RANK TESTS FOR TIME SERIES ANALYSIS A SURVEY

By

Marc Hallin

and

Madan L. Puri

IMA Preprint Series # 876

October 1991

This article will appear in the IMA Volumes in Mathematics and its Applications, Proceedings on Time Series.

RANK TESTS FOR TIME SERIES ANALYSIS A SURVEY

MARC HALLIN* AND MADAN L. PURI†

Abstract. Rank-based testing procedures have proven quite efficient in classical linear models with independent observations; they long ago have entered daily practice in such fields as biostatistics or experimental planning. Still, despite the fact that many "historical" rank statistics (runs, signs of differences, turning point, Wald and Wolfowitz's rank autocorrelation coefficient, ...) actually were devised for time series situations, and despite the recognized need for nonparametric, robust or non Gaussian methods in the area, most time series analysts completely ignore rank–based procedures. Our objective with the present survey is to show that rank–based techniques very successfully can handle most of the testing problems occurring in time series context, such as testing for white noise, testing ARMA (p, q) dependence against ARMA $(p + d, q + d)$ dependence, or testing linear hypotheses about the coefficients of an ARMA model—always, under unspecified innovation density. The resulting tests of course are distribution–free. But they also are as powerful as (often, strictly more powerful than) their classical, correlogram–based counterparts. In addition, they are considerably more robust: whereas classical parametric methods can yield extremely misleading diagnostic information when the data have outliers, atypical startup behavior or heavy–tailed distributions, rank–based tests exhibit much better resistance to aberrations of this type. All these properties should make them extremely attractive, e.g. in the identification and diagnostic checking process, where the conclusions drawn from rank–based techniques are likely to be much more reliable than those resulting from an inspection of traditional correlograms.

We start (Section 1) with a general introduction, where we show how invariance arguments naturally lead to (signed or unsigned) rank–based methods. Section 2 provides a bibliographical survey of rank–based methods in serial dependence problems. The so–called linear serial rank statistics and their asymptotic distributions are introduced in Section 3. Section 4 deals with the basic theoretical result from which the optimal testing procedures described in the subsequent sections follow: local asymptotic normality of ARMA likelihood families and the local sufficiency of ranks. Locally most powerful and locally maximin rank–based tests are investigated in Section 5. Their asymptotic performance is discussed; explicit ARE (asymptotic relative efficiency) values (with respect to Gaussian parametric methods) are provided. Finally, Section 6 deals with the important problem of aligned rank tests: asymptotic invariant aligned rank tests, asymptotically most stringent within the class of all asymptotically similar tests at given probability level are derived for linear hypotheses about the coefficients of an ARMA model (with unspecified innovation density). The ARE of these tests with respect to the corresponding Gaussian Lagrange multiplier method is still the same as in Section 5, and can be as high as 2, e.g. when using median–test (or Laplace) scores under double–exponential innovation density.

AMS(MOS) subject classifications. Primary 62G10, 62M10

Key words and phrases. Rank tests, signed rank tests, invariance, aligned rank tests, permutation tests, nonhomogeneous white noise, linear trend, serial and nonserial linear rank statistics, local asymptotic normality, local asymptotic sufficiency, rank autocorrelations, locally optimal rank tests, asymptotic relative efficiencies, asymptotic invariance, asymptotically similar tests.

---

*Département de Mathématique and Institut de Statistique, C.P. 210, Université Libre de Bruxelles, 1050 Bruxelles, Belgium

†Department of Mathematics, Indiana University, Swain Hall East, Bloomington, Indiana 47405 U.S.A.
Research supported by the Office of Naval Research Contract N00014–91–J–1020

§1. INVARIANCE, THE HYPOTHESIS OF WHITE NOISE AND RANKS

1.1 Testing for white noise: invariance arguments. Consider the null hypothesis H under which a series $\mathbf{X} = (X_1, \ldots, X_n)$ of observations is white noise. More precisely, let H denote the hypothesis under which observations $X_1, \ldots, X_t, \ldots, X_n$ are independent and identically distributed, with unspecified density (all densities here are with respect to the Lebesgue measure on $\mathbb{R}$). As an alternative, take the extremely vast and general collection K of all n-dimensional distributions under which the observed X_t's are not white noise anymore—either because of distributional heterogeneity, serial dependence, or both.

The whole theory of rank tests originates in this almost trivial observation that, letting $g : \mathbb{R} \to \mathbb{R}$ denote an arbitrary, continuous, one–to–one (hence monotonic: without any loss of generality, it can be assumed order–preserving) transformation, $\mathbf{X} = (X_1, \ldots, X_n)$ is white noise if and only if $g(\mathbf{X}) = (g(X_1), \ldots, g(X_n))$ is. More rigorously, it is easy to see that the problem of testing H against K is invariant with respect to the group $\mathcal{G}$ of continuous, order–preserving transformations $g : \mathbf{x} \mapsto g(\mathbf{x})$.

The classical attitude in such a situation consists in restricting oneself to the consideration of invariant tests, i.e. tests which are measurable with respect to some maximal invariant statistic. And a maximal invariant statistic for H consists of the vector of ranks $\mathbf{R} = (R_1, \ldots, R_n)$, where R_t denotes the rank of X_t among $X_1, \ldots, X_n$. This is how ranks come into the picture: invariance is the theoretical cornerstone of rank–based inference. Distribution–freeness (hence similarity and unbiasedness), increased power, robustness, ..., however attractive they are, are just by–products, and generally can be reached through other, more specific techniques.

As a consequence, there is no point in using ranks if, under the null hypothesis to be tested, the quantities from which ranks are computed do not constitute a white noise series (or, at least, an exchangeable series).

The above invariance argument is well accepted in such statistical areas as life testing or experimental planning. It probably sounds highly unfamiliar, theoretical and abstract to most time series analysts. Time series analysis indeed is very deeply marked with hidden Gaussian assumptions — more deeply, perhaps, than any other statistical area: the $\mathcal{L}^2$ approach, pervasive use of autocovariances and autocorrelations, Gaussian likelihood or least squares estimates, Gaussian Lagrange multipliers, ... all arise from (explicit or implicit) Gaussian assumptions. Though most of such Gaussian methods generally remain *asymptotically* valid under fairly general assumptions, it is usually agreed that the data preferably should be at least *approximately* Gaussian. Whenever they really do not look Gaussian , various preliminary transformations are generally performed before the analysis is started: logarithms, Box–Cox transformations, etc... Now, if the objective is "to normalize" (under H) a series $X_1, \ldots, X_n$, the exact transformation to be applied is $\Phi^{-1} \circ F$, where F denotes the cumulated distribution function of the X_t's and Φ stands for

the standard normal distribution function. Of course, the trouble is that F in practice is unknown. However, if the empirical distribution function $\widehat{F}$ is substituted for the unknown one, $\Phi^{-1} \circ \widehat{F}$ computed at X_t is nothing else than $\Phi^{-1}(R_t/(n+1))$, where R_t is the rank of X_t among $X_1, \ldots, X_n$; $\widehat{F}(X_t)$ here is taken as $R_t/(n+1)$ in order to avoid trivial complications in the definition of $\Phi^{-1} \circ \widehat{F}$. The transformed series then is measurable with respect to the vector of ranks, and any subsequent inference procedure (of the van der Waerden type, in rank–order theory terminology) will be rank–based, hence invariant. The idea of invariance thus is not that remote from time series practice where the parameters e.g. of a Box–Cox transformation are also estimated from the data.

1.2 Testing for white noise: unbiasedness arguments. Still, one might be reluctant in enforcing the invariance principle underlying the use of rank–based tests. A more commonly accepted hypothesis testing principle then is the principle of unbiasedness. It is well known that (under fairly general continuity assumptions) a necessary condition for unbiasedness is similarity. Whenever a sufficient, boundedly complete statistic exists for the submodel consisting of the common border $\overline{H} \cap \overline{K}$ between the null hypothesis and the alternative (in any topology for which expectation is continuous with respect to the distributions in H and K), α-similarity in turn is equivalent to Neyman's α–structure property, with respect to the latter statistic. And test statistics possessing this latter property are obtained by conditioning upon the sufficient statistic at hand, then considering conditional tests.

Clearly, in the present problem of testing H against K , $\overline{H} \cap \overline{K} = H$, and a sufficient, complete statistic is the order statistic $\mathbf{X}_{(\cdot)} = (X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)})$. Conditioning upon $\mathbf{X}_{(\cdot)}$ yields the class of *permutation tests*.

Whereas invariance arguments lead to rank–based tests, unbiasedness arguments thus lead to permutation tests. Since rank tests are a particular case of permutation tests, unbiasedness conditions are uniformly weaker than invariance ones.

As we shall see, the class of rank–based tests and hence, a fortiori, that of permutation tests are locally asymptotically essentially complete for problems dealing e.g. with ARMA models, in the sense that they both contain locally asymptotically optimal tests for all usual testing problems in the area.

1.3 Hypotheses reducing to white noise. Aligned ranks. The hypothesis of white noise of course plays a very fundamental role, in its own right, in a variety of classical problems, such as testing against location shifts in two–or k-sample situations, testing against trend or serial dependence, etc. This role is even more fundamental in view of the host of problems which after adequate transformation reduce to that of testing for white noise.

In most statistical models, an observed series $\mathbf{X} = (X_1, \ldots, X_n)$ indeed is such that, denoting by $\boldsymbol{\theta}$ some parameter, the transformed series $\mathbf{Z}(\boldsymbol{\theta}) = (Z_1, \ldots, Z_m) = \mathcal{F}_{\boldsymbol{\theta}}(\mathbf{X})$— call it a residual series—is white noise. In a "parametric" setting, the distribution of

this residual series is assumed to be known (in most cases, it is assumed to be Gaussian); very often however, a "nonparametric" approach, where the residual density remains unspecified, would be much more realistic. The problem of testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ then reduces to that of testing for white noise (with unspecified density) in terms of the residual series $\mathbf{Z}(\boldsymbol{\theta}_0)$.

Accordingly, if invariance arguments are to be considered, rank tests should be used, where the ranks R_t are computed from the residual series $\mathbf{Z}(\boldsymbol{\theta}_0)$. Similarly, permutation tests, satisfying weaker unbiasedness requirements, should be based on the $m!$ permutations of $\mathbf{Z}(\boldsymbol{\theta}_0)$.

In the specific area of time series analysis, ARMA models, or linear models with ARMA error terms, or bilinear models, fall within the above class of statistical models whenever the innovation density remains unspecified. Consider for example the case of an autoregressive process with linear trend.

Example 1.1. AR(1) model with linear trend. The observations X_t, $t = 0, 1, \ldots, n$, satisfy

$$X_t = \alpha + \beta c_t + e_t, \qquad t = 0, 1, \ldots, n$$

where α and β are regression parameters, $c_0, c_1, \ldots, c_n$ are known regression constants, and e_t is some solution of

$$e_t - \rho e_{t-1} = \varepsilon_t \ , \qquad t = 1, \ldots, n \ ,$$

with $\{\varepsilon_t\}$ an iid process with unspecified density, mean zero and variance one, and $|\rho| < 1$. Clearly, the null hypothesis

$$H_0 : (\alpha, \beta, \rho) = (\alpha_0, \beta_0, \rho_0)$$

is equivalent to

$$H : \mathbf{Z}(\alpha_0, \beta_0, \rho_0) \quad \text{is white noise (unspecified density)}$$

where

$$\mathbf{Z}(\alpha_0, \beta_0, \rho_0) = (Z_1, \ldots, Z_n) \ ,$$

with

$$Z_t = X_t - \rho_0 X_{t-1} - \alpha_0(1 - \rho_0) - \beta_0(c_t - \rho_0 e_{t-1}) \ , \quad t = 1, \ldots, n \ .$$

Still, hypotheses to be tested in practice seldom are of the simple form $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$, under which the parameter $\boldsymbol{\theta}$ has a completely specified value $\boldsymbol{\theta}_0$. In most problems, null hypotheses take the more general form $H_0 : \boldsymbol{\theta} \in \Theta_0$, where Θ_0 is some subset of the parametric space Θ. Common examples in time series analysis are: testing an ARMA (p, q) model (with unspecified innovation density and unspecified autoregressive and moving average

coefficients) against alternatives of ARMA $(p+\pi, q+\pi)$ dependence (π a given positive integer), testing for the absence of trend in an ARMA (p,q) process with unspecified covariance structure and unspecified innovation density, testing for ARMA (p,q) dependence (specified orders p and q, unspecified coefficients, unspecified innovation density) against bilinear dependence, etc.

Though a transformation $\mathcal{F}_{\boldsymbol{\theta}}$ still exists which transforms the observed series $\mathbf{X}$ into a residual white noise $\mathbf{Z}(\boldsymbol{\theta})$, the value of $\boldsymbol{\theta}$ remains unspecified also under H_0.

A tempting, intuitively natural idea then consists in substituting some "good" estimate $\widehat{\boldsymbol{\theta}}$ for the unknown paramether $\boldsymbol{\theta}$, and an "estimated" residual series $\mathbf{Z}(\widehat{\boldsymbol{\theta}})$ for the "true" one $\mathbf{Z}(\boldsymbol{\theta})$. By a "good" estimate $\widehat{\boldsymbol{\theta}}$, it is meant that $\widehat{\boldsymbol{\theta}}$ has values in Θ_0 and enjoys, under H_0, such standard properties as being root n-consistent. Estimated residuals can be expected to be close to the exact ones. Treating them as the latter yields *aligned ranks* $\widehat{R}_t$ and *aligned rank tests*, or *aligned permutation tests*.

The big trouble with aligned ranks is that they generally do not enjoy (even "approximately") the fundamental invariance properties of exact ranks: the estimation of $\boldsymbol{\theta}$ indeed most of time destroys the exchangeability structure of $\mathbf{Z}(\boldsymbol{\theta})$ which constitutes the justification of rank–based methods, so that from a decision–theoretical point of view, there is no point anymore in using (aligned) ranks. By the way, the distribution–freeness property of rank–based statistics under the null hypothesis, which was a consequence of invariance, also disappears, hence the similarity and unbiasedness features of (genuinely distribution-free) rank tests. As a consequence, the computation of exact critical values for aligned rank tests is generally impossible. Even worse: the asymptotic results available for rank statistics (based on exact ranks R_t) generally cannot be used for the corresponding aligned ones (based on the aligned ranks $\widehat{R}_t$) since in general exact and aligned rank statistics are not even asymptotically equivalent.

From a theoretical point of view as from the point of view of practice, aligned rank methods (as well as aligned permutation tests) thus are settling quite a number of nontrivial problems. Section 6 below is devoted to providing some methodology in solving these problems in time series context.

1.4 Testing for symmetric white noise. The density of the observations X_t (of the residuals Z_t in section 1.3) under the null hypothesis H so far has been assumed to be completely unspecified (except perhaps for moment conditions such as, e.g. the existence of a finite variance or finite Fisher information). In many practical problems, this density further can be assumed to be symmetric (with respect to some known median, which can be set equal to zero without any loss of generality). Denote by H' this hypothesis of symmetric white noise.

The vector of ranks $\mathbf{R}$ then loses its maximal invariant status for the benefit of the couple $(\mathbf{s}, \mathbf{R}_+)$, where $\mathbf{s} = (s_1, \ldots, s_n)$ is the vector of signs $s_t = \text{sign}\,(X_t)$ and $\mathbf{R}_+ = (R_{+,1}, \ldots, R_{+,n})$ is the vector of the ranks $R_{+,t}$ of the absolute values $|X_t|$ among

$|X_1| \ldots |X_n|$ (invariance here is with respect to the group of continuous, even, order-preserving transformations).

Statistics which are measurable with respect to this maximal invariant are known as *signed–rank statistics*.

The unbiasedness approach to the same problem leads to conditioning upon the order statistic of absolute values $|\mathbf{X}|_{(\cdot)} = (|X|_{(1)} \leq \cdots \leq |X|_{(n)})$, which is sufficient complete here, yielding a class of sign-assignment – absolute value permutation tests, based on the $2^n n!$ possible combination of a permutation of $|\mathbf{X}|_{(\cdot)}$ and a vector of n signs.

1.5 Testing for nonhomogeneous white noise. A further case frequently occurs in applications, in which the observations, or adequate residual values, under the null hypotheses to be tested, are still independent and symmetrically distributed, though their distributions possibly might not be identical anymore—denote this by H''. This hypothesis H'' of nonhomogeneous symmetric white noise allows for heteroskedasticity, contamination, etc. Symmetric discrete distributions are also allowed.

A maximal invariant (with respect to the group of componentwise continuous, even and order–preserving transformations of $\mathbf{X}$) is the vector $\mathbf{s}$ of signs, leading to sign tests.

The unbiasedness approach—conditioning upon the vector of absolute values $(|X_1|, \ldots, |X_n|)$—yields the broader class of tests which are *conditionally* measurable with respect to the vector signs.

§2. NONPARAMETRIC TESTS FOR SERIAL DEPENDENCE: A BIBLIOGRAPHICAL SURVEY

The nonparametric, rank–based approach to time series analysis problems actually has a pretty long history, since rank tests against serial dependence and runs tests, which are a particular case of rank tests, can be traced back to the very beginnings of rank–based inference. Wald and Wolfowitz (1943) already suggest to substitute the ranks, or some function thereof, for the observations in the problem of testing randomness against serial dependence. Tests based on runs, runs up and down, signs of first differences or turning points already had been considered, for the same problem, in Fisher (1926), Kermack and McKendrick (1937), Mood (1940) and Wallis and Moore (1941); they subsequently have been developed in Moore and Wallis (1943), Levene and Wolfowitz (1944), Wolfowitz (1944), David (1947), Goodman (1958), Edgington (1961) and Granger (1963).

Jogdeo (1968) derives asymptotical results for a very general class of rank statistics, which includes the rank autocorrelation coefficients to be used repeatedly in subsequent sections. The conditions he puts on the score functions however are too restrictive for most purposes (excluding, e.g. the so–called van der Waerden autocorrelations).

In none of these early papers is any particular alternative considered, nor any optimality question addressed. The first attempts to investigate the power of serial rank

procedures against specific alternatives are due to Beran (1972), who introduces rank analogues of integrated periodogram spectral processes, and, in a more applied context, to Knoke (1977), who conducts a Monte Carlo study of the asymptotic relative efficiencies of several tests based on serial rank statistics (namely, the Wald–Wolfowitz rank autocorrelation coefficient, the turning point statistic and a Kolmogorov–Smirnov one) with respect to the classical first–order sample autocorrelation coefficient, for first–order autoregressive alternatives. This, in some sense, was a first step towards the introduction of rank methods in time series practice—but still the statistics studied there are not new, are not specifically devised against any particular alternative, and cannot handle the more general problem of testing, e.g. an ARMA model against other ARMA models (as described in section 1.3).

Locally most powerful rank tests for randomness against Gaussian autoregressive or moving average alternatives are derived in Gupta and Govindarajulu (1980). The test statistics depend on ranks via the expected values of products of order statistics, and are asymptotically equivalent to our van der Waerden autocorrelation coefficients. Aiyar (1981) also investigates first–order statistics of the Wald–Wolfowitz and van der Waerden types, and derives their asymptotic relative efficiencies with respect to the corresponding Gaussian procedure, still under first–order autoregressive alternatives.

In a similar situation, Dufour (1981) suggests (traditional nonserial) signed–rank tests for H' based on the signs and ranks of products of the form $X_t X_{t-k}$. This way of reducing a serial problem to a nonserial one however may result in a loss of relevant information and the ranks used there are not those following from invariance arguments (except in the case of nonhomogeneous white noise—see Section 1.5). In a more applied context, Bartels (1982) introduces a rank–based version of von Neumann's ratio statistic which is asymptotically equivalent to the Wald–Wolfowitz autocorrelation coefficient of order one, and investigates the power of the resulting test through Monte Carlo techniques.

Runs also have been used for testing the absence of cross–correlation between two series: see Goodman and Grunfeld (1961) and Yang and Schreckengost (1981). In the more complex domain of dynamic econometric models, Campbell and Dufour (1991) present a promising rank-based approach to the Mankiw–Shapiro rational expectation model. Switzer (1984) proposes a Wald–Wolfowitz version of the spatial variogram—a rather isolated attempt to introduce ranks in the area of spatial processes.

In a somewhat different, nonserial context, rank tests against trend alternatives have been intensively investigated (see Mann (1945) or Savage (1957) for early papers on the subject, Aiyar, Guillier and Albers (1979) for a more recent one), as well as the asymptotic behavior of nonserial rank statistics, including multivariate ones, under various mixing conditions (Serfling (1968), Albers (1978), Tran (1988), Harel (1988) and Harel and Puri (1989a,b; 1990a; 1991; 1992)).

For more details, we refer to the review papers by Dufour, Lepage and Zeidan (1982) and Bhattacharyya (1984); see also Govindarajulu (1983).

