Noname manuscript No.
(will be inserted by the editor)

Detecting the direction of a signal on high-dimensional spheres

Non-null and Le Cam optimality results

Davy Paindaveine · Thomas Verdebout

Received: date / Accepted: date

Abstract We consider one of the most important problems in directional statistics, namely the problem of testing the null hypothesis that the spike direction $\boldsymbol{\theta}$ of a Fisher–von Mises–Langevin distribution on the p-dimensional unit hypersphere is equal to a given direction $\boldsymbol{\theta}_0$. After a reduction through invariance arguments, we derive local asymptotic normality (LAN) results in a general high-dimensional framework where the dimension p_n goes to infinity at an arbitrary rate with the sample size n, and where the concentration κ_n behaves in a completely free way with n, which offers a spectrum of problems ranging from arbitrarily easy to arbitrarily challenging ones. We identify various asymptotic regimes, depending on the convergence/divergence properties of (κ_n), that yield different contiguity rates and different limiting experiments. In each regime, we derive Le Cam optimal tests under specified κ_n and we compute, from the Le Cam third lemma, asymptotic powers of the classical Watson test under contiguous alternatives. We further establish LAN results with respect to both spike direction and concentration, which allows us to discuss optimality also under unspecified κ_n. To investigate the non-null behavior of the Watson test outside the parametric framework above, we derive its local

Research is supported by the Program of Concerted Research Actions (ARC) of the Université libre de Bruxelles, by a research fellowship from the Francqui Foundation, and by the Crédit de Recherche J.0134.18 of the FNRS (Fonds National pour la Recherche Scientifique), Communauté Française de Belgique.

Davy Paindaveine
Université libre de Bruxelles, ECARES and Département de Mathématique, Avenue F.D. Roosevelt, 50, ECARES, CP114/04, B-1050, Brussels, Belgium
Tel.: +3226503845
E-mail: dpaindav@ulb.ac.be

Thomas Verdebout
Université libre de Bruxelles, ECARES F.D. Département de Mathématique, Avenue F.D. Roosevelt, 50, ECARES, CP114/04, B-1050, Brussels, Belgium
Tel.: +3226505892
E-mail: tverdebo@ulb.ac.be

asymptotic powers through martingale CLTs in the broader, semiparametric, model of rotationally symmetric distributions. A Monte Carlo study shows that the finite-sample behaviors of the various tests remarkably agree with our asymptotic results.

Keywords High-dimensional statistics · invariance · Le Cam's asymptotic theory of statistical experiments · local asymptotic normality · rotationally symmetric distributions

Mathematics Subject Classification (2010) 62H11 · 62F05 · 62G10

1 Introduction

In directional statistics, the sample space is the unit sphere $\mathcal{S}^{p-1} = \{\mathbf{x} \in \mathbb{R}^p : \|\mathbf{x}\|^2 = \mathbf{x}'\mathbf{x} = 1\}$ in $\mathbb{R}^p$. By far the most classical distributions on $\mathcal{S}^{p-1}$ are the Fisher–von Mises–Langevin (FvML) ones; see, e.g., [27] or [28]. We say that the random vector $\mathbf{X}$ with values in $\mathcal{S}^{p-1}$ has an $\mathrm{FvML}_p(\boldsymbol{\theta}, \kappa)$ distribution, with $\boldsymbol{\theta} \in \mathcal{S}^{p-1}$ and $\kappa \in (0, \infty)$, if it admits the density (throughout, densities on the unit sphere are with respect to the surface area measure)

$$\mathbf{x} \mapsto \frac{c_{p,\kappa}}{\omega_{p-1}} \exp(\kappa \mathbf{x}'\boldsymbol{\theta}), \tag{1.1}$$

where, denoting as $\Gamma(\cdot)$ the Euler Gamma function and as $\mathcal{I}_\nu(\cdot)$ the order-ν modified Bessel function of the first kind, $\omega_p := (2\pi^{p/2})/\Gamma\big(\frac{p}{2}\big)$ is the surface area of $\mathcal{S}^{p-1}$ and

$$c_{p,\kappa} := 1 \Big/ \int_{-1}^{1} (1-t^2)^{(p-3)/2} \exp(\kappa t)\, dt = \frac{(\kappa/2)^{(p/2)-1}}{\sqrt{\pi}\,\Gamma\big(\frac{p-1}{2}\big)\mathcal{I}_{\frac{p}{2}-1}(\kappa)}.$$

Clearly, $\boldsymbol{\theta}$ is a location parameter ($\boldsymbol{\theta}$ is the modal location on the sphere), that identifies the spike direction of the hyperspherical signal. In contrast, κ is a scale or *concentration* parameter: the larger κ, the more concentrated the distribution is about the modal location $\boldsymbol{\theta}$. As κ converges to zero, $c_{p,\kappa}$ converges to $c_p := \Gamma\big(\frac{p}{2}\big)/(\sqrt{\pi}\,\Gamma\big(\frac{p-1}{2}\big))$ and the density in (1.1) converges to the density $\mathbf{x} \mapsto 1/\omega_p$ of the uniform distribution over $\mathcal{S}^{p-1}$. The other extreme case, obtained for arbitrarily large values of κ, provides distributions that converge to a point mass in $\boldsymbol{\theta}$. Of course, it is expected that the larger κ, the easier it is to conduct inference on $\boldsymbol{\theta}$ — that is, the more powerful the tests on $\boldsymbol{\theta}$ and the smaller the corresponding confidence zones.

In this paper, we consider inference on $\boldsymbol{\theta}$ and focus on the generic testing problem for which the null hypothesis $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$, for a fixed $\boldsymbol{\theta}_0 \in \mathcal{S}^{p-1}$, is to be tested against $\mathcal{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ on the basis of a random sample $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$ from the $\mathrm{FvML}_p(\boldsymbol{\theta}, \kappa)$ distribution — the triangular array notation anticipates non-standard setups where p (hence, also $\boldsymbol{\theta}$) and/or κ will depend on n. Inference problems on $\boldsymbol{\theta}$ in the low-dimensional case have been considered among others in [8], [9], [17], [19], [22], [26], [31] and [37]. The related spherical regression problem has been tackled in [14] and [33], while testing for location on axial frames has been considered in [2].

Letting $\bar{\mathbf{X}}_n := \frac{1}{n}\sum_{i=1}^n \mathbf{X}_{ni}$, the most classical test for the testing problem above is the Watson [37] test rejecting the null at asymptotic level α whenever

$$W_n := \frac{n(p-1)\bar{\mathbf{X}}_n'(\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0')\bar{\mathbf{X}}_n}{1-\frac{1}{n}\sum_{i=1}^n(\mathbf{X}_{ni}'\boldsymbol{\theta}_0)^2} > \chi^2_{p-1,1-\alpha}, \tag{1.2}$$

where $\mathbf{I}_\ell$ stands for the ℓ-dimensional identity matrix and $\chi^2_{\ell,1-\alpha}$ denotes the α-upper quantile of the chi-square distribution with ℓ degrees of freedom. In the classical setup where p and κ are fixed, the asymptotic properties of the Watson test are well-known, both under the null and under local alternatives; see, e.g., [28] or [37]. Optimality properties in the Le Cam sense have been studied in [30]. In the non-standard setup where $\kappa = \kappa_n$ converges to zero, [31] investigated the asymptotic null and non-null behaviors of the Watson test. Interestingly, irrespective of the rate at which κ_n converges to zero (that is, irrespective of how fast the inference problem becomes more challenging as a function of n), the Watson test keeps meeting the asymptotic nominal level constraint and maintains strong optimality properties; see [31] for details. In the other non-standard, high-concentration, setup where κ_n diverges to infinity, [32] showed the Watson test also enjoys strong optimality properties.

For a fixed dimension p, this essentially settles the investigation of the properties of the Watson test and the study of the corresponding hypothesis testing problem. Nowadays, however, increasingly many applications lead to considering high-dimensional directional data: tests of uniformity on high-dimensional spheres have been studied in [7], [10], [11] and [12], while high-dimensional FvML distributions (or mixtures of high-dimensional FvML distributions) have been considered in magnetic resonance, gene-expression, and text mining; see, among others, [3], [4] and [15]. This motivates considering the high-dimensional spherical location problem, based on a random sample $\mathbf{X}_{n1},\ldots,\mathbf{X}_{nn}$ from the $\mathrm{FvML}_{p_n}(\boldsymbol{\theta}_n,\kappa_n)$ distribution, with (p_n) diverging to infinity (the dimension of $\boldsymbol{\theta}_n$ then depends on n, which justifies the notation). In this context, it was proved in [25] that the Watson test is robust to high-dimensionality in the sense that, as p_n goes to infinity with n, this test still has asymptotic size α under $\mathcal{H}_0^{(n)} : \boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0}$. This does not require any condition on the concentration sequence (κ_n) nor on the rate at which p_n goes to infinity, hence covers arbitrarily easy problems (κ_n large) and arbitrarily challenging ones (κ_n small), as well as moderately high dimensions and ultra-high dimensions. On its own, however, this null robustness result is obviously far from sufficient to motivate using the Watson test in high dimensions, as it might very well be that robustness under the null is obtained at the expense of power (in the extreme case, the Watson test, in high dimensions, might actually asymptotically behave like the trivial α-level test that randomly rejects the null with probability α).

These considerations raise many interesting questions, among which: are there alternatives under which the Watson test is consistent in high dimensions? What are the less severe alternatives (if any) under which the Watson test exhibits non-trivial asymptotic powers? Is the Watson test rate-optimal or, on the contrary, are there tests that show asymptotic powers under less severe alternatives than those detected by the Watson test? Does the Watson test enjoy optimality properties in high dimensions? As we will show, answering these questions will require considering several regimes fixing how the concentration κ_n behaves as a function of the dimension p_n and sample size n. Our results, that will crucially depend on the regime considered, are extensive in the sense that they answer the questions above in all possible regimes.

Our results will rely on two different approaches. (a) The first approach is based on Le Cam's asymptotic theory of statistical experiments. While this theory is very general, it does not directly apply in the present context since the high-dimensional spherical location problem involves a parametric space, namely $\{(\boldsymbol{\theta}_n, \kappa) : \boldsymbol{\theta}_n \in \mathcal{S}^{p_n-1}, \kappa \in (0, \infty)\}$, that depends on n (through p_n). We solve this by exploiting the invariance properties of the testing problem considered. In the image of the model by the corresponding maximal invariant, indeed, the parametric space does not depend on n anymore, which opens the door to studying the problem through the Le Cam approach. We derive stochastic second-order expansions of the resulting log-likelihood ratios, which is the main technical ingredient to establish the local asymptotic normality (LAN) of the invariant model. The LAN property takes different forms and involves different contiguity rates depending on the regime that is considered. In each regime, we determine the Le Cam optimal test for the problem considered and apply the Le Cam third lemma to obtain the asymptotic powers of this test and of the Watson test. This allows us to determine the regime(s) in which the Watson test is Le Cam optimal, or only rate-optimal, or not even rate-optimal. While this is first done under specified concentration κ_n, we further provide LAN results with respect to both location and concentration to be able to discuss optimality under unspecified κ_n. (b) While our investigation in (a) will fully characterize the asymptotic optimality properties of the Watson test in the FvML case, it will not provide any insight on the non-null behavior of this test outside this stringent parametric framework. This motivates complementing our investigation by a second approach, based on martingale CLTs. We will consider a broad, semiparametric, model, namely the class of rotationally symmetric distributions, and will identify the alternatives (if any) under which the Watson test will show non-trivial asymptotic powers in high dimensions. Again, this requires considering various regimes according to the concentration pattern.

The outline of the paper is as follows. In Section 2, we consider the high-dimensional version of the FvML spherical location problem. In Section 2.1, we describe the invariance approach that allows us to later rely on Le Cam's asymptotic theory of statistical experiments. In Section 2.2, we provide a stochastic second-order expansion of the resulting invariant log-likelihood ratios and prove, in various regimes that we identify, that these invariant models are locally asymptotically normal. This allows us to derive the corresponding Le Cam optimal tests for the specified concentration problem and to study the non-null asymptotic behavior of the Watson test in the light of these results.

In Section 2.3, we tackle the unspecified concentration problem through the derivation of LAN results that are with respect to both location and concentration. In Section 3, we conduct a systematic investigation of the non-null asymptotic properties of the Watson test in the broader context of rotationally symmetric distributions. This in particular confirms the FvML results obtained in Section 2. In Section 4, we conduct a Monte Carlo study to investigate how well the finite-sample behaviors of the various tests reflect our theoretical asymptotic results. In Section 5, we summarize the results obtained in the paper and shortly discuss research perspectives. Finally, an appendix contains all proofs.

2 Invariance and Le Cam optimality

As already mentioned in the introduction, the high-dimensional spherical location problem requires considering triangular arrays of observations of the form $\mathbf{X}_{ni}$, $i = 1, \ldots, n$, $n = 1, 2, \ldots$ For any sequence $(\boldsymbol{\theta}_n)$ such that $\boldsymbol{\theta}_n$ belongs to $\mathcal{S}^{p_n-1}$ for any n and any sequence (κ_n) in $(0, \infty)$, we denote as $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n}$ the hypothesis under which $\mathbf{X}_{ni}$, $i = 1, \ldots, n$, form a random sample from the $\mathrm{FvML}_{p_n}(\boldsymbol{\theta}_n, \kappa_n)$ distribution. The resulting sequence of statistical models is then associated with

$$\mathcal{P}^{(n)} = \left\{ \mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa} : (\boldsymbol{\theta}_n, \kappa) \in \boldsymbol{\Theta}_n := \mathcal{S}^{p_n-1} \times (0, \infty) \right\} \tag{2.3}$$

(the index in the parameter $\boldsymbol{\theta}_n$ in principle is superfluous but is used here to stress the dependence of this parameter on p_n, hence on n). The spherical location problem consists in testing the null hypothesis $\mathcal{H}_0^{(n)} : \boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0}$ against the alternative $\mathcal{H}_0^{(n)} : \boldsymbol{\theta}_n \neq \boldsymbol{\theta}_{n0}$, where $(\boldsymbol{\theta}_{0n})$ is a fixed sequence such that $\boldsymbol{\theta}_{0n}$ belongs to $\mathcal{S}^{p_n-1}$ for any n. Clearly, $\boldsymbol{\theta}_n$ is the parameter of interest, whereas κ_n plays the role of a nuisance. Our main objective in this section is to derive Le Cam optimality results for this problem, referring to sequences of local alternatives of the form $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n}$, with $\boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0} + \nu_n \boldsymbol{\tau}_n$, where the sequence (ν_n) and the bounded sequence $(\boldsymbol{\tau}_n)$, respectively in $(0, \infty)$ and $\mathbb{R}^{p_n}$, are such that $\boldsymbol{\theta}_n \in \mathcal{S}^{p_n-1}$ for any n, which imposes that

$$\boldsymbol{\theta}'_{n0} \boldsymbol{\tau}_n = -\frac{1}{2} \nu_n \|\boldsymbol{\tau}_n\|^2 \tag{2.4}$$

for any n; throughout, "the sequence $(\boldsymbol{\tau}_n)$ in $\mathbb{R}^{p_n}$ is bounded" means that $\boldsymbol{\tau}_n \in \mathbb{R}^{p_n}$ for any n and that $\|\boldsymbol{\tau}_n\| = O(1)$ as $n \to \infty$. The obvious lack of identifiability of ν_n and $\boldsymbol{\tau}_n$ will be no problem in the sequel (only the locally perturbed parameter values $\boldsymbol{\theta}_n = \boldsymbol{\theta}_0 + \nu_n \boldsymbol{\tau}_n$ are of interest, hence not the individual quantities ν_n and $\boldsymbol{\tau}_n$ themselves) and this form of local alternatives is actually the standard one in the Le Cam theory; see, e.g., Chapter 6 in [23] or Definition 7.14 in [36]. Whenever local asymptotic powers will be considered below, we will assume that $\|\boldsymbol{\tau}_n\|$ is $O(1)$ without being $o(1)$, so that ν_n will characterize (the rate of) the severity of the local alternatives $\boldsymbol{\theta}_n = \boldsymbol{\theta}_0 + \nu_n \boldsymbol{\tau}_n$ (the slower ν_n goes to zero, the more severe the corresponding local alternatives).

Since the sequence of "statistical experiments" associated with (2.3) involves parametric spaces $\boldsymbol{\Theta}_n$ that depend on n, applying Le Cam's theory will require the following reduction of the problem through invariance arguments.

2.1 Reduction through invariance

Denoting as $SO_p(\boldsymbol{\theta})$ the collection of $p \times p$ orthogonal matrices satisfying $\mathbf{O}\boldsymbol{\theta} = \boldsymbol{\theta}$, the null hypothesis $\mathcal{H}_0^{(n)}$ is invariant under the group $\mathcal{G}_{n,\circ}$ collecting the transformations

$$(\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}) \mapsto g_{n\mathbf{O}}(\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}) = (\mathbf{O}\mathbf{X}_{n1}, \ldots, \mathbf{O}\mathbf{X}_{nn}),$$

with $\mathbf{O} \in SO_{p_n}(\boldsymbol{\theta}_{n0})$. The transformation $g_{n\mathbf{O}}$ induces a transformation of the parametric space $\boldsymbol{\Theta}_n$ defined through $(\boldsymbol{\theta}_n, \kappa) \mapsto (\mathbf{O}\boldsymbol{\theta}_n, \kappa)$. The orbits of the resulting induced group are $\mathcal{C}_{u,\kappa}(\boldsymbol{\theta}_{n0}) := \{\boldsymbol{\theta}_n \in \mathcal{S}^{p_n-1} : \boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0} = u\} \times \{\kappa\}$, with $u \in [-1, 1]$ and $\kappa \in (0, \infty)$. In such a context, the invariance principle (see, e.g., [24], Chapter 6) leads to restricting to tests ϕ_n that are invariant with respect to the group $\mathcal{G}_{n,\circ}$. Denoting as $\mathbf{T}_n = \mathbf{T}_n(\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn})$ a maximal invariant statistic for $\mathcal{G}_{n,\circ}$, the class of invariant tests coincides with the class of $\mathbf{T}_n$-measurable tests. Invariant tests thus are to be defined in the image

$$\mathcal{P}^{(n)\mathbf{T}_n} = \left\{ \mathrm{P}_{u,\kappa}^{(n)\mathbf{T}_n} : (u, \kappa) \in \boldsymbol{\Psi} := [-1, 1] \times (0, \infty) \right\} \tag{2.5}$$

of the model $\mathcal{P}^{(n)}$ by $\mathbf{T}_n$, where $\mathrm{P}_{u,\kappa}^{(n)\mathbf{T}_n}$ denotes the common distribution of $\mathbf{T}_n$ under any $\mathrm{P}_{\boldsymbol{\theta}_n,\kappa}^{(n)}$ with $(\boldsymbol{\theta}_n, \kappa) \in \mathcal{C}_{u,\kappa}(\boldsymbol{\theta}_{n0})$. Unlike the original sequence of statistical experiments in (2.3), the invariant one in (2.5) involves a fixed parametric space $\boldsymbol{\Psi}$, which makes it in principle possible to rely on Le Cam's asymptotic theory.

Now, the original local log-likelihood ratios $\log(d\mathrm{P}_{\boldsymbol{\theta}_n,\kappa_n}^{(n)}/d\mathrm{P}_{\boldsymbol{\theta}_{n0},\kappa_n}^{(n)})$ associated with the generic local alternatives $\boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0} + \nu_n \boldsymbol{\tau}_n$ above correspond, in view of (2.4), to the invariant local log-likelihood ratios

$$\Lambda_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n}^{(n)\mathrm{inv}} := \log \frac{d\mathrm{P}_{1-\nu_n^2\|\boldsymbol{\tau}_n\|^2/2,\kappa_n}^{(n)\mathbf{T}_n}}{d\mathrm{P}_{1,\kappa_n}^{(n)\mathbf{T}_n}}. \tag{2.6}$$

Deriving local asymptotic normality (LAN) results requires investigating the asymptotic behavior of such invariant log-likelihood ratios, which in turn requires evaluating the corresponding likelihoods. While obtaining a closed-form expression for $\mathbf{T}_n$ and its distribution is a very challenging task, these likelihoods can be obtained from Lemma 2.5.1 in [16], which, denoting as m_n the surface area measure on $\mathcal{S}^{p_n-1} \times \ldots \times \mathcal{S}^{p_n-1}$ (n times), yields

$$\begin{aligned}
\frac{d\mathrm{P}^{(n)\mathbf{T}_n}_{1-\nu_n^2\|\boldsymbol{\tau}_n\|^2/2,\kappa_n}}{dm_n} &= \int_{SO_{p_n}(\boldsymbol{\theta}_{n0})} \frac{d\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n}}{dm_n}(\mathbf{O}\mathbf{X}_{n1},\ldots,\mathbf{O}\mathbf{X}_{nn})\, d\mathbf{O} \\
&= \int_{SO_{p_n}(\boldsymbol{\theta}_{n0})} \prod_{i=1}^{n} \bigg(\frac{c_{p_n,\kappa_n}}{\omega_{p_n-1}} \exp\big(\kappa_n (\mathbf{O}\mathbf{X}_{ni})'\boldsymbol{\theta}_n\big) \bigg)\, d\mathbf{O} \\
&= \frac{c^n_{p_n,\kappa_n}}{\omega^n_{p_n-1}} \int_{SO_{p_n}(\boldsymbol{\theta}_{n0})} \exp\big(n\kappa_n \bar{\mathbf{X}}_n \mathbf{O}'\boldsymbol{\theta}_n\big)\, d\mathbf{O}, \qquad (2.7)
\end{aligned}$$

where integration is with respect to the Haar measure on $SO_{p_n}(\boldsymbol{\theta}_{n0})$. Note that (2.7) shows that the invariant null probability measure $\mathrm{P}^{(n)\mathbf{T}_n}_{1,\kappa_n}$ coincides with the original null probability measure $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$. In other words, it is only for non-null probability measures that the invariance reduction above is non-trivial.

2.2 Optimal testing under specified κ_n

The main ingredient needed to obtain LAN results is Theorem 1 below, that provides a stochastic second-order expansion of the invariant log-likelihood ratios in (2.6). To state this theorem, we need to introduce the following notation. We will refer to the decomposition $\mathbf{X}_{ni} = U_{ni}\boldsymbol{\theta}_{n0} + V_{ni}\mathbf{S}_n$, with

$$U_{ni} = \mathbf{X}'_{ni}\boldsymbol{\theta}_{n0}, \quad V_{ni} = (1-U^2_{ni})^{1/2} \quad \text{and} \quad \mathbf{S}_{ni} = \frac{(\mathbf{I}_{p_n} - \boldsymbol{\theta}_{n0}\boldsymbol{\theta}'_{n0})\mathbf{X}_{ni}}{\|(\mathbf{I}_{p_n} - \boldsymbol{\theta}_{n0}\boldsymbol{\theta}'_{n0})\mathbf{X}_{ni}\|},$$

as the tangent-normal decomposition of $\mathbf{X}_{ni}$ with respect to $\boldsymbol{\theta}_{n0}$. Under the hypothesis $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$, U_{ni} has probability density function

$$u \mapsto c_{p_n,\kappa_n}(1-u^2)^{(p_n-3)/2}\exp(\kappa_n u)\,\mathbb{I}[u\in[-1,1]], \qquad (2.8)$$

where $\mathbb{I}[A]$ denotes the indicator function of A, $\mathbf{S}_{ni}$ is uniformly distributed over the "equator" $\{\mathbf{x}\in\mathcal{S}^{p_n-1} : \mathbf{x}'\boldsymbol{\theta}_{n0}=0\}$, and U_{ni} and $\mathbf{S}_{ni}$ are mutually independent. Throughout, we will denote as $e_{n\ell} = \mathrm{E}[U^\ell_{ni}]$ and $\tilde{e}_{n\ell} = \mathrm{E}[(U_{ni} - e_{n1})^\ell]$, $\ell = 1,2,\ldots$ the non-central and central moments of U_{ni} under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$, and as $f_{n\ell} = \mathrm{E}[V^\ell_{ni}]$ the corresponding non-central moments of V_{ni}. Although this is not stressed in the notation, these moments clearly depend on p_n and κ_n; for instance,

$$e_{n1} = \frac{\mathcal{I}_{\frac{p_n}{2}}(\kappa_n)}{\mathcal{I}_{\frac{p_n}{2}-1}(\kappa_n)}, \quad \tilde{e}_{n2} = 1 - \frac{p_n-1}{\kappa_n}\,e_{n1} - e^2_{n1} \quad \text{and} \quad f_{n2} = \frac{p_n-1}{\kappa_n}\,e_{n1} \quad (2.9)$$

(this readily follows from (2)–(3) in [34] by using the standard properties of exponential families; see also Lemma S.2.1 in [13]). We can now state the stochastic second-order expansion result of the invariant log-likelihood ratios in (2.6).

Theorem 1 *Let (p_n) be a sequence of integers that diverges to infinity and (κ_n) be an arbitrary sequence in $(0,\infty)$. Let $(\boldsymbol{\theta}_{n0})$, (ν_n) and $(\boldsymbol{\tau}_n)$ be sequences such that $\boldsymbol{\theta}_{n0}$ and $\boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0} + \nu_n \boldsymbol{\tau}_n$ belong to $\mathcal{S}^{p_n-1}$ for any n, with $(\boldsymbol{\tau}_n)$ bounded and (ν_n) such that*

$$\nu_n^2 = O\Big(\frac{\sqrt{p_n}}{n\kappa_n e_{n1}}\Big). \tag{2.10}$$

Then, letting

$$Z_n := \frac{\sqrt{n}(\bar{\mathbf{X}}_n'\boldsymbol{\theta}_0 - e_{n1})}{\sqrt{\tilde{e}_{n2}}} \quad \textit{and} \quad \widetilde{W}_n := \frac{W_n - (p_n-1)}{\sqrt{2(p_n-1)}},$$

we have that

$$\begin{aligned}\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n} = &-\frac{1}{2}\sqrt{n}\kappa_n\nu_n^2\sqrt{\tilde{e}_{n2}}\,\|\boldsymbol{\tau}_n\|^2 Z_n + \frac{n\kappa_n\nu_n^2 e_{n1}}{\sqrt{2}p_n^{1/2}}\|\boldsymbol{\tau}_n\|^2\Big(1-\frac{1}{4}\nu_n^2\|\boldsymbol{\tau}_n\|^2\Big)\widetilde{W}_n \\ &-\frac{1}{8}n\kappa_n\nu_n^4 e_{n1}\|\boldsymbol{\tau}_n\|^4 - \frac{n^2\kappa_n^2\nu_n^4 e_{n1}^2}{4p_n}\|\boldsymbol{\tau}_n\|^4\Big(1-\frac{1}{4}\nu_n^2\|\boldsymbol{\tau}_n\|^2\Big)^2 + o_{\mathrm{P}}(1),\end{aligned}$$

as $n\to\infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$.

