TESTING UNIFORMITY ON HIGH-DIMENSIONAL SPHERES AGAINST CONTIGUOUS ROTATIONALLY SYMMETRIC ALTERNATIVES

BY CHRISTINE CUTTING, DAVY PAINDAVEINE* AND THOMAS VERDEBOUT

Université libre de Bruxelles

We consider the problem of testing for uniformity on high-dimensional unit spheres. We are primarily interested in non-null issues. To this end, we consider rotationally symmetric alternatives and identify alternatives that are contiguous to the null of uniformity. This reveals a Locally and Asymptotically Normality (LAN) structure, which, for the first time, allows to use Le Cam's third lemma in the high-dimensional setup. Under very mild assumptions, we derive the asymptotic non-null distribution of the high-dimensional Rayleigh test and show that this test actually exhibits slower consistency rates. All (n, p)-asymptotic results we derive are "universal", in the sense that the dimension p is allowed to go to infinity in an arbitrary way as a function of the sample size n. Part of our results also cover the low-dimensional case, which allows to explain heuristically the high-dimensional non-null behavior of the Rayleigh test. A Monte Carlo study confirms our asymptotic results.

1. Introduction. Problems involving a number p of variables that is large compared to the number n of individuals are of course very common in modern statistics. In various such problems, only the relative magnitude of the p variables is important, so that it may be assumed that the observations belong to the unit sphere $\mathcal{S}^{p-1} := \{\mathbf{x} \in \mathbb{R}^p : \|\mathbf{x}\| = \sqrt{\mathbf{x}'\mathbf{x}} = 1\}$, with p large. This found applications in magnetic resonance, gene-expression, and text mining, among others; see Dryden (2005), Banerjee et al. (2003), and Banerjee et al. (2005), respectively. Developing inference procedures for data on high-dimensional

*Research is supported by an A.R.C. contract from the Communauté Française de Belgique and by the IAP research network grant nr. P7/06 of the Belgian government (Belgian Science Policy).

MSC 2010 subject classifications: Primary 62H11, 62G20; secondary 62H15

Keywords and phrases: Contiguity, high-dimensional statistics, local and asymptotic normality, rotationally symmetric distributions, spherical statistics, tests of uniformity

spheres, in an (n, p)-asymptotic framework, is therefore a legitimate objective, which was already considered, e.g., in Dryden (2005), Cai and Jiang (2012), Cai, Fan and Jiang (2013), Paindaveine and Verdebout (2015), and Ley, Paindaveine and Verdebout (2015).

In this paper, we consider hypothesis testing for high-dimensional spherical data and restrict to the most fundamental problem in the field, namely the problem of testing for uniformity over the unit sphere. More precisely, we assume that observations take the form of a triangular array of random vectors $\mathbf{X}_{ni}$, $i = 1, \ldots, n$, $n = 1, 2, \ldots$, where, for any n, the $\mathbf{X}_{ni}$'s are mutually independent and share a common distribution on the unit sphere $\mathcal{S}^{p_n-1}$, and, as in Cai and Jiang (2012), Cai, Fan and Jiang (2013), and Paindaveine and Verdebout (2015), we want to test the null hypothesis $\mathcal{H}_{0n}$ that the common distribution of the $\mathbf{X}_{ni}$'s, $i = 1, \ldots, n$ is the uniform over $\mathcal{S}^{p_n-1}$. We are primarily interested in the high-dimensional case ($p_n \to \infty$); yet some of our results will also apply to the low-dimensional case (the sequence (p_n) is bounded), hence to the classical fixed p-case ($p_n = p$ for n large enough). This will allow us to compare the low- and high-dimensional cases.

Whenever p_n-dimensional observations $\mathbf{X}_{in}$, $i = 1, \ldots, n$ are available, the most classical test of uniformity is the Rayleigh (1919) test, that rejects $\mathcal{H}_{0n}$ for large values of

$$R_n := np_n \|\bar{\mathbf{X}}_n\|^2 = p_n + \frac{2p_n}{n} \sum_{1 \leq i < j \leq n} \mathbf{X}'_{ni} \mathbf{X}_{nj},$$

where $\bar{\mathbf{X}}_n := \frac{1}{n} \sum_{i=1}^n \mathbf{X}_{ni}$. For fixed p, the test is based on the null asymptotic χ^2_p distribution of R_n. In the high-dimensional setup, Paindaveine and Verdebout (2015) showed that, after appropriate standardization, the Rayleigh statistic is asymptotically normal under the null. More precisely, they established the following result.

THEOREM 1.1 (Paindaveine and Verdebout (2015)). *Let p_n be a sequence of positive integers converging to ∞ as $n \to \infty$. Assume that $\mathbf{X}_{ni}$, $i = 1, \ldots, n$, $n = 1, 2, \ldots$, is a triangular array of random vectors such that, for any n, $\mathbf{X}_{n1}, \mathbf{X}_{n2}, \ldots, \mathbf{X}_{nn}$ are mutually independent and are all uniformly distributed on $\mathcal{S}^{p_n-1}$. Then*

$$R_n^{\text{St}} := \frac{R_n - p_n}{\sqrt{2p_n}} = \frac{\sqrt{2p_n}}{n} \sum_{1 \leq i < j \leq n} \mathbf{X}'_{ni} \mathbf{X}_{nj} \tag{1.1}$$

converges weakly to the standard normal as $n \to \infty$.

The resulting high-dimensional Rayleigh test, that rejects $\mathcal{H}_{0n}$ at asymptotic level α whenever R_n^{St} exceeds the α-upper standard Gaussian quantile, has excellent asymptotic null properties. Quite remarkably, indeed, the asymptotic result above is "universal" in the sense that it does not impose any condition on the way p_n goes to infinity with n, so that the test can be applied as soon as p_n and n are large, without bothering about their relative magnitude. This is in sharp contrast with most results in high-dimensional statistics that typically impose conditions such as $p_n/n \to c$ for some $c > 0$ (for another recent work that defines a universal high-dimensional test, see Wang, Peng and Li (2015)). Moreover, a common asymptotic null distribution is obtained for each (n, p)-regime, unlike, e.g., for the tests proposed in Cai and Jiang (2012) and Cai, Fan and Jiang (2013).

On their own, however, the asymptotic result stated in Theorem 1.1 is not sufficient to justify resorting to the Rayleigh test : the trivial test, that would discard the data and reject $\mathcal{H}_{0n}$ with probability α, is also "universally" asymptotically valid under the null, yet is a very poor test since its power function is uniformly equal to the nominal level α. Showing that the Rayleigh test is to be recommended therefore requires a careful investigation of its power behaviour, which is one of the main objectives of this paper.

Throughout, we will actually consider a specific, semiparametric, class of alternatives, associated with the so-called *rotationally symmetric distributions*; see Section 2 for a definition. While they may at first seem arbitrary, these alternatives are actually the analog, in the directional problem of testing for uniformity, of the "spiked" alternatives that are very often considered in the Euclidean case for tests on covariance matrices; see, e.g., Berthet and Rigollet (2013), Onatski, Moreira and Hallin (2013, 2014), and Wang, Berthet and Samworth (2014). Our first main contribution is to identify sequences of rotationally symmetric alternatives that are contiguous to the null of uniformity. More : we actually show that Local and Asymptotical Normality (LAN) holds in the vicinity of uniformity, which, to the best of our knowledge, is the first instance of LAN in high-dimensional statistics. This LAN result is universal, still in the sense that it does not impose any restriction on the way p_n goes to infinity with n (throughout, we use the word *universal* in this sense).

As usual, the LAN structure allows to use sophisticated asymptotic results such as Le Cam's third lemma. When applied to the Rayleigh test, this well-known asymptotic result reveals that the Rayleigh test has, under sequences of contiguous alternatives, asymptotic powers that are equal to the nominal level α. Under such alternatives, thus, this test is no better than the trivial test. To have a complete understanding of the asymptotic power behavior of the Rayleigh test, we then derive its asymptotic distribution under virtually arbitrary rotationally symmetric alternatives. Our result, which, at least under the so-called Fisher-von Mises Langevin (FvML) alternatives, is universal, allows to identify sequences of rotationally symmetric alternatives along which the Rayleigh test achieves non-trivial limiting powers. While it is not asymptotically optimal (not even rate-optimal), the Rayleigh test therefore is of interest in the high-dimensional setup considered.

The outline of the paper is as follows. In Section 2, we define the class of rotationally symmetric alternatives and we identify the corresponding contiguous alternatives. There, we also provide the first LAN result in the high-dimensional setup. In Section 3, we describe some (infeasible) "oracle" tests that achieve Le Cam optimality, we explain their link with the Rayleigh test, and we apply Le Cam's third lemma to the latter. In Section 4, we derive the asymptotic distribution of the Rayleigh test statistic under general rotationally symmetric alternatives and study the resulting limiting powers. In Section 5, we illustrate our asymptotic results through simulations. We summarize the main findings of the paper in Section 6. Finally, the appendix collects most of the proofs (the remaining proofs, that require original bounds on modified Bessel functions ratios, are reported in the supplementary article Cutting, Paindaveine and Verdebout (2015)).

2. Contiguous alternatives and local asymptotic normality. As already mentioned, we will consider specific alternatives to the null uniformity over the p-dimensional unit sphere $\mathcal{S}^{p-1}$, namely rotationally symmetric alternatives. A p-dimensional vector $\mathbf{X}$ is said to be *rotationally symmetric about* $\boldsymbol{\theta}(\in \mathcal{S}^{p-1})$ if and only if $\mathbf{OX}$ is equal in distribution to $\mathbf{X}$ for any orthogonal $p \times p$ matrix $\mathbf{O}$ satisfying $\mathbf{O}\boldsymbol{\theta} = \boldsymbol{\theta}$; see, e.g., Saw (1978). Such distributions are fully characterized by the location parameter $\boldsymbol{\theta}$ and the cumulative distribution function F of $\mathbf{X}'\boldsymbol{\theta}$. The null of uniformity (under which $\boldsymbol{\theta}$ is not identifiable) is associated with

$$F_p(t) := c_p \int_{-1}^{t} (1-s^2)^{(p-3)/2}\, ds, \qquad \text{with } c_p := \frac{\Gamma\big(\frac{p}{2}\big)}{\sqrt{\pi}\,\Gamma\big(\frac{p-1}{2}\big)}, \tag{2.1}$$

where $\Gamma(\cdot)$ is the Euler Gamma function. Particular alternatives are given, e.g., by the so-called Fisher-von Mises-Langevin (FvML) distributions, that are obtained for

$$F_{p,\kappa}^{\mathrm{FvML}}(t) := c_{p,\kappa}^{\mathrm{FvML}} \int_{-1}^{t} (1-s^2)^{(p-3)/2} \exp(\kappa s)\, ds, \qquad \text{with } c_{p,\kappa}^{\mathrm{FvML}} := \frac{(\kappa/2)^{\frac{p}{2}-1}}{\sqrt{\pi}\,\Gamma(\frac{p-1}{2})\mathcal{I}_{\frac{p}{2}-1}(\kappa)}, \tag{2.2}$$

where $\mathcal{I}_\nu(\cdot)$ is the order-ν modified Bessel function of the first kind and $\kappa(>0)$ is a *concentration* parameter (the larger the value of κ, the more concentrated about $\boldsymbol{\theta}$ the distribution is on $\mathcal{S}^{p-1}$); see Mardia and Jupp (2000) for further details.

In this section, we will actually restrict to rotationally symmetric distributions that are absolutely continuous (with respect to the surface area measure on $\mathcal{S}^{p-1}$) and for which the corresponding densities, in the spirit of FvML distributions, involve a concentration parameter. More precisely, we consider densities of the form

$$\mathbf{x} \mapsto c_{p,\kappa,f}\, f(\kappa\, \mathbf{x}'\boldsymbol{\theta}), \qquad \mathbf{x} \in \mathcal{S}^{p-1}, \tag{2.3}$$

where the location parameter $\boldsymbol{\theta}$ belongs to $\mathcal{S}^{p-1}$, the concentration parameter $\kappa(>0)$ plays the same role as for FvML distributions, and the function $f : \mathbb{R} \to \mathbb{R}^+$ satisfies $f(0) = 1$ and admits a positive derivative at 0 ($f'(0) > 0$). These restrictions on f guarantee identifiability of $\boldsymbol{\theta}$, κ and f. Irrespective of f, the boundary value $\kappa = 0$ corresponds to the uniform distribution over $\mathcal{S}^{p-1}$. Finally, it is well-known that, if $\mathbf{X}$ has density (2.3), then $\mathbf{X}'\boldsymbol{\theta}$ has density

$$t \mapsto c_{p,\kappa,f}(1-t^2)^{(p-3)/2} f(\kappa t)\, \mathbb{I}[t \in [-1,1]]$$

(throughout $\mathbb{I}[A]$ stands for the indicator function of the set or condition A). This is compatible with the cumulative distribution functions in (2.1)-(2.2), and shows that

$$c_{p,\kappa,f} = \Big(\int_{-1}^{1} (1-t^2)^{(p-3)/2} f(\kappa t)\, dt \Big)^{-1}.$$

To address the high-dimensional case, we will consider triangular arrays of observations $\mathbf{X}_{ni}$, $i = 1, \ldots, n$, $n = 1, 2, \ldots$ where the random vectors $\mathbf{X}_{ni}$, $i = 1, \ldots, n$ take values in $\mathcal{S}^{p_n-1}$. More specifically, for any $\boldsymbol{\theta}_n \in \mathcal{S}^{p_n-1}$, $\kappa_n > 0$ and f as above, we will denote as $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$ the hypothesis under which $\mathbf{X}_{ni}$, $i = 1, \ldots, n$ are mutually independent and share the common density $\mathbf{x} \mapsto c_{p_n,\kappa_n,f}\, f(\kappa_n\, \mathbf{x}'\boldsymbol{\theta}_n)$. In line with our interpretation of concentration parameters, larger values of κ_n provide increasingly severe deviations from the null of uniformity, which is obtained as κ_n goes to zero. Denoting the null hypothesis as $\mathrm{P}^{(n)}_0$, it is then natural to wonder whether or not "appropriately small" sequences κ_n make $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$ and $\mathrm{P}^{(n)}_0$ mutually contiguous. The following result answers this question (see Appendix A for a proof).