Up to this point, and in spite of a long history, the subject of rank–based inference

for time series problems thus remained largely unexplored. Quite a number of partial results were scattered around, but no coherent and structured theory was available. The few existing central limit theorems for serial rank statistics were too restrictive for most purposes; except for a few exceptions, the only problem that had been considered was that of testing for white noise; optimality problems remained essentially untouched... the rank–order section of the time series analyst's tool kit was pretty limited and nearly empty.

A systematic and coherent treatment of time series problems, based on a LeCam–Hájek approach, since then has been undertaken by the authors in a series of papers, starting with Hallin, Ingenbleek and Puri (1985, 1987), where the problem of testing for white noise against ARMA alternatives is considered in its full generality (yielding locally maximin, rank–based portmanteau tests). The long term objective is to obtain a logically consistent methodology for classical time series problems (identification, diagnostic checking,...), relying on adequate, rank based substitutes for time series analysis familiar tools such as correlograms, partial correlograms, Lagrange multipliers, etc.

Locally asymptotically maximin rank–based tests for testing an ARMA model (with unspecified innovation density) against other ARMA models are derived in Hallin and Puri (1988). The particular case of first–order autoregressive models is treated in detail in Dufour and Hallin (1987). The small sample performance of rank–based tests is investigated in Hallin and Mélard (1989), and appears to be surprisingly good, even for pretty short series. Signed rank techniques, for ARMA models with symmetric, otherwise unspecified innovation densities, are considered in Hallin and Puri (1991a) and Hallin, Laforet and Mélard (1989). Optimal runs tests, allowing for nonhomogeneous innovation processes, are treated in Dufour and Hallin (1990a), the problem of testing multivariate white noise against alternatives of (multivariate) ARMA dependence in Hallin, Ingenbleek and Puri (1989) and Hallin and Puri (1991b), and that of testing white noise against first–order, superdiagonal bilinear dependence in Benghabrit and Hallin (1992). Finally, the very general case of ARMA models with a regression trend (also known as the dynamic regression model) is studied in Hallin and Puri (1991c), where locally optimal, aligned, rank and signed–rank, tests are derived for a variety of problems, including that of testing an ARMA (p, q) model with unspecified trend component, unspecified ARMA coefficients and unspecified innovation density, against ARMA $(p + \pi, q + \pi)$ alternatives.

A review of the main results in the above papers is the subject of the present survey.

Related results on the asymptotic distribution of serial rank statistics under dependence can be found in Harel and Puri (1990b,c), Tran (1990), Nieuwenhuis and Ruymgaart (1990). Permutation tests against serial dependence are considered in David and Fix (1966), Ghosh (1954), Dufour and Roy (1985, 1986), Dufour and Hallin (1990b) and Hallin, Mélard and Milhaud (1992). A somewhat hybrid test, with mixed parametric and nonparametric (rank-based) features has been considered by Kreiss (1990a) for testing AR models.

Rank–based techniques have proven very efficient in classical linear model problems

(regression, trend, analysis of variance,...), and have entered daily practice in biostatistics and experimental planning. When the required technology is available, there is no reason for time series analysis to escape this rule—even more so in view of the recognized need for robust and non–Gaussian methods in the area.

§3. LINEAR RANK STATISTICS

3.1 Serial and nonserial linear rank statistics. Denote by $\mathbf{Z}^{(n)} = Z_1^{(n)}, \ldots, Z_t^{(n)}, \ldots, Z_n^{(n)})$ a series of length n. Depending on the problem, $\mathbf{Z}^{(n)}$ either may be the original observation, or some residual series. Let $R_t^{(n)}$ be the rank of $Z_t^{(n)}$ among $Z_1^{(n)}, \ldots, Z_n^{(n)}$. If $\mathbf{Z}^{(n)}$ is white noise, then, with probability one, $\mathbf{R}^{(n)} = (R_1^{(n)}, \ldots, R_n^{(n)})$ is uniformly distributed over the $n!$ permutations of $(1, 2, \ldots, n)$, whatever the underlying density of the $Z_t^{(n)}$'s.

It is well known (Hájek and Šidák 1967; Puri and Sen 1971 and 1985) that locally asymptotically best tests for linear models (two or k samples, regression, analysis of variance,...) can be based upon the class of *(simple nonserial) linear rank statistics*, of the form

$$S^{(n)} = n^{-1} \sum_{t=1}^{n} c_t a^{(n)}(R_t^{(n)}) \ , \tag{3.1}$$

where $a^{(n)}$ denotes some *score function*, and the c_t's are known *(regression)* constants. In a time series context, they also can be used in testing against trend (Mann 1945; Savage 1957). Such nonserial statistics however cannot capture serial dependence features: actually, it can be shown that their asymptotic distribution is the same under white noise as under local alternatives of serial dependence. Hallin et al. (1985) therefore suggest the consideration of *serial linear rank statistics*, of the form

$$S^{(n)} = (n-k)^{-1} \sum_{t=k+1}^{n} a^{(n)}(R_t^{(n)}, R_{t-1}^{(n)}, \ldots, R_{t-k}^{(n)}) \ , \tag{3.2}$$

where the score function $a^{(n)}$ depends on the values of $(k+1)$ successive ranks: (3.2) is a serial linear rank statistic *of order* k, and can be expected to be sensitive to serial dependencies of orders 1 through k. Nonserial statistics can be considered as a particular case ($k = 0$) of the serial ones (or linear combinations thereof).

Classical examples are the *runs statistic* (where runs are taken with respect to the median), or order one, with scores

$$a^{(n)}(i_1, i_2) = I[(2i_1 - n - 1)(2i_2 - n - 1) < 0] \ ,$$

the *turning point* statistic, or order three, with scores

$$a^{(n)}(i_1, i_2, i_3) = I[i_1 > i_2 < i_3 \quad \text{or} \quad i_1 < i_2 > i_3]$$

or Wald and Wolfowitz' rank autocorrelation coefficient of order k (a serial version of Spearman's statistic, here, unlike in Wald and Wolfowitz (1943)'s paper, in a noncircular version) with scores

$$a^{(n)}(i_1, \ldots, i_{k+1}) = i_1 i_{k+1} \ . \tag{3.3}$$

The mean $m^{(n)}$ of (3.2) under the assumption that $\mathbf{Z}^{(n)}$ is white noise is easy to compute:

$$m^{(n)} = [n(n-1)\ldots(n-k)]^{-1} \sum_{i_1 \neq \cdots \neq i_{k+1}} \sum a^{(n)}(i_1, \ldots, i_{k+1}) \ , \tag{3.4}$$

where $\sum_{i_1 \neq \cdots \neq i_{k+1}} \ldots \sum$ stands for a summation running over all $(k+1)$–tuples of *distinct* integers between 1 and n. The variance of (3.2) also is easy to obtain from combinatorial arguments—though a general closed form is rather tedious.

Most serial statistics of practical interest can be decomposed into a linear combination of simpler statistics of the form

$$S^{(n)} = (n-k)^{-1} \sum_{t=k+1}^{n} a^{(n)}(R_t^{(n)}) b^{(n)}(R_{t-k}^{(n)}) \ . \tag{3.5}$$

The runs statistic for example, is of the form $S_1^{(n)} + S_2^{(n)}$, where $S_1^{(n)}$ and $S_2^{(n)}$ rely on the scores

$$a_1^{(n)}(i_1, i_2) = I[2i_1 < n+1].I[2i_2 > n+1]$$

and

$$a_2^{(n)}(i_1, i_2) = I[2i_2 > n+1].I[2i_1 < n+1] \ .$$

A statistic of the form (3.5) will be called a *simple* serial linear rank statistic of order k. The Wald–Wolfowitz autocorrelation coefficient, as well as the f–rank autocorrelation coefficients to be introduced later, are *simple* rank statistics.

Letting

$$S_{\ell m}^{(n)} = \sum_{i=1}^{n} [a^{(n)}(i)]^{\ell} [b^{(n)}(i)]^{m} \ ,$$

we have, for the mean $m^{(n)}$ of the simple statistic $S^{(n)}$ and the variance $(\sigma^{(n)})^2$ of $(n-k)^{\frac{1}{2}} S^{(n)}$ under white noise,

$$m^{(n)} = [n(n-1)]^{-1} [S_{10}^{(n)} S_{01}^{(n)} - S_{11}^{(n)}] \tag{3.6}$$

and

$$(\sigma^{(n)})^2 = [n(n-1)]^{-1} [S_{20}^{(n)} S_{02}^{(n)} - S_{22}^{(n)}] \tag{3.7}$$

$$+ 2\frac{(n-2k)^+}{n-k}[n(n-1)(n-2)]^{-1}[S_{10}^{(n)}S_{01}^{(n)}S_{11}^{(n)} - S_{21}^{(n)}S_{01}^{(n)} - S_{12}^{(n)}S_{10}^{(n)} - (S_{11}^{(n)})^2 + 2S_{22}^{(n)}]$$
$$+ \frac{(n-k)(n-k-1) - 2(n-2k)^+}{n-k}[n(n-1)(n-2)(n-3)]^{-1}$$
$$\times [(S_{10}^{(n)}S_{01}^{(n)})^2 + 2(S_{11}^{(n)})^2 + S_{20}^{(n)}S_{02}^{(n)} - 6S_{22}^{(n)}$$
$$- 4S_{11}^{(n)}S_{10}^{(n)}S_{01}^{(n)} - S_{20}^{(n)}(S_{01}^{(n)})^2 - S_{02}^{(n)}(S_{10}^{(n)})^2 + 4S_{21}^{(n)}S_{01}^{(n)} + 4S_{12}^{(n)}S_{10}^{(n)}]$$
$$- (n-k)(m^{(n)})^2, \qquad \text{where } (\cdot)^+ = \max(.,0).$$

Accordingly, $(n-k)^{\frac{1}{2}}(S^{(n)})/\sigma^{(n)}$ is exactly standardized (still, under the hypothesis that $\mathbf{Z}^{(n)}$ is white noise).

3.2 Serial and nonserial linear signed–rank statistics. As mentioned in Section 1.4, whenever symmetry assumptions can be made, the vector of ranks loses its maximal invariance properties for the benefit of $(\mathbf{s}^{(n)}, \mathbf{R}_+^{(n)})$, where $\mathbf{s}^{(n)} = (s_1, \ldots, s_n)$ is the vector of signs $s_t = \text{sign}(Z_t^{(n)}) = Z_t^{(n)}/|Z_t^{(n)}|$ (with the convention $0/0 = 1$), and $\mathbf{R}_+^{(n)}$ denotes the vector of the ranks $R_{+,t}^{(n)}$ of absolute values $|Z_t^{(n)}|$ among $|Z_1^{(n)}|, \ldots, |Z_n^{(n)}|$. If $\mathbf{Z}^{(n)}$ is absolutely continuous symmetric white noise, then $\mathbf{s}^{(n)}$ and $\mathbf{R}_+^{(n)}$ are independently distributed, $\mathbf{s}^{(n)}$ uniformly over the 2^n elements of $\{-1,1\}^n$, and $\mathbf{R}_+^{(n)}$ uniformly over the $n!$ permutations of $\{1, \ldots, n\}$.

Here again, locally best tests for linear models with independent and identically distributed symmetric error (unspecified density) terms can be based on (nonserial) linear, signed–rank statistics, of the form

$$S_+^{(n)} = n^{-1}\sum_{t=1}^{n} c_t a_+^{(n)}(s_t R_{+,t}^{(n)}) \,, \tag{3.8}$$

where the c_t's are known constants, and $a_+^{(n)}$ denotes a score function, defined over $\{\pm 1, \pm 2, \ldots, \pm n\}$. Time series problems require the more general class of serial linear signed–rank statistics (Hallin and Puri 1991a), of the form

$$S_+^{(n)} = (n-k)^{-1}\sum_{t=k+1}^{n} a_+^{(n)}(s_t R_{+,t}^{(n)}, s_{t-1}R_{+,t-1}^{(n)}, \ldots, s_{t-k}R_{+,t-k}^{(n)}). \tag{3.9}$$

Simple examples are Goodman (1958)'s simplified runs test (with $a_+^{(n)}(i_1, i_2) = I[i_1 i_2 < 0]$), which coincides (up to additive and multiplicative constants) with Dufour (1981)'s runs test, or the signed version of the Spearman–Wald–Wolfowitz serial correlation tests, with $a_+^{(n)}(i_1, \ldots, i_{k+1}) = i_1 i_{k+1}$.

A linear signed–rank statistic will be called *simple* if it can be written as

$$S_+^{(n)} = (n-k)^{-1} \sum_{t=k+1}^{n} s_t s_{t-k} a_+^{(n)}(R_{+,t}^{(n)}) b_+^{(n)}(R_{+,t-k}^{(n)}) \; . \tag{3.10}$$

The signed Spearman–Wald–Wolfowitz autocorrelations, as well as the signed f–rank autocorrelations to be described later, belong to this family of simple serial statistics. The mean of (3.10) under the hypothesis of symmetric white noise obviously is zero, whereas, due to the fact that all cross–products have expectation zero, the variance of $(n-k)^{\frac{1}{2}} S_+^{(n)}$ takes the very simple form

$$\begin{aligned} (\sigma_+^{(n)})^2 &= [n(n-1)]^{-1} \sum \sum_{i_1 \neq i_2} (a_+^{(n)}(i_1) b_+^{(n)}(i_2))^2 \\ &= [n(n-1)]^{-1} \left\{ \sum_i [a_+^{(n)}(i)]^2 \sum_i [b_+^{(n)}(i)]^2 - \sum_i [a_+^{(n)}(i) b_+^{(n)}(i)]^2 \right\} \; . \end{aligned} \tag{3.11}$$

Here again, $(n-k)^{\frac{1}{2}} S_+^{(n)} / \sigma_+^{(n)}$ is exactly standardized (provided that $\mathbf{Z}^{(n)}$ is symmetric white noise).

3.3 Asymptotic normality. Assuming that the series $\mathbf{Z}^{(n)}$ from which the ranks are taken is white noise, consider the serial statistic $S^{(n)}$ in (3.2). A real–valued function J, defined over the $(k+1)$-dimensional unit square $(0,1)^{k+1}$, will be called a *score-generating function* for $S^{(n)}$ if

$$\int_{(0,1)^{k+1}} J^{2+\delta}(u_1, \ldots, u_{k+1}) du_1 \ldots du_{k+1} < \infty \tag{3.12}$$

for some $\delta > 0$ and

$$\lim_{n \to \infty} E\left[(a^{(n)}(R_{U,1}^{(n)}, \ldots, R_{U,k+1}^{(n)}) - J(U_1, \ldots, U_{k+1}))^2 \right] = 0 \; , \tag{3.13}$$

where $R_{U,1}^{(n)} \ldots, R_{U,k+1}^{(n)}$ denote the ranks of $U_1, \ldots U_{k+1}$ in a random sample $U_1, \ldots, U_n$ of independent and identically distributed, uniform (over [0,1]) random variables.

Define

$$\begin{aligned} J^*(u_1, \ldots, u_{k+1}) = J(u_1, \ldots, u_{k+1}) &- \sum_{i=1}^{k+1} E[J(U_1, \ldots, U_{i-1}, u_1, U_i, \ldots, U_k)] \\ &+ k \; E[J(U_1, \ldots, U_{k+1})] \; . \end{aligned} \tag{3.14}$$

Then, relying on a central limit theorem (e.g., Yoshihara, 1976) for U–statistics under absolutely regular processes, the following asymptotic normality result has been proved (cf. Hallin et al., 1985).

PROPOSITION 3.1. *Let J denote a score–generating function for the serial rank statistic $S^{(n)}$ in (3.2). Then $(n-k)^{\frac{1}{2}}(S^{(n)} - m^{(n)})$ is asymptotically normal (under white noise), with mean zero and variance V^2, where*

$$V^2 = \int_{[0,1]^{k+1}} [J^*(u_1, \ldots, u_{k+1})]^2 du_1 \ldots du_{k+1} + 2\sum_{i=1}^{k} \int_{[0,1]^{k+i+1}} J^*(u_1, \ldots, u_{k+1}) J^*(u_{1+i}, \ldots, u_{k+1+i}) du_1 \ldots du_{k+1+i} \; . \tag{3.15}$$

This result allows for explicit normal approximations, under the null hypotheses of white noise, for P–values and critical points of tests based on linear serial rank statistics.

Example 3.1. The Spearman autocorrelation coefficient of order one is often defined (cf. e.g. Kendall and Stuart, 1968) as

$$r_S^{(n)} = \frac{(n-1)^{-1}\Sigma_{t=2}^{n} R_t^{(n)} R_{t-1}^{(n)} - (n+1)^2/4}{(n^2-1)/12} \; ,$$

though an exactly standardized version (see Hallin and Mélard 1988) is more convenient for practical purposes. Letting $S^{(n)} = (n-1)^{-1} \sum_{t=2}^{n} R_t^{(n)} R_{t-1}^{(n)}/(n+1)^2$, with $m^{(n)} = (3n+2)/12(n+1)$, it is easy to check that

$$r_S^{(n)} - 12(S^{(n)} - m^{(n)}) = 24(n-1)^{-1}(S^{(n)} - m^{(n)}) - (n-1)^{-1} \; . \tag{3.16}$$

$S^{(n)}$ is a linear serial rank statistic of order one. A score–generating function for $S^{(n)}$ is $J(u,v) = uv$, yielding

$$J^*(u,v) = uv - u + \frac{1}{4} \; .$$

It follows from Proposition 3.1 that $(n-1)^{\frac{1}{2}}(S^{(n)} - m^{(n)})$ is asymptotically normal, with mean zero and variance

$$V^2 = \int_{[0,1]^2} \left(uv - u + \frac{1}{4}\right)^2 dudv = 1/144 \; .$$

This, with (3.16), entails that $r_S^{(n)} - 12(S^{(n)} - m^{(n)})$ is $o_P(n^{-\frac{1}{2}})$, as $n \to \infty$, and confirms the classical result that $(n-1)^{\frac{1}{2}} r_S^{(n)}$ (as well as $12(n-1)^{\frac{1}{2}}(S^{(n)} - m^{(n)})$) is asymptotically standard normal.

As in the nonserial case, (3.13) can be shown to hold for J satisfying (3.12) and scores defined by

$$a^{(n)}(i_1,\ldots,i_{k+1}) = E\Big[J(U_1,\ldots,U_{k+1})\big|R_1^{(n)}=i_1,\ldots,R_{k+1}^{(n)}=i_{k+1}\Big]$$

(*exact scores*) or, for J monotone with respect to each argument, satisfying (3.12), and scores

$$a^{(n)}(i_1,\ldots,i_{k+1}) = J\left(\frac{i_1}{n+1},\ldots,\frac{i_{k+1}}{n+1}\right)$$

(*approximate scores*).

The case of signed–rank statistics is roughly similar. A *score–generating function* J^+ for $S_+^{(n)} = (n-k)^{-1}\Sigma a_+^{(n)}(s_tR_t^{(n)},\ldots,s_{t-k}R_{t-k}^{(n)})$ is a real–valued function, defined over the $(k+1)$–dimensional open square $(-1,1)^{k+1}$, such that, denoting by $V_1,\ldots,V_n$ an n–tuple of independent and identically distributed rectangular $[-1,1]$ variables,

$$E[|J_+(V_1,V_2,\ldots,V_{k+1})|^{2+\delta}] < \infty \tag{3.17}$$

for some $\delta > 0$, and

$$\lim_{n\to\infty} E\{[a_+^{(n)}(\mathrm{sgn}(V_1)R_{+,1}^{(n)},\ldots,\mathrm{sgn}(V_{k+1})R_{+,k+1}^{(n)}) - J_+(V_1,\ldots,V_{k+1})]^2\} = 0\ , \tag{3.18}$$

as $n\to\infty$. $R_{+,i}^{(n)}$ here denotes the rank of $|V_i|$ among $|V_1|,\ldots,|V_n|$; the notation $\mathrm{sgn}(V_i)$ is used in an obvious fashion. Associated with J_+, define J_+^* as

$$\begin{aligned} J_+^*(v_1,\ldots,v_{k+1}) &= J_+(v_1,\ldots,v_{k+1}) \\ &\quad - 2^{-(k+1)}\sum_{s\in\{-1,1\}}\sum_{\ell=0}^{k}\int_{[-1,1]^k} J_+(w_1,\ldots,w_\ell,sv_1,w_{\ell+1},\ldots,w_k)d\mathbf{w} \\ &\quad + 2^{-(k+1)}k\int_{[-1,1]^k} J_+(w_1,w_2,\ldots,w_{k+1})d\mathbf{w}\ , \end{aligned} \tag{3.19}$$

where $\int_{[-1,1]^k}(\ldots)d\mathbf{v}$ stands for $\int_{-1}^{1}\ldots\int_{-1}^{1}(\ldots)\,\mathrm{sgn}(v_1)\ldots\mathrm{sgn}(v_k)dv_1\ldots dv_k$.