Recalling that the log-likelihood ratio $\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n}$ refers to the local perturbation $\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0} = 1 - \nu_n^2\|\boldsymbol{\tau}_n\|^2/2$ of the null reference value $\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_{n0} = 1$, the result in Theorem 1 essentially shows that the invariant model considered enjoys a *local asymptotic quadraticity* (LAQ) structure in the vicinity of the null hypothesis $\mathcal{H}_0^{(n)} : \boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0}$; see, e.g., [23], page 120. Actually, quadraticity, which is supposed to be in the increment $-\nu_n^2\|\boldsymbol{\tau}_n\|^2/2$, only holds for arbitrarily small values of this increment, hence only in regimes where ν_n will converge to zero (in regimes below where, in contrast, ν_n will be constant, the non-flat manifold structure of the hypersphere actually prevents a standard quadraticity property). This LAQ result hints that optimal testing for the specified-κ_n problem at hand is obtained by rejecting the null for small values of Z_n (that is, when $\bar{\mathbf{X}}_n$ and $\boldsymbol{\theta}_{n0}$ project far from each other onto the axis $\pm\boldsymbol{\theta}_{n0}$), for large values of $\widetilde{W}_n$ (that is, when $\bar{\mathbf{X}}_n$ and $\boldsymbol{\theta}_{n0}$ project far from each other onto the orthogonal complement to $\boldsymbol{\theta}_{n0}$ in $\mathbb{R}^{p_n}$), or, more generally, for large values of a hybrid test statistic of the form

$$Q_n^{\mu,\lambda} = \mu\widetilde{W}_n + \lambda(-Z_n),$$

with non-negative weights μ and λ. While any $Q_n^{\mu,\lambda}$ provides a reasonable test statistic for the problem at hand, only one set of weights will yield a Le Cam optimal test and, interestingly, this set of weights depends on the way κ_n behaves with p_n and n. This will be one of the many consequences of the following LAN result.

Theorem 2 *Let (p_n) be a sequence of integers that diverges to infinity, (κ_n) be a sequence in $(0,\infty)$, and $(\boldsymbol{\theta}_{n0})$ be a sequence such that $\boldsymbol{\theta}_{n0}$ belongs to $\mathcal{S}^{p_n-1}$ for any n. Then, there exist a sequence (ν_n) in $(0,\infty)$ and a sequence of random variables (Δ_n) that is asymptotically normal with zero mean and variance Γ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$ such that, for any bounded sequence $(\boldsymbol{\tau}_n)$ such that $\boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0} + \nu_n\boldsymbol{\tau}_n$ belongs to $\mathcal{S}^{p_n-1}$ for any n,*

$$\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n} = \|\boldsymbol{\tau}_n\|^2\Delta_n - \frac{1}{2}\|\boldsymbol{\tau}_n\|^4\Gamma + o_{\mathrm{P}}(1)$$

as $n\to\infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$. If (i) $\kappa_n/p_n\to\infty$, then

$$\nu_n = \frac{p_n^{1/4}}{\sqrt{n\kappa_n}}, \quad \Delta_n = \frac{\widetilde{W}_n}{\sqrt{2}}, \quad \textit{and} \quad \Gamma = \frac{1}{2};$$

if (ii) $\kappa_n/p_n\to\xi>0$, then, letting $c_\xi := \frac{1}{2}+\sqrt{\frac{1}{4}+\xi^2}$,

$$\nu_n = \frac{\sqrt{c_\xi}\,p_n^{3/4}}{\sqrt{n\kappa_n}}, \quad \Delta_n = \frac{\widetilde{W}_n}{\sqrt{2}}, \quad \textit{and} \quad \Gamma = \frac{1}{2};$$

if (iii) $\kappa_n/p_n\to 0$ with $\sqrt{n}\kappa_n/p_n\to\infty$, then

$$\nu_n = \frac{p_n^{3/4}}{\sqrt{n\kappa_n}}, \quad \Delta_n = \frac{\widetilde{W}_n}{\sqrt{2}}, \quad \textit{and} \quad \Gamma = \frac{1}{2};$$

if (iv) $\sqrt{n}\kappa_n/p_n\to\xi>0$, then

$$\nu_n = \frac{p_n^{3/4}}{\sqrt{n\kappa_n}}, \quad \Delta_n = \frac{\widetilde{W}_n}{\sqrt{2}} - \frac{Z_n}{2\xi}, \quad \textit{and} \quad \Gamma = \frac{1}{2}+\frac{1}{4\xi^2};$$

if (v) $\sqrt{n}\kappa_n/p_n\to 0$ with $\sqrt{n}\kappa_n/\sqrt{p_n}\to\infty$, then

$$\nu_n = \frac{p_n^{1/4}}{n^{1/4}\sqrt{\kappa_n}}, \quad \Delta_n = -\frac{Z_n}{2}, \quad \textit{and} \quad \Gamma = \frac{1}{4};$$

if (vi) $\sqrt{n}\kappa_n/\sqrt{p_n}\to\xi>0$, then

$$\nu_n = 1, \quad \Delta_n = -\frac{\xi Z_n}{2}, \quad \textit{and} \quad \Gamma = \frac{\xi^2}{4};$$

finally, if (vii) $\sqrt{n}\kappa_n/\sqrt{p_n}\to 0$, then, even with $\nu_n=1$, the invariant log-likelihood ratio $\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n}$ is $o_{\mathrm{P}}(1)$ as $n\to\infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$.

In the image model (2.5), the spherical location problem consists in testing $\mathcal{H}_0^{(n)} : u = 1$ against $\mathcal{H}_0^{(n)} : u < 1$. In the localized at $u = u_0 = 1$ experiments, parametrized by $u = u_0 - \frac{1}{2}\nu_n^2\|\boldsymbol{\tau}_n\|^2$ as in (2.6), this reduces to testing $\mathcal{H}_0^{(n)} : \|\boldsymbol{\tau}_n\| = 0$ against $\mathcal{H}_1^{(n)} : \|\boldsymbol{\tau}_n\| > 0$. In any given regime (i)–(vii) from Theorem 2, it directly follows from this theorem that a locally asymptotically most powerful test for this problem — hence, *locally asymptotically most powerful invariant* test for the original spherical location problem — rejects the null at asymptotic level α whenever

$$\Delta_n/\sqrt{\Gamma} > \Phi^{-1}(1-\alpha), \tag{2.11}$$

where Φ denotes the cumulative distribution function of the standard normal distribution (in the rest of the paper, the term "*optimal*" will refer to this particular Le Cam optimality concept). A routine application of the Le Cam third lemma then shows that, in each regime, the asymptotic distribution of Δ_n, under the corresponding contiguous alternatives $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n,\kappa_n}$ with $\|\boldsymbol{\tau}_n\| \to t$, is normal with mean Γt^2 and variance Γ, so that the resulting asymptotic power of the optimal test in (2.11) is

$$\lim_{n\to\infty} \mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n,\kappa_n}\big[\Delta_n/\sqrt{\Gamma} > \Phi^{-1}(1-\alpha)\big] \\ = 1 - \Phi\Big(\Phi^{-1}(1-\alpha) - \sqrt{\Gamma}t^2\Big). \tag{2.12}$$

In each regime (i)–(vii), ν_n is the contiguity rate, which implies that the least severe alternatives under which a test may have non-trivial asymptotic powers are of the form $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n,\kappa_n}$, with a sequence $(\|\boldsymbol{\tau}_n\|)$ that is $O(1)$ but not $o(1)$. Theorem 2 shows that this contiguity rate depends on the regime considered and does so in a monotonic fashion, which is intuitively reasonable: the larger κ_n (that is, the easier the inference problem), the faster ν_n goes to zero, that is, the less severe the alternatives that can be detected by rate-consistent tests. Because the unit sphere $\mathcal{S}^{p_n-1}$ has a fixed diameter, $\nu_n = 1$ characterizes the most severe alternatives that can be considered. In regime (vi), no tests will therefore be consistent under such most severe alternatives, while, in regime (vii), the distribution is so close to the uniform distribution on $\mathcal{S}^{p_n-1}$ that no tests can show non-trivial asymptotic powers under such alternatives, so that even the trivial α-test is optimal.

One of the most striking consequences of Theorem 2 is that the optimal test depends on the regime considered. In regimes (v)–(vii), the optimal test in (2.11) rejects the null when $Z_n < \Phi^{-1}(\alpha)$; of course, this optimality is degenerate in regime (vii), where any invariant test with asymptotic level α would also be optimal. In contrast, the optimal α-level test in regimes (i)–(iii) rejects the null when

$$\widetilde{W}_n = \frac{W_n - (p_n - 1)}{\sqrt{2(p_n-1)}} > \Phi^{-1}(1-\alpha).$$

Since the chi-square distribution with $p-1$ degrees of freedom converges, after standardization via its mean $p-1$ and standard deviation $\sqrt{2(p-1)}$, to the standard normal distribution as p diverges to infinity, this test is asymptotically equivalent to the Watson test in (1.2), based obviously on the dimension $p = p_n$ at hand. This shows that, in regimes (i)–(iii), the traditional, low-dimensional, Watson test is optimal in high dimensions. In regime (iv), which is at the frontier between these regimes where the optimal test is the Watson test and those where the optimal test is based on Z_n, the optimal test is quite naturally based on a linear combination of $\widetilde{W}_n$ and Z_n.

Finally, the Le Cam third lemma allows us to derive the asymptotic non-null behavior of the Watson test under the contiguous alternatives considered in any regime (i)–(vii). In regimes (i)–(iv), the limiting powers under contiguous alternatives of the form $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n,\kappa_n}$, with $\|\boldsymbol{\tau}_n\| \to t$, are given by

$$1 - \Phi\bigg(\Phi^{-1}(1-\alpha) - \frac{t^2}{\sqrt{2}}\bigg). \tag{2.13}$$

In regimes (i)–(iii), the Watson test is the optimal test and these asymptotic powers are equal to those in (2.12), whereas in regime (iv), the Watson test is only rate-consistent, as the corresponding asymptotic powers of the optimal test are

$$1 - \Phi\bigg(\Phi^{-1}(1-\alpha) - t^2\sqrt{\frac{1}{2} + \frac{1}{4\xi^2}}\,\bigg). \tag{2.14}$$

In regimes (v)–(vi), the Le Cam third lemma shows that the limiting powers of the Watson test, still under the corresponding contiguous alternatives, are equal to the nominal level α, so that the Watson test is not even rate-consistent in those regimes. Finally, as already discussed, the Watson test is optimal in regime (vii), but trivially so since the trivial α-test there also is.

2.3 Optimal testing under unspecified κ_n

The optimal test in regimes (i)–(iii), namely the Watson test, is a genuine test in the sense that it can be applied on the basis of the observations only. In contrast, the optimal tests in regimes (iv)–(vi) are "oracle" tests since they require knowing the values of e_{n1} and $\tilde{e}_{n2}$, or equivalently (see (2.9)), the value of the concentration κ_n. This concentration, however, can hardly be assumed to be specified in practice, so that it is natural to wonder what is the optimal test, in regimes (iv)–(vi), when κ_n is treated as a nuisance parameter.

We first focus on regime (iv). There, the concentration κ_n is asymptotically of the form $\kappa_n = p_n\xi/\sqrt{n}$ for some $\xi > 0$. Within regime (iv), ξ, obviously, is a perfectly valid alternative concentration parameter. Inspired by the classical treatment of asymptotically optimal inference in the presence of nuisance parameters (see, e.g., [5]), this suggests studying the asymptotic behavior of invariant log-likelihood ratios of the form

$$\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n,\kappa_{n,s}/\boldsymbol{\theta}_{n0},\kappa_n} := \log \frac{d\mathrm{P}^{(n)\mathbf{T}_n}_{1-\nu_n^2\|\boldsymbol{\tau}_n\|^2/2,\kappa_{n,s}}}{d\mathrm{P}^{(n)\mathbf{T}_n}_{1,\kappa_n}},$$

where $\kappa_{n,s} := p_n(\xi + \vartheta_n s)/\sqrt{n}$ is a suitable sequence of perturbed concentrations. We have the following result.

Theorem 3 *Let (p_n) be a sequence of integers that diverges to infinity with $p_n = o(n^2)$ as $n \to \infty$. Let $\kappa_n := p_n\xi/\sqrt{n}$, with $\xi > 0$, and $\kappa_{n,s} := p_n(\xi + s/\sqrt{p_n})/\sqrt{n}$, where s is such that $\xi + s/\sqrt{p_n} > 0$ for any n. Let the sequence $(\boldsymbol{\theta}_{n0})$ in $\mathcal{S}^{p_n-1}$ and the bounded sequence $(\boldsymbol{\tau}_n)$ in $\mathbb{R}^{p_n}$ be such that $\boldsymbol{\theta}_{n0}$ and $\boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0} + \nu_n\boldsymbol{\tau}_n$, with the ν_n below, belong to $\mathcal{S}^{p_n-1}$ for any n. Then, putting $\mathbf{t}_n := (\|\boldsymbol{\tau}_n\|^2, s)'$,*

$$\nu_n := \frac{p_n^{3/4}}{\sqrt{n}\kappa_n}, \quad \boldsymbol{\Delta}_n := \begin{pmatrix} \frac{\widetilde{W}_n}{\sqrt{2}} - \frac{Z_n}{2\xi} \\ Z_n \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\Gamma} := \begin{pmatrix} \frac{1}{2} + \frac{1}{4\xi^2} & -\frac{1}{2\xi} \\ -\frac{1}{2\xi} & 1 \end{pmatrix},$$

we have

$$\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n,\kappa_{n,s}/\boldsymbol{\theta}_{n0},\kappa_n} = \mathbf{t}_n'\boldsymbol{\Delta}_n - \frac{1}{2}\mathbf{t}_n'\boldsymbol{\Gamma}\mathbf{t}_n + o_{\mathrm{P}}(1) \tag{2.15}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$, where $\boldsymbol{\Delta}_n$, under the same sequence of hypotheses, is asymptotically normal with mean zero and covariance matrix $\boldsymbol{\Gamma}$.

Theorem 3 shows that, in regime (iv), the sequence of high-dimensional FvML experiments is jointly LAN in the location and concentration parameters. The corresponding Fisher information matrix $\boldsymbol{\Gamma} = (\Gamma_{ij})$ is not diagonal, which entails that the unspecification of the concentration parameter has asymptotically a positive cost when performing inference on the location parameter. In the present joint LAN framework, Le Cam optimal inference for location under unspecified concentration is to be based (see again [5]) on the residual of the regression (in the limiting Gaussian shift experiment) of the location part Δ_{n1} of the central sequence $\boldsymbol{\Delta} = (\Delta_{n1}, \Delta_{n2})'$ with respect to the concentration part Δ_{n2}, that is, is to be based on the *efficient central sequence*

$$\Delta^*_{n1} := \Delta_{n1} - \frac{\Gamma_{12}}{\Gamma_{22}}\Delta_{n2}. \tag{2.16}$$

Under the null, Δ^*_{n1} is asymptotically normal with mean zero and variance $\Gamma^*_{11} = \Gamma_{11} - \Gamma^2_{12}/\Gamma_{22}$, and the Le Cam optimal location test under unspecified κ_n rejects the null at asymptotical level α when

$$\Delta^*_{n1}/\sqrt{\Gamma^*_{11}} = \widetilde{W}_n > \Phi^{-1}(1-\alpha).$$

As a corollary, provided that $p_n = o(n^2)$, the unspecified-κ_n optimal test in regime (iv) is the Watson test. Consequently, the difference between the local asymptotic powers in (2.13) and (2.14), associated with the Watson test and the specified-κ_n optimal test in regime (iv), respectively, can be interpreted as the asymptotic cost of the unspecification of the concentration when performing inference on location in the regime considered. Note that the optimal specified-κ_n test and optimal unspecified-κ_n test exhibit the same consistency rates, so that the cost of not knowing κ_n lies in the difference of powers these tests show under contiguous alternatives.

We now turn to regime (vi), where the concentration κ_n is asymptotically of the form $\kappa_n = \sqrt{p_n}\xi/\sqrt{n}$. In this regime, taking $\nu_n = 1$ (as in Theorem 2) and perturbed concentrations of the form $\kappa_{n,s} := \sqrt{p_n}(\xi + s)/\sqrt{n}$, it is easy to show, by working along the same lines as in the proof of Theorem 3, that the sequence of experiments is also jointly LAN in location and concentration, this time without any condition on p_n. The corresponding central sequence and Fisher information matrix are

$$\boldsymbol{\Delta}_n := \begin{pmatrix} \Delta_{n1} \\ \Delta_{n2} \end{pmatrix} = \begin{pmatrix} -Z_n/2 \\ Z_n \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Gamma} := \begin{pmatrix} 1/4 & -1/2 \\ -1/2 & 1 \end{pmatrix}. \tag{2.17}$$

The collinearity between the location part Δ_{n1} and concentration part Δ_{n2} of the central sequence implies that the efficient central sequence Δ^*_{n1} is zero in regime (vi). As a result, for the unspecified concentration problem, no test can detect alternatives in $\nu_n = 1$ in regime (vi), which is in line with the corresponding trivial asymptotic powers of the Watson test in Section 2.2. Since $\nu_n = 1$ provides the most severe location alternatives than can be considered, we conclude that, for the unspecified concentration problem, no test in regime (vi) can do asymptotically better than the trivial α-level test that randomly rejects the null with probability α. Under unspecified κ_n, thus, the Watson test is optimal in regime (vi), too, even if it is in a degenerate way.

Finally, we consider regime (v), where the situation is more complicated. This regime is associated with $\kappa_n = p_n r_n \xi/\sqrt{n}$, where $\xi > 0$ and (r_n) is a positive sequence satisfying $r_n = o(1)$ and $r_n\sqrt{p_n} \to \infty$. If one takes $\nu_n = p_n^{1/4}/(n^{1/4}\sqrt{\kappa_n})$ (still as in Theorem 2) and considers perturbed concentrations of the form $\kappa_{n,s} = p_n r_n(\xi + s/(\sqrt{p_n}r_n))/\sqrt{n}$, then it can be shown that, provided that $p_n = o(n^2 r_n^{-6})$, the resulting sequence of experiments is still jointly LAN in location and concentration, with the same central sequence and Fisher information matrix as in (2.17). Consequently, the corresponding efficient central sequence Δ^*_{n1} is zero again, so that no unspecified-κ_n test can detect deviations from the null hypothesis at the ν_n-rate in regime (v). Unlike in regime (vi), however, alternatives that are more severe than the contiguous ones can be considered in regime (v). As a consequence, several important questions are left wide open in regime (v) for the unspecified-κ_n problem: (1) are there alternatives that can be detected by an unspecified-κ_n test? (2) If so, what are the least severe ones that can be detected by such a test and (3) what is the Le Cam optimal test (if any)? (4) Are there alternatives that can be detected by the Watson test? (5) Does this test enjoy any Le Cam optimality property in this regime?

To answer these questions, one needs to orthogonalize the parameter of interest $u = \boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0}$ and concentration parameter κ. In regime (iv), this orthogonalization was achieved, within the LAN framework of Theorem 3, by the efficient central sequence in (2.16). In regime (v), where the consistency rates of the Z_n and $\widetilde{W}_n$ tests do not match, this approach does not work and it is needed to perform orthogonalization by introducing explicitly a new parametrization (such an orthogonalization through reparametrization is suitable when Fisher information matrices are singular; see, e.g., [18]). The following LAN result relates to this new parametrization of the statistical experiments at hand, that involves the same parameter of interest $u = \boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0}$ and the alternative concentration parameter $\bar{\kappa}_n = \kappa_n/u$ (of course, this reparametrization requires restricting to the hemisphere associated with $u > 0$, which still allows us to consider "local" alternatives).

Theorem 4 *Let (p_n) be a sequence of integers that diverges to infinity with $p_n = o(n^2 r_n^{-4})$ as $n \to \infty$, where (r_n) is a positive real sequence such that $r_n = o(1)$ and $\sqrt{p_n} r_n \to \infty$. Let $(\boldsymbol{\theta}_{n0})$ be a sequence in $\mathcal{S}^{p_n-1}$ and $(\boldsymbol{\tau}_n)$ be a bounded sequence in $\mathbb{R}^{p_n}$ such that $\boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0} + \nu_n \boldsymbol{\tau}_n$, with the ν_n below, belongs to $\mathcal{S}^{p_n-1}$ for any n. Let $\kappa_n := p_n r_n \xi/\sqrt{n}$, with $\xi > 0$ and*

$$\kappa_{n,s,\boldsymbol{\tau}_n} = \frac{p_n r_n(\xi + s/(\sqrt{p_n} r_n))}{\sqrt{n}(1 - \frac{1}{2}\nu_n^2\|\boldsymbol{\tau}_n\|^2)} =: \frac{\rho_n p_n r_n}{\sqrt{n}}(\xi + s/(\sqrt{p_n} r_n)),$$

where s is such that $\xi + s/(\sqrt{p_n} r_n) > 0$ for any n. Assume that, still with the ν_n below, $\frac{1}{2}\nu_n^2\|\boldsymbol{\tau}_n\|^2$ is upper-bounded by $1 - \delta$ for some $\delta > 0$. Then, putting

$$(a) \qquad \nu_n = \frac{p_n^{3/4}}{\sqrt{n}\kappa_n}, \qquad \mathbf{C}_n = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\Delta}_n = \begin{pmatrix} \frac{\widetilde{W}_n}{\sqrt{2}} \\ Z_n \end{pmatrix}, \quad \text{and } \boldsymbol{\Gamma} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix},$$

$$(b) \quad \nu_n = 1, \ \mathbf{C}_n = \begin{pmatrix} \xi^2\big(1 - \frac{\|\boldsymbol{\tau}_n\|^2}{4}\big) & 0 \\ 0 & 1 \end{pmatrix}, \ \boldsymbol{\Delta}_n = \begin{pmatrix} \frac{\widetilde{W}_n}{\sqrt{2}} \\ Z_n \end{pmatrix}, \ \text{and } \boldsymbol{\Gamma} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix},$$

or

$$(c) \qquad \nu_n = 1, \qquad \mathbf{C}_n = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\Delta}_n = \begin{pmatrix} 0 \\ Z_n \end{pmatrix}, \quad \text{and } \boldsymbol{\Gamma} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

depending on whether (a) $\rho_n p_n^{1/4} r_n \to \infty$, (b) $\rho_n p_n^{1/4} r_n \to 1$, or (c) $\rho_n p_n^{1/4} r_n = o(1)$, respectively, we have, with $\mathbf{t}_n := (\|\boldsymbol{\tau}_n\|^2, s)'$,

$$\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n,\kappa_{n,s,\boldsymbol{\tau}_n}/\boldsymbol{\theta}_{n0},\kappa_n} = \mathbf{t}_n'\mathbf{C}_n\boldsymbol{\Delta}_n - \frac{1}{2}\mathbf{t}_n'\mathbf{C}_n^2\boldsymbol{\Gamma}\mathbf{t}_n + o_{\mathrm{P}}(1) \tag{2.18}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$, where $\boldsymbol{\Delta}_n$, under the same sequence of hypotheses, is asymptotically normal with mean zero and covariance matrix $\boldsymbol{\Gamma}$.

The block-diagonality of the three Fisher information matrices $\boldsymbol{\Gamma}$ in this result confirms that the new parametrization achieves orthogonalization in regime (v). More importantly, Theorem 4 allows us to answer the open questions above. In this purpose, the key observation is that the problem of testing the null hypothesis $\mathcal{H}_0 : u = 1$ against the alternative $\mathcal{H}_1 : u < 1$ under unspecified κ_n in the original parametrization is strictly equivalent to the problem of testing the null hypothesis $\mathcal{H}_0 : u = 1$ against the alternative $\mathcal{H}_1 : u < 1$ under unspecified $\bar{\kappa}_n$ in the new parametrization. Therefore, the $s \equiv 0$ version of Theorem 4 establishes the following: in regime (v$_a$), which refers to case (a) in this result, the Watson test is Le Cam optimal for the unspecified-κ_n problem and will show non-trivial asymptotic powers under alternatives associated with $p_n^{3/4}/(\sqrt{n}\kappa_n)$ (the Le Cam third lemma readily implies that these asymptotic powers are equal to those in (2.13)). In regime (v$_c$), no unspecified-κ_n test can detect even the most severe alternatives associated with $\nu_n = 1$. In the boundary case of regime (v$_b$), the situation is more complex, as the sequence of statistical experiments there is not LAN. Yet, the result shows that the least severe alternatives that can be detected by an unspecified-κ_n test are those associated with $\nu_n = 1$ and that the Watson test is rate-consistent. Theorem 4(b) also shows that the Watson test is Le Cam optimal for small departures $\boldsymbol{\tau}_n$ of the null hypothesis (this follows from the fact that the usual LAN property is obtained for small $\|\boldsymbol{\tau}_n\|$); we refer to Theorem 4.1(iii) in [29] for a similar phenomenon in low dimensions. This thoroughly answers the questions (1)–(5) raised above.

Wrapping up, we proved that the Watson test is optimal in regimes (i)–(iii) only for the specified concentration problem and that it is optimal in *all* regimes in the more important unspecified concentration one (in regimes (iv)–(v$_a$), optimality requires a constraint on p_n that is at most $p_n = o(n^2)$, and optimality is only local in $\boldsymbol{\tau}_n$ in regime (v$_b$)). The asymptotic cost due to the unspecification of the concentration is nil in regimes (i)–(iii) (and (vii)), affects limiting powers but not consistency rates in regime (iv), and is in terms of consistency rates in regimes (v)–(vi). Table 1 provides a summary of the optimality results we obtained both for the specified-κ_n and unspecified-κ_n problems.

