^* \hat{b}_{n;t} = n^{-1/2} \mathbf{Y}_n^{*\prime} \hat{\mathbf{b}}_n,$$

the asymptotic behavior [under $H^n(\mathbf{Q})$] of which we now investigate.

3.2. *Main results.* The main results of this section consist of three theorems. The first one (Theorem 3.1) provides an asymptotic representation, uniform over $\alpha_n \leq \alpha \leq 1 - \alpha_n$, where

$$\alpha_n := n^{-1}(\log n)^2(\log\log n)^2,$$

of the autoregression quantiles $\hat{\boldsymbol{\rho}}_{\mathbf{Q};n}(\alpha)$, along with the corresponding rate of consistency. This result constitutes a nontrivial extension of Lemma 2.2 in Koul and Saleh (1995), which only applies to bounded intervals of the form

$\varepsilon \le \alpha \le 1 - \varepsilon$, $0 < \varepsilon \le 1/2$. This limitation rules out, for instance, the optimal van der Waerden tests, which uniformly dominate the traditional parametric correlogram-based techniques, whereas our Theorem 3.1 allows for the consideration of such unbounded score functions, hence for the locally asymptotically optimal van der Waerden tests described in Section 4. On the other hand, Koul and Saleh's results apply to a larger class of innovation densities, so that neither of the two results is strictly more general than the other.

For all $0 < \alpha < 1$ and $z \in \mathbb{R}$, let

$$\sigma_\alpha = \sigma_{\alpha;f} := \frac{(\alpha(1-\alpha))^{1/2}}{f(F^{-1}(\alpha))}, \qquad \varepsilon_{t;\alpha} := \varepsilon_t - F^{-1}(\alpha),$$
$$\psi_\alpha(x) := \alpha - I[z \le 0]$$

and

$$\Sigma_n^* := n^{-1} \sum_{t=1}^{n} \mathbf{y}_{t-1}^* \mathbf{y}_{t-1}^{*\prime}.$$

THEOREM 3.1. *Under $H^n(\mathbf{Q})$, the autoregression quantile $\hat{\boldsymbol{\rho}}_{\mathbf{Q};n}(\alpha) = (\hat{\rho}_0(\alpha), (\hat{\boldsymbol{\rho}}_1'(\alpha))'$ satisfies*

$$\sup_{\alpha_n \le \alpha \le 1-\alpha_n} \sigma_\alpha^{-1} \left\| \begin{pmatrix} \hat{\rho}_0(\alpha) - F^{-1}(\alpha) \\ \mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha) - \boldsymbol{\theta} \end{pmatrix} \right\| = O_{\mathrm{P}}(n^{-1/2} C_n),$$

where $C_n := C(\log\log n)^{1/2}$, $0 < C < \infty$, and admits the asymptotic representation

$$n^{1/2}\sigma_\alpha^{-1}(\hat{\rho}_0(\alpha) - F^{-1}(\alpha)) = n^{-1/2}(\alpha(1-\alpha))^{-1/2} \sum_{t=1}^{n} \psi_\alpha(\varepsilon_{t;\alpha}) + o_{\mathrm{P}}(1),$$

$$\begin{aligned} &n^{1/2}\sigma_\alpha^{-1}(\mathbf{Q}\hat{\boldsymbol{\rho}}_1(\alpha) - \boldsymbol{\theta}) \\ &\quad = n^{-1/2}(\alpha(1-\alpha))^{-1/2}(\Sigma_n^*)^{-1} \sum_{t=1}^{n} \mathbf{y}_{t-1}^* \psi_\alpha(\varepsilon_{t;\alpha}) \\ &\qquad + o_{\mathrm{P}}(1) \end{aligned}$$

uniformly in $\alpha \in [\alpha_n, 1-\alpha_n]$, as $n \to \infty$.

The second theorem of this section provides an approximation of the autoregression rank score process by an empirical process under null hypotheses of the form $H_f^n(\mathbf{Q})$; this approximation is uniform over $[0, 1]$.

Denote by

$$\bar{\mathbf{y}}_n^* := n^{-1} \sum_{t=1}^{n} \mathbf{y}_{t-1}^* = n^{-1} \mathbf{Y}_n^{*\prime} \mathbf{1}_n \tag{3.5}$$

the arithmetic mean of the $\mathbf{y}_{t-1}^*$'s, and by

$$\boldsymbol{\Pi}_{\mathbf{YQ}} := \mathbf{Y}_n^* \mathbf{Q}(\mathbf{Q}'\mathbf{Y}_n^{*\prime}\mathbf{Y}_n^*\mathbf{Q})^{-1}\mathbf{Q}'\mathbf{Y}_n^{*\prime} \tag{3.6}$$

the (random) projection matrix projecting $\mathbb{R}^n$ onto the linear space spanned by the columns of $\mathbf{Y}_n^* \mathbf{Q}$. Put

$$\hat{\mathbf{Y}}_n^* := \mathbf{\Pi}_{\mathbf{YQ}} \mathbf{Y}_n^* := \left(\hat{\mathbf{y}}_0^*, \dots, \hat{\mathbf{y}}_{n-1}^*\right)', \qquad \bar{\mathbf{y}}_n^0 := n^{-1} \sum_{t=1}^{n} \hat{\mathbf{y}}_{t-1}^* = n^{-1} \hat{\mathbf{Y}}_n^{*\prime} \mathbf{1}_n,$$
$$\mathbf{y}_t^{\perp} := \mathbf{y}_t^* - \hat{\mathbf{y}}_t^* - \left(\bar{\mathbf{y}}_n^* - \bar{\mathbf{y}}_n^0\right) \quad \text{and} \quad \mathbf{Y}_n^{\perp} := \left(\mathbf{y}_0^{\perp}, \dots, \mathbf{y}_{n-1}^{\perp}\right)'. \tag{3.7}$$

THEOREM 3.2. *Let* $\tilde{a}_t(\alpha) := I[\varepsilon_t > F^{-1}(\alpha)]$. *Then, under* $H^n(\mathbf{Q})$,

$$\sup_{0 \le \alpha \le 1} \left\| \sum_{t=1}^{n} \left[(\mathbf{y}_{t-1}^* - \bar{\mathbf{y}}_n^*) \hat{a}_t(\alpha) - \mathbf{y}_t^{\perp} \tilde{a}_t(\alpha) \right] \right\| = o_{\mathrm{P}}(n^{1/2}) \tag{3.8}$$

as $n \to \infty$.

The proofs of Theorems 3.1 and 3.2 mainly rely on a crucial uniform quadratic approximation (Lemma 3.1) of the criterion function to be minimized in (2.6).

LEMMA 3.1. *For all* $\mathbf{z} \in \mathbb{R}^{p+1}$, *define*

$$r_n(\mathbf{z}, \alpha) := \sigma_\alpha^{-1} \sum_{t=1}^{n} \left[h_\alpha\left(\varepsilon_{t;\alpha} - n^{-1/2} \sigma_\alpha \mathbf{z}' \mathbf{y}_{t-1}\right) - h_\alpha(\varepsilon_{t;\alpha}) \right]$$
$$+ n^{-1/2} \mathbf{z}' \sum_{t=1}^{n} \mathbf{y}_{t-1} \psi_\alpha(\varepsilon_{t;\alpha}) - \tfrac{1}{2}\left(\alpha(1-\alpha)\right)^{1/2} \mathbf{z}' \mathbf{\Sigma}_n \mathbf{z}, \qquad \mathbf{z} \in \mathbb{R}^{p+1}, \tag{3.9}$$

where $\mathbf{\Sigma}_n := n^{-1} \sum_{t=1}^{n} \mathbf{y}_{t-1} \mathbf{y}_{t-1}'$. *Then, under any* $H_f^n(\boldsymbol{\theta})$ ($\boldsymbol{\theta} \in \boldsymbol{\Theta}$, $f \in \mathscr{F}$), *as* $n \to \infty$,

$$\sup\left\{ |r_n(\mathbf{z}, \alpha)| \,\middle|\, \alpha_n \le \alpha \le 1 - \alpha_n, \|\mathbf{z}\| \le C_n \right\} = o_{\mathrm{P}}(n^{-1/4}). \tag{3.10}$$

For the proof, see the Appendix, Section A.2.

PROOF OF THEOREM 3.1. Theorem 3.1 follows from Lemma 3.1, via the linear reparametrization described in Section 2.2 and a convexity argument [see Heiler and Willers (1988) or Pollard (1991)], along the same steps as in the proof of Theorem 3.1 in Gutenbrunner, Jurečková, Koenker and Portnoy (1993); we thus omit the details. □

The proof of Theorem 3.2 still requires another lemma.

LEMMA 3.2. *Under* $H^n(\mathbf{Q})$, *as* $n \to \infty$, *we have*

$$\sup_{\substack{\alpha_n \le \alpha \le 1 - \alpha_n \\ \|\mathbf{z}\| \le C_n}} n^{-1/2} \left\| \sum_{t=1}^{n} \mathbf{y}_{t-1}^{\perp} \left[\psi_\alpha\left(\varepsilon_{t;\alpha} - n^{-1/2} \sigma_\alpha \mathbf{z}' \mathbf{y}_{t-1}\right) - \psi_\alpha(\varepsilon_{t;\alpha}) \right] \right\|$$
$$= o_{\mathrm{P}}\left(n^{-1/4} C_n^{1/2}\right), \tag{3.11}$$

where the $\sup_{\mathbf{z}}$ *is taken over all vectors of the form* $\mathbf{z} = (z_0, \mathbf{z}_1^{*\prime})' \in \mathbb{R} \times \mathscr{L}(\mathbf{Q})$.

For the proof of the lemma and Theorem 3.2, see the Appendix, Sections A.3 and A.4.

We now turn back to the behavior of the linear *auto*regression rank score statistic $\mathbf{S}_{\varphi;n}$ defined in (3.4). The third theorem in this section provides an asymptotic representation of $\mathbf{S}_{\varphi;n}$ under the hypothesis $H^n(\mathbf{Q})$.

THEOREM 3.3. *Let φ: $(0,1) \mapsto \mathbb{R}$ be a nondecreasing square-integrable score function such that $\varphi'(u)$ exists in some domain of the form $(0, \alpha_0) \cup (1-\alpha_0, 1)$, $0 < \alpha_0 \le 1/2$. Assume furthermore that φ' in this domain satisfies the Chernoff–Savage condition*

$$|\varphi'(u)| \le c(u(1-u))^{-1-\delta}, \qquad 0 < \delta < \tfrac{1}{4}. \tag{3.12}$$

Then, under $H_f^n(\mathbf{Q})$, with $f \in \mathscr{F}$, the autoregression rank score statistic $\mathbf{S}_{\varphi;n}$ defined in (3.4) *admits the asymptotic representation*

$$\mathbf{S}_{\varphi;n} = n^{-1/2} \sum_{t=1}^{n} \mathbf{y}_{t-1}^{\perp} \varphi(F(\varepsilon_t)) + o_{\mathrm{P}}(1)$$

as $n \to \infty$. Moreover,

$$n^{1/2} K_\varphi^{-1} \left[\mathbf{Y}_n^{*\prime} [\mathbf{I}_n - \mathbf{\Pi}_{\mathbf{YQ}}] \mathbf{Y}_n^{*} - (\bar{\mathbf{y}}_n - \bar{\mathbf{y}}_n^0)(\bar{\mathbf{y}}_n - \bar{\mathbf{y}}_n^0)' \right]^{-1/2} \mathbf{S}_{\varphi;n} \xrightarrow{\mathscr{D}} \mathscr{N}(\mathbf{0}, \mathbf{I}_p),$$

as $n \to \infty$, where $K_\varphi^2 := \int_0^1 \varphi^2(u)\,du$, and $\mathbf{M}^{1/2}$ denotes any symmetric square root of a positive definite matrix $\mathbf{M}$.