PROPOSITION 3.2. *Let J_+ denote a score–generating function for the signed–rank statistic $S_+^{(n)}$ in (3.9). Then $(n-k)^{\frac{1}{2}}S_+^{(n)}$ is asymptotically normal (under symmetric white noise), with mean zero and variance V_+^2, where (same notation as in (3.18))*

$$V_+^2 = E[(J_+^*(V_1,\ldots,V_{k+1}))^2] + 2\sum_{i=1}^{k} E[J_+^*(V_1,\ldots,V_{k+1})J_+^*(V_{1+i},\ldots,V_{k+1+i})]. \tag{3.20}$$

Example 3.2. Parallel to the unsigned statistic considered in Example 3.1, a signed version of the Spearman autocorrelation coefficient of order one can be defined as

$$r_{+,S}^{(n)} = \frac{(n-1)^{-1}\Sigma_{t=2}^{n} s_t s_{t-1} R_{+,t}^{(n)} R_{+,t-1}^{(n)}}{(n+1)(2n+1)/6}$$

(though an exactly standardized version might be preferable for practical purposes—see Hallin et al. 1990). Letting

$$S_{+}^{(n)} = (n-1)^{-1}\Sigma_{t=2}^{n} s_t s_{t-1} R_{+,t}^{(n)} R_{+,t-1}^{(n)}/(n+1)^2,$$

we have $r_{+,S}^{(n)} = 6S_{+}^{(n)}(n+1)^2/(n+1)(2n+1)$. A score–generating function for $S_{+}^{(n)}$ is $J_{+}(u,v) = uv$. Since $E[V_1 V_2 \big| |V_\ell| = |v|] = 0, \ell = 1, 2, v \in [0,1]$ and $E[V_1 V_2] = 0$, $J_{+}^{*}(u,v) = J_{+}(u,v)$ (under symmetric white noise), and

$$V_{+}^{2} = \frac{1}{4} \int_{[-1,1]^2} (uv)^2 du\ dv = 1/9\ .$$

It follows that $(n-1)^{\frac{1}{2}} S^{(n)}$ is asymptotically normal, with mean zero and variance 1/9, and that

$$r_{+,S}^{(n)} - 3S_{+}^{(n)} = S_{+}^{(n)} \left[6\frac{(n+1)^2}{(n+1)(2n+1)} - 3 \right] = o_P(n^{-\frac{1}{2}})\ .$$

Consequently, $(n-1)^{\frac{1}{2}} r_{+,S}^{(n)}$ is asymptotically standard normal (under symmetric white noise). This statistic, to the best of our knowledge, never had been considered in the literature. Its asymptotic performance is investigated in Hallin et al. (1990).

As in the unsigned case, (3.18) holds, for $(2+\delta)$-integrable score-generating functions J_{+} and *exact scores*

$$a_{+}^{(n)}(i_1, \ldots, i_{k+1}) = E\left[J_{+}(V_1, \ldots, V_{k+1}) \big| \operatorname{sgn}(V_1) R_{+,1}^{(n)} = i_1, \ldots, \operatorname{sgn}(V_{k+1}) R_{+,k+1}^{(n)} \right],$$

or, provided in addition that J_{+} is monotone with respect to each argument, with *approximate scores*

$$a_{+}^{(n)}(i_1, \ldots, i_{k+1}) = J_{+}\left(\frac{i_1}{n+1}, \ldots, \frac{i_{k+1}}{n+1} \right)\ .$$

§4. LOCAL ASYMPTOTIC NORMALITY OF ARMA PROCESSES AND THE LOCAL SUFFICIENCY OF RANKS

4.1 Local asymptotic normality results for ARMA models. The asymptotic results of Section 3 are valid under white noise assumptions, and thus allow for constructing invariant, distribution–free tests based on signed or unsigned serial linear rank statistics, for a variety of time series problems. If, however, power or optimality issues are to be

addressed, more information is needed on the structure of the specific problem at hand. Since uniformly most powerful tests cannot be expected to exist—such strong optimality results in general are not available, in time series analysis, even in a more restricted parametric, Gaussian context—weaker optimality properties have to be considered. These weaker properties, of a local and asymptotic nature, rely on the local structure of the families of likelihood functions involved.

This local structure has been studied by several authors, who under various technical assumptions have established the locally asymptotically normal (LAN) structure (LeCam, 1960) of families of ARMA likelihood functions. Davies (1973) and Dzhaparidze (1986) investigate the LAN property for Gaussian ARMA processes. Akritas and Johnson (1982) deal with AR processes only. Hallin et al. (1985) and Hallin and Puri (1987) established a LAN result for general ARMA processes, using classical, Cramér-type, technical assumptions. Swensen (1985) considers the case of AR processes with a regression trend, and Garel (1989) that of MA ones, still with trend. Kreiss (1987) deals with ARMA processes without trend, and Kreiss (1990b) with AR(∞) process. The results of Swensen and Kreiss rely on martingale central limit theorems, and those by Akritas and Johnson on quadratic mean differentiability conditions which are less restrictive than the Cramér–type assumptions made by Hallin et al. (1985), Hallin and Puri (1987) and Garel (1989).

We shall not attempt here to describe in detail the technical conditions under which the LAN property holds. The assumptions we are giving here are not the weakest; but they are quite simple, and they are satisfied by most densities used in practice (not all of them: e.g. Cauchy).

4.2 Notation and main assumptions. All densities below—denoted by f, g, with distribution functions F, G—are of the form

$$g(x) = g_\sigma(x) = \sigma^{-1} g_1(x/\sigma) > 0 \ , \quad x \in \mathbb{R}$$

with

$$\int x \ g_1(x)dx = 0 \quad \text{and} \quad \int x^2 g_1(x)dx = 1 \ .$$

The variance σ^2 will always remain unspecified; since however it plays no role in the sequel, it will be dropped in the notation. Still for simplicity, we shall assume that g is absolutely continuous, so that the derivative $\dot{g}(x) = dg(x)/dx$ exists for almost all x. Defining $\varphi_g(x) = -\dot{g}(x)/g(x)$, we also assume that the Fisher information

$$\int \varphi_g^2(x) g(x) dx = \sigma^2 I(g) = I(g_1) = \int \varphi_{g_1}^2(x) g_1(x) dx$$

is finite. Since $g(x)$ is always strictly positive, G is strictly increasing, and the inverse G^{-1} is well defined.

Consider the stochastic difference equation (ARMA (p_1, q_1) model)

$$X_t - \sum_{i=1}^{p_1} A_i X_{t-i} = \varepsilon_t + \sum_{i=1}^{q_1} B_i \varepsilon_{t-i} \ , \tag{4.1}$$

in short $A(L)X_t = B(L)\varepsilon_t$, where L stands for the lag operator and $A(z) = 1 - \Sigma A_i z^i$, $B(z) = 1 + \Sigma B_i z^i$. In all ARMA models below, it is always assumed that $A(z)$ and $B(z)$, $z \in \mathbb{C}$ have no common roots, and satisfy the usual invertibility and causality conditions. We denote by $H_g^{(n)}(A, B)$ the hypothesis under which an observed series $\mathbf{X}^{(n)} = (X_1^{(n)}, \ldots, X_t^{(n)}, \ldots, X_n^{(n)})$ is a finite realization of some solution of (4.1), where $\{\varepsilon_t\}$ is a white noise process (i.e., an iid process) with density g. The notation $H^{(n)}(A, B) = \underset{g}{\cup} H_g^{(n)}(A, B)$ is used whenever g remains completely unspecified (except for the few general technical conditions above); the notation $H_+^{(n)}(A, B) = \underset{g}{\cup}\{H_g^{(n)}(A, B) |\ g$ symmetric with respect to zero$\}$ is used whenever g is symmetric but otherwise unspecified (still, apart from the few general technical conditions).

4.3 Signed and unsigned f–rank autocorrelations. Just as parametric autocorrelation coefficients play an essential role in the classical parametric analysis of time series models, a special case of serial rank statistics will play an essential role in our nonparametric approach. These nonparametric counterparts of usual autocorrelation coefficients have been introduced in Hallin et al. (1987) as *f–rank autocorrelations* and in Hallin and Puri (1991a) as *signed f–rank autocorrelations*.

Denote by $f, \varphi = -\dot{f}/f$ and F a probability density function, the corresponding score function and distribution function, respectively. The f–rank and signed f–rank autocorrelations (see below for a definition) being invariant with respect to scale transformations, f can be chosen as f_1, so that $\int x^2 f(x)dx = 1$. The (unsigned) f–rank autocorrelation of lag i is then defined as

$$r_{i;f}^{(n)} = (n-i)^{-1} \left[\sum_{t=i+1}^{n} \varphi\left(F^{-1}\left(\frac{R_t^{(n)}}{n+1} \right) \right) F^{-1}\left(\frac{R_{t-i}^{(n)}}{n+1} \right) - m^{(n)} \right] / s^{(n)} \ , \tag{4.2}$$

where $m^{(n)}$ and $s^{(n)}$ are given by (3.6) and (3.7) with

$$S_{\ell m}^{(n)} = \sum_{i=1}^{n} \varphi\left(F^{-1}\left(\frac{i}{n+1} \right) \right)^{\ell} \left(F^{-1}\left(\frac{i}{n+1} \right) \right)^{m} .$$

Particular cases are

(a) the *van der Waerden autocorrelation coefficients*, associated with Gaussian densities,

$$r_{i;vdW}^{(n)} = \left[(n-i)^{-1} \sum_{t=i+1}^{n} \mathbf{\Phi}^{-1}\left(\frac{R_t^{(n)}}{n+1} \right) \mathbf{\Phi}^{-1}\left(\frac{R_{t-i}^{(n)}}{n+1} \right) - m_{vdW}^{(n)} \right] / s_{vdW}^{(n)}$$

(as usual $\Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} e^{-y^2/2} dy$ denotes the standard normal distribution function; note that, due to the symmetry of the Gaussian distribution, quite a number of terms in (3.6) and (3.7) cancel out),

(b) the *Wilcoxon autocorrelation coefficients*, associated with logistic densities,

$$r_{i;W}^{(n)} = \left[(n-i)^{-1} \sum_{t=i+1}^{n} \left(\frac{R_t^{(n)}}{n+1} - \frac{1}{2} \right) \log \frac{R_{t-i}^{(n)}}{n+1-R_{t-i}^{(n)}} - m_W^{(n)} \right] / s_W^{(n)} ,$$

(c) the *Laplace autocorrelation coefficients*, associated with double exponential densities,

$$r_{i;L}^{(n)} = \left[(n-i)^{-1} \sum_{t=i+1}^{n} \operatorname{sgn} \left(\frac{R_t^{(n)}}{n+1} - \frac{1}{2} \right) \left[\log(2 \frac{R_{t-i}^{(n)}}{n+1}) I \left[R_t^{(n)} \geq \frac{n+1}{2} \right] \right. \right.$$
$$\left. \left. - \log \left(2 - 2 \frac{R_{t-i}^{(n)}}{n+1} \right) I \left[R_t^{(n)} < \frac{n+1}{2} \right] \right] - m_L^{(n)} \right] / s_L^{(n)} .$$

Note that only the van der Waerden statistic can be considered, *stricto sensu*, as an autocorrelation coefficient. Actually, up to normalizing constants, $r_{i;vdW}^{(n)}$ is the classical autocorrelation coefficient resulting from substituting the transformed observations $(\Phi^{-1}(R_t^{(n)}/n+1))$ for the original one ($X_t^{(n)}$ or $Z_t^{(n)}$). The symmetric structure of this measure of serial dependence, where the past and future play exchangeable roles, is a consequence of the fact that (except for a few, very particular MA cases) Gaussian ARMA processes are the only time–reversible ones (Weiss, 1975; Hallin et al., 1988). The past and future, in non Gaussian processes, do not play symmetric roles; accordingly, they do not play symmetric roles in the definition of f–rank autocorrelations either.

The importance of f–rank autocorrelations will follow from the asymptotic decomposition of log likelihood ratios (section 4.4 below).

Under symmetric densities, a signed version of f–rank autocorrelations, involving signed ranks, can be used with the same notation as above. Letting $F_+ = 2F - 1$, define the signed f–rank autocorrelation of lag i as

$$(4.3)\quad r_{i;f}^{(n)+} = (n-i)^{-1} \sum_{t=i+1}^{n} \operatorname{sgn}(Z_t^{(n)} Z_{t-i}^{(n)}) \varphi \left(F_+^{-1} \left(\frac{R_{+,t}^{(n)}}{n+1} \right) \right) F_+^{-1} \left(\frac{R_{+,t-i}^{(n)}}{n+1} \right) / s_{+,f}^{(n)}$$

where $\{Z_t^{(n)}\}$ is the series from which the ranks are computed, and

$$(s_{+,f}^{(n)})^2 = [n(n-1)]^{-1} \left\{ \sum_{i=1}^{n} \varphi \left(F_+^{-1} \left(\frac{i}{n+1} \right) \right)^2 \sum_{i=1}^{n} \left(F_+^{-1} \left(\frac{i}{n+1} \right) \right)^2 \right.$$
$$\left. - \sum_{i=1}^{n} \left[\varphi \left(F_+^{-1} \left(\frac{i}{n+1} \right) \right) F_+^{-1} \left(\frac{i}{n+1} \right) \right]^2 \right\}.$$

Here again, provided the $Z_t^{(n)}$'s are symmetric white noise, $(n-i)^{\frac{1}{2}} r_{i;f}^{(n)+}$ is exactly standardized. Particular cases are

(d) the *signed van der Waerden autocorrelation coefficients*

$$r_{i;vdW}^{(n)+} = (n-i)^{-1} \sum_{t=i+1}^{n} \operatorname{sgn}(Z_t^{(n)} Z_{t-i}^{(n)}) \Phi^{-1}\left(\frac{n+1+R_{+,t}^{(n)}}{2(n+1)}\right) \Phi^{-1}\left(\frac{n+1+R_{+,t-i}^{(n)}}{2(n+1)}\right)$$
$$\times [n(n-1)]^{\frac{1}{2}} \left\{ \left[\sum_{i=1}^{n} (\Phi^{-1}(i/(n+1)))^2\right]^2 - \sum_{i=1}^{n} (\Phi^{-1}(i/(n+1)))^4 \right\}^{-\frac{1}{2}},$$

(e) the *signed Wilcoxon autocorrelation coefficients*

$$r_{i;W}^{(n)+} = (n-i)^{-1} \sum_{t=i+1}^{n} \operatorname{sgn}(Z_t^{(n)} Z_{t-i}^{(n)}) R_{+,t}^{(n)} \log\left(\frac{n+1+R_{+,t-i}^{(n)}}{n+1-R_{+,t-i}^{(n)}}\right) / s_{+,W}^{(n)},$$

(f) the *signed Laplace autocorrelation coefficients*

$$r_{i;L}^{(n)+} = -(n-i)^{-1} \sum_{t=i+1}^{n} \operatorname{sgn}(Z_t^{(n)} Z_{t-i}^{(n)}) \log\left(1 - \frac{R_{+,t-i}^{(n)}}{n+1}\right) \Big/ \left\{ n^{-1} \sum_{i=1}^{n} \left[\log\left(\frac{i}{n+1}\right)\right]^2 \right\}^{\frac{1}{2}}.$$

A signed van der Waerden autocorrelation is thus the usual autocorrelation computed from the series of "signed standard normal quantiles" $\operatorname{sgn}(Z_t^{(n)}) \Phi^{-1}\left(\frac{1}{2} + R_t^{(n)}/2(n+1)\right)$ associated with the residual series $\mathbf{Z}^{(n)}$. Wilcoxon and Laplace autocorrelations constitute weighted versions of the traditional Wilcoxon signed rank and runs test statistics respectively.

The following asymptotic results can be established for all f–rank autocorrelations. If $H^{(n)}(A,B)$ (resp. $H_+^{(n)}(A,B)$) is the hypothesis of interest, the ranks (resp. signed ranks) to be used are those of the residuals $Z_t^{(n)} = [A(L)/B(L)]X_t^{(n)}$ (the "starting values" used in the inversion of $B(L)$ can be chosen arbitrarily.)

PROPOSITION 4.1. (Hallin et al., 1987; Hallin and Puri, 1991).

(i) Under $H^{(n)}(A,B)$, $(n-i)^{\frac{1}{2}} r_{i;f}^{(n)}$ *and* $(n-j)^{\frac{1}{2}} r_{j;f}^{(n)}$, $i \neq j$, *are asymptotically jointly normal, with mean zero and unit covariance matrix.*

(ii) Under $H_+^{(n)}(A,B)$, $(r_{i;f}^{(n)+} - r_{i;f}^{(n)})$ *is* $o_P(n^{-\frac{1}{2}})$; *accordingly,* $(n-i)^{\frac{1}{2}} r_{i;f}^{(n)+}$ *and* $(n-j)^{\frac{1}{2}} r_{j;f}^{(n)+}$, $i \neq j$ *also are asymptotically jointly normal, with mean zero and unit covariance matrix.*

4.4 Local asymptotic normality. Let $p_2 \geq p_1$, $q_2 \geq q_1$, $\mathbf{A} = (A_1 \dots A_{p_1} 0 \dots 0)' \in \mathbb{R}^{p_2}$, $\mathbf{B} = (B_1 \dots B_{q_1} 0 \dots 0)' \in \mathbb{R}^{q_2}$, $\boldsymbol{\theta} = (\mathbf{A}', \mathbf{B}')' \in \mathbb{R}^{p_2+q_2}$. Denote by $\boldsymbol{\Theta}$ the (open) subset of $\mathbb{R}^{p_2+q_2}$ for which (4.1) constitutes a valid, causal and invertible, ARMA model of orders p_1 and q_1 (i.e. $A_{p_1} \neq 0 \neq B_{q_1}$, no common roots, causality and invertibility). Similarly, let $\boldsymbol{\gamma} \in \mathbb{R}^{p_2}$, $\boldsymbol{\delta} = \mathbb{R}^{q_2}$, $\boldsymbol{\tau} = (\boldsymbol{\gamma}', \boldsymbol{\delta}')' \in \mathbb{R}^{p_2+q_2}$. $H_g^{(n)}(\boldsymbol{\theta} + n^{-\frac{1}{2}}\boldsymbol{\tau})$ accordingly constitutes a sequence of hypotheses, approaching, in some sense to be made clearer in the sequel, to $H_g^{(n)}(\boldsymbol{\theta})$. Under this sequence $\mathbf{X}^{(n)}$ is generated by some ARMA (p, q) model, $p_1 \leq p \leq p_2$, $q_1 \leq q \leq q_2$, with coefficients of the form $\mathbf{A} + n^{-\frac{1}{2}}\boldsymbol{\gamma}, \mathbf{B} + n^{-\frac{1}{2}}\boldsymbol{\delta}$ and innovation density g.

Denote by $L_{\boldsymbol{\theta};g}^{(n)}$ the likelihood function of $\mathbf{X}^{(n)}$ under $H_g^{(n)}(\boldsymbol{\theta})$, hence by $L_{\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau};g}^{(n)}$ the likelihood of $\mathbf{X}^{(n)}$ under $H_g^{(n)}(\boldsymbol{\theta} + n^{-\frac{1}{2}}\boldsymbol{\tau})$. Define the random variable

$$\Lambda_{\boldsymbol{\theta};\boldsymbol{\tau};g}^{(n)}(\mathbf{X}^{(n)}) = \log\left[L_{\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau};g}^{(n)}(\mathbf{X}^{(n)})/L_{\boldsymbol{\theta};g}^{(n)}(\mathbf{X}^{(n)})\right]$$

(whenever $L_{\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau};g}^{(n)}$ and $L_{\boldsymbol{\theta};g}^{(n)}$ are zero, Λ can be left arbitrary).