THEOREM 2.1. *Let (p_n) be a sequence in $\{2, 3, \ldots\}$. Let $(\boldsymbol{\theta}_n)$ be an arbitrary sequence with $\boldsymbol{\theta}_n \in \mathcal{S}^{p_n-1}$ for all n, (κ_n) be a positive sequence such that $\kappa_n^2 = O(\frac{p_n}{n})$, and assume that f is twice differentiable in 0. Then, the sequence of alternative hypotheses $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$ and the null sequence $\mathrm{P}^{(n)}_0$ are mutually contiguous.*

This contiguity result covers both the low- and high-dimensional cases. In the low-dimensional case, the usual parametric rate $\kappa_n \sim 1/\sqrt{n}$ provides contiguous alternatives, which implies that, irrespective of f, there exist no consistent tests for $\mathcal{H}_{0n} : \{\mathrm{P}^{(n)}_0\}$ against $\mathcal{H}_{1n} : \{\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}\}$ if $\kappa_n = \tau/\sqrt{n}$, $\tau > 0$. The high-dimensional case is more interesting. First, we stress that the contiguity result in Theorem 2.1 is universal, hence in particular applies when (a) $p_n/n \to c$ for some $c \in (0, \infty)$ or (b) $p_n/n \to \infty$. Interestingly, the result shows that contiguity in cases (a)-(b) can be achieved for sequences (κ_n) that do not converge to zero : a constant sequence (κ_n) ensures contiguity in case (a), whereas contiguity in case (b) may even be obtained for a sequence (κ_n) that converges to infinity in a suitable way. In both cases, there then exist no consistent tests for $\mathcal{H}_{0n} : \{\mathrm{P}^{(n)}_0\}$ against the corresponding sequence of alternatives $\mathcal{H}_{1n} : \{\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}\}$, despite the fact that the sequences (κ_n) are not $o(1)$. This may be puzzling at first since such sequences are expected to lead to severe alternatives to uniformity; it actually makes sense, however, that the fast increase of the dimension p_n, despite the favorable sequences (κ_n), makes the problem difficult enough to prevent the existence of consistent tests.

The next result states that the model considered is *Locally and Asymptotically Normal (LAN)*, with contiguity rate $\kappa_n = O(\sqrt{p_n/n})$. As Theorem 2.1, the result covers both the low- and high-dimensional cases. In low dimensions, it shows that the usual parametric contiguity rate $\kappa_n \sim 1/\sqrt{n}$ is obtained. More importantly, it provides to the best of our knowledge, the first instance of the LAN structure in high dimensions (see Appendix A for a proof).

THEOREM 2.2. *Let (p_n) be a sequence in $\{2, 3, \ldots\}$. Let $(\boldsymbol{\theta}_n)$ be an arbitrary sequence with $\boldsymbol{\theta}_n \in \mathcal{S}^{p_n-1}$ for all n, $\kappa_n = \tau_n\sqrt{p_n/n}$, where the positive sequence (τ_n) is $O(1)$ but not $o(1)$, and assume that f is twice differentiable in 0. Let*

$$\Delta^{(n)}_{\boldsymbol{\theta}_n,f} := \sqrt{np_n}\, f'(0)\bar{\mathbf{X}}_n'\boldsymbol{\theta}_n, \quad \textit{with } \bar{\mathbf{X}}_n := \frac{1}{n}\sum_{i=1}^n \mathbf{X}_{ni}, \tag{2.4}$$

and $\Gamma_f := (f'(0))^2$*. Then, as* $n \to \infty$ *under* $\mathrm{P}_0^{(n)}$*, we have that*

$$\text{(i)}\ \log \frac{d\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}}{d\mathrm{P}^{(n)}_0} = \tau_n \Delta^{(n)}_{\boldsymbol{\theta}_n,f} - \frac{1}{2}\Gamma_f \tau_n^2 + o_{\mathrm{P}}(1) \tag{2.5}$$

and that (ii) $\Delta^{(n)}_{\boldsymbol{\theta}_n,f}$ *is asymptotically normal with zero mean and variance* Γ_f*. In other words, the model* $\{\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa,f} : \kappa \geq 0\}$ *is locally and asymptotically normal (LAN) at* $\kappa = 0$ *with* central sequence $\Delta^{(n)}_{\boldsymbol{\theta}_n,f}$, Fisher information Γ_f, *and* contiguity rate $\sqrt{p_n/n}$.

As usual, local asymptotic normality paves the way to the construction of (locally and asymptotically) optimal tests. The corresponding asymptotic powers provide the natural benchmark to evaluate the performance of the Rayleigh test, both in the low- and high-dimensional cases. These points are addressed in the next section.

3. Optimal testing and performances of the Rayleigh test against contiguous alternatives. Fix $(\boldsymbol{\theta}_n)$, (κ_n) and f as in Theorem 2.2, and consider the problem of testing $\{\mathrm{P}^{(n)}_0\}$ (uniformity over $\mathcal{S}^{p_n-1}$) against $\{\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}\}$. Denoting by Φ the cumulative distribution function of the standard normal, the test $\phi^{(n)}_{\boldsymbol{\theta}_n}$ that rejects the null whenever

$$\frac{\Delta^{(n)}_{\boldsymbol{\theta}_n,f}}{\sqrt{\Gamma_f}} = \sqrt{np_n}\, \bar{\mathbf{X}}_n' \boldsymbol{\theta}_n > z_\alpha, \qquad \text{with } z_\alpha := \Phi^{-1}(1-\alpha), \tag{3.1}$$

is locally and asymptotically most powerful at level α. This test does not depend on f (whence the notation $\phi^{(n)}_{\boldsymbol{\theta}_n}$), hence is also locally and asymptotically most powerful at level α when testing $\{\mathrm{P}^{(n)}_0\}$ against $\cup_{f\in\mathcal{F}}\{\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}\}$, where $\mathcal{F}$ stands for the collection of functions $f : \mathbb{R} \to \mathbb{R}^+$ that are twice differentiable at 0 and satisfy $f(0) = 1$ and $f'(0) > 0$.

Le Cam's third lemma implies that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$, with $\kappa_n = \tau\sqrt{p_n/n}$, $\Delta^{(n)}_{\boldsymbol{\theta}_n,f}$ is asymptotically normal with mean $\Gamma_f\tau$ and variance Γ_f. Consequently, the corresponding asymptotic power of $\phi^{(n)}_{\boldsymbol{\theta}_n}$ is

$$\lim_{n\to\infty} \mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}\left[\frac{\Delta^{(n)}_{\boldsymbol{\theta}_n,f}}{\sqrt{\Gamma_f}} > z_\alpha\right] = 1 - \Phi\big(z_\alpha - f'(0)\tau\big). \tag{3.2}$$

This applies both to the low- and high-dimensional cases.

Of course, the "oracle" test above is infeasible since $\boldsymbol{\theta}_n$ is unspecified in practice (all the more so that, under the null of uniformity, $\boldsymbol{\theta}_n$ is not identifiable). Of course, it is natural to replace $\boldsymbol{\theta}_n$ with an estimator, such as the so-called *spherical mean* $\hat{\boldsymbol{\theta}}_n = \bar{\mathbf{X}}_n/\|\bar{\mathbf{X}}_n\|$. The resulting test rejects the null of uniformity for large values of

$$\frac{\Delta^{(n)}_{\hat{\boldsymbol{\theta}}_n,f}}{\sqrt{\Gamma_f}} = \sqrt{np_n}\,\bar{\mathbf{X}}_n'\hat{\boldsymbol{\theta}}_n = \sqrt{np_n}\,\|\bar{\mathbf{X}}_n\|,$$

or equivalently, for large values of the Rayleigh test statistic $R_n = np_n\|\bar{\mathbf{X}}_n\|^2$. This suggests that the Rayleigh test may be locally and asymptotically most powerful at level α when testing $\{\mathrm{P}^{(n)}_0\}$ against $\cup_{f\in\mathcal{F}}\{\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}\}$ We now investigate whether this is the case or not, both in the low- and high-dimensional cases.

We start with the low-dimensional case. We actually restrict to the fixed-p case, since it is needed that $p_n \to p$ for the Rayleigh test statistic to have a non-trivial asymptotic distribution under contiguous alternatives. Consider then the contiguous alternatives $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$, with $\kappa_n = \tau_n\sqrt{p/n}$, where the sequence (τ_n) converges to some $\tau \in (0,\infty)$; compare with the local alternatives from Theorem 2.2. Denoting by $\chi^2_k(\delta)$ the non-central chi-square distribution with k degrees of freedom and non-centrality parameter δ, Le Cam's third lemma allows to show that, as $n\to\infty$ under the sequence of alternatives above,

$$R_n \xrightarrow{\mathcal{D}} \chi^2_p\big((f'(0)\tau)^2\big), \tag{3.3}$$

where $\xrightarrow{\mathcal{D}}$ denotes weak convergence (for the sake of completeness, we provide a proof in Section 1 of the supplementary article Cutting, Paindaveine and Verdebout (2015)). Consequently, the asymptotic power of the Rayleigh test under this sequence of alternatives is

$$\mathrm{P}\big[Y > \Psi_p^{-1}(1-\alpha)\big], \qquad \text{with } Y \sim \chi^2_p\big((f'(0)\tau)^2\big), \tag{3.4}$$

where $\Psi_p(\cdot)$ denotes the cumulative distribution function of the χ^2_p distribution. This shows that, in the fixed-p case, the Rayleigh test has non-trivial asymptotic powers against the contiguous alternatives from Theorem 2.2, but that it is not locally and asymptotically most powerful at level α (it can indeed be checked that, irrespective of $f'(0)(>0)$, $\tau(>0)$, and $\alpha(\in(0,1))$, the powers in (3.4) are strictly smaller than those in (3.2)).

The story for the high-dimensional case is very different, as it can be guessed from the fixed-p result in (3.3), by adopting the following heuristic reasoning. In view of (3.3), we have that, as $n\to\infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$, with $\kappa_n = \tau_n\sqrt{p/n}$, where the sequence (τ_n) converges to some $\tau \in (0,\infty)$,

$$R_n^{\mathrm{St}} = \frac{R_n - p}{\sqrt{2p}} \xrightarrow{\mathcal{D}} \frac{\chi^2_1\big((f'(0)\tau)^2\big) - 1}{\sqrt{2p}} + \frac{\chi^2_{p-1} - (p-1)}{\sqrt{2p}},$$

where both chi-square terms are independent. When both n and p are large, it is therefore expected that, under the same sequence of alternatives,

$$R_n^{\rm St} \approx \mathcal{N}\bigg(\frac{(f'(0)\tau)^2}{\sqrt{2p}}\,, 1 + \frac{2(f'(0)\tau)^2}{p}\bigg),$$

where $Z_n \approx \mathcal{L}$ means "the distribution of Z_n is close to $\mathcal{L}$". Thus, in the high-dimensional case (where $p = p_n \to \infty$), $R_n^{\rm St}$ is expected to be standard normal under these alternatives, which would imply that the Rayleigh test has asymptotic powers equal to the nominal level α (Theorem 1.1 indeed states that the asymptotic null distribution of $R_n^{\rm St}$ is also standard normal).

Thanks to our high-dimensional LAN result in Theorem 2.2, this heuristics can be confirmed rigorously. Theorem 2.2 readily yields that, as $n \to \infty$ under ${\rm P}_0^{(n)}$, and with $\kappa_n = \tau_n\sqrt{p_n/n}$ (where τ_n is $O(1)$),

$$\begin{aligned}
{\rm Cov}\bigg[R_n^{\rm St}, \log \frac{d{\rm P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}}{d{\rm P}_0^{(n)}}\bigg] &= {\rm Cov}\big[R_n^{\rm St}, \Delta^{(n)}_{\boldsymbol{\theta}_n,f}\big]\tau + o(1) \\
&= \frac{\sqrt{2}p_n}{n^{3/2}} f'(0)\tau \sum_{i=1}^n \sum_{1\leq k<\ell\leq n} {\rm E}[(\mathbf{X}'_{ni}\boldsymbol{\theta}_n)(\mathbf{X}'_{nk}\mathbf{X}_{n\ell})] + o(1) = o(1),
\end{aligned}$$

so that Le Cam's third lemma implies that $R_n^{\rm St}$ remains — (n,p)-universally — asymptotically standard normal under ${\rm P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$, where $\kappa_n = \tau_n\sqrt{p_n/n}$, with $\tau_n = O(1)$. This confirms that, unlike in the low-dimensional case, the Rayleigh test does not show any power under the high-dimensional contiguous alternatives from Theorem 2.2. Equivalently, this shows that the Rayleigh test fails to be *rate-consistent* in high dimensions.

This of course raises natural questions on the Rayleigh test : in the high-dimensional case, can this test asymptotically detect non-trivial alternatives? If the answer is positive, how close are these alternatives to the contiguous alternatives in Theorem 2.2? We thoroughly address these questions in the next section.

4. Asymptotic non-null behavior of the Rayleigh test. In this section, we derive the (n,p)-asymptotic distribution of the Rayleigh test under distributions that encompass those considered in Section 2, namely under *general* rotationally symmetric distributions. We do not require that the rotationally symmetric alternatives considered are absolutely continuous with respect to the surface area measure on the unit sphere, nor that they involve a concentration parameter κ. Yet one of our objectives is to interpret the results we derive in this section in the light of the contiguity/LAN/rate-consistency results obtained in Sections 2 and 3.

More specifically, the sequences of alternatives we consider in this section are described by triangular arrays of observations $\mathbf{X}_{ni}$, $i = 1, \ldots, n$, $n = 1, 2, \ldots$ such that, for any n, $\mathbf{X}_{n1}, \mathbf{X}_{n2}, \ldots, \mathbf{X}_{nn}$ are mutually independent and share a common rotationally symmetric distribution on $\mathcal{S}^{p_n-1}$. We will denote by $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n, F_n}$ the corresponding hypothesis in the case where $\mathbf{X}_{ni}$ is rotationally symmetric about $\boldsymbol{\theta}_n$ and $\mathbf{X}'_{ni}\boldsymbol{\theta}_n$ has cumulative distribution function F_n. Since the Rayleigh test statistic is invariant under rotations, we will, without any loss of generality, restrict to the case for which $\boldsymbol{\theta}_n$, for any n, coincides with the first vector of the canonical basis of $\mathbb{R}^{p_n}$. The corresponding sequence of hypotheses will then simply be denoted as $\mathrm{P}^{(n)}_{F_n}$.

Under the null of uniformity $\mathrm{P}^{(n)}_0$, the test statistic R^{St}_n in (1.1) has mean zero and variance $\frac{n-1}{n}(\to 1)$. Rotationally symmetric alternatives are expected to have an impact on the asymptotic mean and variance of R^{St}_n. Exact values are obtained in the following result (see Appendix B.1 for a proof).

