Affine-invariant Aligned Rank Tests for the Multivariate General Linear Model with VARMA Errors

Marc Hallin* and Davy Paindaveine*

I.S.R.O., E.C.A.R.E.S., and Département de Mathématique
Université Libre de Bruxelles, Belgium

Abstract

We develop optimal rank-based procedures for testing affine-invariant linear hypotheses on the parameters of a multivariate general linear model with elliptical VARMA errors. We propose a class of optimal procedures that are based either on residual (pseudo-)Mahalanobis signs and ranks, or on *absolute interdirections* and *lift-interdirection ranks*, i.e., on hyperplane-based signs and ranks. The Mahalanobis versions of these procedures are strictly affine-invariant, while the hyperplane-based ones are asymptotically affine-invariant. Both versions generalize the univariate signed rank procedures proposed by Hallin and Puri (1994), and are locally asymptotically most stringent under correctly specified radial densities. Their AREs with respect to Gaussian procedures are shown to be convex linear combinations of the AREs obtained in Hallin and Paindaveine (2002a, 2002b) for the pure location and purely serial models, respectively. The resulting test statistics are provided under closed form for several important particular cases, including multivariate Durbin-Watson tests, VARMA order identification tests, etc. The key technical result is a multivariate asymptotic linearity result proved in Hallin and Paindaveine (2004b).

AMS 1980 subject classification : 62M10, 62G10, 62H15, 62G35.

Key words and phrases : Multivariate ranks and signs, Affine-invariant inference, VARMA models.

1 Introduction.

1.1 Ranks, multivariate ranks, and time series analysis.

Time series analysis, certainly in the multivariate context, is deeply marked by explicit or implicit Gaussian assumptions. The pervasive supremacy of correlogram- or cross-correlation-based methods, for instance, is a direct consequence of such assumptions—although their validity does not necessarily require normality; see, e.g., Hallin and Werker (1999).

If Gaussian assumptions are abandoned, the model takes the form of a semiparametric model where the innovation density plays the role of a nuisance. In such a situation, it has been shown in Hallin and Werker (2003) that, under quite general assumptions, conditioning on residual ranks leads to semiparametrically efficient—hence, in adaptive models, parametrically efficient—inference methods. For univariate ARMA models (which are adaptive), this is the

*Research supported by a P.A.I. contract of the Belgian federal Government, and an A.R.C. contract of the Communauté française de Belgique.

strategy adopted in a series of papers (Hallin and Puri 1988, 1991, and 1994), where a fairly complete toolbox of optimal testing procedures based on ranks or signed ranks is constructed.

Besides their efficiency properties, rank tests enjoy highly desirable distribution-freeness (implying wider applicability, similarity, and unbiasedness) and robustness features. Such features are even more desirable in the multivariate context. Yet, and despite the recognized need for non-Gaussian and robust methods in the area, little progress had been made until recently, due, mainly, to the lack of an appropriate multivariate generalization of ranks and signs. Some efforts have been made in the late eighties (Hallin, Ingenbleek, and Puri 1989; see also Hallin and Puri 1995), extending to problems of serial dependence the componentwise-rank approach developed for models involving independent observations (see the monograph by Puri and Sen (1971) for an extensive account of these methods). Componentwise ranks however do not meet the invariance properties one would expect from an extension of univariate ranks (componentwise-rank statistics are not even distribution-free); and they do not yield the semiparametric efficiency benefits of univariate ranks.

Componentwise ranks are thus inadequate, and are to be abandoned. Several alternative concepts have been proposed, mainly by Randles, Hettmansperger, Oja, and their collaborators: see Marden (1999), or Oja (1999) for a review and exhaustive reference lists. The sign and/or rank tests described in Randles (1989, 2000), Peters and Randles (1990), Hettmansperger, Nyblom, and Oja (1994), Jan and Randles (1994), Möttönen, and Oja (1995), Hettmansperger, Möttönen, and Oja (1997), Randles and Um (1998), to quote only a few, are dealing with independent observations: location and analysis of variance models, essentially. They are mainly based on heuristic and robustness arguments, paying little attention to optimality.

Emphasizing invariance and optimality (in the Le Cam sense), we have started a systematic study of (signed) rank-based inference for general linear models with VARMA errors under elliptic innovation densities. This model of course encompasses all models that have been studied previously (one- and two-sample location, analysis of variance, regression), but also VARMA time series models. The ultimate objective is to provide locally asymptotically optimal tests for affine invariant linear hypotheses on the parameters of this very general model, based on *pseudo-Mahalanobis signs* or a modified version (*absolute interdirections*) of *Randles' interdirections*, and *pseudo-Mahalanobis ranks* or the Oja-Paindaveine *hyperplane-based ranks*—all generalizing the classical concept of signed ranks while preserving their role in semiparametric efficiency (see Section 4 for precise definitions). Achieving this objective requires a number of nontrivial intermediate steps; the present paper is the final one of a series that eventually achieves this objective.

A first step in that direction was taken in Hallin and Paindaveine (2002a and c) for the simple location problem (fully specified location under the null), and in Hallin and Paindaveine (2002b and 2004a) for the simple VARMA problem (fully specified VARMA equation under the null). There, the adequate rank-based test statistics (the nonserial and the serial ones) are derived for these two problems, asymptotic representation and asymptotic normality results are proved, and asymptotic relative efficiencies (with respect to classical Gaussian methods, such as Hotelling or the usual correlogram-based tests) are obtained. In more realistic problems, however, the value of the parameter is never fully specified under the null, and *aligned signs* and *ranks* have to be substituted for the *exact* ones that cannot be computed from the observations. Handling this alignment device requires an asymptotic linearity property that is established in Hallin and Paindaveine (2004b). Building on these previous results, we provide here the optimal rank-based tests for affine-invariant linear hypotheses, and characterize their relative (with respect to their

Gaussian counterparts) asymptotic performances.

1.2 Serial and nonserial rank-based test statistics.

A precise description of the test statistics to be used is difficult at this stage, as it involves a number of preliminary definitions, and this is postponed to Section 5. Very roughly, though, let us assume that the residuals (the innovations) $\mathbf{Z}_t$, $t = 1, \ldots, n$ can be computed from the observations: the ranks used throughout are a reconstruction of the ranks R_t of the moduli $(\mathbf{Z}_t'\mathbf{\Sigma}^{-1}\mathbf{Z}_t)^{1/2}$ of these residuals, in the metric defined by the shape matrix $\mathbf{\Sigma}$ of the underlying elliptical innovation density. Denoting by $\mathbf{U}_t$ the unit vectors pointing out into the direction of the *sphericized residuals* $\mathbf{\Sigma}^{-1/2}\mathbf{Z}_t$, the multivariate signs (pseudo-Mahalanobis signs or *absolute* interdirections) allow for reconstructing the cosines $(\mathbf{U}_t)_i$ of the angles between the residuals $\mathbf{Z}_t$ and the axes. These ranks and cosines allow for computing nonserial statistics of the form

$$\frac{1}{n-i}\sum_{t=i+1}^{n} J_0\Big(\frac{R_t}{n+1}\Big)\mathbf{U}_t\mathbf{x}_{t-i}',$$

and serial ones of the form (a sign-and-rank-based measure of residual cross-correlations at lag i)

$$\frac{1}{n-i}\sum_{t=i+1}^{n} J_1\Big(\frac{R_t}{n+1}\Big)J_2\Big(\frac{R_{t-i}}{n+1}\Big)\mathbf{U}_t\mathbf{U}_{t-i}'$$

(J_0, J_1, and J_2 denote adequate score functions; the $\mathbf{x}_t$'s are the covariates in the trend part of the model). Plugged into the adequate quadratic forms (depending on the null hypothesis to be tested), these statistics yield a rank-based version of the locally optimal test statistics derived from the local asymptotic normality (LAN) structure of the model under study.

The main problem, of course, is that the residuals $\mathbf{Z}_t$ cannot be recovered from the observations (the parameter of the model is not entirely specified under the null hypothesis, unlike in Hallin and Paindaveine 2004a), and that the shape matrix $\mathbf{\Sigma}$ in practice is not known. Those ranks and cosines accordingly cannot be computed. The pseudo-Mahalanobis ranks or the hyperplane-based Oja-Paindaveine ranks on one hand, the pseudo-Mahalanobis signs or the (absolute) interdirections on the other, are ingenious devices for reconstructing these ranks and cosines. Moreover, they are evaluated at estimated residuals (the *alignment* problem). The major part of the paper consists in proving that such a reconstruction is still possible.

The benefits of the methods we propose here are the same as those of using ranks in the traditional univariate setting: distribution-freeness or asymptotic distribution-freeness, robustness, and efficiency. It has been shown, for instance (Hallin and Paindaveine 2002a and b), that the celebrated Chernoff-Savage result that the van der Waerden version of our test statistics has asymptotic relative efficiency uniformly larger than or equal to one with respect to the everyday practice Gaussian procedures, still holds here, under arbitrary dimension.

1.3 Outline of the paper.

The paper is organized as follows. The first three sections are mainly preparation and notation. In Sections 1.4, we describe the model to be considered throughout. Section 1.5 discusses the linear null hypotheses we are testing; delicate identification problems indeed are to be fixed, due to the fact that the orders of VARMA models are not constant over linear restrictions of the parameter space. In Section 1.6, we describe three particular cases that will be treated in detail in the sequel: a multivariate version of the classical Durbin-Watson problem, the test of the order of a VAR model, and the detection of a switching location regime. Section 2 regroups, for convenient reference, all assumptions that are required at various places. In Section 3, we state the uniform local asymptotic normality (LAN) result we are considering throughout. The concepts of multivariate ranks and signs, and the signed rank statistics (serial and nonserial) we are using are described in Section 4, along with their asymptotic behavior and equivariance/invariance properties. Finally, in Section 5.1, we provide the exact form, and the asymptotic performance, of the optimal test statistics. For the purpose of comparison, the optimal parametric Gaussian procedures our signed rank tests are competing with are described in Section 5.2. Asymptotic relative efficiencies are derived in Section 5.3. Section 6 is devoted to a detailed study of the three particular cases introduced in Section 1.6. Proofs are concentrated in Section 7.

1.4 The multivariate general linear model with VARMA error terms.

The general model we are considering throughout this paper is

$$\mathbf{Y}^{(n)} = \mathbf{X}^{(n)}\boldsymbol{\beta} + \mathbf{U}^{(n)}, \tag{1}$$

where

$$\mathbf{X}^{(n)} := \begin{pmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,m} \\ \vdots & \vdots & & \vdots \\ x_{n,1} & x_{n,2} & \dots & x_{n,m} \end{pmatrix} := \begin{pmatrix} \mathbf{x}'_1 \\ \vdots \\ \mathbf{x}'_n \end{pmatrix} \quad \text{and} \quad \boldsymbol{\beta} := \begin{pmatrix} \beta_{1,1} & \beta_{1,2} & \dots & \beta_{1,k} \\ \vdots & \vdots & & \vdots \\ \beta_{m,1} & \beta_{m,2} & \dots & \beta_{m,k} \end{pmatrix} := \begin{pmatrix} \boldsymbol{\beta}'_1 \\ \vdots \\ \boldsymbol{\beta}'_m \end{pmatrix}$$

denote an $n \times m$ matrix of constants (the design matrix), and the $m \times k$ regression parameter, respectively. Instead of the traditional assumption that the error term

$$\mathbf{U}^{(n)} := \begin{pmatrix} U_{1,1} & U_{1,2} & \dots & U_{1,k} \\ \vdots & \vdots & & \vdots \\ U_{n,1} & U_{n,2} & \dots & U_{n,k} \end{pmatrix} := \begin{pmatrix} \mathbf{U}'_1 \\ \vdots \\ \mathbf{U}'_n \end{pmatrix}$$

is white noise, we rather assume $(\mathbf{U}_t,\ t = 1, \dots, n)$ to be a finite realization (of length n) of a solution of the multivariate linear stochastic difference equation of the form

$$\mathbf{A}(L)\,\mathbf{U}_t = \mathbf{B}(L)\,\boldsymbol{\varepsilon}_t, \qquad t \in \mathbb{Z}, \tag{2}$$

where $\mathbf{A}(L) := \mathbf{I}_k - \sum_{i=1}^{p_1} \mathbf{A}_i L^i$ and $\mathbf{B}(L) := \mathbf{I}_k + \sum_{i=1}^{q_1} \mathbf{B}_i L^i$ for some $(p_1 + q_1)$-tuple of $k \times k$ real matrices $(\mathbf{A}_1, \dots, \mathbf{A}_{p_1},\ \mathbf{B}_1, \dots, \mathbf{B}_{q_1})$; $\mathbf{I}_k$ stands for the k-dimensional identity matrix, L for the lag operator, and $\{\boldsymbol{\varepsilon}_t \,|\, t \in \mathbb{Z}\}$ is a k-dimensional white-noise process. Under this model, the observation

$$\mathbf{Y}^{(n)} := \begin{pmatrix} Y_{1,1} & Y_{1,2} & \dots & Y_{1,k} \\ \vdots & \vdots & & \vdots \\ Y_{n,1} & Y_{n,2} & \dots & Y_{n,k} \end{pmatrix} := \begin{pmatrix} \mathbf{Y}'_1 \\ \vdots \\ \mathbf{Y}'_n \end{pmatrix}$$

is the realization of a k-variate VARMA process $\{\mathbf{Y}_t,\ t \in \mathbb{Z}\}$, of orders smaller than or equal to p_1 and q_1, with trend $\mathrm{E}[\mathbf{Y}_t] = \boldsymbol{\beta}'\mathbf{x}_t$.

1.5 Linear hypotheses.

1.5.1 Linear restrictions, stationarity, invertibility, and identifiability.

Denote by

$$\boldsymbol{\theta} := \Big((\mathrm{vec}\,\boldsymbol{\beta}')', (\mathrm{vec}\,\mathbf{A}_1)', \ldots, (\mathrm{vec}\,\mathbf{A}_{p_1})', (\mathrm{vec}\,\mathbf{B}_1)', \ldots, (\mathrm{vec}\,\mathbf{B}_{q_1})'\Big)' \in \mathbb{R}^K,$$

where $K := km + k^2(p_1 + q_1)$, the parameter of the model described in the previous section. The null hypotheses we are considering are imposing some linear constraints on $\boldsymbol{\theta}$, of the form

$$\boldsymbol{\theta} \in \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}), \tag{3}$$

where $\mathcal{M}(\boldsymbol{\Upsilon})$ denotes the vector subspace of $\mathbb{R}^K$ spanned by the columns of some full-rank $(K \times r)$ matrix $\boldsymbol{\Upsilon}$, and $\boldsymbol{\theta}_0 \in \mathbb{R}^K$. Some precautions however are to be taken

(i) about the stationarity-invertibility-identifiability properties of the VARMA models characterized by $\boldsymbol{\theta} \in \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$, particularly when those linear restrictions imply actual orders p_0 and q_0 that are strictly less than p_1 and/or q_1;

(ii) about the affine-invariance properties of the linear restrictions to be tested.

For any linear restriction of the form (3), let

$$p_0 = p_0(\boldsymbol{\theta}_0, \boldsymbol{\Upsilon}) := \min_{0 \le p \le p_1} \Big\{ p \,\Big|\, (\boldsymbol{\theta}_0)_i = 0 = (\boldsymbol{\Upsilon})_{ij},\ km + k^2 p < i \le km + k^2 p_1,\ 1 \le j \le r \Big\}$$

and

$$q_0 = q_0(\boldsymbol{\theta}_0, \boldsymbol{\Upsilon}) := \min_{0 \le q \le q_1} \Big\{ q \,\Big|\, (\boldsymbol{\theta}_0)_i = 0 = (\boldsymbol{\Upsilon})_{ij},\ km + k^2(p_1 + q) < i \le K,\ 1 \le j \le r \Big\}.$$

Let us first assume that p_0 and q_0 are both strictly positive. Then, all $\boldsymbol{\theta}$ satisfying (3) are characterizing VARMA models with

$$(\mathbf{A}_1, \ldots, \mathbf{A}_{p_1}) = (\mathbf{A}_1, \ldots, \mathbf{A}_{p_0}, \mathbf{0}, \ldots, \mathbf{0}) \quad \text{and} \quad (\mathbf{B}_1, \ldots, \mathbf{B}_{q_1}) = (\mathbf{B}_1, \ldots, \mathbf{B}_{q_0}, \mathbf{0}, \ldots, \mathbf{0}), \tag{4}$$

where $\mathbf{A}_{p_0} \neq \mathbf{0} \neq \mathbf{B}_{q_0}$. This is not sufficient, however, for the corresponding VARMA model being a well-identified model of orders p_0 and q_0. Therefore, let $\boldsymbol{\Theta}_{p_0,q_0}$ denote the set of all $\boldsymbol{\theta}$'s such that (4) holds, and

(a) $|\mathbf{A}_{p_0}| \neq 0 \neq |\mathbf{B}_{q_0}|$;

(b) all solutions of the determinantal equations $\det(\mathbf{I}_k - \sum_{i=1}^{p_0} \mathbf{A}_i z^i) = 0$ and $\det(\mathbf{I}_k + \sum_{i=1}^{q_0} \mathbf{B}_i z^i) = 0$ lie outside the unit ball in $\mathbb{C}$;

(c) the greatest common left divisor of $\mathbf{I}_k - \sum_{i=1}^{p_0} \mathbf{A}_i z^i$ and $\mathbf{I}_k + \sum_{i=1}^{q_0} \mathbf{B}_i z^i$ is the identity matrix $\mathbf{I}_k$.

Write $(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$ for $(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})) \cap \boldsymbol{\Theta}_{p_0(\theta_0,\boldsymbol{\Upsilon}),q_0(\theta_0,\boldsymbol{\Upsilon})}$. The model characterized by (2) and $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$ then is a well-identified stationary and invertible VARMA model of orders p_0 and q_0 (see, for instance, Brockwell and Davis 1987 or Dunsmuir and Hannan 1976).

It may happen however that, for some $(\boldsymbol{\theta}_0, \boldsymbol{\Upsilon})$, the intersection $(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})) \cap \boldsymbol{\Theta}_{p_0(\theta_0,\boldsymbol{\Upsilon}),q_0(\theta_0,\boldsymbol{\Upsilon})}$ is empty. If, for instance, $\boldsymbol{\theta}_0 = \mathbf{0}$ and all entries, in $\boldsymbol{\Upsilon}$'s rows $km + k^2 p_0 - k + 1$ through $km + k^2 p_0$ are zero, we have, for all $\boldsymbol{\theta} \in \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$, $(\mathbf{A}_{p_0})_{1k} = \ldots = (\mathbf{A}_{p_0})_{kk} = 0$, so that $|\mathbf{A}_{p_0}| = 0$ and $(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^* = \emptyset$.

Finally, if $p_0 = 0$ and/or $q_0 = 0$, Model (2) for $\boldsymbol{\theta} \in \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$ either describes a pure MA(q_0), a pure VAR(p_0), or a white noise process. The definition of $\boldsymbol{\Theta}$ in such cases is adapted in an obvious way: (a), (b), and (c) either only apply to the MA(q_0) or the VAR(p_0) operators, or they are void.

1.5.2 Linear restrictions and affine invariance.

We conform the somewhat loose terminology used in the robust-statistic literature by calling *affine* a linear transformation (rather than the combination of a linear transformation and a translation), i.e., any transformation $\mathbf{x} \mapsto \mathbf{Mx}$ of $\mathbb{R}^k$, where $\mathbf{M}$ is a full-rank $k \times k$ matrix; *affine-invariance*, *affine-equivariance*, etc. throughout are to be understood with that particular acceptation.

Affine-invariant testing methods only can deal with affine-invariant null hypotheses. As the concepts of multivariate ranks and signs we are using are affine-invariant, the linear restrictions to be tested also should be invariant under affine transformations, in the following sense. For any $k \times k$ full-rank matrix $\mathbf{M}$, the affine transformation $\boldsymbol{\varepsilon}_t \mapsto \mathbf{M}\boldsymbol{\varepsilon}_t$ of the noise induces the transformation

$$(\boldsymbol{\beta}, \mathbf{A}_1, \ldots, \mathbf{A}_{p_1}, \mathbf{B}_1, \ldots, \mathbf{B}_{q_1}) \mapsto (\boldsymbol{\beta}\,\mathbf{M}', \mathbf{MA}_1\mathbf{M}^{-1}, \ldots, \mathbf{MA}_{p_1}\mathbf{M}^{-1}, \mathbf{MB}_1\mathbf{M}^{-1}, \ldots, \mathbf{MB}_{q_1}\mathbf{M}^{-1})$$

of the parameter. In terms of $\boldsymbol{\theta}$, this induced transformation is $\boldsymbol{\theta} \mapsto \mathbf{g}_{\mathbf{M}}^{(m,p_1+q_1)}\boldsymbol{\theta}$, where

$$\mathbf{g}_{\mathbf{M}}^{(r_1,r_2)} := \begin{pmatrix} \mathbf{I}_{r_1} \otimes \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{r_2} \otimes (\mathbf{M}'^{-1} \otimes \mathbf{M}) \end{pmatrix}.$$

Letting $\mathcal{G}_{r_2}^{r_1}(k) := \{\mathbf{g}_{\mathbf{M}}^{(r_1,r_2)}, \mathbf{M} \text{ of full-rank}\}$, we say that the linear restriction $\boldsymbol{\theta} \in \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$ is invariant under affine transformations iff

$$\mathbf{g}_{\mathbf{M}}^{(m,p_1+q_1)}\,(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})) = \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}), \quad \text{for all} \quad \mathbf{g}_{\mathbf{M}}^{(m,p_1+q_1)} \in \mathcal{G}_{p_1+q_1}^{m}(k). \tag{5}$$

Let $\mathbf{l}_k := \operatorname{vec}\mathbf{I}_k$, $\mathbf{L}_k := \frac{1}{k}\mathbf{l}_k\mathbf{l}_k'$, and denote by $\mathbf{P}_k$ the $k^2 \times (k^2-1)$ array obtained by deleting the last column in $\mathbf{I}_{k^2} - \mathbf{L}_k$. Then Hallin and Paindaveine (2003a) showed that the linear restriction $\boldsymbol{\theta} \in \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$ is invariant under affine transformations iff $\boldsymbol{\theta}_0$ and $\boldsymbol{\Upsilon}$ are of the form

$$\boldsymbol{\theta}_0 = \begin{pmatrix} \mathbf{0} \\ \mathbf{w} \otimes \mathbf{l}_k \end{pmatrix} + \boldsymbol{\Upsilon}\,\boldsymbol{\omega},$$

and

$$\boldsymbol{\Upsilon} = \begin{pmatrix} \boldsymbol{\Upsilon}_I & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Upsilon}_{II} \end{pmatrix} = \begin{pmatrix} \mathbf{Z} \otimes \mathbf{I}_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \otimes \mathbf{P}_k & \mathbf{W} \otimes \mathbf{l}_k \end{pmatrix} \mathbf{G},$$

respectively, where $\mathbf{w}$ and $\boldsymbol{\omega}$ denote arbitrary vectors with dimensions p_1+q_1 and r, respectively, $\mathbf{Z}$, $\mathbf{V}$, and $\mathbf{W}$ are (possibly void) full-rank matrices with dimensions $m \times r_Z$, $(p_1+q_1) \times r_V$, and $(p_1+q_1) \times r_W$, respectively, and (letting $r = r_I + r_{II}$, where $r_I := r_Z k$ and $r_{II} := r_V(k^2-1) + r_W$) $\mathbf{G}$ is an invertible $r \times r$ matrix. Since $\mathcal{M}(\boldsymbol{\Upsilon}) = \mathcal{M}(\boldsymbol{\Upsilon}\mathbf{G})$ for any such $\mathbf{G}$, we may assume, without loss of generality, that $\mathbf{G} = \mathbf{I}_r$ in the sequel. In case $p_0 < p_1$ and/or $q_0 < q_1$, the matrices $\mathbf{w}$, $\mathbf{V}$ and $\mathbf{W}$ have only zeros in rows $p_0+1, \ldots, p_1$ and rows $p_1+q_0+1, \ldots, p_1+q_1$. Finally, note that $(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$ is affine-invariant, in the sense that it also satisfies (5), iff $(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))$ is.

1.6 Three examples.

The class of affine-invariant linear restrictions covers a wide range of problems of practical interest. The following particular cases will be treated in detail in Section 6.

(a) The multivariate Durbin-Watson problem, which corresponds to $\boldsymbol{\theta}_0 = \mathbf{0}$, $\boldsymbol{\Upsilon}_I = \mathbf{I}_{km}$, and $\boldsymbol{\Upsilon}_{II} = \emptyset$, where $\emptyset$ denotes the void matrix. This consists in testing serial independence of the error term in an unspecified linear model versus VARMA errors of orders less than or equal to p_1 and q_1 (the linear model structure of the trend plays the role of the nuisance).

(b) Testing the orders of VARMA errors. In this second example, we consider the problem of testing a VARMA(p_0, q_0) model versus a higher-order VARMA(p_1, q_1). This is obtained by letting $\boldsymbol{\theta}_0 = \mathbf{0}$, $\boldsymbol{\Upsilon}_I = \emptyset$, and

$$\boldsymbol{\Upsilon}_{II} = \begin{pmatrix} \mathbf{I}_{k^2 p_0} & \mathbf{0}_{k^2 p_0 \times k^2 q_0} \\ \mathbf{0}_{k^2(p_1-p_0)\times k^2 p_0} & \mathbf{0}_{k^2(p_1-p_0)\times k^2 q_0} \\ \mathbf{0}_{k^2 q_0 \times k^2 p_0} & \mathbf{I}_{k^2 q_0} \\ \mathbf{0}_{k^2(q_1-q_0)\times k^2 p_0} & \mathbf{0}_{k^2(q_1-q_0)\times k^2 q_0} \end{pmatrix}$$

(here again, the linear model structure of the trend plays the role of the nuisance). The particular case where $p_1 - p_0 = q_1 - q_0 = 1$ plays an important role in several model identification procedures (see, e.g., Pötscher 1983, 1985, or Garel and Hallin 1999 for the univariate case). For the sake of notational simplicity, we restrict to $p_1 - p_0 = 1$, $q_1 = q_0 = 0$ in the sequel.

(c) Testing against switching location regime. Let $(t_i^{(n)})$, $i = 1, \ldots, m-1$, be a sequence of $(m-1)$-tuples such that $t_0^{(n)} := 0 < t_1^{(n)} < \ldots < t_{m-1}^{(n)} < t_m^{(n)} := n$ for all n. Denoting by $\mathbf{e}_i^{(m)}$ the ith vector of the canonical basis in $\mathbb{R}^m$, consider the design matrix defined by

$$\mathbf{x}_t^{(n)} = \mathbf{e}_i^{(m)}, \quad \text{for } t_{i-1}^{(n)} < t \leq t_i^{(n)}.$$

The resulting model is a VARMA(p_1, q_1) one, with time-dependent trend (more precisely, with mean $\boldsymbol{\beta}_i$ for t between $t_{i-1}^{(n)} + 1$ and $t_i^{(n)}$). In this setup, the testing problem associated with $\boldsymbol{\Upsilon}_I = (1, \ldots, 1)' \otimes \mathbf{I}_k$, $\boldsymbol{\Upsilon}_{II} = \mathbf{I}_{k^2(p_1+q_1)}$ corresponds to the problem of testing the absence of different regimes, i.e., to the null hypothesis under which $\boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_m$. The coefficients of the VARMA operators here are nuisance parameters. Note that if there is no serial component in the model, then this reduces to the standard m-sample problem, i.e., to the most standard testing problem in analysis of variance.