Denote by g_u and h_u the Green's functions associated with the difference operators $A(L)$ and $B(L)$ respectively (i.e. characterized by

$$\sum_{u=0}^{\infty} g_u L^u = [A(L)]^{-1} \quad \text{and} \quad \sum_{u=0}^{\infty} h_u L^u = [B(L)]^{-1}) \ .$$

Let

$$a_i = \sum_{j=1}^{\min(p_2, i+p_1-1)} \gamma_j g_{i-j} \tag{4.4a}$$

and

$$b_i = \sum_{j=1}^{\min(q_2, i+q_1-1)} \delta_j h_{i-j}, \tag{4.4b}$$

so that

$$a(L) = \sum_{i=1}^{\infty} a_i L^i = \left[\sum_{i=1}^{p_2} \gamma_i L^i\right] / A(L)$$

and

$$b(L) = \sum_{i=1}^{\infty} b_i L^i = \left(\sum_{i=1}^{q_2} \delta_i L^i\right) / B(L).$$

It follows from the causality and invertibility properties of (4.1) that the sequences $\mathbf{a} = (a_i)$ and $\mathbf{b} = (b_i)$ are absolutely summable: denote by $\|\mathbf{a}+\mathbf{b}\|$ their ℓ^2 norm $\left[\sum_{i=1}^{\infty}(a_i+b_i)^2\right]^{\frac{1}{2}}$. Also, letting $\pi = \max(p_2 - p_1, q_2 - q_1)$, introduce the $(\pi + p_1 + q_1) \times (p_2 + q_2)$ matrix

$$(4.5)\quad \mathbf{M}(\boldsymbol{\theta}) = \begin{pmatrix} 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ g_1 & 1 & & & h_1 & 1 & & \\ \vdots & & \ddots & \vdots & \vdots & & \ddots & \vdots \\ g_{p_2-1} & & \dots & 1 & h_{q_2-1} & & \dots & 1 \\ g_{p_2} & & \dots & g_1 & h_{q_2} & & \dots & h_1 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ g_{\pi+p_1+q_1-1} & & \cdots & g_{\pi+p_1+q_1-p_2} & h_{\pi+p_1+q_1-1} & & \cdots & h_{\pi+p_1+q_1-q_2} \end{pmatrix}.$$

Finally, denote by $\{\psi_t^{(1)} \dots \psi_t^{(p_1+q_1)}\}$ an arbitrary fundamental system of solutions of the homogeneous equation (or order $p_1 + q_1$) $A(L)B(L)\psi_t = 0$, $t \in \mathbf{Z}$; by $\mathbf{C}_\psi$ the Casorati matrix

$$(4.6)\quad \mathbf{C}_\psi = \begin{pmatrix} \psi_{\pi+1}^{(1)} & \dots & \psi_{\pi+1}^{(p_1+q_1)} \\ \psi_{\pi+2}^{(1)} & \dots & \psi_{\pi+2}^{(p_1+q_1)} \\ \vdots & & \vdots \\ \psi_{\pi+p_1+q_1}^{(1)} & \dots & \psi_{\pi+p_1+q_1}^{(p_1+q_1)} \end{pmatrix}$$

and by $n^{\frac{1}{2}}\mathbf{T}_{\psi;g}^{(n)}$ the vector of rank statistics

$$(4.7)\quad n^{\frac{1}{2}}\mathbf{T}_{\psi;g}^{(n)} = \begin{pmatrix} (n-1)^{\frac{1}{2}}\ r_{1;g}^{(n)} \\ \vdots \\ (n-\pi)^{\frac{1}{2}}\ r_{\pi;g}^{(n)} \\ \sum_{i=\pi+1}^{n-1} (n-i)^{\frac{1}{2}}\psi_i^{(1)}\ r_{i;g}^{(n)} \\ \vdots \\ \sum_{i=\pi+1}^{n-1} (n-i)^{\frac{1}{2}}\ \psi_i^{(p_1+q_1)}\ r_{i;g}^{(n)} \end{pmatrix},$$

where the ranks are those of the residuals $Z_t^{(n)} = [A(L)/B(L)]X_t^{(n)}$. The covariance matrix of $n^{\frac{1}{2}}\mathbf{T}_{\psi;g}^{(n)}$ under $H^{(n)}(A,B)$ is

$$\mathbf{W}_{\psi}^{(n)} = \left(\begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{w}_{\psi}^{(n)} \end{array}\right), \tag{4.8}$$

with $$w_{\psi;k\ell}^{(n)} = \sum_{i=\pi+1}^{n-1} \psi_i^{(k)}\psi_i^{(\ell)} \ , \ k,\ell = 1,\ldots,p_1+q_1 \ ,$$

which converges, as $n \to \infty$, to

$$\mathbf{W}_{\psi} = \left(\begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{w}_{\psi} \end{array}\right), \tag{4.9}$$

with $w_{\psi,k\ell} = \sum_{i=\pi+1}^{\infty} \psi_i^{(k)}\psi_i^{(\ell)}$ (convergence again follows from the causality and invertibility of $A(L)$ and $B(L)$).

We may now state the main LAN result.

PROPOSITION 4.2. *For each $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ (and under mild technical assumptions on g)*

(i) $$\Lambda_{\boldsymbol{\theta};\boldsymbol{\tau}^{(n)};g}^{(n)}(\mathbf{X}^{(n)}) = n^{\frac{1}{2}}(\boldsymbol{\tau}^{(n)})'\mathbf{M}'(\boldsymbol{\theta}) \left(\begin{array}{c|c} \mathbf{I}_{\pi\times\pi} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{C}_{\psi}'^{-1} \end{array}\right) \mathbf{T}_{\psi;g}^{(n)}(I(g_1))^{\frac{1}{2}}$$

$$-\frac{1}{2}\,(\boldsymbol{\tau}^{(n)})'\mathbf{M}'(\boldsymbol{\theta}) \left(\begin{array}{c|c} \mathbf{I}_{\pi\times\pi} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{C}_{\psi}'^{-1}\mathbf{w}_{\psi}\mathbf{C}_{\psi}^{-1} \end{array}\right) \mathbf{M}(\boldsymbol{\theta})\boldsymbol{\tau}^{(n)}\ I(g_1) + o_P(1) \ ,$$

under $H_g^{(n)}(\boldsymbol{\theta})$, as $n \to \infty$, for $\boldsymbol{\tau}^{(n)}$ such that $\sup_n (\boldsymbol{\tau}^{(n)})'\boldsymbol{\tau}^{(n)} < \infty$.

(ii) *for all $\boldsymbol{\tau} \in \mathbb{R}^{p_2+q_2}$, $n^{\frac{1}{2}}\boldsymbol{\tau}'\mathbf{M}'(\boldsymbol{\theta}) \left(\begin{array}{c|c} \mathbf{I}_{\pi\times\pi} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{C}_{\psi}'^{-1} \end{array}\right) \mathbf{T}_{\psi;g}^{(n)}$ is asymptotically normal, under $H^{(n)}(\boldsymbol{\theta})$, as $n \to \infty$, with mean zero and variance*

$$\boldsymbol{\tau}'\mathbf{M}'(\boldsymbol{\theta}) \left(\begin{array}{c|c} \mathbf{I}_{\pi\times\pi} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{C}_{\psi}'^{-1}\mathbf{w}_{\psi}\mathbf{C}_{\psi}^{-1} \end{array}\right) \mathbf{M}(\boldsymbol{\theta})\boldsymbol{\tau}\ I(g_1) = \|\mathbf{a}+\mathbf{b}\|^2 I(g_1).$$

(iii) *Denote by* $r_{i;g}^{(n)}(\boldsymbol{\tau})$ *the* f*-rank autocorrelations computed from the residuals*

$$Z_t^{(n)}(\boldsymbol{\tau}) = \left[\sum_{i=1}^{p_2}(A_i + n^{-\frac{1}{2}}\gamma_i)L^i / \sum_{i=1}^{q_2}(B_i + n^{-\frac{1}{2}}\delta_i)L^i\right] X_t^{(n)} \ .$$

Then, under $H_f^{(n)}(\boldsymbol{\theta})$, as $n \to \infty$,

$$\begin{aligned}(4.10)\quad & n^{\frac{1}{2}}(r_{i;g}^{(n)}(\boldsymbol{\tau}) - r_{i,g}^{(n)}) = \\ & -(a_i+b_i)\left\{\int_0^1 \varphi_f(F^{-1}(u))\varphi_g(G^{-1}(u))du \int_0^1 F^{-1}(u)G^{-1}(u)du\right\} I(g_1)^{-\frac{1}{2}} + o_P(1) \\ & = -(a_i+b_i)I(g|f) + o_P(1) \\ & \text{with } I(g|f) = \left\{\int_0^1 \varphi_f(F^{-1}(u))\varphi_g(G^{-1}(u))du \int F^{-1}(u)G^{-1}(u)du\right\} I(g_1)^{-\frac{1}{2}} \ .\end{aligned}$$

It follows from (i) and (ii) that $\Lambda_{\boldsymbol{\theta};\boldsymbol{\tau};g}^{(n)}(\mathbf{X}^{(n)})$ is asymptotically normal, under $H_g^{(n)}(\boldsymbol{\theta})$, with mean $-\frac{1}{2}\|\mathbf{a}+\mathbf{b}\|^2\mathbf{I}(g_1)$ and variance $\|\mathbf{a}+\mathbf{b}\|^2 I(g_1)$.

The family of likelihoods $\{L_{\boldsymbol{\theta};g},\ \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ is thus locally asymptotically normal (more precisely, the family $\{L_{\boldsymbol{\theta};g}, \boldsymbol{\theta} \in \mathbb{R}^{p_2+q_2}\}$ is *restricted locally asymptotically normal* for $\boldsymbol{\theta} \in \boldsymbol{\Theta}$) in the sense of LeCam (1960)'s conditions (DN1 to DN6), and $\mathbf{T}_{\psi;f}^{(n)}$—hence the corresponding f–rank autocorrelations, or the ranks themselves are locally sufficient. Note that the dimension of the locally sufficient statistic $\mathbf{T}_{\psi;f}^{(n)}$ is $\pi+p_1+q_1 = \max(p_1+q_2, p_2+q_1)$, whereas the dimension of the parameter space is p_2+q_2. This corresponds to the well-known fact that the information matrix of an ARMA model is singular.

In view of Proposition 4.1, Proposition 4.2 remains valid, under symmetric densities, if signed ranks and signed rank autocorrelations are substituted for the unsigned ones. This latter fact is of particular interest.

The type of ranks (signed or unsigned) to be adopted—if rank–based techniques are to be considered—indeed depends, as explained in Section 1, on the invariance features of the testing problem at hand. Now, the effectiveness of the choice between signed unsigned ranks, in the classical symmetric i.i.d. case, is obscured, in practice, by the fact that (unsigned) ranks are totally insensitive to a variety of alternatives. Testing the slope of a regression line e.g. can be achieved (in a strictly unbiased manner) by means of either unsigned or signed rank techniques, whereas testing the intercept using unsigned ranks is impossible. A highly perverse consequence of this fact is that the type of alternative at hand in most cases apparently dictates which type of ranks (unsigned ones for the slope, signed ones for the intercept) should be adopted, and thus which invariance argument is relevant—see e.g. Puri and Sen (1985, Sections 5.2 and 5.3) for a typical example

of this questionable attitude. Serial dependence problems and, more particularly, that of testing an ARMA model with unspecified innovation density, provide an interesting instance where an effective choice between signed and unsigned ranks cannot be eluded, and where power considerations do not supersede the much more fundamental invariance principles underlying this choice.

The following consequences of Proposition 4.2 will be particularly useful in the sequel.

PROPOSITION 4.3. *(i) Let J denote a score-generating function for the serial rank statistic $S^{(n)}$ in (3.2). Then, under $H_g^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau})$, $(n-k)^{\frac{1}{2}}(S^{(n)}-m^{(n)})$ is asymptotically normal, as $n \to \infty$, with mean $-\frac{1}{2}\sum_{i=1}^{k}(a_i+b_i)C_i$, where*

$$C_i = \sum_{j=0}^{k-i} \int_{[0,1]^{k+1}} J^*(u_1,\ldots,u_{k+1})\varphi_g(G^{-1}(u_{1+j}))G^{-1}(u_{1+j+i})du_1\ldots du_{k+1} \ ,$$

and variance V^2 given in (3.15).

(ii) Let J_+ denote a score-generating function for the serial signed-rank statistic $S_+^{(n)}$ in (3.9). Then, under $H_g^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau})$ (g, a symmetric density function), $(n-k)^{\frac{1}{2}}S_+^{(n)}$ is asymptotically normal, as $n \to \infty$, with mean $\sum_{i=1}^{k}(a_i+b_i)C_i^+$, where

$$C_i^+ = \sum_{j=0}^{k-i} \int_{[0,1]^{k+1}} J_+(2u_1-1,\ldots,2u_{k+1}-1)\varphi_g(G^{-1}(u_{1+j}))G^{-1}(u_{1+j+i})du_1\ldots du_{k+1} \ ,$$

and variance V_+^2 given in (3.18) (due to the fact that

$$\int_0^1 [J_+(2u-1,\ldots,2u_{k+1}-1)-J_+^*(2u_1-1,\ldots,2u_{k+1}-1)\varphi_g(G^{-1}(u_\ell))G^{-1}(u_{\ell'})du_1\ldots du_{k+1}=0$$

for all $1 \le \ell \ne \ell' \le k+1$, whether J_+ or J_+^* is used does not affect the value of C_i^+).

(iii) Both $(n-i)^{\frac{1}{2}}r_{i;f}^{(n)}$ and $(n-i)^{\frac{1}{2}}r_{i,f}^{(n)+}$ (which are asymptotically standard normal under $H^{(n)}(\boldsymbol{\theta})$ and $H_+^{(n)}(\boldsymbol{\theta})$, respectively), are asymptotically normal under $H_g^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau})$, as $n \to \infty$, with mean $(a_i+b_i)I(f|g)$, where

$$I(f|g) = \left\{\int_0^1 \varphi_f(F^{-1}(u))\varphi_g(G^{-1}(u))du \int_0^1 F^{-1}(u)G^{-1}(u)du\right\}[I(f_1)]^{-\frac{1}{2}}$$

and variance one (for $r_{i;f}^{(n)+}$, of course, f and g are to be symmetric).

PROPOSITION 4.4. *(i) $n^{\frac{1}{2}}\mathbf{T}_{\psi;f}^{(n)}$ is asymptotically normal, as $n \to \infty$, with mean* **0** *under $H^{(n)}(\boldsymbol{\theta})$, mean*

$$\begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_\pi + b_\pi \\ \sum_{i=\pi+1}^{\infty} (a_i + b_i)\psi_i^{(1)} \\ \vdots \\ \sum_{i=\pi+1}^{\infty} (a_i + b_i)\psi_i^{(p_1+q_1)} \end{pmatrix} I(f|g) = \boldsymbol{\mu}_\psi(\boldsymbol{\tau})I(f|g) \tag{4.11}$$

under $H_g^{(n)}(\boldsymbol{\theta} + n^{-\frac{1}{2}}\boldsymbol{\tau})$ and full–rank covariance matrix $\mathbf{W}_\psi$ (see (4.9)) under both.

(ii) The quadratic rank statistic $n(\mathbf{T}_{\psi;f}^{(n)})' \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{w}_\psi^{-1} \end{pmatrix} \mathbf{T}_{\psi;f}^{(n)}$ does not depend on the particular fundamental system $\{\psi_t^{(1)}, \ldots, \psi_t^{(p_1+q_1)}\}$ adopted; it is asymptotically chi–square, with $\pi + p_1 + q_1 = \max(p_1 + q_2, p_2 + q_1)$ degrees of freedom under $H^{(n)}(\boldsymbol{\theta})$, and asymptotically noncentral chi–square, still with $\max(p_1 + q_2, p_2 + q_1)$ degrees of freedom, but with noncentrality parameter $\frac{1}{2}[\|\mathbf{a} + \mathbf{b}\|I(f|g)]^2$ under $H_g^{(n)}(\boldsymbol{\theta} + n^{-\frac{1}{2}}\boldsymbol{\tau})$.

The consequences of Proposition 4.2 are of primary importance for all inference problems in the area: hypothesis testing, estimation, model selection and identification, among others. In the hypothesis testing context, which we develop in some detail in Sections 5 and 6, Proposition 4.2, in a very intuitive interpretation, implies that, asymptotically and locally, testing problems about $\boldsymbol{\theta}$ can be treated as testing problems about the mean $\boldsymbol{\mu}_\psi$ of the central sequence of statistics $\mathbf{T}_{\psi;f}^{(n)}$. A complex problem about ARMA parameters $\boldsymbol{\theta}$ thus turns out to be asymptotically equivalent to a hopefully much simpler one about the mean of a multinormal statistic with locally constant covariance structure.

§5. LOCALLY ASYMPTOTICALLY OPTIMAL RANK TESTS

5.1 Locally asymptotically most powerful tests. The theoretical results of Section 4.4 allow for the construction of locally asymptotically optimal rank tests for a variety of problems. The problems treated in the present section are those dealing with ARMA models which, under the null hypothesis and except for the innovation density, are completely specified. Depending upon the alternative (which may have a one–dimensional or

multidimensional structure, locally (one–sided) most powerful or locally maximin rank–based tests can be derived.

Denote by $H^{(n)}$ and $K^{(n)}$ two sequences of hypotheses, i.e. two sequences of non-overlapping subsets of some parameter space. A sequence $\phi_*^{(n)}$ of tests is called *asymptotically most powerful for* $H^{(n)}$ against $K^{(n)}$ at probability level α if

$$\limsup_{n\to\infty}\left[E_{\boldsymbol{\theta}^{(n)}}(\phi_*^{(n)})-\alpha\right]\leq 0\ ,\quad \boldsymbol{\theta}^{(n)}\in H^{(n)} \tag{5.1}$$

and, for any sequence $\phi^{(n)}$ satisfying (5.1),

$$\liminf_{n\to\infty}\left[E_{\boldsymbol{\theta}^{(n)}}(\phi_*^{(n)}-\phi^{(n)})\right]\geq 0\ ,\quad \boldsymbol{\theta}^{(n)}\in K^{(n)}\ . \tag{5.2}$$

Whenever $K^{(n)}$ can be considered as a *local alternative* with respect to $H^{(n)}$ (e.g. $K^{(n)}$ contiguous to $H^{(n)}$), $\phi_*^{(n)}$ will also be termed *locally asymptotically most powerful.*

PROPOSITION 5.1. *The (sequence of) rank–based test(s)*

$$\phi^{(n)}=1\quad\text{iff}\sum_{i=1}^{n-1}(n-i)^{\frac{1}{2}}(a_i+b_i)r_{i;f}^{(n)}>k_{1-\alpha}\left[\sum_{i=1}^{n-1}(a_i+b_i)^2\right]^{\frac{1}{2}}\ , \tag{5.3}$$

where $k_{1-\alpha}=\Phi^{-1}(1-\alpha)$ *denotes the* $(1-\alpha)$*–standard normal quantile,*

(i) is locally asymptotically most powerful (at probability level α*) for testing* $H^{(n)}(\boldsymbol{\theta})$ *against* $H_f^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}k\boldsymbol{\tau})$, $k>0$ *arbitrary.*

(ii) has asymptotic power

$$1-\Phi(k_{1-\alpha}-\|\mathbf{a}+\mathbf{b}\|\ I(f|g)) \tag{5.4}$$

against $H_g^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau})$.

Note that (as expected) the asymptotic power of the rank–based test (5.3) against $H_f^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau})$ —namely, $1-\Phi(k_{1-\alpha}-\|\mathbf{a}+\mathbf{b}\|[I(f_1)]^{\frac{1}{2}})$ —equals that of the Neyman test $\phi_{NP}^{(n)}$ (for $H_f^{(n)}(\boldsymbol{\theta})$ against $H_f^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau})$). The latter indeed consists in rejecting $H_f^{(n)}(\boldsymbol{\theta})$ whenever $\Lambda_{\boldsymbol{\theta};\boldsymbol{\tau};f}^{(n)}$ is "too large", i.e., in view of Proposition 4.2, asymptotically reduces to

$$\phi_{NP}^{(n)}=1\ \text{iff}\ \Lambda_{\boldsymbol{\theta};\boldsymbol{\tau};f}^{(n)}+\frac{1}{2}\|\mathbf{a}+\mathbf{b}\|^2I(f_1)>k_{1-\alpha}\|\mathbf{a}+\mathbf{b}\|[I(f_1)]^{\frac{1}{2}}\ . \tag{5.5}$$

The asymptotic distribution of $\Lambda_{\boldsymbol{\theta};\boldsymbol{\tau};f}^{(n)}$ under $H_f^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau})$ can be shown to be normal, with mean $\frac{1}{2}\ \|\mathbf{a}+\mathbf{b}\|^2I(f_1)$ and the same variance as under $H_f^{(n)}(\boldsymbol{\theta})$. The asymptotic power

of (5.5) therefore is $1-\Phi(k_{1-\alpha}-\|\mathbf{a}+\mathbf{b}\|[I(f_1)]^{\frac{1}{2}})$. The norm $\|\mathbf{a}+\mathbf{b}\|$ accordingly can be interpreted as a "natural" distance between the sequences $H^{(n)}(\boldsymbol{\theta})$ and $H_f^{(n)}(\boldsymbol{\theta}+n^{\frac{1}{2}}\boldsymbol{\tau})$, equivalent to the L^1–distance $2\sup_\alpha\{1-\alpha-\Phi(k_{1-\alpha}-\|\mathbf{a}+\mathbf{b}\|[I(f_1)]^{-\frac{1}{2}})\}$.