For the proof, see the Appendix, Section A.4.

4. Optimal tests based on autoregression rank scores.

4.1. *Testing linear restrictions on AR(p) parameters.* Consider the problem of testing the null hypothesis $H^n(\mathbf{Q})$ under which $\boldsymbol{\theta}$ belongs to the intersection between $\boldsymbol{\Theta}$ and the linear space $\mathscr{L}(\mathbf{Q})$ spanned by the columns of $\mathbf{Q}$ against the alternative

$$\bigcup_f \bigcup_{\boldsymbol{\theta} \notin \mathscr{L}(\mathbf{Q})} H_f^n(\boldsymbol{\theta}).$$

We propose the simple test statistic

$$T_{\varphi;n} = T_{\varphi;n}(\mathbf{Q}) := n K_\varphi^{-2} \mathbf{S}_{\varphi;n}' (\mathbf{Y}_n^{*\prime} \mathbf{Y}_n^{*})^{-1} \mathbf{S}_{\varphi;n}, \tag{4.1}$$

where $\mathbf{S}_{\varphi;n}$ is the linear regression rank score statistic defined in (3.4).

The following theorem shows that the tests based on statistics of the type (4.1) are asymptotically distribution free, and that, for an adequate choice of φ, they are asymptotically equivalent to the *ranked residuals* tests of Kreiss (1990) [see also Hallin and Werker (1998)]. Contrary to the latter, however, they do not require any preliminary estimation of $\boldsymbol{\theta}$. Moreover, the class of tests based on (4.1) contains a locally asymptotically optimal element against any alternative of the form $\bigcup_{\boldsymbol{\theta} \notin \mathscr{L}(\mathbf{Q})} H_g^n(\boldsymbol{\theta})$ associated with any symmetric scale family of densities $g \in \mathscr{F}$. For local asymptotic optimality, we refer to Strasser (1985), Le Cam (1986), Kreiss (1987) or Hallin and Puri (1994).

Note that the symmetry assumption made on g in part (iii) of the theorem is motivated by the skew-symmetry assumption on φ. In practice, all innovation densities under which optimality properties are expected (Gaussian, logistic, double exponential, ...) are symmetric. On the other hand, we insist that this assumption of symmetry is never made on the actual innovation density f, which remains entirely unspecified within $\mathscr{F}$.

THEOREM 4.1. *Let $T_{\varphi;n}(\mathbf{Q})$ be the autoregression rank score statistic defined in* (4.1), *with the score function φ satisfying* (3.11). *Then*:

(i) *Under $H^n(\mathbf{Q})$, as $n \to \infty$, $T_{\varphi;n}(\mathbf{Q})$ is asymptotically χ^2, with $p-q$ degrees of freedom.*

(ii) *Under $H_f^n(\boldsymbol{\theta} + n^{-1/2}\mathbf{h})$, $f \in \mathscr{F}$, $\boldsymbol{\theta} \in \mathscr{L}(\mathbf{Q})$ and $\mathbf{h} \notin \mathscr{L}(\mathbf{Q})$, $T_{\varphi;n}(\mathbf{Q})$ is asymptotically noncentral χ^2, with $p-q$ degrees of freedom and noncentrality parameter*

$$c_f(\mathbf{h}, \boldsymbol{\theta}) = K_\varphi^{-2} C(\mathbf{h}, \boldsymbol{\theta}) \left(\int_{-\infty}^{\infty} \varphi(F(x)) f'(x)\, dx \right)^2,$$

where the constant $C(\mathbf{h}, \boldsymbol{\theta})$ depends only on $\mathbf{h}$ and $\boldsymbol{\theta}$, not on φ and f [see Kreiss (1990) *for an explicit form*].

(iii) *Let $g \in \mathscr{F}$ be symmetric with respect to* 0; *denote by G_1 the distribution function associated with the standardized version g_1 of g, and assume that φ: $u \mapsto -g_1'(G_1^{-1}(u))/g_1(G^{-1}(u))$, $0 < u < 1$ satisfies* (3.12). *Then, the test rejecting $H^n(\mathbf{Q})$ whenever*

$$T_{\varphi;n}(\mathbf{Q}) \geq \chi^2_{p-q;1-\alpha}, \qquad 0 < \alpha < 1,$$

is locally asymptotically maximin for $H^n(\mathbf{Q})$ against $\bigcup_{\boldsymbol{\theta} \notin \mathscr{L}(\mathbf{Q})} H_g^n(\boldsymbol{\theta})$, at asymptotic level α; $\chi^2_{p-q;1-\alpha}$ denotes the $(1-\alpha)$-quantile of the central χ^2 distribution with $p-q$ degrees of freedom.

PROOF. Assumptions (2.2), (2.3) and (F1a) guarantee the LAN structure of the problem for innovation densities $f \in \mathscr{F}$ [Kreiss (1987)]. Letting $\boldsymbol{\Pi}_{\mathbf{Y}^*} := \mathbf{Y}_{\mathrm{n}}^*(\mathbf{Y}_n^{*\prime}\mathbf{Y}_n^*)^{-1}\mathbf{Y}_n^{*\prime}$, rewrite $T_{\varphi;n}$ as

$$\begin{aligned} T_{\varphi;n} &= nK_\varphi^{-2} \left\| \mathbf{Y}_n^* (\mathbf{Y}_n^{*\prime}\mathbf{Y}_n^*)^{-1} \mathbf{S}_{\varphi;n} \right\|^2 = K_\varphi^{-2} \left\| \mathbf{Y}_n^* (\mathbf{Y}_n^{*\prime}\mathbf{Y}_n^*)^{-1} \mathbf{Y}_n^{*\prime} \left(\hat{b}_{n;1}, \dots, \hat{b}_{n;n} \right)' \right\|^2 \\ &= K_\varphi^{-2} \left\| \boldsymbol{\Pi}_{\mathbf{Y}^*} \left(\hat{b}_{n;1}, \dots, \hat{b}_{n;n} \right)' \right\|^2. \end{aligned}$$

It follows from Theorem 3.3 that $T_{\varphi;n}$ under $H_f^n(\mathbf{Q})$ is asymptotically equivalent to

$$K_\varphi^{-2} \left\| \boldsymbol{\Pi}_{\mathbf{Y}^*}[\mathbf{I}_n - \boldsymbol{\Pi}_{\mathbf{YQ}}]\left[\mathbf{I}_n - n^{-1}\mathbf{1}_n\mathbf{1}_n'\right]\left(\varphi(F(\varepsilon_1)), \dots, \varphi(F(\varepsilon_n))\right)' \right\|^2 \tag{4.2}$$

$$= K_\varphi^{-2} \left\| \boldsymbol{\Pi}_{\mathbf{Y}^*}[\mathbf{I}_n - \boldsymbol{\Pi}_{\mathbf{YQ}}]\left(\varphi(F(\varepsilon_1)), \dots, \varphi(F(\varepsilon_n))\right)' \right\|^2 + o_{\mathrm{P}}(1), \tag{4.3}$$

where $\Pi_{\mathbf{YQ}}$ is given in (3.6). Part (i) of the theorem readily follows from (4.2) and theorem 3.3. Lemma 4.2 of Kreiss (1990) in turn implies that (4.3), still under $H_f^n(\mathbf{Q})$, is asymptotically equivalent to

$$K_\varphi^{-2}\left\|\Pi_{\mathbf{Y}^*}[\mathbf{I}_n - \Pi_{\mathbf{YQ}}]\left(\varphi\left(\frac{R_{n;1}}{n+1}\right),\dots,\varphi\left(\frac{R_{n;n}}{n+1}\right)\right)'\right\|^2, \tag{4.4}$$

where $R_{n;t}$ denotes the rank of the estimated residual $\hat{\varepsilon}_t = X_t - \sum_{i=1}^p \hat{\boldsymbol{\theta}}_{n;i} X_{t-i}$ associated with a $\sqrt{n}$-consistent [under $H_f^n(\mathbf{Q})$], asymptotically discrete sequence $\hat{\boldsymbol{\theta}}_n$ of estimates of $\boldsymbol{\theta}$. Now, the statistic in (4.4) is precisely the statistic on which the Kreiss *ranked residuals* test is based: hence, our tests based on $T_{\varphi;n}$ enjoy the same asymptotic properties. Part (ii) of the theorem thus follows, as well as [part (iii)] the local asymptotic minimaxity under innovation density g, for score generating functions of the form

$$\varphi(u) = -\frac{g'(G^{-1}(u))}{g(G^{-1}(u))}.$$

Note that φ actually can be taken as $\varphi(u) = -g_1'(G_1^{-1}(u))/g_1(G_1^{-1}(u)) = -\sigma g_\sigma'(G_\sigma^{-1}(u))/g_\sigma(G_\sigma^{-1}(u))$; the influence of σ indeed is annihilated, in this case, through the standardizing factor $K_\varphi^2 = \mathscr{I}(g_1) = \sigma^2\mathscr{I}(g_\sigma)$, and thus can be safely ignored. □

4.2. *Testing* AR($p-1$) *against* AR(p) *dependence*. Due to its special role in the identification of the order of an AR model, the problem of testing AR($p-1$) against AR(p) dependence is an important special case of the problem treated in the previous section. The null hypothesis here includes all AR($p-1$) models [$\boldsymbol{\theta}$ of the form $(\theta_1,\dots,\theta_{p-1},0)'$] with unspecified innovation density $f \in \mathscr{F}$ and thus corresponds to the matrix $\mathbf{Q} = \binom{\mathbf{I}_{p-1}}{0\cdots 0}$, with rank $p-1$. The test can be based on the simple one-dimensional autoregression rank score statistic

$$S_{\varphi;n} = n^{-1/2}\sum_{t=1}^n X_{t-p}\hat{b}_{n;t};$$

the scores $\hat{b}_{n;t}$ defined in (3.3), are computed from the autoregression rank scores $\hat{\mathbf{a}}_n$ resulting from the linear program

$$\begin{aligned}
&\mathbf{X}'\mathbf{a} := \max,\\
&\sum_{t=1}^n a_t = n(1-\alpha),\\
&\sum_{t=1}^n X_{t-i}a_t = (1-\alpha)\sum_{t=1}^n X_{t-i}, \qquad i = 1,\dots,p-1,\\
&\mathbf{a} \in [0,1]^n, \qquad 0 \le \alpha \le 1.
\end{aligned} \tag{4.5}$$