Proposition 4.1. *Under $\mathrm{P}^{(n)}_{F_n}$, the mean and variance of R^{St}_n are given by*

$$\mathrm{E}[R^{\mathrm{St}}_n] = \frac{(n-1)\sqrt{p_n}}{\sqrt{2}}\, e^2_{n1}$$

and

$$\mathrm{Var}[R^{\mathrm{St}}_n] = \frac{n-1}{n}\Big(p_n \tilde{e}^2_{n2} + \frac{p_n}{p_n-1}\, f^2_{n2} + 2(n-1)p_n e^2_{n1}\tilde{e}_{n2}\Big),$$

where we let $e_{n\ell} := \mathrm{E}[(\mathbf{X}'_{ni}\boldsymbol{\theta}_n)^\ell]$ and $\tilde{e}_{n\ell} := \mathrm{E}[(\mathbf{X}'_{ni}\boldsymbol{\theta}_n - e_{n1})^\ell]$ denote the ℓth-order non-central and central moments associated with F_n and where $f_{n\ell} := \mathrm{E}[(1-(\mathbf{X}'_{ni}\boldsymbol{\theta}_n)^2)^{\ell/2}]$.

Note that, under $\mathrm{P}^{(n)}_0$, we have $e_{1n} = 0$ and $\tilde{e}_{2n} = e_{2n} = 1/p_n$, which provides the null values of $\mathrm{E}[R^{\mathrm{St}}_n]$ and $\mathrm{Var}[R^{\mathrm{St}}_n]$ stated above. Now, as soon as p_n goes to infinity with n, the asymptotic variance in the previous result is equivalent to the simpler quantity

$$\sigma^2_n := p_n \tilde{e}^2_{n2} + 2np_n e^2_{n1}\tilde{e}_{n2} + f^2_{n2}, \tag{4.1}$$

which we use in the sequel. Parallel to to the null case (see Theorem 1.1), it can be expected that, when properly standardized by using the mean and variance in Proposition 4.1 (or the asymptotically equivalent variance σ_n^2), the Rayleigh test statistic will be asymptotically standard normal under a broad class of rotationally symmetric alternatives. This is confirmed in the following result (see Appendix B.2 for a proof).

THEOREM 4.1. *Let (p_n) be a sequence of positive integers converging to ∞ as $n \to \infty$. Assume that the sequence $(\mathrm{P}_{F_n}^{(n)})$ is such that, as $n \to \infty$,*

$$(i)\ \min\Big(\frac{p_n \tilde{e}_{n2}^2}{f_{n2}^2}, \frac{\tilde{e}_{n2}}{n e_{n1}^2}\Big) = o(1), \quad (ii)\ \tilde{e}_{n4}/\tilde{e}_{n2}^2 = o(n), \quad \textit{and} \quad (iii)\ f_{n4}/f_{n2}^2 = o(n).$$

Then, under $\mathrm{P}_{F_n}^{(n)}$,

$$\frac{R_n^{\mathrm{St}} - \mathrm{E}[R_n^{\mathrm{St}}]}{\sigma_n} = \frac{\sqrt{2p_n}}{n\sigma_n} \sum_{1 \le i < j \le n} \big(\mathbf{X}_{ni}' \mathbf{X}_{nj} - e_{n1}^2\big)$$

converges weakly to the standard normal distribution as $n \to \infty$.

This result applies under very mild assumptions, that in particular do not impose absolute continuity nor any other regularity conditions. The only structural assumptions are the conditions (i)-(iii) above. These, however, may only be violated for rotationally symmetric distributions that are very far from the null of uniformity (hence, for alternatives under which there is in practice no need for a test of uniformity). Indeed, a necessary — yet far from sufficient — condition for (i)-(iii) to be violated is that $\mathbf{X}_{ni}' \boldsymbol{\theta}_n$ converges in probability to some constant $c(\in [-1,1])$, which is also quite pathological in the high-dimensional context considered. Moreover, if one restricts to FvML rotationally symmetric distributions (see Section 2), then (i)-(iii) *always* hold, that is, they hold without any constraint on the concentration κ_n nor on the way the dimension p_n goes to infinity with n (the proof of this statement is very lengthy and requires original results on modified Bessel functions ratios, hence is provided in the supplementary article; see Section 2 of Cutting, Paindaveine and Verdebout (2015)). This shows that the non-null asymptotic result in Theorem 4.1, parallel to the null one in Theorem 1.1, may be considered universal.

Convergence rates in the asymptotic normality in Theorem 4.1 can be obtained by deriving appropriate Berry-Esseen bounds. This relatively easily follows from a classical result from Heyde and Brown (1970) and the estimates provided in the proof of Theorem 4.1. More precisely, we establish the following result in Appendix B.3.

Theorem 4.2. *Let (p_n) be a sequence of positive integers converging to ∞ as $n \to \infty$. Assume that the sequence $(\mathrm{P}^{(n)}_{F_n})$ satisfies the assumptions of Theorem 4.1. Then, there exist a constant C and a positive sequence (s_n) converging to one as $n \to \infty$ such that*

$$\sup_{z\in\mathbb{R}}\bigg|\mathrm{P}^{(n)}_{F_n}\bigg[\frac{R_n^{\mathrm{St}}-\mathrm{E}[R_n^{\mathrm{St}}]}{\sigma_n}\leq s_n z\bigg]-\Phi(z)\bigg| \leq C\bigg(\min\Big(\frac{p_n\tilde{e}^2_{n2}}{f^2_{n2}},\frac{\tilde{e}_{n2}}{ne^2_{n1}}\Big)+\frac{\tilde{e}_{n4}}{n\tilde{e}^2_{n2}}+\frac{f_{n4}}{nf^2_{n2}}+\frac{1}{p_n}\bigg)^{1/5}$$

for n large enough (the exact expression of s_n is given in the appendix; see (B.7)).

Theorems 4.1 and 4.2 allow to compute the asymptotic power of the Rayleigh test under appropriate sequences of alternatives. As mentioned above, the null of uniformity $\mathcal{H}_{0n}$ provides $e_{1n} = 0$ and $\tilde{e}_{2n} = 1/p_n$. Here, we therefore consider "local" departures from uniformity of the form $\mathcal{H}_{1n} : \{\mathrm{P}^{(n)}_{F_n} : e_{n1} = 0 + \nu_n\tau_1, \tilde{e}_{n2} = (1/p_n) + \xi_n\tau_2\}\cdot$ A natural question is : in the high-dimensional case, what are the rates ν_n and ξ_n at which the Rayleigh test can discriminate between the null and these local alternatives? The following result is also proved in Appendix B.3.

Theorem 4.3. *Let (p_n) be a sequence of positive integers converging to ∞ as $n \to \infty$. Consider a sequence $(\mathrm{P}^{(n)}_{F_n})$ that satisfies the assumptions of Theorem 4.1 and provides*

$$e_{n1} = \frac{\tau_1}{n^{1/2}p_n^{1/4}} + o\bigg(\frac{1}{n^{1/2}p_n^{1/4}}\bigg) \qquad \text{and} \qquad \tilde{e}_{n2} = \frac{1}{p_n} + o\Big(\frac{1}{p_n}\Big). \tag{4.2}$$

Then, under $\mathrm{P}^{(n)}_{F_n}$, the asymptotic power of the Rayleigh test is given by $1-\Phi\big(z_\alpha-\tau_1^2/\sqrt{2}\big)$.

Clearly, it is of interest to investigate how severe are the local alternatives in (4.2) compared to the contiguous alternatives in Theorem 2.2, under which the Rayleigh test does not show any power in the high-dimensional case; see Section 3. To do so, note that, as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$, with $\kappa_n = \tau_n\sqrt{p_n/n}$, where the positive sequence (τ_n) is $o(\sqrt{n})$, we have

$$\begin{aligned}
e_{1n} &= c_{p_n,\kappa_n,f}\int_{-1}^{1} s(1-s^2)^{(p_n-3)/2} f(\kappa_n s)\, ds \\
&= \Big(\frac{c_{p_n}}{c_{p_n,\kappa_n,f}}\Big)^{-1}\frac{c_{p_n}}{\kappa_n}\int_{-1}^{1}(1-s^2)^{(p_n-3)/2}\,\kappa_n s f(\kappa_n s)\, ds \\
&= \bigg(1+\frac{\kappa_n^2}{2p_n}f''(0)+o\Big(\frac{\kappa_n^2}{p_n}\Big)\bigg)^{-1}\bigg(\frac{\kappa_n}{p_n}f'(0)+o\Big(\frac{\kappa_n}{p_n}\Big)\bigg), \qquad (4.3)
\end{aligned}$$

where we used twice Lemma A.1. Under the same sequence of alternatives, we obtain similarly

$$\begin{aligned} e_{2n} &= c_{p_n,\kappa_n,f} \int_{-1}^{1} s^2(1-s^2)^{(p_n-3)/2} f(\kappa_n s)\, ds \\ &= \Big(\frac{c_{p_n}}{c_{p_n,\kappa_n,f}}\Big)^{-1} \frac{c_{p_n}}{\kappa_n^2} \int_{-1}^{1} (1-s^2)^{(p_n-3)/2} (\kappa_n s)^2 f(\kappa_n s)\, ds \\ &= \Big(1+\frac{f''(0)\kappa_n^2}{2p_n} + o\Big(\frac{\kappa_n^2}{p_n}\Big)\Big)^{-1} \Big(\frac{1}{p_n} + o\Big(\frac{1}{p_n}\Big)\Big). \end{aligned} \tag{4.4}$$

The contiguous alternatives in Theorem 2.2 are of the form $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$, with $\kappa_n = \tau\sqrt{p_n/n}$ (for the sake of simplicity, we consider here the case $\tau_n = \tau$ for all n). For such alternatives, (4.3)-(4.4) provide

$$e_{n1} = \frac{f'(0)\tau}{\sqrt{np_n}} + o\Big(\frac{1}{\sqrt{np_n}}\Big) \qquad \text{and} \qquad \tilde{e}_{n2} = \frac{1}{p_n} + o\Big(\frac{1}{p_n}\Big), \tag{4.5}$$

which clearly corresponds to (slightly) less severe deviations from the null than the local alternatives in (4.2).

Interestingly, we might have guessed that, in the high-dimensional case, the local alternatives in (4.2) are those that can be detected by the Rayleigh test. Recall indeed that heuristic arguments in Section 3 suggested that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$, with $\kappa_n = \tau\sqrt{p/n}$, the distribution of R_n^{St} is close to $\mathcal{N}\big(\tau^2/\sqrt{2p}, 1+2(f'(0)\tau)^2/p\big)$ for n large. Consequently, to obtain, in the high-dimensional case, an asymptotic non-null distribution that is different from the limiting null (standard normal) one, we need to consider alternatives of the form $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$, with $\kappa_n = \tau p_n^{3/4}/\sqrt{n}$, under which, the distribution of R_n^{St} should approximately be $\mathcal{N}\big(\tau^2/\sqrt{2}, 1\big)$ for n large. At least if $p_n = o(n^2)$ (a constraint that is superfluous in the FvML case), these alternatives, in view of (4.3)-(4.4), lead to

$$e_{n1} = \frac{f'(0)\tau}{n^{1/2}p_n^{1/4}} + o\Big(\frac{1}{n^{1/2}p_n^{1/4}}\Big) \qquad \text{and} \qquad \tilde{e}_{n2} = \frac{1}{p_n} + o\Big(\frac{1}{p_n}\Big), \tag{4.6}$$

which coincides with the local alternatives in (4.2).

5. A Monte-Carlo study.

In this section, we present the results of a Monte Carlo study we conducted to check the validity of our asymptotic results. We performed two simulations. In the first one, we generated independent random samples of the form

$$\mathbf{X}_{i;j}^{(\ell)} \quad i=1,\ldots,n, \quad j=1,2, \quad \ell=0,1,2,3,4. \tag{5.1}$$

For $\ell = 0$, the common distribution of the $\mathbf{X}_{i;j}^{(\ell)}$'s is the uniform distribution on the unit sphere $\mathcal{S}^{p-1}$, while, for $\ell > 0$, the $\mathbf{X}_{i;j}^{(\ell)}$'s have an FvML distribution on $\mathcal{S}^{p-1}$ with location $\boldsymbol{\theta} = (1, 0, \ldots, 0)' \in \mathbb{R}^p$ and concentration $\kappa_j^{(\ell)}$, with

$$\kappa_1^{(\ell)} = 0.6\ell \sqrt{\frac{p}{n}} \quad \text{and} \quad \kappa_2^{(\ell)} = 0.6\ell \, \frac{p^{3/4}}{\sqrt{n}} \cdot$$

In the second simulation, we considered again independent random samples of the form (5.1), still with $\mathbf{X}_{i;j}^{(0)}$'s that are uniform over $\mathcal{S}^{p-1}$. Here, however, the $\mathbf{X}_{i;j}^{(\ell)}$'s, for $\ell = 1, 2, 3, 4$, are rotationally symmetric with location $\boldsymbol{\theta} = (1, 0, \ldots, 0)' \in \mathbb{R}^p$ and are such that the $\boldsymbol{\theta}'\mathbf{X}_{i;j}^{(\ell)}$'s are beta with mean $e_{1;j}^{(\ell)}$ and variance $\tilde{e}_{2;j} = 1/p$, where we let

$$e_{1;1}^{(\ell)} = \frac{0.6\ell}{\sqrt{np}} \quad \text{and} \quad e_{1;2}^{(\ell)} = \frac{0.6\ell}{n^{1/2}p^{1/4}} \cdot$$

In both simulations, the value $\ell = 0$ corresponds to the null hypothesis of uniformity, while $\ell = 1, 2, 3, 4$ provide increasingly severe alternatives. The case $j = 1$ corresponds to the contiguous alternatives under which local asymptotic normality holds (see Theorem 2.2 and (4.5)), whereas $j = 2$ is associated with the alternatives under which the Rayleigh test shows non-trivial asymptotic powers in the high-dimensional setup (see (4.2) and the discussion above (4.6)).

For any $(n, p) \in C \times C$, with $C := \{30, 100, 400\}$, any $j \in \{1, 2\}$, and any $\ell \in \{0, 1, 2, 3, 4\}$, we generated $M = 2{,}500$ independent random samples $\mathbf{X}_{i;j}^{(\ell)}$, $i = 1, \ldots, n$, as described above, and evaluated the rejection frequencies of the following two tests, conducted at nominal level 5%: (i) the oracle test $\phi_{\boldsymbol{\theta}_n}^{(n)}$ in (3.1) and (ii) the high-dimensional Rayleigh test (that is, the test that rejects the null of uniformity whenever the statistic R^{St} in (1.1) exceeds the 95%-quantile of the standard normal distribution). Rejection frequencies are plotted in Figures 1 and 2, for FvML and "beta" alternatives, respectively. In each figure, we also plot the corresponding asymptotic powers, obtained from (3.2), Theorem 4.3, and the facts that (i) the high-dimensional Rayleigh test does not show any asymptotic power against ($j = 1$)-alternatives and that (ii) the oracle test is consistent against ($j = 2$)-alternatives.

