

# Testing uniformity on high-dimensional spheres: The non-null behaviour of the Bingham test

Christine Cutting<sup>a,\*</sup>, Davy Paindaveine<sup>a,b,†</sup> and Thomas Verdebout<sup>a,‡</sup>

<sup>a</sup>ECARES and Department of Mathematics, Université libre de Bruxelles, Brussels, Belgium.

*E-mail:* \**Christine.Cutting@ulb.ac.be;* †*Davy.Paindaveine@ulb.be;* ‡*Thomas.Verdebout@ulb.be* 

<sup>b</sup>Toulouse School of Economics, Université Toulouse Capitole, Toulouse, France

Received 10 February 2020; revised 28 November 2020; accepted 12 March 2021

**Abstract.** In the framework of axial data, the most classical test of uniformity on the unit sphere of  $\mathbb{R}^p$  is the Bingham (*Ann. Statist.* **2** (1974) 1201–1225) test. In this work, we study the non-null behaviour of this test in high-dimensional asymptotic scenarios where  $p = p_n$  diverges to infinity with the sample size *n*. We consider a semiparametric class of alternatives that includes Watson alternatives and we derive a local asymptotic normality property. Le Cam's third lemma reveals that the Bingham test is blind to the corresponding contiguous alternatives, though. By using martingale central limit theorems, we therefore study the behaviour of the Bingham test under more severe alternatives. Far from restricting to the aforementioned semiparametric alternatives, our results cover a broad class of rotationally symmetric alternatives, which allows us to consider non-axial alternatives, too. In every distributional framework we consider, the "detection threshold" of the Bingham test is identified and a comparison with the classical test of uniformity for non-axial data, namely the Rayleigh (*Philos. Mag.* **37** (1919) 321–346) test, is made possible. In the framework of axial data, we derive a lower bound on the minimax separation rate and establish that the Bingham test is minimax rate-optimal in the class of Watson distributions.

**Résumé.** Pour des données axiales, le test d'uniformité le plus classique sur la sphère unité de  $\mathbb{R}^p$  est le test de Bingham (*Ann. Statist.* **2** (1974) 1201–1225). Dans cet article, nous étudions la puissance de ce test dans des scénarios asymptotiques de grande dimension pour lesquels  $p = p_n$  diverge vers l'infini avec la taille d'échantillon *n*. Nous considérons une classe semiparamétrique de contre-hypothèses incluant les distributions de Watson et nous établissons une propriété de normalité locale asymptotique. Le troisième lemme de Le Cam révèle cependant que le test Bingham ne détecte pas les contre-hypothèses contiguës associées. En utilisant un théorème central limite pour des différences de martingales, nous étudions donc le comportement du test de Bingham sous des contre-hypothèses plus sévères. Loin de nous restreindre aux contre-hypothèses semiparamétriques mentionnées ci-dessus, nos résultats couvrent une large classe de contre-hypothèses à symétrie rotationnelle, ce qui nous permet de considérer également des contre-hypothèses non axiales. Dans chaque contexte distributionnel, nous obtenons le "seuil de détection" du test de Bingham, ce qui rend possible une comparaison avec le test classique d'uniformité pour des données non axiales, c'est-à-dire le test de Rayleigh (*Philos. Mag.* **37** (1919) 321–346). Dans le cas axial, nous déterminons une borne inférieure pour le taux de séparation minimax et établissons que le test de Bingham est minimax optimal dans la classe des distributions de Watson.

MSC2020 subject classifications: Primary 62H11; 62F05; secondary 62E20

*Keywords:* Directional statistics; High-dimensional statistics; Local asymptotic normality; Minimax separation rates; Rotationally symmetric distributions; Tests of uniformity

# 1. Introduction

We consider the problem of testing uniformity over the *p*-dimensional unit sphere  $S^{p-1} := {\mathbf{x} \in \mathbb{R}^p : ||\mathbf{x}||^2 = \mathbf{x}'\mathbf{x} = 1}$ . More specifically, on the basis of mutually independent observations  $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$  with the same distribution  $\mathcal{D}_n$  on  $S^{p_n-1}$  (it will become clear below why we adopt a triangular array notation), we consider the problem of testing the null hypothesis that  $\mathcal{D}_n$  is the uniform distribution. This is a fundamental problem in directional statistics that has been considered a lot recently; for instance, [30] considered the problem in the presence of noisy data, [19] used nearest neighbor distances, [42] obtained Bayesian optimality properties of some tests while [26] and [29] provided new Sobolev tests. This problem is also discussed at length in strictly all textbooks in the field; see, among many others, [20,31], and [32]. For a most recent account, we refer to the review paper [23].

In directional statistics, one often discriminates between axial and non-axial distributions. An axial distribution is antipodally symmetric, in the sense that it attributes equal probability masses to antipodal regions on the sphere. Such distributions need be considered when facing *axial data*, that is, in applications that do not offer genuine locations on the sphere but rather undirected axes (an example is given by the directions of optical axes in quartz crystals; see, e.g., [32]). There is a large literature considering axial distributions and inference for axial data; see, among many others, [2,6,8,18, 28,41,43], and [44]. For the problem of testing uniformity on the unit sphere, monographs discriminate between tests of uniformity for axial data and tests of uniformity for non-axial data, that are designed to show power primarily against axial alternatives or non-axial alternatives, such as Watson distributions or Fisher–von Mises–Langevin (FvML) distributions, respectively. In the classical setup where  $p_n \equiv p$ , the textbook test of uniformity for axial data is the Bingham [8] test, that rejects the null hypothesis of uniformity at asymptotic level  $\alpha$  when

$$Q_n = \frac{np_n(p_n+2)}{2} \left( \operatorname{tr} \left[ \mathbf{S}_n^2 \right] - \frac{1}{p_n} \right) > \chi^2_{d_{p_n}, 1-\alpha},$$

where  $\mathbf{S}_n := (1/n) \sum_{i=1}^n \mathbf{X}_{ni} \mathbf{X}'_{ni}$  is the covariance matrix of the observations (using the centre of the sphere as a specified location), tr[**A**] is the trace of the matrix **A**, and  $\chi^2_{d_p,1-\alpha}$  denotes the upper  $\alpha$ -quantile of the chi-square distribution with  $d_p := p(p+1)/2 - 1$  degrees of freedom. A modification of this test to bring the null distribution close to its asymptotic distribution has been proposed in [28], while [22] used it to build tests for rotational symmetry. For Euclidean data, this test can also be seen as the sign test of sphericity from [24].

Nowadays, it is more and more common to face high-dimensional spherical data, for instance in brain shape modeling ([18]) or text mining ([3,4]). In line with this, inference on high-dimensional spheres has been much considered in the last decades. While other inference problems have also been considered (see, e.g., [37] for high-dimensional location problems), most of the focus has been on the problem of testing uniformity on the unit sphere. An asymptotic investigation of this problem requires that the dimension  $p_n$  diverges to infinity with n (see, e.g., [9] and [15]), which makes it necessary to adopt the triangular framework above (some other works, including [11,12], actually rather consider a fixed-n large-p asymptotic scenario; see also the monograph [13]). Further relevant references include [14], that defines a test of uniformity that can be applied in high dimensions, or [27], where uniformity testing on high-dimensional spheres finds applications in outlier detection.

In the present work, we consider the problem of testing uniformity on  $S^{p_n-1}$ , in a double asymptotic scenario where  $p_n$  diverges to infinity with *n*. A result from [35] (see their Theorem 2.5) implies that, under the null hypothesis of uniformity,

$$Q_n^{\text{St}} = \frac{Q_n - d_{p_n}}{\sqrt{2d_{p_n}}} = \frac{p_n}{n} \sum_{\substack{i < j \\ i, j = 1}}^n \left\{ \left( \mathbf{X}'_{ni} \mathbf{X}_{nj} \right)^2 - \frac{1}{p_n} \right\} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1),$$
(1)

as soon as  $p_n$  diverges to infinity with n, where  $\stackrel{\mathcal{D}}{\rightarrow}$  denotes weak convergence and  $\mathcal{N}(0, 1)$  is the standard normal distribution. Therefore, denoting as  $z_{\alpha}$  the upper  $\alpha$ -quantile of this distribution, the Bingham test, that equivalently rejects the null hypothesis of uniformity at asymptotic level  $\alpha$  when

$$Q_n^{\text{St}} = \frac{Q_n - d_{p_n}}{\sqrt{2d_{p_n}}} > \frac{\chi_{d_{p_n}, 1-\alpha}^2 - d_{p_n}}{\sqrt{2d_{p_n}}} \left( = z_\alpha + o(1) \right), \tag{2}$$

has asymptotic size  $\alpha$  under the null hypothesis, irrespective of the rate at which  $p_n$  diverges to infinity with n. This shows that the original, low-dimensional, Bingham test does not need any modification to achieve asymptotically the target null size in high dimensions, hence is, in this sense, robust to high-dimensionality. The resulting test is therefore a natural test of uniformity in high dimensions, particularly so in the context of axial data – just as the Rayleigh [39] test is a natural test of uniformity in high dimensions in the framework of non-axial data; see [15].

Now, the robustness to high-dimensionality above is a null result, that might be achieved at the expense of power. To date, there is unfortunately no guarantee that the Bingham test exhibits power in high dimensions. This is in contrast with (a) low dimensions – where the non-null behaviour of the Bingham test is well understood (see, e.g., [17] and [32]) – and with (b) other classical tests in high dimensions ([15] extensively studied the non-null behaviour of the Rayleigh test of uniformity in high dimensions). In the present paper, we therefore study the non-null behaviour of the Bingham test of uniformity in a double asymptotic scenario where the dimension  $p_n$  diverges to infinity with n. We consider both axial and non-axial alternatives to uniformity.

Since the Bingham test is primarily designed for axial alternatives, we first consider a semiparametric class of alternatives of this type, that includes the classical Watson alternatives. In this particular class, we obtain a local asymptotic normality (LAN) result and show that the Bingham test is unfortunately blind to the corresponding contiguous alternatives (Section 2). It is then natural to investigate whether or not this test shows power against more severe alternatives. To tackle this question, we conduct a systematic investigation, relying on martingale central limit theorems, of the non-null behaviour of the Bingham test under a very broad class of rotationally symmetric alternatives. This identifies the "detection threshold" of the Bingham test, that discriminates between alternatives to which the test will be blind and those under which it will be consistent (Section 3). We then apply our general non-null results to describe the non-null behaviour of the Bingham test in the aforementioned semiparametric class of axial alternatives; we complement our analysis by deriving a lower bound on the corresponding minimax separation rate and by proving that the Bingham test is minimax rate-optimal in the class of Watson distributions (Section 4). Since the results of Section 3 do not require the underlying alternatives to be axial, we also describe the non-null behaviour of the Bingham test in a semiparametric class of non-axial alternatives, that contains the classical FvML alternatives (Section 5). Both in the axial and non-axial cases, our results allow for a comparison with the Rayleigh test in terms of consistency rates and asymptotic power. Throughout, our asymptotic findings are confirmed via Monte Carlo exercises. Finally, a wrap up is provided (Section 6). An Appendix collects the proofs.

# 2. A LAN result in the high-dimensional axial case

In this section, we consider a semiparametric extension of the class of Watson distributions, containing axial distributions admitting a density of the form (throughout, densities over  $S^{p-1}$  are with respect to the surface area measure)

$$\mathbf{x} \mapsto \frac{c_{p,\kappa,f} \Gamma(\frac{p-1}{2})}{2\pi^{(p-1)/2}} f\left(\kappa\left(\mathbf{x}'\boldsymbol{\theta}\right)^2\right), \quad \text{with } c_{p,\kappa,f} = 1/\int_{-1}^{1} (1-s^2)^{(p-3)/2} f\left(\kappa s^2\right) ds, \tag{3}$$

where  $\theta \in S^{p-1}$ ,  $\kappa \in \mathbb{R}$ , and f belongs to the class  $\mathcal{F}$  of functions from  $\mathbb{R}$  to  $\mathbb{R}^+$  that are monotone increasing, twice differentiable at 0, and satisfy f(0) = f'(0) = 1 (these restrictions on f ensure that  $\kappa$  and f are identifiable); here,  $\Gamma(\cdot)$ denotes the usual Euler Gamma function. The expression of the normalizing constant  $c_{p,\kappa,f}$  results from the fact that if  $\mathbf{X}$  admits the density in (3), then  $\mathbf{X}'\theta$  admits the density  $s \mapsto c_{p,\kappa,f}(1-s^2)^{(p-3)/2}f(\kappa s^2)\mathbb{I}[s \in [-1, 1]]$ , where  $\mathbb{I}[A]$ is the indicator function of A. The density above attributes equal probabilities to antipodal regions on the sphere, hence is suitable for axial data, that is, data where only undirected axes, or equivalently pairs of antipodal points, are observed. In line with this, the particular case obtained with  $f(s) = \exp(s)$  provides the Watson distributions, which are the most commonly used rotationally symmetric axial distributions. The parameter  $\kappa$  is a concentration parameter: increasingly large positive values of  $\kappa$  provide *bipolar distributions*, that are more and more concentrated (symmetrically) about both poles  $\pm \theta$ , whereas increasingly large negative values of  $\kappa$  are associated with *girdle distributions*, that more and more concentrate over the hyper-equator { $\mathbf{x} \in S^{p-1} : \mathbf{x}'\theta = 0$ }. Irrespective of f, the value  $\kappa = 0$  corresponds to the uniform distribution over  $S^{p-1}$ . In the semiparametric model associated with the densities in (3), testing uniformity versus nonuniformity over the sphere consists in testing  $\mathcal{H}_0 : \kappa = 0$  versus  $\mathcal{H}_1 : \kappa \neq 0$ .

The distributional framework above provides natural alternatives to uniformity. Accordingly, for any sequence of positive integers  $(p_n)$ , sequence  $(\theta_n)$  such that  $\theta_n \in S^{p_n-1}$  for any n, real sequence  $(\kappa_n)$ , and  $f \in \mathcal{F}$ , we will denote as  $P_{\theta_n,\kappa_n,f}^{(n)}$  the hypothesis under which  $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$  form a random sample from the density (3) with the corresponding parameters  $\theta_n$ ,  $\kappa_n$ , and f (to keep the notation as light as possible, we do not stress the dimension  $p_n$  in the notation). For simplicity, we write  $P_0^{(n)} := P_{\theta_n,0,f}^{(n)}$  for the uniform particular case. In the high-dimensional framework considered in this work, we then have the following LAN result (see Appendix A for a proof).

**Theorem 2.1.** Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$  and  $(\boldsymbol{\theta}_n)$  be a sequence such that  $\boldsymbol{\theta}_n \in S^{p_n-1}$  for any n. Let  $\kappa_n = \tau_n p_n / \sqrt{n}$ , where  $(\tau_n)$  is O(1), and fix  $f \in \mathcal{F}$ . Then, as  $n \to \infty$  under  $P_0^{(n)}$ ,

$$\Lambda_n = \log \frac{d\mathbf{P}_{\boldsymbol{\theta}_n,\kappa_n,f}^{(n)}}{d\mathbf{P}_0^{(n)}} = \tau_n \Delta_{\boldsymbol{\theta}_n} - \tau_n^2 + o_{\mathbf{P}}(1),$$

where  $\Delta_{\theta_n} := \sqrt{n}(p_n \theta'_n \mathbf{S}_n \theta_n - 1)$  is asymptotically normal with mean zero and variance two (here,  $\mathbf{S}_n$  refers to the covariance matrix defined in the Introduction).

This result shows that, for any  $f \in \mathcal{F}$ , the sequence of models  $\{P_{\theta_n,\kappa,f}^{(n)} : \kappa \in \mathbb{R}\}$  is locally asymptotically normal at  $\kappa = 0$  with central sequence  $\Delta_{\theta_n}$ , Fisher information  $\Gamma = 2$ , and contiguity rate  $p_n/\sqrt{n}$ . As a direct corollary, the sequences of

null hypotheses  $(\mathbf{P}_0^{(n)})$  and alternative hypotheses  $(\mathbf{P}_{\theta_n,\kappa_n,f}^{(n)})$  with  $\kappa_n = O(p_n/\sqrt{n})$  are mutually contiguous; that is, there do no exist consistent tests when testing uniformity against such alternatives. The result also readily implies that the Le Cam optimal test of uniformity in this model rejects the null hypothesis at asymptotic level  $\alpha$  whenever  $|\Delta_{\theta_n}| > \sqrt{2}z_{\alpha/2}$ . A routine application of Le Cam's third lemma then entails that, under  $\mathbf{P}_{\theta_n,\kappa_n,f}^{(n)}$  with  $\kappa_n = \tau_n p_n/\sqrt{n}$  and  $\tau_n \to \tau$ , the central sequence  $\Delta_{\theta_n}$  is asymptotically normal with mean  $2\tau$  and variance 2, which provides the asymptotic power

$$\lim_{n \to \infty} \mathcal{P}_{\boldsymbol{\theta}_n, \kappa_n, f}^{(n)} \left[ |\Delta_{\boldsymbol{\theta}_n}| > \sqrt{2} z_{\alpha/2} \right] = 2 - \sum_{s \in \{-1, 1\}} \Phi(z_{\alpha/2} + s\sqrt{2}\tau), \tag{4}$$

where  $\Phi$  denotes the cumulative distribution function of the standard normal distribution. This optimal test can thus detect these contiguous alternatives and shows a symmetric power pattern against girdle alternatives ( $\tau < 0$ ) and bipolar alternatives ( $\tau > 0$ ).

These results show that optimal testing of uniformity for high-dimensional axial data is based on  $S_n$ , at least in the considered model. This is to be compared with the non-axial case investigated in [15], where optimal testing of uniformity in high dimensions is rather based on  $\bar{X}_n := n^{-1} \sum_{i=1}^n X_{ni}$ . Another difference between the axial and non-axial cases materializes in the corresponding contiguity rates: the contiguity rate  $p_n/\sqrt{n}$  in the axial case above is worse than the contiguity rate  $\sqrt{p_n}/\sqrt{n}$  obtained in the non-axial case; see Theorem 5.1 below. In high dimensions, it is therefore more challenging to detect axial departures from uniformity than non-axial ones.

Now, it is important to note that the optimal test above is of little practical relevance because it requires the polar axis  $\pm \theta_n$  to be known, which is almost never the case in applications. In contrast, the Bingham test does not require such knowledge. It is therefore natural to derive the asymptotic power of this test against the contiguous alternatives considered in (4). Interestingly, this can be done via Le Cam's third lemma. We have the following result (see Appendix A for a proof).

**Proposition 2.1.** Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$  and  $(\boldsymbol{\theta}_n)$  be a sequence such that  $\boldsymbol{\theta}_n \in S^{p_n-1}$  for any n. Let  $\kappa_n = \tau_n p_n / \sqrt{n}$ , with  $(\tau_n) \to \tau$ , and fix  $f \in \mathcal{F}$ . Then,  $\operatorname{Cov}[Q_n^{\operatorname{St}}, \Lambda_n] = o(1)$  as  $n \to \infty$  under  $\mathbb{P}_0^{(n)}$ , so that Le Cam's third lemma implies that  $Q_n^{\operatorname{St}}$  remains asymptotically standard normal under  $\mathbb{P}_{\boldsymbol{\theta}_n,\kappa_n,f}^{(n)}$ .

A direct corollary is that, unlike the optimal test above, the Bingham test does not show power against the contiguous alternatives considered in (4). In other words, the Bingham test is not rate consistent in high dimensions, in contrast to what happens in low dimensions; see [17]. Consequently, a natural question, that we tackle in the next two sections, is whether or not there exist more severe alternatives that can be detected by the Bingham test in high dimensions. The LAN framework is not suitable to investigate this point, which leads us to perform a direct asymptotic analysis relying on martingale central limit theorems.

#### 3. Non-null results under general rotationally symmetric alternatives

We now study the non-null asymptotic behaviour of the Bingham test under general rotationally symmetric alternatives. Recall that the distribution of a random vector **X** with values in  $S^{p-1}$  is said to be *rotationally symmetric about*  $\theta \in S^{p-1}$  if **OX** is equal in distribution to **X** for any  $p \times p$  orthogonal matrix **O** such that  $O\theta = \theta$ . In this case,  $\mathbf{S} := (\mathbf{I}_p - \theta\theta')\mathbf{X}/\|(\mathbf{I}_p - \theta\theta')\mathbf{X}\|$  is uniformly distributed over the hyper-equator  $\{\mathbf{x} \in S^{p-1} : \mathbf{x}'\theta = 0\}$  and is independent of  $\mathbf{X}'\theta$ . Therefore, the distribution of **X** is then completely characterized by  $(\theta, F)$ , where *F* is the cumulative distribution function of  $\mathbf{X}'\theta$ . Depending on the particular rotationally symmetric distribution considered,  $\theta$  and *F* may be unidentifiable, but this possible lack of identifiability will not play any role below in the context of testing uniformity over  $S^{p-1}$ .

In the rest of this section, we consider sequences of alternatives that are described by triangular arrays of observations  $\mathbf{X}_{ni}$ , i = 1, ..., n, n = 1, 2, ... such that, for any n,  $\mathbf{X}_{n1}$ ,  $\mathbf{X}_{n2}$ , ...,  $\mathbf{X}_{nn}$  are mutually independent and share a common rotationally symmetric distribution on  $S^{p_n-1}$ . We denote by  $P_{\theta_n, F_n}^{(n)}$  the corresponding hypothesis when  $\mathbf{X}_{n1}$  is rotationally symmetric about  $\theta_n$  and  $u_{n1} := \mathbf{X}'_{n1}\theta_n$  has cumulative distribution function  $F_n$ . The following result provides the expectation and variance of  $Q_n^{\text{St}}$  under such sequences of hypotheses (see Appendix B for a proof).

**Proposition 3.1.** Let  $(p_n)$  be a sequence of positive integers,  $(\boldsymbol{\theta}_n)$  be a sequence such that  $\boldsymbol{\theta}_n \in S^{p_n-1}$  for any n, and  $(F_n)$  be a sequence of cumulative distribution functions on [-1, 1]. Write  $e_{n\ell} := e_{n\ell}(F_n) := \mathbb{E}[u_{n1}^{\ell}]$  and  $f_{n\ell} := f_{n\ell}(F_n) := \mathbb{E}[v_{n1}^{\ell}]$  for the  $\ell$ th moments of  $u_{n1}$  and  $v_{n1} = \sqrt{1 - u_{n1}^2}$  under  $\mathbb{P}_{\boldsymbol{\theta}_n, F_n}^{(n)}$ . Then, letting  $g_{n2} := e_{n2} - (1/p_n)$ ,

$$E[Q_n^{St}] = \frac{p_n^2(n-1)}{2(p_n-1)}g_{n2}^2$$
(5)

and

$$\operatorname{Var}[\mathcal{Q}_{n}^{\operatorname{St}}] = \frac{(n-1)p_{n}^{2}}{2n} \left\{ e_{n4}^{2} + \frac{6}{p_{n}-1}(e_{n2}-e_{n4})^{2} + \frac{3f_{n4}^{2}}{p_{n}^{2}-1} - \left(\frac{p_{n}}{p_{n}-1}g_{n2}^{2} + \frac{1}{p_{n}}\right)^{2} \right\} + \frac{(n-1)(n-2)p_{n}^{4}}{n(p_{n}-1)^{2}}(e_{n4}-e_{n2}^{2})g_{n2}^{2}$$
(6)

under  $\mathbf{P}_{\boldsymbol{\theta}_n, F_n}^{(n)}$ .

Under the null hypothesis of uniformity (which we still denote as  $P_0^{(n)}$ ), the random variable  $u_{n1} = \mathbf{X}'_{n1}\boldsymbol{\theta}_n$  has, irrespective of  $\boldsymbol{\theta}_n$ , a symmetric distribution over [-1, 1] that is such that  $u_{n1}^2$  is Beta distributed with parameters 1/2 and  $(p_n - 1)/2$  (see, e.g., [33], Theorem 1.5.7(ii)), so that

$$e_{n2} = \frac{1}{p_n}$$
 and  $e_{n4} = \frac{3}{p_n(p_n+2)}$ 

Using these values along with the identities  $f_{n2} = 1 - e_{n2}$  and  $f_{n4} = 1 - 2e_{n2} + e_{n4}$ , Proposition 3.1 shows that  $\mathbb{E}[Q_n^{St}] = 0$ and  $\operatorname{Var}[Q_n^{St}] = (n-1)(p_n-1)/(n(p_n+2))$  under  $\mathbb{P}_0^{(n)}$ , which is compatible with the null asymptotic normality result in (1). A key step to study the non-null asymptotic behaviour of the Bingham test in high dimensions is to extend this asymptotic normality result to the rotationally symmetric alternatives considered in this section. This is the topic of the following theorem (the proof, that is long and tedious, relies on martingale central limit theorems; see Appendix B).

**Theorem 3.1.** Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$  and  $(\theta_n)$  be a sequence such that  $\theta_n \in S^{p_n-1}$  for any n. Assume that the sequence  $(F_n)$  is such that, as  $n \to \infty$ , (a)  $e_{n4} = o(1/p_n)$ , (b)  $e_{n8} = o(n^{2/3}/p_n^2)$ , and (c)  $g_{n2} = O(1/\sqrt{np_n})$ . Then,

$$\frac{Q_n^{\text{St}} - \mathbb{E}[Q_n^{\text{St}}]}{\sqrt{\text{Var}[Q_n^{\text{St}}]}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1)$$

as  $n \to \infty$  under  $\mathbf{P}_{\boldsymbol{\theta}_n, F_n}^{(n)}$ .