Writing $\mathbf{Y}_n^* = (\mathbf{Y}_{n;I} \vdots \mathbf{Y}_{n;II})$, with $\mathbf{Y}_{n;I}$ of order $(n \times p-1)$ and $\mathbf{Y}_{n;II} := (X_{-p+1},\dots,X_{n-p})'$ of order $(n\times 1)$, let

$$t_{\varphi;n} := n^{1/2}K_\varphi^{-1}D_{\mathbf{Y}}^{-1}S_{\varphi;n},$$

where

$$D_{\mathbf{Y}}^2 := \mathbf{Y}_{n;II}' \left[\mathbf{I}_n - \mathbf{Y}_{n;I}(\mathbf{Y}_{n;I}'\mathbf{Y}_{n;I})^{-1}\mathbf{Y}_{n;I}'\right]\mathbf{Y}_{n;II}.$$

THEOREM 4.2. *Let the score function φ satisfy the assumptions of Theorem 3.3. Denote, as usual, by Φ and $z_\alpha = \Phi^{-1}(1-\alpha)$ the standard normal distribution function and quantiles. Then the test rejecting the null hypothesis of AR$(p-1)$ dependence (with unspecified innovation density in $\mathscr{F}$) whenever $|t_{\varphi;n}| > z_{\alpha/2}$:*

(i) *has asymptotic level α, and*
(ii) *has asymptotic power*

$$1 - F_{\chi_1^2}\left(z_{\alpha/2}^2; \delta_{\varphi;f}^2(\tau)\right) \tag{4.6}$$

against AR(p) alternatives with parameter value $(\theta_1, \dots, \theta_{p-1}, n^{-1/2}\tau)'$ and innovation density $f \in \mathscr{F}$, where $F_{\chi_1^2}(\cdot; \delta^2)$ stands for the noncentral chi-square distribution function with one degree of freedom and noncentrality parameter δ^2, and

$$\delta_{\varphi;f}^2(\tau) = K_\varphi^{-2}\left[\int_0^1 \varphi(u)\phi_f\left(F^{-1}(u)\right)du\right]^2 \tau^2 \quad \textit{with } \phi = -f'/f. \tag{4.7}$$

PROOF. The projection matrix $\mathbf{\Pi}_{\mathbf{YQ}}$ here takes the form $\mathbf{\Pi}_{\mathbf{Y}_I} = \mathbf{Y}_I(\mathbf{Y}_I'\mathbf{Y}_I)^{-1}\mathbf{Y}_I'$, and $[\mathbf{I} - \mathbf{\Pi}_{\mathbf{Y}_I}]\mathbf{Y}^{*\prime} = (\mathbf{0} \vdots \mathbf{Y}_{II})'$. Hence, the test statistic $T_{\varphi;n}(\mathbf{Q})$ given in (4.1), which, from Theorem 3.3, is asymptotically equivalent to $K_\varphi^{-2}\hat{\mathbf{b}}_n'[\mathbf{I} - \mathbf{\Pi}_{\mathbf{Y}_I}]\mathbf{Y}^*(\mathbf{Y}^{*\prime}\mathbf{Y}^*)^{-1}\mathbf{Y}^{*\prime}[\mathbf{I} - \mathbf{\Pi}_{\mathbf{Y}_I}]\hat{\mathbf{b}}_n$, is also asymptotically equivalent to $K_\varphi^{-2}\hat{\mathbf{b}}_n'(\mathbf{0}\mathbf{Y}_{II})(\mathbf{Y}^{*\prime}\mathbf{Y}^*)^{-1}(\mathbf{0} \vdots \mathbf{Y}_{II})'\hat{\mathbf{b}}_n$. The form of the test then follows from Theorem 4.1 by writing

$$(\mathbf{Y}^{*\prime}\mathbf{Y}^*)^{-1} = \begin{pmatrix} \mathbf{Y}_I'\mathbf{Y}_I & \mathbf{Y}_I'\mathbf{Y}_{II} \\ \mathbf{Y}_{II}'\mathbf{Y}_I & \mathbf{Y}_{II}'\mathbf{Y}_{II} \end{pmatrix}^{-1}$$

under partitioned form $\left(\begin{smallmatrix} \mathbf{A} & \mathbf{a} \\ \mathbf{a}' & c \end{smallmatrix}\right)$, with $\mathbf{A}$ of order $(p \times p)$ and

$$c = \left[\mathbf{Y}_{n;II}'\mathbf{Y}_{n;II} - \mathbf{Y}_{n;II}'\mathbf{Y}_{n;I}'(\mathbf{Y}_{n;I}'\mathbf{Y}_{n;I})^{-1}\mathbf{Y}_{n;I}'\mathbf{Y}_{n;II}\right]^{-1} = (D_{\mathbf{Y}})^{-2}.$$

The local asymptotic power (4.6) follows [after some routine algebra for the exact form of the noncentrality parameter (4.7)] either from Kreiss (1990), Theorem 4.1 or from Hallin and Puri (1994), Proposition 6.1.

Hallin, Jurečková, Kalvová, Picek and Zahaf (1997) provide a simulation study of the tests described in Theorem 4.2. The same tests are used there and also in Kalvová, Jurečková, Nemešová and Picek (1998), in the analysis of meteorological data.

4.3. *Comparison with existing procedures*: *AREs*. In this section, we compute the asymptotic relative efficiencies (AREs) of the testing procedures described in Theorems 3.1 and 3.2 [based on $T_{\varphi;n}(\mathbf{Q})$ and $t_{\varphi;n}$, respectively], with respect to a variety of existing methods. All AREs are derived under innovation densities f belonging to $\mathscr{F}$.

1. ARE with respect to Kreiss's (1990) test (4.5) [*general score test* based on the score function Ψ]:

$$ARE_1 = \frac{K_\Psi^2}{K_\varphi^2}\left[\frac{\int_0^1 \varphi(u)\,\phi_{f_1}\bigl(F_1^{-1}(u)\bigr)\,du}{\int_0^1 \Psi(u)\,\phi_f\bigl(F^{-1}(u)\bigr)\,du}\right]^2.$$

2. ARE with respect to Kreiss's (1990) test (4.4) [*ranked residual test* based on the score function φ]:

$$ARE_2 = 1.$$

3. ARE with respect to Hallin and Puri's (1994) test (3.3) [*aligned rank test* based on the rank autocorrelation coefficients associated with innovation density g, where g is assumed to satisfy the assumptions of part (iii) of Theorem 4.1; the notation g_1, G_1, $\phi_{g_1}, \dots$ is used in an obvious fashion]:

$$ARE_3 = \frac{\int_0^1 \phi_{g_1}^2\bigl(G_1^{-1}(u)\bigr)\,du}{K_\varphi^2} \times \left[\frac{\int_0^1 \varphi(u)\,\phi_{f_1}\bigl(F_1^{-1}(u)\bigr)\,du}{\int_0^1 F_1^{-1}(u)G_1^{-1}(u)\,du \int_0^1 \phi_{f_1}\bigl(F_1^{-1}(u)\bigr)\phi_{g_1}\bigl(G_1^{-1}(u)\bigr)\,du}\right]^2$$

which, on account of the Cauchy–Schwarz inequality, reduces to

$$ARE_3' = \left[\int_0^1 F_1^{-1}(u)G_1^{-1}(u)\,du\right]^{-2} \geq 1,$$

for $\varphi(u) = \phi_{g_1}(G_1^{-1}(u))$, with equality at $F_1 = G_1$ only.

4. ARE with respect to traditional (i.e., Gaussian) Lagrange multiplier tests [cf., for example, Godfrey (1979), Hosking (1980), Pötscher (1983)]:

$$ARE_4 = \frac{1}{K_\varphi^2}\left[\int_0^1 \varphi(u)\,\phi_{f_1}\bigl(F_1^{-1}(u)\bigr)\,du\right]^2,$$

a value which, for $\varphi(u) = \Phi^{-1}(u)$ (the standard normal quantile function), is uniformly larger than one, with equality at $F_1 = \Phi$ only [Chernoff and Savage (1958)].

These ARE values again follow easily from inspecting the noncentrality parameters associated with the various noncentral chi-square distributions under local alternatives, as given in Kreiss (1990) or Hallin and Puri (1988, 1994). All these noncentrality parameters indeed only differ by multiplicative constants, the ratios of which yield the desired results.

Some numerical values for ARE_4 are provided in Hallin and Werker (1998).

APPENDIX

A.1. Some properties of exponentially tailed densities. Before turning to the proof of Lemma 3.1, let us summarize some of the consequences of assumptions (F1) and (F2).

LEMMA A.1. *Let $f \in \mathcal{F}$. Then*

$$\text{(i)} \quad \lim_{u \to 0} \frac{f(F^{-1}(u))}{u(-\log u)^{1-1/r}} = \lim_{u \to 1} \frac{f(F^{-1}(u))}{(1-u)(-\log(1-u))^{1-1/r}} = rb^{1/r};$$

$$\text{(ii)} \quad \lim_{x \to -\infty} \left(-\frac{f'(x)}{f(x)} |x|^{1-r} \right) = \lim_{x \to \infty} \left(-\frac{f'(x)}{f(x)} x^{1-r} \right) = rb;$$

$$\text{(iii)} \quad \lim_{u \to 0} \frac{-F^{-1}(u)}{(-\log u)^{1/r}} = \lim_{u \to 1} \frac{F^{-1}(u)}{(-\log(1-u))^{1/r}} = (1/b)^{1/r}.$$

The proof follows straightforwardly from (F1a, b), (F2) and l'Hospital rule.

LEMMA A.2. *Denote by $F_{\boldsymbol{\theta}}^X$ the marginal distribution of the stationary solution of* (2.1), *where $\boldsymbol{\theta}$ and the innovation density f satisfy the causality assumption and* ($F2$), *respectively. Then, there exists $b^* > 0$ such that*

$$\text{(A.1)} \quad \liminf_{x \to -\infty} \frac{-\log F_{\boldsymbol{\theta}}^X(x)}{b^*|x|^r} \geq 1 \quad \textit{and} \quad \liminf_{x \to \infty} \frac{-\log(1 - F_{\boldsymbol{\theta}}^X(x))}{b^* x^r} \geq 1.$$

PROOF. The stationary solution of (2.1) can be written as $X_t = \sum_{k=0}^{\infty} c_k \varepsilon_{t-k}$, with $|c_k| < A_\theta \lambda^k$ for some positive constant $A_{\boldsymbol{\theta}}$ and $0 < \lambda < 1$. Denote by μ_k, $k = 0, 1, \ldots$ an arbitrary sequence of positive weights such that $\sum_{k=0}^{\infty} \mu_k = 1$. Then, for all $x > 0$,

$$\begin{aligned} \mathrm{P}\left[\sum_{k=0}^{\infty} c_k \varepsilon_{t-k} > x \right] &\leq \mathrm{P}\left[\bigcup_{k=0}^{\infty} \{ c_k \varepsilon_{t-k} > \mu_k x \} \right] \\ &\leq \sum_{k=0}^{\infty} \left\{ \max\left[\left(1 - F\left(\frac{\mu_k}{|c_k|} x \right) \right), F\left(-\frac{\mu_k}{|c_k|} x \right) \right] \right\}. \end{aligned}$$

Let $\mu_k^* := B^{1/r}(k+1)^{1/r}|c_k|$, where B is such that $\sum_{k=0}^{\infty} \mu_k^* = 1$; such a value of B exists, since

$$0 < \sum_{k=0}^{\infty} (k+1)^{1/r} |c_k| < A_{\boldsymbol{\theta}} \sum_{k=0}^{\infty} (k+1)^{1/r} \lambda^k < A_{\boldsymbol{\theta}} \frac{\lambda}{1-\lambda} < \infty.$$