3 Non-null investigation via martingale CLTs

The results above thoroughly describe the asymptotic non-null and optimality properties of the Watson test in the FvML case and provide a strong motivation to use this test in this specific parametric framework. While the Watson test remains valid (in the sense that it still meets the asymptotic nominal level constraint) under much broader distributional assumptions, it is unclear how well this test behaves under high-dimensional non-FvML alternatives (we refer to [30], [31] and [32] for an extensive study of the low-dimensional case). In this section, we therefore investigate, through a different approach relying on martingale CLTs, the non-null behavior of the Watson test under general rotationally symmetric distributions.

Recall that the distribution of a random vector $\mathbf{X}$ with values in $\mathcal{S}^{p-1}$ is *rotationally symmetric about* $\boldsymbol{\theta}(\in \mathcal{S}^{p-1})$ if $\mathbf{OX}$ and $\mathbf{X}$ share the same distribution for any $\mathbf{O} \in SO_{\boldsymbol{\theta}}(p)$, and that it is *rotationally symmetric* if it is rotationally

#	Regime	κ_n *specified*	κ_n *unspecified*
(i)	$\kappa_n/p_n \to \infty$	$\widetilde{W}_n$	$\widetilde{W}_n$
(ii)	$\kappa_n/p_n \to \xi > 0$	$\widetilde{W}_n$	$\widetilde{W}_n$
(iii)	$\kappa_n/p_n \to 0$ with $\sqrt{n}\kappa_n/p_n \to \infty$	$\widetilde{W}_n$	$\widetilde{W}_n$
(iv)	$\sqrt{n}\kappa_n/p_n \to \xi > 0$	$\frac{\widetilde{W}_n}{\sqrt{2}} - \frac{Z_n}{2\xi}$	$\widetilde{W}_n$ $(\star)$
(v$_a$)	$\sqrt{n}\kappa_n/p_n \to 0$ with $\sqrt{n}\kappa_n/p_n^{3/4} \to \infty$	Z_n	$\widetilde{W}_n$ $(\star)$
(v$_b$)	$\sqrt{n}\kappa_n/p_n^{3/4} \to \xi > 0$	Z_n	$\widetilde{W}_n$ (†)
(v$_c$)	$\sqrt{n}\kappa_n/p_n^{3/4} \to 0$ with $\sqrt{n}\kappa_n/\sqrt{p_n} \to \infty$	Z_n	$\emptyset$
(vi)	$\sqrt{n}\kappa_n/\sqrt{p_n} \to \xi > 0$	Z_n	$\emptyset$
(vii)	$\sqrt{n}\kappa_n/\sqrt{p_n} \to 0$	$\emptyset$	$\emptyset$

Table 1 The test statistics on which locally asymptotically optimal tests are based in the various asymptotic regimes for both the specified-κ_n and unspecified-κ_n problems. The symbol $\emptyset$ means that no test can detect even the most severe alternatives associated with $\nu_n = 1$. The symbol $\star$ indicates that the result is obtained provided that $p_n = o(n^2)$ (for the $\star$ in regime (v$_a$), the constraint is actually milder than $p_n = o(n^2)$; see Theorem 4 for details). The symbol † stresses that, in the non-standard limiting experiment obtained in regime (v$_b$) for unspecified κ_n, Le Cam optimality is achieved only locally in $\boldsymbol{\tau}_n$.

symmetric about some $\boldsymbol{\theta}$ in $\mathcal{S}^{p-1}$. Clearly, if $\mathbf{X}$ has an FvML$_p(\boldsymbol{\theta}, \kappa)$ distribution, then it is rotationally symmetric about $\boldsymbol{\theta}$, so that the distributional context considered in this section will encompass the one in Section 2. Parallel to what was done there, we will refer to the decomposition $\mathbf{X} = U\boldsymbol{\theta} + V\mathbf{S}$, with $U = \mathbf{X}'\boldsymbol{\theta}$, $V = \sqrt{1-U^2}$ and $\mathbf{S} = (\mathbf{I}_p - \boldsymbol{\theta}\boldsymbol{\theta}')\mathbf{X}/\|(\mathbf{I}_p - \boldsymbol{\theta}\boldsymbol{\theta}')\mathbf{X}\|$, as the tangent-normal decomposition of $\mathbf{X}$ with respect to $\boldsymbol{\theta}$. If $\mathbf{X}$ is rotationally symmetric about $\boldsymbol{\theta}$, then $\mathbf{S}$ is uniformly distributed over $\{\mathbf{x} \in \mathcal{S}^{p-1} : \mathbf{x}'\boldsymbol{\theta} = 0\}$ and is independent of U. The distribution of $\mathbf{X}$ is then fully determined by $\boldsymbol{\theta}$ and by the cumulative distribution function F of U, which justifies denoting the corresponding distribution as Rot$_p(\boldsymbol{\theta}, F)$. In the sequel, we tacitly restrict to classes of rotationally symmetric distributions making $\boldsymbol{\theta}$ identifiable, which typically only excludes distributions satisfying Rot$_p(-\boldsymbol{\theta}, F) = $ Rot$_p(\boldsymbol{\theta}, F)$.

We consider then a triangular array of observations of the form $\mathbf{X}_{ni}$, $i = 1, \ldots, n$, $n = 1, 2, \ldots$, where $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$ form a random sample from the rotationally symmetric distribution Rot$_{p_n}(\boldsymbol{\theta}_n, F_n)$. The corresponding hypothesis, that will be denoted as $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n, F_n}$ involves a sequence of integers (p_n) diverging to infinity, a sequence $(\boldsymbol{\theta}_n)$ such that $\boldsymbol{\theta}_n \in \mathcal{S}^{p_n - 1}$ for any n, and a sequence (F_n) of cumulative distribution functions over $[-1, 1]$. In this framework, the spherical location problem consists in testing $\mathcal{H}_0^{(n)} : \boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0}$ against

$\mathcal{H}_1^{(n)} : \boldsymbol{\theta}_n \neq \boldsymbol{\theta}_{n0}$, where $(\boldsymbol{\theta}_{n0})$ is a fixed null parameter sequence. Parallel to the notation that was used in the FvML case, we will write $e_{n\ell}$ and $\tilde{e}_{n\ell}$, $\ell = 1, 2, \ldots$ for the non-central and central moments of F_n, respectively. These are the moments, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n, F_n}$, of the quantity $U_{n1} = \mathbf{X}'_{n1}\boldsymbol{\theta}_n$ in the tangent-normal decomposition of $\mathbf{X}_{n1}$ with respect to $\boldsymbol{\theta}_n$. The corresponding non-central moments of $V_{n1} = \sqrt{1 - U_{n1}^2}$ will still be denoted as $f_{n\ell}$.

Using the notation V_{ni} and $\mathbf{S}_{ni}$ from the tangent-normal decomposition of $\mathbf{X}_{ni}$ with respect to the null location $\boldsymbol{\theta}_{n0}$, the Watson test statistic rewrites

$$\widetilde{W}_n = \frac{W_n - (p_n - 1)}{\sqrt{2(p_n - 1)}} = \frac{\sqrt{2(p_n - 1)}}{\sum_{i=1}^n V_{ni}^2} \sum_{1 \leq i < j \leq n} V_{ni} V_{nj} \mathbf{S}'_{ni} \mathbf{S}_{nj},$$

where W_n denotes the Watson test statistic in (1.2) based on the null location $\boldsymbol{\theta}_{n0}$. Under the null and under appropriate local alternatives, it is expected that $\widetilde{W}_n$ is asymptotically equivalent in probability to

$$W_n^* := \frac{\sqrt{2(p_n - 1)}}{n f_{n2}} \sum_{1 \leq i < j \leq n} V_{ni} V_{nj} \mathbf{S}'_{ni} \mathbf{S}_{nj},$$

so that an important step in the investigation of the non-null properties of $\widetilde{W}_n$ is the study of the non-null behavior of W_n^*. A classical martingale central limit theorem (see, e.g., Theorem 35.12 in [6]) provides the following result.

Theorem 5 *Let (p_n) be a sequence of integers that diverges to infinity and $(\boldsymbol{\theta}_{n0})$ be a sequence such that $\boldsymbol{\theta}_{n0}$ belongs to $\mathcal{S}^{p_n-1}$ for any n. Let (F_n) be a sequence of cumulative distribution functions on $[-1, 1]$ such that (a) $f_{n2} > 0$ for any n, (b) $f_{n4}/f_{n2}^2 = o(n)$ and (c) $\sqrt{p_n} e_{n2} = o(1)$. Then, we have the following, where, in each case, $(\boldsymbol{\tau}_n)$ refers to an arbitrary sequence such that $\boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0} + \nu_n \boldsymbol{\tau}_n$ belongs to $\mathcal{S}^{p_n-1}$ for any n and such that $(\|\boldsymbol{\tau}_n\|)$ converges to $t (\in [0, \infty))$:*

(i)–(iii) if (i) $\sqrt{n} e_{n1} \to \infty$, if (ii) $\sqrt{n} e_{n1} \to \xi > 0$, or if (iii) $\sqrt{n} e_{n1} \to 0$ with $\sqrt{n} p_n^{1/4} e_{n1} \to \infty$, then

$$W_n^* \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{t^2}{\sqrt{2}}, 1\right)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0} + \nu_n \boldsymbol{\tau}_n, F_n}$, with $\nu_n = \sqrt{f_{n2}}/(\sqrt{n} p_n^{1/4} e_{n1})$; in cases (i)–(ii), the constraint (c) above is superfluous;

(iv) if $\sqrt{n} p_n^{1/4} e_{n1} \to \xi > 0$, then

$$W_n^* \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\xi^2 t^2}{\sqrt{2}} \left(1 - \frac{t^2}{4}\right), 1\right)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0} + \nu_n \boldsymbol{\tau}_n, F_n}$, with $\nu_n = 1$;

(v) if $\sqrt{n}p_n^{1/4}e_{n1} = o(1)$*, then*

$$W_n^* \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n,F_n}$*, with* $\nu_n = 1$*.*

To obtain the corresponding non-null results for the Watson test statistic $\widetilde{W}_n$, we need to prove that $\widetilde{W}_n$ and W_n^* are indeed asymptotically equivalent in probability. The following result does so in the, possibly non-null, general rotationally symmetric context considered (in the *FvML* case, the *null* version of this result was established when proving the results of Section 2; see the proof of Lemma 2).

Theorem 6 *Let* (p_n) *be a sequence of integers that diverges to infinity and* $(\boldsymbol{\theta}_{n0})$ *be a sequence such that* $\boldsymbol{\theta}_{n0}$ *belongs to* $\mathcal{S}^{p_n-1}$ *for any* n*. Let* (F_n) *be a sequence of cumulative distribution functions on* $[-1,1]$ *such that (a)* $f_{n2} > 0$ *for any* n *and (b)* $f_{n4}/f_{n2}^2 = o(n)$*. Then, with* (ν_n) *and* $(\boldsymbol{\tau}_n)$ *as in Theorem 5, we have that, in each regime (i)–(v) considered there,*

$$\widetilde{W}_n = W_n^* + o_{\mathrm{P}}(1)$$

as $n \to \infty$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n,F_n}$*.*

Of course, Theorem 6 readily implies that Theorem 5 still holds if one substitutes $\widetilde{W}_n$ for W_n^*. Rather than restating the result explicitly, we present the following corollary, which focuses on the FvML case.

Corollary 1 *Let* (p_n) *be a sequence of integers that diverges to infinity,* (κ_n) *be a sequence in* $(0,\infty)$*, and* $(\boldsymbol{\theta}_{n0})$ *be a sequence such that* $\boldsymbol{\theta}_{n0}$ *belongs to* $\mathcal{S}^{p_n-1}$ *for any* n*. Then, we have the following, where in each case* $(\boldsymbol{\tau}_n)$ *refers to an arbitrary sequence such that* $\boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0} + \nu_n\boldsymbol{\tau}_n$ *belongs to* $\mathcal{S}^{p_n-1}$ *for any* n *and such that* $(\|\boldsymbol{\tau}_n\|)$ *converges to* $t(\in[0,\infty))$*:*

(i) if $\sqrt{n}\kappa_n/p_n^{3/4} \to \infty$*, then*

$$W_n^* \xrightarrow{\mathcal{D}} \mathcal{N}\Big(\frac{t^2}{\sqrt{2}}, 1\Big)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n,\kappa_n}$*, with* $\nu_n = p_n^{3/4}/(\sqrt{n}\kappa_n\sqrt{f_{n2}}\,)$*;*

(ii) if $\sqrt{n}\kappa_n/p_n^{3/4} \to \xi > 0$*, then*

$$W_n^* \xrightarrow{\mathcal{D}} \mathcal{N}\Big(\frac{\xi^2 t^2}{\sqrt{2}}\Big(1 - \frac{t^2}{4}\Big), 1\Big)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n,\kappa_n}$*, with* $\nu_n = 1$*;*

(iii) if $\sqrt{n}\kappa_n/p_n^{3/4} = o(1)$*, then*

$$W_n^* \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n,\kappa_n}$*, with* $\nu_n = 1$*.*

It is interesting to comment on how this relates to the results of the previous section: Corollary 1(i) covers the regimes (i)–(iv) and (v_a). In view of the asymptotic behavior of f_{n2} in these regimes (see Lemma 4), Corollary 1(i) confirms the consistency rates of the Watson test in Theorem 2–4, as well as the corresponding asymptotic powers obtained in (2.13) through the Le Cam third lemma. Corollary 1(ii) relates to regime (v_b), where the Watson test can only see the "fixed" alternatives associated with $\nu_n = 1$, with limiting power

$$1 - \Phi\Big(\Phi^{-1}(1-\alpha) - \frac{\xi^2 t^2}{\sqrt{2}}\Big(1 - \frac{t^2}{4}\Big)\Big) \tag{3.19}$$

(note that this limiting power can be obtained both by using Corollary 1(ii) or by applying the Le Cam third lemma in Theorem 4, even if the second approach will provide the result only for alternatives associated with $t < \sqrt{2}$, that is, for alternatives in the open hemisphere centered at the null location).

The limiting power in (3.19) increases monotonically from the nominal level α (for $t = 0$, where the underlying location is the null one) to its maximal value (achieved at $t = \sqrt{2}$, that is, when the true location is orthogonal to the null one), then decreases monotonically to α (this limiting value being obtained when the true location is antipodal to the null location). This non-monotonic pattern of the asymptotic power in this regime is a direct consequence of the nature of the Watson test that, as already mentioned, rejects the null when $\bar{\mathbf{X}}_n$ and $\boldsymbol{\theta}_{n0}$ project far from each other onto the orthogonal complement to $\boldsymbol{\theta}_{n0}$ in $\mathbb{R}^{p_n}$. Finally, Corollary 1(iii) indicates that, for $\sqrt{n}\kappa_n/p_n^{3/4} = o(1)$, there are no alternatives under which the Watson test can show asymptotic powers larger than the nominal level α, which is perfectly in line with results obtained in the previous section for the corresponding regimes, namely for regimes (v_c), (vi) and (vii).

4 Simulations

This section reports the results of a Monte Carlo study we conducted to see how well the finite-sample behavior of the various tests reflect the asymptotic findings in Theorems 2–4 and Corollary 1. To compare the results for different values of p/n (note that most aforementioned asymptotic findings allow p_n to go to infinity at an arbitrary rate), we conducted three simulations, for $(n,p) = (800,200)$, $(n,p) = (400,400)$, and $(n,p) = (200,800)$, respectively. In each simulation, we generated, for every combination of $r = (i),\ldots,(iv),(v_a),(v_b),(vi),(vii)$ and $\ell = 0,1,\ldots,L = 5$, a collection of $M =$

1, 000 independent random samples of size n from the p-variate FvML distribution with location

$$\begin{aligned}\boldsymbol{\theta}_{n,r,\ell} &:= (1,0,\ldots,0)' + \nu_{n,r}\Big(-\frac{2\nu_{n,r}\ell^2}{L^2}, \frac{2\ell}{L}\Big(1-\frac{\nu_{n,r}^2\ell^2}{L^2}\Big)^{1/2}, 0,\ldots,0\Big)' \\ &=: \boldsymbol{\theta}_{n0} + \nu_{n,r}\boldsymbol{\tau}_{n,r,\ell} \in \mathcal{S}^{p_n-1}\end{aligned}$$

and concentration $\kappa_{n,r}$. The index r allows to consider the various regimes from Theorem 2 (associated with the $\kappa_{n,r}$ used). In each case, we considered the corresponding local alternatives (associated with $\nu_{n,r}$) from the same theorem. More precisely, we used

- $\kappa_{n,(i)} := p_n^2$, $\nu_{n,(i)} = p_n^{1/4}/\sqrt{n\kappa_n}$,
- $\kappa_{n,(ii)} := p_n$, $\nu_{n,(ii)} = \sqrt{c_1}p_n^{3/4}/(\sqrt{n}\kappa_n)$,
- $\kappa_{n,(iii)} := p_n/n^{1/4}$, $\nu_{n,(iii)} = p_n^{3/4}/(\sqrt{n}\kappa_n)$,
- $\kappa_{n,(iv)} := p_n/\sqrt{n}$, $\nu_{n,(iv)} = p_n^{3/4}/(\sqrt{n}\kappa_n)$,
- $\kappa_{n,(v_a)} := p_n^{7/8}/\sqrt{n}$, $\nu_{n,(v_a)} = p_n^{1/4}/(n^{1/4}\sqrt{\kappa_n})$,
- $\kappa_{n,(v_b)} := p_n^{3/4}/\sqrt{n}$, $\nu_{n,(v_b)} = p_n^{1/4}/(n^{1/4}\sqrt{\kappa_n})$,
- $\kappa_{n,(vi)} := \sqrt{p_n}/\sqrt{n}$, $\nu_{n,(vi)} = 1$, and
- $\kappa_{n,(vii)} := p_n^{1/4}/\sqrt{n}$, $\nu_{n,(vii)} = 1$.

The value $\ell = 0$ corresponds to the null hypothesis $\mathcal{H}_0^{(n)} : \boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0}$, whereas the values $\ell = 1,\ldots,5$ provide increasingly severe alternatives. For each sample, we performed three tests, all at asymptotic level $\alpha = 5\%$, namely (a) the Watson test rejecting the null when

$$W_n = \frac{n(p-1)\bar{\mathbf{X}}_n'(\mathbf{I}_p - \boldsymbol{\theta}_{n0}\boldsymbol{\theta}_{n0}')\bar{\mathbf{X}}_n}{1-\frac{1}{n}\sum_{i=1}^n(\mathbf{X}_{ni}'\boldsymbol{\theta}_{n0})^2} > \chi^2_{p-1,1-\alpha},$$

(b) the Z_n-test rejecting the null when

$$Z_n = \frac{\sqrt{n}(\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} - e_{n1})}{\sqrt{\tilde{e}_{n2}}} < \Phi^{-1}(\alpha),$$

and (c) the hybrid test rejecting the null when

$$H_n := \Big(\frac{\widetilde{W}_n}{\sqrt{2}} - \frac{Z_n}{2\xi_n}\Big)\Big/\sqrt{\frac{1}{2}+\frac{1}{4\xi_n^2}} > \Phi^{-1}(1-\alpha),$$

where $\xi_n := \sqrt{n}\kappa_n/p_n$ is based on the (unknown) concentration κ_n depending on the regime r at hand. In each regime from Theorem 2, this hybrid test is clearly expected to behave as the corresponding optimal specified-κ_n test. We stress that the tests (b)–(c) address the specified-κ_n problem only, whereas the Watson test (a) addresses both the specified-κ_n and unspecified-κ_n problems.

Plots of the resulting rejection frequencies are provided in Figures 1 to 3, for $(n, p) = (800, 200)$, $(n, p) = (400, 400)$ and $(n, p) = (200, 800)$, respectively.

In each case, the asymptotic powers, obtained from (2.12)–(2.14), are also plotted. Clearly, irrespective of the three values of p/n considered, the rejection frequencies of the tests are in an excellent agreement with the corresponding asymptotic powers. Also, the results confirm the adaptive nature of the hybrid test, that throughout is the most powerful test.

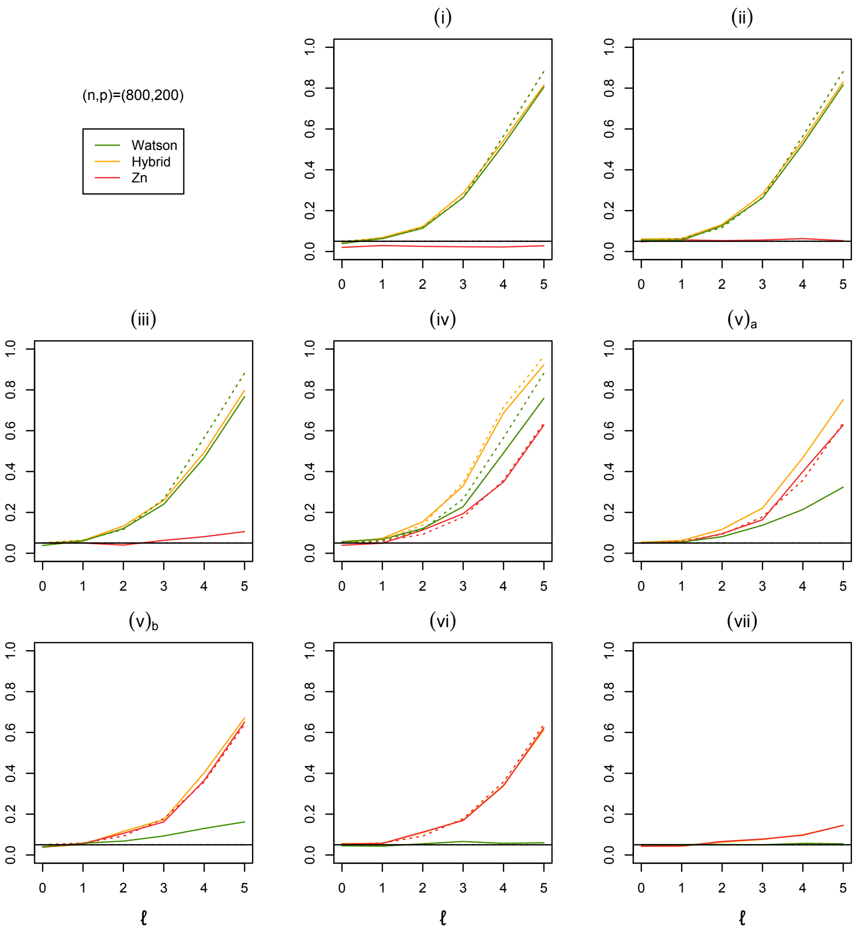

Fig. 1 Rejection frequencies (solid lines), out of $M = 1{,}000$ independent replications, of the Watson test (green), the hybrid test (orange) and the Z_n-based test (red) for $\mathcal{H}_0^{(n)}$: $\boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0} = (1, 0, \ldots, 0)' \in \mathbb{R}^p$, under the null ($\ell = 0$) and under increasingly severe p-dimensional FvML alternatives ($\ell = 1, \ldots, 5$); here, the sample size is $n = 800$ and the dimension is $p = 200$. The regimes $(i), \ldots, (vii)$ fix the way the underlying concentration κ_n is chosen as a function of n and p. In each regime, the corresponding contiguous alternatives from Theorem 2 are used; see Section 4 for details. The corresponding asymptotic powers are plotted in each case (dashed lines).

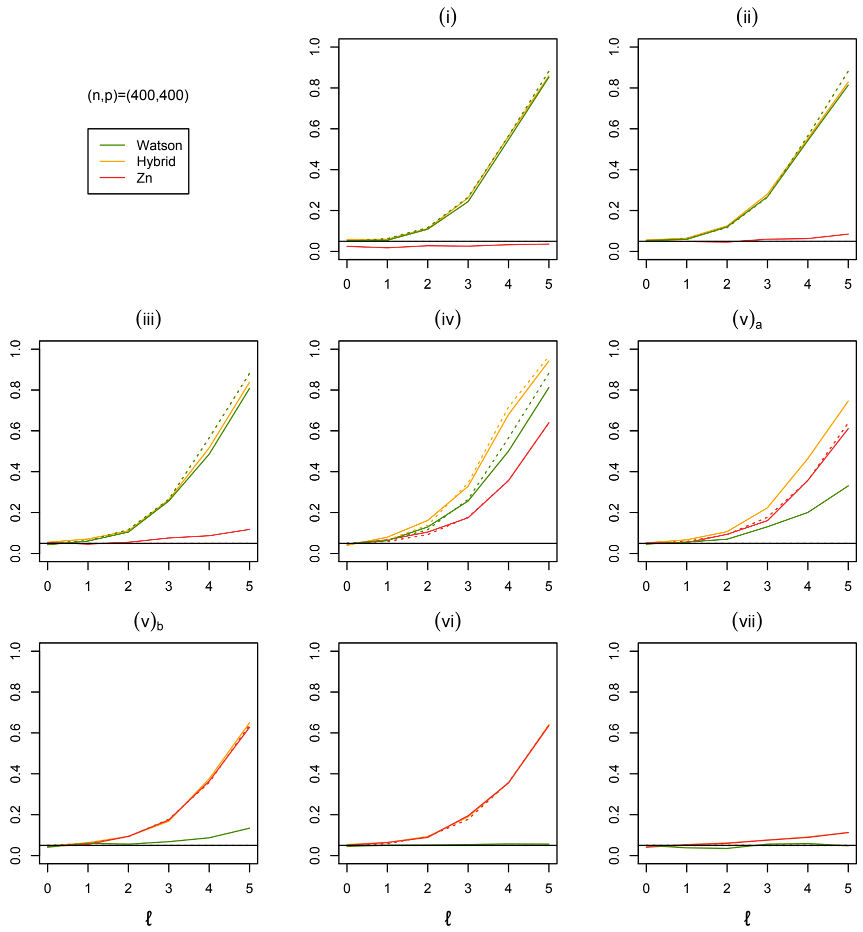

Fig. 2 Same results as in Figure 1, but for sample size $n = 400$ and dimension $p = 400$.

To illustrate similarly the results of Theorem 4 and Corollary 1, we focused on the regimes (v_a)–(v_b) above, but considered the corresponding more severe alternatives. More precisely, we here took

- $\kappa_{n,(v_a)} := p_n^{7/8}/\sqrt{n}$, $\nu_{n,(v_a)} = p_n^{3/4}/(\sqrt{n}\kappa_n)$, and
- $\kappa_{n,(v_b)} := p_n^{3/4}/\sqrt{n}$, $\nu_{n,(v_b)} = 1$.