This general asymptotic normality result extends the null one in (1), since Proposition 3.1 implies that  $E[Q_n^{St}] = 0$ and  $Var[Q_n^{St}] = 1 + o(1)$  under  $P_0^{(n)}$ . It is worth pointing out that, like the result in (1), Theorem 3.1 in principle does not impose restrictions on the rate at which  $p_n$  diverges to infinity with n (some restrictions may arise when considering some particular alternatives, though; see, e.g., Theorems 4.1 and 5.2 below). More importantly, we can now state the main result of this paper, that describes the non-null behaviour of the Bingham test statistic under general rotationally symmetric alternatives (see Appendix B for a proof).

**Theorem 3.2.** Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$  and  $(\theta_n)$  be a sequence such that  $\theta_n \in S^{p_n-1}$  for any *n*. Assume that the sequence  $(F_n)$  is such that, as  $n \to \infty$ , (a)  $e_{n4} = o(1/p_n)$  and (b)  $e_{n8} = o(n^{2/3}/p_n^2)$ . Then, we have the following: (i) if (c)  $g_{n2} = o(1/\sqrt{np_n})$ , then

$$Q_n^{\operatorname{St}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

as  $n \to \infty$  under  $\mathbb{P}_{\theta_n, F_n}^{(n)}$ ; (ii) if (c)  $g_{n2} = \xi_n / \sqrt{np_n}$  with  $(\xi_n) \to \xi \neq 0$ , then

$$Q_n^{\operatorname{St}} \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\xi^2}{2}, 1\right)$$

as  $n \to \infty$  under  $\mathbb{P}_{\theta_n, F_n}^{(n)}$ ; (iii) if (c)  $\sqrt{np_n}|g_{n2}| \to \infty$ , then, for any real number M,

$$\mathbf{P}_{\boldsymbol{\theta}_n,F_n}^{(n)} \left[ Q_n^{\mathrm{St}} > M \right] \to 1 \tag{7}$$

as  $n \to \infty$  ((7) still holds if (a)–(b) are replaced with the single condition  $e_{n4} = o(ng_{n2}^2)$ , which, in case (iii), is weaker than (a)–(b)).

Since  $g_{n2} = e_{n2} - 1/p_n = 0$  under the null hypothesis of uniformity,  $|g_{n2}|$  can be read as a measure of the severity of the alternatives at hand. In this context, Theorem 3.2 states that the Bingham test is blind to alternatives for which  $g_{n2} =$  $\xi_n/\sqrt{np_n}$  with  $\xi_n \to 0$  (Part (i) of the result) and is consistent under alternatives for which  $g_{n2} = \xi_n/\sqrt{np_n}$  with  $|\xi_n| \to \infty$ (Part (iii) of the result). Therefore, using the terminology from [5], the "detection threshold" of the Bingham test in high dimensions is associated with alternatives in  $g_{n2} = \xi_n / \sqrt{np_n}$ , with  $\xi_n \to \xi \neq 0$ , under which the Bingham test achieves the non-trivial asymptotic power

$$\lim_{n \to \infty} \mathsf{P}_{\boldsymbol{\theta}_n, F_n}^{(n)} \left[ \mathcal{Q}_n^{\mathrm{St}} > z_\alpha \right] = 1 - \Phi\left( z_\alpha - \frac{\xi^2}{2} \right) \tag{8}$$

(Part (ii) of the result). We stress that the results of this section are very general: first, they do not require that the considered rotationally symmetric distributions admit a density of a specific form on the sphere (such as the Watsontype densities of the previous section), nor even that they admit a density at all. Second, they do not restrict to axial distributions, that is, they do not assume that  $F_n$  is the cumulative distribution function of a symmetric distribution over [-1, 1]; see Section 5.

We conducted the following simulation exercise to check the validity of Theorem 3.2. For each  $n \in \{100, 400, 800\}$ and each corresponding dimension  $p_n = \lfloor n^a \rfloor$  with  $a \in \{\frac{1}{2}, \frac{3}{4}, 1, \frac{5}{4}\}$  (leading to 12 combinations of *n* and  $p_n$ ), we generated M = 2500 independent random samples of size *n* from five rotationally symmetric distributions over  $S^{p_n}$ . Each rotationally symmetric distribution uses  $\theta_n = (1, 0, \dots, 0)' \in \mathbb{R}^{p_n}$  and  $F_n$  that is the cumulative distribution function of the Beta( $\alpha_n, \beta_n$ ) distribution, where the parameters

$$\alpha_n = \frac{1}{2} - \frac{p_n^2 g_{n2}}{2p_n(g_{n2} - 1) + 2} \quad \text{and} \quad \beta_n = \frac{p_n - 1}{2}$$
(9)

are based on

- (0)  $g_{n2} = 0$
- (i)  $g_{n2} = 1/(np_n)$
- (ii)<sub>a</sub>  $g_{n2} = \xi / \sqrt{np_n}$ , with  $\xi = 2$
- (ii)<sub>b</sub>  $g_{n2} = \xi / \sqrt{np_n}$ , with  $\xi = 3$  (iii)  $g_{n2} = 1/(n^{1/4} \sqrt{p_n})$ .

As the notation suggests,  $\alpha_n$  and  $\beta_n$  are such that  $g_{n2}$  in (9) is equal to  $e_{n2} - (1/p_n)$ , where  $e_{n2}$  is the second moment associated with  $F_n$ , so that this quantity  $g_{n2}$  coincides with the one in Theorem 3.2. Case (0) yields the null hypothesis of uniformity over  $S^{p_n-1}$ , whereas cases (i)–(iii) provide increasingly severe alternatives. For each of the 12 combinations  $(n, p_n)$  and each of these five cases, Figure 1 then reports kernel density estimates (obtained from the R command density with default parameter values) of the resulting M = 2500 values of the Bingham statistic  $Q_{n}^{\text{St}}$  (as well as raw histograms in case (ii)<sub>a</sub>). The figure also provides the densities of the corresponding asymptotic distributions in cases (0)–  $(ii)_b$ , that are obtained from Theorem 3.2(i)–(ii). Clearly, empirical results are in perfect agreement with the theory, not only for the matching between finite-sample and asymptotic distributions in cases (0)-(ii)<sub>b</sub> but also for the consistency behaviour in case (iii) (since kernel density estimates in this case shift to infinity as expected).

#### 4. Non-null results under Watson-type alternatives

We consider again the Watson-type alternatives to uniformity introduced in Section 2. As Proposition 2.1 showed, the Bingham test is blind to alternatives of the form  $\kappa_n = \tau_n p_n / \sqrt{n}$ , with  $(\tau_n) \to \tau$ . The objective of the present section is to apply the results of Section 3 to identify the Watson-type alternatives that can be detected by the Bingham test in high dimensions. In order to do so, we need to study the asymptotic behaviour of  $e_{n2}$ ,  $e_{n4}$  and  $e_{n8}$  in the semiparametric model at hand. This is the topic of the following proposition (see Appendix C for a proof).

**Proposition 4.1.** Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$ ,  $(\theta_n)$  be a sequence such that  $\theta_n \in S^{p_n-1}$ for any n, and  $(\kappa_n)$  be a real sequence that is  $o(p_n)$  as  $n \to \infty$ . Fix  $f \in \mathcal{F}$ . Then, under  $P_{\theta_n,\kappa_n,f}^{(n)}$ ,

$$e_{n2} = \frac{1}{p_n} + \frac{2\kappa_n}{p_n^2} + o\left(\frac{\kappa_n}{p_n^2}\right) \quad and \quad e_{n4} = \frac{3}{p_n^2} + o\left(\frac{1}{p_n^2}\right)$$

as  $n \to \infty$ , so that  $e_{n8} = O(1/p_n^2)$  as  $n \to \infty$ .



Fig. 1. For any sample size  $n \in \{100, 400, 800\}$  and any corresponding dimensions  $p = p_n = \lfloor n^a \rfloor$  with  $a \in \{\frac{1}{2}, \frac{3}{4}, 1, \frac{5}{4}\}$ , kernel estimates of the density of  $Q_n^{\text{St}}$  (solid lines) obtained from M = 2500 independent samples from the rotationally symmetric distributions associated with  $\theta_n = (1, 0, ..., 0)' \in \mathbb{R}^p$  and the cumulative distribution functions  $F_n$  associated with (0)–(iii) in page 572. In case (ii)<sub>a</sub>, raw histograms are provided. In cases (0)–(ii)<sub>b</sub>, the corresponding asymptotic densities, obtained from Theorem 3.2(i)–(ii), are also plotted (dashed lines).

The following result is then a corollary of Theorem 3.2 (see Appendix C for a proof).

**Theorem 4.1.** Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$  and  $(\boldsymbol{\theta}_n)$  be a sequence such that  $\boldsymbol{\theta}_n \in S^{p_n-1}$  for any n. Fix  $f \in \mathcal{F}$ . Then, we have the following: (i) if  $\kappa_n = o(p_n^{3/2}/\sqrt{n})$  and  $p_n = O(n)$  (or more generally, if  $\kappa_n = o(p_n^{3/2}/\sqrt{n})$  and  $\kappa_n = o(p_n)$ ), then

$$Q_n^{\operatorname{St}} \xrightarrow{D} \mathcal{N}(0,1)$$

as 
$$n \to \infty$$
 under  $\mathbb{P}_{\theta_n,\kappa_n,f}^{(n)}$ ; (ii) if  $\kappa_n = \tau_n p_n^{3/2} / \sqrt{n}$  with  $(\tau_n) \to \tau (\neq 0)$  and  $p_n = o(n)$ , then  
 $Q_n^{\text{St}} \xrightarrow{\mathcal{D}} \mathcal{N}(2\tau^2, 1)$ 
(11)

as  $n \to \infty$  under  $\mathbb{P}^{(n)}_{\theta_n,\kappa_n,f}$ ; (iii) if  $\sqrt{n}|\kappa_n|/p_n^{3/2} \to \infty$  and  $\kappa_n = o(p_n)$ , then, for any real number M,

$$\mathbf{P}_{\boldsymbol{\theta}_n,\kappa_n,f}^{(n)} \left[ Q_n^{\mathrm{St}} > M \right] \to 1$$

as  $n \to \infty$ .

Part (i) of this result confirms that the Bingham test is blind to alternatives in  $\kappa_n = o(p_n/\sqrt{n})$ , irrespective of the rate at which  $p_n$  diverges to infinity with n. Under some mild assumption on this rate, this extends to alternatives in  $\kappa_n = o(p_n^{3/2}/\sqrt{n})$ , whereas Part (iii) of the result shows that the Bingham test is consistent under alternatives such that  $\sqrt{n}|\kappa_n|/p_n^{3/2} \to \infty$ . For Watson-like alternatives, the detection threshold is thus  $\kappa_n \sim p_n^{3/2}/\sqrt{n}$ ; the corresponding asymptotic power, under  $P_{\theta_n,\kappa_n,f}^{(n)}$ , with  $\kappa_n = \tau_n p_n^{3/2}/\sqrt{n}$  and  $(\tau_n) \to \tau$ , is

$$\lim_{n\to\infty} \mathbf{P}_{\boldsymbol{\theta}_n,\kappa_n,f}^{(n)} \big[ Q_n^{\mathrm{St}} > z_{\alpha} \big] = 1 - \Phi \big( z_{\alpha} - 2\tau^2 \big).$$

These powers are symmetric for girdle-type alternatives ( $\tau < 0$ ) and bipolar alternatives ( $\tau > 0$ ), as it was already the case in low dimensions; see [17].

We conducted the following simulation. For any  $n \in \{100, 400, 800\}$ , we generated M = 2500 independent random samples of size *n* from the Watson distribution with dimension  $p_n = \lfloor \sqrt{n} \rfloor$ , location  $\boldsymbol{\theta}_n = (1, 0, \dots, 0)' \in \mathbb{R}^{p_n}$ , and concentration

- (i)  $\kappa_n = \tau p_n / \sqrt{n}$ , (ii)  $\kappa_n = \tau p_n^{3/2} / \sqrt{n}$ , (iii)  $\kappa_n = \tau p_n^{7/4} / \sqrt{n}$ ,

in each case with  $\tau = 0, 0.4, \dots, 2$ . Figure 2 reports the rejection frequencies of the Bingham test as well as the corresponding asymptotic powers obtained from Theorem 4.1. The figure also provides the rejection frequencies and asymptotic powers of the Rayleigh [39] test (these asymptotic powers are obtained from Theorem 5.1 in [15]). Clearly, the results are in excellent agreement with the theory. In particular, the Bingham test is blind to the alternatives associated with (i) and is consistent under those in (iii). Under the threshold ones in (ii), this test shows rejection frequencies that are close to the corresponding asymptotic powers. The Rayleigh test is blind to all alternatives considered, which is also in line with the theory: since  $e_{n1} = 0$  under any hypothesis of the form  $P_{\theta_n,\kappa_n,f}^{(n)}$ , Theorem 5.1 in [15] implies that the Rayleigh test will be blind to *all* Watson-type alternatives, irrespective of their severity.

While this Monte Carlo study nicely confirms the asymptotic results in Theorem 4.1, an important question remains open: is the Bingham test rate-optimal for the present problem, that is, is it so that any feasible test will be blind to alternatives that are less severe than those seen by the Bingham test in Theorem 4.1(ii)? The following result shows that this is indeed the case.

**Theorem 4.2.** Let  $(\kappa_n)$  be a real sequence such that  $\kappa_n = o(p_n^{3/2}/\sqrt{n})$  and  $\kappa_n = o(p_n)$ , and assume that  $p_n/\sqrt[3]{n} \to \infty$ (in particular, this covers the setup where  $\kappa_n = o(p_n^{3/2}/\sqrt{n})$ , with  $p_n/\sqrt[3]{n} \to \infty$  and  $p_n = O(n)$ ). Let  $f_W(s) = \exp(s)$  be the function f providing Watson distributions. Then, any test  $\phi_n$  satisfies

$$\inf_{\boldsymbol{\theta}_n \in \mathcal{S}^{p_n-1}} \mathbb{E}_{\mathsf{P}_{\boldsymbol{\theta}_n, \kappa_n, f_W}^{(n)}} [\boldsymbol{\phi}_n] \le \mathbb{E}_{\mathsf{P}_0^{(n)}} [\boldsymbol{\phi}_n] + o(1)$$

as  $n \to \infty$ .

This result shows that, as soon as  $p_n/\sqrt[3]{n} \to \infty$ , any test of uniformity  $\phi_n$  whose asymptotic null size is smaller than or equal to  $\alpha$  will exhibit, under the Watson version of the alternatives considered in Theorem 4.1(i), a worst power that is asymptotically upper-bounded by  $\alpha$  (as we explain below the proof of Lemma C.1, performing higher-order expansions of confluent hypergeometric functions of matrix arguments than the second-order one we perform in the proof of Theorem 4.2 would probably allow us to weaken the condition  $p_n/\sqrt[3]{n} \to \infty$  into  $p_n/n^{\delta} \to \infty$  for an arbitrary  $\delta > 0$ ).



Fig. 2. For any  $n \in \{100, 400, 800\}$ , rejection frequencies of the Bingham test (solid green curve) and of the Rayleigh test (solid red curve), obtained from a collection of M = 2500 independent random samples of size *n* from the Watson distribution with dimension  $p_n = \lfloor \sqrt{n} \rfloor$ , location  $\theta_n = (1, 0, \dots, 0)' \in \mathbb{R}^{p_n}$ , and concentration  $\kappa_n = \tau p_n / \sqrt{n}$  (top),  $\kappa_n = \tau p_n^{3/2} / \sqrt{n}$  (middle), or  $\kappa_n = \tau p_n^{7/4} / \sqrt{n}$  (bottom). The corresponding asymptotic powers are also plotted (dashed curves).

Now, since the Bingham test is invariant under the group of rotations  $g(\mathbf{x}_1, \dots, \mathbf{x}_n) = (\mathbf{O}\mathbf{x}_1, \dots, \mathbf{O}\mathbf{x}_n)$ , where **O** is an arbitrary  $p_n \times p_n$  orthogonal matrix, its power under any alternative  $P_{\boldsymbol{\theta}_n,\kappa_n,f}^{(n)}$  does not depend on  $\boldsymbol{\theta}_n$ , so that Theorem 4.1(ii) ensures that, if  $\kappa_n = \tau_n p_n^{3/2} / \sqrt{n}$  and  $(\tau_n) \to \tau \neq 0$ , then

$$\lim_{n\to\infty}\inf_{\boldsymbol{\theta}_n\in\mathcal{S}^{p_n-1}}\mathsf{P}_{\boldsymbol{\theta}_n,\kappa_n,f}^{(n)}\big[\mathcal{Q}_n^{\mathrm{St}}>z_\alpha\big]=1-\Phi\big(z_\alpha-2\tau^2\big)>\alpha.$$

As a corollary, the Bingham test is minimax rate-optimal in the parametric Watson model. As a final remark, note that the lower bound on the minimax separation rate provided in Theorem 4.2 in the parametric Watson model trivially also holds in the semiparametric model where f would remain unspecified in any class of functions containing the Watson one  $f_W$ .

# 5. Non-null results under non-axial alternatives

As mentioned in Section 3, the general result in Theorem 3.2 does not restrict to axial distributions. It can thus also be used to investigate the non-null behaviour of the Bingham test under, e.g., Fisher–von Mises–Langevin (FvML) alternatives. In this section, we consider a semiparametric model that includes FvML distributions, in the same way the semiparametric model considered in Section 2 contains the Watson ones. More precisely, we consider distributions admitting a density of the form

$$\mathbf{x} \mapsto \frac{\tilde{c}_{p,\kappa,f} \Gamma(\frac{p-1}{2})}{2\pi^{(p-1)/2}} f(\kappa \mathbf{x}' \boldsymbol{\theta}), \quad \text{with } \tilde{c}_{p,\kappa,f} = 1 / \int_{-1}^{1} (1 - s^2)^{(p-3)/2} f(\kappa s) \, ds, \tag{12}$$

with  $\theta \in S^{p-1}$ ,  $\kappa \in \mathbb{R}^+$ , and  $f \in \mathcal{F}$ . The particular case of FvML distributions is obtained with  $f(s) = \exp(s)$ . Unlike the densities in (3), these new densities do not attribute equal probabilities to antipodal regions on the sphere but rather are monotone, in the sense that the probability mass decreases monotonically when going from the modal location  $\theta$  to its antipodal location  $-\theta$ . The parameter  $\kappa$  is still a concentration parameter: typically, the larger  $\kappa$ , the more concentrated (about  $\theta$ ) the probability mass. Irrespective of f, the value  $\kappa = 0$  provides the uniform distribution over  $S^{p-1}$ , so that testing uniformity versus non-uniformity over the sphere consists in testing  $\mathcal{H}_0 : \kappa = 0$  versus  $\mathcal{H}_1 : \kappa > 0$  (negative values of  $\kappa$  would only swap the roles of  $\theta$  and  $-\theta$ , hence should be discarded in this model). The semiparametric model above was used in [15] for the problem of testing uniformity, as well as in [36,38] for low- and high-concentration location problems, respectively.

As in Section 2, the densities in (12) provide natural alternatives to uniformity, which leads us to consider the following sequences of hypotheses. For any sequence of positive integers  $(p_n)$ , sequence  $(\theta_n)$  such that  $\theta_n \in S^{p_n-1}$  for any n, nonnegative real sequence  $(\kappa_n)$ , and  $f \in \mathcal{F}$ , we denote as  $\tilde{P}_{\theta_n,\kappa_n,f}^{(n)}$  the hypothesis under which  $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$  form a random sample from the density (12) with the corresponding parameters  $\theta_n, \kappa_n$ , and f. We still write  $P_0^{(n)} := \tilde{P}_{\theta_n,0,f}^{(n)}$  for the uniform particular case. The following LAN result was obtained in [15].

**Theorem 5.1.** Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$  and  $(\theta_n)$  be a sequence such that  $\theta_n \in S^{p_n-1}$  for any n. Let  $\kappa_n = \tau_n \sqrt{p_n/n}$ , where  $(\tau_n)$  is O(1), and fix  $f \in \mathcal{F}$ . Then, as  $n \to \infty$  under  $\mathbb{P}_0^{(n)}$ ,

$$\Lambda_n = \log \frac{d\tilde{\mathsf{P}}_{\theta_n,\kappa_n,f}^{(n)}}{d\mathsf{P}_0^{(n)}} = \tau_n \Delta_{\theta_n} - \frac{1}{2}\tau_n^2 + o_{\mathsf{P}}(1),$$

where  $\Delta_{\boldsymbol{\theta}_n} := \sqrt{np_n} \bar{\mathbf{X}}'_n \boldsymbol{\theta}_n$  is asymptotically standard normal (here,  $\bar{\mathbf{X}}_n$  denotes the sample average of  $\mathbf{X}_{ni}$ , i = 1, ..., n).

It then easily follows from Le Cam's third lemma that the Bingham test statistic  $Q_n^{\text{St}}$  remains asymptotically standard normal under  $\tilde{P}_{\theta_n,\kappa_n,f}^{(n)}$ , with  $\kappa_n = \tau_n \sqrt{p_n/n}$  and  $(\tau_n) \to \tau$ . Consequently, the Bingham test is blind to such contiguous alternatives. Again, it is natural to study the non-null behaviour of the Bingham test to see whether or not this test can detect more severe non-axial alternatives. In order to apply our general result in Theorem 3.2, we need to study the asymptotic behaviour of the quantities  $e_{n\ell}$ ,  $\ell = 2, 4, 8$ , under the above non-axial hypotheses (see Appendix D for a proof).

**Proposition 5.1.** Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$ ,  $(\boldsymbol{\theta}_n)$  be a sequence such that  $\boldsymbol{\theta}_n \in S^{p_n-1}$  for any n, and  $(\kappa_n)$  be a nonnegative real sequence that is  $o(\sqrt{p_n})$  as  $n \to \infty$ . Fix  $f \in \mathcal{F}$ . Then, under  $\tilde{P}_{\boldsymbol{\theta}_n,\kappa_n,f}^{(n)}$ ,

$$e_{n2} = \frac{1}{p_n} + \frac{\kappa_n^2}{p_n^2} f''(0) + o\left(\frac{\kappa_n^2}{p_n^2}\right) \quad and \quad e_{n4} = \frac{3}{p_n^2} + o\left(\frac{1}{p_n^2}\right)$$

as  $n \to \infty$ , so that  $e_{n8} = O(1/p_n^2)$  as  $n \to \infty$ .

The following theorem then results from Theorem 3.2 (see Appendix D for a proof).

**Theorem 5.2.** Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$  and  $(\boldsymbol{\theta}_n)$  be a sequence such that  $\boldsymbol{\theta}_n \in S^{p_n-1}$  for any n. Fix  $f \in \mathcal{F}$ . Then, we have the following: (i) if  $\kappa_n = o(p_n^{3/4}/n^{1/4})$  and  $p_n = O(n)$  (or more generally, if  $\kappa_n = o(p_n^{3/4}/n^{1/4})$  and  $\kappa_n = o(\sqrt{p_n})$ ), then

$$Q_n^{\text{St}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \tag{13}$$

as  $n \to \infty$  under  $\tilde{P}^{(n)}_{\theta_n,\kappa_n,f}$ ; (ii) if  $\kappa_n = \tau_n p_n^{3/4} / n^{1/4}$  with  $(\tau_n) \to \tau (>0)$  and  $p_n = o(n)$ , then

$$Q_n^{\text{St}} \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\tau^4}{2} \left(f''(0)\right)^2, 1\right)$$
(14)

as  $n \to \infty$  under  $\tilde{P}^{(n)}_{\theta_n,\kappa_n,f}$ ; (iii) if  $n^{1/4}\kappa_n/p_n^{3/4} \to \infty$ ,  $\kappa_n = o(\sqrt{p_n})$  and  $f''(0) \neq 0$ , then, for any real number M,

$$\tilde{\mathsf{P}}_{\boldsymbol{\theta}_n,\kappa_n,f}^{(n)} \big[ Q_n^{\mathrm{St}} > M \big] \to 1$$

as  $n \to \infty$ .

Interpreting this result in the same way as Theorem 4.1, we learn that the detection threshold of the Bingham test under FvML-type alternatives is  $\kappa_n \sim p_n^{3/4}/n^{1/4}$ , with resulting asymptotic powers

$$\lim_{n \to \infty} \tilde{\mathsf{P}}_{\boldsymbol{\theta}_n, \kappa_n, f}^{(n)} \left[ Q_n^{\text{St}} > z_\alpha \right] = 1 - \Phi \left( z_\alpha - \frac{\tau^4}{2} \left( f''(0) \right)^2 \right)$$

under  $\tilde{P}_{\theta_n,\kappa_n,f}^{(n)}$ , with  $\kappa_n = \tau_n p_n^{3/4} / n^{1/4}$  and  $(\tau_n) \to \tau$ . This should be compared to the detection threshold of the Rayleigh test, that problem, which is in line with the fact that the Bingham test is primarily designed for axial data whereas the Rayleigh one aims at non-axial data. As mentioned in Section 4, however, the Rayleigh test will be blind to arbitrarily severe Watsontype alternatives, whereas the Bingham test will show power under both FvML-type and Watson-type alternatives.