Clearly, for both simulations, rejection frequencies match extremely well the corresponding asymptotic powers, irrespective of the tests and types of alternatives considered (the only possible exception is the oracle test under ($\ell = 1, j = 1$)-alternatives; this, however, is obviously only a consequence of the lack of continuity of the corresponding asymptotic power curve). Quite remarkably, this agreement is also very good for relatively small sample size n and dimension p. Beyond validating our asymptotic results, this Monte Carlo study therefore also shows that these results are relevant for practical values of n and p.

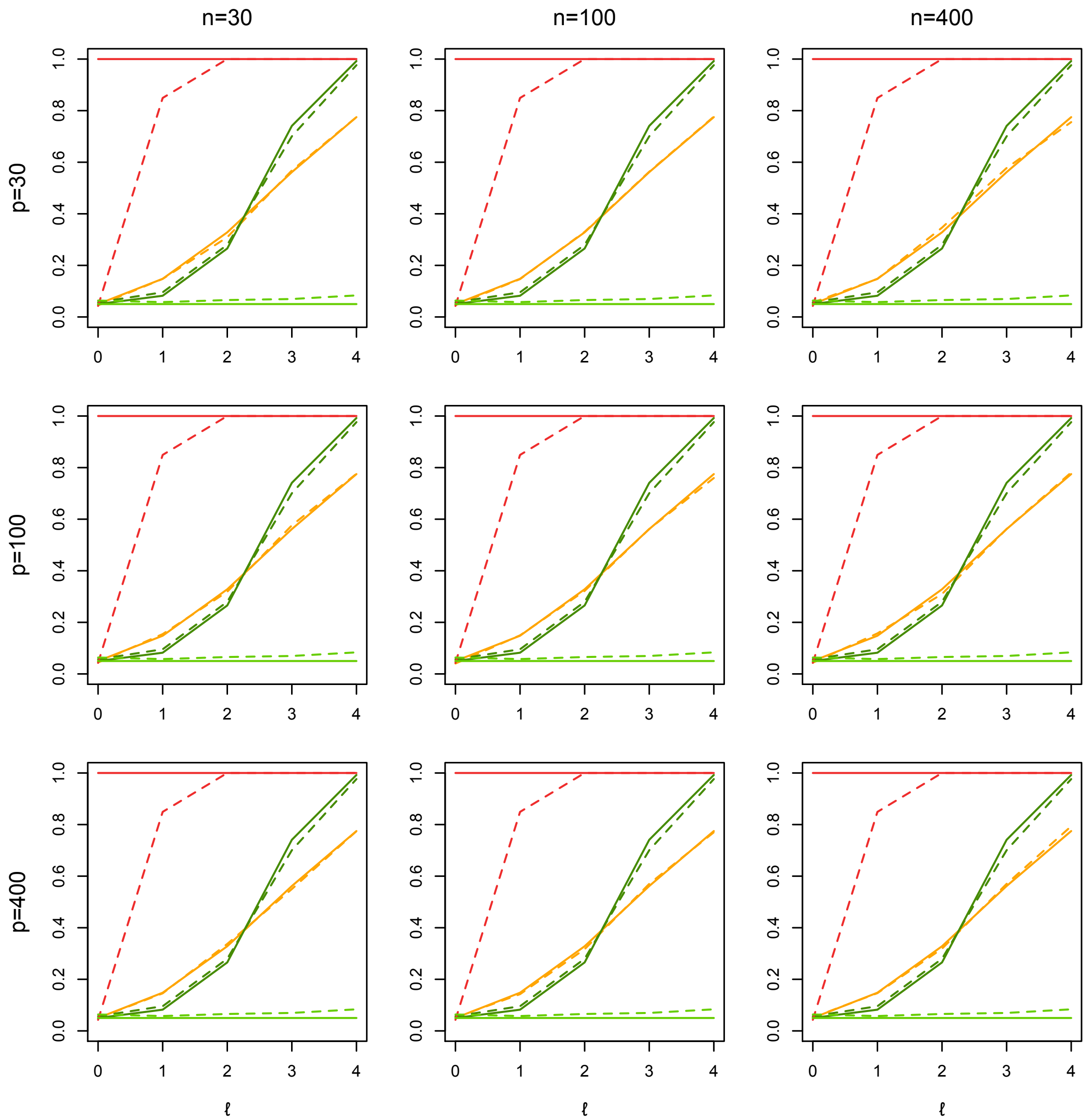

Fig 1. *Rejection frequencies (dashed) and asymptotic powers (solid), under the null of uniformity over the p-dimensional unit sphere ($\ell = 0$) and increasingly severe FvML alternatives ($\ell = 1, 2, 3, 4$), of the oracle test from Section 3 (red/orange) and the high-dimensional Rayleigh test (green). Light colors (orange and light green) are asociated with contiguous alternatives under which LAN holds, whereas dark colors (red and dark green) correspond to the more severe alternatives under which the Rayleigh test shows non-trivial (n, p)-asymptotic powers; see Section 5 for details.*

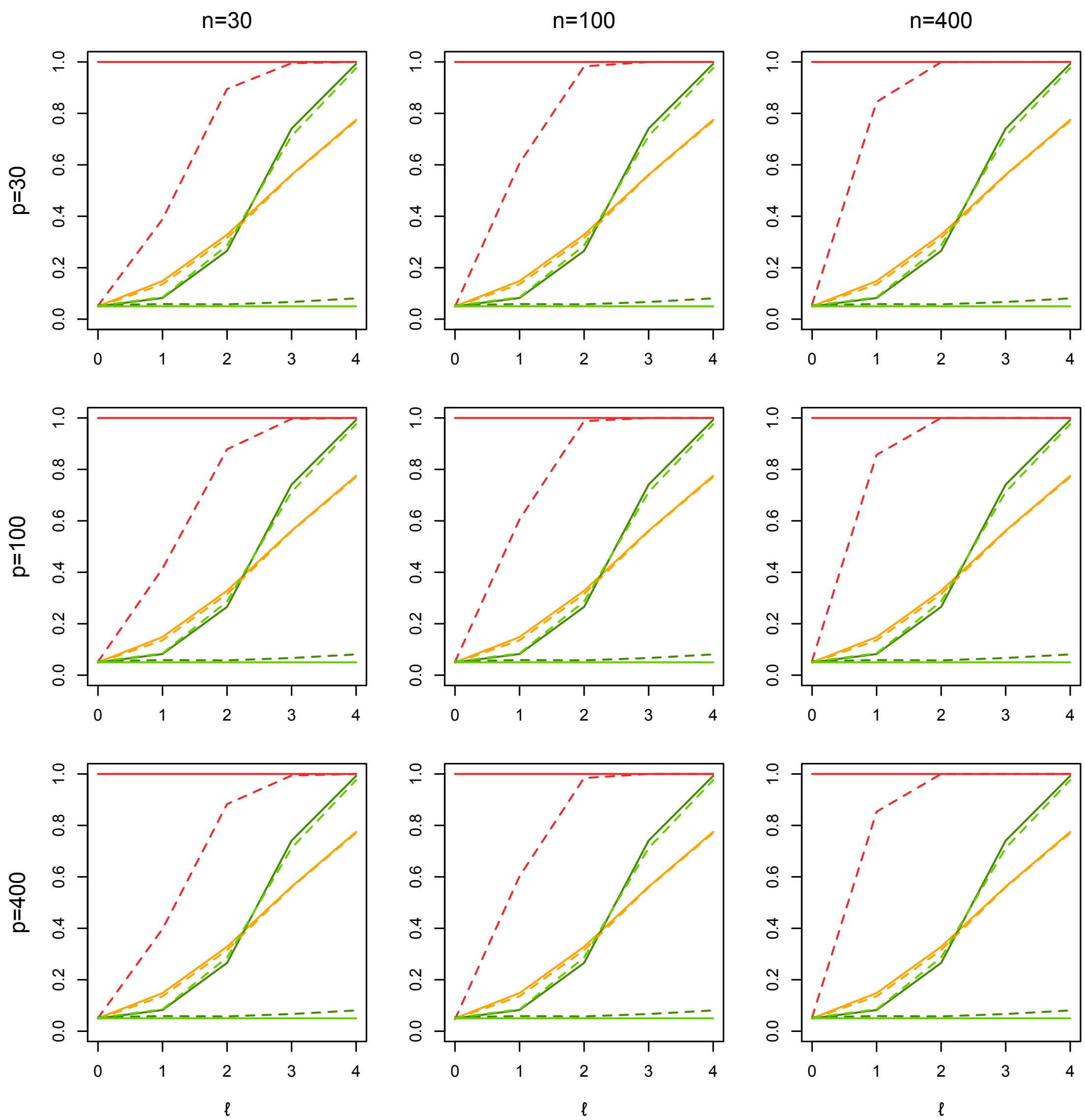

FIG 2. *Rejection frequencies (dashed) and asymptotic powers (solid), under the null of uniformity over the p-dimensional unit sphere* ($\ell = 0$) *and increasingly severe "beta" rotationally symmetric alternatives* ($\ell = 1, 2, 3, 4$)*, of the oracle test from Section 3 (red/orange) and the high-dimensional Rayleigh test (green). Light colors (orange and light green) are associated with contiguous alternatives under which LAN holds, whereas dark colors (red and dark green) correspond to the more severe alternatives under which the Rayleigh test shows non-trivial* (n, p)*-asymptotic powers; see Section 5 for details.*

6. Conclusions. We summarize the results derived in the paper. In the problem of testing uniformity on the unit sphere against rotationally symmetric alternatives, we identified contiguous alternatives and showed that the model considered is locally and asymptotically normal (LAN) in the vicinity of the null hypothesis. This LAN structure allows to determine the maximal local asymptotic powers that can be achieved and to define "oracle" tests achieving these. Oracle tests are infeasible since they require the location parameter $\boldsymbol{\theta}_n$ to be known. When replacing $\boldsymbol{\theta}_n$ with the spherical mean estimator $\bar{\mathbf{X}}_n/\|\bar{\mathbf{X}}_n\|$, the oracle test actually reduces to the Rayleigh test. We thoroughly studied the asymptotic behaviour of the latter test under general rotationally symmetric alternatives.

Throughout, both the low-dimensional and high-dimensional cases are covered and actually lead to very different conclusions. In the low-dimensional case, the contiguity rate is the classical $1/\sqrt{n}$ rate and the Rayleigh test shows non-trivial asymptotic powers against the corresponding alternatives, even though it is not asymptotically optimal. In the high-dimensional case, contiguity rates are of the form $\sqrt{p_n/n}$, irrespective of the speed at which p_n goes to infinity with n. Under such local alternatives, the Rayleigh test has powers equal to the nominal level α, so that the cost of estimating $\boldsymbol{\theta}_n$ is more severe in the high-dimensional case than in the low-dimensional one. While the Rayleigh test is not rate-optimal in high dimensions, we identified less severe local alternatives, with rate $p_n^{3/4}/\sqrt{n}$, that can be detected asymptotically by this test. Simulation results are in remarkable agreement with our asymptotic results, even for moderate values of n and p.

APPENDIX A: PROOFS FOR SECTION 2

The proofs of Theorem 2.1 and Theorem 2.2 require the following preliminary result.

LEMMA A.1. *Let $g : \mathbb{R} \to \mathbb{R}$ be twice differentiable in 0. Let κ_n be a positive sequence that is $o(\sqrt{p_n})$ as $n \to \infty$. Then*

$$R_n(g) := c_{p_n} \int_{-1}^{1} (1 - s^2)^{(p_n-3)/2} g(\kappa_n s)\, ds = g(0) + \frac{\kappa_n^2}{2p_n} g''(0) + o\Big(\frac{\kappa_n^2}{p_n}\Big).$$

PROOF OF LEMMA A.1. Write first

$$R_n(g) = g(0) + c_{p_n} \int_{-1}^{1} (1 - s^2)^{(p_n-3)/2} (g(\kappa_n s) - g(0) - \kappa_n s g'(0))\, ds.$$

Using the identity

$$c_{p_n}\int_{-1}^{1} s^2(1-s^2)^{(p_n-3)/2}\,ds=\frac{1}{p_n}$$

and letting $t=\kappa_n s$ then provides

$$R_n(g)=g(0)+\frac{\kappa_n^2}{p_n}\int_{-\kappa_n}^{\kappa_n} h_n(t)\Big(\frac{g(t)-g(0)-tg'(0)}{t^2}\Big)\,dt,$$

where h_n is defined through

$$t\mapsto h_n(t)=\frac{(\frac{t}{\kappa_n})^2(1-(\frac{t}{\kappa_n})^2)^{(p_n-3)/2}}{\int_{-\kappa_n}^{\kappa_n}(\frac{s}{\kappa_n})^2(1-(\frac{s}{\kappa_n})^2)^{(p_n-3)/2}\,ds}\,\mathbb{I}[|t|\leq\kappa_n].$$

By using the fact that $\kappa_n=o(\sqrt{p_n})$, it is easy to check that the h_n's form an *approximate δ-sequence*, in the sense that

$$\int_{-\infty}^{\infty} h_n(t)\,dt=1\quad\forall n\qquad\text{and}\qquad\int_{-\varepsilon}^{\varepsilon} h_n(t)\,dt\to 1$$

for any $\varepsilon>0$. It follows that

$$\begin{aligned}R_n(g) &= g(0)+\frac{\kappa_n^2}{p_n}\lim_{t\to 0}\Big(\frac{g(t)-g(0)-tg'(0)}{t^2}\Big)+o\Big(\frac{\kappa_n^2}{p_n}\Big)\\ &= g(0)+\frac{\kappa_n^2}{2p_n}\lim_{t\to 0}\Big(\frac{g'(t)-g'(0)}{t}\Big)+o\Big(\frac{\kappa_n^2}{p_n}\Big),\end{aligned}$$

where we used L'Hôpital's rule. This yields the result. □

PROOF OF THEOREM 2.1. In this proof, all expectations and variances are taken under the null of uniformity $\mathrm{P}_0^{(n)}$ and all stochastic convergences and o_{P}'s are as $n\to\infty$ under $\mathrm{P}_0^{(n)}$. Consider then the local log-likelihood ratio

$$\begin{aligned}\Lambda_n:=\log\frac{d\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}}{d\mathrm{P}_0^{(n)}} &= \sum_{i=1}^{n}\log\frac{c_{p_n,\kappa_n,f}\,f(\kappa_n\mathbf{X}_{ni}'\boldsymbol{\theta}_n)}{c_{p_n}}\\ &= n\Big(\log\frac{c_{p_n,\kappa_n,f}}{c_{p_n}}+E_{n1}\Big)+\sum_{i=1}^{n}\big(\log f(\kappa_n\mathbf{X}_{ni}'\boldsymbol{\theta}_n)-E_{n1}\big)\\ &=: L_{n1}+L_{n2};\end{aligned}\tag{A.1}$$

throughout, we write $\ell_{f,k}:=(\log f)^k$ and $E_{nk}:=\mathrm{E}\big[\ell_{f,k}(\kappa_n\mathbf{X}_{ni}'\boldsymbol{\theta}_n)\big]$ (E_{nk} actually depends on κ_n, p_n and f, but we simply write E_{nk} to avoid a heavy notation).