Here again, in case the innovation densities are restricted to symmetric ones, signed ranks and signed autocorrelation coefficients can be substituted for the unsigned ones: Proposition 5.1 can be formulated without modification with $r_{i;f}^{(n)+}$ instead of $r_{i;f}^{(n)}$ (and, of course, symmetric densities f and g). If a parametric, Gaussian approach, is adopted (all densities then, under the null hypothesis as well as under the alternative, are assumed to be Gaussian), Proposition 5.1 takes the form

PROPOSITION 5.2. *Denote by $r_i^{(n)}$ the classical autocorrelation coefficient of order i. The parametric test*

$$\phi^{(n)}=1 \text{ iff } \sum_{i=1}^{n-1}(n-i)^{\frac{1}{2}}(a_i+b_i)\,r_i^{(n)} > k_{1-\alpha}\left[\sum_{i=1}^{n-1}(a_i+b_i)^2\right]^{\frac{1}{2}}$$

(i) is locally most powerful (at probability level α) for testing $H^{(n)}(\boldsymbol{\theta})$ (or $H_+^{(n)}(\boldsymbol{\theta})$) against $H_{\mathcal{N}}^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}k\boldsymbol{\tau})$ (where $\mathcal{N}$ stands for a normal density with mean zero and arbitrary variance), $k>0$ arbitrary.

(ii) has asymptotic power

$$1-\Phi(k_{1-\alpha}-\|\mathbf{a}+\mathbf{b}\|) \tag{5.6}$$

against $H_g^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau})$, whatever g may be.

This confirms the analogy between f–rank autocorrelations (under innovation density f) and usual autocorrelation coefficients (under Gaussian innovation densities). The eventual superiority of f–rank autocorrelations comes from the $\mathbf{I}(\mathbf{f}|\mathbf{g})$ factor appearing in (5.4), but not in (5.6).

Example 5.1. Consider the problem of testing the null hypotheses $H^{(n)}(0)$ that $\mathbf{X}^{(n)}$ is white noise (with unspecified density) against an alternative of possible ARMA (p,q) dependence, $\max(p,q)=1$, i.e. $X_t-AX_{t-1}=\varepsilon_t+B\varepsilon_{t-1}$, $A+B>0$. Here $Z_t^{(n)}=X_t^{(n)}$ and the ranks are those of the observed series $\mathbf{X}^{(n)}$ itself. The rank–based test rejecting $H^{(n)}(0)$ whenever $(n-1)^{\frac{1}{2}}r_{1;f}^{(n)}>k_{1-\alpha}$ is locally most powerful (against $X_t-n^{-\frac{1}{2}}\alpha X_{t-1}=\varepsilon_t+n^{-\frac{1}{2}}\beta\varepsilon_{t-1}$, $\alpha+\beta>0$ arbitrary) if ε_t, under the alternative, has density f (with arbitrary variance). If Gaussian (logistic, double exponential) alternatives are to be privileged or are to be feared most, then a van der Waerden (Wilcoxon, Laplace)

autocorrelation $r^{(n)}_{1;vdW}$ $(r^{(n)}_{1;W}, r^{(n)}_{1;L})$ should be adopted. If the density under $H^{(n)}(0)$ can be assumed symmetric, then signed ranks and signed autocorrelations can be substituted for the unsigned ones.

Example 5.2. Consider the problem of testing $\rho = \rho_0$ against $\rho > \rho_0$ in the $AR(1)$ model $X_t - \rho X_t = \varepsilon_t$ with unspecified but symmetric innovation density. A signed–rank, genuinely distribution–free test can be performed, which is locally most powerful against Gaussian alternatives, among all tests at (asymptotic) probability level α, by rejecting $\rho = \rho_0$ whenever

$$(1-\rho_0^2)^{\frac{1}{2}} \sum_{i=1}^{n-1} (n-i)^{\frac{1}{2}} \rho_0^{i-1} r^{(n)+}_{i;f} > k_{1\alpha} \ .$$

The signs and ranks here are those of the residuals $Z^{(n)}_t = X^{(n)}_t - \rho_0 X^{(n)}_{t-1}$ and their absolute values $|Z^{(n)}_t|$ (assume $\mathbf{X}^{(n)} = (X^{(n)}_0, X^{(n)}_1, \ldots, X^{(n)}_n)$). If the density (under $\rho = \rho_0$) cannot be assumed symmetric, unsigned ranks and unsigned autocorrelations should be used instead of the signed ones.

The asymptotic relative efficiencies of the signed and unsigned rank procedures described in this section with respect to each other, and with respect to their parametric counterparts, are computed in Section 5.3.

5.2 Locally maximin tests. The local alternatives considered in Section 5.2 are basically one–dimensional, one–sided alternatives: the tests provided by Propositions 5.1 and 5.2 are optimal against a specified $(\boldsymbol{\gamma}, \boldsymbol{\delta})$ direction in the parameter space, along with specified innovation density. In many situations of practical interest, the alternative is inherently multidimensional, and no particular $(\boldsymbol{\gamma}, \boldsymbol{\delta})$ direction is to be privileged. A natural idea then consists in considering tests which are *locally asymptotically maximin.*

Informally speaking, a maximin test, at given probability level α, is a test whose worst performance is best within the class of all tests at level α. More precisely, let $H^{(n)}$ and $K^{(n)}$ denote two sequences of (nonoverlapping) hypotheses. The corresponding sequence of *envelope power functions* is

$$\beta(\alpha; H^{(n)}, K^{(n)}) = \sup_{\phi} \inf_{\boldsymbol{\theta}} E_{\boldsymbol{\theta}}(\phi) \ , \quad \alpha \in (0,1) \ ,$$

where the $\sup_{\phi}$ and $\inf_{\boldsymbol{\theta}}$ are taken over all tests ϕ satisfying $E_{\boldsymbol{\theta}}\phi \leq \alpha$, $\boldsymbol{\theta} \in H^{(n)}$, and all values of $\boldsymbol{\theta}$ in $K^{(n)}$, respectively. A sequence $\phi^{(n)}_*$ of tests is called *asymptotically* (*locally*, if $K^{(n)}$ is a local alternative to $H^{(n)}$) maximin for $H^{(n)}$ against $K^{(n)}$, at (asymptotic) level α, if

$$\limsup_{n\to\infty} \left[E_{\boldsymbol{\theta}^{(n)}}(\phi^{(n)}_*) - \alpha \right] \leq 0 \ , \quad \boldsymbol{\theta}^{(n)} \in H^{(n)} \tag{5.7}$$

and

$$\liminf_{n\to\infty}\left[E_{\boldsymbol{\theta}^{(n)}}(\phi_*^{(n)}) - \beta(\alpha; H^{(n)}, K^{(n)})\right] \geq 0\ , \quad \boldsymbol{\theta}^{(n)} \in K^{(n)}\ . \tag{5.8}$$

If however nontrivial maximin tests are to be obtained, $K^{(n)}$ has to be bounded away from $H^{(n)}$. If Indeed the L^1–distance between $H^{(n)}$ and $K^{(n)}$ (for fixed n) would be zero, the envelope power function trivially would reduce to $\beta(\alpha; H^{(n)}, K^{(n)}) = \alpha$. The simplest idea, if $H^{(n)}$ and a local alternative $K^{(n)}$ are to be considered, consists in defining a local alternative whose L^1–distance to $H^{(n)}$ remains (asymptotically) bounded from below by some fixed, strictly positive, constant d. In the problem of testing $H^{(n)}(\boldsymbol{\theta})$ (or $H_+^{(n)}(\boldsymbol{\theta})$) —i.e. a specified ARMA (p_1, q_1) model with unspecified (or symmetric unspecified) innovation density —against unspecified ARMA (p, q) alternatives, with $p_1 \leq p \leq p_2$ and $q_1 \leq q \leq q_2$, such local alternatives with bounded L_1–distance from $H^{(n)}(\boldsymbol{\theta})$ (or $H_+^{(n)}(\boldsymbol{\theta})$) are (for fixed f) of the form

$$K_f^{(n)}(d) = \bigcup\{H_f^{(n)}(\boldsymbol{\theta} + n^{-\frac{1}{2}}\boldsymbol{\tau}) \mid \|\mathbf{a} + \mathbf{b}\| \geq d\}\ , \tag{5.9}$$

where the union is taken over all $\boldsymbol{\tau} \in \mathbb{R}^{p_2+q_2}$ such that $\|\mathbf{a}+\mathbf{b}\|$ (with $\mathbf{a}$ and $\mathbf{b}$ as defined in (4.4) is larger than or equal to d. The alternative (5.9) accordingly can be interpreted as the outside of a L^1–hypersphere. Due to the particular L^1 topology of ARMA likelihoods, $K_f^{(n)}(d)$ however resembles a cylinder rather than a sphere in the $\mathbb{R}^{p_2+q_2}$ space. The choice of d, as we shall see, does not affect the final result and the form of locally maximin tests.

PROPOSITION 5.3. *(i) The (sequence of) rank–based test(s)*

$$\phi^{(n)} = 1 \text{ iff } n(\mathbf{T}_{\psi;f}^{(n)})' \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{w}_\psi^{-1} \end{pmatrix} \mathbf{T}_{\psi;f}^{(n)} > \chi^2_{1-\alpha} \tag{5.10}$$

(which, from Proposition 4.4, do not depend on the particular fundamental system adopted), where $\chi^2_{1-\alpha}$ *denotes the* $(1-\alpha)$*–quantile of a chi–square variable with* $\pi + p_1 + q_1 = \max(p_1 + q_2, p_2 + q_1)$ *degrees of freedom, is locally asymptotically maximin, at probability level* α*, against any alternative of the form* $K_f^{(n)}(d)$, $d > 0$.

(ii) The asymptotic power of (5.10) against $H_g^{(n)}(\boldsymbol{\theta} + n^{-\frac{1}{2}}\boldsymbol{\tau})$ *is*

$$1 - F(\chi^2_{1-\alpha}; \tfrac{1}{2}\|\mathbf{a} + \mathbf{b}\|^2(I(f|g))^2)\ ,$$

where $F(\ ; \lambda)$ *denotes the distribution function of a noncentral chi–square variable with* $\pi + p_1 + q_1$ *degrees of freedom and noncentrality parameter* λ.

(iii) The envelope power function $\beta(\alpha; H^{(n)}(\boldsymbol{\theta}), K_f^{(n)}(d))$ *converges to (same notation as above)* $1 - F(\chi^2_{1-\alpha}; \frac{1}{2}\|\mathbf{a} + \mathbf{b}\|^2 I(f_1))$.

The same result of course still holds if signed ranks are substituted for unsigned ones. Its parametric counterpart involves a parametric version $\mathbf{T}_{\psi;\mathcal{N}}^{(n)}$, say, of $\mathbf{T}_{\psi;f}^{(n)}$, where classical parametric autocorrelations $r_i^{(n)}$ are substituted for the rank-band ones $r_{i,f}^{(n)}$.

PROPOSITION 5.4. *(i) Substituting* $\mathbf{T}^{(n)}_{\psi;\mathcal{N}}$ *for* $\mathbf{T}^{(n)}_{\psi;f}$ *in (5.10) yields a parametric test which is locally maximin, at probability level* α, *for* $H^{(n)}(\boldsymbol{\theta})$ *(for* $H^{(n)}_{+}(\boldsymbol{\theta})$*) against any* $K^{(n)}_{\mathcal{N}}(d)$, $d > 0$ *(where* $\mathcal{N}$ *stands for a normal density with mean zero and arbitrary variance).*

(ii) The asymptotic power of this test against $H^{(n)}_{g}(\boldsymbol{\theta} + n^{-\frac{1}{2}}\boldsymbol{\tau})$ *is (same notation as above)* $1 - F(\chi^2_{1-\alpha}; \frac{1}{2}\|\mathbf{a} + \mathbf{b}\|^2)$.

Example 5.3. As in Example 5.1, consider the problem of testing the null hypothesis $H^{(n)}(0)$ that $\mathbf{X}^{(n)}$ is white noise, with unspecified (or unspecified symmetric) density. The alternative now is the whole class of possible ARMA (p, q) dependencies, with $0 < \max(p, q) \leq \pi$ and unspecified coefficients, but the subset of it at which optimality is desired is that with innovation density f. The *f–rank* (signed or unsigned) *portmanteau test* of order π

$$\phi^{(n)} = 1 \quad \text{if and only if} \quad \sum_{i=1}^{\pi}(n-i)(r^{(n)}_{i;f})^2 > \chi^2_{1-\alpha} ,$$

where $\chi^2_{1-\alpha}$ is the $(1-\alpha)$–quantile of a chi–square variable with π degrees of freedom, is then locally asymptotically maximin. The same property holds for the parametric portmanteau test (where the parametric autocorrelations $r^{(n)}_i$ are substituted for the rank–based $r^{(n)}_{i;f}$) against normal alternatives.

Example 5.4. Consider the problem of testing the null hypothesis that an observed series $\mathbf{X}^{(n)}$ was generated by the ARMA (1,1) model $X_t - \frac{4}{5}X_{t-1} = \varepsilon_t + \frac{1}{2}\varepsilon_{t-1}$, $t \in \mathbb{Z}$, against unspecified ARMA (2,2) alternatives (or, equivalently, against ARMA (2,1) or ARMA (1,2) alternatives). Let $Z^{(n)}_{-1} = Z^{(n)}_0 = 0$ and $Z^{(n)}_t = X^{(n)}_t - \frac{4}{5}X^{(n)}_{t-1} - \frac{1}{2}Z^{(n)}_{t-1}$, $t = 1, \ldots, n$; denote by $r^{(n)}_{i,f}$ the corresponding f–rank autocorrelations. A fundamental system of solutions of

$$\left(1 - \frac{4}{5}L\right)\left(1 + \frac{1}{2}\right)\psi_t = 0 \ , \ t \in \mathbb{Z}$$

is

$$\left\{\psi^{(1)}_t = \left(\frac{4}{5}\right)^{t-2} ; \psi^{(2)}_t = \left(-\frac{1}{2}\right)^{t-2}\right\};$$

here $\pi = \max(p_2 - p_1, q_2 - q_1) = 1$. The covariance matrix $\mathbf{W}^2$ is of the form

$$\mathbf{W}^2 = \begin{pmatrix} \left[1 - \left(\frac{4}{5}\right)^2\right]^{-1} & \left[1 + \frac{2}{5}\right]^{-1} \\ \left[1 + \frac{2}{5}\right]^{-1} & \left[1 - \left(\frac{1}{2}\right)^2\right]^{-1} \end{pmatrix} = \begin{pmatrix} \frac{25}{9} & \frac{5}{7} \\ \frac{5}{7} & \frac{4}{3} \end{pmatrix} .$$

The test statistic of Proposition 5.3(i) is thus

$$Q^{(n)}_{\psi;f} = n(r^{(n)}_{1;f})^2 +$$
$$n\left(\sum_{i=2}^{n-1}\left(\frac{4}{5}\right)^{i-2} r^{(n)}_{i;f}, \sum_{i=2}^{n-1}\left(-\frac{1}{2}\right)^{i-2} r^{(n)}_{i;f}\right)\begin{pmatrix}\frac{25}{9} & \frac{5}{7}\\ \frac{5}{7} & \frac{4}{3}\end{pmatrix}\begin{pmatrix}\sum_{i=2}^{n-1}\left(\frac{4}{5}\right)^{i-2} r^{(n)}_{i;f}\\ \sum_{i=2}^{n-1}\left(-\frac{1}{2}\right)^{i-2} r^{(n)}_{i;f}\end{pmatrix}$$
$$\simeq n(r^{(n)}_{1;f})^2 + (1.436)n\left[\sum_{i=2}^{n-1}\left(\frac{4}{5}\right)^{i} r^{(n)}_{i;f}\right]^2 + (19.600)n\left[\sum_{i=2}^{n-1}\left(-\frac{1}{2}\right)^{i} r^{(n)}_{i;f}\right]^2$$
$$- (3.937)n\left[\sum_{i=2}^{n-1}\left(\frac{4}{5}\right)^{i} r^{(n)}_{i;f}\right]\left[\sum_{i=2}^{n-1}\left(-\frac{1}{2}\right)^{i} r^{(n)}_{i;f}\right] .$$

If ARMA (3,3) (or ARMA (3,2), ARMA (3,1), ARMA (1,3), ...) alternatives are to be considered, $Q^{(n)}_{\psi;f}$ has to be modified to

$$n(r^{(n)}_{1;f})^2 + n(r^{(n)}_{2;f})^2 + (2.243)n\left[\sum_{i=3}^{n-1}\left(\frac{4}{5}\right)^{i} r^{(n)}_{i;f}\right]^2 + (78.400)n\left[\sum_{i=3}^{n-1}\left(-\frac{1}{2}\right)^{i} r^{(n)}_{i;f}\right]^2$$
$$+ (9.844)n\left[\sum_{i=3}^{n-1}\left(\frac{4}{5}\right)^{i} r^{(n)}_{i;f}\right]\left[\sum_{i=3}^{n-1}\left(-\frac{1}{2}\right)^{i} r^{(n)}_{i;f}\right] .$$

5.3 Asymptotic Relative Efficiencies. Asymptotic, local power comparisons between tests are usually made on the basis of asymptotic relative efficiencies (ARE's). Recall that the ARE of a test ψ_1 with respect to another one ψ_2, both at probability level α, can be defined as the limiting value of the ratio N_2/N_1, where N_1 is the number of observations required for the power of ψ_1 to equal the power of ψ_2 based on N_2 observations, as $\min(N_1, N_2) \to \infty$ and the alternative converges to the null hypothesis (see e.g. Puri and Sen 1985, section 3.8, for a more rigorous definition).

ARE computations are greatly simplified in situations where ψ_1 and ψ_2 are based on test statistics which are asymptotically normal or chi–square under the null hypothesis and contiguous consequences of alternatives. If ψ_1 and ψ_2 consist in rejecting the null hypothesis $H^{(n)}$ for "large values of" test statistics, S_1 and S_2 respectively, which are asymptotically normal, $\mathcal{N}(0, \sigma_1)$ and $\mathcal{N}(0, \sigma_2)$ respectively, under $H^{(n)}$, and $\mathcal{N}(\mu_1, \sigma_1)$ and $\mathcal{N}(\mu_2, \sigma_2)$ respectively, under the alternative $K^{(n)}$, then the ARE of ψ_1 with respect to ψ_2 is

$$\text{ARE } (\psi_1/\psi_2) = \left(\frac{\mu_1\sigma_2}{\sigma_1\mu_2}\right)^2 .$$

If S_1 and S_2 are both asymptotically chi–square with d degrees of freedom under $H^{(n)}$, and asymptotically noncentral chi–square with d degrees of freedom and noncentrality parameters λ_1 and λ_2, respectively, then

$$\text{ARE } (\psi_1/\psi_2) = (\lambda_1/\lambda_2)^2 \ .$$

The results of Sections 5.1 and 5.2 thus allow for an explicit computation of the mutual ARE's of the various parametric and nonparametric tests proposed.

PROPOSITION 5.5. *The asymptotic efficiency of the optimal procedures (as described in Propositions 5.1 and (5.3) based on (signed or unsigned) f–rank autocorrelations with respect to their counterparts based on (signed or unsigned) g–rank autocorrelations, against alternatives of the form $H_h^{(n)}(\boldsymbol{\theta} + n^{-\frac{1}{2}}\boldsymbol{\tau})$, is*

$$[I(f|h)/I(g|h)]^2 = \left[\frac{\int_0^1 \phi_f(F^{-1}(u))\phi_h(H^{-1}(u))du \ \int_0^1 F^{-1}(u)H^{-1}(u)du}{\int_0^1 \phi_g(G^{-1}(u))\phi_h(H^{-1}(u))du \int_0^1 G^{-1}(u)H^{-1}(u)du} \right]^2 ,$$

where notations $\phi_f, \phi_g, \phi_h, F, G, H$ are used in an obvious fashion. Under the same conditions, the ARE of procedures based on f–rank autocorrelations with respect to the optimal Gaussian parametric procedures, based on classical parametric autocorrelation coefficients (as described in Propositions 5.2 and 5.4) is

$$\left[\int_0^1 \phi_f(F^{-1}(u))\phi_h(H^{-1}(u))du \int_0^1 F^{-1}(u)H^{-1}(u)du \right]^2 / I(f_1) \ ,$$

which reduces to $I(f_1) \geq 1$ for $h = f$ and $I(f_1) = 1$ if and only if f and h are Gaussian.

Other asymptotic efficiencies can be derived from Proposition 4.3. Explicit numerical values are provided in Table 5.1 below.