In order to illustrate these results, we performed the following FvML version of the Watson simulation exercise conducted in Section 4. For any  $n \in \{100, 400, 800\}$ , we generated M = 2500 independent random samples of size n from the FvML distribution with dimension  $p_n = \lfloor \sqrt{n} \rfloor$ , location  $\theta_n = (1, 0, \dots, 0)' \in \mathbb{R}^{p_n}$ , and concentration

- (i)  $\kappa_n = \tau p_n^{3/4} / \sqrt{n}$ , (ii)  $\kappa_n = \tau p_n^{3/4} / n^{1/4}$ , (iii)  $\kappa_n = \tau p_n^{5/4} / n^{1/4}$ ,

still with  $\tau = 0, 0.4, \dots, 2$  in each case. Figure 3 reports the rejection frequencies of the Bingham test and of the Rayleigh test, as well as the corresponding asymptotic powers (obtained from Theorem 5.2 for the Bingham test and from Theorem 5.1 in [15] for the Rayleigh test). The results fully support the comments from the previous paragraph, for both the Bingham and Rayleigh tests. Strictly speaking, the consistency result in Theorem 5.2(iii) does not apply in the concentration scheme (iii) above, as the condition  $\kappa_n = o(\sqrt{p_n})$  is not met there; in Appendix E, however, we show that this condition is superfluous in the FvML case, so that our theoretical results imply consistency in the concentration scheme (iii), too.

# 6. Wrap up

In this paper, we identified the sequences of local alternatives that, within the considered axial model, are contiguous to the null hypothesis of uniformity on high-dimensional spheres. The Bingham test is blind to these alternatives. By relying on martingale central limit theorems, we derived the first results characterizing the non-null behaviour of this test in high dimensions. We identified the rotationally symmetric alternatives under which the Bingham test will show non-trivial asymptotic powers. We proved that this test will be blind to less severe alternatives and consistent under more severe ones. Our results impose only very mild assumptions on the considered rotationally symmetric alternatives. In particular, they apply to both the axial and non-axial cases, which allowed us to determine the detection threshold of the Bingham test in each case. Our results reveal that although it exhibits slower consistency rates than the Rayleigh test in the nonaxial case, the Bingham test can detect both types of alternatives, whereas the Rayleigh test will be blind to arbitrarily severe axial alternatives. These results are summarized in Table 1. In the axial case, the Bingham test is also minimax rate-optimal in the class of Watson distributions.



Fig. 3. For any  $n \in \{100, 400, 800\}$ , rejection frequencies of the Bingham test (solid green curve) and of the Rayleigh test (solid red curve), obtained from a collection of M = 2500 independent random samples of size *n* from the FvML distribution with dimension  $p_n = \lfloor \sqrt{n} \rfloor$ , location  $\boldsymbol{\theta}_n = (1, 0, \dots, 0)' \in \mathbb{R}^{p_n}$ , and concentration  $\kappa_n = \tau p_n^{3/4} / \sqrt{n}$  (top),  $\kappa_n = \tau p_n^{3/4} / n^{1/4}$  (middle), or  $\kappa_n = \tau p_n^{5/4} / n^{1/4}$  (bottom). The corresponding asymptotic powers are also plotted (dashed curves).

#### Table 1

For the axial (resp., non-axial) case, this summary table provides the contiguity rate associated with the semiparametric class of distributions considered in Sections 2 and 4 (resp., in Section 5), as well as the corresponding detection thresholds of the Bingham and Rayleigh tests. Here,  $\emptyset$  refers to the fact that Rayleigh test is blind to arbitrarily severe axial alternatives

|                                | Axial case                                | Non-axial case                             |
|--------------------------------|---|--|
| Contiguity rate                | $\kappa_n \sim rac{p_n}{\sqrt{n}}$       | $\kappa_n \sim rac{\sqrt{p_n}}{\sqrt{n}}$ |
| Bingham's detection threshold  | $\kappa_n \sim rac{p_n^{3/2}}{\sqrt{n}}$ | $\kappa_n \sim \frac{p_n^{3/4}}{n^{1/4}}$  |
| Rayleigh's detection threshold | Ø   | $\kappa_n \sim \frac{p_n^{3/4}}{\sqrt{n}}$ |

#### **Appendix A: Proofs for Section 2**

We will need two preliminary results. The following result is a high-dimensional version of Lemma A.1 in [17].

**Lemma A.1.** Let  $g : \mathbb{R} \to \mathbb{R}$  be twice differentiable at 0. Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$  and  $(\kappa_n)$  be a real sequence that is  $o(p_n)$ . Then,

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

as  $n \to \infty$ , where we let  $c_p := 1/\int_{-1}^{1} (1-s^2)^{(p-3)/2} ds$ .

**Proof of Lemma A.1.** If **X** is uniformly distributed over  $S^{p_n-1}$ , then, for any  $\boldsymbol{\theta} \in S^{p_n-1}$ , the distribution of  $\boldsymbol{\theta}'\mathbf{X}$  is symmetric about zero and  $(\boldsymbol{\theta}'\mathbf{X})^2 \sim \text{Beta}(1/2, (p_n - 1)/2)$ ; see, e.g., [33], Theorem 1.5.7(ii). Therefore,  $\boldsymbol{\theta}'\mathbf{X}$  has density  $s \mapsto c_{p_n}(1-s^2)^{(p_n-3)/2}\mathbb{I}[|s| \le 1]$ , and we have

$$c_{p_n} \int_{-1}^{1} s^2 (1-s^2)^{(p_n-3)/2} ds = \mathbb{E}[(\boldsymbol{\theta}' \mathbf{X})^2] = \frac{1}{p_n},$$
(15)

and

$$c_{p_n} \int_{-1}^{1} s^4 (1-s^2)^{(p_n-3)/2} ds = \mathbb{E}[(\boldsymbol{\theta}' \mathbf{X})^4] = \frac{3}{p_n(p_n+2)}.$$
(16)

From (15), we can write

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

Note that we can without any loss of generality assume that  $(\kappa_n)$  is a sequence in  $\mathbb{R}_0$ , which allows us to perform the change of variables  $t = |\kappa_n|^{1/2} s$ . Doing so and using (15)–(16) then provides (throughout,  $s_{\kappa_n}$  denotes the sign of  $\kappa_n$ )

$$R_n(g) - g(0) - \frac{\kappa_n}{p_n}g'(0) = \frac{3\kappa_n^2}{p_n(p_n+2)} \int_{-\infty}^{\infty} h_n(t) \left(\frac{g(s_{\kappa_n}t^2) - g(0) - s_{\kappa_n}t^2g'(0)}{t^4}\right) dt,$$

or, equivalently,

$$\frac{R_n(g) - g(0) - \frac{\kappa_n}{p_n}g'(0) - \frac{3\kappa_n^2}{2p_n(p_n+2)}g''(0)}{\frac{3\kappa_n^2}{p_n(p_n+2)}} = \int_{-\infty}^{\infty} h_n(t) \left(\frac{g(s_{\kappa_n}t^2) - g(0) - s_{\kappa_n}t^2g'(0)}{t^4}\right) dt - \frac{1}{2}g''(0),$$
(17)

where  $h_n$  is defined through

$$t \mapsto h_n(t) = \frac{t^4 (1 - \frac{t^2}{|\kappa_n|})^{(p_n - 3)/2} \mathbb{I}[|t| \le |\kappa_n|^{1/2}]}{\int_{-\infty}^{\infty} t^4 (1 - \frac{t^2}{|\kappa_n|})^{(p_n - 3)/2} \mathbb{I}[|t| \le |\kappa_n|^{1/2}] dt}$$

It can be checked that, since  $\kappa_n = o(p_n)$ , the sequence  $(h_n)$  is an *approximate*  $\delta$ -sequence, in the sense that  $\int_{-\infty}^{\infty} h_n(t) dt = 1$  for any n and  $\int_{-\varepsilon}^{\varepsilon} h_n(t) dt \to 1$  for any  $\varepsilon > 0$ . Hence,

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} h_n(t) \left( \frac{g(s_{\kappa_n} t^2) - g(0) - s_{\kappa_n} t^2 g'(0)}{t^4} \right) dt = \lim_{t \to 0} \frac{g(s_{\kappa_n} t^2) - g(0) - s_{\kappa_n} t^2 g'(0)}{t^4},$$

which, by using L'Hôpital's rule, is equal to

$$\lim_{t \to 0} \frac{2s_{\kappa_n} t g'(s_{\kappa_n} t^2) - 2s_{\kappa_n} t g'(0)}{4t^3} = \frac{1}{2} \lim_{t \to 0} \frac{g'(s_{\kappa_n} t^2) - g'(0)}{s_{\kappa_n} t^2} = \frac{1}{2} g''(0)$$

Thus, (17) yields

$$R_n(g) - g(0) - \frac{\kappa_n}{p_n} g'(0) - \frac{3\kappa_n^2}{2p_n(p_n+2)} g''(0) = o\left(\frac{\kappa_n^2}{p_n^2}\right),$$
  
establishes the result

which establishes the result.

**Lemma A.2.** Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$  and  $\kappa_n = \tau_n p_n / \sqrt{n}$ , where  $(\tau_n)$  is O(1) but not o(1). Writing  $E_{nk} := \mathbb{E}[\ell_{f,k}(\kappa_n(\mathbf{X}'_{ni}\boldsymbol{\theta}_n)^2)]$ , with  $\ell_{f,k} := (\log f)^k$ , and  $V_n := \operatorname{Var}[\log f(\kappa_n(\mathbf{X}'_{ni}\boldsymbol{\theta}_n)^2)]$ , define

$$W_{ni} := \frac{1}{\sqrt{nV_n}} \left( \log f\left(\kappa_n \left(\mathbf{X}'_{ni} \boldsymbol{\theta}_n\right)^2\right) - E_{n1} \right), \quad i = 1, \dots, n$$

Then, writing  $s_a$  for the sign of the real number a, we have that

$$d_n := \mathbf{E}\left[\left(\Delta_{\theta_n} - \sqrt{2}s_{\tau_n} \sum_{i=1}^n W_{ni}\right)^2\right] = o(1)$$
(18)

as  $n \to \infty$  under  $\mathbf{P}_0^{(n)}$ .

**Proof of Lemma A.2.** First note that, since  $\mathbf{X}'_{ni}\boldsymbol{\theta}_n$  has density  $s \mapsto c_{p_n}(1-s^2)^{(p_n-3)/2}\mathbb{I}[|s| \le 1]$  under  $\mathbf{P}_0^{(n)}$ , we have

$$E_{n1} = c_{p_n} \int_{-1}^{1} (1 - s^2)^{(p_n - 3)/2} \ell_{f,1}(\kappa_n s^2) ds$$
  
$$= \frac{\kappa_n}{p_n} \ell'_{f,1}(0) + \frac{3\kappa_n^2}{2p_n^2} \ell''_{f,1}(0) + o\left(\frac{\kappa_n^2}{p_n^2}\right) = \frac{\kappa_n}{p_n} + \frac{3\kappa_n^2}{2p_n^2} \left(f''(0) - 1\right) + o\left(\frac{\kappa_n^2}{p_n^2}\right)$$
(19)

and

$$E_{n2} = c_{p_n} \int_{-1}^{1} (1 - s^2)^{(p_n - 3)/2} \ell_{f,2}(\kappa_n s^2) ds$$
  
=  $\frac{\kappa_n}{p_n} \ell'_{f,2}(0) + \frac{3\kappa_n^2}{2p_n^2} \ell''_{f,2}(0) + o\left(\frac{\kappa_n^2}{p_n^2}\right) = \frac{3\kappa_n^2}{p_n^2} + o\left(\frac{\kappa_n^2}{p_n^2}\right),$ 

so that

$$nV_n = n\left(E_{n2} - E_{n1}^2\right) = \frac{2n\kappa_n^2}{p_n^2} + o\left(\frac{n\kappa_n^2}{p_n^2}\right).$$
(20)

Now, write

$$M_{n} := \sqrt{n V_{n}} \left( \Delta_{\boldsymbol{\theta}_{n}} - \sqrt{2} s_{\tau_{n}} \sum_{i=1}^{n} W_{ni} \right)$$
$$= \sum_{i=1}^{n} \left\{ p_{n} \sqrt{V_{n}} \left( \left( \mathbf{X}_{ni}^{\prime} \boldsymbol{\theta}_{n} \right)^{2} - \frac{1}{p_{n}} \right) - \sqrt{2} s_{\tau_{n}} \left( \log f \left( \kappa_{n} \left( \mathbf{X}_{ni}^{\prime} \boldsymbol{\theta}_{n} \right)^{2} \right) - E_{n1} \right) \right\}.$$
(21)

Then using the fact that  $E[(\mathbf{X}'_{n1}\boldsymbol{\theta}_n)^2] = 1/p_n$  and  $E[(\mathbf{X}'_{n1}\boldsymbol{\theta}_n)^4] = 3/(p_n(p_n + 2))$  (see the proof of Lemma A.1) provides  $Var[(\mathbf{X}'_{n1}\boldsymbol{\theta}_n)^2] = 2/p_n^2 + o(1/p_n^2)$ , we obtain

$$E[M_n^2] = nE\left[\left\{p_n\sqrt{V_n}\left(\left(\mathbf{X}_{ni}'\boldsymbol{\theta}_n\right)^2 - \frac{1}{p_n}\right) - \sqrt{2}s_{\tau_n}\left(\log f\left(\kappa_n\left(\mathbf{X}_{ni}'\boldsymbol{\theta}_n\right)^2\right) - E_{n1}\right)\right\}^2\right]$$
  
$$= (4 + o(1))nV_n - 2\sqrt{2V_n}np_ns_{\tau_n}E\left[\left(\left(\mathbf{X}_{n1}'\boldsymbol{\theta}_n\right)^2 - \frac{1}{p_n}\right)\left(\log f\left(\kappa_n\left(\mathbf{X}_{n1}'\boldsymbol{\theta}_n\right)^2\right) - E_{n1}\right)\right]$$
  
$$= 4nV_n - 2\sqrt{2nV_n}s_{\tau_n}\left(\frac{\sqrt{n}p_n}{\kappa_n}E[g\left(\kappa_n\left(\mathbf{X}_{n1}'\boldsymbol{\theta}_n\right)^2\right)] - \sqrt{n}E_{n1}\right) + o(nV_n), \qquad (22)$$

where we let  $g(x) := x \log f(x)$ . Using Lemma A.1 again, we obtain

$$\mathbb{E}[g(\kappa_n(\mathbf{X}'_{n1}\boldsymbol{\theta}_n)^2)] = c_{p_n} \int_{-1}^{1} (1-s^2)^{(p_n-3)/2} g(\kappa_n s^2) \, ds = \frac{3\kappa_n^2}{p_n^2} + o\left(\frac{\kappa_n^2}{p_n^2}\right).$$

Using this jointly with (19) and (20), it follows from (21)-(22) that

$$d_n = 4 - \frac{2\sqrt{2}s_{\tau_n}(\frac{3\sqrt{n\kappa_n}}{p_n} + o(\frac{\sqrt{n\kappa_n}}{p_n}) - \{\frac{\sqrt{n\kappa_n}}{p_n} + \frac{3\sqrt{n\kappa_n^2}}{2p_n^2}(f''(0) - 1) + o(\frac{\sqrt{n\kappa_n^2}}{p_n^2})\})}{\sqrt{\frac{2n\kappa_n^2}{p_n^2} + o(\frac{n\kappa_n^2}{p_n^2})}} + o(1)$$
$$= 4 - \frac{2\sqrt{2}s_{\tau_n}(2\tau_n + o(1))}{\sqrt{2\tau_n^2} + o(1)} = o(1),$$

which establishes the result.

**Proof of Theorem 2.1.** In this proof, all expectations and variances are taken under the null hypothesis of uniformity  $P_0^{(n)}$  and all stochastic convergences and  $o_P$ 's are as  $n \to \infty$  under  $P_0^{(n)}$ . Using the same notation as in Lemma A.2, consider then the local log-likelihood ratio

$$\begin{split} \Lambda_n &= \log \frac{d\mathbf{P}_{\boldsymbol{\theta}_n,\kappa_n,f}^{(n)}}{d\mathbf{P}_0^{(n)}} = \sum_{i=1}^n \log \frac{c_{p_n,\kappa_n,f} f \left(\kappa_n (\mathbf{X}_{ni}^{\prime} \boldsymbol{\theta}_n)^2\right)}{c_{p_n}} \\ &= n \left(\log \frac{c_{p_n,\kappa_n,f}}{c_{p_n}} + E_{n1}\right) + \sum_{i=1}^n \{\log f \left(\kappa_n \left(\mathbf{X}_{ni}^{\prime} \boldsymbol{\theta}_n\right)^2\right) - E_{n1}\} \\ &= n \left(\log \frac{c_{p_n,\kappa_n,f}}{c_{p_n}} + E_{n1}\right) + \sqrt{nV_n} \sum_{i=1}^n W_{ni}. \end{split}$$

Lemma A.1 readily yields

$$\log \frac{c_{p_n,\kappa_n,f}}{c_{p_n}} = -\log \left( c_{p_n} \int_{-1}^{1} (1-s^2)^{(p_n-3)/2} f(\kappa_n s^2) \, ds \right)$$
$$= -\log \left( 1 + \frac{\kappa_n}{p_n} + \frac{3\kappa_n^2}{2p_n^2} f''(0) + o\left(\frac{\kappa_n^2}{p_n^2}\right) \right)$$
$$= -\frac{\kappa_n}{p_n} - \frac{3\kappa_n^2}{2p_n^2} f''(0) + \frac{\kappa_n^2}{2p_n^2} + o\left(\frac{\kappa_n^2}{p_n^2}\right),$$
(23)

so that (19) provides

$$n\left(\log\frac{c_{p_n,\kappa_n,f}}{c_{p_n}}+E_{n1}\right)=-\frac{n\kappa_n^2}{p_n^2}+o\left(\frac{n\kappa_n^2}{p_n^2}\right).$$

Using (20), we thus obtain

$$\Lambda_n = -\frac{n\kappa_n^2}{p_n^2} + \sqrt{\frac{2n\kappa_n^2}{p_n^2} + o\left(\frac{n\kappa_n^2}{p_n^2}\right)} \sum_{i=1}^n W_{ni} + o\left(\frac{n\kappa_n^2}{p_n^2}\right).$$
(24)

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

$$\mathbf{E}[\Lambda_n^2] = \left(\mathbf{E}[\Lambda_n]\right)^2 + \mathbf{Var}[\Lambda_n] = \frac{n^2 \kappa_n^4}{p_n^4} + o\left(\frac{n^2 \kappa_n^4}{p_n^4}\right) + \frac{2n \kappa_n^2}{p_n^2} + o\left(\frac{n \kappa_n^2}{p_n^2}\right).$$

Recall that the theorem assumes that  $\kappa_n = \tau_n p_n / \sqrt{n}$ , with  $\tau_n = O(1)$ . If  $\tau_n = o(1)$ , then we obtain that  $E[\Lambda_n^2] = o(1)$ , so that  $\Lambda_n = o_P(1)$ , which proves the result in this case. We may thus assume that  $(\tau_n)$  is O(1) but not o(1). Then, (24)

rewrites

$$\Lambda_n = -\tau_n^2 + \sqrt{2\tau_n^2 + o(1)} \sum_{i=1}^n W_{ni} + o(1) = -\tau_n^2 + \sqrt{2} |\tau_n| \sum_{i=1}^n W_{ni} + o_{\rm P}(1),$$
(25)

where we used the fact that, since  $(\tau_n)$  is not o(1),

$$\mathbf{E}\left[\left\{\left(\sqrt{2\tau_n^2 + o(1)} - \sqrt{2\tau_n^2}\right)\sum_{i=1}^n W_{ni}\right\}^2\right] = \frac{o(1)}{(\sqrt{2\tau_n^2 + o(1)} + \sqrt{2\tau_n^2})^2} = o(1).$$

Now, applying the Cauchy–Schwarz inequality and the Chebychev inequality, then using Lemma A.1 and (20), provides that, for some positive constant C and any  $\varepsilon > 0$ ,

$$\begin{split} \left(\sum_{i=1}^{n} \mathbb{E}\left[W_{ni}^{2} \mathbb{I}\left[|W_{ni}| > \varepsilon\right]\right]\right)^{2} &\leq n^{2} \mathbb{E}\left[W_{ni}^{4}\right] \mathbb{P}\left[|W_{ni}| > \varepsilon\right] \leq \frac{n^{2}}{\varepsilon^{2}} \mathbb{E}\left[W_{ni}^{4}\right] \operatorname{Var}\left[W_{ni}\right] = \frac{n}{\varepsilon^{2}} \mathbb{E}\left[W_{ni}^{4}\right] \\ &\leq \frac{Cn E_{n4}}{\varepsilon^{2} n^{2} V_{n}^{2}} = \frac{C(\frac{n\kappa_{n}}{p_{n}} \ell'_{f,4}(0) + \frac{3n\kappa_{n}^{2}}{2p_{n}^{2}} \ell''_{f,4}(0) + o(\frac{n\kappa_{n}^{2}}{p_{n}^{2}}))}{\varepsilon^{2}(\frac{2n\kappa_{n}^{2}}{p_{n}^{2}} + o(\frac{n\kappa_{n}^{2}}{p_{n}^{2}}))^{2}} = \frac{o(1)}{\varepsilon^{2}(2\tau_{n}^{2} + o(1))^{2}} = o(1). \end{split}$$

Therefore,  $\sum_{i=1}^{n} W_{ni}$  satisfies the classical Levy–Lindeberg condition, hence is asymptotically standard normal (recall that  $W_{ni}$ , i = 1, ..., n, are mutually independent with mean zero and variance 1/n). The result thus follows from Lemma A.2.

The proof of Proposition 2.1 requires the following lemma.

**Lemma A.3.** If  $\mathbf{X}_1, \ldots, \mathbf{X}_n$  is a random sample from the uniform distribution over  $S^{p-1}$ , then  $\mathbb{E}[\operatorname{tr}[\mathbf{S}_n^2](\mathbf{S}_n - \frac{1}{p}\mathbf{I}_p)] = \mathbf{0}$ , where we let  $\mathbf{S}_n := \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i \mathbf{X}'_i$  and where  $\mathbf{I}_p$  denotes the  $p \times p$  identity matrix.

Proof of Lemma A.3. First note that

$$\operatorname{tr}[\mathbf{S}_{n}^{2}]\left(\mathbf{S}_{n}-\frac{1}{p}\mathbf{I}_{p}\right) = \frac{1}{n^{3}}\sum_{i,j,k=1}^{n}\operatorname{tr}[\mathbf{X}_{i}\mathbf{X}_{j}'\mathbf{X}_{j}\mathbf{X}_{j}']\left(\mathbf{X}_{k}\mathbf{X}_{k}'-\frac{1}{p}\mathbf{I}_{p}\right)$$
$$= \frac{1}{n^{3}}\sum_{i,j,k=1}^{n}\left\{\left(\mathbf{X}_{i}'\mathbf{X}_{j}\right)^{2}\mathbf{X}_{k}\mathbf{X}_{k}'-\frac{1}{p}\left(\mathbf{X}_{i}'\mathbf{X}_{j}\right)^{2}\mathbf{I}_{p}\right\}.$$
(26)

For  $i \neq j$ , we have, irrespective of whether  $k \in \{i, j\}$  or not,

$$\mathbf{E}[(\mathbf{X}'_{i}\mathbf{X}_{j})^{2}\mathbf{X}_{k}\mathbf{X}'_{k}] = \mathbf{E}[\mathbf{E}[(\mathbf{X}'_{i}\mathbf{X}_{j})^{2}\mathbf{X}_{k}\mathbf{X}'_{k}|\mathbf{X}_{k}]] = \frac{1}{p}\mathbf{E}[\mathbf{X}_{k}\mathbf{X}'_{k}] = \frac{1}{p^{2}}\mathbf{I}_{p} = \mathbf{E}\left[\frac{1}{p}(\mathbf{X}'_{i}\mathbf{X}_{j})^{2}\mathbf{I}_{p}\right],$$

whereas, for i = j, we trivially have

$$\mathbf{E}[(\mathbf{X}_{i}'\mathbf{X}_{j})^{2}\mathbf{X}_{k}\mathbf{X}_{k}'] = \mathbf{E}[\mathbf{X}_{k}\mathbf{X}_{k}'] = \frac{1}{p}\mathbf{I}_{p} = \mathbf{E}\left[\frac{1}{p}(\mathbf{X}_{i}'\mathbf{X}_{j})^{2}\mathbf{I}_{p}\right].$$

The result thus follows from (26).