2 Main assumptions.

In this section we collect, for convenient reference, all assumptions we need in the sequel. These assumptions are dealing with the design of the trend part of the model, the linear restrictions to be tested, the innovation density, the score functions to be used in test statistics, and the estimators of unspecified and nuisance parameters.

2.1 Asymptotic behavior of covariates.

We begin with some structural conditions on the covariates involved in the trend part of the model. The following assumptions are standard in the context (see Garel and Hallin 1995).

Assumption (A1). Letting $\mathbf{C}_i^{(n)} := (n-i)^{-1} \sum_{t=i+1}^{n} \mathbf{x}_t^{(n)} \mathbf{x}_{t-i}^{(n)\prime}$, $i = 0, 1, \ldots, n-1$, denote by $\mathbf{D}^{(n)}$ the diagonal matrix with elements $(\mathbf{C}_0^{(n)})_{11}, \ldots, (\mathbf{C}_0^{(n)})_{mm}$.

(i) $(\mathbf{C}_0^{(n)})_{jj} > 0$ for all j.

(ii) Let $\mathbf{R}_i^{(n)} := (\mathbf{D}^{(n)})^{-1/2} \mathbf{C}_i^{(n)} (\mathbf{D}^{(n)})^{-1/2}$. The limits $\lim_{n\to\infty} \mathbf{R}_i^{(n)} =: \mathbf{R}_i$ exist for all i; $\mathbf{R}_0$ is positive definite, and therefore can be factorized into $\mathbf{R}_0 = (\mathbf{K}\,\mathbf{K}')^{-1}$ for some full-rank $m \times m$ matrix $\mathbf{K}$. Letting $\mathbf{K}^{(n)} := (\mathbf{D}^{(n)})^{-1/2} \mathbf{K}$ (defining $\mathbf{K}^{(n)}$, note that $\mathbf{K}^{(n)}$ also has full rank).

(iii) The classical Noether conditions hold : the $(\mathbf{x}_t^{(n)})_j$, $t = 1, \ldots, n$, are not all equal, and, letting $\bar{x}_j^{(n)} := n^{-1} \sum_{t=1}^n (\mathbf{x}_t^{(n)})_j$,

$$\lim_{n\to\infty} \frac{\max_{1\le t\le n} \left((\mathbf{x}_t^{(n)})_j - \bar{x}_j^{(n)}\right)^2}{\sum_{t=1}^n \left((\mathbf{x}_t^{(n)})_j - \bar{x}_j^{(n)}\right)^2} = 0, \quad j = 1, \ldots, m.$$

The description of the asymptotic behavior of the proposed test statistics under local alternatives will require the following reinforcement of (A1).

Assumption (A1′). Same as Assumption (A1), but we further assume that $\lim_{n\to\infty}[\mathbf{D}^{(n)}/\operatorname{tr}\mathbf{D}^{(n)}] =: \mathbf{D}^2$, where $\mathbf{D}$ is a finite, positive definite diagonal matrix.

2.2 Linear restrictions.

As discussed in Section 1.5, the linear restrictions to be considered should be compatible with stationarity, invertibility, and identifiability under the null, and should be affine-invariant. Using the notation and definitions of Section 1.5, these two requirements are summarized in the following assumption.

Assumption (A2). The linear restriction $\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$ is affine-invariant, and $(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^* \neq \emptyset$.

This assumption has crucial implications in the sequel: for instance, it allows for the existence of root-n consistent estimators $\hat{\boldsymbol{\theta}}^{(n)}$ of $\boldsymbol{\theta}$ under null hypotheses of the form $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$.

The same assumption also guarantees that the unobserved starting values, in (2), have no influence on asymptotic results. Denote by $\mathbf{G}_u(\boldsymbol{\theta})$, $u \in \mathbb{N}$, the Green's matrices associated with the autoregressive difference operator $\mathbf{A}(L) = \mathbf{I}_k - \sum_{i=1}^{p_0} \mathbf{A}_i L^i$. These matrices can be defined recursively by $\mathbf{A}(L)\mathbf{G}_u = \mathbf{G}_u - \sum_{i=1}^{\min(p_0,u)} \mathbf{A}_i \mathbf{G}_{u-i} = \delta_{u0}\,\mathbf{I}_k$, where $\delta_{u0} = 1$ if $u = 0$, and $\delta_{u0} = 0$ otherwise. Under null hypotheses of the form $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$, $\mathbf{G}_u(\boldsymbol{\theta})$ also can be defined by means of

$$\sum_{u=0}^{+\infty} \mathbf{G}_u z^u := \left(\mathbf{I}_k - \sum_{i=1}^{p_0} \mathbf{A}_i z^i\right)^{-1}, \qquad z \in \mathbb{C},\ |z| < 1. \tag{6}$$

Similarly, we denote by $\mathbf{H}_u(\boldsymbol{\theta})$, $u \in \mathbb{N}$, the Green's matrices associated with the moving average difference operator $\mathbf{B}(L)$. These matrices play a central role in the statement of the LAN structure of the model (Section 3). Clearly, they all are continuous functions of $\boldsymbol{\theta}$. When no confusion is possible, we will not stress their dependence on $\boldsymbol{\theta}$.

The residuals $(\mathbf{Z}_1^{(n)}(\boldsymbol{\theta}), \ldots, \mathbf{Z}_n^{(n)}(\boldsymbol{\theta}))$ associated with a value $\boldsymbol{\theta}$ of the parameter then can be computed from a set of initial values $\boldsymbol{\varepsilon}_{-q_0+1} \ldots, \boldsymbol{\varepsilon}_0, \mathbf{Y}_{-p_0+1}^{(n)}, \ldots, \mathbf{Y}_0^{(n)}$ and the observed series $(\mathbf{Y}_1^{(n)}, \ldots, \mathbf{Y}_n^{(n)})$ via the recursion (based on $\boldsymbol{\beta}$, the $\mathbf{A}_i$'s, $\mathbf{B}_j$'s, and $\mathbf{H}_u$'s associated with $\boldsymbol{\theta}$)

$$\mathbf{Z}_t^{(n)}(\boldsymbol{\theta}) = \sum_{i=0}^{t-1}\sum_{j=0}^{p_0} \mathbf{H}_i \mathbf{A}_j (\mathbf{Y}_{t-i-j}^{(n)} - \boldsymbol{\beta}' \mathbf{x}_{t-i-j}^{(n)}) \tag{7}$$
$$+(\mathbf{H}_{t+q_0-1} \ldots \mathbf{H}_t) \begin{pmatrix} \mathbf{I}_k & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{B}_1 & \mathbf{I}_k & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_{q_0-1} & \mathbf{B}_{q_0-2} & \ldots & \mathbf{I}_k \end{pmatrix} \begin{pmatrix} \boldsymbol{\varepsilon}_{-q_0+1} \\ \vdots \\ \boldsymbol{\varepsilon}_0 \end{pmatrix}.$$

For $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$, $\{\boldsymbol{\varepsilon}_t\}$ is $\{\mathbf{Y}_t\}$'s innovation process, and $\|\mathbf{H}_t\| = O(\Lambda^t)$ as $t \to \infty$, for some $0 < \Lambda < 1$, so that neither the (generally unobserved) values $(\boldsymbol{\varepsilon}_{-q_0+1}, \ldots, \boldsymbol{\varepsilon}_0)$ of the innovation, nor the initial values $(\mathbf{Y}^{(n)}_{-p_0+1}, \ldots, \mathbf{Y}^{(n)}_0)$, have any influence on asymptotic results; therefore, they all safely can be put to zero in the sequel.

2.3 Elliptically symmetric innovation density.

Throughout, we will assume that the density $\underline{f}$ of the noise $\{\boldsymbol{\varepsilon}_t\}$ is elliptically symmetric. Denote by $\boldsymbol{\Sigma}$ a symmetric positive definite $k \times k$ matrix, and let $f : \mathbb{R}^+_0 \to \mathbb{R}^+$ be such that $f > 0$ a.e. and $\int_0^\infty r^{k-1} f(r)\, dr < \infty$: we will assume throughout that $\{\boldsymbol{\varepsilon}^{(n)}_1, \ldots, \boldsymbol{\varepsilon}^{(n)}_n\}$ is a finite realization of an elliptic white noise process with shape matrix $\boldsymbol{\Sigma}$ and radial density f :

Assumption (B1). The innovation density is of the form

$$\underline{f}(\mathbf{z}; \boldsymbol{\Sigma}, f) := c_{k,f}\, (\det \boldsymbol{\Sigma})^{-1/2}\, f(\|\mathbf{z}\|_{\boldsymbol{\Sigma}}), \quad \mathbf{z} \in \mathbb{R}^k; \tag{8}$$

$\|\mathbf{z}\|_{\boldsymbol{\Sigma}} := (\mathbf{z}' \boldsymbol{\Sigma}^{-1} \mathbf{z})^{1/2}$, as usual, denotes the norm of $\mathbf{z}$ in the metric associated with $\boldsymbol{\Sigma}$, the constant $c_{k,f}$ is the normalization factor $(\omega_k\, \mu_{k-1;f})^{-1}$, where ω_k stands for the $(k-1)$-dimensional Lebesgue measure of the unit sphere $\mathcal{S}^{k-1} \subset \mathbb{R}^k$, and $\mu_{l;f} := \int_0^\infty r^l f(r)\, dr)$.

Note that, despite the notation, $\boldsymbol{\Sigma}$ needs not be a covariance matrix.

Local asymptotic normality requires some further regularity assumptions on the innovation density. The set of assumptions (B1′)-(B3) collects these assumptions.

Assumption (B1′). Same as Assumption (B1), but with $\mu_{k+1,f} < \infty$.
Assumption (B2). The square root $f^{1/2}$ of the radial density f is in the subspace $W^{1,2}(\mathbb{R}^+_0, \mu_{k-1})$ of $L^2(\mathbb{R}^+_0, \mu_{k-1})$ containing all functions admitting a weak derivative that also belongs to $L^2(\mathbb{R}^+_0, \mu_{k-1})$ (where $L^2(\mathbb{R}^+_0, \mu_l)$ stands for the space of all measurable functions $h : \mathbb{R}^+_0 \to \mathbb{R}$ satisfying $\int_0^\infty [h(r)]^2 r^l\, dr < \infty$).

Assumption (B2) is strictly equivalent to the assumption that $\underline{f}^{1/2}$ is differentiable in quadratic mean (see Hallin and Paindaveine 2002a). Denoting by $(f^{1/2})'$ the weak derivative of $f^{1/2}$ in $L^2(\mathbb{R}^+_0, \mu_{k-1})$, let $\varphi_f := -2\frac{(f^{1/2})'}{f^{1/2}}$. Under (B2), the *radial Fisher information* $\mathcal{I}_{k,f} := (\mu_{k-1;f})^{-1} \int_0^\infty [\varphi_f(r)]^2 r^{k-1} f(r)\, dr$ is finite. In the pure location or purely serial problems considered in Hallin and Paindaveine (2002a, b, and 2004a), this was sufficient for LAN. However, as pointed out by Garel and Hallin (1995), LAN, when serial and nonserial features both are present in the model, requires the stronger assumption

Assumption (B3). $\int_0^\infty [\varphi_f(r)]^4 r^{k-1} f(r)\, dr < \infty$.

The joint distribution of the observation $\mathbf{Y}^{(n)}$ under parameter value $\boldsymbol{\theta}$ and innovation density (8) will be denoted as $\mathrm{P}^{(n)}_{\boldsymbol{\theta}, \boldsymbol{\Sigma}, f}$.

2.4 Score functions.

Assumptions (C) and (C′) impose some mild conditions on the score functions J_ℓ, $\ell = 0, 1, 2$, to be used when building rank-based statistics.

Assumption (C). The score functions $J_\ell :]0, 1[\to \mathbb{R}$, $\ell = 0, 1, 2$, are continuous differences of two monotone increasing functions, and satisfy $\int_0^1 [J_\ell(u)]^2\, du < \infty$ $(\ell = 0, 1, 2)$.

The score functions yielding locally and asymptotically optimal procedures are of the form $J_0 = J_1 := \varphi_{f_\star} \circ \tilde{F}_{\star k}^{-1}$ and $J_2 := \tilde{F}_{\star k}^{-1}$, for some radial density $f_\star$ (here $\tilde{F}_{\star k}$ stands for the cdf associated with the radial pdf $\tilde{f}_{\star k}(r) = (\mu_{k-1;f_\star})^{-1} r^{k-1} f_\star(r) I_{[r>0]}$, $r \in \mathbb{R}$). Assumption (C) then takes the form of an assumption on $f_\star$:

Assumption (C′). The radial density $f_\star$ is such that $\varphi_{f_\star}$ is the continuous difference of two monotone increasing functions, $\mu_{k+1;f_\star} < \infty$, and $\int_0^\infty [\varphi_{f_\star}(r)]^2 r^{k-1} f_\star(r)\, dr < \infty$.

2.5 Estimation of nuisance parameters.

The shape matrix $\mathbf{\Sigma}$ in Assumption (B1) is unknown and has to be estimated by some $\widehat{\mathbf{\Sigma}}^{(n)}$. We assume the following.

Assumption (D1). Let $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$ be i.i.d., with density $\underline{f}$ satisfying Assumption (B1). The sequence $\widehat{\mathbf{\Sigma}}^{(n)} = \widehat{\mathbf{\Sigma}}^{(n)}(\mathbf{Z}_1, \ldots, \mathbf{Z}_n)$ of estimators of $\mathbf{\Sigma}$ is such that

(i) $\sqrt{n}(\widehat{\mathbf{\Sigma}}^{(n)} - a\,\mathbf{\Sigma}) = O_{\mathrm{P}}(1)$ as $n \to \infty$ for some positive real a, and

(ii) $\widehat{\mathbf{\Sigma}}^{(n)}$ is invariant under permutations and reflections (with respect to the origin in $\mathbb{R}^k$) of the $\mathbf{Z}_t$'s.

Assumption (D1) will be sufficient for the validity of the proposed procedures. However, their affine-invariance requires the following equivariance assumption on $\widehat{\mathbf{\Sigma}} := \widehat{\mathbf{\Sigma}}^{(n)}$.

Assumption (D2). The estimator $\widehat{\mathbf{\Sigma}}$ is *quasi-affine-equivariant*, in the sense that, for all n, all $k \times k$ full-rank matrix $\mathbf{M}$, $\widehat{\mathbf{\Sigma}}(\mathbf{M}) = d\,\mathbf{M}\widehat{\mathbf{\Sigma}}\mathbf{M}'$, where $\widehat{\mathbf{\Sigma}}(\mathbf{M})$ stands for the statistic $\widehat{\mathbf{\Sigma}}$ computed from the n-tuple $(\mathbf{MZ}_1, \ldots, \mathbf{MZ}_n)$, and d denotes some positive scalar that may depend on $\mathbf{M}$ and $(\mathbf{Z}_1, \ldots, \mathbf{Z}_n)$.

Since the parameter of interest $\boldsymbol{\theta}$ remains partially unspecified under the null, we also need replacing it with some estimate. More precisely, let $\boldsymbol{\theta}_0$ and $\mathbf{\Upsilon}$ satisfy Assumption (A2). For all $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\mathbf{\Upsilon}))^*$, we will assume the existence of an estimator $\hat{\boldsymbol{\theta}} := \hat{\boldsymbol{\theta}}^{(n)}$ satisfying Assumptions (E1) and (E2) below.

Assumption (E1). The sequence of estimators $(\hat{\boldsymbol{\theta}}^{(n)}, n \in \mathbb{N})$ is

(i) *constrained:* $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\mathbf{\Sigma},f}\left[\hat{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta}_0 \in \mathcal{M}(\mathbf{\Upsilon})\right] = 1$ for all n, $\mathbf{\Sigma}$, f, and $\boldsymbol{\theta}$,

(ii) *root-n consistent:* for all $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\mathbf{\Upsilon}))^*$, $(\hat{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta}) = O_{\mathrm{P}}(n^{-1/2})$, as $n \to \infty$, under $\bigcup_{\mathbf{\Sigma}} \bigcup_f \mathrm{P}^{(n)}_{\boldsymbol{\theta},\mathbf{\Sigma},f}$, and

(iii) *locally asymptotically discrete:* for all $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\mathbf{\Upsilon}))^*$, and all $c > 0$, there exists an $M(c) > 0$ such that the number of possible values of $\hat{\boldsymbol{\theta}}^{(n)}$ in balls of the form $\{\mathbf{t} \in \mathbb{R}^K : n^{1/2}\|(\mathbf{t} - \boldsymbol{\theta})\| \leq c\}$ is bounded by M, uniformly as $n \to \infty$.

The root-n consistency requirement in part (ii) of Assumption (E1) is satisfied by all classical estimators (Yule-Walker, least squares, maximum likelihood, ...). Note however that root-n consistency results for M-estimators (maximum likelihood or least squares) in general are proved over compact sets of parameter values, while $(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$ is not a closed set. Notable exceptions, however, exist; see, for instance, the early contributions by Dunsmuir and Hannan (1976) and Deistler, Dunsmuir, and Hannan (1978). As for the local discreteness assumption (E1) (iii), which goes back to Le Cam (1960) or (1986), it is a purely technical requirement, with little practical implications as, for fixed sample size, any estimate can be considered part of a locally asymptotically discrete sequence. The combination of parts (i) and (ii) of Assumption (E1) does not create any additional difficulty: any sequence $\tilde{\boldsymbol{\theta}}^{(n)}$ of unconstrained root-n consistent estimators indeed easily can be turned into a constrained one by means of a simple projection, namely,

$$\hat{\boldsymbol{\theta}}^{(n)} := \boldsymbol{\theta}_0 + \boldsymbol{\Pi}_{\boldsymbol{\Upsilon}}(\tilde{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta}_0), \tag{9}$$

where $\boldsymbol{\Pi}_{\boldsymbol{\Upsilon}} := \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'$ is the projection matrix onto $\mathcal{M}(\boldsymbol{\Upsilon}) \subseteq \mathbb{R}^K$. The resulting $\hat{\boldsymbol{\theta}}^{(n)}$ clearly satisfies parts (i) and (ii) of (E1).

Note that part (i) of Assumption (E1) does not require $\hat{\boldsymbol{\theta}}^{(n)}$ to belong ($\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},f}$-a.s.) to $(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$. However, whenever $\boldsymbol{\theta}$ belongs to $(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$, which is an open subset with respect to $(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))$, (E1) (ii) and (i) jointly imply that the $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},f}$−probability that $\hat{\boldsymbol{\theta}}^{(n)} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$ tends to one as $n \to \infty$.

While Assumption (E1) is classical for both the univariate and the multivariate versions of the testing problem under study, Assumption (E2) below is specific to the multivariate case (it is essentially void for $k = 1$), and is required if affine-invariance is to be achieved; see Section 1.5.2 for notation.

Assumption (E2). For any full-rank $k \times k$ matrix $\mathbf{M}$, denote by $\hat{\boldsymbol{\theta}}(\mathbf{M})$ the value of $\hat{\boldsymbol{\theta}}^{(n)}$ computed from the transformed sample $\mathbf{MY}_1, \ldots, \mathbf{MY}_n$: $\hat{\boldsymbol{\theta}}^{(n)}$ is affine-equivariant, meaning that $\hat{\boldsymbol{\theta}}(\mathbf{M}) = g_{\mathbf{M}}^{(m,p_1+q_1)}\hat{\boldsymbol{\theta}}$, for all $g_{\mathbf{M}}^{(m,p_1+q_1)} \in \mathcal{G}^m_{p_1+q_1}(k)$.

Equivalently, (E2) means that the estimators we are considering are assumed to satisfy $\hat{\boldsymbol{\beta}}(\mathbf{M}) = \hat{\boldsymbol{\beta}}\mathbf{M}'$, $\hat{\mathbf{A}}_i(\mathbf{M}) = \mathbf{M}\hat{\mathbf{A}}_i\mathbf{M}^{-1}$ for all $i = 1, \ldots, p_0$, and $\hat{\mathbf{B}}_j(\mathbf{M}) = \mathbf{M}\hat{\mathbf{B}}_j\mathbf{M}^{-1}$ for all $j = 1, \ldots, q_0$. Note that the corresponding Green's matrices then also are affine-equivariant, i.e., $\mathbf{G}_u(\hat{\boldsymbol{\theta}}(\mathbf{M})) = \mathbf{M}\,\mathbf{G}_u(\hat{\boldsymbol{\theta}})\mathbf{M}^{-1}$ and $\mathbf{H}_u(\hat{\boldsymbol{\theta}}(\mathbf{M})) = \mathbf{M}\,\mathbf{H}_u(\hat{\boldsymbol{\theta}})\mathbf{M}^{-1}$ for all integer u. In the sequel, we write $\hat{\mathbf{G}}^{(n)}_u$ and $\hat{\mathbf{H}}^{(n)}_u$ for $\mathbf{G}_u(\hat{\boldsymbol{\theta}})$ and $\mathbf{H}_u(\hat{\boldsymbol{\theta}})$, respectively.

Clearly, if $\hat{\boldsymbol{\theta}}^{(n)}$ is a constrained estimator satisfying (E2), $\hat{\boldsymbol{\theta}}(\mathbf{M})$ is also constrained, since $\hat{\boldsymbol{\theta}}(\mathbf{M}) = g_{\mathbf{M}}^{(m,p_1+q_1)}\hat{\boldsymbol{\theta}} \in g_{\mathbf{M}}^{(m,p_1+q_1)}(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})) = \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$ for all $\mathbf{M}$ (we restricted to pairs $(\boldsymbol{\theta}_0, \mathcal{M}(\boldsymbol{\Upsilon}))$ for which the null hypothesis is affine-invariant). In other words, affine-equivariance in (E2) and part (i) of (E1) are compatible, provided that the linear restriction in (E1) (i) itself is affine-invariant.

Finally, one can easily check that the affine-invariance of the linear restrictions under consideration implies that projecting an *affine-equivariant* sequence $\tilde{\boldsymbol{\theta}}^{(n)}$ of unconstrained estimators yields an affine-equivariant sequence $\hat{\boldsymbol{\theta}}^{(n)}$ of projected estimators (9). This provides a convenient way to construct sequences of estimators satisfying Assumptions (E1) and (E2) from traditional affine-equivariant ones (such as the Yule-Walker estimators). Similarly, if constrained M-estimation methods (such as constrained maximum likelihood: see, e.g., Reinsel 1997) are adopted, the affine-invariance of the linear restriction entails that of the resulting constrained estimators.

2.6 Linear hypotheses.

It will be convenient to write $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$ for the simple hypothesis $\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},f}\}$. The null hypotheses we are interested in are of the form $\bigcup_{\boldsymbol{\Sigma}} \bigcup_f \mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\Upsilon}}(\boldsymbol{\Sigma}, f) := \bigcup_{\boldsymbol{\Sigma}} \bigcup_f \left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},f} \,\middle|\, \boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*\right\}$ where $f, \boldsymbol{\Sigma}, \boldsymbol{\theta}_0$, and $\boldsymbol{\Upsilon}$ are such that Assumptions (A2) and (B) hold. The notation $\mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\Upsilon}}$ will be used for $\bigcup_{\boldsymbol{\Sigma}} \bigcup_f \mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\Upsilon}}(\boldsymbol{\Sigma}, f)$.

The goal of this paper is to develop testing procedures for $\mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\Upsilon}}$ that

(a) are *non-parametric*, i.e., valid (or, at least, asymptotically valid) under any distribution in $\mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\Upsilon}}$ —in particular, under any elliptically symmetric innovation density (possibly satisfying some moment constraints);

(b) are *locally and asymptotically optimal (LAO)* (locally asymptotically *most stringent*, in this case) at some fixed radial density $f_\star$, that is, against sequences of alternatives of the form $\bigcup_{\boldsymbol{\theta}\notin\boldsymbol{\theta}_0+\mathcal{M}(\boldsymbol{\Upsilon})} \bigcup_{\boldsymbol{\Sigma}} \mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f_\star)$; such a property of course requires the local asymptotic normality (LAN) of the parametric submodel associated with $f_\star$;

(c) *comply with the invariance principle*: we restricted to null hypotheses that are invariant with respect to the group of affine transformations. The hypotheses considered are also invariant with respect to the group of continuous monotone radial transformations (acting on residuals; see Section 4.1 for a precise definition). The proposed procedures should be (at least asymptotically) invariant with respect to these two groups.

3 Uniform local asymptotic normality (ULAN).

In this section, we briefly recall the ULAN (uniform local asymptotic normality) result proved in Hallin and Paindaveine (2004b) for the model under study. Local asymptotic normality for AR or ARMA processes was first established by Kreiss (1987) and Swensen (1985). In the multivariate case, linear models with VARMA errors have been considered by Garel and Hallin (1995). Elliptic symmetry however allows for a more convenient form, which we now describe.

Let $\boldsymbol{\theta}$ belong to some $\boldsymbol{\Theta}_{p_0,q_0}$ $(0 \leq p_0 \leq p_1;\ 0 \leq q_0 \leq q_1)$. Denote by $\boldsymbol{\beta}$, $\mathbf{A}(L)$, and $\mathbf{B}(L)$ the corresponding regression coefficients and VARMA polynomials. The sequences of local alternatives to be considered for LAN at $\boldsymbol{\theta}$ are associated with sequences of models of the form

$$\mathbf{Y}^{(n)} = \mathbf{X}^{(n)}\boldsymbol{\beta}^{(n)} + \mathbf{U}^{(n)}, \qquad \mathbf{A}^{(n)}(L)\,\mathbf{U}^{(n)}_t = \mathbf{B}^{(n)}(L)\,\boldsymbol{\varepsilon}^{(n)}_t, \quad t \in \mathbb{Z}, \tag{10}$$

where $\boldsymbol{\beta}^{(n)} := \boldsymbol{\beta} + n^{-1/2}\mathbf{K}^{(n)}\boldsymbol{\eta}^{(n)}$, $\mathbf{A}^{(n)}(L) := \mathbf{I}_k - \sum_{i=1}^{p_1}(\mathbf{A}_i + n^{-1/2}\boldsymbol{\gamma}^{(n)}_i)\,L^i$, $\mathbf{B}^{(n)}(L) := \mathbf{I}_k + \sum_{i=1}^{q_1}(\mathbf{B}_i + n^{-1/2}\boldsymbol{\delta}^{(n)}_i)\,L^i$, and the sequence

$$\boldsymbol{\tau}^{(n)} := \Big((\mathrm{vec}\,\boldsymbol{\eta}^{(n)\prime})', (\mathrm{vec}\,\boldsymbol{\gamma}^{(n)}_1)', \ldots, (\mathrm{vec}\,\boldsymbol{\gamma}^{(n)}_{p_1})', (\mathrm{vec}\,\boldsymbol{\delta}^{(n)}_1)', \ldots, (\mathrm{vec}\,\boldsymbol{\delta}^{(n)}_{q_1})'\Big)' \in \mathbb{R}^K$$

is bounded as $n \to \infty$. The perturbed parameter is thus

$$\boldsymbol{\theta}^{(n)} := \boldsymbol{\theta} + \boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)} := \boldsymbol{\theta} + n^{-1/2}\begin{pmatrix} \mathbf{K}^{(n)} \otimes \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{k^2(p_1+q_1)} \end{pmatrix}\boldsymbol{\tau}^{(n)}.$$

The corresponding sequence of local alternatives is thus $\mathcal{H}^{(n)}(\boldsymbol{\theta} + \boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}, \boldsymbol{\Sigma}, f)$.