		(1)	(2)	(3)	(4)	(5)	
	Classical	1.000	1.000	1.005	1.634	1.096	normal
(1)	parametric	1.000	0.954	0.911	1.232	1.000	logistic
		1.000	0.816	0.675	0.500	0.790	double exp.
	Signed or	1.000	1.000	1.055	1.634	1.096	normal
(2)	unsigned	1.048	1.000	0.954	1.291	1.048	logistic
	van der Waerden	1.226	1.000	0.827	0.613	0.968	double exp.
	Signed or	0.948	0.948	1.000	1.550	1.091	normal
(3)	unsigned	1.098	1.048	1.000	1.352	1.098	logistic
	Wilcoxon	1.482	1.209	1.000	0.741	1.170	double exp.
	Signed or	0.612	0.612	0.646	1.000	0.671	normal
(4)	unsigned	0.812	0.775	0.740	1.000	0.812	logistic
	Laplace	2.000	1.631	1.350	1.000	1.580	double exp.
	Signed or	0.912	0.912	0.917	1.490	1.000	normal
(5)	unsigned	1.000	0.954	0.911	1.232	1.000	logistic
	Spearman	1.266	1.033	0.855	0.633	1.000	double exp.

Table 5.1. *Mutual ARE's for the various parametric and nonparametric tests described in Sections 5.1 and 5.2, under normal, logistic and double exponential densities, respectively.*

An inspection of Table 5.1 reveals the excellent asymptotic performances of rank tests: the van der Waerden tests perform uniformly and strictly better than the corresponding normal–theory tests—except of course under normal densities, where they perform equally well. The ARE of optimal normal theory tests with respect to the corresponding Laplace procedure can be as low as 0.500. Spearman tests (which are never optimal, since the Spearman autocorrelation coefficients do not belong to the class of f–rank autocorrelations) are uniformly (though not very significantly) dominated by Wilcoxon tests.

As expected from Proposition 4.3, signed and unsigned optimal tests are asymptotically equivalent. The fact that their mutual ARE's are one however does not imply that the advantage of using signed–rank tests instead of unsigned ones (whenever innovation densities are symmetric) is nil, or negligible, neither for short series lengths, nor even asymptotically. Numerical investigations of the performance of signed–rank tests (Hallin et al. 1990) indicate that, for moderate, fixed n, the power of signed–rank procedures can be substantially larger than that of unsigned ones, an empirical finding that should be confirmed by a theoretical investigation of the corresponding deficiencies.

A systematic investigation of the finite sample behavior of (unsigned) rank based tests

for randomness, both under the null hypothesis (tables of exact critical values) and under alternatives of $AR(1)$ serial dependence (Monte Carlo study of the power function) has been conducted in Hallin and Mélard (1988). This study reveals that rank tests often are substantially more powerful than traditional parametric procedures, even for pretty short series ($n = 20$, e.g.); of course, they are much more reliable under the null hypothesis, and theoretically provide exact tests (whereas the probability level of parametric tests is only approximative) even when traditional procedures are not valid — e.g. under Cauchy innovation densities (see Hallin and Puri (1991a) for a numerical illustration). In addition, they also are considerably more robust, and less sensitive to the presence of outliers, atypical startup behavior, etc.

All these properties show how useful a general rank based methodology would be in the identification and validation steps of time series analysis. Such a methodology however requires additional results allowing for the treatment of nuisance parameters. This is the subject of the next section.

§6. ALIGNED RANK TESTS

6.1 Ranks and aligned ranks. The test described in Section 5 are valid if and only if the residuals $Z_t^{(n)}$, the ranks of which are used, under the null hypothesis to be tested, are *exact* residuals. These tests thus allow for testing ARMA models with unspecified innovation densities but completely specified coefficients.

Unfortunately, in most practical problems, one has to test ARMA models with *unspecified* coefficients. This is the case, typically, in diagnostic checking and model validation situations. It is also the case in identification problems, where the main tools (partial correlograms, corner method tables, ...) actually are test statistics—even though the identification process does not exactly reduce to any hypothesis testing problem.

More generally, one might like to test the null hypothesis $H^{(n)}(\mathcal{E}_r)$ (or $H_+^{(n)}(\mathcal{E}_r)$, in the case of unspecified, symmetric innovations) that the parameters $A_1, \ldots, A_{p_1}, B_1, \ldots, B_{q_1}$ of the ARMA (p_1, q_1) model underlying the observed series $\mathbf{X}^{(n)}$ satisfy some given linear constraints (to be precise, $p_1 + q_1 - r$ linearly independent ones) and thus belong to some r-dimensional linear subspace $\mathcal{E}_r$ of $\mathbb{R}^{p_1+q_1}$. Innovation densities in $H^{(n)}(\mathcal{E}_r)$ and $H_+^{(n)}(\mathcal{E}_r)$ remain unspecified; whenever they need to be specified, the notation $H_f^{(n)}(\mathcal{E}_r)$ will be used. The alternative consists of unrestricted ARMA (p, q) models with $p_1 \leq p \leq p_2$, $q_1 \leq q \leq q_2$. More specific alternatives however might be considered when optimality properties are to be described.

In what follows, we assume the existence of a root n consistent sequence $\widehat{\boldsymbol{\theta}}^{(n)} = (\widehat{A}_1^{(n)}, \ldots, \widehat{A}_{p_1}^{(n)}, 0, \ldots, 0, \widehat{B}_1^{(n)}, \ldots, \widehat{B}_{q_1}^{(n)}, 0, \ldots, 0)'$ of constrained estimates of $\boldsymbol{\theta} = (A_1, \ldots, A_{p_1}, 0, \ldots, 0, B_1, \ldots, B_{q_1}, 0, \ldots, 0)'$, i.e. a sequence of statistics such that $(\widehat{A}_1^{(n)}, \ldots, \widehat{A}_{p_1}^{(n)}, \widehat{B}_1^{(n)}, \ldots, \widehat{B}_{q_1}^{(n)})' \in \mathcal{E}_r$ (in the sequel, we also write $\widehat{\boldsymbol{\theta}}^{(n)} \in \mathcal{E}_r$ or $\boldsymbol{\theta} \in \mathcal{E}_r$) and the distributions of $n^{\frac{1}{2}}(\widehat{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta})$ form, under $H^{(n)}(\mathcal{E}_r)$ (or under $H_+^{(n)}(\mathcal{E}_r)$), a relatively compact sequence.

Root n consistency however is not sufficient for establishing the asymptotic results stated below: additional uniformity properties are required, which are satisfied if "smoother" versions of $\widehat{\boldsymbol{\theta}}^{(n)}$ are used (see LeCam 1960, Appendix 1). One of the "smoothed" versions of $\widehat{\boldsymbol{\theta}}^{(n)}$ is as follows (same reference). Denote by $\mathcal{A}^{(n)}$ the σ–field of Borel sets of $\mathbb{R}^{(n)}$, by V an open convex symmetric neighborhood of the origin in $\mathcal{E}_r$, by $\mathcal{A}_*^{(n)}$ the product σ–field of $\mathcal{A}^{(n)}$ by the σ–field of Borel subsets of $\mathcal{E}_r$. On V, letting ν denote a probability measure having a bounded continuous density with respect to the Lebesgue measure, define $\mathbf{v}^{(n)}$ as a random variable having distribution ν on V, and put

$$\boldsymbol{\theta}_*^{(n)} = \widehat{\boldsymbol{\theta}}^{(n)} + n^{-\frac{1}{2}}\mathbf{v}^{(n)} \ .$$

The sequence of estimates $\boldsymbol{\theta}_*^{(n)}$ is still root n consistent, and meets the uniformity requirements that might not hold for $\widehat{\boldsymbol{\theta}}^{(n)}$ and are technically necessary for establishing Propositions 6.1 and 6.2 below.

Other smoothing methods are also described in LeCam (1960).

We do insist however that these smoothing procedures are of little practical relevance, if any at all. First, because they typically have the nature of analytical procedures guaranteeing ad hoc probabilistic limit results, whereas in statistical practice, one is interested in approximation results only: for given V and ν, and for fixed, "reasonably large" n, it makes extremely little difference whether $\boldsymbol{\theta}_*^{(n)}$ or $\widehat{\boldsymbol{\theta}}^{(n)}$ is used. A second reason is that the actual computation of $\widehat{\boldsymbol{\theta}}^{(n)}$, with a finite number of decimal values, *automatically* provides a "smooth", approximate value $\boldsymbol{\theta}_{**}^{(n)}$, say, of $\widehat{\boldsymbol{\theta}}^{(n)}$ (hence of $\boldsymbol{\theta}_*^{(n)}$).

For all these reasons, this smoothing problem should not be given too much attention and, in the sequel, we shall make no notational difference between $\widehat{\boldsymbol{\theta}}^{(n)}$, $\boldsymbol{\theta}_*^{(n)}$ and $\boldsymbol{\theta}_{**}^{(n)}$: the simple notation $\widehat{\boldsymbol{\theta}}^{(n)}$ will be used throughout, and $\widehat{\boldsymbol{\theta}}^{(n)}$ will be assumed "smooth enough".

Denote by $\widehat{Z}_t^{(n)} = Z_t^{(n)}(\widehat{\boldsymbol{\theta}}^{(n)}) = (\Sigma_i \widehat{A}_i^{(n)} L^i / \Sigma_i \widehat{B}_i^{(n)} L^i) X_t^{(n)}$ the "estimated" residuals associated with $\widehat{\boldsymbol{\theta}}^{(n)}$, by $\widehat{R}_t^{(n)} = R_t^{(n)}(\widehat{\boldsymbol{\theta}}^{(n)})$, $\widehat{R}_{+,t}^{(n)} = R_{+,t}^{(n)}(\widehat{\boldsymbol{\theta}}^{(n)})$ and $\widehat{s}_t = s_t(\widehat{\boldsymbol{\theta}}^{(n)})$ the corresponding ranks and signs (namely, the *aligned* ranks, *aligned* signed ranks and *aligned* signs for the problem considered). The notation $\widehat{r}_{i;f}^{(n)} = r_{i;f}^{(n)}(\widehat{\boldsymbol{\theta}}^{(n)})$, $\widehat{r}_{i;f}^{(n)+} = r_{i;f}^{(n)+}(\widehat{\boldsymbol{\theta}}^{(n)})$, $\widehat{a}_i, \widehat{b}_i, \widehat{\psi}_t^{(1)}, \ldots, \widehat{\psi}_t^{(p_1+q_1)}, \widehat{\mathbf{w}}_{\widehat{\psi}}^{(n)}, \widehat{\mathbf{W}}_{\widehat{\psi}}^{(n)}, \widehat{\mathbf{T}}_{\widehat{\psi};f}^{(n)}, \widehat{\mathbf{T}}_{\widehat{\psi};f}^{(n)+}, \ldots$ will be used in a similar fashion.

Since the f–rank autocorrelations $r_{i;f}^{(n)}$ and the rank–based statistics $\mathbf{T}_{\psi;f}^{(n)}$ (or the signed ones $\mathbf{T}_{\psi;f}^{(n)+}$) have been shown locally asymptotically sufficient for $H^{(n)}(\boldsymbol{\theta})$ (or $H_+^{(n)}(\boldsymbol{\theta})$) against $H_f^{(n)}(\boldsymbol{\theta} + n^{-\frac{1}{2}}\boldsymbol{\tau}), \boldsymbol{\tau} \in \mathbb{R}^{p_2+q_2}$, a natural idea would consist in using the aligned autocorrelations $\widehat{r}_{i;f}^{(n)}$ (or $\widehat{r}_{i;f}^{(n)+}$) and aligned rank statistics $\widehat{\mathbf{T}}_{\widehat{\psi};f}^{(n)}$ (or $\widehat{\mathbf{T}}_{\widehat{\psi};f}^{(n)+}$) for $H^{(n)}(\mathcal{E}_r)$ (or $H_+^{(n)}(\mathcal{E}_r)$) against $\bigcup\{H_f^{(n)}(\boldsymbol{\theta} + n^{-\frac{1}{2}}\boldsymbol{\tau}),\ \boldsymbol{\theta} \in \mathcal{E}_r,\ \boldsymbol{\theta} + n^{-\frac{1}{2}}\boldsymbol{\tau} \notin \mathcal{E}_r\}$. A closer look at this apparently simple idea however reveals a number of apparently unredeemable theoretical drawbacks.

First, invariance is lost. Substituting estimated parameters $\widehat{\boldsymbol{\theta}}^{(n)}$ for exact ones $\boldsymbol{\theta}$ induces among the residuals $\widehat{Z}_t^{(n)}$ complex interrelations that destroy their exchange-ability features, hence the invariance property of ranks. This is not just a slight defect in view of the fact, stressed in Section 1, that invariance is the cornerstone of rank–based inference: *if invariance is lost, there is not point in using ranks anymore.*

Second, distribution–freeness, and even asymptotic distribution–freeness, is lost as well: it follows indeed from Proposition 6.1 below that the asymptotic mean of aligned rank autocorrelation coefficients under $H_g^{(n)}(\boldsymbol{\theta})$ depends on g. This again is not just a detail without importance, since distribution-freeness (or, at least, asymptotic distribution-freeness) is crucially necessary if a testing procedure valid under unspecified densities g is to be carried out.

Last and not the least, the local optimality properties of Section 5 obviously cannot be expected to hold anymore in the case of aligned rank tests. Even in the parametric, normal-theory context, additional requirements of unbiasedness or similarity have to be invoked, when nuisance parameters are present, if optimality results are to be obtained. Some form of similarity is likely to be necessary here, too.

6.2 Asymptotic invariance. Since strict invariance apparently is too restrictive a requirement here, some weaker, asymptotic form of it might be more appropriate. Let us define an asymptotically invariant statistic as a statistic asymptotically equivalent (under the null hypothesis considered, hence under contiguous alternatives) to an invariant (hence, in the present context, rank–based) one. More precisely, let $S^{(n)}$ denote a sequence of rank–based statistics such that, for some appropriate centering sequence $m^{(n)}$, $n^{-\frac{1}{2}}(S^{(n)}-m^{(n)})$ is relatively compact under a sequence of null hypotheses $H^{(n)}$. The sequence $\widetilde{S}^{(n)}$ obtained on substituting aligned ranks (aligned signs, aligned signed ranks) for the exact ones in $S^{(n)}$ is said to be *asymptotically invariant* under $H^{(n)}$ if $\widetilde{S}^{(n)} - S^{(n)} = o_P(n^{-\frac{1}{2}})$, under $H^{(n)}$, as $n \to \infty$.

This concept of asymptotic invariance here would be helpful if, e.g. $\widehat{r}_{i;f}^{(n)} - r_{i;f}^{(n)}$ or $\widehat{r}_{i;f}^{(n)+} - r_{i;f}^{(n)+}$ would be $o_P(n^{-\frac{1}{2}})$ under $H^{(n)}(\boldsymbol{\theta})$, i.e. if aligned rank autocorrelation coefficients would be asymptotically equivalent to the genuine ones. Unfortunately, this does not hold, since, from Proposition 4.2, it can be shown that

PROPOSITION 6.1. *Under* $H_g^{(n)}(\boldsymbol{\theta})$,

$$n^{\frac{1}{2}}(\widehat{r}_{i;f}^{(n)} - r_{i;f}^{(n)}) = -(\widehat{a}_i^{(n)} + \widehat{b}_i^{(n)})I(f|g) + o_P(1) \ , \quad n \to \infty \ , \tag{6.1}$$

where $\widehat{a}_i^{(n)}$ *and* $\widehat{b}_i^{(n)}$ *result from substituting* $n^{1/2}(\widehat{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta})$ *for* $\boldsymbol{\tau} = (\boldsymbol{\gamma}', \boldsymbol{\delta}')'$ *in (4.4). The same result also holds for* $\widehat{r}_{i;f}^{(n)+}$.

It follows that $\widehat{r}_{i;f}^{(n)}$ *cannot be asymptotically invariant:* $I(f|g)$ *indeed is not a distribution–free quantity. Nor can* $\widehat{r}_{i;f}^{(n)+}$ *be.*

Though however no aligned–rank autocorrelation individually is asymptotically invariant, some specific linear combinations of them are. More precisely, denote by $\mathbf{T}^{(n)}_{\psi;f}$ the locally (at $\boldsymbol{\theta}$) sufficient statistic associated with the fundamental system $\{\psi_t^{(1)},\ldots,\psi_t^{(p_1+q_1)}\}$, by $\widehat{\psi}_t^{(j)}$ the value at t of the solution of $(\Sigma_i\widehat{A}_iL^i)(\Sigma_i\widehat{B}_iL^i)\psi_t=0$ characterized by the same starting values $\widehat{\psi}_s^{(j)}=\psi_s^{(j)}$, $s=1,2,\ldots,p_1+q_1$ as $\psi_t^{(j)}$, and by $\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f}$ the locally (at $\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}^{(n)}$) sufficient statistic associated with the fundamental system $\{\widehat{\psi}_t^{(n)},\ldots,\widehat{\psi}_t^{(p_1+q_1)}\}$ but computed from the aligned ranks $\widehat{R}_t^{(n)}\cdot\widehat{\mathbf{T}}^{(n)+}_{\widehat{\psi};f}$ can be defined similarly in the signed rank case.

The asymptotic mean under $H_g^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau})$ of $n^{1/2}\mathbf{T}^{(n)}_{\psi;f}$ is $\boldsymbol{\mu}_{\psi;f}(\boldsymbol{\tau})I(f|g)$ (Proposition 4.4) where $\boldsymbol{\mu}_{\psi;f}$ is a linear transform (of rank $\pi+p_1+q_1$) of $\boldsymbol{\tau}$, depending on $\boldsymbol{\theta}$ and the choice of the fundamental system $\{\psi_t^{(j)}\}$, but not on g. When $\boldsymbol{\tau}$ unrestrictedly takes its values in $\mathbb{R}^{p_2+q_2}$,$\boldsymbol{\mu}_{\psi;f}$ can take any value in $\mathbb{R}^{\pi+p_1+q_1}$. If $\boldsymbol{\tau}$ is of the form $(\gamma_1,\ldots,\gamma_{p_1},0,\ldots,0,\delta_1,\ldots,\delta_{q_1},0,\ldots,0)'$, where $(\gamma_1,\ldots,\gamma_{p_1},\delta_1,\ldots,\delta_{q_1})'$ belongs to $\mathcal{E}_r$, thus satisfying p_1+q_1-r independent linear restrictions, it can be shown that $\boldsymbol{\mu}_{\psi;f}(\boldsymbol{\tau})$ also satisfies $\pi+p_1+q_1-r$ independent linear restrictions, of the form

$$\boldsymbol{\Omega}(\boldsymbol{\theta})\boldsymbol{\mu}_{\psi;f}=\mathbf{0}\ , \tag{6.2}$$

where $\boldsymbol{\Omega}(\boldsymbol{\theta})$, a $(\pi+p_1+q_1-r)\times(\pi+p_1+q_1)$ matrix of rank $\pi+p_1+q_1-r=\max(p_2+q_1,p_1+q_2)-r$, is a continuous function of $\boldsymbol{\theta}$. The vector $\boldsymbol{\mu}_{\psi;f}$ accordingly lies in a r-dimensional linear subspace of $\mathbb{R}^{\pi+p_1+q_1}$.

The following result can then be established.

PROPOSITION 6.2. *For any $\boldsymbol{\theta}\in\mathcal{E}_r$, any fundamental system $\{\psi_t^{(i)}\}$ and any f ,*

$$\boldsymbol{\Omega}(\widehat{\boldsymbol{\theta}}^{(n)})\left[\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f}-\mathbf{T}^{(n)}_{\psi;f}\right]=o_P(n^{-\frac{1}{2}}) \tag{6.3}$$

under $H^{(n)}(\boldsymbol{\theta})$, $\boldsymbol{\theta}\in\mathcal{E}_r$, as $n\to\infty$. Accordingly, $\boldsymbol{\Omega}(\widehat{\boldsymbol{\theta}}^{(n)})\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f}$ is a vector of $\pi+p_1+q_1-r$ asymptotically invariant statistics (under $H^{(n)}(\boldsymbol{\theta})$).

A similar result can be proved, under $H_+^{(n)}(\boldsymbol{\theta})$, for $\boldsymbol{\Omega}(\widehat{\boldsymbol{\theta}}^{(n)})\widehat{\mathbf{T}}^{(n)+}_{\widehat{\psi};f}$.