**Proof of Proposition 2.1.** Theorem 2.1 implies that, as  $n \to \infty$  under  $P_0^{(n)}$ ,

$$\operatorname{Cov}\left[Q_{n}^{\operatorname{St}},\Lambda_{n}\right] = \frac{\tau_{n}}{\sqrt{2d_{p_{n}}}} \operatorname{E}\left[Q_{n}\Delta_{\boldsymbol{\theta}_{n}}\right] + o(1) = \frac{np(p+2)\tau_{n}}{2\sqrt{2d_{p_{n}}}} \operatorname{E}\left[\operatorname{tr}\left[\mathbf{S}_{n}^{2}\right]\Delta_{\boldsymbol{\theta}_{n}}\right] + o(1)$$
$$= \frac{n^{3/2}p^{2}(p+2)\tau_{n}}{2\sqrt{2d_{p_{n}}}} (\boldsymbol{\theta}_{n}\otimes\boldsymbol{\theta}_{n})' \operatorname{E}\left[\operatorname{tr}\left[\mathbf{S}_{n}^{2}\right]\operatorname{vec}\left(\mathbf{S}_{n} - \frac{1}{p}\mathbf{I}_{p}\right)\right] + o(1), \tag{27}$$

so that Lemma A.3 yields that this covariance is o(1) as  $n \to \infty$ . Since  $Q_n^{\text{St}}$  is asymptotically standard normal under  $P_0^{(n)}$  (see (1)), Le Cam's third lemma then entails that  $Q_n^{\text{St}}$  remains asymptotically standard normal under  $P_{\theta_n,\kappa_n,f}^{(n)}$ .

## **Appendix B: Proofs for Section 3**

Throughout, we will make use of the *tangent-normal decomposition*  $\mathbf{X}_{ni} = u_{ni}\boldsymbol{\theta}_n + v_{ni}\mathbf{S}_{ni}$  of  $\mathbf{X}_{ni}$ , involving

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

and  $v_{ni} := (1 - u_{ni}^2)^{1/2}$  (recall that  $u_{ni} = \mathbf{X}'_{ni}\boldsymbol{\theta}_n$ ). Under  $\mathbf{P}_{\boldsymbol{\theta}_n,F_n}^{(n)}$ ,  $u_{ni}$  has cumulative distribution function  $F_n$ , the  $p_n$ -vector  $\mathbf{S}_{ni}$  is uniformly distributed over  $\{\mathbf{x} \in S^{p_n-1} : \boldsymbol{\theta}'_n \mathbf{x} = 0\}$ , and  $u_{ni}$  (equivalently,  $v_{ni}$ ) and  $\mathbf{S}_{ni}$  are mutually independent. Recall that  $e_{n\ell} = \mathbf{E}[u_{ni}^{\ell}]$  and  $f_{n\ell} = \mathbf{E}[v_{ni}^{\ell}]$  for any positive integer  $\ell$ .

To prove Proposition 3.1, we need both following preliminary results.

**Lemma B.1.** Under  $P_{\theta_n, F_n}^{(n)}$ , (i)  $E[\mathbf{S}_{ni}\mathbf{S}'_{ni}] = (1/(p_n - 1))(\mathbf{I}_{p_n} - \theta_n \theta'_n)$  for any i; (ii)  $E[(\mathbf{S}'_{ni}\mathbf{S}_{nj})^2] = 1/(p_n - 1)$  for any  $i \neq j$ ; (iii)  $E[(\mathbf{S}'_{ni}\mathbf{S}_{nj})^4] = 3/(p_n^2 - 1)$  for any  $i \neq j$ ; (iv)  $E[(\mathbf{S}'_{ni}\mathbf{S}_{nj})^8] = 105/((p_n^2 - 1)(p_n + 3)(p_n + 5))$  for any  $i \neq j$ .

**Proof of Lemma B.2.** Parts (i)–(iii) were proved in Lemma B.1 from [15], whereas Part (iv) strictly follows from the same argument as in that lemma by using the fact that  $E[(\mathbf{U}'\mathbf{V})^8] = \frac{105}{(p(p+2)(p+4)(p+6))}$  if the independent random *p*-vectors **U**, **V** are uniformly distributed over  $S^{p-1}$  (see Lemma A.1 in [35]).

**Lemma B.2.** Under  $P_{\theta_n,F_n}^{(n)}$ , (i)  $E[(\mathbf{X}'_{ni}\mathbf{X}_{nj})^2] = p_n g_{n2}^2/(p_n - 1) + (1/p_n)$  for any  $i \neq j$ , (ii)  $E[(\mathbf{X}'_{ni}\mathbf{X}_{nj})^4] = e_{n4}^2 + 6(e_{n2} - e_{n4})^2/(p_n - 1) + 3f_{n4}^2/(p_n^2 - 1)$  for any  $i \neq j$ , (iii)  $E[(\mathbf{X}'_{ni}\mathbf{X}_{nj})^2(\mathbf{X}'_{nk}\mathbf{X}_{n\ell})^2] = e_{n4}e_{n2}^2 + 2e_{n2}(e_{n2} - e_{n4})f_{n2}/(p_n - 1) + f_{n2}^2f_{n4}/(p_n - 1)^2$  for any  $i \neq j$  and  $k \neq \ell$  such that  $\{i, j, k, \ell\}$  contains exactly three different indices, and (iv)  $E[(\mathbf{X}'_{ni}\mathbf{X}_{nj})^2(\mathbf{X}'_{nk}\mathbf{X}_{n\ell})^2] = (p_n g_{n2}^2/(p_n - 1) + (1/p_n))^2$  if  $i, j, k, \ell$  are pairwise different.

**Proof of Lemma B.2.** (i) The tangent-normal decomposition provides  $\mathbf{X}'_{ni}\mathbf{X}_{nj} = u_{ni}u_{nj} + v_{ni}v_{nj}(\mathbf{S}'_{ni}\mathbf{S}_{nj})$ , which yields

$$\left(\mathbf{X}_{ni}'\mathbf{X}_{nj}\right)^{2} = u_{ni}^{2}u_{nj}^{2} + v_{ni}^{2}v_{nj}^{2}\left(\mathbf{S}_{ni}'\mathbf{S}_{nj}\right)^{2} + 2u_{ni}v_{ni}u_{nj}v_{nj}\left(\mathbf{S}_{ni}'\mathbf{S}_{nj}\right).$$
(28)

Lemma B.1(ii) then shows that  $E[(\mathbf{X}'_{ni}\mathbf{X}_{nj})^2] = (E[u_{n1}^2])^2 + (E[v_{n1}^2])^2/(p_n - 1) = e_{n2}^2 + f_{n2}^2/(p_n - 1)$  under  $P_{\theta_n,F_n}^{(n)}$ . Part (i) of the lemma thus follows by using successively the identities  $f_{n2} = 1 - e_{n2}$  and  $e_{n2} = g_{n2} + (1/p_n)$ . (ii) Using (28) and the fact that  $E[(\mathbf{S}'_{ni}\mathbf{S}_{nj})^\ell] = 0$  for any odd positive integer  $\ell$ , we obtain

$$E[(\mathbf{X}'_{ni}\mathbf{X}_{nj})^{4}] = E[u^{4}_{ni}u^{4}_{nj} + 6u^{2}_{ni}u^{2}_{nj}v^{2}_{ni}v^{2}_{nj}(\mathbf{S}'_{ni}\mathbf{S}_{nj})^{2} + v^{4}_{ni}v^{4}_{nj}(\mathbf{S}'_{ni}\mathbf{S}_{nj})^{4}]$$
  
$$= (E[u^{4}_{n1}])^{2} + \frac{6}{p_{n}-1}(E[u^{2}_{n1}(1-u^{2}_{n1})])^{2} + \frac{3}{p^{2}_{n}-1}(E[v^{4}_{n1}])^{2},$$

which establishes Part (ii) of the result. (iii) Without any loss of generality, assume that, in  $\{i, j, k, \ell\}$ , only j and k are equal to each other. Proceeding as in (ii), we then have

$$\begin{split} & \mathsf{E}[(\mathbf{X}'_{ni}\mathbf{X}_{nj})^{2}(\mathbf{X}'_{nk}\mathbf{X}_{n\ell})^{2}] \\ &= \mathsf{E}[u_{ni}^{2}u_{nj}^{2}u_{nk}^{2}u_{n\ell}^{2} + u_{ni}^{2}u_{nj}^{2}v_{nk}^{2}v_{n\ell}^{2}(\mathbf{S}'_{nk}\mathbf{S}_{n\ell})^{2} \\ &+ u_{nk}^{2}u_{n\ell}^{2}v_{ni}^{2}v_{nj}^{2}(\mathbf{S}'_{ni}\mathbf{S}_{nj})^{2} + v_{ni}^{2}v_{nj}^{2}v_{nk}^{2}v_{n\ell}^{2}(\mathbf{S}'_{ni}\mathbf{S}_{nj})^{2}(\mathbf{S}'_{nk}\mathbf{S}_{n\ell})^{2}] \\ &= (\mathsf{E}[u_{ni}^{2}])^{2}\mathsf{E}[u_{nk}^{4}] + \frac{2}{p_{n}-1}\mathsf{E}[u_{n1}^{2}]\mathsf{E}[v_{n1}^{2}]\mathsf{E}[u_{n1}^{2}(1-u_{n1}^{2})] + \frac{1}{(p_{n}-1)^{2}}(\mathsf{E}[v_{n1}^{2}])^{2}\mathsf{E}[v_{n1}^{4}], \end{split}$$

where we used the fact that  $E[(\mathbf{S}'_{ni}\mathbf{S}_{nj})^2(\mathbf{S}'_{nj}\mathbf{S}_{n\ell})^2] = E[E[(\mathbf{S}'_{ni}\mathbf{S}_{nj})^2(\mathbf{S}'_{nj}\mathbf{S}_{n\ell})^2|\mathbf{S}_{nj}]] = 1/(p_n - 1)^4$  since  $E[(\mathbf{S}'_{ni}\mathbf{S}_{nj})^2|\mathbf{S}_{nj}] = 1/(p_n - 1)$ . (iv) Since  $i, j, k, \ell$  are pairwise different,  $\mathbf{X}'_{ni}\mathbf{X}_{nj}$  and  $\mathbf{X}'_{nk}\mathbf{X}_{n\ell}$  are mutually independent, so that the result directly follows from Part (i) of the lemma.

Proof of Proposition 3.1. Using Lemma B.2(i), we readily obtain

$$\mathbf{E}[\mathcal{Q}_n^{\text{St}}] = \frac{p_n}{n} \sum_{\substack{i < j \\ i, j = 1}}^n \left( \mathbf{E}[(\mathbf{X}'_{ni}\mathbf{X}_{nj})^2] - \frac{1}{p_n} \right) = \frac{p_n}{n} \times \frac{n(n-1)}{2} \times \frac{p_n}{p_n-1} g_{n2}^2 = \frac{p_n^2(n-1)}{2(p_n-1)} g_{n2}^2$$

Turning to the variance, we have

$$\operatorname{Var}[\mathcal{Q}_{n}^{\operatorname{St}}] = \frac{p_{n}^{2}}{n^{2}} \sum_{1 \leq i < j \leq n} \sum_{1 \leq k < \ell \leq n} \left\{ \operatorname{E}[(\mathbf{X}_{ni}'\mathbf{X}_{nj})^{2}(\mathbf{X}_{nk}'\mathbf{X}_{n\ell})^{2}] - \left(\operatorname{E}[(\mathbf{X}_{ni}'\mathbf{X}_{nj})^{2}]\right)^{2} \right\}$$
$$= \frac{p_{n}^{2}}{n^{2}} \sum_{1 \leq i < j \leq n} \sum_{1 \leq k < \ell \leq n} \left\{ \operatorname{E}[(\mathbf{X}_{ni}'\mathbf{X}_{nj})^{2}(\mathbf{X}_{nk}'\mathbf{X}_{n\ell})^{2}] - \left(\frac{p_{n}}{p_{n}-1}g_{n2}^{2} + \frac{1}{p_{n}}\right)^{2} \right\}.$$

In this last sum,  $\binom{n}{2}$  terms correspond to Lemma B.2(ii) and  $6\binom{n}{4}$  terms (all equal to zero) correspond to Lemma B.2(iv). Thus,  $\binom{n}{2}^2 - \binom{n}{2} - 6\binom{n}{4} = n(n-1)(n-2)$  terms correspond to Lemma B.2(iii), which leads to

$$\begin{aligned} \operatorname{Var}[\mathcal{Q}_{n}^{\mathsf{St}}] &= \frac{p_{n}^{2}}{n^{2}} \bigg[ \frac{n(n-1)}{2} \bigg\{ e_{n4}^{2} + \frac{6(e_{n2} - e_{n4})^{2}}{p_{n} - 1} + \frac{3f_{n4}^{2}}{p_{n}^{2} - 1} - \bigg( \frac{p_{n}}{p_{n} - 1} g_{n2}^{2} + \frac{1}{p_{n}} \bigg)^{2} \bigg\} \\ &+ n(n-1)(n-2) \bigg\{ e_{n4} e_{n2}^{2} + \frac{2e_{n2}(e_{n2} - e_{n4})f_{n2}}{p_{n} - 1} + \frac{f_{n2}^{2}f_{n4}}{(p_{n} - 1)^{2}} - \bigg( \frac{p_{n}}{p_{n} - 1} g_{n2}^{2} + \frac{1}{p_{n}} \bigg)^{2} \bigg\} \bigg] \\ &= \frac{(n-1)p_{n}^{2}}{2n} \bigg\{ e_{n4}^{2} + \frac{6}{p_{n} - 1}(e_{n2} - e_{n4})^{2} + \frac{3f_{n4}^{2}}{p_{n}^{2} - 1} - \bigg( \frac{p_{n}}{p_{n} - 1} g_{n2}^{2} + \frac{1}{p_{n}} \bigg)^{2} \bigg\} \\ &+ \frac{(n-1)(n-2)p_{n}^{4}}{n(p_{n} - 1)^{2}} \big( e_{n4} - e_{n2}^{2} \big) g_{n2}^{2}, \end{aligned}$$

which establishes the result.

The proof of Theorem 3.1 still requires both following lemmas.

**Lemma B.3.** Let **U** be uniformly distributed on  $S^{p-1}$ . Then,

$$\mathbb{E}[(\mathbf{v}'\mathbf{U})^2(\mathbf{w}'\mathbf{U})^2] = \frac{2(\mathbf{v}'\mathbf{w})^2 + 1}{p(p+2)}$$

for any  $\mathbf{v}, \mathbf{w} \in \mathcal{S}^{p-1}$ .

**Proof of Lemma B.3.** Let  $\mathbf{K}_{\ell}$  be the  $\ell^2 \times \ell^2$  commutation matrix and define  $\mathbf{J}_{\ell} := (\text{vec } \mathbf{I}_{\ell})(\text{vec } \mathbf{I}_{\ell})'$ , where vec  $\mathbf{A}$  is the vector stacking the columns of  $\mathbf{A}$  on top of each other (recall that  $\mathbf{I}_{\ell}$  is the  $\ell \times \ell$  identity matrix). Then,

$$\mathbf{E}[(\mathbf{v}'\mathbf{U})^{2}(\mathbf{w}'\mathbf{U})^{2}] = \mathbf{E}[\mathbf{v}'\mathbf{U}\mathbf{U}'\mathbf{v}\mathbf{w}'\mathbf{U}\mathbf{U}'\mathbf{w}] = (\mathbf{v}\otimes\mathbf{v})'\mathbf{E}[\operatorname{vec}(\mathbf{U}\mathbf{U}')\operatorname{vec}'(\mathbf{U}\mathbf{U}')](\mathbf{w}\otimes\mathbf{w})$$

Lemma A.2(iii) from [35] then yields

$$\mathbf{E}[(\mathbf{v}'\mathbf{U})^2(\mathbf{w}'\mathbf{U})^2] = \frac{1}{p(p+2)}(\mathbf{v}\otimes\mathbf{v})'(\mathbf{I}_{p^2} + \mathbf{J}_p + \mathbf{K}_p)(\mathbf{w}\otimes\mathbf{w}) = \frac{2(\mathbf{v}'\mathbf{w})^2 + 1}{p(p+2)},$$

where we used the identities  $(\text{vec } \mathbf{A})'(\text{vec } \mathbf{B}) = \text{tr}[\mathbf{A}'\mathbf{B}]$  and  $\mathbf{K}_p(\mathbf{w} \otimes \mathbf{w}) = (\mathbf{w} \otimes \mathbf{w})\mathbf{K}_1 = (\mathbf{w} \otimes \mathbf{w})$ .

**Lemma B.4.** Under  $P_{\theta_n, F_n}^{(n)}$ ,

(*i*) 
$$\mathrm{E}[(\mathbf{X}'_{ni}\mathbf{X}_{n\ell})^2|\mathbf{X}_{ni}] = \frac{p_n}{p_n - 1}g_{n2}u_{ni}^2 + \frac{f_{n2}}{p_n - 1}$$

*for any*  $i \neq \ell$ *,* 

(*ii*) 
$$E[(\mathbf{X}'_{ni}\mathbf{X}_{n\ell})^{2}(\mathbf{X}'_{nj}\mathbf{X}_{n\ell})^{2}|\mathbf{X}_{ni},\mathbf{X}_{nj}]$$

$$= \frac{e_{n2} - e_{n4}}{p_{n} - 1} + \frac{p_{n}e_{n4} - e_{n2}}{p_{n} - 1}u_{ni}^{2}u_{nj}^{2}$$

$$+ \left(\frac{f_{n4}}{p_{n}^{2} - 1} - \frac{e_{n2} - e_{n4}}{p_{n} - 1}\right)v_{ni}^{2}v_{nj}^{2} + \frac{2f_{n4}}{p_{n}^{2} - 1}v_{ni}^{2}v_{nj}^{2}(\mathbf{S}'_{ni}\mathbf{S}_{nj})^{2} + \frac{4(e_{n2} - e_{n4})}{p_{n} - 1}u_{ni}v_{ni}u_{nj}v_{nj}(\mathbf{S}'_{ni}\mathbf{S}_{nj})$$

for any indices  $i, j, \ell$  such that  $\ell \notin \{i, j\}$ , and

(*iii*) 
$$\mathrm{E}[(\mathbf{X}'_{ni}\mathbf{X}_{n\ell})^2 u_{n\ell}^2 | \mathbf{X}_{ni}] = \frac{e_{n2} - e_{n4}}{p_n - 1} + \frac{p_n e_{n4} - e_{n2}}{p_n - 1} u_{ni}^2$$

*for any*  $i \neq \ell$ *.* 

**Proof of Lemma B.4.** (i) Using (28) and the fact that  $E[(\mathbf{S}'_{ni}\mathbf{S}_{n\ell})^2|\mathbf{X}_{ni}] = 1/(p_n - 1)$  for any  $i \neq \ell$ ,

$$E[(\mathbf{X}'_{ni}\mathbf{X}_{n\ell})^{2}|\mathbf{X}_{ni}] = E[u_{ni}^{2}u_{n\ell}^{2} + v_{ni}^{2}v_{n\ell}^{2}(\mathbf{S}'_{ni}\mathbf{S}_{n\ell})^{2} + 2u_{ni}v_{ni}u_{n\ell}v_{n\ell}(\mathbf{S}'_{ni}\mathbf{S}_{n\ell})|\mathbf{X}_{ni}]$$
  
$$= e_{n2}u_{ni}^{2} + v_{ni}^{2}\frac{f_{n2}}{p_{n}-1} = \left(e_{n2} - \frac{f_{n2}}{p_{n}-1}\right)u_{ni}^{2} + \frac{f_{n2}}{p_{n}-1} = \frac{p_{n}}{p_{n}-1}g_{n2}u_{ni}^{2} + \frac{f_{n2}}{p_{n}-1},$$

where we used the identities  $v_{ni}^2 = 1 - u_{ni}^2$ ,  $f_{n2} = 1 - e_{n2}$  and  $g_2 = e_2 - (1/p_n)$ . (ii) Fix  $i, j, \ell$  with  $\ell \notin \{i, j\}$ . Since  $E[(\mathbf{S}'_{ni}\mathbf{S}_{n\ell})(\mathbf{S}'_{nj}\mathbf{S}_{n\ell})^2|\mathbf{X}_{ni}, \mathbf{X}_{nj}] = 0$ , we have

$$\begin{split} & \mathsf{E}[(\mathbf{X}'_{ni}\mathbf{X}_{n\ell})^{2}(\mathbf{X}'_{nj}\mathbf{X}_{n\ell})^{2}|\mathbf{X}_{ni},\mathbf{X}_{nj}] \\ &= \mathsf{E}[u_{ni}^{2}u_{nj}^{2}u_{n\ell}^{4} + u_{ni}^{2}v_{nj}^{2}u_{n\ell}^{2}v_{n\ell}^{2}(\mathbf{S}'_{nj}\mathbf{S}_{n\ell})^{2} + v_{ni}^{2}u_{nj}^{2}u_{n\ell}^{2}v_{n\ell}^{2}(\mathbf{S}'_{ni}\mathbf{S}_{n\ell})^{2} \\ &+ v_{ni}^{2}v_{nj}^{2}v_{n\ell}^{4}(\mathbf{S}'_{ni}\mathbf{S}_{n\ell})^{2}(\mathbf{S}'_{nj}\mathbf{S}_{n\ell})^{2} + 4u_{ni}v_{ni}u_{nj}v_{nj}u_{n\ell}^{2}v_{n\ell}^{2}(\mathbf{S}'_{ni}\mathbf{S}_{n\ell})(\mathbf{S}'_{nj}\mathbf{S}_{n\ell})|\mathbf{X}_{ni},\mathbf{X}_{nj}]. \end{split}$$

Therefore, applying Lemma B.3 (in the fourth term of the righthand side) and Lemma B.1(i) (in the fifth one) provides

$$\begin{split} & \mathsf{E}[(\mathbf{X}'_{ni}\mathbf{X}_{n\ell})^{2}(\mathbf{X}'_{nj}\mathbf{X}_{n\ell})^{2}|\mathbf{X}_{ni},\mathbf{X}_{nj}] \\ &= e_{n4}u_{ni}^{2}u_{nj}^{2} + \frac{e_{n2} - e_{n4}}{p_{n} - 1}(u_{ni}^{2}v_{nj}^{2} + v_{ni}^{2}u_{nj}^{2}) \\ &+ \frac{f_{n4}}{p_{n}^{2} - 1}v_{ni}^{2}v_{nj}^{2}(1 + 2(\mathbf{S}'_{ni}\mathbf{S}_{nj})^{2}) + \frac{4(e_{n2} - e_{n4})}{p_{n} - 1}u_{ni}v_{ni}u_{nj}v_{nj}(\mathbf{S}'_{ni}\mathbf{S}_{nj}). \end{split}$$

The result then follows by using the identity  $u_{ni}^2 v_{nj}^2 + v_{ni}^2 u_{nj}^2 = 1 - u_{ni}^2 u_{nj}^2 - v_{ni}^2 v_{nj}^2$  (which result from the fact that  $u_{ni}^2 + v_{ni}^2 = 1$  for any *i*). (iii) Finally,

$$E[(\mathbf{X}'_{ni}\mathbf{X}_{n\ell})^{2}u_{n\ell}^{2}|\mathbf{X}_{ni}] = E[(u_{ni}^{2}u_{n\ell}^{2} + v_{ni}^{2}v_{n\ell}^{2}(\mathbf{S}'_{ni}\mathbf{S}_{n\ell})^{2} + 2u_{ni}v_{ni}u_{n\ell}v_{n\ell}(\mathbf{S}'_{ni}\mathbf{S}_{n\ell}))u_{n\ell}^{2}|\mathbf{X}_{ni}]$$
  
$$= e_{n4}u_{ni}^{2} + \frac{e_{n2} - e_{n4}}{p_{n} - 1}v_{ni}^{2} = \frac{e_{n2} - e_{n4}}{p_{n} - 1} + \frac{p_{n}e_{n4} - e_{n2}}{p_{n} - 1}u_{ni}^{2},$$

where we used again the identity  $v_{ni}^2 = 1 - u_{ni}^2$ .

We can now turn to the proof of Theorem 3.1. First note that under the assumptions of this theorem, we have  $e_{n2} = o(1/\sqrt{p_n})$ , which, jointly with Assumptions (a) and (c), entails that  $\operatorname{Var}[Q_n^{\operatorname{St}}] = 1 + o(1)$ . Therefore, it is sufficient to prove that

$$Q_n^{\text{St}} - \mathbb{E}[Q_n^{\text{St}}] \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$
<sup>(29)</sup>

To do so, let  $\mathcal{F}_{n\ell}$ , for  $\ell = 1, ..., n$ , be the  $\sigma$ -algebra generated by  $\mathbf{X}_{n1}, ..., \mathbf{X}_{n\ell}$ , and let  $\mathcal{F}_{n0}$  be the trivial  $\sigma$ algebra  $\{\emptyset, \Omega\}$ . Denoting as  $\mathbf{E}_{n\ell}$  conditional expectation with respect to  $\mathcal{F}_{n\ell}$ , we have

$$Q_n^{\text{St}} - \mathbb{E}[Q_n^{\text{St}}] = \frac{p_n}{n} \sum_{\substack{i < j \\ i, j = 1}}^n \{ (\mathbf{X}'_{ni} \mathbf{X}_{nj})^2 - \mathbb{E}[(\mathbf{X}'_{ni} \mathbf{X}_{nj})^2] \} = \sum_{\ell=1}^n D_{n\ell},$$

where  $D_{n\ell} := E_{n\ell}[Q_n^{\text{St}}] - E_{n,\ell-1}[Q_n^{\text{St}}]$ ,  $\ell = 1, ..., n$ , is a martingale difference sequence with respect to the filtration  $\mathcal{F}_{n1}, \mathcal{F}_{n2}, ..., \mathcal{F}_{nn}$ . Letting throughout a sum over an empty set of indices be zero, we have

$$D_{n\ell} = \frac{p_n}{n} \sum_{\substack{i < j \\ i, j = 1}}^{n} \left\{ E_{n\ell} \Big[ \big( \mathbf{X}'_{ni} \mathbf{X}_{nj} \big)^2 \big] - E_{n,\ell-1} \Big[ \big( \mathbf{X}'_{ni} \mathbf{X}_{nj} \big)^2 \big] \right\}$$

$$= \frac{p_n}{n} \sum_{i=1}^{\ell-1} \left\{ \big( \mathbf{X}'_{ni} \mathbf{X}_{n\ell} \big)^2 - E \Big[ \big( \mathbf{X}'_{ni} \mathbf{X}_{n\ell} \big)^2 | \mathbf{X}_i \big] \right\}$$

$$+ \frac{p_n}{n} \sum_{j=\ell+1}^{n} \sum_{\ell+1 \le j \le n} \left\{ E \Big[ \big( \mathbf{X}'_{n\ell} \mathbf{X}_{nj} \big)^2 | \mathbf{X}_\ell \big] - E \Big[ \big( \mathbf{X}'_{n\ell} \mathbf{X}_{nj} \big)^2 \big] \right\}$$

$$= \frac{p_n}{n} \sum_{i=1}^{\ell-1} \left\{ \big( \mathbf{X}'_{ni} \mathbf{X}_{n\ell} \big)^2 - \Big( \frac{p_n}{p_n - 1} g_{n2} u_{ni}^2 + \frac{f_{n2}}{p_n - 1} \Big) \right\} + \frac{(n-\ell) p_n^2}{n(p_n - 1)} g_{n2} \big( u_{n\ell}^2 - e_{n2} \big), \tag{30}$$

where the last equality follows from Lemma B.2(i) and Lemma B.4(i). Using (30), one can check that one indeed has  $Q_n^{\text{St}} - \mathbb{E}[Q_n^{\text{St}}] = \sum_{\ell=1}^n D_{n\ell}$ . The proof of Theorem 3.1 is then based on the following central limit theorem for martingale differences.