Lemma A.1 readily yields

$$\begin{aligned}n\log\frac{c_{p_n,\kappa_n,f}}{c_{p_n}} &= -n\log\Big(c_{p_n}\int_{-1}^{1}(1-s^2)^{(p_n-3)/2}f(\kappa_n s)\,ds\Big)\\ &= -n\log\Big(1+\frac{\kappa_n^2}{2p_n}f''(0)+o\Big(\frac{\kappa_n^2}{p_n}\Big)\Big)=-\frac{n\kappa_n^2}{2p_n}f''(0)+o\Big(\frac{n\kappa_n^2}{p_n}\Big).\end{aligned}\tag{A.2}$$

Similarly, for any positive integer k,

$$E_{nk} = c_{p_n} \int_{-1}^{1} (1-s^2)^{(p_n-3)/2} \ell_{f,k}(\kappa_n s)\, ds = \frac{\kappa_n^2}{2p_n} \ell''_{f,k}(0) + o\Big(\frac{\kappa_n^2}{p_n}\Big). \tag{A.3}$$

Combining (A.2) and (A.3), and using the identity $\ell''_{f,1}(0) = f''(0) - (f'(0))^2$ readily yields

$$L_{n1} = \frac{n\kappa_n^2}{2p_n}\Big(-f''(0) + \ell''_{f,1}(0)\Big) + o\Big(\frac{n\kappa_n^2}{p_n}\Big) = -\frac{n\kappa_n^2}{2p_n}(f'(0))^2 + o\Big(\frac{n\kappa_n^2}{p_n}\Big).$$

Turning to L_{n2}, write

$$L_{n2} = \sqrt{nV_n} \sum_{i=1}^n W_{ni} := \sqrt{nV_n} \sum_{i=1}^n \frac{\log f(\kappa_n \mathbf{X}'_{ni}\boldsymbol{\theta}_n) - E_{n1}}{\sqrt{nV_n}},$$

where we wrote $V_n := \mathrm{Var}\big[\log f(\kappa_n \mathbf{X}'_{ni}\boldsymbol{\theta}_n)\big]$. First note that (A.3) provides

$$nV_n = n\big(E_{n2} - E_{n1}^2\big) = \frac{n\kappa_n^2}{2p_n}\ell''_{f,2}(0) + o\Big(\frac{n\kappa_n^2}{p_n}\Big) = \frac{n\kappa_n^2}{p_n}(f'(0))^2 + o\Big(\frac{n\kappa_n^2}{p_n}\Big), \tag{A.4}$$

which leads to

$$\Lambda_n = -\frac{n\kappa_n^2}{2p_n}(f'(0))^2 + \sqrt{\frac{n\kappa_n^2}{p_n}(f'(0))^2 + o\Big(\frac{n\kappa_n^2}{p_n}\Big)} \sum_{i=1}^n W_{ni} + o\Big(\frac{n\kappa_n^2}{p_n}\Big). \tag{A.5}$$

Since the W_{ni}, $i = 1, \ldots, n$ are mutually independent with mean zero and variance $1/n$, we obtain that

$$\mathrm{E}\big[\Lambda_n^2\big] = \big(\mathrm{E}\big[\Lambda_n\big]\big)^2 + \mathrm{Var}\big[\Lambda_n\big] = \frac{n^2\kappa_n^4}{4p_n^2}(f'(0))^4 + o\Big(\frac{n^2\kappa_n^4}{p_n^2}\Big) + \frac{n\kappa_n^2}{p_n}(f'(0))^2 + o\Big(\frac{n\kappa_n^2}{p_n}\Big). \tag{A.6}$$

If $\kappa_n^2 = o(\frac{p_n}{n})$, then (A.6) implies that $\exp(\Lambda_n) \to Z$ in distribution, where $Z \equiv 1$. Since $\mathrm{P}[Z = 0] = 0$ and $\mathrm{E}[Z] = 1$, Le Cam's first lemma yields that $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$ and $\mathrm{P}^{(n)}_0$ are mutually contiguous.

We may therefore assume that $\kappa_n^2 = \tau_n^2 p_n/n$, where the positive sequence (τ_n) is $O(1)$ but not $o(1)$. In this case, (A.5) rewrites as

$$\Lambda_n = -\frac{\tau_n^2}{2}(f'(0))^2 + \sqrt{\tau_n^2(f'(0))^2 + o(1)} \sum_{i=1}^n W_{ni} + o(1).$$

Applying the Cauchy-Schwarz inequality and the Chebychev inequality, then using (A.3) and (A.4), provides that, for some positive constant C,

$$\sum_{i=1}^n \mathrm{E}[W_{ni}^2 \mathbb{I}[|W_{ni}| > \varepsilon]] \leq n\sqrt{\mathrm{E}[W_{ni}^4]\mathrm{P}[|W_{ni}| > \varepsilon]} \leq \frac{n}{\varepsilon}\sqrt{\mathrm{E}[W_{ni}^4]\mathrm{Var}[W_{ni}]} = \frac{\sqrt{n}}{\varepsilon}\sqrt{\mathrm{E}[W_{ni}^4]}$$

$$\leq \frac{Cn^{1/2}E_{n4}^{1/2}}{\varepsilon(nV_n)^2} = \frac{C\Big(\frac{n\kappa_n^2\ell''_{f,4}(0)}{2p_n} + o\big(\frac{n\kappa_n^2}{p_n}\big)\Big)^{1/2}}{\varepsilon\Big(\frac{n\kappa_n^2}{p_n}(f'(0))^2 + o\big(\frac{n\kappa_n^2}{p_n}\big)\Big)^2} = \frac{o(\tau_n)}{\varepsilon\big(\tau_n^2(f'(0))^2 + o(\tau_n^2)\big)^2} = o(1),$$

where we have used the fact that $\ell''_{f,4}(0) = 0$. This shows that $\sum_{i=1}^n W_{ni}$ satisfies the classical Levy-Lindeberg condition, hence is asymptotically standard normal (as already mentioned, W_{ni}, $i = 1, \ldots, n$ are mutually independent with mean zero and variance $1/n$). For any subsequence $(\exp(\Lambda_{n_m}))$ converging in distribution, we must then have

$$\exp(\Lambda_{n_m}) \to \exp(Y), \quad \text{with } Y \sim \mathcal{N}\Big(- \frac{(f'(0))^2}{2} \lim_{n\to\infty} \tau_{n_m}^2, (f'(0))^2 \lim_{n\to\infty} \tau_{n_m}^2 \Big).$$

Mutual contiguity $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$ and $\mathrm{P}^{(n)}_0$ then follows from the fact that $\mathrm{P}[\exp(Y) = 0] = 0$ and $\mathrm{E}[\exp(Y)] = 1$. □

Proof of Theorem 2.2. As in the proof of Theorem 2.1, all expectations and variances in this proof are taken under the null of uniformity $\mathrm{P}^{(n)}_0$ and all stochastic convergences and o_P's are as $n \to \infty$ under $\mathrm{P}^{(n)}_0$. The central limit theorem then directly establishes Part (ii) of the result, since $\mathrm{E}[\Delta^{(n)}_{\boldsymbol{\theta}_n,f}] = 0$ and $\mathrm{Var}[\Delta^{(n)}_{\boldsymbol{\theta}_n,f}] = \frac{p_n}{n}(f'(0))^2 \mathrm{Var}\big[\sum_{i=1}^n \mathbf{X}'_{ni}\boldsymbol{\theta}_n\big] = (f'(0))^2$.

It therefore remains to establish Part (i). Recall that, in the case where (τ_n) is $O(1)$ but not $o(1)$, we have obtained in the proof of Theorem 2.1 that

$$\begin{aligned}\Lambda_n &= -\frac{\tau_n^2}{2}(f'(0))^2 + \sqrt{\tau_n^2(f'(0))^2 + o(1)} \sum_{i=1}^n W_{ni} + o(1) \\ &= -\frac{\tau_n^2}{2}(f'(0))^2 + \tau_n f'(0) \sum_{i=1}^n W_{ni} + o_\mathrm{P}(1),\end{aligned}$$

where $\sum_{i=1}^n W_{ni} = (1/\sqrt{nV_n}) \sum_{i=1}^n (\log f(\kappa_n \mathbf{X}'_{ni}\boldsymbol{\theta}_n) - E_{n1})$ is asymptotically standard normal. To establish the result, it is therefore sufficient to show that $\tau_n[(\sum_{i=1}^n W_{ni}) - (1/f'(0))\Delta^{(n)}_{\boldsymbol{\theta}_n,f}]$ converges to zero in quadratic mean. To do so, write

$$\tau_n\Big(\sum_{i=1}^n W_{ni}\Big) - \frac{\tau_n}{f'(0)}\Delta^{(n)}_{\boldsymbol{\theta}_n,f} = \frac{\tau_n}{\sqrt{nV_n}} \sum_{i=1}^n \big(\log f(\kappa_n \mathbf{X}'_{ni}\boldsymbol{\theta}_n) - E_{n1} - \sqrt{p_n V_n}\, \mathbf{X}'_{ni}\boldsymbol{\theta}_n \big) =: \frac{M_n}{\sqrt{nV_n}}.$$

Then using $\mathrm{E}[\mathbf{X}'_{n1}\boldsymbol{\theta}_n] = 0$ and $\mathrm{E}[(\mathbf{X}'_{n1}\boldsymbol{\theta}_n)^2] = 1/p_n$, we obtain

$$\begin{aligned}\mathrm{E}\big[M_n^2\big] &= n\tau_n^2\, \mathrm{E}\big[\big(\log f(\kappa_n \mathbf{X}'_{ni}\boldsymbol{\theta}_n) - E_{n1} - \sqrt{p_n V_n}\, \mathbf{X}'_{ni}\boldsymbol{\theta}_n\big)^2\big] \\ &= n\tau_n^2\, (2V_n - 2\sqrt{p_n V_n}\, \mathrm{E}[\mathbf{X}'_{ni}\boldsymbol{\theta}_n(\log f(\kappa_n \mathbf{X}'_{ni}\boldsymbol{\theta}_n) - E_{n1})]) \\ &= 2n\tau_n^2 V_n - 2\tau_n n^{3/2}\sqrt{V_n}\, \mathrm{E}[\kappa_n \mathbf{X}'_{ni}\boldsymbol{\theta}_n \log f(\kappa_n \mathbf{X}'_{ni}\boldsymbol{\theta}_n)],\end{aligned}$$

which, letting $g(x) := x(\log f(x))$, provides

$$\mathrm{E}\Big[\Big(\tau_n\Big(\sum_{i=1}^n W_{ni}\Big) - \frac{\tau_n}{f'(0)}\Delta^{(n)}_{\boldsymbol{\theta}_n,f}\Big)^2\Big] = 2\tau_n^2 - \frac{2\tau_n n}{\sqrt{nV_n}}\, \mathrm{E}[g(\kappa_n \mathbf{X}'_{ni}\boldsymbol{\theta}_n)]. \tag{A.7}$$

Using Lemma A.1,

$$\begin{aligned}\mathrm{E}[g(\kappa_n \mathbf{X}'_{ni}\boldsymbol{\theta}_n)] &= c_{p_n}\int_{-1}^{1}(1-s^2)^{(p_n-3)/2} g(\kappa_n s)\,ds \\ &= \frac{\kappa_n^2}{2p_n} g''(0) + o\Big(\frac{\kappa_n^2}{p_n}\Big) = \frac{\kappa_n^2}{p_n} f'(0) + o\Big(\frac{\kappa_n^2}{p_n}\Big).\end{aligned}$$

Plugging in (A.7) and using (A.4) then yields

$$\mathrm{E}\Big[\Big(\tau_n\Big(\sum_{i=1}^n W_{ni}\Big) - \frac{\tau_n}{f'(0)}\Delta^{(n)}_{\boldsymbol{\theta}_n,f}\Big)^2\Big] = 2\tau_n^2 - \frac{2\tau_n\Big(\frac{n\kappa_n^2}{p_n}f'(0) + o\Big(\frac{n\kappa_n^2}{p_n}\Big)\Big)}{\Big(\frac{n\kappa_n^2}{p_n}(f'(0))^2 + o\Big(\frac{n\kappa_n^2}{p_n}\Big)\Big)^{1/2}} = o(1),$$

as was to be showed. □

APPENDIX B: PROOFS FOR SECTION 4

In this second appendix, we establish Proposition 4.1, and Theorems 4.1, 4.2 and 4.3. We start with some preliminary results and the proof of Proposition 4.1.

B.1. Preliminary lemmas and proof of Proposition 4.1.

Define the quantities

$$u_{ni} := \mathbf{X}'_{ni}\boldsymbol{\theta}_n \quad \text{and} \quad v_{ni} := \sqrt{1-u_{ni}^2}$$

that are associated with the *tangent-normal decomposition* $\mathbf{X}_{ni} = u_{ni}\boldsymbol{\theta}_n + v_{ni}\mathbf{S}_{ni}$ of $\mathbf{X}_{ni}$, where

$$\mathbf{S}_{ni} := \begin{cases} \dfrac{\mathbf{X}_{ni} - (\mathbf{X}'_{ni}\boldsymbol{\theta}_n)\boldsymbol{\theta}_n}{\|\mathbf{X}_{ni} - (\mathbf{X}'_{ni}\boldsymbol{\theta}_n)\boldsymbol{\theta}_n\|} & \text{if } \mathbf{X}_{ni} \neq \boldsymbol{\theta}_n \\ \mathbf{0} & \text{otherwise.}\end{cases}$$

With this notation, $e_{n\ell} = \mathrm{E}[u_{ni}^\ell]$ and $f_{n\ell} = \mathrm{E}[v_{ni}^\ell]$ (see Proposition 4.1). We start with the following lemma.

LEMMA B.1. *Under* $\mathrm{P}^{(n)}_{F_n}$, *we have that*

(i) $\mathrm{E}[\mathbf{X}'_{ni}\mathbf{X}_{nj}] = e_{n1}^2$ *for any* $i<j$.

(ii) $\mathrm{E}[(\mathbf{X}'_{ni}\mathbf{X}_{nj})^2] = e_{n2}^2 + f_{n2}^2/(p_n-1)$ *for any* $i<j$.