The rejection frequencies of the same three tests as above, still based on $M = 1{,}000$ independent replications, are provided in Figure 4. For the Watson test, the agreement between rejection frequencies and asymptotic powers is perfect in regime (v_b) (where the non-monotonic asymptotic power pattern is confirmed), but is less so in regime (v_a); at the finite dimensions / sample sizes considered, this may be explained by the fact that the regimes (v_a)–(v_b)

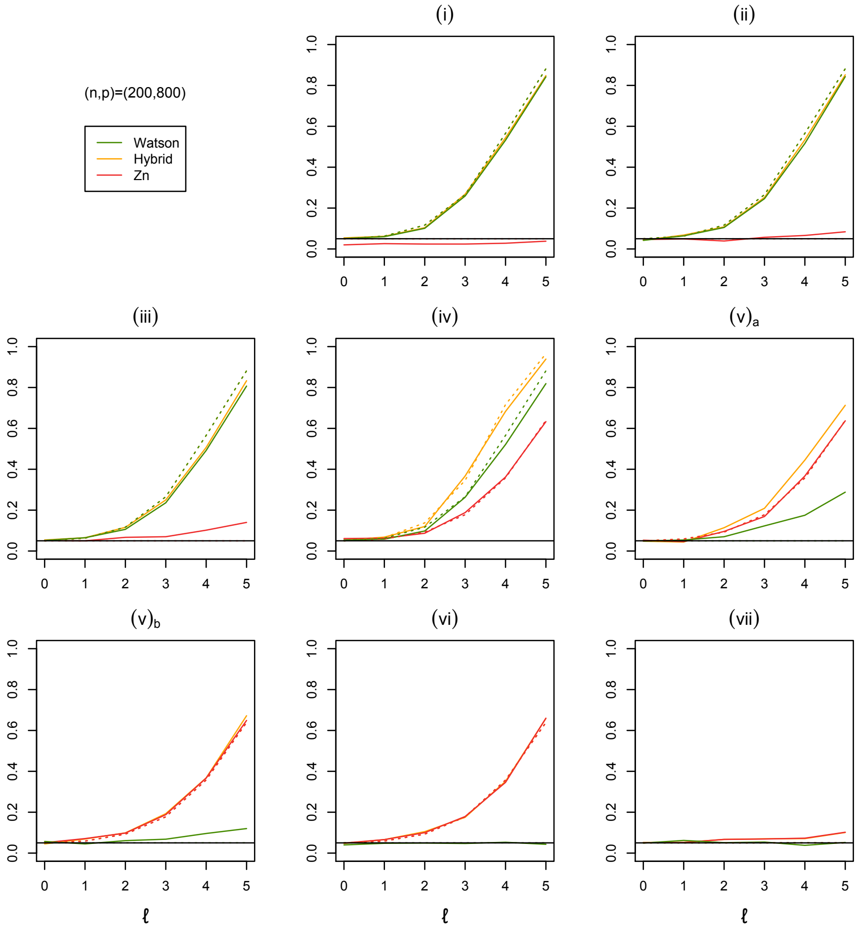

Fig. 3 Same results as in Figures 1–2, but for sample size $n = 200$ and dimension $p = 800$.

are close to each other, so that the empirical powers of the Watson test in regime (v_a) tends to be pulled to the ones in regime (vi).

5 Summary and research perspectives

In the present paper, we tackled the problem of testing, in high dimensions, the null hypothesis that the spike direction $\boldsymbol{\theta}$ of a rotationally symmetric distribution is equal to a given direction $\boldsymbol{\theta}_0$. Under FvML distributional assumptions, we showed that, after resorting to the invariance principle, the sequence of statistical experiments at hand is LAN. More precisely, we identified seven regimes, according to the way the underlying concentration parameter κ_n depends on n and p_n, each leading to a specific limiting experiment, with its

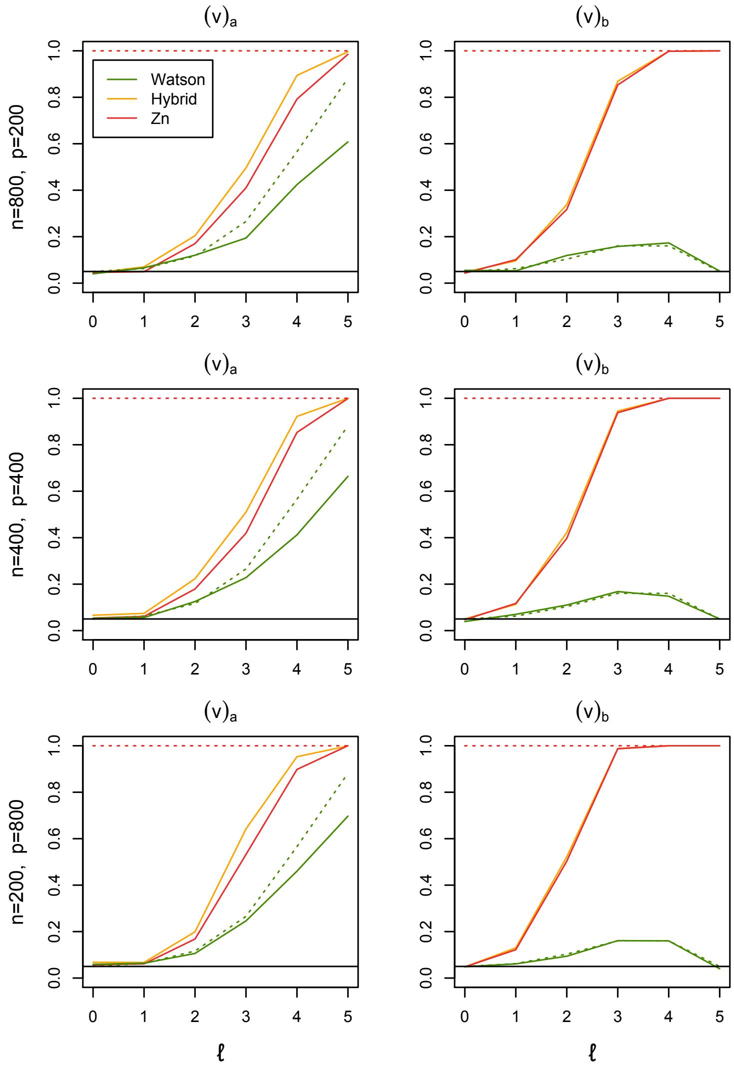

Fig. 4 Rejection frequencies (solid lines), out of $M = 1{,}000$ independent replications, of the Watson test (green), the hybrid test (orange) and the Z_n-based test (red) for $\mathcal{H}_0^{(n)} : \boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0} = (1, 0, \ldots, 0)' \in \mathbb{R}^p$, under the null ($\ell = 0$) and under increasingly severe p-dimensional FvML alternatives ($\ell = 1, \ldots, 5$); the couples (n, p) used are those from Figures 1–3. Here, we focus on the regimes (v$_a$)–(v$_b$) and consider the more severe alternatives associated with Theorem 4 and Corollary 1; see Section 4 for details. The corresponding asymptotic powers are plotted in each case (dashed lines).

own central sequence, Fisher information and contiguity rate (interestingly, these heterogeneous contiguity rates precisely quantify how difficult the problem gets for low concentration situations). As a result, the Le Cam optimal test (more precisely, the locally asymptotically most powerful invariant test) depends on the regime considered. In regimes where $\sqrt{n}\kappa_n/p_n \to \infty$, the classical Watson test is optimal, whereas in regimes where $\sqrt{n}\kappa_n/p_n = O(1)$, the optimal test is an oracle test that explicitly involves the unknown value of the underlying concentration κ_n. If $\sqrt{n}\kappa_n/p_n \to \xi > 0$, then the Watson test fails to be optimal but is still rate-consistent, whereas if $\sqrt{n}\kappa_n/p_n = o(1)$, then it is not even rate-consistent. In all cases, we obtained from the Le Cam third Lemma the asymptotic powers of the corresponding optimal tests and of the Watson test under contiguous alternatives. All results above allow the dimension p_n to go to infinity arbitrarily slowly or arbitrarily fast as a function of n, hence cover moderately high dimensions as well as ultra-high dimensions.

Optimality above refers to the specified-κ_n version of the testing problem considered. Since the concentration κ_n can hardly be assumed to be known in practice, however, optimality results for the corresponding unspecified-κ_n problem are more relevant. For this problem, the Watson test of course remains optimal in regimes where $\sqrt{n}\kappa_n/p_n \to \infty$. But remarkably, for unspecified κ_n, the Watson test is also optimal in regimes where $\sqrt{n}\kappa_n/p_n = O(1)$, sometimes under the condition that $p_n = o(n^2)$ (on an even weaker condition on p_n); we refer to Table 1 and to Theorems 3–4 for details.

Our work opens several perspectives for future research. (a) First, while we derived non-null results for the Watson test also outside the FvML distributional setup, all our optimality results are limited to the FvML case. A natural question is therefore whether or not the strong optimality properties of the Watson test extend away from the FvML case. The low-dimensional investigation conducted in [31] leads us to conjecture that optimality would also hold away from the FvML case, at least in low concentration patterns. Establishing this would require expanding invariant log-likelihood ratios taking a much more complicated form than in the FvML case. This calls for entirely different techniques, hence is beyond the scope of the present paper. (b) Second, we would like to mention that our results are also relevant in a Euclidean (i.e., non-directional) context. They indeed characterize the asymptotic efficiency of sign tests for the direction $\boldsymbol{\theta}$ of a skewed single-spiked distribution in $\mathbb{R}^p$, that is, a distribution whose projection along $\boldsymbol{\theta}$ is skewed and whose projection onto the orthogonal complement to $\boldsymbol{\theta}$ is spherically symmetric. This skewed version of the corresponding classical, elliptical, problem is natural in a signal detection framework, where the signal at hand is quite naturally maximal in direction $\boldsymbol{\theta}$ and minimal in the opposite direction $-\boldsymbol{\theta}$. While our results exhaustively address the question of efficiency of sign tests for this problem (that is, of tests that involve the observations only through their direction form the center of the distribution), it would be of interest to also consider the efficiency of more general testing procedures.

A Technical proofs for Section 2

The proof of Theorem 1 requires the following preliminary results.

Lemma 1 *Let (p_n) be a sequence of integers diverging to infinity and (κ_n) be an arbitrary sequence in $(0,\infty)$. Let $L_n := \sum_{i=1}^n V_{ni}^2/(nf_{n2})$, where we used the notation $V_{ni} = (1-(\mathbf{X}_{ni}'\boldsymbol{\theta}_{n0})^2)^{1/2}$ and $f_{n2} = \mathrm{E}[V_{n1}^2]$. Then, $\mathrm{E}\big[(L_n-1)^2\big] = o(p_n^{-1})$ as $n\to\infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$.*

Proof of Lemma 1. Since

$$\mathrm{E}\Bigg[\bigg(\frac{\sum_{i=1}^n V_{ni}^2}{nf_{n2}}-1\bigg)^2\Bigg] = \frac{1}{f_{n2}^2}\,\mathrm{E}\Bigg[\bigg(\frac{1}{n}\sum_{i=1}^n V_{ni}^2 - \mathrm{E}[V_{n1}^2]\bigg)^2\Bigg]$$
$$= \frac{1}{f_{n2}^2}\,\mathrm{Var}\Bigg[\frac{1}{n}\sum_{i=1}^n V_{ni}^2\Bigg] = \frac{\mathrm{Var}[V_{n1}^2]}{nf_{n2}^2} = \frac{f_{n4}-f_{n2}^2}{nf_{n2}^2}$$

(recall that $f_{n4} := \mathrm{E}[V_{n1}^4]$), it is sufficient to prove that

$$\frac{f_{n4}-f_{n2}^2}{f_{n2}^2} = O(p_n^{-1}). \tag{A.20}$$

Now, the expression for f_{n4}/f_{n2}^2 in page 82 of [25] yields

$$\bigg|\frac{f_{n4}-f_{n2}^2}{f_{n2}^2}\bigg| = \bigg|\frac{(p_n+1)\mathcal{I}_{\frac{p_n}{2}+1}(\kappa_n)\mathcal{I}_{\frac{p_n}{2}-1}(\kappa_n)}{(p_n-1)(\mathcal{I}_{\frac{p_n}{2}}(\kappa_n))^2}-1\bigg|$$
$$= \bigg|\frac{(p_n+1)(\mathcal{I}_{\frac{p_n}{2}+1}(\kappa_n)\mathcal{I}_{\frac{p_n}{2}-1}(\kappa_n)-(\mathcal{I}_{\frac{p_n}{2}}(\kappa_n))^2)}{(p_n-1)(\mathcal{I}_{\frac{p_n}{2}}(\kappa_n))^2}+\frac{2}{p_n-1}\bigg|$$
$$\leq \frac{3|\mathcal{I}_{\frac{p_n}{2}+1}(\kappa_n)\mathcal{I}_{\frac{p_n}{2}-1}(\kappa_n)-(\mathcal{I}_{\frac{p_n}{2}}(\kappa_n))^2|}{(\mathcal{I}_{\frac{p_n}{2}}(\kappa_n))^2}+\frac{2}{p_n-1}.$$

Since $|\mathcal{I}_{\frac{p_n}{2}+1}(\kappa_n)\mathcal{I}_{\frac{p_n}{2}-1}(\kappa_n)-(\mathcal{I}_{\frac{p_n}{2}}(\kappa_n))^2| \leq (\mathcal{I}_{\frac{p_n}{2}}(\kappa_n))^2/(\frac{p_n}{2}+1)$ (see (3.1)–(3.2) in [21]), the result follows. □

Lemma 2 *Let (p_n) be a sequence of integers that diverges to infinity and (κ_n) be an arbitrary sequence in $(0,\infty)$. Let $(\boldsymbol{\theta}_{n0})$ be a sequence such that $\boldsymbol{\theta}_{n0}$ belongs to $\mathcal{S}^{p_n-1}$ for any n. Consider the random variables $\widetilde{W}_n$ and Z_n introduced in Theorem 1. Then, $(\widetilde{W}_n, Z_n)'$ is asymptotically standard bivariate normal under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$.*

Proof of Lemma 2. Throughout the proof, expectations and variances are under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$ and stochastic convergences are as $n\to\infty$ under the same sequence of hypotheses, whereas U_{ni}, V_{ni} and $\mathbf{S}_{ni}$ refer to the tangent-normal decomposition of $\mathbf{X}_{ni}$ with respect to $\boldsymbol{\theta}_{n0}$. Letting then

$$W_n^* = \frac{\sqrt{2(p_n-1)}}{nf_{n2}}\sum_{1\leq i<j\leq n} V_{ni}V_{nj}\mathbf{S}_{ni}'\mathbf{S}_{nj},$$

assume that $(W_n^*, Z_n)'$ is asymptotically standard bivariate normal. Then,

$$\widetilde{W}_n - W_n^* = \left[\frac{\sqrt{2(p_n-1)}}{\sum_{i=1}^n V_{ni}^2} - \frac{\sqrt{2(p_n-1)}}{nf_{n2}} \right] \sum_{1\leq i<j\leq n} V_{ni}V_{nj}\mathbf{S}_{ni}'\mathbf{S}_{nj} \tag{A.21}$$

$$= \left[1 - \frac{\sum_{i=1}^n V_{ni}^2}{nf_{n2}} \right] \times \frac{nf_{n2}}{\sum_{i=1}^n V_{ni}^2} \times \left(\frac{\sqrt{2(p_n-1)}}{nf_{n2}} \sum_{1\leq i<j\leq n} V_{ni}V_{nj}\mathbf{S}_{ni}'\mathbf{S}_{nj} \right)$$

$$= \frac{1-L_n}{L_n} W_n^*,$$

where L_n was introduced in Lemma 1. This lemma implies that $L_n - 1$, hence also $(1 - L_n)/L_n$, is $o_{\rm P}(1)$. If $(W_n^*, Z_n)'$ is indeed asymptotically standard bivariate normal, then we conclude that $\widetilde{W}_n - W_n^*$ is $o_{\rm P}(1)$, so that $(\widetilde{W}_n, Z_n)'$ itself is asymptotically standard bivariate normal.

It is therefore sufficient to show that $(W_n^*, Z_n)'$ is asymptotically standard bivariate normal. We will do this by fixing γ and η such that $\gamma^2+\eta^2=1$ and by using a classical martingale Central Limit Theorem to show that $D_n := \gamma W_n^* + \eta Z_n$ is asymptotically standard normal. To do so, let $\mathcal{F}_{n\ell}$ be the σ-algebra generated by $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{n\ell}$ and denote by $\mathrm{E}_{n\ell}[.]$ the conditional expectation with respect to $\mathcal{F}_{n\ell}$. Define $D_{n\ell} := \mathrm{E}_{n\ell}[D_n] - \mathrm{E}_{n,\ell-1}[D_n]$ for $\ell = 1, \ldots, n$ and $D_{n\ell} = 0$ for $\ell > n$. It is then easy to check that $D_{n\ell} = \gamma W_{n\ell}^* + \eta Z_{n\ell}$, with

$$W_{n\ell}^* := \frac{\sqrt{2(p_n-1)}}{nf_{n2}} \sum_{i=1}^{\ell-1} V_{ni}V_{n\ell}\mathbf{S}_{ni}'\mathbf{S}_{n\ell} \quad \text{and} \quad Z_{n\ell} := \frac{U_{n\ell} - e_{n1}}{\sqrt{n\tilde{e}_{n2}}}$$

for $\ell = 1, \ldots, n$ and $W_{n\ell}^* = 0 = Z_{n\ell}$ for $\ell > n$ (W_{n1}^* is also to be understood as zero). To conclude from the martingale Central Limit Theorem in Theorem 35.12 from [6] that $D_n = \sum_{\ell=1}^\infty D_{n\ell}$ is indeed asymptotically standard normal, we need to show that (a) $\sum_{\ell=1}^n \sigma_{n\ell}^2 \to 1$ in probability, with $\sigma_{n\ell}^2 := \mathrm{E}_{n,\ell-1}[D_{n\ell}^2]$, and that (b) $\sum_{\ell=1}^n \mathrm{E}[D_{n\ell}^2\, \mathbb{I}[|D_{n\ell}| > \varepsilon]] \to 0$ for any $\varepsilon > 0$. Clearly, for $\ell = 1, \ldots, n$,

$$\sigma_{n\ell}^2 = \gamma^2 \mathrm{E}_{n,\ell-1}[(W_{n\ell}^*)^2] + \eta^2 \mathrm{E}_{n,\ell-1}[Z_{n\ell}^2] + 2\gamma\eta \mathrm{E}_{n,\ell-1}[W_{n\ell}^* Z_{n\ell}]$$

$$= \gamma^2 \mathrm{E}_{n,\ell-1}[(W_{n\ell}^*)^2] + \frac{\eta^2}{n}, \tag{A.22}$$

so that (a) follows from Lemma A.1 in [25]. We may thus focus on (b). Since

$$\mathrm{E}_{n,\ell-1}[(W_{n\ell}^*)^2] = 2(n^2 f_{n2})^{-1} \sum_{i,j=1}^{\ell-1} V_{ni}V_{nj}\mathbf{S}_{ni}'\mathbf{S}_{nj},$$

we obtain $\mathrm{Var}[D_{n\ell}] = \mathrm{E}[\sigma_{n\ell}^2] = \gamma^2 \mathrm{E}[(W_{n\ell}^*)^2] + (\eta^2/n) = 2\gamma^2(\ell-1)/n^2 + (\eta^2/n) \leq 2/n$, which yields that there exists a constant C such that, for any $\varepsilon > 0$,

$$\sum_{\ell=1}^n \mathrm{E}[D_{n\ell}^2\, \mathbb{I}[|D_{n\ell}| > \varepsilon]] \leq \sum_{\ell=1}^n \sqrt{\mathrm{E}[D_{n\ell}^4]}\, \sqrt{\mathrm{P}[|D_{n\ell}| > \varepsilon]}$$

$$\leq \frac{1}{\varepsilon} \sum_{\ell=1}^n \sqrt{\mathrm{E}[D_{n\ell}^4]}\, \sqrt{\mathrm{Var}[D_{n\ell}]} \leq \frac{\sqrt{2}}{\sqrt{n}\varepsilon} \sum_{\ell=1}^n \sqrt{\mathrm{E}[D_{n\ell}^4]}$$

$$\leq \frac{C}{\sqrt{n}\varepsilon} \sum_{\ell=1}^n \sqrt{\mathrm{E}[(W_{n\ell}^*)^4]} + \frac{C}{\sqrt{n}\varepsilon} \sum_{\ell=1}^n \sqrt{\mathrm{E}[Z_{n\ell}^4]}.$$

From (A.9) in [25], we then obtain

$$\sum_{\ell=1}^{n} \mathrm{E}[D_{n\ell}^2\, \mathbb{I}[|D_{n\ell}| > \varepsilon]] \leq \frac{C}{\sqrt{n}\varepsilon} \sum_{\ell=1}^{n} \sqrt{\frac{12}{n^4}\Big(\ell\, \frac{f_{n4}^2}{f_{n2}^4} + \ell^2\, \frac{f_{n4}}{f_{n2}^2}\Big)} + \frac{C\sqrt{n}}{\varepsilon} \sqrt{\frac{\tilde{e}_{n4}}{n^2\tilde{e}_{n2}^2}}$$

$$\leq \frac{\sqrt{12}C}{\varepsilon} \sqrt{\Big(\frac{f_{n4}}{nf_{n2}^2}\Big)^2 + \frac{f_{n4}}{nf_{n2}^2}} + \frac{C}{\varepsilon} \sqrt{\frac{\tilde{e}_{n4}}{n\tilde{e}_{n2}^2}}.$$

The result therefore follows from the fact that both f_{n4}/f_{n2}^2 and $\tilde{e}_{n4}/\tilde{e}_{n2}^2$ are upper-bounded by a universal constant; see Theorem S.2.1 in [13]. □

Lemma 3 *Let (ν_n) be a sequence in $(0,\infty)$ that diverges to ∞, (a_n), (b_n) be sequences in $(0,\infty)$ such that $\liminf a_n > 0$, $b_n/\nu_n \to \xi \in [0,\infty)$ and $b_n^6 = o(a_n^4\nu_n^5)$. Let T_n be a sequence of random variables that is $O_{\mathrm{P}}(1)$. Then, writing,*

$$H_\nu(x) := \frac{\int_{-1}^{1}(1-t^2)^{\nu-\frac{1}{2}}\exp(xt)\,dt}{\int_{-1}^{1}(1-t^2)^{\nu-\frac{1}{2}}\,dt} = \frac{\Gamma(\nu+1)\mathcal{I}_\nu(x)}{(x/2)^\nu},$$

we have that

$$a_n^2 \log H_{\nu_n}\Big(\frac{b_nT_n}{a_n}\Big) = \frac{b_n^2T_n^2}{4\nu_n} - \frac{b_n^4T_n^4}{32\nu_n^3a_n^2} + \frac{\xi^2T_n^2}{4} + o_{\mathrm{P}}(1)$$

as $n\to\infty$.