**Theorem B.1 ([7], Theorem 35.12).** Let  $D_{n\ell}$ ,  $\ell = 1, ..., n, n = 1, 2, ...,$  be a triangular array of random variables such that, for any  $n, D_{n1}, D_{n2}, ..., D_{nn}$  is a martingale difference sequence with respect to some filtration  $\mathcal{F}_{n1}, \mathcal{F}_{n2}, ..., \mathcal{F}_{nn}$ . Assume that, for any  $n, \ell$ ,  $D_{n\ell}$  has a finite variance. Letting  $\sigma_{n\ell}^2 = \mathbb{E}[D_{n\ell}^2|\mathcal{F}_{n,\ell-1}]$  (with  $\mathcal{F}_{n0}$  being the trivial  $\sigma$ -algebra { $\emptyset$ ,  $\Omega$ } for all n), further assume that, as  $n \to \infty$ ,

$$\sum_{\ell=1}^{n} \sigma_{n\ell}^2 = 1 + o_{\mathbf{P}}(1) \quad and \quad \sum_{\ell=1}^{n} \mathbb{E}\left[D_{n\ell}^2 \mathbb{I}\left[|D_{n\ell}| > \varepsilon\right]\right] \to 0$$

for any  $\varepsilon > 0$ . Then  $\sum_{\ell=1}^{n} D_{n\ell}$  is asymptotically standard normal.

We establish Theorem 3.1 by proving both following propositions.

**Proposition B.1.** Under the assumptions of Theorem 3.1, (i)  $\sum_{\ell=1}^{n} \mathbb{E}[\sigma_{n\ell}^2] = 1$  and (ii)  $\operatorname{Var}[\sum_{\ell=1}^{n} \sigma_{n\ell}^2] = o(1)$  as  $n \to \infty$ , so that  $\sum_{\ell=1}^{n} \sigma_{n\ell}^2 = 1 + o_P(1)$  as  $n \to \infty$ .

**Proposition B.2.** Under the assumptions of Theorem 3.1,  $\sum_{\ell=1}^{n} \mathbb{E}[D_{n\ell}^2 \mathbb{I}[|D_{n\ell}| > \varepsilon]] \to 0$  for any  $\varepsilon > 0$ .

**Proof of Proposition B.1.** (i) Using (30),  $\sigma_{n\ell}^2 = E_{n,\ell-1}[D_{n\ell}^2]$  takes the form

$$\sigma_{n\ell}^{2} = \frac{p_{n}^{2}}{n^{2}} \sum_{i,j=1}^{\ell-1} \left\{ E_{n,\ell-1} \Big[ (\mathbf{X}_{ni}' \mathbf{X}_{n\ell})^{2} (\mathbf{X}_{nj}' \mathbf{X}_{n\ell})^{2} \Big] \\ - \Big( \frac{p_{n}}{p_{n}-1} g_{n2} u_{ni}^{2} + \frac{f_{n2}}{p_{n}-1} \Big) \Big( \frac{p_{n}}{p_{n}-1} g_{n2} u_{nj}^{2} + \frac{f_{n2}}{p_{n}-1} \Big) \Big\} \\ + \frac{(n-\ell)^{2} p_{n}^{4}}{n^{2} (p_{n}-1)^{2}} g_{n2}^{2} \operatorname{Var} \Big[ u_{n\ell}^{2} \Big] + \frac{2(n-\ell) p_{n}^{3}}{n^{2} (p_{n}-1)} g_{n2} \sum_{i=1}^{\ell-1} E_{n,\ell-1} \Big[ (\mathbf{X}_{ni}' \mathbf{X}_{n\ell})^{2} (u_{n\ell}^{2} - e_{n2}) \Big]$$

$$=: \frac{p_n^2}{n^2} \sum_{i,j=1}^{\ell-1} A_{n,ij} + \frac{(n-\ell)^2 p_n^4}{n^2 (p_n-1)^2} g_{n2}^2 \operatorname{Var} \left[ u_{n\ell}^2 \right] + \frac{2(n-\ell) p_n^3}{n^2 (p_n-1)} g_{n2} \sum_{i=1}^{\ell-1} B_{n,i}.$$

Using the identity  $u_{ni}^2 + u_{ni}^2 = 1 + u_{ni}^2 u_{nj}^2 - v_{ni}^2 v_{nj}^2$ , Lemma B.4(ii) implies that

$$\begin{aligned} A_{n,ij} &= \frac{e_{n2} - e_{n4}}{p_n - 1} + \frac{p_n e_{n4} - e_{n2}}{p_n - 1} u_{ni}^2 u_{nj}^2 + \left(\frac{f_{n4}}{p_n^2 - 1} - \frac{e_{n2} - e_{n4}}{p_n - 1}\right) v_{ni}^2 v_{nj}^2 \\ &+ \frac{2f_{n4}}{p_n^2 - 1} v_{ni}^2 v_{nj}^2 \left(\mathbf{S}_{ni}' \mathbf{S}_{nj}\right)^2 + \frac{4(e_{n2} - e_{n4})}{p_n - 1} u_{ni} v_{ni} u_{nj} v_{nj} \left(\mathbf{S}_{ni}' \mathbf{S}_{nj}\right) \\ &- \frac{p_n^2}{(p_n - 1)^2} g_{n2}^2 u_{ni}^2 u_{nj}^2 - \left(\frac{f_{n2}}{p_n - 1}\right)^2 - \frac{p_n f_{n2} g_{n2}}{(p_n - 1)^2} \left(1 + u_{ni}^2 u_{nj}^2 - v_{ni}^2 v_{nj}^2\right), \end{aligned}$$

which, after some algebra, rewrites

$$A_{n,ij} = \frac{p_n(e_{n4} - e_{n2}^2)}{p_n - 1} u_{ni}^2 u_{nj}^2 + \left(\frac{f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2} + \frac{e_{n4} - e_{n2}^2}{p_n - 1}\right) v_{ni}^2 v_{nj}^2 + \frac{2f_{n4}}{p_n^2 - 1} v_{ni}^2 v_{nj}^2 \left(\mathbf{S}_{ni}' \mathbf{S}_{nj}\right)^2 + \frac{4(e_{n2} - e_{n4})}{p_n - 1} u_{ni} v_{ni} u_{nj} v_{nj} \left(\mathbf{S}_{ni}' \mathbf{S}_{nj}\right) - \frac{e_{n4} - e_{n2}^2}{p_n - 1}.$$

Similarly, after some algebra, Lemma B.4(i) and (iii) yield

$$B_{n,i} = \frac{e_{n2} - e_{n4}}{p_n - 1} + \frac{p_n e_{n4} - e_{n2}}{p_n - 1} u_{ni}^2 - e_{n2} \left( \frac{p_n}{p_n - 1} g_{n2} u_{ni}^2 + \frac{f_{n2}}{p_n - 1} \right)$$
$$= \frac{p_n (e_{n4} - e_{n2}^2)}{p_n - 1} \left( u_{ni}^2 - \frac{1}{p_n} \right).$$

Therefore, we conclude that

$$\sigma_{n\ell}^{2} = \frac{p_{n}^{2}}{n^{2}} \sum_{i,j=1}^{\ell-1} \left\{ \frac{p_{n}(e_{n4} - e_{n2}^{2})}{p_{n} - 1} u_{ni}^{2} u_{nj}^{2} + \left( \frac{f_{n4}}{p_{n}^{2} - 1} - \frac{f_{n2}^{2}}{(p_{n} - 1)^{2}} + \frac{e_{n4} - e_{n2}^{2}}{p_{n} - 1} \right) v_{ni}^{2} v_{nj}^{2} \right. \\ \left. + \frac{2f_{n4}}{p_{n}^{2} - 1} v_{ni}^{2} v_{nj}^{2} \left( \mathbf{S}_{ni}' \mathbf{S}_{nj} \right)^{2} + \frac{4(e_{n2} - e_{n4})}{p_{n} - 1} u_{ni} v_{ni} u_{nj} v_{nj} \left( \mathbf{S}_{ni}' \mathbf{S}_{nj} \right) \right\} \\ \left. - \frac{(\ell - 1)^{2} p_{n}^{2}}{n^{2} (p_{n} - 1)} \left( e_{n4} - e_{n2}^{2} \right) + \frac{(n - \ell)^{2} p_{n}^{4}}{n^{2} (p_{n} - 1)^{2}} \left( e_{n4} - e_{n2}^{2} \right) g_{n2}^{2} \\ \left. + \frac{2(n - \ell) p_{n}^{4}}{n^{2} (p_{n} - 1)^{2}} \left( e_{n4} - e_{n2}^{2} \right) g_{n2} \sum_{i=1}^{\ell-1} \left( u_{ni}^{2} - \frac{1}{p_{n}} \right).$$

$$(31)$$

Splitting the double sum over *i*, *j* according to whether i = j or  $i \neq j$ , taking expectation yields

$$\begin{split} \mathbf{E}\left[\sigma_{n\ell}^{2}\right] &= (\ell-1)(\ell-2)\frac{p_{n}^{2}}{n^{2}}\left\{\frac{p_{n}(e_{n4}-e_{n2}^{2})}{p_{n}-1}e_{n2}^{2} + \left(\frac{f_{n4}}{p_{n}^{2}-1} - \frac{f_{n2}^{2}}{(p_{n}-1)^{2}} + \frac{e_{n4}-e_{n2}^{2}}{p_{n}-1}\right)f_{n2}^{2} \\ &+ \frac{2f_{n4}f_{n2}^{2}}{(p_{n}-1)^{2}(p_{n}+1)}\right\} + (\ell-1)\frac{p_{n}^{2}}{n^{2}}\left\{\frac{p_{n}(e_{n4}-e_{n2}^{2})}{p_{n}-1}e_{n4} + \left(\frac{f_{n4}}{p_{n}^{2}-1} - \frac{f_{n2}^{2}}{(p_{n}-1)^{2}} + \frac{e_{n4}-e_{n2}^{2}}{p_{n}-1}\right)f_{n4} \\ &+ \frac{2f_{n4}^{2}}{p_{n}^{2}-1} + \frac{4(e_{n2}-e_{n4})^{2}}{p_{n}-1}\right\} - \frac{(\ell-1)^{2}p_{n}^{2}}{n^{2}(p_{n}-1)}(e_{n4}-e_{n2}^{2}) + \frac{(n-\ell)^{2}p_{n}^{4}}{n^{2}(p_{n}-1)^{2}}(e_{n4}-e_{n2}^{2})g_{n2}^{2} \\ &+ \frac{2(\ell-1)(n-\ell)p_{n}^{4}}{n^{2}(p_{n}-1)^{2}}(e_{n4}-e_{n2}^{2})g_{n2}^{2}, \end{split}$$

which eventually provides

$$\begin{split} \sum_{\ell=1}^{n} \mathbf{E}[\sigma_{n\ell}^{2}] &= \frac{(n-1)(n-2)p_{n}^{2}}{3n} \left\{ \frac{p_{n}(e_{n4}-e_{n2}^{2})}{p_{n}-1} e_{n2}^{2} + \left(\frac{f_{n4}}{p_{n}^{2}-1} - \frac{f_{n2}^{2}}{(p_{n}-1)^{2}} + \frac{e_{n4}-e_{n2}^{2}}{p_{n}-1}\right) f_{n2}^{2} \\ &+ \frac{2f_{n4}f_{n2}^{2}}{(p_{n}-1)^{2}(p_{n}+1)} \right\} + \frac{(n-1)p_{n}^{2}}{2n} \left\{ \frac{p_{n}(e_{n4}-e_{n2}^{2})}{p_{n}-1} e_{n4} + \left(\frac{f_{n4}}{p_{n}^{2}-1} - \frac{f_{n2}^{2}}{(p_{n}-1)^{2}} + \frac{e_{n4}-e_{n2}^{2}}{p_{n}-1}\right) f_{n4} \\ &+ \frac{2f_{n4}^{2}}{p_{n}^{2}-1} + \frac{4(e_{n2}-e_{n4})^{2}}{p_{n}-1} \right\} - \frac{(n-1)(2n-1)p_{n}^{2}}{6n(p_{n}-1)} (e_{n4}-e_{n2}^{2}) + \frac{(n-1)(2n-1)p_{n}^{4}}{6n(p_{n}-1)^{2}} (e_{n4}-e_{n2}^{2}) g_{n2}^{2} \\ &+ \frac{(n-1)(n-2)p_{n}^{4}}{3n(p_{n}-1)^{2}} (e_{n4}-e_{n2}^{2}) g_{n2}^{2} = 1, \end{split}$$

where the last equality is obtained after painful, yet straightforward, algebra. (ii) Note that (31) implies that

$$\sigma_{n\ell}^{2} = \frac{p_{n}^{2}}{n^{2}} \sum_{i,j=1}^{\ell-1} \left\{ \frac{p_{n}(e_{n4} - e_{n2}^{2})}{p_{n} - 1} u_{ni}^{2} u_{nj}^{2} + \left(\frac{f_{n4}}{p_{n}^{2} - 1} - \frac{f_{n2}^{2}}{(p_{n} - 1)^{2}} + \frac{e_{n4} - e_{n2}^{2}}{p_{n} - 1}\right) v_{ni}^{2} v_{nj}^{2} + \frac{2f_{n4}}{p_{n}^{2} - 1} v_{ni}^{2} v_{nj}^{2} \left(\mathbf{S}_{ni}' \mathbf{S}_{nj}\right)^{2} + \frac{4(e_{n2} - e_{n4})}{p_{n} - 1} u_{ni} v_{ni} u_{nj} v_{nj} \left(\mathbf{S}_{ni}' \mathbf{S}_{nj}\right)\right\} \\ + \frac{2(n - \ell)p_{n}^{4}}{n^{2}(p_{n} - 1)^{2}} \left(e_{n4} - e_{n2}^{2}\right) g_{n2} \sum_{i=1}^{\ell-1} u_{ni}^{2} + C_{n\ell}$$

for some real constant  $C_{n\ell}$ . Therefore,

$$\operatorname{Var}\left[\sum_{\ell=1}^{n}\sigma_{n\ell}^{2}\right] \leq 2\left(\operatorname{Var}[T_{1n}] + \operatorname{Var}[T_{2n}]\right),$$

where we let

$$T_{1n} := \frac{p_n^2}{n^2} \sum_{\ell=1}^n \sum_{i,j=1}^{\ell-1} \left\{ \frac{p_n(e_{n4} - e_{n2}^2)}{p_n - 1} u_{ni}^2 u_{nj}^2 + \left(\frac{f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2} + \frac{e_{n4} - e_{n2}^2}{p_n - 1}\right) v_{ni}^2 v_{nj}^2 + \frac{2f_{n4}}{p_n^2 - 1} v_{ni}^2 v_{nj}^2 \left(\mathbf{S}_{ni}' \mathbf{S}_{nj}\right)^2 + \frac{4(e_{n2} - e_{n4})}{p_n - 1} u_{ni} v_{ni} u_{nj} v_{nj} \left(\mathbf{S}_{ni}' \mathbf{S}_{nj}\right) \right\}$$

and

$$T_{2n} := \frac{2p_n^4}{n^2(p_n-1)^2} (e_{n4} - e_{n2}^2) g_{n2} \sum_{\ell=1}^n (n-\ell) \sum_{i=1}^{\ell-1} u_{ni}^2.$$

In the rest of the proof, C is a positive constant that may change from line to line. Let us start with the variance of  $T_{2n}$ . Since

$$\sum_{\ell=1}^{n} (n-\ell) \sum_{i=1}^{\ell-1} u_{ni}^2 = \sum_{i=1}^{n-1} c_{ni} u_{ni}^2$$

for positive constants  $c_{ni}$  that are upper-bounded by  $n^2$ , we have

$$\operatorname{Var}[T_{2n}] \le C \frac{p_n^8}{n^4 (p_n - 1)^4} \left( e_{n4} - e_{n2}^2 \right)^2 g_{n2}^2 (n - 1) n^4 \left( e_{n4} - e_{n2}^2 \right) \le C n p_n^4 \left( e_{n4} - e_{n2}^2 \right)^3 g_{n2}^2 = o(1)$$

by Assumptions (a) and (c). We turn to  $T_{1n}$ , that can be split into  $T_{1n}^{(a)} + T_{1n}^{(b)}$ , where

$$T_{1n}^{(a)} = \frac{p_n^2}{n^2} \sum_{i=1}^{n-1} (n-i) \left\{ \frac{p_n(e_{n4} - e_{n2}^2)}{p_n - 1} u_{ni}^4 + \left( \frac{3f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2} + \frac{e_{n4} - e_{n2}^2}{p_n - 1} \right) v_{ni}^4 \right\}$$

$$+\frac{4(e_{n2}-e_{n4})}{p_n-1}u_{ni}^2v_{ni}^2\bigg\}$$

and

$$T_{1n}^{(b)} = \frac{2p_n^2}{n^2} \sum_{\substack{i < j \\ i, j = 1}}^{n-1} (n-j) \left\{ \frac{p_n(e_{n4} - e_{n2}^2)}{p_n - 1} u_{ni}^2 u_{nj}^2 + \left( \frac{f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2} + \frac{e_{n4} - e_{n2}^2}{p_n - 1} \right) v_{ni}^2 v_{nj}^2 \right. \\ \left. + \frac{2f_{n4}}{p_n^2 - 1} v_{ni}^2 v_{nj}^2 \left( \mathbf{S}_{ni}' \mathbf{S}_{nj} \right)^2 + \frac{4(e_{n2} - e_{n4})}{p_n - 1} u_{ni} v_{ni} u_{nj} v_{nj} \left( \mathbf{S}_{ni}' \mathbf{S}_{nj} \right) \right\} \\ \left. := \frac{p_n^2}{n^2} \sum_{\substack{i < j \\ i, j = 1}}^{n-1} (n-j) w_{n,ij}.$$

Firstly, since  $\operatorname{Var}[u_{n1}^r v_{n1}^s] \le \operatorname{E}[u_{n1}^{2r} v_{n1}^{2s}] \le e_{n,2r}$  for any  $r, s \ge 0$  (here, we let  $e_{n0} = 1$ ),

$$\begin{aligned} \operatorname{Var}\left[T_{1n}^{(a)}\right] \\ &\leq \frac{p_n^4}{n} \operatorname{Var}\left[\frac{p_n(e_{n4} - e_{n2}^2)}{p_n - 1}u_{n1}^4 + \left(\frac{3f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2} + \frac{e_{n4} - e_{n2}^2}{p_n - 1}\right)v_{n1}^4 + \frac{4(e_{n2} - e_{n4})}{p_n - 1}u_{n1}^2v_{n1}^2\right] \\ &\leq \frac{Cp_n^4}{n} \left\{\frac{p_n^2(e_{n4} - e_{n2}^2)^2}{(p_n - 1)^2}e_{n8} + \left(\frac{3f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2} + \frac{e_{n4} - e_{n2}^2}{p_n - 1}\right)^2 + \frac{16(e_{n2} - e_{n4})^2}{(p_n - 1)^2}e_{n4}\right\} \\ &\leq \frac{C}{n} \left\{p_n^4(e_{n4} - e_{n2}^2)^2e_{n8} + f_{n4}^2 + f_{n2}^4 + p_n^2(e_{n4} - e_{n2}^2)^2 + p_n^2(e_{n2} - e_{n4})^2e_{n4}\right\} = o\left(\frac{1}{n^{1/3}}\right) \end{aligned}$$

under Assumptions (a)-(b). Secondly,

$$\operatorname{Var}[T_{1n}^{(b)}] = \frac{Cp_n^4}{n^4} \sum_{\substack{i < j \\ i, j = 1}}^n \sum_{\substack{k < \ell \\ k, \ell = 1}}^n (n-j)(n-\ell) \operatorname{Cov}[w_{n,ij}, w_{n,k\ell}].$$

Using the same argument as in the proof of Proposition 3.1, this sum over  $i, j, k, \ell$  contains  $\binom{n}{2}$  terms corresponding to  $\operatorname{Var}[w_{n,ij}]$  and n(n-1)(n-2) terms corresponding to  $\operatorname{Cov}[w_{n,ij}, w_{n,k\ell}]$  with  $i \neq j, k \neq \ell$  and  $\#\{i, j, k, \ell\} = 3$  (terms for which the four indices are pairwise different are equal to zero since  $w_{n,ij}$  and  $w_{n,k\ell}$  are then mutually independent). Therefore,

$$\operatorname{Var}[T_{1n}^{(b)}] \le Cp_n^4(\operatorname{Var}[w_{n,12}] + n\operatorname{Cov}[w_{n,12}, w_{n,13}]).$$
(32)

Since  $\operatorname{Var}[Z] \leq \operatorname{E}[Z^2]$  for any random variable Z,

$$\begin{aligned} \operatorname{Var}[w_{n,12}] \\ &\leq C \bigg\{ \frac{p_n^2 (e_{n4} - e_{n2}^2)^2}{(p_n - 1)^2} \operatorname{Var}[u_{n1}^2 u_{n2}^2] + \bigg( \frac{f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2} + \frac{e_{n4} - e_{n2}^2}{p_n - 1} \bigg)^2 \operatorname{Var}[v_{n1}^2 v_{n2}^2] \\ &+ \frac{4f_{n4}^2}{(p_n^2 - 1)^2} \operatorname{Var}[v_{n1}^2 v_{n2}^2 (\mathbf{S}_{n1}' \mathbf{S}_{n2})^2] + \frac{16(e_{n2} - e_{n4})^2}{(p_n - 1)^2} \operatorname{Var}[u_{n1} v_{n1} u_{n2} v_{n2} (\mathbf{S}_{n1}' \mathbf{S}_{n2})] \bigg\} \\ &\leq C \bigg\{ \frac{p_n^2 (e_{n4} - e_{n2}^2)^2}{(p_n - 1)^2} e_{n4}^2 + \bigg( \frac{f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2} + \frac{e_{n4} - e_{n2}^2}{p_n - 1} \bigg)^2 f_{n4}^2 \\ &+ \frac{12f_{n4}^4}{(p_n^2 - 1)^3} + \frac{16(e_{n2} - e_{n4})^4}{(p_n - 1)^3} \bigg\}, \end{aligned}$$

which is  $o(1/p_n^4)$ , since

$$\frac{f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2} = \frac{4e_{n2} - 2 + (p - 1)e_{n4} - (p + 1)e_{n2}^2}{(p^2 - 1)(p - 1)} = o\left(\frac{1}{p_n^2}\right)$$

under Assumption (a). Now,

$$\begin{aligned} \operatorname{Cov}[w_{n,12}, w_{n,13}] \\ &= \frac{p_n^2 (e_{n4} - e_{n2}^2)^2}{(p_n - 1)^2} e_{n4} e_{n2}^2 + \left(\frac{f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2} + \frac{e_{n4} - e_{n2}^2}{p_n - 1}\right)^2 f_{n4} f_{n2}^2 \\ &+ \frac{4f_{n4}^3 f_{n2}^2}{(p_n^2 - 1)^2 (p_n - 1)^2} + \frac{2p_n (e_{n4} - e_{n2}^2)}{p_n - 1} \left(\frac{f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2} + \frac{e_{n4} - e_{n2}^2}{p_n - 1}\right) (e_{n2} - e_{n4}) e_{n2} f_{n2} \\ &+ \frac{4p_n (e_{n4} - e_{n2}^2) f_{n4}}{(p_n - 1)^2 (p_n^2 - 1)} (e_{n2} - e_{n4}) e_{n2} f_{n2} + \frac{4f_{n4}^2 f_{n2}^2}{(p_n^2 - 1) (p_n - 1)} \left(\frac{f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n^2 - 1) - (p_n - 1)^2} + \frac{e_{n4} - e_{n2}^2}{p_n - 1}\right) \\ &- \left\{ \frac{p_n (e_{n4} - e_{n2}^2)}{p_n - 1} e_{n2}^2 + \left(\frac{f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2} + \frac{e_{n4} - e_{n2}^2}{p_n - 1}\right) f_{n2}^2 + \frac{2f_{n4} f_{n2}^2}{(p_n^2 - 1) (p_n - 1)} \right\}^2. \end{aligned}$$

Tedious computations provide

$$\operatorname{Cov}[w_{n,12}, w_{n,13}] = \frac{p_n^4}{(p_n - 1)^4} (e_{n4} - e_{n2}^2)^3 g_{n2}^2,$$

which is  $o(n^{-1}p_n^{-4})$  under Assumptions (a) and (c). From (32), we conclude that  $Var[T_{1n}^{(b)}]$  is o(1), which establishes the result.