Decompose $\mathbf{Z}_t(\boldsymbol{\theta}) := \mathbf{Z}_t^{(n)}(\boldsymbol{\theta})$ into $\mathbf{Z}_t(\boldsymbol{\theta}) = d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})\, \boldsymbol{\Sigma}^{1/2}\mathbf{U}_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})$, where $d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}) = d_t^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) := \|\mathbf{Z}_t(\boldsymbol{\theta})\|_{\boldsymbol{\Sigma}}$ and $\mathbf{U}_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}) = \mathbf{U}_t^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) := \boldsymbol{\Sigma}^{-1/2}\mathbf{Z}_t(\boldsymbol{\theta})/d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})$. Under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$, the $d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})$'s are i.i.d. with density $\tilde{f}_k(r) = (\mu_{k-1;f})^{-1}\, r^{k-1} f(r)\, I_{[r>0]}$ and distribution function $\tilde{F}_k$. As we will see, the central sequences involved in the ULAN result are linear combinations of (the entries of) the generalized cross-covariance matrices

$$\boldsymbol{\Gamma}^{(n)}_{i;\boldsymbol{\Sigma},f}(\boldsymbol{\theta}) := (n-i)^{-1}\, \boldsymbol{\Sigma}^{\prime -1/2} \left(\sum_{t=i+1}^{n} \varphi_f(d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}))\, d_{t-i}(\boldsymbol{\theta}, \boldsymbol{\Sigma})\ \mathbf{U}_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})\mathbf{U}^{\prime}_{t-i}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) \right) \boldsymbol{\Sigma}^{\prime 1/2}, \tag{11}$$

and the matrices of nonserial statistics

$$\boldsymbol{\Lambda}^{(n)}_{i;\boldsymbol{\Sigma},f}(\boldsymbol{\theta}) := (n-i)^{-1}\, \boldsymbol{\Sigma}^{\prime -1/2} \sum_{t=i+1}^{n} \varphi_f(d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}))\, \mathbf{U}_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})\, \mathbf{x}^{(n)\prime}_{t-i}\mathbf{K}^{(n)}, \tag{12}$$

which therefore contain all the relevant information (in the local and asymptotic sense) about $\boldsymbol{\theta}$. The coefficients of these linear combinations are rather complicated, though, and require some further notation, mainly connected with the algebra of linear difference equations.

Associated with any k-dimensional linear difference operator of the form $\mathbf{C}(L) := \sum_{i=0}^{\infty} \mathbf{C}_i\, L^i$ (letting $\mathbf{C}_i = \mathbf{0}$ for $i > s$, this includes, of course, the operators with finite order s), define, for any integers u and v, the $k^2u \times k^2v$ matrices

$$\mathbf{C}^{(l)}_{u,v} := \begin{pmatrix} \mathbf{C}_0 \otimes \mathbf{I}_k & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{C}_1 \otimes \mathbf{I}_k & \mathbf{C}_0 \otimes \mathbf{I}_k & \dots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{C}_{v-1} \otimes \mathbf{I}_k & \mathbf{C}_{v-2} \otimes \mathbf{I}_k & \dots & \mathbf{C}_0 \otimes \mathbf{I}_k \\ \vdots & & & \vdots \\ \mathbf{C}_{u-1} \otimes \mathbf{I}_k & \mathbf{C}_{u-2} \otimes \mathbf{I}_k & \dots & \mathbf{C}_{u-v} \otimes \mathbf{I}_k \end{pmatrix} \tag{13}$$

and

$$\mathbf{C}^{(r)}_{u,v} := \begin{pmatrix} \mathbf{I}_k \otimes \mathbf{C}_0 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{I}_k \otimes \mathbf{C}_1 & \mathbf{I}_k \otimes \mathbf{C}_0 & \dots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{I}_k \otimes \mathbf{C}_{v-1} & \mathbf{I}_k \otimes \mathbf{C}_{v-2} & \dots & \mathbf{I}_k \otimes \mathbf{C}_0 \\ \vdots & & & \vdots \\ \mathbf{I}_k \otimes \mathbf{C}_{u-1} & \mathbf{I}_k \otimes \mathbf{C}_{u-2} & \dots & \mathbf{I}_k \otimes \mathbf{C}_{u-v} \end{pmatrix}, \tag{14}$$

respectively; write $\mathbf{C}^{(l)}_u$ for $\mathbf{C}^{(l)}_{u,u}$ and $\mathbf{C}^{(r)}_u$ for $\mathbf{C}^{(r)}_{u,u}$. With this notation, note that $\mathbf{G}^{(l)}_u$, $\mathbf{G}^{(r)}_u$, $\mathbf{H}^{(l)}_u$, and $\mathbf{H}^{(r)}_u$ are the inverses of $\mathbf{A}^{(l)}_u$, $\mathbf{A}^{(r)}_u$, $\mathbf{B}^{(l)}_u$, and $\mathbf{B}^{(r)}_u$, respectively. Denoting by $\mathbf{C}^{\prime(l)}_{u,v}$ and $\mathbf{C}^{\prime(r)}_{u,v}$ the matrices associated with the transposed operator $\mathbf{C}^{\prime}(L) := \sum_{i=0}^{\infty} \mathbf{C}^{\prime}_i\, L^i$, we also have $\mathbf{G}^{\prime(l)}_u = (\mathbf{A}^{\prime(l)}_u)^{-1}$, $\mathbf{H}^{\prime(l)}_u = (\mathbf{B}^{\prime(l)}_u)^{-1}$, etc. We will use the notation $\bar{\mathbf{C}}^{(l)}_{u,v}, \bar{\mathbf{C}}^{(r)}_{u,v}, \bar{\mathbf{C}}^{(l)}_u$, etc. when the identity matrices involved in (13) and (14) are m-dimensional rather than k-dimensional.

Let $\pi := \max(p_1 - p_0, q_1 - q_0)$ and $\pi_0 := \pi + p_0 + q_0$, and define the $k^2\pi_0 \times k^2(p_1 + q_1)$ matrix

$$\mathbf{M}_{\boldsymbol{\theta}} := \left(\mathbf{G}^{\prime(l)}_{\pi_0,p_1} \vdots \mathbf{H}^{\prime(l)}_{\pi_0,q_1} \right); \tag{15}$$

under Assumption (A2), $\mathbf{M}_{\boldsymbol{\theta}}$, for $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$, is of full rank.

Consider the operator $\mathbf{D}(L) := \mathbf{I}_k + \sum_{i=1}^{p_0+q_0} \mathbf{D}_i L^i$ (just as $\mathbf{M}_{\boldsymbol{\theta}}$, $\mathbf{D}(L)$ and most quantities defined below depend on $\boldsymbol{\theta}$, but, for simplicity, we are dropping this reference to $\boldsymbol{\theta}$), where, putting $\mathbf{G}_{-1} = \mathbf{G}_{-2} = ... = \mathbf{G}_{-p_0+1} = \mathbf{0} = \mathbf{H}_{-1} = \mathbf{H}_{-2} = ... = \mathbf{H}_{-q_0+1}$,

$$\begin{pmatrix} \mathbf{D}_1' \\ \vdots \\ \mathbf{D}_{p_0+q_0}' \end{pmatrix} := - \begin{pmatrix} \mathbf{G}_{q_0} & \mathbf{G}_{q_0-1} & \dots & \mathbf{G}_{-p_0+1} \\ \mathbf{G}_{q_0+1} & \mathbf{G}_{q_0} & \dots & \mathbf{G}_{-p_0+2} \\ \vdots & & \ddots & \vdots \\ \mathbf{G}_{p_0+q_0-1} & \mathbf{G}_{p_0+q_0-2} & \dots & \mathbf{G}_0 \\ \mathbf{H}_{p_0} & \mathbf{H}_{p_0-1} & \dots & \mathbf{H}_{-q_0+1} \\ \mathbf{H}_{p_0+1} & \mathbf{H}_{p_0} & \dots & \mathbf{H}_{-q_0+2} \\ \vdots & & \ddots & \vdots \\ \mathbf{H}_{p_0+q_0-1} & \mathbf{H}_{p_0+q_0-2} & \dots & \mathbf{H}_0 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{G}_{q_0+1} \\ \vdots \\ \mathbf{G}_{p_0+q_0} \\ \mathbf{H}_{p_0+1} \\ \vdots \\ \mathbf{H}_{p_0+q_0} \end{pmatrix}.$$

Note that $\mathbf{D}(L)\mathbf{G}_t' = \mathbf{0}$ for $t = q_0+1, \dots, p_0+q_0$, and $\mathbf{D}(L)\mathbf{H}_t' = \mathbf{0}$ for $t = p_0+1, \dots, p_0+q_0$.

Let $\{\boldsymbol{\Psi}_t^{(1)}, \dots, \boldsymbol{\Psi}_t^{(p_0+q_0)}\}$ be a set of $k \times k$ matrices forming a fundamental system of solutions of the homogeneous linear difference equation associated with $\mathbf{D}(L)$ (such a system can be obtained, for instance, from the Green's matrices of the operator $\mathbf{D}(L)$: see Hallin 1986). Define

$$\bar{\boldsymbol{\Psi}}_m(\boldsymbol{\theta}) := \begin{pmatrix} \boldsymbol{\Psi}_{\pi+1}^{(1)} & \dots & \boldsymbol{\Psi}_{\pi+1}^{(p_0+q_0)} \\ \boldsymbol{\Psi}_{\pi+2}^{(1)} & \dots & \boldsymbol{\Psi}_{\pi+2}^{(p_0+q_0)} \\ \vdots & & \vdots \\ \boldsymbol{\Psi}_{m}^{(1)} & \dots & \boldsymbol{\Psi}_{m}^{(p_0+q_0)} \end{pmatrix} \otimes \mathbf{I}_k \qquad (m > \pi),$$

$$\mathbf{P}_{\boldsymbol{\theta}} := \begin{pmatrix} \mathbf{I}_{k^2\pi} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_\Psi^{-1} \end{pmatrix}, \quad \text{and} \quad \mathbf{Q}_{\boldsymbol{\theta}}^{(n)} := \mathbf{H}_{n-1}^{(r)} \mathbf{B}_{n-1}'^{(l)} \begin{pmatrix} \mathbf{I}_{k^2\pi} & \mathbf{0} \\ \mathbf{0} & \bar{\boldsymbol{\Psi}}_{n-1} \end{pmatrix}, \tag{16}$$

where $\mathbf{C}_\Psi$ is the Casorati matrix $\bar{\boldsymbol{\Psi}}_{\pi_0}$.

Finally, put (with $\boldsymbol{\Lambda}_{i;\boldsymbol{\Sigma},f}^{(n)}$ defined in (12))

$$\mathbf{S}_{I;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}) := \left(n^{1/2} (\text{vec}\,\boldsymbol{\Lambda}_{0;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}))', \dots, (n-i)^{1/2} (\text{vec}\,\boldsymbol{\Lambda}_{i;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}))', \dots, (\text{vec}\,\boldsymbol{\Lambda}_{n-1;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}))'\right)',$$

$$n^{1/2}\mathbf{T}_{I;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}) := \mathbf{L}_{\boldsymbol{\theta}}^{(n)\prime} \mathbf{S}_{I;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}), \quad \text{and} \quad \mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}} := \lim_{n\to+\infty} \mathbf{L}_{\boldsymbol{\theta}}^{(n)\prime} (\boldsymbol{K}_n \otimes \boldsymbol{\Sigma}^{-1})\, \mathbf{L}_{\boldsymbol{\theta}}^{(n)}, \tag{17}$$

where $\mathbf{L}_{\boldsymbol{\theta}}^{(n)} := \bar{\mathbf{H}}_n^{(r)}(\boldsymbol{\theta}) \bar{\mathbf{A}}_{n,1}^{(r)}(\boldsymbol{\theta})$, and where $\boldsymbol{K}_{l,\tilde{l}}$ denotes the $lm \times \tilde{l}m$ matrix with block $\mathbf{K}'\,\mathbf{R}_{|i-j|}\,\mathbf{K}$ in position (i,j) $(i = 1, \dots, l,\ j = 1, \dots, \tilde{l})$. We write $\boldsymbol{K}_l$ instead of $\boldsymbol{K}_{l,l}$. Similarly, for the serial part, let (with $\boldsymbol{\Gamma}_{i;\boldsymbol{\Sigma},f}^{(n)}$ defined in (11))

$$\mathbf{S}_{II;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}) := \left((n-1)^{1/2} (\text{vec}\,\boldsymbol{\Gamma}_{1;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}))', \dots, (n-i)^{1/2} (\text{vec}\,\boldsymbol{\Gamma}_{i;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}))', \dots, (\text{vec}\,\boldsymbol{\Gamma}_{n-1;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}))'\right)',$$

$$n^{1/2}\mathbf{T}_{II;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}) := \mathbf{Q}_{\boldsymbol{\theta}}^{(n)\prime} \mathbf{S}_{II;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}), \quad \text{and} \quad \mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}} := \lim_{n\to+\infty} \mathbf{Q}_{\boldsymbol{\theta}}^{(n)\prime} \left[\mathbf{I}_{n-1} \otimes (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}^{-1})\right] \mathbf{Q}_{\boldsymbol{\theta}}^{(n)} \tag{18}$$

(convergence in (17) and (18) follows from the exponential decrease, as $u \to \infty$, under (A2), of the Green's matrices $\mathbf{G}_u$ and $\mathbf{H}_u$).

We now can state the ULAN proved in Hallin and Paindaveine (2004b).

Proposition 1 (ULAN) *Assume that* $\boldsymbol{\theta}$ *belongs to some* $\boldsymbol{\Theta}_{p_0,q_0}$ ($0 \leq p_0 \leq p_1$; $0 \leq q_0 \leq q_1$). *Let Assumptions (A1), (B1′), (B2), and (B3) hold, and consider a sequence* $\boldsymbol{\theta}_n$ *such that* $\boldsymbol{\theta}_n - \boldsymbol{\theta} = O(n^{-1/2})$ *as* $n \to \infty$. *Then, for any bounded sequence* $\boldsymbol{\tau}^{(n)}$, *the logarithm* $L^{(n)}_{\boldsymbol{\theta}_n+\boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}/\boldsymbol{\theta}_n;\boldsymbol{\Sigma},f}$ *of the likelihood ratio associated with the sequence of local alternatives* $\mathcal{H}^{(n)}(\boldsymbol{\theta}_n + \boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}, \boldsymbol{\Sigma}, f)$ *with respect to* $\mathcal{H}^{(n)}(\boldsymbol{\theta}_n, \boldsymbol{\Sigma}, f)$ *is such that*

$$L^{(n)}_{\boldsymbol{\theta}_n+\boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}/\boldsymbol{\theta}_n;\boldsymbol{\Sigma},f}(\mathbf{Y}^{(n)}) = (\boldsymbol{\tau}^{(n)})'\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\Sigma},f}(\boldsymbol{\theta}_n) - \frac{1}{2}(\boldsymbol{\tau}^{(n)})'\boldsymbol{\Gamma}_{\boldsymbol{\Sigma},f}(\boldsymbol{\theta})\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1),$$

as $n \to \infty$, *under* $\mathcal{H}^{(n)}(\boldsymbol{\theta}_n, \boldsymbol{\Sigma}, f)$, *with the central sequence*

$$\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\Sigma},f}(\boldsymbol{\theta}_n) := \begin{pmatrix} \boldsymbol{\Delta}^{(n)}_{I;\boldsymbol{\Sigma},f}(\boldsymbol{\theta}_n) \\ \boldsymbol{\Delta}^{(n)}_{II;\boldsymbol{\Sigma},f}(\boldsymbol{\theta}_n) \end{pmatrix} := n^{1/2} \begin{pmatrix} \mathbf{I}_{km} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}'_{\boldsymbol{\theta}_n}\mathbf{P}'_{\boldsymbol{\theta}_n} \end{pmatrix} \begin{pmatrix} \mathbf{T}^{(n)}_{I;\boldsymbol{\Sigma},f}(\boldsymbol{\theta}_n) \\ \mathbf{T}^{(n)}_{II;\boldsymbol{\Sigma},f}(\boldsymbol{\theta}_n) \end{pmatrix}, \tag{19}$$

and the information matrix

$$\boldsymbol{\Gamma}_{\boldsymbol{\Sigma},f}(\boldsymbol{\theta}) := \begin{pmatrix} \boldsymbol{\Gamma}_{I;\boldsymbol{\Sigma},f}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma}_{II;\boldsymbol{\Sigma},f}(\boldsymbol{\theta}) \end{pmatrix},$$

where $\boldsymbol{\Gamma}_{I;\boldsymbol{\Sigma},f}(\boldsymbol{\theta}) := \frac{1}{k}\mathcal{I}_{k,f}\mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}$ *and* $\boldsymbol{\Gamma}_{II;\boldsymbol{\Sigma},f}(\boldsymbol{\theta}) := \frac{\mu_{k+1;f}\,\mathcal{I}_{k,f}}{k^2\,\mu_{k-1;f}}\mathbf{N}_{\boldsymbol{\theta},\boldsymbol{\Sigma}}$, *with* $\mathbf{N}_{\boldsymbol{\theta},\boldsymbol{\Sigma}} := \mathbf{M}'_{\boldsymbol{\theta}}\mathbf{P}'_{\boldsymbol{\theta}}\mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}}\mathbf{P}_{\boldsymbol{\theta}}\mathbf{M}_{\boldsymbol{\theta}}$. *Moreover,* $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\Sigma},f}(\boldsymbol{\theta}_n)$, *still under* $\mathcal{H}^{(n)}(\boldsymbol{\theta}_n, \boldsymbol{\Sigma}, f)$, *is asymptotically* $\mathcal{N}_K(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\Sigma},f}(\boldsymbol{\theta}))$.

Note that the asymptotic information matrix $\boldsymbol{\Gamma}_{\boldsymbol{\Sigma},f}(\boldsymbol{\theta})$ may be singular (such a singularity occurs as soon as $p_1 > p_0$ and $q_1 > q_0$). In such a case, a careful treatment, involving generalized inverses, will be required in the derivation of the asymptotic distributions of test statistics.

4 Multivariate signs and ranks, serial and nonserial signed rank statistics.

4.1 Multivariate signs and ranks.

The generalized cross-covariances (11) and nonserial statistics (12) are measurable with respect to the spherical distances $d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}) = \|\boldsymbol{\Sigma}^{-1/2}\mathbf{Z}(\boldsymbol{\theta})\|$ between the residuals $\mathbf{Z}_t(\boldsymbol{\theta})$ and the origin in $\mathbb{R}^k$, and the "multivariate signs" $\mathbf{U}_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}) = \boldsymbol{\Sigma}^{-1/2}\mathbf{Z}_t(\boldsymbol{\theta})/\|\boldsymbol{\Sigma}^{-1/2}\mathbf{Z}_t(\boldsymbol{\theta})\|$.

For each $\boldsymbol{\Sigma}$ and n, the group of *continuous monotone radial transformations* $\mathcal{G}^{(n)}_{\boldsymbol{\Sigma}} = \{\mathcal{G}^{(n)}_g\}$, acting on $(\mathbb{R}^k)^n$ and characterized by

$$\mathcal{G}^{(n)}_g(\mathbf{Z}_1(\boldsymbol{\theta}), \ldots, \mathbf{Z}_n(\boldsymbol{\theta})) := \Big(g(d_1(\boldsymbol{\theta}, \boldsymbol{\Sigma}))\,\boldsymbol{\Sigma}^{1/2}\mathbf{U}_1(\boldsymbol{\theta}, \boldsymbol{\Sigma}), \ldots, g(d_n(\boldsymbol{\theta}, \boldsymbol{\Sigma}))\,\boldsymbol{\Sigma}^{1/2}\mathbf{U}_n(\boldsymbol{\theta}, \boldsymbol{\Sigma})\Big), \tag{20}$$

where $g : \mathbb{R}^+ \to \mathbb{R}^+$ is a continuous monotone increasing function such that $g(0) = 0$ and $\lim_{r\to\infty} g(r) = \infty$, is a generating group for $\bigcup_f \mathcal{H}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$. Along with the signs $(\mathbf{U}_1(\boldsymbol{\theta}, \boldsymbol{\Sigma}), \ldots, \mathbf{U}_n(\boldsymbol{\theta}, \boldsymbol{\Sigma}))$, the ranks $(R_1(\boldsymbol{\theta}, \boldsymbol{\Sigma}), \ldots, R_n(\boldsymbol{\theta}, \boldsymbol{\Sigma}))$ of the distances $d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ constitute a maximal invariant for that group $\mathcal{G}^{(n)}_{\boldsymbol{\Sigma}}$ of radial transformations.

Because the true value of the shape matrix is unkown, the *genuine* ranks $R_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ and signs $\mathbf{U}_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ cannot be computed from the residuals $\mathbf{Z}_1(\boldsymbol{\theta}), \ldots, \mathbf{Z}_n(\boldsymbol{\theta})$, but the following alternative quantities can.

4.2 Pseudo-Mahalanobis signs and ranks.

The pseudo-Mahalanobis signs are defined as $\mathbf{W}_t(\boldsymbol{\theta}) = \mathbf{W}_t^{(n)}(\boldsymbol{\theta}) := \widehat{\boldsymbol{\Sigma}}^{-1/2}\mathbf{Z}_t(\boldsymbol{\theta})/\|\widehat{\boldsymbol{\Sigma}}^{-1/2}\mathbf{Z}_t(\boldsymbol{\theta})\|$, where $\widehat{\boldsymbol{\Sigma}}$ is the estimator in Assumptions (D1)-(D2). Similarly, the pseudo-Mahalanobis ranks $\widehat{R}_t(\boldsymbol{\theta}) := \widehat{R}_t^{(n)}(\boldsymbol{\theta})$ are defined as the ranks of the pseudo-Mahalanobis distances $d_t(\boldsymbol{\theta}, \hat{\boldsymbol{\Sigma}}) = \|\hat{\boldsymbol{\Sigma}}^{-1/2}\mathbf{Z}_t(\boldsymbol{\theta})\|$. The terminology *Mahalanobis signs* and *ranks* will be used when $\widehat{\boldsymbol{\Sigma}}$ is the empirical covariance matrix.

4.3 Hyperplane-based signs and ranks.

Pseudo-Mahalanobis signs and ranks are based on an estimation of the underlying shape matrix. A completely different approach can be based on counts of hyperplanes, and leads to a modification of Randles's interdirections (namely, the *absolute interdirections*) for multivariate signs, and to Oja and Paindaveine (2004)'s concept of *lift-interdirection ranks* for multivariate ranks.

Write $\mathcal{Q} := (i_1, i_2, \ldots, i_{k-1})$ $(1 \leq i_1 < i_2 < \ldots < i_{k-1} \leq n)$ and $\mathcal{P} := (j_1, j_2, \ldots, j_k)$ $(1 \leq j_1 < j_2 < \ldots < j_k \leq n)$ for arbitrary ordered sets of indices with sizes $(k-1)$ and k, respectively. Denote by $\mathbf{e}_{\mathcal{Q}}$ and $(d_{0\mathcal{P}}, \mathbf{d}'_{\mathcal{P}})'$ the vectors whose components are the cofactors of the last column in the arrays

$$(\mathbf{Z}_{i_1}(\boldsymbol{\theta}), \ldots, \mathbf{Z}_{i_{k-1}}(\boldsymbol{\theta}), \mathbf{z}) \quad \text{and} \quad \begin{pmatrix} 1 & 1 & \ldots & 1 & 1 \\ \mathbf{Z}_{j_1}(\boldsymbol{\theta}) & \mathbf{Z}_{j_2}(\boldsymbol{\theta}) & \ldots & \mathbf{Z}_{j_k}(\boldsymbol{\theta}) & \mathbf{z} \end{pmatrix},$$

respectively. The vector $\mathbf{e}_{\mathcal{Q}}$ (resp., $\mathbf{d}_{\mathcal{P}}$) is orthogonal to the hyperplane $\Pi(\mathcal{Q})$ spanned by $\mathbf{Z}_{i_1}(\boldsymbol{\theta}), \ldots, \mathbf{Z}_{i_{k-1}}(\boldsymbol{\theta})$ (resp., the hyperplane $\Pi(\mathcal{P})$ going through $\mathbf{Z}_{i_1}(\boldsymbol{\theta}), \ldots, \mathbf{Z}_{i_k}(\boldsymbol{\theta})$), and the sign of $\mathbf{e}'_{\mathcal{Q}}\mathbf{z}$ (resp. of $d_{0\mathcal{P}} + \mathbf{d}'_{\mathcal{P}}\mathbf{z}$) indicates on which side of $\Pi(\mathcal{Q})$ (resp. of $\Pi(\mathcal{P})$) the point $\mathbf{z}$ lies.

The *absolute interdirection* associated with residual $\mathbf{Z}_i(\boldsymbol{\theta})$ in the n-tuple $(\mathbf{Z}_1(\boldsymbol{\theta}), \ldots, \mathbf{Z}_n(\boldsymbol{\theta}))$ is defined as

$$\mathbf{V}_i(\boldsymbol{\theta}) = \mathbf{V}_i^{(n)}(\boldsymbol{\theta}) := (\cos(\pi p_{i;1}^{(n)}(\boldsymbol{\theta})), \ldots, \cos(\pi p_{i;k}^{(n)}(\boldsymbol{\theta})))',$$

with $p_{i;l}^{(n)}(\boldsymbol{\theta}) := \binom{n}{k-1}^{-1} c(\widehat{\boldsymbol{\Sigma}}^{1/2}\mathbf{e}_l^{(k)}, \mathbf{Z}_i(\boldsymbol{\theta}))$, where $\mathbf{e}_l^{(k)}$ denotes the lth unit vector of the canonical basis of $\mathbb{R}^k$, and $c(\mathbf{v}, \mathbf{w})$ is the hyperplane-based empirical angular distance

$$c(\mathbf{v}, \mathbf{w}) := \frac{1}{2}\sum_{\mathcal{Q}}\{1 - \text{sign}(\mathbf{e}'_{\mathcal{Q}}\mathbf{v})\ \text{sign}(\mathbf{e}'_{\mathcal{Q}}\mathbf{w})\}.$$

Note that the statistics $q_{ij}^{(n)}(\boldsymbol{\theta}) := c(\mathbf{Z}_i(\boldsymbol{\theta}), \mathbf{Z}_j(\boldsymbol{\theta}))$ are the so-called Randles' interdirections (Randles 1989); $q_{ij}^{(n)}$ is—up to a small-sample correction—the number of hyperplanes in $\mathbb{R}^k$ passing through the origin and $(k-1)$ out of the $(n-2)$ points $\mathbf{Z}_1(\boldsymbol{\theta}), \ldots, \mathbf{Z}_{i-1}(\boldsymbol{\theta}), \mathbf{Z}_{i+1}(\boldsymbol{\theta}), \ldots, \mathbf{Z}_{j-1}(\boldsymbol{\theta}), \mathbf{Z}_{j+1}(\boldsymbol{\theta}), \ldots, \mathbf{Z}_n(\boldsymbol{\theta})$ that separate $\mathbf{Z}_i(\boldsymbol{\theta})$ and $\mathbf{Z}_j(\boldsymbol{\theta})$.