In view of (F2), for all $\eta > 0$, there exists $M_1(\eta)$ such that, for $x > M_1$,

$$\begin{aligned} \max\left[\left(1 - F\left(\frac{\mu_k^*}{|c_k|} x \right) \right), F\left(-\frac{\mu_k^*}{|c_k|} x \right) \right] &\leq \exp\left\{ -(1-\eta) b \left(\frac{\mu_k^*}{|c_k|} x \right)^r \right\} \\ &= \left(\exp\{ -(1-\eta) b B x^r \} \right)^{k+1}, \\ &\qquad k = 0, 1, \ldots . \end{aligned}$$

Hence

$$\mathrm{P}\left[\sum_{k=0}^{\infty} c_k \varepsilon_{t-k} > x\right] \le \sum_{k=0}^{\infty} \left(\exp\{-b(1-\eta)Bx^r\}\right)^{k+1}$$
$$= \frac{\exp\{-b(1-\eta)Bx^r\}}{1-\exp\{-b(1-\eta)Bx^r\}}.$$

For x larger than some $M_2(\eta)$, $1-\exp\{-b(1-\eta)Bx^r\} > 1-\eta$, so that, for all $x > \max(M_1(\eta), M_2(\eta))$,

$$\mathrm{P}\left[\sum_{k=0}^{\infty} c_k \varepsilon_{t-k} > x\right] \le \frac{1}{1-\eta}\exp\{-b(1-\eta)Bx^r\}.$$

This holds for any $\eta > 0$; taking logarithms and changing signs, we thus obtain

$$\liminf_{x\to\infty} \frac{-\log \mathrm{P}[\sum_{k=0}^{\infty} c_k \varepsilon_{t-k} > x]}{bBx^r} \ge 1.$$

The proof is entirely similar for left tails. The lemma follows, with $b^* = bB$. □

COROLLARY A.1. *Denote by $F_{\boldsymbol{\theta}}^{\mathbf{y}}$ the distribution function of $\|\mathbf{y}_t\|$ defined in* (2.5). *Under the same assumptions as in Lemma* A.2,

$$\liminf_{y\to\infty} \frac{-\log(1-F_{\boldsymbol{\theta}}^{\mathbf{y}}(y))}{b^{**}y^r} \ge 1 \tag{A.2}$$

for some $b^{**} > 0$.

The proof follows immediately from Lemma A.2, since $\|\mathbf{y}_t\|^2$ is the sum of p successive squared values of X_t.

A.2. Proof of Lemma 3.1. Since the convergence in (3.10), uniform over $\alpha_0 \le \alpha \le 1-\alpha_0$ and $\|\mathbf{z}\| \le C$ for fixed $\alpha_0 \in (0, \frac{1}{2})$ and $C \in (0,\infty)$, follows from Koul and Saleh (1995), Lemma 2.2, we may concentrate mainly on values of α which are close to α_n or $1-\alpha_n$. For given $\mathbf{z} \in \mathbb{R}^{p+1}$, let

$$\delta_{t;\alpha} = \delta_{n;t;\alpha}(\mathbf{z}) := n^{-1/2}\sigma_\alpha \mathbf{z}'\mathbf{y}_{t-1}, \qquad t = 0,1,2,\dots. \tag{A.3}$$

By Lemma A.1,

$$\sigma_\alpha(\alpha(1-\alpha))^{1/2}(-\log(\alpha(1-\alpha)))^{1-1/r} \to (rb^{1/r})^{-1} \tag{A.4}$$

as $\alpha \to 0$ or 1; hence, for α sufficiently close to α_n,

$$|\delta_{t;\alpha}| \le 2b^{-1/r}(\log n)^{(1/r)-2}(\log\log n)^{-1} C_n \max_{1\le t\le n} \|\mathbf{y}_{t-1}\|. \tag{A.5}$$

As a stationary AR process, X_t is strongly mixing [see, e.g., Pham and Tran (1985)]. Hence, $\mathbf{y}_t$ and $\|\mathbf{y}_t\|$ also are strongly mixing [see, e.g., Theorem 5.2 in O'Brien (1987), under the terminology 1-*strongly mixing*]. It follows from Theorem 5.1 (same reference) that $\|\mathbf{y}_t\|$ has *asymptotic independence of maxima relative to* $\{r_n\}$ ($AIM(r_n)$) for any real sequence $\{r_n\}$ such that $\lim_{n\to\infty}, r_n = \infty$.

Let

$$R_n := (\log n)^{1/r}(\log\log n)^{1/4}. \tag{A.6}$$

It follows from Corollary A.1 that, for any $\eta > 0$, there exists a $N_0 = N_0(\eta)$ such that

$$\log\left(1 - \left(F_{\boldsymbol{\theta}}^{\mathbf{y}}(R_n)\right)^n\right) \le \log n - (1-\eta)b^{**}(R_n)^r$$

for all $n \ge N_0$. Hence, for n sufficiently large,

$$1 - \left(F_{\boldsymbol{\theta}}^{\mathbf{y}}(R_n)\right)^n \le n^{-(1-\eta)b^{**}(\log\log n)^{r/4}+1} < n^{-b^{**}(\log\log n)^{r/8}+1}. \tag{A.7}$$

O'Brien's (1987) Theorem 2.1 applies to the $AIM(R_n)$ sequence $\|\mathbf{y}_t\|$, where R_n, defined in (A.6) satisfies O'Brien's condition (2.3). It follows that

$$\mathrm{P}\left[\max_{1\le t\le n}\|\mathbf{y}_{t-1}\| > R_n\right] < \left(1 - \left(F_{\boldsymbol{\theta}}^{\mathbf{y}}(R_n)\right)^n\right)(1+o(1)) = o(1) \tag{A.8}$$

as $n \to \infty$. Hence,

$$\tau_n^* = \max_{1\le t\le n}\|\mathbf{y}_{t-1}\| = O_{\mathrm{P}}(R_n) = O_{\mathrm{P}}\left((\log n)^{1/r}(\log\log n)^{1/4}\right) \tag{A.9}$$

as $n \to \infty$, and, for any $K > 0$ and $n \ge N_0(K)$, $P[\tau_n^* > R_n] \le \frac{1}{2}n^{-K}$. It follows that

$$\sup\left\{\sigma_\alpha^{-1}|\,\delta_{t;\alpha}(\mathbf{z})|: 0 \le \alpha \le 1, \|\mathbf{z}\| \le C_n\right\} = O_{\mathrm{P}}\left(n^{-1/2}C_n\tau_n^*\right).$$

Let

$$Q_t(\mathbf{z},\alpha) = Q_t := \sigma_\alpha^{-1}\left\{h_\alpha(\varepsilon_{t;\alpha} - \delta_{t;\alpha}(\mathbf{z})) - h_\alpha(\varepsilon_{t;\alpha}) + \delta_{t;\alpha}(\mathbf{z})\psi_\alpha(\varepsilon_{t;\alpha})\right\}.$$

Obviously,

$$Q_t = \sigma_\alpha^{-1}\{(\varepsilon_{t;\alpha} - \delta_{t;\alpha})I[\delta_{t;\alpha} < \varepsilon_{t;\alpha} \le 0] + (\delta_{t;\alpha} - \varepsilon_{t;\alpha})I[0 < \varepsilon_{t;\alpha} < \delta_{t;\alpha}]\}.$$

Denote by $\mathscr{A}_t$ the σ-algebra generated by $\{X_{-p+1},\dots,X_0;\varepsilon_s|s \le t\}$. For each t, ε_t is independent of $\mathscr{A}_{t-1}$, and the sequence

$$\sum_{t=1}^{n}\left[Q_t - E(Q_t|\mathscr{A}_{t-1})\right] := \sum_{t=1}^{n}\zeta_{n;t} \tag{A.10}$$

forms a square-integrable martingale with respect to the filtration $\{\mathscr{A}_t\}$. For $\delta_{n;t;\alpha} = \delta > 0$, elementary algebra yields

$$\begin{aligned}
\mathrm{E}(Q_t|\mathscr{A}_{t-1}) &= \sigma_\alpha^{-1}\int_{F^{-1}(\alpha)}^{F^{-1}(\alpha)+\delta}\left(\delta - x + F^{-1}(\alpha)\right)\left(f(x) - f\left(F^{-1}(\alpha)\right)\right)dx \\
&\quad + \frac{\delta^2 f\left(F^{-1}(\alpha)\right)}{2\sigma_\alpha} \\
&= \frac{1}{2n}(\alpha(1-\alpha))^{1/2}(\mathbf{z}'\mathbf{y}_{t-1})^2 \\
&\quad + \sigma_\alpha^{-1}\int_{F^{-1}(\alpha)}^{F^{-1}(\alpha)+\delta}\left(\delta - x + F^{-1}(\alpha)\right)\int_{F^{-1}(\alpha)}^{x} f'(z)\,dz\,dx.
\end{aligned}$$

It follows from (F1a, b) and (F2) that

$$\lim_{x\to\pm\infty}\frac{|f'(x)|}{f^2(x)}F(x)(1-F(x)) = 1.$$

Hence, for $\delta_{n;t;\alpha}=\delta>0$ and α close to zero, and for all $\eta>0$, there exists n_0 such that, for all $n\ge n_0$,

$$\begin{aligned}
&\left|\mathrm{E}(Q_t|\mathscr{A}_{t-1}) - \frac{1}{2n}(\alpha(1-\alpha))^{1/2}(\mathbf{z}'\mathbf{y}_{t-1})^2\right|\\
&\qquad\le(1+\eta)\sigma_\alpha^{-1}\\
&\qquad\quad\times\left|\int_{F^{-1}(\alpha)}^{F^{-1}(\alpha)+\delta}(\delta-x+F^{-1}(\alpha))\int_\alpha^{F(x)}\frac{f(F^{-1}(u))}{u(1-u)}\,du\,dx\right|\\
&\qquad\le(1+\eta)^2\sigma_\alpha^{-1}\\
&\qquad\quad\times\left|\int_{F^{-1}(\alpha)}^{F^{-1}(\alpha)+\delta}(\delta-x+F^{-1}(\alpha))\int_\alpha^{F(x)}|F^{-1}(u)|^{r-1}\,du\,dx\right|\\
&\qquad\le K_2\sigma_\alpha^{-1}|F^{-1}(\alpha)|^{r-1}f(F^{-1}(\alpha)+\delta)|\delta|^3.
\end{aligned}\tag{A.11}$$

Similar approximations are valid for $\delta_{n;t;\alpha}<0$ and/or for α close to $1-\alpha_n$. Furthermore, by (A.4), (A.5) and (A.9) we have, for any $\eta>0$, $|\alpha-\alpha_n|$ sufficiently close to zero and n sufficiently large,

$$\max_{1\le t\le n}|\delta_{n;t;\alpha}|\le|F^{-1}(\alpha)|(\log n)^{-1}C_n\tau_n^*(1+\eta).\tag{A.12}$$

Note that, since α is close to α_n, $F^{-1}(\alpha)<0$. If $\delta_{t;\alpha}<0$, then $F^{-1}(\alpha)+\delta_{t;\alpha}<F^{-1}(\alpha)$, hence $f(F^{-1}(\alpha)+\delta_{t;\alpha})\le f((F^{-1}(\alpha))$. If $\delta_{t;\alpha}>0$, then, by (A.12),

$$F^{-1}(\alpha)+\delta_{t;\alpha}\le F^{-1}(\alpha)(1-\Delta_n\tau_n^*(1+\eta)),\tag{A.13}$$

where $\Delta_n := C_n(\log n)^{-1}$. Provided that $\tau_n^*<R_n$, (A.13) and Lemma A.1 entail

$$\begin{aligned}
f(F^{-1}(\alpha)+\delta_{t;\alpha}) &\le f\left(F^{-1}(\alpha)(1-\Delta_n\tau_n^*(1+\eta))^r\right)\\
&\le r(-\log\alpha)^{1-1/r}\alpha^{(1-\Delta_n\tau_n^*(1+\eta))^r}(1-\Delta_n\tau_n^*(1+\eta))^{r-1}\\
&\le r(-\log\alpha)^{1-1/r}\alpha^{1-r\Delta_n\tau_n^*(1+\eta)}(1-\Delta_n\tau_n^*(1+\eta))^{r-1},
\end{aligned}$$

where we used the fact that $(1-x)^r\ge 1-rx$ for $r\ge1$ and $0<x<1$.