**Proof of Proposition B.2.** Since  $\operatorname{Var}[D_{n\ell}] \leq \operatorname{E}[D_{n\ell}^2] = \operatorname{E}[\sigma_{n\ell}^2]$ , Proposition B.1(i) shows that  $\sum_{\ell=1}^{n} \operatorname{Var}[D_{n\ell}] \leq \sum_{\ell=1}^{n} \operatorname{E}[\sigma_{n\ell}^2] = 1$ . Therefore, applying Cauchy–Schwarz inequality, Chebychev inequality, then Cauchy–Schwarz inequality again, yields

$$\sum_{\ell=1}^{n} \mathbb{E}\left[D_{n\ell}^{2} \mathbb{I}\left[|D_{n\ell}| > \varepsilon\right]\right] \leq \sum_{\ell=1}^{n} \sqrt{\mathbb{E}\left[D_{n\ell}^{4}\right]} \sqrt{\mathbb{P}\left[|D_{n\ell}| > \varepsilon\right]} \leq \frac{1}{\varepsilon} \sum_{\ell=1}^{n} \sqrt{\mathbb{E}\left[D_{n\ell}^{4}\right]} \sqrt{\operatorname{Var}[D_{n\ell}]}$$
$$\leq \frac{1}{\varepsilon} \sqrt{\sum_{\ell=1}^{n} \mathbb{E}\left[D_{n\ell}^{4}\right]} \sqrt{\sum_{\ell=1}^{n} \operatorname{Var}[D_{n\ell}]} \leq \frac{1}{\varepsilon} \sqrt{\sum_{\ell=1}^{n} \mathbb{E}\left[D_{n\ell}^{4}\right]}.$$
(33)

Letting

$$Y_{ni\ell} := \left(\mathbf{X}'_{ni}\mathbf{X}_{n\ell}\right)^2 - \frac{p_n}{p_n - 1}g_{n2}u_{ni}^2 - \frac{f_{n2}}{p_n - 1},$$

(30) provides

$$\mathbb{E}[D_{n\ell}^{4}] \leq \frac{Cp_{n}^{4}}{n^{4}} \mathbb{E}\left[\left(\sum_{i=1}^{\ell-1} Y_{ni\ell}\right)^{4}\right] + \frac{C(n-\ell)^{4}p_{n}^{8}}{n^{4}(p_{n}-1)^{4}}g_{n2}^{4}\mathbb{E}[\left(u_{n\ell}^{2}-e_{n2}\right)^{4}],$$

hence

$$\sum_{\ell=1}^{n} \mathbb{E}[D_{n\ell}^{4}] \leq \frac{Cp_{n}^{4}}{n^{4}} \sum_{\ell=1}^{n} \sum_{i,j,r,s=1}^{\ell-1} \mathbb{E}[Y_{ni\ell}Y_{nj\ell}Y_{nr\ell}Y_{ns\ell}] + Cnp_{n}^{4}g_{n2}^{4}e_{n8}$$
$$= \frac{Cp_{n}^{4}}{n^{4}} \sum_{\ell=1}^{n} \sum_{i,j,r,s=1}^{\ell-1} \mathbb{E}[Y_{ni\ell}Y_{nj\ell}Y_{nr\ell}Y_{ns\ell}] + o\left(\frac{1}{n^{1/3}}\right).$$
(34)

Now, in the sum over *i*, *j*, *r*, *s*, there are  $\ell - 1 \le n$  terms for which  $\{i, j, r, s\}$  has cardinality one; for these terms, (28) yields

$$E[|Y_{ni\ell}Y_{nj\ell}Y_{nr\ell}Y_{ns\ell}|] = E[Y_{n1\ell}^{4}] \le CE\left[\left(\mathbf{X}_{n1}'\mathbf{X}_{n\ell}\right)^{8} + \frac{p_{n}^{4}}{(p_{n}-1)^{4}}g_{n2}^{4}u_{n1}^{8} + \frac{f_{n2}^{4}}{(p_{n}-1)^{4}}\right]$$
$$\le C\left\{E\left[u_{n1}^{8}u_{n\ell}^{8} + v_{n1}^{8}v_{n\ell}^{8}\left(\mathbf{S}_{n1}'\mathbf{S}_{n\ell}\right)^{8} + u_{n1}^{4}v_{n\ell}^{4}u_{n\ell}^{4}v_{n\ell}^{4}\left(\mathbf{S}_{n1}'\mathbf{S}_{n\ell}\right)^{4}\right] + g_{n2}^{4}e_{n8} + \frac{f_{n2}^{4}}{p_{n}^{4}}\right\}$$
$$\le C\left\{e_{n8}^{2} + \frac{105}{(p_{n}^{2}-1)(p_{n}+3)(p_{n}+5)} + \frac{3e_{n4}^{2}}{p_{n}^{2}-1} + g_{n2}^{4}e_{n8} + \frac{f_{n2}^{4}}{p_{n}^{4}}\right\} = o\left(\frac{n^{4/3}}{p_{n}^{4}}\right), \tag{35}$$

where we used Lemma B.1(iv), the identities  $v_{n1}^4$ ,  $v_{n\ell}^4 \leq 1$ , and Assumptions (a)–(c). In the sum over *i*, *j*, *r*, *s*, there are  $3(\ell - 1)(\ell - 2) \leq 3n^2$  for which  $\{i, j, r, s\}$  has cardinality two and contains two pairs of equal indices. For such terms, Lemma B.2(ii) yields

$$\mathbb{E}\left[|Y_{ni\ell}Y_{nj\ell}Y_{nr\ell}Y_{ns\ell}|\right] = \left(\mathbb{E}\left[Y_{n1\ell}^{2}\right]\right)^{2} \leq C \left\{\mathbb{E}\left[\left(\mathbf{X}_{ni}'\mathbf{X}_{n\ell}\right)^{4}\right] + \frac{p_{n}^{2}}{(p_{n}-1)^{2}}g_{n2}^{2}e_{n4} + \frac{f_{n2}^{2}}{(p_{n}-1)^{2}}\right\}^{2} \\ \leq C \left\{e_{n4}^{2} + \frac{2(e_{n2}-e_{n4})^{2}}{p_{n}-1} + \frac{3f_{n4}^{2}}{p_{n}^{2}-1} + g_{n2}^{2}e_{n4} + \frac{f_{n2}^{2}}{p_{n}^{2}}\right\}^{2} = O\left(\frac{1}{p_{n}^{4}}\right).$$

Similarly, the sum over i, j, r, s in (34) contains no more than  $C_1 n^2$  terms (where  $C_1$  does not depend on  $\ell$ ) such that  $\{i, j, r, s\}$  has cardinality two and contains a triple of equal indices. For such terms, Lemma B.4(i) yields

$$E[Y_{ni\ell}Y_{nj\ell}Y_{nr\ell}Y_{ns\ell}] = E[E[Y_{n1\ell}^{3}Y_{n2\ell}|\mathbf{X}_{n\ell}]] = E[E[Y_{n1\ell}^{3}|\mathbf{X}_{n\ell}]E[Y_{n2\ell}|\mathbf{X}_{n\ell}]]$$
  
$$= \frac{p_n}{p_n - 1}g_{n2}E[E[Y_{n1\ell}^{3}|\mathbf{X}_{n\ell}](u_{n\ell}^2 - e_{n2})] = \frac{p_n}{p_n - 1}g_{n2}E[Y_{n1\ell}^{3}(u_{n\ell}^2 - e_{n2})],$$

so that the Hölder inequality provides

$$\begin{aligned} \left| \mathbb{E}[Y_{ni\ell}Y_{nj\ell}Y_{nr\ell}Y_{ns\ell}] \right| &\leq Cg_{n2}\mathbb{E}[Y_{n1\ell}^{3}(u_{n\ell}^{2} - e_{n2})] \leq Cg_{n2} \left(\mathbb{E}[Y_{nr\ell}^{4}]\right)^{3/4} \left(\mathbb{E}[(u_{n\ell}^{2} - e_{n2})^{4}]\right)^{1/4} \\ &\leq Cg_{n2} \left(\mathbb{E}[Y_{nr\ell}^{4}]\right)^{3/4} e_{n8}^{1/4} = O\left(\frac{1}{\sqrt{np_{n}}}\right) o\left(\frac{n}{p_{n}^{3}}\right) o\left(\frac{n^{1/6}}{\sqrt{p_{n}}}\right) = o\left(\frac{n^{2/3}}{p_{n}^{4}}\right), \end{aligned}$$

where we used (35). The sum over i, j, r, s in (34) contains no more than  $C_2 n^3$  (where  $C_2$  does not depend on  $\ell$ ) terms such that  $\{i, j, r, s\}$  has cardinality three. Proceeding as above, the corresponding terms are seen to satisfy

$$\begin{split} \mathbf{E}[Y_{ni\ell}Y_{nj\ell}Y_{nr\ell}Y_{ns\ell}] &= \mathbf{E}\Big[\mathbf{E}\Big[Y_{n1\ell}^2|\mathbf{X}_{n\ell}\Big]\mathbf{E}[Y_{n2\ell}|\mathbf{X}_{n\ell}]\mathbf{E}[Y_{n3\ell}|\mathbf{X}_{n\ell}]\Big] \\ &= \frac{p_n^2}{(p_n-1)^2}g_{n2}^2\mathbf{E}\Big[\mathbf{E}\Big[Y_{n1\ell}^2|\mathbf{X}_{n\ell}\Big]\big(u_{n\ell}^2 - e_{n2}\big)^2\Big] = \frac{p_n}{p_n-1}g_{n2}^2\mathbf{E}\Big[Y_{n1\ell}^2\big(u_{n\ell}^2 - e_{n2}\big)^2\Big], \end{split}$$

which, by using the Cauchy-Schwarz inequality, yields

$$\begin{aligned} |\mathrm{E}[Y_{ni\ell}Y_{nj\ell}Y_{nr\ell}Y_{ns\ell}]| &\leq Cg_{n2}^{2}\mathrm{E}[Y_{n1\ell}^{2}\left(u_{n\ell}^{2}-e_{n2}\right)^{2}] \leq Cg_{n2}^{2}\sqrt{\mathrm{E}[Y_{n1\ell}^{4}]}\sqrt{\mathrm{E}[\left(u_{n\ell}^{2}-e_{n2}\right)^{4}]} \\ &\leq Cg_{n2}^{2}\sqrt{\mathrm{E}[Y_{nr\ell}^{4}]}\sqrt{e_{n8}} = O\left(\frac{1}{np_{n}}\right)o\left(\frac{n^{2/3}}{p_{n}^{2}}\right)o\left(\frac{n^{1/3}}{p_{n}}\right) = o\left(\frac{1}{p_{n}^{4}}\right).\end{aligned}$$

Finally, there obviously are less than  $(\ell - 1)^4 \le n^4$  terms such that  $\{i, j, r, s\}$  has cardinality four, and these terms are such that

$$\begin{split} \mathsf{E}[Y_{ni\ell}Y_{nj\ell}Y_{nr\ell}Y_{ns\ell}] &= \left|\mathsf{E}\big[\mathsf{E}[Y_{n1\ell}|\mathbf{X}_{n\ell}]\mathsf{E}[Y_{n2\ell}|\mathbf{X}_{n\ell}]\mathsf{E}[Y_{n3\ell}|\mathbf{X}_{n\ell}]\mathsf{E}[Y_{n4\ell}|\mathbf{X}_{n\ell}]\big]\right| \\ &= \frac{p_n^4}{(p_n-1)^4}g_{n2}^4\mathsf{E}\big[\big(u_{n\ell}^2 - e_{n2}\big)^4\big] \le Cg_{n2}^4e_{n8} = O\left(\frac{1}{n^2p_n^2}\right)o\left(\frac{n^{2/3}}{p_n^2}\right) = o\left(\frac{1}{n^{4/3}p_n^4}\right). \end{split}$$

Altogether, (34) thus yields

$$\sum_{\ell=1}^{n} \mathbb{E}[D_{n\ell}^{4}] = \frac{Cnp_{n}^{4}}{n^{4}} \left\{ no\left(\frac{n^{4/3}}{p_{n}^{4}}\right) + 3n^{2}O\left(\frac{1}{p_{n}^{4}}\right) + C_{1}n^{2}o\left(\frac{n^{2/3}}{p_{n}^{4}}\right) + C_{2}n^{3}o\left(\frac{1}{p_{n}^{4}}\right) + n^{4}o\left(\frac{1}{n^{4/3}p_{n}^{4}}\right) \right\} + o\left(\frac{1}{n^{1/3}}\right) = o(1).$$

From (33), we thus conclude that

$$\sum_{\ell=1}^{n} \mathbb{E}\left[D_{n\ell}^{2} \mathbb{I}\left[|D_{n\ell}| > \varepsilon\right]\right] \leq \frac{1}{\varepsilon} \sqrt{\sum_{\ell=1}^{n} \mathbb{E}\left[D_{n\ell}^{4}\right] = o(1)},$$

which establishes the result.

**Proof of Theorem 3.2.** (i)–(ii) In these cases, we have  $\sqrt{np_n}g_{n2} \rightarrow \xi$ , with  $\xi = 0$  in case (i) and  $\xi \neq 0$  in case (ii). Under Assumption (a) and  $g_{n2} = O(1/\sqrt{np_n})$ , Proposition 3.1 then yields

$$\mu_n := \mathbf{E}[Q_n^{\text{St}}] = \frac{p_n^2(n-1)}{2(p_n-1)}g_{n2}^2 = \frac{\xi^2}{2} + o(1)$$

and

$$\sigma_n^2 := \operatorname{Var}[Q_n^{\operatorname{St}}] = \frac{(n-1)p_n^2}{2n} \left\{ \frac{3f_{n4}^2}{p_n^2 - 1} - \left(O\left(\frac{1}{np_n}\right) + \frac{1}{p_n}\right)^2 \right\} + o(1) = 1 + o(1),$$

so that Theorem 3.1 and Slutzky's lemma provide

$$Q_n^{\text{St}} = \sigma_n \left( \frac{Q_n^{\text{St}} - \text{E}[Q_n^{\text{St}}]}{\sqrt{\text{Var}[Q_n^{\text{St}}]}} \right) + \mu_n \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\xi^2}{2}, 1\right),$$

as was to be shown. (iii) Assume that  $\sqrt{np_n}|g_{n2}| \to \infty$  and fix M > 0 (clearly, it is enough to prove the result for M > 0). Proposition 3.1 then ensures that  $\mu_n$  diverges to infinity, so that there exists  $n_0$  such that  $\mu_n > M$  for any  $n \ge n_0$ . For any such n, the Chebychev inequality yields

$$P_{\theta_n,F_n}^{(n)} [Q_n^{\text{St}} \le M] = P_{\theta_n,F_n}^{(n)} [Q_n^{\text{St}} - \mu_n \le M - \mu_n] \\ \le P_{\theta_n,F_n}^{(n)} [|Q_n^{\text{St}} - \mu_n| \ge \mu_n - M] \le \frac{\sigma_n^2}{(\mu_n - M)^2} = \frac{\sigma_n^2}{\mu_n^2} (1 + o(1)).$$

Now, if  $e_{n4} = o(ng_{n2}^2)$ , then we also have  $p_n e_{n2}^2 = o(np_n g_{n2}^2)$  and  $p_n e_{n2}^2 = o((np_n g_{n2}^2)^2)$ . Therefore,

$$\begin{split} \frac{\sigma_n^2}{\mu_n^2} &\leq \frac{Cp_n^2}{(np_ng_{n2}^2)^2} \left\{ e_{n4}^2 + \frac{6}{p_n - 1} (e_{n2} - e_{n4})^2 + \frac{3f_{n4}^2}{p_n^2 - 1} - \left(\frac{p_n}{p_n - 1}g_{n2}^2 + \frac{1}{p_n}\right)^2 \right\} \\ &+ \frac{Cnp_n^2}{(np_ng_{n2}^2)^2} (e_{n4} - e_{n2}^2)g_{n2}^2 \\ &\leq \frac{C}{(np_ng_{n2}^2)^2} \left\{ p_n^2 e_{n4}^2 + 12p_n(e_{n2} - e_{n4})^2 \right\} + \frac{Cp_n}{np_ng_{n2}^2} (e_{n4} - e_{n2}^2) + o(1) = o(1), \end{split}$$

which implies that  $P_{\theta_n, F_n}^{(n)}[Q_n^{\text{St}} \le M] \to 0$ , hence establishes the result.

## **Appendix C:** Proofs for Section 4

**Proof of Proposition 4.1.** Since  $\kappa_n$  is assumed to be  $o(p_n)$  as  $n \to \infty$ , Lemma A.1 provides

$$e_{n2} = \frac{1}{\kappa_n} \left(\frac{c_{p_n}}{c_{p_n,\kappa_n,f}}\right)^{-1} c_{p_n} \int_{-1}^{1} (1-s^2)^{(p_n-3)/2} (\kappa_n s^2) f(\kappa_n s^2) ds$$
$$= \frac{\frac{1}{\kappa_n} (\frac{\kappa_n}{p_n} + \frac{3\kappa_n^2}{p_n^2} + o(\frac{\kappa_n^2}{p_n^2}))}{1 + \frac{\kappa_n}{p_n} + o(\frac{\kappa_n}{p_n})} = \frac{\frac{1}{p_n} + \frac{3\kappa_n}{p_n^2} + o(\frac{\kappa_n}{p_n^2})}{1 + \frac{\kappa_n}{p_n} + o(\frac{\kappa_n}{p_n})} = \frac{1}{p_n} + \frac{2\kappa_n}{p_n^2} + o\left(\frac{\kappa_n}{p_n^2}\right),$$

which proves the result for  $e_{n2}$ . The same lemma also yields

$$e_{n4} = \frac{1}{\kappa_n^2} \left( \frac{c_{p_n}}{c_{p_n,\kappa_n,f}} \right)^{-1} c_{p_n} \int_{-1}^1 (1 - s^2)^{(p_n - 3)/2} (\kappa_n s^2)^2 f(\kappa_n s^2) ds$$
$$= \frac{\frac{1}{\kappa_n^2} (\frac{3\kappa_n^2}{p_n^2} + o(\frac{\kappa_n^2}{p_n^2}))}{1 + \frac{\kappa_n}{p_n} + o(\frac{\kappa_n}{p_n})} = \frac{\frac{3}{p_n^2} + o(\frac{1}{p_n^2})}{1 + \frac{\kappa_n}{p_n} + o(\frac{\kappa_n}{p_n})} = \frac{3}{p_n^2} + o\left(\frac{1}{p_n^2}\right).$$

The claim for  $e_{n8}$  directly follows from the identity  $e_{n8} \le e_{n4}$ .

**Proof of Theorem 4.1.** First note that, in all cases (i)–(iii), we have  $\kappa_n = o(p_n)$ , so that Proposition 4.1 applies and ensures that conditions (a)–(b) in Theorem 3.2 are fulfilled. Let us then treat cases (i)–(iii) separately. (i) Since  $\kappa_n =$  $o(p_n^{3/2}/\sqrt{n})$ , Proposition 4.1 implies that

$$g_{n2} = e_{n2} - \frac{1}{p_n} = O\left(\frac{\kappa_n}{p_n^2}\right) = o\left(\frac{1}{\sqrt{np_n}}\right),$$

so that Theorem 3.2(i) shows that  $Q_n^{\text{St}}$  is asymptotically standard normal. (ii) Since  $\sqrt{n\kappa_n}/p_n^{3/2} \rightarrow \tau (\neq 0)$  and  $p_n = o(n)$ , Proposition 4.1 provides

$$\sqrt{np_n}g_{n2} = \frac{2\sqrt{n\kappa_n}}{p_n^{3/2}} + o\left(\frac{\sqrt{n\kappa_n}}{p_n^{3/2}}\right) = 2\tau + o(1),$$

Theorem 3.2(ii) shows that  $Q_n^{\text{St}} \xrightarrow{\mathcal{D}} \mathcal{N}(\xi^2/2, 1)$ , with  $\xi = 2\tau$ , which establishes the result. (iii) The claim follows from Theorem 3.2(iii) since

$$\sqrt{np_n}|g_{n2}| = \frac{2\sqrt{n\kappa_n}}{p_n^{3/2}} + o\left(\frac{\sqrt{n\kappa_n}}{p_n^{3/2}}\right)$$

`

diverges to infinity.

We now turn to the proof of Theorem 4.2, that requires both following technical results providing asymptotic expansions of confluent hypergeometric functions of matrix arguments; see, e.g., Section 7.3 of [33] for the definition of these functions.

**Lemma C.1.** Let  $\mathcal{A}$  be the collection of  $\ell \times \ell$  real symmetric matrices whose eigenvalues all belong to (-1, 1). Then, (i) for any  $\mathbf{A} \in \mathcal{A}$ ,

$${}_{1}F_{1}\left(a,\frac{m}{2},\frac{m}{2}\mathbf{A}\right)$$
  
=  $|\mathbf{I}_{\ell} - \mathbf{A}|^{-a}\left\{1 - \frac{a}{2m}\left(\left(\operatorname{tr}[\mathbf{A}_{1}]\right)^{2} + (2a+1)\operatorname{tr}[\mathbf{A}_{1}^{2}]\right) + \frac{1}{m^{2}}P_{a}(\mathbf{A}_{1}) + O\left(\frac{1}{m^{3}}\right)\right\}$ 

as  $m \to \infty$ , where we let  $\mathbf{A}_1 := \mathbf{A}(\mathbf{I}_{\ell} - \mathbf{A})^{-1}$  and where  $\mathbf{B} \mapsto P_a(\mathbf{B})$  is a continuous function that, for any  $\ell \times \ell$  matrix  $\mathbf{B}$ , satisfies  $P_a(\lambda \mathbf{B}) = O(\lambda^2)P_a(\mathbf{B})$  as  $\lambda \to 0$ . Moreover, (ii) this expansion is uniform over any compact subset of  $\mathcal{A}$  in the sense that

$$\sup_{\mathbf{A}\in\mathcal{K}} \left| |\mathbf{I}_{\ell} - \mathbf{A}|_{1}^{a} F_{1}\left(a, \frac{m}{2}, \frac{m}{2}\mathbf{A}\right) - \left\{ 1 - \frac{a}{2m} \left( \left( \operatorname{tr}[\mathbf{A}_{1}]\right)^{2} + (2a+1) \operatorname{tr}[\mathbf{A}_{1}^{2}] \right) + \frac{1}{m^{2}} P_{a}(\mathbf{A}_{1}) \right\} \right| = O\left(\frac{1}{m^{3}}\right)$$

for any compact subset  $\mathcal{K}$  of  $\mathcal{A}$ .

**Proof of Lemma C.1.** For any partition  $\kappa = (k_1, k_2, ..., k_\ell)$  of the nonnegative integer k (that is,  $k = k_1 + k_2 + \cdots + k_\ell$ , with  $k_1 \ge k_2 \ge \cdots \ge k_\ell \ge 0$ ), we let  $(a)_{\kappa} := \prod_{i=1}^{\ell} (a - \frac{1}{2}(i-1))_{k_i}$ , where  $(c)_r := c(c+1) \dots (c+r-1)$  for r > 0 and  $(c)_0 := 1$ , and we denote as  $C_{\kappa}(\mathbf{B})$  the zonal polynomial associated with  $\kappa$  evaluated at the  $\ell \times \ell$  matrix  $\mathbf{B}$ ; see, e.g., Section 7.2 of [33]. With this standard notation,

$${}_{1}F_{1}\left(a,\frac{m}{2},\frac{m}{2}\mathbf{A}\right) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a)_{\kappa}}{(\frac{m}{2})_{\kappa}} \frac{C_{\kappa}(\frac{m}{2}\mathbf{A})}{k!} = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a)_{\kappa}}{(\frac{m}{2})_{\kappa}} \frac{(\frac{m}{2})^{k}C_{\kappa}(\mathbf{A})}{k!},$$

where the sum in  $\kappa$  is over all partitions  $\kappa$  of k (the last equality results from the homogeneity properties of zonal polynomials). Expanding  $(\frac{m}{2})_{\kappa}/(\frac{m}{2})^k$  then yields

$${}_{1}F_{1}\left(a,\frac{m}{2},\frac{m}{2}\mathbf{A}\right) = \sum_{k=0}^{\infty} \sum_{\kappa} (a)_{\kappa} \left(1 - \frac{1}{m}a_{1}(\kappa) + \frac{1}{6m^{2}}\left(3a_{1}^{2}(\kappa) + a_{2}(\kappa) - k\right) + O\left(\frac{1}{m^{3}}\right)\right) \frac{C_{\kappa}(\mathbf{A})}{k!}$$

as  $m \to \infty$ , with  $a_1(\kappa) = \sum_{i=1}^{\ell} k_i(k_i - i)$  and  $a_2(\kappa) = \sum_{i=1}^{\ell} k_i(4k_i^2 - 6k_ii + 3i^2)$ ; see (5.8) in [21]. Using the identities (1.6) and (2.18) from the same paper then provides

$${}_{1}F_{1}\left(a,\frac{m}{2},\frac{m}{2}\mathbf{A}\right) = |\mathbf{I}_{\ell} - \mathbf{A}|^{-a} \left\{ 1 - \frac{a}{2m} \left( \left( \operatorname{tr}[\mathbf{A}_{1}]\right)^{2} + (2a+1)\operatorname{tr}[\mathbf{A}_{1}^{2}] \right) \right\} + \sum_{k=0}^{\infty} \sum_{\kappa} (a)_{\kappa} \left( \frac{1}{6m^{2}} \left( 3a_{1}^{2}(\kappa) + a_{2}(\kappa) - k \right) + O\left(\frac{1}{m^{3}}\right) \right) \frac{C_{\kappa}(\mathbf{A})}{k!}$$

as  $m \to \infty$ . Part (i) of the result then follows from the other identities in Lemma 3 from [21] (it is readily checked that the terms in  $P_a(\mathbf{B})$  of the form  $c_a \operatorname{tr}[\mathbf{B}]$ , which are the only ones that would prevent  $P_a(\lambda \mathbf{B})$  to be  $O(\lambda^2)P_a(\mathbf{B})$  as  $\lambda \to 0$ , cancel out). Since Part (ii) readily follows by noting that the  $O(1/m^3)$  quantity above does not depend on  $\mathbf{A}$ , this establishes the result.