(iii) $\mathrm{E}[(\mathbf{X}'_{ni}\mathbf{X}_{nk})(\mathbf{X}'_{n\ell}\mathbf{X}_{nj})] = e_{n2}e_{n1}^2$ *for any* $i<j$ *and* $k<\ell$ *such that there are exactly three different indices in* $\{i,j,k,\ell\}$.

(iv) $\mathrm{E}[(\mathbf{X}'_{ni}\mathbf{X}_{nj})(\mathbf{X}'_{nk}\mathbf{X}_{n\ell})] = e_{n1}^4$ *for any* $i<j$ *and* $k<\ell$ *such that there are exactly four different indices in* $\{i,j,k,\ell\}$.

PROOF. The first part of the lemma directly follows from

$$\mathbf{X}'_{ni}\mathbf{X}_{nj} = (u_{ni}\boldsymbol{\theta}_n + v_{ni}\mathbf{S}_{ni})'(u_{nj}\boldsymbol{\theta}_n + v_{nj}\mathbf{S}_{nj}) = u_{ni}u_{nj} + v_{ni}v_{nj}\mathbf{S}'_{ni}\mathbf{S}_{nj}.$$

For the remaining claims, write

$$\begin{aligned}(\mathbf{X}'_{ni}\mathbf{X}_{nj})(\mathbf{X}'_{nk}\mathbf{X}_{n\ell}) &= (u_{ni}u_{nj} + v_{ni}v_{nj}\mathbf{S}'_{ni}\mathbf{S}_{nj})(u_{nk}u_{n\ell} + v_{nk}v_{n\ell}\mathbf{S}'_{nk}\mathbf{S}_{n\ell}) \\ &= u_{ni}u_{nj}u_{nk}u_{n\ell} + u_{ni}u_{nj}v_{nk}v_{n\ell}\mathbf{S}'_{nk}\mathbf{S}_{n\ell} \\ &\qquad + v_{ni}v_{nj}u_{nk}u_{n\ell}\mathbf{S}'_{ni}\mathbf{S}_{nj} + v_{ni}v_{nj}v_{nk}v_{n\ell}(\mathbf{S}'_{ni}\mathbf{S}_{nj})(\mathbf{S}'_{nk}\mathbf{S}_{n\ell}),\end{aligned}$$

which, for $i < j$ and $k < \ell$, entails

$$\text{(B.1)}\quad \mathrm{E}[(\mathbf{X}'_{ni}\mathbf{X}_{nj})(\mathbf{X}'_{nk}\mathbf{X}_{n\ell})] = \mathrm{E}[u_{ni}u_{nj}u_{nk}u_{n\ell}] + \mathrm{E}[v_{ni}v_{nj}v_{nk}v_{n\ell}]\mathrm{E}[(\mathbf{S}'_{ni}\mathbf{S}_{nj})(\mathbf{S}'_{nk}\mathbf{S}_{n\ell})].$$

Part 2 of the result then follows from the fact that $\mathrm{E}[(\mathbf{S}'_{ni}\mathbf{S}_{nj})^2] = 1/(p_n - 1)$. For Parts 3-4 of the result, there is always one of the indices i, j, k, ℓ that is different from the other three indices, which implies that $\mathrm{E}[(\mathbf{S}'_{ni}\mathbf{S}_{nj})(\mathbf{S}'_{nk}\mathbf{S}_{n\ell})] = 0$. The result readily follows. □

Lemma B.1 allows to prove Proposition 4.1.

PROOF OF PROPOSITION 4.1. Since the expectation readily follows from Lemma B.1(i), we can focus on the variance. Using Lemma B.1(i) again, we obtain

$$\begin{aligned}\mathrm{Var}_{F_n}[R^{st}_{p,n}] &= \frac{2p_n}{n^2} \sum_{1\le i<j\le n} \sum_{1\le k<\ell\le n} \mathrm{Cov}[\mathbf{X}'_{ni}\mathbf{X}_{nj}, \mathbf{X}'_{nk}\mathbf{X}_{n\ell}] \\ &= \frac{2p_n}{n^2} \sum_{1\le i<j\le n} \sum_{1\le k<\ell\le n} \Big(\mathrm{E}[(\mathbf{X}'_{ni}\mathbf{X}_{nj})(\mathbf{X}'_{nk}\mathbf{X}_{n\ell})] - e^4_{n1}\Big).\end{aligned}$$

In this sum, there are $\binom{n}{2}$ terms corresponding to Lemma B.1(ii). There are $6\binom{n}{4}$ terms corresponding to Lemma B.1(iv) (these terms do not contribute to the sum). Therefore, there are

$$\binom{n}{2}^2 - \binom{n}{2} - 6\binom{n}{4} = n(n-1)(n-2)$$

terms corresponding to Lemma B.1(iii). Consequently,

$$\begin{aligned}\mathrm{Var}_{F_n}[R^{st}_{p,n}] &= \frac{2p_n}{n^2}\bigg\{\binom{n}{2}\Big(e^2_{n2} + f^2_{n2}/(p_n - 1) - e^4_{n1}\Big) + n(n-1)(n-2)\Big(e_{n2}e^2_{n1} - e^4_{n1}\Big)\bigg\} \\ &= \frac{p_n(n-1)}{n}\bigg\{\Big(e^2_{n2} + f^2_{n2}/(p_n - 1) - e^4_{n1}\Big) + 2(n-2)\Big(e_{n2}e^2_{n1} - e^4_{n1}\Big)\bigg\} \\ &= \frac{n-1}{n}\Big\{p_n\tilde{e}^2_{n2} + \frac{p_n}{p_n - 1}f^2_{n2} + 2(n-1)p_n e^2_{n1}\tilde{e}_{n2}\Big\},\end{aligned}$$

which establishes the result. □

Both following lemmas are needed to establish Theorems 4.1 and 4.2.

LEMMA B.2. *Under* $\mathrm{P}^{(n)}_{F_n}$, *we have that*

(i) $\mathrm{E}\big[\big(\mathbf{X}_{ni}-e_{n1}\boldsymbol{\theta}_n\big)\big(\mathbf{X}_{ni}-e_{n1}\boldsymbol{\theta}_n\big)'\big]=\tilde{e}_{n2}\boldsymbol{\theta}_n\boldsymbol{\theta}_n'+\frac{f_{n2}}{p_n-1}(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}_n')$.

(ii) $\mathrm{Var}\big[\big(\mathbf{X}_{ni}'\boldsymbol{\theta}_n-e_{n1}\big)\big(\mathbf{X}_{nj}'\boldsymbol{\theta}_n-e_{n1}\big)\big]=\tilde{e}_{n4}-\tilde{e}_{n2}^2$ *for* $i=j$ *and* $\tilde{e}_{n2}^2$ *for* $i\neq j$.

(iii) $\mathrm{E}\big[\mathbf{X}_{ni}'(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{nj}\big]=f_{n2}$ *for* $i=j$ *and 0 for* $i\neq j$.

(iv) $\mathrm{Var}\big[\mathbf{X}_{ni}'(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{nj}\big]=f_{n4}-f_{n2}^2$ *for* $i=j$ *and* $f_{n2}^2/(p_n-1)$ *for* $i\neq j$.

PROOF. (i) Using the tangent-normal decomposition, we obtain

$$\mathrm{E}\Big[\big(\mathbf{X}_{ni}-e_{n1}\boldsymbol{\theta}_n\big)\big(\mathbf{X}_{ni}-e_{n1}\boldsymbol{\theta}_n\big)'\Big]=\mathrm{E}\Big[\big((u_{ni}-e_{n1})\boldsymbol{\theta}_n+v_{ni}\mathbf{S}_{ni}\big)\big((u_{ni}-e_{n1})\boldsymbol{\theta}_n+v_{ni}\mathbf{S}_{ni}\big)'\Big]$$

$$=\mathrm{E}\big[(u_{ni}-e_{n1})^2\big]\boldsymbol{\theta}_n\boldsymbol{\theta}_n'+f_{n2}\,\mathrm{E}\big[\mathbf{S}_{ni}\mathbf{S}_{ni}'\big]=\tilde{e}_{n2}\boldsymbol{\theta}_n\boldsymbol{\theta}_n'+\frac{f_{n2}}{p_n-1}(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}_n').$$

(ii)-(iv) The results readily follow from the fact that $\mathbf{X}_{ni}'\boldsymbol{\theta}_n-e_{n1}=u_{ni}-e_{n1}$ and $\mathbf{X}_{ni}'(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{nj}=v_{ni}v_{nj}\mathbf{S}_{ni}'\mathbf{S}_{nj}$ (and the identity $\mathrm{E}[(\mathbf{S}_{ni}'\mathbf{S}_{nj})^2]=1/(p_n-1)$). □

LEMMA B.3. *Consider expectations of the form* $c_{ijrs}=\mathrm{E}\left[\Delta_{i\ell}\Delta_{j\ell}\Delta_{r\ell}\Delta_{s\ell}\right]$ *taken under* $\mathrm{P}^{(n)}_{F_n}$, *with* $\Delta_{i\ell}:=(\mathbf{X}_{ni}-e_{n1}\boldsymbol{\theta}_n)'(\mathbf{X}_{n\ell}-e_{n1}\boldsymbol{\theta}_n)$ *and* $i\leq j\leq r\leq s<\ell$. *Then*

(i) $c_{ijrs}=\tilde{e}_{n4}^2+\frac{6}{p_n-1}\big(\mathrm{E}\big[v_{ni}^2(u_{ni}-e_{n1})^2\big]\big)^2+\frac{3f_{n4}^2}{p_n^2-1}$ *if* $i=j=r=s$.

(ii) $c_{ijrs}=\tilde{e}_{n2}^2\tilde{e}_{n4}+\frac{2\tilde{e}_{n2}f_{n2}}{p_n-1}\mathrm{E}\big[v_{ni}^2(u_{ni}-e_{n1})^2\big]+\frac{f_{n2}^2f_{n4}}{(p_n-1)^2}$ *if* $i=j<r=s$.

(iii) $c_{ijrs}=0$ *otherwise.*

PROOF. We start with the proof of (iii). Assume that $j=r$, so that we are not in case (ii). Since case (i) is excluded, we have $i<j$ or $r<s$. In both cases, one of the four indices i,j,r,s is different from the other three indices. Since $\mathrm{E}[\Delta_{i\ell}]=0$, we obtain that $c_{ijrs}=0$, which establishes (iii). Turning to the proof of (i)-(ii), we use the tangent-normal decomposition again to write $\Delta_{j\ell}$ as $(u_{nj}-e_{n1})(u_{n\ell}-e_{n1})+v_{nj}v_{n\ell}(\mathbf{S}_{nj}'\mathbf{S}_{n\ell})$. Since

$\mathrm{E}[(\mathbf{S}'_{nj}\mathbf{S}_{n\ell})^k] = 0$ for any odd integer k, this leads to decomposing c_{jjrr} into

$$
\begin{aligned}
c_{jjrr} = \mathrm{E}\left[(u_{nj}-e_{n1})^2(u_{nr}-e_{n1})^2(u_{n\ell}-e_{n1})^4\right] \\
+2\mathrm{E}\left[(u_{nr}-e_{n1})^2(u_{n\ell}-e_{n1})^2 v_{nj}^2 v_{n\ell}^2(\mathbf{S}'_{nj}\mathbf{S}_{n\ell})^2\right] \\
+4\mathrm{E}\left[(u_{nj}-e_{n1})(u_{nr}-e_{n1})(u_{n\ell}-e_{n1})^2 v_{nj}v_{nr}v_{n\ell}^2(\mathbf{S}'_{nj}\mathbf{S}_{n\ell})(\mathbf{S}'_{nr}\mathbf{S}_{n\ell})\right] \\
+\mathrm{E}\left[v_{nj}^2 v_{nr}^2 v_{n\ell}^4(\mathbf{S}'_{nj}\mathbf{S}_{n\ell})^2(\mathbf{S}'_{nr}\mathbf{S}_{n\ell})^2\right].
\end{aligned}
$$

The result then follows from the identities $\mathrm{E}[(\mathbf{S}'_{nj}\mathbf{S}_{n\ell})^2] = 1/(p_n-1)$ and $\mathrm{E}[(\mathbf{S}'_{nj}\mathbf{S}_{n\ell})^4] = 3/(p_n^2-1)$. □

B.2. Proof of Theorem 4.1. The proof is based on the following central limit theorem for martingale differences.

THEOREM B.1 (Billingsley 1995, Theorem 35.12). *Let $D_{n\ell}$, $\ell = 1,\ldots,n$, $n = 1,2,\ldots$, be a triangular array of random variables such that, for any n, $D_{n1}, D_{n2},\ldots,D_{nn}$ is a martingale difference sequence with respect to some filtration $\mathcal{F}_{n1},\mathcal{F}_{n2},\ldots,\mathcal{F}_{nn}$. Assume that, for any n,ℓ, $D_{n\ell}$ has a finite variance. Letting $\sigma_{n\ell}^2 = \mathrm{E}\big[D_{n\ell}^2\,|\,\mathcal{F}_{n,\ell-1}\big]$ (with $\mathcal{F}_{n0}$ being the trivial σ-algebra $\{\emptyset,\Omega\}$ for all n), further assume that, as $n\to\infty$,*

$$\sum_{\ell=1}^n \sigma_{n\ell}^2 \xrightarrow{\mathrm{P}} 1 \tag{B.2}$$

(where $\xrightarrow{\mathrm{P}}$ denotes convergence in probability), and

$$\sum_{\ell=1}^n \mathrm{E}\big[D_{n\ell}^2\,\mathbb{I}[|D_{n\ell}|>\varepsilon]\big] \to 0. \tag{B.3}$$

Then $\sum_{\ell=1}^n D_{n\ell}$ is asymptotically standard normal.