In the same time, a hyperplane-based empirical distance between a vector $\mathbf{v}$ and the origin in $\mathbb{R}^k$ can be defined as

$$l^{(n)}(\mathbf{v}) := \frac{1}{2}\sum_{\mathcal{P}}\left(1 - \text{sign}(d_{0\mathcal{P}} + \mathbf{d}'_{\mathcal{P}}\mathbf{v})\,\text{sign}(d_{0\mathcal{P}} - \mathbf{d}'_{\mathcal{P}}\mathbf{v})\right),$$

i.e., as the number of hyperplanes in $\mathbb{R}^k$ passing through k out of the n points $\mathbf{Z}_1(\boldsymbol{\theta}), \ldots, \mathbf{Z}_n(\boldsymbol{\theta})$ that are separating $\mathbf{v}$ and its reflection $-\mathbf{v}$. For symmetry reasons, however, we rather consider the symmetrized distances

$$\underline{l}^{(n)}(\mathbf{v}) := \frac{1}{2} \sum_{\mathcal{P}} \sum_{\mathbf{s}} \left(1 - \text{sign}(d_{0\mathcal{P}}(\mathbf{s}) + \mathbf{d}'_{\mathcal{P}}(\mathbf{s})\mathbf{v}) \, \text{sign}(d_{0\mathcal{P}}(\mathbf{s}) - \mathbf{d}'_{\mathcal{P}}(\mathbf{s})\mathbf{v})\right),$$

where, for $\mathcal{P} = (j_1, \ldots, j_k)$ and $\mathbf{s} \in \{-1, 1\}^k$, $(d_{0\mathcal{P}}(\mathbf{s}), \mathbf{d}'_{\mathcal{P}}(\mathbf{s}))'$ stands for the vector of cofactors associated with the last column in the array

$$\begin{pmatrix} 1 & 1 & \ldots & 1 & 1 \\ s_1 \mathbf{Z}_{j_1}(\boldsymbol{\theta}) & s_2 \mathbf{Z}_{j_2}(\boldsymbol{\theta}) & \ldots & s_k \mathbf{Z}_{j_k}(\boldsymbol{\theta}) & \mathbf{z} \end{pmatrix}$$

(see Oja and Paindaveine 2004). The *lift-interdirection ranks* are the ranks $\underline{R}_i := \underline{R}_i^{(n)}$ of the symmetrized lift-interdirections $\underline{l}_i^{(n)} := \underline{l}^{(n)}(\mathbf{Z}_i(\boldsymbol{\theta}))$, $i = 1, \ldots, n$ among $\underline{l}_1^{(n)}, \ldots, \underline{l}_n^{(n)}$.

4.4 Serial and nonserial signed rank statistics.

The nonparametric (signed rank) J-score versions of the serial and nonserial statistics (11) and (12) are, in the serial case,

$$\underset{\sim}{\boldsymbol{\Gamma}}{}_{i;J}^{(n)}(\boldsymbol{\theta}) := \widehat{\boldsymbol{\Sigma}}'^{-1/2} \left(\frac{1}{n-i} \sum_{t=i+1}^{n} J_1\Big(\frac{\widehat{R}_t(\boldsymbol{\theta})}{n+1}\Big) J_2\Big(\frac{\widehat{R}_{t-i}(\boldsymbol{\theta})}{n+1}\Big) \, \mathbf{W}_t(\boldsymbol{\theta}) \, \mathbf{W}'_{t-i}(\boldsymbol{\theta}) \right) \widehat{\boldsymbol{\Sigma}}'^{1/2}, \qquad (21)$$

and, in the nonserial case,

$$\underset{\sim}{\boldsymbol{\Lambda}}{}_{i;J}^{(n)}(\boldsymbol{\theta}) := (n-i)^{-1} \, \widehat{\boldsymbol{\Sigma}}'^{-1/2} \sum_{t=i+1}^{n} J_0\Big(\frac{\widehat{R}_t(\boldsymbol{\theta})}{n+1}\Big) \, \mathbf{W}_t(\boldsymbol{\theta}) \, \mathbf{x}_{t-i}^{(n)\prime} \mathbf{K}^{(n)}, \qquad (22)$$

where the score functions J_ℓ ($\ell = 0, 1, 2$) are as in Assumption (C). Here we used pseudo-Mahalanobis signs and ranks. But every combination of a concept of multivariate signs (either Mahalanobis signs, pseudo-Mahalanobis signs, or absolute interdirections) with a concept of multivariate ranks (Mahalanobis, pseudo-Mahalanobis, or lift-interdirection ranks) may be considered and actually yields the same asymptotic representation results, as shown by the following proposition (see Hallin and Paindaveine (2004b) for a proof). Note however that their equivariance properties may be different (see the next subsection).

Proposition 2 *Assume that $\boldsymbol{\theta}$ belongs to some $\boldsymbol{\Theta}_{p_0,q_0}$ ($0 \leq p_0 \leq p_1$; $0 \leq q_0 \leq q_1$). Let Assumptions (A1), (B1), (C), and (D1) hold. Then, defining*

$$\widetilde{\boldsymbol{\Gamma}}_{i;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}) := \boldsymbol{\Sigma}'^{-1/2} \left(\frac{1}{n-i} \sum_{t=i+1}^{n} J_1(\tilde{F}_k(d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}))) J_2(\tilde{F}_k(d_{t-i}(\boldsymbol{\theta}, \boldsymbol{\Sigma}))) \; \mathbf{U}_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}) \mathbf{U}'_{t-i}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) \right) \boldsymbol{\Sigma}'^{1/2} \qquad (23)$$

and (see Assumption (D1) for the definition of a)

$$\widetilde{\boldsymbol{\Lambda}}_{i;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}) := (n-i)^{-1} \, a^{-1/2} \, \boldsymbol{\Sigma}'^{-1/2} \sum_{t=i+1}^{n} J_0(\tilde{F}_k(d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}))) \; \mathbf{U}_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}) \, \mathbf{x}_{t-i}^{(n)\prime} \mathbf{K}^{(n)}, \qquad (24)$$

(i) $\text{vec}\,(\underset{\sim}{\boldsymbol{\Lambda}}{}_{i;J}^{(n)}(\boldsymbol{\theta}) - \widetilde{\boldsymbol{\Lambda}}_{i;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}))$ *and* $\text{vec}\,(\underset{\sim}{\boldsymbol{\Gamma}}{}_{i;J}^{(n)}(\boldsymbol{\theta}) - \widetilde{\boldsymbol{\Gamma}}_{i;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}))$ *are $o_{\mathrm{P}}(n^{-1/2})$ for all i, as $n \to \infty$;*

(ii) the same result still holds if in $\underset{\sim}{\boldsymbol{\Lambda}}{}_{i;J}^{(n)}(\boldsymbol{\theta})$ and $\underset{\sim}{\boldsymbol{\Gamma}}{}_{i;J}^{(n)}(\boldsymbol{\theta})$ the pseudo-Mahalanobis signs $\mathbf{W}_t(\boldsymbol{\theta})$ are replaced by the corresponding absolute interdirections $\mathbf{V}_t(\boldsymbol{\theta})$, and/or the pseudo-Mahalanobis ranks $\widehat{R}_t(\boldsymbol{\theta})$ are replaced by the lift-interdirection ranks $\underline{R}_t(\boldsymbol{\theta})$.

Let $D_k(J;f) := \int_0^1 J(u)\,\tilde{F}_k^{-1}(u)\,du$ and $C_k(J;f) := \int_0^1 J(u)\,\varphi_f \circ \tilde{F}_k^{-1}(u)\,du$, where J denotes some score function defined over $]0,1[$. When J is the score associated with some radial density f_1 (namely, when $J_0 = J_1 = \varphi_{f_1} \circ \tilde{F}_{1k}^{-1}$ and $J_2 = \tilde{F}_{1k}^{-1}$), we write $D_k(f_1, f_2)$ and $C_k(f_1, f_2)$ for $D_k(\tilde{F}_{1k}^{-1}; f_2)$ and $C_k(\varphi_{f_1} \circ \tilde{F}_{1k}^{-1}; f_2)$, respectively; for simplicity, we also write $C_k(f)$ and $D_k(f)$ instead of $C_k(f,f)$ and $D_k(f,f)$. The asymptotic behavior of the nonparametric statistics (21) and (22) readily follows from Proposition 2 and the following lemma (see Hallin and Paindaveine 2004b).

Lemma 1 *Let the assumptions of Proposition 2 hold. For all couples of integers $(l, \tilde{l})$, the vector*

$$\Big(n^{1/2}\,(\mathrm{vec}\,\widetilde{\mathbf{\Lambda}}^{(n)}_{0;J;\mathbf{\Sigma},f}(\boldsymbol{\theta}))', \ldots, (n-l+1)^{1/2}\,(\mathrm{vec}\,\widetilde{\mathbf{\Lambda}}^{(n)}_{l-1;J;\mathbf{\Sigma},f}(\boldsymbol{\theta}))',$$
$$(n-1)^{1/2}\,(\mathrm{vec}\,\widetilde{\mathbf{\Gamma}}^{(n)}_{1;J;\mathbf{\Sigma},f}(\boldsymbol{\theta}))', \ldots, (n-\tilde{l})^{1/2}\,(\mathrm{vec}\,\widetilde{\mathbf{\Gamma}}^{(n)}_{\tilde{l};J;\mathbf{\Sigma},f}(\boldsymbol{\theta}))'\Big)'$$

is asymptotically normal, with mean $\mathbf{0}$ and mean

$$\begin{pmatrix} \frac{1}{k}\,C_k(J_0;f)(\mathbf{I}_{lm} \otimes \mathbf{\Sigma}^{-1})[\lim_{n\to\infty}(\boldsymbol{K}_{l,n} \otimes \mathbf{I}_k)\,\mathbf{L}^{(n)}_{\boldsymbol{\theta}}]\,(\mathrm{vec}\,\boldsymbol{\eta}') \\ \frac{1}{k^2}\,C_k(J_1;f)\,D_k(J_2;f)\,[\mathbf{I}_{\tilde{l}} \otimes (\mathbf{\Sigma} \otimes \mathbf{\Sigma}^{-1})]\,\mathbf{Q}^{(\tilde{l}+1)}_{\boldsymbol{\theta}}\,\mathbf{P}_{\boldsymbol{\theta}}\,\mathbf{M}_{\boldsymbol{\theta}}\,((\mathrm{vec}\,\boldsymbol{\gamma})', (\mathrm{vec}\,\boldsymbol{\delta})')' \end{pmatrix}$$

under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \mathbf{\Sigma}, f)$ and $\mathcal{H}^{(n)}(\boldsymbol{\theta} + \boldsymbol{\nu}(n)\boldsymbol{\tau}, \mathbf{\Sigma}, f)$, respectively, and covariance matrix

$$\begin{pmatrix} \frac{1}{k}\,\mathrm{E}[J_0^2(U)]\,(\boldsymbol{K}_l \otimes \mathbf{\Sigma}^{-1}) & \mathbf{0} \\ \mathbf{0} & \frac{1}{k^2}\,\mathrm{E}[J_1^2(U)]\,\mathrm{E}[J_2^2(U)]\;[\mathbf{I}_{\tilde{l}} \otimes (\mathbf{\Sigma} \otimes \mathbf{\Sigma}^{-1})] \end{pmatrix}$$

under both.

Letting $\mathbf{h}_j = \mathbf{h}_j(\boldsymbol{\theta}) := \mathbf{H}_j(\boldsymbol{\theta}) - \sum_{i=1}^{\min(p_0,j)} \mathbf{H}_{j-i}(\boldsymbol{\theta})\mathbf{A}_i(\boldsymbol{\theta})$, $j = 0, 1, 2, \ldots$, note that

$$\lim_{n\to\infty}(\boldsymbol{K}_{l,n} \otimes \mathbf{I}_k)\,\mathbf{L}^{(n)}_{\boldsymbol{\theta}} = \begin{pmatrix} \sum_{j=0}^{\infty}(\mathbf{K}'\,\mathbf{R}_{|j|}\,\mathbf{K}) \otimes \mathbf{h}_j \\ \vdots \\ \sum_{j=0}^{\infty}(\mathbf{K}'\,\mathbf{R}_{|i-j|}\,\mathbf{K}) \otimes \mathbf{h}_j \\ \vdots \\ \sum_{j=0}^{\infty}(\mathbf{K}'\,\mathbf{R}_{|l-j-1|}\,\mathbf{K}) \otimes \mathbf{h}_j \end{pmatrix}.$$

Also, defining

$$\mathbf{a}_i(\boldsymbol{\tau};\boldsymbol{\theta}) := \sum_{j=1}^{\min(p_1,i)} \sum_{l=0}^{i-j} \sum_{k=0}^{\min(q_0,i-j-l)} (\mathbf{G}_{i-j-l-k}(\boldsymbol{\theta})\mathbf{B}_k(\boldsymbol{\theta}) \otimes \mathbf{H}'_l(\boldsymbol{\theta}))'\,\mathrm{vec}\,\boldsymbol{\gamma}_j,$$

and

$$\mathbf{b}_i(\boldsymbol{\tau};\boldsymbol{\theta}) := \sum_{j=1}^{\min(q_1,i)} (\mathbf{I}_k \otimes \mathbf{H}_{i-j}(\boldsymbol{\theta}))\,\mathrm{vec}\,\boldsymbol{\delta}_j,$$

one can easily check that

$$\begin{pmatrix} \mathbf{a}_1(\boldsymbol{\tau};\boldsymbol{\theta}) + \mathbf{b}_1(\boldsymbol{\tau};\boldsymbol{\theta}) \\ \vdots \\ \mathbf{a}_{\tilde{l}}(\boldsymbol{\tau};\boldsymbol{\theta}) + \mathbf{b}_{\tilde{l}}(\boldsymbol{\tau};\boldsymbol{\theta}) \end{pmatrix} = \mathbf{Q}^{(\tilde{l}+1)}_{\boldsymbol{\theta}}\,\mathbf{P}_{\boldsymbol{\theta}}\,\mathbf{M}_{\boldsymbol{\theta}} \begin{pmatrix} \mathrm{vec}\,\boldsymbol{\gamma} \\ \mathrm{vec}\,\boldsymbol{\delta} \end{pmatrix}.$$

This allows for a direct comparison between Lemma 1 and the corresponding univariate result (Proposition 4.3 in Hallin and Puri 1994).

4.5 Equivariance/invariance properties.

In this section, we use hats to indicate that all parameters involved are estimated. Consider the original sample $(\mathbf{Y}_1, \ldots, \mathbf{Y}_n)$ and the transformed sample $(\mathbf{M}\mathbf{Y}_1, \ldots, \mathbf{M}\mathbf{Y}_n)$, where $\mathbf{M}$ is a full-rank $k \times k$ matrix, and denote by $T(\mathbf{M})$ (resp., T) the value of a statistic T computed from the transformed (resp., original) sample. Assumption (E2) ensures that the residual sample of the $\hat{\mathbf{Z}}_i(\mathbf{M}) = \mathbf{Z}_i(\hat{\boldsymbol{\theta}}(\mathbf{M}))$'s is affine-equivariant, meaning that

$$(\hat{\mathbf{Z}}_1(\mathbf{M}), \ldots, \hat{\mathbf{Z}}_n(\mathbf{M})) = (\mathbf{M}\hat{\mathbf{Z}}_1, \ldots, \mathbf{M}\hat{\mathbf{Z}}_n).$$

Under Assumption (D2), $\widehat{\boldsymbol{\Sigma}}^{-1/2}$ enjoys the equivariance property

$$\widehat{\boldsymbol{\Sigma}}^{-1/2}(\mathbf{M}) = d^{-1/2}\mathbf{O}\widehat{\boldsymbol{\Sigma}}^{-1/2}\mathbf{M}^{-1}, \tag{25}$$

for some $k \times k$ orthogonal matrix $\mathbf{O}$ (recall that $\widehat{\boldsymbol{\Sigma}}(\mathbf{M})$ and $\widehat{\boldsymbol{\Sigma}}$ are computed from the residual samples $(\hat{\mathbf{Z}}_1(\mathbf{M}), \ldots, \hat{\mathbf{Z}}_n(\mathbf{M}))$ and $(\hat{\mathbf{Z}}_1, \ldots, \hat{\mathbf{Z}}_n)$, respectively). The affine-invariance/equivariance properties of pseudo-Mahalanobis signs and ranks easily follow. More precisely, denoting by $\hat{\mathbf{W}}_t(\mathbf{M})$ and $\widehat{R}_t(\mathbf{M})$ the pseudo-Mahalanobis signs and ranks computed from the transformed residuals $(\hat{\mathbf{Z}}_1(\mathbf{M}), \ldots, \hat{\mathbf{Z}}_n(\mathbf{M}))$, we have

$$\hat{\mathbf{W}}_t(\mathbf{M}) = \mathbf{O}\hat{\mathbf{W}}_t, \quad \widehat{R}_t(\mathbf{M}) = \widehat{R}_t,$$

where $\mathbf{O}$ is the orthogonal matrix in (25).

As for hyperplane-based signs and ranks, absolute interdirections are only asymptotically affine-equivariant, i.e., under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$,

$$\hat{\mathbf{V}}_t(\mathbf{M}) = \mathbf{O}\hat{\mathbf{V}}_t + o_{\mathrm{P}}(1), \quad \text{as } n \to \infty, \tag{26}$$

still with the orthogonal matrix $\mathbf{O}$ in (25), whereas lift-interdirection ranks $\hat{\underline{R}}_t := \underline{R}_t(\hat{\boldsymbol{\theta}})$ are strictly affine-invariant (see Oja and Paindaveine 2004).

This entails, for the nonparametric statistics $\hat{\underset{\sim}{\boldsymbol{\Lambda}}}{}^{(n)}_{i;J}$ and $\hat{\underset{\sim}{\boldsymbol{\Gamma}}}{}^{(n)}_{i;J}$, the following equivariance properties.

Lemma 2 *Assume that Assumptions (D2) and (E2) hold. Denote by* $\hat{\underset{\sim}{\boldsymbol{\Lambda}}}{}^{(n)}_{i;J}(\mathbf{M})$ *and* $\hat{\underset{\sim}{\boldsymbol{\Gamma}}}{}^{(n)}_{i;J}(\mathbf{M})$ *the statistics* $\hat{\underset{\sim}{\boldsymbol{\Lambda}}}{}^{(n)}_{i;J}$ *and* $\hat{\underset{\sim}{\boldsymbol{\Gamma}}}{}^{(n)}_{i;J}$ *computed from the n-tuple* $(\mathbf{M}\mathbf{Y}_1, \ldots, \mathbf{M}\mathbf{Y}_n)$*, where* $\mathbf{M}$ *is a* $k \times k$ *full-rank matrix. Then,*

$$\hat{\underset{\sim}{\boldsymbol{\Lambda}}}{}^{(n)}_{i;J}(\mathbf{M}) = d^{-1/2}\mathbf{M}^{-1\prime}\,\hat{\underset{\sim}{\boldsymbol{\Lambda}}}{}^{(n)}_{i;J} \quad \text{and} \quad \hat{\underset{\sim}{\boldsymbol{\Gamma}}}{}^{(n)}_{i;J}(\mathbf{M}) = \mathbf{M}^{-1\prime}\,\hat{\underset{\sim}{\boldsymbol{\Gamma}}}{}^{(n)}_{i;J}\mathbf{M}';$$

the same result still holds if in $\hat{\underset{\sim}{\boldsymbol{\Lambda}}}{}^{(n)}_{i;J}$ *and* $\hat{\underset{\sim}{\boldsymbol{\Gamma}}}{}^{(n)}_{i;J}$ *the pseudo-Mahalanobis ranks* $\widehat{R}_t(\hat{\boldsymbol{\theta}})$ *are replaced by the lift-interdirection ranks* $\underline{R}_t(\hat{\boldsymbol{\theta}})$*.*

Proof. The result directly follows from the equivariance and invariance properties of pseudo-Mahalanobis signs and ranks. □

If the pseudo-Mahalanobis signs $\mathbf{W}_t(\hat{\boldsymbol{\theta}})$ in $\hat{\underset{\sim}{\boldsymbol{\Lambda}}}{}^{(n)}_{i;J}$ and $\hat{\underset{\sim}{\boldsymbol{\Gamma}}}{}^{(n)}_{i;J}$ are replaced by the corresponding absolute interdirections (in combination with any type of ranks), then it is clear from (26) that $\hat{\underset{\sim}{\boldsymbol{\Lambda}}}{}^{(n)}_{i;J}$ and $\hat{\underset{\sim}{\boldsymbol{\Gamma}}}{}^{(n)}_{i;J}$ can only be asymptotically affine-equivariant. The resulting hyperplane-based test statistics will accordingly be only asymptotically affine-invariant (see Section 5 and the proof of Proposition 3).

5 Aligned rank tests.

5.1 The proposed rank-based procedures.

Considering the linear restriction characterized by $(\boldsymbol{\theta}_0, \boldsymbol{\Upsilon})$, assume that (A2) holds, and that $\boldsymbol{\theta}$ belongs to $(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$. Let $n^{1/2}\,\underset{\sim}{\mathbf{T}}{}^{(n)}_{J}(\boldsymbol{\theta})$ be given by

$$\begin{pmatrix} n^{1/2}\,\underset{\sim}{\mathbf{T}}{}^{(n)}_{I;J}(\boldsymbol{\theta}) \\ n^{1/2}\,\underset{\sim}{\mathbf{T}}{}^{(n)}_{II;J}(\boldsymbol{\theta}) \end{pmatrix} := \begin{pmatrix} \mathbf{L}^{(n)\prime}_{\boldsymbol{\theta}}\,\underset{\sim}{\mathbf{S}}{}^{(n)}_{I;J}(\boldsymbol{\theta}) \\ \mathbf{Q}^{(n)\prime}_{\boldsymbol{\theta}}\,\underset{\sim}{\mathbf{S}}{}^{(n)}_{II;J}(\boldsymbol{\theta}) \end{pmatrix} := \begin{pmatrix} \mathbf{L}^{(n)\prime}_{\boldsymbol{\theta}}(n^{1/2}(\mathrm{vec}\,\underset{\sim}{\boldsymbol{\Lambda}}{}^{(n)}_{0;J}(\boldsymbol{\theta}))', \ldots, (\mathrm{vec}\,\underset{\sim}{\boldsymbol{\Lambda}}{}^{(n)}_{n-1;J}(\boldsymbol{\theta}))')' \\ \mathbf{Q}^{(n)\prime}_{\boldsymbol{\theta}}((n-1)^{1/2}\,(\mathrm{vec}\,\underset{\sim}{\boldsymbol{\Gamma}}{}^{(n)}_{1;J}(\boldsymbol{\theta}))', \ldots, (\mathrm{vec}\,\underset{\sim}{\boldsymbol{\Gamma}}{}^{(n)}_{n-1;J}(\boldsymbol{\theta}))')' \end{pmatrix},$$

and define

$$\mathbf{J}^{(n)}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}} := \mathbf{L}^{(n)\prime}_{\boldsymbol{\theta}}\,(\boldsymbol{K}_n \otimes \boldsymbol{\Sigma}^{-1})\,\mathbf{L}^{(n)}_{\boldsymbol{\theta}} \quad \text{and} \quad \mathbf{J}^{(n)}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}} := \mathbf{Q}^{(n)\prime}_{\boldsymbol{\theta}}\left[\mathbf{I}_{n-1} \otimes (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}^{-1})\right]\mathbf{Q}^{(n)}_{\boldsymbol{\theta}}.$$

Denote by $\dot{\mathbf{M}}_{\boldsymbol{\theta}}$ the full-rank $k^2\pi_0 \times k^2(p_0+q_0)$ matrix resulting from $\mathbf{M}_{\boldsymbol{\theta}}$ by deleting columns $k^2p_0+1, \ldots, k^2p_1$ and $k^2(p_1+q_0)+1, \ldots, k^2(p_1+q_1)$. Similarly, let $\dot{\boldsymbol{\Upsilon}}_{II}$ be the $k^2(p_0+q_0) \times r_{II}$ array resulting from $\boldsymbol{\Upsilon}_{II}$ by deleting lines $k^2p_0+1, \ldots, k^2p_1$ and $k^2(p_1+q_0)+1, \ldots, k^2(p_1+q_1)$. Note that $\mathbf{M}_{\boldsymbol{\theta}}\,\boldsymbol{\Upsilon}_{II} = \dot{\mathbf{M}}_{\boldsymbol{\theta}}\,\dot{\boldsymbol{\Upsilon}}_{II}$. Finally, let

$$\bar{\mathbf{Q}}^{(n)}_{I;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}) := \frac{k}{\mathrm{E}[J_0^2(U)]}\Big[(\mathbf{J}^{(n)}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}})^{-1} - (\mathbf{K}^{(n)} \otimes \mathbf{I}_k)^{-1}\,\boldsymbol{\Upsilon}_I$$
$$\left(\boldsymbol{\Upsilon}'_I\,(\mathbf{K}^{(n)} \otimes \mathbf{I}_k)'^{-1}\,\mathbf{J}^{(n)}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,(\mathbf{K}^{(n)} \otimes \mathbf{I}_k)^{-1}\,\boldsymbol{\Upsilon}_I\right)^{-1}\boldsymbol{\Upsilon}'_I\,(\mathbf{K}^{(n)} \otimes \mathbf{I}_k)'^{-1}\Big],$$

and, denoting by $\mathbf{A}^-$ an arbitrary generalized inverse of $\mathbf{A}$,

$$\bar{\mathbf{Q}}^{(n)}_{II;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}) := \frac{k^2}{\mathrm{E}[J_1^2(U)]\mathrm{E}[J_2^2(U)]}\Big[(\mathbf{J}^{(n)}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}})^{-1} - \mathbf{P}_{\boldsymbol{\theta}}\dot{\mathbf{M}}_{\boldsymbol{\theta}}\,\dot{\boldsymbol{\Upsilon}}_{II}$$
$$\left(\dot{\boldsymbol{\Upsilon}}'_{II}\,\dot{\mathbf{M}}'_{\boldsymbol{\theta}}\,\mathbf{P}'_{\boldsymbol{\theta}}\mathbf{J}^{(n)}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,\mathbf{P}_{\boldsymbol{\theta}}\dot{\mathbf{M}}_{\boldsymbol{\theta}}\,\dot{\boldsymbol{\Upsilon}}_{II}\right)^{-}\dot{\boldsymbol{\Upsilon}}'_{II}\,\dot{\mathbf{M}}'_{\boldsymbol{\theta}}\,\mathbf{P}'_{\boldsymbol{\theta}}\Big].$$

Then the J-score version of the proposed test statistics is

$$\widehat{\mathcal{W}}^{(n)}_{J} = \widehat{\mathcal{W}}^{(n)}_{J}(\hat{\boldsymbol{\theta}}) := n\left(\underset{\sim}{\mathbf{T}}{}^{(n)}_{J}(\hat{\boldsymbol{\theta}})\right)'\begin{pmatrix} \bar{\mathbf{Q}}_{I;J;\widehat{\boldsymbol{\Sigma}}}(\hat{\boldsymbol{\theta}}) & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{Q}}_{II;J;\widehat{\boldsymbol{\Sigma}}}(\hat{\boldsymbol{\theta}}) \end{pmatrix}\underset{\sim}{\mathbf{T}}{}^{(n)}_{J}(\hat{\boldsymbol{\theta}}) =: \widehat{\mathcal{W}}^{(n)}_{I;J} + \widehat{\mathcal{W}}^{(n)}_{II;J}$$
$$:= n\left(\underset{\sim}{\mathbf{T}}{}^{(n)}_{I;J}(\hat{\boldsymbol{\theta}})\right)'\bar{\mathbf{Q}}^{(n)}_{I;J;\widehat{\boldsymbol{\Sigma}}}(\hat{\boldsymbol{\theta}})\,\underset{\sim}{\mathbf{T}}{}^{(n)}_{I;J}(\hat{\boldsymbol{\theta}}) + n\left(\underset{\sim}{\mathbf{T}}{}^{(n)}_{II;J}(\hat{\boldsymbol{\theta}})\right)'\bar{\mathbf{Q}}^{(n)}_{II;J;\widehat{\boldsymbol{\Sigma}}}(\hat{\boldsymbol{\theta}})\,\underset{\sim}{\mathbf{T}}{}^{(n)}_{II;J}(\hat{\boldsymbol{\theta}}),$$

where the estimators $\widehat{\boldsymbol{\Sigma}} = \widehat{\boldsymbol{\Sigma}}^{(n)}$ and $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}^{(n)}$ satisfy Assumptions (D1)-(D2) and (E1)-(E2), respectively. The scores allowing for local asymptotic optimality at radial density $f_\star$ are $J_0 = J_1 := \varphi_{f_\star} \circ \tilde{F}^{-1}_{\star k}$ and $J_2 := \tilde{F}^{-1}_{\star k}$. The corresponding statistics will be denoted by $\widehat{\mathcal{W}}^{(n)}_{f_\star}$.