6.3 Locally asymptotically similar tests. The following definition of local asymptotic similarity has been proposed by LeCam (1960) (under the terminology *differential asymptotic similarity*). A sequence $\phi^{(n)}$ of tests is said to be locally asymptotically similar under $H_f^{(n)}(\mathcal{E}_r)$, at asymptotic probability level α, if for every $\boldsymbol{\theta}\in\mathcal{E}_r$ and every bounded $B\subset\mathcal{E}_r$

$$\lim_{n\to\infty}\sup_{\boldsymbol{\tau}\in B}\left|E_{H_f^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau})}(\phi^{(n)})-\alpha\right|=0\ . \tag{6.4}$$

LeCam (loc. cit.) then shows how locally optimal tests, asymptotically enjoying, within the class of locally asymptotically similar tests, the same optimal properties (stringency or minimaxity) as the usual normal theory tests for Gaussian linear hypotheses, follow on applying the Gaussian likelihood ratio principle at the local asymptotical level.

For the particular problem of testing $H_f^{(n)}(\mathcal{E}_r)$, this leads to a test statistic of the form

$$
\begin{aligned}
Q_f^{(n)} &= \inf_{\widetilde{\boldsymbol{\mu}}:\boldsymbol{\Omega}(\widehat{\boldsymbol{\theta}}^{(n)})\widetilde{\boldsymbol{\mu}}=\mathbf{0}} [(n^{\frac{1}{2}}\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f} - \widetilde{\boldsymbol{\mu}})'\widehat{\mathbf{W}}^{-1}_{\widehat{\psi}}(n^{\frac{1}{2}}\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f} - \widetilde{\boldsymbol{\mu}})] \\
&= n(\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f})'\left[\widehat{\mathbf{W}}^{-1}_{\widehat{\psi}} - \widehat{\mathbf{W}}^{-1}_{\widehat{\psi}}\mathbf{K}(\mathbf{K}'\widehat{\mathbf{W}}^{-1}_{\widehat{\psi}}\mathbf{K})^{-1}\mathbf{K}'\widehat{\mathbf{W}}^{-1}_{\widehat{\psi}}\right]\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f} ,
\end{aligned}
\tag{6.5}
$$

where $\mathbf{K} = \mathbf{K}(\widehat{\boldsymbol{\theta}}^{(n)})$ is any $(\pi + p_1 + q_1) \times r$ full rank matrix such that $\boldsymbol{\Omega}(\widehat{\boldsymbol{\theta}}^{(n)})\mathbf{K} = \mathbf{0}$ (so that $\boldsymbol{\Omega}(\widehat{\boldsymbol{\theta}}^{(n)})\widetilde{\boldsymbol{\mu}} = \mathbf{0}$ if and only if $\widetilde{\boldsymbol{\mu}} = \mathbf{K}\boldsymbol{\lambda}$ for some $\boldsymbol{\lambda} \in \mathbb{R}^r$, and $\mathbf{N}\ \mathbf{K} = \mathbf{0}$ if and only if $\mathbf{N} = \mathbf{M}\boldsymbol{\Omega}(\widehat{\boldsymbol{\theta}}^{(n)})$ for some matrix $\mathbf{M}$ of appropriate dimension).

Denoting by $\widehat{\mathbf{W}}^{\frac{1}{2}}_{\widehat{\psi}}$ an arbitrary symmetric square root of $\widehat{\mathbf{W}}_{\widehat{\psi}}$, (6.5) also takes the form

$$
\begin{aligned}
&n(\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f})'\widehat{\mathbf{W}}^{-\frac{1}{2}}_{\widehat{\psi}}\left[\mathbf{I} - \widehat{\mathbf{W}}^{-\frac{1}{2}}_{\widehat{\psi}}\mathbf{K}(\mathbf{K}'\widehat{\mathbf{W}}^{-1}_{\widehat{\psi}}\mathbf{K})^{-1}\mathbf{K}'\widehat{\mathbf{W}}^{-\frac{1}{2}}_{\widehat{\psi}}\right]\widehat{\mathbf{W}}^{-\frac{1}{2}}_{\widehat{\psi}}\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f} \\
&\quad = n(\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f})'\widehat{\mathbf{W}}^{-\frac{1}{2}}_{\widehat{\psi}}\mathbf{P}\widehat{\mathbf{W}}^{-\frac{1}{2}}_{\widehat{\psi}}\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f} ,
\end{aligned}
\tag{6.6}
$$

where $\mathbf{P}$ is a symmetric, idempotent matrix of rank $\pi + p_1 + q_1 - r$. $Q_f^{(n)}$ therefore is asymptotically chi–square under $H_f^{(n)}(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \mathcal{E}_r$. But, since $\mathbf{P}\ \widehat{\mathbf{W}}^{-\frac{1}{2}}_{\widehat{\psi}}\mathbf{K} = \mathbf{0}$, $Q_f^{(n)}$ is also of the form

$$
n(\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f})'\boldsymbol{\Omega}'(\widehat{\boldsymbol{\theta}}^{(n)})\mathbf{M}'\ \mathbf{M}\boldsymbol{\Omega}(\widehat{\boldsymbol{\theta}}^{(n)})\widehat{\mathbf{T}}^{(n)}_{\widehat{\psi};f}
$$

(for some matrix $\mathbf{M}$, the explicit form of which is not needed here); hence, in view of Proposition 6.2, $\mathbf{Q}_f^{(n)}$ is asymptotically invariant under $H^{(n)}(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \mathcal{E}_r$. Accordingly, $Q_f^{(n)}$ is also asymptotically chi–square, with $\pi + p_1 + q_1 - r$ degrees of freedom, under any hypothesis of the form $H_g^{(n)}(\boldsymbol{\theta})$, hence under $H^{(n)}(\boldsymbol{\theta})$.

A similar result holds for the signed version $Q_f^{(n)+}$ of $Q_f^{(n)}$.

6.4 Optimal aligned rank tests. Summing up the findings of Sections 6.2 and 6.3, we may state the following result.

PROPOSITION 6.3. *The sequence of aligned–rank tests rejecting $H^{(n)}(\mathcal{E}_r)$ whenever*

$$
Q_f^{(n)} > \chi^2_{1-\alpha} ,
$$

where $\chi^2_{1-\alpha}$ denotes the $(1-\alpha)$–quantile of a chi–square variable with $\pi + p_1 + q_1 - r$ degrees of freedom and $Q_f^{(n)}$ is given in (6.5)

(i) is asymptotically invariant

(ii) is locally asymptotically similar under $H^{(n)}(\mathcal{E}_r)$, at probability level α

(iii) is asymptotically most stringent against $\bigcup\{H_f^{(n)}(\boldsymbol{\theta}+n^{-\frac{1}{2}}\boldsymbol{\tau})|\boldsymbol{\theta}\in\mathcal{E}_r,\ \boldsymbol{\tau}\notin\mathcal{E}_r\}$, within the class of all asymptotically similar (under $H_f^{(n)}(\mathcal{E}_r)$, at probability level α) tests.

A similar proposition holds in the signed–rank case.

The test thus proposed possesses all the asymptotic optimality properties one can expect when testing $H_f^{(n)}(\mathcal{E}_r)$ (with specified innovation density f). In addition, it is asymptotically invariant, similar and distribution–free under the much broader null hypothesis $H^{(n)}(\mathcal{E}_r)$ (with unspecified innovation density); it has much better robustness features than its usual Gaussian parametric competitors, with ARE values (with respect to the latter) as shown in Table 5.1. The amount of computation involved is not heavier than in the case e.g. of traditional Lagrange multiplier tests (see also Kreiss (1990 a) for optimal parametric AR procedures).

For all these reasons, they should be very attractive in a variety of practical problems when innovation densities are suspected to be severely non–Gaussian, or contaminated by possible outliers. For instance, it advantageously could be substituted for parametric Lagrange multiplier statistics in Pötscher (1983, 1985)'s recursive identification procedure.

6.5 Example: testing AR (1) versus ARMA (2,1) dependence. As an illustration, consider the problem of testing a null hypothesis of first–order autoregressive dependence

$$X_t-\rho X_{t-1}=\varepsilon_t\ ,\quad 0<|\rho|<1\ ,$$

where the parameter ρ and density of ε_t remain unspecified. Denote by $\widehat{\rho}^{(n)}$ a root n consistent estimate of ρ (the least square estimate, for instance), and let $\widehat{Z}_t^{(n)}=X_t^{(n)}-\widehat{\rho}^{(n)}X_{t-1}^{(n)}$, $t=1,\ldots,n$ (for simplicity, assume that $X_0^{(n)}$ is available). Assume that an ARMA (2,1) alternative is considered. Suppose however that, for some reason, one is willing to be particularly powerful against ARMA processes with innovation density f (f specified up to a scale transformation). The (aligned) ranks $\widehat{R}_t^{(n)}$ here are those of the residuals $\widehat{Z}_t^{(n)}$, and the autocorrelation coefficients to be used are the (aligned) f–rank autocorrelations $\widehat{r}_{i;f}^{(n)}$.

A simple fundamental system (a fundamental system here consists of any nonidentically zero solution, since the dimension of the solution space is one) of solutions of $(1-\widehat{\rho}^{(n)}L)\psi_t=0$ is $\widehat{\psi}_t=(\widehat{\rho}^{(n)})^{t-2}$, and the corresponding (aligned) locally sufficient statistic is ($p_1=1$, $q_1=0$; hence $\pi=1$ and $r=1$)

$$n^{\frac{1}{2}}\widehat{\mathbf{T}}_{\widehat{\rho};f}^{(n)}=\begin{pmatrix}(n-1)^{\frac{1}{2}}\ \widehat{r}_{1;f}^{(n)}\\ \sum_{j=2}^{n-1}(n-j)^{\frac{1}{2}}(\widehat{\rho}^{(n)})^{j-2}\widehat{r}_{j;f}^{(n)}\end{pmatrix}\ .$$

The corresponding, *exact* $\mathbf{T}^{(n)}_{\psi;f}$ (not a statistic since ρ is unknown) is

$$n^{\frac{1}{2}}\mathbf{T}^{(n)}_{\rho;f} = \begin{pmatrix} (n-1)^{\frac{1}{2}}\ r^{(n)}_{1;f} \\ \\ \sum_{j=2}^{n-1}(n-j)^{\frac{1}{2}}\ \rho^{j-2} r^{(n)}_{j;f} \end{pmatrix}$$

which is asymptotically bivariate normal, with mean zero and covariance matrix

$$\mathbf{W}_{\psi} = \begin{pmatrix} 1 & 0 \\ \\ 0 & (1-\rho^2)^{-1} \end{pmatrix}$$

under $H^{(n)}(\rho)$.

Local ARMA (2,1) alternatives $H^{(n)}_g(\rho, \boldsymbol{\gamma}, \delta)$ are of the form

$$X_t - (\rho + n^{-\frac{1}{2}}\gamma_1)X_{t-1} - n^{-\frac{1}{2}}\gamma_2 X_{t-2} = \varepsilon_t + n^{-\frac{1}{2}}\delta\ \varepsilon_{t-1}\ ,$$

where $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)'$ and the innovation density is g ; $n^{\frac{1}{2}}\mathbf{T}^{(n)}_{\psi;f}$ under such an alternative is asymptotically bivariate normal, still with covariance matrix $\mathbf{W}$, and with mean

$$\boldsymbol{\mu}_{\rho;f}(\boldsymbol{\gamma}, \delta)I(f|g) = \begin{pmatrix} \gamma_1 + \delta \\ \\ [\rho(\gamma_1 + \delta) + \gamma_2]/(1-\rho^2) \end{pmatrix} I(f|g)$$

since $$a_i = \begin{cases} \gamma_1 & \text{if} \quad i = 1 \\ \\ \rho^{i-2}(\rho\gamma_1 + \gamma_2) & \text{if} \quad i > 1 \end{cases} \quad \text{and} \quad b_i = \delta\rho^{i-1}\ ,\ i = 1, 2, \ldots\ .$$

Clearly, if $\gamma_2 = \delta = 0$, $\boldsymbol{\mu}_{\rho;f}(\boldsymbol{\gamma}, \delta)$ is of the form $\gamma_1(1 \quad \rho/(1-\rho^2))'$, and accordingly satisfies the (unique, since $\pi + p_1 + q_1 - r = 1 + 1 + 0 - 1 = 1$) linear constraint

$$(-\rho \qquad 1-\rho^2)\boldsymbol{\mu}_{\rho;f}(\boldsymbol{\gamma}, \delta) = 0\ .$$

Here $\boldsymbol{\Omega}(\boldsymbol{\theta}) = \boldsymbol{\Omega}(\rho)$ is the 1×2 row matrix $(-\rho \quad 1-\rho^2)$. It follows that

$$\begin{aligned} &(-\widehat{\rho}^{(n)} \qquad 1 - (\widehat{\rho}^{(n)})^2)\ \widehat{\mathbf{T}}^{(n)}_{\rho;f} \\ &\quad = -(n-1)^{\frac{1}{2}}\widehat{\rho}^{(n)}\widehat{r}^{(n)}_{1;f} + [1 - (\widehat{\rho}^{(n)})^2]\sum_{j=2}^{n-1}(n-j)^{\frac{1}{2}}(\widehat{\rho}^{(n)})^{j-2}\widehat{r}^{(n)}_{j;f} \end{aligned} \tag{6.7}$$

is asymptotically invariant, and asymptotically equivalent to its exact, non-aligned counterpart. The optimal, asymptotically similar aligned rank test of Proposition 6.3 consists in rejecting the null hypothesis whenever the quadratic statistic

$$Q^{(n)}_f = \left\{ -(n-1)^{\frac{1}{2}}\widehat{\rho}^{(n)}\widehat{r}^{(n)}_{1;f} + [1 - (\widehat{\rho}^{(n)})^2]\sum_{j=2}^{n-1}(n-j)^{\frac{1}{2}}(\widehat{\rho}^{(n)})^{j-2}\ \widehat{r}^{(n)}_{j;f} \right\}^2 \tag{6.8}$$

exceeds the $(1-\alpha)$ quantile of a chi–square variable with one degree of freedom (since the asymptotic variance of (6.7) is one).

The form of the test statistic (6.8) is not quite familiar, and its relation with classical, parametric time series procedures does not straightforwardly appear from (6.8). Denote by $Q_{\mathcal{N}}^{(n)}$ the Gaussian counterpart of $Q_f^{(n)}$

$$Q_{\mathcal{N}}^{(n)} = \left\{ -(n-1)^{\frac{1}{2}} \widehat{\rho}^{(n)} \widehat{r}_1^{(n)} + [1-(\widehat{\rho}^{(n)})^2] \sum_{j=2}^{n-1} (n-j)^{\frac{1}{2}} (\widehat{\rho}^{(n)})^{j-2} \widehat{r}_j^{(n)} \right\}^2 , \tag{6.9}$$

where $\widehat{\rho}^{(n)}$ denotes the Gaussian maximum likelihood estimator of the AR (1) parameter ρ, and $\widehat{r}_j^{(n)}$ stands for the corresponding classical residual autocorrelations. The Gaussian Lagrange multiplier approach to the problem (Godfrey 1979; Hosking 1980) leads to the test statistic (Hosking 1980, Theorem 1)

$$Q_{\mathcal{L}}^{(n)} = n \left\{ \sum_{j=2}^{n-1} (\widehat{\rho}^{(n)})^{j-2} \widehat{r}_j^{(n)} \right\}^2 , \tag{6.10}$$

to be compared also with the quantiles of a chi–square variable with one degree of freedom.

$Q_{\mathcal{N}}^{(n)}$ and $Q_{\mathcal{L}}^{(n)}$ apparently are distinct test statistics. However, it follows from asymptotic linear relationships among (parametric) estimated residual autocorrelations (McLeod 1978, Theorem 1) that

$$\sum_{j=1}^{n-1} (\widehat{\rho}^{(n)})^{j-1} \widehat{r}_j^{(n)} = o_P(n^{-\frac{1}{2}}) , \qquad n \to \infty . \tag{6.11}$$

Accordingly, our locally optimal Gaussian statistic $Q_{\mathcal{N}}^{(n)}$ satisfies

$$Q_{\mathcal{N}}^{(n)} = n(\widehat{r}_1^{(n)})^2 + o_P(1) , \qquad n \to \infty . \tag{6.12}$$

Similarly, for the Lagrange multiplier statistic, (6.11) implies

$$Q_{\mathcal{L}}^{(n)} = n(\widehat{r}_1^{(n)})^2 + o_P(1) , \qquad n \to \infty , \tag{6.13}$$

which in turn entails

$$Q_{\mathcal{N}}^{(n)} - Q_{\mathcal{L}}^{(n)} = o_P(1) , \qquad n \to \infty .$$

The parametric, Gaussian counterpart of our rank–based statistic (6.8) is thus asymptotically equivalent to the Gaussian Lagrange multiplier test statistic. It follows that the

ARE's of rank tests based on (6.8) with respect either to their Gaussian parametric counterpart (based on (6.9)) or the more familiar Lagrange multiplier test (based on (6.10)) are still those provided in Table 5.1.

Whether an asymptotic equivalence of the form (6.11) or (6.12) also holds for aligned rank–based residual autocorrelations is not known. The asymptotic invariance of (6.7) implies that (6.8) is asymptotically equivalent to its exact unaligned version $\overline{Q}_f^{(n)}$, say. The van der Waerden version of the latter, under Gaussian densities, in turn is asymptotically equivalent to the parametric $Q_{\mathcal{N}}^{(n)}$, hence to $n(\widehat{r}_1^{(n)})^2$. Note however that, whereas $Q_{vdW}^{(n)}$, $Q_{\mathcal{N}}^{(n)}$ and $Q_{\mathcal{L}}^{(n)}$ are asymptotically equivalent to the corresponding quantities computed from exact residuals, these equivalences do not hold for $n(\widehat{r}_1^{(n)})^2$ which is asymptotically distinct from $n(r_1^{(n)})^2$, where $r_1^{(n)}$ denotes the exact first–order residual autocorrelation. Nor is $\widehat{r}_{1;vdW}^{(n)}$ asymptotically invariant. The linear relation (6.11) indeed does not hold for exact residual autocorrelations.

Acknowledgement. This work has greatly benefited from the pleasant environment and peaceful landscapes of La Martinie-Sergeac. Marc Hallin would like to thank Mr. and Mrs. Baudoux for their warm hospitality during the summer of 1990, as well as Dominique, Marie, Catherine and Françoise for their enduring patience.

REFERENCES

AIYAR, R.J. (1981), *Asymptotic efficiency of rank tests of randomness against autocorrelation*, Ann. Inst. Statist. Math **33 A**, 255–262.

AIYAR, R.J., GUILLIER, C.L., and ALBERS, W. (1979), *Asymptotic relative efficiencies of rank tests for trend alternatives*, Jour. Amer. Statist. Assoc. **74**, 226–231.

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

ALBERS, W. (1978), *One–sample rank tests under autoregressive dependence*, Ann. Statist. **6**, 836–845.

BARTELS, R. (1982), *The rank version of von Neumann's ratio test for randomness*, Jour. Amer. Statist. Assoc. **77**, 40–46.

BENGHABRIT, Y. and HALLIN, M. (1992), *Optimal rank–based tests against first–order superdiagonal bilinear dependence*, Jour. Statist. Plan. and Inf., to appear.

BERAN, R.J. (1972), *Rank spectral processes and tests for serial dependence*, Ann. Math. Statist. **43**, 1749–1766.

BHATTACHARYYA, G.K. (1984), *Tests for randomness against trend or serial correlations*, In Handbook of Statistics, **4**, (P.R. Krishnaiah and P.K. Sen, eds.), 89–111. North–Holland, Amsterdam and New York.

BROCKWELL, P.J., and DAVIS, R.A. (1987), *Time Series: Theory and Applications*, Springer Verlag, New York.

CAMPBELL, P.J., and DUFOUR, J.–M. (1991), *Over-rejections in rational expectations models: a non-parametric approach to the Mankiw–Shapiro problem*, Economics Letters **35**, 285–290.

DAVID, F.N. (1947), *A power function for tests of randomness in a sequence of alternatives*, Biometrika **34**, 335–339.

DAVID, F.N. and FIX, E. (1966), *Randomization and the serial correlation coefficient*, In Research Papers in Statistics, Festschrift for J. Neyman, (F.N. David, ed.), 461–468. J. Wiley, New York.

DAVIES, R.B. (1973), *Asymptotic inference in stationary Gaussian time series*, Adv. Appl. Prob. **5**, 469–497.

DUFOUR, J.–M. (1981), *Rank tests for serial dependence*, Jour. Time Series Anal. **2**, 117–128.

DUFOUR, J.–M. and HALLIN, M. (1987), *Tests non paramétriques optimaux pour le modèle autorégressif d'ordre un*, Ann. Econ. Stat. **5**, 411–434.

DUFOUR, J.–M. and HALLIN, M. (1990a), *Runs tests for ARMA dependence*, Technical Report, Département de Sciences Economiques, Université de Montréal, Montréal.

DUFOUR, J.–M. and HALLIN, M. (1990b), *An exponential bound for the permutational distribution of a first–order autocorrelation coefficient*, Statistique et Analyse des Données **15**, 45–56.