Then, by (A.11), (A.4) and (A.5),

$$\begin{aligned}
&\left|\sum_{t=1}^{n} \mathrm{E}(Q_{n;t}|\mathscr{A}_{t-1}) - \tfrac{1}{2}(\alpha(1-\alpha))^{1/2}\mathbf{z}'\mathbf{\Sigma}_n\mathbf{z}\right| \\
&\qquad \le K_2 C_n^3 \alpha^{-r\Delta_n\tau_n^*(1-\eta)} n^{-1/2} n^{-1} \sum_{t=1}^{n} \|\mathbf{y}_{t-1}\|^3 \\
&\qquad \le K_3 C_n^3 n^{-(1/2)+r\Delta_n\tau_n^*(1-\eta)}(C + o_{\mathrm{P}}(1))
\end{aligned} \tag{A.14}$$

as $|\alpha - \alpha_n| \to 0$ and $n \to \infty$; indeed, it follows from Corollary A.1 that $\|\mathbf{y}_t\|$ has finite moments of all orders, and thus, as $n \to \infty$,

$$\mathbf{\Sigma}_n = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}^* \end{pmatrix} + o_{\mathrm{P}}(1) := \mathbf{\Sigma} + o_{\mathrm{P}}(1) \quad \text{and} \quad n^{-1}\sum_{t=1}^{n}\|\mathbf{y}_{t-1}\|^3 \le C + o_{\mathrm{P}}(1),$$

where $\mathbf{\Sigma}^*$ is the $p \times p$ autocovariance matrix in the distribution of the stationary solution of (2.1), and $C > 0$ is some constant. Putting

$$B_n := C_n^{5/2} n^{-(1/2)+2r\Delta_n R_n}(\log n)^2, \tag{A.15}$$

we have, by (A.8), for any $\lambda > 0$ and $n \ge n_0$,

$$\mathrm{P}\left\{\left|\sum_{t=1}^{n}\zeta_{n;t}\right| \ge \lambda B_n\right\} \le \mathrm{P}\left\{\left|\sum_{t=1}^{n}\zeta_{n;t}\right| \ge \lambda B_n,\, \tau_n^* \le R_n\right\} + \tfrac{1}{2}n^{-K}. \tag{A.16}$$

Moreover, (A.6) and (A.9) imply that $-a_n^2 \le \zeta_{n;t} \le a_n$ for $n \ge n_0$, with

$$a_n := C_n n^{-1/2}(\log n)^{1/r}(\log\log n)^{1/4}.$$

Hence, applying Hoeffding's (1963) inequality to the martingale

$$\sum_{t=1}^{n}\frac{\zeta_{n;t} + a_n^2}{a_n(1+a_n)}, \qquad t = 1, \dots, n$$

with bounded increments $0 \le (\zeta_{n;t} + a_n^2)/(a_n(1+a_n)) \le 1$, we obtain, for any $\lambda > 0$ and $n \ge n_0$,

$$\begin{aligned}
&\mathrm{P}\left\{\left|\sum_{t=1}^{n}\zeta_{n;t}\right| \ge \lambda B_n,\, \tau_n^* \le R_n\right\} \\
&\quad \le \mathrm{P}\left\{\left|\sum_{t=1}^{n}\left[\frac{\zeta_{n;t} + a_n^2}{a_n(1+a_n)} - \mathrm{E}\left(\frac{\zeta_{n;t} + a_n^2}{a_n(1+a_n)}\right)\right]\right| \ge \frac{\lambda B_n}{a_n(1+a_n)}\right\} \\
&\quad \le \exp\left\{-\frac{\lambda^2}{2}\frac{B_n^3}{a_n^3}\right\} \le \frac{1}{2}n^{-\lambda^2/2}.
\end{aligned} \tag{A.17}$$

Combining (A.8), (A.16) and (A.17), and taking $K > \lambda^2/2$, we obtain

$$\mathrm{P}\left\{\left|\sum_{t=1}^{n}[Q_t - \mathrm{E}(Q_t|\mathscr{A}_{t-1})]\right| \ge \lambda B_n\right\} \le n^{-\lambda^2/2} \tag{A.18}$$

for $n \geq n_0$. Finally, (A.18), along with (A.14), yields

$$
\begin{aligned}
&\mathrm{P}\left\{\left|\sum_{t=1}^{n} Q_t - \tfrac{1}{2}\big(\alpha(1-\alpha)\big)^{1/2}\mathbf{z}'\boldsymbol{\Sigma}\mathbf{z}\right| \geq (\lambda+1)B_n\right\} \\
&\quad \leq \mathrm{P}\left\{\left|\sum_{t=1}^{n} Q_t - \mathrm{E}\big(Q_t|\mathscr{A}_{t-1}\big)\right| \geq \lambda B_n\right\} \\
&\qquad + \mathrm{P}\left\{\left|\sum_{t=1}^{n} \mathrm{E}\big(Q_t|\mathscr{A}_{t-1}\big) - \tfrac{1}{2}\big(\alpha(1-\alpha)\big)^{1/2}\mathbf{z}'\boldsymbol{\Sigma}\mathbf{z}\right| \geq B_n\right\} \qquad \text{(A.19)} \\
&\quad \leq \tfrac{1}{2}n^{-\lambda^2/2} + \tfrac{1}{2}n^{-K} + \mathrm{P}\left\{\left|\sum_{t=1}^{n} \mathrm{E}\big(Q_t|\mathscr{A}_{t-1}\big) - \tfrac{1}{2}\big(\alpha(1-\alpha)\big)^{1/2}\mathbf{z}'\boldsymbol{\Sigma}\mathbf{z}\right|\right. \\
&\qquad\qquad \left. \geq B_n,\ \tau_n^* \leq R_n\right\} \\
&\quad = \tfrac{1}{2}n^{-\lambda^2/2} + \tfrac{1}{2}n^{-K} + 0 \leq n^{-\lambda^2/2}
\end{aligned}
$$

for $K > \lambda^2/2$ and $n \geq n_0$; the vanishing of the last probability term in (A.19) follows from the fact that, by (A.14),

$$
\begin{aligned}
\left|\sum_{t=1}^{n} \mathrm{E}\big(Q_t|\mathscr{A}_{t-1}\big) - \tfrac{1}{2}\big(\alpha(1-\alpha)\big)^{1/2}\mathbf{z}'\boldsymbol{\Sigma}\mathbf{z}\right| &\leq K_2 C_n^3 n^{-(1/2)+r\Delta_n R_n(1-\eta)} R_n^3 \\
&\leq B_n = C_n^{5/2} n^{-(1/2)+2r\Delta_n R_n} (\log n)^2.
\end{aligned}
$$

Now, let us choose a collection of intervals $[\iota_\nu, \iota_{\nu+1}]$, of length n^{-5}, covering $[\alpha_n, 1-\alpha_n]$ and a collection of balls of radius n^{-5} covering $\{\mathbf{z}\colon \|\mathbf{z}\| \leq C_n\}$. Let $(\alpha_1, \alpha_2) \subset (\iota_\nu, \iota_{\nu+1})$, and let $\mathbf{z}_1, \mathbf{z}_2$ lie in the same ball. Then, by (A.4),

$$
\big|\big(\sigma_{\alpha_1}/\sigma_{\alpha_2}\big) - 1\big| = O\big(n^{-4}(\log n)^{2-1/r}\big). \qquad \text{(A.20)}
$$

For fixed t, $1 \leq t \leq n$, we have

$$
\begin{aligned}
|Q_t(\mathbf{z}_2, \alpha_2) - Q_t(\mathbf{z}_1, \alpha_1)| \leq & |Q_t(\mathbf{z}_2, \alpha_2) - Q_t(\mathbf{z}_1, \alpha_2)| \\
& + |Q_t(\mathbf{z}_1, \alpha_2) - Q_t(\mathbf{z}_1, \alpha_1)|.
\end{aligned} \qquad \text{(A.21)}
$$

Consider the two terms on the right-hand side separately. Starting with the first one,

$$
|Q_t(\mathbf{z}_2, \alpha_2) - Q_t(\mathbf{z}_1, \alpha_2)| \leq n^{-1/2}|(\mathbf{z}_2 - \mathbf{z}_1)'\mathbf{y}_{t-1}| = O_{\mathrm{P}}\big(n^{-5.5}\|\mathbf{y}_{t-1}\|\big).
$$

For the corresponding centering term, we obtain the bound

$$
\tfrac{1}{2}\big(\alpha_2(1-\alpha_2)\big)^{1/2} n^{-1}\Big|\big(\mathbf{z}_2'\mathbf{y}_{t-1}\big)^2 - \big(\mathbf{z}_1'\mathbf{y}_{t-1}\big)^2\Big| = O_{\mathrm{P}}\big(n^{-6}C_n\|\mathbf{y}_{t-1}\|^2\big). \qquad \text{(A.22)}
$$

Next, consider the second term on the right-hand side of (A.21) (denoted as Q^* for the sake of brevity). We should distinguish two cases.