Finally, in order to describe the asymptotic behavior of $\widehat{\mathcal{W}}^{(n)}_{J}$ under local alternatives, define

$$r_{\boldsymbol{\theta},\boldsymbol{\Sigma}}(\boldsymbol{\eta}) := \left(\mathrm{vec}\,\boldsymbol{\eta}'\right)' \Big[\mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}} - \mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,(\mathbf{K}^{-1}\mathbf{D}\otimes\mathbf{I}_k)\,\boldsymbol{\Upsilon}_I$$

$$\times\left(\boldsymbol{\Upsilon}_I'\,(\mathbf{K}^{-1}\mathbf{D}\otimes\mathbf{I}_k)'\,\mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,(\mathbf{K}^{-1}\mathbf{D}\otimes\mathbf{I}_k)\,\boldsymbol{\Upsilon}_I\right)^{-1}\boldsymbol{\Upsilon}_I'\,(\mathbf{K}^{-1}\mathbf{D}\otimes\mathbf{I}_k)'\mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}\Big]\left(\mathrm{vec}\,\boldsymbol{\eta}'\right)$$

and

$$s_{\boldsymbol{\theta},\boldsymbol{\Sigma}}(\boldsymbol{\gamma},\boldsymbol{\delta}) := \begin{pmatrix}\mathrm{vec}\,\boldsymbol{\gamma}\\ \mathrm{vec}\,\boldsymbol{\delta}\end{pmatrix}' \left[\mathbf{N}_{\boldsymbol{\theta},\boldsymbol{\Sigma}} - \mathbf{N}_{\boldsymbol{\theta},\boldsymbol{\Sigma}}\,\boldsymbol{\Upsilon}_{II}\,(\boldsymbol{\Upsilon}_{II}'\mathbf{N}_{\boldsymbol{\theta},\boldsymbol{\Sigma}}\,\boldsymbol{\Upsilon}_{II})^{-}\,\boldsymbol{\Upsilon}_{II}'\mathbf{N}_{\boldsymbol{\theta},\boldsymbol{\Sigma}}\right]\begin{pmatrix}\mathrm{vec}\,\boldsymbol{\gamma}\\ \mathrm{vec}\,\boldsymbol{\delta}\end{pmatrix},$$

where $\mathbf{D}$ is the array involved in Assumption (A1′) and $\mathbf{N}_{\boldsymbol{\theta},\boldsymbol{\Sigma}}$ is defined in Proposition 1. We now can state the main result of this paper.

Proposition 3 *Assume that (A1), (A2), (B1′), (B2), (B3), (C), (D1), (D2), (E1), and (E2) hold. Consider the sequence of aligned rank tests* $\phi_J^{(n)}$ *(resp.,* $\phi_{f_\star}^{(n)}$*) that reject the null hypothesis* $\mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\Upsilon}}$ *whenever* $\widehat{\mathcal{W}}_J^{(n)}$ *(resp.,* $\widehat{\mathcal{W}}_{f_\star}^{(n)}$*) exceeds the* α*-upper quantile* $\chi^2_{km+k^2\pi_0-r,1-\alpha}$ *of a chi-square distribution with* $km+k^2\pi_0-r$ *degrees of freedom. Then,*

(i) $\widehat{\mathcal{W}}_J^{(n)}$ *is strictly affine-invariant (only asymptotically so, if absolute interdirections are used as multivariate signs), and asymptotically invariant with respect to the group of continuous monotone radial transformations;*

(ii) $\widehat{\mathcal{W}}_J^{(n)}$ *is asymptotically chi-square with* $km+k^2\pi_0-r$ *degrees of freedom under* $\mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\Upsilon}}$ *(so that* $\phi_J^{(n)}$ *has asymptotic level* α*);*

(iii) $\widehat{\mathcal{W}}_J^{(n)}$ *is asymptotically noncentral chi-square, still with* $km+k^2\pi_0-r$ *degrees of freedom, and with noncentrality parameter*

$$\frac{1}{k}\,\frac{C_k^2(J_0;f)}{\mathrm{E}[J_0^2(U)]}\,r_{\boldsymbol{\theta},\boldsymbol{\Sigma}}(\boldsymbol{\eta}) + \frac{1}{k^2}\,\frac{C_k^2(J_1;f)}{\mathrm{E}[J_1^2(U)]}\,\frac{D_k^2(J_2;f)}{\mathrm{E}[J_2^2(U)]}\,s_{\boldsymbol{\theta},\boldsymbol{\Sigma}}(\boldsymbol{\gamma},\boldsymbol{\delta})$$

under $\mathcal{H}^{(n)}(\boldsymbol{\theta}+\boldsymbol{\nu}(n)\boldsymbol{\tau},\boldsymbol{\Sigma},f)$*,* $\boldsymbol{\theta}\in(\boldsymbol{\theta}_0+\mathcal{M}(\boldsymbol{\Upsilon}))^\star$ *and* $\boldsymbol{\nu}(n)\boldsymbol{\tau}\notin\mathcal{M}(\boldsymbol{\Upsilon})$*, provided that (A1) is reinforced into (A1′);*

(iv) for any $f_\star$ *satisfying Assumptions (B1′), (B2), (B3) and (C′), the sequence of tests* $\phi_{f_\star}^{(n)}$ *is locally asymptotically most stringent for* $\mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\Upsilon}}$ *against* $\bigcup_{\boldsymbol{\theta}\notin\boldsymbol{\theta}_0+\mathcal{M}(\boldsymbol{\Upsilon})}\bigcup_{\boldsymbol{\Sigma}}\mathcal{H}^{(n)}(\boldsymbol{\theta},\boldsymbol{\Sigma},f_\star)$*, at asymptotic probability level* α*.*

The proof of this proposition is based on the following asymptotic linearity property; proofs are given in the Appendix.

Lemma 3 *Let* $\hat{\boldsymbol{\theta}}^{(n)}$ *and* $\boldsymbol{\theta}$ *denote a sequence of estimators and a parameter value such that, under* $\mathcal{H}^{(n)}(\boldsymbol{\theta},\boldsymbol{\Sigma},f)$*, as* $n\to\infty$*,* $\hat{\boldsymbol{\theta}}^{(n)}$ *satisfies at* $\boldsymbol{\theta}$ *the root-n consistency and local asymptotic discreteness properties given in parts (ii) and (iii) of Assumption (E1). Assume that (A1), (B1′), (B2), (B3), (C), and (D1) hold, and partition* $\boldsymbol{\theta}$ *into* $(\boldsymbol{\theta}_I',\boldsymbol{\theta}_{II}')'\in\mathbb{R}^{km}\times\mathbb{R}^{k^2(p_1+q_1)}$*. Then,*

$$n^{1/2}(\underset{\sim}{\mathbf{T}}{}^{(n)}_{I;J}(\hat{\boldsymbol{\theta}}) - \underset{\sim}{\mathbf{T}}{}^{(n)}_{I;J}(\boldsymbol{\theta})) + \frac{1}{k}\,C_k(J_0;f)\,\mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,(\mathbf{K}^{(n)}\otimes\mathbf{I}_k)^{-1}\,n^{1/2}(\hat{\boldsymbol{\theta}}_I^{(n)}-\boldsymbol{\theta}_I)$$

and

$$n^{1/2}(\underset{\sim}{\mathbf{T}}^{(n)}_{II;J}(\hat{\boldsymbol{\theta}}) - \underset{\sim}{\mathbf{T}}^{(n)}_{II;J}(\boldsymbol{\theta})) + \frac{1}{k^2}\,D_k(J_2;f)\,C_k(J_1;f)\,\mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,\mathbf{P}_{\boldsymbol{\theta}}\,\mathbf{M}_{\boldsymbol{\theta}}\,n^{1/2}(\hat{\boldsymbol{\theta}}^{(n)}_{II} - \boldsymbol{\theta}_{II})$$

are $o_{\mathrm{P}}(1)$*, still under* $\mathcal{H}^{(n)}(\boldsymbol{\theta},\boldsymbol{\Sigma},f)$*, as* $n\to\infty$*.*

Note that $\underset{\sim}{\mathbf{T}}^{(n)}_{J}(\hat{\boldsymbol{\theta}})$, $\dot{\mathbf{M}}_{\hat{\boldsymbol{\theta}}}$, $\dot{\boldsymbol{\Upsilon}}_{II}$, and π_0—hence also the test statistic $\widehat{\mathcal{W}}^{(n)}_J$ and $\phi^{(n)}_J$— depend on p_0, p_1, q_0 and q_1 through $\pi = \max(p_1-p_0, q_1-q_0)$ only. Also, it should be stressed that $\widehat{\mathcal{W}}^{(n)}_J$ does not depend on the particular choice of the fundamental system $\widehat{\boldsymbol{\Psi}} := \{\boldsymbol{\Psi}^{(1)}_t(\hat{\boldsymbol{\theta}}),\ldots,\boldsymbol{\Psi}^{(p_0+q_0)}_t(\hat{\boldsymbol{\theta}})\}$ in $\mathbf{P}_{\hat{\boldsymbol{\theta}}}$ and $\mathbf{Q}^{(n)}_{\hat{\boldsymbol{\theta}}}$. Indeed, for any fundamental system $\widehat{\boldsymbol{\Phi}} := \{\boldsymbol{\Phi}^{(1)}_t(\hat{\boldsymbol{\theta}}),\ldots,\boldsymbol{\Phi}^{(p_0+q_0)}_t(\hat{\boldsymbol{\theta}})\}$, there exists an invertible matrix $\bar{\boldsymbol{\Lambda}}$ such that $\mathbf{Q}_{\hat{\boldsymbol{\theta}};\hat{\boldsymbol{\Phi}}} = \mathbf{Q}_{\hat{\boldsymbol{\theta}};\hat{\boldsymbol{\Psi}}}\,\bar{\boldsymbol{\Lambda}}$ (see the proof of Proposition 4(i) in Hallin and Paindaveine (2004a)). It easily follows that $\underset{\sim}{\mathbf{T}}^{(n)}_{II;J;\hat{\boldsymbol{\Phi}}}(\hat{\boldsymbol{\theta}}) := \bar{\boldsymbol{\Lambda}}'\,\underset{\sim}{\mathbf{T}}^{(n)}_{II;J;\hat{\boldsymbol{\Psi}}}(\hat{\boldsymbol{\theta}})$, $\mathbf{J}^{(n)}_{II;\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}};\hat{\boldsymbol{\Phi}}} = \bar{\boldsymbol{\Lambda}}'\,\mathbf{J}^{(n)}_{II;\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}};\hat{\boldsymbol{\Psi}}}\,\bar{\boldsymbol{\Lambda}}$, and $\mathbf{P}_{\hat{\boldsymbol{\theta}};\hat{\boldsymbol{\Phi}}} = \bar{\boldsymbol{\Lambda}}^{-1}\,\mathbf{P}_{\hat{\boldsymbol{\theta}};\hat{\boldsymbol{\Phi}}}$. Since the nonserial part $\widehat{\mathcal{W}}^{(n)}_{I;J}$ of $\widehat{\mathcal{W}}^{(n)}_J$ clearly is not affected by the choice of $\widehat{\boldsymbol{\Psi}}$, we obtain that $\widehat{\mathcal{W}}^{(n)}_{J;\hat{\boldsymbol{\Phi}}} = \widehat{\mathcal{W}}^{(n)}_{J;\hat{\boldsymbol{\Psi}}}$.

5.2 The Gaussian procedure.

Let the same assumptions hold (about $\boldsymbol{\theta}_0, \boldsymbol{\Upsilon}$, etc.) as in Section 5.1. In order to compute asymptotic relative efficiencies, we now provide the Gaussian parametric counterparts of the rank-based procedures developed in the previous section. Define

$$\mathbf{J}^{(n)}_{I;\mathcal{N};\boldsymbol{\theta}} := \mathbf{L}^{(n)\prime}_{\boldsymbol{\theta}}\left(\boldsymbol{K}_n \otimes (\mathbf{S}^{(n)}_{\boldsymbol{\theta}})^{-1}\right)\mathbf{L}^{(n)}_{\boldsymbol{\theta}} \quad \text{and} \quad \mathbf{J}^{(n)}_{II;\mathcal{N};\boldsymbol{\theta}} := \mathbf{Q}^{(n)\prime}_{\boldsymbol{\theta}}\left[\mathbf{I}_{n-1}\otimes\hat{\boldsymbol{\Gamma}}^{(n)}_{\boldsymbol{\theta}}\right]\mathbf{Q}^{(n)}_{\boldsymbol{\theta}},$$

where $\mathbf{S}^{(n)}_{\boldsymbol{\theta}} := n^{-1}\sum_{t=1}^n \mathbf{Z}_t(\boldsymbol{\theta})\mathbf{Z}'_t(\boldsymbol{\theta})$ is a consistent estimator, under $\mathcal{H}^{(n)}(\boldsymbol{\theta},\boldsymbol{\Sigma},f)$, of the innovation covariance $(\mathrm{E}[(\tilde{F}_k^{-1}(U))^2]/k)\,\boldsymbol{\Sigma}$, and

$$\hat{\boldsymbol{\Gamma}}^{(n)}_{\boldsymbol{\theta}} := (n-1)^{-1}\sum_{t=2}^{n}\mathrm{vec}\left(\mathbf{Z}_t(\boldsymbol{\theta})\mathbf{Z}'_{t-1}(\boldsymbol{\theta})\right)\left(\mathrm{vec}\left(\mathbf{Z}_t(\boldsymbol{\theta})\mathbf{Z}'_{t-1}(\boldsymbol{\theta})\right)\right)'$$

is consistent for $(\mathrm{E}[(\tilde{F}_k^{-1}(U))^2]/k)^2\,\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}^{-1}$ under the same sequence of hypotheses. Let

$$\begin{aligned}\bar{\mathbf{Q}}^{(n)}_{I;\mathcal{N}}(\boldsymbol{\theta}) := (\mathbf{J}^{(n)}_{I;\mathcal{N};\boldsymbol{\theta}})^{-1} - (\mathbf{K}^{(n)}\otimes\mathbf{I}_k)^{-1}\,\boldsymbol{\Upsilon}_I \\ \times\left(\boldsymbol{\Upsilon}'_I\,(\mathbf{K}^{(n)}\otimes\mathbf{I}_k)'^{-1}\,\mathbf{J}^{(n)}_{I;\mathcal{N};\boldsymbol{\theta}}\,(\mathbf{K}^{(n)}\otimes\mathbf{I}_k)^{-1}\,\boldsymbol{\Upsilon}_I\right)^{-1}\boldsymbol{\Upsilon}'_I\,(\mathbf{K}^{(n)}\otimes\mathbf{I}_k)'^{-1},\end{aligned}$$

and

$$\bar{\mathbf{Q}}^{(n)}_{II;\mathcal{N}}(\boldsymbol{\theta}) := (\mathbf{J}^{(n)}_{II;\mathcal{N};\boldsymbol{\theta}})^{-1} - \mathbf{P}_{\boldsymbol{\theta}}\dot{\mathbf{M}}_{\boldsymbol{\theta}}\,\dot{\boldsymbol{\Upsilon}}_{II}\left(\dot{\boldsymbol{\Upsilon}}'_{II}\,\dot{\mathbf{M}}'_{\boldsymbol{\theta}}\,\mathbf{P}'_{\boldsymbol{\theta}}\mathbf{J}^{(n)}_{II;\mathcal{N};\boldsymbol{\theta}}\,\mathbf{P}_{\boldsymbol{\theta}}\dot{\mathbf{M}}_{\boldsymbol{\theta}}\,\dot{\boldsymbol{\Upsilon}}_{II}\right)^{-}\dot{\boldsymbol{\Upsilon}}'_{II}\,\dot{\mathbf{M}}'_{\boldsymbol{\theta}}\,\mathbf{P}'_{\boldsymbol{\theta}}.$$

Then the Gaussian parametric test statistic is

$$\widehat{\mathcal{W}}^{(n)}_{\mathcal{N}} := n\left(\mathbf{T}^{(n)}_{I;\mathbf{S},\phi}(\hat{\boldsymbol{\theta}})\right)'\bar{\mathbf{Q}}^{(n)}_{I;\mathcal{N}}(\hat{\boldsymbol{\theta}})\,\mathbf{T}^{(n)}_{I;\mathbf{S},\phi}(\hat{\boldsymbol{\theta}}) + n\left(\mathbf{T}^{(n)\prime}_{II;\mathbf{S},\phi}(\hat{\boldsymbol{\theta}})\right)'\bar{\mathbf{Q}}^{(n)}_{II;\mathcal{N}}(\hat{\boldsymbol{\theta}})\,\mathbf{T}^{(n)}_{II;\mathbf{S},\phi}(\hat{\boldsymbol{\theta}}),$$

where $\mathbf{T}^{(n)}_{I;\mathbf{S},\phi}(\boldsymbol{\theta})$ and $\mathbf{T}^{(n)}_{II;\mathbf{S},\phi}(\boldsymbol{\theta})$ are defined in (17) and (18) respectively, $\mathbf{S} := \mathbf{S}^{(n)}_{\boldsymbol{\theta}}$, and $\phi(r) := \exp(-r^2/2)$ stands for the Gaussian radial density. Note that $\mathbf{T}^{(n)}_{I;\mathbf{S},\phi}(\boldsymbol{\theta})$ and $\mathbf{T}^{(n)}_{II;\mathbf{S},\phi}(\boldsymbol{\theta})$ are based on Gaussian statistics of the form

$$\boldsymbol{\Lambda}^{(n)}_{i;\mathbf{S},\phi}(\boldsymbol{\theta}) = (\mathbf{S}^{(n)}_{\boldsymbol{\theta}})^{-1}\left(\frac{1}{n-i}\sum_{t=i+1}^{n}\mathbf{Z}_t(\boldsymbol{\theta})\,\mathbf{x}^{(n)\prime}_{t-i}\mathbf{K}^{(n)}\right) \text{ and } \boldsymbol{\Gamma}^{(n)}_{i;\mathbf{S},\phi}(\boldsymbol{\theta}) = (\mathbf{S}^{(n)}_{\boldsymbol{\theta}})^{-1}\left(\frac{1}{n-i}\sum_{t=i+1}^{n}\mathbf{Z}_t(\boldsymbol{\theta})\,\mathbf{Z}'_{t-i}(\boldsymbol{\theta})\right), \tag{27}$$

respectively.

Proposition 4 *Assume that (A1), (A2), (B1′), (B2), (B3), (E1) and (E2) hold. Consider the sequence of parametric Gaussian tests* $\phi^{(n)}_{\mathcal{N}}$ *that reject the null hypothesis* $\mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\Upsilon}}$ *whenever* $\widehat{\mathcal{W}}^{(n)}_{\mathcal{N}}$ *exceeds the* α*-upper quantile* $\chi^2_{km+k^2\pi_0-r,1-\alpha}$ *of a chi-square distribution with* $km + k^2\pi_0 - r$ *degrees of freedom. Then,*

(i) $\widehat{\mathcal{W}}^{(n)}_{\mathcal{N}}$ *is strictly affine-invariant;*

(ii) $\widehat{\mathcal{W}}^{(n)}_{\mathcal{N}}$ *is asymptotically chi-square with* $km + k^2\pi_0 - r$ *degrees of freedom under* $\mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\Upsilon}}$ *(so that* $\phi^{(n)}_{\mathcal{N}}$ *has asymptotic level* α*);*

(iii) $\widehat{\mathcal{W}}^{(n)}_{\mathcal{N}}$ *is asymptotically noncentral chi-square, still with* $km + k^2\pi_0 - r$ *degrees of freedom but with noncentrality parameter*

$$\frac{k}{D_k(f)}\,r_{\boldsymbol{\theta},\boldsymbol{\Sigma}}(\boldsymbol{\eta}) + s_{\boldsymbol{\theta},\boldsymbol{\Sigma}}(\boldsymbol{\gamma},\boldsymbol{\delta}),$$

under $\mathcal{H}^{(n)}(\boldsymbol{\theta} + \boldsymbol{\nu}(n)\boldsymbol{\tau}, \boldsymbol{\Sigma}, f)$, $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$ *and* $\boldsymbol{\nu}(n)\boldsymbol{\tau} \notin \mathcal{M}(\boldsymbol{\Upsilon})$, *provided that (A1) is reinforced into (A1′);*

(iv) the sequence of tests $\phi^{(n)}_{\mathcal{N}}$ *is locally asymptotically most stringent for* $\mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\Upsilon}}$ *against*

$$\bigcup_{\boldsymbol{\theta}\notin\boldsymbol{\theta}_0+\mathcal{M}(\boldsymbol{\Upsilon})}\ \bigcup_{\boldsymbol{\Sigma}}\mathcal{H}^{(n)}(\boldsymbol{\theta},\boldsymbol{\Sigma},\phi),$$

at asymptotic probability level α*.*

Again, the test statistics $\widehat{\mathcal{W}}^{(n)}_{\mathcal{N}}$ do not depend on the particular choice of the fundamental system $\{\boldsymbol{\Psi}^{(1)}_t(\hat{\boldsymbol{\theta}}), \ldots, \boldsymbol{\Psi}^{(p_0+q_0)}_t(\hat{\boldsymbol{\theta}})\}$, and, for given values of p_0 and q_0, depend on p_1 and q_1 through $\pi = \max(p_1 - p_0, q_1 - q_0)$ only.

The proof of Proposition 4 follows along the same lines as for Proposition 3. The key ingredient is again an asymptotic linearity result, which, in this parametric Gaussian context, takes the following form (the proof of Lemma 3 readily extends to this situation).

Lemma 4 *Assume that (A1), (B1′), (B2), (B3) and (D1) hold, and partition* $\boldsymbol{\theta}$ *into* $(\boldsymbol{\theta}'_I, \boldsymbol{\theta}'_{II})' \in \mathbb{R}^{km} \times \mathbb{R}^{k^2(p_1+q_1)}$. *Let* $\hat{\boldsymbol{\theta}}^{(n)}$ *and* $\boldsymbol{\theta}$ *denote a sequence of estimators and a parameter value such that, under* $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},f}$, *as* $n \to \infty$, $\hat{\boldsymbol{\theta}}^{(n)}$ *satisfies at* $\boldsymbol{\theta}$ *the root-*n *consistency and local asymptotic discreteness properties given in Assumptions (E1) (ii) and (iii). Then,*

$$n^{1/2}(\mathbf{T}^{(n)}_{I;\mathbf{S},\phi}(\hat{\boldsymbol{\theta}}) - \mathbf{T}^{(n)}_{I;\mathbf{S},\phi}(\boldsymbol{\theta})) + \frac{k}{D_k(f)}\,\mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,(\mathbf{K}^{(n)} \otimes \mathbf{I}_k)^{-1}\,n^{1/2}(\hat{\boldsymbol{\theta}}^{(n)}_I - \boldsymbol{\theta}_I)$$

and

$$n^{1/2}(\mathbf{T}^{(n)}_{II;\mathbf{S},\phi}(\hat{\boldsymbol{\theta}}) - \mathbf{T}^{(n)}_{II;\mathbf{S},\phi}(\boldsymbol{\theta})) + \mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,\mathbf{P}_{\boldsymbol{\theta}}\,\mathbf{M}_{\boldsymbol{\theta}}\,n^{1/2}(\hat{\boldsymbol{\theta}}^{(n)}_{II} - \boldsymbol{\theta}_{II})$$

are $o_{\mathrm{P}}(1)$ *as* $n \to \infty$, *under* $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$.

5.3 Asymptotic relative efficiencies.

We finally turn to asymptotic relative efficiencies of the rank-based tests $\phi_J^{(n)}$ with respect to their Gaussian counterparts $\phi_{\mathcal{N}}^{(n)}$. The ARE values in the following proposition directly follow as the ratios of the noncentrality parameters in the asymptotic distributions of the various test statistics under local alternatives (see Propositions 3 and 4).

Proposition 5 *Assume that (A1′), (A2), (B1′), (B2), (B3), (C), (D1), (D2), (E1) and (E2) hold. Then, the asymptotic relative efficiency of $\phi_J^{(n)}$ with respect to the Gaussian test $\phi_{\mathcal{N}}^{(n)}$, under radial density f, is*

$$\mathrm{ARE}_{k,f}(\phi_J^{(n)}/\phi_{\mathcal{N}}^{(n)}) = (1-\lambda_{\boldsymbol{\theta},\boldsymbol{\Sigma},f}(\boldsymbol{\tau}))\,\frac{1}{k^2}\,D_k(f)\,\frac{C_k^2(J_0;f)}{\mathrm{E}[J_0^2(U)]} + \lambda_{\boldsymbol{\theta},\boldsymbol{\Sigma},f}(\boldsymbol{\tau})\,\frac{1}{k^2}\,\frac{D_k^2(J_2;f)}{\mathrm{E}[J_1^2(U)]}\,\frac{C_k^2(J_1;f)}{\mathrm{E}[J_2^2(U)]},$$

where $\lambda_{\boldsymbol{\theta},\boldsymbol{\Sigma},f}(\boldsymbol{\tau}) := (D_k(f)s_{\boldsymbol{\theta},\boldsymbol{\Sigma}}(\boldsymbol{\gamma},\boldsymbol{\delta}))/(k\,r_{\boldsymbol{\theta},\boldsymbol{\Sigma}}(\boldsymbol{\eta}) + D_k(f)\,s_{\boldsymbol{\theta},\boldsymbol{\Sigma}}(\boldsymbol{\gamma},\boldsymbol{\delta})) \in [0,1]$.

Denoting by $\mathrm{ARE}_{k,f}^{(\mathrm{loc})}(\phi_J^{(n)}/\phi_{\mathcal{N}}^{(n)})$ and $\mathrm{ARE}_{k,f}^{(\mathrm{ser})}(\phi_J^{(n)}/\phi_{\mathcal{N}}^{(n)})$ the AREs achieved in the pure location and purely serial problems (see Hallin and Paindveine (2002a and b)), respectively, we have

$$\mathrm{ARE}_{k,f}(\phi_J^{(n)}/\phi_{\mathcal{N}}^{(n)}) = (1-\lambda_{\boldsymbol{\theta},\boldsymbol{\Sigma},f}(\boldsymbol{\tau}))\,\mathrm{ARE}_{k,f}^{(\mathrm{loc})}(\phi_J^{(n)}/\phi_{\mathcal{N}}^{(n)}) + \lambda_{\boldsymbol{\theta},\boldsymbol{\Sigma},f}(\boldsymbol{\tau})\,\mathrm{ARE}_{k,f}^{(\mathrm{ser})}(\phi_J^{(n)}/\phi_{\mathcal{N}}^{(n)}).$$

Thus, the asymptotic relative efficiencies of the proposed procedures with respect to the parametric Gaussian procedure are convex linear combinations of the corresponding asymptotic relative efficiencies in the pure location and purely serial models (see Hallin and Paindaveine 2002a and 2002b, respectively). As a corollary, the generalized Chernoff-Savage results obtained in Hallin and Paindaveine (2002a and b) still hold here: the asymptotic relative efficiencies of our procedures, when van der Waerden scores ($J_0(u) = J_1(u) = J_2(u) = (\Psi_k^{-1}(u))^{1/2}$, where Ψ_k stands for the chi-square distribution function with k degrees of freedom) are used are always larger than or equal to one with respect to the Gaussian procedure, irrespective of the radial density f and the dimension k of the observation space. For the same reason, the generalized serial version (we refer to Proposition 7 of Hallin and Paindaveine 2002b for details) of the celebrated Hodges-Lehmann result also holds here.