The expansion in Lemma C.1 is a second-order refinement of the first-order expansion in (A.6.17) from [13]. Without this refinement, the condition  $p_n/\sqrt[3]{n} \to \infty$  in Theorem 4.2 would have to be reinforced into  $p_n/\sqrt{n} \to \infty$ . We conjecture that even higher-order expansions would allow us to weaken the condition  $p_n/\sqrt[3]{n} \to \infty$  into  $p_n/n^{\delta} \to \infty$  for an arbitrary  $\delta > 0$ , but we did not try to prove this as obtaining such higher-order expansions is extremely laborious.

The following result, that will be needed in the proof of Theorem 4.2, is now a rather direct corollary of Lemma C.1.

**Lemma C.2.** Let  $(p_n)$  and  $(\kappa_n)$  satisfy the assumptions of Theorem 4.2. Let  $\mathbf{M}_z$  be a  $p_n \times p_n$  matrix whose only non-zero entries are  $(\mathbf{M}_z)_{11} = (1+z)/2$  and  $(\mathbf{M}_z)_{22} = (1-z)/2$ , with  $z \in [-1, 1]$ . Then, with  $t_n := \kappa_n/p_n$ ,

$$\left({}_{1}F_{1}\left(\frac{1}{2},\frac{p_{n}}{2},t_{n}p_{n}\right)\right)^{2n} = (1-2t_{n})^{-n}(1+o(1))$$

and

$$\left({}_{1}F_{1}\left(\frac{1}{2},\frac{p_{n}}{2},2t_{n}p_{n}\mathbf{M}_{z}\right)\right)^{n} = \left((1-2t_{n})^{2}-4t_{n}^{2}z^{2}\right)^{-n/2}\left(1+o(1)\right)$$

as  $n \to \infty$ .

**Proof of Lemma C.2.** In this proof, all o's and O's are as  $n \to \infty$ . Using Lemma C.1(ii), we have

$${}_{1}F_{1}\left(\frac{1}{2}, \frac{p_{n}}{2}, t_{n}p_{n}\right) = \frac{1}{(1-2t_{n})^{1/2}} \left(1 - \frac{3t_{n}^{2}}{p_{n}(1-2t_{n})^{2}} + O\left(\frac{t_{n}^{2}}{p_{n}^{2}}\right) + O\left(\frac{1}{p_{n}^{3}}\right)\right)$$
$$= \frac{1}{(1-2t_{n})^{1/2}} \left(1 + O\left(\frac{t_{n}^{2}}{p_{n}}\right) + O\left(\frac{1}{p_{n}^{3}}\right)\right).$$

Since  $t_n^2 = o(p_n/n)$  and  $n/p_n^3 = o(1)$ , this yields

$$(1-2t_n)^n \left( {}_1F_1\left(\frac{1}{2}, \frac{p_n}{2}, t_n p_n\right) \right)^{2n} = \left(1+o\left(\frac{1}{n}\right)\right)^{2n} = 1+o(1),$$

which establishes the first result. For the second one, first note that the definition of hypergeometric functions implies that  ${}_{1}F_{1}(\frac{1}{2}, \frac{p_{n}}{2}, 2t_{n}p_{n}\mathbf{M}_{z}) = {}_{1}F_{1}(\frac{1}{2}, \frac{p_{n}}{2}, 2t_{n}p_{n}\mathbf{N}_{z})$ , where  $\mathbf{N}_{z} = \text{diag}((1+z)/2, (1-z)/2)$  is the 2 × 2 upper-left block of  $\mathbf{M}_z$ . Lemma C.1(ii) then similarly yields

$${}_{1}F_{1}\left(\frac{1}{2},\frac{p_{n}}{2},2t_{n}p_{n}\mathbf{N}_{z}\right) = \frac{1}{|\mathbf{I}_{2}-4t_{n}\mathbf{N}_{z}|^{1/2}}\left(1-\frac{1}{4p_{n}}\left(\left(\mathrm{tr}[\mathbf{A}_{n,z}]\right)^{2}+2\,\mathrm{tr}[\mathbf{A}_{n,z}^{2}]\right)+O\left(\frac{t_{n}^{2}}{p_{n}^{2}}\right)+O\left(\frac{1}{p_{n}^{3}}\right)\right),$$

with  $\mathbf{A}_{n,z} := 4t_n \mathbf{N}_z (\mathbf{I}_2 - 4t_n \mathbf{N}_z)^{-1}$ . Direct computations show that  $|\mathbf{I}_2 - 4t_n \mathbf{N}_z| = (1 - 2t_n)^2 - 4t_n^2 z^2$  for *n* large enough and that  $(\operatorname{tr}[\mathbf{A}_{n,z}])^2 = O(t_n^2) = \operatorname{tr}[\mathbf{A}_{n,z}^2]$ . Therefore, proceeding as for the first result,

$$\left((1-2t_n)^2 - 4t_n^2 z^2\right)^{n/2} \left({}_1F_1\left(\frac{1}{2}, \frac{p_n}{2}, 2t_n p_n \mathbf{M}_z\right)\right)^n = \left(1 + O\left(\frac{t_n^2}{p_n}\right) + O\left(\frac{1}{p_n^3}\right)\right)^n = \left(1 + O\left(\frac{1}{n}\right)\right)^n = 1 + o(1),$$
  
n establishes the second result.

which establishes the second result.

We can now prove Theorem 4.2.

**Proof of Theorem 4.2.** We proceed along the same lines as in the proof of Theorem 1 from [10]. Consider the set  $V_n :=$  $\{\pm 1\}^{p_n}$ , that has cardinality  $2^{p_n}$ , and the mixture distribution

$$\mathbf{P}_{1}^{(n)} := \frac{1}{2^{p_n}} \sum_{\mathbf{v}_n \in \mathcal{V}_n} \mathbf{P}_{\mathbf{v}_n/\sqrt{p_n},\kappa_n,f_W}^{(n)}.$$

For any test  $\phi_n$ , we then have

$$\begin{split} & \operatorname{E}_{\mathsf{P}_{0}^{(n)}}[\phi_{n}] + \sup_{\theta_{n} \in \mathcal{S}^{p_{n-1}}} \operatorname{E}_{\mathsf{P}_{\theta_{n},\kappa_{n},f_{W}}}[1-\phi_{n}] \\ &= \sup_{\theta_{n} \in \mathcal{S}^{p_{n-1}}} \left( \operatorname{E}_{\mathsf{P}_{0}^{(n)}}[\phi_{n}] + \operatorname{E}_{\mathsf{P}_{\theta_{n},\kappa_{n},f_{W}}^{(n)}}[1-\phi_{n}] \right) \\ &\geq \inf_{\psi_{n}} \sup_{\theta_{n} \in \mathcal{S}^{p_{n-1}}} \left( \operatorname{E}_{\mathsf{P}_{0}^{(n)}}[\psi_{n}] + \operatorname{E}_{\mathsf{P}_{\theta_{n},\kappa_{n},f_{W}}^{(n)}}[1-\psi_{n}] \right) \geq \inf_{\psi_{n}} \frac{1}{2^{p_{n}}} \sum_{\mathbf{v}_{n} \in \mathcal{V}_{n}} \left( \operatorname{E}_{\mathsf{P}_{0}^{(n)}}[\psi_{n}] + \operatorname{E}_{\mathsf{P}_{n,\kappa_{n},f_{W}}^{(n)}}[1-\psi_{n}] \right) \\ &= \inf_{\psi_{n}} \left( \operatorname{E}_{\mathsf{P}_{0}^{(n)}}[\psi_{n}] + \operatorname{E}_{\mathsf{P}_{1}^{(n)}}[1-\psi_{n}] \right) = 1 - \inf_{\psi_{n}} \left( \operatorname{E}_{\mathsf{P}_{1}^{(n)}}[\psi_{n}] - \operatorname{E}_{\mathsf{P}_{0}^{(n)}}[\psi_{n}] \right) = 1 - \frac{1}{2} \left\| \operatorname{P}_{1}^{(n)} - \operatorname{P}_{0}^{(n)} \right\|_{1}, \end{split}$$

where  $||Q - P||_1$  is the  $L_1$  distance between P and Q. Since this rewrites

$$\inf_{\theta_n \in \mathcal{S}^{p_n-1}} \mathbb{E}_{\mathsf{P}_{\theta_n,\kappa_n,f_{\mathrm{W}}}^{(n)}}[\phi_n] \le \mathbb{E}_{\mathsf{P}_0^{(n)}}[\phi_n] + \frac{1}{2} \|\mathsf{P}_1^{(n)} - \mathsf{P}_0^{(n)}\|_1$$

we have, as  $n \to \infty$ ,

$$\inf_{\theta_n \in \mathcal{S}^{p_n-1}} \mathbb{E}_{\mathsf{P}_{\theta_n,\kappa_n,f_{\mathsf{W}}}^{(n)}}[\phi_n] \le \mathbb{E}_{\mathsf{P}_0^{(n)}}[\phi_n] + \frac{1}{2} \left\| \mathsf{P}_1^{(n)} - \mathsf{P}_0^{(n)} \right\|_1 + o(1)$$

$$\leq \mathbf{E}_{\mathbf{P}_{0}^{(n)}}[\phi_{n}] + \frac{1}{2} \left( \mathbf{E}_{\mathbf{P}_{0}^{(n)}} \left[ \left( \frac{d\mathbf{P}_{1}^{(n)}}{d\mathbf{P}_{0}^{(n)}} - 1 \right)^{2} \right] \right)^{1/2} + o(1)$$
$$= \mathbf{E}_{\mathbf{P}_{0}^{(n)}}[\phi_{n}] + \frac{1}{2} \left( \mathbf{E}_{\mathbf{P}_{0}^{(n)}} \left[ \left( \frac{d\mathbf{P}_{1}^{(n)}}{d\mathbf{P}_{0}^{(n)}} \right)^{2} \right] - 1 \right)^{1/2} + o(1).$$

Therefore, it is sufficient to prove that

$$E_{P_0^{(n)}} \left[ \left( \frac{dP_1^{(n)}}{dP_0^{(n)}} \right)^2 \right] = 1 + o(1),$$
(36)

as  $n \to \infty$ . To do so, note that, from (3), the Watson distributions admit the density

$$\mathbf{x} \mapsto C_{p,\kappa} \exp\left(\kappa \left(\mathbf{x}'\boldsymbol{\theta}\right)^2\right) := \frac{c_{p,\kappa,f_W} \Gamma(\frac{p-1}{2})}{2\pi^{(p-1)/2}} \exp\left(\kappa \left(\mathbf{x}'\boldsymbol{\theta}\right)^2\right),$$

with

$$c_{p,\kappa,f_W} = 1/\int_{-1}^{1} (1-s^2)^{(p-3)/2} \exp(\kappa s^2) \, ds = \frac{\Gamma(\frac{p}{2})}{\sqrt{\pi} \Gamma(\frac{p-1}{2})_1 F_1(\frac{1}{2}, \frac{p}{2}, \kappa)}.$$

Since this yields

$$\frac{d\mathbf{P}_{1}^{(n)}}{d\mathbf{P}_{0}^{(n)}} = \frac{1}{2^{p_{n}}} \sum_{\mathbf{v}_{n} \in \mathcal{V}_{n}} \frac{d\mathbf{P}_{\mathbf{v}_{n}/\sqrt{p_{n},\kappa_{n}}}^{(n)}}{d\mathbf{P}_{0}^{(n)}} = \frac{C_{p_{n},\kappa_{n}}^{n}}{2^{p_{n}}C_{p_{n},0}^{n}} \sum_{\mathbf{v}_{n} \in \mathcal{V}_{n}} \exp\left(\kappa_{n} \sum_{i=1}^{n} (\mathbf{X}_{ni}'\mathbf{v}_{n})^{2}/p_{n}\right),$$

we have

$$\left(\frac{d\mathbf{P}_{1}^{(n)}}{d\mathbf{P}_{0}^{(n)}}\right)^{2} = \frac{C_{p_{n},\kappa_{n}}^{2n}}{4^{p_{n}}C_{p_{n},0}^{2n}}\sum_{\mathbf{v}_{n},\mathbf{w}_{n}\in\mathcal{V}_{n}}\exp\left(\kappa_{n}\sum_{i=1}^{n}(\mathbf{X}_{ni}'\left\{\left(\mathbf{v}_{n}\mathbf{v}_{n}'+\mathbf{w}_{n}\mathbf{w}_{n}'\right)/p_{n}\right\}\mathbf{X}_{ni}\right)\right),$$

which provides

$$\mathbf{E}_{\mathbf{P}_{0}^{(n)}}\left[\left(\frac{d\mathbf{P}_{1}^{(n)}}{d\mathbf{P}_{0}^{(n)}}\right)^{2}\right] = \frac{C_{p_{n},\kappa_{n}}^{2n}}{4^{p_{n}}C_{p_{n},0}^{n}}\sum_{\mathbf{v}_{n},\mathbf{w}_{n}\in\mathcal{V}_{n}}\left(\int_{\mathcal{S}^{p_{n}-1}}\exp(\kappa_{n}\mathbf{x}'\{(\mathbf{v}_{n}\mathbf{v}_{n}'+\mathbf{w}_{n}\mathbf{w}_{n}')/p_{n}\}\mathbf{x})\,d\sigma(\mathbf{x})\right)^{n}.$$

It is easy to check that the eigenvalues of  $(\mathbf{v}_n \mathbf{v}'_n + \mathbf{w}_n \mathbf{w}'_n)/p_n$  are  $\lambda_{1n} = 1 + (\mathbf{v}'_n \mathbf{w}_n)/p_n$ ,  $\lambda_{2n} = 1 - (\mathbf{v}'_n \mathbf{w}_n)/p_n$ , and  $\lambda_{3n} = \cdots = \lambda_{p_n n} = 0$ . Therefore, letting  $t_n := \kappa_n/p_n$ ,  $\mathbf{\Lambda}(z) = \mathbf{\Lambda}_n(z) := \text{diag}(1 + z/p_n, 1 - z/p_n, 0, \dots, 0)$  and  $\mathbf{B}_{\mathbf{v}} = \mathbf{B}_{n,\mathbf{v}} := \text{diag}((1 + \bar{\mathbf{v}})/2, (1 - \bar{\mathbf{v}})/2, 0, \dots, 0)$ , where  $\bar{\mathbf{v}}$  is the average of  $\mathbf{v}$ 's entries, we have

$$\begin{aligned} \mathbf{E}_{\mathbf{P}_{0}^{(n)}} \left[ \left( \frac{d\mathbf{P}_{1}^{(n)}}{d\mathbf{P}_{0}^{(n)}} \right)^{2} \right] &= \frac{C_{p_{n},\kappa_{n}}^{2n}}{4^{p_{n}}C_{p_{n},0}^{n}} \sum_{\mathbf{v}_{n}\in\mathcal{V}_{n}} \sum_{\mathbf{v}_{n}\in\mathcal{V}_{n}} \left( \int_{\mathcal{S}^{p_{n}-1}} \exp(\kappa_{n}\mathbf{x}'\mathbf{\Lambda}(\mathbf{v}_{n}'\mathbf{w}_{n})\mathbf{x}) \, d\sigma(\mathbf{x}) \right)^{n} \\ &= \frac{C_{p_{n},\kappa_{n}}^{2n}}{2^{p_{n}}C_{p_{n},0}^{n}} \sum_{\mathbf{v}_{n}\in\mathcal{V}_{n}} \left( \int_{\mathcal{S}^{p_{n}-1}} \exp(\kappa_{n}\mathbf{x}'\mathbf{\Lambda}(\mathbf{v}_{n}'\mathbf{1})\mathbf{x}) \, d\sigma(\mathbf{x}) \right)^{n} \\ &= \frac{C_{p_{n},\kappa_{n}}^{2n}}{2^{p_{n}}C_{p_{n},0}^{2n}} \sum_{\mathbf{v}_{n}\in\mathcal{V}_{n}} \left( C_{p_{n},0} \int_{\mathcal{S}^{p_{n}-1}} \exp(2t_{n}p_{n}\mathbf{x}'\mathbf{B}_{\mathbf{v}_{n}}\mathbf{x}) \, d\sigma(\mathbf{x}) \right)^{n} \\ &= \frac{C_{p_{n},\kappa_{n}}^{2n}}{2^{p_{n}}C_{p_{n},0}^{2n}} \sum_{\mathbf{v}_{n}\in\mathcal{V}_{n}} \left( 1F_{1}\left(\frac{1}{2},\frac{p_{n}}{2},2t_{n}p_{n}\mathbf{B}_{\mathbf{v}_{n}}\right) \right)^{n} \end{aligned} \tag{37}$$

$$=\frac{\mathrm{E}[({}_{1}F_{1}(\frac{1}{2},\frac{p_{n}}{2},2t_{n}p_{n}\mathbf{B}_{\mathbf{V}_{n}}))^{n}]}{({}_{1}F_{1}(\frac{1}{2},\frac{p_{n}}{2},t_{n}p_{n}))^{2n}},$$
(38)

where  $\mathbf{V}_n = (V_{n1}, \dots, V_{np_n})$  is uniformly distributed over  $\mathcal{V}_n$ , hence has mutually independent Rademacher marginals; here, the equality in (37) results from page 1208 of [34] or page 4 from [25], whereas the one in (38) follows from the fact that

$$C_{p_n,\kappa_n} = \frac{c_{p_n,\kappa_n} \Gamma(\frac{p_n-1}{2})}{2\pi^{(p_n-1)/2}} = \left(\frac{\Gamma(\frac{p_n}{2})}{\sqrt{\pi} \Gamma(\frac{p_n-1}{2})_1 F_1(\frac{1}{2},\frac{p_n}{2},\kappa_n)}\right) \frac{\Gamma(\frac{p_n-1}{2})}{2\pi^{(p_n-1)/2}} = \frac{\Gamma(\frac{p_n}{2})}{2\pi_1^{p_n/2} F_1(\frac{1}{2},\frac{p_n}{2},\kappa_n)}$$

yields  $C_{p_n,\kappa_n}/C_{p_n,0} = {}_1F_1(\frac{1}{2}, \frac{p_n}{2}, 0)/{}_1F_1(\frac{1}{2}, \frac{p_n}{2}, \kappa_n) = {}_1/{}_1F_1(\frac{1}{2}, \frac{p_n}{2}, t_n p_n).$ By applying Lemma C.2 to (38), we obtain that, for a sequence  $\eta_n$  that converges to one,

$$\mathbf{E}_{\mathbf{P}_{0}^{(n)}}\left[\left(\frac{d\mathbf{P}_{1}^{(n)}}{d\mathbf{P}_{0}^{(n)}}\right)^{2}\right] = \eta_{n} \mathbf{E}\left[\frac{\left((1-2t_{n})^{2}-4t_{n}^{2}\bar{\mathbf{V}}_{n}^{2}\right)^{-n/2}}{(1-2t_{n})^{-n}}\right] = \eta_{n} \mathbf{E}\left[\left(1-s_{n}\bar{\mathbf{V}}_{n}^{2}\right)^{-n/2}\right],$$

where we let  $s_n := 4t_n^2/(1-2t_n)^2$  and  $\bar{\mathbf{V}}_n := \frac{1}{p_n} \sum_{\ell=1}^{p_n} V_{n\ell}$ . Now, since  $0 \le (1-x)^{-1/(2x)} \le 2$  for any  $x \in [0, \frac{1}{2}]$  and  $s_n \le 1/2$  for *n* large enough, we then have, for *n* large enough,

$$E_{P_0^{(n)}} \left[ \left( \frac{dP_1^{(n)}}{dP_0^{(n)}} \right)^2 \right] = \eta_n E[\left( \left( 1 - s_n \bar{\mathbf{V}}_n^2 \right)^{-1/(2s_n \bar{\mathbf{V}}_n^2)} \right)^{ns_n \bar{\mathbf{V}}_n^2} \right] \le \eta_n E[2^{ns_n \bar{\mathbf{V}}_n^2}] = \eta_n E[\exp(ns_n \bar{\mathbf{V}}_n^2 \log 2)]$$
  
=  $\eta_n \int_0^\infty P[\exp(ns_n \bar{\mathbf{V}}_n^2 \log 2) \ge u] \, du = \eta_n \left( 1 + \int_1^\infty P\left[ \bar{\mathbf{V}}_n^2 \ge \frac{\log u}{(\log 2)ns_n} \right] \, du \right).$ 

We then use Hoeffding's inequality, which, applied to the mutually independent Rademacher random variables  $V_{n\ell}$ , provides  $P[\bar{\mathbf{V}}_n^2 \ge z] \le 2 \exp(-2p_n z)$  for any z > 0; see [10], page 2371. This yields

$$\mathsf{E}_{\mathsf{P}_{0}^{(n)}}\left[\left(\frac{d\mathsf{P}_{1}^{(n)}}{d\mathsf{P}_{0}^{(n)}}\right)^{2}\right] \leq \eta_{n}\left(1+2\int_{1}^{\infty}\exp\left(-\frac{2p_{n}\log u}{(\log 2)ns_{n}}\right)du\right) = \eta_{n}\left(1+2\int_{1}^{\infty}u^{-\frac{2p_{n}}{(\log 2)ns_{n}}}du\right)$$

for n large enough. Since the assumptions of Theorem 4.2 ensure that

$$\frac{p_n}{ns_n} = \frac{p_n(1-2t_n)^2}{4nt_n^2} = \frac{p_n(1+o(1))}{4nt_n^2} = \left(\frac{4n\kappa_n^2}{p_n^3(1+o(1))}\right)^{-1} \to \infty,$$

we have that, still for *n* large enough,

$$\mathsf{E}_{\mathsf{P}_{0}^{(n)}}\left[\left(\frac{d\mathsf{P}_{1}^{(n)}}{d\mathsf{P}_{0}^{(n)}}\right)^{2}\right] \leq \eta_{n}\left(1+2\left[\frac{u^{1-\frac{2p_{n}}{(\log 2)ns_{n}}}}{1-\frac{2p_{n}}{(\log 2)ns_{n}}}\right]_{1}^{\infty}\right) = \eta_{n}\left(1+\frac{2}{\frac{2p_{n}}{(\log 2)ns_{n}}-1}\right) \to 1.$$

Since

$$1 = \left( E_{P_0^{(n)}} \left[ \frac{dP_1^{(n)}}{dP_0^{(n)}} \right] \right)^2 \le E_{P_0^{(n)}} \left[ \left( \frac{dP_1^{(n)}}{dP_0^{(n)}} \right)^2 \right]$$

this establishes (36), which proves the result.

#### Appendix D: Proofs for Section 5

The following result is a higher-order extension of Lemma A.1 from [15] and is needed to prove Proposition 5.1.

**Lemma D.1.** Let  $g : \mathbb{R} \to \mathbb{R}$  be four times differentiable at 0. Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$  and  $(\kappa_n)$  be a real sequence that is  $o(\sqrt{p_n})$ . Then,

$$R_n(g) := c_{p_n} \int_{-1}^{1} (1 - s^2)^{(p_n - 3)/2} g(\kappa_n s) \, ds = g(0) + \frac{\kappa_n^2}{2p_n} g''(0) + \frac{\kappa_n^4}{8p_n^2} g^4(0) + o\left(\frac{\kappa_n^4}{p_n^2}\right)$$

as  $n \to \infty$ , still with  $c_p := 1/\int_{-1}^{1} (1-s^2)^{(p-3)/2} ds$ .