Writing $\mathrm{E}_{n\ell}$ for the conditional expectation with respect to the σ-algebra $\mathcal{F}_{n\ell}$ generated by $\mathbf{X}_{n1},\ldots,\mathbf{X}_{n\ell}$, we have

$$\mathrm{E}_{n\ell}\big[\bar{R}_n^{\mathrm{St}}\big] = \frac{\sqrt{2p_n}}{n\sigma_n}\Bigg\{\sum_{1\le i<j\le\ell}\big(\mathbf{X}'_{ni}\mathbf{X}_{nj}-e_{n1}^2\big)+(n-\ell)e_{n1}\sum_{i=1}^{\ell}\big(\mathbf{X}'_{ni}\boldsymbol{\theta}_n-e_{n1}\big)\Bigg\}.$$

Let

$$
\begin{aligned}
D_{n\ell} &:= \mathrm{E}_{n\ell}\big[\bar{R}_n^{\mathrm{St}}\big]-\mathrm{E}_{n,\ell-1}\big[\bar{R}_n^{\mathrm{St}}\big] \\
&= \frac{\sqrt{2p_n}}{n\sigma_n}\Bigg\{\sum_{i=1}^{\ell-1}(\mathbf{X}_{ni}-e_{n1}\boldsymbol{\theta}_n)+(n-1)e_{n1}\boldsymbol{\theta}_n\Bigg\}'\big(\mathbf{X}_{n\ell}-e_{n1}\boldsymbol{\theta}_n\big), \qquad \text{(B.4)}
\end{aligned}
$$

for any $\ell = 1, 2, \ldots$ (throughout, sums over empty set of indices are defined as being equal to zero). It can be checked that $\bar{R}_n^{\rm St} = \sum_{\ell=1}^n D_{n\ell}$. The following lemmas then take care of the conditions required in Theorem B.1.

Lemma B.4. *Let the assumptions of Theorem 4.1 hold. Then, under* $\mathrm{P}_{F_n}^{(n)}$, *(i)* $\sum_{\ell=1}^n \mathrm{E}[\sigma_{n\ell}^2]$ *converges to one as* $n \to \infty$, *and* $\mathrm{Var}[\sum_{\ell=1}^n \sigma_{n\ell}^2]$ *converges to zero as* $n \to \infty$.

Lemma B.5. *Let the assumptions of Theorem 4.1 hold and fix* $\varepsilon > 0$. *Then, under* $\mathrm{P}_{F_n}^{(n)}$, $\sum_{\ell=1}^n \mathrm{E}\big[(D_{n\ell})^2\, \mathbb{I}[|D_{n\ell}| > \varepsilon]\big] \to 0$ *as* $n \to \infty$.

In the rest of the paper, C is a positive constant that may change from line to line.

Proof of Lemma B.4. (i) Note that

$$\sigma_{n\ell}^2 = \frac{2p_n}{n^2\sigma_n^2}\Big\{\sum_{i=1}^{\ell-1}(\mathbf{X}_{ni} - e_{n1}\boldsymbol{\theta}_n) + (n-1)e_{n1}\boldsymbol{\theta}_n\Big\}' \mathrm{E}\Big[\big(\mathbf{X}_{n\ell} - e_{n1}\boldsymbol{\theta}_n\big)\big(\mathbf{X}_{n\ell} - e_{n1}\boldsymbol{\theta}_n\big)'\Big] \times \Big\{\sum_{j=1}^{\ell-1}(\mathbf{X}_{nj} - e_{n1}\boldsymbol{\theta}_n) + (n-1)e_{n1}\boldsymbol{\theta}_n\Big\}.$$

Using the fact that

$$\mathrm{E}\Big[\big(\mathbf{X}_{n\ell} - e_{n1}\boldsymbol{\theta}_n\big)\big(\mathbf{X}_{n\ell} - e_{n1}\boldsymbol{\theta}_n\big)'\Big] = \tilde{e}_{n2}\boldsymbol{\theta}_n\boldsymbol{\theta}_n' + \frac{f_{n2}}{p_n - 1}(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n')$$

(see Lemma B.2), we obtain

$$\sigma_{n\ell}^2 = \frac{2p_n\tilde{e}_{n2}}{n^2\sigma_n^2}\Big\{\sum_{i,j=1}^{\ell-1}(\mathbf{X}_{ni}'\boldsymbol{\theta}_n - e_{n1})(\mathbf{X}_{nj}'\boldsymbol{\theta}_n - e_{n1}) + 2(n-1)e_{n1}\sum_{i=1}^{\ell-1}(\mathbf{X}_{ni}'\boldsymbol{\theta}_n - e_{n1}) + (n-1)^2 e_{n1}^2\Big\}$$
$$+ \frac{2p_n f_{n2}}{(p_n-1)n^2\sigma_n^2}\sum_{i,j=1}^{\ell-1}\mathbf{X}_{ni}'(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{nj}. \tag{B.5}$$

Therefore

$$\mathrm{E}[\sigma_{n\ell}^2] = \frac{2p_n\tilde{e}_{n2}}{n^2\sigma_n^2}\Big\{(\ell-1)\tilde{e}_{n2} + 0 + (n-1)^2 e_{n1}^2\Big\} + \frac{2p_n(\ell-1)f_{n2}^2}{(p_n-1)n^2\sigma_n^2}, \tag{B.6}$$

where we have used Lemma B.2(iii). This yields

$$s_n^2 := \sum_{\ell=1}^n \mathrm{E}[\sigma_{n\ell}^2] = \frac{(n-1)p_n\tilde{e}_{n2}^2}{n\sigma_n^2} + \frac{2p_n\tilde{e}_{n2}}{n\sigma_n^2}(n-1)^2 e_{n1}^2 + \frac{(n-1)p_n f_{n2}^2}{(p_n-1)n\sigma_n^2} \to 1 \tag{B.7}$$

as $n \to \infty$, as was to be shown.

(ii) From (B.5), we obtain

$$\mathrm{Var}\Big[\sum_{\ell=1}^{n}\sigma_{n\ell}^2\Big] \leq C\big(\mathrm{Var}\big[A_n\big]+\mathrm{Var}\big[B_n\big]+\mathrm{Var}\big[C_n\big]\big),$$

where

$$A_n := \frac{p_n\tilde{e}_{n2}}{n^2\sigma_n^2}\sum_{\ell=1}^{n}\sum_{i,j=1}^{\ell-1}(\mathbf{X}_{ni}'\boldsymbol{\theta}_n-e_{n1})(\mathbf{X}_{nj}'\boldsymbol{\theta}_n-e_{n1}),$$

$$B_n := \frac{p_n e_{n1}\tilde{e}_{n2}}{n\sigma_n^2}\sum_{\ell=1}^{n}\sum_{i=1}^{\ell-1}(\mathbf{X}_{ni}'\boldsymbol{\theta}_n-e_{n1})$$

and

$$C_n := \frac{f_{n2}}{n^2\sigma_n^2}\sum_{\ell=1}^{n}\sum_{i,j=1}^{\ell-1}\mathbf{X}_{ni}'(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{nj}.$$

We establish the result by showing that, under the assumptions considered, $\mathrm{Var}[A_n]$, $\mathrm{Var}[B_n]$ and $\mathrm{Var}[C_n]$ all are $o(1)$ as $n\to\infty$. We start with A_n, which we split into

$$\begin{aligned}A_n &= \frac{p_n\tilde{e}_{n2}}{n^2\sigma_n^2}\sum_{\ell=1}^{n}\sum_{i=1}^{\ell-1}(\mathbf{X}_{ni}'\boldsymbol{\theta}_n-e_{n1})^2+\frac{2p_n\tilde{e}_{n2}}{n^2\sigma_n^2}\sum_{\ell=1}^{n}\sum_{1\leq i<j\leq \ell-1}(\mathbf{X}_{ni}'\boldsymbol{\theta}_n-e_{n1})(\mathbf{X}_{nj}'\boldsymbol{\theta}_n-e_{n1})\\ &= \frac{p_n\tilde{e}_{n2}}{n^2\sigma_n^2}\sum_{i=1}^{n-1}(n-i)(\mathbf{X}_{ni}'\boldsymbol{\theta}_n-e_{n1})^2+\frac{2p_n\tilde{e}_{n2}}{n^2\sigma_n^2}\sum_{1\leq i<j\leq n-1}(n-j)(\mathbf{X}_{ni}'\boldsymbol{\theta}_n-e_{n1})(\mathbf{X}_{nj}'\boldsymbol{\theta}_n-e_{n1})\\ &=: A_n^{(1)}+A_n^{(2)},\end{aligned}$$

say. Clearly,

$$\begin{aligned}\mathrm{Var}\big[A_n^{(1)}\big] &= \frac{p_n^2\tilde{e}_{n2}^2}{n^4\sigma_n^4}\sum_{i=1}^{n-1}(n-i)^2\,\mathrm{Var}\big[(\mathbf{X}_{ni}'\boldsymbol{\theta}_n-e_{n1})^2\big]\\ &\leq C\,\frac{p_n^2\tilde{e}_{n2}^2\big(\tilde{e}_{n4}-\tilde{e}_{n2}^2\big)}{n\sigma_n^4}\leq C\,\frac{p_n^2\tilde{e}_{n2}^2\big(\tilde{e}_{n4}-\tilde{e}_{n2}^2\big)}{n(p_n\tilde{e}_{n2}^2)^2}=C\left(\frac{\tilde{e}_{n4}}{n\tilde{e}_{n2}^2}-\frac{1}{n}\right),\end{aligned}$$

which, by assumption, is $o(1)$ as $n\to\infty$. Since $(\mathbf{X}_{ni}'\boldsymbol{\theta}_n-e_{n1})(\mathbf{X}_{nj}'\boldsymbol{\theta}_n-e_{n1})$, $i<j$, and $(\mathbf{X}_{nk}'\boldsymbol{\theta}_n-e_{n1})(\mathbf{X}_{n\ell}'\boldsymbol{\theta}_n-e_{n1})$, $k<\ell$, are uncorrelated as soon as $(i,j)\neq(k,\ell)$, we obtain

$$\mathrm{Var}\big[A_n^{(2)}\big]=\frac{4p_n^2\tilde{e}_{n2}^2}{n^4\sigma_n^4}\sum_{1\leq i<j\leq n-1}(n-j)^2\,\mathrm{Var}\big[(\mathbf{X}_{ni}'\boldsymbol{\theta}_n-e_{n1})(\mathbf{X}_{nj}'\boldsymbol{\theta}_n-e_{n1})\big]\leq C\,\frac{p_n^2\tilde{e}_{n2}^4}{\sigma_n^4}.$$

In view of the majorations

$$\frac{p_n^2\tilde{e}_{n2}^4}{\sigma_n^4}\leq C\frac{p_n^2\tilde{e}_{n2}^4}{(2np_ne_{n1}^2\tilde{e}_{n2})^2}=C\Big(\frac{\tilde{e}_{n2}}{ne_{n1}^2}\Big)^2\quad\text{and}\quad\frac{p_n^2\tilde{e}_{n2}^4}{\sigma_n^4}\leq C\Big(\frac{p_n\tilde{e}_{n2}^2}{f_{n2}^2}\Big)^2,$$

$\mathrm{Var}\big[A_n^{(2)}\big]$, by assumption, is $o(1)$ as $n\to\infty$. Therefore, $\mathrm{Var}[A_n]$ is indeed $o(1)$ as $n\to\infty$.

Turning to B_n,

$$\mathrm{Var}[B_n] = \frac{p_n^2 e_{n1}^2 \tilde{e}_{n2}^2}{n^2\sigma_n^4} \mathrm{Var}\Big[\sum_{i=1}^{n-1}(n-i)(\mathbf{X}_{ni}'\boldsymbol{\theta}_n - e_{n1})\Big] = \frac{p_n^2 e_{n1}^2 \tilde{e}_{n2}^2}{n^2\sigma_n^4}\sum_{i=1}^{n-1}(n-i)^2\tilde{e}_{n2} \leq C\frac{np_n^2e_{n1}^2\tilde{e}_{n2}^3}{\sigma_n^4},$$

which can be upper-bounded by

$$C\frac{np_n^2e_{n1}^2\tilde{e}_{n2}^3}{(2np_ne_{n1}^2\tilde{e}_{n2})^2} = C\frac{\tilde{e}_{n2}}{ne_{n1}^2} \quad \text{and by} \quad C\frac{np_n^2e_{n1}^2\tilde{e}_{n2}^3}{np_ne_{n1}^2\tilde{e}_{n2}f_{n2}^2} = C\frac{p_n\tilde{e}_{n2}^2}{f_{n2}^2}\cdot$$

We conclude that $\mathrm{Var}[B_n]$ is also $o(1)$ as $n \to \infty$.

Finally, we consider C_n. Proceeding as for A_n, we split C_n into

$$\begin{aligned} C_n &= \frac{f_{n2}}{n^2\sigma_n^2}\sum_{\ell=1}^{n}\sum_{i=1}^{\ell-1}\mathbf{X}_{ni}'(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{ni} + \frac{2f_{n2}}{n^2\sigma_n^2}\sum_{\ell=1}^{n}\sum_{1\leq i<j\leq \ell-1}\mathbf{X}_{ni}'(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{nj} \\ &= \frac{f_{n2}}{n^2\sigma_n^2}\sum_{i=1}^{n-1}(n-i)\mathbf{X}_{ni}'(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{ni} + \frac{2f_{n2}}{n^2\sigma_n^2}\sum_{1\leq i<j\leq n-1}(n-j)\mathbf{X}_{ni}'(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{nj} \\ &=: C_n^{(1)} + C_n^{(2)}, \end{aligned}$$

say. Clearly,

$$\begin{aligned} \mathrm{Var}\big[C_n^{(1)}\big] &= \frac{f_{n2}^2}{n^4\sigma_n^4}\sum_{i=1}^{n-1}(n-i)^2\,\mathrm{Var}\big[\mathbf{X}_{ni}'(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{ni}\big] \\ &\leq C\frac{f_{n2}^2(f_{n4} - f_{n2}^2)}{n\sigma_n^4} \leq C\frac{f_{n4} - f_{n2}^2}{nf_{n2}^2} = C\Big(\frac{f_{n4}}{nf_{n2}^2} - \frac{1}{n}\Big) = o(1) \end{aligned}$$

as $n \to \infty$. Since $\mathbf{X}_{ni}'(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{nj}$, $i<j$, and $\mathbf{X}_{nk}'(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{n\ell}$, $k<\ell$, are uncorrelated as soon as $(i,j) \neq (k,\ell)$, we obtain

$$\mathrm{Var}\big[C_n^{(2)}\big] = \frac{4f_{n2}^2}{n^4\sigma_n^4}\sum_{1\leq i<j\leq n-1}(n-j)^2\,\mathrm{Var}\big[\mathbf{X}_{ni}'(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n')\mathbf{X}_{nj}\big] \leq C\frac{f_{n2}^4}{\sigma_n^4(p_n-1)} \leq \frac{C}{p_n}\cdot$$

Therefore, $\mathrm{Var}[C_n]$ is also $o(1)$ as $n \to \infty$, which establishes the result. □

PROOF OF LEMMA B.5. Using first Cauchy-Schwarz inequality and then Chebychev's inequality, we obtain

$$\text{(B.8)}\quad \sum_{\ell=1}^{n}\mathrm{E}\big[D_{n\ell}^2\mathbb{I}[|D_{n\ell}|>\varepsilon]\big] \leq \sum_{\ell=1}^{n}\sqrt{\mathrm{E}\big[D_{n\ell}^4\big]\mathrm{P}\big[|D_{n\ell}|>\varepsilon\big]} \leq \frac{1}{\varepsilon}\sum_{\ell=1}^{n}\sqrt{\mathrm{E}\big[D_{n\ell}^4\big]\mathrm{Var}\big[D_{n\ell}\big]}.$$