6 Examples.

6.1 A multivariate Durbin-Watson test.

The generalized Durbin-Watson testing problem corresponds to $\boldsymbol{\theta}_0 = \mathbf{0}$, $\boldsymbol{\Upsilon}_I = \mathbf{I}_{km}$, and $\boldsymbol{\Upsilon}_{II} := \emptyset$. Here $\pi = \max(p_1, q_1)$. One easily checks that $\widehat{\mathcal{W}}_{I;J}^{(n)} = 0$, $n^{1/2}\,\underset{\sim}{\mathbf{T}}_{II;J}^{(n)}(\boldsymbol{\theta}) = \underset{\sim}{\mathbf{S}}_{II;\pi+1}^{(n)}(\boldsymbol{\theta})$, and $\mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}}^{(n)} = \mathbf{I}_\pi \otimes (\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}^{-1})$, so that

$$\widehat{\mathcal{W}}_J^{(n)} = \widehat{\mathcal{W}}_{II;J}^{(n)} = \frac{k^2}{\mathrm{E}[J_1^2(U)]\mathrm{E}[J_2^2(U)]}\sum_{i=1}^{\pi}(n-i)^{-1} \times\sum_{s,t=i+1}^{n} J_1\Big(\frac{\hat{R}_s(\hat{\boldsymbol{\beta}})}{n+1}\Big)J_1\Big(\frac{\hat{R}_t(\hat{\boldsymbol{\beta}})}{n+1}\Big)J_2\Big(\frac{\hat{R}_{s-i}(\hat{\boldsymbol{\beta}})}{n+1}\Big)J_2\Big(\frac{\hat{R}_{t-i}(\hat{\boldsymbol{\beta}})}{n+1}\Big)\;\mathbf{W}'_{s-i}(\hat{\boldsymbol{\beta}})\mathbf{W}_{t-i}(\hat{\boldsymbol{\beta}})\;\mathbf{W}'_s(\hat{\boldsymbol{\beta}})\mathbf{W}_t(\hat{\boldsymbol{\beta}}) \qquad (28)$$

(if there is no trend part in the model, the test statistic (28) is the Mahalanobis version of the test statistic based on pseudo-Mahalanobis ranks and interdirections proposed in Hallin and Paindaveine (2002b) for the problem of testing for serial randomness). The resulting Durbin-Watson test consists (at asymptotic level α) in rejecting the null hypothesis of independent noise as soon as $\widehat{\mathcal{W}}_J^{(n)}$ exceeds the α-upper quantile of a chi-square distribution with $k^2\pi$ degrees of freedom. One could also obtain purely hyperplane-based Durbin-Watson tests (that are *strictly* affine-invariant in this case) by replacing the pseudo-Mahalanobis ranks $\hat{R}_t(\hat{\boldsymbol{\beta}})$ and the pseudo-Mahalanobis angles $\mathbf{W}_s'(\hat{\boldsymbol{\beta}})\mathbf{W}_t(\hat{\boldsymbol{\beta}})$ by lift-interdirection ranks $\underline{R}_t(\hat{\boldsymbol{\beta}})$ and the cosines based on Randles' interdirections $q_{st}(\hat{\boldsymbol{\beta}})$, respectively.

6.2 Testing the order of a VAR model.

For the problem of testing VAR(p_0) against VAR(p_0+1) dependence, the proposed tests consist (at asymptotic level α) in rejecting the null hypothesis as soon as

$$\widehat{\mathcal{W}}_J^{(n)} = \widehat{\mathcal{W}}_{II;J}^{(n)} = n \left(\underset{\sim}{\mathbf{T}}_{II;J}^{(n)}(\hat{\boldsymbol{\theta}}) \right)' \bar{\mathbf{Q}}_{II;J;\widehat{\boldsymbol{\Sigma}}}^{(n)}(\hat{\boldsymbol{\theta}})\, \underset{\sim}{\mathbf{T}}_{II;J}^{(n)}(\hat{\boldsymbol{\theta}}) \tag{29}$$

exceeds the α-upper quantile of a chi-square distribution with k^2 degrees of freedom, where, letting $\mathbf{v}^{(n)}(\boldsymbol{\theta}) := (\mathbf{v}_1^{(n)\prime}(\boldsymbol{\theta}), \ldots, \mathbf{v}_{p_0}^{(n)\prime}(\boldsymbol{\theta}))'$ and $\mathbf{v}_i^{(n)}(\boldsymbol{\theta}) := \sum_{t=\max(i,2)}^{n-1} (n-t)^{1/2} (\mathbf{G}_{t-i}(\boldsymbol{\theta}) \otimes \mathbf{I}_k)$ $(\mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}_{t;J}^{(n)}(\boldsymbol{\theta}))$,

$$n^{1/2}\, \underset{\sim}{\mathbf{T}}_{II;J}^{(n)}(\boldsymbol{\theta}) := \begin{pmatrix} (n-1)^{1/2}\,(\mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}_{1;J}^{(n)}(\boldsymbol{\theta})) \\ \mathbf{v}^{(n)}(\boldsymbol{\theta}) \end{pmatrix},$$

and

$$\bar{\mathbf{Q}}_{II;J;\boldsymbol{\Sigma}}^{(n)}(\boldsymbol{\theta}) := \frac{k^2}{\mathrm{E}[J_1^2(U)]\mathrm{E}[J_2^2(U)]} \left[\begin{pmatrix} \boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{w}^2 \end{pmatrix}^{-1} - \begin{pmatrix} \mathbf{I}_{k^2} & \mathbf{0} \\ & \mathbf{I}_{k^2 p_0} \end{pmatrix} \mathbf{W}^{-2} \begin{pmatrix} \mathbf{I}_{k^2} & \mathbf{0} \\ & \mathbf{I}_{k^2 p_0} \end{pmatrix}' \right].$$

Above, $\mathbf{w}^2$ and $\mathbf{W}^2$ stand for the $k^2p_0 \times k^2p_0$ arrays with blocks $\sum_{t=\max(i,j,2)}^{n-1} \mathbf{G}_{t-i}(\boldsymbol{\theta})\boldsymbol{\Sigma}\mathbf{G}_{t-j}'(\boldsymbol{\theta}) \otimes \boldsymbol{\Sigma}^{-1}$ and $\sum_{t=\max(i,j)}^{n-1} \mathbf{G}_{t-i}(\boldsymbol{\theta})\boldsymbol{\Sigma}\mathbf{G}_{t-j}'(\boldsymbol{\theta}) \otimes \boldsymbol{\Sigma}^{-1}$, respectively, in position (i,j) $(i,j = 1, \ldots, p_0)$. Note that $\mathbf{W}^2 = \mathbf{w}^2 + \mathbf{e}_1^{(p_0)}\mathbf{e}_1^{(p_0)\prime} \otimes (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}^{-1})$ only differs from $\mathbf{w}^2$ through the block in position $(1,1)$.

The test statistic (29) has the same algebraic structure as in the univariate case (see Hallin and Puri (1994), or Garel and Hallin (1999)). However, it should be pointed out that the test statistic associated with the problem of testing MA(q_0) dependence versus MA(q_0+1) dependence is much more complex here than in the univariate case. This is due to the presence of the factors $\mathbf{H}_{n-1}^{(r)}$ and $\mathbf{B}_{n-1}^{\prime(l)}$ in $\mathbf{Q}_{\boldsymbol{\theta}}^{(n)}$ which cancel each other in the univariate case *only*. In the multivariate case, they do not, yielding in $n^{1/2}\, \underset{\sim}{\mathbf{T}}_{II;J}^{(n)}(\boldsymbol{\theta})$ quite intricate linear combinations of the cross-covariance matrices $\underset{\sim}{\boldsymbol{\Gamma}}_{t;J}^{(n)}(\boldsymbol{\theta})$.

6.3 Detecting switching location regimes.

We finally consider the problem of detecting the presence of different "location regimes" in a VAR(1) series with a time-dependent trend (with mean $\boldsymbol{\beta}_i$ for $t \in C_i = C_i^{(n)} := \{t_{i-1}^{(n)} +$

$1, \ldots, t_i^{(n)}\}$). More precisely, the null hypothesis $\mathcal{H}_0 : \boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_m$ we are considering here is associated with $\boldsymbol{\Upsilon}_I = (1, \ldots, 1)' \otimes \mathbf{I}_k$, $\boldsymbol{\Upsilon}_{II} = \mathbf{I}_{k^2}$. Letting $\boldsymbol{\lambda}^{(n)} := ((\lambda_1^{(n)})^{1/2}, \ldots, (\lambda_m^{(n)})^{1/2})'$, with $\lambda_j^{(n)} := n_j/n := (t_j^{(n)} - t_{j-1}^{(n)})/n$, the test statistic is

$$\widehat{\mathcal{W}}_J^{(n)} = \widehat{\mathcal{W}}_{I;J}^{(n)} = n \left(\underset{\sim}{\mathbf{T}}_{I;J}^{(n)}(\hat{\boldsymbol{\theta}}) \right)' \bar{\mathbf{Q}}_{I;J;\hat{\boldsymbol{\Sigma}}}^{(n)}(\hat{\boldsymbol{\theta}}) \, \underset{\sim}{\mathbf{T}}_{I;J}^{(n)}(\hat{\boldsymbol{\theta}}),$$

where, denoting by L the backshift operator and with the convention that $J_0\Big(\frac{\hat{R}_{n+1}(\boldsymbol{\theta})}{n+1}\Big)\hat{\boldsymbol{\Sigma}}^{-1/2'}$ $\mathbf{W}_{n+1}(\boldsymbol{\theta}) = \mathbf{0}$,

$$n^{1/2} \underset{\sim}{\mathbf{T}}_{I;J}^{(n)}(\boldsymbol{\theta}) = \left[\mathbf{I}_{km} - \sqrt{\frac{n}{n-1}} (\mathbf{I}_m \otimes \mathbf{A}') L^{-1} \right] \begin{pmatrix} \frac{1}{\sqrt{n_1}} \sum_{t \in C_1} J_0\Big(\frac{\hat{R}_t(\boldsymbol{\theta})}{n+1}\Big) \hat{\boldsymbol{\Sigma}}^{-1/2'} \mathbf{W}_t(\boldsymbol{\theta}) \\ \frac{1}{\sqrt{n_2}} \sum_{t \in C_2} J_0\Big(\frac{\hat{R}_t(\boldsymbol{\theta})}{n+1}\Big) \hat{\boldsymbol{\Sigma}}^{-1/2'} \mathbf{W}_t(\boldsymbol{\theta}) \\ \vdots \\ \frac{1}{\sqrt{n_m}} \sum_{t \in C_m} J_0\Big(\frac{\hat{R}_t(\boldsymbol{\theta})}{n+1}\Big) \hat{\boldsymbol{\Sigma}}^{-1/2'} \mathbf{W}_t(\boldsymbol{\theta}) \end{pmatrix},$$

and

$$\bar{\mathbf{Q}}_{I;J;\boldsymbol{\Sigma}}^{(n)}(\boldsymbol{\theta}) := \frac{k}{\mathrm{E}[J_0^2(U)]} \left[(\mathbf{I}_m - \boldsymbol{\lambda}^{(n)} \boldsymbol{\lambda}^{(n)'}) \otimes [\boldsymbol{\Sigma}^{-1} - \mathbf{A}'\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{A} + \mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A}]^{-1} \right].$$

If there is no serial part in the model (i.e., when the errors are independent white noise), the problem reduces to the m-sample location problem (classical MANOVA), and the test statistic takes the simpler form (just put $\mathbf{A} = \mathbf{0}$)

$$\widehat{\mathcal{W}}_J^{(n)} = \frac{k}{\mathrm{E}[J_0^2(U)]} \left[\sum_{j=1}^m \frac{1}{n_j} \sum_{i, \tilde{\imath} \in C_j} J_0\Big(\frac{\hat{R}_i(\hat{\boldsymbol{\beta}})}{n+1}\Big) J_0\Big(\frac{\hat{R}_{\tilde{\imath}}(\hat{\boldsymbol{\beta}})}{n+1}\Big) \mathbf{W}_i'(\hat{\boldsymbol{\beta}}) \mathbf{W}_{\tilde{\imath}}(\hat{\boldsymbol{\beta}}) \right.$$
$$\left. - \frac{1}{n} \sum_{j, \tilde{\jmath}=1}^m \sum_{i \in C_j} \sum_{\tilde{\imath} \in C_{\tilde{\jmath}}} J_0\Big(\frac{\hat{R}_i(\hat{\boldsymbol{\beta}})}{n+1}\Big) J_0\Big(\frac{\hat{R}_{\tilde{\imath}}(\hat{\boldsymbol{\beta}})}{n+1}\Big) \mathbf{W}_i'(\hat{\boldsymbol{\beta}}) \mathbf{W}_{\tilde{\imath}}(\hat{\boldsymbol{\beta}}) \right],$$

i.e., a purely pseudo-Mahalanobis version of Randles and Um (1998)'s test statistic. Again, a strictly affine-invariant purely hyperplane-based version of $\widehat{\mathcal{W}}_J^{(n)}$ can be obtained in the same way as for the Durbin-Watson tests, just by plugging in lift-interdirection ranks and Randles' interdirections.

7 Appendix

7.1 Proof of Lemma 3.

The proof of Lemma 3 is based on the following asymptotic linearity result for the individual nonserial and serial statistics $\underset{\sim}{\boldsymbol{\Lambda}}_{i;J}^{(n)}$ and $\underset{\sim}{\boldsymbol{\Gamma}}_{i;J}^{(n)}$ (see Hallin and Paindaveine 2004b).

Proposition 6 *Assume that (A1), (B1′), (B2), (B3), (C), and (D1) hold. Then,*

$$(n-i)^{1/2}\,\mathrm{vec}\,(\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\boldsymbol{\theta}+\boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)})-\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\boldsymbol{\theta}))+\frac{1}{k}\,C_k(J_0;f) \times(\mathbf{I}_m\otimes\mathbf{\Sigma}^{-1})\left(\sum_{j=0}^{\infty}(\mathbf{K}'\,\mathbf{R}_{|i-j|}\,\mathbf{K})\otimes\mathbf{h}_j\right)\left(\mathrm{vec}\,\boldsymbol{\eta}^{(n)'}\right)=o_{\mathrm{P}}(1) \tag{30}$$

and

$$(n-i)^{1/2}\,\mathrm{vec}\,(\underset{\sim}{\mathbf{\Gamma}}{}^{(n)}_{i;J}(\boldsymbol{\theta}+\boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)})-\underset{\sim}{\mathbf{\Gamma}}{}^{(n)}_{i;J}(\boldsymbol{\theta}))+\frac{1}{k^2}\,D_k(J_2;f)\,C_k(J_1;f) \times(\mathbf{\Sigma}\otimes\mathbf{\Sigma}^{-1})\left[\mathbf{a}_i(\boldsymbol{\tau}^{(n)};\boldsymbol{\theta})+\mathbf{b}_i(\boldsymbol{\tau}^{(n)};\boldsymbol{\theta})\right]=o_{\mathrm{P}}(1) \tag{31}$$

as $n\to\infty$, under $\mathcal{H}^{(n)}(\boldsymbol{\theta},\mathbf{\Sigma},f)$.

Proof of Lemma 3. Let us first prove the first statement in Lemma 3. Clearly,

$$\begin{aligned} n^{1/2}(\underset{\sim}{\mathbf{T}}{}^{(n)}_{I;J}(\hat{\boldsymbol{\theta}})-\underset{\sim}{\mathbf{T}}{}^{(n)}_{I;J}(\boldsymbol{\theta})) &= \mathbf{L}^{(n)'}_{\hat{\boldsymbol{\theta}}}\,\underset{\sim}{\mathbf{S}}{}^{(n)}_{I;J}(\hat{\boldsymbol{\theta}})-\mathbf{L}^{(n)'}_{\boldsymbol{\theta}}\,\underset{\sim}{\mathbf{S}}{}^{(n)}_{I;J}(\boldsymbol{\theta}) \\ &= \sum_{i=0}^{n-1}(n-i)^{1/2}\left[(\mathbf{I}_m\otimes\hat{\mathbf{h}}_i')\,\mathrm{vec}\,\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}})-(\mathbf{I}_m\otimes\mathbf{h}_i')\,\mathrm{vec}\,\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\boldsymbol{\theta})\right]. \end{aligned}$$

Now, for some fixed integer s (and $n>s+1$),

$$\begin{aligned} n^{1/2}(\underset{\sim}{\mathbf{T}}{}^{(n)}_{I;J}(\hat{\boldsymbol{\theta}})-\underset{\sim}{\mathbf{T}}{}^{(n)}_{I;J}(\boldsymbol{\theta})) &= \sum_{i=0}^{s}(n-i)^{1/2}\left[\left(\mathbf{I}_m\otimes(\hat{\mathbf{h}}_i'-\mathbf{h}_i')\right)\,\mathrm{vec}\,\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}})\right] \\ &+\sum_{i=0}^{s}\left[(\mathbf{I}_m\otimes\mathbf{h}_i')\left((n-i)^{1/2}\mathrm{vec}\,(\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}})-\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\boldsymbol{\theta}))\right)\right] \\ &+\sum_{i=s+1}^{n-1}(n-i)^{1/2}\left[(\mathbf{I}_m\otimes\hat{\mathbf{h}}_i')\,\mathrm{vec}\,\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}})-(\mathbf{I}_m\otimes\mathbf{h}_i')\,\mathrm{vec}\,\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\boldsymbol{\theta})\right]. \end{aligned} \tag{32}$$

Next, the local discreteness of $\hat{\boldsymbol{\theta}}^{(n)}$ (see Assumption (D1)(iii)) allows to replace $\boldsymbol{\theta}^{(n)}=\boldsymbol{\theta}+\boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}$ with $\hat{\boldsymbol{\theta}}^{(n)}$ in (30) (see Kreiss 1987, Lemma 4.4). Since $\boldsymbol{\beta}^{(n)}=\boldsymbol{\beta}+n^{-1/2}\mathbf{K}^{(n)}\boldsymbol{\eta}^{(n)}$ can be written under the form $n^{1/2}\mathrm{vec}\,(\boldsymbol{\beta}^{(n)'}-\boldsymbol{\beta}')=(\mathbf{K}^{(n)}\otimes\mathbf{I}_k)(\mathrm{vec}\,\boldsymbol{\eta}^{(n)'})$, this yields

$$(n-i)^{1/2}\,\mathrm{vec}\left(\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}}^{(n)})-\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\boldsymbol{\theta})\right)=-\frac{1}{k}\,C_k(J_0;f) (\mathbf{I}_m\otimes\mathbf{\Sigma}^{-1})\left(\sum_{j=0}^{\infty}(\mathbf{K}'\,\mathbf{R}_{|i-j|}\,\mathbf{K})\otimes\mathbf{h}_j\right)(\mathbf{K}^{(n)}\otimes\mathbf{I}_k)^{-1}n^{1/2}(\hat{\boldsymbol{\theta}}^{(n)}_I-\boldsymbol{\theta}_I)+\mathbf{R}^{(n)}_i, \tag{33}$$

where $\mathbf{R}^{(n)}_i$ is $o_{\mathrm{P}}(1)$ as $n\to\infty$, under $\mathcal{H}^{(n)}(\boldsymbol{\theta},\mathbf{\Sigma},f)$. Substituting in (32), we obtain

$$\begin{aligned} n^{1/2}(\underset{\sim}{\mathbf{T}}{}^{(n)}_{I;J}(\hat{\boldsymbol{\theta}})-\underset{\sim}{\mathbf{T}}{}^{(n)}_{I;J}(\boldsymbol{\theta})) &= \sum_{i=0}^{s}(n-i)^{1/2}\left[\left(\mathbf{I}_m\otimes(\hat{\mathbf{h}}_i-\mathbf{h}_i)'\right)\,\mathrm{vec}\,\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}})\right] \\ -\frac{1}{k}\,C_k(J_0;f)\sum_{i=0}^{s}&\left[(\mathbf{I}_m\otimes\mathbf{h}_i'\mathbf{\Sigma}^{-1})\left(\sum_{j=0}^{\infty}(\mathbf{K}'\,\mathbf{R}_{|i-j|}\,\mathbf{K})\otimes\mathbf{h}_j\right)(\mathbf{K}^{(n)}\otimes\mathbf{I}_k)^{-1}n^{1/2}(\hat{\boldsymbol{\theta}}^{(n)}_I-\boldsymbol{\theta}_I)\right] \\ &+\sum_{i=s+1}^{n-1}(n-i)^{1/2}\left[(\mathbf{I}_m\otimes\hat{\mathbf{h}}_i')\,\mathrm{vec}\,\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}})-(\mathbf{I}_m\otimes\mathbf{h}_i')\,\mathrm{vec}\,\underset{\sim}{\mathbf{\Lambda}}{}^{(n)}_{i;J}(\boldsymbol{\theta})\right]+\sum_{i=0}^{s}\mathbf{R}^{(n)}_i. \end{aligned}$$

This finally yields the decomposition

$$n^{1/2}(\underset{\sim}{\mathbf{T}}^{(n)}_{I;J}(\hat{\boldsymbol{\theta}}) - \underset{\sim}{\mathbf{T}}^{(n)}_{I;J}(\boldsymbol{\theta})) + \frac{1}{k}\, C_k(J_0; f)$$

$$\times\Big[\sum_{i,j=0}^{\infty} (\mathbf{I}_m \otimes \mathbf{h}_i^{'}) \left((\mathbf{K}^{'}\, \mathbf{R}_{|i-j|}\, \mathbf{K}) \otimes \boldsymbol{\Sigma}^{-1}\right) (\mathbf{I}_m \otimes \mathbf{h}_j)\Big](\mathbf{K}^{(n)} \otimes \mathbf{I}_k)^{-1} n^{1/2}(\hat{\boldsymbol{\theta}}_I^{(n)} - \boldsymbol{\theta}_I)$$

$$= \mathbf{T}_1^{(n,s)} + \mathbf{T}_2^{(n,s)},$$

say, where

$$\mathbf{T}_1^{(n,s)} := \sum_{i=0}^{s} (n-i)^{1/2} \left[\left(\mathbf{I}_m \otimes (\hat{\mathbf{h}}_i - \mathbf{h}_i)^{'}\right) \text{ vec } \underset{\sim}{\boldsymbol{\Lambda}}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}})\right] + \sum_{i=0}^{s} \mathbf{R}_i^{(n)},$$

and

$$\mathbf{T}_2^{(n,s)} := \sum_{i=s+1}^{n-1} (n-i)^{1/2} \left[(\mathbf{I}_m \otimes \hat{\mathbf{h}}_i^{'}) \text{ vec } \underset{\sim}{\boldsymbol{\Lambda}}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}}) - (\mathbf{I}_m \otimes \mathbf{h}_i^{'}) \text{ vec } \underset{\sim}{\boldsymbol{\Lambda}}^{(n)}_{i;J}(\boldsymbol{\theta})\right]$$

$$+\frac{1}{k}\, C_k(J_0; f) \sum_{i=s+1}^{\infty} \left[(\mathbf{I}_m \otimes \mathbf{h}_i^{'} \boldsymbol{\Sigma}^{-1}) \left(\sum_{j=0}^{\infty} (\mathbf{K}^{'}\, \mathbf{R}_{|i-j|}\, \mathbf{K}) \otimes \mathbf{h}_j\right)\right] (\mathbf{K}^{(n)} \otimes \mathbf{I}_k)^{-1} n^{1/2}(\hat{\boldsymbol{\theta}}_I^{(n)} - \boldsymbol{\theta}_I).$$

Since $\sum_{i,j=0}^{\infty} (\mathbf{I}_m \otimes \mathbf{h}_i^{'}) \left((\mathbf{K}^{'}\, \mathbf{R}_{|i-j|}\, \mathbf{K}) \otimes \boldsymbol{\Sigma}^{-1}\right) (\mathbf{I}_m \otimes \mathbf{h}_j) = \mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}$, the first statement in Lemma 3 takes the form

$$n^{1/2}(\underset{\sim}{\mathbf{T}}^{(n)}_{I;J}(\hat{\boldsymbol{\theta}}) - \underset{\sim}{\mathbf{T}}^{(n)}_{I;J}(\boldsymbol{\theta})) + \frac{1}{k}\, C_k(J_0; f)$$

$$\Big[\sum_{i,j=0}^{\infty} (\mathbf{I}_m \otimes \mathbf{h}_i^{'}) \left((\mathbf{K}^{'}\, \mathbf{R}_{|i-j|}\, \mathbf{K}) \otimes \boldsymbol{\Sigma}^{-1}\right) (\mathbf{I}_m \otimes \mathbf{h}_j)\Big](\mathbf{K}^{(n)} \otimes \mathbf{I}_k)^{-1} n^{1/2}(\hat{\boldsymbol{\theta}}_I^{(n)} - \boldsymbol{\theta}_I) = o_{\mathrm{P}}(1)$$

as $n \to \infty$ under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$. Now, it follows, from the continuity of $\boldsymbol{\theta} \mapsto \mathbf{h}_i(\boldsymbol{\theta})$ and the boundedness (in probability, under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$; see (33)) of $(n-i)^{1/2}\text{vec } \underset{\sim}{\boldsymbol{\Lambda}}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}})$, that $\mathbf{T}_1^{(n,s)}$ is $o_{\mathrm{P}}(1)$ under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$, for any fixed s, as $n \to \infty$. On the other hand, the exponential decrease in i of the $\mathbf{h}_i$'s and the root-n consistency of $\hat{\boldsymbol{\theta}}$ imply that $\mathbf{T}_2^{(n,s)}$ is $o_{\mathrm{P}}(1)$ under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$, as $s \to \infty$, uniformly in n.

Now, $\mathrm{P}\left[\|\mathbf{T}_1^{(n,s)} + \mathbf{T}_2^{(n,s)}\| > \delta\right] \leq \mathrm{P}\left[\|\mathbf{T}_1^{(n,s)}\| > \delta/2\right] + \mathrm{P}\left[\|\mathbf{T}_2^{(n,s)}\| > \delta/2\right]$, for all s and n. For any $\varepsilon > 0$, one can always choose $s = S$ sufficiently large so that $\mathrm{P}\left[\|\mathbf{T}_2^{(n,S)}\| > \delta/2\right] < \varepsilon$ uniformly in n. Since $\mathbf{T}_1^{(n,S)}$ is $o_{\mathrm{P}}(1)$ as $n \to \infty$, it is possible to find a integer $N = N(\varepsilon)$ such that $\mathrm{P}\left[\|\mathbf{T}_1^{(n,S)}\| > \delta/2\right] < \varepsilon$ for all $n \geq N$. Consequently, for all $\varepsilon > 0$, $N = N(\varepsilon)$ is such that

$$\mathrm{P}\Bigg[\Big\|n^{1/2}(\underset{\sim}{\mathbf{T}}^{(n)}_{I;J}(\hat{\boldsymbol{\theta}}) - \underset{\sim}{\mathbf{T}}^{(n)}_{I;J}(\boldsymbol{\theta})) + \frac{1}{k}\, C_k(J_0; f)$$

$$\Big[\sum_{i,j=0}^{\infty} (\mathbf{I}_m \otimes \mathbf{h}_i^{'}) \left((\mathbf{K}^{'}\, \mathbf{R}_{|i-j|}\, \mathbf{K}) \otimes \boldsymbol{\Sigma}^{-1}\right) (\mathbf{I}_m \otimes \mathbf{h}_j)\Big](\mathbf{K}^{(n)} \otimes \mathbf{I}_k)^{-1} n^{1/2}(\hat{\boldsymbol{\theta}}_I^{(n)} - \boldsymbol{\theta}_I)\Big\| > \delta\Bigg] < 2\varepsilon$$

for all $n \geq N$. The result follows.