Proof of Lemma 3. The proof is based on the bounds

$$S_{\nu+\frac{1}{2},\nu+\frac{3}{2}}(x) \leq \log H_\nu(x) \leq S_{\nu,\nu+2}(x)$$

for any $x>0$, with $S_{\alpha,\beta}(x) := \sqrt{x^2+\beta^2} - \beta - \alpha\log((\alpha+\sqrt{x^2+\beta^2})/(\alpha+\beta))$; see (5) in [20]. Consider

$$G_\nu(x) := \log H_\nu(x) - \frac{x^2}{4\nu} + \frac{x^4}{32\nu^3} + \frac{x^2}{4\nu^2},$$

along with its resulting lower and upper bounds

$$G_\nu^{\mathrm{low}}(x) := S_{\nu+\frac{1}{2},\nu+\frac{3}{2}}(x) - \frac{x^2}{4\nu} + \frac{x^4}{32\nu^3} + \frac{x^2}{4\nu^2}$$

and

$$G_\nu^{\mathrm{up}}(x) := S_{\nu,\nu+2}(x) - \frac{x^2}{4\nu} + \frac{x^4}{32\nu^3} + \frac{x^2}{4\nu^2}.$$

We prove the lemma by establishing that

$$a_n^2 G_{\nu_n}^{\mathrm{low/up}}\Big(\frac{b_nT_n}{a_n}\Big) = o_{\mathrm{P}}(1). \tag{A.23}$$

To do so, we expand the log term in $G_\nu^{\mathrm{low/up}}(x)$ as $\log x = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3c^3}(x-1)^3$ with $c\in(1,x)$ (note that the argument of these log terms is larger than or equal to one), and we write $G_\nu^{\mathrm{low/up}}(x) = G_\nu^{\mathrm{low/up},1}(x) + G_\nu^{\mathrm{low/up},2}(x)$, with

$$G_\nu^{\mathrm{low},1}(x) := \sqrt{x^2+(\nu+\tfrac{3}{2})^2} - (\nu+\tfrac{3}{2})$$
$$-(\nu+\tfrac{1}{2})\Bigg[\Bigg(\frac{(\nu+\frac{1}{2})+\sqrt{x^2+(\nu+\frac{3}{2})^2}}{2(\nu+1)} - 1\Bigg)$$
$$-\frac{1}{2}\Bigg(\frac{(\nu+\frac{1}{2})+\sqrt{x^2+(\nu+\frac{3}{2})^2}}{2(\nu+1)} - 1\Bigg)^2\Bigg] - \frac{x^2}{4\nu} + \frac{x^4}{32\nu^3} + \frac{x^2}{4\nu^2},$$

$$G_\nu^{\mathrm{up},1}(x) := \sqrt{x^2+(\nu+2)^2}-(\nu+2)$$
$$-\nu\Bigg[\bigg(\frac{\nu+\sqrt{x^2+(\nu+2)^2}}{2(\nu+1)}-1\bigg)$$
$$-\frac{1}{2}\bigg(\frac{\nu+\sqrt{x^2+(\nu+2)^2}}{2(\nu+1)}-1\bigg)^2\Bigg]-\frac{x^2}{4\nu}+\frac{x^4}{32\nu^3}+\frac{x^2}{4\nu^2},$$

$$G_\nu^{\mathrm{low},2}(x) := -\frac{\nu+\frac{1}{2}}{3(c^{\mathrm{low}})^3}\bigg(\frac{(\nu+\frac{1}{2})+\sqrt{x^2+(\nu+\frac{3}{2})^2}}{2(\nu+1)}-1\bigg)^3,$$

and

$$G_\nu^{\mathrm{up},2}(x) := -\frac{\nu}{3(c^{\mathrm{up}})^3}\bigg(\frac{\nu+\sqrt{x^2+(\nu+2)^2}}{2(\nu+1)}-1\bigg)^3.$$

Routine yet tedious computations allow to show that

$$G_\nu^{\mathrm{low},1}(x) = \frac{x^2}{4\nu^2(\nu+1)}+\frac{(4\nu^2+5\nu+2)x^4}{32\nu^3(\nu+1)^2(\nu+2)}$$
$$+\Bigg(1-\frac{4(1+\frac{2}{\nu})(1+\frac{3}{2\nu})}{\Big((1+\frac{3}{2\nu})+\sqrt{(\frac{x}{\nu})^2+(1+\frac{3}{2\nu})^2}\Big)^2}\Bigg)\frac{x^4}{32(\nu+1)^2(\nu+2)} \tag{A.24}$$

and

$$G_\nu^{\mathrm{up},1}(x) = \frac{x^2}{4\nu^2(\nu+1)}+\frac{(4\nu^2+5\nu+2)x^4}{32\nu^3(\nu+1)^2(\nu+2)}$$
$$+\Bigg(1-\frac{4}{\Big(1+\sqrt{(\frac{x}{\nu+2})^2+1}\Big)^2}\Bigg)\frac{x^4}{32(\nu+1)^2(\nu+2)}. \tag{A.25}$$

Since both c^{low} and c^{up} are larger than one, we easily obtain

$$|G_\nu^{\mathrm{low},2}(x)| \leq \Big((1+\tfrac{3}{2\nu})+\sqrt{(\tfrac{x}{\nu})^2+(1+\tfrac{3}{2\nu})^2}\Big)^{-3}\frac{(\nu+\frac{1}{2})x^6}{24\nu^3(\nu+1)^3} \tag{A.26}$$

and

$$|G_\nu^{\mathrm{up},2}(x)| \leq \Big((1+\tfrac{2}{\nu})+\sqrt{(\tfrac{x}{\nu})^2+(1+\tfrac{2}{\nu})^2}\Big)^{-3}\frac{x^6}{24\nu^2(\nu+1)^3}. \tag{A.27}$$

Using the mean value theorem to control the last term in the righthand sides of (A.24)–(A.25), it directly follows from (A.24)–(A.27) that, under the assumptions of the lemma,

$$a_n^2 G_{\nu_n}^{\mathrm{low/up},1}\Big(\frac{b_nT_n}{a_n}\Big)=o_{\mathrm{P}}(1) \quad \text{and} \quad a_n^2 G_{\nu_n}^{\mathrm{low/up},2}\Big(\frac{b_nT_n}{a_n}\Big)=o_{\mathrm{P}}(1),$$

which proves (A.23), hence establishes the result. □

Proof of Theorem 1. Throughout the proof, distributions and expectations are under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$ and stochastic convergences are as $n\to\infty$ under the same sequence of hypotheses. By using the fact that $\mathbf{O}\boldsymbol{\theta}_{n0}=\boldsymbol{\theta}_{n0}=\mathbf{O}'\boldsymbol{\theta}_{n0}$ for any $\mathbf{O}\in SO_{p_n}(\boldsymbol{\theta}_{n0})$ and by decomposing $\boldsymbol{\tau}_n$ into $(\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0})\boldsymbol{\theta}_{n0} + \Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n$, with $\Pi_{\boldsymbol{\theta}_{n0}} := \mathbf{I}_{p_n} - \boldsymbol{\theta}_{n0}\boldsymbol{\theta}_{n0}'$, (2.7) yields

$$\frac{d\mathrm{P}^{(n)\mathbf{T}_n}_{1-\frac{1}{2}\nu_n^2\|\boldsymbol{\tau}_n\|^2,\kappa_n}}{dm_n} = \frac{c^n_{p_n,\kappa_n}}{\omega^n_{p_n-1}} \int_{SO_{p_n}(\boldsymbol{\theta}_{n0})} \exp\big(n\kappa_n \bar{\mathbf{X}}_n'\mathbf{O}'(\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n)\big)\, d\mathbf{O}$$

$$= \frac{c^n_{p_n,\kappa_n}}{\omega^n_{p_n-1}} \exp\big(n\kappa_n \bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0}\big) \int_{SO_{p_n}(\boldsymbol{\theta}_{n0})} \exp\big(n\kappa_n\nu_n \bar{\mathbf{X}}_n'\mathbf{O}'\boldsymbol{\tau}_n\big)\, d\mathbf{O}$$

$$= \frac{c^n_{p_n,\kappa_n}}{\omega^n_{p_n-1}} \exp\big(n\kappa_n \bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0}\big) \int_{SO_{p_n}(\boldsymbol{\theta}_{n0})} \exp\big(n\kappa_n\nu_n \bar{\mathbf{X}}_n'[(\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0})\boldsymbol{\theta}_{n0} + \mathbf{O}'\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n]\big)\, d\mathbf{O}$$

$$= \frac{c^n_{p_n,\kappa_n}}{\omega^n_{p_n-1}} \exp\big(n\kappa_n (1+\nu_n(\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0}))\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0}\big) \int_{SO_{p_n}(\boldsymbol{\theta}_{n0})} \exp\big(n\kappa_n\nu_n \bar{\mathbf{X}}_n'\mathbf{O}'\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\big)\, d\mathbf{O}.$$

Now, since $\mathbf{O}'\Pi_{\boldsymbol{\theta}_{n0}} = \mathbf{O}'\Pi^2_{\boldsymbol{\theta}_{n0}} = \Pi_{\boldsymbol{\theta}_{n0}}\mathbf{O}'\Pi_{\boldsymbol{\theta}_{n0}}$,

$$\int_{SO_{p_n}(\boldsymbol{\theta}_{n0})} \exp\big(n\kappa_n\nu_n \bar{\mathbf{X}}_n'\mathbf{O}'\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\big)\, d\mathbf{O}$$

$$= \int_{SO_{p_n}(\boldsymbol{\theta}_{n0})} \exp\big(n\kappa_n\nu_n \bar{\mathbf{X}}_n'\Pi_{\boldsymbol{\theta}_{n0}}\mathbf{O}'\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\big)\, d\mathbf{O}$$

$$= \int_{SO_{p_n}(\boldsymbol{\theta}_{n0})} \exp\bigg(n\kappa_n\nu_n \|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|\|\Pi_{\boldsymbol{\theta}_{n0}}\bar{\mathbf{X}}_n\|\Big(\frac{\Pi_{\boldsymbol{\theta}_{n0}}\bar{\mathbf{X}}_n}{\|\Pi_{\boldsymbol{\theta}_{n0}}\bar{\mathbf{X}}_n\|}\Big)'\Big(\mathbf{O}'\frac{\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n}{\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|}\Big)\bigg)\, d\mathbf{O}$$

$$= \mathrm{E}\big[\exp\big(n\kappa_n\nu_n \|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|\|\Pi_{\boldsymbol{\theta}_{n0}}\bar{\mathbf{X}}_n\|\mathbf{v}_n'\mathbf{S}\big) \,|\, \mathbf{X}_{n1},\ldots,\mathbf{X}_{nn}\big],$$

where $\mathbf{S}$ is uniformly distributed over $\mathcal{S}^{p_n-1}_{\boldsymbol{\theta}_{n0}} := \{\mathbf{x}\in\mathcal{S}^{p_n-1} : \mathbf{x}'\boldsymbol{\theta}_{n0} = 0\}$ and where $\mathbf{v}_n \in \mathcal{S}^{p_n-1}_{\boldsymbol{\theta}_{n0}}$ is arbitrary. Since $\mathbf{v}_n'\mathbf{S}$ has density $t\mapsto c_{p_n-1}(1-t^2)^{\frac{p_n-4}{2}}\,\mathbb{I}[t\in[-1,1]]$, with $c_{p_n-1} = 1/\int_{-1}^1 (1-t^2)^{\frac{p_n-4}{2}}\, dt$, this yields

$$\int_{SO_{p_n}(\boldsymbol{\theta}_{n0})} \exp\big(n\kappa_n\nu_n \bar{\mathbf{X}}_n'\mathbf{O}'\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\big)\, d\mathbf{O}$$

$$= c_{p_n-1}\int_{-1}^1 \exp\big(n\kappa_n\nu_n \|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|\|\Pi_{\boldsymbol{\theta}_{n0}}\bar{\mathbf{X}}_n\| t\big)(1-t^2)^{\frac{p_n-4}{2}}\, dt$$

$$= H_{\frac{p_n-3}{2}}(n\kappa_n\nu_n \|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|\|\Pi_{\boldsymbol{\theta}_{n0}}\bar{\mathbf{X}}_n\|).$$

Summing up,

$$\frac{d\mathrm{P}^{(n)\mathbf{T}_n}_{1-\frac{1}{2}\nu_n^2\|\boldsymbol{\tau}_n\|^2,\kappa_n}}{dm_n} = \frac{c^n_{p_n,\kappa_n}}{\omega^n_{p_n-1}} \exp\big(n\kappa_n \bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0}\big)$$
$$\times \exp\big(n\kappa_n\nu_n(\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0})\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0}\big) H_{\frac{p_n-3}{2}}(n\kappa_n\nu_n \|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|\|\Pi_{\boldsymbol{\theta}_{n0}}\bar{\mathbf{X}}_n\|). \quad \text{(A.28)}$$

Now, with the quantity L_n introduced in Lemma 1, we have

$$T_n := 1 + \frac{\sqrt{2}\,\widetilde{W}_n}{\sqrt{p_n-1}} = \frac{W_n}{p_n-1} = \frac{n^2}{\sum_{i=1}^n V_{ni}^2}\bar{\mathbf{X}}_n'(\mathbf{I}_{p_n} - \boldsymbol{\theta}_{n0}\boldsymbol{\theta}_{n0}')\bar{\mathbf{X}}_n$$

$$= \frac{n}{f_{n2}L_n}\|\Pi_{\boldsymbol{\theta}_{n0}}\bar{\mathbf{X}}_n\|^2 = \frac{n\kappa_n}{(p_n-1)e_{n1}L_n}\|\Pi_{\boldsymbol{\theta}_{n0}}\bar{\mathbf{X}}_n\|^2,$$

where we used the identity $f_{n2} = (p_n - 1)e_{n1}/\kappa_n$; see (2.9). Therefore, (A.28) yields

$$\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n} = \log \frac{d\mathrm{P}^{(n)\mathbf{T}_n}_{1-\frac{1}{2}\nu_n^2\|\boldsymbol{\tau}_n\|^2,\kappa_n}}{d\mathrm{P}^{(n)\mathbf{T}_n}_{1,\kappa_n}}$$

$$= n\kappa_n\nu_n(\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0})\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} + \log H_{\frac{p_n-3}{2}}(n\kappa_n\nu_n\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|\|\Pi_{\boldsymbol{\theta}_{n0}}\bar{\mathbf{X}}_n\|)$$

$$= n\kappa_n\nu_n(\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0})\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} + \log H_{\frac{p_n-3}{2}}(n^{1/2}(p_n-1)^{1/2}\kappa_n^{1/2}\nu_n e_{n1}^{1/2}\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|L_n^{1/2}T_n^{1/2}).$$

Since $\widetilde{W}_n = \sqrt{(p_n-1)/2} \times (T_n - 1)$ is asymptotically standard normal (Lemma 2), we have that $T_n = 1 + o_\mathrm{P}(1)$. Moreover, it directly follows from Lemma 1 that $L_n = 1 + o_\mathrm{P}(1)$. Consequently, Lemma 3 shows that, if ν_n satisfies (2.10), then

$$\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n} = n\kappa_n\nu_n(\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0})\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} + \frac{p_n-1}{2(p_n-3)}n\kappa_n\nu_n^2 e_{n1}\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|^2 L_n T_n$$
$$- \frac{(p_n-1)^2}{4(p_n-3)^3}n^2\kappa_n^2\nu_n^4 e_{n1}^2\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|^4 L_n^2 T_n^2 + o_\mathrm{P}(1).$$

Using (2.10), Lemma 1 and the fact that $T_n = 1 + o_\mathrm{P}(1)$ yields

$$\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n} = n\kappa_n\nu_n(\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0})\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} + \frac{1}{2}n\kappa_n\nu_n^2 e_{n1}\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|^2 T_n - \frac{n^2\kappa_n^2\nu_n^4 e_{n1}^2}{4p_n}\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|^4 + o_\mathrm{P}(1). \tag{A.29}$$

Using the definitions of Z_n and T_n, we obtain

$$\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n} = n\kappa_n\nu_n(\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0})\Big(e_{n1} + \frac{\sqrt{\tilde{e}_{n2}}}{n^{1/2}}Z_n\Big)$$
$$+\frac{1}{2}n\kappa_n\nu_n^2 e_{n1}\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|^2\Big(1 + \frac{\sqrt{2}\,\widetilde{W}_n}{\sqrt{p_n-1}}\Big) - \frac{n^2\kappa_n^2\nu_n^4 e_{n1}^2}{4p_n}\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|^4 + o_\mathrm{P}(1)$$
$$= \sqrt{n}\kappa_n\nu_n\sqrt{\tilde{e}_{n2}}(\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0})Z_n + \frac{n\kappa_n\nu_n^2 e_{n1}}{\sqrt{2}(p_n-1)^{1/2}}\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|^2\widetilde{W}_n$$
$$+n\kappa_n\nu_n(\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0})e_{n1} + \frac{1}{2}n\kappa_n\nu_n^2 e_{n1}\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|^2 - \frac{n^2\kappa_n^2\nu_n^4 e_{n1}^2}{4p_n}\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|^4 + o_\mathrm{P}(1).$$

Using the identities

$$\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0} = -\frac{1}{2}\nu_n\|\boldsymbol{\tau}_n\|^2 \tag{A.30}$$

and

$$\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|^2 = \|\boldsymbol{\tau}_n\|^2 - (\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0})^2 = \|\boldsymbol{\tau}_n\|^2\Big(1 - \frac{1}{4}\nu_n^2\|\boldsymbol{\tau}_n\|^2\Big) \tag{A.31}$$

provides

$$\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n} = -\frac{1}{2}\sqrt{n}\kappa_n\nu_n^2\sqrt{\tilde{e}_{n2}}\|\boldsymbol{\tau}_n\|^2 Z_n$$
$$+\frac{n\kappa_n\nu_n^2 e_{n1}}{\sqrt{2}(p_n-1)^{1/2}}\|\boldsymbol{\tau}_n\|^2\Big(1 - \frac{1}{4}\nu_n^2\|\boldsymbol{\tau}_n\|^2\Big)\widetilde{W}_n - \frac{1}{2}n\kappa_n\nu_n^2 e_{n1}\|\boldsymbol{\tau}_n\|^2$$
$$+\frac{1}{2}n\kappa_n\nu_n^2 e_{n1}\|\boldsymbol{\tau}_n\|^2\Big(1 - \frac{1}{4}\nu_n^2\|\boldsymbol{\tau}_n\|^2\Big) - \frac{n^2\kappa_n^2\nu_n^4 e_{n1}^2}{4p_n}\|\boldsymbol{\tau}_n\|^4\Big(1 - \frac{1}{4}\nu_n^2\|\boldsymbol{\tau}_n\|^2\Big)^2 + o_\mathrm{P}(1).$$

The result then follows from (2.10) and from the tightness of $\widetilde{W}_n$ (Lemma 2). □

Lemma 4 *Let (p_n) be a sequence of integers that diverges to infinity and (κ_n) be an arbitrary sequence in $(0,\infty)$. Let e_{n1} (resp., $\tilde{e}_{n2}$) be the expectation (resp., the variance) of the distribution with probability density function (2.8). Then, we have the following: (i) if $\kappa_n/p_n \to \infty$, then*

$$e_{n1} = 1 + o(1), \quad \tilde{e}_{n2} = O\Big(\frac{p_n}{\kappa_n^2}\Big) \quad \text{and} \quad f_{n2} = \frac{p_n}{\kappa_n} + o\Big(\frac{p_n}{\kappa_n}\Big);$$

(ii) if $\kappa_n/p_n \to \xi > 0$, then, letting $c_\xi := \frac{1}{2} + \sqrt{\frac{1}{4} + \xi^2}$,

$$e_{n1} \to \frac{\xi}{c_\xi} + o(1), \quad \tilde{e}_{n2} = O\Big(\frac{1}{p_n}\Big) \quad \text{and} \quad f_{n2} = \frac{1}{c_\xi} + o(1);$$

(iii) if $\kappa_n/p_n \to 0$, then

$$e_{n1} = \frac{\kappa_n}{p_n} + O\Big(\frac{\kappa_n^3}{p_n^3}\Big), \quad \tilde{e}_{n2} = \frac{1}{p_n} + o\Big(\frac{1}{p_n}\Big) \quad \text{and} \quad f_{n2} = 1 + o(1).$$

Proof of Lemma 4. Denoting again as $\mathcal{I}_\nu(\cdot)$ the order-ν modified Bessel function of the first kind, we recall (see (2.9)) that

$$e_{n1} = \frac{\mathcal{I}_{\frac{p_n}{2}}(\kappa_n)}{\mathcal{I}_{\frac{p_n}{2}-1}(\kappa_n)}, \quad \tilde{e}_{n2} = 1 - \frac{p_n - 1}{\kappa_n} e_{n1} - e_{n1}^2 \quad \text{and} \quad f_{n2} = \frac{p_n - 1}{\kappa_n} e_{n1}.$$

In each case (i)–(iii), the claim for f_{n2} directly follows from the result on e_{n1}, so that it is sufficient to prove the results for e_{n1} and $\tilde{e}_{n2}$. To do so, we will use the bounds

$$R_\nu^{\text{low}}(z) := \frac{z}{\nu + 1 + \sqrt{(\nu+1)^2 + z^2}} \le \frac{\mathcal{I}_{\nu+1}(z)}{\mathcal{I}_\nu(z)} \le \frac{z}{\nu + \sqrt{(\nu+2)^2 + z^2}} =: R_\nu^{\text{up}}(z) \quad \text{(A.32)}$$

and

$$\tilde{R}_\nu^{\text{low}}(z) := \frac{z}{\nu + \frac{1}{2} + \sqrt{(\nu + \frac{3}{2})^2 + z^2}} \le \frac{\mathcal{I}_{\nu+1}(z)}{\mathcal{I}_\nu(z)}; \quad \text{(A.33)}$$

see (11) and (16) in [1], respectively. (i) The lower bound in (A.32) provides

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

which, since $e_{n1} \le 1$, establishes the result for e_{n1}. Making use of the bound in (A.33), we can write

$$\tilde{e}_{n2} \le 1 - \frac{p_n - 1}{\kappa_n} \tilde{R}^{\text{low}}_{\frac{p_n}{2}-1}(\kappa_n) - \big(\tilde{R}^{\text{low}}_{\frac{p_n}{2}-1}(\kappa_n)\big)^2.$$

Lengthy yet quite straightforward computations allow to rewrite this as

$$\tilde{e}_{n2} \le \frac{p_n}{\kappa_n^2 \Big(\frac{p_n - 1}{2\kappa_n} + \sqrt{\big(\frac{p_n+1}{2\kappa_n}\big)^2 + 1}\Big)^2}.$$

It readily follows that $\kappa_n^2 \tilde{e}_{n2}/p_n$ is $O(1)$, as was to be showed. Let us turn to the proof of (iii). The bounds in (A.32) readily yield

$$\frac{1}{\frac{1}{2} + \sqrt{\big(\frac{1}{2}\big)^2 + \big(\frac{\kappa_n}{p_n}\big)^2}} \le \frac{e_{n1}}{\kappa_n/p_n} \le \frac{1}{\frac{1}{2} - \frac{1}{p_n} + \sqrt{\big(\frac{1}{2} + \frac{1}{p_n}\big)^2 + \big(\frac{\kappa_n}{p_n}\big)^2}}, \quad \text{(A.34)}$$

which provides

$$\frac{-\big(\frac{\kappa_n}{p_n}\big)^2}{\Big(\frac{1}{2} + \sqrt{\big(\frac{1}{2}\big)^2 + \big(\frac{\kappa_n}{p_n}\big)^2}\Big)^2} \le \frac{e_{n1}}{\kappa_n/p_n} - 1 \le \frac{-\big(\frac{\kappa_n}{p_n}\big)^2}{\frac{1}{2} - \frac{1}{p_n} + \sqrt{\big(\frac{1}{2} + \frac{1}{p_n}\big)^2 + \big(\frac{\kappa_n}{p_n}\big)^2}}.$$

This proves the result for e_{n1}. Turning to $\tilde{e}_{n2}$, the bounds in (A.32) lead to

$$1-\frac{p_n-1}{\kappa_n}R^{\rm up}_{\frac{p_n}{2}-1}(\kappa_n)-\big(R^{\rm up}_{\frac{p_n}{2}-1}(\kappa_n)\big)^2\leq \tilde{e}_{n2}\leq 1-\frac{p_n-1}{\kappa_n}R^{\rm low}_{\frac{p_n}{2}-1}(\kappa_n)-\big(R^{\rm low}_{\frac{p_n}{2}-1}(\kappa_n)\big)^2.$$

As above, heavy but rather straightforward computations allow to rewrite this as

$$\frac{\frac{3}{2}+\frac{1}{p_n}-\sqrt{\big(\frac{1}{2}+\frac{1}{p_n}\big)^2+\big(\frac{\kappa_n}{p_n}\big)^2}}{p_n\Big(\frac{1}{2}-\frac{1}{p_n}+\sqrt{\big(\frac{1}{2}+\frac{1}{p_n}\big)^2+\big(\frac{\kappa_n}{p_n}\big)^2}\Big)^2}\leq \tilde{e}_{n2}\leq \frac{1}{p_n\Big(\frac{1}{2}+\sqrt{\big(\frac{1}{2}\big)^2+\big(\frac{\kappa_n}{p_n}\big)^2}\Big)}, \tag{A.35}$$

which establishes that $p_n\tilde{e}_{n2}=1+o(1)$. Finally, the result in (ii) readily follows from (A.34) and from the upper bound in (A.35). □

Proof of Theorem 2. Stochastic convergences throughout the proof are as $n\to\infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$. Assume that (i) $\kappa_n/p_n\to\infty$, (ii) $\kappa_n/p_n\to\xi>0$, or (iii) $\kappa_n/p_n\to 0$ with $\sqrt{n}\kappa_n/p_n\to\infty$, and let (ν_n) be the corresponding sequence in the statement of the theorem. Using Lemma 4 and the identity $\kappa_n f_{n2}=(p_n-1)e_{n1}$, it is then easy to check that ν_n satisfies (2.10), is such that $\nu_n=o(1)$, and is asymptotically equivalent to $\tilde{\nu}_n=p_n^{3/4}/(\sqrt{n}\kappa_n\sqrt{f_{n2}})$ in the sense that $\tilde{\nu}_n/\nu_n\to 1$. Theorem 1 thus applies and yields

$$\begin{aligned}\Lambda^{(n)\rm inv}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n}&=-\frac{p_n^{3/2}\sqrt{\tilde{e}_{n2}}}{2\sqrt{n}\kappa_n f_{n2}}\|\boldsymbol{\tau}_n\|^2 Z_n+\frac{p_n e_{n1}}{\sqrt{2}\kappa_n f_{n2}}\|\boldsymbol{\tau}_n\|^2\widetilde{W}_n\\&\qquad-\frac{p_n^3 e_{n1}}{8n\kappa_n^3 f_{n2}^2}\|\boldsymbol{\tau}_n\|^4-\frac{p_n^2 e_{n1}^2}{4\kappa_n^2 f_{n2}^2}\|\boldsymbol{\tau}_n\|^4+o_{\rm P}(1).\end{aligned}$$

Using again the identity $\kappa_n f_{n2}=(p_n-1)e_{n1}$, we then obtain

$$\begin{aligned}\Lambda^{(n)\rm inv}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n}&=-\frac{p_n^{3/2}\sqrt{\tilde{e}_{n2}}}{2\sqrt{n}(p_n-1)e_{n1}}\|\boldsymbol{\tau}_n\|^2 Z_n+\frac{1}{\sqrt{2}}\|\boldsymbol{\tau}_n\|^2\widetilde{W}_n\\&\qquad-\frac{p_n^3}{8n\kappa_n(p_n-1)^2 e_{n1}}\|\boldsymbol{\tau}_n\|^4-\frac{1}{4}\|\boldsymbol{\tau}_n\|^4+o_{\rm P}(1).\end{aligned} \tag{A.36}$$

The result in cases (i)–(iii) then follows from the fact that Lemma 4 implies that, in each case, the first and third term of the righthand side of (A.36) are $o_{\rm P}(1)$.