Recalling that $\sigma_{n\ell}^2 = \mathrm{E}\big[D_{n\ell}^2\,|\,\mathcal{F}_{n,\ell-1}\big]$, (B.6) yields

$$\begin{aligned} \mathrm{Var}[D_{n\ell}] \leq \mathrm{E}\big[D_{n\ell}^2\big] = \mathrm{E}[\sigma_{n\ell}^2] &\leq \frac{2p_n}{n\sigma_n^2}\Big(\tilde{e}_{n2}^2 + ne_{n1}^2\tilde{e}_{n2} + \frac{f_{n2}^2}{p_n-1}\Big) \\ &\leq C\Big(\frac{p_n\tilde{e}_{n2}^2}{np_n\tilde{e}_{n2}^2} + \frac{p_ne_{n1}^2\tilde{e}_{n2}^2}{2np_ne_{n1}^2\tilde{e}_{n2}} + \frac{f_{n2}^2}{nf_{n2}^2}\Big) \leq \frac{C}{n}\cdot \end{aligned}$$

Besides, using (B.4), the inequality $(a+b)^4 \leq 8(a^4+b^4)$ and the fact that $\sigma_n^2 \geq 2np_n e_{n1}^2 \tilde{e}_{n2}$, we obtain

$$\begin{aligned}
\mathrm{E}\big[D_{n\ell}^4\big] &\leq \frac{Cp_n^2}{n^4\sigma_n^4}\Bigg(\mathrm{E}\Bigg[\bigg(\sum_{i=1}^{\ell-1}(\mathbf{X}_{ni}-e_{n1}\boldsymbol{\theta}_n)'(\mathbf{X}_{n\ell}-e_{n1}\boldsymbol{\theta}_n)\bigg)^4\Bigg]+n^4e_{n1}^4\mathrm{E}\big[(\mathbf{X}_{n\ell}'\boldsymbol{\theta}_n-e_{n1})^4\big]\Bigg) \\
&\leq \frac{Cp_n^2}{n^4\sigma_n^4}\,\mathrm{E}\Bigg[\bigg(\sum_{i=1}^{\ell-1}(\mathbf{X}_{ni}-e_{n1}\boldsymbol{\theta}_n)'(\mathbf{X}_{n\ell}-e_{n1}\boldsymbol{\theta}_n)\bigg)^4\Bigg]+\frac{C\tilde{e}_{n4}}{n^2\tilde{e}_{n2}^2}. \qquad \text{(B.9)}
\end{aligned}$$

Applying Lemma B.3, we have

$$\begin{aligned}
\mathrm{E}\Bigg[\bigg(\sum_{i=1}^{\ell-1}(\mathbf{X}_{ni}-e_{n1}\boldsymbol{\theta}_n)'(\mathbf{X}_{n\ell}-e_{n1}\boldsymbol{\theta}_n)\bigg)^4\Bigg] &= (\ell-1)\left(\tilde{e}_{n4}^2+\frac{6}{p_n-1}\mathrm{E}\big[v_{ni}^2(u_{ni}-e_{n1})^2\big]^2+\frac{3f_{n4}^2}{p_n^2-1}\right) \\
&\quad +3(\ell-1)(\ell-2)\left(\tilde{e}_{n2}^2\tilde{e}_{n4}+\frac{2\tilde{e}_{n2}f_{n2}}{p_n-1}\mathrm{E}\big[v_{ni}^2(u_{ni}-e_{n1})^2\big]+\frac{f_{n2}^2f_{n4}}{(p_n-1)^2}\right),
\end{aligned}$$

By Cauchy-Schwarz, this yields

$$\begin{aligned}
&\frac{p_n^2}{n^4\sigma_n^4}\,\mathrm{E}\Bigg[\bigg(\sum_{i=1}^{\ell-1}(\mathbf{X}_{ni}-e_{n1}\boldsymbol{\theta}_n)'(\mathbf{X}_{n\ell}-e_{n1}\boldsymbol{\theta}_n)\bigg)^4\Bigg] \\
&\quad\leq \frac{1}{n^3\sigma_n^4}\left(p_n^2\tilde{e}_{n4}^2+6p_nf_{n4}\tilde{e}_{n4}+3f_{n4}^2\right)+\frac{3}{n^2\sigma_n^4}\left(p_n^2\tilde{e}_{n2}^2\tilde{e}_{n4}+2p_n\tilde{e}_{n2}f_{n2}f_{n4}^{1/2}\tilde{e}_{n4}^{1/2}+f_{n2}^2f_{n4}\right) \\
&\quad\leq \frac{C}{n^3}\left(\frac{\tilde{e}_{n4}^2}{\tilde{e}_{n2}^4}+\frac{f_{n4}\tilde{e}_{n4}}{f_{n2}^2\tilde{e}_{n2}^2}+\frac{f_{n4}^2}{f_{n2}^4}\right)+\frac{C}{n^2}\left(\frac{\tilde{e}_{n4}}{\tilde{e}_{n2}^2}+\left(\frac{f_{n4}\tilde{e}_{n4}}{f_{n2}^2\tilde{e}_{n2}^2}\right)^{1/2}+\frac{f_{n4}}{f_{n2}^2}\right).
\end{aligned}$$

Plugging into (B.9), we conclude that

$$\begin{aligned}
\mathrm{E}\big[D_{n\ell}^4\big] &\leq \frac{C}{n^3}\left(\frac{\tilde{e}_{n4}^2}{\tilde{e}_{n2}^4}+\frac{f_{n4}\tilde{e}_{n4}}{f_{n2}^2\tilde{e}_{n2}^2}+\frac{f_{n4}^2}{f_{n2}^4}\right)+\frac{C}{n^2}\left(\frac{\tilde{e}_{n4}}{\tilde{e}_{n2}^2}+\left(\frac{f_{n4}\tilde{e}_{n4}}{f_{n2}^2\tilde{e}_{n2}^2}\right)^{1/2}+\frac{f_{n4}}{f_{n2}^2}\right) \\
&\leq \frac{C}{n^3}\left(\frac{\tilde{e}_{n4}}{\tilde{e}_{n2}^2}+\frac{f_{n4}}{f_{n2}^2}\right)^2+\frac{C}{n^2}\left(\frac{\tilde{e}_{n4}^{1/2}}{\tilde{e}_{n2}}+\frac{f_{n4}^{1/2}}{f_{n2}}\right)^2\leq\frac{C}{n}\left(\frac{\tilde{e}_{n4}}{n\tilde{e}_{n2}^2}+\frac{f_{n4}}{nf_{n2}^2}\right), \qquad \text{(B.10)}
\end{aligned}$$

which, by assumption, is $o(1/n)$ as $n\to\infty$.

All majorations and o's above being uniform in ℓ, we finally obtain that

$$\sum_{\ell=1}^{n}\sqrt{\mathrm{E}\big[D_{n\ell}^4\big]\mathrm{Var}\big[D_{n\ell}\big]}\leq C\Big(n\max_{\ell=1,\ldots,n}\mathrm{E}\big[D_{n\ell}^4\big]\Big)^{1/2}\to 0$$

as $n\to\infty$, which, in view of (B.8), establishes the result. □

B.3. Proof of Theorems 4.2 and 4.3.

Proof of Theorem 4.2. Applying the main result from Heyde and Brown (1970) with $\delta = 1$, we obtain that, for n large enough,

$$\text{(B.11)}\quad \sup_{z\in\mathbb{R}}\left|\mathrm{P}_{F_n}\left[\frac{R_n^{\mathrm{St}} - \mathrm{E}[R_n^{\mathrm{St}}]}{\sigma_n} \leq s_n z\right] - \Phi(z)\right| \leq \frac{C}{s_n^4}\left(\sum_{\ell=1}^n \mathrm{E}[D_{n\ell}^4] + \mathrm{Var}\left[\sum_{\ell=1}^n \sigma_{nl}^2\right]^2\right)^{1/5}.$$

An inspection of the proof of Lemma B.4(ii) reveals that

$$\mathrm{Var}\left[\sum_{\ell=1}^n \sigma_{nl}^2\right] \leq C\left(\frac{\tilde{e}_{n4}}{n\tilde{e}_{n2}^2} + \left(\min\left(\frac{p_n\tilde{e}_{n2}^2}{f_{n2}^2}, \frac{\tilde{e}_{n2}}{ne_{n1}^2}\right)\right)^2 + \min\left(\frac{p_n\tilde{e}_{n2}^2}{f_{n2}^2}, \frac{\tilde{e}_{n2}}{ne_{n1}^2}\right) + \frac{f_{n4}}{nf_{n2}^2} + \frac{1}{p_n}\right).$$

Plugging this and (B.10) in (B.11) and using the fact that s_n converges to one as $n \to \infty$ provides the result. □

Proof of Theorem 4.3. First, note that in view of Theorem 4.2,

$$\begin{aligned}
\left|\mathrm{P}_{F_n}^{(n)}[R_n^{\mathrm{St}} > z_\alpha] - \left(1 - \Phi\left(\frac{z_\alpha - \mathrm{E}[R_n^{\mathrm{St}}]}{s_n\sigma_n}\right)\right)\right| &= \left|\mathrm{P}_{F_n}^{(n)}[R_n^{\mathrm{St}} \leq z_\alpha] - \Phi\left(\frac{z_\alpha - \mathrm{E}[R_n^{\mathrm{St}}]}{s_n\sigma_n}\right)\right| \\
&\leq \left|\mathrm{P}_{F_n}^{(n)}\left[\frac{R_n^{\mathrm{St}} - \mathrm{E}[R_n^{\mathrm{St}}]}{s_n\sigma_n} \leq \frac{z_\alpha - \mathrm{E}[R_n^{\mathrm{St}}]}{s_n\sigma_n}\right] - \Phi\left(\frac{z_\alpha - \mathrm{E}[R_n^{\mathrm{St}}]}{s_n\sigma_n}\right)\right| \\
&\leq \sup_{z\in\mathbb{R}}\left|\mathrm{P}_{F_n}^{(n)}\left[\frac{R_n^{\mathrm{St}} - \mathrm{E}[R_n^{\mathrm{St}}]}{s_n\sigma_n} \leq z\right] - \Phi(z)\right| \to 0
\end{aligned}$$

as $n \to \infty$. Recalling that $s_n \to 1$ as $n \to \infty$, this implies that

$$\lim_{n\to\infty}\mathrm{P}_{F_n}^{(n)}[R_n^{\mathrm{St}} > z_\alpha] = 1 - \Phi\left(\lim_{n\to\infty}\frac{R_n^{\mathrm{St}} - \mathrm{E}[R_n^{\mathrm{St}}]}{\sigma_n}\right) = 1 - \Phi\left(z_\alpha - \frac{\tau_1^2}{\sqrt{2}}\right),$$

as was to be proved. □

SUPPLEMENTARY MATERIAL

Supplement to "Testing Uniformity on High-Dimensional Spheres against Contiguous Rotationally Symmetric Alternatives"
(doi: completed by the typesetter; .pdf). In this supplementary article, we derive the fixed-p asymptotic non-null distribution of the Rayleigh test statistic in (3.3), and we show that the conditions (i)-(iii) of Theorem 4.1 always hold under FvML distributions.

REFERENCES

Banerjee, A., Dhillon, I., Ghosh, J. and Sra, S. (2003). Generative model-based clustering of directional data. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining* 19–28.

Banerjee, A., Dhillon, I., Ghosh, J. and Sra, S. (2005). Clustering on the unit hypersphere using von Mises-Fisher distributions. *J. Mach. Learn. Res.* **6** 1345–1382.

Berthet, Q. and Rigollet, P. (2013). Optimal detection of sparse principal components in high dimension. *Ann. Statist.* **41** 1780–1815.

Cai, T., Fan, J. and Jiang, T. (2013). Distributions of angles in random packing on spheres. *J. Mach. Learn. Res.* **14** 1837–1864.

Cai, T. and Jiang, T. (2012). Phase transition in limiting distributions of coherence of high-dimensional random matrices. *J. Multivariate Anal.* **107** 24–39.

Cutting, C., Paindaveine, D. and Verdebout, T. (2015). Supplement to "Testing uniformity on high-dimensional spheres against contiguous rotationally symmetric alternatives". Submitted.

Dryden, I. L. (2005). Statistical analysis on high-dimensional spheres and shape spaces. *Ann. Statist.* **33** 1643–1665.

Heyde, C. C. and Brown, B. M. (1970). On the departure from normality of a certain class of martingales. *Ann. Mathem. Statist.* **41** 2161–2165.

Ley, C., Paindaveine, D. and Verdebout, T. (2015). High-dimensional tests for spherical location and spiked covariance. *J. Multivariate Anal.,* under revision.

Mardia, K. V. and Jupp, P. E. (2000). *Directional Statistics*. John Wiley & Sons.

Onatski, A., Moreira, M. J. and Hallin, M. (2013). Asymptotic power of sphericity tests for high-dimensional data. *Ann. Statist.* **41** 1204–1231.

Onatski, A., Moreira, M. J. and Hallin, M. (2014). Signal detection in high dimension: the multispiked case. *Ann. Statist.* **42** 225–254.

Paindaveine, D. and Verdebout, T. (2015). Universal asymptotics for high-dimensional sign tests. *Bernoulli,* to appear.

Rayleigh, L. (1919). On the problem of random vibrations and random flights in one, two and three dimensions. *Phil. Mag.* **37** 321–346.

Saw, J. G. (1978). A family of distributions on the m-sphere and some hypothesis tests. *Biometrika* **65** 69–73.

Wang, T., Berthet, Q. and Samworth, R. J. (2014). Statistical and computational trade-offs in estimation of sparse principal components. *arXiv:1408.5369*.

Wang, L., Peng, B. and Li, R. (2015). A high-dimensional nonparametric multivariate test for mean vector. *J. Amer. Statist. Assoc.,* to appear.

Université libre de Bruxelles
Département de Mathématique
Bld du Triomphe, Campus Plaine, CP210
B-1050, Brussels
Belgium
E-mail: Christine.Cutting@ulb.ac.be
tverdebo@ulb.ac.be
URL: http://tverdebo.ulb.ac.be

Université libre de Bruxelles
ECARES and Département de Mathématique
Avenue F.D. Roosevelt, 50
ECARES, CP114/04
B-1050, Brussels
Belgium
E-mail: dpaindav@ulb.ac.be
URL: http://homepages.ulb.ac.be/˜dpaindav