Turning to the proof of the serial part of the lemma, denote by $\mathbf{Q}_{i,j} = \mathbf{Q}_{i,j}^{(n)}$ (resp., $\hat{\mathbf{Q}}_{i,j} = \hat{\mathbf{Q}}_{i,j}^{(n)}$) the $k^2 \times k^2$ block in position (i,j) ($i = 1, \ldots, n-1$, $j = 1, \ldots, \pi_0$) in $\mathbf{Q}_{\boldsymbol{\theta}}^{(n)}$ (resp., in $\mathbf{Q}_{\hat{\boldsymbol{\theta}}}^{(n)}$). Then,

$$n^{1/2}(\underset{\sim}{\mathbf{T}}{}_{II;J}^{(n)}(\hat{\boldsymbol{\theta}}) - \underset{\sim}{\mathbf{T}}{}_{II;J}^{(n)}(\boldsymbol{\theta})) = \mathbf{Q}_{\hat{\boldsymbol{\theta}}}^{(n)\prime} \underset{\sim}{\mathbf{S}}{}_{II;J}^{(n)}(\hat{\boldsymbol{\theta}}) - \mathbf{Q}_{\boldsymbol{\theta}}^{(n)\prime} \underset{\sim}{\mathbf{S}}{}_{II;J}^{(n)}(\boldsymbol{\theta})$$

$$= \sum_{i=1}^{n-1} (n-i)^{1/2} \left[\begin{pmatrix} \hat{\mathbf{Q}}_{i,1}' \\ \vdots \\ \hat{\mathbf{Q}}_{i,\pi_0}' \end{pmatrix} \mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}{}_{i;J}^{(n)}(\hat{\boldsymbol{\theta}}) - \begin{pmatrix} \mathbf{Q}_{i,1}' \\ \vdots \\ \mathbf{Q}_{i,\pi_0}' \end{pmatrix} \mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}{}_{i;J}^{(n)}(\boldsymbol{\theta}) \right].$$

The same decomposition as for the trend part then yields, for some fixed integer s (and still for $n > s+1$),

$$n^{1/2}(\underset{\sim}{\mathbf{T}}{}_{II;J}^{(n)}(\hat{\boldsymbol{\theta}}) - \underset{\sim}{\mathbf{T}}{}_{II;J}^{(n)}(\boldsymbol{\theta})) = \sum_{i=1}^{s} (n-i)^{1/2} \left[\begin{pmatrix} \hat{\mathbf{Q}}_{i,1}' \\ \vdots \\ \hat{\mathbf{Q}}_{i,\pi_0}' \end{pmatrix} - \begin{pmatrix} \mathbf{Q}_{i,1}' \\ \vdots \\ \mathbf{Q}_{i,\pi_0}' \end{pmatrix} \right] \mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}{}_{i;J}^{(n)}(\hat{\boldsymbol{\theta}}) \tag{34}$$

$$+ \sum_{i=1}^{s} \left[\begin{pmatrix} \mathbf{Q}_{i,1}' \\ \vdots \\ \mathbf{Q}_{i,\pi_0}' \end{pmatrix} \left((n-i)^{1/2} \mathrm{vec}\, (\underset{\sim}{\boldsymbol{\Gamma}}{}_{i;J}^{(n)}(\hat{\boldsymbol{\theta}}) - \underset{\sim}{\boldsymbol{\Gamma}}{}_{i;J}^{(n)}(\boldsymbol{\theta})) \right) \right]$$

$$+ \sum_{i=s+1}^{n-1} (n-i)^{1/2} \left[\begin{pmatrix} \hat{\mathbf{Q}}_{i,1}' \\ \vdots \\ \hat{\mathbf{Q}}_{i,\pi_0}' \end{pmatrix} \mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}{}_{i;J}^{(n)}(\hat{\boldsymbol{\theta}}) - \begin{pmatrix} \mathbf{Q}_{i,1}' \\ \vdots \\ \mathbf{Q}_{i,\pi_0}' \end{pmatrix} \mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}{}_{i;J}^{(n)}(\boldsymbol{\theta}) \right].$$

Again, the local discreteness of $\hat{\boldsymbol{\theta}}^{(n)}$ and (31) yield

$$(n-i)^{1/2}\, \mathrm{vec}\, \left(\underset{\sim}{\boldsymbol{\Gamma}}{}_{i;J}^{(n)}(\hat{\boldsymbol{\theta}}^{(n)}) - \underset{\sim}{\boldsymbol{\Gamma}}{}_{i;J}^{(n)}(\boldsymbol{\theta}) \right) \tag{35}$$

$$= -\frac{1}{k^2} D_k(J_2; f)\, C_k(J_1; f)\, (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}^{-1}) \left[\mathbf{a}_i(n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}); \boldsymbol{\theta}) + \mathbf{b}_i(n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}); \boldsymbol{\theta}) \right] + \mathbf{R}_i^{(n)}$$

$$= -\frac{1}{k^2} D_k(J_2; f)\, C_k(J_1; f)\, (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}^{-1})\, (\mathbf{Q}_{i,1} \ldots \mathbf{Q}_{i,\pi_0})\, \mathbf{P}_{\boldsymbol{\theta}}\, \mathbf{M}_{\boldsymbol{\theta}}\, n^{1/2}(\hat{\boldsymbol{\theta}}_{II} - \boldsymbol{\theta}_{II}) + \mathbf{R}_i^{(n)},$$

where $\mathbf{R}_i^{(n)}$ is $o_{\mathrm{P}}(1)$ (as $n \to \infty$, under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$). Hence, (34) becomes

$$n^{1/2}(\underset{\sim}{\mathbf{T}}^{(n)}_{II;J}(\hat{\boldsymbol{\theta}}) - \underset{\sim}{\mathbf{T}}^{(n)}_{II;J}(\boldsymbol{\theta})) = \sum_{i=1}^{s} (n-i)^{1/2} \left[\begin{pmatrix} \hat{\mathbf{Q}}'_{i,1} \\ \vdots \\ \hat{\mathbf{Q}}'_{i,\pi_0} \end{pmatrix} - \begin{pmatrix} \mathbf{Q}'_{i,1} \\ \vdots \\ \mathbf{Q}'_{i,\pi_0} \end{pmatrix} \right] \mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}})$$

$$-\frac{1}{k^2} D_k(J_2; f)\, C_k(J_1; f) \sum_{i=1}^{s} \left[\begin{pmatrix} \mathbf{Q}'_{i,1} \\ \vdots \\ \mathbf{Q}'_{i,\pi_0} \end{pmatrix} (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}^{-1})\, (\mathbf{Q}_{i,1} \ldots \mathbf{Q}_{i,\pi_0}) \right] \mathbf{P}_{\boldsymbol{\theta}}\, \mathbf{M}_{\boldsymbol{\theta}}\, n^{1/2}(\hat{\boldsymbol{\theta}}^{(n)}_{II} - \boldsymbol{\theta}_{II})$$

$$+ \sum_{i=s+1}^{n-1} (n-i)^{1/2} \left[\begin{pmatrix} \hat{\mathbf{Q}}'_{i,1} \\ \vdots \\ \hat{\mathbf{Q}}'_{i,\pi_0} \end{pmatrix} \mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}}) - \begin{pmatrix} \mathbf{Q}'_{i,1} \\ \vdots \\ \mathbf{Q}'_{i,\pi_0} \end{pmatrix} \mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}^{(n)}_{i;J}(\boldsymbol{\theta}) \right] + \sum_{i=1}^{s} \mathbf{R}_i^{(n)}.$$

Noting that

$$\sum_{i=1}^{s} \left[\begin{pmatrix} \mathbf{Q}'_{i,1} \\ \vdots \\ \mathbf{Q}'_{i,\pi_0} \end{pmatrix} (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}^{-1})\ (\mathbf{Q}_{i,1} \ldots \mathbf{Q}_{i,\pi_0}) \right] = \mathbf{Q}_{\boldsymbol{\theta}}^{(s+1)\prime} \Big[\mathbf{I}_s \otimes (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}^{-1}) \Big]\, \mathbf{Q}_{\boldsymbol{\theta}}^{(s+1)},$$

we finally decompose

$$n^{1/2}(\underset{\sim}{\mathbf{T}}^{(n)}_{II;J}(\hat{\boldsymbol{\theta}}) - \underset{\sim}{\mathbf{T}}^{(n)}_{II;J}(\boldsymbol{\theta})) + \frac{1}{k^2} D_k(J_2; f)\, C_k(J_1; f)\, \mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}}\, \mathbf{P}_{\boldsymbol{\theta}}\, \mathbf{M}_{\boldsymbol{\theta}}\, n^{1/2}(\hat{\boldsymbol{\theta}}^{(n)}_{II} - \boldsymbol{\theta}_{II})$$

into $\mathbf{T}_1^{(n,s)} + \mathbf{T}_2^{(n,s)}$, where

$$\mathbf{T}_1^{(n,s)} := \sum_{i=1}^{s} (n-i)^{1/2} \left[\begin{pmatrix} \hat{\mathbf{Q}}'_{i,1} \\ \vdots \\ \hat{\mathbf{Q}}'_{i,\pi_0} \end{pmatrix} - \begin{pmatrix} \mathbf{Q}'_{i,1} \\ \vdots \\ \mathbf{Q}'_{i,\pi_0} \end{pmatrix} \right] \mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}}) + \sum_{i=1}^{s} \mathbf{R}_i^{(n)}$$

and

$$\mathbf{T}_2^{(n,s)} := \frac{1}{k^2} D_k(J_2; f)\, C_k(J_1; f)\ \Big[\mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}} - \mathbf{Q}_{\boldsymbol{\theta}}^{(s+1)\prime} [\mathbf{I}_s \otimes (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}^{-1})]\, \mathbf{Q}_{\boldsymbol{\theta}}^{(s+1)} \Big]\ \mathbf{P}_{\boldsymbol{\theta}}\, \mathbf{M}_{\boldsymbol{\theta}}$$

$$n^{1/2}(\hat{\boldsymbol{\theta}}^{(n)}_{II} - \boldsymbol{\theta}_{II}) + \sum_{i=s+1}^{n-1} (n-i)^{1/2} \left[\begin{pmatrix} \hat{\mathbf{Q}}'_{i,1} \\ \vdots \\ \hat{\mathbf{Q}}'_{i,\pi_0} \end{pmatrix} \mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}}) - \begin{pmatrix} \mathbf{Q}'_{i,1} \\ \vdots \\ \mathbf{Q}'_{i,\pi_0} \end{pmatrix} \mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}^{(n)}_{i;J}(\boldsymbol{\theta}) \right].$$

As for the trend part, the continuity in $\boldsymbol{\theta}$ of the Green's matrices, the fact that $(n-i)^{1/2}\mathrm{vec}\, \underset{\sim}{\boldsymbol{\Gamma}}^{(n)}_{i;J}(\hat{\boldsymbol{\theta}})$ is $O_{\mathrm{P}}(1)$ (as $n \to \infty$, under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$), and the root-$n$ consistency of $\hat{\boldsymbol{\theta}}$, entail that $\mathbf{T}_1^{(n,s)}$ and $\mathbf{T}_2^{(n,s)}$ are $o_{\mathrm{P}}(1)$ under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$, for fixed s, as $n \to \infty$, and uniformly in n, as $s \to \infty$, respectively. The result follows. □

7.2 Proof of Proposition 3.

(i) We first prove that $\widehat{\mathcal{W}}_J^{(n)}$ is affine-invariant. Clearly, the scalar factor $d^{-1/2}$ in the equivariance relation (25) has no influence on the affine-invariance of $\widehat{\mathcal{W}}_J^{(n)}$; consequently, we can assume, without loss of generality, that $d = 1$. With the notation of Section 4.5, Lemma 2 yields $\hat{\underset{\sim}{\mathbf{S}}}^{(n)}_{I;J}(\mathbf{M}) =$

$g^{(mn,0)}_{\mathbf{M}'^{-1}}\,\hat{\underset{\sim}{\mathbf{S}}}{}^{(n)}_{I;J}$. From Assumption (E2), $\mathbf{L}^{(n)}_{\hat{\boldsymbol{\theta}}}(\mathbf{M}) = g^{(mn,0)}_{\mathbf{M}}\mathbf{L}^{(n)}_{\hat{\boldsymbol{\theta}}}\,g^{(m,0)}_{\mathbf{M}^{-1}}$. Consequently, $\hat{\underset{\sim}{\mathbf{T}}}{}^{(n)}_{I;J}(\mathbf{M}) = g^{(m,0)}_{\mathbf{M}'^{-1}}\,\hat{\underset{\sim}{\mathbf{T}}}{}^{(n)}_{I;J}$. Analogously, $\hat{\underset{\sim}{\mathbf{S}}}{}^{(n)}_{II;J}(\mathbf{M}) = g^{(0,n-1)}_{\mathbf{M}'^{-1}}\,\hat{\underset{\sim}{\mathbf{S}}}{}^{(n)}_{II;J}$, $\mathbf{Q}^{(n)}_{\hat{\boldsymbol{\theta}}}(\mathbf{M}) = g^{(0,n-1)}_{\mathbf{M}}\mathbf{Q}^{(n)}_{\hat{\boldsymbol{\theta}}}\,g^{(0,\pi_0)}_{\mathbf{M}^{-1}}$, and therefore, $\hat{\underset{\sim}{\mathbf{T}}}{}^{(n)}_{II;J}(\mathbf{M}) = g^{(0,\pi_0)}_{\mathbf{M}'^{-1}}\,\hat{\underset{\sim}{\mathbf{T}}}{}^{(n)}_{II;J}$. This implies that $\hat{\underset{\sim}{\mathbf{T}}}{}^{(n)}_{J}(\mathbf{M}) = g^{(m,\pi_0)}_{\mathbf{M}'^{-1}}\,\hat{\underset{\sim}{\mathbf{T}}}{}^{(n)}_{J}$.

For the variances, $\mathbf{J}^{(n)}_{I;\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}}(\mathbf{M}) = g^{(m,0)}_{\mathbf{M}'^{-1}}\mathbf{J}^{(n)}_{I;\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}}\,g^{(m,0)}_{\mathbf{M}^{-1}}$ and $\mathbf{J}^{(n)}_{II;\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}}(\mathbf{M}) = g^{(0,\pi_0)}_{\mathbf{M}'^{-1}}\mathbf{J}^{(n)}_{II;\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}}\,g^{(0,\pi_0)}_{\mathbf{M}^{-1}}$. Since, moreover, $\mathbf{P}_{\hat{\boldsymbol{\theta}}}(\mathbf{M}) = g^{(0,\pi_0)}_{\mathbf{M}}\mathbf{P}_{\hat{\boldsymbol{\theta}}}\,g^{(0,\pi_0)}_{\mathbf{M}^{-1}}$ and $\mathbf{M}_{\hat{\boldsymbol{\theta}}}(\mathbf{M}) = g^{(0,\pi_0)}_{\mathbf{M}}\mathbf{M}_{\hat{\boldsymbol{\theta}}}\,g^{(0,p_1+q_1)}_{\mathbf{M}^{-1}}$, standard algebra shows that $\widehat{\mathcal{W}}^{(n)}_J(\mathbf{M}) = n\,(\underset{\sim}{\mathbf{T}}{}^{(n)}_J(\hat{\boldsymbol{\theta}}))'\boldsymbol{\Psi}'\boldsymbol{\Lambda}\boldsymbol{\Psi}\,\underset{\sim}{\mathbf{T}}{}^{(n)}_J(\hat{\boldsymbol{\theta}})$, where

$$\boldsymbol{\Psi} := \begin{pmatrix} \frac{1}{k}\,\mathrm{E}[J_0^2(U)]\,(\mathbf{J}^{(n)}_{I;\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}})^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{k^2}\,\mathrm{E}[J_1^2(U)]\,\mathrm{E}[J_2^2(U)]\,(\mathbf{J}^{(n)}_{II;\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}})^{-1} \end{pmatrix}^{1/2}$$

and (writing $\boldsymbol{\Pi}_{\boldsymbol{\Upsilon}}$ for the projection onto $\mathcal{M}(\boldsymbol{\Upsilon})$; see (9))

$$\boldsymbol{\Lambda} := \mathbf{I}_{km+k^2\pi_0} - \boldsymbol{\Pi}_{\boldsymbol{\Upsilon}}\left(\begin{pmatrix} \mathbf{J}^{(n)}_{I;\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}^{(n)}_{II;\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}} \end{pmatrix}^{1/2} \begin{pmatrix} (\mathbf{K}^{(n)}\otimes\mathbf{I}_k)^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{\hat{\boldsymbol{\theta}}}\mathbf{M}_{\hat{\boldsymbol{\theta}}} \end{pmatrix} g^{(m,p_1+q_1)}_{\mathbf{M}^{-1}}\,\boldsymbol{\Upsilon}\right).$$

Now, under Assumption (A2), the null hypothesis is affine-invariant, i.e., the couple $(\boldsymbol{\theta}_0, \boldsymbol{\Upsilon})$ is such that $g^{(m,p_1+q_1)}_{\mathbf{M}}(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})) = \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$ for any full-rank matrix $\mathbf{M}$. This implies that $\mathcal{M}(g^{(m,p_1+q_1)}_{\mathbf{M}}\boldsymbol{\Upsilon}) = \mathcal{M}(\boldsymbol{\Upsilon})$ for all such $\mathbf{M}$ (see the proof of Proposition 2 in Hallin and Paindaveine 2003a). The affine-invariance of $\widehat{\mathcal{W}}^{(n)}_J$ follows.

The asymptotic representation result of Proposition 2 will be sufficient (see the proof of (ii), (iii) below) to prove that all versions of $\widehat{\mathcal{W}}^{(n)}_J$ (based on any type of signs and ranks) have the same asymptotic representation, and thus are asymptotically equivalent; the asymptotic affine-invariance of the absolute-interdirection-based version of $\widehat{\mathcal{W}}^{(n)}_J$ follows since we just showed that the pseudo-Mahalanobis version of $\widehat{\mathcal{W}}^{(n)}_J$ is strictly affine-invariant.

Let us now prove that $\widehat{\mathcal{W}}^{(n)}_J$ is asymptotically invariant with respect to the group of continuous monotone radial transformations. Let $n^{1/2}\widetilde{\mathbf{T}}^{(n)}_{J;\boldsymbol{\Sigma}}(\boldsymbol{\theta})$ be given by

$$\begin{pmatrix} n^{1/2}\widetilde{\mathbf{T}}^{(n)}_{I;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}) \\ n^{1/2}\widetilde{\mathbf{T}}^{(n)}_{II;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}) \end{pmatrix} := \begin{pmatrix} \mathbf{L}^{(n)\prime}_{\boldsymbol{\theta}}\,(n^{1/2}(\mathrm{vec}\,\widetilde{\boldsymbol{\Lambda}}^{(n)}_{0;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}))', \ldots, (\mathrm{vec}\,\widetilde{\boldsymbol{\Lambda}}^{(n)}_{n-1;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}))')' \\ \mathbf{Q}^{(n)\prime}_{\boldsymbol{\theta}}\,((n-1)^{1/2}\,(\mathrm{vec}\,\widetilde{\boldsymbol{\Gamma}}^{(n)}_{1;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}))', \ldots, (\mathrm{vec}\,\widetilde{\boldsymbol{\Gamma}}^{(n)}_{n-1;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}))')' \end{pmatrix},$$

where $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$,

$$\widetilde{\boldsymbol{\Lambda}}^{(n)}_{i;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}) := (n-i)^{-1}\,\boldsymbol{\Sigma}'^{-1/2}\sum_{t=i+1}^{n} J_0\Big(\frac{R_t(\boldsymbol{\theta},\boldsymbol{\Sigma})}{n+1}\Big)\,\mathbf{U}_t(\boldsymbol{\theta},\boldsymbol{\Sigma})\,\mathbf{x}^{(n)\prime}_{t-i}\mathbf{K}^{(n)},$$

and

$$\widetilde{\boldsymbol{\Gamma}}^{(n)}_{i;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}) := \boldsymbol{\Sigma}'^{-1/2}\left(\frac{1}{n-i}\sum_{t=i+1}^{n} J_1\Big(\frac{R^{(n)}_t(\boldsymbol{\theta},\boldsymbol{\Sigma})}{n+1}\Big)J_2\Big(\frac{R_{t-i}(\boldsymbol{\theta},\boldsymbol{\Sigma})}{n+1}\Big)\,\mathbf{U}_t(\boldsymbol{\theta},\boldsymbol{\Sigma})\,\mathbf{U}'_{t-i}(\boldsymbol{\theta},\boldsymbol{\Sigma})\right)\boldsymbol{\Sigma}'^{1/2}.$$

Proceeding as in the proof of Proposition 2, one can verify that $\underset{\sim}{\mathbf{T}}_J^{(n)}(\boldsymbol{\theta}) - \widetilde{\mathbf{T}}_{J;\boldsymbol{\Sigma}}^{(n)}(\boldsymbol{\theta})$ is $o_{\mathrm{P}}(n^{-1/2})$ under $\bigcup_f \mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$. Using Lemma 3, this yields

$$n^{1/2}\,\underset{\sim}{\mathbf{T}}_J^{(n)}(\hat{\boldsymbol{\theta}}) = n^{1/2}\widetilde{\mathbf{T}}_{J;\boldsymbol{\Sigma}}^{(n)}(\boldsymbol{\theta}) -$$

$$\frac{C_k(J_0;f)}{k}\begin{pmatrix} \mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,(\mathbf{K}^{(n)} \otimes \mathbf{I}_k)^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{D_k(J_2;f)}{k}\,\mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,\mathbf{P}_{\boldsymbol{\theta}}\,\mathbf{M}_{\boldsymbol{\theta}} \end{pmatrix} n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + o_{\mathrm{P}}(1),$$

under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$. On the other hand, using the continuity of $\bar{\mathbf{Q}}_{I;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta})$ and $\bar{\mathbf{Q}}_{II;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$, we obtain that, under $\bigcup_f \mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$,

$$\widehat{\mathcal{W}}_J^{(n)} = \left(n^{1/2}\,\underset{\sim}{\mathbf{T}}_J^{(n)}(\hat{\boldsymbol{\theta}})\right)' \begin{pmatrix} \bar{\mathbf{Q}}_{I;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{Q}}_{II;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}) \end{pmatrix} \left(n^{1/2}\,\underset{\sim}{\mathbf{T}}_J^{(n)}(\hat{\boldsymbol{\theta}})\right) + o_{\mathrm{P}}(1)$$

(here, and in the sequel, $\bar{\mathbf{Q}}_{I;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta})$ (resp., $\bar{\mathbf{Q}}_{II;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta})$) denotes the array obtained by replacing $\mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}^{(n)}$ by $\mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}$ (resp., $\mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}}^{(n)}$ by $\mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}}$) in $\bar{\mathbf{Q}}_{I;J;\boldsymbol{\Sigma}}^{(n)}(\boldsymbol{\theta})$ (resp. in $\bar{\mathbf{Q}}_{II;J;\boldsymbol{\Sigma}}^{(n)}(\boldsymbol{\theta})$)). Writing $\bar{\mathbf{K}}$ for $\mathbf{K}^{(n)} \otimes \mathbf{I}_k$, and using Lemma 2.2.6 (c) in Rao and Mitra (1971),

$$\begin{pmatrix} \bar{\mathbf{Q}}_{I;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{Q}}_{II;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}) \end{pmatrix} \begin{pmatrix} \mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,\bar{\mathbf{K}}^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{D_k(J_2;f)}{k}\,\mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,\mathbf{P}_{\boldsymbol{\theta}}\,\mathbf{M}_{\boldsymbol{\theta}} \end{pmatrix} \boldsymbol{\Upsilon} =$$

$$\begin{pmatrix} c_I(\bar{\mathbf{K}}^{-1}\boldsymbol{\Upsilon}_I - \bar{\mathbf{K}}^{-1}\,\boldsymbol{\Upsilon}_I(\boldsymbol{\Upsilon}_I'\,\bar{\mathbf{K}}'^{-1}\,\mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,\bar{\mathbf{K}}^{-1}\,\boldsymbol{\Upsilon}_I)^{-1}\boldsymbol{\Upsilon}_I'\,\bar{\mathbf{K}}'^{-1}\,\mathbf{J}_{I;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,\bar{\mathbf{K}}^{-1}\boldsymbol{\Upsilon}_I) \\ c_{II}(\mathbf{P}_{\boldsymbol{\theta}}\mathbf{M}_{\boldsymbol{\theta}}\,\boldsymbol{\Upsilon}_{II} - \mathbf{P}_{\boldsymbol{\theta}}\mathbf{M}_{\boldsymbol{\theta}}\,\boldsymbol{\Upsilon}_{II}(\boldsymbol{\Upsilon}_{II}'\mathbf{M}_{\boldsymbol{\theta}}'\mathbf{P}_{\boldsymbol{\theta}}'\mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,\mathbf{P}_{\boldsymbol{\theta}}\mathbf{M}_{\boldsymbol{\theta}}\,\boldsymbol{\Upsilon}_{II})^{-}\boldsymbol{\Upsilon}_{II}'\mathbf{M}_{\boldsymbol{\theta}}'\mathbf{P}_{\boldsymbol{\theta}}'\,\mathbf{J}_{II;\boldsymbol{\theta},\boldsymbol{\Sigma}}\,\mathbf{P}_{\boldsymbol{\theta}}\mathbf{M}_{\boldsymbol{\theta}}\,\boldsymbol{\Upsilon}_{II}) \end{pmatrix} = \mathbf{0},$$

for some constants c_I, c_{II}. This and the constraints on $\hat{\boldsymbol{\theta}}$ jointly entail that, under $\bigcup_f \mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$, with $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^\star$,

$$\widehat{\mathcal{W}}_J^{(n)} = \left(n^{1/2}\widetilde{\mathbf{T}}_{J;\boldsymbol{\Sigma}}^{(n)}(\boldsymbol{\theta})\right)' \begin{pmatrix} \bar{\mathbf{Q}}_{I;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{Q}}_{II;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta}) \end{pmatrix} \left(n^{1/2}\widetilde{\mathbf{T}}_{J;\boldsymbol{\Sigma}}^{(n)}(\boldsymbol{\theta})\right) + o_{\mathrm{P}}(1),$$

which proves that $\widehat{\mathcal{W}}_J^{(n)}$ is indeed asymptotically invariant with respect to the group $\mathcal{G}_{\boldsymbol{\Sigma}}^{(n)}$, since $n^{1/2}\widetilde{\mathbf{T}}_{J;\boldsymbol{\Sigma}}^{(n)}(\boldsymbol{\theta})$ is strictly invariant with respect to that group.