1. Either $\delta_{t;\,\alpha_2 z_1} < \varepsilon_{t;\,\alpha_2} \le 0$ and $\delta_{t;\,\alpha_1 z_1} < \varepsilon_{t;\,\alpha_1} \le 0$ [or $0 < \varepsilon_{t;\,\alpha_2} < \delta_{t;\,\alpha_2 z_1}$ and $0 < \varepsilon_{t;\,\alpha_1} < \delta_{t;\,\alpha_1 z_1}$]; then,

$$|Q^*| \le 2n^{-1/2}\Big|\big[1 - (\sigma_{\alpha_2}/\sigma_{\alpha_1})\big]\mathbf{z}_1'\mathbf{y}_{t-1}\Big| + \sigma_{\alpha_1}^{-1}\big|F^{-1}(\alpha_2) - F^{-1}(\alpha_1)\big|$$
$$= O_{\mathrm{P}}\big(n^{-4}C_n\|\mathbf{y}_{t-1}\|\big) + O_{\mathrm{P}}(n^{-3}).$$

2. Or $\delta_{t;\,\alpha_2 z_1} < \varepsilon_{t;\,\alpha_2} \le 0$ and $\varepsilon_{t;\,\alpha_1} < \delta_{t;\,\alpha_1 z_1} < 0$ [or $\delta_{t;\,\alpha_1 z_1} < \varepsilon_{t;\,\alpha_1} \le 0$ and $\varepsilon_{t;\,\alpha_2} < \delta_{t;\,\alpha_2 z_1} < 0$]; then,

$$|Q^*| = \sigma_{\alpha_2}^{-1}|\varepsilon_{t;\,\alpha_2} - \delta_{t;\,\alpha_2 z_1}| = \sigma_{\alpha_2}^{-1}\big|\varepsilon_{t;\,\alpha_1} + F^{-1}(\alpha_1) - F^{-1}(\alpha_2) - \delta_{t;\,\alpha_2 z_1}\big|$$

$$\text{(A.23)} \quad \le \sigma_{\alpha}^{-1}\Big(|\delta_{t;\,\alpha_1 z_1} - \delta_{t;\,\alpha_2 z_1}| + \big|F^{-1}(\alpha_2) - F^{-1}(\alpha_1)\big|\Big)$$
$$= O_{\mathrm{P}}\big(n^{-3.5}C_n\|\mathbf{y}_{t-1}\|\big) + O_{\mathrm{P}}(n^{-3}).$$

Moreover, the centering term of Q^* can be estimated as

$$\text{(A.24)} \quad \tfrac{1}{2}n^{-1}|\mathbf{z}_1'\mathbf{y}_{t-1}|^2\Big|\big(\alpha_2(1-\alpha_2)\big)^{1/2} - \big(\alpha_1(1-\alpha_1)\big)^{1/2}\Big| = O_{\mathrm{P}}\big(n^{-6}C_n^2\|\mathbf{y}_{t-1}\|^2\big).$$

Let us fix one set S_ν in the decomposition of $[\alpha_n, 1-\alpha_n] \times \{\mathbf{z}: \|\mathbf{z}\| \le C_n\}$; the number of such sets is at most $(2C_n)^{p+1}n^{5(p+1)}$. It follows from (A.20)–(A.24) that

$$\sup_{S_\nu}|r_n(\mathbf{z}_2, \alpha_2) - r_n(\mathbf{z}_1, \alpha_1)| \le K_1 n^{-3}\sum_{t=1}^{n}\|\mathbf{y}_{t-1}\| + K_2 n^{-3},$$

where K_1 and K_2 are positive constants. By (A.7)–(A.9), for n sufficiently large,

$$\text{(A.25)} \quad \mathrm{P}\Big\{\sup_{S_\nu}|r_n(\mathbf{z}_2, \alpha_2) - r_n(\mathbf{z}_1, \alpha_1)| > \lambda B_n\Big\} \le \mathrm{P}\big\{\tau_n^* \ge \big(\lambda B_n - K_2 n^{-3}\big)K_1^{-1}n^2\big\} \le \tfrac{1}{2}n^{-K}.$$

Hence, from (A.19) and (A.25), $\mathrm{P}\{\sup_{S_\nu}|r_n(\mathbf{z}, \alpha)| \ge 2(\lambda+1)B_n\} \le n^{-\lambda^2/2} + \frac{1}{2}n^{-K} \le \frac{3}{2}n^{-\lambda^2/2}$, and finally,

$$\mathrm{P}\Big\{\sup_{\|\mathbf{z}\|\le C_n,\, \alpha_n \le \alpha \le 1-\alpha_n}|r_n(\mathbf{z}, \alpha)| \ge 2(\lambda+1)n^{-(1/2)+2\kappa}(\log n)^2 C_n^{5/2}\Big\}$$
$$\le \sum_\nu \mathrm{P}\Big\{\sup_{S_\nu}|r_n(\mathbf{z}, \alpha)| \ge 2(\lambda+1)n^{-(1/2)+2\kappa}(\log n)^2 C_n^{5/2}\Big\}$$
$$\le 3C_n^{p+1}n^{5(p+1)}n^{-\lambda^2/2} + n^{-(1/2)+2\kappa}(\log n)^2 C_n^{5/2} = o(1)$$

for all $\kappa > 0$ and $\lambda^2 > 10(p+1)$. This in turn implies (3.10), and completes the proof of the lemma. □

A.3. Proof of Lemma 3.2. The proof proceeds in two steps. The particular case of a unit matrix $\mathbf{Q}$ is considered first; the general case is treated as a second step.

Step 1. Assume that $\mathbf{Q}$ is of the form $\mathbf{Q}_0' := (\mathbf{I}_q \vdots \mathbf{0})$ considered in Section 2.2. Without loss of generality, we may as well assume $q = p - 1$. Then, $\mathbf{Y}_n$ decomposes into

$$\left(\mathbf{1}_n \vdots \tilde{\mathbf{Y}}_n^* \vdots (X_{1-p}, \dots, X_{n-p})'\right) := \left(\tilde{\mathbf{Y}}_n \vdots (X_{1-p}, \dots, X_{n-p})'\right).$$

Similarly, $\mathbf{y}_{t-1}'$ decomposes into $(1, \tilde{\mathbf{y}}_{t-1}^{*\prime}, X_{t-p}) := (\tilde{\mathbf{y}}_{t-1}', X_{t-p})$. Denote by

$$\left(\hat{X}_{1-p}, \dots, \hat{X}_{n-p}\right) := \mathbf{\Pi}_{\hat{\mathbf{Y}}}(X_{1-p}, \dots, X_{n-p})',$$

with $\mathbf{\Pi}_{\tilde{\mathbf{Y}}} := \tilde{\mathbf{Y}}_n(\tilde{\mathbf{Y}}_n'\tilde{\mathbf{Y}}_n)^{-1}\tilde{\mathbf{Y}}_n'$, the projection of $\mathbf{Y}_n$'s last column onto the linear space spanned by the first p ones, and let $(X_{1-p}^{\perp}, \dots, X_{n-p}^{\perp})' := [\mathbf{I}_n - \mathbf{\Pi}_{\tilde{\mathbf{Y}}}]$ $(X_{1-p}, \dots, X_{n-p})'$. It is easily checked that, in the notation introduced in (3.7), we have

$$\mathbf{Y}_n^{\perp} = \left(\mathbf{0}_{n\times(p-1)} \vdots \left(X_{1-p}^{\perp}, \dots, X_{n-p}^{\perp}\right)'\right) \quad \text{hence } \mathbf{y}_{t-1}^{\perp} = \left(0, \dots, 0, X_{t-p}^{\perp}\right).$$

For any $\mathbf{z} := (\tilde{\mathbf{z}}', z_p)' \in \mathbb{R}^{p+1}$, $\mathbf{z}'\mathbf{y}_{t-1}$ decomposes into $\tilde{\mathbf{z}}'\tilde{\mathbf{y}}_{t-1} + z_p X_{t-p}$. Letting $A_n = C_n^{1/2} n^{-1/4}$, Lemma 3.1 yields

$$\sup_{\|\mathbf{z}\|\le C_n,\, \alpha_n\le\alpha\le 1-\alpha_n} \left\{\left| \sigma_\alpha^{-1} \sum_{t=1}^n \left[h_\alpha\left(\varepsilon_{t;\alpha} - n^{-1/2}\sigma_\alpha \mathbf{z}'\mathbf{y}_{t-1}\right) - h_\alpha(\varepsilon_{t;\alpha})\right] \right.\right.$$
$$\left.\left. + n^{-1/2}\mathbf{z}' \sum_{t=1}^n \mathbf{y}_{t-1}\psi_\alpha(\varepsilon_{t;\alpha}) - \tfrac{1}{2}(\alpha(1-\alpha))^{1/2}\mathbf{z}'\mathbf{\Sigma}_n\mathbf{z} \right|\right\} \tag{A.26}$$
$$= o_{\mathrm{P}}(A_n).$$

Hence also

$$\sup_{\|\mathbf{z}\|\le C_n,\, \alpha_n\le\alpha\le 1-\alpha_n} \left\{\left| \sigma_\alpha^{-1} \sum_{t=1}^n \left[h_\alpha\left(\varepsilon_{t;\alpha} - n^{-1/2}\sigma_\alpha \mathbf{z}'\mathbf{y}_{t-1}\right) \right.\right.\right.$$
$$\left. - h_\alpha\left(\varepsilon_{t;\alpha} - n^{-1/2}\sigma_\alpha \tilde{\mathbf{z}}'\tilde{\mathbf{y}}_{t-1}\right)\right] + n^{-1/2} z_p \sum_{t=1}^n X_{t-p}\psi(\varepsilon_{t;\alpha})$$
$$\left.\left. - \tfrac{1}{2}(\alpha(1-\alpha))^{1/2}\left[2n^{-1}z_p\tilde{\mathbf{z}}' \sum_{t=1}^n \tilde{\mathbf{y}}_{t-1}X_{t-p} + n^{-1}z_p^2 \sum_{t=1}^n X_{t-p}^2\right]\right|\right\}$$
$$= o_{\mathrm{P}}(A_n).$$

For $z_p = \delta \in (0, C_n)$, we have the identity

$$\sigma_\alpha^{-1} \sum_{t=1}^n \left[h_\alpha\left(\varepsilon_{t;\alpha} - n^{-1/2}\sigma_\alpha \mathbf{z}'\mathbf{y}_{t-1}\right) - h_\alpha\left(\varepsilon_{t;\alpha} - n^{-1/2}\sigma_\alpha\tilde{\mathbf{z}}'\tilde{\mathbf{y}}_{t-1}\right)\right]$$
$$= -n^{-1/2} \sum_{t=1}^n X_{t-p} \int_0^\delta \psi_\alpha\left(\varepsilon_{t;\alpha} - n^{-1/2}\sigma_\alpha\left(\tilde{\mathbf{z}}'\tilde{\mathbf{y}}_{t-1} + X_{t-p}u\right)\right) du.$$

It follows from (A.26) that, for any $\delta \in (0, C_n)$,

$$\sup_{\|\mathbf{z}\| \le C_n,\, \alpha_n \le \alpha \le 1-\alpha_n} \Bigg| -n^{-1/2} \sum_{t=1}^{n} X_{t-p} \int_0^{\delta} \Big[\psi_\alpha\Big(\varepsilon_{t;\alpha} - n^{-1/2}\sigma_\alpha \big(\tilde{\mathbf{z}}\tilde{\mathbf{y}}_{t-1} + X_{t-p}u \big)\Big) - \psi_\alpha(\varepsilon_{t;\alpha})\Big]\, du$$
$$- (\alpha(1-\alpha))^{1/2} n^{-1} \Bigg\{ \sum_{t=1}^{n} X_{t-p} \tilde{\mathbf{z}}' \tilde{\mathbf{y}}_{t-1} \int_0^{\delta} du + \sum_{t=1}^{n} X_{t-p}^2 \int_0^{\delta} u\, du \Bigg\} \Bigg| = o_{\mathrm{P}}(A_n). \tag{A.27}$$