**Proof of Lemma D.1.** We proceed as in the proof of Lemma A.1. Denoting as  $g^{\ell}$  the  $\ell$ th derivative of g, (15) allows us to write

$$R_n(g) - g(0) - \frac{\kappa_n^2}{2p_n}g''(0) = c_{p_n} \int_{-1}^1 (1 - s^2)^{(p_n - 3)/2} \left\{ g(\kappa_n s) - \sum_{\ell=0}^3 \frac{1}{\ell!} \kappa_n^\ell s^\ell g^\ell(0) \right\} ds.$$

We may assume without any loss of generality that  $(\kappa_n)$  is a sequence in  $\mathbb{R}^+_0$ , so that the change of variables  $t = \kappa_n s$  yields

$$R_n(g) - g(0) - \frac{\kappa_n^2}{2p_n}g''(0) = \frac{3\kappa_n^4}{p_n(p_n+2)} \int_{-\infty}^{\infty} h_n(t) \left\{ \frac{g(t) - \sum_{\ell=0}^3 \frac{1}{\ell!} t^\ell g^\ell(0)}{t^4} \right\} dt,$$

or, equivalently,

$$\frac{R_n(g) - g(0) - \frac{\kappa_n^2}{2p_n}g''(0) - \frac{3\kappa_n^4}{24p_n(p_n+2)}g^4(0)}{\frac{3\kappa_n^4}{p_n(p_n+2)}} = \int_{-\infty}^{\infty} h_n(t) \left\{ \frac{g(t) - \sum_{\ell=0}^3 \frac{1}{\ell!}t^\ell g^\ell(0)}{t^4} \right\} dt - \frac{1}{24}g^4(0),$$
(39)

where  $h_n$  is defined through

$$t \mapsto h_n(t) = \frac{t^4 (1 - \frac{t^2}{\kappa_n^2})^{(p_n - 3)/2} \mathbb{I}[|t| \le \kappa_n]}{\int_{-\infty}^{\infty} t^4 (1 - \frac{t^2}{\kappa_n^2})^{(p_n - 3)/2} \mathbb{I}[|t| \le \kappa_n] dt}.$$

It can be checked that, since  $\kappa_n = o(\sqrt{p_n})$ , the sequence  $(h_n)$  is an *approximate*  $\delta$ -sequence, in the sense that  $\int_{-\infty}^{\infty} h_n(t) dt = 1$  for any n and  $\int_{-\varepsilon}^{\varepsilon} h_n(t) dt \to 1$  for any  $\varepsilon > 0$ . Hence,

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} h_n(t) \left\{ \frac{g(t) - \sum_{\ell=0}^3 \frac{1}{\ell!} t^\ell g^{\ell}(0)}{t^4} \right\} dt = \lim_{t \to 0} \frac{g(t) - \sum_{\ell=0}^3 \frac{1}{\ell!} t^\ell g^{\ell}(0)}{t^4}$$

which, by using three times L'Hôpital's rule, is equal to

$$\lim_{t \to 0} \frac{g^{\underline{3}}(t) - g^{\underline{3}}(0)}{24t} = \frac{1}{24}g^{\underline{4}}(0),$$

where the last equality results from differentiability of  $g^{\underline{3}}$  at zero. Thus, (39) yields

$$R_n(g) - g(0) - \frac{\kappa_n^2}{2p_n}g''(0) - \frac{3\kappa_n^4}{24p_n(p_n+2)}g^4(0) = o\left(\frac{\kappa_n^4}{p_n^2}\right),$$

which establishes the result.

**Proof of Proposition 5.1.** Let us first recall that if **X** admits the density in (12), then  $\mathbf{X}'\boldsymbol{\theta}$  admits the density  $z \mapsto \tilde{c}_{p,\kappa,f}(1-s^2)^{(p-3)/2} f(\kappa s) \mathbb{I}[z \in [-1,1]]$  (which explains the expression of the normalizing constant  $\tilde{c}_{p,\kappa,f}$  in (12)).

Since  $\kappa_n$  is assumed to be  $o(\sqrt{p_n})$  as  $n \to \infty$ , Lemma D.1 provides

$$e_{n2} = \frac{1}{\kappa_n^2} \left( \frac{c_{p_n}}{\tilde{c}_{p_n,\kappa_n,f}} \right)^{-1} c_{p_n} \int_{-1}^{1} (1 - s^2)^{(p_n - 3)/2} (\kappa_n s)^2 f(\kappa_n s) \, ds$$
  
$$= \frac{\frac{1}{\kappa_n^2} \left( \frac{\kappa_n^2}{p_n} + \frac{12\kappa_n^4}{8p_n^2} f''(0) + o(\frac{\kappa_n^4}{p_n^2}) \right)}{1 + \frac{\kappa_n^2}{2p_n} f''(0) + \frac{\kappa_n^4}{8p_n^2} f^{\underline{4}}(0) + o(\frac{\kappa_n^4}{p_n^2})} = \frac{\frac{1}{p_n} + \frac{3\kappa_n^2}{2p_n^2} f''(0) + o(\frac{\kappa_n^2}{p_n^2})}{1 + \frac{\kappa_n^2}{2p_n} f''(0) + \frac{\kappa_n^4}{8p_n^2} f^{\underline{4}}(0) + o(\frac{\kappa_n^2}{p_n^2})}$$
  
$$= \frac{1}{p_n} + \frac{\kappa_n^2}{p_n^2} f''(0) + o\left(\frac{\kappa_n^2}{p_n^2}\right),$$

which proves the result for  $e_{n2}$ . The same lemma also yields

$$e_{n4} = \frac{1}{\kappa_n^4} \left(\frac{c_{p_n}}{\tilde{c}_{p_n,\kappa_n,f}}\right)^{-1} c_{p_n} \int_{-1}^{1} (1 - s^2)^{(p_n - 3)/2} (\kappa_n s)^4 f(\kappa_n s) \, ds$$
  
$$= \frac{\frac{1}{\kappa_n^4} \left(\frac{24\kappa_n^4}{8p_n^2} + o\left(\frac{\kappa_n^4}{p_n^2}\right)\right)}{1 + \frac{\kappa_n^2}{2p_n} f''(0) + \frac{\kappa_n^4}{8p_n^2} f^{\frac{4}{2}}(0) + o\left(\frac{\kappa_n^4}{p_n^2}\right)} = \frac{\frac{3}{p_n^2} + o\left(\frac{1}{p_n^2}\right)}{1 + \frac{\kappa_n^2}{2p_n} f''(0) + \frac{\kappa_n^4}{8p_n^2} f^{\frac{4}{2}}(0) + o\left(\frac{\kappa_n^4}{p_n^2}\right)}$$
  
$$= \frac{3}{p_n^2} + o\left(\frac{1}{p_n^2}\right).$$

The claim for  $e_{n8}$  directly follows from the identity  $e_{n8} \le e_{n4}$ .

**Proof of Theorem 5.2.** First note that, in all cases (i)–(iii), we have  $\kappa_n = o(\sqrt{p_n})$ , so that Proposition 5.1 applies and ensures that conditions (a)–(b) in Theorem 3.2 are fulfilled. Let us then treat cases (i)–(iii) separately. (i) Since  $\kappa_n =$  $o(p_n^{3/4}/n^{1/4})$ , Proposition 5.1 implies that

$$g_{n2} = e_{n2} - \frac{1}{p_n} = O\left(\frac{\kappa_n^2}{p_n^2}\right) = o\left(\frac{1}{\sqrt{np_n}}\right),$$

so that Theorem 3.2(i) shows that  $Q_n^{\text{St}}$  is asymptotically standard normal. (ii) Since  $n^{1/4}\kappa_n/p_n^{3/4} \to \tau \neq 0$  and  $p_n = o(n)$ , Proposition 5.1 provides

$$\sqrt{np_n}g_{n2} = \frac{\sqrt{n\kappa_n^2}}{p_n^{3/2}}f''(0) + o\left(\frac{\sqrt{n\kappa_n^2}}{p_n^{3/2}}\right) = \tau^2 f''(0) + o(1),$$

Theorem 3.2(ii) shows that  $Q_n^{\text{St}} \xrightarrow{\mathcal{D}} \mathcal{N}(\xi^2/2, 1)$ , with  $\xi = \tau^2 f''(0)$ , which establishes the result. (iii) The claim follows from Theorem 3.2(iii) since

$$\sqrt{np_n}|g_{n2}| = \frac{2\sqrt{n\kappa_n}}{p_n^{3/2}} + o\left(\frac{\sqrt{n\kappa_n}}{p_n^{3/2}}\right)$$
  
diverges to infinity.

# Appendix E: Consistency in the FvML case

In this final appendix, we show that the constraint  $\kappa_n = o(\sqrt{p_n})$  in Theorem 5.2(iii) is superfluous in the FvML case, which validates the concentration scheme (iii) in the simulation exercise we conducted in Section 5 (recall that the condition  $\kappa_n = o(\sqrt{p_n})$  is not met in this concentration scheme).

**Proposition E.1.** Let  $(p_n)$  be a sequence of positive integers diverging to  $\infty$  and  $(\theta_n)$  be a sequence such that  $\theta_n \in S^{p_n-1}$ for any n. Fix  $f(z) = \exp(z)$ . Assume that the real non-negative sequence  $(\kappa_n)$  satisfies  $n^{1/4}\kappa_n/p_n^{3/4} \to \infty$ . Then, (i)  $\sqrt{np_n}g_{n2} \to \infty$  as  $n \to \infty$  and (ii)  $e_{n4} = o(ng_{n2}^2)$  as  $n \to \infty$ , so that (iii) for any real number  $M, \tilde{P}_{\theta_n,\kappa_n,f}^{(n)}[Q_n^{\text{St}} > M] \to 1$ as  $n \to \infty$ .

Proof of Proposition E.1. (i) As seen from (2)-(3) in [40] or from Lemma S.2.1 in [16],

$$e_{n1} = \frac{\mathcal{I}_{p_n/2}(\kappa_n)}{\mathcal{I}_{(p_n/2)-1}(\kappa_n)}, \qquad e_{n2} = -\frac{p_n - 1}{\kappa_n} e_{n1} + 1$$
(40)

and

$$e_{n4} = -\frac{(p_n - 1)(p_n(p_n + 1) + 2\kappa_n^2)}{\kappa_n^3} e_{n1} + \frac{(p_n - 1)(p_n + 1) + \kappa_n^2}{\kappa_n^2},$$
(41)

where  $\mathcal{I}_{\nu}(\cdot)$  stands for the order- $\nu$  modified Bessel function of the first kind. It follows from (11) in [1] that

$$\frac{\kappa_n}{\frac{p_n}{2} + \sqrt{\frac{p_n^2}{4} + \kappa_n^2}} \le e_{n1} \le \frac{\kappa_n}{\frac{p_n}{2} - 1 + \sqrt{(\frac{p_n}{2} + 1)^2 + \kappa_n^2}}$$
(42)

Now, note that (40) provides (for  $p_n \ge 2$ )

$$g_{n2} = e_{n2} - \frac{1}{p_n} = 1 - \frac{1}{p_n} - \frac{p_n - 1}{\kappa_n} e_{n1} = \frac{p_n - 1}{p_n} \left( 1 - \frac{e_{n1}}{\kappa_n / p_n} \right) \ge \frac{1}{2} \left( 1 - \frac{e_{n1}}{\kappa_n / p_n} \right).$$

Since (42) implies in particular that

$$1 - \frac{e_{n1}}{\kappa_n/p_n} \ge 1 - \frac{1}{\frac{1}{2} - \frac{1}{p_n} + \sqrt{(\frac{1}{2} + \frac{1}{p_n})^2 + \frac{\kappa_n^2}{p_n^2}}} = \frac{\kappa_n^2/p_n^2}{\frac{1}{2} + \frac{1}{p_n} + \frac{\kappa_n^2}{p_n^2} + \sqrt{(\frac{1}{2} + \frac{1}{p_n})^2 + \frac{\kappa_n^2}{p_n^2}},$$

we thus have

$$2\sqrt{np_n}g_{n2} \ge \frac{\sqrt{n\kappa_n^2/p_n^{3/2}}}{\frac{1}{2} + \frac{1}{p_n} + \frac{\kappa_n^2}{p_n^2} + \sqrt{(\frac{1}{2} + \frac{1}{p_n})^2 + \frac{\kappa_n^2}{p_n^2}}}$$

Therefore, by using the identity  $\sqrt{a^2 + b^2} \le |a| + |b|$ , we obtain that (still for  $p_n \ge 2$ )

$$2\sqrt{np_n}g_{n2} \ge \frac{\sqrt{n\kappa_n^2/p_n^{3/2}}}{1 + \frac{2}{p_n} + \frac{\kappa_n^2}{p_n^2} + \frac{\kappa_n}{p_n}} \ge \frac{\sqrt{n\kappa_n^2/p_n^{3/2}}}{2(1 + \frac{\kappa_n}{p_n})^2} \ge \frac{\sqrt{n\kappa_n^2/p_n^{3/2}}}{2(2\max(1, \frac{\kappa_n}{p_n}))^2}$$
$$= \frac{\sqrt{n\kappa_n^2}}{8p_n^{3/2}}\min\left(1, \frac{p_n^2}{\kappa_n^2}\right) = \frac{1}{8}\min\left(\left(\frac{n^{1/4}\kappa_n}{p_n^{3/4}}\right)^2, \sqrt{np_n}\right),$$
(43)

which diverges to infinity by assumption. Part (i) the result follows. (ii) From (43), we have

$$\frac{e_{n4}}{ng_{n2}^2} = \frac{4p_n e_{n4}}{(2\sqrt{np_n}g_{n2})^2} \le 256p_n e_{n4} \max\left(\frac{p_n^3}{n\kappa_n^4}, \frac{1}{np_n}\right) \le 256 \max\left(\frac{p_n^4 e_{n4}}{n\kappa_n^4}, \frac{1}{n}\right).$$
(44)

Now, by using (41) and (42), we obtain

$$\begin{split} e_{n4} &\leq -\frac{(p_n-1)(p_n(p_n+1)+2\kappa_n^2)}{\kappa_n^3} \times \frac{\kappa_n/p_n}{\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{\kappa_n^2}{p_n^2}}} + \frac{(p_n-1)(p_n+1)+\kappa_n^2}{\kappa_n^2} \\ &= 1+\frac{p_n^2-1}{\kappa_n^2} - \frac{p_n^2-1}{\kappa_n^2(\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{\kappa_n^2}{p_n^2}})} - \frac{2(p_n-1)}{p_n(\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{\kappa_n^2}{p_n^2}})} \\ &= 1+\frac{p_n^2-1}{\kappa_n^2} \left(1-\frac{1}{\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{\kappa_n^2}{p_n^2}}}\right) - \frac{2}{\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{\kappa_n^2}{p_n^2}}} + \frac{2}{p_n(\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{\kappa_n^2}{p_n^2}})} \end{split}$$

$$=1+\frac{p_n^2-1}{p_n^2(\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{\kappa_n^2}{p_n^2}})^2}-\frac{2}{\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{\kappa_n^2}{p_n^2}}}+\frac{2}{p_n(\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{\kappa_n^2}{p_n^2}})},$$

which entails

$$e_{n4} \le 1 + \frac{1}{(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\kappa_n^2}{p_n^2}})^2} - \frac{2}{\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\kappa_n^2}{p_n^2}}} + \frac{2}{p_n}$$
$$= \left(1 - \frac{1}{\frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\kappa_n^2}{p_n^2}}}}\right)^2 + \frac{2}{p_n} = \frac{\kappa_n^4/p_n^4}{(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\kappa_n^2}{p_n^2}})^4} + \frac{2}{p_n} \le \frac{\kappa_n^4}{p_n^4} + \frac{2}{p_n}$$

Therefore,

$$\frac{e_{n4}p_n^4}{n\kappa_n^4} \le \frac{1}{n} + \frac{2p_n^3}{n\kappa_n^4} = o(1)$$

by assumption, so that Part (ii) of the result follows from (44). (iii) In view of Parts (i)–(ii) of the result, the claim directly follows from Theorem 3.2(iii).

# Acknowledgements

The authors would like to thank the anonymous referees, an Associate Editor and the Editor for their constructive comments that improved the quality of this paper. Davy Paindaveine's research is supported by a research fellowship from the Francqui Foundation and by the Program of Concerted Research Actions (ARC) of the Université libre de Bruxelles. Thomas Verdebout's research is supported by the ARC Program of the Université libre de Bruxelles and by the Crédit de Recherche J.0134.18 of the FNRS (Fonds National pour la Recherche Scientifique), Communauté Française de Belgique.

#### References

- D. E. Amos. Computation of modified Bessel functions and their ratios. Math. Comp. 28 (1974) 239–251. MR0333287 https://doi.org/10.2307/ 2005830
- [2] T. W. Anderson and M. A. Stephens. Tests for randomness of directions against equatorial and bimodal alternatives. *Biometrika* 43 (1972) 613–621. MR0334367 https://doi.org/10.1093/biomet/59.3.613
- [3] A. Banerjee, I. Dhillon, J. Ghosh and S. Sra. Clustering on the unit hypersphere using von Mises–Fisher distributions. J. Mach. Learn. Res. 6 (2005) 1345–1382. MR2249858
- [4] A. Banerjee and J. Ghosh. Frequency sensitive competitive learning for scalable balanced clustering on high-dimensional hyperspheres. *IEEE Trans. Neural Netw.* 15 (2004) 702–719.
- [5] B. B. Bhattacharya. Asymptotic distribution and detection thresholds for two-sample tests based on geometric graphs. Ann. Statist. 48 (2020) 2879–2903. MR4152627 https://doi.org/10.1214/19-AOS1913
- [6] A. S. Bijral, M. Breitenbach and G. Grudic. Mixture of Watson distributions: A generative model for hyperspherical embeddings. In Artificial Intelligence and Statistics 35–42, 2007.
- [7] P. Billingsley. Probability and Measure, 3rd edition. Wiley, New York, Chichester, 1995.
- [8] C. Bingham. An antipodally symmetric distribution on the sphere. Ann. Statist. 2 (1974) 1201–1225.
- [9] T. Cai, J. Fan and T. Jiang. Distributions of angles in random packing on spheres. J. Mach. Learn. Res. 14 (2013) 1837–1864.
- [10] T. T. Cai and Z. Ma. Optimal hypothesis testing for high dimensional covariance matrices. Bernoulli 19 (2013) 2359–2388.
- [11] Y. Chikuse. High dimensional limit theorems and matrix decompositions on the Stiefel manifold. J. Multivariate Anal. 36 (1991) 145–162. MR1096663 https://doi.org/10.1016/0047-259X(91)90054-6
- [12] Y. Chikuse. High dimensional asymptotic expansions for the matrix Langevin distributions on the Stiefel manifold. J. Multivariate Anal. 44 (1993) 82–101. MR1208471 https://doi.org/10.1006/jmva.1993.1005
- [13] Y. Chikuse. Statistics on Special Manifolds. Lecture Notes in Statistics 174. Springer, New York, 2003. MR1960435 https://doi.org/10.1007/ 978-0-387-21540-2
- [14] J. A. Cuesta-Albertos, A. Cuevas and R. Fraiman. On projection-based tests for directional and compositional data. *Stat. Comput.* **19** (2009) 367–380.
- [15] C. Cutting, D. Paindaveine and T. Verdebout. Testing uniformity on high-dimensional spheres against monotone rotationally symmetric alternatives. Ann. Statist. 45 (2017) 1024–1058. MR3662447 https://doi.org/10.1214/16-AOS1473
- [16] C. Cutting, D. Paindaveine and T. Verdebout. Supplement to "Testing uniformity on high-dimensional spheres against monotone rotationally symmetric alternatives". Ann. Statist. 45 (2017) 1024–1058. MR3662447 https://doi.org/10.1214/16-AOS1473

- [17] C. Cutting, D. Paindaveine and T. Verdebout. On the power of axial tests of uniformity. *Electron. J. Stat.* 14 (2020) 2123–2154. MR4097051 https://doi.org/10.1214/20-EJS1716
- [18] I. L. Dryden. Statistical analysis on high-dimensional spheres and shape spaces. Ann. Statist. 33 (2005) 1643–1665. MR2166558 https://doi.org/10. 1214/009053605000000264
- [19] B. Ebner, N. Henze and J. E. Yukich. Multivariate goodness-of-fit on flat and curved spaces via nearest neighbor distances. J. Multivariate Anal. 165 (2018) 231–242. MR3768763 https://doi.org/10.1016/j.jmva.2017.12.009
- [20] N. I. Fisher, T. Lewis and B. J. Embleton. Statistical Analysis of Spherical Data. Cambridge Univ. Press press, Cambridge, 1987.
- [21] Y. Fujikoshi. Asymptotic expansions of the distributions of test statistics in multivariate analysis. J. Sci. Hiroshima Univ. Ser. A-1 34 (1970) 73–144.
- [22] E. García-Portugués, D. Paindaveine and T. Verdebout. On optimal tests for rotational symmetry against new classes of hyperspherical distributions. J. Amer. Statist. Assoc. 115 (2020) 1873–1887. MR4189764 https://doi.org/10.1080/01621459.2019.1665527
- [23] E. García-Portugués and T. Verdebout. An overview of uniformity tests on the hypersphere. ArXiv preprint. Available at arXiv:1804.00286 (2020).
- [24] M. Hallin and D. Paindaveine. Semiparametrically efficient rank-based inference for shape. I. Optimal rank-based tests for sphericity. Ann. Statist. 34 (2006) 2707–2756.
- [25] G. Hillier. The density of a quadratic form in a vector uniformly distributed on the n-sphere. Econometric Theory 17 (2001) 1–28.
- [26] S. R. Jammalamadaka, S. Meintanis and T. Verdebout. On new Sobolev tests of uniformity on the circle with extension to the sphere. *Bernoulli* 26 (2020) 2226–2252. MR4091107 https://doi.org/10.3150/19-BEJ1191
- [27] J. Juan and F. J. Prieto. Using angles to identify concentrated multivariate outliers. *Technometrics* 43 (2001) 311–322. MR1943186 https://doi.org/10.1198/004017001316975907
- [28] P. E. Jupp. Modifications of the Rayleigh and Bingham tests for uniformity of directions. J. Multivariate Anal. 77 (1) (2001) 1–20. MR1838711 https://doi.org/10.1006/jmva.2000.1922
- [29] P. E. Jupp. Data-driven Sobolev tests of uniformity on compact Riemannian manifolds. Ann. Statist. 36 (3) (2008) 1246–1260. MR2418656 https://doi.org/10.1214/009053607000000541
- [30] C. Lacour and T. M. Pham Ngoc. Goodness-of-fit test for noisy directional data. *Bernoulli* 20 (4) (2014) 2131–2168. MR3263101 https://doi.org/10.3150/13-BEJ553
- [31] C. Ley and T. Verdebout. Modern Directional Statistics. Chapman and Hall/CRC, Boca Raton, 2017. MR4011950 https://doi.org/10.1080/ 01621459.2019.1665527
- [32] K. V. Mardia and P. E. Jupp. Directional Statistics. John Wiley & Sons, Chichester, 2000.
- [33] R. J. Muirhead. Aspects of Multivariate Statistical Theory. John Wiley & Sons, New York, 2005. MR0652932 https://doi.org/10.1080/01621459. 2019.1665527
- [34] A. Onatski, M. J. Moreira and M. Hallin. Asymptotic power of sphericity tests for high-dimensional data. Ann. Statist. 41 (2013) 1204–1231. MR3113808 https://doi.org/10.1214/13-AOS1100
- [35] D. Paindaveine and T. Verdebout. On high-dimensional sign tests. Bernoulli 22 (2016) 1745–1769.
- [36] D. Paindaveine and T. Verdebout. Inference on the mode of weak directional signals: A Le Cam perspective on hypothesis testing near singularities. Ann. Statist. 45 (2017) 800–832. MR3650401 https://doi.org/10.1214/16-AOS1468
- [37] D. Paindaveine and T. Verdebout. Detecting the direction of a signal on high-dimensional spheres: Non-null and Le Cam optimality results. Probab. Theory Related Fields 176 (2020) 1165–1216. MR4087491 https://doi.org/10.1007/s00440-019-00937-x
- [38] D. Paindaveine and T. Verdebout. Inference for spherical location under high concentration. Ann. Statist. 48 (2020) 2982–2998. MR4152631 https://doi.org/10.1214/19-AOS1918
- [39] L. Rayleigh. On the problem of random vibrations and random flights in one, two and three dimensions. *Philos. Mag.* 37 (1919) 321–346.
- [40] G. Schou. Estimation of the concentration parameter in von Mises-Fisher distributions. *Biometrika* 65 (2) (1978) 369–377. MR0513935 https://doi.org/10.1093/biomet/65.2.369
- [41] S. Sra and D. Karp. The multivariate Watson distribution: Maximum-likelihood estimation and other aspects. J. Multivariate Anal. 114 (2013) 256–269. MR2993885 https://doi.org/10.1016/j.jmva.2012.08.010
- [42] S. Z. Sun and R. A. Lockhart. Bayesian optimality for Beran's class of tests of uniformity around the circle. J. Statist. Plann. Inference 198 (2019) 79–90. MR3809397 https://doi.org/10.1016/j.jspi.2018.03.006
- [43] D. E. Tyler. Statistical analysis for the angular central Gaussian distribution on the sphere. Biometrika 74 (3) (1987) 579–589.
- [44] G. S. Watson. Equatorial distributions on a sphere. Biometrika 52 (1/2) (1965) 193–201. MR0207115 https://doi.org/10.1093/biomet/52.1-2.193