(ii), (iii) Part *(i)* of the Proposition, and the continuity of $\bar{\mathbf{Q}}_{I;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta})$ and $\bar{\mathbf{Q}}_{II;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$ again, imply that $\widehat{\mathcal{W}}_J^{(n)}$ has the same asymptotic behavior as

$$n\,(\,\underset{\sim}{\mathbf{T}}_{I;J}^{(n)}(\boldsymbol{\theta}))'\,\bar{\mathbf{Q}}_{I;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta})\;\underset{\sim}{\mathbf{T}}_{I;J}^{(n)}(\boldsymbol{\theta}) + n\,(\,\underset{\sim}{\mathbf{T}}_{II;J}^{(n)}(\boldsymbol{\theta}))'\,\bar{\mathbf{Q}}_{II;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta})\;\underset{\sim}{\mathbf{T}}_{II;J}^{(n)}(\boldsymbol{\theta}) \tag{36}$$

under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$, $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$ as under the sequence of local alternatives $\mathcal{H}^{(n)}(\boldsymbol{\theta} + \boldsymbol{\nu}(n)\boldsymbol{\tau}, \boldsymbol{\Sigma}, f)$, with $\boldsymbol{\theta} \in (\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}))^*$ and $\boldsymbol{\nu}(n)\boldsymbol{\tau} \notin \mathcal{M}(\boldsymbol{\Upsilon})$. On the other hand, Proposition 2 implies that (36) behaves as

$$n\,\left(\widetilde{\mathbf{T}}_{I;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta})\right)'\,\bar{\mathbf{Q}}_{I;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta})\,\widetilde{\mathbf{T}}_{I;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}) + n\,\left(\widetilde{\mathbf{T}}_{I;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta})\right)'\,\bar{\mathbf{Q}}_{II;J;\boldsymbol{\Sigma}}(\boldsymbol{\theta})\,\widetilde{\mathbf{T}}_{II;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}), \tag{37}$$

where we let $n^{1/2}\widetilde{\mathbf{T}}_{I;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}) := \mathbf{L}_{\boldsymbol{\theta}}^{(n)\prime}(n^{1/2}\,(\mathrm{vec}\,\widetilde{\boldsymbol{\Lambda}}_{0;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}))', \ldots, (\mathrm{vec}\,\widetilde{\boldsymbol{\Lambda}}_{n-1;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}))')'$ and $n^{1/2}\widetilde{\mathbf{T}}_{II;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}) := \mathbf{Q}_{\boldsymbol{\theta}}^{(n)\prime}((n-1)^{1/2}\,(\mathrm{vec}\,\widetilde{\boldsymbol{\Gamma}}_{1;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}))', \ldots, (\mathrm{vec}\,\widetilde{\boldsymbol{\Gamma}}_{n-1;J;\boldsymbol{\Sigma},f}^{(n)}(\boldsymbol{\theta}))')'$.

Now, $n^{1/2}\widetilde{\mathbf{T}}^{(n)}_{I;J;\mathbf{\Sigma},f}(\boldsymbol{\theta})$ is asymptotically km-normal, with mean $\mathbf{0}$ under $\mathcal{H}^{(n)}(\boldsymbol{\theta},\mathbf{\Sigma},f)$ (with $\boldsymbol{\theta}\in(\boldsymbol{\theta}_0+\mathcal{M}(\mathbf{\Upsilon}))^*$), mean $(C_k(J_0;f)/k)\,\mathbf{J}_{I;\boldsymbol{\theta},\mathbf{\Sigma}}\,(\mathrm{vec}\,\boldsymbol{\eta}')$ under the sequence of local alternatives under consideration, and variance $(\mathrm{E}[J_0^2(U)]/k)\,\mathbf{J}_{I;\boldsymbol{\theta},\mathbf{\Sigma}}$ under both.

Since $(\mathrm{E}[J_0^2(U)]/k)\,\mathbf{J}^{1/2}_{I;\boldsymbol{\theta},\mathbf{\Sigma}}\,\bar{\mathbf{Q}}_{I;J;\mathbf{\Sigma}}(\boldsymbol{\theta})\,\mathbf{J}^{1/2}_{I;\boldsymbol{\theta},\mathbf{\Sigma}}$ is a symmetric idempotent matrix with rank $km-r_I$, this implies that the first term in (37) is asymptotically chi-square with $km-r_I$ degrees of freedom under $\mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\mathbf{\Upsilon}}$, and asymptotically noncentral chi-square, still with $km-r_I$ degrees of freedom but with noncentrality parameter $(C_k^2(J_0;f)/(k\,\mathrm{E}[J_0^2(U)]))\,r_{\boldsymbol{\theta},\mathbf{\Sigma}}(\boldsymbol{\eta})$ under $\mathcal{H}^{(n)}(\boldsymbol{\theta}+\boldsymbol{\nu}(n)\boldsymbol{\tau},\mathbf{\Sigma},f)$ with $\boldsymbol{\theta}\in(\boldsymbol{\theta}_0+\mathcal{M}(\mathbf{\Upsilon}))^*$, and $\boldsymbol{\nu}(n)\boldsymbol{\tau}\notin\mathcal{M}(\mathbf{\Upsilon})$.

For the serial part in (37), $n^{1/2}\widetilde{\mathbf{T}}^{(n)}_{II;J;\mathbf{\Sigma},f}(\boldsymbol{\theta})$ is asymptotically $k^2\pi_0$-normal, with mean $\mathbf{0}$ under $\mathcal{H}^{(n)}(\boldsymbol{\theta},\mathbf{\Sigma},f)$, $\boldsymbol{\theta}\in(\boldsymbol{\theta}_0+\mathcal{M}(\mathbf{\Upsilon}))^*$, mean

$$\frac{1}{k^2}\,D_k(J_2;f)\,C_k(J_1;f)\,\mathbf{J}_{II;\boldsymbol{\theta},\mathbf{\Sigma}}\,\mathbf{P}_{\boldsymbol{\theta}}\,\mathbf{M}_{\boldsymbol{\theta}}\begin{pmatrix}\mathrm{vec}\,\boldsymbol{\gamma}\\ \mathrm{vec}\,\boldsymbol{\delta}\end{pmatrix}$$

under the sequence of alternatives under consideration, and variance $(\mathrm{E}[J_1^2(U)]\mathrm{E}[J_2^2(U)]/k^2)\,\mathbf{J}_{II;\boldsymbol{\theta},\mathbf{\Sigma}}$ under both. The result follows from the fact that $(\mathrm{E}[J_1^2(U)]\mathrm{E}[J_2^2(U)]/k^2)\,\mathbf{J}^{1/2}_{II;\boldsymbol{\theta},\mathbf{\Sigma}}\,\bar{\mathbf{Q}}_{II;J;\mathbf{\Sigma}}(\boldsymbol{\theta})\,\mathbf{J}^{1/2}_{II;\boldsymbol{\theta},\mathbf{\Sigma}}$ is a symmetric idempotent matrix with rank

$$\begin{aligned}
&\mathrm{tr}\,(\mathbf{I}_{k^2\pi_0}-\mathbf{J}^{1/2}_{II;\boldsymbol{\theta},\mathbf{\Sigma}}\mathbf{P}_{\boldsymbol{\theta}}\mathbf{M}_{\boldsymbol{\theta}}\,\mathbf{\Upsilon}_{II}(\mathbf{\Upsilon}'_{II}\mathbf{M}'_{\boldsymbol{\theta}}\mathbf{P}'_{\boldsymbol{\theta}}\mathbf{J}_{II;\boldsymbol{\theta},\mathbf{\Sigma}}\,\mathbf{P}_{\boldsymbol{\theta}}\mathbf{M}_{\boldsymbol{\theta}}\,\mathbf{\Upsilon}_{II})^{-}\mathbf{\Upsilon}'_{II}\mathbf{M}'_{\boldsymbol{\theta}}\mathbf{P}'_{\boldsymbol{\theta}}\mathbf{J}^{1/2}_{II;\boldsymbol{\theta},\mathbf{\Sigma}})\\
&\quad= k^2\pi_0-\mathrm{tr}\,((\mathbf{\Upsilon}'_{II}\mathbf{M}'_{\boldsymbol{\theta}}\mathbf{P}'_{\boldsymbol{\theta}}\mathbf{J}_{II;\boldsymbol{\theta},\mathbf{\Sigma}}\mathbf{P}_{\boldsymbol{\theta}}\mathbf{M}_{\boldsymbol{\theta}}\,\mathbf{\Upsilon}_{II})(\mathbf{\Upsilon}'_{II}\mathbf{M}'_{\boldsymbol{\theta}}\mathbf{P}'_{\boldsymbol{\theta}}\mathbf{J}_{II;\boldsymbol{\theta},\mathbf{\Sigma}}\,\mathbf{P}_{\boldsymbol{\theta}}\mathbf{M}_{\boldsymbol{\theta}}\,\mathbf{\Upsilon}_{II})^{-})\\
&\quad= k^2\pi_0-\mathrm{rank}(\mathbf{\Upsilon}'_{II}\mathbf{M}'_{\boldsymbol{\theta}}\mathbf{P}'_{\boldsymbol{\theta}}\mathbf{J}_{II;\boldsymbol{\theta},\mathbf{\Sigma}}\mathbf{P}_{\boldsymbol{\theta}}\mathbf{M}_{\boldsymbol{\theta}}\,\mathbf{\Upsilon}_{II})\\
&\quad= k^2\pi_0-\mathrm{rank}(\dot{\mathbf{M}}_{\boldsymbol{\theta}}\,\dot{\mathbf{\Upsilon}}_{II})=k^2\pi_0-\min(k^2(p_0+q_0),r_{II})=k^2\pi_0-r_{II},
\end{aligned}$$

since $\mathbf{M}_{\boldsymbol{\theta}}\,\mathbf{\Upsilon}_{II}=\dot{\mathbf{M}}_{\boldsymbol{\theta}}\,\dot{\mathbf{\Upsilon}}_{II}$ is the product of two full-rank matrices.

(iv) Adapted to the current context, Hallin and Puri (1994)'s general Lemma 5.12 shows that the test $\underline{\phi}^{(n)}_{\mathbf{\Sigma},f_\star}$ that rejects the null hypothesis whenever

$$\begin{aligned}
&\mathbf{\Delta}^{(n)\prime}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta})\Big[\mathbf{I}-\mathbf{\Gamma}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta})\mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{\Gamma}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta})\mathbf{\Upsilon})^{-}\mathbf{\Upsilon}'\mathbf{\Gamma}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta})(\mathbf{\Gamma}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta}))^{-}\Big]'(\mathbf{\Gamma}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta}))^{-}\\
&\qquad\Big[\mathbf{I}-\mathbf{\Gamma}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta})\mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{\Gamma}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta})\mathbf{\Upsilon})^{-}\mathbf{\Upsilon}'\mathbf{\Gamma}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta})(\mathbf{\Gamma}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta}))^{-}\Big]\,\mathbf{\Delta}^{(n)}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta})>\chi^2_{s,1-\alpha},
\end{aligned}$$

where $s:=\mathrm{rank}(\mathbf{\Gamma}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta}))-\mathrm{rank}(\mathbf{\Upsilon}'\mathbf{\Gamma}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta})\mathbf{\Upsilon})$, is locally and asymptotically most stringent for $\mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\mathbf{\Upsilon}}(\mathbf{\Sigma},f_\star)$ against $\bigcup_{\boldsymbol{\theta}\notin\boldsymbol{\theta}_0+\mathcal{M}(\mathbf{\Upsilon})}\bigcup_{\mathbf{\Sigma}}\mathcal{H}^{(n)}(\boldsymbol{\theta},\mathbf{\Sigma},f_\star)$, at asymptotic probability level α. Of course, the same optimality property holds for the asymptotically equivalent (under $\mathcal{H}^{(n)}_{\boldsymbol{\theta}_0,\mathbf{\Upsilon}}(\mathbf{\Sigma},f_\star)$, and under contiguous alternatives) test $\underline{\phi}^{(n)}_{f_\star}$ that rejects the null hypothesis whenever

$$\begin{aligned}
\underline{\mathcal{W}}^{(n)}_{f_\star}:=\;&\underset{\sim}{\mathbf{\Delta}}^{(n)\prime}_{f_\star}(\hat{\boldsymbol{\theta}})\Big[\mathbf{I}-\hat{\mathbf{\Gamma}}^{(n)}_{f_\star}(\hat{\boldsymbol{\theta}})\mathbf{\Upsilon}(\mathbf{\Upsilon}'\hat{\mathbf{\Gamma}}^{(n)}_{f_\star}(\hat{\boldsymbol{\theta}})\mathbf{\Upsilon})^{-}\mathbf{\Upsilon}'\hat{\mathbf{\Gamma}}^{(n)}_{f_\star}(\hat{\boldsymbol{\theta}})(\hat{\mathbf{\Gamma}}^{(n)}_{f_\star}(\hat{\boldsymbol{\theta}}))^{-}\Big]'(\hat{\mathbf{\Gamma}}^{(n)}_{f_\star}(\hat{\boldsymbol{\theta}}))^{-}\\
&\Big[\mathbf{I}-\hat{\mathbf{\Gamma}}^{(n)}_{f_\star}(\hat{\boldsymbol{\theta}})\mathbf{\Upsilon}(\mathbf{\Upsilon}'\hat{\mathbf{\Gamma}}^{(n)}_{f_\star}(\hat{\boldsymbol{\theta}})\mathbf{\Upsilon})^{-}\mathbf{\Upsilon}'\hat{\mathbf{\Gamma}}^{(n)}_{f_\star}(\hat{\boldsymbol{\theta}})(\hat{\mathbf{\Gamma}}^{(n)}_{f_\star}(\hat{\boldsymbol{\theta}}))^{-}\Big]\,\underset{\sim}{\mathbf{\Delta}}^{(n)}_{f_\star}(\hat{\boldsymbol{\theta}})>\chi^2_{s,1-\alpha},
\end{aligned}\tag{38}$$

where

$$\underset{\sim}{\mathbf{\Delta}}^{(n)}_{f_\star}(\boldsymbol{\theta}):=n^{1/2}\begin{pmatrix}\mathbf{I}_{km} & \mathbf{0}\\ \mathbf{0} & \mathbf{M}'_{\boldsymbol{\theta}}\,\mathbf{P}'_{\boldsymbol{\theta}}\end{pmatrix}\underset{\sim}{\mathbf{T}}^{(n)}_{J}(\boldsymbol{\theta}),$$

with $J_0 = J_1 := \varphi_{f_\star} \circ \tilde{F}_{\star k}^{-1}$ and $J_2 = \tilde{F}_{\star k}^{-1}$, and

$$\hat{\mathbf{\Gamma}}_{f_\star}^{(n)}(\boldsymbol{\theta}) := \begin{pmatrix} \frac{1}{k}\,\mathcal{I}_{k,f_\star}\,\mathbf{J}_{I;\boldsymbol{\theta},\hat{\mathbf{\Sigma}}}^{(n)} & \mathbf{0} \\ \mathbf{0} & \frac{\mu_{k+1;f_\star}\,\mathcal{I}_{k,f_\star}}{k^2\,\mu_{k-1;f_\star}}\,\mathbf{M}_{\boldsymbol{\theta}}'\mathbf{P}_{\boldsymbol{\theta}}'\mathbf{J}_{II;\boldsymbol{\theta},\hat{\mathbf{\Sigma}}}^{(n)}\mathbf{P}_{\boldsymbol{\theta}}\mathbf{M}_{\boldsymbol{\theta}} \end{pmatrix} = \mathbf{\Gamma}_{\mathbf{\Sigma},f_\star}(\boldsymbol{\theta}) + o_{\mathrm{P}}(1)$$

under $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \mathbf{\Sigma}, f_\star)$.

If we can assume that $\underline{\mathcal{W}}_{f_\star}^{(n)} = \widehat{\mathcal{W}}_{f_\star}^{(n)}$, then, by *(ii)*, we have $s = km + k^2\pi_0 - r$, so that $\underline{\phi}_{f_\star}^{(n)}$ and $\phi_{f_\star}^{(n)}$ actually coincide. The result then follows from the invariance properties of $\phi_{f_\star}^{(n)}$. In order to complete the proof, it is thus sufficient to show that indeed $\underline{\mathcal{W}}_{f_\star}^{(n)} = \widehat{\mathcal{W}}_{f_\star}^{(n)}$. The block-diagonal structure of the quadratic form in the definition of $\underline{\mathcal{W}}_{f_\star}^{(n)}$ allows for a decomposition of the form $\underline{\mathcal{W}}_{I;f_\star}^{(n)} + \underline{\mathcal{W}}_{II;f_\star}^{(n)}$, where $\underline{\mathcal{W}}_{I;f_\star}^{(n)}$ (resp., $\underline{\mathcal{W}}_{II;f_\star}^{(n)}$) deals with the trend part (resp., the serial part). While routine algebra yields $\underline{\mathcal{W}}_{I;f_\star}^{(n)} = \widehat{\mathcal{W}}_{I;f_\star}^{(n)}$, the situation for the serial part is more intricate, mainly due to the presence of generalized inverses. Write $\hat{\mathbf{P}}$, $\hat{\mathbf{M}}$, $\hat{\mathbf{J}}_{II}$, and $\hat{\mathbf{N}}$ for $\mathbf{P}_{\hat{\boldsymbol{\theta}}}$, $\mathbf{M}_{\hat{\boldsymbol{\theta}}}$, $\mathbf{J}_{II;\hat{\boldsymbol{\theta}},\hat{\mathbf{\Sigma}}}^{(n)}$, and $\mathbf{M}_{\hat{\boldsymbol{\theta}}}'\mathbf{P}_{\hat{\boldsymbol{\theta}}}'\mathbf{J}_{II;\hat{\boldsymbol{\theta}},\hat{\mathbf{\Sigma}}}^{(n)}\mathbf{P}_{\hat{\boldsymbol{\theta}}}\mathbf{M}_{\hat{\boldsymbol{\theta}}}$, respectively. Standard calculation yields

$$\underline{\mathcal{W}}_{II;f_\star}^{(n)} = \frac{nk^2\,\mu_{k-1;f_\star}}{\mu_{k+1;f_\star}\,\mathcal{I}_{k,f_\star}}\,\underset{\sim}{\mathbf{T}}_{II;J}^{(n)\prime}(\hat{\boldsymbol{\theta}})\,\hat{\mathbf{P}}\,\hat{\mathbf{M}}\left\{\hat{\mathbf{N}}^-\left[\mathbf{I} - \hat{\mathbf{N}}\,\mathbf{\Upsilon}_{II}\left(\mathbf{\Upsilon}_{II}'\,\hat{\mathbf{N}}\,\mathbf{\Upsilon}_{II}\right)^-\mathbf{\Upsilon}_{II}'\,\hat{\mathbf{N}}\,\hat{\mathbf{N}}^-\right]\right\}\hat{\mathbf{M}}'\,\hat{\mathbf{P}}'\,\underset{\sim}{\mathbf{T}}_{II;J}^{(n)}(\hat{\boldsymbol{\theta}}),$$

that is, in view of Lemma 2.2.6(c) in Rao and Mitra (1971),

$$\underline{\mathcal{W}}_{II;f_\star}^{(n)} = \frac{nk^2\,\mu_{k-1;f_\star}}{\mu_{k+1;f_\star}\,\mathcal{I}_{k,f_\star}}\,\underset{\sim}{\mathbf{T}}_{II;J}^{(n)\prime}(\hat{\boldsymbol{\theta}})\left[\hat{\mathbf{P}}\,\hat{\mathbf{M}}\,\hat{\mathbf{N}}^-\,\hat{\mathbf{M}}'\,\hat{\mathbf{P}}' - \hat{\mathbf{P}}\,\hat{\mathbf{M}}\,\mathbf{\Upsilon}_{II}\left(\mathbf{\Upsilon}_{II}'\,\hat{\mathbf{N}}\,\mathbf{\Upsilon}_{II}\right)^-\mathbf{\Upsilon}_{II}'\,\hat{\mathbf{M}}'\,\hat{\mathbf{P}}'\right]\underset{\sim}{\mathbf{T}}_{II;J}^{(n)}(\hat{\boldsymbol{\theta}}).$$

This implies that $\underline{\mathcal{W}}_{II;f_\star}^{(n)} = \widehat{\mathcal{W}}_{II;f_\star}^{(n)}$, since $\hat{\mathbf{M}}\,\mathbf{\Upsilon}_{II} = \dot{\hat{\mathbf{M}}}\,\dot{\mathbf{\Upsilon}}_{II}$, and since, from Lemma 2.2.5(c) in Rao and Mitra (1971), $\hat{\mathbf{P}}\,\hat{\mathbf{M}}\,(\hat{\mathbf{M}}'\hat{\mathbf{P}}'\hat{\mathbf{J}}_{II}\,\hat{\mathbf{P}}\,\hat{\mathbf{M}})^-\,\hat{\mathbf{M}}'\hat{\mathbf{P}}' = \hat{\mathbf{J}}_{II}^{-1}$. Consequently, $\underline{\mathcal{W}}_{f_\star}^{(n)} = \widehat{\mathcal{W}}_{f_\star}^{(n)}$, which completes the proof. □

Acknowledgement. The authors are most grateful to an anonymous referee for his very careful reading of the initial version of this paper, and his pertinent remarks, that helped correcting several inaccuracies and greatly improved the presentation.

References

[1] Brockwell, P. J., and R. A. Davis (1987). *Time Series : Theory and Methods*, Springer, New York.

[2] Dunsmuir, W., and E.J. Hannan (1976). Vector linear time series, *Advances in Applied Probability* **8**, 339–364.

[3] Deistler, M., W. Dunsmuir, and E.J. Hannan (1978). Vector linear time series: corrections and extensions, *Advances in Applied Probability* **10**, 360–372.

[4] Garel, B., and M. Hallin (1995). Local asymptotic normality of multivariate ARMA processes with a linear trend, *Annals of the Institute of Statistical Mathematics* **47**, 551-579.

[5] Garel, B. and M. Hallin (1999). Rank-based AR order identification, *Journal of the American Statistical Association* **94**, 1357-1371.

[6] Hallin, M. (1986). Non-stationary q-dependent processes and time-varying moving-average models : invertibility properties and the forecasting problem, *Advances in Applied Probability* **18**, 170-210.

[7] Hallin, M., J.-Fr Ingenbleek, and M.L. Puri (1989). Asymptotically most powerful rank tests for multivariate randomness against serial dependence. *Journal of Multivariate Analysis* **30**, 34-71.

[8] Hallin, M., and D. Paindaveine (2002a). Optimal tests for multivariate location based on interdirections and pseudo-Mahalanobis ranks, *Annals of Statistics* **30**, 1103-1133.

[9] Hallin, M., and D. Paindaveine (2002b). Optimal procedures based on interdirections and pseudo-Mahalanobis ranks for testing multivariate elliptic white noise against ARMA dependence, *Bernoulli* **8**, 787-816.

[10] Hallin, M., and D. Paindaveine (2002c). Multivariate signed ranks : Randles' interdirections or Tyler's angles? In Y. Dodge, Ed., *Statistical data analysis based on the L1-norm and related methods*, Birkhäuser, Basel, 271-282.

[11] Hallin, M., and D. Paindaveine (2003a). Affine invariant linear hypotheses for the multivariate general linear model with VARMA error terms. In M. Moore, S. Froda, and Chr. Léger, Eds, *Mathematical Statistics and Applications: Festschrift for Constance van Eeden*, I.M.S. Lecture Notes-Monograph Series, I.M.S., Hayward, California, 417-434.

[12] Hallin, M., and D. Paindaveine (2004a). Rank-based optimal tests of the adequacy of an elliptic VARMA model, *Annals of Statistic*, to appear.

[13] Hallin, M., and D. Paindaveine (2004b). Asymptotic linearity of serial and nonserial multivariate signed rank statistics. Submitted.

[14] Hallin, M., and M.L. Puri (1988). Optimal rank-based procedures for time-series analysis: testing an ARMA model against other ARMA models, *Annals of Statistics* **16**, 402-432.

[15] Hallin, M., and M.L. Puri (1991). Time-series analysis via rank-order theory: signed-rank tests for *ARMA* models. *Journal of Multivariate Analysis* **39**, 1-29.

[16] Hallin, M., and M.L. Puri (1994). Aligned rank tests for linear models with autocorrelated error terms. *Journal of Multivariate Analysis* **50**, 175-237.

[17] Hallin, M., and M.L. Puri (1995). A multivariate Wald-Wolfowitz rank test against serial dependence. *Canadian Journal of Statistics* **23**, 55-65.

[18] Hallin, M., and B. J. M. Werker (1999). Optimal testing for semi-parametric AR models: from Gaussian Lagrange multipliers to autoregression rank scores and adaptive tests. In S. Ghosh, Ed., *Asymptotics, Nonparametrics, and Time Series*, 295-350. M. Dekker, New York.

[19] Hallin, M., and B. J. M. Werker (2003). Semiparametric efficiency, distribution-freeness, and invariance, *Bernoulli* **9**, 137-165.

[20] Hannan, E. J. (1970). *Multiple Time Series*. J. Wiley, New York.

[21] Hettmansperger, T. P., J. Nyblom, and H. Oja (1994). Affine invariant multivariate one-sample sign tests, *Journal of the Royal Statistical Society Series B* **56**, 221-234.

[22] Hettmansperger, T. P., J. Möttönen, and H. Oja (1997). Affine invariant multivariate one-sample signed-rank tests, *Journal of the American Statistical Association* **92**, 1591-1600.

[23] Jan, S.-L., and R. H. Randles (1994). A multivariate signed-sum test for the one-sample location problem, *Journal of Nonparametric Statistics* **4**, 49-63.

[24] Kreiss, J.P. (1987). On adaptive estimation in stationary ARMA processes, *Annals of Statistics* **15**, 112-133.

[25] Le Cam, L. (1960). Locally asymptotically normal families of distributions, *University of California Publications in Statistics* **3**, 37-98.

[26] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory.* Springer-Verlag, New York.

[27] Marden, J. (1999). Multivariate rank tests. In S. Ghosh, Ed., *Multivariate Analysis, Design of Experiments, and Sampling Survey*, 401-432. M. Dekker, New York.

[28] Möttönen, J., and H. Oja (1995). Multivariate spatial sign and rank methods, *Journal of Nonparametric Statistics* **5**, 201-213.

[29] Oja, H. (1999). Affine invariant multivariate sign and rank tests and corresponding estimates : a review, *Scandinavian Journal of Statistics* **26**, 319-343.

[30] Oja, H., and D. Paindaveine (2004). Optimal signed-rank tests based on hyperplanes, *Journal of Statistical Planning and Inference*, to appear.

[31] Peters, D., and R. H. Randles (1990). A multivariate signed-rank test for the one-sample location problem, *Journal of the American Statistical Association* **85**, 552-557.

[32] Pötscher, B. M. (1983). Order estimation in ARMA models by Lagrangian multiplier tests, *Annals of Statistics* **11**, 872-885.

[33] Pötscher, B. M. (1985). The behaviour of the Lagrange multiplier test in testing the orders of an ARMA model, *Metrika* **32**, 129-150.

[34] Puri, M. L., and P. K. Sen (1971). *Nonparametric Methods in Multivariate Analysis.* J. Wiley, New York.

[35] Randles, R. H. (1989). A distribution-free multivariate sign test based on interdirections, *Journal of the American Statistical Association* **84**, 1045-1050.

[36] Randles, R. H. (2000). A simpler, affine-invariant, multivariate, distribution-free sign test, *Journal of the American Statistical Association* **95**, 1263-1268.

[37] Randles, R. H. and Y. Um (1998). Nonparametric tests for the multivariate multi-sample location problem, *Statistica Sinica* **8**, 801-812.

[38] Rao, C. R., and S. K. Mitra (1971). *Generalized Inverses of Matrices and its Applications.* J. Wiley, New York.

[39] Reinsel, G. C. (1997). *Elements of multivariate Time Series Analysis.* Springer-Verlag, New York.

[40] Swensen, A. R. (1985). The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend, *Journal of Multivariate Analysis* **16**, 54-70.

Marc Hallin
Département de Mathématique,
I.S.R.O., and E.C.A.R.E.S.
Université Libre de Bruxelles
Campus de la Plaine CP 210
B-1050 Bruxelles Belgium
mhallin@ulb.ac.be

Davy Paindaveine
Département de Mathématique,
I.S.R.O., and E.C.A.R.E.S.
Université Libre de Bruxelles
Campus de la Plaine CP 210
B-1050 Bruxelles Belgium
dpaindav@ulb.ac.be