We turn to case (iv), for which $\sqrt{n}\kappa_n/p_n=\xi$ (so that, like for all subsequent cases, $\kappa_n=o(p_n)$). Then, the same argument as above allows to check that $\nu_n=p_n^{3/4}/(\sqrt{n}\kappa_n\sqrt{f_{n2}})$ still satisfies (2.10) and is such that $\nu_n=o(1)$, so that, jointly with Lemma 4, Theorem 1 provides

$$\begin{aligned}\Lambda^{(n)\rm inv}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n}&=-\frac{p_n^{3/2}\tilde{e}_{n2}^{1/2}}{2\sqrt{n}\kappa_n f_{n2}}\|\boldsymbol{\tau}_n\|^2 Z_n+\frac{p_n e_{n1}}{\sqrt{2}\kappa_n f_{n2}}\|\boldsymbol{\tau}_n\|^2\widetilde{W}_n\\&\qquad-\frac{p_n^3 e_{n1}}{8n\kappa_n^3 f_{n2}^2}\|\boldsymbol{\tau}_n\|^4-\frac{p_n^2 e_{n1}^2}{4\kappa_n^2 f_{n2}^2}\|\boldsymbol{\tau}_n\|^4+o_{\rm P}(1)\\&=-\frac{1}{2\xi}\|\boldsymbol{\tau}_n\|^2 Z_n+\frac{1}{\sqrt{2}}\|\boldsymbol{\tau}_n\|^2\widetilde{W}_n-\frac{1}{8\xi^2}\|\boldsymbol{\tau}_n\|^4-\frac{1}{4}\|\boldsymbol{\tau}_n\|^4+o_{\rm P}(1),\end{aligned}$$

as was to be shown. Consider now case (v), under which $\sqrt{n}\kappa_n/p_n\to 0$ with $\sqrt{n}\kappa_n/\sqrt{p_n}\to\infty$, which still ensures that $\nu_n=p_n^{1/4}/(n^{1/4}\sqrt{\kappa_n})$ is $o(1)$ and satisfies (2.10). Theorem 1 applies and, by using Lemma 4 again, yields

$$\begin{aligned}\Lambda^{(n)\rm inv}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n}&=-\frac{1}{2}\sqrt{p_n\tilde{e}_{n2}}\|\boldsymbol{\tau}_n\|^2 Z_n+\frac{\sqrt{n}e_{n1}}{\sqrt{2}}\|\boldsymbol{\tau}_n\|^2\widetilde{W}_n-\frac{p_n e_{n1}}{8\kappa_n}\|\boldsymbol{\tau}_n\|^4-\frac{n e_{n1}^2}{4}\|\boldsymbol{\tau}_n\|^4+o_{\rm P}(1)\\&=-\frac{1}{2}\|\boldsymbol{\tau}_n\|^2 Z_n-\frac{1}{8}\|\boldsymbol{\tau}_n\|^4+o_{\rm P}(1),\end{aligned}$$

which establishes the result in case (v). If $\sqrt{n}\kappa_n/\sqrt{p_n} = \xi$ (case (vi)), then Lemma 4 implies that $\nu_n = 1$ satisfies (2.10). Theorem 1 then provides

$$\begin{aligned}\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n} &= -\frac{\sqrt{n}\kappa_n}{2\sqrt{p_n}}\sqrt{p_n\tilde{e}_{n2}}\,\|\boldsymbol{\tau}_n\|^2 Z_n + \frac{n\kappa_n e_{n1}}{\sqrt{2}p_n^{1/2}}\|\boldsymbol{\tau}_n\|^2\Big(1-\frac{1}{4}\|\boldsymbol{\tau}_n\|^2\Big)\widetilde{W}_n \\ &\quad -\frac{1}{8}n\kappa_n e_{n1}\|\boldsymbol{\tau}_n\|^4 - \frac{n^2\kappa_n^2 e_{n1}^2}{4p_n}\|\boldsymbol{\tau}_n\|^4\Big(1-\frac{1}{4}\|\boldsymbol{\tau}_n\|^2\Big)^2 + o_{\mathrm{P}}(1) \qquad \text{(A.37)} \\ &= -\frac{\xi}{2}\|\boldsymbol{\tau}_n\|^2 Z_n - \frac{\xi^2}{8}\|\boldsymbol{\tau}_n\|^4 + o_{\mathrm{P}}(1),\end{aligned}$$

where we used Lemma 4. Finally, if $\sqrt{n}\kappa_n/\sqrt{p_n} = o(1)$ (case (vii)), then (2.10) again holds for $\nu_n = 1$. Therefore, Theorem 1 shows that $\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_n}$ satisfies the first equality of (A.37), hence is $o_{\mathrm{P}}(1)$. □

Proof of Theorem 3. First note that, since $p_n = o(n^2)$, Lemma 4(iii) entails that

$$\begin{aligned}Z_n &= \frac{\sqrt{n}(\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} - e_{n1})}{\sqrt{\tilde{e}_{2n}}} = \frac{\sqrt{n}(\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} - \xi/\sqrt{n} + O(n^{-3/2}))}{\sqrt{\frac{1}{p_n} + o(\frac{1}{p_n})}} \\ &= \frac{\sqrt{p_n}(\sqrt{n}\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} - \xi + O(1/n))}{\sqrt{1+o(1)}} = \sqrt{np_n}\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} - \sqrt{p_n}\xi + o_{\mathrm{P}}(1) \qquad \text{(A.38)}\end{aligned}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$. Write then

$$\begin{aligned}\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n,\kappa_{n,s}/\boldsymbol{\theta}_{n0},\kappa_n} &= \Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n,\kappa_{n,s}/\boldsymbol{\theta}_{n0},\kappa_{n,s}} + \Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_{n0},\kappa_{n,s}/\boldsymbol{\theta}_{n0},\kappa_n} \\ &= \Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n,\kappa_{n,s}/\boldsymbol{\theta}_{n0},\kappa_{n,s}} + \log\frac{d\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_{n,s}}}{d\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}} \\ &=: L_{n1} + L_{n2}.\end{aligned}$$

Using (A.38), we obtain

$$\begin{aligned}L_{n2} &= n\big(\log(c_{p_n,\kappa_{n,s}}) - \log(c_{p_n,\kappa_n})\big) + n(\kappa_{n,s} - \kappa_n)\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} \\ &= n\Big[\log\Big(\frac{c_{p_n,0}}{c_{p_n,\kappa_n}}\Big) - \log\Big(\frac{c_{p_n,0}}{c_{p_n,\kappa_{n,s}}}\Big)\Big] + s\sqrt{np_n}\,\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} \\ &= n\Big[\log H_{\frac{p_n}{2}-1}(\kappa_n) - \log H_{\frac{p_n}{2}-1}(\kappa_{n,s})\Big] + s\sqrt{p_n}\xi + sZ_n + o_{\mathrm{P}}(1) \\ &=: \tilde{L}_{n2} + sZ_n + o_{\mathrm{P}}(1)\end{aligned}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$. Since $p_n = o(n^2)$, we can apply Lemma 3 with $a_n = \sqrt{n}$ and $T_n \equiv 1$. This yields

$$\begin{aligned}\tilde{L}_{n2} &= \Big(\frac{n\kappa_n^2}{4(\frac{p_n}{2}-1)} - \frac{n\kappa_n^4}{32(\frac{p_n}{2}-1)^3}\Big) - \Big(\frac{n\kappa_{n,s}^2}{4(\frac{p_n}{2}-1)} - \frac{n\kappa_{n,s}^4}{32(\frac{p_n}{2}-1)^3}\Big) + s\sqrt{p_n}\xi + o(1) \\ &= -\frac{n(\kappa_{n,s}^2 - \kappa_n^2)}{2p_n - 4} + \frac{n(\kappa_{n,s}^4 - \kappa_n^4)}{4(p_n-2)^3} + s\sqrt{p_n}\xi + o(1) \\ &= -\frac{p_n^2((\xi + s/\sqrt{p_n})^2 - \xi^2)}{2p_n - 4} + \frac{p_n^4((\xi + s/\sqrt{p_n})^4 - \xi^4)}{4n(p_n-2)^3} + s\sqrt{p_n}\xi + o(1) \\ &= -\frac{s^2}{2} + o(1)\end{aligned}$$

as $n \to \infty$. Therefore,

$$L_{n2} = sZ_n - \frac{s^2}{2} + o_{\rm P}(1) \tag{A.39}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$, which implies that the sequences of probability measures $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_{n,s}}$ and $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$ are mutually contiguous (this results from the Le Cam first lemma).

Now, denote as $e_{n1,s}$ and $\tilde{e}_{2n,s}$, respectively, the values of e_{n1} and $\tilde{e}_{2n}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_{n,s}}$. Then, proceeding as in (A.38) and using the fact that contiguity implies that (A.38) also holds under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_{n,s}}$, one obtains

$$Z_{n,s} := \frac{\sqrt{n}(\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} - e_{n1,s})}{\sqrt{\tilde{e}_{2n,s}}} = \sqrt{np_n}\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} - \sqrt{p_n}\xi - s + o_{\rm P}(1) = Z_n - s + o_{\rm P}(1)$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_{n,s}}$. Consequently, Theorem 2(iv) implies that

$$\begin{aligned} L_{1n} =& \|\boldsymbol{\tau}_n\|^2\left(\frac{\widetilde{W}_n}{\sqrt{2}} - \frac{Z_{n,s}}{2\xi}\right) - \frac{1}{2}\|\boldsymbol{\tau}_n\|^4\left(\frac{1}{2} + \frac{1}{4\xi^2}\right) + o_{\rm P}(1) \\ =& \|\boldsymbol{\tau}_n\|^2\left(\frac{\widetilde{W}_n}{\sqrt{2}} - \frac{Z_n}{2\xi}\right) + \frac{\|\boldsymbol{\tau}_n\|^2 s}{2\xi} - \frac{1}{2}\|\boldsymbol{\tau}_n\|^4\left(\frac{1}{2} + \frac{1}{4\xi^2}\right) + o_{\rm P}(1) \end{aligned} \tag{A.40}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_{n,s}}$, hence, from contiguity, also under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$. Combining (A.44) and (A.40) establishes the local asymptotic quadraticity result in (2.18). Finally, the asymptotic normality result of $\boldsymbol{\Delta}_n$ trivially follows from Lemma 2. □

The proof of Theorem 4 requires both following lemmas.

Lemma 5 *Let (p_n) be a sequence of integers that diverges to infinity. Let (κ_n) and (κ_{n*}) be sequences in $(0,\infty)$ that are $o(p_n)$ and write e_{n1} and $\tilde{e}_{n2}$ (resp., e_{n1*} and $\tilde{e}_{n2*}$) for the corresponding moments based on κ_n (resp., on κ_{n*}). Let $(\boldsymbol{\theta}_{n0})$, (ν_n) and $(\boldsymbol{\tau}_n)$ be sequences such that $\boldsymbol{\theta}_{n0}$ and $\boldsymbol{\theta}_n = \boldsymbol{\theta}_{n0} + \nu_n\boldsymbol{\tau}_n$ belong to $\mathcal{S}^{p_n-1}$ for any n, with $(\boldsymbol{\tau}_n)$ bounded and (ν_n) such that*

$$\nu_n^2 = O\Big(\frac{\sqrt{p_n}}{n\kappa_{n*}e_{n1*}}\Big). \tag{A.41}$$

Then, with the same Z_n and $\widetilde{W}_n$ as in Theorem 1, we have that

$$\begin{aligned} \Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_{n*}} =& -\frac{1}{2}\sqrt{n\kappa_{n*}\nu_n^2}\sqrt{\tilde{e}_{n2}}\,\|\boldsymbol{\tau}_n\|^2 Z_n + \frac{n\kappa_{n*}\nu_n^2 e_{n1*}}{\sqrt{2}p_n^{1/2}}\|\boldsymbol{\tau}_n\|^2\Big(1 - \frac{1}{4}\nu_n^2\|\boldsymbol{\tau}_n\|^2\Big)\widetilde{W}_n \\ & -\frac{1}{8}n\kappa_{n*}\nu_n^4 e_{n1*}\|\boldsymbol{\tau}_n\|^4 - \frac{n^2\kappa_{n*}^2\nu_n^4 e_{n1*}^2}{4p_n}\|\boldsymbol{\tau}_n\|^4\Big(1 - \frac{1}{4}\nu_n^2\|\boldsymbol{\tau}_n\|^2\Big)^2 \\ & +\frac{1}{2}n\kappa_{n*}\nu_n^2(e_{n1*} - e_{n1})\|\boldsymbol{\tau}_n\|^2 + o_{\rm P}(1), \end{aligned}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$.

Proof of Lemma 5. Since κ_n and κ_{n*} are both $o(n)$, Lemma 4 ensures that $f_{n2}/f_{n2*} = 1 + o(1)$, where f_{n2*} denotes the quantity f_{n2} based on κ_{n*}. Using this, it can be showed along the exact same lines as in the proof of (A.29) in Theorem 1 that, as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$,

$$\begin{aligned} \Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n/\boldsymbol{\theta}_{n0};\kappa_{n*}} =& n\kappa_{n*}\nu_n(\boldsymbol{\tau}_n'\boldsymbol{\theta}_{n0})\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} + \frac{1}{2}n\kappa_{n*}\nu_n^2 e_{n1*}\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|^2 T_n \\ & -\frac{n^2\kappa_{n*}^2\nu_n^4 e_{n1*}^2}{4p_n}\|\Pi_{\boldsymbol{\theta}_{n0}}\boldsymbol{\tau}_n\|^4 + o_{\rm P}(1), \end{aligned}$$

where $T_n := 1+\sqrt{2}\,\widetilde{W}_n/\sqrt{p_n-1}$. If one replaces T_n by this expression and $\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0}$ by $e_{n1}+\sqrt{\tilde{e}_{n2}}Z_n/\sqrt{n}$, then the result follows by using (A.30), (A.31), (A.41), and the tightness of $\widetilde{W}_n$. □

The second lemma reinforces the variance result in Lemma 4(iii).

Lemma 6 *Let (p_n) be a sequence of integers that diverges to infinity and (κ_n) be a sequence in $(0,\infty)$ that is $o(p_n)$. Denote as $\tilde{e}_{n2}$ the variance of the distribution with probability density function (2.8). Then, $\sqrt{p_n\tilde{e}_{n2}}-1=O(\kappa_n^2/p_n^2)$ as $n\to\infty$.*

Proof of Lemma 6. In this proof, C denotes a generic constant that may differ from line to line. Since (A.35) yields

$$\frac{\sqrt{\frac{3}{2}+\frac{1}{p_n}-\sqrt{\left(\frac{1}{2}+\frac{1}{p_n}\right)^2+\left(\frac{\kappa_n}{p_n}\right)^2}}}{\frac{1}{2}-\frac{1}{p_n}+\sqrt{\left(\frac{1}{2}+\frac{1}{p_n}\right)^2+\left(\frac{\kappa_n}{p_n}\right)^2}}-1\leq\sqrt{p_n\tilde{e}_{n2}}-1\leq\frac{1}{\sqrt{\frac{1}{2}+\sqrt{\left(\frac{1}{2}\right)^2+\left(\frac{\kappa_n}{p_n}\right)^2}}}-1\leq 0,$$

we have

$$\begin{aligned}
|\sqrt{p_n\tilde{e}_{n2}}-1| &\leq 1-\frac{\sqrt{\frac{3}{2}+\frac{1}{p_n}-\sqrt{\left(\frac{1}{2}+\frac{1}{p_n}\right)^2+\left(\frac{\kappa_n}{p_n}\right)^2}}}{\frac{1}{2}-\frac{1}{p_n}+\sqrt{\left(\frac{1}{2}+\frac{1}{p_n}\right)^2+\left(\frac{\kappa_n}{p_n}\right)^2}}\\
&\leq C\Bigg\{\left(\frac{1}{2}-\frac{1}{p_n}+\sqrt{\left(\frac{1}{2}+\frac{1}{p_n}\right)^2+\left(\frac{\kappa_n}{p_n}\right)^2}\right)-\sqrt{\frac{3}{2}+\frac{1}{p_n}-\sqrt{\left(\frac{1}{2}+\frac{1}{p_n}\right)^2+\left(\frac{\kappa_n}{p_n}\right)^2}}\Bigg\}\\
&\leq C\Bigg\{\left(\frac{1}{2}-\frac{1}{p_n}+\sqrt{\left(\frac{1}{2}+\frac{1}{p_n}\right)^2+\left(\frac{\kappa_n}{p_n}\right)^2}\right)^2-\left(\frac{3}{2}+\frac{1}{p_n}-\sqrt{\left(\frac{1}{2}+\frac{1}{p_n}\right)^2+\left(\frac{\kappa_n}{p_n}\right)^2}\right)\Bigg\}.
\end{aligned}$$

Standard computations allow us to rewrite this upper-bound as

$$\begin{aligned}
|\sqrt{p_n\tilde{e}_{n2}}-1| &\leq C\Bigg\{\frac{2(p_n-1)}{p_n}\sqrt{\left(\frac{p_n+2}{2p_n}\right)^2+\left(\frac{\kappa_n}{p_n}\right)^2}-\frac{(p_n-1)(p_n+2)-\kappa_n^2}{p_n^2}\Bigg\}\\
&\leq C\Bigg\{\frac{4(p_n-1)^2}{p_n^2}\left(\left(\frac{p_n+2}{2p_n}\right)^2+\left(\frac{\kappa_n}{p_n}\right)^2\right)-\frac{((p_n-1)(p_n+2)-\kappa_n^2)^2}{p_n^4}\Bigg\}=C\left(6-\frac{6}{p_n}-\frac{\kappa_n^2}{p_n^2}\right)\frac{\kappa_n^2}{p_n^2}.
\end{aligned}$$

which, for n large, is upper-bounded by $C\kappa_n^2/p_n^2$, as was to be proved. □

Proof of Theorem 4. Since $\kappa_n=o(p_n)$, Lemma 4(iii) entails that

$$\begin{aligned}
Z_n=\frac{\sqrt{n}(\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0}-e_{n1})}{\sqrt{\tilde{e}_{2n}}}&=\frac{\sqrt{n}(\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0}-\xi r_n/\sqrt{n}+O(r_n^3n^{-3/2}))}{\sqrt{\frac{1}{p_n}+o(\frac{1}{p_n})}}\\
=\frac{\sqrt{p_n}(\sqrt{n}\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0}-\xi r_n+O(r_n^3n^{-1}))}{\sqrt{1+o(1)}}&=\sqrt{np_n}\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0}-\xi\sqrt{p_n}r_n+O_{\rm P}\Big(\frac{\sqrt{p_n}r_n^3}{n}\Big) \qquad \text{(A.42)}
\end{aligned}$$

as $n\to\infty$ under ${\rm P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$. Write then

$$\begin{aligned}
\Lambda^{(n){\rm inv}}_{\boldsymbol{\theta}_n,\kappa_{n,s},\boldsymbol{\tau}_n/\boldsymbol{\theta}_{n0},\kappa_n}&=\Lambda^{(n){\rm inv}}_{\boldsymbol{\theta}_n,\kappa_{n,s},\boldsymbol{\tau}_n/\boldsymbol{\theta}_{n0},\kappa_{n,s},\boldsymbol{\tau}_n}+\Lambda^{(n){\rm inv}}_{\boldsymbol{\theta}_{n0},\kappa_{n,s},\boldsymbol{\tau}_n/\boldsymbol{\theta}_{n0},\kappa_n}\\
&=\Lambda^{(n){\rm inv}}_{\boldsymbol{\theta}_n,\kappa_{n,s},\boldsymbol{\tau}_n/\boldsymbol{\theta}_{n0},\kappa_{n,s},\boldsymbol{\tau}_n}+\log\frac{d{\rm P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_{n,s},\boldsymbol{\tau}_n}}{d{\rm P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}}\\
&=: L_{n1}+L_{n2}.
\end{aligned}$$

Letting $\rho_n := (1 - \frac{1}{2}\nu_n^2\|\boldsymbol{\tau}_n\|^2)^{-1}$ and using (A.42), we obtain

$$
\begin{aligned}
L_{n2} &= n\big(\log(c_{p_n,\kappa_{n,s,\boldsymbol{\tau}_n}}) - \log(c_{p_n,\kappa_n})\big) + n(\kappa_{n,s,\boldsymbol{\tau}_n} - \kappa_n)\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} \\
&= n\Big[\log\Big(\frac{c_{p_n,0}}{c_{p_n,\kappa_n}}\Big) - \log\Big(\frac{c_{p_n,0}}{c_{p_n,\kappa_{n,s,\boldsymbol{\tau}_n}}}\Big)\Big] + \rho_n(s\sqrt{np_n} + \tfrac{1}{2}\xi\sqrt{n}p_n r_n\nu_n^2\|\boldsymbol{\tau}_n\|^2)\,\bar{\mathbf{X}}_n'\boldsymbol{\theta}_{n0} \\
&= n\Big[\log H_{\frac{p_n}{2}-1}(\kappa_n) - \log H_{\frac{p_n}{2}-1}(\kappa_{n,s,\boldsymbol{\tau}_n})\Big] \\
&\qquad + \rho_n(s\sqrt{np_n} + \tfrac{1}{2}\xi\sqrt{n}p_n r_n\nu_n^2\|\boldsymbol{\tau}_n\|^2)\left(\frac{Z_n}{\sqrt{np_n}} + \frac{\xi r_n}{\sqrt{n}} + O_{\rm P}\Big(\frac{r_n^3}{n^{3/2}}\Big)\right)
\end{aligned}
$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$. Since $\sqrt{p_n}r_n^3$ and $p_n r_n^4\nu_n^2$ are $o(n)$, this yields

$$
\begin{aligned}
L_{n2} &= n\Big[\log H_{\frac{p_n}{2}-1}(\kappa_n) - \log H_{\frac{p_n}{2}-1}(\kappa_{n,s,\boldsymbol{\tau}_n})\Big] \\
&\quad + \rho_n\big(s + \tfrac{1}{2}\xi\sqrt{p_n}r_n\nu_n^2\|\boldsymbol{\tau}_n\|^2\big)Z_n + \xi s\rho_n r_n\sqrt{p_n} + \tfrac{1}{2}\xi^2\rho_n p_n r_n^2\nu_n^2\|\boldsymbol{\tau}_n\|^2 + o_{\rm P}(1) \\
&=: \tilde{L}_{n2} + \rho_n(s + \tfrac{1}{2}\xi\sqrt{p_n}r_n\nu_n^2\|\boldsymbol{\tau}_n\|^2)Z_n + o_{\rm P}(1)
\end{aligned}
$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$. Since $p_n = o(n^2 r_n^{-4})$, we can apply Lemma 3 with $a_n = \sqrt{n}$ and $T_n \equiv 1$, which provides

$$
\begin{aligned}
\tilde{L}_{n2} &= \Big(\frac{n\kappa_n^2}{4(\frac{p_n}{2}-1)} - \frac{n\kappa_n^4}{32(\frac{p_n}{2}-1)^3}\Big) - \Big(\frac{n\kappa_{n,s,\boldsymbol{\tau}_n}^2}{4(\frac{p_n}{2}-1)} - \frac{n\kappa_{n,s,\boldsymbol{\tau}_n}^4}{32(\frac{p_n}{2}-1)^3}\Big) \\
&\qquad + \xi s\rho_n r_n\sqrt{p_n} + \tfrac{1}{2}\xi^2\rho_n p_n r_n^2\nu_n^2\|\boldsymbol{\tau}_n\|^2 + o(1) \\
&= -\frac{n(\kappa_{n,s,\boldsymbol{\tau}_n}^2 - \kappa_n^2)}{2p_n - 4} + \frac{n(\kappa_{n,s,\boldsymbol{\tau}_n}^4 - \kappa_n^4)}{4(p_n-2)^3} + \xi s\rho_n r_n\sqrt{p_n} + \tfrac{1}{2}\xi^2\rho_n p_n r_n^2\nu_n^2\|\boldsymbol{\tau}_n\|^2 + o(1) \\
&= -\frac{p_n^2 r_n^2\{\rho_n^2(\xi + s/(\sqrt{p_n}r_n))^2 - \xi^2\}}{2p_n - 4} + S_n + \xi s\rho_n r_n\sqrt{p_n} + \tfrac{1}{2}\xi^2\rho_n p_n r_n^2\nu_n^2\|\boldsymbol{\tau}_n\|^2 + o(1),
\end{aligned}
$$

where, since $1 - \rho_n = -\frac{1}{2}\rho_n\nu_n^2\|\boldsymbol{\tau}_n\|^2$,

$$
S_n = \frac{p_n^4 r_n^4\{\rho_n^4(\xi + s/(\sqrt{p_n}r_n))^4 - \xi^4\}}{4n(p_n-2)^3} = \frac{p_n^4 r_n^4}{4n(p_n-2)^3}\Big((\rho_n^4 - 1)\xi^4 + \textstyle\sum_{\ell=1}^4 \binom{4}{\ell}\frac{s^\ell\xi^{4-\ell}}{(\sqrt{p_n}r_n)^\ell}\Big)
$$

$$
= O\Big(\frac{p_n r_n^4}{n}(\rho_n - 1)(\rho_n + 1)(\rho_n^2 + 1)\Big) + O\Big(\frac{\sqrt{p_n}r_n^3}{n}\Big) = O\Big(\frac{p_n r_n^4\nu_n^2}{n}\Big) + O\Big(\frac{\sqrt{p_n}r_n^3}{n}\Big) = o(1).
$$

Thus, using the identities $1 - \rho_n = -\frac{1}{2}\rho_n\nu_n^2\|\boldsymbol{\tau}_n\|^2$ and $\rho_n^2 - 1 = \rho_n^2\nu_n^2\|\boldsymbol{\tau}_n\|^2 - \frac{1}{4}\rho_n^2\nu_n^4\|\boldsymbol{\tau}_n\|^4$, we have

$$
\begin{aligned}
\tilde{L}_{n2} &= -\frac{(\rho_n^2 - 1)\xi^2 p_n^2 r_n^2}{2p_n - 4} - \frac{s^2\rho_n^2 p_n}{2p_n - 4} - \frac{2\xi s\rho_n^2 p_n^{3/2} r_n}{2p_n - 4} + \xi s\rho_n r_n\sqrt{p_n} + \tfrac{1}{2}\xi^2\rho_n p_n r_n^2\nu_n^2\|\boldsymbol{\tau}_n\|^2 + o(1) \\
&= -\frac{1}{2}\xi^2\rho_n^2 p_n r_n^2\nu_n^2\|\boldsymbol{\tau}_n\|^2 + \frac{1}{8}\xi^2\rho_n^2 p_n r_n^2\nu_n^4\|\boldsymbol{\tau}_n\|^4 - \frac{1}{2}s^2\rho_n^2 \qquad \text{(A.43)} \\
&\qquad - \xi s\rho_n^2\sqrt{p_n}r_n + \xi s\rho_n r_n\sqrt{p_n} + \tfrac{1}{2}\xi^2\rho_n p_n r_n^2\nu_n^2\|\boldsymbol{\tau}_n\|^2 + o(1) \\
&= \frac{1}{2}\xi^2\rho_n(1 - \rho_n)p_n r_n^2\nu_n^2\|\boldsymbol{\tau}_n\|^2 + \frac{1}{8}\xi^2\rho_n^2 p_n r_n^2\nu_n^4\|\boldsymbol{\tau}_n\|^4 - \frac{1}{2}s^2\rho_n^2 + \xi s\rho_n(1 - \rho_n)\sqrt{p_n}r_n + o(1) \\
&= -\frac{1}{2}s^2\rho_n^2 - \frac{1}{2}\xi s\rho_n^2\sqrt{p_n}r_n\nu_n^2\|\boldsymbol{\tau}_n\|^2 - \frac{1}{8}\xi^2\rho_n^2 p_n r_n^2\nu_n^4\|\boldsymbol{\tau}_n\|^4 + o(1)
\end{aligned}
$$

as $n \to \infty$. Therefore, we proved that, as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_n}$,

$$L_{n2} = \rho_n(s + \tfrac{1}{2}\xi\sqrt{p_n}r_n\nu_n^2\|\boldsymbol{\tau}_n\|^2)Z_n \qquad \text{(A.44)}$$
$$-\frac{1}{2}s^2\rho_n^2 - \frac{1}{2}\xi s\rho_n^2\sqrt{p_n}r_n\nu_n^2\|\boldsymbol{\tau}_n\|^2 - \frac{1}{8}\xi^2\rho_n^2 p_n r_n^2\nu_n^4\|\boldsymbol{\tau}_n\|^4 + o_\mathrm{P}(1).$$

We turn to L_{1n}. Write $c_{n,s,\boldsymbol{\tau}_n} := n\nu_n^2\kappa_{n,s,\boldsymbol{\tau}_n}e_{n1,s,\boldsymbol{\tau}_n}/\sqrt{p_n}$, where $e_{n1,s,\boldsymbol{\tau}_n}$ and $\tilde{e}_{2n,s,\boldsymbol{\tau}_n}$ denote the values of e_{n1} and $\tilde{e}_{2n}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0},\kappa_{n,s,\boldsymbol{\tau}_n}}$. Since $\sqrt{p_n}r_n^2\nu_n^2 = O(1)$, (A.48) below ensures that $c_{n,s,\boldsymbol{\tau}_n}$ is $O(1)$. Therefore, Lemma 5 yields

$$L_{1n} = L_{1n}^Z + \tilde{L}_{1n}^Z + \bar{L}_{1n}^Z + L_{1n}^W + o_\mathrm{P}(1), \qquad \text{(A.45)}$$

where we let

$$L_{1n}^Z := -\frac{1}{2}\sqrt{n}\kappa_{n,s,\boldsymbol{\tau}_n}\nu_n^2\sqrt{\tilde{e}_{n2}}\,\|\boldsymbol{\tau}_n\|^2 Z_n, \quad \tilde{L}_{1n}^Z := -\frac{1}{8}\sqrt{p_n}\nu_n^2 c_{n,s,\boldsymbol{\tau}_n}\|\boldsymbol{\tau}_n\|^4,$$
$$\bar{L}_{1n}^Z := \frac{1}{2}n\kappa_{n,s,\boldsymbol{\tau}_n}\nu_n^2(e_{n1,s,\boldsymbol{\tau}_n} - e_{n1})\|\boldsymbol{\tau}_n\|^2,$$

and

$$L_{1n}^W := \frac{1}{\sqrt{2}}c_{n,s,\boldsymbol{\tau}_n}\|\boldsymbol{\tau}_n\|^2\Big(1 - \frac{1}{4}\nu_n^2\|\boldsymbol{\tau}_n\|^2\Big)\widetilde{W}_n - \frac{1}{4}c_{n,s,\boldsymbol{\tau}_n}^2\|\boldsymbol{\tau}_n\|^4\Big(1 - \frac{1}{4}\nu_n^2\|\boldsymbol{\tau}_n\|^2\Big)^2.$$

Lemma 6 provides

$$L_{1n}^Z = -\frac{\sqrt{n}\nu_n^2}{2\sqrt{p_n}}\|\boldsymbol{\tau}_n\|^2\Big(\frac{\rho_n p_n r_n}{\sqrt{n}}\Big(\xi + \frac{s}{\sqrt{p_n}r_n}\Big)\Big)\sqrt{p\tilde{e}_{n2}}\,Z_n$$
$$= -\frac{1}{2}\rho_n\sqrt{p_n}r_n\nu_n^2\|\boldsymbol{\tau}_n\|^2\Big(\xi + \frac{s}{\sqrt{p_n}r_n}\Big)\Big(1 + O\Big(\frac{\kappa_n^2}{p_n^2}\Big)\Big)Z_n$$
$$= -\frac{1}{2}\xi\rho_n\sqrt{p_n}r_n\nu_n^2\|\boldsymbol{\tau}_n\|^2 Z_n - \frac{1}{2}s\rho_n\nu_n^2\|\boldsymbol{\tau}_n\|^2 Z_n + o_\mathrm{P}(1), \qquad \text{(A.46)}$$

where we used the fact that $\sqrt{p_n}r_n^3\nu_n^2$ is $o(n)$.