Now, splitting X_{t-p} into $\hat{X}_{t-p} + X_{t-p}^{\perp}$, $t = 1, \dots, n$, and taking into account the orthogonality in $\mathbb{R}^n$ of $(\hat{X}_{1-p}, \dots, \hat{X}_{n-p})'$ and $(X_{1-p}^{\perp}, \dots, X_{n-p}^{\perp})'$, (A.27) still holds if X_{t-p} is replaced either by $\hat{X}_{t-p}$ or by $X_{t-p}^{\perp}$. Since moreover the definition of $X_{t-p}^{\perp}$ implies that $\sum_{t=1}^{n} X_{t-p}^{\perp} \tilde{\mathbf{y}}_{t-1} = \mathbf{0}$, it follows from (A.23) that, for all $\delta \in (0, C_n)$, all $\epsilon > 0$ and $\eta > 0$, there exists $N_0 = N_0(\delta, \epsilon, \eta)$ such that

$$\sup_{\|\mathbf{z}\| \le C_n,\, \alpha_n \le \alpha \le 1-\alpha_n} \Bigg| -n^{-1/2} \sum_{t=1}^{n} X_{t-p}^{\perp} \int_0^{\delta} \Big[\psi_\alpha\Big(\varepsilon_{t;\alpha} - n^{-1/2}\sigma_\alpha \big(\tilde{\mathbf{z}}'\tilde{\mathbf{y}}_{t-1} + X_{t-p}u \big)\Big) - \psi_\alpha(\varepsilon_{t;\alpha})\Big]\, du$$
$$- (\alpha(1-\alpha))^{1/2} n^{-1} \sum_{t=1}^{n} \big(X_{t-p}^{\perp}\big)^2 \int_0^{\delta} u\, du \Bigg| \le A_n \epsilon \tag{A.28}$$

with probability at least equal to $1 - \eta$ for all $n \ge N_0$. A similar statement can be made for the integral running over $(-\delta, 0)$; ψ_α is a nondecreasing function, and $n^{-1}\sum_{t=1}^{n} (X_{t-p}^{\perp})^2$ is bounded by $n^{-1}\sum_{t=1}^{n} (X_{t-p})^2$, which converges in probability to the variance of the stationary solution of (2.1), so that, for all $\delta \in (0, C_n)$, and uniformly in $\|\mathbf{z}\| \le C_n$ and $\alpha_n \le \alpha \le 1 - \alpha_n$,

$$\begin{aligned} -n^{-1/2} &\sum_{t=1}^{n} X_{t-p}^{\perp} \Big[\psi_\alpha\Big(\varepsilon_{t;\alpha} - n^{-1/2}\sigma_\alpha \tilde{\mathbf{z}}'\tilde{\mathbf{y}}_{t-1}\Big) - \psi_\alpha(\varepsilon_{t;\alpha})\Big] \\ &\le \frac{1}{\delta} \int_0^{\delta} \Bigg\{ -n^{-1/2} \sum_{t=1}^{n} X_{t-p}^{\perp} \Big[\psi_\alpha\Big(\varepsilon_{t;\alpha} - n^{-1/2}\sigma_\alpha \big(\tilde{\mathbf{z}}'\tilde{\mathbf{y}}_{t-1} + X_{t-p}u \big) - \psi_\alpha(\varepsilon_{t;\alpha})\Big] \Bigg\}\, du \\ &\le (\alpha(1-\alpha))^{1/2} n^{-1} \sum_{t=1}^{n} \big(X_{t-p}^{\perp}\big)^2 \frac{\delta}{2} + \frac{A_n \epsilon}{\delta} \le K\delta + \frac{A_n \epsilon}{\delta} \end{aligned}$$

with probability larger than or equal to $1 - \eta$, for $n \ge N_0$. Similarly, we obtain that

$$-n^{-1/2} \sum_{t=1}^{n} X_{t-p}^{\perp} \Big[\psi_\alpha\Big(\varepsilon_{t;\alpha} - n^{-1/2}\sigma_\alpha \tilde{\mathbf{z}}'\tilde{\mathbf{y}}_{t-1}\Big) - \psi_\alpha(\varepsilon_{t;\alpha})\Big] \ge -K\delta - \frac{A_n \epsilon}{\delta},$$

still with probability larger than or equal to $1 - \eta$, for $n \geq N_0$. Hence, putting $\delta = (A_n \epsilon)^{1/2}$, we obtain

$$\mathrm{P}\left\{ \sup_{\|\mathbf{z}\| \leq C_n,\, \alpha_n \leq \alpha \leq 1-\alpha_n} \left| n^{-1/2} \sum_{t=1}^{n} X_{t-p}^{\perp} \left[\psi_\alpha\left(\varepsilon_{t;\alpha} - n^{-1/2} \sigma_\alpha \tilde{\mathbf{z}}' \tilde{\mathbf{y}}_{t-1} \right) - \psi_\alpha(\varepsilon_{t;\alpha}) \right] \right| > (K+1)(A_n \epsilon)^{1/2} \right\} < \eta$$

for $n \geq N_0$. This first part of the proof is thus complete, since, for $\mathbf{z} = (z_0, \mathbf{z}_1^*)$, $\mathbf{z}_1^* \in \mathscr{L}(\mathbf{Q}_0)$, we have $\mathbf{z}'\mathbf{y}_{t-1} = \tilde{\mathbf{z}}'\tilde{\mathbf{y}}_{t-1}$.

Step 2. We now turn to the general case of an arbitrary matrix $\mathbf{Q}$. Consider the linear reparametrization described in Section 2.2 (with $\check{\mathbf{y}}_{t-1}^* = \mathbf{A}^{-1\prime}\mathbf{y}_{t-1}^*$, $\|\mathbf{A}\| = 1$). Lemma 3.1 is totally unaffected, in the sense that, writing $r_n(\mathbf{z}, \alpha; \mathbf{Y}_n)$ instead of $r_n(\mathbf{z}, \alpha)$, we have $r_n(\mathbf{z}, \alpha; \check{\mathbf{Y}}_n) = r_n(\mathbf{A}\mathbf{z}, \alpha; \mathbf{Y}_n)$ and $\|\mathbf{A}\mathbf{z}\| \leq \|\mathbf{z}\|$. Step 1 of the proof implies that

$$\sup_{\|\check{\mathbf{z}}\| \leq C_n,\, \alpha_n \leq \alpha \leq 1-\alpha_n} \left\| n^{-1/2} \sum_{t=1}^{n} \check{\mathbf{y}}_{t-1}^{\perp} \left[\psi_\alpha\left(\varepsilon_{t;\alpha} - n^{-1/2} \sigma_\alpha \check{\mathbf{z}}' \check{\mathbf{y}}_{t-1} \right) - \psi_\alpha(\varepsilon_{t;\alpha}) \right] \right\| = o_{\mathrm{P}}(A_n) \tag{A.29}$$

for all $\check{\mathbf{z}}$ of the form $(\check{z}_0, \check{z}_1, \dots, \check{z}_q, 0, \dots, 0)'$. The desired result (3.11) then follows from letting $(z_1, \dots, z_p)' = \mathbf{A}^{-1}(\check{z}_0, \check{z}_1, \dots, \check{z}_q, 0, \dots, 0)' \in \mathscr{L}(\mathbf{Q})$ in (A.29) and noting again that $\|\mathbf{A}\| = 1$. □

A.4. Proof of Theorem 3.2. Inserting $\mathbf{z}' = n^{1/2}\sigma_\alpha^{-1}(\hat{\rho}_0(\alpha) - F^{-1}(\alpha), \hat{\boldsymbol{\rho}}_1'(\alpha)\mathbf{Q}' - \boldsymbol{\theta}')$ [a random vector of the form $\mathbf{z} = (z_0, \mathbf{z}_1^{*\prime})' \in \mathbb{R} \times \mathscr{L}(\mathbf{Q})$, which, in view of Theorem 3.1, is $O_{\mathrm{P}}(C_n)$] into (3.11) yields

$$\sup_{\alpha_n \leq \alpha \leq 1-\alpha_n} n^{-1/2} \left\| \sum_{t=1}^{n} \mathbf{y}_{t-1}^{\perp} \left[\left(I\left[\varepsilon_t > \hat{\rho}_0(\alpha) + (\hat{\boldsymbol{\rho}}_1'(\alpha)\mathbf{Q}' - \boldsymbol{\theta}')\mathbf{y}_{t-1}^* \right] - (1-\alpha) \right) - \left(I\left[\varepsilon_t > F^{-1}(\alpha) \right] - (1-\alpha) \right) \right] \right\| = o_{\mathrm{P}}\left(n^{-1/4} C_n^{1/2} \right). \tag{A.30}$$

Now, from (2.9), we have that $I[\varepsilon_t > \hat{\rho}_0(\alpha) + (\hat{\boldsymbol{\rho}}_1'(\alpha)\mathbf{Q}' - \boldsymbol{\theta}')\mathbf{y}_{t-1}^*] = 1$ iff $\hat{a}_t(\alpha) = 1$, and, in view of (A.9),

$$\begin{aligned} \sup_{\alpha_n \leq \alpha \leq 1-\alpha_n} n^{-1/2} \left\| \sum_{t=1}^{n} \mathbf{y}_{t-1}^{\perp} I\left[\varepsilon_t = \hat{\rho}_0(\alpha) + (\hat{\boldsymbol{\rho}}_1'(\alpha)\mathbf{Q}' - \boldsymbol{\theta}')\mathbf{y}_{t-1}^* \right] \right\| &= O_{\mathrm{P}}\left(n^{-1/2} \max_{1 \leq t \leq n} \|\mathbf{y}_{t-1}^{\perp}\| \right) \\ &= o_{\mathrm{P}}\left(n^{-1/4} C_n^{1/2} \right). \end{aligned} \tag{A.31}$$

On the other hand, since $\hat{\mathbf{Y}}_n^{*\prime} = \mathbf{Y}_n^* \mathbf{Y}_n^{*\prime}\mathbf{Q}(\mathbf{Q}'\mathbf{Y}_n^{*\prime}\mathbf{Y}_n^*\mathbf{Q})^{-1}\mathbf{Q}'\mathbf{Y}_n^{*\prime}$, the linear constraints in (2.8) imply that, for all α and n, $\hat{\mathbf{Y}}_n^{*\prime}(\hat{\mathbf{a}}_{Q;n}(\alpha) - (1-\alpha)\mathbf{1}_n) = \mathbf{0}$. Moreover, $\sum_{t=1}^n(\hat{a}_{n;t}(\alpha) - (1-\alpha)) = 0$, so that

$$\sum_{t=1}^{n}\left(\hat{\mathbf{y}}_{t-1}^* + \left(\bar{\mathbf{y}}_n^* - \bar{\mathbf{y}}_n^0\right)\right)\left(\hat{a}_{n;t}(\alpha) - (1-\alpha)\right) = 0.$$

Going back to (A.30), we obtain

$$\sup_{\alpha_n\le\alpha\le 1-\alpha_n} n^{-1/2}\left\| \sum_{t=1}^{n}\left[\mathbf{y}_{t-1}^*\left(\hat{a}_{n;t}(\alpha) - (1-\alpha)\right) - \mathbf{y}_{t-1}^{\perp}\left(\tilde{a}_{n;t}(\alpha) - (1-\alpha)\right)\right]\right\| = O_{\mathrm{P}}(n^{-1/4}) \tag{A.32}$$

as $n \to \infty$. Thus, it is sufficient to consider the behavior of the process (3.8) on the intervals $[0, \alpha_n]$ and $[1-\alpha_n, 1]$, where we have

$$\begin{aligned}\sup_{0\le\alpha\le\alpha_n} n^{-1/2}\left\|\sum_{t=1}^{n}\mathbf{y}_{t-1}^{\perp}\hat{a}_{n;t}(\alpha)\right\| &= \sup_{0\le\alpha\le\alpha_n} n^{-1/2}\left\|\sum_{t=1}^{n}\mathbf{y}_{t-1}^{\perp}(1-\hat{a}_{n;t}(\alpha))\right\| \\ &\le \tau_n^* n^{-1/2}\sup_{\alpha}\sum_{t=1}^{n}(1-\hat{a}_{n;t}(\alpha)) \\ &= O_{\mathrm{P}}\left(\tau_n^* n^{1/2}\alpha_n\right) \\ &= O_{\mathrm{P}}\left(n^{-1/2}(\log n)^{2+1/r_1}(\log\log n)^{2+1/4}\right) \\ &= o_{\mathrm{P}}(n^{-1/4}).\end{aligned} \tag{A.33}$$