DUFOUR, J.–M., and HALLIN, M., (1991), *Improved Eaton bounds for linear combinations of bounded random variables, with statistical applications*, Technical Report, C.R.D.E., Université de Montréal, Montréal.

DUFOUR, J.–M., LEPAGE, Y. and ZEIDAN, H. (1982), *Nonparametric testing for time series: a bibliography*, Canad. Jour. Statist., **10**, 1–38.

DUFOUR, J.–M., and ROY, R. (1985), *Some robust exact results on sample autocorrelations and tests of randomness*, Journal of Econometrics **29** , 257–273 and **41**, 279–281.

DUFOUR, J.–M., and ROY, R.(1986), *Generalized portmanteau statistics and tests for randomness*, Communications in Statistics A, Theory and Methods **15**, 2953–2972.

DZHAPARIDZE, K.O. (1986), *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*, Springer Verlag, New York.

EDGINGTON, E.S. (1961), *Probability table for number of runs of signs of first differences in ordered series*, J. Amer. Statist. Assoc. **56**, 156–159.

FISHER, R.A. (1926), *On the random sequence*, Quart. J. Roy. Meteorol. Soc. **52**, 250.

GAREL, B. (1989), *The asymptotic distribution of the likelihood ratio for MA processes with a regression trend*, Statist. Decisions **7**, 167–184.

GHOSH, M.N. (1954), *Asymptotic distribution of serial statistics and applications to problems of nonparametric tests of hypotheses*, Ann. Math. Statist. **25**, 218–251.

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

GOODMAN, L.A. (1958), *Simplified runs tests and likelihood ratio tests for Markov chains*, Biometrika **45**, 181–197.

GOODMAN, L.A. and GRUNFELD, Y. (1961), *Some nonparametric tests for comovements between time series*, J. Amer. Statist. Assoc. **56**, 11–26.

GOVINDARAJULU, Z. (1983), *Rank tests for randomness against autocorrelated alternatives*, In Time Series Analysis: Theory and Practice 4. (O.D. Anderson, ed.) 65–73. North–Holland, Amsterdam and New York.

GRANGER, C.W.J. (1963), *A quick test for serial correlation suitable for use with non–stationary time series*, Jour. Amer. Statist. Assoc. **58**, 728–736.

GUPTA, G.D. and GOVINDARAJULU, Z. (1980), *Nonparametric tests of randomness against autocorrelated normal alternatives*, Biometrika **67**, 375–379.

HÁJEK, J. and ŠIDÁK, Z. (1967), *Theory of Rank Tests*, Academic Press, New York.

HALLIN, M. (1986), *Nonstationary q-dependent processes and time-varying moving-average models: invertibility properties and the forecasting problem*, Advances in Applied Probability **18**, 170–210.

HALLIN, M., INGENBLEEK, J.–FR. and PURI, M.L. (1985), *Linear serial rank tests for randomness against ARMA alternatives*, Ann. Statist. **13**, 1156–1181.

HALLIN, M., INGENBLEEK, J.–FR. and PURI, M.L. (1987), *Linear and quadratic serial rank tests for randomness against serial dependence*, Jour. Time Series Anal. **8**, 409–424.

HALLIN, M., INGENBLEEK, J.–FR. and PURI, M.L. (1989), *Asymptotically most powerful rank tests for multivariate randomness against serial dependence*, Jour. Multivariate Analysis **30**, 34–71.

HALLIN, M., LAFORET, A. and MÉLARD, G. (1990), *Distribution–free tests against serial dependence: signed or unsigned ranks?*, Jour. Statist. Plan. Inf. **24**, 151–165.

HALLIN, M., LEFEVRE, C. and PURI, M.L. (1988), *On time–reversibility and the uniqueness of moving average representations for non–Gaussian stationary time series*, Biometrika **75**, 170–171.

HALLIN, M. and MÉLARD, G. (1988), *Rank–based tests for randomness against first order serial dependence*, Jour. Amer. Statist. Assoc. **83**, 1117–1129.

HALLIN, M., MÉLARD, G. and MILHAUD, X. (1992), *Permutational extreme values of autocorrelation coefficients and a Pitman test against serial dependence*, Ann. Statist. **20**, to appear.

HALLIN, M. and PURI, M.L. (1988), *On locally asymptotically maximin tests for ARMA processes*, in Statistical Theory and Data Analysis II, K. Matusita, Ed., North–Holland, Amsterdam and New-York, 495–500.

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. (1991a), *Time series analysis via rank–order theory: signed–rank tests for ARMA models*, Jour. Multivariate Analysis **39**, 1–29.

HALLIN, M. and PURI, M.L. (1991b), *A multivariate Wald–Wolfowitz rank test against serial dependence*, Technical Report.

HALLIN, M. and PURI, M.L. (1991c), *Aligned rank tests for linear models with autocorrelated errors and time series with a linear trend*, Technical Report.

HAREL, M. (1988), *Weak convergence of multidimensional rank statistics under φ–mixing conditions*, Jour. Statist. Plan. Inf. **20**, 41–63.

HAREL, M. and PURI, M.L. (1989a), *Limiting behavior of U–statistics, V–statistics and one sample rank order statistics for nonstationary absolutely regular processes*, Jour. Multivariate Analysis **30**, 181–204.

HAREL, M. and PURI, M.L. (1989b), *Weak convergence of the U–statistic and weak invariance of the one sample rank order statistic for Markov processes and ARMA models*, Jour. Multivariate Analysis **31**, 258–265.

HAREL, M. and PURI, M.L. (1990a), *Limiting behavior of one sample rank–order statistics with unbounded scores for nonstationary absolutely regular processes*, Jour. Statist. Plan. Inf. **27**, 1–23.

HAREL, M. and PURI, M.L. (1990b), *Weak convergence of serial rank statistics under dependence with applications in time series and Markov processes*, Ann. Prob. **18**, 1361–1387.

HAREL, M. and PURI, M.L.(1990c), *Weak convergence of the serial linear rank statistic with unbounded scores and regression constants under mixing conditions*, Jour. Statist. Plan. Inf. **25**, 163–186.

HAREL, M. and PURI, M.L. (1991), *Weak invariance of the multidimensional rank statistic for nonstationary absolutely regular processes*, Jour. Multivariate Analysis **36**, 204–221.

HAREL, M. and PURI, M.L. (1992), *Weak convergence of the simple linear rank statistic under mixing conditions in the nonstationary case*, Teor. Veroyatnost. i Premenen., to appear.

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

JOGDEO, K. (1968), *Asymptotic normality in nonparametric methods*, Ann. Math. Statist. **39**, 905–922.

KENDALL, M., and STUART, A. (1968), *The advanced Theory of Statistics*, Vol. 3, Griffin, London.

KERMACK, W.O. and MCKENDRICK, A.G. (1937), *Tests for randomness in a series of numerical observations*, Proc. Roy. Soc. Edinburgh **57**, 228–240.

KNOKE, D.J. (1977), *Testing for randomness against autocorrelation: alternative tests*, Biometrika **64**, 523–529.

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

KREISS, J.–P. (1990a), *Testing linear hypotheses in autoregressions*, Annals of Statistics **18**, 1470–1482.

KREISS, J.–P. (1990b), *Local asymptotic normality for autoregression with infinite order*, Jour. of Statist. Plan. and Inf., **26**, 185–219.

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

LECAM, L.(1986), *Asymptotic Methods in Statistical Decision Theory*, Springer–Verlag, New York.

LEVENE, H. and WOLFOWITZ, J. (1944), *The covariance matrix of runs up and down*, Ann. Math. Statist. **15**, 58–69.

MANN, H.B. (1945), *Nonparametric test against trend*, Econometrica **13**, 245–259.

MC.LEOD, A.I. (1978), *On the distribution of residual autocorrelations in Box–Jenkins models*, J.R.S.S. ser. B, **40**, 296–302.

MOOD, A.M. (1940), *The distribution theory of runs*, Ann. Math. Statist. **11**, 367–392.

MOORE, G.H. and WALLIS, W.A. (1943), *Time–series significance tests based on signs of differences*, Jour. Amer. Statist. Assoc. **38**, 153–164.

NIEUWENHUIS, G. and RUYMGAART, F. (1990), *Some stochastic inequalities and asymptotic normality of serial rank statistics in general linear processes*, Jour. Statist. Plan. Inf. **25**, 53–80.

PHAM, D.T. (1987), *Exact maximum likelihood estimate and Lagrange multiplier test statistic for ARMA models*, Jour. Time Series Anal. **8**, 61–78.

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

PÖTSCHER, B.M. (1985), *The behaviour of the Lagrange multiplier test in testing the orders of an ARMA model*, Metrika, **32**, 129–150.

PURI, M.L. and SEN, P.K. (1971), *Nonparametric Methods in Multivariate Analysis*, Wiley, New York.

PURI, M.L. and SEN, P.K. (1985), *Nonparametric Methods in General Linear Models*, Wiley, New York.

SAVAGE, I.R. (1957), *Contributions to the theory of rank order statistics—the "trend" case*, Ann. Math. Statist. **28**, 968–977.

SERFLING, R. (1968), *The Wilcoxon two–sample statistic on strong mixing process*, Ann. Math. Statist. **39**, 1202–1209.

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

SWITZER, P. (1984), *Inference for spatial autocorrelation functions*, In Geostatistics for Natural Resources Characterization, (G. Verly et al., eds.). D. Reidel, Dordrecht,127–140.

TRAN, L.T. (1988), *Rank order statistics for time series models*, Annals Inst. Stat. Math. **40**, 247–260.

TRAN, L.T. (1990), *Rank statistics for testing randomness against serial dependence*, Jour. Statist. Plan. and Inf. **24**, 215–232.

WALD, A. (1943), *Tests of statistical hypotheses concerning several parameters when the number of observations is large*, Trans. Amer. Math. Soc. **54**, 426–482.

WALD, A. and WOLFOWITZ, J. (1943), *An exact test for randomness in the nonparametric case based on serial correlation*, Ann. Math. Statist. **14**, 378–388.

WALLIS, W.A. and MOORE, G.H. (1941), *A significance test for time series*, Jour. Amer. Statist. Assoc. **36**, 401–409.

WEISS, G. (1975), *Time-series reversibility of linear stochastic processes*, Jour. of Appl. Probability, **12**, 831–836.

WOLFOWITZ, J. (1944), *Asymptotic distribution of runs up and down*, Ann. Math. Statist. **15**, 163–172.

YANG, M.C.K. and SCHRECKENGOST, J.F. (1981), *Difference sign test for comovements between two time series*, Comm. Statist. Theor. Meth. A10 **4**, 355–369.

Recent IMA Preprints

#	Author/s Title
801	**Hi Jun Choe**, Regularity for solutions of nonlinear variational inequalities with gradient constraints
802	**Peter Shi and Yongzhi Xu**, Quasistatic linear thermoelasticity on the unit disk
803	**Satyanad Kichenassamy and Peter J. Olver**, Existence and non-existence of solitary wave solutions to higher order model evolution equations
804	**Dening Li**, Regularity of solutions for a two-phase degenerate Stefan Problem
805	**Marek Fila, Bernhard Kawohl and Howard A. Levine**, Quenching for quasilinear equations
806	**Yoshikazu Giga, Shun'ichi Goto and Hitoshi Ishii**, Global existence of weak solutions for interface equations coupled with diffusion equations
807	**Mark J. Friedman and Eusebius J. Doedel**, Computational methods for global analysis of homoclinic and hetero clinic orbits: a case study
808	**Mark J. Friedman**, Numerical analysis and accurate computation of heteroclinic orbits in the case of center manifolds
809	**Peter W. Bates and Songmu Zheng**, Inertial manifolds and inertial sets for the phase-field equations
810	**J. López Gómez, V. Márquez and N. Wolanski**, Global behavior of positive solutions to a semilinear equation with a nonlinear flux condition
811	**Xinfu Chen and Fahuai Yi**, Regularity of the free boundary of a continuous casting problem
812	**Eden, A., Foias, C., Nicolaenko, B. and Temam, R.**, Inertial sets for dissipative evolution equations Part I: Construction and applications
813	**Jose–Francisco Rodrigues and Boris Zaltzman**, On classical solutions of the two-phase steady-state Stefan problem in strips
814	**Viorel Barbu and Srdjan Stojanovic**, Controlling the free boundary of elliptic variational inequalities on a variable domain
815	**Viorel Barbu and Srdjan Stojanovic**, A variational approach to a free boundary problem arising in electro-photography
816	**B.H. Gilding and R. Kersner**, Diffusion-convection-reaction, free boundaries, and an integral equation
817	**Shoshana Kamin, Lambertus A. Peletier and Juan Luis Vazquez**, On the Barenblatt equation of elasto-plastic filtration
818	**Avner Friedman and Bei Hu**, The Stefan problem with kinetic condition at the free boundary
819	**M.A. Grinfeld**, The stress driven instabilities in crystals: mathematical models and physical manifestations
820	**Bei Hu and Lihe Wang**, A free boundary problem arising in electrophotography: solutions with connected toner region
821	**Yongzhi Xu, T. Craig Poling, and Trent Brundage**, Direct and inverse scattering of time harmonic acoustic waves in an inhomogeneous shallow ocean
822	**Steven J. Altschuler**, Singularities of the curve shrinking flow for space curves
823	**Steven J. Altschuler and Matthew A. Grayson**, Shortening space curves and flow through singularities
824	**Tong Li**, On the Riemann problem of a combustion model
825	**L.A. Peletier & W.C. Troy**, Self-similar solutions for diffusion in semiconductors
826	**C.J. van Duijn, L.A. Peletier & R.J. Schotting**, On the analysis of brine transport in porous media
827	**Minkyu Kwak**, Finite dimensional description of convective reaction-diffusion equations
828	**Minkyu Kwak**, Finite dimensional inertial forms for the 2D Navier–Stokes equations
829	**Victor A. Galaktionov and Sergey A. Posashkov**, On some monotonicity in time properties for a quasilinear parabolic equation with source
830	**Victor A. Galaktionov**, Remark on the fast diffusion equation in a ball
831	**Hi Jun Choe and Lihe Wang**, A regularity theory for degenerate vector valued variational inequalities
832	**Vladimir I. Oliker and Nina N. Uraltseva**, Evolution of nonparametric surfaces with speed depending on curvature, II. The mean curvature case.
833	**S. Kamin and W. Liu**, Large time behavior of a nonlinear diffusion equation with a source
834	**Shoshana Kamin and Juan Luis Vazquez**, Singular solutions of some nonlinear parabolic equations
835	**Bernhard Kawohl and Robert Kersner**, On degenerate diffusion with very strong absorption
836	**Avner Friedman and Fernando Reitich**, Parameter identification in reaction-diffusion models
837	**E.G. Kalnins, H.L. Manocha and Willard Miller, Jr.**, Models of q-algebra representations I. Tensor products of special unitary and oscillator algebras
838	**Robert J. Sacker and George R. Sell**, Dichotomies for linear evolutionary equations in Banach spaces
839	**Oscar P. Bruno and Fernando Reitich**, Numerical solution of diffraction problems: a method of variation of boundaries
840	**Oscar P. Bruno and Fernando Reitich**, Solution of a boundary value problem for Helmholtz equation via variation of the boundary into the complex domain
841	**Victor A. Galaktionov and Juan L. Vazquez**, Asymptotic behaviour for an equation of superslow diffusion. The Cauchy problem
842	**Josephus Hulshof and Juan Luis Vazquez**, The Dipole solution for the porous medium equation in several

space dimensions

843 **Shoshana Kamin and Juan Luis Vazquez**, The propagation of turbulent bursts

844 **Miguel Escobedo, Juan Luis Vazquez and Enrike Zuazua**, Source-type solutions and asymptotic behaviour for a diffusion-convection equation

845 **Marco Biroli and Umberto Mosco**, Discontinuous media and Dirichlet forms of diffusion type

846 **Stathis Filippas and Jong-Shenq Guo**, Quenching profiles for one-dimensional semilinear heat equations

847 **H. Scott Dumas**, A Nekhoroshev-like theory of classical particle channeling in perfect crystals

848 **R. Natalini and A. Tesei**, On a class of perturbed conservation laws

849 **Paul K. Newton and Shinya Watanabe**, The geometry of nonlinear Schrödinger standing waves

850 **S.S. Sritharan**, On the nonsmooth verification technique for the dynamic programming of viscous flow

851 **Mario Taboada and Yuncheng You**, Global attractor, inertial manifolds and stabilization of nonlinear damped beam equations

852 **Shigeru Sakaguchi**, Critical points of solutions to the obstacle problem in the plane

853 **F. Abergel, D. Hilhorst and F. Issard-Roch**, On a dissolution-growth problem with surface tension in the neighborhood of a stationary solution

854 **Erasmus Langer**, Numerical simulation of MOS transistors

855 **Haïm Brezis and Shoshana Kamin**, Sublinear elliptic equations in $\mathbb{R}^n$

856 **Johannes C.C. Nitsche**, Boundary value problems for variational integrals involving surface curvatures

857 **Chao–Nien Chen**, Multiple solutions for a semilinear elliptic equation on $\mathbb{R}^N$ with nonlinear dependence on the gradient

858 **D. Brochet, X. Chen and D. Hilhorst**, Finite dimensional exponential atttractor for the phase field model

859 **Joseph D. Fehribach**, Mullins-Sekerka stability analysis for melting-freezing waves in helium-4

860 **Walter Schempp**, Quantum holography and neurocomputer architectures

861 **D.V. Anosov**, An introduction to Hilbert's 21st problem

862 **Herbert E Huppert and M Grae Worster**, Vigorous motions in magma chambers and lava lakes

863 **Robert L. Pego and Michael I. Weinstein**, A class of eigenvalue problems, with applications to instability of solitary waves

864 **Mahmoud Affouf**, Numerical study of a singular system of conservation laws arising in enhanced oil reservoirs

865 **Darin Beigie, Anthony Leonard and Stephen Wiggins**, The dynamics associated with the chaotic of tangles two dimensional quasiperiodic vector fields: theory and applications

866 **Gui–Qiang Chen and Tai–Ping Liu**, Zero relaxation and dissipation limits for hyperbolic conservation laws

867 **Gui–Qiang Chen and Jian–Guo Liu**, Convergence of second–order schemes for isentropic gas dynamics

868 **Aleksander M. Simon and Zbigniew J. Grzywna**, On the Larché–Cahn theory for stress-induced diffusion

869 **Jerzy Łuczka, Adam Gadomski and Zbigniew J. Grzywna**, Growth driven by diffusion

870 **Mitchell Luskin and Tsorng-Whay Pan**, Nonplanar shear flows for nonaligning nematic liquid crystals

871 **Mahmoud Affouf**, Unique global solutions of initial-boundary value problems for thermodynamic phase transitions

872 **Richard A. Brualdi, Keith L. Chavey and Bryan L. Shader**, Rectangular L-matrices

873 **Xinfu Chen, Avner Friedman and Bei Hu**, The thermistor problem with zero–one conductivity II

874 **Raoul LePage**, Controlling a diffusion toward a large goal and the Kelly principle

875 **Raoul LePage**, Controlling for optimum growth with time dependent returns

876 **Marc Hallin and Madan L. Puri**, Rank tests for time series analysis a survey

877 **V.A. Solonnikov**, Solvability of an evolution problem of thermocapillary convection in an infinite time interval

878 **Horia I. Ene and Bogdan Vernescu**, Viscosity dependent behaviour of viscoelastic porous media

879 **Kaushik Bhattacharya**, Self-accommodation in martensite

880 **D. Lewis, T. Ratiu, J.C. Simo and J.E. Marsden**, The heavy top: a geometric treatment

881 **Leonid V. Kalachev**, Some applications of asymptotic methods in semiconductor device modeling

882 **David C. Dobson**, Phase reconstruction via nonlinear least-squares

883 **Patricio Aviles and Yoshikazu Giga**, Minimal currents, geodesics and relaxation of variational integrals on mappings of bounded variation

884 **Patricio Aviles and Yoshikazu Giga**, Partial regularity of least gradient mappings

885 **Charles R. Johnson and Michael Lundquist**, Operator matrices with chordal inverse patterns

886 **B.J. Bayly**, Infinitely conducting dynamos and other horrible eigenproblems

887 **Charles M. Elliott and Stefan Luckhaus**, 'A generalised diffusion equation for phase separation of a multi-component mixture with interfacial free energy'

888 **Christian Schmeiser and Andreas Unterreiter**, The derivation of analytic device models by asymptotic methods

889 **LeRoy B. Beasley and Norman J. Pullman**, Linear operators that strongly preserve the index of imprimitivity

890 **Jerry Donato**, The Boltzmann equation with lie and cartan

891 **Thomas R. Hoffend Jr., Peter Smereka and Roger J. Anderson**, Method for resolving the laser induced local heating of moving magneto-optical recording media