Now, Lemma 4(iii) yields

$$e_{n1,s,\boldsymbol{\tau}_n} = \frac{\kappa_{n,s,\boldsymbol{\tau}_n}}{p_n} + O\Big(\frac{\kappa_{n,s,\boldsymbol{\tau}_n}^3}{p_n^3}\Big) = \frac{\rho_n r_n}{\sqrt{n}}\Big(\xi + \frac{s}{\sqrt{p_n}r_n}\Big) + O\Big(\frac{r_n^3}{n^{3/2}}\Big)$$
$$= \frac{\xi\rho_n r_n}{\sqrt{n}} + \frac{s\rho_n}{\sqrt{np_n}} + O\Big(\frac{r_n^3}{n^{3/2}}\Big), \qquad \text{(A.47)}$$

so that

$$c_{n,s,\boldsymbol{\tau}_n} = \frac{n\nu_n^2}{\sqrt{p_n}}\Big(\frac{\rho_n p_n r_n}{\sqrt{n}}\Big(\xi + \frac{s}{\sqrt{p_n}r_n}\Big)\Big)\Big(\frac{\xi\rho_n r_n}{\sqrt{n}} + \frac{s\rho_n}{\sqrt{np_n}} + O\Big(\frac{r_n^3}{n^{3/2}}\Big)\Big)$$
$$= \rho_n^2\sqrt{p_n}r_n^2\nu_n^2\Big(\xi + \frac{s}{\sqrt{p_n}r_n}\Big)\Big(\xi + \frac{s}{\sqrt{p_n}r_n} + O\Big(\frac{r_n^2}{n}\Big)\Big), \qquad \text{(A.48)}$$

which in turn implies that

$$\tilde{L}_{1n}^Z = -\frac{1}{8}\sqrt{p_n}\nu_n^2\Big(\xi^2\rho_n^2\sqrt{p_n}r_n^2\nu_n^2 + 2\xi s\rho_n^2 r_n\nu_n^2 + s^2\frac{\rho_n^2\nu_n^2}{\sqrt{p_n}} + O\Big(\frac{\sqrt{p_n}r_n^4\nu_n^2}{n}\Big)\Big)\|\boldsymbol{\tau}_n\|^4$$
$$= -\frac{1}{8}\xi^2\rho_n^2 p_n r_n^2\nu_n^4\|\boldsymbol{\tau}_n\|^4 - \frac{1}{4}\xi s\rho_n^2\sqrt{p_n}r_n\nu_n^4\|\boldsymbol{\tau}_n\|^4 - \frac{1}{8}s^2\rho_n^2\nu_n^4\|\boldsymbol{\tau}_n\|^4 + o(1). \qquad \text{(A.49)}$$

Using (A.47) and applying Lemma 4(iii) again, we obtain

$$
\begin{aligned}
e_{n1,s,\boldsymbol{\tau}_n} - e_{n1} &= \Big(\frac{\xi\rho_n r_n}{\sqrt{n}} + \frac{s\rho_n}{\sqrt{np_n}} + O\Big(\frac{r_n^3}{n^{3/2}}\Big)\Big) - \frac{\kappa_n}{p_n} + O\Big(\frac{\kappa_n^3}{p_n^3}\Big) \\
&= \frac{\xi(\rho_n - 1) r_n}{\sqrt{n}} + \frac{s\rho_n}{\sqrt{np_n}} + O\Big(\frac{r_n^3}{n^{3/2}}\Big) = \frac{\xi\rho_n r_n \nu_n^2}{2\sqrt{n}}\|\boldsymbol{\tau}_n\|^2 + \frac{s\rho_n}{\sqrt{np_n}} + O\Big(\frac{r_n^3}{n^{3/2}}\Big),
\end{aligned}
$$

so that

$$
\begin{aligned}
\bar{L}_{1n}^Z &= \frac{1}{2} n \Big(\frac{\rho_n p_n r_n}{\sqrt{n}}\Big(\xi + \frac{s}{\sqrt{p_n} r_n}\Big)\Big)\nu_n^2 \Big(\frac{\xi\rho_n r_n \nu_n^2}{2\sqrt{n}}\|\boldsymbol{\tau}_n\|^2 + \frac{s\rho_n}{\sqrt{np_n}} + O\Big(\frac{r_n^3}{n^{3/2}}\Big)\Big)\|\boldsymbol{\tau}_n\|^2 \\
&= \frac{1}{2}\rho_n p_n r_n \sqrt{n}\nu_n^2 \|\boldsymbol{\tau}_n\|^2 \xi \Big(\frac{\xi\rho_n r_n \nu_n^2}{2\sqrt{n}}\|\boldsymbol{\tau}_n\|^2 + \frac{s\rho_n}{\sqrt{np_n}} + O\Big(\frac{r_n^3}{n^{3/2}}\Big)\Big) \\
&\qquad + \frac{1}{2}\rho_n p_n r_n \sqrt{n}\nu_n^2 \|\boldsymbol{\tau}_n\|^2 \frac{s}{\sqrt{p_n} r_n}\Big(\frac{\xi\rho_n r_n \nu_n^2}{2\sqrt{n}}\|\boldsymbol{\tau}_n\|^2 + \frac{s\rho_n}{\sqrt{np_n}} + O\Big(\frac{r_n^3}{n^{3/2}}\Big)\Big) \\
&= \frac{1}{4}\xi^2 \rho_n^2 p_n r_n^2 \nu_n^4 \|\boldsymbol{\tau}_n\|^4 + \frac{1}{2}\xi s \rho_n^2 \sqrt{p_n} r_n \nu_n^2 \|\boldsymbol{\tau}_n\|^2 \\
&\qquad + \frac{1}{4}\xi s \rho_n^2 \sqrt{p_n} r_n \nu_n^4 \|\boldsymbol{\tau}_n\|^4 + \frac{1}{2} s^2 \rho_n^2 \nu_n^2 \|\boldsymbol{\tau}_n\|^2 + o(1),
\end{aligned}
\tag{A.50}
$$

where we used the facts that $p_n r_n^4 \nu_n^2$ and $\sqrt{p_n} r_n^3 \nu_n^2$ are $o(n)$.

Jointly with (A.44), (A.46), (A.49), and (A.50), this shows that

$$
\begin{aligned}
\Lambda^{(n)\mathrm{inv}}_{\boldsymbol{\theta}_n,\kappa_n,s,\boldsymbol{\tau}_n/\boldsymbol{\theta}_{n0},\kappa_n} - L_{1n}^W &= s\rho_n(1 - \tfrac{1}{2}\nu_n^2\|\boldsymbol{\tau}_n\|^2) Z_n - \frac{1}{2}s^2\rho_n^2(1 - \nu_n^2\|\boldsymbol{\tau}_n\|^2 + \tfrac{1}{4}\nu_n^4\|\boldsymbol{\tau}_n\|^4) + o_{\mathrm{P}}(1) \\
&= sZ_n - \frac{1}{2}s^2 + o_{\mathrm{P}}(1).
\end{aligned}
$$

The result thus follows from the definition of L_{1n}^W and the fact that (A.48) implies that $c_{n,s,\boldsymbol{\tau}_n} = 1 + o(1)$ in case (a), $c_{n,s,\boldsymbol{\tau}_n} = \xi^2 + o(1)$ in case (b), and $c_{n,s,\boldsymbol{\tau}_n} = o(1)$ in case (c) (in each case, the asymptotic normality result of $\boldsymbol{\Delta}_n$ follows from Lemma 2). □

B Technical proofs for Section 3

The proof of Theorem 5 requires the following lemma.

Lemma 7 *Let $\mathbf{M}_n := \boldsymbol{\theta}_n\boldsymbol{\theta}_n' - \boldsymbol{\theta}_{n0}\boldsymbol{\theta}_{n0}'$, where $(\boldsymbol{\theta}_n)$ and $(\boldsymbol{\theta}_{n0})$ belong to $\mathcal{S}^{p_n-1}$. Then, for any real numbers a, b, c, d, we have that $\mathrm{tr}\big[\mathbf{M}_n^\ell(a\boldsymbol{\theta}_n\boldsymbol{\theta}_n' + b(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n'))\mathbf{M}_n^\ell(c\boldsymbol{\theta}_n\boldsymbol{\theta}_n' + d(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n'))\big]$ is equal to $(ad + bc)(1 - (\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2) + (a-b)(c-d)(1 - (\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2)^2$ for $\ell = 1$ and to $(ac + bd)(1 - (\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2)^2$ for $\ell = 2$.*

Proof of Lemma 7. Direct computations yield

$$
\mathbf{M}_n^2 = \boldsymbol{\theta}_n\boldsymbol{\theta}_n' + \boldsymbol{\theta}_{n0}\boldsymbol{\theta}_{n0}' - (\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0})\boldsymbol{\theta}_{n0}\boldsymbol{\theta}_n' - (\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0})\boldsymbol{\theta}_n\boldsymbol{\theta}_{n0}' \quad \text{and} \quad \mathbf{M}_n^4 = (1 - (\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2)\mathbf{M}_n^2.
$$

This provides $\mathrm{tr}[\mathbf{M}_n^2] = 2(1 - (\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2)$ and $\mathrm{tr}[\mathbf{M}_n^4] = 2(1 - (\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2)^2$, and allows to show that $\boldsymbol{\theta}_n'\mathbf{M}_n^2\boldsymbol{\theta}_n = 1 - (\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0})^2$ and $\boldsymbol{\theta}_n'\mathbf{M}_n^4\boldsymbol{\theta}_n = (1 - (\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2)^2$. Since $\boldsymbol{\theta}_n'\mathbf{M}_n\boldsymbol{\theta}_n =$

$1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2$, this yields

$$\begin{aligned}
&\mathrm{tr}\Big[\mathbf{M}_n(a\boldsymbol{\theta}_n\boldsymbol{\theta}'_n+b(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}'_n))\mathbf{M}_n(c\boldsymbol{\theta}_n\boldsymbol{\theta}'_n+d(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}'_n))\Big]\\
&\quad=\mathrm{tr}\Big[\mathbf{M}_n(b\mathbf{I}_{p_n}+(a-b)\boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\mathbf{M}_n(d\mathbf{I}_{p_n}+(c-d)\boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\Big]\\
&\quad=bd\,\mathrm{tr}[\mathbf{M}_n^2]+b(c-d)\boldsymbol{\theta}'_n\mathbf{M}_n^2\boldsymbol{\theta}_n+(a-b)d\boldsymbol{\theta}'_n\mathbf{M}_n^2\boldsymbol{\theta}_n+(a-b)(c-d)(\boldsymbol{\theta}'_n\mathbf{M}_n\boldsymbol{\theta}_n)^2\\
&\quad=2bd\boldsymbol{\theta}'_n\mathbf{M}_n\boldsymbol{\theta}_n+\{b(c-d)+(a-b)d\}\boldsymbol{\theta}'_n\mathbf{M}_n\boldsymbol{\theta}_n+(a-b)(c-d)(\boldsymbol{\theta}'_n\mathbf{M}_n\boldsymbol{\theta}_n)^2\\
&\quad=(ad+bc)(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)+(a-b)(c-d)(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{tr}\Big[\mathbf{M}_n^2(a\boldsymbol{\theta}_n\boldsymbol{\theta}'_n+b(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}'_n))\mathbf{M}_n(c\boldsymbol{\theta}_n\boldsymbol{\theta}'_n+d(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}'_n))\Big]\\
&\quad=\mathrm{tr}\Big[\mathbf{M}_n^2(b\mathbf{I}_{p_n}+(a-b)\boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\mathbf{M}_n^2(d\mathbf{I}_{p_n}+(c-d)\boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\Big]\\
&\quad=bd\,\mathrm{tr}[\mathbf{M}_n^4]+b(c-d)\boldsymbol{\theta}'_n\mathbf{M}_n^4\boldsymbol{\theta}_n+(a-b)d\boldsymbol{\theta}'_n\mathbf{M}_n^4\boldsymbol{\theta}_n+(a-b)(c-d)(\boldsymbol{\theta}'_n\mathbf{M}_n^2\boldsymbol{\theta}_n)^2\\
&\quad=2bd(\boldsymbol{\theta}'_n\mathbf{M}_n\boldsymbol{\theta}_n)^2+\{b(c-d)+(a-b)d\}(\boldsymbol{\theta}'_n\mathbf{M}_n\boldsymbol{\theta}_n)^2+(a-b)(c-d)(\boldsymbol{\theta}'_n\mathbf{M}_n\boldsymbol{\theta}_n)^2\\
&\quad=\{ad+bc+(a-b)(c-d)\}(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2\\
&\quad=(ac+bd)(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2,
\end{aligned}$$

as was to be showed. □

Proof of Theorem 5. All expectations and variances below are taken under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,F_n}$, with $\boldsymbol{\theta}_n=\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n$, and stochastic convergences are under the corresponding sequence of hypotheses. We have

$$\mathrm{E}[\mathbf{X}_{ni}]=e_{n1}\boldsymbol{\theta}_n \quad\text{and}\quad \mathrm{E}[\mathbf{X}_{ni}\mathbf{X}'_{ni}]=e_{n2}\boldsymbol{\theta}_n\boldsymbol{\theta}'_n+\frac{f_{n2}}{p_n-1}(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}'_n);$$

see the proof of Lemma B.3 in [12]. Writing $\mathbf{W}_{ni}:=(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\mathbf{X}_{ni}$, this implies that

$$\mathrm{E}[\mathbf{W}_{ni}]=\mathbf{0} \quad\text{and}\quad \mathrm{E}[\mathbf{W}_{ni}\mathbf{W}'_{ni}]=\frac{f_{n2}}{p_n-1}(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}'_n). \tag{B.51}$$

Writing $\mathbf{M}_n=\boldsymbol{\theta}_n\boldsymbol{\theta}'_n-\boldsymbol{\theta}_{n0}\boldsymbol{\theta}'_{n0}$ as in Lemma 7 and $\mathbf{Y}_{ni}:=(\mathbf{I}_{p_n}-\boldsymbol{\theta}_{n0}\boldsymbol{\theta}'_{n0})\mathbf{X}_{ni}$, we have $\mathbf{W}_{ni}=\mathbf{Y}_{ni}-\mathbf{M}_n\mathbf{X}_{ni}$. This allows to decompose W_n^* as

$$\begin{aligned}
W_n^*&:=\frac{\sqrt{2(p_n-1)}}{nf_{n2}}\sum_{1\le i<j\le n}\mathbf{Y}'_{ni}\mathbf{Y}_{nj}\\
&=\frac{\sqrt{2(p_n-1)}}{nf_{n2}}\sum_{1\le i<j\le n}(\mathbf{W}'_{ni}\mathbf{W}_{nj}+\mathbf{X}'_{ni}\mathbf{M}_n\mathbf{W}_{nj}+\mathbf{W}'_{ni}\mathbf{M}_n\mathbf{X}_{nj}+\mathbf{X}'_{ni}\mathbf{M}_n^2\mathbf{X}_{nj})\\
&=:W_{n0}^*+W_{na}^*+W_{nb}^*+W_{nc}^*.
\end{aligned}$$

From (B.51), $\mathrm{E}[W^*_{na}] = \mathrm{E}[W^*_{nb}] = 0$. Now,

$$\begin{aligned}\mathrm{Var}[W^*_{na}] &= \frac{2(p_n-1)}{n^2 f_{n2}^2} \sum_{1\le i<j\le n} \sum_{1\le r<s<n} \mathrm{E}[\mathbf{X}'_{ni}\mathbf{M}_n\mathbf{W}_{nj}\mathbf{X}'_{nr}\mathbf{M}_n\mathbf{W}_{ns}] \\ &= \frac{2(p_n-1)}{n^2 f_{n2}^2} \sum_{1\le i<j\le n} \sum_{1\le r<s<n} \mathrm{tr}[\mathrm{E}[\mathbf{M}_n\mathbf{X}_{nr}\mathbf{X}'_{ni}\mathbf{M}_n\mathbf{W}_{nj}\mathbf{W}'_{ns}]] \\ &= \frac{2(p_n-1)}{n^2 f_{n2}^2} \sum_{1\le i<j\le n} \sum_{1\le r<s<n} c_{n,ijrs}.\end{aligned}$$

Clearly, $c_{n,ijrs} = 0$ if $s \neq j$. Lemma 7 entails that for $s = j$ and $r \neq i$, we have

$$\begin{aligned}c_{n,ijrs} &= \mathrm{tr}[\mathbf{M}_n\mathrm{E}[\mathbf{X}_{nr}\mathbf{X}'_{ni}]\mathbf{M}_n\mathrm{E}[\mathbf{W}_{nj}\mathbf{W}'_{nj}]] \\ &= \mathrm{tr}\Big[\mathbf{M}_n(e_{n1}^2\boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\mathbf{M}_n\Big(\frac{f_{n2}}{p_n-1}(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\Big)\Big] \\ &= \frac{e_{n1}^2 f_{n2}}{p_n-1}(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2) - \frac{e_{n1}^2 f_{n2}}{p_n-1}(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2\end{aligned}$$

and that, for $s = j$ and $r = i$, we have

$$\begin{aligned}c_{n,ijrs} &= \mathrm{tr}[\mathbf{M}_n\mathrm{E}[\mathbf{X}_{ni}\mathbf{X}'_{ni}]\mathbf{M}_n\mathrm{E}[\mathbf{W}_{nj}\mathbf{W}'_{nj}]] \\ &= \mathrm{tr}\Big[\mathbf{M}_n\Big(e_{n2}\boldsymbol{\theta}_n\boldsymbol{\theta}'_n + \frac{f_{n2}}{p_n-1}(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\Big)\mathbf{M}_n\Big(\frac{f_{n2}}{p_n-1}(\mathbf{I}_{p_n}-\boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\Big)\Big] \\ &= \frac{e_{n2} f_{n2}}{p_n-1}(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2) - \frac{(p_n e_{n2}-1) f_{n2}}{(p_n-1)^2}(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2.\end{aligned}$$

We conclude that

$$\begin{aligned}\mathrm{Var}[W^*_{na}] = \frac{2(p_n-1)}{n^2 f_{n2}^2}\Bigg[&\frac{n(n-1)(n-2)}{3}\Big(\frac{e_{n1}^2 f_{n2}}{p_n-1}(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2) - \frac{e_{n1}^2 f_{n2}}{p_n-1}(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2\Big) \\ &+\frac{n(n-1)}{2}\Big(\frac{e_{n2} f_{n2}}{p_n-1}(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2) - \frac{(p_n e_{n2}-1) f_{n2}}{(p_n-1)^2}(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2\Big)\Bigg] \\ = \frac{n-1}{3n}\Bigg[&\frac{2(n-2)e_{n1}^2+3e_{n2}}{f_{n2}}(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2) - \Big(\frac{2(n-2)e_{n1}^2}{f_{n2}} + \frac{3(p_n e_{n2}-1)}{(p_n-1)f_{n2}}\Big)(1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2\Bigg].\end{aligned}$$

Since $\boldsymbol{\theta}'_n\boldsymbol{\theta}_{n0} = (\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n)'\boldsymbol{\theta}_{n0} = 1+\nu_n(\boldsymbol{\tau}'_n\boldsymbol{\theta}_{n0}) = 1-\frac{1}{2}\nu_n^2\|\boldsymbol{\tau}_n\|^2$, we have $1-(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2 = O(\nu_n^2)$, which, by using the fact that $\nu_n = O(1)$, yields

$$\mathrm{Var}[W^*_{na}] = \frac{np_n e_{n1}^2 + p_n e_{n2} + 1}{p_n f_{n2}}\, O(\nu_n^2).$$

The same computations provide $\mathrm{Var}[W^*_{nb}] = \mathrm{Var}[W^*_{na}]$. Turning to W^*_{nc}, we directly obtain

$$\begin{aligned}\mathrm{E}[W^*_{nc}] &= \frac{\sqrt{2(p_n-1)}}{nf_{n2}} \times \frac{n(n-1)}{2}\, e_{n1}^2\boldsymbol{\theta}'_n\mathbf{M}_n^2\boldsymbol{\theta}_n \\ &= \frac{(n-1)(p_n-1)^{1/2}e_{n1}^2}{\sqrt{2}f_{n2}}\,(1-(\boldsymbol{\theta}'_n\boldsymbol{\theta}_{n0})^2) \\ &= \frac{np_n^{1/2}e_{n1}^2}{\sqrt{2}f_{n2}}\,\nu_n^2\|\boldsymbol{\tau}_n\|^2(1-\tfrac{1}{4}\nu_n^2\|\boldsymbol{\tau}_n\|^2)(1+o(1));\end{aligned}$$

see the proof of Lemma 7. As for the variance,

$$\begin{aligned}\mathrm{Var}[W^*_{nc}] &= \frac{2(p_n-1)}{n^2 f_{n2}^2} \sum_{1\le i<j\le n} \sum_{1\le r<s\le n} \mathrm{Cov}[\mathbf{X}'_{ni}\mathbf{M}_n^2\mathbf{X}_{nj}, \mathbf{X}'_{nr}\mathbf{M}_n^2\mathbf{X}_{ns}] \\ &= \frac{2(p_n-1)}{n^2 f_{n2}^2} \sum_{1\le i<j\le n} \sum_{1\le r<s\le n} \Big(\mathrm{E}[\mathbf{X}'_{ni}\mathbf{M}_n^2\mathbf{X}_{nj}\mathbf{X}'_{nr}\mathbf{M}_n^2\mathbf{X}_{ns}] - (e_{n1}^2\boldsymbol{\theta}'_n\mathbf{M}_n^2\boldsymbol{\theta}_n)^2\Big) \\ &= \frac{2(p_n-1)}{n^2 f_{n2}^2} \sum_{1\le i<j\le n} \sum_{1\le r<s\le n} \Big(\mathrm{tr}[\mathrm{E}[\mathbf{X}'_{ni}\mathbf{M}_n^2\mathbf{X}_{nj}\mathbf{X}'_{ns}\mathbf{M}_n^2\mathbf{X}_{nr}]] - e_{n1}^4(1-(\boldsymbol{\theta}'_n\boldsymbol{\theta}_{n0})^2)^2\Big) \\ &= \frac{2(p_n-1)}{n^2 f_{n2}^2} \sum_{1\le i<j\le n} \sum_{1\le r<s\le n} \Big(d_{n,ijrs} - e_{n1}^4(1-(\boldsymbol{\theta}'_n\boldsymbol{\theta}_{n0})^2)^2\Big).\end{aligned}$$