Similarly,

$$\sup_{0\le\alpha\le\alpha_n} n^{-1/2}\left\|\sum_{t=1}^{n}\mathbf{y}_{t-1}^{\perp}\tilde{a}_{n;t}(\alpha)\right\| = o_{\mathrm{P}}(n^{-1/4}).$$

The treatment is entirely the same for $1-\alpha_n \le \alpha \le 1$. □

A.5. Proof of Theorem 3.3. First, note that $\hat{a}_{n;t}(\alpha) - \tilde{a}_{n;t}(\alpha) = 0$ at $\alpha = 0$ and $\alpha = 1$. Integrating by parts [note that the integrals involved in the definition converge: see (3.3)], we obtain

$$-\int_0^1\varphi(u)\,d\left(\hat{a}_{n;t}(u) - \tilde{a}_{n;t}(u)\right) = \int_0^1\left(\hat{a}_{n;t}(u) - \tilde{a}_{n;t}(u)\right)d\varphi(u).$$

By Theorem 3.2 and the dominated convergence theorem,

$$n^{-1/2}\sum_{t=1}^{n}\mathbf{y}_{t-1}^{\perp}\int_{\alpha_0}^{1-\alpha_0}\left(\hat{a}_{n;t}(u) - \tilde{a}_{n;t}(u)\right)d\varphi(u) = o_{\mathrm{P}}(1).$$

For $u \in [\alpha_n, \alpha_0]$, by (3.12) and (A.32),

$$
\begin{aligned}
n^{-1/2}&\left\| \sum_{t=1}^{n} \mathbf{y}_{t-1}^{\perp} \int_{\alpha_n}^{\alpha_0} (\hat{a}_{n;t}(u) - \tilde{a}_{n;t}(u))\, d\varphi(u) \right\| \\
&\le c \int_{\alpha_n}^{\alpha_0} (u(1-u))^{-1-\delta} n^{-1/2} \left\| \sum_{t=1}^{n} \mathbf{y}_{t-1}^{\perp} (\hat{a}_{n;t}(u) - \tilde{a}_{n;t}(u)) \right\| du \\
&\le [u^{-\delta}]_{\alpha_n}^{\alpha_0} O_{\mathrm{P}}(n^{-1/4}) = O_{\mathrm{P}}\Big(n^{-(1/4)+\delta} (\log n \log\log n)^{-2\delta}\Big) = o_{\mathrm{P}}(1).
\end{aligned}
$$

Finally, for $u \in (0, \alpha_n]$, the inequality

$$
n^{-1/2}\left\| \sum_{t=1}^{n} \mathbf{y}_{t-1}^{\perp} \hat{a}_{n;t}(u) \right\| = n^{-1/2}\left\| \sum_{t=1}^{n} \mathbf{y}_{t-1}^{\perp} (1 - \hat{a}_{n;t}(u)) \right\| \le n^{1/2} \max_{1 \le t \le n} \|\mathbf{y}_{t-1}^{\perp}\| u
$$

leads to

$$
\begin{aligned}
n^{-1/2}&\left\| \sum_{t=1}^{n} \mathbf{y}_{t-1}^{\perp} \int_{0}^{\alpha_n} \hat{a}_{n;t}(u)\, d\varphi(u) \right\| \\
&\le n^{-1/2} \int_{0}^{\alpha_n} (u(1-u))^{-1-\delta} \left\| \sum_{t=1}^{n} \mathbf{y}_{t-1}^{\perp} \hat{a}_{n;t}(u) \right\| du \\
&\le n^{1/2} \max_{1 \le t \le n} \|\mathbf{y}_{t-1}^{\perp}\| \int_{0}^{\alpha_n} u^{-\delta} (1-u)^{-1-\delta}\, du \\
&= O_{\mathrm{P}}\Big(n^{-(1/4)+\delta} (\log n \log\log n)^{2(1-\delta)}\Big) = o_{\mathrm{P}}(1).
\end{aligned}
\tag{A.34}
$$

Similarly,

$$
\begin{aligned}
n^{-1/2}&\left\| \sum_{t=1}^{n} \mathbf{y}_{t-1}^{\perp} \int_{0}^{\alpha_n} \tilde{a}_{n;t}(u)\, d\varphi(u) \right\| \\
&\le n^{-1/2} \int_{0}^{\alpha_n} (u(1-u))^{-1-\delta} \left\| \sum_{t=1}^{n} \mathbf{y}_{t-1}^{\perp} I[\varepsilon_t > F^{-1}(u)] \right\| du \\
&= o_{\mathrm{P}}\Big(n^{-(1/4)+\delta} (\log n \log\log n)^{-2\delta}\Big)
\end{aligned}
\tag{A.35}
$$

as $n \to \infty$. The treatment, again, is entirely similar for $1 - \alpha_0 < \alpha < 1$. Asymptotic normality readily follows. □

Acknowledgment. The authors gratefully acknowledge the suggestions of the Associate Editor and the very careful reading and constructive remarks of an anonymous referee, which led to a definite improvement of the manuscript.

REFERENCES

BAHADUR, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37** 557–580.

BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series: Theory and Methods*, 2nd ed. Springer, New York.

CHANDA, K. C., PURI, M. L. and RUYMGAART, F. (1990). Asymptotic normality of L-statistics based on $m(n)$-decomposable time series. *J. Multivariate Anal.* **35** 260–275.

CHERNOFF, H. and SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric tests. *Ann. Math. Statist.* **29** 972–994.

GODFREY, L. G. (1979). Testing the adequacy of a time series model. *Biometrika* **66** 67–72.

GUTENBRUNNER, C. and JUREČKOVÁ, J. (1992). Regression rank scores and regression quantiles. *Ann. Statist.* **20** 305–330.

GUTENBRUNNER, C., JUREČKOVÁ, J., KOENKER, R. and PORTNOY, S. (1993). Tests of linear hypotheses based on regression rank scores. *J. Nonparametric Statist.* **2** 307–331.

HALLIN, M. (1994). On the Pitman-nonadmissability of correlogram-based methods. *J. Time Ser. Anal.* **15** 607–612.

HALLIN, M., JUREČKOVÁ, J., KALVOVÁ, J., PICEK, J. and ZAHAF, T. (1997). Non-parametric tests in AR models, with applications to climatic data. *Environmetrics* **8** 651–660.

HALLIN, M. and PURI, M. L. (1988). Optimal rank-based procedures for time-series analysis: testing an ARMA model against other ARMA models. *Ann. Statist.* **16** 402–432.

HALLIN, M. and PURI, M. L. (1994). Aligned rank tests for linear models with autocorrelated errors. *J. Multivariate Anal.* **50** 175–237.

HALLIN, M. and WERKER, B. J. M. (1998). Optimal testing for semiparametric AR models: from Lagrange multipliers to autoregression rank scores and adaptive tests. In *Asymptotics, Nonparametrics and Time Series* (S. Ghosh, ed.), 295–350. Dekker, New York.

HEILER, S. and WILLERS, R. (1988). Asymptotic normality of R-estimates in the linear model. *Statistics* **19** 173–184.

HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variable. *J. Amer. Statist. Assoc.* **58** 13–30.

HOSKING, J. R. M. (1980). Lagrange multiplier tests of time series model. *J. Roy. Statist. Soc. Ser. B* **42** 170–181.

JUREČKOVÁ, J. (1984). Regression quantiles and trimmed least squares estimator under a general design. *Kybernetika* **20** 345–357.

JUREČKOVÁ, J. (1991). Tests of Kolmogorov–Smirnov type based on regression rank scores. In *Trans. Eleventh Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes* (J. Á. Višek, ed.) 41–49. Academia, Prague.

JUREČKOVÁ, J. (1999). Regression rank scores against heavy-tailed alternatives. *Bernoulli* **5** 1–18.

JURĚKOVÁ, J. and SEN, P. K. (1996). *Robust Statistical Inference: Asymptotics and Interrelations*. Wiley, New York.

KALVOVÁ, J., JUREČKOVÁ, J., NEMEŠOVÁ, I. and PICEK, J. (1998). On the order of autoregression models in temperature data. Unpublished manuscript.

KOENKER, R. and BASSETT, G. (1978). Regression quantiles. *Econometrica* **46** 33–50.

KOENKER, R. and D'OREY, V. (1987). Computing regression quantiles. *Appl. Statist.* **36** 383–393.

KOENKER, R. and D'OREY, V. (1994). A remark on algorithm AS 229: computing dual regression quantiles and regression rank scores. *Appl. Statist.* **43** 410–414.

KOUL, H. L. and SALEH, A. K. MD. E. (1995). Autoregression quantiles and related rank scores processes. *Ann. Statist.* **23** 670–689.

KOUL, H. L. and SCHICK, A. (1997). Efficient estimation in nonlinear autoregressive time-series models. *Bernoulli* **3** 247–277.

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

KREISS, J. P. (1990). Testing linear hypotheses in autoregression. *Ann. Statist.* **18** 1470–1482.

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

O'BRIEN, G. L. (1987). Extreme values for stationary and Markov sequences. *Ann. Probab.* **15** 281–291.

OSBORNE, M. R. (1992). An effective method for computing regression quantiles. *IMA J. Numer. Anal.* **12** 151–166.

PHAM, D. T. and TRAN, L. T. (1985). Some mixing properties of time series models. *Stochastic Process Appl.* **19** 297–303.

POLLARD, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econometric Theory* **7** 186–200.

PORTNOY, S. (1991). Asymptotic behavior of regression quantiles in non-stationary, dependent cases. *J. Multivariate Anal.* **38** 100–113.

PÖTSCHER, B. M. (1983). Order estimation in ARMA models by Lagrangian multiplier tests. *Ann. Statist.* **11** 872–885.

RUPPERT, D. and CARROLL, R. J. (1980). Trimmed least squares estimation in the linear model. *J. Amer. Statist. Assoc.* **75** 828–838.

SCHEFFÉ, H. (1959). *The Analysis of Variance*. Wiley, New York.

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

SWENSEN, A. R. (1985). The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend. *J. Multivariate Anal.* **16** 54–70.

INSTITUT DE STATISTIQUE
ET DE RECHERCHE OPERATIONNELLE
UNIVERSITÉ LIBRE DE BRUXELLES
CAMPUS PLAINE CP 210
BOULEVARD DU TRIOMPHE
B-1050 BRUXELLES
BELGIUM

DEPARTMENT OF PROBABILITY AND STATISTICS
CHARLES UNIVERSITY
SOKOLOVSKÁ 83
CZ-18675 PRAGUE 8
CZECH REPUBLIC