We consider three cases. (1) If i,j,r,s contain two pairs of equal indices (equivalently, if $r=i$ and $s=j$), then

$$\begin{aligned}d_{n,ijrs} &= \mathrm{tr}[\mathbf{M}_n^2\mathrm{E}[\mathbf{X}_{ni}\mathbf{X}'_{ni}]\mathbf{M}_n^2\mathrm{E}[\mathbf{X}_{nj}\mathbf{X}'_{nj}]] \\ &= \mathrm{tr}\Big[\mathbf{M}_n^2\Big(e_{n2}\boldsymbol{\theta}_n\boldsymbol{\theta}'_n + \frac{f_{n2}}{p_n-1}(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\Big)\mathbf{M}_n^2\Big(e_{n2}\boldsymbol{\theta}_n\boldsymbol{\theta}'_n + \frac{f_{n2}}{p_n-1}(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\Big)\Big] \\ &= \Big(e_{n2}^2 + \frac{f_{n2}^2}{(p_n-1)^2}\Big)(1-(\boldsymbol{\theta}'_n\boldsymbol{\theta}_{n0})^2)^2.\end{aligned}$$

(2) If i,j,r,s contain exactly one pair of equal indices, then

$$\begin{aligned}d_{n,ijrs} &= \mathrm{tr}[\mathbf{M}_n^2\mathrm{E}[\mathbf{X}_{ni}\mathbf{X}'_{ni}]\mathbf{M}_n^2\mathrm{E}[\mathbf{X}_{nj}\mathbf{X}'_{ns}]] \\ &= \mathrm{tr}\Big[\mathbf{M}_n^2\Big(e_{n2}\boldsymbol{\theta}_n\boldsymbol{\theta}'_n + \frac{f_{n2}}{p_n-1}(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\Big)\mathbf{M}_n^2(e_{n1}^2\boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\Big] \\ &= e_{n1}^2 e_{n2}(1-(\boldsymbol{\theta}'_n\boldsymbol{\theta}_{n0})^2)^2.\end{aligned}$$

(3) If the indices i,j,r,s are pairwise different, then

$$\begin{aligned}d_{n,ijrs} &= \mathrm{tr}[\mathbf{M}_n^2\mathrm{E}[\mathbf{X}_{nr}\mathbf{X}'_{ni}]\mathbf{M}_n^2\mathrm{E}[\mathbf{X}_{nj}\mathbf{X}'_{ns}]] \\ &= \mathrm{tr}\Big[\mathbf{M}_n^2(e_{n1}^2\boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\mathbf{M}_n^2(e_{n1}^2\boldsymbol{\theta}_n\boldsymbol{\theta}'_n)\Big] \\ &= e_{n1}^4(1-(\boldsymbol{\theta}'_n\boldsymbol{\theta}_{n0})^2)^2.\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{Var}[W^*_{nc}] = \frac{2(p_n-1)}{n^2 f_{n2}^2}\Big[&\frac{n(n-1)}{2}\Big(e_{n2}^2 + \frac{f_{n2}^2}{(p_n-1)^2}\Big) + n(n-1)(n-2)e_{n1}^2 e_{n2} \\ &+ \frac{n(n-1)(n-2)(n-3)}{4}\,e_{n1}^4 - \frac{n^2(n-1)^2}{4}\,e_{n1}^4\Big](1-(\boldsymbol{\theta}'_n\boldsymbol{\theta}_{n0})^2)^2 \\ = \frac{(n-1)(p_n-1)\tilde{e}_{n2}(\tilde{e}_{n2}+2(n-1)e_{n1}^2)}{n f_{n2}^2}&(1-(\boldsymbol{\theta}'_n\boldsymbol{\theta}_{n0})^2)^2 + o(1).\end{aligned}$$

This finally yields

$$\mathrm{Var}[W^*_{nc}] = \frac{p_n\tilde{e}_{n2}^2 + np_n e_{n1}^2\tilde{e}_{n2}}{f_{n2}^2}\,O(\nu_n^4).$$

Summarizing, $W_n^* = W_{n0}^* + W_{na}^* + W_{nb}^* + W_{nc}^*$, where W_{n0}^* is asymptotically standard normal (see Theorem 3.1 from [25]),

$$\mathrm{E}[W_{na}^*] = \mathrm{E}[W_{nb}^*] = 0, \quad \mathrm{E}[W_{nc}^*] = \frac{np_n^{1/2}e_{n1}^2}{\sqrt{2}f_{n2}}\,\nu_n^2\|\boldsymbol{\tau}_n\|^2(1-\tfrac{1}{4}\nu_n^2\|\boldsymbol{\tau}_n\|^2)(1+o(1)),$$

$$\mathrm{Var}[W_{na}^*] = \mathrm{Var}[W_{nb}^*] = \frac{np_n e_{n1}^2 + p_n e_{n2} + 1}{p_n f_{n2}}\,O(\nu_n^2)$$

and

$$\mathrm{Var}[W_{nc}^*] = \frac{p_n\tilde{e}_{n2}^2 + np_n e_{n1}^2\tilde{e}_{n2}}{f_{n2}^2}\,O(\nu_n^4).$$

We can now consider the several cases of the theorem. In cases (i)–(iii), the sequence (ν_n) involved, namely $\nu_n = \sqrt{f_{n2}}/(\sqrt{n}p_n^{1/4}e_{n1})$, is $o(1)$, so that $\mathrm{E}[W_{nc}^*] = t^2/\sqrt{2}+o(1)$. In all three cases, one checks that $\mathrm{Var}[W_{n\ell}^*] = o(1)$ for $\ell = a,b,c$ (note that in cases (ii)–(iii), the fact that $e_{n2} \leq e_{n1}$ implies that both e_{n2} and $\tilde{e}_{n2}$ are $o(1)$), which establishes that $W_n^* \xrightarrow{\mathcal{D}} \mathcal{N}\big(\frac{t^2}{\sqrt{2}},1\big)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n,F_n}$. In case (iv), we have, with $\nu_n = 1$, $\mathrm{E}[W_{nc}^*] = \frac{\xi^2t^2}{\sqrt{2}}\big(1-\frac{t^2}{4}\big)+o(1)$. Since $\sqrt{p_n}e_{n2} = o(1)$ by assumption, one can check that $\mathrm{Var}[W_{n\ell}^*] = o(1)$ for $\ell = a,b,c$, which yields $W_n^* \xrightarrow{\mathcal{D}} \mathcal{N}\big(\frac{\xi^2t^2}{\sqrt{2}}\big(1-\frac{t^2}{4}\big),1\big)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n,F_n}$, as was to be showed. Finally, in case (v), still with $\nu_n = 1$, we have $\mathrm{E}[W_{nc}^*] = o(1)$. One can again check that $\mathrm{Var}[W_{n\ell}^*] = o(1)$ for $\ell = a,b,c$, which yields that W_n^* is asymptotically standard normal. This establishes the result. □

We turn to the proof of Theorem 6, that will make use of the following lemma.

Lemma 8 *Under* $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,F_n}$,

$$\mathrm{E}[(\mathbf{X}_{n1}'\boldsymbol{\theta}_{n0})^2] = e_{n2}(\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2 + \frac{f_{n2}}{p_n-1}(1-(\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2)$$

and

$$\mathrm{E}[(\mathbf{X}_{n1}'\boldsymbol{\theta}_{n0})^4] = e_{n4}(\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0})^4 + \frac{6(e_{n2}-e_{n4})}{p_n-1}(\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0})^2(1-(\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2) + \frac{3f_{n4}}{p_n^2-1}(1-(\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2)^2.$$

Proof of Lemma 8. All computations in this proof are performed under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,F_n}$, which leads us to consider the tangent-decomposition $\mathbf{X}_{n1} = U_{n1}\boldsymbol{\theta}_n + V_{n1}\mathbf{S}_{n1}$ of $\mathbf{X}_{n1}$ with respect to $\boldsymbol{\theta}_n$. Since $\mathbf{X}_{n1}$ is rotationally symmetric with respect to $\boldsymbol{\theta}_n$, $\mathbf{S}_{n1}$ is equal in distribution to $\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}\mathbf{U}_n$, where $\mathbf{U}_n$ is uniformly distributed over the unit sphere $\mathcal{S}^{p_n-2}$ in $\mathbb{R}^{p_n-1}$ and where $\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}$ is an arbitrary $p_n\times(p_n-1)$ matrix whose columns form an orthonormal basis of the orthogonal complement of $\boldsymbol{\theta}_n$ in $\mathbb{R}^{p_n}$ (so that $\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}'\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n} = \mathbf{I}_{p_n-1}$ and $\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}' = \mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n'$). In particular,

$$\mathrm{E}[\mathbf{S}_{n1}] = \mathbf{0} \quad \text{and} \quad \mathrm{E}[\mathbf{S}_{n1}\mathbf{S}_{n1}'] = \frac{1}{p_n-1}(\mathbf{I}_{p_n} - \boldsymbol{\theta}_n\boldsymbol{\theta}_n').$$

This readily yields

$$\begin{aligned}\mathrm{E}[(\mathbf{X}_{n1}'\boldsymbol{\theta}_{n0})^2] &= \mathrm{E}[(U_{n1}\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0} + V_{n1}\mathbf{S}_{n1}'\boldsymbol{\theta}_{n0})^2]\\ &= \mathrm{E}[U_{n1}^2](\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0})^2 + 2\mathrm{E}[U_{n1}V_{n1}]\mathrm{E}[\mathbf{S}_{n1}'\boldsymbol{\theta}_{n0}](\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0}) + \mathrm{E}[V_{n1}^2]\boldsymbol{\theta}_{n0}'\mathrm{E}[\mathbf{S}_{n1}\mathbf{S}_{n1}']\boldsymbol{\theta}_{n0}\\ &= e_{n2}(\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2 + \frac{f_{n2}}{p_n-1}(1-(\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2).\end{aligned}$$

Using the identity $U_{n1}^2 V_{n1}^2 = U_{n1}^2 - U_{n1}^4$, we obtain similarly

$$\begin{aligned}
\mathrm{E}[(\mathbf{X}_{n1}'\boldsymbol{\theta}_{n0})^4] &= \mathrm{E}[(U_{n1}\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0} + V_{n1}\mathbf{S}_{n1}'\boldsymbol{\theta}_{n0})^4] \\
&= e_{n4}(\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0})^4 + 6\mathrm{E}[U_{n1}^2 V_{n1}^2 (\mathbf{S}_{n1}'\boldsymbol{\theta}_{n0})^2](\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0})^2 + f_{n4}\mathrm{E}[(\mathbf{S}_{n1}'\boldsymbol{\theta}_{n0})^4] \\
&= e_{n4}(\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0})^4 + 6(e_{n2} - e_{n4})\,(\boldsymbol{\theta}_{n0}'\mathrm{E}[\mathbf{S}_{n1}\mathbf{S}_{n1}']\boldsymbol{\theta}_{n0})(\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0})^2 + f_{n4}\mathrm{E}[(\mathbf{S}_{n1}'\boldsymbol{\theta}_{n0})^4] \\
&= e_{n4}(\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0})^4 + \frac{6(e_{n2} - e_{n4})}{p_n - 1}(\boldsymbol{\theta}_n'\boldsymbol{\theta}_{n0})^2(1 - (\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2) + f_{n4}\mathrm{E}[(\mathbf{S}_{n1}'\boldsymbol{\theta}_{n0})^4]. \qquad (\text{B.52})
\end{aligned}$$

Standard formulas for the Kronecker product yield

$$\mathrm{E}[(\mathbf{S}_{n1}'\boldsymbol{\theta}_{n0})^4] = \mathrm{E}[(\boldsymbol{\theta}_{n0}'\mathbf{S}_{n1}\mathbf{S}_{n1}'\boldsymbol{\theta}_{n0})^2] = (\boldsymbol{\theta}_{n0}\otimes\boldsymbol{\theta}_{n0})'\mathrm{E}[\mathrm{vec}(\mathbf{S}_{n1}\mathbf{S}_{n1}')\mathrm{vec}'(\mathbf{S}_{n1}\mathbf{S}_{n1}')](\boldsymbol{\theta}_{n0}\otimes\boldsymbol{\theta}_{n0})$$

$$= (\boldsymbol{\theta}_{n0}\otimes\boldsymbol{\theta}_{n0})'(\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}\otimes\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n})\mathrm{E}[\mathrm{vec}(\mathbf{U}_n\mathbf{U}_n')\mathrm{vec}'(\mathbf{U}_n\mathbf{U}_n')](\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}'\otimes\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}')(\boldsymbol{\theta}_{n0}\otimes\boldsymbol{\theta}_{n0})$$

$$= \frac{1}{p_n^2 - 1}(\boldsymbol{\theta}_{n0}\otimes\boldsymbol{\theta}_{n0})'(\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}\otimes\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n})\big(\mathbf{I}_{(p_n-1)^2} + \mathbf{K}_{p_n-1} + \mathbf{J}_{p_n-1}\big)(\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}'\otimes\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}')(\boldsymbol{\theta}_{n0}\otimes\boldsymbol{\theta}_{n0}),$$

where $\mathbf{K}_\ell$ is the $\ell\times\ell$ commutation matrix and where we let $\mathbf{J}_\ell = (\mathrm{vec}\,\mathbf{I}_\ell)(\mathrm{vec}\,\mathbf{I}_\ell)'$; see [35], page 244. Using the fact that $\mathbf{K}_\ell(\mathbf{A}\otimes\mathbf{B}) = (\mathbf{A}\otimes\mathbf{B})\mathbf{K}_{\ell'}$ for $\ell\times\ell'$ matrices $\mathbf{A}$ and $\mathbf{B}$, along with the identity $\mathbf{K}_1 = 1$, we obtain

$$\begin{aligned}
\mathrm{E}[(\mathbf{S}_{n1}'\boldsymbol{\theta}_{n0})^4] &= \frac{2}{p_n^2-1}(\boldsymbol{\theta}_{n0}\otimes\boldsymbol{\theta}_{n0})'(\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}\otimes\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n})(\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}'\otimes\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}')(\boldsymbol{\theta}_{n0}\otimes\boldsymbol{\theta}_{n0}) \\
&\quad + \frac{1}{p_n^2-1}(\boldsymbol{\theta}_{n0}\otimes\boldsymbol{\theta}_{n0})'\mathrm{vec}(\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}')\mathrm{vec}'(\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}\boldsymbol{\Gamma}_{\boldsymbol{\theta}_n}')(\boldsymbol{\theta}_{n0}\otimes\boldsymbol{\theta}_{n0}) \\
&= \frac{3}{p_n^2-1}(1 - (\boldsymbol{\theta}_{n0}'\boldsymbol{\theta}_n)^2)^2.
\end{aligned}$$

Plugging this in (B.52) provides the result. □

Proof of Theorem 6. Fix a sequence of hypotheses $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_{n0}+\nu_n\boldsymbol{\tau}_n,F_n}$ associated with a given regime (i) to (v) in Theorem 5. Throughout the proof, stochastic convergences, expectations and variances refer to this sequence of hypotheses. In view of the decomposition $\widetilde{W}_n - W_n^* = L_n^{-1}(1-L_n)W_n^*$ from (A.21), it is sufficient to show that L_n converges to one in quadratic mean (note indeed that Theorem 5 indeed implies that W_n^* is $O_\mathrm{P}(1)$). In order to do so, write

$$\begin{aligned}
\mathrm{E}\big[(L_n-1)^2\big] &= \frac{1}{f_{n2}^2}\,\mathrm{E}\Bigg[\bigg(f_{n2} - \bigg[\frac{1}{n}\sum_{i=1}^n V_{ni}^2\bigg]\bigg)^2\Bigg] = \frac{1}{f_{n2}^2}\,\mathrm{E}\Bigg[\bigg(\bigg[\frac{1}{n}\sum_{i=1}^n(\mathbf{X}_{ni}'\boldsymbol{\theta}_{n0})^2\bigg] - e_{n2}\bigg)^2\Bigg] \\
&= \frac{1}{f_{n2}^2}\,\mathrm{E}\Bigg[\bigg(\bigg[\frac{1}{n}\sum_{i=1}^n(\mathbf{X}_{ni}'\boldsymbol{\theta}_{n0})^2\bigg] - \mathrm{E}[(\mathbf{X}_{n1}'\boldsymbol{\theta}_{n0})^2] + \mathrm{E}[(\mathbf{X}_{n1}'\boldsymbol{\theta}_{n0})^2] - e_{n2}\bigg)^2\Bigg] \\
&\le \frac{2}{f_{n2}^2}\,\mathrm{Var}\bigg[\frac{1}{n}\sum_{i=1}^n(\mathbf{X}_{ni}'\boldsymbol{\theta}_{n0})^2\bigg] + \frac{2}{f_{n2}^2}\big(\mathrm{E}[(\mathbf{X}_{n1}'\boldsymbol{\theta}_{n0})^2] - e_{n2}\big)^2 \\
&\le \frac{2}{nf_{n2}^2}\Big(\mathrm{E}\big[(\mathbf{X}_{n1}'\boldsymbol{\theta}_{n0})^4\big] - \big(\mathrm{E}\big[(\mathbf{X}_{n1}'\boldsymbol{\theta}_{n0})^2\big]\big)^2\Big) + \frac{2}{f_{n2}^2}\big(\mathrm{E}[(\mathbf{X}_{n1}'\boldsymbol{\theta}_{n0})^2] - e_{n2}\big)^2 \\
&=: 2T_{na} + 2T_{nb},
\end{aligned}$$

say. Since $f_{n2} = 1 - e_{n2}$, Lemma 8 provides

$$\mathrm{E}[(\mathbf{X}'_{n1}\boldsymbol{\theta}_{n0})^2] - e_{n2} = \left(\frac{f_{n2}}{p_n - 1} - e_{n2}\right)(1 - (\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2) = \frac{1 - p_n e_{n2}}{p_n - 1}(1 - (\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2),$$

which yields

$$T_{nb} = \frac{(1 - p_n e_{n2})^2}{(p_n - 1)^2 f_{n2}^2}(1 - (\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2 = \frac{1 + p_n^2 e_{n2}^2}{p_n^2 f_{n2}^2}\, O(\nu_n^4).$$

In each of the regimes considered in Theorem 5, we thus obtain that $T_{nb} = o(1)$, irrespective of the fact that $\sqrt{p_n} e_{n2} = o(1)$ or not. Turning to T_{na}, Lemma 8 yields

$$\begin{aligned} n f_{n2}^2 T_{na} &= e_{n4}(\boldsymbol{\theta}'_n\boldsymbol{\theta}_{n0})^4 + \frac{6(e_{n2} - e_{n4})}{p_n - 1}(\boldsymbol{\theta}'_n\boldsymbol{\theta}_{n0})^2(1 - (\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2) + \frac{3f_{n4}}{p_n^2 - 1}(1 - (\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2 \\ &\quad - \left(e_{n2}^2(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^4 + \frac{2e_{n2}f_{n2}}{p_n - 1}(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2(1 - (\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2) + \frac{f_{n2}^2}{(p_n - 1)^2}(1 - (\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2\right) \\ &= (e_{n4} - e_{n2}^2)(\boldsymbol{\theta}'_n\boldsymbol{\theta}_{n0})^4 + \left(\frac{6(e_{n2} - e_{n4})}{p_n - 1} - \frac{2e_{n2}f_{n2}}{p_n - 1}\right)(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2(1 - (\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2) \\ &\quad + \left(\frac{3f_{n4}}{p_n^2 - 1} - \frac{f_{n2}^2}{(p_n - 1)^2}\right)(1 - (\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2. \end{aligned}$$

Using the facts that $e_{n4} - e_{n2}^2 = \mathrm{Var}[U_{n1}^2] = \mathrm{Var}[V_{n1}^2] \leq \mathrm{E}[V_{n1}^4] = f_{n4}$ and that $e_{n2} - e_{n4} = \mathrm{E}[U_{n1}^2(1 - U_{n1}^2)] \leq \mathrm{E}[1 - U_{n1}^2] = f_{n2}$, we then obtain

$$\begin{aligned} T_{na} &\leq \frac{f_{n4}}{n f_{n2}^2}(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2 + \frac{6 - 2e_{n2}}{n(p_n - 1)f_{n2}}(\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2(1 - (\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2) \\ &\qquad + \left(\frac{3f_{n4}}{n(p_n^2 - 1)f_{n2}^2} - \frac{1}{n(p_n - 1)^2}\right)(1 - (\boldsymbol{\theta}'_{n0}\boldsymbol{\theta}_n)^2)^2 \\ &= o(1) + \frac{1}{n p_n f_{n2}}\, O(\nu_n^2) + o(\nu_n^4) = o(1) + \frac{1}{n p_n f_{n2}}\, O(\nu_n^2). \end{aligned}$$

Trivially, we then have $T_{na} = o(1)$ in each of the regime considered in Theorem 5, still irrespective of the fact that $\sqrt{p_n} e_{n2} = o(1)$ or not. This establishes the result. □

References

1. Amos, D.E.: Computation of modified bessel functions and their ratios. Mathematics of Computation **28**, 239–251 (1974)
2. Arnold, R., Jupp, P.: Statistics of orthogonal axial frames. Biometrika **100**, 571–586 (2013)
3. Banerjee, A., Dhillon, I., Ghosh, J., Sra, S.: Generative model-based clustering of directional data. In Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, 19–28 (2003)
4. Banerjee, A., Dhillon, I., Ghosh, J., Sra, S.: Clustering on the unit hypersphere using von Mises–Fisher distributions. J. Mach. Learn. Res. **6**, 1345–1382 (2005)
5. Bickel, P.J., Klaassen, C.A., Ritov, Y., Wellner, J.A.: Efficient and Adaptive Estimation for Semiparametric Models. Springer, New York (1998)
6. Billingsley, P.: Probability and Measure, 3rd edition edn. Wiley, New York, Chichester (1995)
7. Cai, T., Fan, J., Jiang, T.: Distributions of angles in random packing on spheres. J. Mach. Learn. Res. **14**, 1837–1864 (2013)

8. Chan, Y., He, X.: On median-type estimators of direction for the von mises-fisher distributions. Biometrika **80**, 869–875 (1993)
9. Chang, T., Rivest, L.P.: M-estimation for location and regression parameters in group models: A case study using stiefel manifolds. Ann. Statist. **29**, 784–814 (2001)
10. Chikuse, Y.: High dimensional limit theorems and matrix decompositions on the Stiefel manifold. J. Multivariate anal. **36**, 145–162 (1991)
11. Cuesta-Albertos, J.A., Cuevas, A., Fraiman, R.: On projection-based tests for directional and compositional data. Stat. Comput. **19**, 367–380 (2009)
12. Cutting, C., Paindaveine, D., Verdebout, T.: Testing uniformity on high-dimensional spheres against monotone rotationally symmetric alternatives. Ann. Statist. **45**, 1024–1058 (2017a)
13. Cutting, C., Paindaveine, D., Verdebout, T.: Supplement to "testing uniformity on high-dimensional spheres against monotone rotationally symmetric alternatives". Ann. Statist. **45**, 1024–1058 (2017b)
14. Downs, T.: Spherical regression. Biometrika **90**, 655–668 (2003)
15. Dryden, I.L.: Statistical analysis on high-dimensional spheres and shape spaces. Ann. Statist. **33**, 1643–1665 (2005)
16. Giri, N.C.: Group Invariance in Statistical Inference. World Scientific Publishing Company, Singapore (1996)
17. Guttorp, P., Lockhart, R.: Finding the location of a signal: a bayesian analysis. J. Amer. Statist. Assoc. **83**, 322–330 (1988)
18. Hallin, M., Ley, C.: Skew-symmetric distributions and Fisher information: the double sin of the skew-normal. Bernoulli **20**, 1432–1453 (2014)
19. He, X., Simpson, D.: Robust direction estimation. Ann. Statist. **20**, 351–369 (1992)
20. Hornika, K., Grün, B.: movmf: An R package for fitting mixtures of von mises-fisher distributions. J. Statist. Softw. **58** (2014)
21. Joshi, C., Bissu, S.: Some inequalities of bessel and modified bessel functions. J. Austral. Math. Soc. (Series A) **50**, 333–342 (1991)
22. Larsen, P.V., Blæsild, P., Sørensen, M.: Improved likelihood ratio tests on the von mises–fisher distribution. Biometrika **89**, 947–951 (2002)
23. Le Cam, L., Yang, G.L.: Asymptotics in Statistics: Some Basic Concepts. Springer, New York (2000)
24. Lehmann, E., Romano, J.: Testing Statistical Hypotheses. Springer, New York (2005)
25. Ley, C., Paindaveine, D., Verdebout, T.: High-dimensional tests for spherical location and spiked covariance. J. Multivariate Anal., **139**, 79–91 (2015)
26. Ley, C., Swan, Y., Thiam, B., Verdebout, T.: Optimal R-estimation of a spherical location. Statist. Sinica **23**, 305–333 (2013)
27. Ley, C., Verdebout, T.: Modern Directional Statistics. CRC Press (2017)
28. Mardia, K.V., Jupp, P.E.: Directional Statistics. John Wiley & Sons (2000)
29. Paindaveine, D., Remy, J., Verdebout, T.: Testing for principal component directions under weak identifiability. Ann. Statist., to appear (2019)
30. Paindaveine, D., Verdebout, T.: Optimal rank-based tests for the location parameter of a rotationally symmetric distribution on the hypersphere. In: M. Hallin, D. Mason, D. Pfeifer, J. Steinebach (eds.) Mathematical Statistics and Limit Theorems: Festschrift in Honor of Paul Deheuvels, pp. 249–270. Springer (2015)
31. Paindaveine, D., Verdebout, T.: Inference on the mode of weak directional signals: a Le Cam perspective on hypothesis testing near singularities. Ann. Statist. **45**, 800–832 (2017)
32. Paindaveine, D., Verdebout, T.: Inference for spherical location under high concentration. ArXiv preprint arXiv:1901.00359v1 (2019)
33. Rivest, L.P.: Spherical regression for concentrated Fisher–von Mises distributions. Ann. Statist. **17**, 307–317 (1989)
34. Schou, G.: Estimation of the concentration parameter in von Mises–Fisher distributions. Biometrika **65**, 369–377 (1978)
35. Tyler, D.: A distribution-free M-estimator of multivariate scatter. Ann. Statist. **15**, 234–251 (1987)
36. Van der Vaart, A.: Asymptotic Statistics. Cambridge University Press, Cambridge, United Kingdom (1998)

37. Watson, G.S.: Statistics on Spheres. Wiley, New York (1983)